Calculation of Solvation and Binding Free Energy Differences for Folate-Based Inhibitors of the Enzyme Thymidylate Synthase

M. Rami Reddy,*,† Russell J. Bacquet, Dominic Zichi, David A. Matthews, Katherine M. Welsh, Terence R. Jones, and Stephan Freer

Contribution from Agouron Pharmaceuticals, Inc., 3565 General Atomics Court, San Diego, California 92121. *Received January 28, 1992*

Abstract: The thermodynamic cycle perturbation approach, in conjunction with molecular dynamics simulations, has been used to calculate relative binding free energies for the closely related inhibitors 10-propargyl-5,8-dideazafolic acid (PDDF) and 10-formyl-5,8-dideazafolic acid (FDDF) to the binary complex consisting of the enzyme *E. coli* thymidylate synthase and 5-fluoro-2′-deoxyuridylate (FdUMP). The calculated difference in binding free energy (2.9 kcal/mol) is in good agreement with the experimentally observed preference for PDDF (3.75 kcal/mol). Energetically, PDDF is more difficult to desolvate than FDDF, but it more than compensates with favorable interactions in the complex. Entropic effects favor FDDF binding but are smaller in magnitude than energetic ones. The PDDF propargyl group makes several good hydrophobic contacts with protein and FdUMP atoms and terminates in a hydrogen bond to a protein carbonyl group. The FDDF formyl group makes a single hydrogen bond to a bound water molecule and has an electrostatically and sterically unfavorable contact with FdUMP. PDDF is predicted to have a more negative solvation free energy than FDDF, a trend opposite that found computationally and experimentally for but-1-yne and acetaldehyde, which are simple molecular analogues of the different functional groups which distinguish these two dideazafolic acid inhibitors. This result can be rationalized in terms of charge distribution shifts, steric blocking of water, and solvent polarization by the charged species.

I. Introduction

The recognition and subsequent binding of ligands to biological receptors is the most important step in many biochemical processes. Theoretical approaches which predict differences in binding affinity for structurally similar ligands interacting with a common receptor would be useful in the rational design of therapeutic agents. The statistical mechanical technique of thermodynamic cycle perturbation (TCP) has been successfully used to calculate relative binding free energies of ligands with a common receptor.[1-16] In this paper, we present results for such a calculation for inhibitors of the enzyme *E. coli* thymidylate synthase (TS). We also report studies of the solvation free energies of the inhibitors and smaller molecules related to them.

The TS enzyme plays a crucial role in DNA biosynthesis by catalyzing the reductive methylation of deoxyuridine monophosphate to produce thymidine monophosphate. In recent years, TS has been the focus of intense chemical, biological, and clinical research aimed at generating novel antitumor drugs with properties different from those of other antifolates such as methotrexate which acts against dihydrofolate reductase. 10-Propargyl-5,8-dideazafolate (PDDF) was discovered to be a tight binding inhibitor of TS, particularly after polyglutamylation, and was clinically evaluated in the early to mid eighties.[17] 10-Formyl-5,8-dideazafolate (FDDF) is a related compound with modest activity.[21,22] We have carefully measured the binding free energy difference against purified enzyme for this pair of inhibitors. Recently, Matthews and co-workers[18-20] determined high-resolution crystal structures of *E. coli* TS ternary complexes containing 5-fluoro-2′-deoxyuridylate (FdUMP) and either PDDF or FDDF. We have refined the crystallographic data for both complexes. With regard to the calculation of a binding free energy difference by TCP protein simulation methods, the crystal structure containing PDDF is a key input for the starting point, while the crystal structure involving FDDF provides a valuable check of the end point. In order to investigate the binding free energy difference, we utilize both crystal structures in conventional molecular dynamics simulations. Additionally, we have calculated the relative solvation energy for PDDF and FDDF. The related smaller molecules but-1-yne and acetaldehyde were also studied because experimental data were available and the trend in solvation energy was markedly different. The structures of these molecules are given in Figure 1.

† Present address: Gensia Pharmaceuticals, Inc., 4575 Eastgate Mall, San Diego, CA 92121.

(1) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420-1426.
(2) *Computer Simulation of Biomolecular Systems*; Van Gunsteren, Weiner, Eds.; ESCOM Science Publishers: The Netherlands, 1989.
(3) Wong, C. F.; McCammon J. A. *J. Am. Chem Soc.* **1986**, *108*, 3830.
(4) Singh, U. C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4280-4284.
(5) Singh, U. C.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 9519-9523.
(6) Fleischman, S. H.; Brooks, C. A., III *Proteins* **1990**, *7*, 52-61.
(7) Merz, K. M., Jr.; Kollman, P. A. *J. Am. Chem. Soc.* **1989**, *111*, 5649-5658.
(8) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987.
(9) Lybrand, T. P.; McCammon, J. A. *J. Comp. Mol. Design* **1988**, *2*, 259-266.
(10) Kuczera, K.; Gao, J.; Tidor, B.; Karplus, M. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, *87*, 8481-8485.
(11) Beveridge, D. L.; Dicolpua, F. M. *Annu. Rev. Biophys. Biophys. Chem.* **1989**, *18*, 431-492.
(12) Gao, J.; Kuczera, K.; Tidor, B.; Karplus, M. *Science* **1989**, *244*, 1069-1072.
(13) Hansen, L. M.; Kollman, P. A. *J. Comp. Chem.* **1990**, *11*, 994-1002.
(14) Reddy, M. R.; Viswanadhan, V. N.; Weinstein, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 10287-10291.
(15) Hausheer, F. H.; Singh, U. C.; Saxe, J. D.; Flory, J. P.; Tufto, K. B. *J. Am. Chem. Soc.* **1992**, *114*, 5356-5362.
(16) Ferguson, D. M.; Radmer, R. J.; Kollman, P. A. *J. Med. Chem.* **1991**, *34*, 2654-2659.
(17) (a) Jones, T. R.; Calvert, A. H.; Jackman, A. L.; Eakin, M. A.; Smithers, M. J.; Betteridge, R. F.; Newell, D. R.; Hayter, A. J.; Stocker, A.; Harland, S. J.; Davis, L. C.; Harrap, K. R. *J. Med. Chem.* **1985**, *28*, 1468-1476. (b) Nair, M. G.; Nanavati, N. T.; Nair, I. G.; Kialiuk, R. L.; Gaumont, Y.; Hsiao, M. C.; Kalman, T. I. *J. Med. Chem.* **1986**, *29*, 1754-1760.
(18) Matthews, D. A.; Appelt, K.; Oatley, S. J.; Xuong, Ng. H. *J. Mol. Biol.* **1990**, *214*, 923-936.
(19) Matthews, D. A.; Villafranca, J. E.; Janson, C. A.; Smith, W. W.; Welsh, K.; Freer, S. *J. Mol. Biol.* **1990**, *214*, 937-948.
(20) Appelt, K.; Bacquet, R. J.; Bartlett, C. A.; Booth, C. L. J.; Freer, S. T.; Fuhry, M. A. M.; Gehring, M. R.; Herrmann, S. M.; Howland, E. F.; Janson, C. A.; Jones, T. R.; Kan, C.; Kathardekar, V.; Lewis, K. K.; Marzoni, G. P.; Matthews, D. A.; Mohr, C.; Moomaw, E. W.; Morse, C. A.; Oatley, S. J.; Ogden, R. O.; Reddy, M. R.; Reich, S. H.; Schoettlin, W. S.; Smith, W. W.; Varney, M. D.; Villafranca, J. E.; Ward, R. W.; Webber, S.; Webber, S. E.; Welsh, K. M.; White, J. *J. Med. Chem.* **1991**, *34*, 1925-1934.
(21) Hynes, J. B.; Eason, D. E.; Garrett, C. M.; Colvin, P. L., Jr.; Shores, K. E.; Freisheim, J. H. *J. Med. Chem.* **1977**, *20*, 588-591.
(22) Rode, W.; Scanlon, K. J.; Hynes, J. B.; Bertino, J. R. *J. Biol. Chem.* **1979**, *254*, 11538-11543.

0002-7863/92/1514-10117$03.00/0

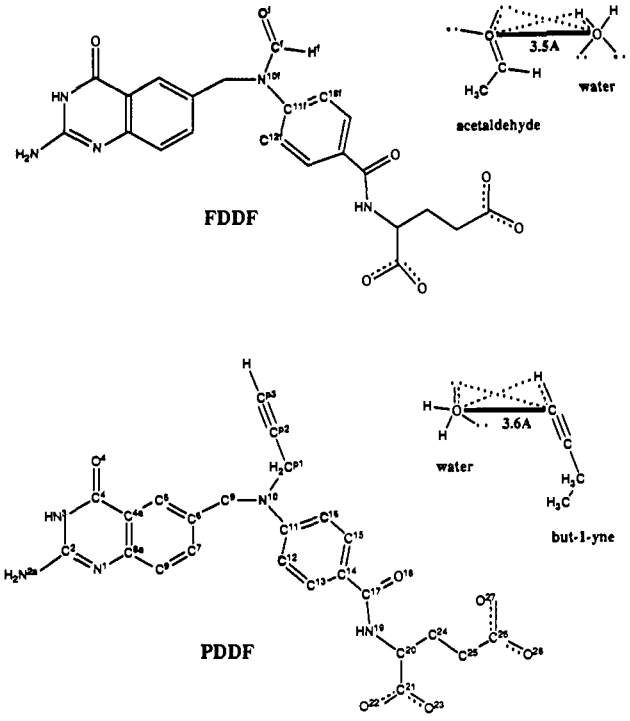

Figure 1. Structures of the TS inhibitors PDDF and FDDF and the simple molecules acetaldehyde and but-1-yne used to test the force field parameters adopted for the formyl and propargyl functional groups. The geometric criteria used to define a hydrogen bond are shown. The maximum donor–acceptor distance (bold line) is 3.5 Å for acetaldehyde and 3.6 Å for but-1-yne. The minimum acceptor–hydrogen–donor and donor–"lone-pair"–acceptor angles (dotted lines) are 100°.

Our analysis of the simulations reported here has the following main results. Energetic interactions underlie the tighter binding of PDDF to TS. The formyl group of FDDF has two possible orientations differing by a 180° rotation; the crystallographically observed orientation is calculated to be 0.9 kcal/mol lower in energy. Excellent agreement with experimental results is obtained for the relative solvation free energy of but-1-yne and acetaldehyde. The relative solvation free energy of PDDF and FDDF has a sign opposite that found for but-1-yne and acetaldehyde, but a semiquantitative explanation is provided by several auxiliary calculations.

II. Simulation Methods

Solvation and binding are complex physical processes that are usually difficult to simulate directly. Nonphysical but computationally more tractable processes can be used to form a thermodynamic cycle which must sum to zero regardless of path since free energy is a state function. Figure 2 illustrates such processes and the solvation and binding cycles for the species of interest in this study. The difference between the free energy changes for two real processes is often the quantity of most relevance, such as $\Delta A_7 - \Delta A_6$ for solvation. Clearly this is equal to the difference between the two nonphysical processes of the zero-sum solvation cycle, $\Delta A_3 - \Delta A_5$, which we will refer to as the relative solvation free energy, $\Delta\Delta A_{solv}$. Expressed in terms of the corresponding experimentally measurable equilibrium constants, $\Delta\Delta A_{solv}$ is given by $-k_BT \ln (k_7/k_6)$, where k_BT is the product of Boltzmann's constant and the temperature. Binding free energy differences are analogous, i.e., we define $\Delta\Delta A_{bind}$ by the computed quantity $\Delta A_4 - \Delta A_3$ and, equivalently, the physically relevant quantity $\Delta A_2 - \Delta A_1$. The determination of ΔA_5, ΔA_3, and ΔA_4 by TCP simulations has been described.[2,14] For all free energy calculations reported here, the entire ligand is defined as the group of atoms to be perturbed and both intramolecular and intermolecular energy terms are included.

We accomplish the required nonphysical transformation of one molecule into a related molecule with the thread procedure.[5] A single topology is defined for those atoms which are identical in both molecules in the sense that force constants and equilibrium geometries are the same (partial charges can vary). For the portion of the molecule which must be transformed, both the starting (reactant) and ending (product) topologies are defined with their correct geometries, one beginning the simulation entirely as dummy atoms and the other ending entirely as dummy atoms. Dummy atoms are identical to real atoms except for their Lennard-Jones parameters and charges which are zero. At intermediate points during the transformation, all atoms of both topologies have fractional Lennard-Jones parameters and charges and both topologies interact with the environment but not with each other. This should be contrasted with the quite different and commonly used procedure of employing a single topology for the entire molecule which has some real and some dummy atoms. In that method, real atoms are turned off (become dummies) and dummy atoms are turned on as necessary while equilibrium internal coordinates are gradually modified from those appropriate to reactant to those appropriate to product.

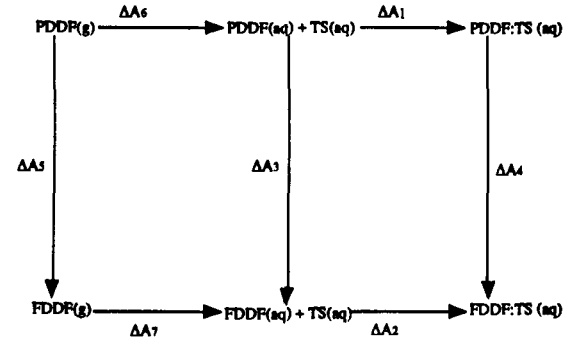

Figure 2. Schematic solvation and binding free energy changes for related substrates PDDF and FDDF and receptor TS. The horizontal free energies correspond to experimental measurements while the nonphysical, vertical ones are calculated.

III. Computational Details

In this section we discuss the model and parameters used in the molecular dynamics (MD) computer simulations for both solvent and protein environments. All MD simulations were performed with a modified form of the AMBER program.[23,24] Standard MD techniques were employed, including use of the Verlet algorithm[25] for integration with a 2 fs time step and use of SHAKE for constraining all bond lengths.[26] Constant temperature (T = 298 K) was maintained by velocity scaling. A residue based cutoff of 9.5 Å was used for nonbonded interactions in general. However, all interactions among solute atoms and atoms of charged residues were included regardless of distance. The aqueous small molecule simulations were performed in a periodic rectangular box dimensioned to allow a 9.5-Å layer of water to surround the solute atoms. For the macromolecule simulations, protein atoms beyond 25.0 Å from the center of the mutating groups were frozen. The protein was immersed in a 26.0-Å sphere of solvent centered at the same point and solvent molecules were subject to a half-harmonic restraint near the boundary to prevent evaporation.

The SPC/E[27] water model, which provides a density, diffusion constant, and dielectric constant in excellent agreement[28] with experiment, was chosen. The protein model adopted is that of the AMBER[24] united atom force field. For small molecules, the AMBER all atom force field[29] was used. Some force constants for the fused rings and the propargyl group were missing from the published force field. They were estimated by linear interpolation of the ab initio bond length to experimental bond lengths and force constants for reference single and double bonds involving those elements.[24] A complete listing of force constants and equilibrium geometries for the small molecules of this study is available upon request from the authors. Point partial charge models for the small molecules were fitted with CHELP[30] to the quantum mechanical electrostatic potential computed from GAUSSIAN88[31] ab initio 3-21G*//

(23) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagoha, G.; Profeta, S., Jr.; Weiner, P. K. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
(24) *AMBER (Version 3.0)*; Singh, U. C.; Weiner, P. K.; Caldwell, J. K.; Kollman, P. A.; University of California: San Francisco, 1986.
(25) Verlet, L. *Phys. Rev.* **1967**, *159*, 98–103.
(26) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327–341.
(27) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269.
(28) Reddy, M. R.; Berkowitz, M. L. *Chem. Phys. Lett.* **1989**, *155*, 173–176.
(29) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comp. Chem.* **1986**, *7*, 230–252.
(30) Chirlian, L. E.; Francl, M. M. *J. Comp. Chem.* **1987**, *8*, 894–905.

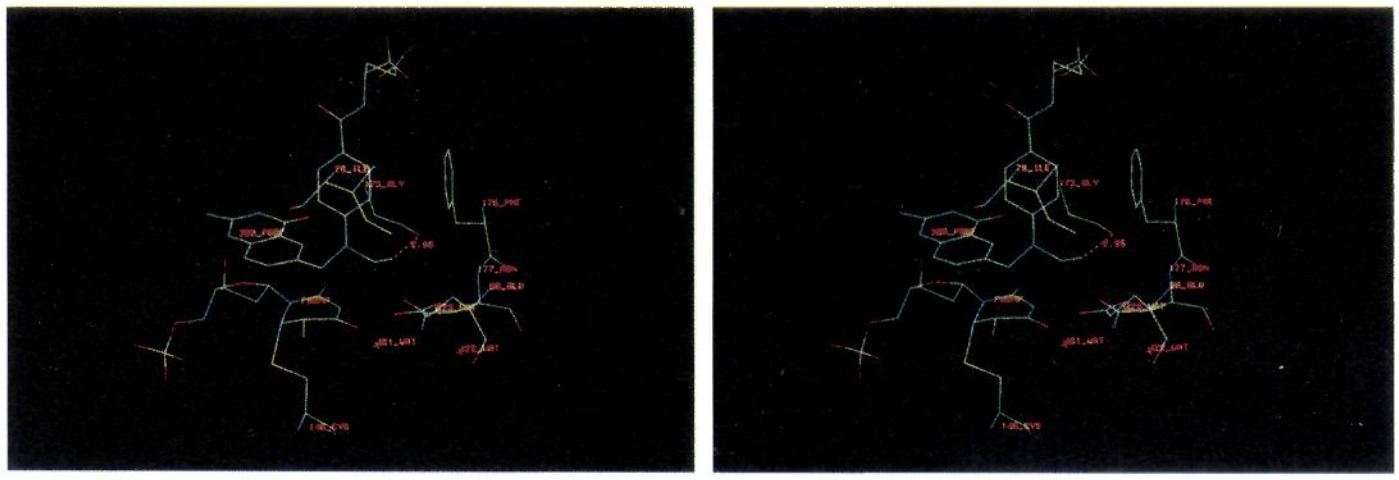

Figure 3. Stereodrawing showing the ternary complex of PDDF, FdUMP, and nearby residues of *E. coli* TS. Protein and ligand atoms are shown in green (carbon), blue (nitrogen), red (oxygen), yellow (sulfur), and dark green (fluorine). Water molecules are represented as red crosses. The distance between the terminal propargyl carbon of PDDF and the carbonyl oxygen of Gly173 is given.

6-31G* wave functions and are given in Table I.

In all cases, the mutation propargyl → formyl was carried out with the thread method[5] described earlier. The propargyl group of PDDF and the formyl group of FDDF are threaded together at the N10 position. In the case of acetaldehyde → but-1-yne, the mutating groups are threaded at the methyl carbon. For all simulations involving FDDF, PDDF, or TS, a two-stage procedure is used. During stage 1 the charges of the reactant atoms are turned off while the Lennard-Jones parameters of the product atoms are turned on. During stage 2 the Lennard-Jones parameters of the reactant atoms are turned off while the charges of the product atoms are turned on.

In all free energy simulations, the system was initially equilibrated for 20 ps followed by 1 ps of equilibrium and 2 ps of data collection for each window. A total of 51 windows were used for each complete mutation of but-1-yne → acetaldehyde, while 102 windows (51 for each of the two stages described above) were used for all simulations of PDDF → FDDF, with or without TS. While these simulations are long, they are expected to achieve no more than a reasonably representative sampling of microstates. Recent work[32,33] has shown that some fairly simple systems require even longer runs for full convergence. Error bars are estimated for each window by dividing the window statistics into four groups and computing the standard deviation. The root mean square of these window errors is reported in Table II as a measure of the statistical uncertainty in the result for each complete mutation.

Conventional molecular dynamics simulations are used to estimate and analyze energetic contributions to free energy changes and to determine average structural and geometric features of the systems.[34] These 200-ps simulations of the solvated and protein-bound states are identical with the free energy simulations except that $H(\lambda)$ is fixed at the value λ = 0 or 1. For the solvated small molecules PDDF, FDDF, but-1-yne, and acetaldehyde, we compute the average total system energy, total solute–water energy, and the propargyl or formyl H-bond energy. The H-bond energy is defined here as the sum of all AMBER force field pairwise energy terms involving atoms of the H-bonded waters interacting with either the C, O, and H of a formyl group or the two triple bonded C's and H of a propargyl group. H-bonded waters are defined by the geometric criteria given in Figure 1. The difference between free energy changes and system energy changes yields $-T\Delta S$. For the protein-bound simulations of PDDF and FDDF, we compute only the average interaction of all inhibitor atoms with all other atoms of the system, that is, with all solvent and protein atoms. We expect this to be the largest contribution to energy changes.

IV. Determination of Inhibition Constants and Crystallographic Refinement

Standard steady state analysis was utilized to determine the inhibition constants of FDDF against *E. coli* TS in the presence of saturating dUMP. The resultant observed inhibition constants were seen to be mixed noncompetitive with K_i slope of 1.6 μM and K_i intercept of 14 μM.[35] These values were calculated by fitting the observed data to a mixed noncompetitive inhibition scheme with EZ-FIT, a nonlinear regression analysis program.[36]

The compound PDDF proved to be a much tighter binding inhibitor. The inhibition kinetics were such that the order of magnitude of the inhibition constant approached the concentration of the enzyme in the reaction mixture and free inhibitor could no longer be equated with total inhibitor.[37,38] Consequently, the assays were modified to accommodate tight binding kinetics. The resulting data were analyzed with equations described by Morrison[39] using a nonlinear regression analysis program described by Jackson et al.[40] The independence of the observed inhibition constant, 2.8 nM, on the concentration of the variable substrate [(6*R*,6*S*)-5,10-methylenetetrahydrofolate] revealed a noncompetitive inhibition pattern as described by Henderson.[41]

It should be noted that all binding affinity measurements were made in the presence of dUMP, while FdUMP was used for

Table I. Fitted ab Initio[a] Atomic Charges for PDDF and FDDF[b] (electron units)

atom	charge	atom	charge	atom	charge
N1	−0.80878	CP1	−0.05590	C21	0.71000
C2	1.04976	H	0.09755	O22	−0.79050
N2a	−1.07155	H	0.09755	O23	−0.79050
H	0.43501	CP2	0.04273	C24	−0.32600
H	0.43501	CP3	−0.46399	H	0.08900
N3	−0.76850	H	0.30618	H	0.08900
H	0.36517	C11	0.26617	C25	−0.21000
C4	0.79222	C12	−0.33011	H	0.04300
O4	−0.57181	H	0.22234	H	0.04300
C4a	−0.35304	C13	−0.05718	C26	0.97500
C5	−0.19901	H	0.11105	O27	−0.87400
H	0.17181	C14	−0.16780	O28	−0.87400
C6	0.18910	C15	−0.05718	N10f	−0.20597
C7	−0.19221	H	0.11105	Cf	0.57500
H	0.15907	C16	−0.33011	Hf	−0.01000
C8	−0.31857	H	0.22234	Of	−0.51000
H	0.18157	C17	0.59756	C11f	0.29176
C8a	0.55307	O18	−0.57243	C12f	−0.27527
C9	−0.21120	N19	−0.83300	Hf	0.22234
H	0.08144	H	0.35900	C16f	−0.27527
H	0.08144	C20	0.41200	Hf	0.22234
N10	−0.03982	H	−0.02200		

[a] 3-21G*//6-31G* with GAUSSIAN88 and CHELP. [b] PDDF = 10-propargyl-5,8-dideazafolic acid; FDDF = 10-formyl-5,8-dideazafolic acid; attached hydrogens follow each heavy atom; see Figure 1 for atom names.

(31) *GAUSSIAN 88*; Frisch, M. J.; Head-Gordon, M.; Schlegel, H. B.; Raghavachari, K.; Binkley, J. S.; Gonzalez, C.; Defrees, D. J.; Fox, D. J.; Whiteside, R. A.; Seeger, R.; Melius, C. F.; Baker, J.; Martin, R.; Kahn, L. R.; Stewart, J. J. P.; Fluder, E. M.; Topiol, S.; Pople, J. A.; Gaussian, Inc.: Pittsburgh, PA, 1988.

(32) Mitchell, M. J.; McCammon, J. A. *J. Comp. Chem.* **1991**, *12*, 271–275.

(33) Pearlman, D. A.; Kollman, P. A. *J. Chem. Phys.* **1991**, *94*, 4532–4545.

(34) Reddy, M. R.; Bacquet, R. J.; Varney, M. *J. Chim. Phys. Phys.-Chim. Biol.* **1991**, *88*, 2605–2615.

(35) Cleland, W. W. *Biochim. Biophys. Acta* **1963**, *67*, 173–187.

(36) *EZ-FIT* (*Version 2*); Perella, F. W.; Perella Scientific, Inc.: Springfield, PA, 1989.

(37) Webb, J. L. *Enzyme and Metabolic Inhibitors*; Academic Press: New York, 1963; p 72.

(38) Straus, O. H.; Goldstein, A. *J. Gen. Physiol.* **1943**, *26*, 559–585.

(39) Morrison, J. F. *Biochim. Biophys. Acta* **1969**, *185*, 269–286.

(40) Jackson, R. C.; Fry, D. W.; Boritzki, T. J.; Besserer, J. A.; Leopold, W. R.; Sloan, B. J.; Elslager, E. F. *Adv. Enzyme Regul.* **1984**, *22*, 187–206.

(41) Henderson, P. J. F. *Biochem. J.* **1972**, *127*, 321–333.

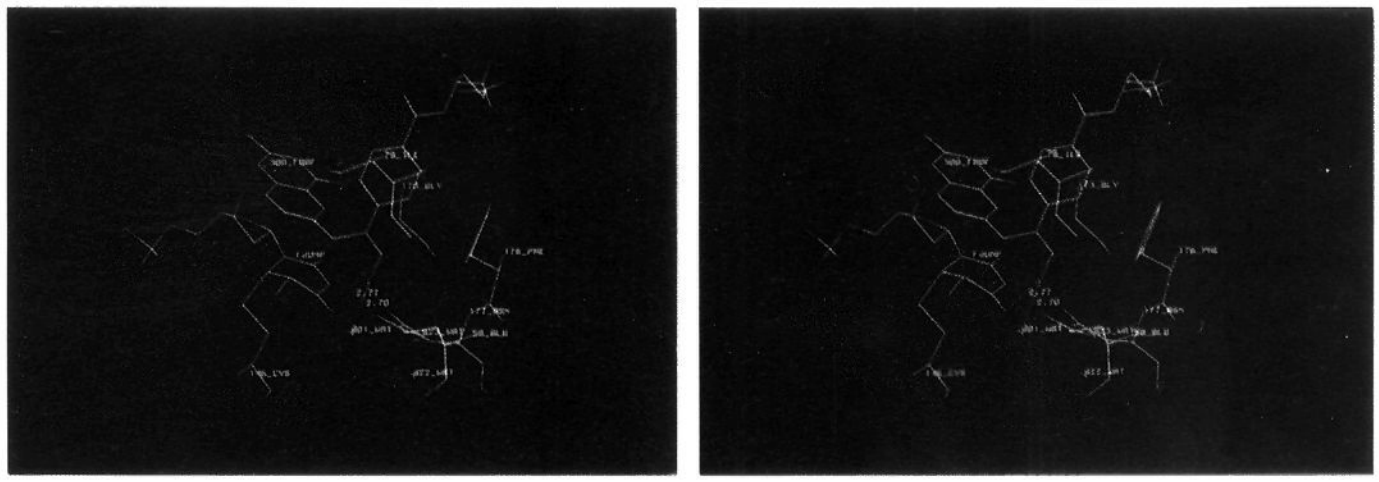

Figure 4. Stereodrawing showing the ternary complex of FDDF, FdUMP, and nearby residues of *E. coli* TS. Protein and ligand atoms are shown in green (carbon), blue (nitrogen), red (oxygen), yellow (sulfur), and dark green (fluorine). Water molecules are represented as red crosses. The distances between the formyl oxygen of FDDF and the oxygens of water 401 and FdUMP are given.

Table II. Relative Energy, Entropy, and Free Energy Differences (kcal/mol)

changes of state[a]	$\Delta\Delta E$ calc	$-T\Delta\Delta S$ calc	$\Delta\Delta A$ calc	$\Delta\Delta G$ expt
BUT(aq) + ACE(g) → BUT(g) + ACE(aq)	−5.5 ± 2.2	2.1 ± 2.3	−3.4 ± 0.6	−3.3[c]
PDDF(aq) + FDDF(g) → PDDF(g) + FDDF(aq)	9.9 ± 3.1	−6.3 ± 3.2	3.6 ± 0.7	
TS:PDDF(aq) + FDDF(aq) → PDDF(aq) + TS:FDDF(aq)	8.6 ± 3.0[b]	−5.7 ± 3.1[b]	2.9 ± 0.8	3.75 ± 0.1[d]

[a] "aq" and "g" indicate aqueous and gas phase environments, respectively; BUT = but-1-yne; ACE = acetaldehyde; TS = thymidylate synthase; PDDF = 10-propargyl-5,8-dideazafolic acid; FDDF = 10-formyl-5,8-dideazafolic acid. [b] For this change of state only, the relative energy difference reported is the interaction energy of all solute atoms with all other atoms of the system for both the aqueous and protein simulations; the value given for $-T\Delta\Delta S$ is obtained from $\Delta\Delta A - \Delta\Delta E$ and includes any water–water, water–protein, and protein–protein contributions omitted from the reported value of $\Delta\Delta E$. [c] From ref 43. [d] This work, see Section IV.

Table III. Number of Hydrogen Bonds, Hydrogen Bond Energies, and Interaction Energy of Solute with Solvent (kcal/mol)

	between waters and the formyl or propargyl group		
system[a]	av no. of H bonds	av energy of H bonds	interaction energy between all solute and all solvent atoms
ACE	1.98 ± 0.05	−4.20 ± 0.11	−14.98 ± 0.45
BUT	0.98 ± 0.04	−1.33 ± 0.05	−12.18 ± 0.36
FDDF	1.49 ± 0.04	−3.30 ± 0.10	−435.3 ± 1.20
PDDF	1.01 ± 0.03	−1.43 ± 0.06	−441.8 ± 1.30

[a] BUT = but-1-yne; ACE = acetaldehyde; PDDF = 10-propargyl-5,8-dideazafolic acid; FDDF = 10-formyl-5,8-dideazafolic acid.

Table IV. Contributions to Relative Energy Differences (kcal/mol)

	$\Delta\Delta E_{\text{solute–water}}$			
system[a]	formyl–water and propargyl–water H-bond energy differences	remainder of solute–water interaction differences	$\Delta\Delta E_{\text{water–water}}$	$\Delta\Delta E_{\text{total}}$
ACE–BUT	−2.87 ± 0.12	0.07	−2.7	−5.5
FDDF–PDDF	−1.87 ± 0.11	8.37	3.4	9.9

[a] BUT = but-1-yne; ACE = acetaldehyde; PDDF = 10-propargyl-5,8-dideazafolic acid; FDDF = 10-formyl-5,8-dideazafolic acid.

crystallographic work and subsequent molecular dynamics simulations. A negligible effect on inhibitor binding is expected for this substitution.

The crystallographic TS:FdUMP:PDDF structure obtained by Matthews et al.[18] was refined by restrained least squares[42] with programs PROTIN and PROFFT.[43] Disordered water was modeled by the procedure for Bolin et al.,[44] and all measured reflections were used in the refinement. The crystallographic R factor after least squares cycle 72 was 19.2% for all measured reflections to 2.3-Å resolution. This trial structure was used as the starting point of the TCP simulation.

Subsequent to the simulations described in this paper the PDDF ternary complex was further refined to an R factor of 16.2% and the FDDF ternary complex was refined to an R factor of 15.6%.[45] No significant structural changes occurred for TS:FdUMP:PDDF during the additional refinement; the root-mean-square shift in a binding region of 20.0-Å diameter was 0.2 Å. The refined TS:FdUMP:FDDF coordinates are very similar to those of the PDDF complex. The root-mean-square difference between equivalent atoms for these structures is 0.3 Å. The binding site regions of these complexes are shown in Figures 3 and 4.

V. Solvation Free Energy Results

The mutation PDDF → FDDF was carried out in water and in the gas phase in order to obtain the relative solvation free energy. All relative free energies reported in this paper are in the direction propargyl → formyl, thus a positive number indicates that propargyl is preferred. The calculated $\Delta\Delta A_{\text{solv}}$ is 3.6 kcal/mol but no experimental value is known for comparison. However, the related smaller compounds but-1-yne and acetaldehyde have a measured relative solvation free energy of −3.3 kcal/mol,[46] differing even in sign. In order to test the force field parameters describing the formyl and propargyl groups, a calculation of the relative solvation free energy of but-1-yne and acetaldehyde was performed. The result of −3.4 kcal/mol agreed closely with ex-

(42) Hendrickson, W. A.; Konnert, J. H. *Computing in Crystallography*; Diamond, Ramaseshan, Venkatesan, Eds.; Indian Academy of Sciences: Bangalore, 1980; pp 13.01–13.23.
(43) Finzel, B. C. *J. Appl. Crystallogr.* **1987**, *20*, 53–55.
(44) Bolin, J. T.; Filman, D. J.; Matthews, D. A.; Hamlin, R. C.; Kraut, J. *J. Biol. Chem.* **1982**, *257*, 13650–13662.
(45) Freer, S.; Matthews, D. A., unpublished results.
(46) Hine, J.; Mookerjee, P. K. *J. Org. Chem.* **1975**, *40*, 292–298.

periment, confirming the suitability of the parameters and supporting our result for the inhibitors PDDF and FDDF.

Because it could be argued that 3-(dimethylamino)propyne and dimethyl formamide are better simple models of PDDF and FDDF, we also computed their relative solvation free energy. The result of −2.8 kcal/mol was similar to that of but-1-yne → acetaldehyde. Experimental solvation data for 3-(dimethylamino)propyne was unavailable and no further work was done for these compounds.

Decompositions into $\Delta\Delta E$ and $-T\Delta\Delta S$ were accomplished with the aid of conventional molecular dynamics trajectories for each of the four solutes and are listed in Table II. For these simulations some further structural and energetic analysis of the trajectories was done as described in Section III. Those results are given in Tables III and IV. Experimental data for $\Delta\Delta E$ were unavailable for these precise systems. However, some validation of our results is provided by measurements for acetone and but-1-yne, which have a $\Delta\Delta H_{solvation}$ of −6.0 kcal/mol and a $\Delta\Delta G_{solvation}$ of −3.7 kcal/mol.[47] Energy changes are found to dominate the free energy changes for both PDDF → FDDF and BUT → ACE.

For BUT → ACE, the majority of the energy change can be traced to hydrogen bond differences between water and the propargyl or formyl group. The acetaldehyde oxygen accepts on the average two good hydrogen bonds from water, which together contribute −4.2 kcal/mol, whereas the carbon of but-1-yne donates one weak hydrogen bond of −1.33 kcal/mol, consistent with other studies of acetylenic hydrogens.[48] A dispersion interaction between non-hydrogen bonding groups would be considerably weaker, perhaps −0.5 kcal/mol. The contribution to $\Delta\Delta E$ arising from all other interactions between the solute and waters is less than 0.1 kcal/mol.

For PDDF → FDDF, the situation is quite different. Here the energy difference is primarily due to the 8.37 kcal/mol arising from solute:water interactions other than H bonding between waters and the propargyl or formyl groups. These "other interactions" consist of dispersion and H bonding by the non-propargyl and non-formyl atoms of the solute as well as longer-ranged solvent polarization by the solute's overall charge distribution and magnitude. The −2 net charge of these inhibitors strongly polarizes water. The rearrangement of highly polarized water induced by the mutation PDDF → FDDF may contribute substantially to the large energy difference even though the net charge remains constant. In order to explore this idea we mutated neutralized PDDF (both carboxylates protonated) into neutralized FDDF in both the aqueous and gas phase environments. The relative solvation free energy was −1.15 kcal/mol, differing from the PDDF → FDDF result by −4.75 kcal/mol. This calculation supports the energetic importance of rearrangements of highly polarized water but should be interpreted with caution. For example, the conformations of the neutralized solutes may be different from those found for the −2 charge states.

It is also instructive to compare directly the results for formyl-containing molecules and to do the same for propargyl-containing molecules. In acetaldehyde, the formyl group is fully exposed to solvent and accepts an average of two good hydrogen bonds with water, whereas in FDDF the formyl group accepts an average of 1.5 hydrogen bonds from water. The steric bulk of FDDF interferes with full solvation of the formyl group, resulting in a loss in hydrogen bond energy compared to acetaldehyde. For the propargyl group of PDDF, no hindrance in hydrogen bonding occurs since the acidic hydrogen is pointing directly into solvent. The propargyl groups of PDDF and but-1-yne both donate one hydrogen bond to water.

Another factor contributing to differences in solvation free energy trends for PDDF → FDDF and but-1-yne → acetaldehyde can be seen in the partial charges listed in Table I. The charges

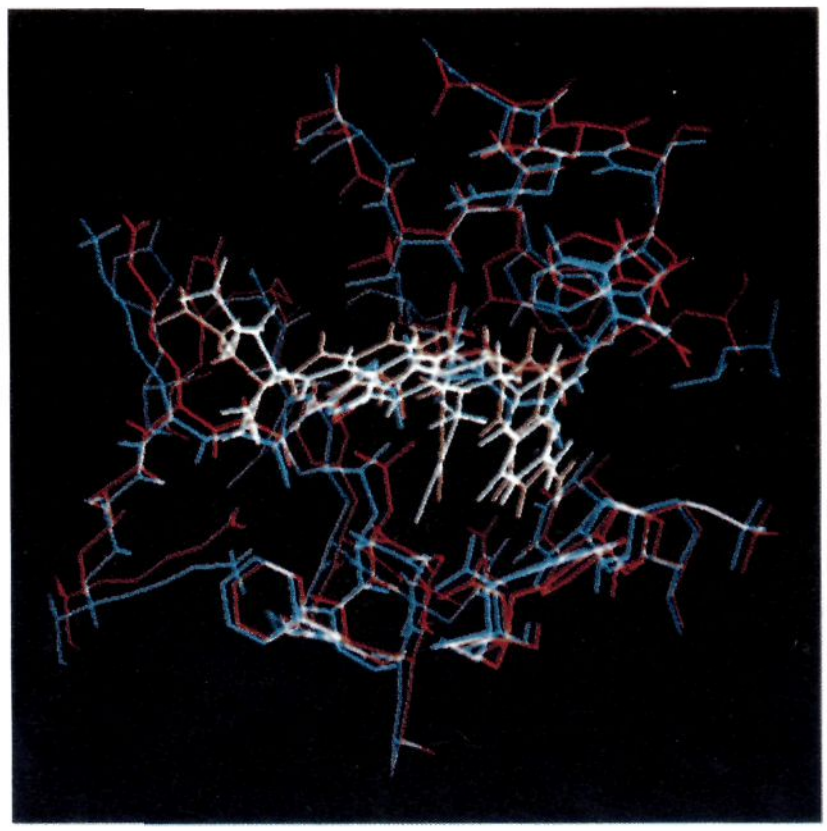

Figure 5. 20-ps molecular dynamics averaged structure of protein TS (red) and ligand PDDF (yellow). X-ray structure of protein TS (blue) and ligand PDDF (white).

on the nitrogen and three ring carbons closest to the mutation change significantly during PDDF → FDDF and have an impact on solute–water interactions. This effect is entirely absent during the mutation but-1-yne → acetaldehyde. In order to quantify this contribution, the conversion of PDDF to FDDF was carried out in two stages. In the first stage, the propargyl group atoms and their charges are converted into formyl group atoms and charges, but the charges on the nearby nitrogen and three closest ring carbons remain fixed at their original values (except for a small change in the nitrogen charge to maintain overall neutrality). The relative solvation free energy change for this stage, $\Delta\Delta A_1$, is 1.29 kcal/mol. In the second stage, the charges of the nearby nitrogen and closest three ring carbons are changed to the values which correspond to our FDDF model. The solvation free energy change in this stage, $\Delta\Delta A_2$, is 2.16 kcal/mol, a sizable contribution not present in the but-1-yne to acetaldehyde case. The total solvation free energy change, $\Delta\Delta A_1 + \Delta\Delta A_2$, is 3.45 kcal/mol, in good agreement with the 3.6 kcal/mol obtained earlier for the direct conversion of PDDF to FDDF.

VI. TS Binding Free Energy Results

The high quality, 2.3 Å resolution refined X-ray structure for the TS:FdUMP:PDDF complex and bound waters was used as a starting configuration for the TCP simulation (see Figure 3). Histidine protonation at one or both ring nitrogens was deduced from H bonding and other features of the environment. A bias toward protonation of both histidine nitrogens was applied in order to reduce the net charge on the dimeric protein to −4. FDDF, PDDF, and FdUMP each have a −2 charge. Since two inhibitor molecules and two FdUMP molecules are present in the TS dimer, the net charge on the system is −12. No counterions or changes in the customary charge of protein residues were used. While such an electrostatic model is far from ideal, alternatives sometimes adopted have their own drawbacks. The development and testing of better methods is an active and crucial area of research.

To check our procedure and model for the complex, an average structure was computed from a 20-ps trajectory collected after a 20-ps equilibration of the minimized TS:FdUMP:PDDF system. A comparison with the experimental X-ray structure gave root-mean-square deviations of 1.1 Å for backbone atoms and 1.6 Å for side chain atoms, indicating close agreement. The binding site region of these structures is displayed in Figure 5. Similar agreement was obtained for the TS:FdUMP:FDDF complex.

The ΔA's computed from TCP simulations of the inhibitor in water and of the TS protein ternary complex are combined to give

(47) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981**, *10*, 563–595.

(48) (a) Berkovitch-Yellin, Z.; Leiserowitz, L. *Acta Crystallogr.* **1984**, *B40*, 159. (b) Copley, M. J.; Holley, C. E., Jr. *J. Am. Chem. Soc.* **1939**, *61*, 1599. (c) DeLaat, A. M.; Ault, B. S. *J. Am. Chem. Soc.* **1987**, *109*, 4232. (d) West, R.; Kraihanzel, C. S. *J. Am. Chem. Soc.* **1961**, *83*, 765.

the relative free energy change for the binding process, $\Delta\Delta A_{bind}$. As before, all relative free energies are in the direction propargyl → formyl, so that a positive number indicates a preference for propargyl. Our net calculated value for $\Delta\Delta A_{bind}$ is 2.9 kcal/mol, reproducing very well the experimental result of 3.8 kcal/mol. These and other relative free energies are reported in Table II along with an approximate decomposition.

The binding energy change is obtained by averaging over conventional dynamics trajectories. For the ternary complex, we report in Table II only the relative interaction of the inhibitors with the rest of the system, a quantity collected during the simulation. That is, differences in water–water, water–protein, and protein–protein interactions are neglected. By computing ΔE in water the same way, we obtain a relative binding energy, $\Delta\Delta E_{bind}$, of 8.6 kcal/mol. As long as the neglected parts of $\Delta\Delta E_{bind}$ are small in magnitude or positive, it follows that the relative binding free energy change is controlled by relative energy changes.

A more detailed breakdown of the time-averaged energy would be needed to identify with confidence the energetic reasons for tighter binding of the propargyl group. However, the following comments, some stated previously,[18] can be made based on the structures obtained by X-ray crystallography (see Figures 3 and 4). Clearly, a weak hydrogen bond exists between the terminal hydrogen of the propargyl group[48,49] and the carbonyl oxygen of Gly173, the donor–acceptor distance being 2.95 Å and the carbonyl carbon–oxygen–propargyl carbon angle being 117°. In addition, the propargyl atoms fit well in a TS:FdUMP pocket which may yield more dispersion interactions than would exist in water. This would arise from the more dense protein environment which offers a greater number of dispersion interactions compared to water, where strong hydrogen bonding leads to a more open structure. Such an effect may be partly offset however by the conformational entropy loss implied by burial in the protein. Good complementarity between the hydrophobic part of the propargyl group (atoms CP1 and CP2, see Figure 1 and Table I) and hydrophobic atoms of the TS:FdUMP binary complex (phenyl ring of Phe176 and pyrimidine ring of FdUMP) contributes to tight binding. While the propargyl finds itself in a favorable environment, the opposite holds for the bound formyl group. The formyl oxygen has a repulsive 2.8-Å interaction with a carbonyl oxygen of the substrate analogue FdUMP and has no direct interaction with protein atoms. The formyl oxygen accepts a hydrogen bond from a bound water molecule but would average 1.5 hydrogen bonds if free in bulk water (see Table III).

Can MD free energy computations be used to reliably predict binding conformations of a ligand? Examination of the FDDF:FdUMP:TS crystal structure suggests that the formyl group could be accommodated in either of two orientations related by a 180° rotation about the N10f–Cf bond. Although electron density maps for the complex detect only one orientation, we carried out a TCP simulation of the other FDDF conformation which yielded a binding free energy relative to PDDF of 3.8 kcal/mol. Along with our earlier result this places the unobserved conformer 0.9 kcal/mol higher in calculated free energy than the observed one. Since the corresponding population ratio is 5:1, visibility of the minor component would not be expected in our 2.3 Å resolution electron density map. Modeling based on the refined X-ray structure suggests that the higher energy of the unobserved orientation is due in part to loss of the single hydrogen bond of the formyl oxygen. Additionally, although the repulsive interaction with FdUMP is relieved, the formyl group now has bad steric contacts with an intramolecular ortho hydrogen and with Phe176.

(49) Such an interaction was postulated without protein structural data in: Jones, T. R.; Betteridge, R. F.; Neidle, S.; Jackman, A. L.; Calvert, A. H. *Anti-Cancer Drug Design* **1989**, *3*, 243–248.

VII. Conclusions

We have calculated the free energy differences of both TS binding and solvation for the inhibitors PDDF and FDDF by means of the TCP molecular dynamics method. Solvation free energies were also determined for several related smaller molecules. Insight was provided by analysis of energetic contributions collected during conventional molecular dynamics.

The solvation free energy calculation for PDDF → FDDF showed that the propargyl-containing compound is favored in water by 3.6 kcal/mol, despite the greater polarity of the formyl group itself relative to the propargyl group. An opposite trend of about the same magnitude (−3.4 kcal/mol) is computed and measured experimentally for but-1-yne and acetaldehyde, resulting in a difference in the relative free energy changes of −7.0 kcal/mol. The actual conversion of PDDF into but-1-yne and FDDF into acetaldehyde was not carried out, but three other calculations provide semiquantitative support for the following explanations. There is a shift in charge distribution for the nitrogen and three ring carbons during the propargyl → formyl conversion which contributes about −2 kcal/mol. Neutralizing the −2 net charge of both PDDF and FDDF may change $\Delta\Delta A$ by roughly −5 kcal/mol. Finally, removing steric bulk allows an additional H bond for the formyl group which has an energy of approximately −1 kcal/mol. These three calculations are not precisely equivalent to the conversion of interest but they are similar. That their sum is −8 kcal/mol supports our rationalization of the computed difference in the relative free energy changes for PDDF → FDDF and but-1-yne → acetaldehyde.

PDDF is predicted to have the stronger binding affinity for TS in full accord with experiment. With the reasonable assumption that contributions to $\Delta\Delta E$ other than inhibitor–environment ones are small, we find that energetic contributions underlie the tighter binding of PDDF to TS. The propargyl of PDDF interacts with a backbone carbonyl and makes good hydrophobic contacts with protein atoms. The formyl of FDDF does not interact with protein atoms, accepts a single solvent H bond, and has a repulsive interaction with a carbonyl oxygen of FdUMP.

Acknowledgment. This work was supported in part by Grant GM-39599 from the National Institutes of Health. We thank Peter Johnson, Dr. David Henry, Dr. Robert Jackson, Dr. J. E. Villafranca, and Dr. Mike Varney for their support and encouragement of this work. We thank Dottie Olson for expert assistance in preparing the manuscript.

Registry No. PDDF, 76849-19-9; FDDF, 61038-31-1.