Mechanism of Dihydroneopterin Aldolase: A Molecular Dynamics Study of the Apo Enzyme and Its Product Complex

Lishan Yao,[†,‡] Honggao Yan,[*,§,‡] and Robert I. Cukier[*,†,‡]

Department of Chemistry, Department of Biochemistry and Molecular Biology, and MSU Quantitative Biology and Modeling Initiative, Michigan State University, East Lansing, Michigan 48824

Received: August 27, 2005; In Final Form: November 21, 2005

Dihydroneopterin aldolase (DHNA), an enzyme in the pathway that generates folic acid in bacteria, is investigated by a series of molecular dynamics simulations in its free form and complexed with its product, 6-hydroxymethyl-7,8-dihydropterin (HP). The active sites in DHNA are formed at the interface between pairs of protomers in this octameric protein. On the basis of root-mean-square deviation and root-mean-square fluctuation analyses of the trajectories, which take advantage of the presence of eight active sites, flexible regions of the apo protein surrounding the active site are identified and, upon binding HP, show that the active site is rigidified. Specific residues, associated with binding and the catalytic mechanism of DHNA, are associated with these flexible regions, and their interactions with HP account for most of the binding energy. A Principal Component Analysis shows rigidification of DHNA upon HP binding and that only a few modes of motion capture most of the atomic fluctuations in both apo and HP-bound forms. HP is pushed out of the active site in a series of simulations with different restrained positions between HP and DHNA to obtain a view of the exit pathway and energetic barrier to product release. The chosen pathway leads to a minimal disturbance of the system and provides a barrier consistent with the experimentally determined rate of product release. An analysis of the various components that contribute to the exit path energy and entropy provides insight into the energy−entropy compensation for product release.

I. Introduction

Dihydroneopterin aldolase (DHNA) catalyzes the conversion of 7,8-dihydroneopterin into 6-hydroxymethyl-7,8-dihydropterin (HP) and glycolaldehyde that may occur by a retroaldol-based, carbon−carbon bond cleavage mechanism.[1] Plants and most microorganisms obtain folate cofactors by de novo synthesis, while vertebrates must depend on nutritional sources.[2,3] Therefore, DHNA, along with other enzymes in the folate pathway, is a potential target for antimicrobial and anti-parasitic drugs.[3] The particular enzyme under study, DHNA from *Staphylococcus aureus*, has recently been explored for this purpose.[4]

S. aureus DHNA crystallizes as a D_4-symmetric homo-octamer.[5,6] A solution NMR study also leads to an octameric protein structure.[7] Structures are available for the apo and HP-bound forms of DHNA.[5] There are eight active sites that form at the interfaces between pairs of adjacent protomers. The bottom of the active site has a glutamate (Glu74) that serves as an "anchor" for binding and there are residues lining the active site that stabilize HP. Why DHNA and at least one other folate pathway enzyme are octameric proteins is not clear. There is a suggestion that because DHNA can be a partner in a bifunctional enzyme, an active site formed from a noncovalent interface between protomers may aid in product/reactant transport between the active sites.[8]

In this article, DHNA and its product-bound complex are studied by molecular dynamics (MD) simulations and complemented by a series of simulations based on pushing HP out of its active site. The octameric structure with its eight active sites provides the advantage that for a given simulation length, assuming independence of the active sites (in apo and HP-bound forms), there is eight times the data that would be available from a monomeric protein simulation. Comparing statistical measures of protein residue fluctuations that are based either on the entire octamer or on the individual dimeric units provides a convenient way to study the active sites of DHNA. In particular, the issue of the flexibility of residues forming the active site and their rigidification by HP binding can be addressed. The binding of HP does indeed rigidify DHNA both in DHNA's flexible regions and in its more rigid regions. Another view of such changes in DHNA can be obtained by Principal Component Analysis (PCA),[9] which decomposes the fluctuations of a protein into, typically, a relatively small set of modes that account for a great deal of the motion and a remainder, large set of small fluctuations that contribute little to the total fluctuation.[10] As applied to apo and HP-bound DHNA, the PCA analysis shows that in both structures there are small sets of modes that do dominate the total protein fluctuation. Furthermore, HP binding substantially reduces the total fluctuation relative to the apo form, again indicating the rigidification of the protein induced by HP. The character of the PCA modes and their differences between apo and HP bound forms provide insight into the nature of the binding site.

As noted above, the active site is formed by the noncovalent interaction of protomer pairs and that provides impetus for exploring the exit pathway of bound HP. Such a study requires determination of an exit pathway and construction of the potential of mean force along the path. These are both challenging tasks since the true de-binding path most likely is

* Address correspondence to these authors. R.I.C.: e-mail cukier@cem.msu.edu; phone 517-355-9715, ext 263; fax 517-353-1793. H.Y.: e-mail yanh@msu.edu; phone 517-353-5282; fax 517-353-9334.
† Department of Chemistry.
‡ MSU Quantitative Biology and Modeling Initiative.
§ Department of Biochemistry and Molecular Biology.

10.1021/jp054854n CCC: $33.50
Published on Web 12/30/2005

a high dimensional one that is not determinable exogenously and, once chosen, constructing a potential of mean force to cover the bound to free form is computationally expensive.[11−14] Here, a compromise approach will be used where, by examination of the crystal structure bound form, a possible exit pathway is found and implemented by restraining the distance between one atom in DHNA and one in HP. Provided that the atoms are suitably chosen, and the distance between the atoms is increased at a sufficiently slow rate, the protein structure should not be distorted by the added force required to push out the ligand. By using a restraining potential between only a pair of atoms, HP should be free to translate and rotate to find an exit path from the bound state that should resemble the true path.

Instead of trying to determine directly the free energy along the exit pathway, only the energetics of the HP-DHNA interaction will be evaluated with averages over time windows that are chosen to obtain reliable energy averages. This energetics will serve to show which residues are involved in gradually releasing HP from the binding pocket. The internal energy of HP along the exit pathway is also evaluated since it could provide a substantial energetic contribution from the release of strain energy as HP exits from its binding site. An estimate of the free energy can be obtained by incorporating entropic contributions from the increasing freedom of HP as it is released from the binding site and the difference in the DHNA entropy of the apo and HP-bound forms. The summation of these entropic and energetic terms then provides an estimate of the kinetic barrier to HP release.

II. Methods

Starting coordinates for the MD simulations of the apo and HP product complex forms of DHNA were taken from crystal structures.[5] The protein atom coordinates of the two forms are essentially the same, with a ∼0.3 Å root-mean-square deviation (RMSD) for all CA atoms. The N5 of HP was deprotonated in accordance with the spectroscopic evidence[15] and all residues were assumed to be in their standard ionization states at pH 7. Atom-centered partial charges were derived by using the AMBER antechamber program (RESP methodology).[16] The RESP charges for HP were obtained at the HF/6-31G* level.

MD simulations for apo and HP complex forms were performed by using the SANDER module in the AMBER 7.0 program, with the amber94 force field. All crystallographic waters were removed and the system solvated with ∼19 500 TIP3P water molecules and 32 Na^+ ions were added to neutralize the system. The MD simulations were run at constant volume for 5 ps to heat up the system from 100 to 300 K, then at constant temperature and pressure for ∼25 ps to get the proper density, and then at constant temperature and volume for 2 ns. The simulation box is approximately a cube of side 90 Å, and the octameric protein is cylindrical in shape with diameter and major axis approximately 70 Å. All velocities were assigned according to the Maxwell−Boltzmann distribution at 100 K; and the system was coupled to a temperature bath with a time constant of 0.5 ps.[17] The Particle-Mesh-Ewald method[18] was used to evaluate the contributions of the long-range electrostatic interactions. A nonbonded pair list cutoff of 8.0 Å was used and the nonbonded pair list was updated every 25 steps. All bonds to hydrogen atoms were constrained by using the SHAKE algorithm[19] allowing a time step of 0.002 ps. The coordinates were saved every 2 ps, and the trajectories after 500 ps were analyzed with the Moil-view program and the PTRAJ module of AMBER 7.0. Hydrogen bonds were assigned when the distance of two heavy atoms (O or N) is less than 3.2 Å and the angle (heavy atom−hydrogen−heavy atom) is greater than 120°.

The starting structure for pushing HP out of the active site was taken from a snapshot of the HP-DHNA complex simulation. A harmonic restraint was added between the backbone N of Arg118 and C10 of HP. The reasons to choose the N of Arg118 are the following: (1) Arg118 is in the C-terminal β sheet, which is quite rigid (see discussion below). (2) The vector between Arg118 N and HP C10 points along the channel between protomer pairs that is quite open and water accessible. Different force constant values were tested in the simulation. Using too small a constant may not push out the ligand efficiently, while using too large a constant may take too long to equilibrate the system. The force constant finally used was 20 kcal/(mol·Å^2). The equilibrium distances were changed by 0.2 Å every 50 ps. Since the force constant used corresponds to rms fluctuations around the equilibrium distance of ∼0.25 Å, this pushing rate should give us continuous sampling along the exit path. HP was pushed out by 8 Å in 2 ns. It moves out of the active site and stays on the surface of the protein. A simulation was also run that starts with all sites occupied by HP but only pushes out one of them, otherwise using the same protocol and parameter values. The HP protein interaction energetics along the trajectory was found to be similar to that for pushing all out at the same time. The energetic fluctuations for pushing out one protomer were greater than when averaged over the eight protomers, as expected.

Root-mean-square fluctuations (RMSF) were analyzed in two ways, which are based on different averaging procedures, stimulated by the feature that DHNA is a homooctamer. Denote the vector location of atom r for the tth configuration in the ith protomer as $\mathbf{r}_t^i = \{x_t^i, y_t^i, z_t^i\}$. The trajectory average over the N_C configurations ($t = 1, 2, ..., N_C$) for the ith protomer is

$$\langle \mathbf{r}_t^i \rangle_C = \frac{1}{N_C} \sum_{t=1}^{N_C} \mathbf{r}_t^i \tag{1}$$

and the further average over the N_P protomers ($i = 1, 2, ..., N_P$) is

$$\langle \langle \mathbf{r}_t^i \rangle_C \rangle_P = \frac{1}{N_P} \sum_{i=1}^{N_P} \langle \mathbf{r}_t^i \rangle_C \tag{2}$$

Also denote deviations from these two averages as $\delta_i \mathbf{r}_t^i = \mathbf{r}_t^i - \langle \mathbf{r}_t^i \rangle_C$ and $\delta \mathbf{r}_t^i = \mathbf{r}_t^i - \langle \langle \mathbf{r}_t^i \rangle_C \rangle_P$, with δ_i distinguishing which protomer average is used in the former definition. The first RMSF we define is

$$\mathrm{RMSF1} = \langle \langle (\delta_i \mathbf{r}_t^i)^2 \rangle_C^{1/2} \rangle_P \equiv \langle \mathrm{RMSF}_i \rangle_P \tag{3}$$

RMSF1 is the average over protomers of the ith protomer RMSF_i and will be referred to as the average protomer RMSF. The second RMSF we define is

$$\mathrm{RMSF2} = \langle \langle (\delta \mathbf{r}_t^i)^2 \rangle_C \rangle_P^{1/2} \tag{4}$$

RMSF2 is an RMSF obtained by treating the trajectory as one record over the entire $(N_P)*(N_C)$ data, and will be referred to as the octamer RMSF. RMSF2 does not distinguish atom r among the different protomers. To the extent that the averages in these deviations are different, there is no rigorous connection between the two RMSFs. If, however, the protomer averages are

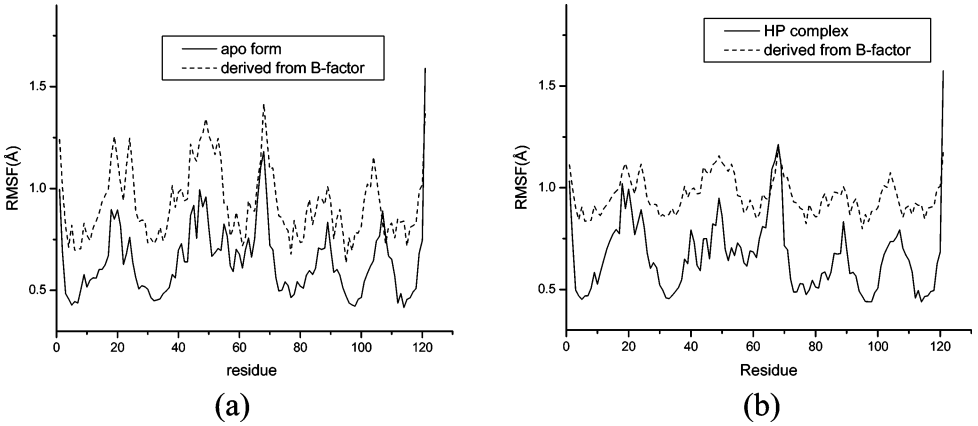

Figure 1. CA atom RMSDs from the simulations and corresponding crystal structure B factors: (a) Apo form and (b) HP bound form.

independent of protomer, $\langle \mathbf{r}_t^i \rangle_C = \langle \langle \mathbf{r}_t^i \rangle_C \rangle_P$ ($i = 1, ..., N_P$), then they can be related as follows. Using the protomer RMSF$_i$ definition in eq 3, the difference between RMSF2 and RMSF1 can be expressed as

$$\Delta\text{RMSF} \equiv [\text{RMSF2}^2 - \text{RMSF1}^2]^{1/2} = \langle \text{RMSF}_i^2 \rangle_P - \langle \text{RMSF}_i \rangle_P^2 > 0 \quad (5)$$

Thus, ΔRMSF is the RMS fluctuation over different protomers of the individual protomer RMSF$_i$s. This indicates that, to the extent that the average values of the vector positions are independent of protomer, as they will be approximately (when referenced to the same coordinate origin), the RMSF2s will be larger than the RMSF1s. This difference should be magnified for the atoms that are intrinsically more mobile, and this feature will be useful for distinguishing fluctuations of given atoms in the apo versus bound forms of DHNA. When the protomer averages are dependent on protomer, the inequality in eq 5 may be violated, and RMSF1 could be slightly larger than RMSF2.

The principal component analysis was carried out with the gcovar and ganaeig modules of GROMACS 3.0.[20] The PCA method diagonalizes the covariance matrix $\sigma_{ij} = \langle \delta x_i \delta x_j \rangle$ of the CA atom fluctuations from their trajectory-averaged $\langle ... \rangle$ values, where $\delta x_i = x_i - \langle x_i \rangle$ and x_i denotes the Cartesian components of the ith CA atom.[10] For this analysis, $\langle ... \rangle = \langle \langle ... \rangle_C \rangle_P$; the entire trajectory is used and corresponds to the octamer-based RMSF approach. The overall translational/rotational motion of all the snapshots is removed relative to a reference structure before matrix diagonalization. The sum of the eigenvalues of the covariance matrix corresponds to the total variance of the motion over the trajectory. Ordering the eigenvalues from large to small can, in favorable cases, provide a small set of modes that captures most of the protein's fluctuations. Another application of PCA is to the evaluation of entropy.[21] Assuming that the covariance matrix σ_{ij} describes the protein dynamics permits the entropy change in DHNA for binding of HP to be estimated as:

$$\Delta S = \frac{k_B}{2} \ln\left[\frac{\det(\sigma_a)}{\det(\sigma_b)}\right] \quad (6)$$

where k_B is Boltzmann's constant, and σ_a and σ_b are the above covariance matrices in the bound and free forms, respectively. After diagonalization of these two matrices, the entropy change can be obtained since the determinants in eq 6 are expressible as the products of their PCA eigenvalues.

The entropy of HP was estimated by using the quasih module of AMBER 7.0 to obtain the normal modes that can be used in standard formulas from statistical thermodynamics.

III. Results and Discussion

1. Apo DHNA and Its HP Complex. *A. Flexibility Analysis of DHNA and Its HP Complex Based on the Octamer and Its Protomers.* DHNA is an octamer, with its eight active sites that bind HP formed by interfaces between pairs of adjacent protomers. There is no cooperation among the active sites based on kinetic studies[22,23] and, therefore, they can be treated as statistically independent units, indicating that more conformational space can be explored than in a usual MD simulation of a given duration. The RMSDs of the CA atoms of the instantaneous MD structures from the initial X-ray structure for the apo DHNA and its HP complex are ca. 1.1−1.5 Å between 0.5 and 2.0 ns (data not shown). For this comparison, the superposition of structures for the RMSD evaluation was based on the octamer providing an effective simulation time of 12 ns (8 × 1.5 ns of time after 0.5 ns for equilibration).

The RMS fluctuations (RMSF) of the CA atoms for each residue (Figure 1a,b) were first calculated by the average protomer RMSF method, as discussed in the Methods section. This approach evaluates each protomer RMSF and averages them over the eight protomers to obtain the average protomer RMSF. Overall, the protein is quite rigid with an average RMSF of ∼0.6 Å. The binding of HP apparently does not change the RMSF of the protein on the MD time scale. The RMSF variations along the amino acid sequence obtained from the MD simulations and experimental crystal B-factors are quite similar, with a correlation coefficient 0.79 for the apo form and 0.77 for the HP complex.

The MD trajectory of the DHNA octamer was then partitioned into the coordinate sets for the individual protomers and combined to obtain a data record of length N_P*N_C with averages and fluctuations from averages based on all the data for a given CA. This octamer-based RMSF is the second approach discussed in the Methods section and, as noted there, will tend to emphasize the fluctuations. These octamer RMSF values, compared with the previous average protomer RMSF, are plotted in Figure 2. The new RMSFs display a pattern similar to the previous ones, but with larger variations, especially for the apo enzyme (Figure 2a), in some regions. Strikingly, all these regions are around the active sites, at the interface of two subunits (Figure 3), including residues 15−25, 45*−55*, 68−74, and 100−110. (We will henceforth denote residues in the adjacent subunit that together form the active site with an asterisk, "*",

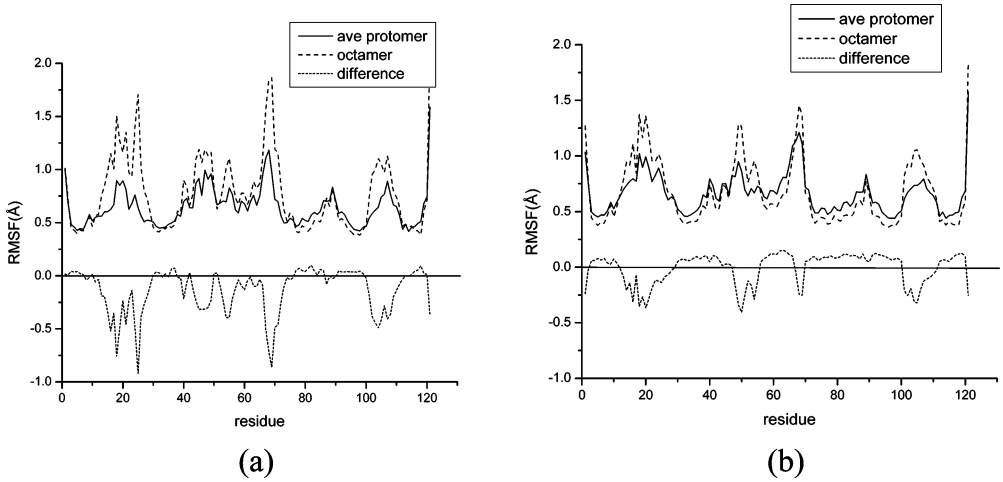

Figure 2. Comparison of the RMSFs derived from the average protomer RMSF and the octamer RMSF (see text for the difference in these RMSFs): (a) Apo form and (b) HP bound form.

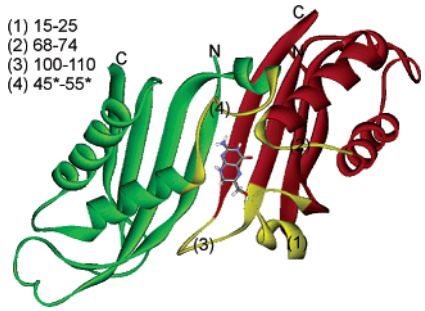

Figure 3. A representative snapshot of two adjacent subunits of apo DHNA from the MD simulation. HP is placed for the identification of the active site. One subunit is in green, and the other in red. The regions corresponding to large RMSF differences are highlighted in gold color. For clarity, only one set of the regions is highlighted.

as in the notation 45*−55*). This difference reflects the limited 2 ns sampling space around the initial subunit structures. The "extra" flexibility around the active site might assist in ligand binding and release, as discussed below.

The increase in RMSF on an octamer versus average protomer basis is unanimous in the apo enzyme simulation. In the HP complex simulation, the analogous increase is smaller, which indicates that HP binding rigidifies the enzyme. This feature can be seen clearly in Figure 4a, where the apo and HP-bound RMSF differences are compared for octamer and average protomer methods. Thus, while RMSF differences based on the average protomer RMSF would not support the hypothesis that the binding of HP rigidifies the protein, the differences based on the octamer RMSF show clearly that the binding of HP rigidifies DHNA.

While most parts of DHNA are rigidified by binding HP, residues 49−51 seem more flexible after HP binding (Figure 4a, dotted line). Further investigation indicates that the ψ dihedral angle of Asp*50 and the φ dihedral angle of Thr*51 can change cooperatively to give a different conformation for Asp*50 and Thr*51. Figure 5 shows the time evolution of these two dihedral angles in the different subunits displayed sequentially as one trajectory. In subunits 1, 2, 3, 4, 7, and 8, φ and ψ remain at ∼0° and −140°, respectively; but in subunits 5 and 6 (especially 6), ψ changes from 0° to −40°, at the same time that φ changes from −140° to −70°. These simultaneous changes of dihedral angles may be caused by the new interactions between the Thr*51 side chain and its surroundings (see discussion below). After excluding subunit 6 from the trajectory, the new RMSF differences between apo enzyme and HP bound enzyme show no positive spike around residue 50 (Figure 4b).

B. Principal Component Analysis. The rigidification of DHNA upon HP binding can also be investigated by the PCA method, which attempts to decompose the total variance of the atom RMS fluctuations over the MD trajectory into a contribution from a small set of modes that have a relatively large contribution to this variance and a large set of modes with a small summed contribution.[10] Because the above analysis does show that there are flexible regions of DHNA that become rigidified upon HP binding, a comparison by PCA of the two forms should be instructive. The trajectory of subunit 6 was excluded in the PCA analysis, because Asp*50 and Thr*51 in subunit 6 have backbone conformations quite different from those in the other subunits. To highlight the flexible regions in the PCA analysis, only CAs of rigid regions (with RMSF less than 0.7 Å in the free form) were fitted to remove the overall translation/rotational motion. The new fitting gives a total variance of 96.6 Å^2 for all CAs in the apo form and 67.2 Å^2 in the bound form. The first four modes contribute 56% of the variance in the apo and 39% in the bound form. The remaining modes contribute ∼43 Å^2 in the apo form and ∼41 Å^2 in the bound form. Thus, as indicated in Figure 6, the first four modes incorporate most of the flexibility difference between the apo and bound forms. That a few modes capture most of the difference in RMS fluctuations between the apo and bound forms reinforces the data obtained above that a limited number of residues are responsible for the changes in DHNA flexibility upon HP binding. The first mode has a variance of ∼25 Å^2 in the free form but only ∼9 Å^2 in the HP bound form.

The RMS fluctuations of each CA corresponding to the trajectory based on mode 1 for the apo and HP-bound forms are displayed in Figure 7. They show that the enhanced flexibility of the apo relative to bound form is concentrated in three regions, 45*−55*, 68−74, and 100−110, which were identified in Section A. These data, along with the data for modes 2 and 3, suggest that examination of three distances between three residues, one in each of these regions, should provide a refined perspective on the active site fluctuations. Labeling the Tyr*54-Ile105 distance as d_1, the Tyr*54-Ala69 distance as d_2, and the Ala69-Ile105 distance as d_3 and parametrically plotting the data for all eight active sites in 3D reveals that the most interesting differences are in the d_2 and d_3 directions, as displayed in Figure 8. The fluctuations in distance

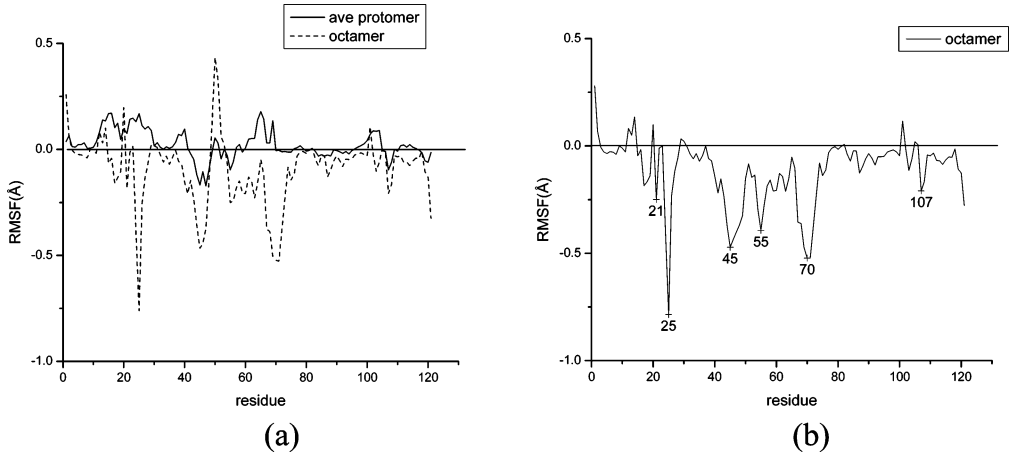

Figure 4. (a) RMSF differences between the apo DHNA and HP complex simulations. The differences based on the average protomer and octamer RMSF methods are displayed with solid and dashed lines, respectively. (b) RMSF differences between the apo DHNA and HP complex simulations excluding protomer 6, based on the octamer RMSF method.

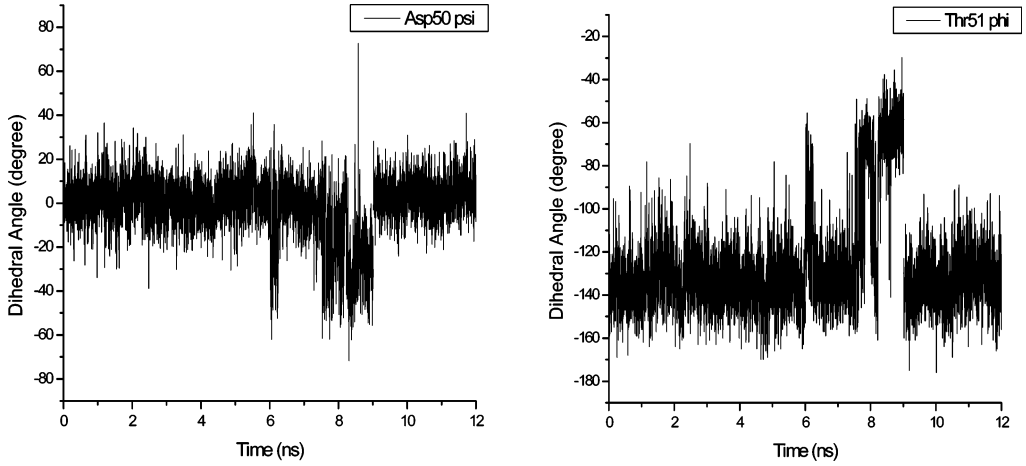

Figure 5. Time evolution of the ψ dihedral angle of Asp 50 and the φ of Thr51 in the eight subunits displayed sequentially as one trajectory.

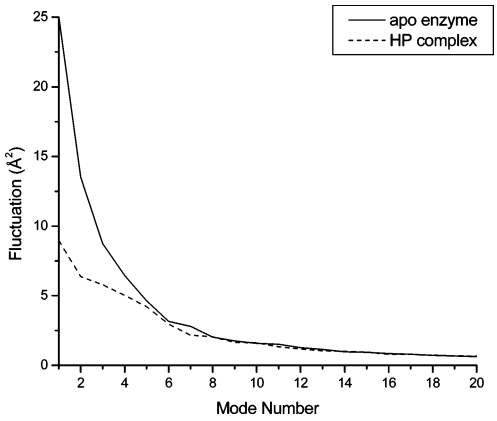

Figure 6. Principal Component Analysis eigenvalues of the first 20 modes in the apo and HP complex simulations. The total fluctuation over all the modes is 96.6 Å^2 for the apo and 67.2 Å^2 for the complex simulation.

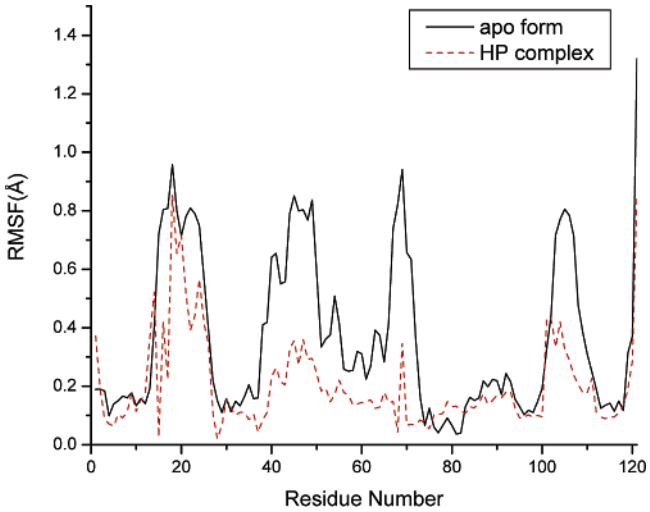

Figure 7. The CA atom RMSF projections onto the first principal component eigenvector for the apo and HP bound forms. The enhanced fluctuation regions of apo versus HP complex are evident and similar to those found as displayed in Figure 4.

along the d_1 direction are quite similar for the apo and HP-bound forms, and its scale is about half of the total d_2 and d_3 apo fluctuation scale. The apo form has a major and a minor state, while the HP-bound form has only one state, which is confined within the major apo state. The apo motion is correlated in the sense that on average an increase in d_2 is accompanied by an increase in d_3, indicating a kind of breathing motion (both distances have a common residue) that is, approximately, an equally weighted combination of the two distances.

The estimated entropy loss of DHNA due to HP binding for CA atoms is about 4.9 cal/(mol·deg) based on the first four modes and about 13.1 cal/(mol·deg) including all the 357 PCA-determined modes, when evaluated from eq 6. It is quite interesting that even though the first several modes contribute most of the flexibility changes in the binding, ~60% of the

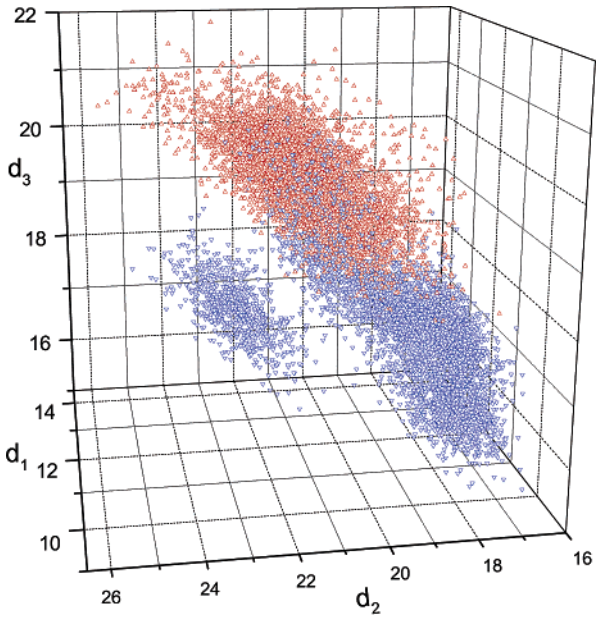

Figure 8. The CA distances: d_1 (Tyr*54 to Ile105), d_2 (Tyr*54 to Ala69), and d_3 (Ala69 to Ile105) parametric on the trajectory for the apo (blue) and HP bound (red) forms. There is a major and minor component for the apo form. The HP bound form distance fluctuations are smaller and fairly well confined to a part of the apo major state volume.

entropic loss comes from the small eigenvalue (high frequency) contribution. This is consistent with the fact that the binding of HP rigidifies not only the flexible regions but also the rigid regions.

C. Active Site Analysis. Since the most significant flexibility changes are localized in the active site region, it is useful to analyze the interactions between the active site residues and between the active site residues and bound HP. In the active site of apo DHNA (Figure 9a), the carboxylate of Glu22 forms a salt bridge with Lys100 and hydrogen bonds with the backbone amides of Ala18 and Leu19 and with the side chain amide of Gln27. The carboxylate of Glu74 forms hydrogen bonds with the hydroxyl of Thr*51 and the phenoxyl of Tyr*54 from the adjacent subunit and the backbone amides of Leu73 and Glu74 from the same subunit. The strong hydrogen bond networks formed by the carboxylates of Glu22 and Glu74 with their surroundings indicate that these two residues are restrained in the apo form. By comparing the hydrogen bond interactions in different active sites (Table 1), the hydrogen bond strengths are shown to be more diverse around Glu22 than Glu74, which may be caused by the difference in flexibility between these two regions. The different active sites between different subunits show quite consistent hydrogen bond interactions (Table 1), which suggests that the active sites are well defined in the apo form MD simulation. All the hydrogen bonds in Table 1 are similar to those found in the crystal structure, with the one exception that the Y54* to Glu74 hydrogen bond, which spans two subunits, is longer in the crystal structure.

The hydrogen bond interactions after binding HP are listed in Tables 2 and 3 and a typical snapshot displayed in Figure 9b. The results are in good agreement with those from the crystal structure. Compared with the apo form, the strong hydrogen bond between the amide and carboxyl of Glu74 is maintained, as well as the hydrogen bond between the side chain hydroxyl

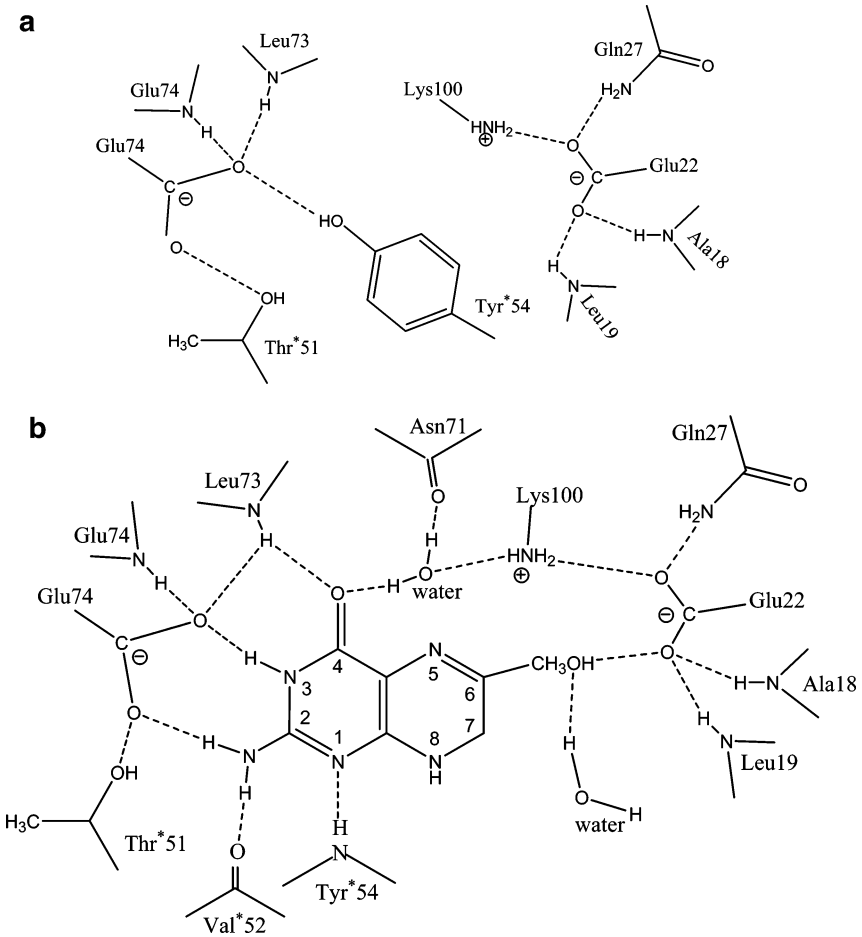

Figure 9. Active site interactions based on the MD simulations of apo enzyme (a) and HP complex (b). Residues in one protomer are distinguished from the other protomer by asterisks.

TABLE 1: Active Site Interactions in Apo DHNA

	subunit =							
	1	2	3	4	5	6	7	8
Glu74-OE1...	2.71 (0.13)	2.70 (0.12)	2.68 (0.13)	2.76 (0.16)	2.68 (0.13)	2.71 (0.13)	2.68 (0.12)	2.71 (0.15)
Thr*51-OGH	100%	100%	100%	100%	100%	100%	100%	100%
Glu74-OE2...	2.67 (0.12)	2.63 (0.10)	2.93 (0.25)	2.62 (0.10)		2.66 (0.12)	2.65 (0.12)	2.62 (0.10)
Tyr*54-OH	100%	87%	35%	100%		100%	100%	100%
Glu74-OE2...	3.23 (0.16)	3.21 (0.16)	3.24 (0.16)	3.28 (0.15)	3.27 (0.16)	3.24 (0.15)	3.17 (0.16)	3.25 (0.15)
Leu73-NH	72%	82%	61%	41%	37%	75%	76%	63%
Glu74-OE2...	2.78 (0.09)	2.76 (0.08)	2.76 (0.09)	2.78 (0.10)	2.77 (0.09)	2.78 (0.09)	2.77 (0.09)	2.76 (0.09)
Glu74-NH	100%	100%	100%	100%	100%	100%	100%	100%
Glu22-OE1...	2.77 (0.11)	2.80 (0.12)	2.74 (0.12)	2.90 (0.18)	2.93 (0.22)	2.75 (0.10)	2.74 (0.09)	2.86 (0.16)
Lys100-NZ[a]	100%	98%	100%	100%	77%	100%	100%	99%
Glu22-OE2...	3.12 (0.20)	3.05 (0.21)	2.96 (0.18)	2.88 (0.15)	3.06 (0.22)	3.12 (0.19)	2.96 (0.17)	3.01 (0.20)
Ala18-NH	51%	60%	81%	97%	100%	100%	91%	78%
Glu22-OE2...	3.26 (0.16)		3.04 (0.18)	3.13 (0.19)	2.99 (0.16)	2.98 (0.16)	2.98 (0.15)	3.15 (0.18)
Leu19-NH	34%		91%	36%	93%	99%	99%	73%
Glu22-OE1...	2.94 (0.17)	2.88 (0.15)		2.90 (0.17)	2.82 (0.12)	2.81 (0.11)	2.80 (0.11)	2.90 (0.15)
Gln27-NEH	96%	97%		53%	99%	100%	100%	97%

[a] Only distance criterion was used for the salt bridge analysis.

TABLE 2: Hydrogen Bonds between HP and Its Surroundings in the HP Complex[a]

	subunit =							
	1	2	3	4	5	6	7	8
N1...	3.15 (0.16)			3.25 (0.15)	3.21 (0.15)	3.25 (0.15)	3.23 (0.15)	
Tyr*54-NH	94%			31%	84%	25%	26%	
N2H1...	2.80 (0.10)	2.80 (0.10)	2.80 (0.10)	2.75 (0.09)	2.81 (0.10)	2.78 (0.10)	2.80 (0.11)	2.80(0.10)
Glu74-OE1	100%	100%	100%	100%	100%	100%	100%	100%
N2H2...	3.06 (0.19)	3.15 (0.20)	3.07 (0.20)	2.97 (0.15)	3.04 (0.18)	3.13 (0.19)	3.04 (0.18)	3.08 (0.20)
Val*52-CO	87%	47%	61%	98%	95%	59%	95%	52%
N3H...	2.79 (0.08)	2.80 (0.09)	2.80 (0.09)	2.99 (0.17)	2.91 (0.14)	2.81 (0.10)	2.83 (0.13)	2.79 (0.09)
Glu74-OE2	100%	100%	100%	98%	100%	100%	99%	100%
O4...	2.96 (0.15)	2.99 (0.16)	2.98 (0.16)	3.17 (0.18)	2.97 (0.16)	3.03 (0.16)	3.03 (0.17)	2.90 (0.13)
Leu73-NH	99%	99%	97%	81%	97%	98%	95%	100%
1′ OH...	2.79 (0.23)	2.62 (0.12)	2.80 (0.26)	2.64 (0.11)	2.67 (0.17)	2.69 (0.22)	2.75 (0.23)	2.63 (0.10)
Glu22-OE1	91%	100%	82%	95%	88%	67%	73%	100%
Glu22-OE2...	2.94 (0.27)	3.17 (0.19)	2.93 (0.30)	3.14 (0.35)	3.07 (0.34)	2.77 (0.31)	2.97 (0.26)	3.24 (0.16)
1′OH	78%	36%	69%	18.4%	57%	45%	54%	54%
O4...	2.85 (0.22)	2.80 (0.18)	2.88 (0.20)	2.80 (0.20)	2.85 (0.22)	2.79 (0.17)	2.77 (0.16)	2.86 (0.19)
H_2O	103%	100%	60%	106%	106%	94%	94%	91%
Asn71-O...	2.85 (0.20)	2.84 (0.19)	3.05 (0.25)	2.74 (0.13)	2.84 (0.19)	2.88 (0.18)	2.79 (0.17)	2.85 (0.19)
H_2O	90%	99%	11%	100%	86%	77%	97%	96%
Lys100-NZ...	2.86 (0.14)	2.92 (0.15)	2.95 (0.19)	2.86 (0.11)	2.88 (0.14)	2.88 (0.13)	2.86 (0.13)	2.87 (0.12)
H_2O	98%	98%	61%	100%	85%	100%	100%	100%
1′OH...	2.85 (0.22)	2.91 (0.22)	2.90 (0.22)	2.85 (0.18)	2.91 (0.21)	3.21 (0.22)	2.83 (0.19)	2.80 (0.17)
H_2O	101%	68%	136%	55%	35%	8%	69%	113%

[a] Some percentages are greater than 100% since more than one atom can hydrogen bond at a given time.

TABLE 3: Active Site Interactions in the HP Complex

	subunit =							
	1	2	3	4	5	6	7	8
Glu74-OE1...	2.64 (0.10)	2.65 (0.11)	2.68 (0.11)	2.98 (0.26)	2.83 (0.21)	2.64 (0.10)	2.68 (0.14)	2.65 (0.11)
Thr*51-OGH	100%	100%	100%	63%	95%	100%	100%	100%
Glu74-OE2...				2.90 (0.23)	3.11 (0.25)		2.96 (0.24)	
Tyr*54-OH				75%	74%		11%	
Glu74-OE2...	3.29 (0.13)	3.32 (0.12)	3.30 (0.13)	3.29 (0.14)	3.27 (0.14)	3.31 (0.13)	3.29 (0.16)	3.32 (0.13)
Leu73-NH	56%	43%	51%	55%	59%	46%	54%	40%
Glu74-OE2...	2.74 (0.08)	2.76 (0.08)	2.78 (0.09)	2.86 (0.11)	2.80 (0.10)	2.76 (0.08)	2.78 (0.09)	2.75 (0.08)
Glu74-NH	100%	100%	100%	100%	100%	100%	100%	100%
Glu22-OE1...	2.84 (0.18)	2.84 (0.15)	2.90 (0.22)	2.88 (0.16)	2.88 (0.20)	2.86 (0.18)	2.80 (0.12)	2.82 (0.13)
Lys100-NZ	67%	99%	22%	91%	100%	100%	100%	100%
Glu22-OE2...	2.94 (0.17)	2.98 (0.19)	2.92 (0.16)	2.97 (0.18)			2.96 (0.18)	2.90 (0.14)
Ala18-NH	32%	62%	91%	86%			83%	98%
Glu22-OE2...	3.09 (0.22)	3.12 (0.19)	3.06 (0.19)	3.16 (0.19)			3.03 (0.18)	3.01 (0.17)
Leu19-NH	21%	65%	88%	60%			88%	96%
Glu22-OE1...		2.91 (0.17)	3.02 (0.20)	2.87 (0.14)	2.84 (0.14)	2.84 (0.12)	2.86 (0.14)	2.89 (0.15)
Gln27-NEH		97%	39%	87%	95%	93%	99%	98%

of Thr*51 and the Glu74 carboxyl group. The Glu74 carboxyl forms two new strong and stable hydrogen bonds with N2H1 and N3H of HP. This suggests that Glu74, which sits at the bottom of the active site and is well restrained by the hydrogen bonds with surrounding residues, acts as an anchor to fix the orientation of HP. The side chains of His*53 and Tyr*54

TABLE 4: Hydrogen Bonds between Asp*46, Asn71, and Leu72 in the HP Complex

	subunit =							
	1	2	3	4	5	6	7	8
Asp*46-OD1...	2.85 (0.14)	2.85 (0.14)	2.91 (0.16)	2.87 (0.15)	2.84 (0.14)	2.86 (0.14)	2.87 (0.15)	2.86 (0.14)
Asn71-ND2H1	96%	88%	43%	81%	99%	89%	96%	94%
Asp*46-OD1...	3.17 (0.18)	3.13 (0.18)	3.06 (0.18)	3.12 (0.18)	3.12 (0.18)	3.09 (0.18)	3.09 (0.18)	3.08 (0.18)
Asn71-NH	85%	91%	96%	92%	88%	93%	95%	95%
Asp*46-OD2...	2.92 (0.14)	2.98 (0.17)	2.94 (0.15)	2.91 (0.13)	2.94 (0.15)	2.97 (0.16)	2.93 (0.14)	2.98 (0.16)
Asn71-NH	100%	99%	100%	100%	100%	100%	99%	99%
Asp*46-OD2...	2.95 (0.15)	2.95 (0.19)	2.95 (0.15)	2.93 (0.14)	2.98 (0.16)	2.95 (0.14)	2.99 (0.16)	2.96 (0.15)
Leu72-NH	95%	99%	99%	100%	99%	99%	99%	99%

TABLE 5: Hydrogen Bonds with Glu24 and Ile25 in the HP Complex

	subunit =							
	1	2	3	4	5	6	7	8
Ser20-O...	3.05(0.19)	3.06(0.20)	3.00(0.19)	3.10(0.19)	3.00(0.17)	3.17(0.19)	3.06(0.19)	3.19(0.19)
Glu24-NH	88%	79%	86%	77%	96%	50%	84%	61%
Ala21-O...	3.17(0.18)	3.11(0.17)	3.20(0.18)	3.13(0.18)	3.22(0.18)	3.12(0.18)	3.11(0.18)	3.04(0.17)
Glu24-NH	65%	82%	60%	75%	56%	83%	85%	94%
Ala21-O...	3.17(0.18)	3.02(0.18)	3.00(0.17)	3.07(0.18)	3.24(0.16)	3.18(0.18)	3.07(0.19)	3.08(0.18)
Ile25-NH	57%	95%	83%	79%	46%	58%	84%	90%
Glu22-O...	3.22(0.17)			3.25(0.19)	3.32(0.14)	3.30(0.15)		
Ile25-NH	53%			11%	23%	15%		

sandwich the HP pterin ring forming a $\pi-\pi$ stacking interaction in some active sites, while in others the His53 imidazole ring is angled away. In the crystal structure, the Tyr*54 phenol is stacked with the pterin ring, but the His*53 imidazole is not. HP also forms hydrogen bonds with the carbonyl of Val*52 and amide of Tyr*54 from the adjacent subunit.

The O4 atom of HP forms a stable hydrogen bond with a water molecule that is hydrogen bonded to the carbonyl of Asn71 and the side chain amine of Lys100 (Table 2, Figure 9b), which is similar to the bound water interaction found in the HP crystal structure. Since all crystallographic water molecules were removed before the start of the MD simulation, the appearance and consistent presence of water in this position in all the active sites suggests that it is a stable point for water, and that the simulation time was sufficient to organize proper active sites. All of the above-noted interactions that do not exist in the free form enzyme contribute to make the complex more stable and rigid.

It is sensible that HP binding rigidifies the region around residue 70 (Figure 3) considering that HP forms two hydrogen bonds with Glu74, one with Leu73, and one hydrogen bond network with Asn71 through the trapped water. Since this region is quite close to residues around 45 from the adjacent subunit (Figure 3), these interactions may also contribute to rigidifying the region around residue 45*. Further investigation indicates that the amides of Asn71 and Leu72 form a hydrogen bond network with the carboxyl of Asp*46 from the adjacent subunit (Table 4), which might be the reason that residues around 45* are also rigidified.

The strong hydrogen bonds between HP and Glu22 certainly stabilize the region around residue 22. Interestingly, the largest rigidity change due to HP binding comes from residue 25, which does not interact with HP directly. Ile25 shows great flexibility in the free form simulation with RMSF equal to 1.7 Å that might occur because Ile25 is at the end of helix 1 (residues 20−24) and close to the beginning of β sheet 3 (residues 27−37). In the bound form, the Ile25 RMSF is only about 0.9 Å. Further inspection shows that the carbonyl of Ile25 forms one hydrogen bond with the amide of Ala21 and one weak hydrogen bond with the amide of Glu22, while Glu24 forms two hydrogen bonds with Ser20 and Ala21 (Table 5). Clearly, the rigidity increase around Glu22 also stabilizes Glu24 and Ile25.

D. Product Binding Energy Decomposition. The nature of the binding pocket can be defined by examining which residues contribute the most to stabilizing HP. As shown in Figure 9b, the direct interactions with HP typically include hydrogen bonds with the carboxyl groups of Glu22 and Glu74, the carbonyl oxygen of Val*52, and the amide nitrogens of Tyr*54 and Leu73. To assess how the individual residues contribute to the binding of HP, the interaction energies between HP and residues within 5 Å of HP, averaged over the 2 ns trajectory and the eight active sites, are shown in Figure 10. These residues contribute about −80 kcal/mol of electrostatic energy (97% of the total over all residues) and about −23 kcal/mol of van der Waals energy (95% of the total over all residues).

The highly conserved residues Glu22, Glu74, and Lys100 contribute about −68 kcal/mol of the electrostatic energy (82% of the total over all residues electrostatic energy) and clearly are essential to fix the orientation of HP within the active site. The electrostatic interaction (−35 kcal/mol) between Glu74 and HP is the largest among all individual residues due in part to the two strong hydrogen bonds between the carboxyl group and HP. Glu22 contributes −11 kcal/mol of electrostatic stabilization in a similar way. Though Lys100 does not form a direct hydrogen bond with HP, being water-mediated, the positively charged amine group of Lys100 does have a strong interaction with the partial negative charges of O4 and N5 of HP, which contributes the second largest electrostatic energy of −23 kcal/mol. His*53 and Tyr*54 contribute −9.3 kcal/mol (37%) van der Waals interaction due to the fact that the HP pterin ring is sandwiched by the side chains of these two residues in four of the active sites. In the other four active sites, the pterin ring is tilted slightly and moves closer to Lys100 and Glu22. Tyr54 is maintained in a similar position but His53 moves away. This might be caused by the strong interaction between HP and Lys100 and Glu22. The solvent−HP interaction provides about −14 kcal/mol electrostatic and −1 kcal/mol van der Waals stabilization energy. A major contributor (6−7 kcal/mol) to this energy is the hydrogen bond between the HP O3 atom and the water bound to Lys100.

2. Exit Path of HP from DHNA. *A. Pushing HP Out of the Active Site*: *Fluctuation Analysis.* To investigate how the HP release process changes the structure and flexibility of DHNA, HP was pushed out by 8 Å in all the active sites. The pushing

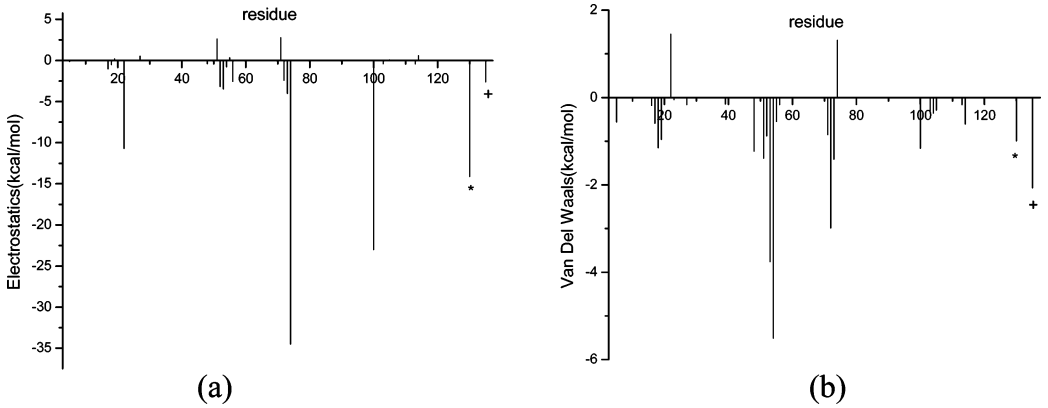

Figure 10. (a) Electrostatic and (b) van der Waals interaction energies between HP and all residues within 5 Å of HP. The energies are averages over the 2 ns simulation and the protomers.

TABLE 6: RMSF of Arg118 Backbone N Atom in the Regular and Pushing MD Simulations

	subunit =							
	1	2	3	4	5	6	7	8
regular MD (Å)	0.47	0.49	0.48	0.48	0.51	0.45	0.52	0.54
pushing MD (Å)	0.50	0.61	0.51	0.64	0.56	0.61	0.56	0.50

force was imposed between the backbone N of Arg118 and C10 of HP. The trajectory average RMSD of 1.56 Å is quite stable and slightly larger than the 1.42 Å of the regular HP complex simulation, indicating that no significant changes of the overall backbone structure occurred during the push. The average RMSD of the N of Arg118 is 1.29 Å during this pushing simulation compared with 1.10 Å in the regular HP simulation. As shown in Table 6, the RMSF of this N atom in the eight different subunits increases slightly overall in the pushing MD simulation, with the average increase ~0.07 Å. Both RMSD and RMSF measures show that the backbone N of Arg118 does not move much during the pushing simulation, which confirms that the restraint between this N atom and C10 of HP and the time scale used led to a sensible product release trajectory.

Figure 11 shows the residue RMSF comparison (based on superposition of individual subunits) between the pushing and regular MD simulations. The product release does not change the backbone protein structure and flexibility significantly, though slight changes occur around the flexible regions, residues 15−25, 45*−55*, 68−74, and 100−110, which may reflect the disturbance that the exit of HP creates in these regions. The total fluctuation of the CA atoms is 84.9 Å^2, larger than the 65.6 Å^2 of the regular HP complex simulation, but comparable with that of apo enzyme simulation's 86.5 Å^2. This is consistent with the previous simulation result that product binding rigidifies the protein; therefore, the release of HP should increase the enzyme flexibility.

To understand how protein structure and flexibility change along the HP release trajectory, the data were also analyzed based on individual time windows. The average protomer structure in each window (an average over all subunits over the simulation time window of 50 ps) was compared with the average protomer structure (average over all subunits over the 1.5 ns simulation time) in the regular HP complex simulation (Figure 12). The first window corresponds to the first 50 ps of the regular simulation. The CA RMSD changes from 0.2 Å to 0.4 Å during the first 2 Å of pushing, and then fluctuates around 0.4 Å. The initial window difference of 0.2 Å is a consequence of the 50 ps versus 1.5 ns averaging time. The 0.2 Å subsequent RMSD increase again illustrates the minimal overall protein disturbance, considering that the CA RMSD between the average

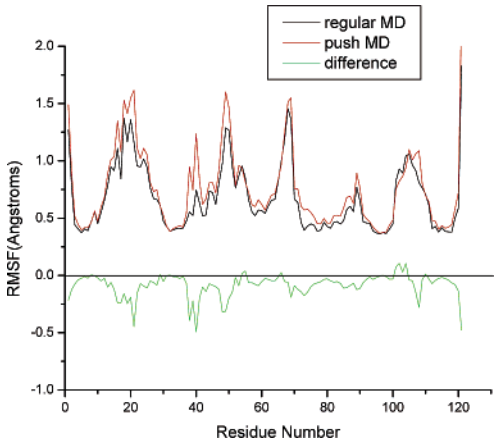

Figure 11. CA RMSF comparison between the push and regular simulations, indicating the small disturbance in the protein fluctuations from the exit of HP for the chosen force constant and pulling rate.

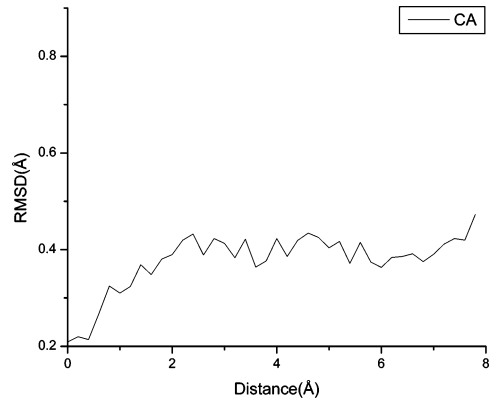

Figure 12. CA RMSD of the 50 ps averaged structure in each window relative to the average structure in the HP complex regular simulation.

structure of the regular HP complex and that of the apo form simulation is only 0.44 Å.

The CA fluctuation in each window is plotted along the exit path of HP in Figure 13. Unlike the average structure plot (Figure 12), the fluctuation increases from ~65 Å^2 to ~92 Å^2 along the HP exit path from 1 to 6 Å, and then stabilizes after 6 Å. Interestingly, the average fluctuation between 6 and 8 Å is slightly larger than that of the apo form of ~87 Å^2. This extra flexibility may be caused by the weak DHNA-HP interaction perturbing the energy surface of the protein and assisting it to access more conformation space. Note that the difference in fluctuation magnitude is probably not a sampling deficiency, though it is hard to get converged fluctuations in

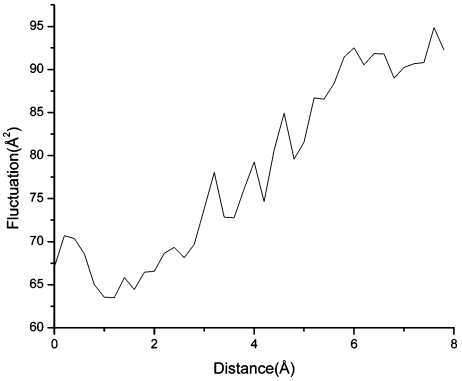

Figure 13. Total CA fluctuation of DHNA along the exit trajectory of HP showing a range of protein fluctuations that span the bound to free form results.

short time simulations, because the fluctuation in the first window of the HP push is 67.0 Å^2, which is consistent with the 65.6 Å^2 in the regular HP complex enzyme simulation.

The window RMSD calculations were also carried out on the CA of the four fragments that correspond to the flexible regions: 15−25, 45*−55*, 68−74, and 100−110, as displayed in Figure 14. The CA RMSD of residues 15−25 increases from 0.3 Å to 0.7 Å (Figure 14a) over the first 2 Å push, but then drops back to 0.3 Å around the distance 3.5 Å and thereafter fluctuates around 0.5 Å. This region may undergo a conformational change during the exit of HP associated with its strong interaction with Glu22. The CA RMSD of residues 45−55 (Figure 14b) increases from 0.2 Å to 0.5 Å, and then fluctuates before increasing to 0.8 Å around the distance 7−8 Å. This residue range includes Thr*54 that is π-stacked with HP. The CA RMSD of residues 68−74 (Figure 14c) increases from 0.2 Å to 0.5 Å, and then goes back to 0.3 Å at the distance 6 Å, thereafter fluctuating around 0.4 Å. The RMSD of CAs of residues 100−110 (Figure 14d) goes up to 0.5 Å, and then fluctuates.

While the RMSD data provide information about the drift of protein structure during the HP exit, the fluctuation data indicate how flexibility changes in these regions. Residues 15−25 show the largest CA fluctuation changes, from ∼10 to ∼22 Å^2 (Figure 15a), while residues 45*−55*, 68−74, and 100−110 show slight increases of flexibility along the path (Figure 15b,c,d). By summing the fluctuations in these four regions, as shown in Figure 15e, the overall fluctuation increases from ∼30 Å^2 to ∼50 Å^2, which contributes ∼75% of the increase in protein flexibility. Therefore, the major increase in flexibility comes from these more mobile regions around the active site.

B. Pushing HP Out of the Active Site: *Energetic Analysis.* The local nature of the HP-DHNA interactions suggests that an exit path should show a relatively fast transition between bound and weakly bound behavior. Figure 16a plots the electrostatic interactions between HP and Glu22, Glu74 and Lys100. The local nature of the electrostatic interactions between HP and the protein noted in Section 1D is consistent with the fairly well defined transition behavior in this plot. The magnitude of the electrostatic interaction decreases mainly due to the breaking of hydrogen bonds between Glu22, Glu74, and HP and the increase of distance between Lys100 and HP. Figure 16a also shows that the hydrogen bonds between HP and Glu22 and Glu74 break when HP is 2−3 Å away from its original

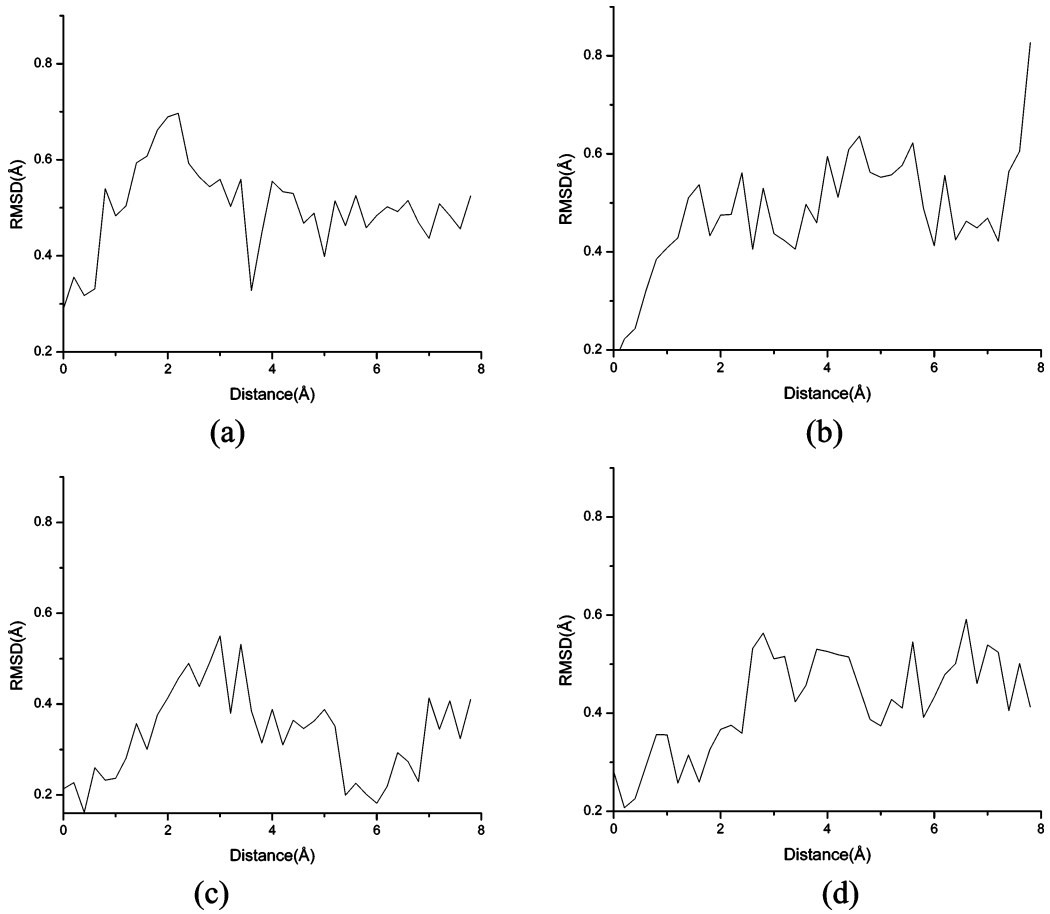

Figure 14. CA RMSD of the flexible regions of the 50 ps averaged structure in each window relative to the average fragments in the HP complex regular simulation: (a) residues 15−25; (b) residues 45*−55*; (c) residues 68−74; and (d) residues 100−110.

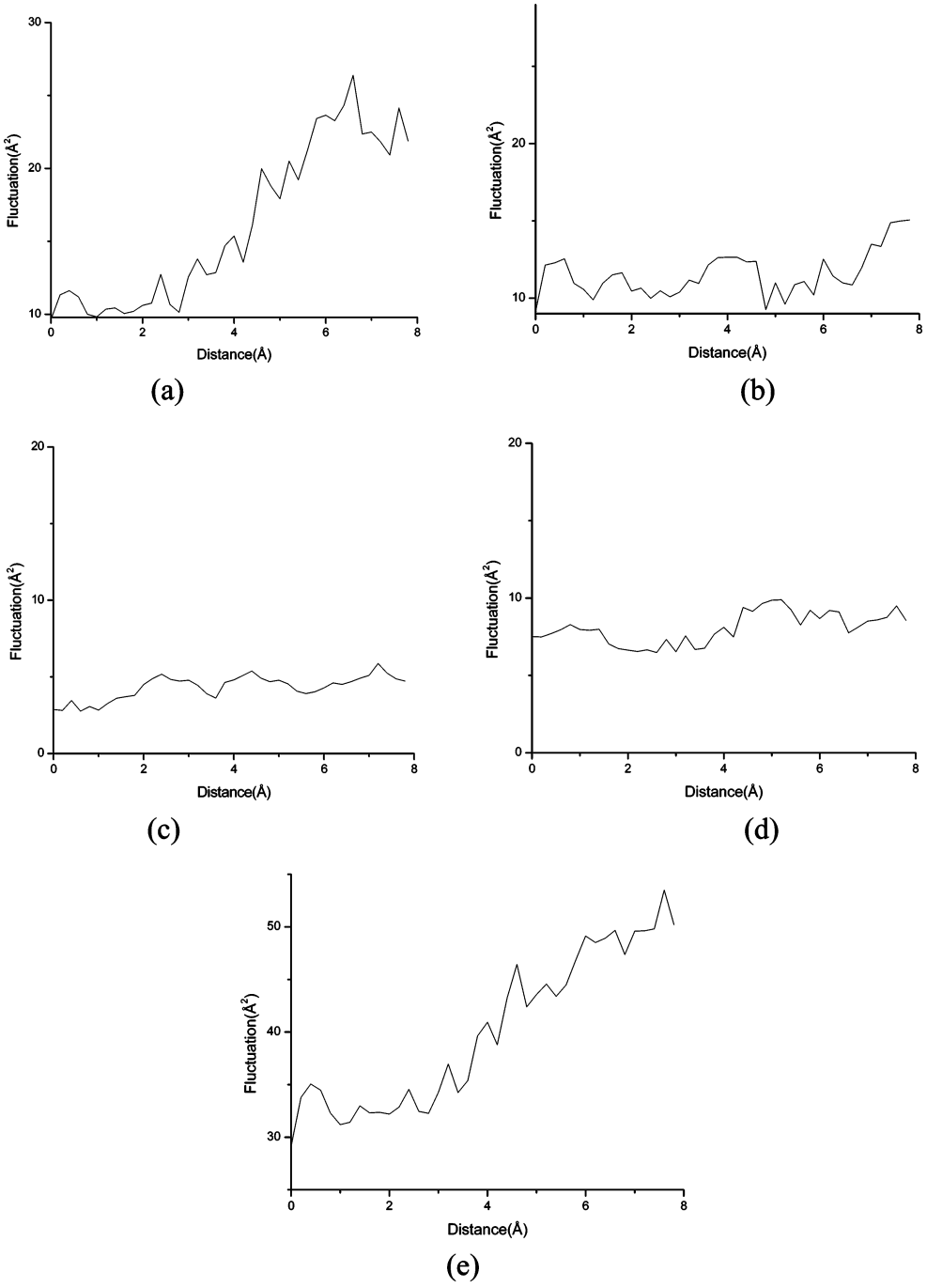

Figure 15. CA fluctuations of flexible regions along the exit path of HP: (a) residues 15−25; (b) residues 45*−55*; (c) residues 68−74; (d) residues 100−110; and (e) sum of (a), (b), (c), and (d).

bound position. Interestingly, the interaction between Glu22 and HP changes from −10 to 6 kcal/mol during the first 4 Å of HP exit, indicating that the motion of HP pushes this residue away. Then, this interaction energy decreases to −4 kcal/mol, which suggests that Glu22 favorably interacts with HP afterward. Figure 16b displays the interaction energies of the flexible regions with HP. The interaction between HP and residues 15−25 shows a pattern similar to that between HP and Glu22, with a lower maximum (−2 kcal/mol) at a distance of 4 Å, and strong interaction of −10 kcal/mol after 6 Å. This feature suggests that the interaction between HP and residues 15−25 can assist in breaking the hydrogen bond between HP and Glu22 and also that this region undergoes a conformational change to accommodate the release of HP. The trend in energy for residues 15−25 as a function of distance is consistent with the RMSD trend summarized by Figure 14a. The interaction energy between HP and residues 45*−55* fluctuates around −20 kcal/mol until the distance increases to 6 Å. The strong stabilizing van der Waals interaction as well as electrostatics between this region and HP (Figure 10) was maintained in the first 4 Å of pushing, with the residues moving with HP. Then, the van der Waals energy increases and the electrostatic energy drops to maintain the overall energy until about 6 Å. The residue regions 68−74 and 100−110 show similar interaction energy changes as those of Glu74 and Lys100, respectively. Along the exit path, the interaction between HP and DHNA weakens from −108 to −33 kcal/mol while the HP solvent interaction strengthens from −16 to −61 kcal/mol (mainly electrostatic energy), the latter compensating for the loss of interactions between HP and the enzyme (Figure 16b).

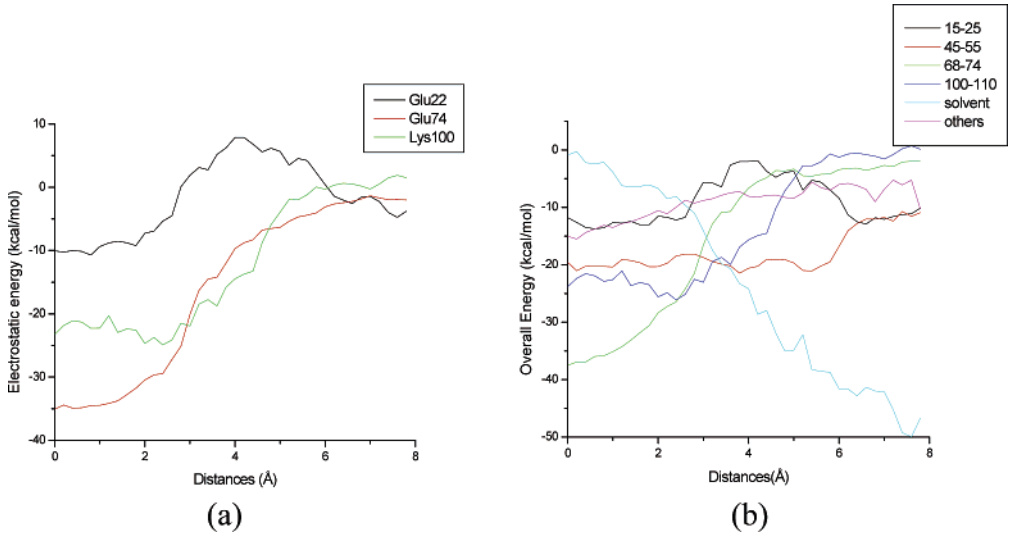

Figure 16. The interaction energies between HP and its surroundings over the exit pathway. (a) Electrostatic interactions between HP, and Glu22, Glu74 and Lys100. (The solvation energy of the starting point (−14.5 kcal/mol) was set to zero, to give a better view for this figure.) (b) Overall interactions between HP and residues 15−25, 45*−55*, 68−74, and 100−110, solvent, and the other parts of the protein.

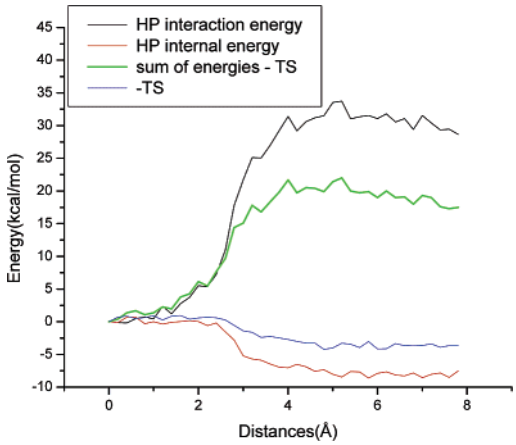

Figure 17. The intramolecular energetic and entropic contributions of HP, and its interaction energy with the environment and solvent, along the exit pathway.

The total interaction energy between HP and its environment along the product release path is plotted in Figure 17, which would imply a barrier to HP release of ∼32 kcal/mol. That is a rather large barrier to product release; however, it does not account for the internal energy of HP. If the internal energy of HP along the path is also included (averaging the HP internal energy over the 50 ps windows) the barrier drops by ∼7 kcal/mol and lowers the ligand release barrier to ∼25 kcal/mol. The decrease of internal energy of HP helps to compensate for the loss of interaction between HP and the enzyme relative to its solvation energy. Note that the reaction coordinate is defined by the N of Arg118 to C10 of HP distance, so that 0 Å corresponds to 17.5 Å. When HP is pushed out by restraining this distance, the Glu74 to HP distance, for example, does not necessarily immediately respond. Furthermore, various orientations are being sampled at each distance. That the energetic change does not occur immediately (between 0 and 2 Å) can be attributed to such effects. Most of the decrease in internal energy occurs between 2 and 4 Å and is a combination of the loss of interaction energy of HP with the protein and the release of strain energy of the bound product.

C. Pushing HP Out of the Active Site: *Free Energy Analysis.* A binding constant is of course related to the free energy difference between bound and solvated states, and a kinetic constant for dissociation is related to the free energy barrier separating bound and unbound states. Above, we focused on the energy of pushing HP out of the binding pocket. In principle, the exit pathway strategy pursued here could be used to obtain directly the free energy along the chosen coordinate. However, that would require an averaging period for each of the 40 windows that cannot be practically reached in current MD simulations. An approximate free energy profile may be obtained by including the entropy associated with HP along the exit path and the corresponding protein entropy. The HP entropy may be estimated by a quasiharmonic analysis since, with the exception of the ring NH_2 and the CH_2OH tail, it is a quite rigid structure. The line in Figure 17 labeled as -TS is this vibrational contribution to the HP entropy. The ΔS of ∼13 cal/(mol·deg) obtained in Section 1B for DHNA between bound and apo forms provides a less than 4 kcal/mol decrease in free energy for unbinding. Including these two entropic contributions leads to a barrier to HP release of ∼15 kcal/mol around 4 Å.

IV. Conclusions

The eight catalytic active sites of DHNA are formed by the noncovalent association of protomers. The presence of eight active sites permits a more efficient investigation by molecular dynamics of the binding site, and of how HP is released from DHNA. The reliability of the simulations was assessed based on RMSD from the crystal structures, the active site hydrogen bonding patterns, and the conformations of the active sites compared with those of the crystal structures. The RMSD and RMSF data confirm that the simulations maintain the basic DHNA conformation. The overall DHNA structure is quite rigid, and the binding of HP does not cause significant changes in the structure.

By analyzing the trajectory data with the two RMSF methods denoted as average protomer RMSF and octamer RMSF, the more flexible regions of DHNA were identified. In particular, the four regions spanning residues 15−25, 45*−55*, 68−74, and 100−110 are the more flexible regions of apo DHNA. These regions include residues that are part of the active site. The binding of HP rigidifies DHNA, not only in the above flexible regions but also in the more rigid regions that do not directly participate in HP binding. All these conclusions relied on a comparison of the octamer and average protomer method

changes in residue RMSFs for the apo and HP-bound forms; it would not have been evident from just the average protomer-based method. The network of hydrogen bonds formed by HP and the protein, and also induced in the protein by HP, are schematized in Figure 9b, and specified, for each dimer, in Tables 2−5. These interactions between HP and active site residues might explain the increase in rigidity of the enzyme when HP is bound.

The PCA results of apo and HP-bound DHNA complements those of the RMSF analyses. The rigidification, on average, of DHNA upon HP binding is evident in the total PCA variance of 86.5 Å^2 for the apo and 65.6 Å^2 for the bound form. The data in Figure 6 show that the first four modes not only capture a significant fraction of the total fluctuation in both apo and bound form, but most of the effect of rigidification upon HP binding is captured in those four modes, suggesting that a limited number of residues are responsible for the change in DHNA flexibility upon HP binding. That picture is reinforced by the projections of the atom motions onto the first few principal components, as shown for the first mode in Figure 7, and leads to the identification of residues in the three regions 45*−55*, 68−74, and 100−110, which can provide a contrast between the fluctuations in the apo and HP-bound forms, as displayed in Figure 8.

An energetic analysis of the residues that contribute to binding HP reveals that the binding is localized, as summarized by the data in Figure 10. It is remarkable that the HP binding energy is so local, even though it is dominated by the nominally long-range electrostatic interaction energy contribution. Residues within 5 Å of HP contribute about 97% of the total electrostatic and 95% of the total van der Waals energy. Just three residues, Glu22, Glu74, and Lys100, contribute most (82%) of the electrostatic energy, which emphasizes the importance of these residues for binding HP and for creating a catalytically competent structure. Glu74, which has the largest interaction with HP (−35 kcal/mol), sits at the bottom of the active site, and may act as an electrostatic attractor for HP. Both Glu22 and Lys100 are implicated in the catalytic mechanism of HP. Note that Lys100 is not directly hydrogen bonded to HP; but there is one water molecule that bridges Lys100 and O4 of HP, which is also hydrogen bonded to Asn71. This water is found in the crystal structure. In the simulation, which is carried out in the absence of crystallographic waters, a water molecule from the solvent migrates to this position and persists in all eight active sites (Table 2). This is about half the total HP water interaction energy, which indicates the importance of this water for binding HP. The hydrogen bond between Lys100 and the trapped water may also serve to decrease the pK_a of the water molecule as part of the catalytic mechanism of DHNA. The electrostatic interaction energy between Lys100 and HP of −23 kcal/mol may be important for binding the pterin ring, versus modifying the N5 pK_a, as noted in a Raman spectroscopic study that found N5 to be deprotonated even at pH 6.5.[15]

The flexibility of DHNA is 87 Å^2 in the apo and 65 Å^2 in the HP-bound form, providing a measure of how HP binding increases the rigidity of DHNA, which complements the view obtained from specific interactions as detailed in this work. It would be possible, of course, for HP to rigidify DHNA locally but still lead to an overall increase in DHNA flexibility. The increase in rigidity that is found is qualitatively related to an entropic loss upon binding that needs to be compensated by a sufficient increase in DHNA−HP interaction energy. It is interesting that during the course of the HP exit path simulation, DHNA increases its flexibility to 92 Å^2, which is even larger than the free form value. This "extra" flexibility may be due to the surface-bound HP interacting with the protein and helping it sample more conformation space. Much (75%) of the flexibility increase comes from the four flexible regions around the active site that have been identified in the simulation. Among them, residues 15−25 contribute the most, about 40%, to the increase. Since the binding of HP is so local, it is sensible that the flexibility decrease upon binding should also be concentrated in this active site region.

The direction and speed of pushing that we chose for the HP exit path did not significantly disturb the overall DHNA structure, as summarized in Figures 11 and 12. As noted above, HP is bound strongly by Glu74 at the bottom of the binding site, and examination of HP relative to its surroundings along the exit pathway shows that it moves out with the pterin ring first parallel with its bound position followed by a gradual upward rotation as it exits. That the active site is formed by the noncovalent association of two protomers indicates that it is mainly HP-residue hydrogen bonds that are being broken in the release process. The HP−DHNA interaction energies, as displayed in Figures 16 and 17, show a fairly abrupt transition region around the 2−4 Å range, which, again, is consistent with the local binding energetics (Figure 10) mainly involving Glu22, Glu74, and Lys100. It appears (Figure 16a) that Lys100 must be pushed aside as HP exits the active site. The local binding energetics of HP may be important to the release mechanism because it lets HP easily leave once those local interactions are broken.

The large decrease in HP−DHNA stabilization energy during the exit of HP needs to be compensated, if the barrier to HP release is not rendered too large. Partly, this comes from the increase of the HP−solvent interaction, as shown in Figure 16b. At the same time, the internal energy of HP also decreases (Figure 17), which suggests that HP relaxes a conformationally based strain during the release. Thus, HP binding destabilizes itself, which may be viewed as a tradeoff associated with positioning HP in the active site to permit catalytic activity.

There are, of course, entropic terms that contribute to the free energy barrier for HP release. The contribution associated with the HP internal fluctuations during the exit path as obtained from a quasiharmonic analysis is shown in Figure 17, and clearly lowers the barrier to and contributes a stabilizing influence on HP release. In view of the enhanced freedom of HP when it is in the solvent, relative to its bound state, an entropic penalty of binding is expected. The change in protein entropy that is obtained from the PCA eigenvalues (based on the CA atoms) provides less than 4 kcal/mol of free energy protein stabilization in favor of the apo protein. The change in protein entropy found most likely is an upper bound because the initial state is all HP bound and the final state is all HP free. If fewer than eight ligands were released, the change in protein entropy should be reduced. The experimental data[23] indicate that the rate constant for product release is ∼10 s^{-1}, which is consistent with the barrier to HP release obtained here.

Acknowledgment. This work was supported by the Quantitative Biology Modeling Initiative of Michigan State University, the Michigan Center for Biological Information of the Michigan Life Sciences Corridor, and NIH grants (GM58221 to H.Y. and GM47274 to R.I.C.).

References and Notes

(1) Illarionova, V.; Eisenreich, W.; Fischer, M.; Haussmann, C.; Romisch, W.; Richter, G.; Bacher, A. *J. Biol. Chem.* **2002**, *277*, 28841.

(2) Henderson, G. B.; Huennekens, F. M. *Methods Enzymol.* **1986**, *122*, 260.

(3) Bermingham, A.; Derrick, J. P. *Bioessays* **2002**, *24*, 637.

(4) Sanders, W. J.; Nienaber, V. L.; Lerner, C. G.; McCall, J. O.; Merrick, S. M.; Swanson, S. J.; Harlan, J. E.; Stoll, V. S.; Stamper, G. F.; Betz, S. F.; Condroski, K. R.; Meadows, R. P.; Severin, J. M.; Walter, K. A.; Magdalinos, P.; Jakob, C. G.; Wagner, R.; Beutel, B. A. *J. Med. Chem.* **2004**, *47*, 1709.

(5) Hennig, M.; Dale, G. E.; D'Arcy, A.; Danel, F.; Fischer, S.; Gray, C. P.; Jolidon, S.; Muller, F.; Page, M. G. P.; Pattison, P.; Oefner, C. *J. Mol. Biol.* **1999**, *287*, 211.

(6) Bauer, S.; Schott, A. K.; Illarionova, V.; Bacher, A.; Huber, R.; Fischer, M. *J. Mol. Biol.* **2004**, *339*, 967.

(7) Salzmann, M.; Pervushin, K.; Wider, G.; Senn, H.; Wuthrich, K. *J. Am. Chem. Soc.* **2000**, *122*, 7543.

(8) Lopez, P.; Lacks, S. A. *J. Bacteriol.* **1993**, *175*, 2214.

(9) Cox, T. F.; Cox, M. A. A. *Multidimensional scaling*, 2nd ed.; Chapman & Hall: Boca Raton, FL, 2001.

(10) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins*: *Struct., Funct., Genet.* **1993**, *17*, 412.

(11) Kollman, P. *Chem. Rev.* **1993**, *93*, 2395.

(12) Gilson, M. K.; Given, J. A.; Bush, B. L.; McCammon, J. A. *Biophys. J.* **1997**, *72*, 1047.

(13) Boresch, S.; Archontis, G.; Karplus, M. *Proteins* **1994**, *20*, 25.

(14) Boresch, S.; Tettinger, F.; Leitgeb, M.; Karplus, M. *J. Phys. Chem. B* **2003**, *107*, 9535.

(15) Deng, H.; Callender, R.; Dale, G. E. *J. Biol. Chem.* **2000**, *275*, 30139.

(16) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E.; Debolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1.

(17) Berendsen, H. H. C.; Postma, J. P. M.; Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.

(18) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, G. L. *J. Chem. Phys.* **1995**, *103*, 8577.

(19) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.

(20) Lindahl, E.; Hess, B.; van der Spoel, D. *J. Mol. Model.* **2001**, *7*, 306.

(21) Levy, R. M.; Karplus, M.; Kushick, J.; Perahia, D. *Macromolecules* **1984**, *17*, 1370.

(22) Haussmann, C.; Rohdich, F.; Schmidt, E.; Bacher, A.; Richter, F. *J. Biol. Chem.* **1998**, *273*, 17418.

(23) Wang, Y.; Li, Y.; Wu, Y.; Yan, H. Unpublished.