Molecular Dynamics/Free Energy Perturbation Study on the Relative Affinities of the Binding of Reduced and Oxidized NADP to Dihydrofolate Reductase

Peter L. Cummins,[†,‡] K. Ramnarayan,[‡] U. C. Singh,[‡] and Jill E. Gready*[,§]

Contribution from the Department of Molecular Biology, Scripps Clinic and Research Foundation, La Jolla, California 92037, and Department of Biochemistry, University of Sydney, Sydney, NSW 2006, Australia. *Received November 1, 1990*

Abstract: The free energy perturbation (FEP) method, implemented within the molecular dynamics (MD) simulation scheme, has been used to investigate the difference between the binding of reduced and oxidized nicotinamide adenine dinucleotide phosphate (NADP) cofactor to *Escherichia coli* dihydrofolate reductase (DHFR) (binary complex) and *E. coli* DHFR bound to the substrate dihydrofolate (ternary complex) in aqueous solution. The FEP results for the binary complex predict relative equilibrium binding constants for the reduced and oxidized forms of NADP to *E. coli* DHFR in good agreement with the available experimental data, suggesting that reduced NADP binds some 10^2 times more strongly than oxidized NADP. The FEP results also predict reduced NADP in the ternary complex with dihydrofolate to bind more strongly than oxidized NADP. However, this differential is calculated to be 10^2–10^3 times greater than in the binary complex. Although there is no direct experimental information for binding in the active ternary complex with which to compare these results, available results are discussed in the context of the molecular form of active complexes seen kinetically compared with those in the theoretical simulations. The question of the influence of the choice of initial enzyme coordinates and configuration space sampling in these simulations is also discussed. The stronger binding affinity of reduced NADP and the differences between NADP binding strengths computed for the binary and ternary complexes are correlated with solvation effects and structural differences between the complexes. Analysis of the MD structures and available crystallographic data suggests that the positioning of a mobile loop (the "Met-20 loop") plays a key role in determining the relative cofactor binding strengths.

Introduction

With the advent of modern supercomputers, it has recently become feasible to predict with some degree of confidence the physical and chemical properties of biological molecules by using theoretical techniques.[1] One such technique, the free energy perturbation (FEP) method, has a potentially very important application in the area of drug design, since it can be used to compute the relative affinities of binding of different ligands to macromolecules in aqueous solution.[2-7] Several of these studies[5-7] have used the FEP method to determine the relative binding strengths of different modifications of trimethoprim, a known drug, to dihydrofolate reductase. In the general case (Figure 1), one usually considers a ligand A, which may be a known drug compound, and a modified form of the ligand, B. The quantity of interest is the magnitude of the difference between the equilibrium constants for the binding of A and B to a particular protein molecule, E. The ratio of the equilibrium constants for the binding of A and B to E is related to the free energy changes in Figure 1 by

$$K_2/K_1 = \exp[-(\Delta G_2 - \Delta G_1)/RT] \quad (1)$$

However, the free energy difference in eq 1 is also simply equal to the difference in the free energies for the nonphysical processes ΔG_3 and ΔG_4, i.e.

$$\Delta G_2 - \Delta G_1 = \Delta G_4 - \Delta G_3 \quad (2)$$

Consequently, with use of the FEP method, the ratio K_2/K_1 can be obtained by computing the free energy changes for a transformation or "mutation" of A to B in both free and unbound forms, rather than by attempting the computationally more difficult task of calculating the ΔG's for the physical binding processes.

In this paper, we have further tested the FEP method for the general scheme proposed in Figure 1 by performing calculations where E is the enzyme dihydrofolate reductase (DHFR), an important target for a range of chemotherapeutic agents,[8-11] and A and B are respectively the reduced and oxidized forms of the cofactor nicotinamide adenine dinucleotide phosphate (NADP). Calculations were carried out for the binary complex of NADP and *E. coli* DHFR and also for the ternary complex of NADP and the substrate dihydrofolate (DHF) bound to *E. coli* DHFR. Experimental evidence for the binary complexes suggests that the reduced form of NADP (NADPH) binds approximately 10^2 times more strongly than the oxidized form ($NADP^+$):[12-19] this represents a free energy change of only ~3 kcal/mol and is therefore a good test of the accuracy of the theory. Similar differences between $NADP^+$ and NADPH binding have also been reported for *Lactobacillus casei* DHFR.[20-24] While equilibrium binding

[†] Permanent address: Department of Biochemistry, University of Sydney.
[‡] Scripps Clinic and Research Foundation.
[§] University of Sydney.

(1) Van Gunsteren, W. F., Weiner, P. K., Eds. *Computer Simulation of Biomolecular Systems*; ESCOM: Leiden, The Netherlands, 1988.
(2) Tembe, B. L.; McCammon, J. A. *Comput. Chem.* **1984**, *8*, 281–283.
(3) Bash, P. A.; Singh, U. C.; Brown, F. K.; Langridge, R.; Kollman, P. A. *Nature* **1987**, *235*, 574–576.
(4) Singh, U. C. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 4280–4284.
(5) Brooks, C. L. *Int. J. Quantum Chem., Quantum Biol. Symp.* **1988**, *15*, 221–234.
(6) Brooks, C. L.; Fleischman, S. H. *J. Am. Chem. Soc.* **1990**, *112*, 3307–3312.
(7) Fleischman, S. H.; Brooks, C. L. *Proteins: Struc., Funct., Genet.* **1990**, *7*, 52–61.
(8) Hitchings, G. H.; Smith, S. L. *Adv. Enzyme Regul.* **1979**, *18*, 349–370.
(9) Gready, J. E. *Adv. Pharmacol. Chemother.* **1980**, *17*, 37–102.
(10) Blakley, R. L. In *Molecular Actions and Targets for Cancer Chemotherapeutic Agents*; Sartorelli, A. C., Lazo, J. S., Eds.; Academic Press: New York, 1981; pp 303–332.
(11) Freisheim, J. H.; Matthews, D. A. In *Folate Antagonists as Therapeutic Agents*; Sirotnak, F. M., Burchall, J. J., Emsinger, W. B., Montgomery, J. A., Eds.; Academic Press: New York, 1984; Vol. 1, pp 69–131.
(12) Williams, M. N.; Greenfield, N. J.; Hoogsteen, K. *J. Biol. Chem.* **1973**, *248*, 6380–6386.
(13) Poe, M.; Greenfield, N. J.; Williams, M. N. *J. Biol. Chem.* **1974**, *249*, 2710–2716.
(14) Chen, J.-T.; Mayer, R. J.; Fierke, C. A.; Benkovic, S. J. *J. Cell Biochem.* **1985**, *29*, 73–82.
(15) Fierke, C. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 4085–4092.
(16) Stone, S. R.; Morrison, J. F. *Biochemistry* **1982**, *21*, 3757–3765.
(17) Stone, S. R.; Mark, A.; Morrison, J. F. *Biochemistry* **1984**, *23*, 4340–4346.
(18) Penner, M. H.; Frieden, C. *J. Biol. Chem.* **1987**, *262*, 5366–5369.
(19) Appleman, J. R.; Howell, E. E.; Kraut, J.; Blakley, R. L. *J. Biol. Chem.* **1990**, *265*, 5579–5584.
(20) Dunn, S. M. J.; Batchelor, J. G.; King, R. W. *Biochemistry* **1978**, *17*, 2356–2364.
(21) Dunn, S. M. J.; King, R. W. *Biochemistry* **1980**, *19*, 766–773.

0002-7863/91/1513-8247$02.50/0

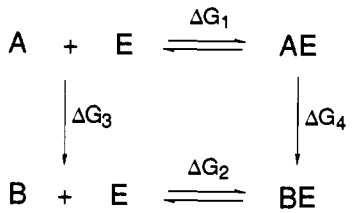

Figure 1. Thermodynamic cycle for the binding of ligands A and B to an enzyme E. ΔG_1 and ΔG_2 are the free energy changes for the physical binding processes. ΔG_3 and ΔG_4 are the free energy changes for the nonphysical mutation of A to B in unbound and enzyme-bound forms, respectively.

in the active ternary complex with dihydrofolate cannot be studied directly, some estimates of an effective binding constant have been made recently in the course of development of complete kinetic schemes for DHFRs from several sources,[25-27] including *E. coli*.[15,18] However, because several molecular changes occur during binding and before the catalytic events, including probable protonation of enzyme and/or substrate as well as an isomerization of the active complex,[28] the molecular form of the initial nonequilibrium ternary complex seen kinetically[15,18] is unknown. In the simulations, however, the molecular form is specified in the starting conditions so that a direct computation of relative binding energies for ternary complexes differing only in the cofactor oxidation state is possible.

By analyzing the resulting MD structures and available crystallographic data, we have attempted to correlate the calculated differences between $NADP^+$ and NADPH binding strengths with structural differences in the binary and ternary complexes. Also, the nature of possible sampling difficulties inherent in the simulations is discussed with reference to the relatively large changes in protein structure observed for DHFR complexed with different ligands, especially the changes between inhibitor and substrate complexes.[29-32]

We have also examined the effect of applying different solvent boundary conditions to the calculated ΔG_3 value in Figure 1 for the cofactor analogues A = *trans*-1-methyldihydronicotinamide (1-Me-H_2Nic) and B = *trans*-1-methylnicotinamide (1-Me-$HNic^+$), which were used to construct models for the reduced and oxidized nicotinamide fragments of the NADP cofactor.

Computational Procedure

Free Energy Perturbation Method. The FEP method is essentially based on the statistical perturbation theory of Zwanzig.[33] If the Hamiltonian for the unperturbed system is H_0 and that for a small perturbation is H_1, the Gibbs free energy difference between the perturbed ($H = H_0 + H_1$) and unperturbed states is given by

$$\Delta G = -RT[\ln (\exp(-H_1/RT))_0] \quad (3)$$

where $H_1 = H - H_0$ and the ensemble average $(\ )_0$ is over the unperturbed system with Hamiltonian H_0. The Gibbs free energy change between two states with Hamiltonians H_A and H_B may then be calculated by defining a hypothetical intermediate Hamiltonian $H(\lambda)$ as a mixture of H_A and H_B via a coupling parameter λ according to the formula

$$H(\lambda) = \lambda H_A + (1 - \lambda)H_B \qquad 0 \le \lambda \le 1 \quad (4)$$

One method of obtaining the total free energy change involves dividing λ into discrete values, λ_i, or so-called "windows", and computing ΔG for each window from the perturbation formula, eq 3, where the perturbation Hamiltonian is given by $H_1 = H(\lambda_i) - H(\lambda_{i\pm1})$. This ensures a smooth transformation from A to B, and the total free energy change is just the sum of contributions from all windows, i.e.

$$\Delta G_{AB} = \sum_i \Delta G_i \quad (5)$$

Evaluating the perturbation free energies ΔG_i in both the forward and backward directions ($\lambda_{i\pm1}$) is a check for possible hysteresis in the calculation and provides a statistical error for the total free energy change.

The FEP method described above has been implemented within the framework of both molecular dynamics (MD) and Monte Carlo simulation schemes by several authors.[34,35] The present calculations were performed with the MD program AMBER (version 3.1), which is a fully vectorized version of AMBER (version 3.0) by Singh et al.,[36] on the Cray X-MP/116se supercomputer at the Scripps Clinic and Research Foundation.

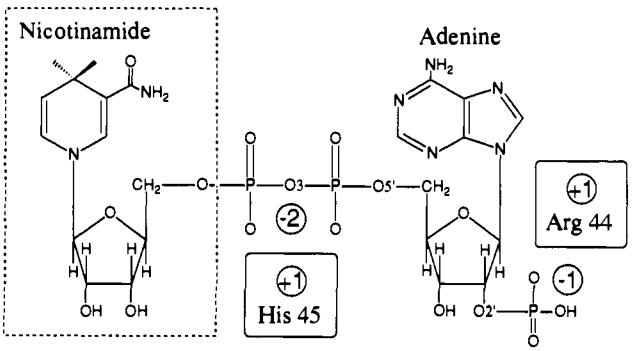

Figure 2. Schematic diagram showing the NADPH cofactor and the salt bridges formed between 5'-phosphate and His-45 and between 2'-phosphate and Arg-44 in *E. coli* DHFR. The perturbed group in the FEP calculations is shown enclosed in dashed lines.

Parameter Assignment. The NADP cofactor (Figure 2) was constructed from adenine–ribose 2'-phosphate, diphosphate, and nicotinamide–ribose fragments by using the PREP module of AMBER. Except for those of the nicotinamide moiety, atomic partial charges and bond, angle, and dihedral parameters were derived from standard literature values[37,38] for the appropriate molecular fragments (e.g., adenine, ribose, etc.). The nicotinamide parameters were obtained by first optimizing the geometry of the cofactor analogues, 1-Me-H_2Nic and 1-Me-$HNic^+$, at the ab initio 6-31G level,[39] followed by single-point calculations at the 6-31G* level[40] in order to obtain the atomic partial charges from the molecular electrostatic potential.[41] The ab initio calculations were carried out by using the program QUEST.[42] In order to balance the total cofactor charge, the net positive charge of 0.2 au calculated to be on the methyl group in 1-Me-HN^+ was distributed evenly among all the ribose atoms of the nicotinamide–ribose fragment in $NADP^+$, i.e. the ribose atom charges in $NADP^+$ were set equal to the values of the corresponding charges in

(22) Birdsall, B.; Burgen, A. S. V.; Roberts, G. C. K. *Biochemistry* **1980**, *19*, 3723–3731.
(23) Hyde, E. I.; Birdsall, B.; Roberts, G. C. K.; Feeney, J.; Burgen, A. S. V. *Biochemistry* **1980**, *19*, 3746–3754.
(24) Blakley, R. L. In *Folates and Pterins*; Blakley, R. L., Benkovic, S. J., Eds.; Wiley: New York, 1984; Vol. 1, pp 230–231 and references therein.
(25) Andrews, J.; Fierke, C. A.; Birdsall, B.; Ostler, G.; Feeney, J.; Roberts, G. C. K.; Benkovic, S. J. *Biochemistry* **1989**, *28*, 5743–5750.
(26) Thillet, J.; Adams, J. A.; Benkovic, S. J. *Biochemistry* **1990**, *29*, 5195–5202.
(27) Appleman, J. R.; Beard, W. A.; Delcamp, T. J.; Prendergast, N. J.; Freisheim, J. H.; Blakley, R. L. *J. Biol. Chem.* **1990**, *265*, 2740–2748.
(28) Beard, W. A.; Appleman, J. R.; Delcamp, T. J.; Freisheim, J. H.; Blakley, R. L. *J. Biol. Chem.* **1989**, *264*, 9391–9399.
(29) Bolin, T. J.; Filman, D. J.; Matthews, D. A.; Hamlin, R. C.; Kraut, J. *J. Biol. Chem.* **1982**, *257*, 13650–13662.
(30) Filman, D. J.; Bolin, T. J.; Matthews, D. A.; Kraut, J. *J. Biol. Chem.* **1982**, *257*, 13663–13672.
(31) Howell, E. E.; Villafranca, J. E.; Warren, M. S.; Oatley, S. J.; Kraut, J. *Science* **1986**, *231*, 1123–1128.
(32) Bystroff, C.; Oatley, S. J.; Kraut, J. *Biochemistry* **1990**, *29*, 3263–3277.
(33) Zwanzig, R. W. *J. Chem. Phys.* **1954**, *22*, 1420–1426.
(34) Jorgensen, W. L.; Ravimohan, C. *J. Chem. Phys.* **1985**, *83*, 3050–3054.
(35) Singh, U. C.; Brown, F. K.; Bash, P. A.; Kollman, P. A. *J. Am. Chem. Soc.* **1987**, *109*, 1607–1614.
(36) Singh, U. C.; Weiner, P. K.; Caldwall, J. W.; Kollman, P. A. *AMBER (Version 3.0)*; University of California: San Francisco, CA, 1986.
(37) Weiner, S. J.; Kollman, P. A.; Nguyen, D. T.; Case, D. A. *J. Comput. Chem.* **1986**, *7*, 230–252.
(38) Weiner, S. J.; Kollman, P. A.; Case, D. A.; Singh, U. C.; Ghio, C.; Alagona, G.; Profeta, S.; Weiner, P. K. *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
(39) Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, *56*, 2257–2261.
(40) Hariharan, P. C.; Pople, J. A. *Chem. Phys. Lett.* **1972**, *16*, 217–219.
(41) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–148.
(42) Singh, U. C.; Kollman, P. A. *QUEST (Version 1.1)*; University of California: San Francisco, CA, 1986.

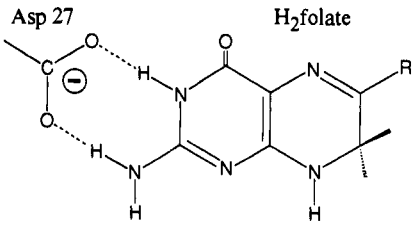

Figure 3. H-bonding between Asp-27 and the pterin ring part of dihydrofolate. R is the methylene(*p*-aminobenzoyl)-L-glutamate side chain.

NADPH plus 0.0118 au. The 6-31G-optimized bond lengths and angles were input to PREP, and the remaining bond, angle, and dihedral energy parameters were taken from the interpolated values generated by PREP, except that the barrier height for rotation of the carboxamide side chain was estimated from previous ab initio 3-21G calculations:[43] 3-21G and 6-31G basis sets yield similar ring and carboxamide side chain geometries for nicotinamide. Standard nonbonded van der Waals (vdW) parameters were used for the nicotinamide moiety.[37,38] Only the nicotinamide–ribose fragment of the cofactor was taken as the perturbed group in the FEP calculations. In a manner similar to the nicotinamide calculations described above, the substrate molecule, DHF, was constructed from 6-methyl-7,8-dihydropterin, *p*-aminobenzamide, and glutarate fragments by using 6-31G optimization and 6-31G* single-point calculations. The values of all derived parameters are given in the supplementary material. A standard united atom force field model[38] was used for the *E. coli* DHFR molecule, except for the active-site residue Phe-31, which was modeled by using an all-atom force field.[37] The all-atom model was used for Phe-31 in order to improve the description of possible ring-stacking interactions between the residue side chain and the dihydropterin ring of the substrate. All water molecules were assigned the TIP3P force field parameters of Jorgensen et al.[44]

Initial Coordinates. The simulations require a set of initial coordinates for the enzyme structure. In the present study, the initial coordinates for the wild-type enzyme and crystallographic water were taken from an X-ray crystal structure determination[31] of the binary complex of *E. coli* DHFR with the inhibitor methotrexate (MTX). This structure is more refined than the earlier *E. coli* structure of Bolin et al.[29,30] Since the calculations were undertaken, structures for *E. coli* DHFR·$NADP^+$ and *E. coli* DHFR·folate·$NADP^+$ complexes have been published, both at lower resolution than the *E. coli* DHFR·MTX structure and with the DHFR·$NADP^+$ structure being incomplete due to disorder of some sections of the protein and cofactor.[32] There are no X-ray structures with bound NADPH or DHF for the *E. coli* enzyme. As DHFR catalyzes the hydride ion transfer reaction between NADPH and DHF to give products $NADP^+$ and tetrahydrofolate, a crystal structure of the reactive ternary complex is unlikely to be obtainable.

The MTX coordinates in the initial structure were deleted, and the NADP and DHF coordinates were obtained from a least-squares superposition of the X-ray DHFR coordinates and the DHFR coordinates of a ternary structure that had previously been refined by molecular dynamics.[45] Any crystallographic water molecules within a 3-Å radius of any atom in the substrate and cofactor were then removed from the active site. Charged amino acid residues not involved in the formation of salt bridges were neutralized by locating Na^+ and Cl^- counterions at distances of about 3 Å from charged amino acid side chains. Both NMR[46] and crystallographic[32] studies indicate salt-bridge formation between His-45 and the nicotinamide 5′-phosphate of NADP, and crystallographic studies[32] have also established the existence of a salt bridge between Arg-44 and the adenine 2′-phosphate. Consequently, protonated Arg-44 and His-45 residues were neutralized by the phosphate groups of NADP, as indicated schematically in Figure 2. Although experiment indicates a dianion in DHFR·NADP complexes,[47] the 2′-phosphate was left singly ionized to avoid having to fit an additional counterion into the active site. Residue Asp-27 is known from X-ray crystal structure analysis to form a highly specific H-bond interaction with folate in the $NADP^+$ ternary complex[32] and also in the binary complex with MTX.[29,31] Similar interactions are conserved in the chicken liver DHFR·$NADP^+$·biopterin complex,[48] the *L. casei* DHFR·NADPH·MTX inhibitor complex,[29] and the human DHFR·folate and DHFR·MTX binary complexes.[49,50] It was assumed that DHF in the reduced and oxidized NADP complexes of *E. coli* DHFR would bind in a similar way to the pterin ring in the folate or biopterin complexes, as indicated by the interaction shown in Figure 3. Note also that Asp-27 is probably ionized, although in the calculations it was not neutralized by adding a counterion since the coordination site is occupied by DHF. Asp-27 was also left unneutralized in the binary complex in order to maintain the charge balance between the binary and ternary complex simulations. The resulting model for the DHFR molecule without ligands, but with its associated Na^+ and Cl^- counterions, has a net charge of +1 (His-45 + Arg-44 + Asp-27) localized in the active-site region. The total charge on both the binary and ternary complexes with the reduced form of NADP bound is thus −2.

Table I. Systems for Which FEP Calculations Were Performed

system	solute	charge[a]	solvent[b]	atoms[b]
a[c]	1-MeH_2Nic → 1-MeHNic$^+$	0	875	2646
b[d]	1-MeH_2Nic → 1-MeHNic$^+$	0	1074	3243
c[c]	folded NADPH → $NADP^+$	−3	1368	4180
d[d]	folded NADPH·Na^+ → $NADP^+$·Na^+	−2	2063	6266
e[d]	unfolded NADPH·Na^+ → $NADP^+$·Na^+	−2	2051	6230
f[d]	NADPH·DHFR → $NADP^+$·DHFR	−2	4340	14734
g[d]	NADPH·DHFR·DHF → $NADP^+$·DHFR·DHF	−2	4343	14691

[a] The solute charge (cofactor + counterions + DHFR) is given for the initial state, i.e. the reduced cofactor analogue or NADPH state. [b] The total number of solvent molecules and the total number of atoms including those of the solute (united atom model for DHFR) are also indicated. [c] Periodic boundary conditions (constant temperature and pressure). [d] Spherical-cap boundary conditions (constant temperature).

Solvated Structures. The solutes a–g, for which results of the FEP calculations will be presented, are summarized in Table I. Crystal structure analysis[29,30,32,48] reveals that NADP binds to DHFR in an extended conformation. However, since NMR experiments suggest that free NADP in solution exists in both folded and unfolded conformations,[51] calculations were carried out for both types of initial conformations as indicated. In simulations d and e, it was necessary to include a positively charged counterion (Na^+) in order that the NADP solutions have the same net charge as the binary and ternary complex solutions, i.e. in order that charge be conserved in the physical binding processes (Figure 1). The placement of this charge, although somewhat arbitrary, is nevertheless very important. It should be coordinated close to NADP to be analogous to the positively charged active site of the DHFR model. This would help to ensure that no artificial electrostatic perturbations are introduced in going from unbound to enzyme-bound NADP. Consequently, the counterion was initially positioned roughly where the His-45 would reside in the complex with DHFR (Figure 2). A structure in which a counterion was placed near the analogous Arg-44 position in Figure 2 was eliminated on energetic grounds, since in the folded conformation the repulsive interaction between Na^+ and $NADP^+$ was found to be exceedingly large and led to a dissolution of the ion from $NADP^+$ in the subsequent dynamics. The chosen placement of the counterion gives rise to similar charge environments for the nicotinamide moiety in bound and unbound solution states: Na^+ mimics His-45, while the negatively charged 2′-phosphate group that is left unneutralized effectively mimics the charge on Asp-27.

Each of the solutes a–g were solvated by using the EDIT module of AMBER by initially placing them at the center of a box constructed from cubes each containing 216 water molecules. For those systems where periodic boundary conditions were to be applied (a and c), water molecules further than a distance of 12–14 Å along the *x*, *y*, and *z* directions from any solute atom were discarded to form an approximately cubic box. For the remaining systems, b and d–g, periodic boundary conditions were not applied but a spherical shell of water was constructed with its center positioned at the center of mass of the solute atoms. The cutoff radii for the neglect of water molecules were 20 Å for b, 25 Å for d and e, and 35 Å for f and g. The initial structure for the solvated binary complex was obtained from the solvated ternary complex by deleting DHF and

(43) Cummins, P. L.; Gready, J. E. *THEOCHEM* **1989**, *183*, 161–174.
(44) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926–935.
(45) Rao, B. G.; Ramnarayan, K.; Singh, U. C. To be published.
(46) Poe, M.; Hoogsteen, K.; Matthews, D. A. *J. Biol. Chem.* **1979**, *254*, 8143–8152.
(47) Birdsall, B.; Roberts, G. C. K.; Feeney, J.; Burgen, A. S. V. *FEBS Lett.* **1977**, *80*, 313–316.
(48) Matthews, D. A.; Oatley, S. J.; Kraut, J. Unpublished results.
(49) Oefner, C.; D'Arcy, A.; Winkler, F. K. *Eur. J. Biochem.* **1988**, *174*, 377–385.
(50) Davies, J. F.; Delcamp, T. J.; Prendergast, N. J.; Ashford, V. A.; Freisheim, J. H.; Kraut, J. *Biochemistry* **1990**, *29*, 9467–9479.
(51) Oppenheimer, N. J.; Arnold, N. J.; Kaplan, N. O. *Proc. Natl. Acad. Sci. U.S.A.* **1971**, *68*, 3200–3204.

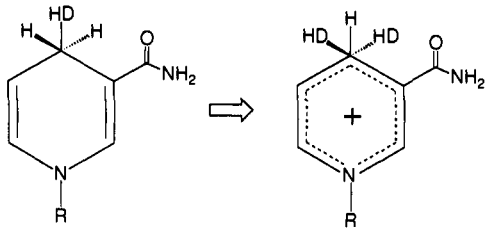

Figure 4. Schematic diagram showing the mapping of NADPH onto $NADP^+$ by a direct atom-to-atom assignment during the perturbation calculations. In the cofactor analogue simulations, a and b, R = CH_3. HD are dummy hydrogen atoms with zero partial charges and vdW energy parameters (see text).

Table II. Atomic Rms Differences (Å) between the Initial X-ray DHFR Coordinates and the DHFR Coordinates Obtained After Computations Were Performed on the Solvated Structures

	rms deviations from X-ray structure[a]	
computed structure	only C_α atoms	all DHFR atoms
energy-minimized binary	0.43	0.60
energy-minimized ternary	0.37	0.53
molecular dynamics binary	1.46 (2.84)	1.89 (3.58)
molecular dynamics ternary	1.21 (2.54)	1.65 (3.04)

[a] NADPH complexes. Values in parentheses refer to $NADP^+$ complexes.

resolvating the active site. The total number of water molecules and the total number of atoms including the solute (note that a united-atom model is used for most of the enzyme) are also given in Table I.

Free Energy Perturbation Calculations. Periodic boundary conditions were applied to a and c and the calculations carried out at a constant temperature of 300 K and at 1 atm pressure. An 8-Å cutoff radius was used for the nonbonded interactions. In b, an 8-Å cutoff radius was also used for the nonbonded interactions. In d–g, a 9-Å nonbonded cutoff radius was chosen for all nonbonded interactions except those involved with the NADP cofactor, which was allowed to interact with all atoms in the system. In all calculations, a constant dielectric factor of $\epsilon = 1$ was used. All systems were first energy-minimized by using the BORN module and then equilibrated by using the GIBBS module before the FEP calculations were started. The minimizations were done in several stages by using a combination of steepest descent and conjugate gradient methods as follows. For a–e, solvent and counterions were minimized first, while the remaining solute atoms were kept fixed, for 800–1000 cycles, followed by a further 800 cycles in which the whole system was minimized. For f and g, the solvent and counterions were minimized for 2000 cycles, followed by 2000 cycles for the system consisting of solvent, counterions, and ligand(s), followed by another 2000 cycles for the complete system including DHFR. In the final stage, all systems were minimized for 100 cycles with all bond lengths kept fixed at their equilibrium values by using the SHAKE algorithm.[52] The systems were then equilibrated with MD using a time step of 0.002 ps for a period of 6.5 ps for a–e and 8.5 ps for f and g. In the initial equilibration stage and the subsequent FEP calculations, the SHAKE algorithm[52] was used to constrain bond lengths at their equilibrium values.

The equilibrated structures were input to the GIBBS module, which performs the FEP calculations. The total perturbed group is shown in Figure 2. The mutation from the reduced to oxidized cofactor was achieved by a direct atom-to-atom assignment[53] as shown in Figure 4. In the simulations, the dummy hydrogen atoms HD have zero partial charge and vdW parameters but have the same bonded parameters as H. Starting from the initial reduced state ($\lambda = 1$), HD is mutated to H, and the two H's are mutated to two HD's. The perturbation was carried out by using the window method with a coupling parameter (λ) divided into 101 windows. At each window, a–e were equilibrated for 0.6 ps and data for the free energy calculations collected for a further 0.6 ps, while the equilibration and data collection times for f and g were 0.5 ps each. A time step of 0.002 ps was used in all calculations.

Results and Discussion

Comparison of Initial X-ray and Computed DHFR Coordinates. Atomic rms differences between the initial X-ray and computed

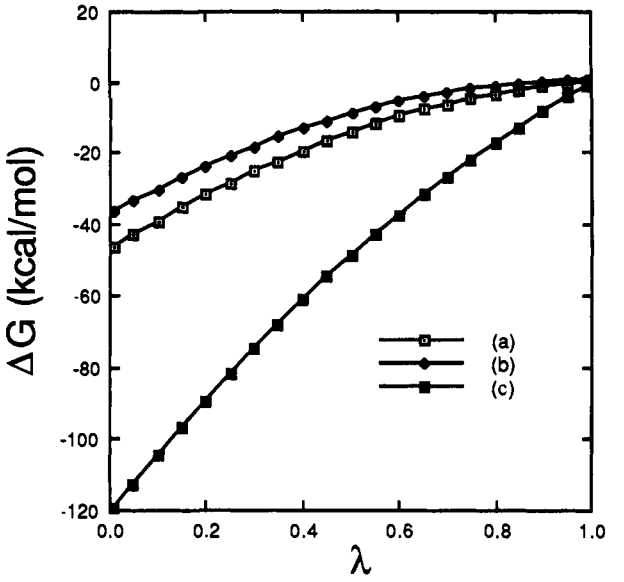

Figure 5. Free energy changes calculated as a function of the coupling parameter λ for systems a–c. (See Table I: a and b are for cofactor analogues and c is for NADP; note the initial state is at $\lambda = 1$ and corresponds to the reduced cofactor.)

structures are given in Table II. The essential features of the crystallographic data are faithfully reproduced in the energy-minimized structures, since the rms deviations in the α-carbon (C_α) backbone atoms are less than 0.5 Å for both the binary and ternary complexes. Note that if all protein atoms are included in the rms fit, the deviation increases by only 0.1–0.2 Å. The binary and ternary energy-minimized coordinates are practically identical, differing by less than 0.1 Å. Note that the NADPH and DHF coordinates were initially energy-minimized while the DHFR coordinates were kept fixed at the X-ray values. This allows relaxation of overlap interactions prior to full minimization, thereby ensuring that the essential features of the X-ray structure are preserved. The MD structures naturally show significantly larger deviations from the initial X-ray coordinates. Due to the prohibitive computational expense of including bulk water in large protein simulations, there are few literature results for comparison. However, a recently reported MD calculation[54] on a somewhat smaller protein (bovine pancreatic trypsin inhibitor) in solution gives an rms deviation in the range 1–2 Å, which is consistent with our simulations for the reduced DHFR complexes, while our oxidized-complex deviations are almost twice as large. The explanation for the increased deviation in the oxidized complex may be, in part, connected with increased hydrophilic interactions; i.e. the charge on $NADP^+$ provides an additional driving force for the solvent to interact with the protein molecule. It should also be noted that the oxidized-complex structure was obtained after a simulation time of 100 ps, compared with only 8.5 ps for the reduced complex.

Computed Free Energy Changes. The changes in free energies as a function of the coupling parameter (λ) for systems a, b, and c are given in Figure 5: the total free energy change, ΔG_{AB}, is given by $\Delta G(\lambda = 0)$. Since the ionic charge varies linearly with λ, the curved, monotonically decreasing plots in Figure 5 are consistent with the Born model for the free energy of solvation of a spherical ion:[55]

$$\Delta G = -(\epsilon_r - 1)(Ze)^2/(8\pi\epsilon_0\epsilon_r R) \quad (6)$$

where ϵ_r is the dielectric constant of the solvent relative to the vacuum value ϵ_0, Ze the charge, and R a characteristic ionic radius. ΔG for c is, as expected, much more negative than for either of

(52) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(53) Rao, B. G.; Singh, U. C. *J. Am. Chem. Soc.* **1989**, *111*, 3125–3133.
(54) Clementi, E.; Corongiu, G.; Lie, G. C.; Niesar, U.; Procacci, P. In *Modern Techniques in Computational Chemistry*; Clementi, E., Ed.; ESCOM: Leiden, The Netherlands, 1989; pp 409–414.
(55) Born, M. *Z. Phys.* **1920**, *1*, 45.

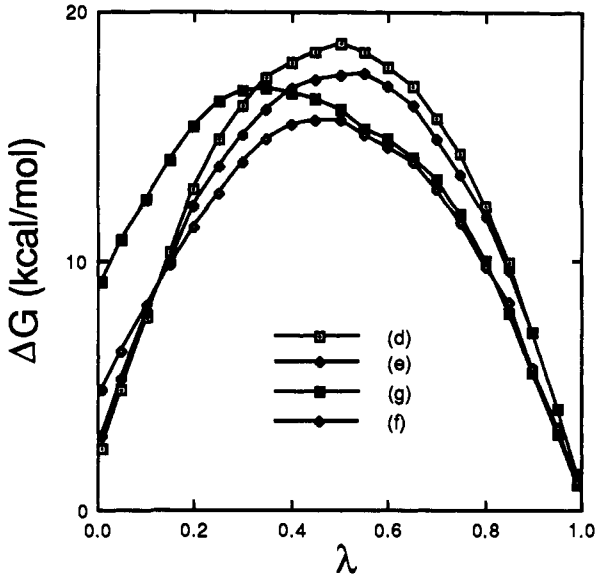

Figure 6. Free energy changes calculated as a function of the coupling parameter (λ) for systems d–g. (See Table I: d and e are for free NADP, f is for the binary complex, and g is for the ternary complex; note the initial state is at λ = 1 and corresponds to the reduced cofactor.)

the cofactor analogue simulations a and b, due to the favorable electrostatic interactions between the negatively charged 2′- and 5′-phosphate groups and the positively charged nicotinamide moiety in NADP. By comparison of the results for a and b, it is evident that the perturbation (b) carried out in a spherical shell of solvent at constant temperature gives rise to a more positive ΔG than that carried out under periodic boundary conditions at a constant temperature and pressure (a). In fact, plot b is actually slightly positive for $\lambda > 0.9$. Clearly, this observation cannot be rationalized in terms of the simple Born formula, which always predicts a negative ΔG, although the bulk of the difference between a and b (10 kcal/mol) must arise from differences in the dielectric behavior of the solvent due to pressure variations when spherical boundary conditions are applied. The important consequence of this difference for the scheme proposed in Figure 1 is that the same boundary conditions should be applied in all simulations, particularly where solvation effects are expected to dominate, i.e. when a neutral or nonpolar moiety is mutated to a charged or highly polar one.

The plots of $\Delta G(\lambda)$ for systems d–g are given in Figure 6. As can be seen, $\Delta G(\lambda)$ describes a parabolic-like path with a maximum at $\lambda \sim 0.5$. The departure from a monotonically decreasing $\Delta G(\lambda)$ for d and e is due partly to the spherical-shell boundary condition, as we have seen in simulation b, but must be due largely to the repulsive interaction between the positive nicotinamide moiety and the Na^+ counterion. Given the similarity among the plots of $\Delta G(\lambda)$ in Figure 6, the net positive charge on the enzyme in the active site appears to take part in an analogous repulsive interaction with $NADP^+$ in the binary (f) and ternary (g) complex simulations, thereby making the total free energy change quite small. The simulation free energies in Table III predict that NADPH has a higher binding affinity for DHFR than $NADP^+$ in both the binary and the ternary complex. Note there is some variation (<0.5 kcal/mol) in the calculated ΔG's, depending on whether the NADP solution state is taken to be folded or unfolded. The calculations predict NADPH to bind 10^4–10^5 times more strongly than $NADP^+$ in the ternary complex with DHF, compared with a factor of at most 10^2 in the binary complex.

In order to gain further insight into the origin of the computed differences between NADPH and $NADP^+$ binding to DHFR in the binary and ternary complexes, we have partitioned the static energies of the initially equilibrated (NADPH) and final perturbed ($NADP^+$) structures into interactions between various parts of the system. In Table IV, we give the energy differences between the oxidized ($NADP^+$) and reduced (NADPH) states for the interaction between the nicotinamide–ribose fragment (N), i.e.

Table III. Differences in the Free Energies of Binding of NADPH and $NADP^+$ in the Binary and Ternary Complexes

method[a]	binary complex	ternary complex
calcd—folded	2.31 ± 0.63[b]	6.64 ± 0.55[b]
calcd—unfolded	1.86 ± 0.79[b]	6.19 ± 0.71[b]
expt ($\Delta G_2 - \Delta G_1$)	2.9 ± 0.4[c]	*k, l*
	2.54 ± 0.21[d]	
	2.84 ± 0.34[e]	
	2.18[f]	
	1.51[g]	
	2.1 ± 0.1[h]	
	i, j	

[a] The calculated free energy values (kcal/mol) are $\Delta G_4 - \Delta G_3$, and the experimental values are $\Delta G_2 - \Delta G_1$ (Figure 1). The calculated results are given for ΔG_3 obtained from both the folded and unfolded initial NADPH solute states. [b] Mean and standard deviation of forward and backward perturbation (see text). [c] $K_{NADPH} = 0.6 \pm 0.1$ μM and 80 ± 40 μM from Poe et al.[13] (Table I (equilib)). [d] $K_{NADPH} = 0.33 \pm 0.06$ μM and $K_{NADP^+} = 24 \pm 4$ μM from Fierke et al.[15] (Table III (equilib)); $K_2/K_1 = 72.7 \pm \sim 20$. [e] $K_{NADPH} = 0.18 \pm 0.04$ μM and $K_{NADP^+} = 22 \pm 7$ μM from Fierke et al.[15] (Table III (kinet)); $K_2/K_1 = 122.2 \pm \sim 60$. [f] $K_{NADPH} = 0.03$ μM and $K_{NADP^+} = 1.2$ μM from Penner and Freiden[18] (Table II (equilib)); $K_2/K_1 = 40.0$. [g] $K_{NADPH} = 0.18$ μM and $K_{NADP^+} = 2.3$ μM from Penner and Freiden[18] (Table II (kinet)); $K_2/K_1 = 12.8$. [h] $K_{NADPH} = 0.51 \pm 0.01$ μM and $K_{NADP^+} = 18 \pm 3$ μM from Stone and Morrison[16] (Table V (equilib)); $K_2/K_1 = 35.3 \pm 7$. [i] Other values for K_{NADPH} only: 0.15 ± 0.02 μM from Appleman et al.[19] (Table I (equilib)) and 0.169 ± 0.02 μM from Appleman et al.[19] (Table I (kinet for E_t)). [j] Other values for K_{NADP^+} only: 39 ± 7 μM from Stone et al.[17] (Table III (equilib)). [k] $K_{NADP^+} = 11.6$ μM from Fierke et al.[15] (Table I (kinet; pH 6.0)). [l] $K_{NADP^+} = 17 \pm 1$ μM from Stone et al.[17] (Table III (equilib)).

Table IV. Results of Energy Partitioning: Energy Differences (kcal/mol) between the Oxidized and Reduced Systems, $\Delta E = E_{NADP^+} - E_{NADPH}$ for Simulations d–g[a]

interacting pair[b]	ΔE(d)	ΔE(e)	ΔE(f)	ΔE(g)
N–C	−45.6	−49.9	−34.2	−36.3
N–R^+	+35.9	+35.7	+49.4	+28.3
N–(C + R^+)	−9.7	−14.2	+15.2	−8.0
N–W			−51.4	−21.9
C–C	−40.9	−58.3	−60.3	−42.6
C–R^+	−12.9	+33.0	+123.6	+16.0
C–(C + R^+)	−53.8	−25.3	+63.3	−26.6
C–W			−73.5	−30.1

[a] See Table I. [b] N = nicotinamide–ribose fragment of NADP (i.e. the fragment perturbed in the FEP calculations); C = complete NADP cofactor, including fragment N; $R^+ = Na^+$ for d and e, DHFR + counterions for f, DHFR + counterions + DHF for g; W = active-site water (molecules less than 4 Å from atoms in N or C).

the perturbed fragment, and the systems consisting of the complete NADP cofactor (C), the active-site water molecules (W), and the positively charged group (R^+), which consists of Na^+ in systems d and e, DHFR + counterions in the binary complex (f), and DHFR + counterions + DHF in the ternary complex (g). Also included are the energy differences for the analogous set of interactions involving the complete cofactor (C). Note that a negative energy difference indicates that $NADP^+$ binds more strongly than NADPH, while a positive difference indicates the reverse. As may be seen, the overall energy differences (i.e. those between N or C and C + R^+) are of similar sign and magnitude for the two free-cofactor systems, (d) and (e), and the ternary complex (g). However, there are large differences between the binary and ternary complexes which stem predominantly from the interactions with DHFR (R^+). Overall, the nicotinamide–ribose fragment (N) or $NADP^+$ binds more strongly to DHFR when DHF is bound, but less when DHF is absent: this is also true if the complete cofactor (C) interactions are considered. The decrease in stabilizing interactions between $NADP^+$ and DHFR in the binary complex appears to be partly compensated by more favorable interactions with neighboring active-site waters. These results seem to suggest that oxidation to $NADP^+$ brings about conformational changes in DHFR (which we will discuss in some

detail later when comparing active-site structures) that allow water molecules a greater role in stabilizing the positive charge on nicotinamide, thus accounting for the more negative ΔG in the binary complex simulation.

Comparison of Simulation Binding Energies with Experiment. The kinetics and thermodynamics of ligand binding to DHFR have been extensively studied for enzymes from both bacterial, including *E. coli*, and vertebrate sources in binary and ternary complexes with oxidized or reduced cofactor and with substrates or inhibitors.[12-28] A comparison of relative binding energies ($\Delta\Delta G$'s) for the *E. coli* NADPH and $NADP^+$ binary complexes obtained from our simulations and from experiment are shown in Table III. These results are in good agreement with experimental values in the range 1.5–2.8 kcal/mol, i.e. a ratio of binding constants K_{NADPH}/K_{NADP^+} of $\sim 10^2$. Note from Table III that the absolute experimental values of K_{NADPH} and K_{NADP^+} differ considerably but the relative values are similar. While some of these differences may be due to sequence variations in wild-type enzymes used by different workers, it is also clear that experimental conditions such as ionic strength affect the values very considerably and therefore it is essential to compare K_d's obtained only under the same conditions. This is precisely why the total charge seen by the nicotinamide moiety has to be kept at the same constant value in all simulations. Note also that the absolute values of the dissociation constants determined by the kinetic method (i.e. as the ratio of the dissociation to association rate constants k_{off}/k_{on}) differ by about a factor of 2 from those determined by equilibrium methods because the apoenzyme exists as an approximately 50:50 mixture of conformers, only one of which binds ligands tightly.[15,19,56-58] Also shown in the footnotes in Table III are experimental dissociation constants for dissociation of $NADP^+$ from the inactive ternary complex.

While such experimental study of binding in binary and nonreactive substrate ternary complexes is straightforward, direct study of the kinetics or thermodynamics of binding in the active ternary complex (i.e. DHFR·NADPH·DHF) is not possible nor is X-ray structural information feasible. In devising kinetic schemes for simulating the full time course of the reaction for *E. coli*, *L. casei*, human, and mouse enzymes, several authors[15,18,25-27] have developed estimates of the dissociation and association rate constants for the active complex. The usual procedure was either direct measurement of an association rate constant for formation of a reactive complex using transient-state kinetics or direct measurement for unreactive or less reactive model complexes such as DHFR·NADPH·folate or DHFR·TNADPH·DHF (where TNADPH is the thionicotinamide derivative), complemented by estimates for dissociation rate constants using these model ternary complexes or self-consistent guesses. From the values for *E. coli* DHFR used by Fierke et al.[15] (Scheme IV), a dissociation constant of 0.34 μM can be calculated for NADPH dissociating from DHFR·NADPH·DHF, with values of the same magnitude deduced by other workers for the *L. casei*, human, and mouse enzymes.[25-27]

Using the experimental data of Fierke et al.[15] for the inactive DHFR·$NADP^+$·DHF ternary complex shown in Table III (11.6 μM), it is possible to estimate a value for the relative energy for binding of NADP in the ternary complexes of 2.1 kcal/mol. In their full kinetic analysis, Penner and Frieden[18] used assumed and other derived values for rate constants (Table I) for the inactive and active ternary complexes from which dissociation constants for NADP(H) from the inactive and active ternary complexes of 1.25 and 0.12 μM can be calculated, giving an estimate for the relative binding of 1.4 kcal/mol. Clearly, there is a major discrepancy between these values and those obtained from our simulations: while the simulations predict NADPH to bind 10^4–10^5 times more strongly than $NADP^+$ in the ternary complex, experiment predicts a factor of only about 10–10^2, i.e. slightly less than in the binary complexes. Note, however, that the theoretical factor is consistent with results for differential binding of $NADP^+$ and NADPH in *L. casei* DHFR ternary complexes with methotrexate and trimethoprim inhibitors.[20-24] As the theoretical simulations provide only a relative binding energy, it is not possible to say definitely whether the discrepancy arises from the active or inactive complex or from both. However, apart from questions, noted above, raised by the use of model complexes in determining the experimental rate constants for the active complex, several lines of evidence suggest other possible sources of error. Variable-viscosity studies by Stone and Morrison[59] aimed at determining the "stickiness" of substrates in the active ternary complex, i.e. the rate at which a complex reacts to give products relative to the rate at which substrate dissociates, found that NADPH *is* sticky in the active complex, in disagreement with the treatment of Fierke et al.[15] Also, Beard et al.[28] found an obligatory fast isomerization of the reactive (or possibly product) complex not seen by Fierke et al.[15] Note that Fierke et al.[15] investigated the possibility of isomerization of the inactive ternary complex (DHFR·$NADP^+$·DHF) and other inactive folate ternary complexes using a combination of transient-state kinetic and ligand competition experiments but found none. The implications of these studies is that the estimated rate constant for NADPH dissociating from the active complex is probably too low and that the complex observed by Fierke et al.[15] is a nonequilibrium one. This would imply a much lower value for the dissociation rate constant and consequently a higher factor for the differential in binding in the active and inactive complexes in the same direction as predicted by theory.

Table V. Numbers of Active-Site Water Molecules[a]

	binary complex		ternary complex	
radius, Å	NADPH	$NADP^+$	NADPH	$NADP^+$
4	6	21	6	10
8	66	79	49	58
12	209	205	189	203
16	551	581	502	554

[a] Numbers are for total water molecules within a given radius of the atoms in the nicotinamide–ribose fragment in the NADPH/$NADP^+$ binary and ternary complexes (see Figures 7–10).

The other factor to consider is the molecular form of the complex with respect to protonation of the enzyme and substrate: in the simulations, these have been fixed to one configuration, as indicated in Figure 3. The mechanism of the reaction requires the addition of both a proton and a hydride ion (from NADPH) to dihydrofolate, and it is generally accepted that protonation precedes hydride ion transfer and activates the substrate.[60] Consequently, from the experimental results, it is not clear at which stage the proton transfer is occurring, whether concurrently with initial association of the active ternary complex, during the isomerization stage, or after the isomerization but before the catalytic event. Thus it is not possible to say whether the observed complexes are in the DHFR·NADPH·DHF or DHFR·NADPH·H^+–DHF forms or whether the enzyme is protonated or deprotonated on Asp-27. This uncertainty also casts doubt on the validity of the use of model folate ternary complexes for estimating the dissociation rate constant of the active complex, as available evidence suggests folate or tetrahydrofolate in such complexes is unprotonated under the experimental conditions and there is no evidence for isomerizations of these complexes.

Comparison of Active-Site Structures. We have attempted to rationalize the calculated free energy changes in terms of structural differences between the complexes. Figures 7 and 8 are detailed stereo images of the active-site regions of the binary complex with NADPH and $NADP^+$, respectively, while Figures 9 and 10 depict the analogous structures for the ternary complexes. The NADPH complexes are those obtained after the initial equilibration phase (8.5 ps of MD), and the $NADP^+$ complexes those obtained after

(56) Cayley, P. J.; Dunn, S. M. J.; King, R. W. *Biochemistry* **1981**, *20*, 874–879.
(57) Penner, M. H.; Frieden, C. *J. Biol. Chem.* **1985**, *260*, 5366–5369.
(58) Dunn, S. M. J.; Lanigan, T. M.; Howell, E. E. *Biochemistry* **1990**, *29*, 8569–8576.
(59) Stone, S. R.; Morrison, J. F. *Biochemistry* **1988**, *27*, 5493–5499.
(60) Gready, J. E. *Biochemistry* **1985**, *24*, 4761–4766.

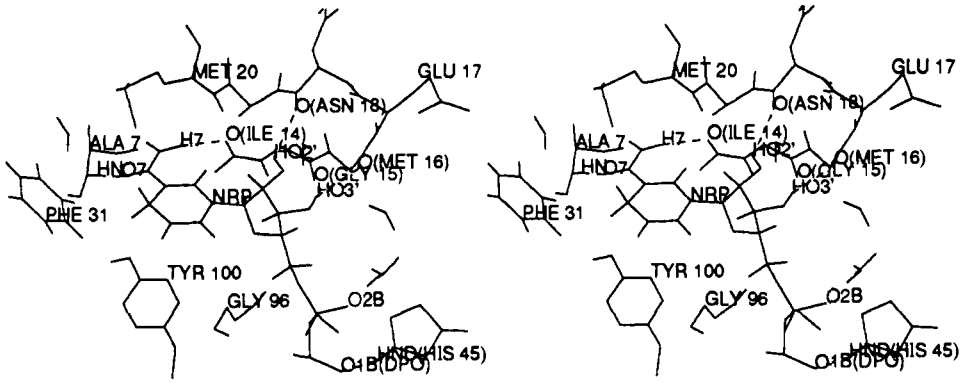

Figure 7. Stereoview of the active site in the DHFR·NADPH binary complex obtained by MD. NRP = reduced nicotinamide–ribose and DPO = 5′-phosphate fragments of NADP. The relevant distances (Å) are as follows: HN(ALA 7)–O7(NRP) = 2.15; O(ILE 14)–H7(NRP) = 2.32; O(MET 16)–HO3′(NRP) = 2.35; O(ASN 18)–HO2′(NRP) = 2.42; HND(HIS 45)–O1B(DPO) = 1.76.

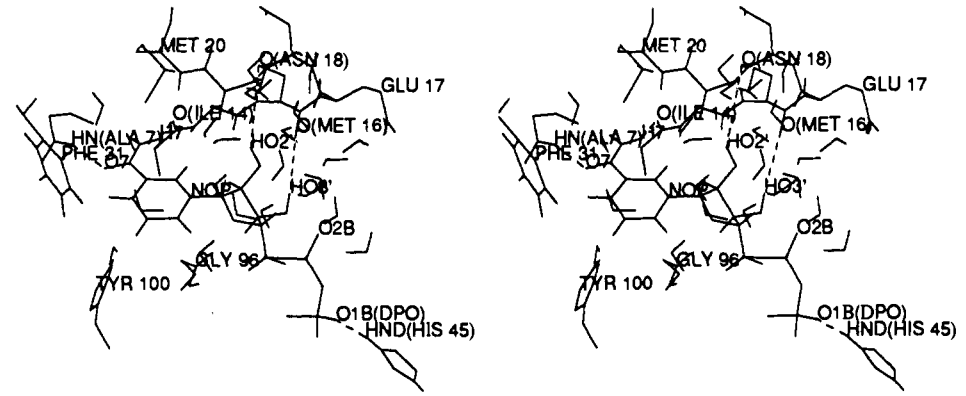

Figure 8. Stereoview of the active site in the DHFR·NADP$^+$ binary complex obtained by MD. NOP = oxidized nicotinamide–ribose and DPO = 5′-phosphate fragments of NADP. The relevant distances (Å) are as follows: HN(ALA 7)–O7(NOP) = 2.08; O(ILE 14)–H7(NOP) = 2.41; O(MET 16)–HO3′(NOP) = 7.20; O(ASN 18)–HO2′(NOP) = 8.44; HND(HIS 45)–O1B(DPO) = 1.69.

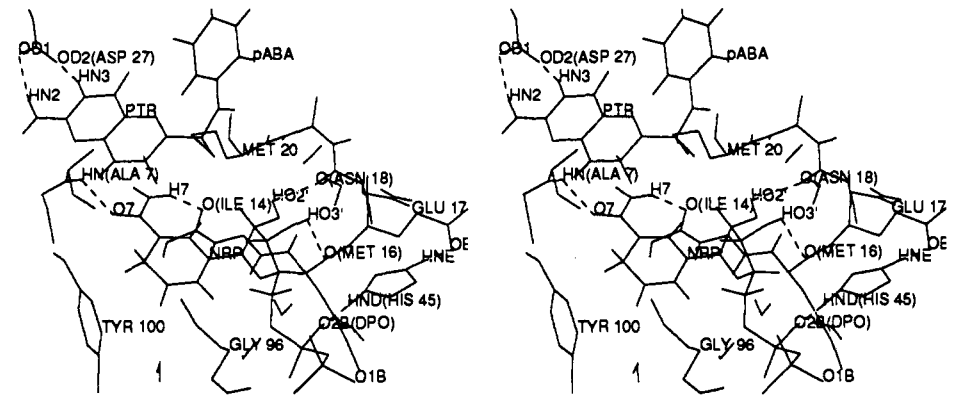

Figure 9. Stereoview of the active site in the DHFR·NADPH·DHF ternary complex obtained by MD. NRP = reduced nicotinamide–ribose and DPO = 5′-phosphate fragments of NADP; PTR = pteridine and pABA = (*p*-aminobenzoyl)glutamate fragments of dihydrofolate. The relevant distances (Å) are as follows: NH(ALA 7)–O7(NRP) = 2.10; O(ILE 14)–H7(NRP) = 2.31; O(MET 16)–HO3′(NRP) = 4.25; O(ASN 18)–HO2′(NRP) = 5.27; HND(HIS 45)–O2B(DPO) = 1.63; HNE(HIS 45)–OE(GLU 17) = 1.66; OD1(ASP 27)–HN2(PTR) = 2.18; OD2(ASP 27)–HN3(PTR) = 1.96.

completion of the perturbation calculations (over 100 ps of MD). Only solvent molecules within a 4-Å radius of the nicotinamide–ribose fragment are shown. As a further indication of the extent of solvation of the active site, the numbers of water molecules within a given radius of the nicotinamide–ribose fragment are listed in Table V.

Figures 7 and 9 show that NADPH is tightly enclosed by the active-site residues of DHFR and that relatively few nearby solvent molecules are present. However, comparing the two structures reveals several differences. In particular, residues Met-16 to Met-20 of the Met-20 loop (residues 9–24 form a loop structure that Bystroff et al.[32] call the "Met-20 loop") and also residue Phe-31, which is shown in Figure 7, make way for the binding of DHF in the ternary complex (Figure 9). Consequently, H-bonds between the nicotinamide–ribose hydroxyl groups (HO2′ and HO3′) and carbonyl oxygens of Met-16 and Asn-18 are present in the binary NADPH complex but absent in the ternary NADPH complex: in the ternary complex, the narrow space formed between ribose and the Met-20 loop is occupied by two water molecules. In the NADP$^+$ binary complex (Figure 8), the residues closest to the bulk solvent, i.e. residues in the Met-20 loop, have completely dissociated from the ribose ring and effectively "dissolved away" from the active-site region leaving an 8 Å wide channel into the nicotinamide binding site. The main consequence of this dissociation is that the nicotinamide–ribose moiety also becomes partially solvated, but the nicotinamide moiety remains essentially bound to the bulk of the enzyme as in the NADPH binary complex, since the H-bond interactions between the carboxamide side chain of nicotinamide and the Ala-7 and Ile-14 residues are preserved. The NADP$^+$ ternary complex structure

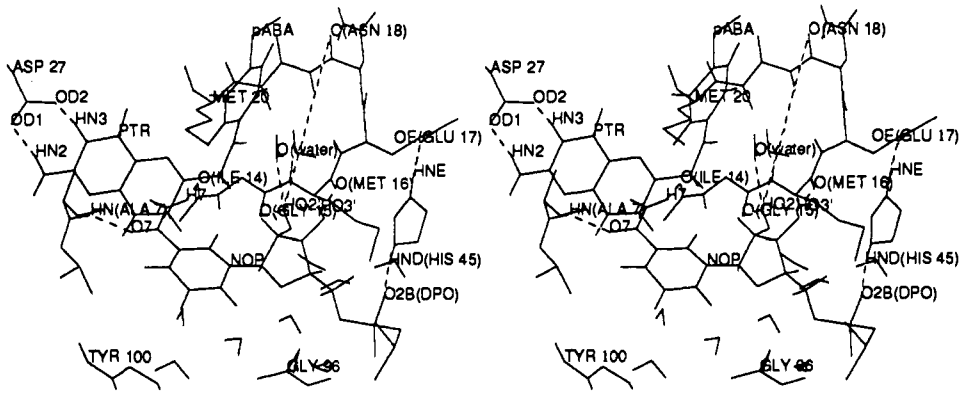

Figure 10. Stereoview of the active site in the DHFR·NADP⁺·DHF ternary complex obtained by MD. NOP = reduced nicotinamide–ribose and DPO = 5′-phosphate fragments of NADP; PTR = pteridine and pABA = (*p*-aminobenzoyl)glutamate fragments of dihydrofolate. The relevant distances (Å) are as follows: HN(ALA 7)–O7(NOP) = 1.97; O(ILE 14)–H7(NOP) = 2.10; O(GLY 15)–HO2′(NOP) = 3.01; O(MET 16)–HO3′(NOP) = 2.08; O(water)–HO2′(NOP) = 2.53; O(ASN 18)–NO2′(NOP) = 8.48; HND(HIS 45)–O2B(DPO) = 1.57; HNE(HIS 45)–OE(GLU 17) = 1.66; OD1(ASP 27)–HN2(PTR) = 2.15; OD2(ASP 27)–HN3(PTR) = 1.99.

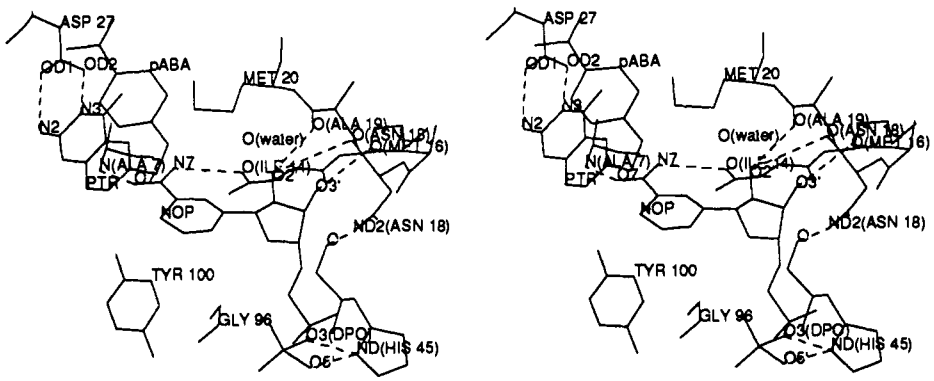

Figure 11. Stereoview of the active site in the DHFR·NADP⁺·folate ternary complex obtained from X-ray crystal structure analysis.[32] NOP = reduced nicotinamide–ribose and DPO = 5′-phosphate fragments of NADP; PTR = pteridine and pABA = (*p*-aminobenzoyl)glutamate fragments of dihydrofolate. The relevant distances (Å) are as follows: N(ALA 7)–O7(NOP) = 2.99; O(ILE 14)–N7(NOP) = 3.42; O(MET 16)–O3′(NOP) = 3.48; O(ASN 18)–O2′(NOP) = 4.07; O(ALA 19)–O2′(NOP) = 4.04; ND(HIS 45)–O3(DPO) = 2.80; ND(HIS 45)–O5′(DPO) = 2.11; O(HIS 45)–ND2(ASN 18) = 3.46; OD1(ASP 27)–N2(PTR) = 3.18; OD2(ASP 27)–N3(PTR) = 2.64.

(Figure 10) shows that there are also significant structural changes in the active-site region on going from NADPH to NADP⁺, but fewer water molecules have access to solvate the positively charged nicotinamide moiety than in the NADP⁺ binary complex. In the ternary complex, the interaction that appears to prevent complete dissociation of the loop, and hence limit solvation of the active site, is the H-bond formed between the oppositely charged Glu-17 and His-45 side chains. The His-45 side chain is in turn H-bonded to O2B of the 5′-phosphate link (DPO) in the cofactor, thus anchoring the Met-20 loop. In the case of *L. casei* DHFR ternary complexes,[29,30] our molecular modeling studies suggest that a similar pattern of interactions between Asp-16, Arg-44, and DPO may also be possible: note that His-45 is conserved as a residue with a basic side chain in all DHFRs and most dehydrogenases. In the binary complex (Figures 7 and 8), His-45 is in a position (H-bonded to O1B rather than O2B) and orientation less favorable for a second interaction with Glu-17.

The active-site structures we have obtained using MD may also be compared with analogous structures determined from a recent crystallographic study[32] of the *E. coli* DHFR·NADP⁺ binary complex and the *E. coli* ternary complex with folate. In the crystal state,[32] residues 16–20 of the Met-20 loop and the nicotinamide–ribose part of NADP⁺ are observed to be disordered in the binary complex but ordered in the ternary complex with folate. In solution, this mobility of the Met-20 loop would allow solvent water into the active site of the binary complex, an effect that has thus been well described in our MD calculations (Figure 8). As seen more clearly in Table V, for both the binary and ternary complex simulations, the number of neighboring active-site water molecules (within ~4 Å) increases on going from NADPH to NADP⁺, creating a large pocket (~20 molecules) of water in the active site of the binary complex. Thus, due largely to mobility of the Met-20 loop, the binary complex allows far more effective solvation of NADP⁺, resulting in a more negative free energy change ΔG_{AB}.

The X-ray structure of the active site in the ternary complex with folate is shown in Figure 11. Note that in the NADPH complex (Figure 9) the loop does not form any H-bonds with ribose because of intervening water, but in both the MD and X-ray NADP⁺ complexes (Figures 10 and 11), HO3′ of ribose is H-bonded to the carbonyl oxygen of Met-16. However, in the NADP⁺ structure obtained by MD (Figure 10), Asn-18 is displaced 8–9 Å from HO2′, with water molecules occupying the intervening space, whereas in the X-ray crystal structure, the top portion of the loop folds over the cofactor to make closer (but still not H-bonded) contacts with the ribose ring. The position of the loop in the X-ray structure favors an H-bond interaction between the backbone carbonyl oxygen of His-45 and the Asn-18 side chain, whereas the MD ternary structures are H-bonded through the charged side chains of His-45 and Glu-17. Note also that the His-45 side chain makes a bifurcated H-bond with the O5′ and O3 oxygens of 5′-phosphate in the X-ray structure, whereas single H-bonds are seen in the MD structures.

The superposition of X-ray and MD structures in Figure 12 illustrates more graphically the differences in the positioning of the Met-20 loop relative to the nicotinamide ring of the cofactor. The movement of the loop from a folded position over the nicotinamide site in the reduced MD binary complex (D) to an open structure in the oxidized MD binary complex (B) is clearly seen, whereas much less movement is apparent between the corresponding MD ternary complexes (E and C). A superposition of the DHFR backbone coordinates for the two X-ray structures (A and F) shows that the only significant differences between them occur in the cofactor-binding region, particularly in residues 15–23

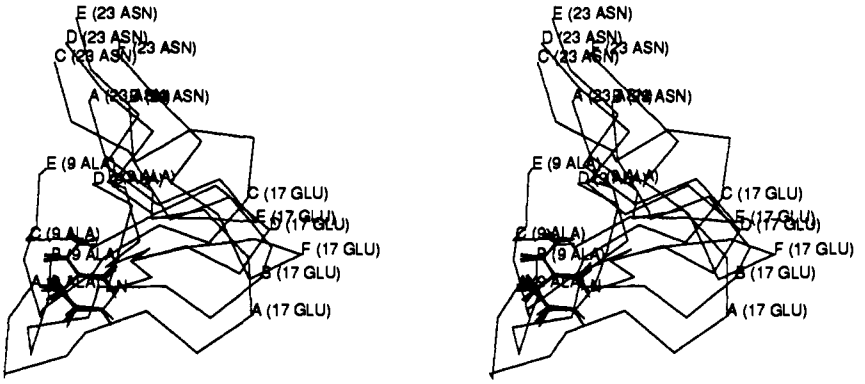

Figure 12. Stereoplot comparing C_α backbone structures of the Met-20 loop for initial X-ray (energy-minimized), MD, and DHFR·NADP$^+$·folate X-ray complexes. The orientation of the nicotinamide ring (N) is similar to that shown in Figures 7–11. A = initial (DHFR·MTX) X-ray;[31] B = DHFR·NADP$^+$ MD; C = DHFR·NADP$^+$·DHF MD; D = DHFR·NADPH MD; E = DHFR·NADPH·DHF MD; F = DHFR·NADP$^+$·folate X-ray.[32]

of the Met-20 loop: the rms deviation for residues 15–23 is 2.7 Å compared with an overall deviation of 0.90 Å. As seen in Figure 12, the essential difference is that the Met-20 loop in DHFR·NADP$^+$·folate (F) folds over nicotinamide, whereas the initial structure used in the calculations (A) is more open. Note also the similarity between D and F (both folded) and between B and A (both open). Thus, although the loop conformations seen in the binary complexes closely resemble known X-ray structural forms, the fact that the MD ternary complex structures appear quite distinct merits further scrutiny, particular in light of the possible disagreement with experimental ΔG's for the ternary complexes discussed previously.

Assessment of the Theoretical Treatment. A possible, although unlikely, source of error in the theoretical treatment may arise from unidentified errors in the force field, particularly that part which describes the interactions between ligands and protein. These types of inaccuracies could conceivably lead to artifactual structures for MD complexes and hence erroneous ΔG's. However, a much more obvious difficulty in the prediction of free energies is apparent from the variety of conformational forms exhibited by the Met-20 loop in both the MD and X-ray structures. It is possible that the MD simulations may have sampled only a subset of experimental structures which may or may not have been observed by experiment so far. For instance, the initial complex represented in the kinetic model of Fierke et al.[15] may be in a transient "unrelaxed" conformation which is not well represented by the MD simulation of the ternary complex: the results of Beard et al.[28] showing an isomerization of the active complex support this suggestion. In light of the structural comparisons of the last section, it is also possible that the choice of initial coordinates for the ternary complex has led to the MD structure being trapped in a relatively high energy conformation. The structural evidence for this can be seen in the strong H-bonds between Glu-17 and His-45 (Figures 9 and 10), which may represent energy barriers preventing the structure from relaxing to a conformation of lower free energy.

To sample the full range of possible loop conformations would require many ligand-binding simulations. Clearly, if the position of the Met-20 loop is a major determinant of the relative binding strengths of the cofactor in both binary and ternary complexes as we suggest, many such simulations may be required in order to gain an adequate impression of sampling for a free energy determination. This represents a daunting computational task, particularly for a fully solvated protein as we have modeled in the present work. X-ray structure determination may give only an indication as to the likely conformational states in solution. In the present study, we used the then available DHFR·MTX coordinates. However, the recently published DHFR·NADP$^+$·folate X-ray structure has a very different Met-20 loop conformation compared with the MTX binary complex and may represent a more appropriate analogue starting structure for the ternary complex with DHF. Unfortunately, the oxidized NADP$^+$ binary complex offers no detailed structural information due to disorder in both nicotinamide and Met-20 loop fragments, and the reduced NADPH binary complex structure has not been resolved. Note also that simulations could be carried out on relative cofactor binding in actual inhibitor ternary complexes with both *E. coli* and *L. casei* DHFR, for which a considerable body of structural and thermodynamic data is already available.[20-24,29,30,61]

It is worth noting that Fleischman and Brooks[7] have also identified possible sampling difficulties in recent free energy calculations on the binding of trimethoprim derivatives to DHFR (from chicken). However, there the problem was not associated with protein conformational states but with different binding modes for the various drug derivatives. The ability of conventional MD procedures to adequately sample configuration space where small but significant energy barriers are involved is a major concern for the computation of free energy differences from the simulations. Various schemes to improve sampling efficiency during free energy simulations have been proposed,[62,63] but their effectiveness in large-scale simulations of biochemical systems has yet to be tested.

Conclusions

The FEP method, implemented within the MD simulation scheme, has been used to predict and rationalize the difference in solution between the affinities of the binding of reduced and oxidized NADP to *E. coli* DHFR in both the binary complex and the ternary complex with dihydrofolate. The FEP results predict a relative binding constant (K_{NADPH}/K_{NADP^+}) for the binary complex of $\sim 10^2$, in good agreement with experiment.[12-19] For the ternary complex with dihydrofolate, this differential is predicted to be greater by a factor of 10^2–10^3. Although there is no accurate experimental binding information for the active ternary complex that would allow this predicted differential to be confirmed, the available kinetic and X-ray structural data suggest that the simulations for the ternary complex may have sampled only one of a possibly large number of structures. Consequently, further study is required to address all aspects of the sampling problems that may be inherent in simulations of this nature and to evaluate their influence on the computed ΔG's.

An analysis of the MD structures and interaction energies suggests the differences in computed binding strengths can be rationalized in terms of a simple solvation–desolvation mechanism. NADPH binds the enzyme more effectively than NADP$^+$ because the active-site environment cannot stabilize the positively charged nicotinamide moiety to the same extent as the bulk solvent. The degree to which NADPH binds more effectively than NADP$^+$ therefore depends to a large extent on how effectively the enzyme desolvates the active site. Consequently, compared with the case of the binary complex, NADPH binds even more effectively than

(61) Champness, J. N.; Stammers, D. K.; Beddell, C. R. *FEBS Lett.* **1986**, *199*, 61–67.
(62) Tobias, D. J.; Brooks, C. L.; Fleischman, S. H. *Chem. Phys. Lett.* **1989**, *156*, 256–260.
(63) Straatsma, T. P.; McCammon, J. A. *J. Chem. Phys.* **1989**, *90*, 3300–3304.

$NADP^+$ in the ternary complex, since the nicotinamide moiety is screened from the solvent via a loop-closing mechanism. This loop closure, and consequent screening of nicotinamide from the solvent, is maintained by strong H-bond interactions between the charged Glu-17 side chain of the Met-20 loop and His-45, which appear to result from structural changes in the enzyme induced by DHF binding.

The mere existence of a multiplicity of protein conformational states makes the straightforward application of the FEP method to the prediction of cofactor-binding energies difficult at best. Nevertheless, MD simulation may provide a powerful means for obtaining detailed structural and energetic information on protein conformations not accessible by X-ray crystallography, some of which may be analogous to conformers seen in kinetics studies or by NMR spectroscopy. For example, three different conformers for the *L. casei* DHFR·$NADP^+$·folate complex are seen by NMR spectroscopy at different pHs, but there is no structural description of the differences in binding, especially as to whether they are local to the folate binding site or involve more distant regions of the protein.[64] It also needs to be stressed that theoretical simulation potentially offers a unique window on the active ternary complex which is not available or only indirectly available from experiment. In this connection, the surprising results of Beard et al.[28] indicating very variable rate constants for isomerization of the active ternary complexes in vertebrate enzymes, despite the fact that the active sites are very similar, is interesting. It may be that these isomerizations involve conformational changes in regions further from the active site, such as in the Met-20 loop region, which show greater sequence variation, and that the structural details of such species differences could be studied by simulation. Also of interest in the current discussion of possibly trapped higher energy conformations in the simulations is an X-ray structure report[49] of two different conformational states for two highly homologous vertebrate DHFRs (mouse and human) complexed with the inhibitor methotrexate, one structure[65] being determined from crystals formed by ligand displacement in preformed crystals and the other from crystals grown from solution:[49] the former complexes are suggested to be in an "unrelaxed" conformation.[49]

Although the numerical accuracy of the computed ΔG's needs to be confirmed, a significantly important result emerges from an analysis of the MD structures, that is, the implication of the Met-20 loop as a key structural factor in determining the relative binding strengths of $NADP^+$ and NADPH in both binary and ternary DHFR complexes. According to our solvation model, the differential cofactor bindings in the binary complexes and the ternary complexes involving DHF would be of similar magnitudes if, in the ternary complex, the Met-20 loop had an open structure similar to the binary-complex loop. Conversely, the cooperativity observed for some inhibitor complexes[20-24] could be explained by loop closure. Thus, the above qualitative argument based on a solvation mechanism remains consistent with the available experimental results. The results of the present calculations provide strong support for this model with regard to the binary complex; however, more work will be required to establish its general applicability, particularly with respect to the ternary complexes.

The existence of loop structures in proteins is well-known. However, it is only now that their exact functional roles are being clearly elucidated, as discussed in a recent paper on the flexible loop in triosephosphate isomerase.[66] Moreover, an independent MD simulation study also suggests a possible ligand-induced mechanism for loop closure in triosephosphate isomerase.[67] As far as we are aware, the present work is the first report to use purely theoretical, i.e. FEP/MD simulation, methods to describe a loop function in ligand binding to DHFR. Similar theoretical studies could be applied not only to the other DHFRs whose structures are known but also to the dehydrogenases where loop structures in active sites have also been identified[68] and where, interestingly, isomerizations of the reactant and product complexes for the reversible reactions catalyzed by lactate dehydrogenase have been found.[69]

Acknowledgment. We are grateful to Dr. Richard A. Lerner for providing the computational facilities for this work at the Scripps Clinic and Research Foundation. Dr. J. E. Villafranca kindly provided the initial DHFR X-ray coordinates used in the calculations, while Professor J. Kraut made available the X-ray coordinates of DHFR·$NADP^+$·folate. Dr. B. G. Rao assisted with performing some aspects of the calculations. Support from the Australian Research Council and the National Health and Medical Research Council is gratefully acknowledged.

Supplementary Material Available: AMBER PREP files for the dihydrofolate and cofactor fragments, containing the internal coordinate parameters and partial atomic charges for dihydrofolate, NADPH, and $NADP^+$, and the AMBER PARMDAT file for DHFR, dihydrofolate, and cofactor, containing atomic masses, internal coordinate and associated energy parameters, and nonbonded van der Waals and hydrogen-bond parameters for the protein (DHFR), dihydrofolate, NADPH, and $NADP^+$ (25 pages). Ordering information is given on any current masthead page.

(64) Birdsall, B.; Feeney, J.; Tendler, S. J. B.; Hammond, S. J.; Roberts, G. C. K. *Biochemistry* **1989**, *28*, 2297–2305.
(65) Stammers, D. K.; Champness, J. N.; Beddell, C. R.; Dann, J. G.; Eliopoulis, E.; Geddes, A. J.; Ogg, D.; North, A. T. C. *FEBS Lett.* **1987**, *218*, 178–184.
(66) Pompliano, D. L.; Peyman, A.; Knowles, J. R. *Biochemistry* **1990**, *29*, 3186–3194.
(67) Joseph, D.; Petsko, G. A.; Karplus, M. *Science* **1990**, *249*, 1425–1428.
(68) Clarke, A. R.; Wigley, D. B.; Chia, W. N.; Barstow, D.; Atkinson, T.; Holbrook, J. J. *Nature* **1986**, *324*, 699–702.
(69) Holbrook, J. J.; Gutfreind, H. *FEBS Lett.* **1973**, *31*, 157–169.