in energy, to enhance the charge transfer from Ni d_{z^2} to CO_2 π^*, and neutralizes the positive charge of Ni^I, to reduce the charge–dipole repulsion between the bending CO_2 and Ni^I. The reduction of Ni^{II} to Ni^I also destabilizes the Ni d orbitals in energy, which favors the charge transfer from Ni d_{z^2} to CO_2 π^*. These factors enable CO_2 to coordinate to $Ni^IF(NH_3)_4$ in the η^1-C coordination mode. In this stable $Ni^IF(NH_3)_4(\eta^1\text{-}CO_2)$, the HOMO consists of the Ni d_σ and deformed π^* orbital of CO_2 in which the O p_π orbital is enhanced in size and the C p_π orbital is diminished. This feature of the HOMO comes from the orbital mixing of Ni d_{z^2} and CO_2 π and π^* orbitals, and results in the considerably large electron density on O of CO_2. Consequently, the reactivity for proton attack is enhanced from viewpoints of frontier and charge controls. Proton attack yields $[Ni^IF(NH_3)_4(\eta^1\text{-}CO_2H)]^+$ which has the structure of a typical M–CO_2H complex. This species easily undergoes the second one-electron reduction, while the unprotonated species $Ni^IF(NH_3)_4(\eta^1\text{-}CO_2)$ has difficulty undergoing the second reduction. The second reduction weakens the C–OH bond very much. As a result, the OH^- anion easily dissociates from $NiF(NH_3)_4(\eta^1\text{-}CO_2H)$, yielding $[Ni^{II}F(CO)(NH_3)_4]^+$. This Ni^{II}–carbonyl complex takes a triplet $^3A_2(d_{z^2}{}^1d_{xy}{}^1)$ state. The CO binding energy of this complex is very small, 10 kcal/mol at the SD-CI level, which indicates that CO easily dissociates from Ni^{II}. All these results offer strong support to the reaction mechanism proposed by Sauvage et al.

Also, these results suggest how to improve the catalytic system of CO_2 electrochemical reduction by Ni^{II}(cyclam): (1) coexistence of such a coordinating anionic ligand as Cl^- or use of an anode adsorbing metal complex, to stabilize the η^1-CO_2 coordination, (2) use of a protic solvent, to make protonation facile, and (3) use of an appropriate buffer solution, to neutralize OH^- produced with CO (high concentration of OH^- suppresses protonation).

A Theoretical Approach to Drug Design. 3. Relative Thermodynamics of Inhibitor Binding by *E. coli* Dihydrofolate Reductase to Ethyl Derivatives of Trimethoprim Substituted at the 3′, 4′, and 5′ Positions

Joseph J. McDonald† and Charles L. Brooks III*

Contribution from the Department of Chemistry, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213. *Received April 17, 1991*

Abstract: The relative binding thermodynamics of trimethoprim [2,4-diamino-5-(3′,4′,5′-methoxybenzyl)pyrimidine] congeners to *E. coli* dihydrofolate reductase have been determined using free energy simulation methods. The thermodynamics associated with methoxy to ethyl substitutions at the 3′, 4′, and 5′ positions on the benzyl ring of trimethoprim have been calculated. The simulations have been carried out for both the binary DHFR/inhibitor and ternary DHFR/NADPH/inhibitor complexes to examine the effects of the cofactor on inhibitor binding. A model structure was computed for the *E. coli* DHFR ternary complex based upon crystallographic structures of the *E. coli* DHFR/trimethoprim and *L. casei* DHFR/NADPH/methotrexate complexes. The conformation of residues 12–25 in the modeled ternary complex, known to undergo a conformational transition upon cofactor binding, reproduces the conformation seen in a recently solved structure of the *E. coli* DHFR/NADP(+)/folate complex. In six of the seven tested congeners, the transformation from trimethoprim to ethylated derivative is preferred in the DHFR/inhibitor system over the DHFR/inhibitor/NADPH system. Further, the presence of cofactor and the conformational differences in residues 12–25 of the ternary complex have a significant effect on the magnitude of energetic and entropic components associated with the relative binding thermodynamics. The protein environment differences between the binary and ternary complexes affect the overall relative binding free energies in a complex manner which appears to be related to both the degree of inhibitor ethylation and the solvent exposure of the transformed functional groups.

Introduction

Dihydrofolate reductase catalyzes the NADPH-linked reduction of 7,8-dihydrofolate to 5,6,7,8-tetrahydrofolate. Trimethoprim, a substituted benzylpyrimidine, is a competitive inhibitor of dihydrofolate reductase with pharmacological applications to treatment of bacterial infections.[1,2] A strong positive cooperativity in the binding of trimethoprim has been noted for the *E. coli* form of dihydrofolate reductase, dependent on the presence of NADPH cofactor, which varies according to the degree of benzyl methoxy substitution.[3] This cooperativity is believed, in part, to confer the approximately 3000 times greater specificity for trimethoprim in *E. coli* dihydrofolate reductase over a mammalian form of the enzyme. In addition, a variety of close structural analogues of trimethoprim have been studied which differ at the methoxy substituents on the benzyl ring, and show antibacterial affinities which vary with the different substitutions.[4-7] The results of these antibacterial studies have been used to formulate structural and chemical requirements for effective dihydrofolate reductase ligands based on QSAR methodologies.[8,9]

The relative binding affinities of the different trimethoprim congeners are dependent on changes in the relative free energies of inhibitor binding. Differences in binding free energies of related

* Author to whom correspondence should be addressed.
† Current address: Monsanto Company, Department of Biological Chemistry, 700 Chesterfield Village Parkway, Chesterfield, MO 63198.

(1) Hitchings, G. H.; Roth, B. In *Enzyme Inhibitors as Drugs*; Sandler, M., Ed.; Macmillan: New York, 1980, pp 263–280.
(2) Hitchings, G. H. *Agnew. Chem., Int. Ed. Engl.* **1989**, *28*, 879.
(3) Baccanari, D. P.; Daluge, S.; King, R. W. *Biochemistry* **1975**, *21*, 5068.
(4) Roth, B.; Aig, E.; Lane, K.; Rauchkman, B. S. *J. Med. Chem.* **1980**, *23*, 535.
(5) Roth, B.; Aig, E. *J. Med. Chem.* **1987**, *30*, 1998.
(6) Roth, B.; Rauckman, B. S.; Ferone, R.; Baccanari, D. P.; Champness, J. N.; Hyde, R. M. *J. Med. Chem.* **1987**, *30*, 348.
(7) Li, R.-L.; Poe, M. *J. Med. Chem.* **1988**, *31*, 366.
(8) Dietrich, S. W.; Blaney, J. M.; Reynolds, M. A.; Jow, P. Y. C.; Hansch, C. *J. Med. Chem.* **1980**, *23*, 1205.
(9) Selassie, C. D.; Fang, Z.-X.; Li, R.-L.; Hansch, C.; Debnath, G.; Klein, T. E.; Langridge, R.; Kaufman, B. T. *J. Med. Chem.* **1989**, *32*, 1895.

0002-7863/92/1514-2062$03.00/0

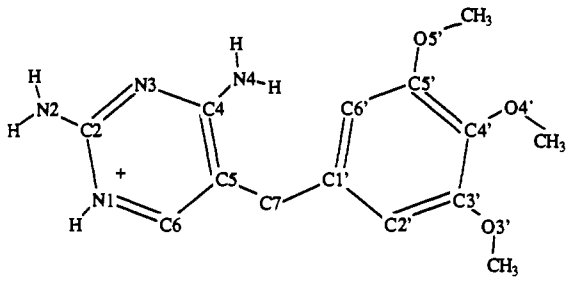


Figure 1. Structural formula for trimethoprim.

ligands to a common receptor arise from changes in the solvation/desolvation thermodynamics of the inhibitor and protein, energetic differences associated with different inhibitor–protein contacts, and conformation entropy changes to the protein and inhibitor.[10] While binding thermodynamics are accessible via experimental techniques,[11,12] theoretical methods such as the thermodynamic cycle perturbation (TCP) approach may be used to provide both thermodynamic and structural insights into the binding of related ligands to a receptor.[13-17] In this article, free energy simulation methods are used to probe the changes in binding thermodynamics of ethylated congeners relative to trimethoprim. Crystallographic information for both the *E. coli* DHFR/trimethoprim and *E. coli* DHFR/trimethoprim/NADPH complexes[18,19] has made possible the calculation of relative binding affinities for the ethylated congeners in both the binary and ternary complexes. The ability to examine thermodynamic data from both complexes yields not only the relative binding differences of ligands to a common receptor, but also the relative thermodynamic differences of a common ligand to different receptors. Therefore, the thermodynamic effects of cofactor on the methoxy to ethyl substitutions in the active site may be considered using these techniques. This is the first report in which a complete series of related inhibitors have been used to examine the thermodynamics of inhibitor binding in different active site environments.

The calculation of the relative free energies of inhibitor binding to a common receptor has two components, one of which is the solvation/desolvation thermodynamics of the drug, while the other is the thermodynamic cost of transforming one inhibitor into another in the enzyme bound state. Previous work by Brooks and Fleischman has used free energy simulation methods to calculate the relative solvation thermodynamics for all combinations of 3′,4′,5′ and mono-, di-, and trisubstituted methoxy to ethyl substitutions on the benzyl ring of trimethoprim (see Figure 1).[20] This study considers the relative thermodynamics of transforming each of the inhibitors at the active site, and assimilates the solvation free energy data in computing the relative binding free energies. Thermodynamic differences among congeners in a common enzyme complex are compared in order to understand the role of position dependence of the transformed ethyl groups on the relative binding thermodynamics. The different balance of dominant energetic and entropic changes are used to examine the thermodynamic effects associated with the presence of cofactor on relative binding affinities.

Methods

The crystallographic structure of the *E. coli* DHFR/trimethoprim complex served as the initial structure for both the binary and ternary complex simulations (personal communication, D. Matthews). This crystallographic structure was appropriate for thermodynamic simulations of the binary complex but required removal of short contacts prior to preparation of the system for molecular dynamics. This was achieved by subjecting the structure (including crystallographic waters) to 120 steps of Powell minimization. All nonhydrogen atoms were constrained harmonically to crystallographic positions. An initial force constant of 24 kcal mol^{-1} A^{-2} was reduced gradually to zero, in 20-cycle increments, over 120 steps.

In the absence of available coordinates for the *E. coli* DHFR/trimethoprim/NADPH ternary complex, a model was constructed based on the binary complex. Residues 13–23 of the enzyme constitute the major conformational difference observed crystallographically between the binary and ternary complexes, and is referred to as the "teen" loop.[19] We chose to use the loop structure from an *L. casei* form of DHFR for our modeled ternary complex, based on a reported small root-mean-square (rms) difference between superimposed loops from the *E. coli* DHFR/trimethoprim/NADPH crystal structure and the *L. casei* DHFR/methotrexate/NADPH crystal structure (1.0 Å over C_α atoms).[19,21] The chicken ternary complex also possesses a strong structural fit to the *E. coli* ternary complex (1.1 Å over C_α atoms). However, the *L. casei* species possesses a 55% sequence homology with the *E. coli* enzyme compared to 33% for the chicken species over the sequence range encompassing this loop. A combination of molecular mechanics and molecular dynamics was used to construct a loop structure consistent with the *L. casei* target structure. In a first step, the backbone and side chain dihedral angles from loop region residues in *L. casei* DHFR (ϕ, ψ, χ_1) were used as target values for constrained minimization of the binary *E. coli* structure, using a dihedral angle force constant of 200 kcal mol^{-1} deg^{-2}. This, however, produced a loop conformation which was qualitatively different from the *L. casei* structure. As a second step, constrained molecular dynamics were employed to refine the loop structure. Inter-C_α distances over residues 10 to 25 were measured for the loop structure of the *L. casei* form of the enzyme, to be used as target constraints in the *E. coli* complex. Only constraints between residues i and greater than or equal to residue $i + 3$ were actually employed. In order to preserve the tertiary structure of the loop, additional inter-C_α atom distance constraints were used between loop residues and residues Gly96 and Gly97. Residues 96 and 97 comprise the N-terminal portion of the αF helix and are homologous to Gly99 and Ala100 in *L. casei* DHFR. Only loop residues 10 to 25 were allowed to move during 12 000 steps of dynamics (18 ps). The resulting loop was qualitatively similar to the *L. casei* loop and showed good visual agreement with the reported structure of the ternary *E. coli* DHFR/trimethoprim/NADPH complex crystal structure. However, during constrained dynamics, peptide bonds between residues Val10/Asp11 and Asp11/Arg12 moved to a cis configuration. Molecular mechanics was employed to change these peptide bonds to a trans configuration. After least-squares superposition of the modeled loop and that from *L. casei*, the RMS difference over C_α positions was 0.2 Å.

After the loop conformation was modeled, the NADPH molecule was placed into the complex and the conformation of the bound trimethoprim molecule was remodeled. The *L. casei* crystal structure was used as a template for the cofactor structure, given the small rms deviation of 0.9 Å between superimposed NADPH structures in the ternary *L. casei* and *E. coli* complexes, noted by Champness and co-workers.[19] These authors also note that protein-cofactor contacts in the *E. coli* enzyme are highly conserved relative to other ternary complexes of DHFR. Residues in the *L. casei* structure, implicated in direct protein–cofactor interctions, were identified, as well as the homologous residues in the *E. coli* model structure. Mainchain atoms in this set of residues were used to least squares fit the ternary *L. casei* DHFR complex (with NADPH) to the *E. coli* modeled structure. Some side chains in the resulting *E. coli* ternary complex model were manipulated to produce cofactor–protein interactions consistent with other observed complexes. In a final step, the NADPH structure was energy minimized in the presence of the protein under the constraints of cofactor–protein hydrogen bond distances obtained from the *L. casei* crystal structure. All protein atoms were constrained to initial positions, with the exception of side chains involved in the ligation of the cofactor, and the cofactor itself.

The conformational differences between the bound trimethoprim molecule in the binary and ternary crystal complexes have been noted by Champness and workers.[19] The θ_1, θ_2 dihedral angles for the inhib-

(10) Beveridge, D. L.; DiCapua, F. M. *Annu. Rev. Biophys. Chem.* **1989**, *18*, 431.
(11) Mares-Guia, M.; Nelson, D. L.; Rogana, E. *J. Am. Chem. Soc.* **1977**, *99*, 2336.
(12) Craik, C. S.; Langman, C.; Fletcher, T.; Rocznick, S.; Barr, P. J.; Fletterick, R.; Rutter, W. J. *Science* **1985**, *228*, 291.
(13) Fleischman, S. H.; Brooks, C. L. *J. Chem. Phys.* **1987**, *87*, 3029.
(14) Fleischman, S. H.; Brooks, C. L. *Proteins* **1990**, *7*, 52.
(15) Wong, C. F.; McCammon, J. A. *J. Am. Chem. Soc.* **1986**, *108*, 3830.
(16) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Science* **1987**, *235*, 574.
(17) McDonald, J. J.; Brooks, C. L. *J. Am. Chem. Soc.* **1991**, *113*, 2295.
(18) Matthews, D. A.; Bolin, J. T.; Burridge, J. M.; Filman, D. J.; Volz, K. W.; Kaufman, B. T.; Beddell, C. R.; Champness, J. N.; Stammers, D. K.; Kraut, J. *J. Biol. Chem.* **1985**, *260*, 381.
(19) Champness, J. N.; Stammers, D. K.; Beddell, C. R. *FEBS Lett.* **1986**, *199*, 61.
(20) Fleischman, S. H.; Brooks, C. L. *J. Am. Chem. Soc.* **1990**, *112*, 3307.
(21) Filman, D. J.; Bolin, J. T.; Matthews, D. A.; Kraut, J. *J. Biol. Chem.* **1982**, *257*, 13663.

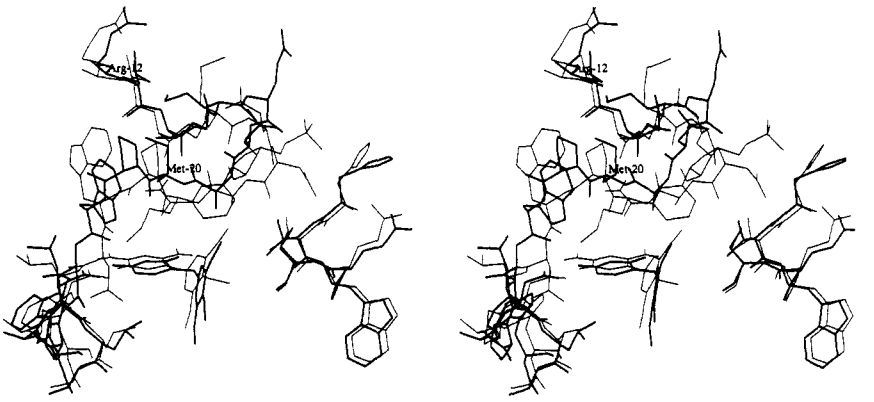


Figure 2. Stereoview of a selected region of atoms taken from a least-squares fit of the *E. coli* binary complex crystal structure and the modeled ternary complex prior to thermodynamic simulations. Thick lines indicate the binary complex and thin lines indicate the ternary complex.

itors, which comprise atoms C6–C5–C7–C1′ and C5–C7–C1′–C6′, respectively, are (−3°, −104°) for the binary complex and (12°, −124°) for the ternary complex. This difference is easily accomodated in the active site cavity by repositioning the benzyl portion of trimethoprim. This excercise is trivial, due to the absence of steric conflict between the protein and trimethoprim benzyl group. Figure 2 illustrates a superposition of the active site regions of the binary crystal structure and the modeled ternary complex.

The ethylated congeners of trimethoprim were transformed from trimethoprim using a hybrid Hamiltonian, of the form, $\mathcal{H}(\lambda) = (1 - \lambda)\mathcal{H}_A + \lambda\mathcal{H}_B + \mathcal{H}_{env}$, in which $\mathcal{H}_A$ are "reactant" methoxy group atoms and $\mathcal{H}_B$ are "product" ethyl group atoms. The reactant state was defined by $\lambda = 0$, and the product state was defined by $\lambda = 1$. $\mathcal{H}_{env}$ represents all solvent atoms (protein, cofactor, and water) and solute atoms chemically equivalent between states A and B. All inhibitor transformations used trimethoprim as the reference state.

All calculations were performed using a highly modified version of the CHARMM program running on a Cray Y-MP machine under the UNICOS operating system. The thermodynamic simulations were carried out using the stochastic boundary molecular dynamics methodology. A 14-Å reaction zone was used, centered at atom C7 of the inhibitor, with a 2-Å buffer region. The stochastic boundary regions were prepared by first solvating and partially equilibrating solvent in a 14-Å reaction zone in the presence of a fixed protein, then equilibrating the entire system, with the atoms beyond 14 Å removed. The details of the procedure may be found in ref 22 and 23. The sizes of the final systems were 798 protein atoms with 87 TIP3P waters for the binary complex, and 819 protein atoms, 64 TIP3P waters plus 55 cofactor atoms for the ternary complex. All thermodynamic simulations were carried out at four λ values to generate an ensemble of configurations for data reduction. The calculations used double-wide sampling at λ = (0.125, 0.5, 0.875, and 0.969) with 7500 steps (11.25 ps) of equilibration and 30 000 steps (45 ps) of production dynamics at each λ value. All calculations were done at 298 ± 5 K using Langevin dynamics for temperature control, by coupling all nonhydrogen atoms to the heat bath. Nonhydrogen buffer region atoms were constrained by harmonic forces derived from crystallographic temperature factors.[23] All heteroatom–hydrogen bond lengths were constrained to equilibrium geometries using the SHAKE algorithm.[24] A nonbonded cutoff of 8.5 Å was employed. Charges for the cofactor and inhibitors have been derived previously and are consistent with the CHARMM19 parameter set, and all protein atoms were represented with the polar hydrogen model.[20,25] Force constants and nonbonded parameters for congeners of trimethoprim were taken from analogous fragments in the CHARMM parameter library.[26] The solvation thermodynamics of the inhibitors, used herein, for the construction of the thermodynamic cycle and the determination of the relative binding free energies have been calculated previously by us and are described fully in ref 20. Time-series analyses of binary and ternary complexes were made based on 7500 steps of stochastic boundary molecular dynamics performed at the end points of thermodynamic simulations. Models of each of the binary and ternary complexes were constructed from an average of 250 coordinate sets collected every 30 steps from the end-point dynamics, and followed by energy minimization.

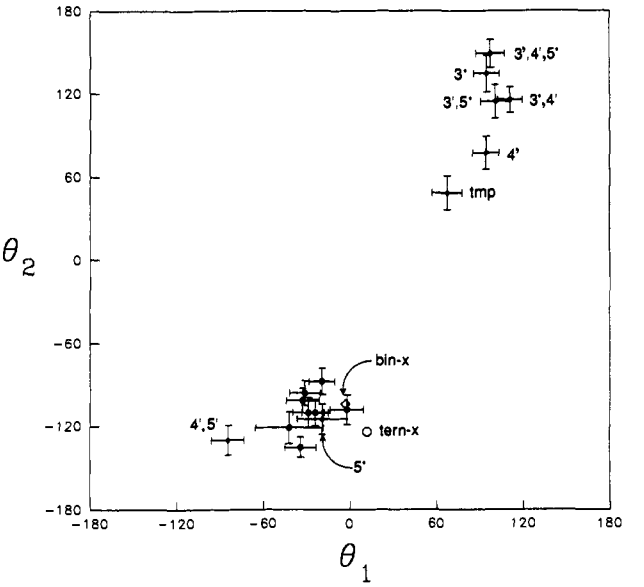


Figure 3. Plot of θ_1 and θ_2 dihedral angles in the inhibitor from molecular dynamics carried out on complexes at the thermodynamic end points. The values represent mean and fluctuations based upon 7500 steps of equilibration and 7500 steps of production dynamics. The θ_1 angle is defined as C6–C5–C7–C1′. The θ_2 dihedral angle is defined as C5–C7–C1′–C6′. Binary complexes denoted by solid diamonds are labeled by derivative type. Ternary complexes are denoted by solid circles. The labeled open symbols denote crystal structure dihedral angles for the binary and ternary complexes.

Results and Discussion

Structural Analysis of Binary Complexes. Two major conformations of the inhibitor have been identified in the binary model complexes, based on mean and fluctuations of trimethoprim θ_1, θ_2 dihedral angles for 7500 steps of molecular dynamics at thermodynamic end points (Figure 3). The figure labels identify the binary complex values. The crystal structure binary *E. coli* complex with trimethoprim demonstrates an inhibitor conformation in which the benzyl ring maintains an approximately perpendicular arrangement to the Phe31 ring, with θ_1 and θ_2 values of −3° and −104°, respectively. ^{1}H NMR measurements describe a similar conformation for the binary bacterial DHFR complex.[27] Taking into account the authors' estimated precision of trimethoprim dihedral angles from NMR, the bounds of their calculated angles are $-6° < \theta_1 < 31°$ and $-129° < \theta_2 < -85°$. The trimethoprim conformation observed in our binary complex

(22) Brooks, C. L.; Karplus, M. *J. Mol. Biol.* **1989**, *208*, 159.
(23) Brooks, C. L.; Brunger, A.; Karplus, M. *Biopolymers* **1985**, *24*, 843.
(24) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(25) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.
(26) Parameters for trimethoprim analogues and NADPH are available from the authors upon request.
(27) Birdsall, B.; Roberts, G. C. K.; Feeney, J.; Dann, J. G.; Burgen, A. S. V. *Biochemistry* **1983**, *22*, 5597.

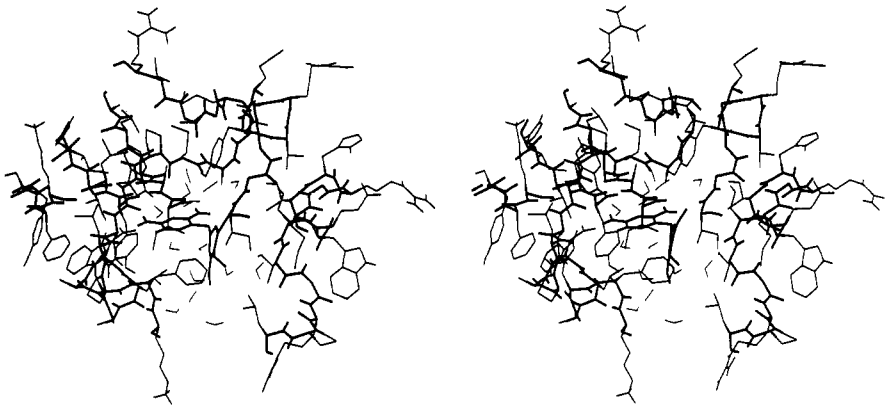


Figure 4. Stereoview of a portion of the active site region used in the thermodynamic simulations of the binary *E. coli* DHFR/trimethoprim complex. This structure was produced by averaging the coordinates over 7500 steps of production dynamics and subsequent energy minimization. Thick lines indicate backbone atoms or trimethoprim. Medium lines indicate side-chain atoms. Thin lines represent positions of water molecules.

Table I. Pyrimidine–Protein Contacts in Model Complexes

	N1		N2		N4		water	
compound	binary	ternary	binary	ternary	binary	ternary	binary	ternary
crystal	Asp-27	N/A	Asp-27	N/A	Ile-5 Ile-94	N/A	N2→Thr-113	N/A
TMP	Asp-27	Asp-27	Thr-113	Asp-27 Thr-113	Ile-5 Ile-94	Ile-5 Ile-94		N2→Tyr-111
3′-Et	Asp-27	Asp-27	Thr-113	Asp-27 Thr-113		Ile-5 Ile-94	N3→Tyr-111 N4→Ile-94	
4′-Et	Asp-27	Asp-27	Thr-113	Asp-27 Thr-113		Ile-5 Ile-94	N4→Ile-94 N3→Tyr-111	
5′-Et		Asp-27	Thr-113	Asp-27 Thr-113	Ile-94	Ile-5 Ile-94	N1→Asp-27 N2→Tyr-111	N2→Tyr-111
3′,4′-Et_2	Asp-27	Asp-27	Thr-113	Asp-27 Thr-113	Ile-5	Ile-5 Ile-94	N4→Ile-94 N2→Tyr-111	
4′,5′-Et_2	Asp-27	Asp-27	Asp-27 Thr-113	Asp-27 Thr-113		Ile-5 Ile-94	N4→Ile-94 N4→Tyr-111	
3′,5′-Et_2	Asp-27	Asp-27	Asp-27 Thr-113	Asp-27 Thr-113		Ile-5 Ile-94	N4→Ile-94 N2→Tyr-111	
3′,4′,5′-Et_3	Asp-27	Asp-27	Thr-113	Asp-27 Thr-113		Ile-5 Ile-94	N4→Ile-94 N2→Tyr-111	N2→Tyr-111

with enzyme exhibits an approximately similar conformation with θ_1 and θ_2 values of 67 ± 11° and −131 ± 12°, respectively. This class of inhibitor conformations is also observed with the 5′-ethyl and 4′,5′-diethyl derivatives. The other observed conformation differs with a benzyl group that is approximately stacked with the Phe31 ring. This second grouping is closer to the inhibitor conformation found in the crystal structure of the chicken DHFR/trimethoprim/NADPH complex (θ_1 = 95°, θ_2 = 102°).[28] It is important to note that both arrangements are sterically accessible in the binary complex active site, and transitions between the conformations are observed during the course of the thermodynamic simulations. To our knowledge, there is no experimental evidence to refute or support the notion of multiple benzyl ring positioning in ethylated trimethoprim derivatives binding to *E. coli* DHFR.

Water molecules play a major role in interactions with the drug in the active site. However, nonspecific water–inhibitor contacts, classified as interactions without apparent hydrogen bonding to inhibitor, are not generally observed in crystal structures because of their mobility. Solvation of the binary complex active site with TIP3P waters for our simulations reveals a highly solvated inhibitor in which nonspecific waters are observed adjacent to both the pyrimidine and benzyl portions of the molecule. The benzyl group of all inhibitors is clearly exposed to solvent, with one to two water shells between the faces of the benzyl group and protein atoms (Figure 4). A number of nonspecific water–inhibitor contacts also occur in the active site space between the "teen" loop protein atoms and the face of the pyrimidine ring. This is common to all the binary models. The opposite face of the pyrimidine group is contacted by the benzyl side chain of residue Phe31.

Specific water–inhibitor contacts, classified in the molecular models as hydrogen-bonded contacts to the inhibitor, were observed in the pyrimidine portion of the inhibitor, mediating interactions with protein functional groups. We have summarized these interactions in Table I. A feature common to six of the seven model complexes is the presence of a water molecule which displaces a direct hydrogen bond between the 4-amino group of the pyrimidine ring of the inhibitor and the carbonyl group of Ile94. Also, a water intervening between the 2-amino group of the inhibitor and the side chain of Thr113 in the crystal structure is positioned slightly different in the models. We observed a water mediating the interaction between the 2-amino or N3 position of the pyrimidine ring and the carbonyl group of Tyr111. These new interactions are sterically possible with minimum rearrangement of the protein. However, the position of the pyrimidine ring in the active site pocket is displaced slightly in order to accommodate the altered interactions. An example of this type of displacement is seen in Figure 5, which compares time-averaged, energy-minimized structures of the triethyl derivatives from binary and ternary complexes. This displacement is absent in the ternary model.

The interactions between the protein and inhibitor can also be understood in terms of specific and nonspecific interactions. The specific interactions in the crystal structure are summarized in Table I, with hydrogen bonding between Asp27 and the protonated N1 group, Asp27 and the 2-amino group, and between the carbonyl groups of residues Ile5 and Ile94 and the 4-amino group. In all but one model, the Asp27 carboxyl group maintains its hydrogen-bonded arrangement with the N1 group of the inhibitor.

(28) Matthews, D. A.; Bolin, J. T.; Burridge, J. M.; Filman, D. J.; Volz, K. W.; Kraut, J. *J. Biol. Chem.* **1985**, *260*, 392.

Figure 5. Stereoview of a least-squares fit of averaged, minimized structures from endpoint molecular dynamics of the 3′,4′,5′-triethyl derivatives taken from the binary and ternary complexes. Thick lines indicate the ternary complex and thin lines indicate the binary complex. The orientation of the figure is such that the "teen" loop is oriented in front of the inhibitor molecules.

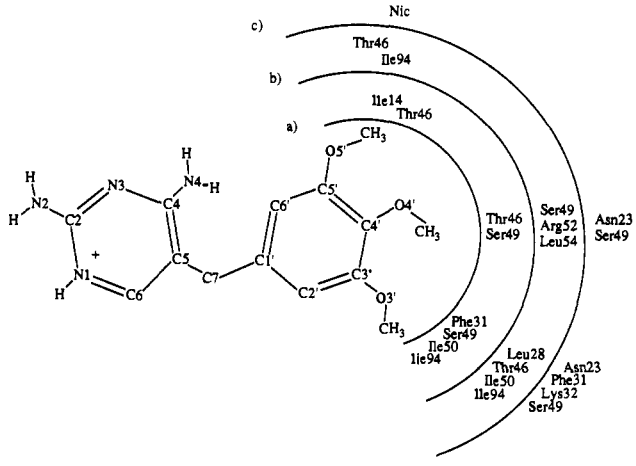


Figure 6. A summary of interactions observed between inhibitor and protein in model structures: (a) inhibitor structures in the binary complexes characterized by a perpendicular juxtaposition of inhibitor benzyl ring and Phe31; (b) model binary complexes characterized by a stacked benzyl ring–Phe31 arrangement; (c) interactions observed in ternary model complexes.

The displacement of a water group from the interaction with the Thr113 side chain, as mentioned above, allows this side chain to hydrogen bond directly to the 2-amino group.

Atom–atom contacts between the benzyl methoxy or ethyl groups and protein atoms are minimal in the model structures. Only two interactions of benzyl functional groups with protein atoms are recorded with an atom–atom distance of less than 3.5 Å across all eight models. On average, each model maintains three interactions between the protein and the benzyl functional groups of the inhibitor in the range of 3.5 to 4.0 Å. For those inhibitor–protein complexes characterized by an approximately perpendicular arrangement with Phe31, the following denotes the totaled interactions at each of the substituted benzyl positions: C3′ (Phe31, Ser49, Ile50, Ile94), C4′ (Thr46, Ser49), C5′ (Ile14, Thr46). The same interaction summary for complexes denoted by an approximately stacked arrangement with Phe31 is as follows: C3′ (Leu28, Thr46, Ile50, Ile94), C4′ (Ser49, Arg52, Leu54), C5′ (Thr46, Ile94). Matthews and co-workers identify trimethoxy benzyl contacts with protein at residues Leu28, Phe31, Ser49, and Ile50, a subset of those observed in our models.[18] A graphic summarizing these contacts is shown in Figure 6.

Structural Analysis of Ternary Complexes. For the ternary model complexes, one class of conformations is found for the inhibitors, based on time-series averages and fluctuations of the θ_1, θ_2 inhibitor dihedral angles. The benzyl ring lies approximately perpendicular to the Phe31 aromatic ring. The functional groups bounding the benzyl group including the Phe31 residue, the nicotinamide group, and the position of the Met20 side chain appear to sterically allow only the one arrangement. A model of a time-averaged, energy-minimized structure for the ternary complex with the trimethoprim inhibitor is shown in Figure 7. This is similar to the arrangement for the reported crystal structure of the *E. coli* DHFR/NADPH/trimethoprim complex.[19]

The modeled protein structures of the ternary complexes are compared with the crystal of the *E. coli* DHFR/NADP(+)/folate complex,[29] the only available *E. coli* ternary complex coordinate set. A single major protein conformational difference at the active site is noted in the ternary complex of DHFR with trimethoprim or folate, compared to the binary complex. This change is observed in the conformation of the "teen" loop comprising residues 13–23, such that the volume of the active site is smaller in the ternary complex (Figure 2). The result of this conformational change is fewer water molecules in the active site compared to the binary complex. The "teen" loop structures of the binary and ternary crystal structures share two common backbone–backbone hydrogen bonds to a neighboring β-strand which appear to stabilize the initial portion of the loop. These are (Arg12(O), Phe125(H)) and (Gly15(O), Asp122(H)). Hydrogen bonds are not observed

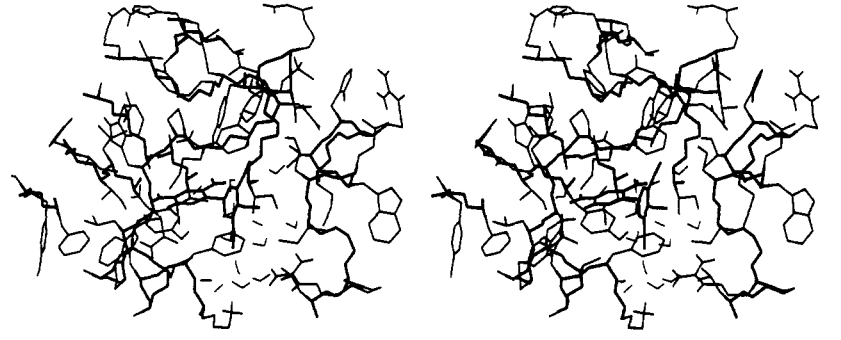


Figure 7. Stereoview of a portion of the active site region used in the thermodynamic simulations of the ternary *E. coli* DHFR/NADPH/trimethoprim complex. This structure was produced by averaging the coordinates over 7500 steps of production dynamics and subsequent energy minimization. Thick lines indicate backbone atoms or trimethoprim. Medium lines indicate side-chain atoms or the nicotinamide nucleotide. Thin lines represent positions of water molecules.

Figure 8. Stereoview of the "teen" loop, cofactor and neighboring residues taken from a least-squares fit of the modeled ternary complex and a recent crystal structure of the *E. coli* DHFR/NADP(+)/folate ternary complex. Thick lines indicate the crystal structure and thin lines indicate the modeled ternary complex.

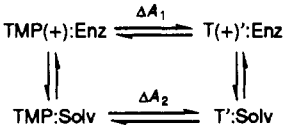
in the binary complex in the range of residues 16 to 22. However, a chain reversal is clearly observed in residues 17 and 18, such that residues 15–17 are loosely antiparallel to residues 18–21. The ternary complex "teen" loop structure is stabilized by two additional hydrogen bonds not found in the binary complex. These are (Met16(H), Ala19(O)) and (Met16(O), Ala19(H)) (see Figure 8). They are also noted in the ternary complex of *L. casei* DHFR with methotrexate and NADPH, from which our initial ternary model was constructed. The (Met16(O), Ala19(H)) hydrogen bond bounds a type I β-turn in the *E. coli* complex and a type II β-turn in the *L. casei* crystal structure. Thus, the ternary "teen" loop structure comprises a short antiparallel β-turn-β structure. This same feature is observed in the time-averaged, energy-minimized structure of the ternary trimethoprim complex, and is also similar to that found for the other congeners.

The mechanics of the chain movement between the binary and ternary crystal structures redirect the position of the side chain at Met20. In the binary complex, the position of Met20 is away from the pyrimidine ring, such that water molecules freely migrate into the space between Met20 and the plane of the pyrimidine ring. The main-chain "teen" loop conformation of the ternary complex reduces the water accessible volume in the active site, particularly around the pyrimidine ring of the inhibitor. In addition, the steric bulk of the Met20 side chain, repositioned in the ternary complex adjacent to the benzyl ring of the inhibitor, further precludes the accessibility of solvent. In all eight of the model ternary complexes considered in this study, the side chain of Met20 points into the void made by the angle formed between the pyrimidine and benzyl portions of the inhibitor. In six of the eight complexes, the side chain of Met20 interacts favorably with the benzyl portion of the inhibitor, at a distance of approximately 4 Å. The exceptions, the 5′-ethyl and 3′,4′,5′-triethyl derivatives, show an average Met20 side-chain conformation in which the C_β atom points in a direction similar to the others noted, but adopts a conformation beyond C_β that points away from the benzyl moiety.

Direct hydrogen bonding between the pyrimidine portion of the inhibitor and protein in the ternary complex environment of the pyrimidine ring is maximized owing to the absence of adjacent waters as well as the restricted environment. Many of the hydrogen bonds noted between the inhibitor and protein in the crystal structure of the binary complex are found among all the model ternary complexes (Table I). This includes hydrogen bonding between the 4-amino group of the inhibitor and the carbonyl moieties of residues Ile5 and Ile94. This interaction is absent in most of the binary model complexes due to a mediating water interaction. Also, the intervening water noted in the binary complex crystal structure between Thr113 and the 2-amino group is absent in these models, and is replaced with a direct hydrogen bond between the protein and inhibitor. However, three of the eight models (trimethoprim, 5′-ethyl and 3′,4′,5′-triethyl complexes) exhibit a trapped water molecule which hydrogen bonds to the 2-amino group of the inhibitor and the carbonyl oxygen of residue Tyr111.

The presence of cofactor and the coincident protein conformational changes yield an active site environment around the benzyl group of the inhibitor which is slightly different from the related binary complex. The smaller active site volume accessible to the benzyl ring results in a smaller set of protein residues which are contacted. The C5′ position of the inhibitor is closest to the nicotinamide ring. Although the methoxy or ethyl group of the inhibitors at the C5′ position does not make direct contact with the nicotinamide ring, six of the eight models show nicotinamide interactions with functional groups at the C5′ or C6′ atoms of slightly less than 4 Å. This suggests significant interactions between the C5′ functional group and the cofactor. The functional group at the C4′ position points out from the active site and is contacted mostly by solvent. However, seven of the eight models show contacts between the C4′ position functional group and Ser49. A contact between Asn23 and the C4′ functional group also appears to be possible, although this interaction was found in only three out of the eight complexes. The C3′ position is adjacent to the aromatic ring of Phe31. Interactions between the functional groups at the C3′ position and Phe31 are noted in six of the eight models. This is the dominant protein interaction for the C3′ functional group. Depending on the θ_1, θ_2 dihedral angles for the different inhibitors, three other residues may also be contacted at the C3′ position (i.e., Asn23, Lys32, or Ser49). The 3′ position shows atom contacts of less than 3.5 Å for the following models (residues in contact are marked parenthetically): 3′,4′,5′-triethyl (Lys32, Ser49), 4′-ethyl (Asn23), 5′-ethyl (Lys32). These interactions are summarized in Figure 6.

Thermodynamics of Inhibitor Binding. Four thermodynamic cycles are relevant to this discussion. The relative binding free energies of the various inhibitors are obtained from the following thermodynamic cycle:



(1)

where TMP and T′ represent trimethoprim and the ethylated congeners, TMP(+) and T(+)′ denote the protonated forms of the inhibitor, which are the relevant species when bound to DHFR as verified by NMR,[30-32] and ΔA_1 and ΔA_2 represent simulations conducted in the protein or solvent environments, respectively. The relative free energy of binding is equivalent to ΔA_1(protein) − ΔA_2(solvent). The second and third thermodynamic cycles of importance reflect the relative contributions of active site envi-

(29) Bystroff, C.; Oatley, S. J.; Kraut, J. *Biochemistry* **1990**, *29*, 3263.
(30) Bevan, A. W.; Roberts, G. C. K.; Feeney, J.; Kuyper, L. *Eur. J. Biophys.* **1985**, *11*, 211.
(31) Roberts, G. C. K.; Feeney, J.; Burgen, A. S. V.; Daluge, S. *FEBS Lett.* **1981**, *131*, 85.
(32) Cocco, L.; Roth, B.; Temple, C.; Montgomery, J. A.; London, R. E.; Blakley, R. L. *Arch. Biochem. Biophys.* **1983**, *226*, 567.

Table II. Thermodynamics of Inhibitor Binding Relative to Trimethoprim in Binary and Ternary Complexes (in kcal/mol)

inhibitor	transformation type	$\Delta\Delta A$ complex		$\Delta\Delta E$ complex		$T\Delta\Delta S$ complex	
		binary	ternary	binary	ternary	binary	ternary
3′-ethyl	(solv)[a]	−1.5 ± 0.2	−1.5 ± 0.2	−12.4 ± 5	−12.4 ± 5	−10.9 ± 5	−10.9 ± 5
	(Enz)[b]	−1.4 ± 0.2	−2.8 ± 0.1	−9.0 ± 5	1.3 ± 6	−7.7 ± 5	4.1 ± 6
	(Enz)−(solv)	0.1 ± 0.3	−1.3 ± 0.2	3.4 ± 7	13.7 ± 8	3.2 ± 7	15.0 ± 8
4′-ethyl	(solv)	−1.5 ± 0.2	−1.5 ± 0.2	−16.5 ± 5	−16.5 ± 5	−15.0 ± 5	−15.0 ± 5
	(Enz)	−2.3 ± 0.1	−1.6 ± 0.1	−1.2 ± 5	2.9 ± 5	1.1 ± 6	4.5 ± 5
	(Enz)−(solv)	−0.8 ± 0.2	−0.1 ± 0.2	15.3 ± 7	19.4 ± 6	16.1 ± 7	19.5 ± 7
5′-ethyl	(solv)	−1.5 ± 0.2	−1.5 ± 0.1	−12.4 ± 5	−12.4 ± 5	−10.9 ± 5	−10.9 ± 5
	(Enz)	−1.7 ± 0.2	−1.5 ± 0.1	9.2 ± 6	−4.1 ± 4	10.9 ± 6	−2.6 ± 4
	(Enz)−(solv)	−0.2 ± 0.3	0.0 ± 0.1	21.6 ± 8	8.3 ± 6	21.8 ± 8	8.3 ± 6
3′,4′-diethyl	(solv)	−2.4 ± 0.2	−2.4 ± 0.2	−18.6 ± 7	−18.6 ± 7	−16.2 ± 7	−16.2 ± 7
	(Enz)	−4.7 ± 0.2	−4.3 ± 0.2	−11.0 ± 7	−36.0 ± 7	−6.3 ± 7	−32.7 ± 7
	(Enz)−(solv)	−2.3 ± 0.3	−1.9 ± 0.3	7.6 ± 10	−17.4 ± 10	9.9 ± 10	−16.5 ± 10
4′,5′-diethyl	(solv)	−2.4 ± 0.2	−2.4 ± 0.2	−18.6 ± 7	−18.6 ± 7	−16.2 ± 7	−16.2 ± 7
	(Enz)	−4.2 ± 0.2	−3.5 ± 0.3	8.8 ± 7	−7.8 ± 8	13.1 ± 7	−4.3 ± 8
	(Enz)−(solv)	−1.7 ± 0.3	−1.1 ± 0.4	27.4 ± 10	10.8 ± 11	29.3 ± 10	11.9 ± 11
3′,5′-diethyl	(solv)	−2.1 ± 0.2	−2.1 ± 0.2	4.1 ± 6	4.1 ± 6	6.2 ± 6	6.2 ± 6
	(Enz)	−4.0 ± 0.2	−3.0 ± 0.2	−6.0 ± 7	−8.9 ± 6	−2.1 ± 6	−6.0 ± 6
	(Enz)−(solv)	−1.9 ± 0.3	−0.9 ± 0.3	−10.1 ± 9	−13.0 ± 8	−8.3 ± 9	−12.2 ± 8
3′,4′,5′-triethyl	(solv)	−4.9 ± 0.3	−4.9 ± 0.3	−2.7 ± 6	−2.7 ± 6	2.2 ± 6	2.2 ± 6
	(Enz)	−6.8 ± 0.4	−5.6 ± 0.2	−17.9 ± 10	3.1 ± 8	−11.0 ± 10	8.7 ± 8
	(Enz)−(solv)	−1.9 ± 0.5	−0.7 ± 0.4	−15.2 ± 12	5.8 ± 10	−13.2 ± 12	6.5 ± 10

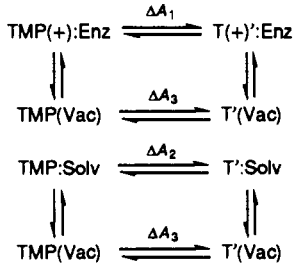
[a] Corresponds to ΔA_2 from eq 1. [b] Corresponds to ΔA_1 from eq 1.

Table III. Active Site Environment and Solvation Contributions to Relative Free Energies of Inhibitor Binding in Binary and Ternary Complexes (in kcal/mol)

inhibitor	transformation type	$\Delta\Delta A$ protein complex binary	$\Delta\Delta A$ protein complex ternary	$\Delta\Delta A$ solvent[a]	$\Delta\Delta E$ protein complex binary	$\Delta\Delta E$ protein complex ternary	$\Delta\Delta E$ solvent	$T\Delta\Delta S$ protein complex binary	$T\Delta\Delta S$ protein complex ternary	$T\Delta\Delta S$ solvent
3′-ethyl	(v)[b]	−3.5 ± 0.2	−3.5 ± 0.2		−2.9 ± 4	−2.9 ± 4		0.6 ± 4	0.6 ± 4	
	(Env)[c]	−1.4 ± 0.2	−2.8 ± 0.1		−9.0 ± 5	1.3 ± 6		−7.7 ± 5	4.1 ± 6	
	(Env)−(v)	2.1 ± 0.3	0.7 ± 0.3	2.0 ± 0.2	−6.1 ± 6	4.2 ± 7	−9.5 ± 6	−8.3 ± 6	3.5 ± 7	−11.5 ± 6
4′-ethyl	(v)	−3.4 ± 0.1	−3.4 ± 0.1		−5.2 ± 2	−5.2 ± 2		−1.8 ± 2	−1.8 ± 2	
	(Env)	−2.3 ± 0.1	−1.6 ± 0.1		−1.2 ± 5	2.9 ± 5		1.1 ± 6	4.5 ± 5	
	(Env)−(v)	1.1 ± 0.1	1.8 ± 0.2	1.9 ± 0.2	4.0 ± 6	8.1 ± 5	−11.3 ± 5	2.9 ± 6	6.3 ± 6	−13.2 ± 5
5′-ethyl	(v)	−3.5 ± 0.2	−3.5 ± 0.2		−2.9 ± 4	−2.9 ± 4		0.6 ± 4	0.6 ± 4	
	(Env)	−1.7 ± 0.2	−1.5 ± 0.1		9.2 ± 6	−4.1 ± 4		10.9 ± 6	−2.6 ± 4	
	(Env)−(v)	1.8 ± 0.3	2.0 ± 0.2	2.0 ± 0.2	12.1 ± 7	−1.2 ± 6	−9.5 ± 6	10.3 ± 7	−3.2 ± 6	−11.5 ± 6
3′,4′-diethyl	(v)	−6.7 ± 0.4	−6.7 ± 0.4		−6.8 ± 9	−6.8 ± 9		−0.1 ± 9	−0.1 ± 9	
	(Env)	−4.7 ± 0.2	−4.3 ± 0.2		−11.0 ± 7	−36 ± 7		−6.2 ± 7	−32.7 ± 7	
	(Env)−(v)	2.0 ± 0.4	2.4 ± 0.5	4.3 ± 0.4	−4.2 ± 12	−29.2 ± 11	−11.8 ± 11	−6.1 ± 12	−32.6 ± 11	−16.1 ± 11
4′,5′-diethyl	(v)	−6.7 ± 0.4	−6.7 ± 0.4		−6.8 ± 9	−6.8 ± 9		−0.1 ± 9	−0.1 ± 9	
	(Env)	−4.2 ± 0.2	−3.5 ± 0.3		8.8 ± 7	−7.8 ± 8		13.0 ± 7	−4.3 ± 8	
	(Env)−(v)	2.5 ± 0.5	3.2 ± 0.5	4.3 ± 0.4	15.6 ± 11	−1.0 ± 12	−11.8 ± 11	13.1 ± 11	−4.2 ± 12	−16.1 ± 11
3′,5′-diethyl	(v)	−3.9 ± 0.2	−3.9 ± 0.2		−4.3 ± 4	−4.3 ± 4		−0.5 ± 4	−0.5 ± 4	
	(Env)	−4.0 ± 0.2	−3.0 ± 0.2		−6.0 ± 7	−8.9 ± 6		−2.1 ± 6	−6.0 ± 6	
	(Env)−(v)	−0.1 ± 0.3	0.9 ± 0.3	1.8 ± 0.3	−1.7 ± 8	−4.6 ± 7	8.4 ± 7	−1.6 ± 8	−5.5 ± 7	6.6 ± 7
3′,4′,5′-triethyl	(v)	−8.8 ± 0.4	−8.8 ± 0.4		−6.8 ± 9	−6.8 ± 9		2.0 ± 9	2.0 ± 9	
	(Env)	−6.8 ± 0.4	−5.6 ± 0.2		−17.9 ± 10	3.1 ± 8		−11.0 ± 10	8.7 ± 8	
	(Env)−(v)	2.0 ± 0.5	3.2 ± 0.5	3.9 ± 0.5	−11.1 ± 13	9.9 ± 12	4.1 ± 11	−13.0 ± 13	6.7 ± 12	0.2 ± 11

[a] Data taken from ref 20. The negative of these values indicate the relative thermodynamics of desolvation. [b] Corresponds to ΔA_3 from eq 2. [c] Env for active site is protein/solvent/cofactor and corresponds to ΔA_1 from eq 1. Env for solvent is bulk water and corresponds to ΔA_2 of eq 1.

ronment or desolvation thermodynamics to inhibitor binding. These are given by eq 2 and 3,



(2)

(3)

where ΔA_3 is the gas-phase process. $\Delta A_1 - \Delta A_3$ (eq 2) represents the thermodynamic contribution of active site environment, and $\Delta A_2 - \Delta A_3$ (eq 3) represents the relative solvation thermodynamics of the inhibitor. It should be noted that we chose to use the unprotonated species of the inhibitor in simulations conducted *in vacuo* or in aqueous solvent (ΔA_3, ΔA_2). Both protonated and unprotonated species of trimethoprim are likely in aqueous solvent at physiological pH, given the pK_a of ~7.2 for protonation at the N1 position. However, we justified our use of the unprotonated species with the observation that ΔA_3 values computed for both species of the 3′,5′-diethyl derivatives fall within the bounds expected for the repeatability of a thermodynamic calculation of a single species from independent initial velocities and starting configurations (±0.4 kcal/mol, unpublished data). The general question of convergence of free energies is an important one, especially in an active site environment. We have followed the free energy profiles and accumulated variances of the free energy in 6 λ windows for the 3′,4′,5′-triethyl derivative in the binary and the ternary complexes. The triethyl derivative was chosen because it represents the perturbation with the greatest number of transformed functional groups. We find that the free energies appear to be well-converged properties, and the accumulated variances are in line with the calculated precision of the free energy numbers reported herein (unpublished data). Tables II and III show the relative free energies of inhibitor binding of both binary

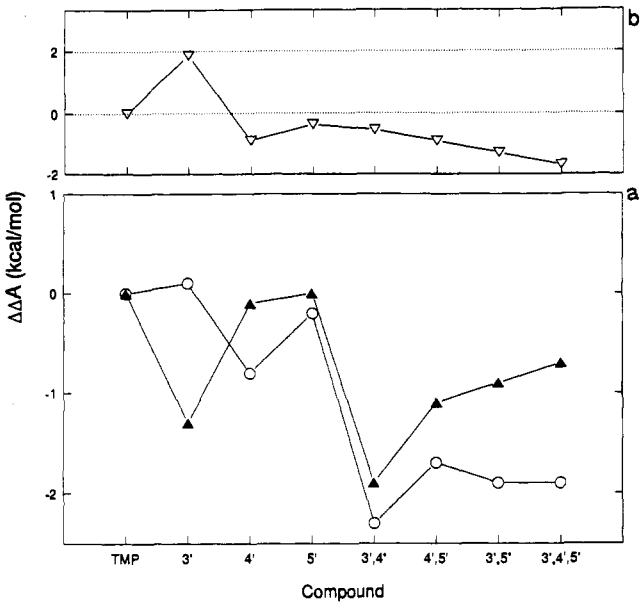


Figure 9. (a) Relative binding free energies of substituted ethyl derivatives of trimethoprim to the *E. coli* form of DHFR. Values indicate computed binding free energies relative to trimethoprim. Closed triangles represent computed values for the DHFR ternary complex. Open circles represent relative free energy values computed for the binary complex. (b) Relative differences in binding free energies between binary and ternary complexes. Negative numbers indicate preferred binding in the binary system.

and ternary complexes and the thermodynamic contributions from active site and solvation, respectively. Although uncertainties in $\Delta\Delta E$ and $\Delta\Delta S$ are large, previous arguments based on statistical criteria have demonstrated the significance of the difference between $\Delta\Delta E$ and $\Delta\Delta S$, thereby allowing the discussion of the relative importance of these components.[20] The free energy of binding an ethylated derivative over trimethoprim in the binary complex versus the ternary complex is the fourth thermodynamic cycle and is given by the equation: ΔA_1(binary) − ΔA_1(ternary), where binary and ternary indicate the different active site environments. The results are summarized graphically in Figure 9.

Thermodynamics of the Binary Complexes. The calculated preference for the binding of ethylated derivatives of trimethoprim to the binary complex is (3′,4′) > (3′,4′,5′) > (3′,5′) > (4′,5′) > 4′ > 5′ > TMPH > 3′. The protein environment contribution to the free energies is positive and unfavorable or near zero for all the congeners, while the desolvation component favors binding for all seven congeners (Table III). Although the protein environment free energy component is generally unfavorable, the calculations demonstrate that the protein environment is a generally more favorable environment for the transformation than solvent. This is consistent with independent thermodynamic simulations demonstrating the favorability of aliphatic groups in an apolar environment over a methoxy or hydroxyl group in an apolar environment.[33] It is this domination by the desolvation term, therefore, that accounts for the overall favorable relative free energy of binding.

The 3′-ethyl derivative appears as the exception to the general rule of favorable relative binding free energies with a positive 0.1 kcal/mol relative binding free energy, and is dominated by an unfavorable protein environment contribution to the overall relative binding free energy (Table III). The basis for the slightly different thermodynamic behavior of the 3′-ethyl derivative may lie in the exposure of the 3′-ethyl group to solvent molecules outside the active site. The time series analysis of computed θ_1, θ_2 inhibitor dihedral angles is consistent with a 3′-ethyl moiety that is highly exposed to solvent outside the active site. The thermodynamic simulations show that the active site relative free energies are similar to the solvation free energies (Table II). In addition, the

(33) Brooks, C. L. *Int. J. Quantum Chem. QBS* **1988**, *15*, 221.

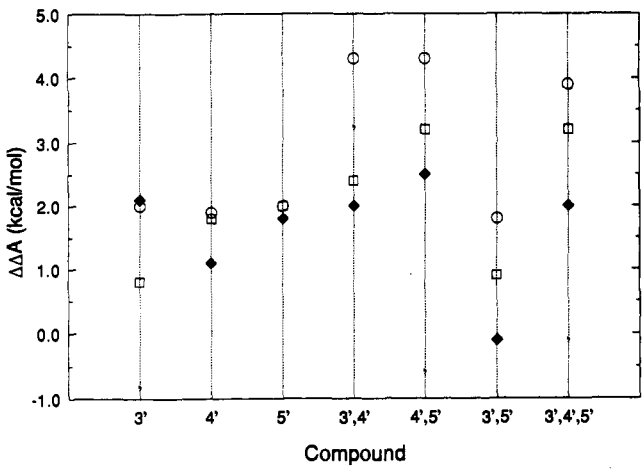


Figure 10. Active site and solvation relative free energy contributions (see also Table II). Open circles represent solvation thermodynamics. Open squares and solid diamonds represent the active site contributions from ternary and binary complexes, respectively.

energetic and entropic components from solvation and active site contributions demonstrate similar, though not identical trends. This implies that the 3′-ethyl group in the 3′-ethyl derivative "sees" an environment much like aqueous solvent. This, however, must be viewed cautiously as this study has not undertaken to assign the individual thermodynamic contributions from protein and solvent in the active site simulations.

A major feature of note in the binary complex free energies is the difference in the relative affinities of the drug depending on the degree of ethylation. The monoethyl substituted derivatives have less favorable binding free energy differences than either the di- or triethyl-substituted derivatives (Table II). The basis for this increase may be considered in terms of the relative contributions of protein environment versus drug solvation thermodynamics. Figure 10 shows the solvation and protein thermodynamic contributions to the relative free energy of inhibitor binding. This figure shows a larger difference between the solvation and protein free energy components for the di- and triethyl derivatives compared to the monoethyl-substituted species. The larger difference implies a more dominant solvation effect, given the larger magnitude solvation piece compared to the protein component. A solvation component that is disfavored relative to the protein part favors the desolvation of the drug upon binding to the enzyme and hence contributes favorably to the free energy of drug binding through the relation $\Delta\Delta A_{bind} = \Delta A_{prot} - \Delta A_{solv}$. Therefore, a more favorable desolvation contribution relative to the protein contribution is the basis for the enhanced binding affinity of the di- and triethyl congeners (i.e., desolvation drives binding for the di- and triethyl derivatives).

Figure 10 illustrates that the thermodynamic components for the 3′,5′-diethyl derivative appear to be atypical of the diethyl-substituted species. While the relative binding free energies of the 3′,5′-diethyl derivatives are similar to the other diethyl derivatives (Figure 9), the solvation and active site components are systematically more negative. The thermodynamic cost of transforming this derivative in water or at the active site is similar in magnitude and sign to those obtained for the other diethyl-substituted species (Table II). The difference lies in the gas-phase thermodynamics of this derivative (Table III), which are used to calculate the solvation and protein component thermodynamics. The gas-phase free energy contribution for the 3′,5′-diethyl derivative is very similar to the monoethyl-substituted species. This same effect has also been noted in the additive effect of transforming trimethoprim to the 5′-ethyl derivative, and the 5′-ethyl derivative to the 3′,5′-diethyl derivative, implying that the overall change for the 3′,5′-diethyl derivative is independent on the thermodynamic path chosen (unpublished data). This more positive gas-phase piece, when subtracted from the aqueous component, produces a solvation free energy for the 3′,5′-diethyl derivative which is more favorable than the other diethyl-sub-

stituted species. Previous theoretical calculations and experimental measurements have corroborated that the central methoxy group demonstrates an enhanced hydrophilicity when functional groups are substituted in a diortho arrangement about this group.[20]

Thermodynamics of Ternary Complexes. The trends observed in relative binding free energies for the ternary complex are, in a very general sense, similar to those calculated for the binary complex (Figure 9). Rather than attempting to describe the thermodynamics associated with the ternary complexes alone, and then in a comparative context with the binary complexes, a stronger sense of the "teen" loop and cofactor thermodynamic contributions are understood by directly comparing the ternary and binary complex thermodynamics. There are a number of important similarities and differences in the relative binding free energy thermodynamic trends for both sets of complexes. The similar trends are identified and discussed followed by a discussion of differences.

A strong preference exists for binding the 3′,4′-diethyl derivative over either the 3′- or 4′-ethyl derivatives in both binary and ternary complexes. The origin of this difference yields insight into the relative contributions of drug desolvation thermodynamics and protein-bound thermodynamics in ligand binding. In particular, the relative binding affinity of an inhibitor with two ethyl groups versus one raises the question of additive thermodynamic effects in solvent versus the enzyme active site. Previous work has shown that the solvation free energy of the 3′,4′-diethyl derivative exhibits an approximately additive relationship with the free energies of the 3′- and 4′-monoethyl derivatives (Table III).[20]. The protein components of the ternary complexes involving the 3′-ethyl, 4′-ethyl, and 3′,4′-diethyl derivatives also show approximately additive free energy behavior, with values of 0.7 kcal/mol (3′-Et), 1.8 kcal/mol (4′-Et), and 2.4 kcal/mol (3′,4′-Et_2) (Table III). Together, the relative binding free energy of the 3′,4′-diethyl derivative is approximately the combined thermodynamic effects of the monoethyl derivatives. The binary complex shows more complex behavior, in which the local protein environment differentially influences the transformed functional groups. The protein relative free energy costs for these derivatives are 2.1 kcal/mol (3′-Et), 1.1. kcal/mol (4′-Et), and 2.0 kcal/mol (3′,4′-Et_2). The protein component of the 3′,4′-diethyl derivative is not the additive effect of the 3′- and 4′-monoethyl derivatives. The basis for the absence or presence of an additive effect may lie in the mobility of the benzyl ring in the trimethoprim molecule. This is illustrated by considering Figure 2, which shows the average θ_1, θ_2 values for the binary and ternary complexes. As stated above, two major conformations for the inhibitor are found in the binary complexes. As such, the functional groups involved in the thermodynamic transformations explore a significant portion of the active site environment. We propose that the various ethyl derivatives prefer slightly different active site environments, resulting in nonadditive thermodynamics. The ternary complexes show a tight grouping of θ_1, θ_2 values compared to the binary complexes, due to the imposed steric constraints of the "teen" loop and cofactor. A reduced number of accessible inhibitor conformations results in a greater likelihood of protein–inhibitor interactions which are common to both the mono- and diethyl-substituted species. Under these conditions, approximate thermodynamic additivity would be expected.

The complex influences of local active site environments on the binding thermodynamics are manifested in the 3′,4′-diethyl and the 4′,5′-diethyl derivatives. In both the binary and ternary active site complexes, the 3′,4′-diethyl derivative exhibits the most favorable binding free energy, followed by less favorable binding free energies for the 4′,5′-diethyl derivative. Because the desolvation thermodynamics of the drugs are identical for symmetry reasons, the basis for the preferred binding of the 3′,4′-diethyl derivative lies in the active site contributions to the binding thermodynamics. The calculated binding affinities of the 3′- or 5′-based ethyl substitutions are not the same due to substantially different local active site environments around those groups and suggests that the local environment around the 3′-functional group favors the ethyl substituent to a greater extent than around the 5′-functional group. However, these local environments are not identical between the binary and ternary complexes. Figure 9 illustrates this point with a less favorable environment for the transformation of these two derivatives in the ternary complex compared to the binary complex. It should be expected that there would be some thermodynamic difference between the binary and ternary complexes for the 4′,5′-diethyl derivative. The 5′-functional group is in close proximity to the cofactor and "teen" loop in the ternary complex, while in the binary complex this group is solvated. It is less clear that the 3′ and 4′ functional groups in the 3′,4′-diethyl derivative would be significantly affected by the presence of the cofactor, given that these functional groups do not directly contact either the "teen" loop or cofactor in the ternary complex. However, our end-point molecular dynamics indicate a difference in the mobility and placement of the benzyl ring in the presence of cofactor, such that it indirectly influences the thermodynamics of binding this derivative. These indirect influences, which result in the slight displacement of bound ligand, are difficult to study using conventional statistical-based structure–activity methods which rely on a relatively static local environment around the bound ligand.

The 5′-ethyl derivative in the binary and ternary complexes posseses very similar relative binding free energies. The solvation thermodynamics are identical for the binary and ternary complexes, so that the question of the origin of similar protein contributions to the relative binding free energy is of interest. As stated in the section detailing the structural features of the complexes, the 5′-ethyl group of the inhibitor occupies two basic positions in the active site among all inhibitor complexes, which differ in the placement of the benzyl group. The first position allows a stacked arrangement between benzyl group and Phe31. The second is perpendicular to the Phe31 ring. The inhibitor dihedral angles from end-point dynamics show that the perpendicular arrangement is accessible to both binary and ternary forms of the 5′-ethyl derivatives. The structural differences in the 5′-ethyl protein environment between binary and ternary complexes would suggest dissimilar thermodynamics between binary and ternary complexes for this derivative. This is further implied by dissimilar energetic and entropic contributions to the active site thermodynamics between the binary and ternary complexes (Table III). The observation of similar relative binding free energies, given the domination of thermodynamics by active site solvent interactions with the 5′-ethyl moiety in the binary complex, and both protein and solvent interactions in the ternary complex, is an interesting one. However, there appears to be no obvious basis for this similarity.

An examination of differences between the relative binding free energies of the binary and ternary complexes yields further insights into the effects of active site changes on the relative binding of trimethoprim congeners. The point at which the two complexes show the greatest differences occurs in the transformation of trimethoprim to the 3′-ethyl derivative, where the transformation is significantly more favorable in the ternary complex relative to the binary complex. Because the thermodynamics for the 3′-ethyl derivative were the exception to a generally similar trend between the binary and ternary complexes, suggesting a possible lack of convergence, sampling was doubled at each λ value for both the binary and ternary complex transformations. The additional sampling did not result in any significant changes to the relative binding free energy for the ternary complex. However, additional sampling for the binary complex resulted in a 0.7 kcal/mol more favorable relative binding free energy. The mean θ_1, θ_2 conformation from the end-point dynamics of the first 3′-ethyl derivative simulation was 94.9°, 134.4°, implying a conformational transition from the starting conformation of −3°, −104°. The mean θ_1, θ_2 conformation from the end-point dynamics of a second 3′-ethyl derivative thermodynamic simulation, using an identical starting configuration but different seed for dynamics, was −53°, −103.9°, implying a different sampling of configuration space for the inhibitor compared to the first simulation. The thermodynamic values reported in this paper reflect the total sampling carried out for the active site bound 3′-ethyl congeners. Given the same

Table IV. Relative Energetic and Entropic Components to Binding Free Energies in Binary vs Ternary Complexes (kcal/mol)

complex	$\Delta\Delta E_{binary} - \Delta\Delta E_{ternary}$	$T\Delta\Delta S_{binary} - T\Delta\Delta S_{ternary}$
3′-Et	−10.3	−11.8*
4′-Et	−4.1*	−3.4
5′-Et	13.3	13.5*
3′,4′-Et_2	25.0	26.4*
4′,5′-Et_2	16.6	17.4*
3′,5′-Et_2	2.9	3.9*
3′,4′,5′-Et_3	−21.0*	−19.7

[a] An asterisk denotes the dominant term.

solvation thermodynamics for the 3′-ethyl congener for binary and ternary complexes (ΔA_2 from eq 1), the thermodynamic difference between binary and ternary complexes originates in the more favorable active site contribution in the ternary complex. The energetic and entropic contributions to the active site thermodynamics are quite different between the binary and ternary complex 3′-ethyl derivatives, suggesting slightly different active site environments for the 3′-ethyl moieties. End-point molecular dynamics from initial and additional sampling indicates that both a stacked planar and perpendicular arrangement of the benzyl ring with Phe31 are possible for the 3′-ethyl congener in the binary complex. In both arrangements, the 3′-ethyl moiety is highly exposed to solvent. Multiple positioning of the benzyl ring is not observed in the end-point dynamics of the ternary complexes, and appears to be consistent with active site steric restrictions imposed by the "teen" loop and cofactor. Although much of the 3′-ethyl group is solvated both in the binary and ternary complexes, the 3′-ethyl group appears to remain in close proximity to Phe31 in the ternary complex. This restricted positioning of the benzyl group in the ternary complex together with increased interactions with the protein likely impact the relative binding thermodynamics.

The large free energy differences between the binary and ternary complexes for the 4′,5′-diethyl, 3′,5′-diethyl, and 3′,4′,5′-triethyl congeners are illustrated in Figure 9. It is clear that there is a more favorable environment for transforming trimethoprim to the ethylated derivative in the binary complex environment. Is there a dominant energetic or entropic basis for this trend? The relative energetic or entropic components may be determined by calculating the differences between the binary and ternary complexes (Table IV). The dominant components are marked with an asterisk in the table. No single dominant energetic or entropic term contributes to the relative free energies. The 5′-ethyl, 3′,4′-, 4′,5′-, and 3′,5′-diethyl congeners in the binary complex are favored over similar transformations in the ternary complex for entropic reasons, while the 4′-ethyl and 3′,4′,5′-triethyl derivatives are dominated by favorable energetic criteria. Thus, a subtle balance of energetic and entropic components contributes to the more favored transformation in the binary complex.

Computed Thermodynamic Values Versus Experimental Inhibitory Activity. To the best of our knowledge, only three of the inhibitors considered in this study have been examined experimentally for the *E. coli* ternary complex. Inhibitory potencies required to reduce reaction rate by 50% (IC_{50}) have been reported for trimethoprim, the 3′,5′-diethyl derivative, and the 3′,4′,5′-triethyl derivative with values of 7 ± 3.5 nM, 14 ± 7 nM, and 10 ± 5 nM, respectively.[5,6,34] This suggests that both the diethyl and triethyl derivatives are slightly less potent inhibitors than trimethoprim. K_i values have been reported for trimethoprim and the triethyl derivative, based on IC_{50} measurements, and have values of 8.1 and 8.00, respectively.[27] The relation $\Delta G_{Et} - \Delta G_{tmp(+)} = -RT \ln K_{Et}/K_{tmp(+)}$ yields the free energy of binding the triethyl congener relative to trimethoprim, similar to that calculated via our method. An experimental value of +0.1 kcal/mol is obtained from K_i values. However, this value has limits of −0.53 kcal/mol and +0.82 kcal/mol based on the stated reliability of the IC_{50} values. The free energy perturbation calculations suggest a relative binding free energy of −0.7 ± 0.4 kcal/mol. Thus, our results fall within the bounds determined by experiment, with the qualitative result that the triethyl congener and trimethoprim bind approximately equivalently to the *E. coli* enzyme.

(34) Margins, of error for IC_{50} values were computed based on 90% tolerance limits of ±50% as reported by Roth and co-workers (see ref 5 and references therein).

Summary

Relative free energies of binding ethylated congeners of trimethoprim have been computed for both the binary and ternary complexes of inhibitor with *E. coli* dihydrofolate reductase. Seven ethylated congeners were examined using thermodynamic perturbation techniques. The conformational features of the active site complexes with trimethoprim and congeners are examined based upon molecular dynamics data collected at the thermodynamic end points. The binary complexes are characterized by an "open" active site environment in which water molecules contact both the pyrimidine and benzyl portions of the inhibitor, and no significant benzyl–protein interactions are observed in model structures. Ternary complexes are characterized by a more "closed" active site in which waters contact the inhibitor benzyl moiety at the surface of the active site. Water–inhibitor interactions in the binary complexes are replaced by inhibitor interactions with the "teen" loop and cofactor atoms in the ternary complexes. Most notably, the methionine side chain at residue 20 is in close proximity to the inhibitor benzyl moiety across all eight ternary complex species, suggesting some importance of this residue in inhibitor binding. The number of contacts gained by favorable interactions formed among the benzyl moieties, the Met20 side chain, and the nicotinamide group in the ternary complexes examined suggests a structural basis for the observed positive cooperativity of inhibitor binding associated with the addition of cofactor. Two classes of conformations are found for the inhibitors in the active site. The ternary complexes appear to sterically allow a single benzyl group conformation characterized by a perpendicular arrangement with the Phe31 aromatic group. The binary complexes allow this conformation, as well as one similar to that found for the chicken crystal structure in which the benzyl group maintains an approximately stacked arrangement with Phe31.

Contrasting the behavior of the relative binding free energies of a congener in the binary versus ternary complexes has provided insight into the contribution of benzyl moieties to inhibitor binding. A favorable relative binding free energy for the ethylated congeners is calculated for 12 of the 14 thermodynamic systems considered in this study. The basis for the favorable binding lies in the dominant desolvation contribution to the relative free energy. Thus, the active site environments significantly change the solvation of the drug. An exception, the 3′-ethyl derivative in the binary complex, exhibits active site thermodynamics similar to the solvated derivative, and appears to be consistent with the observed exposure of the 3′-ethyl moiety to solvent in the bound complex. A relative binding free energy of zero for the 5′-ethyl derivative in the ternary complex is not consistent with a solvent exposure argument, based on the proximity of the transformed group to "teen" loop residue and cofactor atoms. Diethyl and triethyl congeners suggest that the environment for the 5′-Et functional group is favored in the binary complex over the ternary complex. A strong preference was noted for the 3′,4′-diethyl derivative over either the 3′- or 4′-ethyl derivatives in both the binary and ternary complexes. This is clearly due to an increased favorable desolvation term for the diethyl derivative. Also common to both the binary and ternary complexes is the decreased relative binding free energy of the 4′,5′-diethyl derivative relative to the 3′,4′-diethyl derivative. The basis for this difference lies in a less favorable active site contribution in the 4′,5′-diethyl derivative in the relative binding free energy. The calculated differences in the energetic and entropic contributions to relative binding free energies between the binary and ternary complexes do not support a single force (i.e., energy or entropy) which dominates the preference of the transformation in the binary active site environment across all congeners. These results, like the previous paper in this series which details the binding of these inhibitors in the chicken DHFR enzyme complex, suggests a subtle balance of

energetic and entropic components which are sensitive to small differences in active site environments.

The total of thermodynamic data gathered for this system, and the ternary complex for the chicken species of DHFR reported in the previous paper in this series, are being used to examine the question of species specificity of trimethoprim for the bacterial species. The results of these studies are to be presented elsewhere.

Acknowledgment. Support from the NIH (Grant 37554) is gratefully acknowledged. C.L.B. is an Alfred P. Sloan Research Fellow. This research was supported in part by a grant from the Pittsburgh Supercomputing Center through the NIH Division of Research Resources Cooperative Agreement U41RR04154 and through a grant from the National Science Foundation Cooperative Agreement ASC-8500650.

Tunneling Dynamics in Isotopically Substituted Malonaldehyde. Comparison between Symmetric and Asymmetric Species

Enric Bosch, Miquel Moreno, and Josê M. Lluch*

Contribution from the Departament de Quîmica, Universitat Autônoma de Barcelona, 08193 Bellaterra, Barcelona, Spain. *Received June 7, 1991*

Abstract: A theoretical study based on bidimensional model surfaces that retains the main features of the potential hypersurfaces has been performed in order to analyze the influence of primary and secondary effects on intramolecular proton transfer tunneling frequencies of malonaldehyde in the gas phase. To this aim, several isotopically substituted species have been considered. The tunneling frequency is calculated and compared for the different species. It is found that substitution of the transferring proton by deuterium or tritium greatly diminishes the tunneling whereas isotopic substitution of hydrogens not directly involved in the transfer has no appreciable effect when the symmetry of the double well is preserved. However, these secondary effects dramatically increase only by the introduction of a slight asymmetry in the double well potential by isotopic substitution. All these results agree with previous experimental data on isotopically substituted malonaldehyde species.

I. Introduction

The possibility of hydrogen tunneling in chemical reactions was recognized almost from the beginning of quantum theory.[1] Some of the most definite information about the part played by the tunnel effect in many processes comes from the study of isotope effects[1-16] on the rate of chemical processes.

Isotope effects can be classified as primary and secondary according to whether they involve isotopic substitution in a bond which is made or broken during reaction (i.e. primary) or in some other part of the reacting species (i.e. secondary). Primary deuterium and tritium isotope effects in proton-transfer reactions have been extensively used, but much less attention has been given to the secondary isotope effects.[4]

Experimentally, malonaldehyde (MA) is the prototype and most thoroughly investigated molecule for the study of proton transfer in polyatomic systems.[2,3,15] Regarding the isotope effects in this intramolecular reaction, the work of Baughcum et al.[2,3] is noteworthy since the primary and secondary effects are studied by the substitution of hydrogens by deuterium in different positions. An important point suggested by the authors is that a small asymmetry in the molecule induced by isotopic substitution could have an appreciable quenching effect when the transferring hydrogen has been substituted by a deuterium and only a partial effect when the transferring atom is a hydrogen.

From a theoretical point of view, a number of authors have dealt with the MA system.[17-24] However, to our knowledge, the isotopic substitution is only considered in the work of Shida et al.[22] where the primary isotope effect is evaluated.

The goal of this paper is to theoretically study the influence of the isotope effects on proton tunneling frequencies by analyzing several species of MA in the gas phase. Special attention will be

(1) Bell, R. P. *The Tunel Effect in Chemistry*; Chapman & Hall: London, 1980.
(2) Baughcum, S. L.; Duerst, R. W.; Rowe, W. F.; Smith, Z.; Wilson, E. B. *J. Am. Chem. Soc.* **1981**, *103*, 6296.
(3) Baughcum, S. L.; Smith, Z.; Wilson, E. B.; Duerst, R. W. *J. Am. Chem. Soc.* **1984**, *106*, 2260. Turner, P.; Baughcum, S. L.; Coy, S. L.; Smith, Z. *J. Am. Chem. Soc.* **1984**, *106*, 2265.
(4) Saunders, W. H., Jr. *J. Am. Chem. Soc.* **1985**, *107*, 164.
(5) Dormans, G. J. M.; Buck, H. M. *J. Am. Chem. Soc.* **1986**, *108*, 3253.
(6) Cha, Y.; Murray, C. J.; Klinman, J. P. *Science* **1989**, *243*, 1325. Grant, K. L.; Klinman, J. P. *Biochemistry* **1989**, *28*, 6597.
(7) Chantranupong, L.; Wildman, T. A. *J. Am. Chem. Soc.* **1990**, *112*, 4151.
(8) Wolfe, S.; Hoz, S.; Kim, C.; Yang, K. *J. Am. Chem. Soc.* **1990**, *112*, 4186.
(9) Amin, M.; Price, R. C.; Saunders, W. H., Jr. *J. Am. Chem. Soc.* **1990**, *112*, 4467.
(10) Redington, R. L. *J. Chem. Phys.* **1990**, *92*, 6447.
(11) Tucker, S. C.; Truhlar, D. G. *J. Am. Chem. Soc.* **1990**, *112*, 3338.
(12) Truong, T. N.; Truhlar, D. G. *J. Chem. Phys.* **1990**, *93*, 2125.
(13) Lu, D.-h.; Maurice, D.; Truhlar, D. G. *J. Am. Chem. Soc.* **1990**, *112*, 6206 and references 15-30 cited therein.
(14) Firth, D. W.; Barbara, P. F.; Trommdorsf, H. P. *Chem. Phys.* **1989**, *136*, 349.
(15) Seliskar, C. J.; Hoffman, R. E. *J. Mol. Spectrosc.* **1981**, *88*, 30; **1982**, *96*, 146.
(16) Swinney, T. C.; Kelley, D. F. *J. Phys. Chem.* **1991**, *95*, 2430.
(17) Fluder, E. M.; De la Vega, J. R. *J. Am. Chem. Soc.* **1978**, *100*, 5265.
(18) De la Vega, J. R. *Acc. Chem. Res.* **1982**, *15*, 185.
(19) Bicerano, J.; Schaefer, H. F., III; Miller, W. H. *J. Am. Chem. Soc.* **1983**, *105*, 2550.
(20) Carrington, T., Jr.; Miller, W. H. *J. Chem. Phys.* **1986**, *84*, 4364.
(21) Hutchinson, J. S. *J. Phys. Chem.* **1987**, *91*, 4495.
(22) Shida, N.; Barbara, P. F.; Almlöf, J. E. *J. Chem. Phys.* **1989**, *91*, 4061.
(23) Makri, N.; Miller, W. H. *J. Chem. Phys.* **1989**, *91*, 4026.
(24) Bosch, E.; Moreno, M.; Lluch, J. M.; Bertrán, J. *J. Chem. Phys.* **1990**, *93*, 5685.

0002-7863/92/1514-2072$03.00/0