Understanding Substrate Specificity in Human and Parasite Phosphoribosyltransferases through Calculation and Experiment†

Jed W. Pitera, Narsimha R. Munagala, Ching C. Wang, and Peter A. Kollman*

Graduate Group in Biophysics and Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, San Francisco, California 94143-0446

Received February 2, 1999; Revised Manuscript Received May 19, 1999

ABSTRACT: We present molecular dynamics (MD) simulations on two enzymes: a human hypoxanthine-guanine-phosphoribosyltransferase (HGPRTase) and its analogue in the protozoan parasite *Tritichomonas foetus*. The parasite enzyme has an additional ability to process xanthine as a substrate, making it a hypoxanthine-guanine-xanthine phosphoribosyltransferase (HGXPRTase) [Chin, M. S., and Wang, C. C. (1994) *Mol. Biochem. Parasitol. 63* (2), 221−229 (*1*)]. X-ray crystal structures of both enzymes complexed to guanine monoribosyl phosphate (GMP) have been solved, and show only subtle differences in the two active sites [Eads et al. (1994) *Cell 78* (2), 325−334 (*2*); Somoza et al. (1996) *Biochemistry 35* (22), 7032−7040 (*3*)]. Most of the direct contacts with the base region of the substrate are made by the protein backbone, complicating the identification of residues significantly associated with xanthine recognition. Our calculations suggest that the broader specificity of the parasite enzyme is due to a significantly more flexible base-binding region, and rationalize the effect of two mutations, R155E and D163N, that alter substrate specificity [Munagala, N. R., and Wang, C. C. (1998) *Biochemistry 37* (47), 16612−16619 (*4*)]. In addition, our simulations suggested a double mutant (D106E/D163N) that might rescue the D163N mutant. This double mutant was expressed and assayed, and its catalytic activity was confirmed. Our molecular dynamics trajectories were also used with a structure-based design program, Pictorial Representation Of Free Energy Changes (PROFEC), to suggest parasite-selective derivatives of GMP. Our calculations here successfully rationalize the parasite-selectivity of two novel inhibitors derived from the computer-aided design of Somoza et al. (*5*) and demonstrate the utililty of PROFEC in the design of species-selective inhibitors.

Phosphoribosyltransferases (PRTases) are enzymes that catalyze the addition of a nucleobase to the α-carbon of α-phosphoribosylpyrophosphate to form a nucleoside monophosphate and pyrophosphate (Figure 1). They also typically catalyze the reverse reaction at high efficiency. For many parasitic organisms, purine PRTases are essential enzymes of the purine salvage pathway. In the protozoan *Tritichomonas foetus* (*T. foetus*), the essential purine PRTase operates with relatively high efficiency on hypoxanthine (H;R2 = H), xanthine (X;R2 = O), and guanine (G;R2 = NH_2). In contrast to the protozoan HGXPRTase, the corresponding human enzyme shows a substantial preference for hypoxanthine and guanine and minimal xanthine activity (human HGPRTase). Competitive inhibition data (K_i) values show that the human HGPRTase has a 100-fold reduced affinity for xanthine (250 μM) versus hypoxanthine or guanine (1.8 and 2.4 μM) (*6*). The kinetics of both enzymes have also been extensively studied (*7*−*9*), and k_{cat}'s, K_m's, and catalytic efficiencies for the forward reactions of both human and parasite enzymes are summarized in Table 1. Detailed data for the reaction of xanthine with the human enzyme have not been reported in

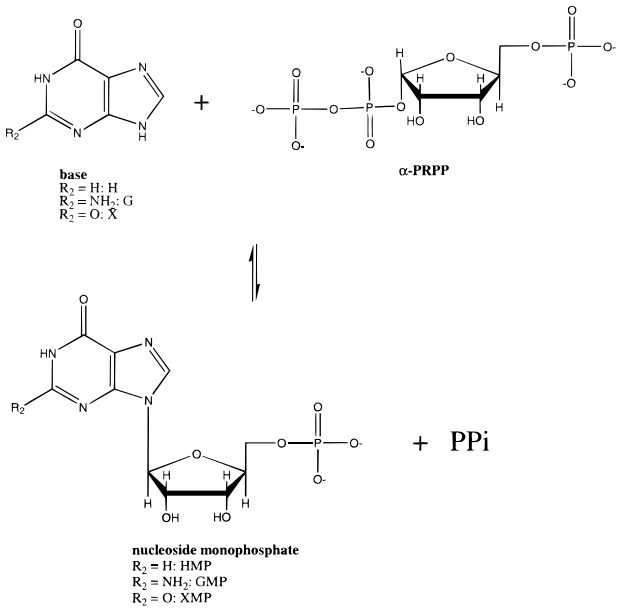

FIGURE 1: Schematic of the reaction carried out by the human and parasite phosphoribosyltransferases.

† This work was supported by the NIH through grants GM-29072 (P.A.K., J.W.P.) and AI-19391 (N.R.M., C.C.W.). J.W.P. received support from a NSF predoctoral fellowship and the University of California Office of the Chancellor.

* To whom correspondence should be addressed at the Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, CA 94143-04446. Phone: (415)476-4637. Fax: (415)-476-0688. E-mail: pak@cgl.ucsf.edu.

10.1021/bi990250w CCC: $18.00
Published on Web 07/17/1999

Table 1: Kinetic Parameters for Human HGPRTase and *T. foetus* HGXPRTase Activity

(A)[a]

enzyme	substrate	$K_{m,app}$ (μM)	K_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})	relative rate
human	hypoxanthine	1.9 ± 0.3	8.4	2.7	1
human	xanthine	NA	NA	NA	NA
human	guanine	3.1 ± 0.9	14.2	7.5	2.7

(B)[b]

enzyme	substrate	K_m (μM)	K_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})	efficiency
parasite	hypoxanthine	3.05 ± 0.54	8.92 ± 0.46	2.92	1
parasite	xanthine	6.08 ± 0.81	4.82 ± 0.8	0.78	0.26
parasite	guanine	2.4 ± 0.74	2.48 ± 0.24	1.03	0.35

(C)[c]

enzyme	substrate	K_m (μM)	K_{cat} (s^{-1})	k_{cat}/K_m (μM^{-1} s^{-1})	efficiency
parasite D163N	hypoxanthine	4.92 ± 0.8	8.95 ± 0.67	1.81	0.62
	xanthine	>300	−	−	−
	guanine	9.13 ± 1.1	2.17 ± 0.13	0.24	0.08
parasite R155E	hypoxanthine	2.06 ± 0.51	10.22 ± 0.8	4.96	1.69
	xanthine	95.8 ± 18.4	33.05 ± 2.6	0.34	0.11
	guanine	3.01 ± 0.23	25.98 ± 1.9	8.63	2.95

[a] Data from (*8*). [b] Data from (*9*). [c] Data from (*4*).

the literature, presumably due to the low affinity of the enzyme for this substrate. Also listed in Table 1 are the properties of two mutant parasite enzymes (R155E, D163N). These two mutants were designed by Munagala and Wang (*4*) to help understand the broad substrate specificity of the parasite enzyme. R155E is particularly interesting since this mutation distant from the active site serves to substantially reduce the ability of the enzyme to process xanthine, XMP and GMP (data not shown). Closer to the substrate, the backbone of residue 163 forms direct hydrogen bonds to the C2 substituent of the base. The D163N mutant is interesting since the side chain of this residue does not form any direct contacts with the substrate. However, the substitution of asparagine for aspartic acid substantially decreases the xanthine activity of the HGXPRTase while only slightly affecting the other two substrates.

X-ray crystal structures of the human HGPRTase complexed with GMP and the parasite HGXPRTase−GMP complex were solved by Eads et al. (*2*) and Somoza et al. (*3*). The two enzymes show about 30% sequence identity. Both active sites are very similar in structure, and use almost identical residues to recognize the substrate. The two complexes are superimposed and shown in Figure 2, with several key residues labeled. The similarity of the two active sites does not answer why the parasite enzyme recognizes xanthine or XMP while the human enzyme does not. To attempt to answer this question, we carried out molecular dynamics (MD) simulations of the human−GMP, human−XMP, parasite−GMP, and parasite−XMP complexes. In addition, we simulated the two mutant parasite complexes (R155E−XMP and D163N−XMP) in order to attempt to understand the effects of these alterations on enzyme specificity. Our molecular dynamics calculations provide a detailed, microscopic picture of the structure and dynamics of each complex, and are discussed in detail below. After this work was initiated, several structures of complete purine phosphoribosyltransferase catalytic complexes were solved (*10*, *11*). These structures show the positions of the substrate base, phosphoribosyl phosphate, and catalytic metal ions just prior to the formation of the nucleotide. As expected, the crucial protein contacts that recognize the base are similar in these new structures and the previous nucleotide complexes which we have simulated.

METHODS

Setup and Model Building. The models of each complex were prepared using the LEAP module of AMBER 5.0 (*12*). The PDB coordinates of the human−GMP (1hmp) and parasite−GMP (1hgx) complex were used as starting points for all calculations. They were modified as necessary to represent alternative substrates (XMP) or mutant structures (D163N, R155E). XMP was built onto the GMP coordinates by hand, while the mutant coordinates were built using the "swapaa" function of MidasPlus (*13*). A sulfate ion (SO_4^{2-}) distant from the ligand binding site was deleted from the parasite structures. All ionizable amino acids were set to their most probable protonation state at pH 7.0, with no exceptions.

It was necessary to build molecular mechanics models of both XMP and GMP for our calculations, since these molecules are not part of the usual AMBER parameter set. While all necessary bond, angle, dihedral, and van der Waals terms were already present in the Cornell et al. (*14*) force field, we had to derive charges for both substrates. The restrained electrostatic potential (RESP) method (*15*) was used to determine a set of point charges for each molecule that best fit their electrostatic potential. These electrostatic potentials were derived from single-point-restricted Hartree−Fock quantum mechanical calculations carried out with the Gaussian 94 computer program (*16*). It was necessary to split the nucleotides into two parts: the base and sugar, which were treated as a single neutral species; and the phosphate group, which was treated as a singly charged anion (*17*). In both calculations, a 6-31G* basis set was used. The two calculations were combined to yield ribonucleoside phosphates each with a net charge of −1.0. The charges and parameters used are available from the authors. In preparing our model of XMP, we considered only the 2-oxo tautomer, as enolic tautomers of the nucleobases are rarely populated in solution (*18*).

Equilibration. Using these molecular mechanics models and the protein structures described above, including any

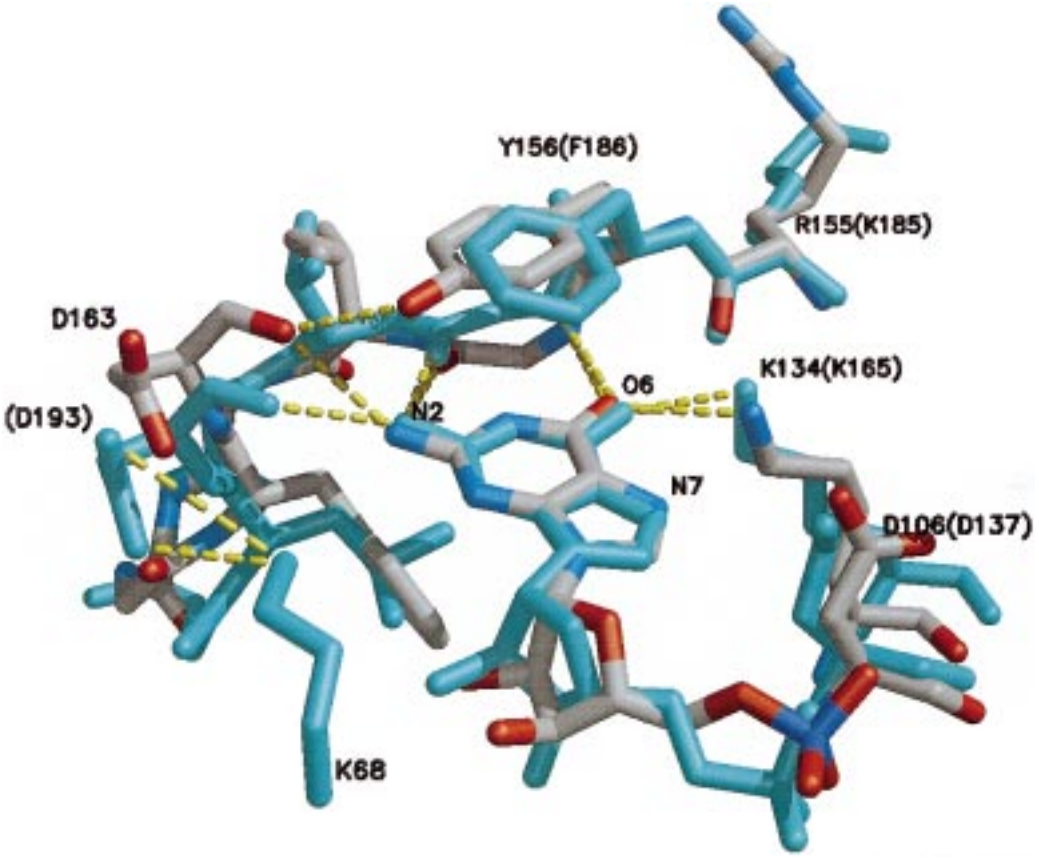

FIGURE 2: Superimposition of the human−GMP and parasite−GMP active sites from their crystal structures. The parasite complex is colored by atom type, and the human complex is displayed in cyan. Key protein residues and substrate atoms mentioned in the text are labeled, with the human residue number in parentheses. Several enzyme−substrate hydrogen bonds are highlighted for both complexes. The additional salt bridge between D193 and K68 in the human structure is also displayed.

counterions and crystallographic waters, we solvated each complex with a 25.0 Å cap of TIP3P (*19*) water molecules centered on the substrate. Though both the human and parasite structures are dimeric enzymes, we concentrated our attentions on a single active site in each case. In fact, only atoms closer than 25 Å from the substrate were permitted to move in each simulation.

Each complex was equilibrated and simulated using an identical protocol and the SANDER molecular dynamics package of AMBER 5.0 (*12*). The molecular mechanics model described above was first subjected to 100 steps of steepest descent minimization followed by 1000 steps of conjugate gradient minimization to fix any errors introduced in the model-building process. After this minimization, a short [2 ps (2 ps), heating from 0 to 300 K] molecular dynamics run was started where only the water molecules were allowed to move while the protein, ligand, and counterions remained fixed. This served to equilibrate the water structure around the complex prior to the production runs. The equilibrated complex was then gradually heated from 0 to 300 K over 8 ps. During this heating period, the protein and ligand heavy atoms were restrained to their original positions with weak (1.0 kcal mol^{-1} $Å^{-2}$) positional restraints. Once the heating was over, the positional restraints were released, and each complex was simulated for a total of 300 ps. For all runs, Berendsen temperature coupling (*20*) was used to maintain the system at its assigned temperature, while the SHAKE algorithm constrained the length of all hydrogen containing bonds to their equilibrium values. The latter serves to permit the use of a 2 fs time step in the dynamics calculations. To minimize computational expense, long-ranged energies and forces (electrostatic and van der Waals) were only calculated out to a fixed (residue-based) cutoff distance. In the past, typical molecular dynamics or Monte Carlo protocols have used 8 or 9 Å cutoffs (*21*). The recent development of Particle Mesh Ewald (PME) electrostatics (*22*) and related algorithms (*23*) allows all long-range contributions to be included at a reasonable cost in periodic systems. Since our complexes were aperiodic, it was necessary to use a cutoff-based approach for our calculations. We initially simulated each complex using a 9 Å cutoff. However, this resulted in rapid distortions of the binding complex including movement of XMP or GMP phosphate groups out of the phosphate-binding loop and into solution. These distortions are commonly seen when using cutoffs in simulations of highly charged systems (*24*). Due to the distortions, these calculations were discarded, and a 14.0 Å cutoff was used for all subsequent calculations. The longer cutoff resulted in much lower structural distortions and relatively stable structures for each complex. A similar effect has also been observed in complexes of isocitrate dehydrogenase (*25*). All calculations reported in this paper used the longer 14 Å cutoff. Attempts to simulate the complex in a vacuum with a 14 Å cutoff and a distance-dependent dielectric of $4R_{ij}$ were unsuccessful due to unrealistic fluctuations of the ligand and Y156 or F186, the protein side chain that stacks atop the base portion of the ligand.

Analysis. The CARNAL trajectory analysis software of AMBER 5.0 was used to calculate the root-mean-square (RMS) deviations and interatomic distances for each trajec-

tory. The MDANAL part of the AMBER 5.0 package was used to calculate the atomic fluctuations of each complex. Structures were displayed and compared using the MidasPlus graphical visualization software (*13*).

PROFEC. The PROFEC free energy extrapolation software (*26*) was used to suggest modifications to GMP or guanine that would increase the parasite selectivity of the resulting compound. Both the parasite−GMP and human−GMP trajectories were analyzed to find locations where the ligand could be favorably derivatized. This is done by calculating the free energy cost of inserting a test particle at various locations near the C2 of GMP. The difference of these two analyses yields a map of positions where GMP could be changed to yield a parasite-selective compound. For these calculations, an uncharged carbonyl oxygen-sized particle was used as the test particle. Once the map was calculated, it was projected on the parasite−XMP complex and displayed using MidasPlus, which also served to superimpose the DOCKed coordinates of the compounds described by Somoza et al. (*5*) The free energy map was contoured at levels of −2.0, 0.0, and +1.0 kcal/mol in order to show regions where modifications to the ligand would be favorable, neutral, or unfavorable, respectively.

Simulations. A total of six simulations were carried out to permit a detailed comparison of the structure and dynamics of human and parasite enzymes bound to each substrate (GMP, XMP) as well as model the behavior of two parasite mutants (D163N, R155E) bound to XMP. In all cases, the binding complex is well maintained, and the overall geometry of the initial model-built complex was preserved.

Site-Directed Mutagenesis. Site-directed mutagenesis and expression and purification of the D106E/D163N double mutant were carried out as previously described (*1*, *4*). Oligonucleotide primers were designed, synthesized, and used with a Stratagene kit for site-directed mutagenesis. The plasmid (pBTfprt) containing the full-length gene encoding *T. foetus* was transformed into an *E. coli* mutant strain Sϕ606. Expression of the mutant *T. foetus* HGXPRTase gene in the plasmid was induced in the low-phosphate culture medium. The recombinant mutant protein was purified to homogeneity from the transformed cells, and steady-state kinetic analysis was performed on the purified enzyme. The kinetic constants were obtained by monitoring the catalysis spectrophotometrically, as previously described (*9*).

RESULTS

Simulations of Wild-Type Complexes. The enzymological data (Table 1) show that the human enzyme substantially prefers guanine over xanthine, whereas the parasite HGXPRTase shows only a slight preference. Our simulations of each wild-type enzyme bound to both substrates were intended to address this issue. The structural changes and fluctuations seen in each simulation help explain the observed substrate specificity.

A rough measure of structural change during each simulation is the root-mean-square (RMS) deviation of the protein backbone from its initial position. Compared to the corresponding parasite complexes, the protein backbones of the human complexes diverge further from the X-ray structure, as measured by RMS deviation (Figure 3, top and middle panels). The average structures of the human−GMP and human−XMP simulations are shown superimposed on the X-ray structure in Figure 4, highlighting a few differences. The salt bridge between D193 and K68 is broken in the human−XMP complex, while it is preserved in the human−GMP complex. In the human−XMP complex, K68 moves to form a salt bridge with D134, which normally interacts with the sugar hydroxyls. In this complex, D193 is also displaced slightly away from the base, minimizing unfavorable interactions between its backbone carbonyl and the C2 carbonyl of XMP. In contrast, this same residue superimposes closely with the crystal structure in our human−GMP simulation, maintaining a good hydrogen bond between its main chain carbonyl and the GMP amino group. Despite these structural differences, both simulations show uniformly low fluctuations of the protein backbone in this ligand-binding region. Also, the phosphate- and sugar-binding regions of both complexes are well preserved, as expected for substrates that differ only at the C2 position.

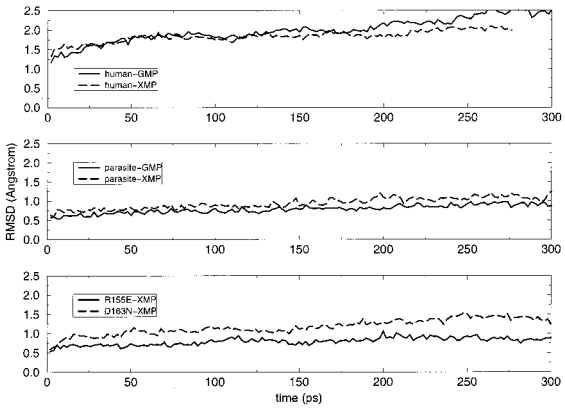

FIGURE 3: Cα root-mean-square deviation of each simulation from its starting structure. Complexes with GMP are solid lines; complexes with XMP are graphed as long dashed lines.

In contrast to the human complexes, the parasite−XMP and parasite−GMP backbones show smaller RMS deviations from the initial X-ray structure (Figure 3, middle panel). However, the two parasite structures are more different from each other than the two human structures are from each other, especially in the C2 pocket region. This is best shown by Figure 5, which superimposes the C2 pockets of each parasite complex. From this figure, one can see that both structures have diverged somewhat from the X-ray crystal structure, primarily due to the loss of the hydrogen bond between the D163 backbone carbonyl and the Y156 hydroxyl group. In the parasite−XMP complex, D163 has moved away from the C2 carbonyl of xanthine, allowing water molecules more access to donate hydrogen bonds to the substrate. With the Y156 hydrogen bond broken, D163 shifts to make a much better hydrogen bond to the GMP NH_2 in the parasite−GMP complex. Again, the phosphate- and sugar-binding regions of both the parasite−GMP and parasite−XMP complexes remain well-structured throughout the calculation.

Simulations of Mutant Complexes. In addition to the wild-type simulations, we carried out simulations on two mutants of the parasite enzyme bound to XMP. During the extensive enzymological and mutagenic studies of *T. foetus* HGXPRTase by Munagala et al. (*4*, *9*), two anomalous mutations were observed. First, the mutation of arginine 155 to a glutamic acid (R155E) disrupts a salt bridge relatively distant (∼15 Å) from the C2 pocket, yet it substantially reduces

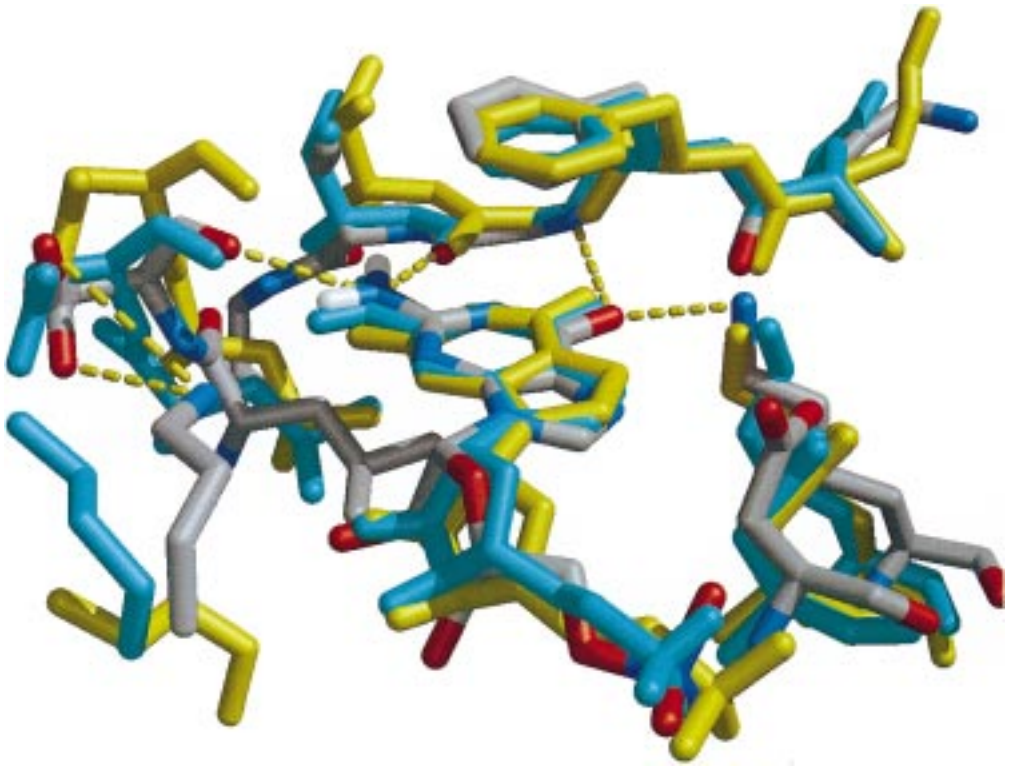

FIGURE 4: Superimposition of the human−GMP crystal structure (colored by atom) with the human−GMP (cyan) and human−XMP (yellow) structures averaged over each simulation. The enzyme−substrate hydrogen bonds shown in Figure 2 are also displayed for the X-ray coordinates, for reference.

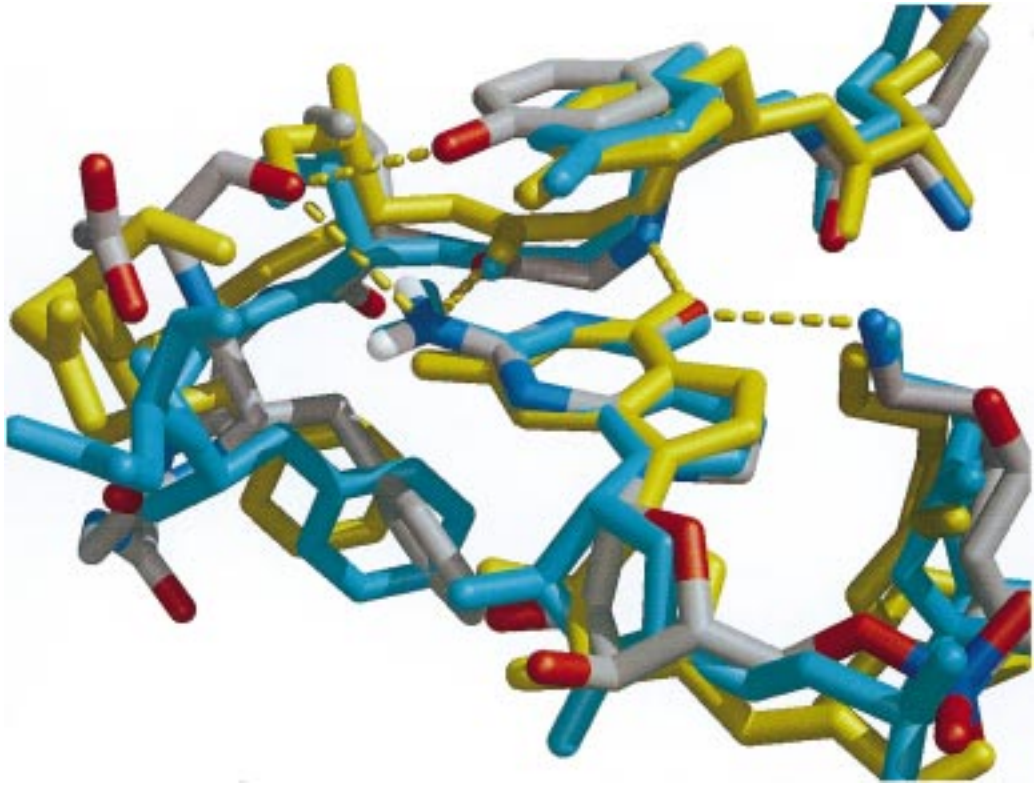

FIGURE 5: Superimposition of the parasite−GMP crystal structure (colored by atom) with the parasite−GMP (cyan) and parasite−XMP (yellow) structures averaged over each simulation. The enzyme−substrate hydrogen bonds shown in Figure 2 are also displayed for the X-ray coordinates, for reference.

the affinity of the enzyme for some C2-substituted substrates (xanthine, XMP, and GMP) while influencing hypoxanthine/IMP binding much less. Our simulation of the parasite R155E−XMP complex shows no large structural differences from the wild-type enzyme, and has a low overall RMS deviation (Figure 3, bottom panel). However, the backbone fluctuations of residue 163 (and the C2 pocket) are substantially decreased in the mutant complex. This suggests that

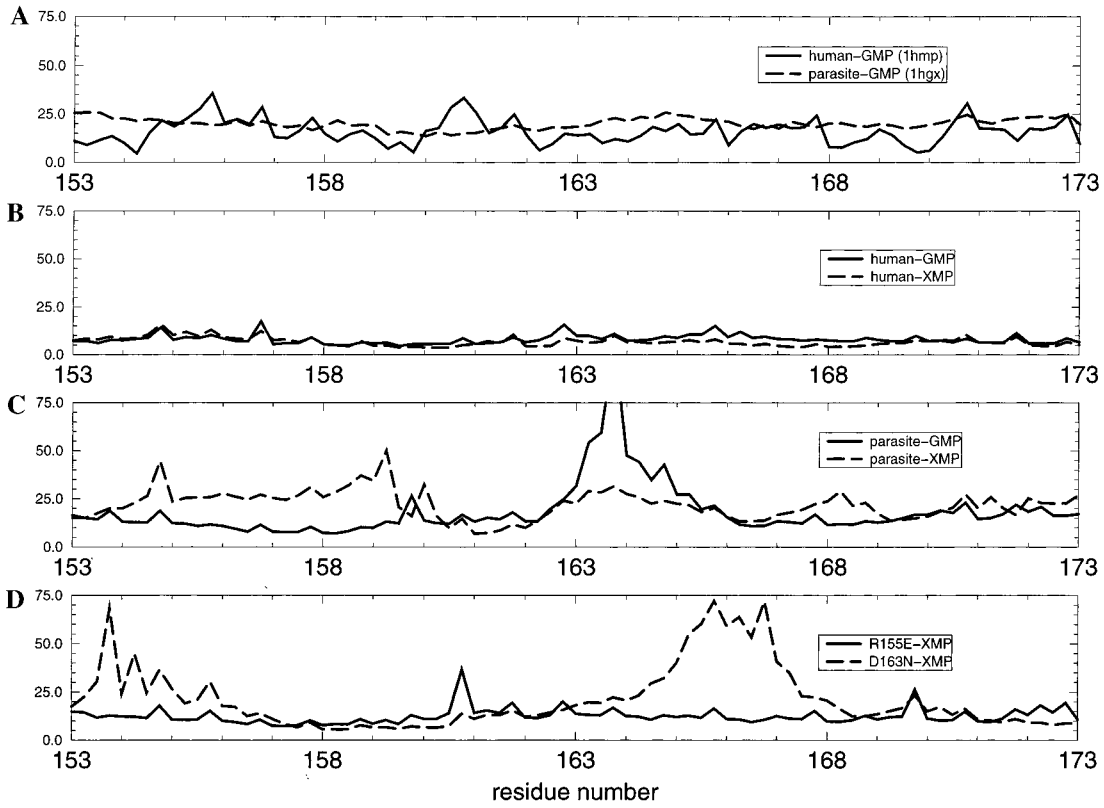

FIGURE 6: Protein backbone atomic fluctuations (*B*-factors) for each set of coordinates. These are only plotted for the residues near the C2 pocket: 153−173 in the parasite and 183−203 in the human enzyme. The parasite numbering is used in this figure, and the human enzyme data have been translated to superimpose corresponding residues. For each residue, *B*-factors are listed for the N, CA, C, and O atoms of the protein backbone.

R155E may exert its effect by reducing the ability of the parasite C2 pocket to reorganize and recognize C2-substituted substrates.

In contrast, the D163N mutant changes the amino acid that forms most of the C2 pocket in the parasite enzyme. Since the C2 pocket is largely formed by the backbone of residue 163, it is interesting that this mutation substantially affects the affinity of the parasite for XMP and xanthine (Table 1). In our simulations, exchange of the negatively charged carboxylic acid for the neutral amide group of asparagine causes substantial structural distortions of the ligand-binding complex. In the wild-type structure, D163 acts as a salt bridge partner for R169. In the D163N mutant, R169 lacks a nearby negatively charged salt bridge partner. Over the course of our simulations, this arginine moves toward the carboxylic acids of D103 and E102 which normally recognize the hydroxyl groups of the ligand. While we do not expect that these specific distortions occur in the actual D163N mutant, it does suggest that the deleterious effect of the D163N mutant is due to a substantial reorganization of the enzyme, rather than a specific contact with the substrate. In support of this observation, we must note that Munagala et al. found the D163E mutant to have almost wild-type activity (*4*). The D163E mutant preserves the negative charge at this position and presumably maintains the D163−R169 salt bridge as well.

Analysis of the Simulations. The backbone (N, CA, C, O) atomic fluctuations (*B*-factors) of each complex are presented in Figure 6. They are shown for residues 153−173 (in the parasite enzyme; 183−203 in the human), which includes the amino acids that recognize the C6, N1, and C2 positions of the base. Figure 6A shows the *B*-factors in this region for the two starting crystal structures. Both the human−GMP and parasite−GMP complexes show low fluctuations in this region according to the crystal structures. For comparison, panels B and C show the fluctuations calculated from our wild-type simulations. Here, both human complexes show the expected low fluctuations in this region. The parasite−GMP and parasite−XMP simulations show much larger fluctuations in this region. This suggests that the parasite enzyme has a far less rigid C2 pocket which permits it to recognize both the amino group of GMP and the carbonyl oxygen of XMP. While the main-chain carbonyl of residue D163/D193 provides an appropriate hydrogen bond acceptor for GMP's amino group, the C2 pocket needs to reorganize to properly recognize XMP's carbonyl group. While buried hydrogen bonds are not necessarily expected to add to binding affinity, the failure to form an expected enzyme−substrate hydrogen bond can cost substantial affinity or stability (*27*). The necessary hydrogen bonds to recognize XMP appear to be provided largely by nearby water molecules in our simulations. In fact, our parasite−XMP simulations shows a very intermittent hydrogen bond between the D163 carbonyl and the hydroxyl of tyrosine Y156. Breaking this hydrogen bond allows the tyrosine and aspartic acid to separate enough that a water molecule can approach the oxygen of XMP from above the plane of the base.

Interestingly, the C2 pocket fluctuations of both mutant parasite complexes are decreased relative to their wild-type counterparts (Figure 6, panel D). In the case of R155E−XMP, this may be the mechanism by which the R155E mutation exerts its effects on some C2-substituted substrates

Table 2: Distances Associated with Enzyme−Substrate Hydrogen Bonds (Å)

simulation/structure	I157/I187 O−R2	D163/D193 O−R2	D163/D193 N−R2
human−GMP (X-ray)	3.0	3.0	4.1
human−GMP	3.5 ± 0.4	3.3 ± 0.4	4.0 ± 0.5
human−XMP	3.8 ± 0.2	3.6 ± 0.3	3.9 ± 0.7
parasite−GMP (X-ray)	3.3	3.2	3.8
parasite−GMP	3.2 ± 0.3	4.7 ± 1.3	4.1 ± 0.7
parasite−XMP	3.6 ± 0.2	4.0 ± 0.4	4.2 ± 0.7
R155E−XMP	3.8 ± 0.2	3.4 ± 0.2	2.1 ± 0.4
D163N−XMP	3.8 ± 0.2	3.4 ± 0.3	2.0 ± 0.3

Table 3: Distances between the Catalytic D106/D136 and the Substrate N7 Nitrogen (Å)

simulation	minimum distance	average	maximum
human−GMP	2.8	3.9	5.4
human−XMP	3.0	4.4	6.0
parasite−GMP	2.8	3.5	4.8
parasite−XMP	2.9	3.5	4.7
R155E−XMP	2.9	3.6	5.0
D163N−XMP	3.0	4.2	6.0

(xanthine, XMP, GMP). In contrast, the D163N mutant shows moderate fluctuations in the backbone of residue 163, but slightly increased fluctuations in residues 153−154 and the region around R165. Since this latter residue shifts substantially during our simulation, its increased fluctuations are expected.

To get a picture of the specific contacts that form the C2 pocket, we monitored the distance between nearby hydrogen bond donors and acceptors during our simulations. There are three major hydrogen bonds that can be formed with the base C2 substituent in both enzymes, as seen in Figure 2. First, the main-chain carbonyl of I157/I187 is seen to accept a hydrogen bond from GMP in both human and parasite crystal structures. Second, the backbone carbonyl of D163/D193 can also accept a hydrogen bond from the amino group of guanine. Third, the amide nitrogen of that same residue can donate a hydrogen bond to the C2 carbonyl of xanthine or XMP if they are present, though this would require some reorganization to achieve the optimal geometry. We monitored the time course of these three distances over each of our simulations, and found that only a few varied significantly over our trajectories. As expected, the D163 backbone−C2 distances are significantly longer in the two wild-type parasite complexes. The averages and standard deviations of each distance are shown in Table 2, along with the values from the two crystal structures for comparison.

Both mutant complexes (R155E, D163N) also show very short distances between the amide nitrogen of residue 163 and the C2 oxygen of XMP, indicating that the base has "tipped down" in the C2 pocket to form this good hydrogen bond.

In addition to contacts along the edges of the base, we also monitored the distance between the N7 atom of each base and the oxygen atoms of aspartic acid D106/D137 (both shown in Figure 2). Proton transfer between this carboxylic acid and the N7 nitrogen has been implicated as the critical catalytic step in the human enzyme (*28*). Distances from the nearest carboxyl oxygen to N7 of each substrate are shown in Table 3.

These distances were calculated over the last 200 ps of each trajectory, to allow for the initial relaxation of each structure. Although we are not necessarily simulating a catalytically competent complex (D106/D137 would need to be protonated to initiate a proton transfer to N7 of the base), the human−XMP and D163N−XMP complexes stand out in the above table as having larger average and maximum distances than the other complexes. Interestingly, these are two of the three complexes that show reduced affinity and catalytic activity (Table 1). The last, R155E, appears to be impaired due to its decreased fluctuations in the C2 pocket, as mentioned above.

Table 4: Kinetic Parameters of the D106E/D163N Double Mutant

enzyme	substrate	K_{cat} (s^{-1})	K_m (μM)	K_{cat}/K_m ($\mu M^{-1}\ s^{-1}$)
wild type	xanthine	4.8 ± 0.8	6.08 ± 0.8	0.78
D163N	xanthine	>300	−	−
D106E/D163N	xanthine	0.39	21.5	0.18

Our observation of an increased D106−XMP N7 distance in the D163N−XMP simulation prompted the design of the double mutant D106E/D163N. The substitution of glutamic acid for aspartic acid at position 106 allows the double mutant to partially compensate for the deleterious effects of the D163N mutation. This double mutant was made, expressed, and assayed by N. Mungala using the previously described protocols for biochemical studies of the parasite HGXPRTase (*1, 4*). The double mutant restores catalytic activity to the D163N enzyme, as shown in Table 4. As well as supporting our model, the activity of the double mutant is additional evidence that the carboxylic acid at position 106 (137 in the human enzyme) is essential to the catalytic mechanism of these enzymes (*28*).

A model of the double mutant was hand-built from the structure of the simulated D163N−XMP complex. The side chain dihedrals of glutamic acid were kept to canonical (rotameric) values (*29*). The D106E substitution allows the catalytic carboxylic acid to be placed within 2.5 Å of the N7 nitrogen without significant distortion of the side chain. In contrast, the corresponding D163N−XMP structure shows a carboxylic acid−N7 distance of over 4 Å.

PROFEC Analysis. The PROFEC free energy estimation software (*26*) was used to suggest how GMP could be modified to yield a parasite-selective ligand. The free energy cost of adding a test particle (or potential modification) at the C2 position of the base was evaluated for both the human−GMP and parasite−GMP trajectories. These two free energy "maps" were then combined to yield a picture of how GMP (or GMP-like ligands) could be modified to improve their parasite-selectivity, either by enhancing their affinity for the HGXPRTase or by impairing their affinity for the human enzyme. Interestingly, there is a large region along the edge of the base by the N2 and N3 positions where our software suggests that modifications be made. This region

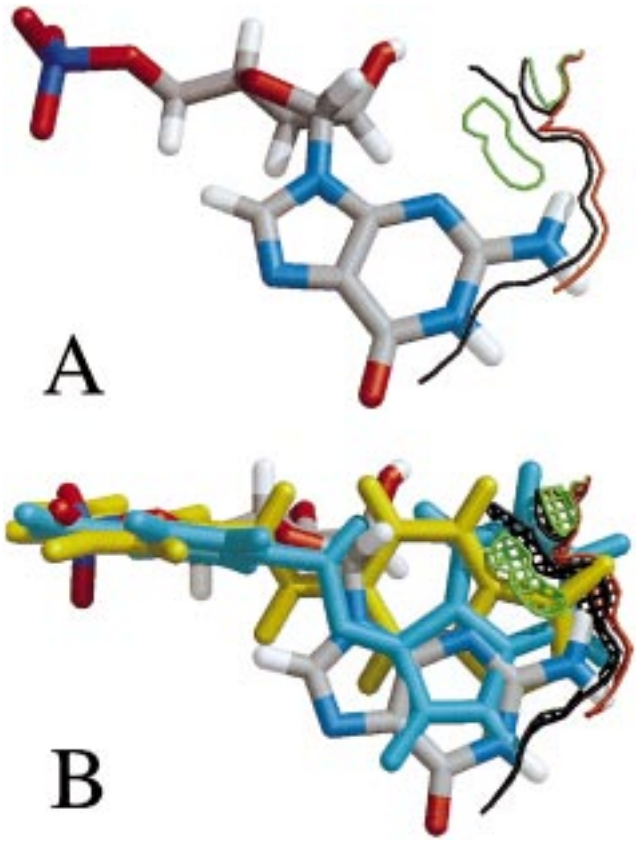

FIGURE 7: (A) PROFEC contour map of parasite selectivity. A slice through the contour map of suggestions provided by the PROFEC software is shown superimposed on the parasite-bound structure of GMP. The green contours correspond to regions where modifications are predicted to enhance parasite selectivity by −2.0 kcal/mol. The black contours correspond to the 0.0 kcal/mol contour, while the red contour is unfavorable by +1.0 kcal/mol. Since this is a contour map, addition of heavy-atom modifications to GMP within the black contour is expected to increase parasite selectivity, while derivatives that place heavy atoms beyond the black contour or in the red regions are expected to show decreased parasite selectivity. (B) PROFEC/DOCK comparison. The DOCKed binding conformations of the phthalic anhydride (yellow) and indol-2-one (cyan) compounds found by Somoza et al. are superimposed on the parasite-bound structure of GMP and the same PROFEC map used above. Again, the contour levels are green (−2.0 kcal/mol), black (0.0 kcal/mol), and red (+1.0 kcal/mol), while several thicknesses of the map have been drawn to aid comparison. Both compounds place several atoms in the favorable region suggested by PROFEC. In addition, the indol-2-one projects significantly beyond the first black contour and into the unfavorable region, while the phthalic anhydride only projects slightly past this region. This is compatible with the higher parasite-selectivity seen for the phthalic anyhydride.

is shown in Figure 7A. Since the parasite−GMP complex is much more solvent-exposed along the N2/N3 edge of the base, it is not surprising that PROFEC picks this region for parasite-selective modifications. Additional atoms here would displace a water molecule in the parasite−GMP complex, while they would probably have a steric clash with D193 or K68 in the slightly better-packed human−GMP complex.

The recent computer-aided design of several parasite-selective ligands here at UCSF (*5*) allows us to test our observations. Somoza et al. found two lead compounds via computer database screening with the DOCK program (*30*). These two compounds, a phthalic anhydride-nitrobenzene and an indol-2-one-nitrobenzene, are shown in Figure 7B, superimposed upon GMP in their DOCKed conformations

Table 5: IC_{50}s of Lead Compounds Studied by Somoza et al. (*5*)

compound	parasite IC_{50} (μM)	human IC_{50} (μM)	ratio (hum/par)
phthalic anhydride	300	>1000	>3
indol-2-one	240	200	0.83

(*31*). Figure 7B also includes the PROFEC contour map showing where GMP could be modified to increase or decrease its parasite selectivity. Interestingly, the two compounds superimpose somewhat with the parasite-favoring region shown in Figure 7A. Moreover, the indol-2-one projects substantially into the region that we expect to be unfavorable for parasite selectivity, while the phthalic anhydride follows the parasite-selective contour much more closely. This only becomes significant when considered in the context of Table 5, which lists the IC_{50}s and selectivity (IC_{50}/IC_{50}) of both compounds as reported by Somoza et al. Here we see that the indol-2-one is in fact relatively nonspecific, while the phthalic anhydride shows a greater than 3-fold preference for the parasite enzyme.

DISCUSSION AND CONCLUSIONS

In any molecular modeling study, one must be aware of the limitations of the model used and how they relate to the questions of interest. In our case, we were interested in the qualitative details of molecular recognition between HGPRTase/HGXPRTase and GMP or XMP. These two substrates only differ by the substitution of a carbonyl (xanthine, XMP) for an amino group (guanine, GMP) at the C2 position. In addition, the two enzymes we studied show only subtle structural differences in their C2-binding pockets. Given these observations, we chose a computational model that allowed us to simulate a number of enzyme−substrate complexes for a reasonable time, and took care to ensure that the representation was sufficient to maintain the structure of each complex relatively well. Since we were interested in the local structure and dynamics of each active site, we chose to model those regions in detail while holding most of each enzyme fixed. While our calculations are relatively short (300 ps each), they appear to be sufficient to highlight the differences between the enzyme−substrate complexes. In each case, enzyme−substrate interactions and geometries are well converged, with the possible exception of our D163N−XMP complex, which shows substantial structural deviations. Given the representation used, we do not expect that we have definitively divined the structure of each simulated complex, nor completely modeled the overall dynamics of either enzyme in solution. Our calculations have, however, allowed us to rationalize the ability of *T. foetus* to process xanthine, explain the effects of two mutations, and suggest the basis for the parasite-selectivity of different ligands.

With the above limitations in mind, our calculations have revealed a plausible model for the ability of the *T. foetus* HGXPRTase to process XMP. The residues of the C2 pocket, specifically D163/D193, show much higher fluctuations in the parasite enzyme when compared to the human enzyme. This is particularly significant given the lower overall RMS deviation of the parasite simulations from their starting structures. The parasite enzyme appears to be able to reorganize to accommodate both the C2 carbonyl of XMP and the C2 amino group of GMP, while the C2 pocket of

the human enzyme is relatively "frozen" in a conformation that only recognizes GMP. In addition, the R155E parasite mutant shows reduced fluctuations in this region, compatible with its reduced ability to process xanthine—it too becomes "frozen" in the GMP-binding conformation. It is important to note (Figure 6) that neither the parasite/GMP nor the human/GMP crystal structures show this difference in their C2 pocket temperature factors. However, the most important observation is probably the difference in human−XMP and parasite−XMP fluctuations, since this helps to explain why the parasite enzyme can recognize XMP while the human cannot. Since there are no crystal structures of either of these complexes, our simulations provide the only structural information of how xanthine and XMP are recognized.

In contrast to R155E, where our simulations show a stable structure with reduced fluctuations from the wild-type enzyme, our D163N model is substantially distorted. The removal of a negative charge associated with the aspartic acid to asparagine mutation removes a salt bridge necessary for structuring R169. Without its salt-bridging partner, R169 is drawn toward the nearby negative charges of E102 and D103. This attraction distorts the structure of the complex. Most importantly, it significantly increases the distance between the N7 nitrogen of XMP and the catalytic carboxylic acid of aspartic acid D106. This model prompted the design and testing of a D106E/D163N double mutant. The increased length of the catalytic carboxylic acid at position 106 restores catalytic activity to the parasite enzyme. While we do not expect that our simulation precisely describes the structure of the D163N/XMP complex, our calculations are substantially validated by the efficacy of the D106E/D163N double mutant. Interestingly, the homologous residues of D163 and R169 form most of the magnesium ion binding site seen in the enzyme−base−PRPP structure solved by Focia et al. (*11*). This led them to conclude, like our study, that the D163N mutation exerts its effects through significant structural changes in the enzyme (distortion of the Mg^{2+} binding site) rather than direct or water-mediated contacts with the base.

In addition to understanding the mechanistic basis of xanthine specificity and providing a model of the specific effect of some parasite mutants, our calculations have shown how GMP could be modified to form a xanthine-selective substrate or inhibitor. These PROFEC calculations indicate a region along the edge of the base, near the C2 substituent and the N3 position, where added derivatives would be preferred in the parasite enzyme but unfavorable in the human. Interestingly, this region is occupied by the parasite-selective ligands recently developed by the Kuntz group at UCSF. As far as we are aware, this represents the first case where free energy extrapolation methods have been used to study species-selective ligand modifications. While PROFEC calculations only yield qualitative suggestions, our results here imply that they will be a useful tool in species-selective drug design.

REFERENCES

1. Chin, M. S., and Wang, C. C. (1994) *Mol. Biochem. Parasitol. 63* (*2*), 221−229.
2. Eads, J. C., Scapin, G., Xu, Y., Grubmeyer, C., and Sacchettini, J. C. (1994) *Cell 78* (2), 325−334.
3. Somoza, J. R., Chin, M. S., Focia, P. J., Wang, C. C., and Fletterick, R. J. (1996) *Biochemistry 35* (22), 7032−7040.
4. Munagala, N. R., and Wang, C. C. (1998) *Biochemistry 37* (47), 16612−16619.
5. Somoza, J. R., Skillman, A. G., Jr., Munagala, N. R., Oshiro, C. M., Knegtel, R. M., Mpoke, S., Fletterick, R. J., Kuntz, I. D., and Wang, C. C. (1998) *Biochemistry 37* (16), 5344−5348.
6. Krenitsky, T., Papaioannou, R., and Elion, G. (1969) *J. Biol. Chem. 244* (5), 1263−1270.
7. Xu, Y., Eads, J., Sacchettini, J. C., and Grubmeyer, C. (1997) *Biochemistry 36* (12), 3700−3712.
8. Keough, D. T., Ng, A. L., Winzor, D. J., Emmerson, B. T., and de Jersey, J. (1999) *Mol. Biochem. Parasitol. 98* (1), 29−41.
9. Munagala, N. R., Chin, M. S., and Wang, C. C. (1998) *Biochemistry 37* (12), 4045−4051.
10. Vos, S., Parry, R., Burns, M., deJersey, J., and Martin, J. (1998) *J. Mol. Biol. 282* (Oct 2), 875−889.
11. Focia, P., Craig, S., and Eakin, A. (1998) *Biochemistry 37* (49), 17120−17127.
12. Case, D. A., Pearlman, D. A., Caldwell, J. W., Cheatham, T. E., Ross, W. S., Simmerling, C. L., Darden, T. A., Merz, K. M., Stanton, R. V., Cheng, A. L., Vincent, J. J., Crowley, M., Ferguson, D. M., Radmer, R. J., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1997) University of California, San Francisco.
13. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *J. Mol. Graphics 6*, 13−27.
14. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1996) *J. Am. Chem. Soc. 118* (9), 2309−2309.
15. Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) *J. Phys. Chem. 97* (40), 10269−10280.
16. Frisch, M., Trucks, G., Schlegel, H., Gill, P., Johnson, B., Robb, M., Cheeseman, J., Keith, T., Petersson, G., Montgomery, J., Raghavachari, K., Al-Laham, M., Zakrzewski, V., Ortiz, J., Foresman, J., Peng, C., Ayala, P., Chen, W., Wong, M., Angres, J., Replogle, E., Gomperts, R., Martin, R., Fox, D., Binkley, J., Defrees, D., Baker, J., Stewart, J., Head-Gordon, M., Gonzales, C., and Pople, J. (1995) Gaussian 94 Software, Revision D.3, Gaussian, Inc., Pittsburgh, PA.
17. Eksterowicz, J. (1997), personal communication.
18. Saenger, W. (1984) *Principles of Nucleic Acid Structure*. *Springer Advanced Texts in Chemistry* (Cantor, C. E., Ed.) Springer-Verlag, New York.
19. Jorgensen, W. L., Chandrasekhar, J., Madura, J., Impey, R. W., and Klein, M. L. (1983) *J. Chem. Phys. 79*, 926.
20. Berendsen, H. J. C., Potsma, J. P. M., van Gunsteren, W. F., DiNola, A. D., and Haak, J. R. (1984) *J. Chem. Phys. 81*, 3684−3690.
21. McCammon, J. A., and Harvey, S. C. (1987) *Dynamics of proteins and nucleic acids*, Cambridge University Press, Cambridge.
22. Darden, T., York, D., and Pedersen, L. (1993) *J. Chem. Phys. 98*, 10089−10092.
23. Luty, B. A., Tironi, I. G., and van Gunsteren, W. F. (1995) *J. Chem. Phys.* 3014−3021.
24. Cheatham, T. E., III, Miller, J. L., Fox, T., Darden, T. A., and Kollman, P. A. (1995) *J. Am. Chem. Soc. 117* (14), 4193−4194.
25. Garcia-Vilorca, M., and Kollman, P. A. (1998) *Int. J. Quantum Chem.* (submitted for publication).
26. Radmer, R. J., and Kollman, P. A. (1998) *J. Comput.-Aided Mol. Des. 12* (3), 215−227.
27. Fersht, A. R. (1985) *Enzyme Structure and Mechanism*, W. H. Freeman and Co., New York.
28. Xu, Y., and Grubmeyer, C. (1998) *Biochemistry 37* (12), 4114−4124.
29. Dunbrack, R. L., and Karplus, M. (1993) *J. Mol. Biol. 230* (2), 543−574.
30. Ewing, T. J. A., and Kuntz, I. D. (1997) *J. Comput. Chem. 18* (9), 1175−1189.
31. Oshiro, C., personal communication.

BI990250W