Insight into the Specificity of Thymidylate Synthase from Molecular Dynamics and Free Energy Perturbation Calculations

Giulio Rastelli,[†,§] Bert Thomas,[‡,⊥] Peter A. Kollman,*[,‡] and Daniel V. Santi*[,†,‡]

Contribution from the Departments of Biochemistry and Biophysics and Pharmaceutical Chemistry, University of California, San Francisco, California 94143-0448

Received August 30, 1994[⊗]

Abstract: Molecular dynamics and free energy perturbation calculations have been used to calculate the relative free energies of binding of 2′-deoxyuridine 5′-monophosphate (dUMP) and 2′-deoxycytidine 5′-monophosphate (dCMP) to thymidylate synthase (TS) and two asparagine 229 mutants. Calculations qualitatively reproduce experimentally observed dissociation constants of the protein−nucleotide complexes. Furthermore, they provide insight into structural aspects of binding and catalysis of these two nucleotides to the protein. The simulations of the wild-type TS complexes with dUMP and dCMP support the key role of asparagine 229 in causing tighter binding of dUMP than dCMP; repulsion between the base of dCMP and the asparagine 229 side chain reduces the ΔG of binding to the protein from that found in aqueous solution and causes the displacement of this nucleotide into a position unsuitable for reaction. The free energy calculations of the aspartate 229 mutant of TS interacting with either dUMP or dCMP suggest a synergism between the aspartate 229 side chain and the vicinal histidine 199 in binding. The best agreement between the calculated and the experimental $\Delta\Delta G$ of binding has been obtained when the aspartate side chain is anionic and the histidine 199 is either protonated or in its δH tautomer. Under these conditions, dUMP and dCMP are both properly positioned for nucleophilic attack. In contrast, calculations with a neutral aspartic acid side chain suggest a strong discriminating power of the neutral 229 side chain in binding the two nucleotides, the preferred one depending on which of the two oxygens of the aspartate is protonated. We speculate that protonation of the aspartate 229 side chain can be the key to rationalizing why the aspartate 229 mutant selectively methylates dCMP. Finally, calculations of the valine 229 mutant demonstrate that substitution of the polar asparagine side chain with a hydrophobic residue does not result in a significant change in the location of the two nucleotides in the active site, except that dUMP seems to be better positioned for nuclephilic attack than dCMP.

Introduction

Thymidylate synthase (TS) catalyzes the reductive methylation of 2′-deoxyuridine 5′-monophosphate (dUMP) by 5,10-methylene-5,6,7,8-tetrahydrofolate (CH_2H_4folate) to give 2′-deoxythymidine 5′-monophosphate (dTMP) and 7,8-dihydrofolate (H_2folate). Its mechanism of action,[1] substrate specificity,[2] and atomic structure[3−5] have been extensively studied over the last few decades. Mutation of specific amino acids in TS has proven to be an important tool in elucidating the intrinsic role of particular residues in binding and catalysis.[6] Recently, the mutation of the conserved asparagine 229 to aspartic acid (N229D) in *Lactobacillus casei* TS was reported to change the TS to an enzyme which methylates 2′-deoxycytidine 5′-monophosphate (dCMP) instead of dUMP.[7] Further studies on Asn229 mutants revealed that Asn229 is not essential in binding and catalysis, but it plays the important role of discriminating between dUMP and dCMP by excluding dCMP from the binding site.[8] With the exception of the N229Q mutant, all Asn229 mutants bind dUMP and dCMP almost equally well, with K_D values approaching that of the wild-type (wt) TS·dUMP binary complex.[8] In the crystal structure of the TS·dUMP binary complex, the Asn229 side chain forms a cyclic hydrogen bonding network with N_3 and O_4 of the uracil base (Figure 1).[9] It has been proposed that the discrimination between dUMP and dCMP results from unfavorable interactions between dCMP and the side chain of Asn229 (Figure 1);[7] rotation of the Asn229 side chain, which could better accommodate dCMP, is presumably too costly since Asn229 is hydrogen bonded to Gln217 and to an ordered water molecule as part of a larger hydrogen bonding network. It is also not clear how the replacement of Asn229 with hydrophobic residues unable to participate in similar hydrogen bonds can preserve the tight binding of dUMP and cause the binding of dCMP to be as tight as that of dUMP. Similar mutagenesis experiments have been performed with Asn177 of *Escherichia coli* TS.[10] Most mutants of Asn177

[†] Department of Biochemistry and Biophysics.
[‡] Department of Pharmaceutical Chemistry.
[§] Current address: Dipartimento di Scienze Farmaceutiche, Universita di Modena, Via Campi 183, 41100 Modena, Italy.
[⊥] Current address: Procept, Inc., Cambridge, MA.
[⊗] Abstract published in *Advance ACS Abstracts,* June 1, 1995.

(1) Santi, D. V.; Danenberg, P. V. In *Folates and Pterins*; Blakley, R. L., Benkovic, S. J., Eds.; John Wiley and Sons: New York, 1984; pp 345−398.
(2) Santi, D. V.; McHenry, C. S.; Raines, R. T.; Ivanetich, K. M. *Biochemistry* **1987**, *26*, 8606−8613.
(3) Hardy, L. W.; Finer-Moore, J. S.; Montford, W. R.; Jones, M. O.; Santi, D. V.; Stroud, R. M. *Science* **1987**, *235*, 448−455.
(4) (a) Montford, W. R.; Perry, K. M.; Fauman, E. B.; Finer-Moore, J. S.; Maley, G. F.; Hardy, L. W.; Maley, F. ; Stroud, R. M. *Biochemistry* **1990**, *29*, 6964−6977. (b) Finer-Moore, J. S.; Montford, W. R.; Stroud, R. M. *Biochemistry* **1990**, *29*, 6977−6986. (c) Perry, K. M.; Fauman, E. B.; Finer-Moore, J. S.; Montford, W. R.; Maley, G. F.; Maley, F.; Stroud, R. M. *Proteins* **1990**, *8*, 315−333.
(5) (a) Matthews, D. A.; Appelt, K.; Oatley, S. J.; Xuong, N. H. *J. Mol. Biol.* **1990**, *214*, 923-936. (b) Matthews, D. A.; Villafranca, J. E.; Janson, P. A.; Smith, W. W.; Welsh, K.; Freer, S. *J. Mol. Biol.* **1990**, *214*, 937−948.
(6) Climie, S.; Ruiz-Perez, L.; Gonzalez-Pacanowska, D.; Prapunwattana, P.; Cho, S. W.; Stroud, R.; Santi, D. V. *J. Biol. Chem.* **1990**, *265*, 18776−18779.
(7) Liu, L.; Santi, D. V. *Biochemistry* **1992**, *31*, 5100−5104.
(8) (a) Liu, L.; Santi, D. V. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8604−8608. (b) Liu, L.; Santi, D. V. *Biochemistry* **1993**, *32*, 9263−9267.
(9) Finer-Moore, J. S.; Fauman, E. B.; Foster, P. G.; Perry, K. M.; Santi, D. V.; Stroud, R. *J. Mol. Biol.* **1993**, *232*, 1101−1116.

0002-7863/95/1517-7213$09.00/0

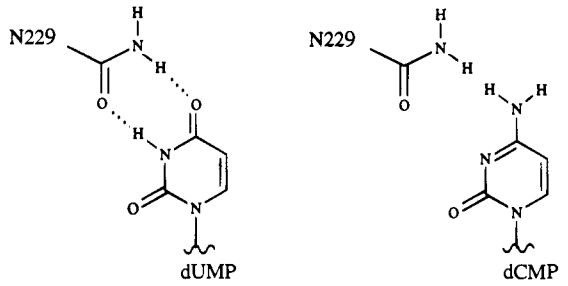

Figure 1. Schematic diagram of the hydrogen bonding between dUMP and the asparagine 229 side chain in the wild-type TS·dUMP complex. Simple static replacement of dUMP with dCMP would cause repulsion with the asparagine side chain.

lowered the activity of *E. coli* TS by at least 100-fold,[11] except for N177D, which converted *E. coli* TS into a dCMP methylase.[12]

In this study we use the free energy perturbation (FEP) method[13] to calculate the relative binding free energies of dUMP and dCMP with wild-type TS and two mutants. The N229D mutant was chosen because of its ability to catalyze the methylation of dCMP. The N229V mutant was chosen to investigate the effect of a hydrophobic residue, valine, on the binding energies of dUMP and dCMP, and because dCMP binds tighter than dUMP to the N229V mutant. These molecular dynamics (MD) simulations provide further insight into the binding modes of dUMP and dCMP to wild-type TS and to the N229D and N229V mutants.

Computational Method

All MD calculations were performed using the AMBER all-atom force field and the AMBER 4.0 molecular dynamics programs.[14] During all simulations, all bond lengths were constrained using the SHAKE[15] algorithm with a tolerance of 0.005 Å, allowing a time step of 2.0 fs. Solute and solvent were coupled to a constant temperature heat bath with a coupling constant of 0.2 ps to maintain a temperature of 300 K. A residue-based cutoff of 8.0 Å was employed. Pairlists were generated every 25 time steps. The entire substrate was used as the perturbed group in the FEP calculations. No intraperturbed group contributions to the free energy were calculated.

The thermodynamic perturbation method was used to calculate the free energy differences.[13] The thermodynamic cycle relevant to this study is shown in Figure 2. The free energies of complexation, ΔG_1 and ΔG_2, of dUMP and dCMP are experimentally determinable values, but they are difficult to calculate. ΔG_{sol} corresponds to mutating dUMP into dCMP in water, while ΔG_{prot} corresponds to changing dUMP into dCMP while bound to TS in water. Since the thermodynamic cycle is closed and free energy is a state function, $\Delta\Delta G_{bind} = \Delta G_1 - \Delta G_2 = \Delta G_{sol} - \Delta G_{prot}$. While ΔG_{prot} and ΔG_{sol} represent physically unrealizable processes, these two simulations are computationally manageable. The experimental free energies of binding for dUMP and dCMP to TS

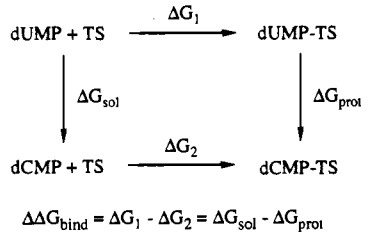

Figure 2. Thermodynamic cycle for complex formations of dUMP and dCMP with TS. The horizontal free energies are experimentally determined, and the vertical ones are calculated.

Table 1. Experimental Free Energies of Binding (ΔG_{bind}) and Relative Differences between dUMP and dCMP ($\Delta\Delta G_{bind}$)[a]

	ΔG_{bind}	$\Delta\Delta G_{bind}$
wt-TS + dUMP	−8.81	−3.60
wt-TS + dCMP	−5.21	
N229D + dUMP	−7.64	-0.02
N229D + dCMP	−7.62	
N229V + dUMP	−5.78	1.50
N229V + dCMP	−7.28	

[a] In kcal/mol. $\Delta\Delta G_{bind} = \Delta G_{bind}(dUMP) - \Delta G_{bind}(dCMP)$.

and the selected mutants are reported in Table 1, as calculated from the dissociation constants (K_D) of their binary complexes.[8]

The crystal structure of the *L. casei* TS·dUMP binary complex was solved to a resolution of 2.55 Å.[9] At the time this work was carried out, the only crystal structure of a binary complex available was that of the TS·dUMP complex. Briefly, it consists of two identical monomers with 316 residues per monomer; components of each monomer contribute to each of the two active sites. One molecule of dUMP is noncovalently bound to each active site. The side chain atoms of Asp111 and Asp111′ are not included in the X-ray coordinates due to disorder; they have been generated using the AMBER internal coordinate database. Hydrogens were added to the protein using the stored internal coordinates of the AMBER all-atom data base and then minimized, keeping the heavy atoms of the protein fixed at their original positions. The initial structures of the mutants were generated from the X-ray structure of the binary complex by replacing the Asn229 with Asp or Val.

All Lys and Arg residues were positively charged and Glu and Asp residues negatively charged, except where noted. The His residues were treated as unprotonated, since the pK_a of histidine is lower than the pH at which the binding assays were performed. The δH or ϵH tautomeric forms of histidines were assigned on the basis of favorable interactions with their environments. In the simulation of the N229D mutant, calculations are presented for the δH, ϵH, and protonated forms of His199 because His199 could act in synergism with Asp229 to determine the effective binding of dUMP and dCMP to the N229D mutant. In TS and the N229V mutant, the δH tautomer of His199 was employed. The parameters for the neutral Asp were taken from ref 16 . Two relative configurations of the neutral Asp229 side chain were initially constructed (Figure 3): configuration 1 places OH of Asp229 next to O_4 of dUMP and O of Asp229 next to HN_3 (a situation optimal for dUMP but unfavorable for dCMP); configuration 2 places OH of Asp next to HN_3 of dUMP and O of Asp next to O_4 (a situation unfavorable for dUMP but optimal for dCMP). Interestingly, each configuration mutally favors the binding of only one nucleotide. Calculations were started by equilibrating the optimal complexes (N229D·dUMP for configuration 1 and N229D·dCMP for configuration 2) and perturbing them in opposite directions (Figure 3).

Five crystallographic water molecules were maintained in the initial structure of TS; these ordered waters connect the dUMP base with the vicinal residues Glu60, His199, Asn229, and Ser232, and are probably important for maintaining the proper orientation of dUMP. The active site was further solvated with a spherical cap of 161 TIP3P[17] water

(10) In *L. casei* TS, the pyrimidine binding residue is N229. TSs from other sources have different numbering systems. In *E. coli* TS, the pyrimidine binding residue is N117. The numbering system for *L. casei* TS will be used in this paper when referring to specific residues unless otherwise stated.

(11) Michaels, M. L.; Kim, C. W.; Matthews, D. A.; Miller, J. H. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 3957−3961.

(12) Hardy, L. W.; Nalivaika, E. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 9725−9729.

(13) (a) Bash, P. A.; Singh, U.; Langridge, R.; Kollman, P. A. *Science* **1987**, *236*, 564−568. (b) McCammon, J. A. *Science* **1987**, *238*, 486−491. (c) Jorgensen, W. L.; Briggs, J. M. *J. Am. Chem. Soc.* **1989**, *111*, 4190−4197. (d) Kollman, P. A. *Chem. Rev.* **1993**, *93*, 2395−2417.

(14) Pearlman, D. A.; Case, D. A.; Caldwell, J. A.; Seibel, G. L.; Singh, U. C.; Weiner, P.; Kollman, P. A. *AMBER 4.0*; University of California, San Francisco, 1991.

(15) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977**, *34*, 1311−1327.

(16) Ferguson, D. M.; Radmer, R. J.; Kollman P. A. *J. Med. Chem.* **1991**, *34*, 2654−2659.

(17) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem.Phys.* **1983**, *79*, 926−935.

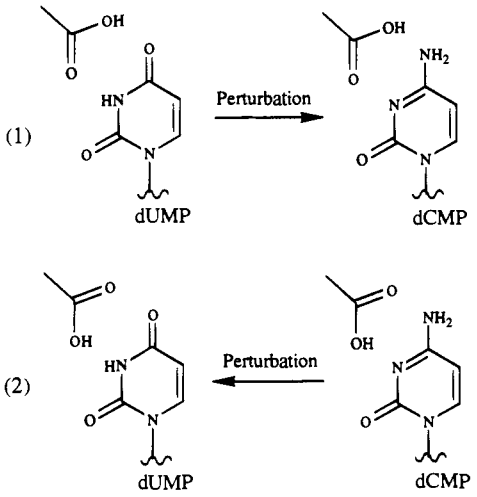

Figure 3. Schematic diagram showing the two relative configurations adopted for dUMP and dCMP interacting with the neutral aspartic acid mutant of TS (N229D). The perturbations have been carried out in the directions of the arrows.

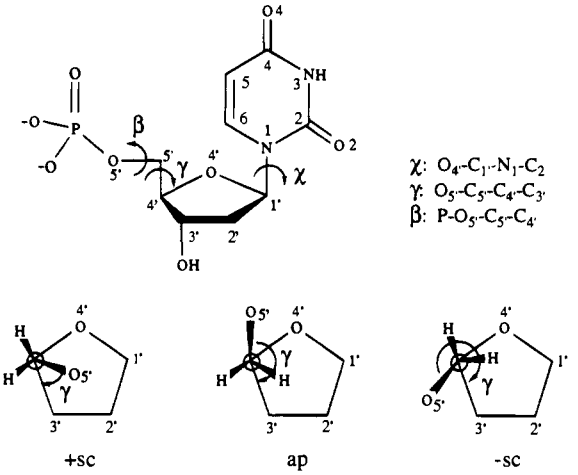

Figure 4. Torsion angles describing the conformation of dUMP.

molecules within 20 Å of the center of mass of dUMP, and harmonic radial forces (1.5 kcal/mol Å^2) were applied to these waters to avoid evaporation. dUMP, all residues within 12 Å of each atom of dUMP, and all the water molecules were allowed to move during the simulations. This selection resulted in a large moveable zone (127 protein residues) since dUMP is a fairly long molecule.

In order to carry out the perturbation of dUMP into dCMP in the binary complex, the initial structure was minimized and equilibrated. The energy minimization was performed by optimizing the water molecules first, keeping the protein and dUMP frozen. Next, all the residues within the belly and all the waters were optimized until the root-mean-square value of the potential gradients was <0.1 kcal/mol Å. The minimized structure was equilibrated for 25 ps at 300 K. Initially, a 5 kcal/mol constraint was applied to the $O_{4(dUMP)}$—H($N_{\delta 2}$/Asn229) and the $HN_{3(dUMP)}$—$O_{\delta 1}$(Asn229) distances; Cartesian coordinate restraints were also applied during the heating to 300 K and gradually removed in order to reduce the initial atomic velocities. The last 10 ps of equilibration was completely free of constraints. The perturbation of dUMP into dCMP bound to TS was carried out in a series of 41 windows with $\Delta\lambda = 0.025$. At each λ, 250 steps each (250/250) of equilibration/data collection (eq/dc) were performed, yielding a total simulation time of 41.0 ps. The N229D·dUMP and N229V·dUMP complexes were minimized and equilibrated following the same procedure used for the TS·dUMP binary complex. The FEP calculations for perturbation of dUMP into dCMP bound to the Asn229 mutants also followed the same procedure used for TS.

In a recent paper by Reddy *et al.*,[18] the authors included all nonbonded interactions among solute atoms and charged protein residues in the calculation of the binding free energy difference of two inhibitors to TS, so as to eliminate fluctuations during the MD simulation. In order to test the dependence of our free energy results on the nonbonded cutoff in presence of charged residues, a dUMP → dCMP perturbation was also performed using an infinite cutoff for the perturbed group (dUMP) and a neutralized protein. At this end, counterions (Na^+, Cl^-) were added to the TS·dUMP binary complex, and the neutralized system was minimized and equilibrated following the procedure already described. The parameters for the Na^+ and Cl^- counterions were taken from the work of Åqvist[19] and Jorgensen,[20] respectively. An 8 Å cutoff was used for the entire system except for dUMP, which has an infinite cutoff. The dUMP → dCMP perturbation in these new conditions was performed with 250/250 steps eq/dc and also with 500/500 steps eq/dc. These last calculations were done with a modified version of the GIBBS module.[21]

Since repulsive interactions between dCMP and Asn229 in wt-TS could be converted into attractive interactions by rotating the Asn229 side chain (Figure 1), a MD simulation was also run on a TS·dCMP complex having this alternative conformation; this simulation was run for 50 ps.

Five different simulations were performed to obtain the relative free energy of solvation of dUMP and dCMP. In simulation 1, 2, 3, and 4, dUMP was solvated in a periodic box of 570 TIP3P waters, resulting in an initial box size of 32.2 Å × 26.4 Å × 24.0 Å. In simulation 5, N_1-methyluracil was solvated in a periodic box of 395 TIP3P waters, resulting in an initial box size of 26.5 Å × 24.5 Å × 21.8 Å. In simulations 1−3 and in all TS simulations, the phosphate groups on dUMP and dCMP were made dianionic since NMR studies indicated that this was the form of dUMP bound to TS.[22] The charges on the phosphate were obtained from an electrostatic potential fit[23] to a STO-3G wave function of $CH_3OPO_3^{2-}$ and then adjusted to give a total 2− net charge on dUMP. A 5 kcal/mol harmonic constraint was applied on the rotation of χ in order to maintain the anti conformation. In each simulation, the perturbation was carried out in a series of 41 windows with $\Delta\lambda = 0.025$. Both forward (dUMP → dCMP) and reverse (dCMP → dUMP) simulations were performed. The reverse simulation was preceded by 5 ps equilibration at $\lambda = 1$. In simulations 1, 2, and 3, 250/250, 500/500, and 1000/1000 steps of eq/dc were performed, respectively, resulting in total simulation times of 41.0, 82.0, and 164 ps. In simulation 4, the charges on the phosphate group were reduced to yield a total charge of 1− on dUMP and dCMP. The constraint on χ was not necessary. 250/250 steps of eq/dc were performed. In simulation 5, N_1-methyluracil was mutated into N_1-methylcytosine. The system was equilibrated using molecular dynamics for 20 ps. Again, 250/250 steps of eq/dc were performed at each λ. In all simulations, the initial solvated system was energy minimized until the root-mean-square value of the potential gradients was <0.1 kcal/mol Å.

The calculations were run on the HP-735 computers in the Structural Biology Group at the University of California, San Francisco. Input files for the calculations are available upon request from the authors.

Results

dUMP and dCMP in Solution. The most relevant torsion angles describing the nucleotide conformation are shown in Figure 4. The torsion angle χ determines the orientation of the base relative to the sugar. It can adopt two main orientations, syn and anti. The torsion angle γ allows $O_{5'}$ to assume different positions relative to the sugar and can adopt the three staggered

(18) Reddy, M. R.; Bacquet, R. J.; Zichi, D.; Matthews, D. A.; Welsh, K. M.; Jones, T. R.; Freer, S. *J. Am. Chem. Soc.* **1992**, *114*, 10117−10122.
(19) Åqvist, J. *J. Phys. Chem.* **1990**, *94*, 8021−8024.
(20) Jorgensen, W. L.; Buckner, J. K.; Huston, S. E.; Rossky, P. J. *J. Am. Chem. Soc.* **1987**, *109*, 1891−1899.
(21) Cieplak, P.; Kollman, P. A. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 291−304.
(22) Beckage, M. J.; Blumenstein, M.; Kisliuk, R. L. *Biochim. Biophys. Acta* **1979**, *571*, 157−161.
(23) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129−144.

Figure 5. Stereoview of the conformation of dUMP in its first hydration shell from the unconstrained MD simulation in water. The base of dUMP rotates in a syn conformation with the constitution of a hydrophilic pocket made of O_2, $O_{4'}$, $O_{5'}$, and PO_3^{2-}.

Table 2. Solvation Free Energy Differences (kcal/mol) between dUMP and dCMP Calculated According to Different Conditions (See Text)

conditions	ΔG_{sol}		
	ΔG_{for} (dUMP → dCMP)	ΔG_{rev} (dUMP → dCMP)	ΔG_{av}
simulation 1 250/250 steps eq/dc 5 kcal constraint on χ	−9.31	9.40	−9.36 ± 0.05
simulation 2 500/500 steps eq/dc 5 kcal constraint on χ	−8.09	10.65	−9.37 ± 1.28
simulation 3 1000/1000 steps eq/dc 5 kcal constraint on χ	−10.23	10.13	−10.18 ± 0.05
simulation 4 250/250 steps eq/dc Charge on the phosphate reduced to 1−; no constraint on χ	−6.16	8.87	−7.52 ± 1.36
simulation 5 N_1-methyluracil into N_1-methylcytosine	−5.77	5.93	−5.85 ± 0.08

positions, +sc, ap, −sc (Figure 4). β is usually trans. χ and γ, as well as sugar puckering, are known to be interdependent in determining the conformations of nucleotides and nucleosides.[24]

In the crystal structure of the dUMP·TS binary complex, dUMP adopts a conformation with χ = anti and γ = −sc; the puckering of the sugar is O_4'-endo. After minimization in the solvent bath, dUMP maintains the same conformation, except that the furanose puckering becomes C_2'-endo. During equilibration with MD, dUMP undergoes conformational changes such that γ = ap and χ turns progressively to a high-anti and syn conformation. It is known that γ tends to assume either +sc or ap positions in nucleosides and nucleotides, the +sc being preferred for nucleotides.[24] In order to check if the γ = +sc conformation could be obtained for dUMP during MD, a constraint that γ = +sc was applied to the dUMP and progressively decreased during the equilibration process; as soon as the constraint was removed, dUMP returned to the γ = ap conformation, which must therefore be the preferred conformation for this torsional angle of dUMP in solution. Conformational studies on nucleosides also revealed that γ assumes the ap conformation preferentially when explicit solvent is used for calculations.[25] Since the γ = −sc conformation of dUMP in the binary complex with TS is rarely observed in isolated nucleotides and nucleosides, it is therefore apparently induced by the conformational requirements for the binding of the nucleotide to the active site of the protein.

The observed anti/syn conformational change of dUMP during long MD simulations is closely related to the large negative charge of the phosphate, which induces the rotation of the base and causes the formation of a hydrophilic pocket consisting of O_2, $O_{4'}$, $O_{5'}$, and PO_3^{2-}. This hydrophilic pocket is shown in Figure 5, along with the first hydration shell. The rotation of χ was not observed when the charge of the phosphate was reduced to 1−, in which simulation χ remains anti and γ is ap. Pyridine nucleotides are known to significantly prefer the anti conformation.[24] Therefore, a 5 kcal/mol harmonic constraint was applied on the rotation of χ in order to maintain the anti conformation. The equilibrium value of χ was set to −150°, which was taken from the unconstrained minimization of dUMP in the solvent bath. This constraint was necessary since the FEP results for dUMP/dCMP perturbation in water proved to be very dependent on the orientation of the base (which is effectively perturbed) with respect to the rest of the solute. This orientation is particularly important since it also determines the position of the base relative to the highly negative phosphate group, which is able to strongly polarize the surrounding water molecules. In this sense, different positions of the base with respect to the phosphate will "see" different distributions of waters. Under the conditions described, 20 ps of equilibration was performed with a 5 kcal/mol constraint on χ. The free energy perturbation of dUMP into dCMP in water was carried out starting from this equilibrated structure, and the constraint on χ was maintained for the whole perturbation.

The results of the FEP calculations for each simulation are given in Table 2. The averaged energies are reported with negative sign according to the sign of the transformation in the direction dUMP to dCMP.

The results from simulation 1 predict dCMP to be better solvated than dUMP by 9.36 kcal/mol. A longer simulation run (82 ps) was carried out using 500/500 of eq/dc at each λ (simulation 2). While the hysteresis is significantly higher in simulation 2, these two simulations yield almost identical values of ΔG_{sol}. In both simulations, γ is always ap and β is always trans (no conformational changes occur at the phosphate moiety). In order to check the convergence of these solvation results, an additional simulation was performed for twice as long as simulation 2 (simulation 3, 164 ps). The free energy of solvation

(24) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1983.
(25) Pearlman, D. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1991**, *113*, 7167−7177.

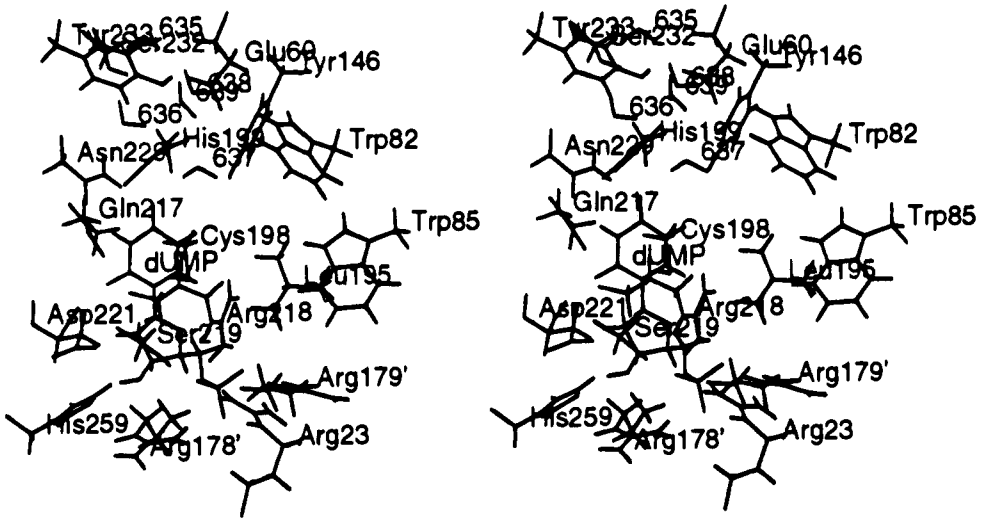

Figure 6. Stereoview of the main residue side chains surrounding dUMP in the crystal structure of the TS·dUMP binary complex. The main hydrogen bonding distances (Å) follow: O_4–H($N_{\delta 2}$)Asn229, 1.93; HN_3–$O_{\delta 1}$Asn229, 2.06; O_4–H(WAT637), 2.01; O_2–HNAsp221, 1.88; $HO_{3'}$–$N_{\epsilon 2}$His259, 1.87; $HO_{3'}$–OHTyr261, 2.99; $O_A(PO_3)$–HN_{21}Arg178′, 1.92; $O_A(PO_3)$–HN_{ϵ}Arg23, 1.82; $O_C(PO_3)$–HN_{ϵ}Arg23, 3.00; $O_C(PO_3)$–HN_{ϵ}-Arg179′, 2.15; $O_C(PO_3)$–HN_{11}Arg179′, 2.01; $O_B(PO_3)$–HN_{11}Arg218, 1.81; $O_C(PO_3)$–HN_{12}Arg218, 1.85; $O_B(PO_3)$–HO_{γ}Ser219, 1.83; $O_{5'}$-HN_{11}Arg23, 1.70.

is −10.18 kcal/mol, and a low hysteresis is obtained (0.05 kcal/mol). It is worth noting that despite the different lengths of these simulations and the need of a constraint on the χ angle, the results from simulations 1−3 fall within a reasonable range.

Simulation 4 was an attempt to estimate the effect of the charge of the phosphate on the result of the perturbation. As already mentioned, reduction of the charge on the phosphate to yield a total charge of 1− on dUMP resulted in χ not undergoing a conformational change, thus no constraint on this torsional angle was necessary for the FEP calculations. In this simulation, the ΔG_{sol} is −7.52 kcal/mol, which is lower than the value obtained for the constrained simulation with the 2− net charge (simulations 1−3). Even if this value cannot be directly compared to those of simulations 1−3 because the total charge on the molecule is different, it demonstrates that the constraint on χ is necessary when the 2− net charge is present on dUMP. In fact, the unconstrained simulations with a 2− net charge on dUMP gave huge solvation free energy differences (∼−19 kcal/mol), which are out of the range commonly associated with such perturbations. The large hysteresis obtained in simulation 4 is probably due to a change in the β torsion angle during the reverse run which does not occur during the forward simulation. During the reverse simulation, the β torsion angle turns from trans, as it was in the forward simulation, to about 70°.

Simulation 5 involves the perturbation of N_1-methyluracil into N_1-methylcytosine. This simulation was run to investigate the effect of the sugar and the phosphate on the result of the perturbation, and also because no conformational changes are present in this simulation. A value of −5.85 kcal/mol was obtained, and a very low hysteresis (0.16 kcal/mol) resulted from forward/reverse simulations. This value is in agreement with previous calculations on similar bases[13a] and experiments.[26] ΔG_{sol} for simulation 5 is smaller than that obtained in the nucleotide simulations. The difference is likely due to the presence of the sugar and the phosphate.

Structures of the TS·dUMP and TS·dCMP Binary Complexes. The amino acid side chains surrounding dUMP in the crystal structure of the *L. casei* binary complex are shown in Figure 6. Five crystallographic waters (635−639) are also shown in Figure 6, along with a description of the hydrogen bonding interactions and distances (Å) between dUMP and the vicinal residues. O_4 of dUMP is hydrogen bonded to the H($N_{\delta 2}$) of Asn229, and the HN_3 of dUMP is hydrogen bonded to $O_{\delta 1}$ of Asn229; O_4 is also hydrogen bonded to WAT637. O_2 of dUMP is hydrogen bonded to the Asp221 backbone, and the 3′-OH of dUMP is hydrogen bonded to $N_{\epsilon 2}$ of His259 and the hydroxyl oxygen of Tyr261. The phosphate of dUMP is surrounded by four arginines (Arg23, 218, 178′, and 179′) and Ser219, which constitute the phosphate binding site.

Table 3. Selected Distances (Å) between dUMP and the Vicinal Side Chains of TS in the X-Ray, Minimized, and Equilibrated Structures

distance	X-ray	min	equil
O_4–Asn229H($N_{\delta 2}$)	1.93	1.96	1.96
HN_3–Asn229$O_{\delta 1}$	2.06	1.91	2.31
O_2–Asp221H(N)	1.88	1.94	1.96
C_6–Cys198S_{γ}	3.52	3.59	4.00
$HO_{3'}$–His259$N_{\epsilon 2}$	1.87	1.97	2.68
$HO_{3'}$–Tyr261OH	2.99	2.52	2.12
$O_A(PO_3)$–Arg178′H(N_{21})	1.92	1.76	1.80
$O_A(PO_3)$–Arg23H(N_{ϵ})	1.82	2.07	3.17
$O_B(PO_3)$–Arg218H(N_{11})	1.81	1.63	1.56
$O_B(PO_3)$–Ser219H(O_{γ})	1.83	1.70	1.96
$O_C(PO_3)$–Arg23H(N_{ϵ})	3.00	2.53	2.46
$O_C(PO_3)$–Arg179′H(N_{ϵ})	2.15	1.74	1.81
$O_C(PO_3)$–Arg179′H(N_{11})	2.01	1.69	1.67
$O_C(PO_3)$–Arg218H(N_{12})	1.85	2.55	3.79
$O_{5'}$–Arg23H(N_{11})	1.70	1.85	2.24
$O_A(PO_3)$–WAT711H_1		1.68	1.80
$O_A(PO_3)$–WAT770H_1			1.91

The interaction distances between dUMP and the nearby side chains in the crystal structure are compared to the minimized and the equilibrated complexes in Table 3. All the hydrogen bonds involving dUMP are conserved in the minimized structure. Recall that water molecules were added to create a solution-like environment in the protein simulation; one of these waters moves to hydrogen bond the phosphate of dUMP. This is consistent with the observation that a water molecule was already present in this position in the X-ray structure.[9] Many other waters solvate the active site, which, in the simulation of the binary complex, is quite open and exposed to solvent. After minimization of the binary complex, the conformation of dUMP is χ = anti, γ = −sc, and β = trans, and the puckering of the sugar is $C_{2'}$-endo.

(26) Cullis, P. M.; Wolfenden, R. *Biochemistry* **1981**, *20*, 3024−3028.

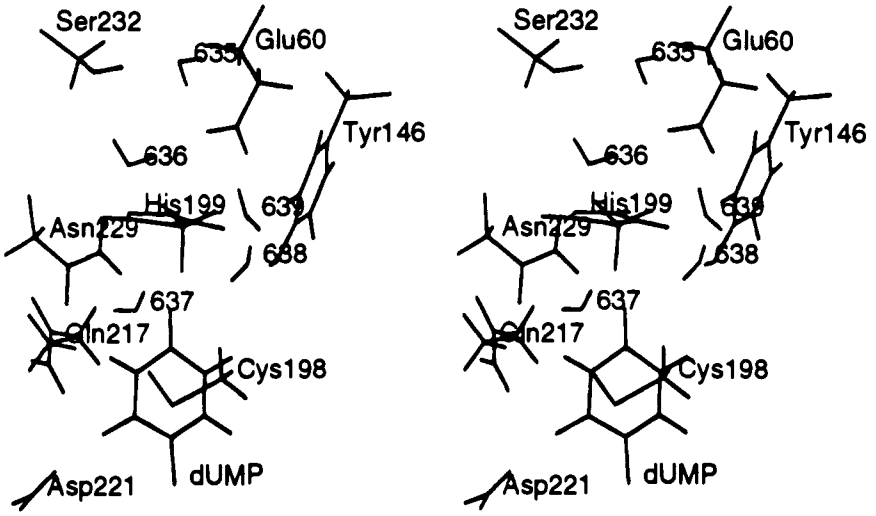

Figure 7. Stereoview of the equilibrated wild-type TS·dUMP complex. A selection of residues surrounding the base of dUMP is shown. The main hydrogen bonding distances (Å) follow: O(636)−$HN_{\delta 1}$Asn229, 2.22; H_1(636)−O_γSer232, 2.23; H_2(636)−$O_{\epsilon 2}$Glu60, 1.85; H_1(635)−$O_{\epsilon 2}$Glu60, 1.76; H_2(637)−$O_{\delta 1}$Asn229, 2.08; H_1(637)−$N_{\epsilon 2}$His199, 2.50; H_2(638)−O_4dUMP, 2.28; H_1(638)−O(639), 1.80; H_2(639)−$O_{\epsilon 2}$Glu60, 2.06; O(637)−HS_γCys198, 2.28; O(635)−HO_γSer232, 1.88; O_2dUMP−HN_3Asp221, 1.96; OHTyr146−HN_δHis199, 2.53; HN_3dUMP−$O_{\delta 1}$Asn229, 2.31; O_4dUMP−$HN_{\delta 2}$Asn229, 1.96; S_γCys198−C_6dUMP, 4.00; O(638)−HN_δHis199, 2.91.

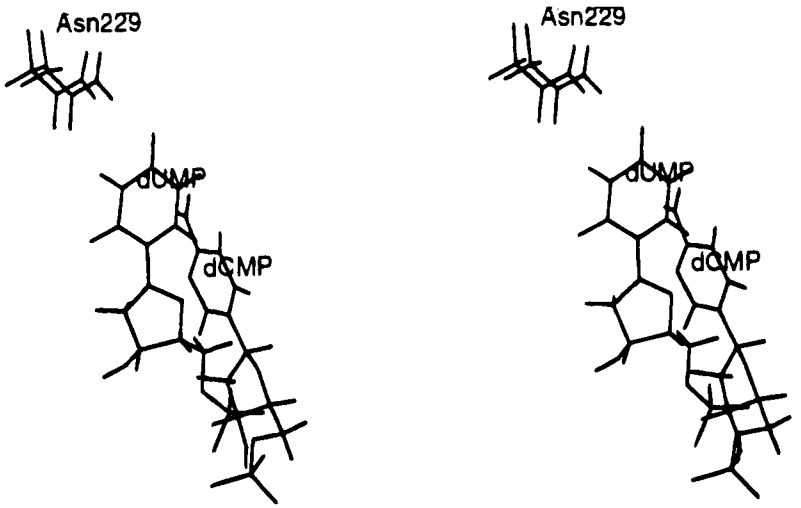

Figure 8. Stereoview of the superimposition of the equilibrated structures of the TS·dUMP and the TS·dCMP complexes. Only the two nucleotides and the corresponding asparagine 229 side chains are shown. Repulsion between dCMP and the asparagine side chain is responsible for the movement of the base away from Asn229.

The binary complex was equilibrated for 25 ps. All the hydrogen bonds between dUMP and the side chains of TS are conserved after equilibration. The 3′-OH of dUMP is hydrogen bonded to the hydroxyl oxygen of Tyr261 (d = 2.12 Å) and to $N_{\epsilon 2}$ of His259 (d = 2.68 Å); Tyr261 and His259 seem to be exchangeable in their ability to provide a strong hydrogen bond the 3′-OH of dUMP, the preferred one probably depending on the conformation of the sugar. Another water molecule of the cap (WAT770) has moved toward the phosphate of dUMP (Table 3). The conformation of dUMP after equilibration is identical to that of the minimized complex.

The analysis of the solvation shell surrounding the uracil base of dUMP after equilibration revealed that WAT635, 636, and 639 are found in the same position as in the minimized complex, but WAT637 becomes hydrogen bonded to the HS_γ of Cys198 and also to $O_{\delta 1}$ of Asn229 (Figure 7). The C_α−C_β bond of Cys198 rotates about 90° in order to make the hydrogen bond with WAT637. It is noteworthy that despite this rotation, the distance between the sulfur of Cys198 and C_6 of dUMP (4.0 Å) and the S_γ−C_6−C_5 angle (90°) remain close to those of the crystal structure. WAT638 replaces WAT637 in hydrogen bonding to O_4 of dUMP. This rearrangement results from the dynamic behavior of the solvent molecules in the cavity. Inspection of the van der Waals volumes of this region of space in the crystal structure revealed that an empty space exists in the crystal structure in the position assumed by WAT637 after equilibration.

The structure of the equilibrated TS·dCMP complex provides an interesting contrast to that of the TS·dUMP complex. Repulsion between dCMP and the Asn229 side chain results in the movement of the base away from Asn229 (Figure 8). Since the active site is quite open and exposed to solvent, dCMP is able to move away from Asn229 while hydrogen bonding to water molecules present in the active site. The main contacts between the base of dCMP and TS are reported in Figure 9. The H_1 hydrogen of the 4-NH_2 group of dCMP is hydrogen bonded with WAT718 and 710. N_3 of dCMP is hydrogen bonded to WAT772. The region of space previously occupied by dUMP in the initial complex is now occupied by water molecules (Figure 9). The position of dCMP in the binary complex results in a C_6−S_γ distance of 7.0 Å and a S_γ−C_6−C_5 angle of 38°. Arg23 and Ser219 do not hydrogen bond to the phosphate of dCMP in this structure; Arg23 moves toward Val316 and is replaced by water molecules.

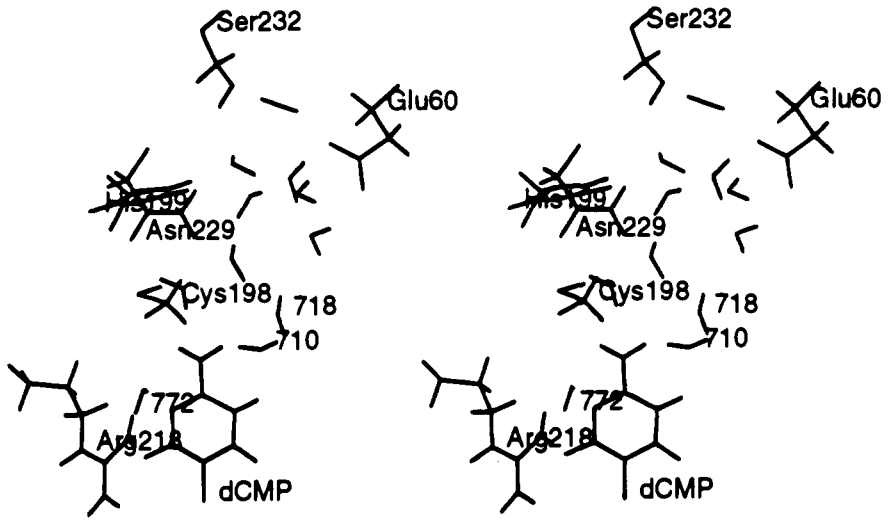

Figure 9. Stereoview of the equilibrated wild-type TS·dCMP complex. The main hydrogen bonding distances (Å) follow: O_2dCMP−HN_2Arg218, 1.95; N_3dCMP−H_1(772), 1.79; H_1(4-NH_2)dCMP−O(718), 2.68; H_1(4-NH_2)dCMP−O(710), 2.32; H_2(4-NH_2)dCMP−S_γCys198, 2.70; C_6dCMP−S_γCys198, 7.00.

Structure of the N229D·dUMP and N229D·dCMP Binary Complexes (Anionic Asp229). Mutation of Asn229 into the anionic form of Asp results in a negative charge distribution close to the His199 side chain, suggesting that prototropic tautomerism and/or protonation could take place at His199 in order to better accommodate Asp229. In order to investigate the effect of the protonation state of His199 of the N229D mutant on the relative binding of dUMP and dCMP, calculations were performed with the δH, ϵH, and protonated forms of His199.

(i) δH Tautomer of His199. The equilibrated structures of the N229D·dUMP and N229D·dCMP complexes with the δH tautomer of His199 are shown in Figure 10. In the N229D·dUMP binary complex, HN_3 of dUMP is hydrogen bonded to $O_{\delta 1}$ of Asp229. WAT702, not seen in the wt-TS−dUMP complex, bridges O_4 of dUMP and $O_{\delta 2}$ of Asp229. O_4 of dUMP remains hydrogen bonded to WAT704 as it was in the TS·dUMP complex (WAT638, Figure 7). WAT704 also hydrogen bonds to H(N_δ) of His199. Asp229 is connected to Ser232 via two water molecules (WAT702 and 639), and WAT635 bridges Ser232 and Glu60. The sulfur of Cys198 is 3.8 Å away from C_6 of dUMP and the $S_\gamma-C_6-C_5$ angle is 72°, placing the sulfur in a favorable position for nucleophilic attack. The hydrogen bonds between the phosphate of dUMP and the enzyme are all preserved.

In the equilibrated structure of the N229D·dCMP complex, repulsive interactions between $O_{\delta 1}$ of Asp229 and N_3 of dCMP cause the Asp229 side chain to move in the direction of Ser232. As a consequence, the hydrogen bonding between Gln217 and Asp229 is broken, and Gln217 rotates in order to better accommodate N_ϵ of His199. The 4-NH_2 hydrogens of dCMP are surrounded by WAT709 and by the $O_{\delta 1}$ of Asp229. The HSγ of Cys198 is hydrogen bonded to N_3 of dCMP. The distance between the sulfur of Cys198 and C_6 of dCMP is 4.3 Å, and the $S_\gamma-C_6-C_5$ angle is 67°. The phosphate−enzyme interactions found in the wt-TS−dUMP complex are preserved.

(ii) ϵH Tautomer of His199. The equilibrated structures of the N229D·dUMP and the N229D·dCMP binary complexes with the ϵH tautomer of His199 are shown in Figure 11. In the N229D·dUMP complex, the ϵH (His199) is hydrogen bonded to $O_{\delta 2}$ of Asp229, as expected. WAT794 bridges O_4 of dUMP and $O_{\delta 2}$ of Asp229 as in the previous structure (Figure 10). Again, HN_3 of dUMP is hydrogen bonded to $O_{\delta 1}$ of Asp229. Interestingly, the water molecule usually found hydrogen bonded to O_4 of dUMP (WAT638 in Figure 7, and WAT704 in Figure 10) is not present in this simulation. The ϵH tautomer of His199 induces a different distribution of water molecules in the region between His199 and dUMP. The N_δ nitrogen of His199 is hydrogen bonded to the hydroxyl of Tyr146, and some water molecules are displaced. Finally, the distance between the sulfur of Cys198 and C_6 of dUMP is 3.8 Å, and the $S_\gamma-C_6-C_5$ angle is 79°. The phosphate binding residues are all positioned to interact with the phosphate.

Analysis of the N229D·dCMP structure using the ϵH tautomer of His199 (Figure 11) reveals that dCMP undergoes a remarkable shift from Asp229. The region previously occupied by dUMP is now occupied by water molecules. O_2 of dCMP is hydrogen bonded to Arg218. Asp229 and His199 are still found hydrogen bonded and connected with Ser232 and Glu60 via water molecules. The phosphate is still hydrogen bonded by the four arginine residues and Ser219.

(iii) Protonated His199. The equilibrated structures of the N229D·dUMP and N229D·dCMP binary complexes obtained using protonated His199 are shown in Figure 12. The presence of both δH and ϵH protons allows His199 to hydrogen bond with Asp229 and with the localized water molecules that are present in the structures obtained with TS and with the δH tautomer of His199. In fact, WAT637 is correctly placed next to O_4 of dUMP and also hydrogen bonds to Tyr146, which in turn hydrogen bonds to His199. Furthermore, the ϵH of His199 bridges $O_{\delta 2}$ of Asp229 and O_4 of dUMP instead of the WAT702 and WAT794 of Figures 10 and 11. Finally, the S_γ of Cys198 and C_6 of dUMP are 4 Å apart, with a $S_\gamma-C_6-C_5$ angle of 79°.

In the N229D·dCMP complex, dCMP is still close to the Asp229 side chain and in a good position for the nucleophilic attack by Cys198 (the $S_\gamma-C_6$ distance is 3.6 Å, and the $S_\gamma-C_6-C_5$ angle is 80°). Importantly, the 4-NH_2 hydrogens of dCMP are hydrogen bonded to the $O_{\delta 1}$ of Asp229 and to WAT795. In this latter complex, the Arg23 usually interacting with the phosphate has been replaced with water molecules and is found close to Val316. Arg178′, which was directly bound to dUMP, hydrogen bonds to the phosphate of dCMP via a water molecule.

Structure of the N229D·dUMP and N229D·dCMP Binary Complexes (Neutral Asp229). The Asp side chain has a pK_a of about 4 and is anionic at physiological pH. However, it has been proposed that methylation of dCMP proceeds through

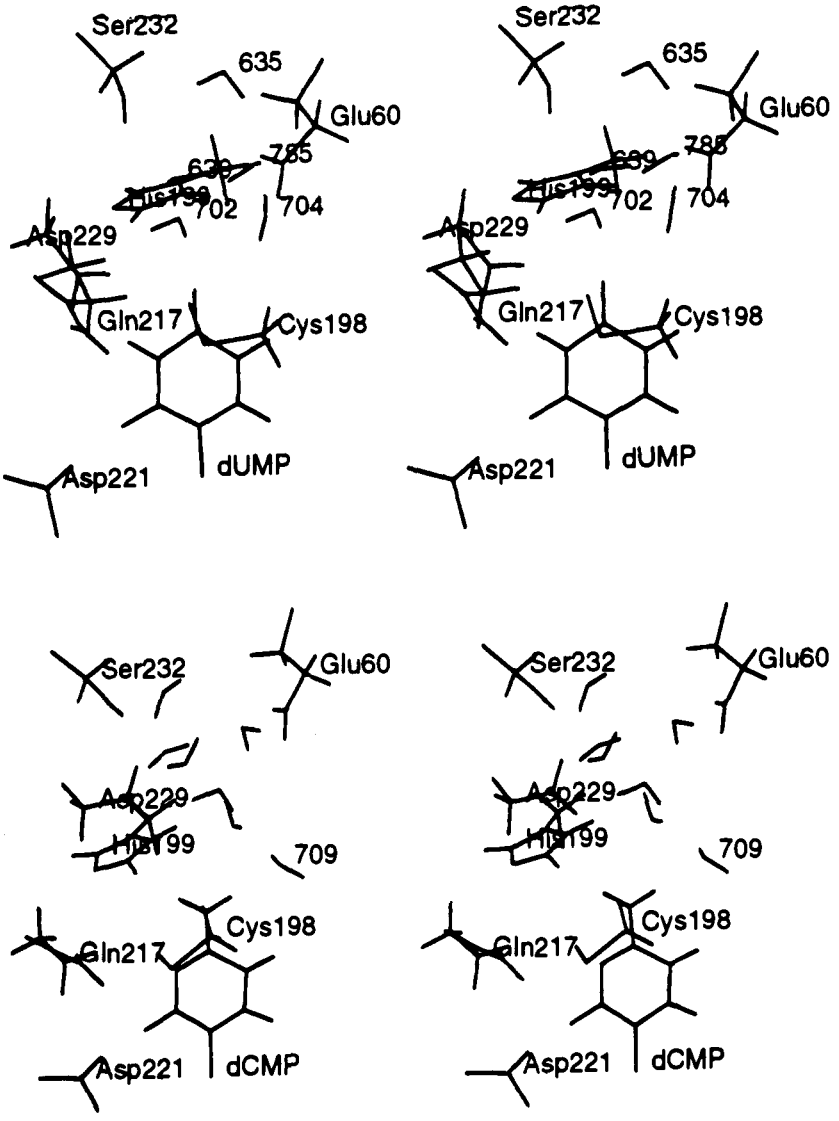

Figure 10. Stereoviews of the equilibrated N229D·dUMP (upper) and N229D·dCMP (lower) complexes obtained with the anionic aspartate 229 side chain and the δH tautomer of histidine 199. The main hydrogen bonding distances (Å) follow. For N229D·dUMP: O_2dUMP−HNAsp221, 2.02; HN_3dUMP−$O_{\delta 1}$Asp229, 1.94; H_2(702)−$O_{\delta 2}$Asp229, 1.84; H_1(702)−O_4dUMP, 2.48; H_2(704)−O_4dUMP, 2.48; H_1(639)−O(702), 2.22; O(639)−HO_γSer232, 2.21; O(704)−HN_δHis199, 1.92; H_1(704)−O(785), 1.88; H_1(785)−$O_{\epsilon 2}$Glu60, 1.75; HS_γCys198−C_6dUMP, 3.76. For N22D·dCMP: O_2dCMP−HNAsp221, 1.89; N_3dCMP−HS_γCys198, 2.26; H_1(4-NH_2)dCMP−O(709), 2.07; H_2(4-NH_2)dCMP−$O_{\delta 2}$Asp229, 3.00; C_6dCMP−S_γCys198, 4.30.

protonated Asp229 with transfer of the proton to N_3 of dCMP during catalysis.[7] In order to investigate the effect that a neutral Asp229 side chain has on the binding of dUMP and dCMP, calculations were also performed for the neutral Asp229 in two different configurations (Figure 3): configuration 1 is optimal for dUMP but unfavorable for dCMP, while configuration 2 is unfavorable for dUMP but optimal for dCMP.

Configuration 1. The equilibrated structures of the N229D·dUMP and N229D·dCMP complexes obtained by using the neutral Asp229 residue in configuration 1 are shown in Figure 13. $O_{\delta 2}$ of Asp229 is hydrogen bonded to HN_3 of dUMP, and $HO_{\delta 1}$ is hydrogen bonded to O_4. Gln217 is hydrogen bonded $O_{\delta 2}$ of Asp229. A water molecule (WAT795) is hydrogen bonded to O_4 of dUMP, as in previous structures, and O_2 of dUMP hydrogen bonds to HN of the Asp221 backbone. HN_δ of His199 is hydrogen bonded to WAT639, which is also hydrogen bonded to WAT795. The 3′-OH of dUMP is hydrogen bonded to $N_{\epsilon 2}$ of His259. Ser219, Arg218, Arg23, and Arg179′ still bind the phosphate of dUMP. The distance between C_6 and S_γ of Cys198 is 3.72 Å, and the S_γ−C_6−C_5 angle is 76°. This situation appears favorable for the nucleophilic attack of Cys198 to C_6 of dUMP.

In configuration 1 of the N229D·dCMP structure, dCMP undergoes a remarkable movement away from Asp229 (Figure13); this movement is the result of the repulsive interactions between the neutral Asp229 side chain and the N_3 and 4-NH_2 of dCMP, as depicted in Figure 3. The Asp229 side chain does not rotate to better accommodate dCMP, and configuration 2 is not obtained from configuration 1. Moreover, the displacement of dCMP from the binding site is very similar to that of dCMP in the wt-TS simulation: in both cases, repulsive interactions with the side chain of residue 229 cause a considerable displacement of the nucleotide. WAT702 hydrogen bonds to H_1(4-NH_2) of dCMP; N_3 of dCMP is hydrogen bonded to the HN backbone of Asp221; H_2(4-NH_2) is close to the S_γ of Cys198. The 3′-OH of dCMP is hydrogen bonded to Tyr24 and a water molecule; only Arg218, Arg23, and Arg179′ bind the phosphate of dCMP in this complex. Arg178′ is displaced. Finally, the distance between C_6 of dCMP and S_γ of Cys198 is 6.3 Å, with a S_γ−C_6−C_5 angle of 45°, unsuitable for nucleophilic attack.

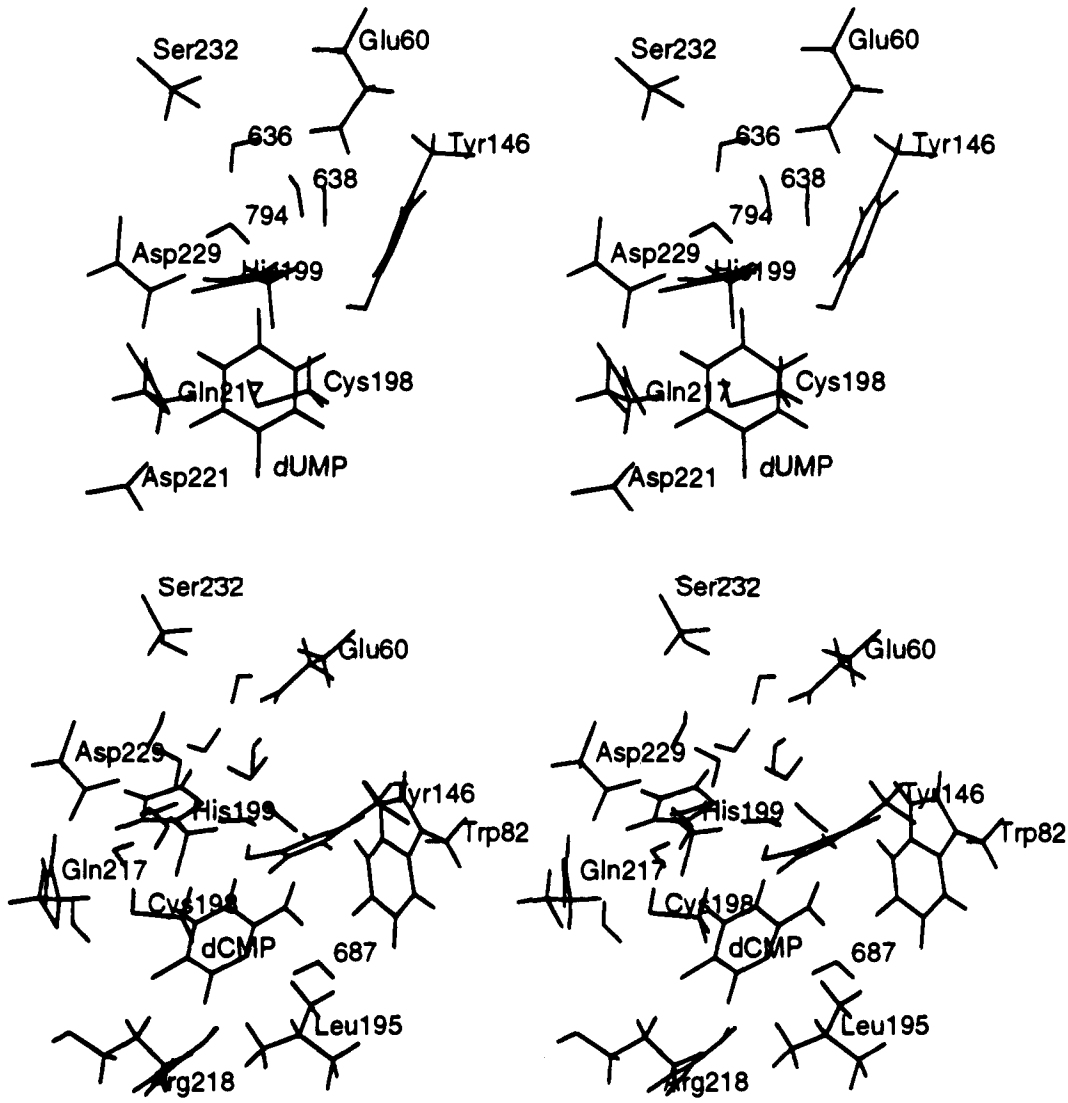

Figure 11. Stereoviews of the equilibrated N229D·dUMP (upper) and N229D·dCMP (lower) complexes obtained with the anionic aspartate 229 side chain and the ϵH tautomer of histidine 199. The main hydrogen bonding distances (Å) follow. For N229·dUMP: O_2dUMP−HNAsp221, 1.94; HN_3dUMP−$O_{\delta 1}$Asp229, 2.17; $H_2(794)$−$O_{\delta 2}$Asp229, 1.82; O_4dUMP−$H_1(794)$, 1.55; $O_{\delta 2}$Asp229−HN_{ϵ}His199, 1.73; $H_2(636)$−O(794), 2.45; O(636)−HO_{γ}Ser232, 2.10; $H_1(636)$−$O_{\epsilon 2}$Glu60, 1.72; $N_{\delta 1}$His199−HOHTyr146, 2.35; $H_1(638)$−$N_{\delta 1}$His199, 1.96; C_6dUMP−S_{γ}Cys198, 3.82. For N229D·dCMP: H_2(4-NH_2)dCMP−O(687), 1.83; N_3dCMP−$H_2(687)$, 2.50; HN_{ϵ}His199−$O_{\delta 2}$Asp229, 1.93; O_2dCMP−HN_{22}Arg218, 2.22; C_6dCMP−S_{γ}Cys198, 4.90.

Configuration 2. The equilibrated structures of the N229D·dCMP and N229D·dUMP complexes obtained starting from configuration 2 are shown in Figure 14. The expected optimal interactions between dCMP and the Asp229 side chain are apparent: $O_{\delta 1}$ of Asp229 is hydrogen bonded to 4-NH_2 of dCMP, and $HO_{\delta 2}$ is hydrogen bonded to N_3 of dCMP. The 4-NH_2 is also hydrogen bonded to a water molecule (WAT637), and O_2 is close to the Asp221 backbone. This complex results in a C_6−Sγ distance of 3.96 Å and a S_{γ}−C_6−C_5 angle of 76°, suitable for nucleophilic attack. The 3′-OH is hydrogen bonded to His259; the binding residues around the phosphate of dCMP are all present.

Perturbation of dCMP into dUMP in configuration 2 causes the displacement of dUMP from Asp229; this situation is exactly the opposite of that obtained with configuration 1, where dCMP is displaced. O_4 of dUMP is hydrogen bonded to WAT635 and O_2 to WAT738. Interestingly, the ϵH hydrogen of His199 is hydrogen bonded to Asp229. The 3′-OH of dUMP remains hydrogen bonded to His259; all the binding residues of the phosphate are present with the exception of Arg23.

Structures of the N229V·dUMP and N229V·dCMP Binary Complexes. The equilibrated structures of the N229V·dUMP and N229V·dCMP binary complexes are shown in Figure 15. These structures are similar to those of the TS complexes. While the substitution of the polar asparagine residue by a hydrophobic one does not result in a significant change in the location of the base in the active site, it does result in some changes in the interactions between the nucleotides and the active site residues.

In the N229V·dUMP binary complex, O_4 of dUMP hydrogen bonds to WAT648 and 705. O_2 of dUMP hydrogen bonds to the backbone of Asp221. The 3′-OH of dUMP is hydrogen bonded to the hydroxyl hydrogen of Tyr261, to N_{ϵ} of His259, and to WAT754. Cys198 lies above the pyrimidine ring, and S_{γ} interacts with the backbone of Ser219. This results in a S_{γ}−C_6(dUMP) distance of 3.89 Å and a S_{γ}−C_6−C_5 angle of 79.4°, suitable for nucleophilic addition. The phosphate of dUMP interacts with Arg178′, Arg179′, Arg218, and Ser219. Five water molecules also interact with the phosphate. Arg23 has moved away significantly from its position in the crystal structure of the wt-TS·dUMP complex.

In the N229V·dCMP complex, WAT648 hydrogen bonds to the hydrogens of the 4-NH_2 group of dCMP. WAT687 hydrogen bonds to the nitrogen of the 4-NH_2 group of dCMP.

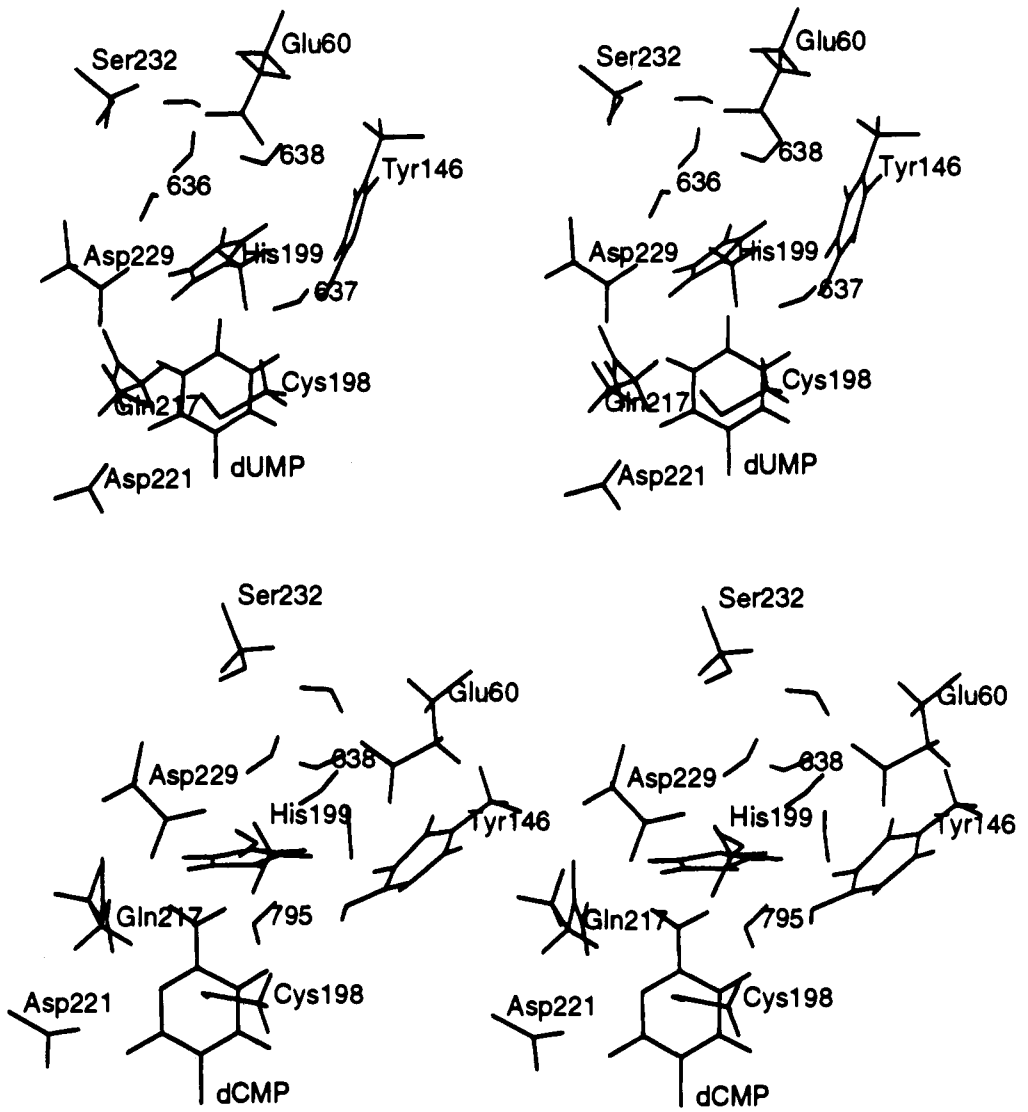

Figure 12. Stereoviews of the equilibrated N229D•dUMP (upper) and N229D•dCMP (lower) complexes obtained with the anionic aspartate 229 side chain and the protonated histidine 199. The main hydrogen bonding distances (Å) follow. For N229D•dUMP: O_2dUMP−HNAsp221, 1.88; HN_3dUMP−$O_{\delta 1}$Asp229, 2.66; O_4dUMP−H_2(637), 1.92; O_4dUMP−HN_ϵHis199, 2.68; O(637)−HOHTyr146, 1.76; OHTyr146−HN_δHis199, 3.00; O(638)−HN_δHis199, 2.05; HN_ϵHis199−$O_{\delta 2}$Asp229, 1.64; O(636)−$O_{\delta 2}$Asp229, 1.78. For N229D•dCMP: O_2dCMP−HNAsp221, 1.78; N_3dCMP−HS_γCys198, 2.47; H_2(4-NH_2)dCMP−$O_{\delta 1}$Asp229, 1.88; H_1(4-NH_2)dCMP−O(795), 2.82; HN_ϵHis199−$O_{\delta 1}$Asp229, 1.77; HN_δHis199−OTyr146, 2.14; O(638)−HN_δHis199, 2.84.

Asp221 hydrogen bonds to O_2 of dCMP in the same manner as in the dUMP binary complex. The 3′-OH of dCMP hydrogen bonds to the hydroxyl hydrogen of Tyr261 and N_ϵ of His259. The S_γ−C_6(dUMP) distance is 4.93 Å, and the S_γ−C_6−C_5 angle is 72.1°. Arg178′, Arg179′, Arg218, Ser219, and five water molecules bind to the phosphate of dCMP. As in the N229V•dUMP binary complex, Arg23 has moved away from the phosphate group.

Energetics from Protein−Ligand Simulations. The free energy results for the forward (dUMP → dCMP) and the reverse (dCMP → dUMP) simulations for the mutation of dUMP into dCMP bound to TS, N229D, and N229V are reported in Table 4. For the TS binary complexes, ΔG_{prot} is −4.61 kcal/mol for the forward simulation and 8.51 kcal/mol for the reverse simulation (250/250 steps eq/dc). The large hystereses for these simulations are probably due to the fact that the X-ray structure of dUMP bound to TS is not accurately regenerated after the reverse simulation. In particular, the cyclic hydrogen bonding network between dUMP and Asn229 was not reestablished. If the ΔG_{prot} obtained from the forward simulation is used to calculate $\Delta\Delta G_{bind}$, a value of −4.75 kcal/mol is obtained. However, if the averaged ΔG_{prot} is used (−6.55 kcal/mol), a value of −2.80 kcal/mol is obtained. Both values of $\Delta\Delta G_{bind}$ are in reasonable agreement with the experimental value of −3.6 kcal/mol, each differing by ∼1 kcal/mol.

Table 4 also reports the results of two test simulations in which the protein was neutralized with counterions and an infinite cutoff was used for the perturbed group (dUMP). These simulations were performed to investigate the dependence of our free energy results and conclusions on the nonbonded cutoff in presence of the charged residues of TS. The 41 ps simulation (250/250 steps eq/dc) under these new conditions gave a ΔG_{for} of −4.90 kcal/mol, which is closely comparable to the −4.61 kcal/mol previously obtained. A longer simulation (82ps, 500/500 steps eq/dc) was also run under these new conditions in order to check the dependence of free energy results on the simulation time. This 82 ps simulation gave a ΔG_{for} of −6.26 kcal/mol, *i.e.* 1.4 kcal/mol more negative than the 41 ps one. These two simulations lead to values of $\Delta\Delta G_{bind}$ of −4.5 kcal/mol (42 ps simulation) and −3.1 kcal/mol (84 ps simulation), comparable to those found with an 8 Å cutoff.

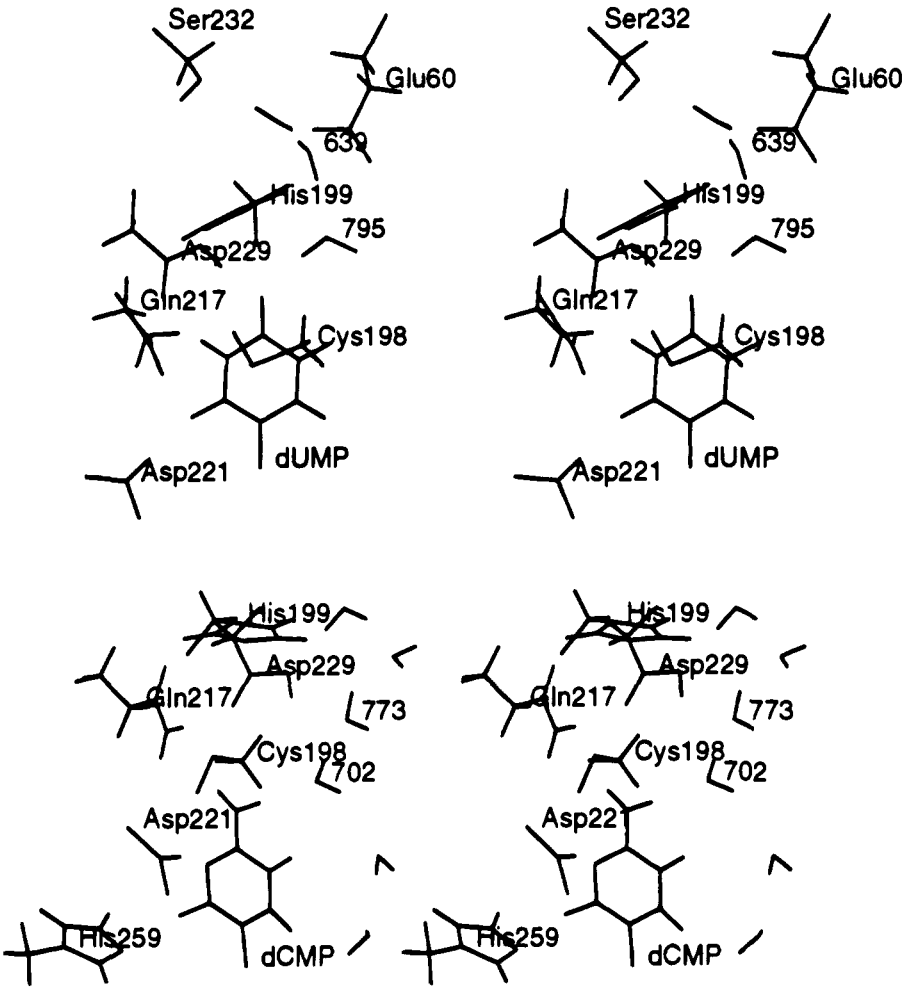

Figure 13. Stereoviews of the equilibrated N229D·dUMP (upper) and N229D·dCMP (lower) complexes obtained with the neutral aspartic acid 229 side chain in configuration 1. The main hydrogen bonding distances (Å) follow. For N229D·dUMP: HN_3dUMP$-O_{\delta 2}$Asp229, 1.92; O_4dUMP$-HO_{\delta 1}$Asp229, 1.88; O_4dUMP$-H_2$(795), 1.92; O_2dUMP$-$HNAsp221, 1.95; O(795)$-H_1$(639), 1.92; O(639)$-$HNδHis199, 2.87; C_6dUMP$-S_\gamma$Cys198, 3.72. For N229D·dCMP: H_1(4-NH_2)dCMP$-$O(702), 2.15; H_2(4-NH_2)dCMP$-S_\gamma$Cys198, 2.91; N_3dCMP$-$HNAsp221, 2.51; C_6dCMP$-S_\gamma$Cys198, 6.30.

For the anionic N229D binary complexes, ΔG_{prot} is similar for the δH tautomer of His199 and protonated His199, but significantly different for the ϵH tautomer. The resulting $\Delta\Delta G_{bind}$ values are -1.07 and -1.13 kcal/mol for the δH tautomer of His199 and protonated His199, respectively. These calculated values of $\Delta\Delta G_{bind}$ are in qualitative agreement with the experimental value (-0.02 kcal/mol, Table 1), again differing by $\sim$1 kcal/mol. The ΔG_{prot} obtained using the ϵH tautomer of His199 is -12.53 kcal/mol, yielding a $\Delta\Delta G_{bind}$ of $+3.17$ kcal/mol, which is large and in the opposite direction compared to experiments.

The free energy results obtained for the N229D complexes with a neutral Asp229 side chain predict a tighter binding of dUMP in configuration 1 ($\Delta\Delta G_{bind} = -3.60$ kcal/mol) and a tighter binding of dCMP in configuration 2 ($\Delta\Delta G_{bind} = 7.62$ kcal/mol). These results show a reversal of binding of dUMP and dCMP depending on which of the oxygens of the Asp229 side chain is protonated (Figure 3). In particular, dCMP is predicted to be preferentially bound by >7 kcal/mol if neutral Asp229 is in configuration 2. However, in both cases, the calculated $\Delta\Delta G_{bind}$ values are not in agreement with the experimental $\Delta\Delta G_{bind}$ of $\sim$0.0 kcal/mol.

For the N229V complexes, a value of ΔG_{prot} of -9.31 kcal/mol is obtained, resulting in a $\Delta\Delta G_{bind}$ of -0.05 kcal/mol. These calculations incorrectly suggest that dUMP is bound slightly tighter than dCMP. The difference between the calculated and experimental values is $\sim$1.5 kcal/mol, somewhat larger than for the other simulations.

Discussion

In the present work, we have assessed the relative binding of dUMP and dCMP to TS and two Asn229 mutants by molecular dynamics and free energy perturbation calculations. The relative binding of dUMP and dCMP with TS ($\Delta\Delta G_{bind}$) can be analyzed as the sum of two components, the free energy difference of solvation of the two ligands in water (ΔG_{sol}) and the free energy difference of binding to the protein (ΔG_{prot}). ΔG_{sol} is -9.4 kcal/mol, indicating that dCMP is better solvated than dUMP in water. Cytosine and uracil each have five potential hydrogen bonding sites and should interact through hydrogen bonds with bulk solvent to approximately the same extent. The negative value of ΔG_{sol} can be attributed in part to the larger dipole moment of dCMP, which causes dCMP to interact more strongly with bulk solvent. Surprisingly, ΔG_{prot} is -6.56 kcal/mol, indicating that TS interacts more strongly with dCMP than dUMP. However, since $\Delta\Delta G_{bind} = \Delta G_{sol} - \Delta G_{prot}$, the effect of the more negative ΔG_{sol} is that dCMP prefers to remain in aqueous solution and dUMP prefers to bind to TS. This suggests that the simple proposal that the TS·dCMP binary complex is inherently less stable than the TS·dUMP binary complex due to repulsive interactions between dCMP and the Asn229 side chain is too cursory. These repulsive

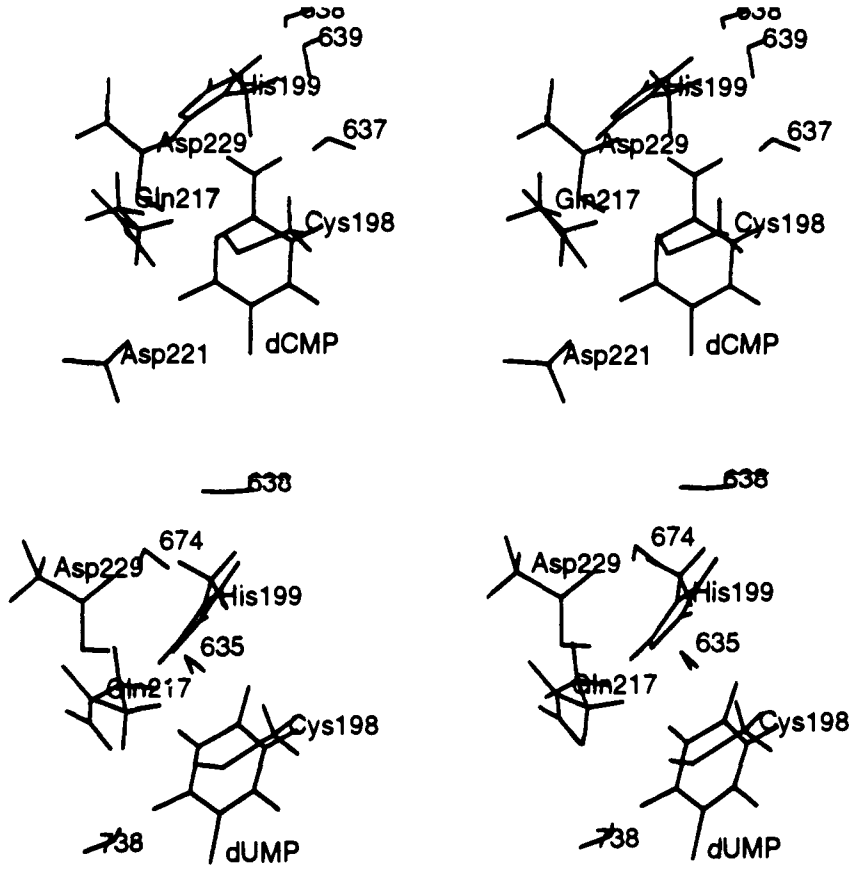

Figure 14. Stereoviews of the equilibrated N229D·dCMP (upper) and N229D·dUMP (lower) complexes obtained with the neutral aspartic acid 229 side chain in configuration 2. The main hydrogen bonding distances (Å) follow. For N229D·dCMP: H_2(4-NH_2)dCMP−$O_{\delta 1}$Asp229, 1.82; H_1(4-NH_2)dCMP−O(637), 1.90; N_3dCMP−$HO_{\delta 2}$Asp229, 1.89; N_3dCMP−HS_γCys198, 2.19; O_2dCMP−HNAsp221, 2.07; O(637)−H_2(639), 2.03; O(639)−H_1(638), 1.90; O(638)−HN_δHis199, 2.01; C_6dCMP−S_γCys198, 3.96. For N229D·dUMP: O_4dUMP−H_1(635), 1.91; O_2dUMP−H_2(738), 1.79; C_6dUMP−S_γCys198, 3.49; $HO_{\delta 2}$Asp229−$N_{\epsilon 2}$His199, 1.96.

interactions are, in fact, present in the simulations of the TS·dCMP complex and result in displacement of the pyrimidine from the normal binding position. However, dCMP still interacts more favorably with the enzyme than dUMP, but the free energy difference (ΔG_{prot}) is not as large as that in water (ΔG_{sol}). The active site of the binary complex is very open to solvent. At any given time, a large number of solvent molecules reside in the active site, and these are accessible to bulk solvent. Only when the cofactor is bound does the enzyme close to sequester the active site from bulk solvent.[4,5,9] Thus, in the binary complex, the substrate can move around in order to reduce repulsive interactions and enhance attractive interactions. Based on the equilibrated structures presented here, it is difficult to rationalize why simulations indicate that TS interacts more strongly with dCMP than dUMP (ΔG_{prot} is negative) since both dUMP and dCMP seem to be well positioned to interact with the enzyme and solvent in the active site. However, simple electrostatics would suggest that the greater dipole moment of dCMP would cause it to be more favorably solvated in any polar media.

The factors determining the specificity of thymidylate synthase are similar to those for staphylococcal nuclease, which is inhibited by dGMP and dTMP. Calculations have shown that dGMP interacts more strongly with the enzyme by ~6 kcal/mol but is better solvated than dTMP by 11.8 kcal/mol.[27] The net result is that dTMP is predicted to bind to staphylococcal nuclease tighter than dGMP, which is in agreement with experimental results. The energy required to desolvate ligands can play an important role in determining the affinity of substrates or inhibitors for an enzyme.

This leaves the question of the role of Asn229. Asn229 does not contribute to the binding of dUMP to TS, nor it is essential to catalysis.[8] Based on experimental data, crystal structures, and these simulations, a major role of Asn229 is in the specificity of dUMP versus dCMP. In the crystal and equilibrated structures of the TS·dUMP binary complex, the S_γ of Cys198 is ~4.0 Å away from C_6 of dUMP and has an attack angle ($C_5-C_6-S_\gamma$) of ~90°.[9] These values are close to those of the MP2/6-31+G* transition structure for the addition of HS^- to acrolein (C−S, 2.639 Å; C−C−S, 116°),[28] implying that very little reorganization of the enzyme is required to reach the transition structure for formation of the covalent enzyme−substrate complex. This is not the case for dCMP, where repulsive interactions between dCMP and Asn229 result in displacement of the pyrimydine from its normal position. Formation of the covalent enzyme−dCMP complex would require significant reorganization of local residues of the enzyme, as indicated by the C_6-S_γ distance of 7.0 Å and the $C_5-C_6-S_\gamma$ angle of 38°, thereby increasing the energy required to reach the transition state structure. The repulsive interactions between dCMP and Asn229 might be converted into attractive interactions by rotating the amide side chain of Asn229 by ~180°. This would result in hydrogen bonding of the N_3 and the 4-NH_2 of dCMP with the rotated Asn229 side chain. A MD simulation was run on this alternative conformation. After heating to 300 K, the simulation was run for over 50 ps. dCMP stayed in a similar position as dUMP in the crystal structure

(27) Caldwell, J. W.; Kollman, P. A. Unpublished results.

(28) Thomas, B. E., IV; Kollman, P. A. Submitted for publication.

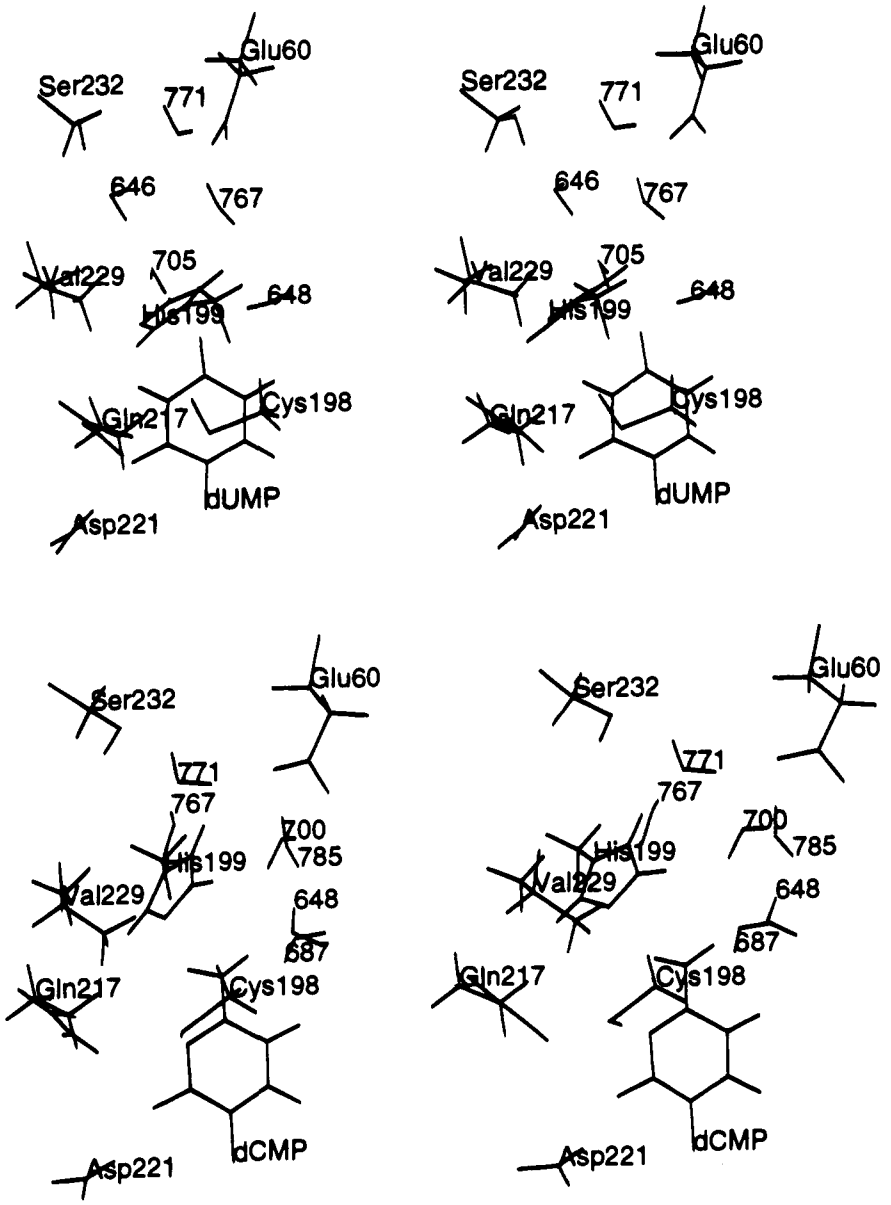

Figure 15. Stereoviews of the equilibrated N229V·dUMP (upper) and N229V·dCMP (lower) complexes. The main hydrogen bonding distances (Å) follow. For N229V·dUMP: O_4dUMP$-H_1$WAT648, 1.99; O_4dUMP$-H_1$(705), 1.81; O_2dUMP$-$HNAsp221, 1.97; C_6dUMP$-S_\gamma$Cys198, 3.89. For N229V·dCMP: H_1(4-NH_2)dCMP$-$O(648), 1.84; N_4dCMP$-H_1$(687), 2.36; N_3dCMP$-HS_\gamma$Cys198, 2.49; C_6dCMP$-S_\gamma$Cys198, 4.93.

and previously described MD simulations. Thus, by rotation of Asn229, it is possible for dCMP to bind in a position suitable for nucleophilic attack. However, there are two main reasons why we think this is not the case: (1) rotation of the Asn229 side chain requires the disruption of the large hydrogen bonded network that exists between various residues and bound waters and (2) rotation of Asn229 would result in a less stable conformation since it led to a repulsive interaction between the NH_2 of Asn229 and the NH_2 of Gln217.

The three different simulations for the anionic N229D mutant provide insight into the binding of dUMP and dCMP. The experimental $\Delta\Delta G_{bind}$ is reproduced by the simulations when the δH tautomer of His199 or a protonated His199 is used. The $\Delta\Delta G_{bind}$ obtained with the ϵH tautomer of His199 is too large and yields a N229D·dCMP structure that is incompatible with formation of a covalent N229D·dCMP complex. The locations of dUMP and dCMP in the active site of N229D are very similar with those of either the δH tautomer of His199 or protonated His199. This raises a problem with rationalization of the change of TS from a dUMP methylase to a dCMP methylase when Asn229 is changed to Asp. It was expected that N229D would help position dCMP for nucleophilic attack while causing dUMP to move to an unproductive part of the active site.

Simulations with the neutral Asp229 side chain of the N229D·dUMP and N229D·dCMP binary complexes provide insight into this problem. The structural and energetic results suggest a strong discriminating power of the neutral Asp229 side chain toward binding of dUMP versus dCMP. The configuration in which the OH of Asp229 is close to O_4 of dUMP and the O of Asp229 is close to HN_3 of dUMP (configuration 1) preferentially binds dUMP and places it properly for methylation. The configuration having O close to O_4 and OH close to HN_3 (configuration 2) preferentially binds dCMP. Although these configurations nicely explain the catalytic preferences, they are in disagreement with the experimental evidence of an equal binding of the two nucleotides to N229D.[8] In the anionic Asp229 simulations (either with δH or with protonated His199), the experimental binding energies are better reproduced and both dUMP and dCMP are in good position for the catalytic reaction to proceed. These findings suggest that no substrate discrimination by N229D occurs at the level of the binary complex. The substrate discrimination

Table 4. Summary of the Calculated Intrinsic Binding Free Energy Differences of dUMP and dCMP Bound to TS and the N229 Mutants[a]

	ΔG_{prot}			
	ΔG_{for} (dUMP → dCMP)	ΔG_{rev} (dCMP → dUMP)	ΔG_{av}	$\Delta\Delta G_{bind}$
	dUMP/dCMP in wt-TS			
250/250 steps eq/dc	−4.61	8.51	−4.61	−4.75
			−6.55 ± 1.35	−2.80 ± 2.00
neutralized protein and infinite cutoff for dUMP				
250/250 steps eq/dc	−4.90			
500/500 steps eq/dc	−6.26			
	dUMP/dCMP in N229D			
anionic Asp229				
250/250 steps eq/dc				
δH tautomer of His199	−7.04	9.53	−8.29 ± 1.25	−1.07 ± 1.30
εH tautomer of His199	−11.39	13.67	−12.53 ± 1.14	3.17 ± 1.19
protonated His199	−8.75	7.71	−8.23 ± 0.52	−1.13 ± 0.57
neutral Asp229				
250/250 steps eq/dc				
configuration 1	−5.76			−3.60
configuration 2		16.98		7.62
	dUMP/dCMP in N229V			
250/250 steps eq/dc	−9.06	9.56	−9.31 ± 0.25	−0.05 ± 0.30

[a] In kcal/mol. $\Delta\Delta G_{bind} = \Delta G_{sol} - \Delta G_{prot}$. $\Delta G_{sol} = -9.36 \pm 0.05$ kcal/mol (from simulation 1).

of dCMP in preference to dUMP could be due to (1) protonation of Asp229 necessary for methylation of dCMP, (2) different electronic assistance of the Asp229 side chain in formation or stabilization of intermediates along the catalytic pathway, or (3) conformational changes upon binding of the folate cofactor which are not considered in the present simulations. Recall that in wild-type TS, binding discrimination between dUMP and dCMP clearly occurs early, i.e., at the level of the binary complex. This is evident from the experimental results and the above simulations of the binary complexes.

It was proposed that a neutral Asp229 side chain is necessary to protonate N_3 of dCMP during methylation.[7,12] However, if this is the case, one of the two residues would require protonation at physiological pH, since both Asp229 and dCMP have a pK_a around 4. Based on the observed synergism between Asp229 and His199, we postulate that His199 can be a source of protons for Asp229; this residue is very close to Asp229, being hydrogen bonded when the Nε is protonated. The unperturbed histidine has a pK_a of 6.1, which implies that ~5% of His199 is protonated at physiological pH, but it is possible that local perturbations could result in a higher pK_a and thus in a larger fraction of protonated His199. Therefore, while His199 is not strictly essential for the catalytic reaction of TS,[6] it could be important for the methylation reaction of dCMP with the N229D mutant.

The calculated $\Delta\Delta G_{bind}$ for the N229V mutant differs from the experimental $\Delta\Delta G_{bind}$ by ~1.5 kcal/mol. In both the N229V·dUMP and the N229V·dCMP complexes, the substrate stays relatively close to the hydrophobic residue. Interestingly, Cys198 is in a better position for nucleophilic attack in the dUMP complex than in the dCMP complex. While this observation agrees well with the finding that N229V catalyzes the methylation of dUMP but not dCMP, it may be fortuitous because of the limited sampling in these MD simulations.

Although the simulations presented here are of rather short duration, they have been useful in providing insight into the relative binding free energies of dUMP and dCMP to TS and mutants, as well as some suggestive ideas on the structural differences in these complexes that may be relevant to catalysis. One should emphasize that "longer is not necessarily better" vis-à-vis MD/FEP simulations, particularly if defects in the force field or representation (e.g., large and unbalanced electrostatics) cause the system to move a large distance from experimentally validated structures. This was illustrated in the necessity of restraints in the simulation of dUMP in solution, to prevent it from moving to an unrealistic syn conformation. Thus, we emphasize that only qualitative and semiquantitative insights are realistic targets of these simulations; the highly accurate predictions of ΔG values currently achievable for small molecule solvation are often difficult to realize for complex enzyme systems.[29]

Conclusions

We have presented molecular dynamics and free energy perturbation calculations on wild-type, N229D, and N229V mutants of TS interacting with dUMP and dCMP. The calculations provide insights into aspects of both binding and catalysis.

First, the calculations have supported the key role of Asn229 in causing tighter binding of dUMP than dCMP to the wild-type TS; the calculated $\Delta\Delta G_{bind}$ is in the range of −3 to −5 kcal/mol, favoring dUMP, and is consistent with the experimental value of −3.6 kcal/mol. The calculations are consistent with experimental results in that $\Delta\Delta G_{bind}$ becomes smaller upon mutation of Asn229 to Asp and to Val. Even though the calculations cannot quantitatively reproduce the $\Delta\Delta G_{bind}$, they reflect the experimental trend nicely.

The predicted structures of the different complexes after molecular dynamics equilibration have implications for the effectiveness of catalysis by the wild-type and mutant enzymes on dUMP and dCMP. For wild-type, the calculations provide an appealing rationale as to why TS methylates dUMP and not dCMP: the sulfur of the catalytic residue of Cys198 is 4 Å from C_6 of dUMP and 7 Å from C_6 in the corresponding dCMP complex.

In the N229D mutant with an *anionic* Asp229 side chain, the free energy calculations suggest that His199 is either cationic or in its δH tautomer. Both dCMP and dUMP are suitably placed with respect to the catalytic sulfur of Cys198 in the anionic N229D mutant; therefore, the calculations do not provide a rationalization of why the N229D mutant prefers to methylate dCMP. In this respect, calculations with the *neutral* Asp229

(29) (a) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050–3054. (b) Sun, Y. X.; Spellmeyer, D.; Pearlman, D.; Kollman, P. A. *J. Am. Chem. Soc.* **1992**, *114*, 6798–6801. (c) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Kollman, P. A. *J. Am. Chem. Soc.* **1993**, *115*, 9620–9631.

side chain highlight a powerful discriminating ability of the neutral Asp toward binding of dUMP and dCMP, which could be the key to rationalizing the selective methylation of dCMP by the N229D mutant. In addition, the observed synergism between anionic Asp229 and His199 lead us to hypothesize that a protonated His199 can serve as proton donor for Asp229 during the catalytic reaction.

The involvement of the aspartic acid side chain in the enzymatic methylation of cytosine residues has also been reported for dCMP–hydroxymethylase[30] and DNA–cytosinemethylase;[31] alignment of dCMP–hydroxymethylase with TS places Asp179 in alignment with Asn229 of TS, suggesting that methylation of cytosine in these enzymes might follow a mechanism similar to that of TS.

Finally, S_γ(Cys198)–C_6 distances obtained from MD equilibration provide a possible rationalization for the observation that the N229V mutant methylates dUMP more efficiently than dCMP. Thus, the calculations have provided insight into both the energetics of substrate binding and some structural aspects of TS important in catalysis.

Acknowledgment. This work was supported by USPHS Grants GM-29072 (P.A.K.) and CA-14394 (D.V.S.). We are glad to acknowledge the use of the UCSF Computer Graphics Lab, supported by RR-1081, T. Ferrin, P.I.

JA942900J

(30) Lamm, N.; Wang, Y.; Mathews, C. K.; Ruger, W. *Eur. J. Biochem.* **1988**, *172*, 553–563.

(31) Klimasauskas, S.; Kumar, R. R. J.; Cheng, X. D. *Cell* **1994**, *76*, 357–369.