Computational Study of IAG-Nucleoside Hydrolase: Determination of the Preferred Ground State Conformation and the Role of Active Site Residues†

Devleena Mazumder-Shivakumar§ and Thomas C. Bruice*

Department of Chemistry and Biochemistry, University of California, Santa Barbara, CA 93106

Received December 10, 2004; Revised Manuscript Received March 9, 2005

ABSTRACT: The mechanism of action of inosine-adenosine-guanosine nucleoside hydrolase (IAG-NH) has been investigated by long-term molecular dynamics (MD) simulation in TIP3P water using stochastic boundary conditions. Special attention has been given to the role of leaving group pocket residues and conformation of the bound substrate at the active site of IAG-NH. We also describe the positioning of the residues of an important flexible loop at the active site, which was previously unobservable by X-ray crystallography due to high *B*-factors. Five MD simulations have been performed with the Enzyme•Substrate complexes: Enzyme•*anti*-Adenosine with Asp40-COOH [E(40H)•Ade(a)], Enzyme•*anti*-Adenosine with Asp40-COO^- [E(40)•Ade(a)], Enzyme•*syn*-Adenosine with Asp40-COOH [E(40H)•Ade(s)], Enzyme•*syn*-Adenosine with Asp40-COO^- [E(40)•Ade(s)], and Enzyme•*anti*-Inosine with Asp40-COO^- [E(40)•Ino(a)]. Overall, the structures generated from the MD simulation of E(40H)•Ade(s) preserve the catalytically important hydrogen bonds as well as electrostatic and hydrophobic interactions to provide a plausible catalytic structure. When deprotonated Asp40 (Asp4-COO^-) is present, the active site is open to water solvent which interferes with the base stacking between Trp83 and nucleobase. A calculation using Poisson−Boltzmann equation module supports that Asp40 indeed has an elevated pK_{app}. Solvent accessible surface area (SASA) calculations on all the five MD structures shows that systems with protonated Asp40, namely, E(40H)•Ade(a) and E(40H)•Ade(s), have zero SASA. It is found that a water molecule is hydrogen-bonded to the N7 of the nucleobase and is probably the essential general acid to protonate the departing nucleobase anion. The N7-bonded water is in turn hydrogen-bonded to waters in a channel, held in place by the residues of the flexible loop, Tyr257, His247, and Cys245. Using normal-mode analysis with elastic network model, we find that the flexible loop explores a conformational space much larger than in the MD trajectory, leading to a "gating"-like motion with respect to the active site.

Millions of people are affected every year by Trypanosomiasis, caused by protozoan parasites of the genus *Trypanosoma*, which enter the bloodstream via the bite of blood-feeding tsetse flies. Most of the parasitic protozoa lack a de novo purine biosynthesis, and thus depend on purine salvage from the hosts for their metabolic pathways and nucleic acids biosynthesis (*1*). Nucleoside hydrolases (NH[1], EC 3.2.2.-) are members of the purine salvage pathway that tend to scavenge purines from their environment. Neither the NH enzymes nor the genes encoding them is found in humans, making it a very important target for antiparasitic chemotherapy (*1*, *2*).

The NH catalyzes the hydrolysis of the *N*-glycosidic bond between the anomeric carbon atom of ribose and the purine or pyrimidine base (Scheme 1). On the basis of the substrate specificity, NH has been divided into three subgroups: inosine-uridine (IU) (*3*), inosine-adenosine-guanosine (IAG) (*4*, *5*), and guanosine-inosine (GI) (*6*). IU-NH has been extensively studied using biochemical (*7*−*9*) and computational tools (*10*, *11*). Schramm and co-workers have established the involvement of a ribooxocarbenium ion-like transition state for NHs (*7*). The transition states for IU-NH and IAG-NH indicate that more bond order remains to the leaving group purine at the transition state of IAG-NH than of IU-NH, establishing an earlier transition state for IAG-NH (*8*). This difference has been attributed to a stronger leaving group interaction in IAG-NH compared to IU-NH, such that half (∼ 8.8 kcal/mol) of the total activation energy (∼17.7 kcal/mol) is provided by leaving group stabilization in IAG-NH (*12*). IU-NH derives only 4.6 kcal/mol from leaving group stabilization and the rest from ribosyl interactions. Therefore, the leaving group interaction of IAG-NH is an important subject for further investigation.

The IAG-NH from *Trypanosoma vivax* is a homodimer of M_r 36 000 subunits (*5*). Scheme 1 depicts the general acid catalysis of departure of the leaving group with formation of a ribooxocarbenium ion-like transition state. The pK_a of the N7 atom of inosine is 2.3, and 8.5 in the product

† Supported by the National Institute of Health-5R37DK9174-41.

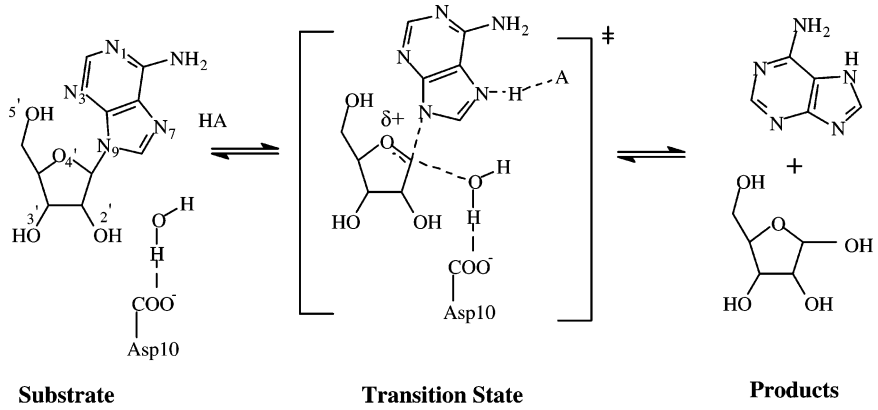
* To whom correspondence should be addressed. Telephone: (805) 893 2044. Fax: (805) 893 2229. E-mail: tcbruice@chem.ucsb.edu.

§ Current address: The Scripps Research Institute, 10550 North Torrey Pines Rd, TPC 15, La Jolla, CA 92037, devleena@scripps.edu.

[1] Abbreviations: NH, nucleoside hydrolase; IU, inosine-uridine; IAG, inosine-adenosine-guanosine; MD, molecular dynamics, rmsd, root-mean-squared deviation; SBMD, stochastic boundary molecular dynamics; SD, steepest descents; ABNR, adopted basis Newton−Raphson; PB, Poisson−Boltzmann; SASA, solvent accessible surface area; NMA, normal-mode analysis.

10.1021/bi047394h CCC: $30.25
Published on Web 05/05/2005

Scheme 1



hypoxanthine. The general acid-mediated protonation of the nucleoside base neutralizes the developing negative charge on the purine ring in the transition state and, thus, facilitates its departure. The pH profile of k_{cat}/K_m for IAG-NH with inosine as a substrate shows an optimum near pH 5.0 and involvement of two groups on the free enzyme with pK_{app}'s of 5.0 ± 0.1 and 7.7 ± 0.1 in protonated states, and a third group with a pK_{app} of 5.1 ± 0.1 in a deprotonated state (*13*). In the *T. vivax* nucleoside hydrolase, however, no enzymatic general acid species has thus far been identified by crystallographic studies and mutagenesis experiments. Initially, the general acid was thought to be part of a flexible loop that has high *B*-factor and, thus, could not be observed in X-ray studies. However, an alanine scanning of the loop did not show a considerable drop in k_{cat} and, therefore, did not support the presence of an operational general acid in the loop (*13*). It should also be pointed out that the rate constants obtained from alanine scanning studies referred to the equilibrium rate constants, and later, it was revealed that product release is the rate-limiting step in IAG-NH catalysis (*14*). On the basis of its location at the active site, the residue Asp40 in the leaving group pocket was thought to be the general acid, but the Asp40Ala mutant did not show any considerable drop in activity. Steyaert and co-workers suggested aromatic base stacking as an alternative to general acid catalysis (*15*). They also made the important observation of a hydrogen bond between a water molecule and the N7 of the nucleobase.

The available crystal structures pose an important question regarding the orientation of the *N*-glycosyl bond [defined by the atoms O4′−C1′−N9−C4] of the substrate at the IAG-NH active site. The X-ray structures of the substrate analogue, 3-deazaadenosine, bound to wild-type (PDB code: 1HP0) (*5*), and Asp10Ala mutant (PDB code: 1KIE) (*13*) IAG-NH show that the base is disposed in a *syn*-conformation with respect to the sugar. However, the X-ray structure of the substrate inosine-bound Asp10Ala mutant of IAG-NH (PDB code: 1KIC) (*13*) shows substrate to be in an *anti*-conformation. This raises an important question: what is the preferred orientation of the substrate bound to the active site of IAG-NH?

To study the IAG-NH catalysis in atomic details we have performed *five* MD simulation studies with different orientations of the glycosyl bond in the substrate and protonation states of Asp40: Enzyme·*anti*-Adenosine with Asp40-COOH [E(40H)·Ade(a)], Enzyme·*anti*-Adenosine with Asp40-COO$^-$ [E(40)·Ade(a)], Enzyme·*syn*-Adenosine with Asp40-COOH [E(40H)·Ade(s)], Enzyme·*syn*-Adenosine with Asp40-COO$^-$ [E(40)·Ade(s)], and Enzyme·*anti*-Inosine with Asp40-COO$^-$ [E(40)·Ino(a)]. Here, we report the detailed analysis of our results.

MATERIALS AND METHODS

MD Setup. The starting coordinates of the substrate analogue, 3-deazaadenosine-bound IAG nucleoside hydrolase homodimer (322 amino acid residues in each subunit) were taken from the entry 1HP0 (2.1 Å resolution) in the Protein Data Bank (*5*). The CHARMM27 topology and parameters were used for protein, inosine, and TIP3P water model (*16*). Hydrogen atoms were added to the X-ray structure via the empirical energy placement protocol H-BUILD (*17*) in the CHARMM program (*16*). Aspartate, glutamate, arginine, and lysine residues were charged, and all the tyrosine residues were neutral, unless otherwise specified. The protonation sites of histidine residues were based on the availability of H-bond donors or acceptors nearby. Both the protonated and deprotonated form of the residue Asp40 were used for our simulations. Partial atomic charges for the substrate ribosyl hydroxyls and the active site Thr137 in the presence of calcium cation were obtained using the Gaussian 98 program (*18*) as described elsewhere (*11*).

Modeling of the Flexible Loop. The loop containing residues 246−257 had high *B*-factor in X-ray structures and, thus, could not be crystallized. The subunit A had residues 248−258 missing, whereas subunit B had residues 248−252 missing from the crystal structure. Since subunit B had only five missing residues, it was an obvious choice for carrying the MD studies. The initial structure for the loop was built using the "Builder" module in the program SYBYL (*19*). Two initial structures (E·Adenosine(*syn*) and E·Adenosine(*anti*)) were generated using the *syn*- and *anti*-conformer of adenosine docked at the active site. Two initial structures were used since the flexible loop was within the nonbonded cutoff distance of the substrate. The enzyme−substrate system was solvated in an equilibrated TIP3P water sphere of 84 Å diameter using the center of mass of the substrate as the origin. Any solvent molecule within 2.8 Å of a heavy atom of the enzyme or crystallographic water molecule was deleted. The 6508 added solvent molecules (19 524 atoms) were minimized using the steepest descents (SD) method followed by adopted basis Newton−Raphson (ABNR) minimization protocol in CHARMM. Stochastic

Boundary Molecular Dynamics (SBMD) procedure was carried out for 3 ns using a 40 Å reaction zone with a 2 Å buffer region (*20*). Eleven sodium ions were added to maintain the electroneutrality of the system. All the protein atoms, substrate, crystal waters, and sodium ion were constrained to their minimized crystallographic coordinates. Only the solvent water molecules and the atoms of the flexible loop were allowed to move during dynamics. Both the models, E·Adenosine(*syn*) and E·Adenosine(*anti*), contained a total of ~30 300 atoms. The same protocols for energy minimization and dynamics were used for both models. The structure averaged from the last 200 ps was used for further MD studies.

Modeling of Enzyme. Substrate Complexes. The structure generated from the above loop building protocol were used as starting structures for Enzyme·*anti*-Adenosine with Asp40-COO^- [E(40)·Ade(a)] and Enzyme·*syn*-Adenosine with Asp40-COO^- [E(40)·Ade(s)]. The adenosine was replaced by inosine in E(40)·Ade(a) to generate the starting structure for Enzyme·*anti*-Inosine with Asp40-COO^- [E(40)·Ino(a)]. A proton was added to the Asp40-COO^- of the above systems to generate the corresponding carboxylic acids, Enzyme·*anti*-Adenosine with Asp40-COOH [E(40H)·Ade(a)] and Enzyme·*syn*-Adenosine with Asp40-COOH [E(40H)·Ade(s)].

SBMD simulation protocol (*20*) was used for the above five systems. The reaction region around an active site was the sphere of radius r of 40 Å; the buffer region was $40 < r < 42$ Å, and the reservoir region corresponded to $r > 42$ Å. Similar protocols for energy minimization and dynamics were used for all models. Before the start of the dynamics, the systems were energy minimized using the SD method followed by ABNR methods for 5000 steps. All atoms in the reservoir region were deleted after the minimization. All five models contained a total of ~30 300 atoms after deletion of reservoir region atoms. The solvent molecules were equilibrated at 300 K for 3 ps keeping the protein, ligand, and the crystallographic water molecules fixed to allow favorable distribution of water molecules on the enzyme surface before the start of free dynamics. Another set of brief minimization was done before the start of dynamics using the SD protocol. During the molecular dynamics simulation, the system was gradually heated by coupling to a heat bath of 150 K for 5 ps, 250 K for the next 15 ps, and 300 K for the rest of the simulation time using a frictional coefficient of 250 ps^{-1} on the heavy atoms of the enzyme. Heavy atoms of the protein in the buffer region were constrained using force constants calculated from their average Debye−Waller factors (*20*). Bonds containing hydrogen were constrained using SHAKE algorithm (*21*). The nonbonded interactions were updated every 20 steps and were cut off at 12 Å by means of a force-shifting function.

The X-ray crystal structure places a nucleophilic crystallographic water molecule Wat1 in position to occupy one of the coordination sites of the octacoordinated calcium ion and at 3.3 Å from the ribosyl 1′-carbon of 3-deaza-adenosine in subunit B (Scheme 1). Preliminary simulation results showed that this water molecule does not exchange with waters from the water pool. However, the calcium−water distance increased from 2.48 to 3 Å during the test dynamics. Thus, a distance constraint was placed between the active site nucleophilic water and calcium ion using a force constant of 10 kcal mol^{-1} $Å^{-1}$. The following MD simulations were carried out: Enzyme·*anti*-Adenosine with Asp40-COOH [E(40H)·Ade(a)] and Enzyme·*anti*-Adenosine with Asp40-COO^- [E(40)·Ade(a)] for 4.5 ns duration; Enzyme·*syn*-Adenosine with Asp40-COOH [E(40H)·Ade(s)], Enzyme·*syn*-Adenosine with Asp40-COO^- [E(40)·Ade(s)], and Enzyme·*anti*-Inosine with Asp40-COO^- [E(40)·Ino(a)] for 5 ns duration.

MD Analysis. The molecular dynamics analysis were carried out after the initial 1 ns for E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a). For all our statistical analysis, only the B-subunit was considered. The "correl" module in CHARMM program was used to extract the data for time-dependent plots from MD trajectories. As a measure of structural stability, root-mean-squared deviation (rmsd) of the MD-simulated structures from the starting structures were calculated on the basis of all the backbone heavy atoms. To reveal the most flexible regions of the protein structure, *B*-factors were calculated for the Cα atoms from the mean square fluctuation

$$\text{(msf) using:} \qquad B = \frac{8\pi^2}{3}(\text{msf}) \tag{1}$$

Calculation of pK_a Shift. The pK_a is related to proton affinity, ΔG, by

$$pK_a = \frac{\Delta G}{k_B T(\ln 10)} \tag{2}$$

By means of our employed force fields we cannot calculate ΔG directly, partly because it involves the solvation energy of a proton and partly because the employed force fields lack the description of (intrinsic) proton affinities. On the other hand, the force fields are capable of describing changes due to altered electrostatic interactions with the environment.

The pK_a shift of Asp40 was calculated using the "PBEQ" module in CHARMM program. ΔG is calculated by first removing all the water molecules and then recalculating the energy with and without a proton added to Asp40-COO^- using a Poisson−Boltzmann (PB) equation module for continuum electrostatics to describe the solvation of the enzyme. The proton affinity can be obtained using the following equation:

$$\Delta G^{\text{calc}} = (E_s^P - E_s^U) - (E_i^P - E_i^U) \tag{3}$$

where, E_s^P and E_i^P are the energy corresponding to the protonated side chain in structure and isolated and E_s^U and E_i^U are the energy corresponding to the unprotonated side chain in structure and isolated. The unprotonated and protonated states refer to Asp40-COO^- and Asp40-COOH, respectively. Absolute pK_a values are obtained by adding the experimental value of 3.8 (for aspartic acid in aqueous solution) to the calculated pK_a shifts.

The atomic radii used for PB calculations to define the location of the solvent−protein dielectric boundary were taken from the work published by Roux and co-workers (*22*, *23*). The PB calculations were done at different dielectric constant ($\epsilon = 2$, 3, 4, 5, and 6) for the enzyme.

Solvent Accessible Surface Area (SASA). The solvent accessible surface area was calculated using the program

NACCESS (*24*). The NACCESS program calculates the atomic accessible surface area defined by rolling a probe of given size (1.4 Å = radius of water) around a van der Waals surface of a macromolecule by implementing the Lee and Richards method (*25*). The calculation makes successive thin slices through the 3D molecular volume to calculate the accessible surface of individual atoms. The SASA was calculated for E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a) using the 3-dimensional coordinate sets from the average structure.

Correlated Motion and Normal-Mode Analysis. A covariance matrix, generated using the CHARMM program, was used to determine the extent of correlated motion between the Cα of the residues in E(40H)·Ade(s). The color-coded matrix was plotted using the program GMT (*26*). MD trajectories were visualized using the program gOpenMol (*27*, *28*) to see the direction of the anti-correlated motion between the desired residues.

Normal-mode analysis (NMA) was done using the elastic network model (*29*). Several studies have shown that this Hookean potential is sufficient to reproduce the low-frequency normal modes of proteins as produced by more complete potential energy functions (*30*). Elastic network model is an efficient way to study the character (directions) of the low-frequency motion but cannot determine the absolute scale of normal mode eigenvalues (frequencies of the motion) correctly. Even though the absolute energy scale or time scale of the molecular motions is distorted in these calculations, the character (directions) of the low-frequency motions is unaffected, thus, providing an efficient way to study the nature of large conformational rearrangements of biological systems that are not accessible via molecular dynamics simulation. In our elastic network model, amino acids were represented by their Cα atoms. The Hessian matrix was constructed using a force constant of 1 kJ·Å^{-2}·mol^{-1} and the cutoff of 8 Å. The plot showing the normal mode vectors was generated using the VMD program (*31*).

RESULTS

Modeling of the Flexible Loop. Modeling of the flexible loop was carried out using two initial structures, E·Adenosine(*syn*) and E·Adenosine(*anti*), with the *syn*- and *anti*-conformer of adenosine docked at the active site. The time-dependent backbone heavy atoms rmsd for the modeled residues of the flexible loop in E·Adenosine(*syn*) and E·Adenosine(*anti*) are 0.6 ± 0.07 and 1.9 ± 0.5 Å, respectively. The superposition of the average structures obtained from the MD simulations of E·Adenosine(*syn*) and E·Adenosine(*anti*) shows that the loop structure is dependent on the *syn*/*anti*-conformation of the nucleoside present at the active site (Figure 1). The loop structure deviates by ∼ 1 Å when the *syn*-conformer is present at the active site compared to when the *anti*-conformer is the resident of the active site. The average structures from the last 200 ps of both the simulations were used for further studies and setting up the next set of simulations, Enzyme·*anti*-Adenosine with Asp40-COOH [E(40H)·Ade(a)], Enzyme·*anti*-Adenosine with Asp40-COO$^-$ [E(40)·Ade(a)], Enzyme·*syn*-Adenosine with Asp40-COOH [E(40H)·Ade(s)], Enzyme·*syn*-Adenosine with Asp40-COO$^-$ [E(40)·Ade(s)], and Enzyme·*anti*-Inosine with Asp40-COO$^-$ [E(40)·Ino(a)].

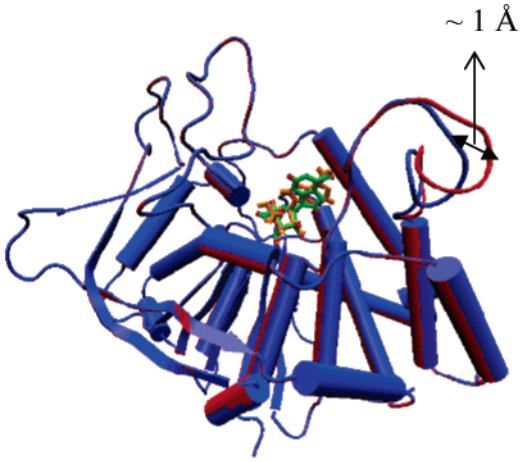


FIGURE 1: Modeling of the flexible loop. Figure showing the overlap of the average structures from the 3 ns MD simulations of E·Adenosine(*syn*) and E·Adenosine(*anti*) in TIP3P Water. The structure of the *syn*-adenosine is shown in green with the enzyme structure in blue. The structure of the *anti*-adenosine is shown in orange with the enzyme structure in red. The loop structure deviates by ∼ 1 Å when *syn*-conformer is present at the active site compared to when *anti*-conformer is the resident of the active site.

Protonation State of Asp40 at the Active Site of IAG-NH. The pK_a values for Asp40, as obtained from PB calculation, are shown in Table S1 (Supporting Information) together with the ΔG^{calc}. In the enzyme dielectric range of 2−6, the pK_a shifts ranges from 3.45 to 4.07 units. Therefore, the absolute pK_a value of Asp40 in the active site of IAG-NH would be ∼ 7.2−7.8. To rationalize the elevated pK_a value of Asp40 and its role in catalysis, further MD studies were done with both protonated as well as unprotonated Asp40 (vide infra).

Stability of Trajectory. The MD simulations of solvated E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a) systems were carried out for 4.5, 4.5, 5, 5, and 5 ns, respectively. As a measure of structural stability, rmsd's of the MD-simulated structures from the starting structures were calculated on the basis of all the backbone heavy atoms of the five systems (Figure S1, Supporting Information). The average rmsd values for the backbone (excluding hydrogen atoms) for the simulations with *anti*-adenosine, E(40H)·Ade(a) and E(40)·Ade(a), are 1.07 ± 0.09 and 1.19 ± 0.13 Å, respectively. For the simulations with *syn*-adenosine, E(40H)·Ade(s) and E(40)·Ade(s), the average backbone rmsd values are 1.22 ± 0.16 and 1.17 ± 0.08 Å, respectively. For the MD simulation with *anti*-inosine, E(40)·Ino(a), the average backbone rmsd is 1.05 ± 0.11 Å. The values fluctuate very little, indicating an equilibrated protein structure.

The *B*-factors display major peaks in the regions of residues 175−183, 248−258, and 298−305. The calculated fluctuations are compared to values obtained from experimental *B*-factors (Figure S2, Supporting Information). Overall, the *B*-factors for all five MD simulations show a good consistency with the experimental *B*-factors from the X-ray crystal structure. The temperature of the system with respect to time remains steady within 300 ± 5 K over the course of the simulations.

Octacoordination of the Calcium Cation. The average calcium to oxygen distances of the amino acids ligated to the octacoordinated calcium ion in the MD simulations of

Table 1: Comparison of the Average Calcium to Ligand Distance during the MD Simulations and X-ray Structure, B-subunit in 1HP0.pdb[a]

	X-ray	E(40H)·Ade(a)	E(40)·Ade(a)	E(40H)·Ade(s)	E(40)·Ade(s)	E(40)·Ino(a)
			Octacoordination of Calcium			
Asp10(OD2)···Cal	2.5	2.36 ± 0.09	2.43 ± 0.15	2.36 ± 0.09	2.33 ± 0.08	2.39 ± 0.08
Asp15(OD1)···Cal	2.4	2.45 ± 0.15	2.41 ± 0.16	2.45 ± 0.15	2.45 ± 0.18	2.48 ± 0.17
Asp15(OD2)···Cal	2.7	2.25 ± 0.08	2.26 ± 0.09	2.25 ± 0.08	2.26 ± 0.09	2.25 ± 0.08
Asp261(OD)···Cal	2.4	2.19 ± 0.06	2.20 ± 0.07	2.21 ± 0.07	2.20 ± 0.06	2.58 ± 0.07
Thr137(O)···Cal	2.4	2.23 ± 0.07	2.22 ± 0.07	2.23 ± 0.07	2.23 ± 0.07	2.22 ± 0.07
O2′···Cal	2.4	2.29 ± 0.08	2.30 ± 0.08	2.29 ± 0.07	2.28 ± 0.07	2.27 ± 0.07
O3′···Cal	2.4	2.54 ± 0.14	2.66 ± 0.21	2.59 ± 0.13	2.66 ± 0.17	2.56 ± 0.12
			Ribose Binding Pocket			
Wat1(OH2)···C1′	3.3	3.23 ± 0.06	3.23 ± 0.11	3.30 ± 0.06	3.23 ± 0.11	3.31 ± 0.10
Wat1(OH2)···Asp10(OD1)	2.6	2.63 ± 0.08	2.59 ± 0.08	2.63 ± 0.08	2.69 ± 0.12	2.59 ± 0.07
Wat1(OH2)···Asn186(OD1)	2.6	2.91 ± 0.15	2.81 ± 0.11	2.80 ± 0.10	2.74 ± 0.09	2.92 ± 0.12
Glu184(OE1)···O5′	2.7	2.69 ± 0.22	2.65 ± 0.11	3.21 ± 0.56	2.64 ± 0.12	3.48 ± 0.70
Asn173(ND2)···O5′	3.0	3.29 ± 0.36	3.13 ± 0.28	3.89 ± 0.53	3.31 ± 0.35	3.84 ± 0.53
Asp261(OD1)···O3′	2.6	2.59 ± 0.08	2.64 ± 0.15	2.78 ± 0.20	2.73 ± 0.17	2.58 ± 0.07
Asn186(ND2)···O3′	3.0	2.95 ± 0.12	3.03 ± 0.18	3.18 ± 0.21	3.18 ± 0.27	3.00 ± 0.15
Asn186(OD1)···O4′	4.1	4.22 ± 0.37	3.82 ± 0.50	4.27 ± 0.26	4.33 ± 0.37	4.38 ± 0.31
Asp14(OD2)···O2′	2.7	2.96 ± 0.33	3.99 ± 0.87	2.65 ± 0.11	3.84 ± 0.80	2.83 ± 0.25
Asp40(OD1)··· Asn12(ND2)	3.1	6.17 ± 0.85	6.49 ± 0.88	3.28 ± 0.87	4.71 ± 1.34	7.19 ± 0.54
			Leaving Group Pocket			
Trp83(CE3)···N9	4.3	5.97 ± 1.33	7.21 ± 1.30	4.32 ± 0.34	4.57 ± 0.32	7.36 ± 1.02
Trp260(CE3)···N9	4.2	4.46 ± 0.28	5.03 ± 0.64	5.00 ± 0.54	5.49 ± 0.29	4.69 ± 0.23
Asp40(OD1)···Phe79(CD2)	3.4	3.87 ± 0.58	4.33 ± 0.58	3.68 ± 0.26	4.42 ± 0.36	5.60 ± 0.23
Trp257(OH)···N7	4.3	3.36 ± 0.24	7.03 ± 1.06	3.98 ± 0.41	4.75 ± 0.72	3.44 ± 0.28
Trp257(HH)···N7	-	3.01 ± 0.30	6.91 ± 0.80	4.06 ± 0.57	4.44 ± 0.78	3.34 ± 0.30

[a] The distances are shown in Å. Note the similarity between the distances obtained from E(40H)·Ade(s) and X-ray structure.

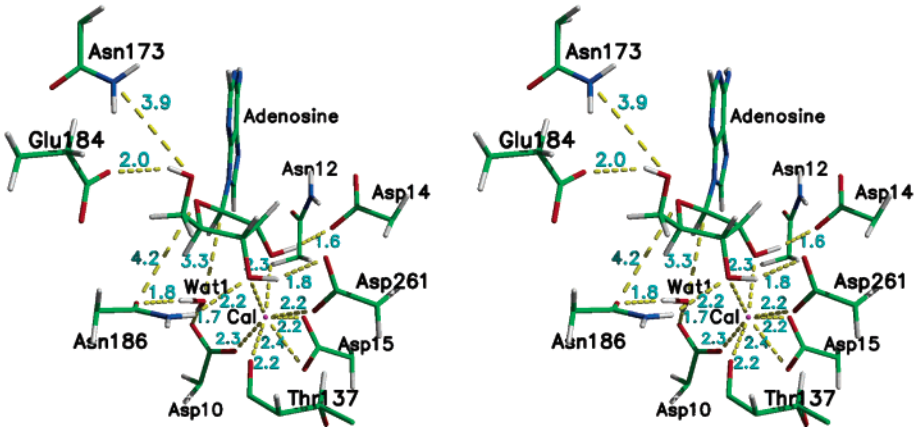


FIGURE 2: Stereoplot showing the structure from the MD simulation of E(40H)·Ade(s). The figure shows the ribose binding region in the active site including substrate adenosine, calcium (magenta sphere), and nucleophilic water (Wat1). The important distances are shown in Å in cyan color.

E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a) are close to that provided by the X-ray crystal structure of the inhibitor bound IAG-NH (Table 1). The octacoordination pattern of calcium is maintained during the MD simulations by the following ligands: Asp10, bidentate Asp15, and the backbone carbonyl oxygen of Thr137 and Asp261 (Figure 2). The remainder of the coordination sites of calcium is satisfied by 2′- and 3′-hydroxyl of the substrate adenosine and the crystallographic water, Wat1. The above calcium−oxygen distances are stable throughout the dynamics.

Ribose-Binding Pocket of IAG-NH. Several electrostatic contacts between the ribose ring of substrate and IAG-NH active site residues of E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a) are shown in Table 1. The structural findings of enzyme−substrate contacts are more easily followed by referring to the structure shown in Figure 2.

The calcium-bound nucleophilic water, Wat1, maintains a stable distance from the electrophilic carbon center C1′ of nucleoside. A hydrogen bond between Wat1 and the proposed general base, Asp10, is persistent throughout the MD simulation. A second hydrogen bond between Wat1 and Asn186 is also observed.

Enzymatic contacts with the ribose ring include hydrogen bonds with all the three hydroxyls, O5′, O3′, and O2′. The hydrogen bond between the OE1 of Glu184 and O5′ of nucleoside is found to be stable during the MD simulations (Table 1). This hydrogen-bonding distance is longer in case of E(40)·Ino(a) and E(40H)·Ade(s) compared to E(40H)·Ade(a), E(40)·Ade(a), and E(40)·Ade(s). The O5′ of nucleoside is also found to hydrogen-bond to the amide oxygen of Asn173. A short hydrogen bond (∼ 2.6 Å) between the carboxylate group, Asp261, and the O3′ of adenosine is observed during all five MD simulations. The residue amide group of Asn186 maintains a hydrogen-bonding distance

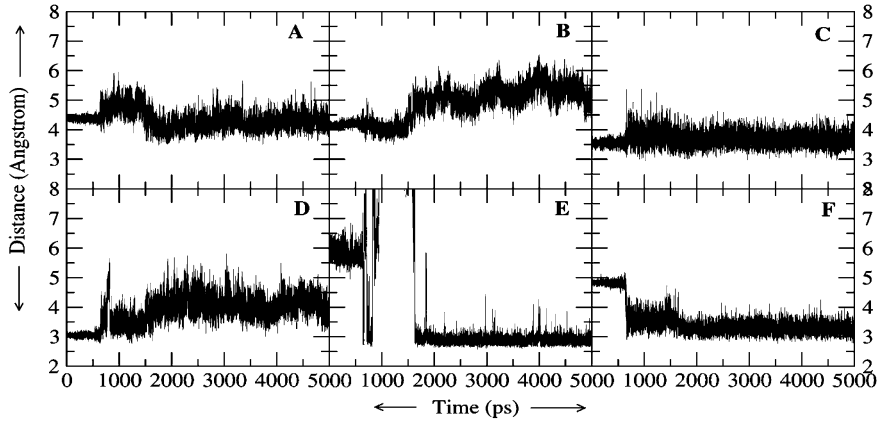


FIGURE 3: Distance trajectories from the MD simulation of E(40H)·Ade(s). The plots correspond to the following distances: Trp83 (CE3)−Adenosine (N9) (A), Trp83 (CE3)−Adenosine (N9) (B), Asp40 (OD1)−Phe79 (CD2) (C), Tyr257 (OH)−Adenosine (N7) (D), Solv6780 (OH2)−Adenosine (N7) (E), and Asn12 (ND2)−Adenosine (C8) (F).

from the O3′ of the ribose ring. The OD1 of Asn186 is located adjacent to O4′ of the ribose ring. The O2′ of the nucleoside is hydrogen-bonded to Asp14 throughout the MD simulations of E(40H)·Ade(a), E(40H)·Ade(s), and E(40)·Ino(a) and is found to be longer in case of E(40)·Ade(a) and E(40)·Ade(s). The Asp40-COOH is also found to make contact with the amide group of Asn12 in E(40H)·Ade(s). This interaction is not observed during the other MD simulations.

Conformational Change in the Ribose Ring. The ribose ring of the nucleoside undergoes a conformational change to O4′-*endo* during the MD simulations of E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a) (Figure S3, Supporting Information). Thereafter, this conformation is maintained for the rest of the simulation time. The average phase angles in E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a) correspond to 100.24 ± 27.46, 102.87 ± 13.48, 104.57 ± 25.11, 107.85 ± 9.66, 106.72 ± 16.02 deg, respectively.

Leaving Group Pocket of IAG-NH. The nucleoside base shows specific interaction with the leaving group pocket (Table 1). The nucleobase is stacked between two tryptophan residues, Trp83 and Trp260. The average distance between Trp83 and the nucleobase exhibits a wider range in the MD simulations of E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a). This average distance is shortest in case of E(40H)·Ade(s). The average distance between Trp260 and the nucleobase exhibits greater stability (comparing the average deviations in Table 1) and does not vary much in all the five MD simulations as compared to the Trp83−nucleobase distance. The important distance trajectories obtained from the MD simulation of the E(40H)·Ade(s) model is shown in Figure 3.

The Asp40-COO(H) is located in the base of the leaving group pocket. When Asp40 is present as Asp40-COOH in E(40H)·Ade(a) and E(40H)·Ade(s), it is shielded from the solvent sphere by the loop−helix consisting of residue 76−85, which includes the tryptophan residue (Trp83) involved in base-stacking interaction with the nucleoside base (Figure 4). The bulky phenyl group of Phe79 moves away when Asp40 is present as Asp40-COO$^-$. In the models E(40)·Ade(s) and E(40)·Ade(a), a water opening is formed during the simulation joining Asp40-COO$^-$ to the solvent sphere (Figure 5). In the MD simulations of E(40H)·Ade(a) and

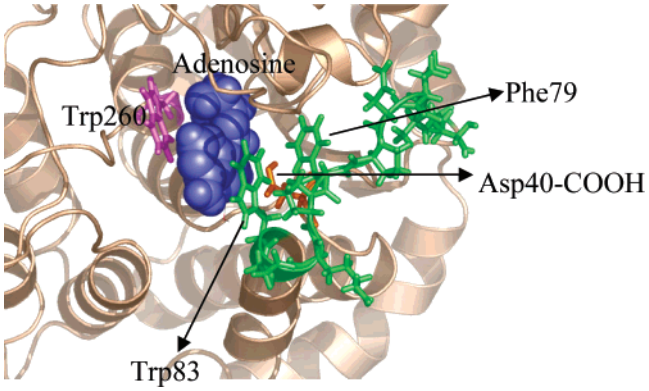


FIGURE 4: Figure showing the location of Asp40-COOH (drawn in orange sticks) in the leaving group pocket during the MD simulation of E(40H)·Ade(a). Asp40-COOH is shielded from the solvent sphere by the loop−helix consisting of residue 76−85 (shown in green), which includes the tryptophan residue (Trp83) involved in base stacking interaction with the nucleoside base. The hydrophobic residue, Phe79, is also shown as a part of this loop. The substrate adenosine is shown in blue.

E(40H)·Ade(s), the residue Asp40-COOH is fully screened from the solvent.

To quantitate the solvent accessible surface area (SASA) of Asp40 carboxylate oxygens throughout the MD simulation, NACCESS program with a probe radius of 1.4 (radius of water) is used (Figure S4, Supporting Information). The average SASA for the carboxylate group of Asp40 is 0 (no contact with a water molecule) throughout the MD simulations of the systems with Asp40-COO$^-$, E(40)·Ade(a), E(40)·Ade(s), and E(40)·Ino(a), and >0 for the systems with Asp40-COOH, E(40H)·Ade(a) and E(40H)·Ade(s).

The N7 of the nucleobase is hydrogen-bonded to a water molecule throughout the MD simulations in all the five models. The hydrogen-bonded water molecule is part of a water channel that reaches to the N7 and is exchanged with the other solvent molecules rapidly. However, in case of E(40H)·Ade(s), the N7 is hydrogen-bonded to a water molecule from solvent pool, which is not exchanged with any other water molecule during the course of the simulation (Figure 3E). Another possible hydrogen bond donor is Tyr257. A movement by about ∼ 3 Å brings the Tyr257-containing flexible loop closer to the N7 of the nucleobase. However, the distance of the hydrogen on Tyr257 is not close

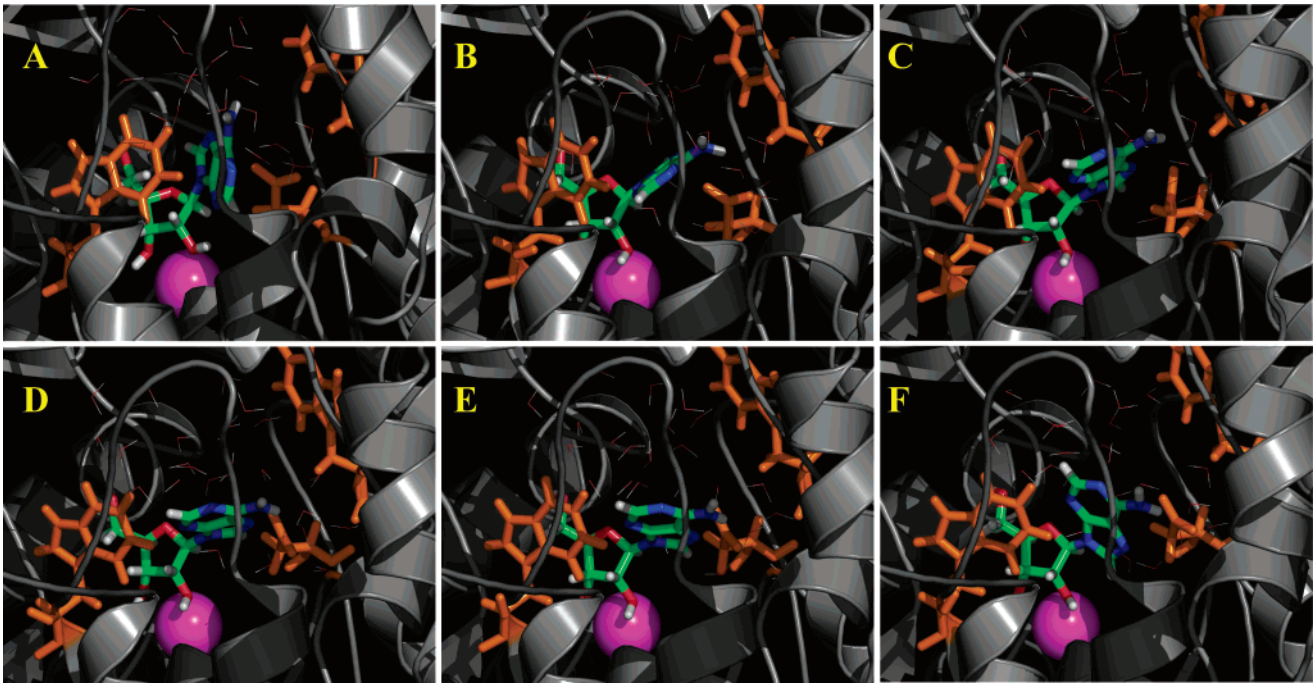


FIGURE 5: Figure showing rotation around the χ-torsion (O4′−C1′−N9−C4) of adenosine. The snapshots were taken from the MD simulation of Enzyme·*anti*-Adenosine with Asp40-COO^-, E(40)·Ade(a). Each panel corresponds to a different simulation time. The starting structure for the MD simulation has adenosine in anti, which undergoes gradual change to *syn*-conformation. The two tryptophans are shown in orange stick diagram. The calcium is shown in magenta sphere.

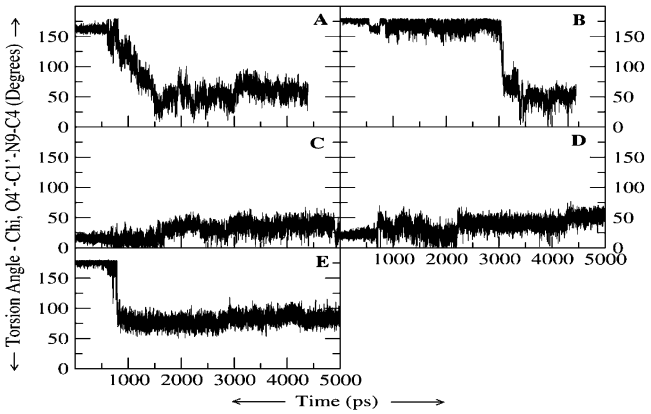


FIGURE 6: Plot showing the time dependent change in glycosyl Chi(χ)-torsion of the purine ring at the active site of IAG-NH. The results were obtained from the MD simulation of (A) E(40H)·Ade(a), (B) E(40)·Ade(a), (C) E(40H)·Ade(s), (D) E(40)·Ade(s), and (E) E(40)·Ino(a).

enough for a direct proton transfer to N7 of the nucleobase (R_{HH}···N7 > 3 Å) (Table 1).

The MD simulations of E(40H)·Ade(a), E(40)·Ade(a), and E(40)·Ino(a) were started with nucleobases in *anti*-conformation which gradually isomerized to *syn*-like structures (Figure 6). The χ-torsion, used to differentiate the *anti*- and *syn*-conformers, is defined as the torsion angle joining four atoms of the nucleobase in the order: O4′−C1′−N9−C4. The average χ-torsion in the MD simulations started with *anti*-nucleoside in E(40H)·Ade(a), E(40)·Ade(a), and E(40)·Ino(a) are 59.35 ± 24.16, 51.32 ± 14.04, and 80.42 ± 9.33 deg, respectively (Figure 7). The average χ-torsion in the MD simulations started with *syn*-nucleoside in E(40H)·Ade(s) and E(40)·Ade(s) are 27.72 ± 18.82 and 32.67 ± 16.16 deg, respectively.

Correlated Motion and Normal-Mode Analysis. Here, we have searched for a coupling between protein dynamics and enzyme catalysis. A direct coupling might, for instance, be through the existence of promoting vibrational modes. A model that includes such a coupling is attractive because it can explain how a local structural change can have a global impact through perturbation of normal modes. A single normal mode generally includes the motion of all of the atoms in the enzyme, and a perturbation of such a mode will therefore be of global character.

We present here two different computational approaches with different levels of resolution and on different time-scales to understand the structuring of the flexible loop at the active site and to analyze the dynamics of this part of the protein. First, we quantitate the coupling between two residues by calculating the covariance between the fluctuations of the two residues from the MD trajectory. Figure 8 shows the covariance matrix for the Cα atoms in E(40H)·Ade(s). Second, we employ normal-mode analysis (NMA) based on elastic network model to get insights, at the atomic level, on the mechanism of large-scale rearrangements of IAG-NH that occur upon ligand binding. Such large-scale motions are difficult to observe using classical MD simulation as they occur at much longer time scales.

The covariance matrix generated from the MD simulation of E(40H)·Ade(s) shows the presence of local motions between active site residues (Figure 8A). The residue Asp40-COOH shows positive correlation with residue Trp83 for 48% of the average time. Trp83 is a part of the helix−loop that shields Asp40-COOH from the solvent (Figure 8B). The residues that show substantial (>0.25) positive correlation coefficient with the catalytic residue Trp260 are 134−138, 214−217, and 314−321. The flexible loop containing residues 247−257 shows substantial positive correlation with two helixes 99−101 and 150−157 which are >20 Å from it. The residues that show negative correlation (Figure 8C) with Asp40-COOH are 174−182, 220−246, and 287−293.

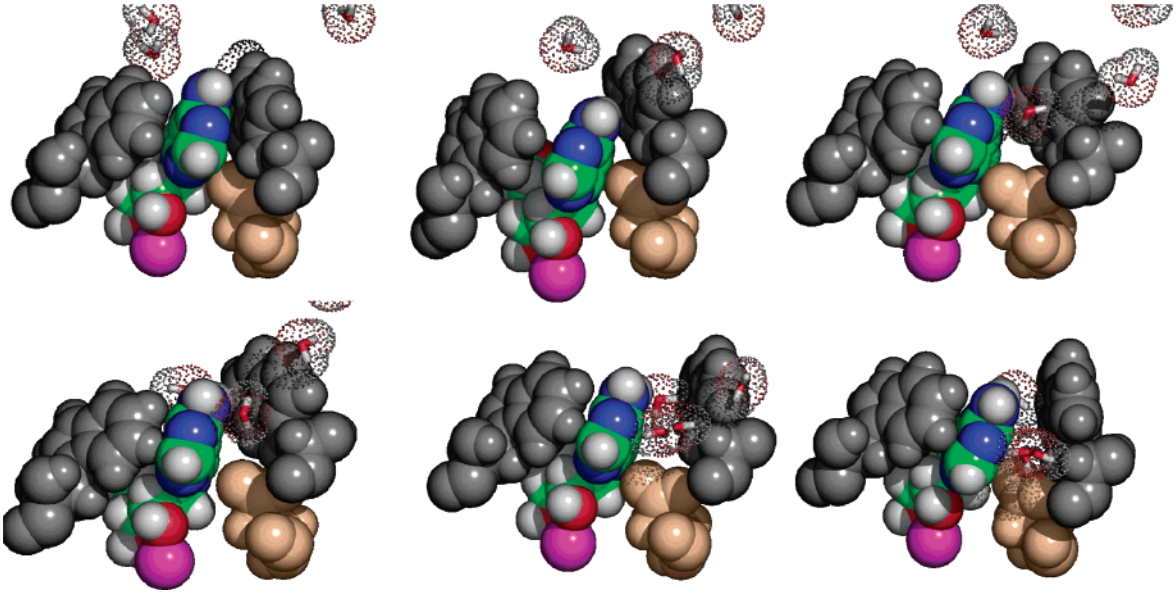


FIGURE 7: Figure showing snapshots of water entry at the active site using a CPK model. The snapshots were taken from the MD simulation of Enzyme•*syn*-Adenosine with Asp40-COO^-, E(40)•Ade(s). The two tryptophan residues (Trp83 and Trp260) are shown in gray, Asp40-COO^- in wheat, and calcium ion in magenta. The substrate and water atoms are colored as carbon (green), nitrogen (blue), oxygen (red), and hydrogen (white). The waters are shown in stick diagram with dot surface representating their van der Waals surface.

The base-stacking residue Trp83 shows anti-correlated motion with two other residues 195 and 226 for ~ 20% of time. The helix 54−65 shows anti-correlation with the flexible loop 242−257. For example, the residue 62 shows anti-correlation with 247 for 42% of simulation time. The residues 186−195 also show negative correlation with residues 242−257. The residues 262−269 show negative correlation with the residues 58−64.

The starting structure for NMA is taken from the coordinates generated from MD simulation of E(40H)•Ade(s), averaged over the final 300 ps duration, followed by minimization with tolerance gradient of $1e^{-11}$. Elastic network model was chosen for NMA calculations since our primary goal was to study the character (directions) of the low-frequency motion and not determine the absolute scale of normal-mode eigenvalues (frequencies of the motion). For our analysis, we have focused primarily on the large-amplitude/low-frequency normal modes, which might be relevant to function.

The normal modes obtained were visualized individually using the program VMD (*31*). The frequencies of the first six normal modes were close to zero and correspond to the rotational and translational modes. Using NMA, we find that the flexible loop explores a conformational space much larger than in the MD trajectory, leading to a "gating"-like motion with respect to the active site (Figure 9). One consequence of this motion is the closure of the cleft between the two domains in which the substrate binding site lies. This closure of the active site brings the flexible loop containing the residues Cys245, His247, Tyr257, and Tyr258 closer to the leaving group of the nucleobase. Such a motion can possibly help to structure the loop over the active site and transfer a proton to the water channel that hydrogen-bonds to the N7 atom of the nucleobase.

DISCUSSION

The nucleoside hydrolase (NH) catalyses the hydrolysis of the *N*-glycosidic bond between the anomeric carbon atom of ribose and the nucleobase (Scheme 1). On the basis of the substrate specificity, NH has been divided into three subgroups: inosine-uridine (IU), inosine-adenosine-guanosine (IAG), and guanosine-inosine (GI). The IU-NH has been extensively studied experimentally by Schramm and co-workers, and computationally by our group (*8*, *10*). Involvement of a ribooxocarbenium ion-like transition state in the catalytic mechanism has been proposed for both IU-NH and IAG-NH. As a continuation of our work on nucleoside hydrolase enzyme, here we discuss results from five long-term MD simulation studies on IAG-NH: Enzyme•*anti*-Adenosine with Asp40-COOH [E(40H)•Ade(a)], Enzyme•*anti*-Adenosine with Asp40-COO^- [E(40)•Ade(a)], Enzyme•*syn*-Adenosine with Asp40-COOH [E(40H)•Ade(s)], Enzyme•*syn*-Adenosine with Asp40-COO^- [E(40)•Ade(s)], and Enzyme•*anti*-Inosine with Asp40-COO^- [E(40)•Ino(a)].

The octacoordination pattern of calcium in IAG-NH is very similar to that of IU-NH (Chart 1). The ribose binding pocket is essentially superimposable in all five MD simulations. The enzymatic residues that coordinate to calcium are Asp10, Asp15 (bidentate), Thr137, and Asp261. The rest of the octacoordination is satisfied by the 2′-OH and 3′-OH of substrate and a crystallographic water molecule, Wat1. The distances between calcium and its coordinating ligands remain steady throughout the MD simulations (Table 1). The crystallographic water, Wat1, maintains a steady and a reasonable distance (~ 3 Å) for a nucleophilic attack on the C1′ of the substrate (adenosine or inosine) and is held in place by two hydrogen bonds to the enzymatic residues, Asp10 and Asn186. The residue Asp10 is proposed as the general base that abstracts the proton from Wat1 for the nucleophilic attack (*13*).

Enzymatic contacts with all three ribose hydroxyl groups were observed. The hydrogen bond between the ribosyl O2′ and Asp14-COO^- is present in all five MD simulations. The O3′ is held strongly by a short hydrogen-bonding interaction with Asp261. The O3′ is also hydrogen-bonded to the amide −NH of Asn186. The O5′ is found to hydrogen-bond to the

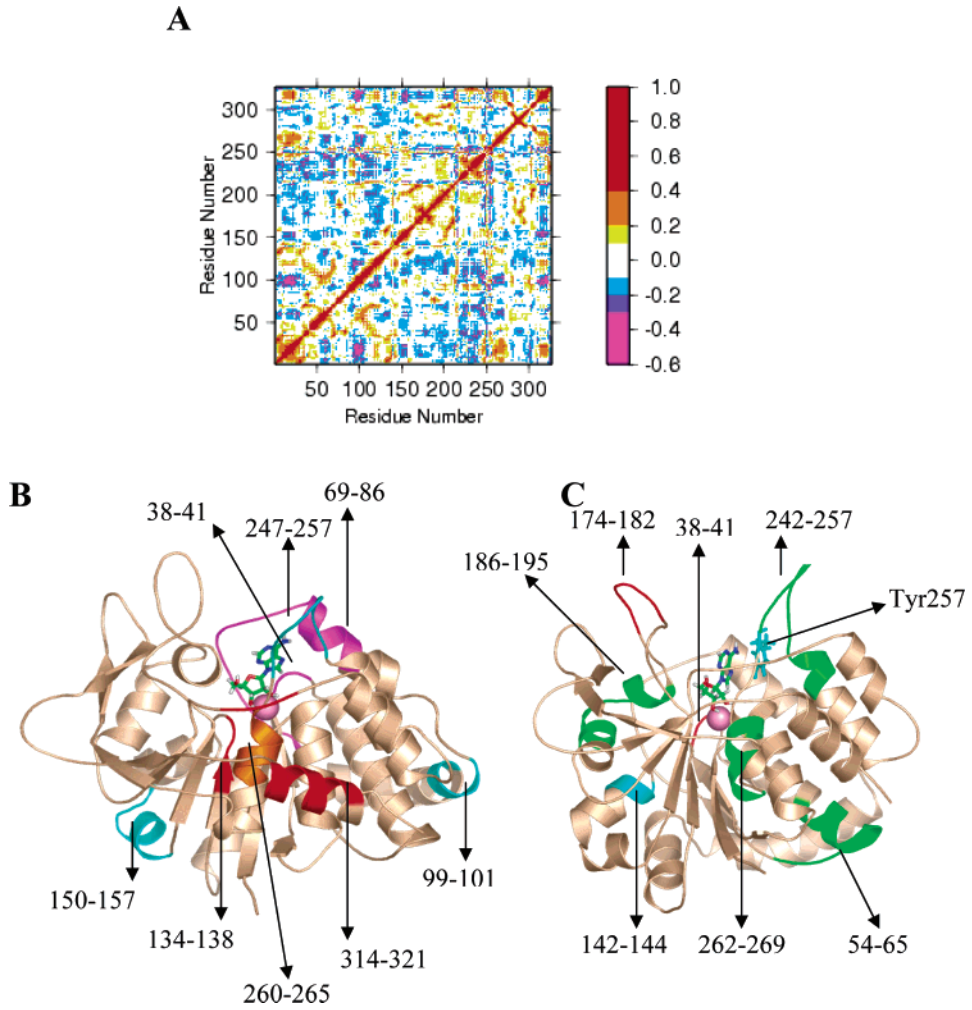


FIGURE 8: Correlation of motion. (A) Covariance matrix generated from the production phase of the MD simulation of E(40H)·Ade(s) system. The scale showing the positive and negative coefficient of correlation is shown on the right. (B and C) The position of the residues with respect to the active site from the MD simulation of E(40H)·Ade(s). Positions of some of the residues displaying positive and negative correlations are shown in B and C, respectively. The residues showing correlated motions with each other are marked with similar colors. For example, in B, the residues 38−41 and 69−86 show anti-correlated motion with each other and are marked with magenta. The calcium is shown in magenta sphere, and the substrate adenosine is shown in green stick diagram.

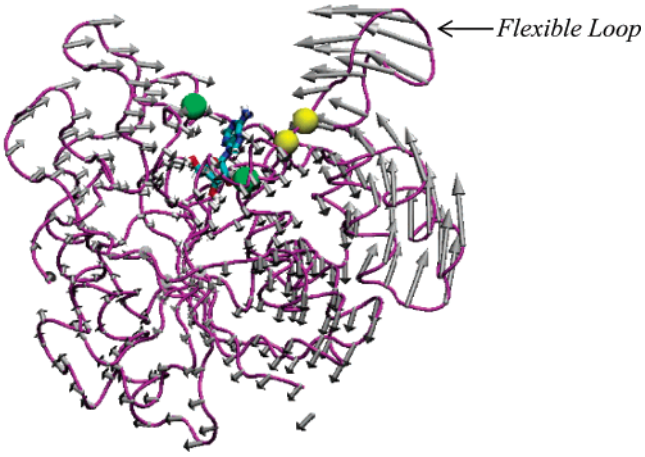
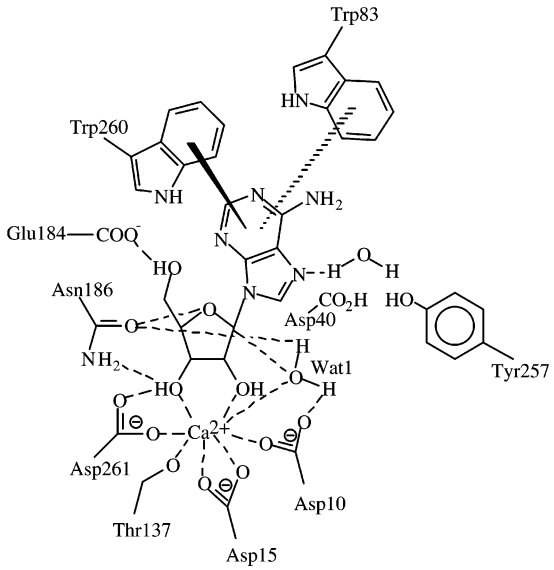


FIGURE 9: Normal-mode analysis on E(40H)·Ade(s) system using elastic network model. Each arrow represents the direction vector for motion of the corresponding enzymatic residues. The two tryptophans, Trp83 and Trp260, are shown in green spheres. The yellow spheres represent Tyr257 and Tyr258. The substrate adenosine is shown in sticks diagram.

Chart 1: Schematic Representation of the Active Site of IAG-NH during MD Simulations



carboxylate of Glu184. This hydrogen bond is found to be a little more flexible in MD simulations of E(40H)·Ade(s) than the other systems. The O5′ is also hydrogen-bonded to the side-chain amide of Asn173.

The amide group of Asn186 is within 4−4.3 Å from the O4′ of the ribose ring of the substrate in all five MD simulations. The residue, Asn168, occupies a similar location in the IU-NH active site in previous MD studies *(10)*. In this published work, the average structures of the ground state- and the transition state-bound IU-NH were used to link the residue Asn168 to transition state stabilization. The residue Asn168 migrated closer to the O4′ in the ribooxocarbenium transition state than in the ground state in IU-NH enzyme. Since Asn186 in IAG-NH occupies a similar position in the active site compared to Asn168 in IU-NH, it may also be looked upon as a catalytic residue contributing toward transition state stabilization. Such an electrostatic interaction contributing toward the stabilization of the ribooxocarbenium ion TS may be looked upon as a general feature for this class of enzymes.

The difference in the MD structures primarily arises from the leaving group pocket. In the MD simulations with E(40H)·Ade(a) and E(40)·Ade(a), the starting structure has the purine nucleobase in an *anti*-conformation (Figure 7). The *anti*-nucleobase in both E(40H)·Ade(a) and E(40)·Ade-(a) undergoes conformational change to adopt a *syn*-conformation during the simulations (Figure 6). The *anti*-adenosine is favored over the *syn*-conformation in aqueous solution by about 4−5 kcal/mol (*32*). However, even with the presence of a moderate energy barrier, both *syn*- and *anti*-forms are present in dynamic equilibrium in aqueous solution. From our MD simulations, it would appear that the enzyme can bind both *syn*- and *anti*-conformations of its substrate. However, it prefers to bind the substrate in *syn*-conformation. When the *anti*-conformation is bound to the active site, it gradually changes to a more *syn*-like conformation (Figure 6).

Large isotope effects for the 5′-^{3}H [1.051] of inosine have been reported previously for this class of enzymes (*7*). Such a high isotope effect is surprising since C5′ is four bonds removed from the reaction center where the bond breaking and making happens. A high isotope effect at 5′-^{3}H may arise from a net loosening of the C5′−H5′ bond in the transition state. This, in turn, could be realized by decreasing the C5′−H5′ bond or by placing H5′ in a less crowded environment in the transition state. A plausible explanation to this high 5′-^{3}H isotope effect can be given on the basis of the present MD studies on IAG-NH. The syn/anti conformational change resulting from the χ-torsion angle variation between the bound and free nucleoside can affect the environment of the H5′. Also, in the *syn* ground state, the environment of H5′ will be more crowded than the transition state. At the transition state, the elongation of the C1′−N9 bond in *syn*-adenosine would dispose the nucleobase away from the H5′, thus making its environment less crowded (Scheme 1). Recently, Schramm et al. have shown that in purine nucleoside phosphorylase enzymes, half of the remote 5′-^{3}H isotope effect is due to the χ-torsion and the other half is due to the transition state distortion (*33*). The ricin toxin chain-A and nucleoside hydrolase both catalyze similar reaction via the ribooxocarbenium ion TS. It has been proposed recently that the glycosidic bond of the susceptible adenosine in the substrate is anti and a base flipping on binding to ricin toxin A-chain would dispose it in a *syn*-orientation prior to catalysis (*34*). Similarly, it can be proposed that even if the substrate is bound in an *anti*-conformation, the IAG-NH active site would dispose it to syn prior to catalysis.

To find out whether the syn/anti conformational change is a feature specific to the nature of the substrate at the active site, we also performed a MD simulation when a different substrate, inosine, is bound to the IAG-NH active site. The results from the MD simulation of Enzyme·Inosine complex, E(40)·Ino(a), also support the similar observation as obtained with Enzyme·Adenosine complexes. The starting structure for the MD simulation of E(40)·Ino(a) has inosine in *anti*-conformation. However, during the MD simulation, a gradual conformational change disposes the nucleobase in a more *syn*-like orientation. Therefore, the MD simulation of E(40)·Ino(a) supports that this syn/anti conformational change is not nucleobase-specific, but rather a general feature of the active site of IAG-NH enzyme. Computational study on the TS-bound IAG-NH can be done as a future work to validate this observation.

On the basis of the X-ray crystal structures of IAG-NH, it was proposed that Asp40 is positioned to act as a general acid to protonate the purine leaving group. However, replacing Asp40 with a neutral residue, Ala, did not have any major effect on k_{cat} and a 3-fold increase in K_m [k_{cat} (wild-type) = 20.81 ± 0.63 s^{-1}, K_m (wild-type) = 35.71 ± 3.77 μM; k_{cat} (Asp40Ala) = 12.16 ± 0.91 s^{-1}, K_m (Asp40Ala) = 113.30 ± 21.38 μM], (*13*) making Asp40's role in catalysis questionable. Keeping these observations from the prior studies in mind, we designed a computational experiment to calculate the pK_a of Asp40 at the active site of IAG-NH.

It is well-known that the microenvironment of enzyme can modulate pK_a of ionizable groups corresponding to aqueous systems (*35*). To find out the effect of the protein microenvironment on the proton affinity of Asp40, we calculated the energies for both Asp40-COO^- and Asp40-COOH models, using the state-of-the-art PB equation module for continuum electrostatics to describe the solvation of the enzyme. The absolute pK_a value of Asp40 is calculated to be ∼ 7.2−7.8 in the enzyme active site in the dielectric range of 2−6. Aspartates with elevated pK_a have been reported earlier for enzymes such as ketosteroids isomerase (*36*) and human thioredoxin (*37*). To rationalize the elevated pK_a value of Asp40 and its role in catalysis, the structures generated from the MD studies with both protonated as well as unprotonated Asp40 were scrutinized.

The solvent accessible surface area calculations with the NACCESS program (Figure S4, Supporting Information) using a probe radius of 1.4 Å show that the neutral Asp40-COOH is virtually inaccessible to the solvent (Figure S5, Supporting Information) compared to the negatively charged Asp40-COO^- at the enzyme active site. During the MD simulations of E(40)·Ade(a) and E(40)·Ade(s), when Asp40 is present as Asp40-COO^-, the loop−helix consisting of residue 76−85 drifts away slightly exposing the electronegative carboxylate oxygens of Asp40-COO^-, and as a result, an opening to water occurs at the active site (Figure 5). This water opening is formed right between Trp83 and the nucleobase, resulting in the distancing of Trp83 from the nucleobase by about 0.6−0.9 Å. However, during the MD simulations of E(40H)·Ade(a) and E(40H)·Ade(s), when Asp40 is present as Asp40-COOH, the phenyl group of Phe79 screens the neutral Asp40-COOH perfectly (Figure 4). As a result, the water opening is not formed. Thus, the

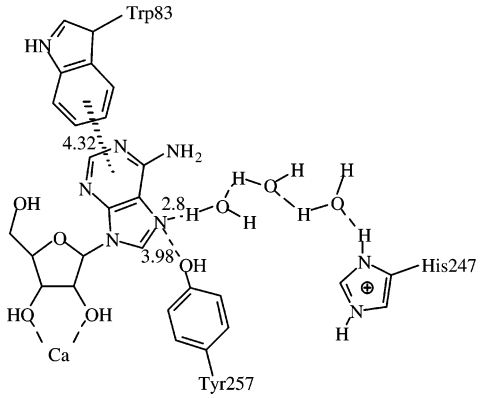
average distance of Trp83 is closer to the nucleobase in the MD simulations of E(40H)·Ade(a) and E(40H)·Ade(s) compared to the MD simulations of E(40)·Ade(a) and E(40)·Ade(s). This, in turn, may explain why the catalytic activity is not affected by Asp40Ala mutation. Upon mutation of Asp40 to a neutral residue (Asp40Ala mutant), it resembles the systems E(40H)·Ade(a) and E(40H)·Ade(s), where the neutral Asp40-COOH form is present. It is likely that, in the mutant enzyme, the water opening separating the nucleobase and Trp83 is not formed. Hence, the catalytic rate constant remains unaffected by the mutation. Even though Asp40-COOH is a good proton donor, it is not involved in any hydrogen-bonding interaction with the ring nitrogens of the nucleobase in the ground state. Therefore, from the present MD studies, Asp40-COOH may be ruled out as a possible general acid. However, Asp40-COOH might play an important role in substrate binding by maintaining the overall electrostatic complementarities (Figure S6, Supporting Information) between the active site of IAG-NH and the substrate without negotiating the catalytically crucial base stacking interaction between Trp83 and the nucleobase of the substrate.

During the catalytic cycle, the flexible loop of the enzyme cycles through different conformations. The conformation of the loop is dependent on the conformation (syn vs anti) of the substrate at the active site (Figure 1). The NMA shows that the loop exhibits significant "gating" motion (Figure 9). The low-frequency vibrational motion in the flexible loop disposes the residues of this loop, such as Tyr257, Tyr258, His247, and Cys245, closer to the N7 of the substrate. Such a motion can help to structure the loop over the active site and possibly transfer a proton to the water channel that hydrogen-bonds to the N7 atom of the nucleobase.

This brings us to the question: What is the general acid that protonates the leaving group ("AH" in Scheme 1) at N7? The experimentally determined k_{cat} versus pH profile shows that deprotonation of a first group with a pK_{app} of 5.6 $\pm$ 0.2 leads to a decrease in k_{cat} by a factor of 10, and deprotonation of a second group with a pK_{app} of 8.6 $\pm$ 0.2 completely abolishes the IAG-NH enzymatic activity; the later may correspond to the required general acid (*13*). A clever computational work by Steyaert et al. suggests that the base stacking interaction between Trp260 and the nucleoside base of the substrate might play a crucial role in modulating the electron distribution of the nucleoside leaving group (*15*). This base stacking interaction is proposed to be involved in electronically activating the leaving group. But there is still a need of a proton donor to neutralize the negative charge of the leaving group. In case of the MD simulation studies on IU-NH, a Tyr residue has been proposed to act as the general acid (*10*). In the present MD simulations of IAG-NH, Tyr257, a member of the flexible loop, is in contact with the N7 of the nucleobase occasionally. However, on examination of the average distance between the phenolic hydrogen (HH) of Tyr257 and N7 ($\sim$ 4 Å), the direct proton transfer seems to be unlikely. This observation explains previous experimental results, where the Tyr257Ala mutant did not show any considerable drop in the steady-state kinetic constants, k_{cat} and k_m (*14*).

Apart from occasional contacts with the Tyr257 residue, the N7 of the nucleobase is also hydrogen-bonded to a water molecule (crystallographic or solvent) throughout the MD simulations (Figure 3E). The water molecule is a part of a channel that reaches to the N7 of the nucleobase. The water-channel wall is lined with polar residues from the flexible loop, such as Cys245, His247, Tyr257, and so forth (Chart 2). Alternative to an enzymatic general acid, this water molecule can also act as a proton donor with assistance from the polar enzymatic residues of the water channel and transfer a proton to the N7 of the electronically activated nucleobase. Therefore, a mechanism can be proposed wherein the N7 abstracts a proton from the solvent water that is hydrogen-bonded to it, which in turn is assisted by other solvent molecules of the water channel and the enzymatic multiple proton donor sources from the flexible loop, such as His245, Tyr257, Tyr258, and so forth.

Chart 2: Role of the Water Channel and the Residues Surrounding It



The Trp83Ala mutation increases the value of K_M for guanosine by 3 orders of magnitude, but has a negligible effect on the catalytic step in IAG-NH (*13*). A possible explanation for no effect on k_{cat} can be explained from our MD studies. The nucleoside base rotates from anti to syn from the side facing Trp83 (Figure 6). On mutation of Trp83 to Ala, the steric strain experienced by the nucleoside base during this rotation is released. However, the catalytic rate does not increase upon mutation, and a possible explanation can be the absence of the energetically favorable stacking interaction between Trp83Ala mutant and nucleoside base.

CONCLUSION

The catalytic mechanism of IAG-NH from *T. vivax* has been computationally studied using five MD simulations on the Enzyme·Substrate complexes: Enzyme·*anti*-Adenosine with Asp40-COOH [E(40H)·Ade(a)], Enzyme·*anti*-Adenosine with Asp40-COO^- [E(40)·Ade(a)], Enzyme·*syn*-Adenosine with Asp40-COOH [E(40H)·Ade(s)], Enzyme·*syn*-Adenosine with Asp40-COO^- [E(40)·Ade(s)], and Enzyme·*anti*-Inosine with Asp40-COO^- [E(40)·Ino(a)]. We have now described the positioning of the residues of an important flexible loop at the active site, which was previously unobservable by X-ray crystallography due to high *B*-factors. Overall, the structures generated from the MD simulation of E(40H)·Ade(s) (with *syn*-adenosine and Asp40-COOH) preserves the catalytically important hydrogen bonds as well as the electrostatic and hydrophobic interactions crucial for the formation of the competent Michaelis complex. Throughout the simulation, adenosine maintains *syn*-conformation. The base-stacking interaction is maintained between the

nucleobase and the two active site catalytic tryptophan residues, Trp83 and Trp260, similar to that observed in the X-ray crystal structures. A solvent water molecule from a water channel is found to be persistently hydrogen-bonded to the N7 of the nucleobase in the structures obtained from the MD simulation of E(40H)·Ade(s) and is in position to play a crucial role in the general acid-catalyzed departure of the leaving group. A similar water channel is present in all the MD simulations and is lined by the residues from the flexible loop involving Tyr257, His247, Cys245, and so forth. A normal-mode analysis using elastic network model identifies one single normal-mode vibration in the protein that can promote catalysis by a "gating" motion of the flexible loop. A water opening is formed in the structures with Asp40-COO^- [E(40)·Ade(a), E(40)·Ade(s), and E(40)·Ino(a)], and this opening reaches to Asp40-COO^-. This water opening interferes with the base-stacking interaction between the nucleobase and the residue, Trp83. When Asp40 is modeled as Asp40-COOH in E(40H)·Ade(s), it has zero solvent accessible surface area, and the water opening separating Trp83 from the nucleobase is not formed. The elevated pK_{app} of Asp40-COOH, as predicted from the MD simulations, is also supported by a set of Poisson−Boltzmann calculations. The preferred conformation of the nucleobase is found to be *syn*-like in the structures generated from the five MD simulations. The *anti*-nucleobase isomerizes to its corresponding *syn*-conformers during the MD simulations.

ACKNOWLEDGMENT

The authors acknowledge National Partnership for Advanced Computational Infrastructure (NPACI) for their generous allocation of the computational resources at the AMD clusters at the University of Michigan Supercomputing Center. D.M.-S. thanks Research Workshop titled "Theory and Computation in Molecular Biological Physics" organized by the center for Theoretical Biological Physics, La Jolla, CA. The authors thank Dr. Swarnalatha Y. Reddy for reading the manuscript and providing helpful comments.

SUPPORTING INFORMATION AVAILABLE

The pK_{app} of Asp40-COOH as calculated from Poisson−Boltzmann (PB) equation module in CHARMM forcefield; backbone heavy atoms rmsd obtained from the MD simulation, comparison of the positional fluctuations of the Cα atoms of subunit B from the X-ray and the average structures from MD simulations, plot showing conformational change in the ribose ring during MD simulations, and SASA throughout the MD simulations of E(40H)·Ade(a), E(40)·Ade(a), E(40H)·Ade(s), E(40)·Ade(s), and E(40)·Ino(a); structural alignment of Asp40-COOH in E(40H)·Ade(s); and electrostatic potential at the active site of E(40H)·Ade(s). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Hammond, D. J., and Gutteridge, W. E. (1984) Purine and pyrimidine metabolism in the Trypanosomatidae, *Mol. Biochem. Parasitol. 13*, 243−261.
2. Berens, R. L., Krug, E. D., and Marr, J. J. (1995) *Purine and Pyrimidine Metabolism*, Academic Press, London, U.K.
3. Parkin, D. W., Horenstein, B. A., Abdulah, D. R., Estupinan, B., and Schramm, V. L. (1991) Nucleoside hydrolase from *Crithidia fasciculata*. Metabolic role, purification, specificity, and kinetic mechanism, *J. Biol. Chem. 266*, 20658−20665.
4. Parkin, D. W. (1996) Purine-specific nucleoside *N*-ribohydrolase from *Trypanosoma brucei* brucei. Purification, specificity, and kinetic mechanism, *J. Biol. Chem. 271*, 21713−21719.
5. Versees, W., Decanniere, K., Pelle, R., Depoorter, J., Brosens, E., Parkin, D. W., and Steyaert, J. (2001) Structure and function of a novel purine specific nucleoside hydrolase from *Trypanosoma vivax*, *J. Mol. Biol. 307*, 1363−1379.
6. Estupinan, B., and Schramm, V. L. (1994) Guanosine-inosine-preferring nucleoside *N*-glycohydrolase from *Crithidia fasciculata*, *J. Biol. Chem. 269*, 23068−23073.
7. Horenstein, B. A., Parkin, D. W., Estupinan, B., and Schramm, V. L. (1991) Transition-state analysis of nucleoside hydrolase from *Crithidia fasciculata*, *Biochemistry 30*, 10788−10795.
8. Schramm, V. L. (2003) Enzymatic transition state poise and transition state analogues, *Acc. Chem. Res. 36*, 588−596.
9. Gopaul, D. N., Meyer, S. L., Degano, M., Sacchettini, J. C., and Schramm, V. L. (1996) Inosine-uridine nucleoside hydrolase from *Crithidia fasciculata*. Genetic characterization, crystallization, and identification of histidine 241 as a catalytic site residue, *Biochemistry 35*, 5963−5970.
10. Mazumder, D., and Bruice, T. C. (2002) Exploring nucleoside hydrolase catalysis in silico: molecular dynamics study of enzyme-bound substrate and transition state, *J. Am. Chem. Soc. 124*, 14591−14600.
11. Mazumder, D., Kahn, K., and Bruice, T. C. (2002) Computer simulations of trypanosomal nucleoside hydrolase: determination of the protonation state of the bound transition-state analogue, *J. Am. Chem. Soc. 124*, 8825−8833.
12. Mazzella, L. J., Parkin, D. W., Tyler, P. C., Furneaux, R. H., and Schramm, V. L. (1996) Mechanistic diagnoses of *N*-ribohydrolases and purine nucleoside phosphorylase, *J. Am. Chem. Soc. 118*, 2111−2112.
13. Versees, W., Decanniere, K., Van Holsbeke, E., Devroede, N., and Steyaert, J. (2002) Enzyme−substrate interactions in the purine-specific nucleoside hydrolase from *Trypanosoma vivax*, *J. Biol. Chem. 277*, 15938−15946.
14. Vandemeulebroucke, A., Versees, W., De Vos, S., Van Holsbeke, E., and Steyaert, J. (2003) Pre-steady-state analysis of the nucleoside hydrolase of *Trypanosoma vivax*. Evidence for half-of-the-sites reactivity and rate-limiting product release, *Biochemistry 42*, 12902−12908.
15. Versees, W., Loverix, S., Vandemeulebroucke, A., Geerlings, P., and Steyaert, J. (2004) Leaving group activation by aromatic stacking: an alternative to general acid catalysis, *J. Mol. Biol. 338*, 1−6.
16. Brooks, B. R., Bruccoleri, R. E., Olafson, B. D., States, D. J., Swaminathan, S., and Karplus, M. (1983) CHARMM: a program for macromolecular energy, minimization and dynamics calculations, *J. Comput. Chem. 4*, 187−217.
17. Brunger, A. T., and Karplus, M. (1988) Polar hydrogen positions in proteins: empirical energy placement and neutron-diffraction comparison, *Proteins 4*, 148−156.
18. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Jr., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., Pople, J. A. (1998) Gaussian 98, Gaussian, Inc., Pittsburgh, PA.
19. SYBYL. 7.0, Tripos Inc., St. Louis, Missouri, 63144.
20. Brooks, C. L., III, Karplus, M. (1989) Solvent effects on protein motion and protein effects on solvent motion, *J. Mol. Biol. 208*, 159−181.
21. Ryckaert, J. P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of *N*-alkanes, *J. Comput. Phys. 23*, 327−341.
22. Nina, M., and Roux, B. (1997) Atomic radii for continuum electrostatics calculations based on molecular dynamics free energy simulations, *J. Phys. Chem. B101*, 5239−5248.

23. Banavali, N. K., and Roux, B. (2002) Atomic radii for continuum electrostatics calculations on nucleic acids, *J. Phys. Chem. B106*, 11026−11035.
24. Hubbard, S. J., and Thornton, J. M. (1993), NACCESS: computer program, Department of Biochemistry and Molecular Biology, University College London, U.K.
25. Lee, B., and Richards, F. M. (1971) The interpretation of protein structures: estimation of static accessibility, *J. Mol. Biol. 55*, 379−380.
26. Wessel, P., and Smith, W. H. F. (1998) New, improved version of the generic mapping tools released, *EOS Trans. AGU 79*, 579−580.
27. Laaksonen, L. (1992) A graphics program for the analysis and display of molecular dynamics trajectories, *J. Mol. Graph. 10*, 33−34.
28. Bergman, D. L., Laaksonen, L., and Laaksonen, A. (1997) Visualization of solvation structures in liquid mixtures, *J. Mol. Graphics Modell. 15*, 301−306.
29. Tirion, M. M. (1996) Large amplitude elastic motions in proteins from a single-parameter, atomic analysis, *Phys. Rev. Lett. 77*, 1905−1908.
30. Tama, F., and Sanejouand, Y. H. (2001) Conformational change of proteins arising from normal mode calculations, *Protein Eng. 14*, 1−6.
31. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD−visual molecular dynamics, *J. Mol. Graphics 14*, 33−38.
32. Saenger, W. (1984) in *Principles of Nucleic Acid Structure*, pp 51−78, Springer-Verlag, New York.
33. Birck, M. R., and Schramm, V. L. (2004) Binding causes the remote [5′-^{3}H]thymidine kinetic isotope effect in human thymidine phosphorylase, *J. Am. Chem. Soc. 126*, 6882−6883.
34. Roday, S., Amukele, T., Evans, G. B., Tyler, P. C., Furneaux, R. H., and Schramm, V. L. (2004) Inhibition of ricin A-chain with pyrrolidine mimics of the oxacarbenium ion transition state, *Biochemistry 43*, 4923−4933.
35. Bruice, T. C., and Schmir, G. L. (1959) The influence of mechanism on the apparent pK_a' of participating groups in enzymic reactions, *J. Am. Chem. Soc. 81*, 4552−4556.
36. Thornburg, L. D., Henot, F., Bash, D. P., Hawkinson, D. C., Bartel, S. D., and Pollack, R. M. (1998) Electrophilic assistance by Asp-99 of 3-oxo-Δ^5-steroid isomerase, *Biochemistry 37*, 10499−10506.
37. Andersen, J. F., Sanders, D. A. R., Gasdaska, J. R., Weichsel, A., Powis, G., and Montfort, W. R. (1997) Human thioredoxin homodimers: regulation by pH, role of aspartate 60, and crystal structure of the aspartate 60 → asparagine mutant, *Biochemistry 36*, 13979−13988.

BI047394H