Energetically Most Likely Substrate and Active-Site Protonation Sites and Pathways in the Catalytic Mechanism of Dihydrofolate Reductase

Peter L. Cummins and Jill E. Gready*

Contribution from the Computational Molecular Biology and Drug Design Group, John Curtin School of Medical Research, Australian National University, P.O. Box 334, Canberra ACT 2601, Australia

Received November 1, 2000

Abstract: Despite much experimental and computational study, key aspects of the mechanism of reduction of dihydrofolate (DHF) by dihydrofolate reductase (DHFR) remain unresolved, while the secondary DHFR-catalyzed reduction of folate has been little studied. Major differences between proposed DHF mechanisms are whether the carboxylate group of the conserved active-site Asp or Glu residue is protonated or ionized during the reaction, and whether there is direct protonation of N5 or a proton shuttle from an initially protonated carboxylate group via O4. We have addressed these questions for *both* reduction steps with a comprehensive set of ab initio quantum chemical calculations on active-site fragment complexes, including the carboxyl side chain and, progressively, all other polar active-site residue groups including conserved water molecules. Addition of two protons in two steps was considered. The polarization effects of the remainder of the enzyme system were approximated by a dielectric continuum self-consistent reaction field (SCRF) model using an effective dielectric constant (ϵ) of 2. Optimized geometries were calculated using the density functional (B3LYP) method and Onsager SCRF model with the 6-31G* basis. Single-point energy calculations were then carried out at the B3LYP/6-311+G** level with either the Onsager or dielectric polarizable continuum model. Additional checking calculations at MP2 and HF levels, or with other basis sets or values of ϵ, were also done. From the results, the conserved water molecule, corresponding to W206 in the *E. coli* DHFR complexes, that is H-bonded to both the OD2 oxygen atom of the carboxyl (Asp) side chain and O4 of the pterin/dihydropterin ring, appears critically important and may determine the protonation site for the enzyme-bound substrates. In the absence of W206, the most stable monoprotonated species are the neutral-pair 4-enol forms of substrates with the carboxyl group OD2 oxygen protonated and H-bonded to N3. If W206 is included, then the most stable forms are still the neutral-pair complexes but now for the N3−H keto forms with the protonated OD2 atom H-bonding with W206. A second proton addition to these complexes gives protonations at N8 (folate) or N5 (DHF). Calculated H-bond distances correlate well with those for the conserved W206 observed in many X-ray structures. For all structures with occluded M20 loop conformations (closed active site), OD2−N3 distances are less than OD2−NA2 distances, which is consistent with those calculated for *protonated* OD2 complexes. Thus, the results (B3LYP; $\epsilon = 2$ calculations) support a mechanism for both folate and DHF reduction in which the OD2 carboxyl oxygen is first protonated, followed by a direct protonation at N8 (folate) and N5 (DHF) to obtain the active cation complexes, i.e., doubly protonated. The results do not support a proposed protonated carboxyl with DHF in the enol form for the Michaelis complex, nor an ionized carboxyl with protonated enol-DHF as a catalytic intermediate. However, as additional calculations for the monoprotonated complete complexes show a reduction in the energy differences between the neutral-pair keto and ion-pair keto (N8- or N5-protonated) forms, we are extending the treatment using combined quantum mechanics and molecular mechanics (QM/MM) and molecular dynamics simulation methods to refine the description of the protein/solvent environment and prediction of the relative stabilization free energies of the various (OD2, O4, N5, and N8) protonation sites.

Introduction

The catalytic mechanism of the enzyme dihydrofolate reductase (DHFR) has been extensively studied by experimentalists and theoreticians.[1] Initially the attention arose largely from its importance as a target for anti-folate cytotoxic drugs, which created a strong structural and biophysical literature[2,3] underpinning structure-based design programs,[4] including our own.[5] However, despite these considerable efforts to define the catalytic chemistry, specifically the origin of the proton required for the overall hydride-ion transfer reaction and the role of a conserved active-site carboxylate residue, key aspects remain unresolved. Several unusual aspects of DHFR chemistry, compared with simple dehydrogenase enzymes catalyzing analogous chemistry, may bear on the apparent subtlety on the mechanism. One is that the net reaction takes place within the

* Corresponding author. E-mail: Jill.Gready@anu.edu.au.

(1) Blakely, R. L. In *Advances in Enzymology*; Meister, A., Ed.; John Wiley: New York, 1995; Vol. 70, p 23.

(2) Kraut, J.; Matthews, D. A. In *Biological Macromolecules and Assemblies*; Jurnak, F. A., McPherson, A., Eds.; John Wiley: New York, 1987; Vol. 3, pp 1−72.

(3) Freisheim, J. H.; Matthews, D. A. In *Antagonists as Therapeutic Agents*; Sirotnak, F. M., Burchall, J. J., Ensminger, W. D., Montgomery, J. A., Eds.; Academic Press: Orlando, 1984; Vol. 1, pp 69−131.

(4) Huennekens, F. M. *Adv. Enzyme Regul.* **1994**, *34*, 397.

(5) Gready, J. E.; Cummins, P. L. In *Free Energy Calculations in Rational Drug Design*; Reddy, M. R., Erion, M. D., Eds.; Kluwer Academic/Plenum Publishers: Boston, in press.

10.1021/ja0038474 CCC: $20.00
Published on Web 03/23/2001

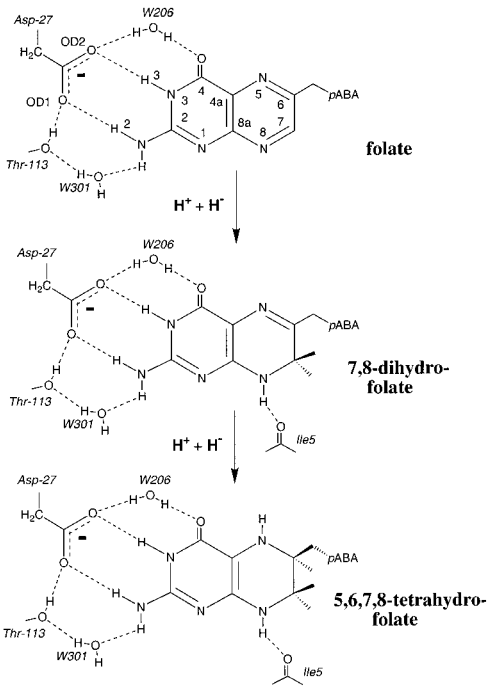

Figure 1. Reduction of folate to 7,8-dihydrofolate (DHF) and 5,6,7,8-tetrahydrofolate (THF) with substrates H-bonded to the carboxylate side chain of the conserved Asp/Glu residue in the active site of DHFR. pABA = *p*-aminobenzoic acid. Other possible H-bond interaction sites, including those with water molecules W206 and W301, are indicated for folate, DHF, and THF based on X-ray structures of *E. coli* DHFR complexes available in the Protein Data Bank (see text).

pyrazine ring of the complex substituted heterobicyclic pteridine ring system, which contains several positions with accessible pK_a values for protonation or deprotonation[6,7] (see Figure 1). A second is that the enzyme catalyzes two reactions, with the fully oxidized form of folate cofactor and with the 7,8-dihydrofolate (DHF) form. Although the first reaction is physiologically much less important in turnover terms, an obligate need to scavenge the oxidized form or initially reduce the oxidized form obtained from nutritional sources in the majority of organisms, including animals, which do not synthesize DHF de novo, may have imposed special constraints on the mechanism.[8] The fact that the majority of interactions between the two substrates and the active site, as revealed by X-ray crystallography, are with the pyrimidine ring, i.e., away from the reaction center, may be related to both these aspects. These interactions exhibit an extensive H-bonded network involving the conserved acidic residue (Asp or Glu) and active-site water molecules, as shown in Figure 1 for the *Escherichia coli* enzyme.[9] Note that, as usual in the DHFR literature, the acidic (Asp27) residue in Figure 1 is depicted in the anionic form, so one of the conserved waters, W206, is assumed to donate a H-bond to both OD2 and O4 of substrate. However, this H-bonding orientation with W206 will change if either OD2 or O4 is protonated.

The pterin rings of folate and DHF substrates have pK_a values of less than 4,[10,11] and thus are unprotonated in solution at physiological pH. Furthermore, the preferred ring protonation site for folate is N1, not N8 as required for the enzymic reduction, although both experiment and computation on pterin analogues indicate N8 protonation is energetically accessible (pK_a of ~1.5 compared with ~2 for N1).[6,7] The transfer of a hydride ion from the cofactor nicotinamide adenine dinucleotide phosphate (NADPH) in the presence of the enzyme requires that a proton be transferred (see Figure 1) to the substrate (S) at N8 (folate) and N5 (DHF) according to

$$S + H^+ + NADPH \rightarrow SH_2 + NADP^+ \qquad (1)$$

For elucidating the catalytic mechanism, particularly the protonation steps, a major limitation of experimental methods is that they can only be performed on binary or analogue ternary ($NADP^+$ or H_2NADPH) complexes, and not on the active NADPH form of the folate or DHF complexes. Direct determination of the substrates' ionization states in the active ternary complexes is thus not possible. Experimental and computational attempts to define the likely tautomer and protonation state of the catalytically active form of DHF have produced ambiguous or contradictory results requiring further interpretations and assumptions. Chen et al.[12] used Raman spectroscopy to determine a pK_a value of 6.5 for N5 in the *E. coli* DHFR·DHF·$NADP^+$ complex, an increase of 2.5−4 from the solution value,[10,11] while a value of less than 4.0 was obtained for the binary complex and DHFR·DHF·H_2NADPH and Asp27Ser-DHFR·DHF·$NADP^+$ ternary complexes. They concluded that direct protonation of N5 is possible in the $NADP^+$ complex, and by inference in the active ternary complex, but that stabilization of N5-protonated DHF requires specific interactions which are not present in the other complexes. Unfortunately, the N5-protonated form of the *E. coli* DHFR·DHF·$NADP^+$ complex is not stable, so its X-ray structure cannot be determined, while the analogous *Lactobacillus casei* DHFR complex is also not stable enough for NMR studies.[13] The Poisson−Boltzmann calculations of Cannon et al.[14] were based on an electrostatic model for the *E. coli* DHFR·DHF·NADPH complex derived from (unprotonated) *E. coli* DHFR·DHF·$NADP^+$ X-ray coordinates.[9] They predicted a 3−4 unit downward shift in pK_a for DHF, suggesting that direct N5 protonation is highly unlikely in the enzyme active site. However, from calculated shifts in the pK_a of Asp27 for the (normal) 4-oxo and 4-hydroxy tautomers of the DHF ternary complexes, they concluded that the latter was more likely, as this gave an *increase* of ~2 units, compatible with an observed pK_a for catalysis of 6.5 and suggesting the active Michaelis complex with protonated Asp27 and enol-DHF. Although Cannon et al.[14] reported quantum chemical calculations of Raman frequencies of active-site fragments of this enol-DHF complex form as being consistent with the data of Chen et al.,[12] the later work of Deng and Callender[15] with quantum chemical calculations of Raman frequencies and additional Raman experiments questioned this interpretation. This work[15] further supported direct protonation of N5 of DHF in the DHFR·DHF·$NADP^+$ complex, although no conclusion on the 4-oxo/hydroxy tautomeric state was possible, but suggested that Asp27 was unprotonated. It also suggested that the immediate environment of N5 was hydro-

(6) Pfleiderer, W. In *Comprehensive Heterocyclic Chemistry—The Structure, Reactions, Synthesis and Uses of Heterocyclic Compounds*; Boulton, A. J., McKillop, A., Eds.; Pergamon Press: New York, 1984; pp 30−80.
(7) Gready, J. E. *J. Comput. Chem.* **1985**, *6*, 377.
(8) Gready, J. E. *Nature* **1979**, *282*, 674.
(9) Bystroff, C.; Oatley, S. J.; Kraut, J. *Biochemistry* **1990**, *29*, 3263.
(10) Poe, M. *J. Biol. Chem.* **1977**, *252*, 3724.
(11) Maharaj, G.; Selinsky, B. S.; Appleman, J. R.; Perlman, M.; London, R. E.; Blakely, R. L. *Biochemistry* **1990**, *29*, 4554.
(12) Chen, Y.-Q.; Kraut, J.; Blakley, R. L.; Callender, R. *Biochemistry* **1994**, *33*, 7021.
(13) Casarotto, M. G.; Basran, J.; Badii, R.; Sze, K. H.; Roberts, G. C. K. *Biochemistry* **1999**, *38*, 8038.
(14) Cannon, W. R.; Garrison, B. J.; Benkovic, S. J. *J. Am. Chem. Soc.* **1997**, *119*, 2386.
(15) Deng, H.; Callender, R., *J. Am. Chem. Soc.* **1998**, *120*, 7730.

phobic (consistent with quoted unpublished X-ray results), thus lacking an immediate water molecule required for the catalytic mechanism suggested by Cannon et al.[14] from the enol-DHF form. On intuitive grounds, one might expect that the pK_a of DHF in the DHFR·DHF·NADP$^+$ complex would be shifted in the opposite direction, i.e., to a lower value, due to the destabilizing effects of two positive charges (DHFH$^+$ and NADP$^+$) in close proximity. However, the fact that the inhibitor methotrexate (MTX) is found to be protonated in the DHFR·MTX·NADP$^+$ complex, with NMR chemical shifts very similar to those for the DHFR·MTX·NADPH complex,[16] indicates such simple arguments are not valid. In summary: rationalization of pK_a shifts in the DHFR active site is problematic, and probably highly dependent on subtle conformational changes and degree of solvation in the various inactive complexes and, by inference, in the active complex.

Due to its conserved H-bond interaction with the substrate, it has generally been assumed that the conserved carboxyl group plays some crucial role in the DHFR catalytic mechanism. However, as for protonation of the substrates, the ionization state of the carboxyl group in the active form of the complexes also remains unclear. The conserved water molecule that forms a H-bond to both OD2 and O4 in crystal structures of DHFR complexes almost certainly affects the structure of complexes and may have a direct influence on the carboxyl group ionization state.[9] The results of Chen et al.[12] using the Asp27Ser *E. coli* DHFR mutant showed that the presence of the carboxyl group had a large effect on the pK_a of DHF in the DHFR·DHF·NADP$^+$ complex. Substitution of Asp27 with Asn in *E. coli* DHFR also leads to a dramatic decrease in catalysis.[17] In contrast, the corresponding mutations in *L. casei* DHFR have been shown to have a relatively small effect on the rate of hydride-ion transfer, consistent with the effect of mutations of more distant residues.[18] Other researchers have focused attention on direct determination of the protonation state of the conserved carboxyl (Asp/Glu) side chain, using NMR[13,19,20] and Raman[21] techniques. Asp26 in the *L. casei* DHFR *apo* and binary complexes with folate and DHF, and Asp27 in *apo E. coli* DHFR, have been shown to be ionized above a pH of 5.[13,19,21] The active-site carboxylate (Glu 30) group also appears to be ionized in recombinant human DHFR.[20]

From the various pieces of information obtained from kinetic, X-ray structure, Raman, and NMR spectroscopic measurements and theoretical calculation, several catalytic mechanisms have been postulated. The conventional mechanism assumes that protonation takes place first, followed by the hydride-ion transfer.[22] We suggested previously that the ionized form of the carboxyl side chain might preferentially stabilize the N8 (folate)- and N5 (DHF)-protonated pterin rings.[23] This suggestion has also been adopted by Chen et al.,[12] who argue on the basis of pK_a determinations that N5 should be directly protonated from solvent in the active site. However, the source of this

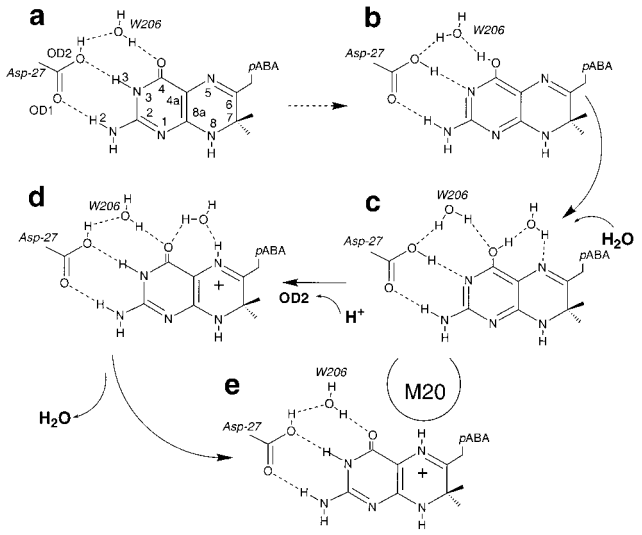

Figure 2. Catalytic mechanism for DHF reduction proposed by Bystroff et al.[9] based on crystallographic structural analysis of the *E. coli* DHFR·folate·NADP$^+$ ternary complex. The conserved water molecule W206 together with a mobile solvent water assist transfer of a proton from OD2 to N5 to form the active complex **e**, in which it is proposed that the flexible loop M20 displaces the solvent water molecule.

proton in the active ternary complex remains unclear. The Asp/Glu residue is the only ionizable residue in the DHFR active site, but its carboxyl group is clearly too distant (see Figure 1) from N5 to allow for direct transfer of a proton. Instead of direct protonation of N5 in the active complex, other mechanisms have been proposed that involve a transient protonation of O4 with solvent-assisted relay of the proton to N5.[9,13,14,24−28] In the mechanism proposed by Bystroff et al.,[9] it is assumed that the conserved active-site carboxyl group remains protonated (i.e., neutral) and that the proton comes from solution, as illustrated in Figure 2. In this mechanism, a neutral carboxyl group is required to promote protonation of the dihydropterin ring. A tightly bound water molecule bridging OD2 and O4, corresponding to water W206 in the *E. coli* DHFR·folate·NADP$^+$ complex, is suggested to stabilize the carboxylic acid form. As shown in Figure 2, the OD2-protonated form of the complex **a** interconverts to an enol form, **c**, with the aid of W206 and a solvent water molecule. A second protonation at OD2 then facilitates the transfer of a proton from O4 to N5, forming **d**, followed by expulsion of the solvent molecule by the closing of the mobile loop (M20) to give the active complex **e**.

However, alternative mechanisms suggest that in the active form of the complex there is a proton on O4. On the basis of their NMR results for *apo*, DHFR·folate, DHFR·DHF, and DHFR·DHF·H_2NADPH complexes which show the Asp26 residue (*L. casei*) in ionized form, Cassarotto et al.[13] argue that the catalytic role of the carboxylate may be to polarize the substrate in such a way as to favor the enol form (O4 protonation) as opposed[23] to the keto form (N5 protonation). It is proposed that the transfer of the proton from O4 to N5 takes place via a solvent water and is concerted with hydride-ion transfer. This concerted proton/hydride transfer mechanism is similar to the mechanism proposed by Cannon et al.,[14] except that here the N3 proton is transferred to the carboxyl group.

(16) Cheung, H. T. A.; Birdsall, B.; Feeney, J. *FEBS Lett.* **1992**, *312*, 147.
(17) Howell, E. E.; Villafranca, J. E.; Warren, M. S.; Oatley, S. J.; Kraut, J. *Science* **1986**, *231*, 1123.
(18) Basran, J.; Casarotto, M. G.; Barsukov, I. L.; Roberts, G. C. K. *Biochemistry* **1995**, *34*, 2872.
(19) Birdsall, B.; Casarotto, M. G.; Cheung, H. T. A.; Basran, J.; Roberts, G. C. K.; Feeney, J. *FEBS Lett.* **1997**, *402*, 157.
(20) Blakley, R. L.; Appleman, J. R.; Freisheim, J. H.; Jablonsky, M. J. *Arch. Biochem. Biophys.* **1993**, 306, 501.
(21) Chen, Y.-Q.; Kraut, J.; Callender, R. *Biophys. J.* **1997**, *72*, 936.
(22) Huennekens, F. M.; Scrimgeour, K. G. In *Pteridine Chemistry*; Pfleiderer, W., Taylor, E. C., Eds.; Pergamon Press: New York, 1964; pp 355−376.
(23) Gready, J. E. *Biochemistry* **1985**, *24*, 4761.
(24) Morrison, J. F.; Stone, S. R. *Biochemistry* **1988**, *27*, 5499.
(25) Uchimaru, T.; Tsuzuki, S.; Tanabe, K.; Benkovic, S. J.; Furukawa, K.; Taira, K. *Biochem. Biophys. Res. Commun.* **1989**, *161*, 64.
(26) Brown, K.; Kraut, J. *Faraday Discuss.* **1992**, *93*, 217.
(27) Reyes, V. M.; Sawaya, M. R.; Kraut, J. *Biochemistry* **1995**, *34*, 2710.
(28) Lee, H.; Reyes, V. M.; Kraut, J. *Biochemistry* **1996**, *35*, 7012.

Such proposed mechanisms, however, ignore the case of folate reduction, where there is no obvious means of protonation at N8 by either the protein or water molecules bound in the active site of DHFR. This complexity prompted our previous suggestion of a possible "non-obvious" mechanism to account for both folate and DHF reductions.[8,29] It should be noted that NMR evidence for folate binding to *L. casei* DHFR in the binary complex and in the catalytically relevant folate·$NADP^+$ complex indicates the 4-oxo form is favored.[19]

Opinions on the mechanism of DHF reduction, therefore, seem to be largely divided between direct protonation of DHF at N5 and indirect protonation via O4. The other main points on which they differ are the ionization state of the conserved active site Asp/Glu and, together with the conserved water (W206 in *E. coli* DHFR), its precise role in the catalytic mechanism, and whether there is a solvent water molecule between O4 and N5 in the active complex.

In this paper we consider the relative stabilities of the various proposed active-complex species and the likelihood of proposed protonation pathways. We have carried out advanced QM calculations including in the *E. coli* DHFR active complexes the conserved water molecule W206, or both W206 and the solvent water molecule. Apart from the interactions with W206 and the solvent water, there are other direct H-bonding contacts with the Asp27 and/or substrate species that may be important. As shown in Figure 1, the crystal structures exhibit H bonds between Thr113 (*E. coli* DHFR) and OD1 and between another bridging water molecule, W301, and the 2-amino group, as well as a possible H-bond interaction between N8−H and the carbonyl backbone of Ile5 (*E. coli* DHFR). Hence, we have also performed calculations on complexes modeling the effects of these interactions explicitly. The polarizing effects of the remaining protein and solvent medium have been approximated using a dielectric continuum self-consistent reaction field (SCRF) method.

Methods

The conserved active-site carboxylate or carboxylic acid side chain was modeled by acetate or acetic acid, and the substrates folate and DHF were modeled by 6-methyl-substituted pterin and dihydropterin. In calculations including H-bond interactions with Thr113 or Ile5 (Figure 1), we modeled the hydroxyl group of Thr113 by a water molecule, and the carbonyl group of Ile5 by formaldehyde. The calculations were done using SCRF methods, i.e., with the model systems embedded in a continuous and isotropic dielectric medium. As proteins are not homogeneous dielectric media, the use of SCRF methods for modeling the local protein environment is not rigorously justified. However, despite the simplifications inherent in the dielectric continuum approach as applied to proteins, we have found that SCRF methods can be of practical use when calibrated against results obtained from microscopic protein simulations.[30,31] As we do not yet have the necessary microscopic simulation data for the present problem, it is important to gauge the likely effects of the remainder of the protein solvent system that is not explicitly included in the model, as even low dielectric constants ($\epsilon \cong 2$) that are typical of nonpolar solvents are found to give rise to significant contributions to the relative stabilities of neutral-pair and ion-pair H-bonded complexes.[31−33] By dispensing with the wider atomic detail of the protein environment, the dielectric continuum SCRF approach allows the most important interactions to be treated at a high level of QM theory with quality basis sets, and including electron correlation and geometry optimization.

The SCRF method requires a value for the macroscopic dielectric constant (ϵ). An appropriate value for the dielectric constant depends on the type of properties under investigation and needs to be chosen carefully with reference to experimental or, as discussed above, calculated simulation data.[31,34−37] It must be remembered, however, that the macroscopic dielectric "constant" for a protein is, in fact, not constant as it varies depending on position in the protein.[37] A guide to the range of dielectric constants that are encountered in protein systems can be obtained from molecular dynamics (MD) simulations on globular proteins.[38,39] These studies have yielded macroscopic dielectric constants ϵ of 2−3 for protein interiors, and up to 11−21 for the whole molecule.[38] The latter higher values are due almost entirely to charged residues on the protein surface, while $\epsilon \cong 10$ is typically found near active sites and individual ionizable groups.[39] Justification for use of a low dielectric constant comes from our previous calculations on DHFR,[40,41] which show small (0.04 electron) overall polarization of the pterin ring by an atom-centered point-charge model for enzyme, indicating that the atomic charges produce small fields in the active-site region, and also from published Poisson−Boltzmann model calculations[14] which show the active site is situated in a region of low electric field strength. At this point, we note that the most important polar interactions are likely to be due to directly H-bonded species that are not adequately treated by a dielectric continuum approximation. Hence, in the present study, the active-site groups directly H-bonded to the carboxyl−substrate complex are successively added until all are ultimately included within the QM region in the largest calculations. As these interactions account for most of the polar groups in the active site, a dielectric $\ll 10$ would seem a reasonable choice for the present SCRF model. Consequently, as we include H bonds with the substrate explicitly, we generally used the lower value of 2, i.e., a "nonpolar" protein interior environment,[38] for the dielectric constant. However, as this choice of an effective ϵ is not calibrated against microscopic simulation data,[30,31] some calculations were also carried out with dielectric constants of $\epsilon = 1$ (no polarization) and $\epsilon = 80$ (strong polarization) for comparison. It should also be stressed that varying ϵ alone will not account for any additional deficiencies in the active-site structural model itself that may affect relative stabilities.

All calculations were performed using the Gaussian 98 program (G98).[42] As SCRF gradients are not available at the MP2 level in G98, we have used a density functional theory (DFT) approach to include the effects of electron correlation. There is now considerable evidence indicating that the B3LYP density functional[43,44] generally yields properties of useful accuracy[45−47] and, for many H-bonded systems, in

(29) Gready, J. E. *Adv. Pharmacol. Chemother.* **1980**, *17*, 37.
(30) Cummins, P. L.; Gready, J. E. *J. Phys. Chem. B* **2000**, 104, 4503.
(31) Cummins, P. L.; Gready, J. E. *J. Mol. Graph. Model.* **2000**, *18*, 42.
(32) Zheng, Y.-J.; Ornstein, R. L. *J. Am. Chem. Soc.* **1996**, *118*, 11237.
(33) Barril, X.; Aleman, C.; Orozco, M.; Luque, F. J. *Proteins: Struct. Funct. Genet.* **1998**, *32*, 67.
(34) Warshel, A.; Russell, S. T. *Q. Rev. Biophys.* **1984**, *17*, 283.
(35) Warshel, A.; Åqvist, J. *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 267.
(36) Sham, Y. Y.; Chu, Z. T.; Warshel, A. *J. Phys. Chem. B* **1997**, *101*, 4458.
(37) Warshel, A.; Papazyan, A. *Curr. Opin. Struct. Biol.* **1998**, *8*, 211.
(38) Simonson, T.; Brooks, C. L. *J. Am. Chem. Soc.* **1996**, *118*, 8452.
(39) King, G.; Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1991**, *95*, 4366.
(40) Greatbanks, S. P.; Gready, J. E.; Limaye, A. C.; Rendell, A. P. *Proteins: Struct., Funct. Genet.* **1999**, *37*, 157.
(41) Greatbanks, S. P.; Gready, J. E.; Limaye, A. C.; Rendell, A. P. *J. Comput. Chem.* **2000**, *21*, 788.
(42) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski, V. G.; Montgomery, J. A., Jr.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A. D.; Kudin, K. N.; Strain, M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.; Pomelli, C.; Adamo, C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98*, Revision A.6; Gaussian, Inc.: Pittsburgh, PA, 1998.
(43) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
(44) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
(45) Koch, W.; Holthausen, M. C. *A Chemist's Guide to Density Functional Theory*; Wiley-VCH: Weinheim, 2000.
(46) Miani, A.; Cane, E.; Palmieri, P.; Trombetti, A.; Handy, N. C. *J. Chem. Phys.* **2000**, *112*, 248.
(47) Cohen, A. J.; Handy, N. C. *Chem. Phys. Lett.* **2000**, *316*, 160.

close agreement with those obtained from MP2 calculations.[48−51] Moreover, DFT methods are far more efficient than conventional ab initio correlated methods and, therefore, are better suited to the larger-scale calculations required in this study. However, as a cautionary note, there appear to some pathological cases, specifically for systems containing anionic species, where the DFT method breaks down.[40,41] Consequently, some single-point calculations at the MP2 level were also performed in order to validate the B3LYP results. The geometry optimizations were carried out at the B3LYP/6-31G* level using the SCRF model ("dipole" option in G98) based on the Onsager equation,[52]

$$\Delta G = -\frac{1-\epsilon}{1+\epsilon}\frac{\mu^2}{R^3} \qquad (2)$$

where μ is the molecular dipole moment inside a cavity of radius R. In the present calculations the cavity radius was set equal to 5.5 Å, as determined in a previous study on similar H-bonded pterin complexes.[31] For the majority of complexes, the relative energies were then obtained at the B3LYP/6-311+G** level using the Onsager model (B3LYP/6-31G*) optimized geometries and the dielectric polarizable continuum model (DPCM)[53−55] ("PCM" option in G98), with the default values for the atomic radii. For the largest complexes studied, i.e., those including all possible H-bonded species (Figure 1), the Onsager model was used for both the B3LYP/6-31G* geometry optimizations and the single-point B3LYP/6-311+G** level calculations for obtaining relative energies. The results of some additional calculations, including geometry optimizations and those using a range of basis sets at both HF and correlated levels, are also reported to illustrate the convergence of the calculated energy differences.

Results

Unbound Pterins. We present first the results for protonation of the 6-methyl-pterin and 6-methyl-7,8-dihydropterin molecules in the absence of H-bonding partners found in the active site (Figure 1). These calculations on the unbound pterins were carried out with a dielectric of 2, but with a smaller radius of 3.5 Å for the Onsager model to allow for a more meaningful comparison with the larger H-bonded complexes. The B3LYP/6-311+G** level energies calculated at the B3LYP/6-31G* optimized geometries give the N8-protonated form of unbound 6-methyl-pterin to be 10.2 kcal/mol more stable than the O4-protonated form, while N5-protonated 6-methyl-dihydropterin is only 1.7 kcal/mol more stable than the corresponding O4-protonated form. An additional calculation on the N1-protonated form of 6-methyl-pterin, which is that preferred in solution,[6] shows that it is marginally (0.5 kcal/mol) less stable than the N8-protonated form. Clearly, this protonation would be highly unlikely in the active site, as no H-bond partner for N1−H exists and it would disrupt the Thr113 and W301 interaction that is observed crystallographically.

Pterins H-Bonded to Carboxyl Group. The most stable protonated forms of the H-bonded complexes formed between acetate and the 6-methyl-pterin and -dihydropterin molecules in the absence of explicit water molecules or other neighboring active-site residue analogues are shown in Figures 3 and 4. Figure 3 shows the most stable structures for the N8- and O4-protonated complexes formed between the acetate ion and folate analogue, i.e., the ion-pair **2** and neutral-pair **3** forms, respectively. Figure 4 shows the stable configurations for the N5- and O4-protonated complexes formed between the acetate ion and DHF analogue, again the ion-pair **7** and neutral-pair **8** forms, respectively. The ion-pair complexes are characterized by relatively short H-bond lengths. In contrast to the results obtained for unbound pterins which favored the N8- or N5-protonated pterin ring, the O4-protonated forms (**3** and **8**) for both folate and DHF analogues are found to be more stable than the N8 or N5 forms (**2** and **7**) by approximately 10 kcal/mol. Note also that on protonation of O4 the hydrogen (H3) that was covalently bonded to N3 in the unbound pterins and initial anion complexes **1** and **6** is covalently bonded to OD2 in the neutral-pair complexes **3** and **8**, i.e., a concurrent proton shift occurs. Moreover, we found no energy minimum at the B3LYP/6-31G* level corresponding to the ion-pair form of the

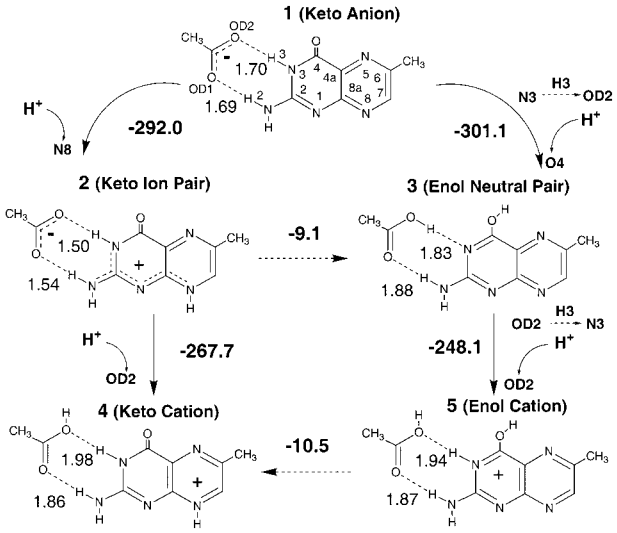

Figure 3. Mono- and diprotonated forms of the H-bonded anion complex formed between acetate and the folate analogue 6-methyl-pterin. All proton affinities and energy differences are given in kilocalories per mole, calculated at the B3LYP(6-311+G**//6-31G*) level with the DPCM model ($\epsilon = 2$). H-bond distances are in angstroms.

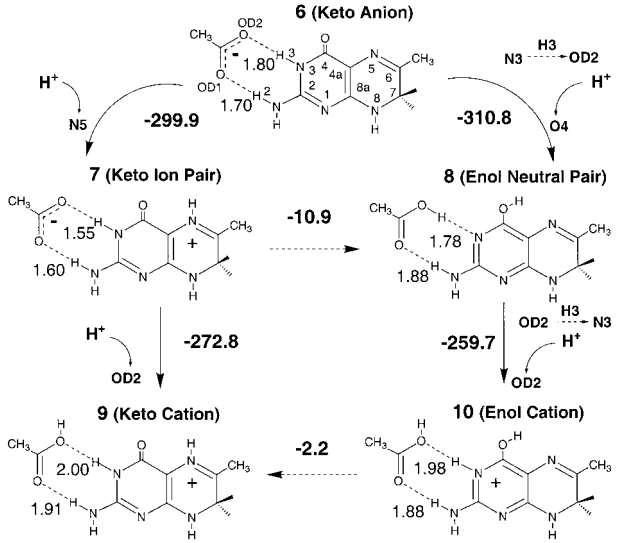

Figure 4. Mono- and diprotonated forms of the H-bonded anion complex formed between acetate and the 7,8-dihydrofolate analogue 6-methyl-7,8-dihydropterin. All proton affinities and energy differences are given in kilocalories per mole, calculated at the B3LYP(6-311+G**//6-31G*) level with the DPCM model ($\epsilon = 2$). H-bond distances are in angstroms.

(48) McAllister, M. A. *Can. J. Chem.* **1997**, *75*, 1195.
(49) McAllister, M. A. *J. Mol. Struct.: THEOCHEM* **1998**, *427*, 39.
(50) Pan, Y. P.; McAllister, M. A. *J. Mol. Struct.: THEOCHEM* **1998**, *427*, 221.
(51) Lozynski, M.; Rusinskaroszak, D.; Mack, H. G. *J. Phys. Chem.* **1998**, *102*, 2899.
(52) Onsager, L. *J. Am. Chem. Soc.* **1936**, *58*, 1486.
(53) Miertus, S.; Scrocco, E.; Tomasi, J.*Chem. Phys.* **1981**, *55*, 117.
(54) Miertus, S.; Tomasi, J. *Chem. Phys.* **1982**, *65*, 239.
(55) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* **1996**, *255*, 327.

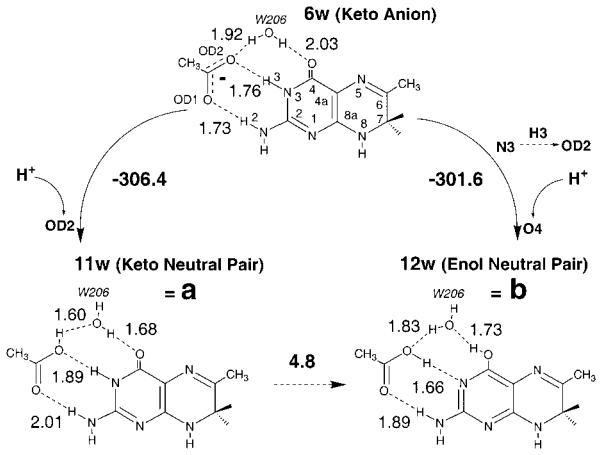

Figure 5. Protonated forms of the H-bonded anion complex formed between acetate and the 7,8-dihydrofolate analogue 6-methyl-7,8-dihydropterin with crystallographic water molecule W206 bound. All proton affinities and energy differences are given in kilocalories per mole, calculated at the B3LYP(6-311+G**//6-31G*) level with the DPCM model ($\epsilon = 2$). H-bond distances are in angstroms. **11w** and **12w** are analogous to the structures **a** and **b** in Figure 2.

O4-protonated complex. We also found low barriers (<2 kcal/mol) for the transfer of the H3 proton in the H-bond between N3 and OD2 for the ion-pair (**2** and **7**) and neutral-pair forms of the keto complexes, similarly to our previous findings for 8-methyl-pterin.[31] Consequently, in a low-dielectric medium, the true representation of the keto state is most likely a hybrid of ion-pair (N3−H3) and neutral-pair (OD2−H3) forms. Clearly, based on the molecular dipole moments alone, the more polar keto ion-pair becomes more stable than the keto neutral-pair formed by migration of H3 to OD2 as the dielectric constant is increased. The ion-pair actually becomes the more stable only for a dielectric constant of $\epsilon \gg 2$ and depends on the quality of the basis sets (ref 31 and unpublished results). Protonation at OD2 of the keto ion-pair and enol neutral-pair complexes produces stable cations for both folate (**4** and **5**) and DHF (**9** and **10**) analogues. For this second protonation, H3 is covalently bound to N3 in all complexes and the states with N8 (**4**) and N5 (**9**) protonated are more stable than those protonated on O4 (**5** and **10**), by 10.5 and 2.2 kcal/mol, respectively.

Pterins H-Bonded to Carboxyl Group, W206, and Solvent Water. The structures in Figures 3 and 4 lack the H-bonds associated with the other neighboring polar groups in the active site (Figure 1). We begin our investigation of these additional H-bond effects by first introducing the conserved water molecule (W206 for *E. coli*) and the mobile solvent water molecule (Figure 2) only. In Figure 5 we show the protonations of the anion complex for DHF with W206 to give the keto and enol neutral-pair forms **11w** and **12w** which are, respectively, the analogues of **a** and **b** of the Bystroff et al.[9] mechanism defined in Figure 2. We find that **12w** is 4.8 kcal/mol higher in energy than **11w**. The protonations of the anion complex for DHF with both W206 and a solvent water molecule are shown in Figure 6. The H-bonding of the solvent water molecule causes an insignificant change (only 0.1 kcal/mol) in the energy difference between tautomers **11ww** and **12ww** compared with that for **11w** and **12w** (Figure 5). Comparison of the structures with **11w** and **12w** in Figure 5 shows that there is also very little change in the H-bond geometry of W206 on binding of the solvent water molecule in the protonated species. However, comparison of **11ww**, **12ww**, and **8ww** with the anion **6ww** shows that there are very significant changes in the H-bonding of the solvent molecule itself on protonation. The H-bond with O4 in the anion **6ww** is absent in the protonated species **11ww** and **12ww** so that there is now only a single H-bond between the water molecule and N5. Changes also occur in the binding of W206 in order to accommodate the proton at OD2. Whereas the H-bond distances in the anion are nearly equal (∼1.7 Å), the OD1···H2 distance is noticeably longer than the OD2···H3 distance in the protonated complexes. However, note that in **8ww** these differences are also very large, although W206 is now less strongly bound to OD2 (distance of 2.05 Å), while the solvent water is bound tightly to the dihydropterin ring mainly via O4 (1.64 Å). Note also that the optimal orientation of W206 in **8ww** is not the same as that proposed for **c** in the mechanism illustrated in Figure 2.

In the mechanism proposed by Bystroff et al.,[9] the enol form **c** converts on protonation at OD2 to the cation form **d**. In Figure 7 we have examined possible energetic pathways for this conversion. Direct protonation of **8ww** (closely analogous to **c**) at OD2 produces the enol form (**10ww**) of the cation complex which may convert to the preferred keto form (**9ww**), which is the analogue of **d**. Thus, **9ww** remains the most stable configuration for H-bonding between the carboxyl group and dihydropterin in the presence of the two water molecules. Alternatively, we considered the case where proton transfer to N5 precedes OD2 protonation. This can be achieved by direct transfer of H3 from OD2 to N3 in **8ww** to produce the enol (**13ww**) ion-pair form, and then via proton transfer through the bound water molecule to the keto (**7ww**) ion-pair complex. Although **13ww** is very similar in energy to **8ww**, the binding of W206 in **13ww** becomes much tighter than that in **8ww**, although not as tight as that in the OD2-protonated complexes: the distance of 2.5 Å between W206 and O4 in **13ww** is rather long.

The keto form of the ion-pair (**7ww**) and not the enol ion-pair **13ww** is the closest in energy (∼3 kcal/mol) to the keto neutral-pair **11ww** (Figure 6), although the energy differences are rather small: note that less than 1 kcal/mol separates **8ww**, **13ww**, and **7ww**. This is in contrast to the results obtained in the absence of bound water (Figures 3 and 4), where the enol neutral-pair form (**8**) is more stable by 10 kcal/mol than the keto ion-pair form (**7**). Note also that the keto form of the cation (**9ww**) is preferred by approximately 5 kcal/mol over the enol cation (**10ww**) in complexes with explicit waters (Figure 7), compared with 2 kcal/mol for the corresponding complexes (**9** and **10**) without water (Figure 4).

As the mobile solvent water is not observed to be bound crystallographically (Figure 1), we also calculated energy differences between the ion-pair and neutral-pair complexes with only W206 bound. Figure 8 illustrates direct protonation of N5 in the H-bonded anion complex **6w** formed between acetate, DHF analogue, and crystallographic water molecule W206 to form the keto ion-pair complex **7w**, compared with the corresponding OD2 protonation to form the keto neutral-pair **11w** (analogue of **a** in Figure 2). The formation of **11w** may then be followed by direct protonation of N5 to form the Michaelis complex **9w** (analogue of **e** in Figure 2). The OD2-protonated form **11w** is almost 10 kcal/mol more stable than the corresponding keto form of the ion-pair (**7w**). Figure 9 shows the results of corresponding calculations for the folate analogue complexes. The OD2-protonated neutral-pair form (**14w**) is 11 kcal/mol more stable than the N8-protonated ion-pair form (**2w**).

Pterins H-Bonded to Carboxyl Group, Thr113, Ile5, W206, and W301. We performed calculations to see how the remaining H-bond interactions with Thr113, W301, and Ile5 (Figure 1) would affect the energy differences between the keto

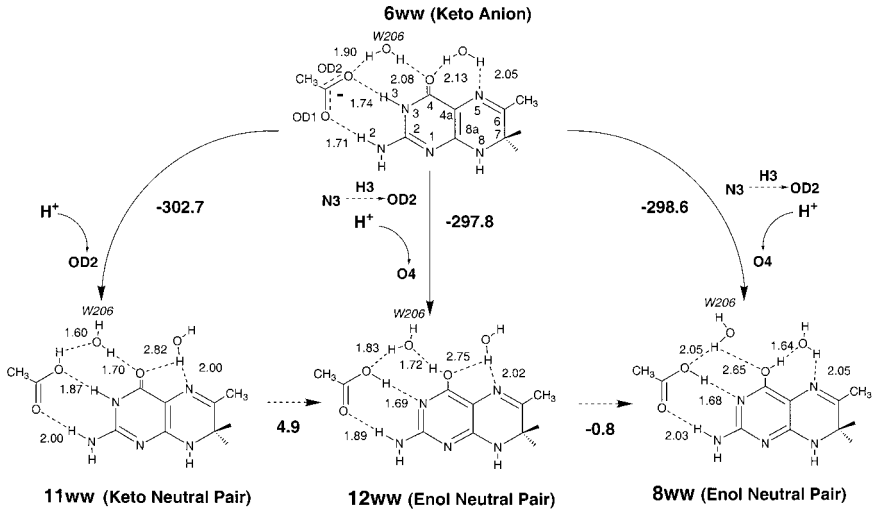

Figure 6. Protonated forms of the H-bonded anion complex formed between acetate and the 7,8-dihydrofolate analogue 6-methyl-7,8-dihydropterin with both the crystallographic water molecule W206 and a solvent water molecule bound. All proton affinities and energy differences are given in kilocalories per mole, calculated at the B3LYP(6-311+G**//6-31G*) level with the DPCM model ($\epsilon = 2$). H-bond distances are in angstroms.

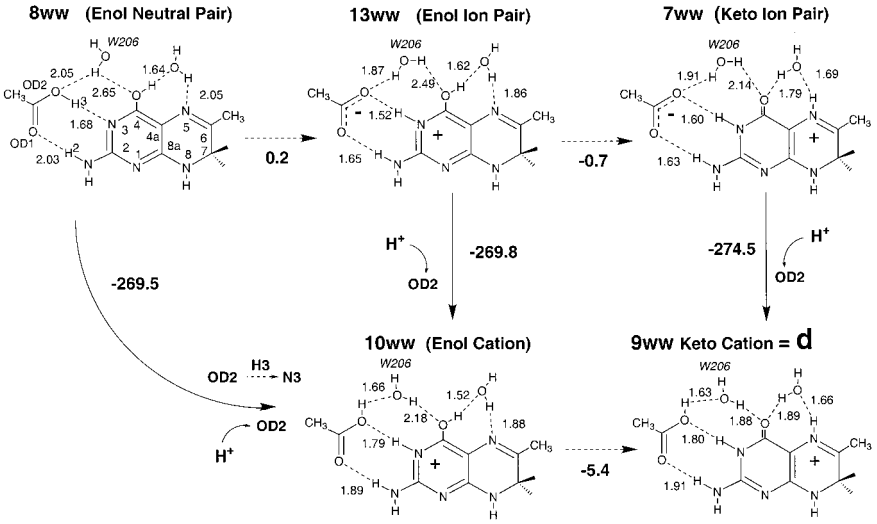

Figure 7. Pathways for the formation of the OD2- and N5-protonated cation complexes between acetate and the 7,8-dihydrofolate analogue 6-methyl-7,8-dihydropterin with both the crystallographic water molecule W206 and a solvent water molecule bound (see analogue **d** in Figure 2). All proton affinities and energy differences are given in kilocalories per mole, calculated at the B3LYP(6-311+G**//6-31G*) level with the DPCM model ($\epsilon = 2$). H-bond distances are in angstroms.

ion-pair and neutral-pair (Figures 8 and 9). The results for the full complement of H-bond interactions are shown in Figure 10. These remaining H-bond interactions reduce the difference in energy between the complexes **2w′** and **14w′** and between **7w′** and **11w′**, compared with **2w** and **14w** in Figure 9, and **7w** and **11w** in Figure 8, respectively. The H-bond distances between the pterins and carboxyl group are also significantly larger compared with the corresponding ones in **2w**, **14w**, **7w**, and **11w**. Nevertheless, the OD2-protonated neutral-pair retains stability over the N5- and N8-protonated ion-pair complexes by 3 kcal/mol for $\epsilon = 2$.

Comparison of Different QM Methods. The results in Table 1 give the energy differences at different levels of QM theory and dielectric constant for the DHF analogue systems **7w′** and **11w′**. We note first that the level of theory at which the geometry is optimized is relatively unimportant. In particular, the polarization of geometries in dielectric media does not have a significant impact on the predictions of relative energies. However, the HF method overestimates the stability of the neutral-pair **11w′** relative to the ion-pair complex **7w′** compared with the correlated (B3LYP and MP2) methods. Most significantly, the B3LYP functional also overestimates the stability of the neutral-pair relative to the ion-pair complex compared with the MP2 method in dielectric ($\epsilon > 1$) media, particularly for strong dielectrics, although the agreement between B3LYP and MP2 in a vacuum ($\epsilon = 1$) remains quite good. To an excellent approximation, the calculated energy differences as a function of $(\epsilon - 1)/(\epsilon + 1)$ are linear (see eq 2). The energy difference for a given value of ϵ was, therefore, obtained by interpolation of the calculated energy differences in Table 1, and we determined that the ion-pair (**7w′**) becomes more stable than the neutral-pair (**11w′**) for $\epsilon > 2.6$ (MP2) and $\epsilon > 21.6$ (B3LYP).

Comparison of H-bond Distances from X-ray Crystallography and Calculation. X-ray crystal structures are available for DHFR from a number of sources, including human. However, the most extensive structural analysis has been done for *E. coli* DHFR.[56] X-ray structural analysis has identified geometries corresponding to open, closed, and occluded conformations of the M20 loop. It is believed that the closed conformation of the M20 loop is the catalytically active one,

(56) Sawaya, M. R.; Kraut, J. *Biochemistry* **1997**, *36*, 586.

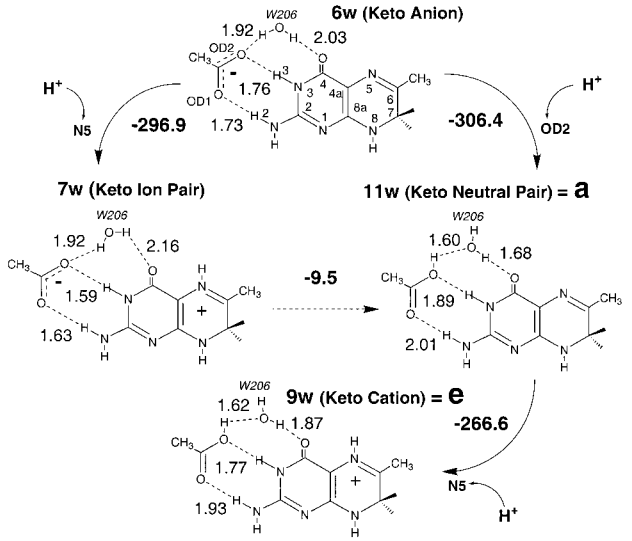

Figure 8. Direct protonation of N5 in the H-bonded anion complex formed between acetate, the 7,8-dihydrofolate analogue 6-methyl-7,8-dihydropterin, and the crystallographic water molecule W206, compared with OD2 protonation (analogue of **a** in Figure 2) followed by direct protonation of N5 to form the Michaelis complex (analogue of **e** in Figure 2). All proton affinities and energy differences are given in kilocalories per mole, calculated at the B3LYP(6-311+G**//6-31G*) level with DPCM model ($\epsilon = 2$). H-bond distances are in angstroms.

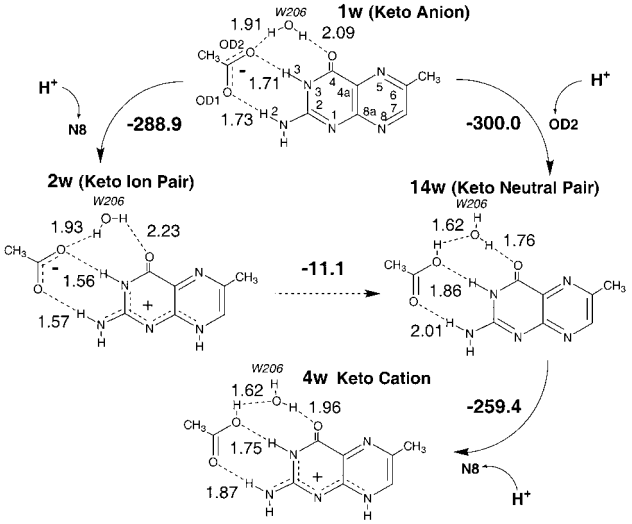

Figure 9. Direct protonation of N8 in the H-bonded anion complex formed between acetate, the folate analogue 6-methyl-pterin, and the crystallographic water molecule W206, compared with OD2 protonation followed by direct protonation of N8 to form the Michaelis complex. All proton affinities and energy differences are given in kilocalories per mole, calculated at the B3LYP(6-311+G**//6-31G*) level with the DPCM model ($\epsilon = 2$). H-bond distances are in angstroms.

while the occluded conformation forms in product complexes.[56] We have extracted the H-bond distances involved in the binding of the conserved water W206 using coordinates obtained from the Protein Data Bank for comparison with the geometries obtained theoretically. The results are shown in Tables 2 and 3 for some *E. coli* DHFR folate, DHF, and THF analogue complexes. As a number of the crystal structures are for 5-deaza complexes, also included in Table 2 are H-bond distances from calculations on 5-deaza complexes. In Table 3, the calculated distances are for the 6-methyl-5-deaza-5,6,7,8-tetrahydropterin analogue of dideazatetrahydrofolate (ddTHF) with protonation and H-bonding as in **11w** (Figure 8). The calculations show that the perturbations to the pterin or dihydropterin rings by replacing N5 with CH produce only minor changes to the H-bond dimensions. Both theory and experiment predict that OD2−W distances are less than O4−W distances (W = water oxygen). The same trend is observed for the human enzyme (PDB files 1dhf and 1drf; results not shown). The theoretical geometries give OD2−W distances of about 2.6 Å for all complexes except the anions. In some ddTHF complexes (occluded M20 conformation) this distance is much lower at 2.3 Å (Table 3), while 2.8 Å is observed for the DHF complex (Table 2), also in the occluded conformation. Except for anion and ion-pair complexes where the calculated carboxyl−pterin H-bond distances for OD2−N3 and OD1−N2 are about equal, we calculate that OD2−N3 H-bonding distances are less than OD1−N2 distances. The FOL crystal structures show some variation for these two distances, but the majority follow the same trend, i.e., OD2−N3 < OD1−N2 as for the calculated distances.

Discussion

The results show that H-bonding with the carboxyl group has a significant effect on the protonation energetics of the pterin and dihydropterin rings. The preference of the O4-protonated (neutral-pair **8**) over the N5-protonated (ion-pair **7**) DHFR analogue complex (Figure 4) is in agreement with the HF/3-21G gas-phase ($\epsilon = 1$) calculations of Cannon et al.[14] Similarly, we found that the neutral-pair O4-protonated complex **3** is more stable than the ion-pair N8-protonated complex **2** for the folate analogue (Figure 3). In both cases, the N3 proton that is bound to the pterin and dihydropterin rings in the unbound form and in the anion complexes with the carboxylate group becomes, on protonation of the anion complexes, covalently bound to the carboxylate, giving effectively the carboxylic acid and enol tautomer of the neutral pterin and dihydropterin. Extension of the model to include the conserved W206, and W206 and solvent water, for the DHF analogue (Figures 5 and 6) again showed preferential protonation of the carboxyl group, although in these cases the keto tautomer (**11w**, **11ww**) is more stable than the enol tautomer (**12w**, **12ww**, **8ww**). Thus, we found no evidence to support the idea that the polarizing effect of the carboxylate preferentially stabilizes either the N8- or N5-protonated keto forms or the O4-protonated enol forms in an ion-pair complex.[13,23] We also attempted to obtain a stable complex between the keto tautomer of the dihydropterin and the OD2-protonated group in the absence of W206, i.e., **11w** without the water molecule. However, the optimization resulted in cleavage of the OD1−NA2 H-bond and formation of a new H-bond between the protonated OD2 and O4 of the pterin ring. Thus, the presence of the conserved water appears crucial to stabilizing the OD1−NA2 H-bond interaction when OD2 is protonated.

However, the effects of the other residues that are directly H-bonded to the carboxyl-pterin or carboxyl-dihydropterin complexes are also significant. From the calculated H-bond geometries, a proton relay from OD2 to O4 to N5 involves considerable changes in the entropy associated with binding of the W206 and solvent waters. The conserved water W206, which is initially tightly bound in **a** = **11w** (Figures 2 and 5), becomes weakly bound in the intermediate state **c** (analogous to **12ww** or **8ww** in Figure 6) of the mechanism proposed by Bystroff et al.[9] and must then regain its tightly bound status in **e** = **9w** (Figure 8). Also, the transition from **b** (**12w** in Figure 5) to **c** involves a reorientation of the O4−H bond through a 180° rotation (Figure 2). Note, however, that the O4−N5 transfer initiated by a second protonation at OD2 as proposed by Bystroff

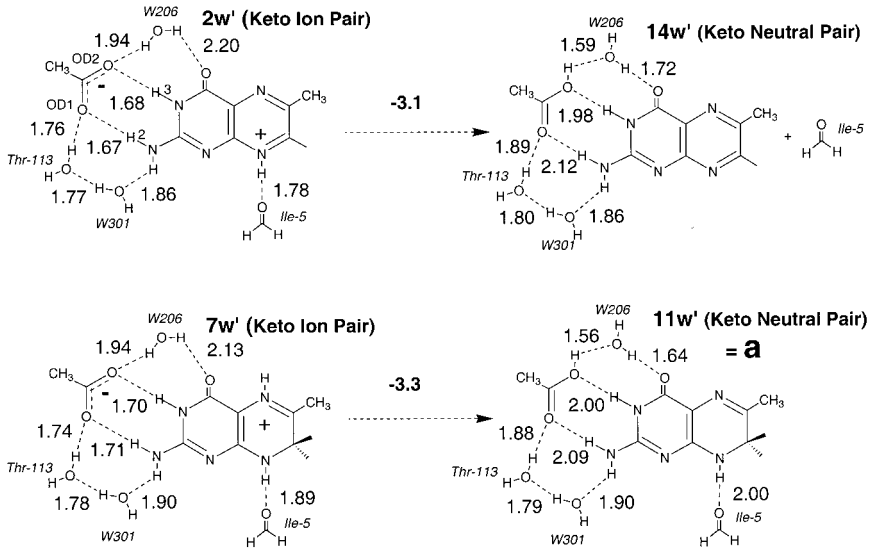

Figure 10. Comparison between N8 (**2w′**)- and OD2 (**14w′**)-protonated (folate analogue) and between N5 (**7w′**)- and OD2 (**11w′**)-protonated (DHF analogue) anion complexes with all possible water and residue H-bond interactions as illustrated in Figure 1. All energy differences are given in kilocalories per mole, calculated at the B3LYP(6-311+G**//6-31G*) level with the Onsager model ($\epsilon = 2$). H-bond distances are in angstroms.

Table 1. Energy Differences (kcal/mol) between N5- and OD2-Protonated (DHF Analogue) Anion Complexes (**7w′** and **11w′** in Figure 10) Calculated at Various Levels of Theory in Vacuum ($\epsilon = 1$), Weak Dielectric ($\epsilon = 2$), and Strong Dielectric ($\epsilon = 80$)

optimization level	energy difference[a]	dielectric constant in SCRF (ϵ)		
		1	2	80
HF/3-21G	HF/3-21G	18.9	16.0	10.7
($\epsilon = 1$)	B3LYP/6-31G*	8.5	5.6	0.1
HF/6-31G*	HF/6-31G*	16.4	12.1	4.4
($\epsilon = 1$)	B3LYP/6-31G*	9.8	6.2	−0.7
	MP2/6-31G*	11.2	3.2	−13.5
DFT/6-31G*	B3LYP/6-31G*	10.4	6.5	−0.7
($\epsilon = 2$)	B3LYP/6-31+G*	6.4	2.3	−5.8
	B3LYP/6-311+G**	7.4	3.3	−4.9

[a] Method used to calculate energy difference.

et al.[9] may not be necessary, as direct protonation of **a** at N5 should also be possible via the solvent water (i.e., from **11ww** in Figure 6), with minimal reorganization of the H-bonding network, and without the energy barriers encountered when differently protonated forms interconvert. Furthermore, the calculated orientation of this solvent molecule in the **a** analogue **11ww** (Figure 6) suggests that O4 is not H-bonded to solvent and, thus, need not be involved in the protonation of N5 by way of transfer from OD2.

The other possibility which has been suggested[12,13,23] is that the active form of the substrate complexes is a singly protonated ion-pair, as illustrated by **2** and **7** (Figures 3 and 4) for the N8- and N5-protonated keto forms[12,23] or **13ww** for the O4-protonated form of dihydropterin,[13] rather than the doubly protonated cation complex **e** = **9w** in the Bystroff et al.[9] mechanism. Our results (Figures 3 and 4) indicate that in the absence of nearest-neighbor H-bond interactions other than the carboxyl side chain, these singly protonated ion-pairs are not the most stable forms of the complexes. Rather, the enol neutral-pair forms with OD2 protonated are more stable. Results in Figures 8, 9, and 10 for ion-pairs singly protonated on N8 (**2w**, **2w′**) and N5 (**7w**, **7w′**) after inclusion in the models of W206, and W206, W301, Thr113, and Ile5, show that they are still less stable than the corresponding neutral-pair complexes **14w** and **14w′**, and **11w** and **11w′** (now in the keto tautomer), although the difference is very much less for the larger model calculations. Also, the results in Figure 7 show that H-bonding of both W206 and the solvent water significantly reduces the difference in energy between the ion-pair (**7ww**) and enol neutral-pair (**8ww**) complexes for the DHF analogue. Thus, although the conserved water W206 alone does not confer sufficient stabilization to the ion-pairs for protonations at N5 and N8, as more H-bonds are included the energy difference between the neutral-pair and ion-pair forms generally becomes smaller. This is easily rationalized in terms of the larger dipole of the ion-pair, and a similar effect can also be achieved by increasing the dielectric constant. On the basis of the MP2 calculations for **7w′** and **11w′** (Table 1), it could be argued that only moderately polar environments ($\epsilon \ll 10$) may be capable of stabilizing the formation of the singly protonated ion-pair complexes. Consequently, it is conceivable that ion-pair complexes analogous to **2** and **7** could be more stable than the corresponding neutral-pair forms in the active site of the enzyme. However, our results (Figures 8 and 9) indicate that double protonation to maintain OD2 protonated produces the N8- and N5-protonated cation complexes **4w** and **9w** = **e** required for the two reactions unambiguously.

It should be emphasized that although the effects of both increasing ϵ and the inclusion of explicit H-bonded neighbors are to reduce the energy difference between ion-pair and neutral-pair forms, they are not quantitatively the same. The results (Figure 7) show that the H-bonded water molecules provide additional stabilization to the keto ion-pair form beyond that given by the dielectric continuum model (Figure 4) and illustrate the importance of modeling residues H-bonded to the substrate and Asp carboxylate explicitly. The NϵH of Trp21, neglected in the present model, also forms a H-bond with the conserved W206 in the crystal structure. While this interaction may require consideration, explicit inclusion of a fragment indole ring would be computationally inefficient. This type of secondary polar interaction could best be described within a combined quantum mechanical and molecular mechanical (QM/MM) model for the enzyme system.

In addition to the electrostatic and polarization effects on the relative stabilities of the protonated complexes discussed above, van der Waals repulsions and packing arrangements with other active-site residues impose additional constraints on the 3-D structure of the H-bonded complexes. These effects too would be accounted for in a QM/MM calculation, but are neglected

Table 2. H-Bond Lengths (in Å) in Folate, N5-Deazafolate, and Dihydrofolate Complexes from X-ray Structures Compared with Results from Calculations

complex[a]	PDB code[b]	L[c]	W[d,e]	OD2−N3		OD1−NA2	OD2−W		O4−W
FOL·NADP$^+$	1ra2	Op	601	2.79	<	2.94	2.63	<	2.95
FOL·NADP$^+$	1rb2 A	Op	479	3.07	≈	3.09	2.44	<	2.65
FOL·NADP$^+$	1rb2 B	Op	469	2.30	≈	2.33	2.49	<	2.57
FOL·NADP$^+$	1rx2	Cl	362	2.62	<	2.97	2.69	<	2.82
FOL·ATP-ribose	1ra8	Op	601	2.81	≈	2.83	2.61	<	3.02
FOL	1rd7 A	Op	422	2.96	<	3.10	2.46	<	3.46
FOL	1rd7 B	Op	nd	2.97	>	2.77	nd		nd
FOL	1re7 A	Op	nd	3.16	>	2.75	nd		nd
FOL	1re7 B	Op	nd	2.64	<	2.82	nd		nd
FOL	1dy1 A	Oc	403	2.59	<	2.92	2.77	<	3.03
FOL	1dy1 B	Oc	403	2.79	<	2.86	2.80	<	3.02
DZF	1dyh A	Oc	403	2.63	<	2.79	2.52	<	2.91
DZF	1dyh B	Oc	403	2.71	<	2.88	2.76	<	2.81
DHF	1rf7	Oc	308	2.69	<	3.05	2.83	<	3.09
				Calculated Complexes					
folate analogue	**1w** (keto anion)		2.76	≈	2.77	2.85	<	3.04	
	2w (keto ion-pair)		2.65	≈	2.64	2.87	<	3.15	
	2w′ (keto ion-pair)		2.71	≈	2.73	2.88	<	3.13	
	14w (keto neutral)		2.88	<	3.01	2.58	<	2.75	
	14w′ (keto neutral)		2.96	<	3.07	2.57	<	2.71	
(5-deaza)	**14w** (keto neutral)		2.86	<	3.06	2.58	<	2.75	
	4w (keto cation)		2.79	<	2.89	2.57	<	2.90	
(5-deaza)	**4w** (keto cation)		2.80	<	2.91	2.57	<	2.88	
DHF analogue	**6w** (keto anion)		2.80	≈	2.78	2.85	<	3.00	
	7w (keto ion-pair)		2.69	≈	2.67	2.86	<	3.10	
	7w′ (keto ion-pair)		2.74	≈	2.76	2.88	<	3.08	
	11w (keto neutral)		2.90	<	3.01	2.57	<	2.68	
	11w′(keto neutral)		2.97	<	3.04	2.55	<	2.65	
(5-deaza)	**11w** (keto neutral)		2.87	<	3.07	2.56	<	2.68	
	9w (keto cation)		2.81	<	2.94	2.57	<	2.84	

[a] FOL, folate; DZF, N5-deazafolate; DHF, dihydrofolate. [b] Molecules A and B of crystallographic dimer. [c] Conformation of M20 loop in *E. coli* DHFR: Op, open; Cl, closed; Oc, occluded. [d] Water corresponding to W206 in the Bystroff et al.[9] structure determination. [e] nd = water not observed in some binary complexes with open loop.

Table 3. H-Bond Lengths (in Å) in 5,10-Dideazatetrahydrofolate from X-ray Structures Compared with Results from Calculations for 6-Methyl-5-deazatetrahydropterin Analogue Complex

complex	PDB code[b]	L[c]	W[d]	OD2−N3		OD1−NA2	OD2−W		O4−W
ddTHF[a]	1rx5	Oc	321	2.81	<	3.14	2.29	<	2.84
ddTHF·ATP-ribose	1rx4	Oc	321	2.87	<	2.93	2.31	<	2.92
ddTHF·NADPH	1rx6	Oc	321	2.87	<	2.91	2.30	<	2.78
ddTHF·NADP$^+$	1rc4	Oc	212	2.72	<	3.04	2.48	<	2.93
ddTHF	1dyj A	Oc	403	2.63	<	2.88	2.55	<	2.93
ddTHF	1dyj B	Oc	403	2.76	<	2.88	2.62	<	2.75
				Calculated Complex[e]					
6-methyl-5-deaza-5,6,7,8-tetrahydropterin				2.87	<	3.11	2.57	<	2.67

[a] ddTHF, 5,10-dideazatetrahydrofolate. [b] Molecules A and B of crystallographic dimer. [c] Conformation of M20 loop in *E. coli* DHFR; Oc, occluded. [d] Water corresponding to W206 in the Bystroff et al.[9] structure determination. [e] Protonation and H-bonding with W206 in the same configuration as the neutral-pair keto form in **11w** (Figure 8).

in the present SCRF model calculations. Nevertheless, comparing the calculated H-bond dimensions with those found in crystal structures of DHFR shows a qualitative correlation with respect to the relative H-bond distances of bound W206 in that the distance between OD2 and W206 is invariably shorter than the O4−W206 distance. There are more significant variations in the H-bonding between the carboxyl group of Asp27 and the pterin and dihydropterin rings that are difficult to rationalize, but may partly be due to the effects of crystal packing and uncertainty in the resolution of diffraction data. It appears that for all structures with occluded M20 loop conformations, the X-ray data yield OD2−N3 < OD2−NA2, consistent with the calculations for protonated OD2 complexes. For both protonated and ionized states of OD2, the calculated OD2−W and O4−W distances follow the same trend, i.e., OD2−W < O4−W in Table 2. This is true also for the other complexes (Figures 5 to 10) not listed in Table 2, the only exception being the enol tautomer **12w** (Figure 5), where the order is reversed. Thus, it is not possible to use the results from calculation to confidently differentiate between protonated and ionized states of OD2 in the X-ray structures. However, it is worth noting that the protonated OD2 complexes give rise to stronger H-bonding between OD2, O4, and the conserved water, compared with those complexes in which OD2 is ionized. One of the ternary FOL·NADP$^+$ complexes (1rb2) exhibits relatively stronger bonding for both OD2−W and O4−W, consistent with the OD2-protonated folate analogue complexes **14w** and **14w′**.

Also in Table 2, the water molecule corresponding to W206 is absent from some binary complexes with folate. Solution NMR studies have revealed that this water is bound differently in the binary and ternary (NADPH) complexes of human DHFR with MTX.[57,58] In both of these complexes the water is long-lived, but in the ternary complex it may have a longer residence

(57) Meiering, E. M.; Wagner, G. *J. Mol. Biol.* **1995**, *247*, 294.

time. It is interesting to note that the missing water molecule occurs for binary complexes of folate with the M20 loop in the open conformation, and experimental studies[13,19−21] show that the carboxyl group is ionized in binary complexes. These observations might be a consequence of weaker binding of this water molecule in binary complexes with substrate so that it is not localized and observable by crystallography.

Conclusions

We studied the relative energies of protonated forms of the enzyme carboxyl group and substrate pterin and dihydropterin rings using high-level ab initio quantum chemical methods. We found that the explicit inclusion in the QM models of directly H-bonded polar groups that are observed in the X-ray crystal structures is essential if a correct assessment of these protonation states is to be made. In particular, the conserved water molecule corresponding to W206 in the *E. coli* DHFR complexes appears to be critically important, and may determine the protonation site for the enzyme-bound substrates.

However, as the SCRF model used in the present study is oversimplistic in its treatment of interactions within proteins, we are currently using molecular dynamics simulation and combined quantum mechanics and molecular mechanics (QM/MM) methodologies to obtain an even more realistic description of the protein/solvent environment. The inclusion of free energy sampling and rigorous treatment of long-range effects should allow a more definitive assessment of the relative stabilities of the OD2-, O4-, N5 (DHF)-, and N8 (folate)-protonated forms in actual DHFR enzyme−substrate complexes.

Acknowledgment. We gratefully acknowledge the Australian National University Supercomputer Facility for generous grants of computer time, and the support of an ANU Strategic Development Grant. We thank a reviewer for helpful comments.

JA0038474

(58) Meiering, E. M.; Li, H.; Delcamp, T. J.; Freisheim, J. H.; Wagner, G. *J. Mol. Biol.* **1995**, *247*, 309.