Insight into Catalytically Relevant Correlated Motions in Human Purine Nucleoside Phosphorylase†

Sara Núñez,‡,§ Corin Wing,‡ Dimitri Antoniou,§ Vern L. Schramm,‡ and Steven D. Schwartz*,‡,§

Department of Biochemistry and Department of Biophysics, Albert Einstein College of Medicine, Bronx, New York 10461

Received: March 11, 2005; In Final Form: May 26, 2005

The catalytic site of the homotrimeric enzyme human purine nucleoside phosphorylase enzyme (hPNP) features residue F200 and the 241−265 loop directly skirting the purine base and a residue belonging to the adjacent monomer, F159, immediately conterminous to the ribosyl moiety. Crystallographic B-factors of apo human purine nucleoside phosphorylase, and hPNP complexed with substrate/transition state (TS) analogues, show that residue E250 is the centroid of a highly mobile loop region. Furthermore, superimposition of apo hPNP and hPNP complexed with TS analogue Immucillin-H shows a tightening of the active site, caused by the ligand-dependent 241−265 loop rearrangement taking place upon substrate/inhibitor binding, suggesting a putative dynamic role of the loop in binding/catalysis. However, crystallographic structures reveal only average atomic positions, and more detailed information is needed to discern the dynamic behavior of hPNP. The Essential Dynamics (ED) method is used here to investigate the existence of *correlated motions* in hPNP and consequently proposes mutagenesis assays to estimate the relative importance of these motions in the phosphorolytic efficiency of the reaction catalyzed by hPNP. We compare the concerted motions obtained from multiple molecular dynamics simulations of apo and Michaelis complex of hPNP both in vacuo and in solution. The results of the principal component analysis for the apo hPNP indicate the existence of strong correlations predominantly in the vicinity of residue F159. However, for the Michaelis complex, concerted motions are seen mostly around both active site residue F200 and loop residue E250. Additionally, for a simulation depicting the relaxation of tight complexed hPNP with a TS analogue, toward its relaxed apo form (after removal of the TS analog), a combination of the apo hPNP and Michaelis complex motions is found, with prominent concerted modes centered around neighboring residues F159, F200, and E250. Finally, we probed the extent to which these concerted motions bear an intrinsic catalytic role by performing experimental site-directed mutagenesis on some residues, followed by kinetic analysis. The F159G and F200G mutants displayed a strong increase in K_M and modest decrease in k_{cat}, suggesting that these concerted motions may provide dynamical roles in substrate binding and/or catalysis. However, further structural data for the hPNP mutants are needed to confirm our hypothesis.

1. Introduction

The process by which enzymes catalyze chemical reactions is still the subject of numerous experimental and theoretical studies. It has been proposed that internal correlated motions of proteins may assist catalysis and protein folding[1−4] or be crucial for substrate specificity.[5−7] Such functional internal motions are concerted motions of, typically, a few amino acid residues. However, it is still a challenge to identify such motions and their biological role from the molecular structure and interactions alone. In other words, crystallographic structures alone may not be enough to infer dynamic behavior in proteins, and additional computational techniques, such as all-atom molecular dynamics (MD) simulations, are needed to understand the dynamics of proteins.

Molecular simulations of macromolecules are good examples of high-dimensional systems where principal component analysis has been successfully employed in the past decade for identifying biologically important collective motions. This method identifies collective motions when they are present in the system, and in our case, we use theory to guide experiment. To our knowledge, the present paper is the first attempt to employ the essential dynamics method to dissect *catalytically relevant* concerted motions. The question addressed in this paper is whether putative correlated motions in hPNP have a participative role in the phosphorolytic reaction, either by creating substrate binding affinity or by accelerating the chemical step. We principally focus on the dynamical aspects of the apo and complexed protein (i.e., Michaelis complex, E·S). In particular, the conformational change in the 241−265 loop is monitored to investigate variations in its orientation, which is crucial in determining the substrate accessibility to the active site.

The outline of this paper is as follows. In the next section, we review the essential dynamics method. Then, we present several molecular dynamics simulations of the apo and E·S forms of hPNP, and we use the essential dynamics technique to probe correlated motions in hPNP. Finally, we report our results from experimental site-directed mutagenesis of a number of residues which our computational study found to have a contribution in correlated motions, to assess their catalytic participation.

† Part of the special issue "Donald G. Truhlar Festschrift".
* Corresponding author. E-mail: sschwartz@aecom.yu.edu
‡ Department of Biochemistry.
§ Department of Biophysics.

10.1021/jp051277u CCC: $33.50
Published on Web 06/29/2005

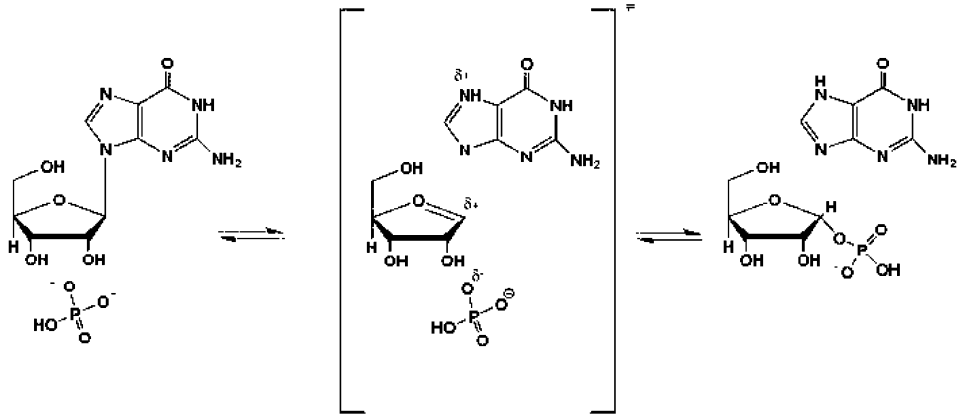

Figure 1. hPNP-catalyzed phosphorolysis of the purine nucleoside guanosine.

2. Essential Dynamics Method

Essential Dynamics (ED), also known as principal component analysis (PCA), is a method commonly used for dissecting the dynamics of proteins and their importance in biological processes, like protein folding or substrate binding. The ED method was first invented in a study of the dynamics of chaotic systems.[8] We briefly summarize the idea. Let $\vec{R} = (R_{1,x}, R_{1,y}, ..., R_{N,z})$ be a $3N$-dimensional position vector that specifies the state of a system consisting of N particles. One assumes that the solution of Newton's equation of motion can be written in the following approximate (Karhunen−Loeve) form:

$$\vec{R}(t) \simeq \sum_{m=1}^{3N} a_m(t)\vec{\eta}_m \tag{1}$$

Subsequently, one tries to find the coefficients and basis functions $a_m, \vec{\eta}_m$ that provide the best approximation (in a least-squares sense) of $\vec{R}(t)$ to the exact solution of Newton's equation of motion. Sirovich proved that this approximation is satisfied by choosing the spatial basis set $\vec{\eta}_m$ to be the eigenfunctions of the covariance matrix

$$C_{ij} = \langle (R_i - \langle R_i \rangle)(R_j - \langle R_j \rangle) \rangle \tag{2}$$

where $\langle \cdots \rangle$ denotes time average over the entire MD trajectory. When a few eigenvalues of the covariance matrix are much bigger than the rest, the corresponding eigenvectors dominate the dynamics. When sorted by their eigenvalues, the ED modes are sorted according to their contribution to the total mean-square fluctuation. These few dominant modes are designed to provide a good *fitting* to the trajectory, and they do not have any more physical meaning than, for example, spline functions that provide a good fitting to a continuous function. Therefore, one should view these dominant ED modes as a way of visualizing a multidimensional trajectory in a low-dimensional subspace. Many uses of ED in analyzing MD simulations are of this type, i.e., of projecting the complicated dynamics to a space spanned by a few principal components (PCs), so that the progress of the trajectory can be easily monitored.[9]

In the early 1990s, the ED method was rediscovered in the chemistry community,[10,11] but with a twist. While earlier ED studies focused on hydrodynamic problems, the application to proteins allowed a new idea to emerge. In a protein, unlike a liquid, besides random thermal motions of atoms, there are concerted motions of groups of atoms (e.g., loop motions) which provide a significant number of covariant matrix elements C_{ij} which are larger than matrix elements that correspond to motions of pairs of atoms that execute random thermal motions, since thermal motions give rise to a narrow distribution around the mean value of the atom positions. Therefore, in cases when there exist large-scale concerted motions in the protein, we can expect these collective motions to appear in the dominant eigenvectors of the covariant matrix. This use of ED is not just a tool of projecting motion to a low-dimensional space but is also a tool for separating the conformational space into an *essential subspace* containing only a few dominant collective modes and a remaining space which contains random atomic fluctuations. In this case, one can simply look at the eigenvectors of the covariance matrix to find which atoms participate in each of these dominant collective modes.[12−17] One should keep in mind that essential dynamics analysis cannot identify collective motions whose relaxation times are longer than the length of the simulation.[18] The ultimate goal of the use of the ED method in this paper is to relate collective protein motions to function.

3. Ligand-Induced Loop Conformational Change in Human PNP

Human purine nucleoside phosphorylase (hPNP) is the only purine salvage enzyme that specifically catalyzes reversible phosphorolysis of 6-oxopurine nucleosides to form α-D-ribose 1-phosphate (Figure 1).[19−24] The enzyme activity is associated with T-cell function, and there is currently an extensive development program for hPNP inhibitors that could moderate responses in T-cell associated disorders such as T-cell leukemia, lymphoma, and rheumatoid arthritis, among others.[25−27] Crystal structures of hPNP complexes have identified the active site residues in contact with the purine nucleoside and the phosphate nucleophile (Figure 2). The direct contacts in the purine binding site indicate that the side chains of E201 and N243 form hydrogen bond contacts with the purine base. Additionally, A116 and F200 contact the two faces of the purine base, with the F200 side-chain phenyl group making a classic herringbone-type interaction with the base. The contacts in the ribose binding site are residues Y88, F159, M219, and H257. In the phosphate binding site, there are four contact residues: S33, R84, H86, and E89, which hold the nucleophile tightly in the active site pocket.

Thus far, the dynamic aspects of correlated motions in hPNP have been inferred from indirect evidence drawn from structural information of X-ray crystal structures. If one assumes that each crystal structure represents a possible conformation in solution, considering several crystals and their relative atomic B-factors, provides a valid experimental view of the conformational flexibility of the protein at an atomic resolution. Figure 3 shows the C_α relative B-factors of apo hPNP and hPNP complexed

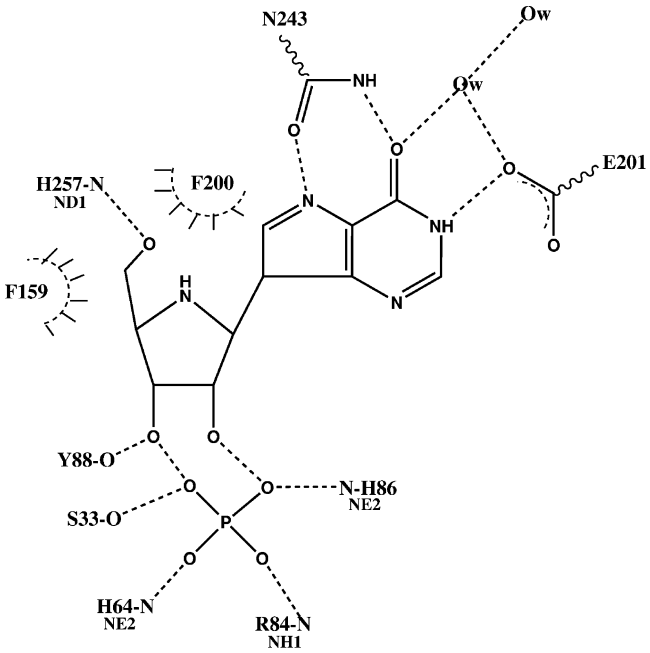

Figure 2. Active site of hPNP with the transition state (TS) analogue ImmH and the phosphate nucleophile.

with the transition state analogue Immucillin-H (ImmH). We can discern that the loop residue E250 is the centroid of a highly mobile region, while residues G63 and E183 are centroids of more localized mobile regions. Note that the relative B-factors stem from different crystal resolutions (2.3 and 2.7 Å for the apo and hPNP·ImmH crystals, respectively); thus a direct comparison between the pair is not to be made. Additionally, from the structural differences between the crystal pairs of apo hPNP-hPNP·guanine (formation of the Michaelis complex) and apo hPNP-hPNP·ImmH (formation of the TS analogue complex), it has been recently questioned whether correlated motions exist to allow the necessary conformational change of the 241−265 loop for substrate binding and turnover. It is accepted that reorganization of this flexible loop, which makes up part of the active site, is essential for allowing different substrates/inhibitors to enter the active site.[28] The loop displacement is proposed to push the substrates toward the active center favoring more reactive configurations. Figure 5 shows the root-mean-square deviation of the C_α geometric difference between the apo hPNP-hPNP·guanine and the apo hPNP-hPNP·ImmH pairs. The same geometrical difference trend follows for any combination of two apo hPNP structures (PDB ID: 1ULA, 1M73) with six substrate/TS analogues (PDB ID: 1PF7, 1PWY, 1V3Q, 1RCT, 1ULB, 1V2H), confirming the substantial loop rearrangement upon substrate analog/transition state analogue binding (data not shown). Figure 4 shows a monomer of hPNP with the substrate guanosine and the loop shown in green.

The conformational change in the loop results in an extension at the N-terminal of the C-terminal helix by about 8 residues (257−265). Thus, in the apo form of hPNP, we observe a much looser, less compact protein. However, upon binding of the substrate/inhibitor, an ordering of the protein active site occurs, which results in the substrate/inhibitor being firmly accommodated into the active site. The radius of gyration is 18.31, 18.29, and 17.94 Å for apo hPNP, hPNP·guanine, and hPNP·ImmH crystals, respectively. Figure 6 shows the backbone superimposition of apo hPNP and complexed hPNP·ImmH, where the "tightening" of the active site, caused by the ligand-dependent 241−265 loop rearrangement taking place upon substrate/inhibitor binding, can be seen. A plausible interpretation could be that this movement is a random uncorrelated thermal motion. Distinguishing between these two possibilities (correlated motions or uncorrelated thermal motions) is the main objective of this study. We use principal component analysis to identify correlated motions in different forms of hPNP, namely, its apo and complexed forms, and assess whether these facilitate the 241−265 loop rearrangement prior to the subsequent phosphorolysis reaction. We compare the principal components (PCs) for the apo and complexed hPNP simulations and scrutinize the different correlated motions for each form of the enzyme, comparing directly to the crystallographic B-factors. Finally, via experimental site-directed mutagenesis, several residues implicated in the correlated motion were mutated, and the kinetic constants k_{cat} and K_M (fingerprints of catalytic efficiency) were measured to weigh the impact of these residues in the phosphorolytic efficiency.

4. Computational Methodology

4a. Model Building. The structural starting point was the 2.5 Å resolution structure of hPNP·Immucillin-H (Figure 2).[29] To make this modeling study computationally feasible, only one monomer of homotrimeric hPNP was used in our calculations, given that the three subunits are identical. We modeled the natural substrate guanosine from ImmH. The protonation state of all residues was assigned for the E·S complex with the BIOPOLYMER module in the commercial package INSIGHTII (Accelrys, San Diego, CA), according to the pK_a of their titratable groups, assuming a pH of 7. The model retained the 25 crystallographic water molecules.

To relax any initial crystallographic strain and to allow the appropriate small geometric changes after the mutation of ImmH into the natural substrate guanosine, we performed minimizations of the three models (APO, ES, and EI-APO, see below), applying a weak restraining harmonic potential of 5 kcal mol^{-1} Å^2 on the backbone atoms in order to maintain them in their crystallographic positions. We used the CHARMM22 software package,[30] with a purely classical potential for all atoms. For the nonbonding energy term, we calculated the energy contributions for pairs of atoms within a cutoff of 12 Å, and we used a smoothing function to gradually reduce to zero this contribution between 12 and 14 Å. A constant dielectric $\epsilon = 1$ was used. This structure was the starting model for the E·S complex simulation (hereafter ES simulation) with the natural substrates guanosine and phosphate. For the simulation of apo hPNP (hereafter APO), we used the crystal structure PDB ID:1M73. For the simulation depicting the relaxation of *tight* complexed hPNP with a TS analogue, toward its *relaxed* apo form (hereafter EI-APO), we started with the geometry of hPNP·ImmH, and subsequently ImmH was extracted from the structure. The minimization procedure followed the same protocol as that for the ES trajectory, yielding the final structure for the subsequent EI-APO the molecular dynamics simulations. As mentioned previously, conformational changes that occur in a slower time scale than the time span of the MD simulations will not appear in the principal component analysis.

4b. Molecular Dynamics Simulations. Current simulation techniques face the challenge of poor conformational sampling. This is of critical relevance when one is interested in computing thermodynamics properties of a macromolecule. This disadvantage arises for any system with a rugged potential energy surface, where conformational barriers exist and may not be crossed with sufficient frequency to obtain an equilibrium distribution in reasonable computation times. Even in lengthy μsecond MD simulations of small proteins, the sampled phase

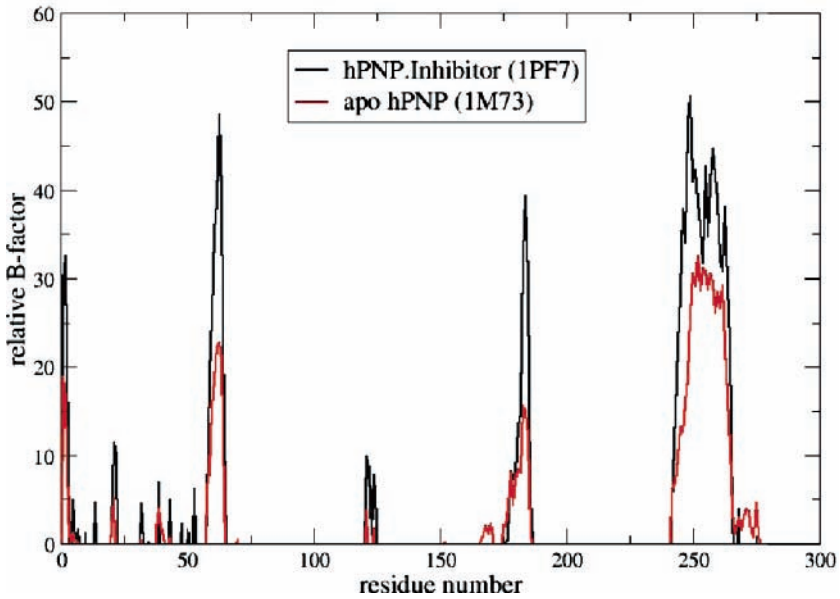

Figure 3. B-factors of the apo hPNP and hPNP complexed with the TS inhibitor Immucillin-H (ImmH). Note: the crystal resolution for the pair of crystals differ; thus a direct comparison of the relative B-factors is not to be made.

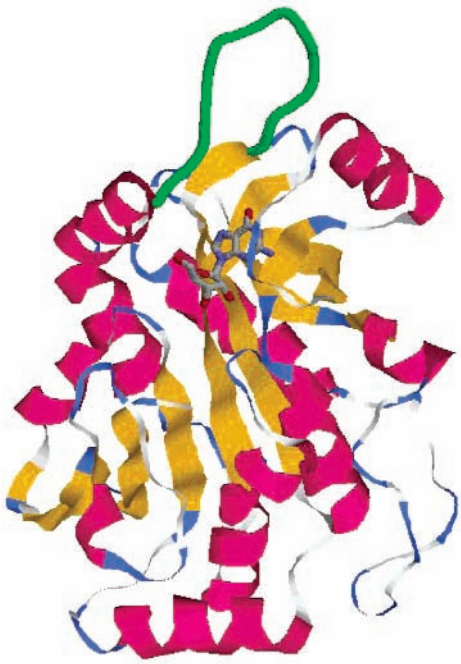

Figure 4. Cartoon representation of a monomer of hPNP with the substrate guanosine and the loop shown in green.

space represents a slim fraction of the available conformational space.[31] Along these lines, collective fluctuations derived from MD simulations can be expected to be affected by limited sampling, imperfections in the interatomic interactions, or the force field used.[32] For this reason, we follow the strategy of Karplus et al.,[9] and we run multiple short simulations of each model for a more efficient description of the existing correlated motions. Subsequently, we monitor the dependence of the concerted motions on the initial conditions and the conformational phase space visited. Thus, three different MD simulations are conducted for the APO and ES models. Additionally, we run triplicated EI-APO simulations, which start with an hPNP geometry of the E·I complex, however stripped of the TS inhibitor. The objective of the EI-APO simulation is to capture putative concerted motions in hPNP during the loop conformational change taking place upon going from an "E·I geometry" toward an "apo hPNP geometry", which reversibly translates into existing motions in hPNP during substrate binding and TS formation. The CHARMM22 all-atom parameter set was used for the standard amino acid residues for the MD simulations performed on the Michaelis complex of hPNP with guanosine and phosphate. MM point charges and parameters based on the CHARMM22 standard values were used for the substrates guanosine and phosphate. The TIP3P model represented the crystallographic water molecules.[33] The SHAKE algorithm, with a tolerance of 10^{-8}, was applied to constrain all bonds involving hydrogen atoms.[34] The time step for the integration was 1 fs, and we sampled the trajectory every 5 fs. The temperature was allowed to increase over the first 10 ps from 0 to 300 K, and this temperature was maintained for the rest of the simulation. Initial velocities were assigned based on a Gaussian distribution. After the heating phase, we allowed the system to equilibrate for 20 ps at 300 K. We used the same nonbonding interactions as for the minimization. A constant dielectric $\epsilon = 1$ was used. Finally, each of the runs are extended for 200 ps, for which only the equilibrated part of the simulation is saved, during which we collected the coordinates every 5 fs for subsequent analysis. For the multiple trajectories of each model, the initial atomic coordinates are identical, but the initial atomic velocity assignment differs. By performing multiple trajectories, we are minimizing the bias inherent in the initial conditions used, which is often difficult to overcome in standard molecular dynamics simulations. Additionally, 1 ns simulations were performed for each model in order to certify that 200 ps is an adequate simulation time to capture the types of concerted motions under investigation, i.e., loop closure in hPNP. Finally, due to the significant increase of CPU cost per dynamics step in explicit solvent, a single solvated simulation under cubic periodic boundary conditions for 200 ps was performed for the APO model in order to verify that existing motions in the in vacuo systems correlate with the motions found in the solvated complexes. Though the majority of the simulations are performed in a vacuum, which inherently enhances flexibility of hydrophobic groups since there is no solvent damping their motions, we expect the analysis of the sampling to provide information that can be carried over to solvated simulations.

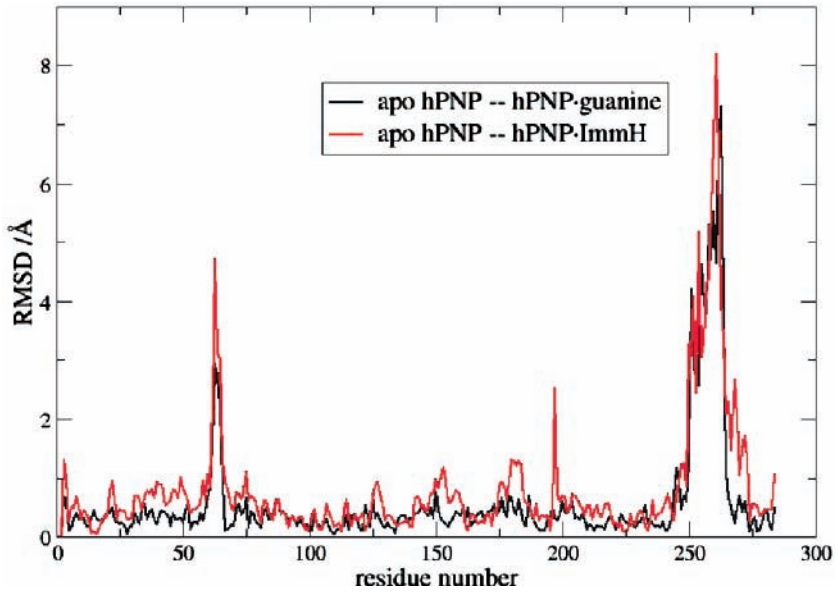

Figure 5. Detailed root-mean-square displacement (RMSD)/Å per residue of the geometric difference between the apo hPNP-hPNP·guanine (black) and apo hPNP-hPNP·ImmH (red) crystal structure pairs. The different RMSDs of both pairs of crystals, evidence that the TS analogue is bound to hPNP much tighter than the substrate analogue.

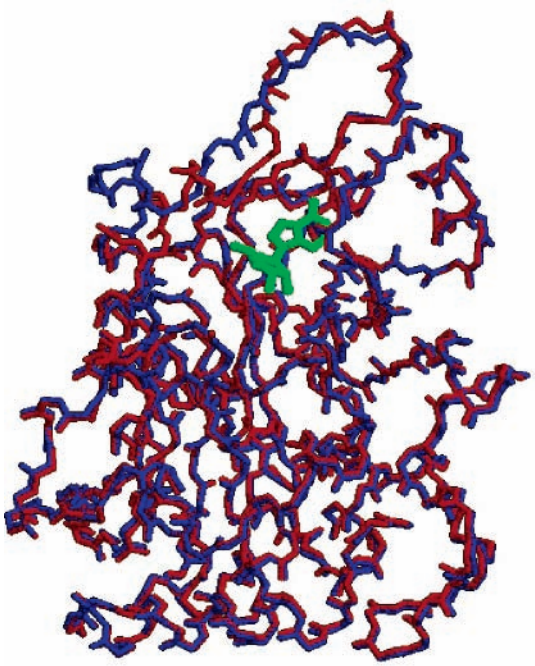

Figure 6. Backbone superimposition of the crystal structures of apo hPNP (blue) and hPNP complexed with the transition state inhibitor ImmH (red). The transition state analogue ImmH is shown in green.

5. Results and Discussion

5a. Stability of the Simulations. The stability of the previous simulations was examined by conventional structural and geometrical analysis. Figure 7 shows the root-mean-square deviation (RMSD) of the C_α atomic positions for the APO, ES, and EI-APO molecular dynamics simulations with respect to the crystal. It can be seen that the simulations have reached an equilibrated state. The RMSDs of the C_α atoms of the three simulations of hPNP suggest that the systems are relatively stable over the explored time scale and do not significantly deviate from their respective crystal structures. The radius of gyration of the apo and complexed protein remains essentially constant (data not shown).

The solvent-accessible-surface area (SASA) was calculated for the non-hydrogen atoms in the 241−265 loop in the context of the protein. The SASA calculation uses a probe radius of 1.4 Å and the Lennard−Jones hard-shell radii to define the surface, as implemented in CHARMM. Figure 8 shows the SASA for the production run of the APO, ES and EI-APO simulations. We can see the solvent-accessible-surface area for the loop region is higher for the APO model than for the ES model. This is an indication that the loop is more exposed to the solvent in the apo form of the enzyme, and that, upon binding of the substrate, a tightening of the loop occurs. Comparison of the substrates' motion during the molecular dynamics simulation of the apo and Michaelis complex of hPNP shows that the substrates appear to be strongly tethered to the active site pocket, with the loop significantly closing in on the active site. This effect can be quantified by calculating the mean distance of the C_α atoms of the residues forming the active site. The average distances are rather constant and considerably lower in the Michaelis complex than in apo hPNP, indicating that the loop and rest of the active site residues tend to embrace the purine nucleoside in a tighter manner (data not shown).

5b. Correlated Motions. Apart from conventional structural and geometrical analysis to assess the stability of the simulations, ED analysis is utilized to identify large concerted motions in hPNP. We build the covariance matrixes of the atomic fluctuations in a molecular dynamics trajectory, in which the overall translation and rotation modes have been removed. Elimination of the overall motion is done at each simulation step, and we ensure the total removal of these irrelevant motions by fitting all the structures to the initial structure. Experience has shown that it is sufficient to use only the C_α atomic positions in building the covariance matrix. The matrix is built considering structures sampled every 5 fs from the production run of the simulations. A total of 282 residues are included in the analysis, and diagonalization of the covariance matrix results in 846 eigenvectors. After diagonalizing the covariance matrix, the robustness of the results is checked by splitting the trajectory in half and performing ED in each half and then comparing the results and also forming inner products between members of the eigenvectors in each half and checking for orthogonality.[14]

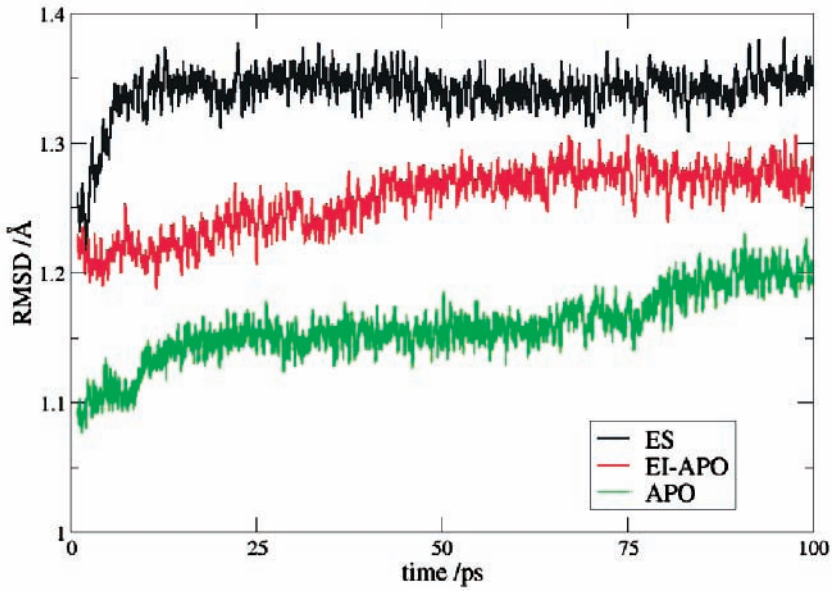

Figure 7. Root-mean-square deviation (RMSD)/Å of the APO (green) and ES (black) and EI-APO (red) molecular dynamics simulations.

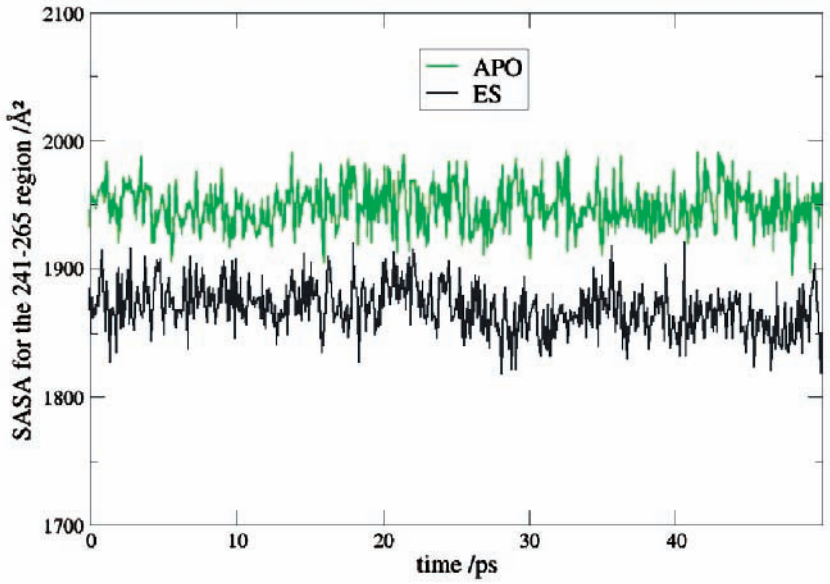

Figure 8. Loop's solvent-accessible-surface area (SASA)/Å^2 of the transient MD structures during the APO (green) and ES (black) molecular dynamics simulations.

TABLE 1: Contribution to the Overall Motion of the First 10, 20, 50, and 100 Eigenvalues for the APO, ES, and EI-APO Simulations[a]

number of eigenvalues	APO	APO2	APO3	ES	ES2	ES3	EI-APO	EI-APO2	EI-APO3
10	22	26	54	42	51	38	41	26	77
20	31	35	59	48	57	45	48	35	81
50	46	50	68	59	66	57	60	49	85
100	61	64	76	70	75	68	72	63	89

[a] APO, APO2, and APO3 correspond to simulations of the empty enzyme; ES, ES2, and ES3 correspond to simulations of the Michaelis complex with the substrates guanosine and phosphate; and EI-APO, EI-APO2, and EI-APO3 are the simulations of the ImmH-stripped hPNP·ImmH complex

The contribution to the overall motion of the first (largest) 10, 20, 50, and 100 eigenvalues of the APO, ES, and EI-APO simulations is shown in Table 1. It can be seen that the first 10 eigenvectors describe a representative percentage of the motion. For other biological problems, e.g., for the protein folding event, it has been found that the first 5 eigenvalues can represent 50−80% of the macromolecule's motion. Figures 9−11 show the first nine eigenvectors of one singled-out simulation for each model and the contribution from each residue in the protein. The first nine eigenvectors of the 200 ps and 1 ns vacuum simulations of the three models and the 200 ps solvated simulation of the APO model are shown in the Supporting Information.

For the APO simulations, from the first few eigenvectors, we observe prominent correlated motions in the neighboring residues F159 and P150, adjacent to the purine substrate.

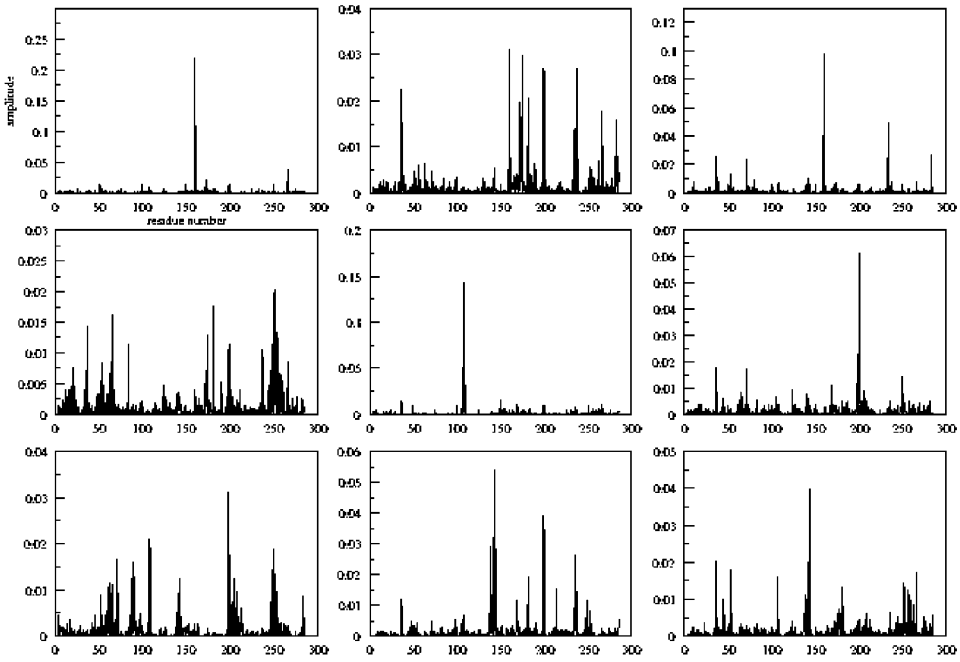

Figure 9. From left to right and top to bottom, the eigenvectors 1−9 of the APO simulation are shown.

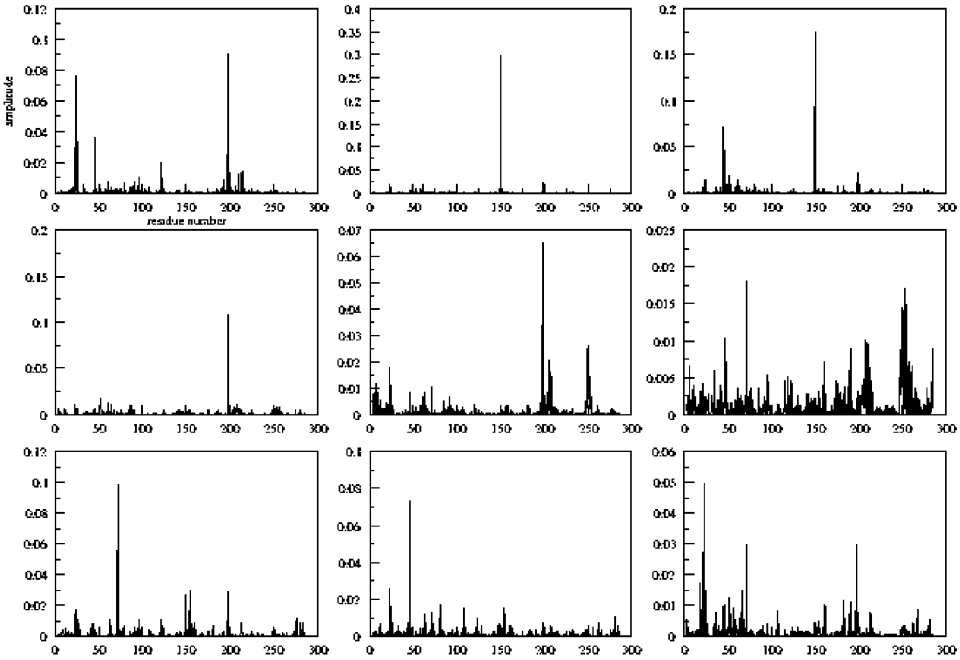

Figure 10. From left to right and top to bottom, the eigenvectors 1−9 of the ES simulation are shown.

Moreover, it is not surprising that G63 and E183 also show strong motions, granted their large B-factors. Additionally, in a lesser extent, the active site residues S33, H84, H86, and F200 display high mobilities. These concerted motions were also observed for the 1 ns vacuum simulations and for the solvated simulation. Overall, the most significant feature of the APO simulations is the solid correlated motions found in the spatially neighboring F159 and P150 residues. In the case of the ES simulations, we discern that there exist robust correlated motions around the F200 and E250 residues, the latter being the epicenter of the loop dynamics. In a lesser magnitude, residues S33 and P150 likewise show concerted behavior. Finally, it is again not surprising that residues H183 and N121 are also involved in dynamics granted their high mobilities based on its high B-factors. Consequently, we already find differences in the concerted motions between the APO and ES models of hPNP: the APO model presents concerted dynamics around the F159 and P150 residues, while, for the simulation of the Michaelis complex (ES), concerted motions are mostly found around the active site residue F200 and loop residue E250. This is in turn suggested that dynamics in the apo enzyme is concentrated around the residues F159 and P150, though once the substrate enters the active site, these become dormant, and active site residue F200 and loop residue E250 (embracing the purine substrate) come into action. For the EI-APO simulations, we note that correlated motions are again found around the neighboring residues F159 and P150, as well as around F200, E250, and S33. This behavior resembles a combination of the APO and ES models. These results suggest these residues are coordinated in the enzyme so that it can successfully accommodate the substrate in the active site and achieve transition state formation.

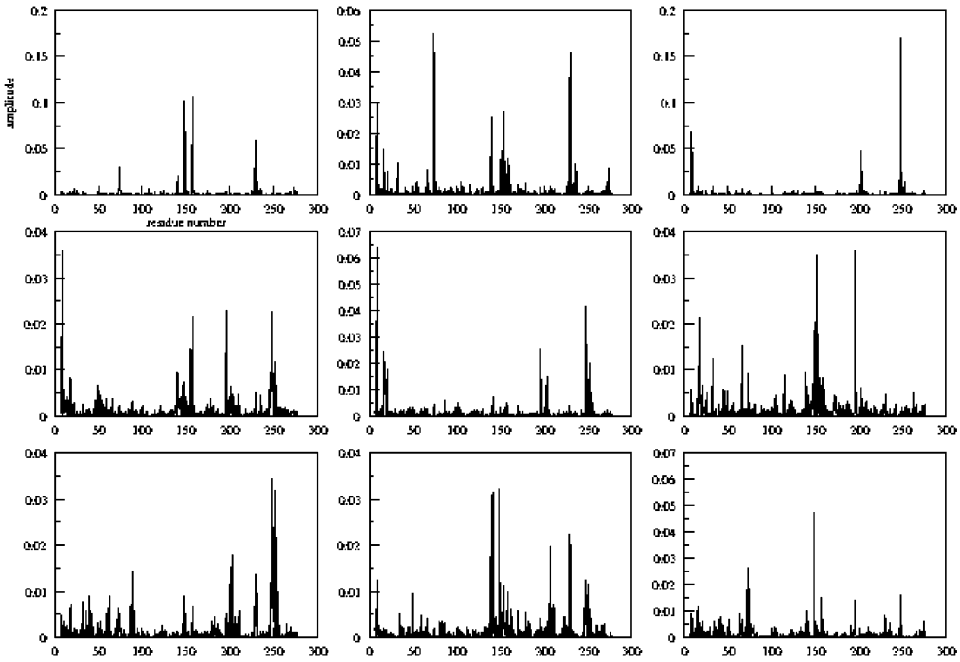

Figure 11. From left to right and top to bottom, the eigenvectors 1−9 of the EI-APO simulation are shown.

We observe that the three computational models (APO, ES, and EI-APO) present distinct characteristics in their correlated motions. Some similarities are found, for example, in the high mobilities of G63 and E183, in agreement with their high crystallographic B-factors. However, other residues which had not been seen as highly mobile based on their B-factors (namely F159, P150, H230, and F200) have been found to be have correlated dynamic modes. This implies that observation of crystallographic structures alone may well not be enough to infer dynamic behavior in proteins and that additional MD studies are needed to identify correlated dynamical modes. Finally, we have been able to successfully recognize correlated motions on the 241−265 loop, particularly present in the simulations of the Michaelis complex, and in a subsequent extent for the EI-APO simulations. This suggests that the loop may have a consequential favorable role in capturing the substrate as well as in the chemical step. Figure 12 highlights the residues that were identified by the principal component analysis. The ED analysis has found significant dynamics in nonloop residues F159 and F200, whose involvement in binding was not obvious from B-factors alone.

We cannot infer from an essential dynamics study alone that specific modes are directly involved in enhancing phosphorolytic efficiency. To test whether the concerted motions identified by ED are involved in catalysis, a few identified key residues were experimentally mutated. Essentially, the ED method found correlated motions around F200, F159, P150, and H230. Interestingly, these residues are spatially close to each other, and moreover, F200 and F159 are directly adjacent to the substrates, suggesting a dynamical role in binding and/or catalysis. Minor concerted motions occur around the residues G63, Y88, N121, and E183. We decided to mutate F200, F159, and H230 to the less bulkier residue glycine, i.e., F159G, F200G, and H230G. Additionally, N121, which is a residue which shows relatively high B-factors and correlated motions in the ES3 simulation, was also mutated into glycine to test whether its dynamics has any effect on catalysis.

5c. Site-Directed Mutagenesis. The X-ray crystal structures indicate that all direct contacts between hPNP and the substrates guanosine and phosphate are through amino acid side chains.

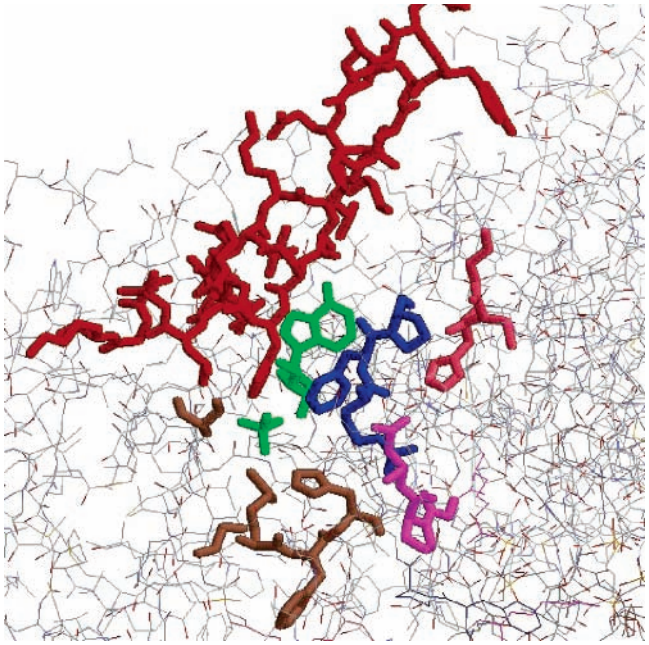

Figure 12. hPNP with the residues which show correlated motions highlighted: S33 (brown), H64 (brown), H86 (brown), P150 (magenta), F159 (blue), H230 (pink), and 241−265 loop (red).

Hence, our site-directed mutagenesis strategy replaced each residue with glycine to asses the contribution of the residue's side chain to substrate binding and/or catalysis. The mutagenesis and purification of his-tagged hPNP mutants are as follows. Site-directed mutagenesis of four amino acid residues was carried out using a QuikChange Site-Directed Mutagenesis kit (Stratagene) using an hPNP template consisting of the hPNP gene inserted into pCRT7/NT-TOPO.[35] The designed primers for each site-directed mutant are shown in Table 2. His-tagged hPNP mutants were purified as previously reported.[35] Assays for the determination of kinetic constants for the hPNP mutants were carried out as previously reported by coupling the production of hypoxanthine to uric acid by xanthine oxidase.[36] Results from the kinetic studies (Table 3) indicate that substrate binding affinity was most sensitive to the F159G and F200G mutations.

TABLE 2: Forward and Reverse Primers Used in the Site-Directed Mutations of Several Residue in hPNP Which Showed Concerted Modes in the ED Analysis

mutation	primer
N121G forward	5′-CAGCAGGAGGGCTGGGGCCCAAGTTTGAGGTTG-3′
N121G reverse	5′-CAACCTCAAACTTGGGCCCCAGCCCTCCTGCTG-3′
F159G forward	5′-AAGGTTTGGAGATCGTGGCCCTGCCATGTCTGATG-3′
F159G reverse	5′-CATCAGACATGGCAGGGCCACGATCTCCAAACCTT-3′
F200G forward	5′-GGCAGGCCCCAGCGGCGAGACTGTGGCAGAA-3′
F200G reverse	5′-TTCTGCCACAGTCTCGCCGCTGGGGCCTGCC-3′
H230G forward	5′-GTTATCGTTGCACGGGGCTGTGGACTTCGAGTC-3′
H230G reverse	5′-GACTCGAAGTCCACAGCCCCGTGCAACGATAAC-3′

TABLE 3: Experimental Kinetic Parameters for Different Mutants Which Show Correlated Motions in hPNP

mutant	WT	F200G	F159G	H230G	N121G
$K_M/\mu M$	76.8	1580	2260	145	64.9
k_{cat}/s^{-1}	32.4	3.8	2.3	2.6	33.0
$k_{cat}/K_M/10^5\ s^{-1}\ M^{-1}$	4.2	0.024	0.01	0.1	5.0

In the phosphorolytic direction, an increase of ca. 400-fold in K_M (and modest change in k_{cat}) is observed for the F159G mutant. Loss of the herringbone-type interaction between the β-face of the ribose and the hydrophobic surface created by this residue is important for substrate binding, thus essential for hPNP activity. Additionally, the H230G mutant showed modest changes in both K_M and k_{cat}, surprising catalytic contribution granted its long distance from the active site. Finally, the N121G mutant did not show any relevant change in catalytic efficiency, regardless of proving substantial concerted dynamic behavior during the Michaelis complex simulation.

One could argue that the decrease in catalytic efficiency we found for the F159, F200, or V260G mutants might be due to a change in the enzyme's structure, i.e., the loss of the side chain−substrate interactions that cause the substrate to bind incorrectly with the concomitant loss in transition state stabilization. Consequently, and due to lack of structural data for the mutants, we calculated the radius of gyration and RMSD of the equilibrated in silico structures for the WT and respective mutants to determine the extent of enzyme distortion and the extent in which the mutations might have caused the observed concomitant loss of catalytic efficiency. In all the three cases analyzed, the radius of gyration and RMS deviation for both WT and mutant structures were small. Therefore, we are able to assert that the alteration of the enzyme's structure is practically negligible, with the mutants conserving structures alike to the WT. We conclude by recalling that the ED method is used here to determine concerted motions in human PNP. However, the computational method alone is not able to quantitate the extent in which mobile residues are involved in catalysis; thus additional experimental kinetic and structural evidence is required.

6. Concluding Remarks

This study is based on the crystallographic observation that the 241−265 loop in hPNP undergoes a conformational change upon substrate binding in the active site. The dynamics of hPNP that may affect the chemical reaction of phosphoryl transfer have been investigated, and essential dynamics analysis on three models of hPNP with and without substrate (namely, APO, ES, and EI-APO) has been carried out to separate large-scale correlated motions from small-scale random harmonic vibrations. A detailed comparison was made between the collective fluctuations observed in the apo and Michaelis complex forms. Concerted motions were also investigated during the ligand-induced geometrical *relaxation* of the loop in the simulation of complexed hPNP when the substrate/inhibitor is taken out from the active site. In summary, the network of conformational changes which plays an important role upon substrate binding has been presented. Molecular dynamics simulations have proved useful to pinpoint the nature of conformational changes in hPNP and the relationship to the presence/absence of the natural substrates guanosine/phosphate in the active site, substantiating that concerted motions in hPNP *effect* catalysis. This implies that MD studies, together with static crystallographic data, are useful to infer dynamic behavior in proteins and identify correlated dynamical modes.

Acknowledgment. The authors thank The National Science Foundation Chemistry Division CHE-0139752 (S.D.S., S.N.), The Office of Naval Research N00014-00-1-0097 (S.D.S., S.N.), and the National Institute of Health GM41916 and GM068036 for support.

Supporting Information Available: Computational details of the Essential Dynamics analysis performed for the APO, ES, and EI-APO. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Pande, V. S.; Rokhsar, D. S. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9062−9067.
(2) Kazmirski, S. L.; Li, A.; Dagget, V. *J. Mol. Biol* **1999**, *290*, 283−304.
(3) Shea, J. E.; Brooks, C. L. *Annu. Rev. Phys. Chem.* **2001**, *52*, 499−535.
(4) Lee, J.; Jang, S.; Pak, Y.; Shin, S. *Bull. Korean Chem. Soc.* **2003**, *24*, 785−791.
(5) Ota, N.; Agard, D. A. *Protein Sci.* **2001**, *10*, 1403−1414.
(6) Peters, G. H.; van Aalten, D. M. F.; Svendsen, A.; Bywater, R. *Protein Eng.* **1997**, *10*, 149−158.
(7) Piana, S.; Carloni, P.; Parrinello, M. *J. Mol. Biol.* **2002**, *319*, 567−583.
(8) Sirovich, I.; Rodriguez, J. *Phys. Lett. A* **1987**, *120*, 211−214.
(9) Caves, L. S.; Evanseck, J. D.; Karplus, M. *Protein Sci.* **1998**, *7*, 649−666.
(10) Garcia, A. E. *Phys. Rev. Lett.* **1992**, *68*, 2696−2699.
(11) Amadei, A.; Linssen, A. B. M.; Berendsen, H. J. C. *Proteins* **1993**, *17*, 412−425.
(12) Scheek, R. M.; van Nuland, N. A. J.; de Groot, B. L.; Amadei, A. *J. Biomol. NMR* **1995**, *6*, 106−111.
(13) Hayward, S.; Kitao, A.; Go, N. *Protein Sci.* **1994**, *3*, 936−943.
(14) Hess, B. *Phys. Rev. E* **2000**, *62*, 8438−8448.
(15) Daidone, I.; Amadei, A.; Roccatano, D.; Nola, A. D. *Biophys. J.* **2003**, *85*, 2865−2871.
(16) Sulpizi, M.; Rothlisberger, U.; Carloni, P. *Biophys. J.* **2003**, *84*, 2207−2215.
(17) de Groot, B. L.; Amadei, A.; Scheek, R. M.; van Nuland, N. A. J.; Berendsen, H. J. C. *Proteins* **1996**, *26*, 314−322.
(18) Balsera, M. A.; Wriggers, W.; Oono, Y.; Schulten, K. *J. Phys. Chem.* **1996**, *100*, 2567−2572.
(19) Stoeckler, J. D.; Cambor, C.; Parks, R. E. J. *Biochemistry* **1980**, *19*, 102−107.
(20) Giblett, E. R.; Ammann, A. J.; Wara, E. W.; Sandman, R.; Diamond, L. K. *Lancet I* **1975**, 1010−1013.

(21) Krenitsky, T. A.; Tuttle, J. V.; Koszalka, G. W.; Chen, I. S.; Beachman, L. M.; Rideout, J. L.; Elion, G. B. *J. Biol. Chem.* **1976**, *251*, 4055−4061.

(22) Carson, D. A.; Kaye, J.; Seegmiller, J. E. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, *74*, 5677−5681.

(23) Mitchell, B. S.; Mejias, E.; Daddona, P. E.; Kelley, W. N. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 5011−5014.

(24) Ullman, C.; Gudas, L. J.; Clift, S. M.; Martin, D. W. *Proc. Natl. Acad. Sci. U.S.A.* **1979**, *76*, 1074−1078.

(25) Schramm, V. L. *Biochim. Biophys. Acta* **2002**, *1587*, 107−117.

(26) Kicska, G. A.; Long, L.; Horig, H.; Fairchild, C.; Tyler, P. C.; Furneaux, R. H.; Schramm, V. L.; Haufman, H. L. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4593−4598.

(27) Bantia, S.; Miller, P. J.; Parker, C. D.; Anath, S. L.; Horn, L. L.; Kilpatrick, J. M.; Morris, P. E.; Hutchinson, T. L.; Montgomery, J. A.; Sandhu, J. S. *Int. Immunopharmacol.* **2001**, *1*, 1199−1210.

(28) Erion, M. D.; Takabayashi, K.; Smith, H.; Kessi, J.; Wagner, S.; Honger, S.; Shames, S.; Ealick, S. E. *Biochemistry* **1997**, *36*, 11725−11734.

(29) Shi, W.; Lewandowicz, A.; Tyler, P. L.; Furneaux, R. H.; Almo, S. C.; Schramm, V. L. On hold for publication, PDB ID: 1RR6.

(30) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.

(31) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740−744.

(32) Clarage, J. B.; Romo, T.; Andrews, B. K.; Pettitt, B. M.; Phillips, G. N. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 3288−3292.

(33) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R.; Klein, M. *J. Chem. Phys.* **1983**, *79*, 926.

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

(35) Lewandowicz, A.; Schramm, V. L. *Biochemistry* **2004**, *43*, 1458−1468.

(36) Lewandowicz, A.; Shi, W.; Evans, G.; Tyles, P.; Furneaux, R.; Basso, L.; Santos, D.; Almo, S.; Schramm, V. *Biochemistry* **2003**, *42*, 6057−6066.