# Protein Dynamics in Enzymatic Catalysis: Exploration of Dihydrofolate Reductase

**Jennifer L. Radkiewicz and Charles L. Brooks, III***

*Contribution from the Department of Molecular Biology, The Scripps Research Institute, La Jolla, California 92037*

*Received April 27, 1999*

**Abstract:** The overall motions and atomic fluctuations of DHFR are calculated using molecular dynamics simulations to explore potential links between catalysis and dynamics in this enzyme system. Calculations of 10 ns duration were performed on three ternary complexes from the DHFR catalytic cycle: DHFR/DHF/NADPH (**DH**), DHFR/THF/$NADP^+$ (**TP**), and DHFR/THF/NADPH (**TH**). The protein maintains a core structure very similar to the initial X-ray model, while several flexible loops undergo conformational changes. Comparison of the dynamics of the protein in the different complexes demonstrates that the ligands affect the behavior of the protein even though the ligands only differ by one to two hydrogen atoms. In particular, strong-coupled motions that appear in the reactive complex **DH** disappear in the product complexes, indicating that these motions may be linked to catalysis. Furthermore, mutants, which have been observed to be debilitating to particular chemical steps in catalysis, occur with high correlation in the regions of the protein structure observed in our simulations to participate in highly coupled motions. We conclude from our analysis that the mutants could be affecting catalysis by altering the protein dynamics.

## Introduction

The source of the rate enhancement of enzymes is still intensely debated, and numerous theories explaining the catalytic power of enzymes have been proposed.[1−5] Part of this debate includes the importance of protein dynamics. The most common theory states that the binding differences between the ground state and the transition state determine the rate of catalysis.[6−8] In this picture, the arrangement of the residues in the active site is the major contributing factor to the rate enhancement, and protein motions have no effect. However, other theories do propose that protein dynamics play a vital role in catalysis.[9−11] In particular, the motions of the protein can act as a "gate" such that some conformations have much lower energy barriers than others.[12] This is not unreasonable, since it is well-known that proteins sample numerous conformational states.[13−15] In addition, vibrational coupling between the protein and substrate could exist. This has been observed for several enzymatic reactions involving electron transfer.[16−18]

Experimental evidence suggests that protein dynamics may play a role in the dihydrofolate reductase catalytic pathway. *Escherichia coli* dihydrofolate reductase (DHFR) catalyzes the NADPH-dependent reduction of 7,8-dihydrofolate (DHF) to 5,6,7,8-tetrahydrofolate (THF). X-ray, NMR, kinetic, mutagenic, and molecular dynamics data indicate that regions of this 159-residue protein are highly flexible and undergo conformational changes on the order of the catalytic cycle time scale.[19] The kinetic pathway consists of five intermediate steps (Scheme 1), two of which are the Michaelis (DHFR/DHF/NADPH) and product (DHFR/THF/$NADP^+$) complexes.[20] A third ternary structure (DHFR/THF/NADPH) is formed during the catalytic cycle. The binding of NADPH increases the loss of THF, which is rate-limiting.

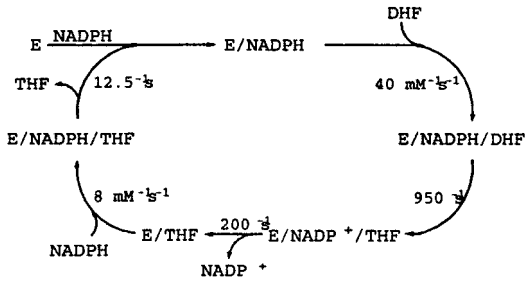
The globular structure of DHFR is comprised of a mixed eight-strand $\beta$-sheet flanked by four α-helices and several flexible loops (Figure 1).[21] The NMR order parameters for the complex of DHFR with folate demonstrate that the main regions of motion are in the M20 loop (residues 14−24), the neighboring FG loop (residues 116−125), the distant CD loop (residues 64−71) and the hinge region (the area connecting the two subdomains of DHFR).[22] In addition, short molecular dynamics simulations on *E. coli* DHFR complexed with folate and with methotrexate, an inhibitor, found that the largest fluctuations occurred in these same regions.[23]

* Coresponding author. E-mail: brooks@scripps.edu. Web: http://www.scripps.edu/brooks.

(1) Post, C. B.; W. J. Ray, J. *Biochemistry* **1995**, *34*, 15881−15885.
(2) Tsou, C. L. *Biochemistry (Moscow)* **1998**, *63*, 253−258.
(3) Cannon, W. R.; Benkovic, S. J. *J. Biol. Chem.* **1998**, *273*, 26257−26260.
(4) Warshel, A. *J. Biol. Chem.* **1998**, *273*, 27035−27038.
(5) Kurzynski, M. *Prog. Biophys. Mol. Biol.* **1998**, *1998*, 23−82.
(6) Jencks, W. P. *Adv. Enzymol.* **1975**, *43*, 219−410.
(7) Schowen, R. L. *Transition States of Biochemical Processes*; Plenum Press: New York, 1978.
(8) Kraut, J. *Science* **1988**, *242*, 533−540.
(9) Karplus, M.; McCammon, J. A. *Annu. Rev. Biochem.* **1983**, *53*, 263−300.
(10) Kurzynski, M. *Biophys. Chem.* **1997**, *1997*, 1−28.
(11) Lau, E. Y.; Bruice, T. C. *J. Am. Chem. Soc.* **1998**, *120*, 12387−12394.
(12) Cannon, W. R.; Singleton, S. F.; Benkovic, S. J. *Nat. Struct. Biol.* **1996**, *3*, 821−833.
(13) Ansari, A.; Berendzen, J.; Bowne, S. F.; Frauenfelder, H.; Iben, I. E. T.; Sauke, T. B.; Shyamsunder, E.; Young, R. D. *Proc. Natl. Acad. Sci. U.S.A.* **1985**, *82*, 5000−5004.
(14) Clarage, J. B.; Romo, T.; Andrews, B. K.; Pettitt, B. M.; Phillips, G. N. J. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *1995*, 3288−3292.
(15) Andrews, B. K.; Romo, T.; Clarage, J. B.; Pettitt, B. M.; Philips, G. N. J. *Structure* **1998**, *6*, 587−594.
(16) Goldstein, R. F.; Bialek, W. *Comments Mol. Cell. Biophys.* **1986**, *1986*, 407−438.
(17) Bialek, W.; Onuchic, J. N. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 5908−5912.
(18) Daizadeh, I.; Medvedev, E. S.; Stuchebrukhov, A. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 3703−3708.
(19) Miller, G. P.; Benkovic, S. J. *Chem. Biol.* **1998**, *5*, R105−R113.
(20) Fierke, C. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 4085−4092.
(21) Sawaya, M. R.; Kraut, J. *Biochemistry* **1997**, *36*, 586−603.

10.1021/ja9913838 CCC: $19.00 
Published on Web 12/28/1999

**Scheme 1**



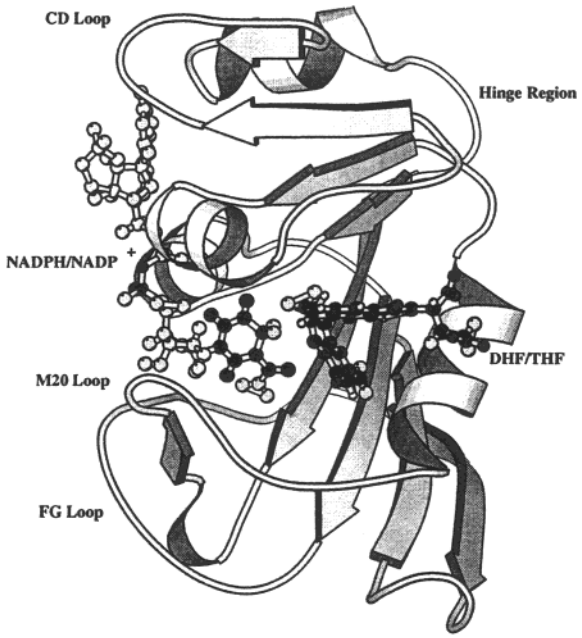




**Figure 1.** Representative structure of the ternary enzyme complex DHFR/DHF/NADPH.

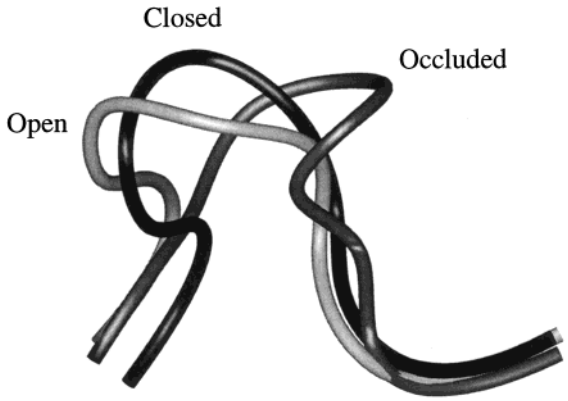



**Figure 2.** M20 loop conformations in DHFR, open, closed, and occluded.

Three distinct M20 loop conformations, occluded, closed, and open, have been observed in X-ray crystal structures and have been suggested to represent important conformations in the catalytic cycle (Figure 2).[21] The nature of bound ligands and the space group of the crystal also contribute to the M20 loop conformation. Thus, several ligand induced conformational changes could transpire during the catalytic cycle. The rate of M20 loop fluctuations, determined from NMR data on the apoenzyme, occur on the same time scale as loss of THF, suggesting that an M20 loop conformational change may limit the rate of THF dissociation.[24] In addition, this loop is not just involved in the regulation of ligand binding but also in catalysis. The replacement of four M20 loop residues (16−19) with a glycine residue result in a 500-fold decrease in the rate of the chemical step, hydride transfer.[25]

Correlation in atomic fluctuations and quasi-harmonic vibrational analysis obtained from molecular dynamics simulations on *Lactobacillus casei* DHFR with NADPH and methotrexate (MTX) indicate the existence of two subdomains.[26] These subdomains are roughly similar to those determined from X-ray structural analysis.[21] The interactions of the NADPH adenosine ring and the (*p*-aminobenzoyl) glutamate section of MTX with different regions of the same subdomain provide long-range coupling between these two groups and may account for the observed cooperativity in ligand binding.

Site-directed mutagenesis has been performed on residues in the active site, in the binding sites of substrate and cofactor, and in positions distant from the active site. The study of these mutants has provided two interesting results. First, residues far from the active site can have a significant effect on the kinetics. In particular, mutations in the FG loop (residues 121 and 122), approximately 17 Å from the active site, cause significant decreases in the rate of hydride transfer.[27−29] According to the theory that binding differences account for enzyme catalysis, mutations far from the active site cause slight alterations in the active-site geometry by communicating structural changes through intervening residues. However, if protein fluctuations are involved in catalysis, mutations far from the active site can affect rates simply by changing the global dynamics of the protein. A similar theory has been proposed for explaining mutational effects on ligand specificity and noncatalytic protein function.[15,30−32] In addition, the effects of distant double mutants in DHFR are not additive.[33−35] These residues must be communicating through the protein scaffold, possibly through protein fluctuations. These experimental results and computations suggest that the DHFR catalytic pathway strongly depends on correlated motions, which can encompass the whole enzyme.

In this study, we explore the protein fluctuations of DHFR using molecular dynamics simulations. DHFR has been the subject of numerous simulations previously; however, most of these focused on protein−inhibitor interactions or used folate as the substrate.[23,26,36−39] The present work covers calculations

(22) Falzone, C. J.; Benkovic, S. J.; Wright, P. E. *Biochemistry* **1990**, *29*, 9667−9677.
(23) Lau, E. Y.; Gerig, J. T. *Biophys. J.* **1997**, *73*, 1579−1592.
(24) Falzone, C. J.; Wright, P. E.; Benkovic, S. J. *Biochemistry* **1994**, *33*, 439−442.
(25) Li, L.; Falzone, C. J.; Wright, P. E.; Benkovic, S. J. *Biochemistry* **1992**, *31*, 7826−7833.
(26) Verma, C. S.; Caves, L. S. D.; Hubbard, R. E.; Roberts, G. C. K. *J. Mol. Biol.* **1997**, *266*, 776−796.
(27) Cameron, C. E.; Benkovic, S. J. *Biochemistry* **1997**, *36*, 15792−15800.
(28) Miller, G. P.; Benkovic, S. J. *Biochemistry* **1998**, *37*, 6327−6335.
(29) Miller, G. P.; Benkovic, S. J. *Biochemistry* **1998**, *37*, 6336−6347.
(30) Feher, V. A.; Cavanagh, J. *Nature* **1999**, *400*, 289−293.
(31) Bahar, I.; Erman, B.; Jernigan, R. L.; Atilgan, A. R.; Covell, D. G. *J. Mol. Biol.* **1999**, *285*, 1023−1037.
(32) Miller, D. W.; Agard, D. A. *J. Mol. Biol.* **1999**, *286*, 267−278.
(33) Wagner, C. R.; Thillet, J.; Benkovic, S. J. *Biochemistry* **1992**, *31*, 7834−7840.
(34) Huang, Z.; Wagner, C. R.; Benkovic, S. J. *Biochemistry* **1994**, *33*, 11576−11585.
(35) Wagner, C. R.; Huang, Z.; Singleton, S.; Benkovic, S. J. *Biochemistry* **1995**, *34*, 15671−15680.
(36) Singh, U. C.; Benkovic, S. J. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 9519−9523.
(37) Fleischman, S. H.; Brooks, C. L., III. *Proteins* **1990**, *1990*, 52−61.
(38) McDonald, J. J.; Brooks, C. L., III. *J. Am. Chem. Soc.* **1992**, *114*, 2062−2072.
(39) Gerber, P. R.; Mark, A. E.; Gunsteren, W. F. v. *J. Comput.-Aided Mol. Des.* **1993**, *7*, 305−323.

on the three ternary complexes from the DHFR catalytic cycle: DHFR/DHF/NADPH (**DH**), DHFR/THF/$NADP^+$ (**TP**), and DHFR/THF/NADPH (**TH**). Comparison of the dynamics of these three complexes will demonstrate the effect ligands can have on the behavior of the protein and possibly which motions are important for catalysis. Ligand-dependent alterations in the protein motions could be necessary for the completion of the catalytic cycle. The ligands only differ by one to two hydrogen atoms. If these small differences can affect dynamics, then mutations, which involve more significant changes, can also influence protein fluctuations.

The motions that we wish to study, however, usually occur on longer time scales than normally covered by molecular dynamics simulations. Thus, our simulations have been carried out to 10 ns, to increase the sampling of slower correlated motions. The extension of these simulations to relatively long time periods raises another question; are the trajectories stable for that length of time. Prior to this, native state simulations had been performed for 5.3 ns on truncated chymotrypsin inhibitor 2 (CI2, 64 residues)[40] and for 100 ns on the villin headpiece subdomain (36 residues)[41,42] (the resulting simulations yielded root-mean-square differences from the target crystallographic or NMR structure of 1.5 Å for both). While the earlier work on CI2 did address concerns about stability, the villin headpiece simulation was more concerned with protein folding and only used the native simulation as a control. Errors in both the methodology and the force field could cause a breakdown of the protein structure over this time period. Hydrogen bonds in secondary structures may shift to neighboring residues, causing the protein to lose its tight packing. Thus, we will also discuss the overall behavior of the protein in our simulations prior to a discussion of the dynamics of the three simulations.

## Methodology

The simulation procedure previously used for segment B1 of protein G was slightly modified for these calculations.[43] The CHARMM macromolecular mechanics program was employed for all molecular dynamics simulations and analysis calculations.[44] The protein was characterized by the CHARMM param22 parameter set,[45] and the water by the TIP3P model.[46] The force field developed by Mackerell and co-workers for NADPH and $NADP^+$ was used to describe the cofactors.[47] The charges for DHF and THF were developed from ab initio calculations. MSI/CHARMM folate parameters[48] were used as a template to develop bonded parameters for DHF and THF.

No crystal structures containing DHF and THF exist; therefore, a model was developed from the crystal structure of DHFR/FOL/$NADPH^+$ with the closed M20 loop conformation.[21] This model was used as the starting structure for all three simulations. After placing the appropriate ligands in their binding sites by superposition onto the folate structure, the complexes were minimized with harmonic restraints to allow the resulting structure to adjust to the new ligands. The minimized structure was then placed in a truncated octahedron of equilibrated water.[43] Overlapping solvent molecules were removed leaving ~3000 water molecules. Before data collection began for analysis, 100 ps of dynamics were run. During the first 20 ps, a harmonic restraint was applied to the protein and ligands, and the temperature was raised to 298 K.

A van der Waals switching function between 8 and 11 Å, an electrostatic shifting function cutoff of 11 Å, a cutoff of 13 Å for the nonbonded list, and periodic boundary conditions were employed during the simulations to reduce errors in the nonbonded forces.[49] All heavy atom−hydrogen atom bonds were held rigid by the SHAKE algorithm.[50] A time step of 0.002 ps was used, and the velocities were reassigned from a Gaussian distribution as necessary to keep the temperature within 5 K of 298 K. During the 10ns, the velocities required adjustment only one time for each simulation. A total of 10 ns of dynamics were run after the equilibration period and snapshots were taken every 200 steps.

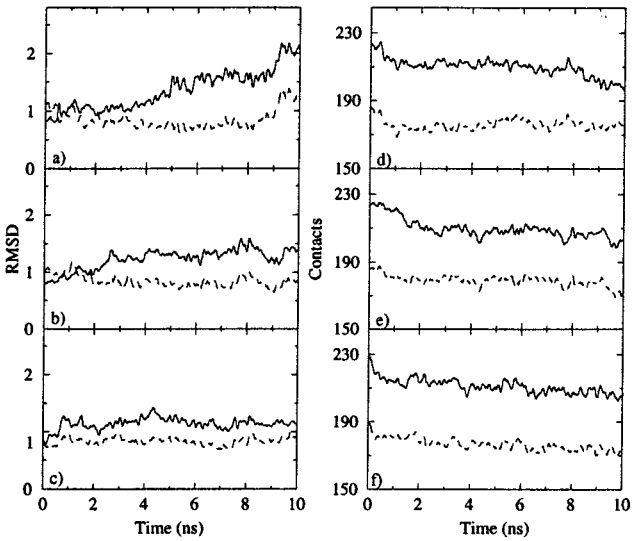



**Figure 3.** Time dependence of structural comparisons in DHFR. (a), (b), and (c) RMSD versus time: RMSD between X-ray crystal structure and trajectory snapshots (solid); RMSD between MD average structure and trajectory snapshots (dashed). (d), (e), and (f) The number of side chain contacts versus time: the number of contacts that overlap with the X-ray crystal structure (solid); the number of overlapping contacts for nonloop residues (dashed). (a) and (d) DHFR/DHF/NADPH, (b) and (e) DHFR/THF/$NADP^+$, (c) and (f) DHFR/THF/NADPH.

## Results and Discussion

**General Characterization of the Models.** Several types of analysis provide information on the stability of a protein during a molecular dynamics trajectory. The time development of the root-mean-square distances (RMSDs) and the static RMSDs with respect to the X-ray structure, as well as the time evolution of side chain contacts and hydrogen bonds, are employed to judge the soundness of the simulations. All of these properties, when combined, furnish a complete picture of the sturdiness of the protein structure in the context of the CHARMM force field.

The X-ray crystal structure and MD average structure were compared to snapshots along the trajectory by calculation of RMSDs (Figure 3). Both the **TP** and **TH** complexes spend an additional period of 2.5 and 0.9 ns, respectively, reaching their equilibrium basins on the potential energy surface. After this, the RMSDs from the X-ray structure oscillate about a mean

(40) Li, A.; Daggett, V. *Protein Eng.* **1995**, *8*, 1117−1128.
(41) Duan, Y.; Wang, L.; Kollman, P. A. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 9897−9902.
(42) Duan, Y.; Kollman, P. A. *Science* **1998**, *282*, 740−744.
(43) Sheinerman, F. B.; Brooks, C. L., III. *Proteins* **1997**, *29*, 193−202.
(44) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187−217.
(45) MacKerell, J. A. D.; Bashford, D.; Bellott, M.; Dunbrack, J. R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Gao, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, I. W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(46) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impley, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(47) Pavelites, J. J.; Gao, J.; Bash, P. A.; Alexander D. Mackerell, J. *J. Comput. Chem.* **1997**, *18*, 221−239.
(48) *QUANTA*; Polygen Corp.: Waltham, MA.
(49) Brooks, C. L., III; Pettitt, B. M.; Karplus, M. *J. Chem. Phys.* **1985**, *83*, 5897−5908.
(50) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.

**Table 1.** Structural Differences between the Reference Crystal Structure (DHFR/FOL/NADP$^+$) and MD Average Structure

| | complex[a] | | |
|---|---|---|---|
| structural element | DHF/NADPH | THF/NADP$^+$ | THF/NADPH |
| backbone | 1.15 | 0.95 | 0.90 |
| nonloop Cα | 0.84 | 0.73 | 0.70 |
| M20 loop | 1.31 | 1.94 | 0.93 |
| CD loop | 2.69 | 1.47 | 1.00 |
| FG loop | 1.80 | 1.28 | 1.41 |
| ligands | 1.50 | 1.37 | 2.18 |

[a] RMSD measured in Å.

value of 1.2 for **TP** and 1.1 for **TH**. For **DH**, periods of fluctuation about an average structure (flat RMSD from the X-ray structure) are interspersed with periods of structural transformations (increasing RMSD from the X-ray structure). These transitions correspond to M20 loop conformational changes as identified by inspection of the snapshots along the trajectory. As noted in the Introduction, such motions are anticipated on time scales greater than a few nanoseconds. Snapshots from the **TP** and **TH** trajectories reveal that the minor long-term fluctuations in the RMSDs from the X-ray structures are also the result of M20 loop conformational changes. In addition, the CD and FG loops undergo slight conformational modifications. The RMSDs from the MD average structures for all three complexes oscillate about a mean of ∼0.85 and are relatively flat except for an increase at ∼9 ns for **DH**.

A comparison of the starting X-ray structure with the MD average structures illustrates the correspondence between the X-ray structure and those adopted within the CHARMM force field and in the presence of differing ligands. The RMSDs between the backbone atoms of these two structures are 1.15 for **DH**, 0.95 for **TP**, and 0.90 for **TH** (Table 1). Such a small amount of movement away from the X-ray structure is remarkable for a 10 ns trajectory. The RMSDs are 0.84, 0.73, and 0.70, respectively, when only the nonloop Cα atoms (i.e. not in the M20, CD, or FG loops) are included in the calculation. This further demonstrates that the core structure of the protein is sustained during the simulation.

The stability of the trajectories can also be assessed by reviewing the time evolution of side chain contacts and hydrogen bonds. This provides a picture of the nature of inter-residue interactions in the complexes. The total number of contacts, whether between side chains or hydrogen bonds, remains constant throughout the simulations, indicating that the protein is not unfolding. In addition, the loss of hydrogen bonds present in the X-ray structure is very small (∼10%), and the hydrogen bonding patterns between all four α-helices and eight *β*-strands are maintained. However, slight but significant decreases in the number of side chain contacts that overlap with contacts in the X-ray structure are observed. Ignoring contacts involving residues in the three loops, M20, CD, and FG, reduces this decrease, and in the case of **DH** completely removes it (Figure 3). Thus, flexible loop conformational changes account for the majority of the changes in contacts compared to the X-ray structure. The core structure of the protein is maintained throughout the simulations.

Our analysis of the RMSDs, hydrogen bonds,and side chain contacts provides evidence for the stability of our simulations. Over such a long trajectory, possible errors in our methodology or force fields could have caused a breakdown in the protein tertiary and secondary structures, but this is not observed. Instead, the core structure of the protein remains particularly close to the X-ray structure, while flexible loops are allowed to change conformation.

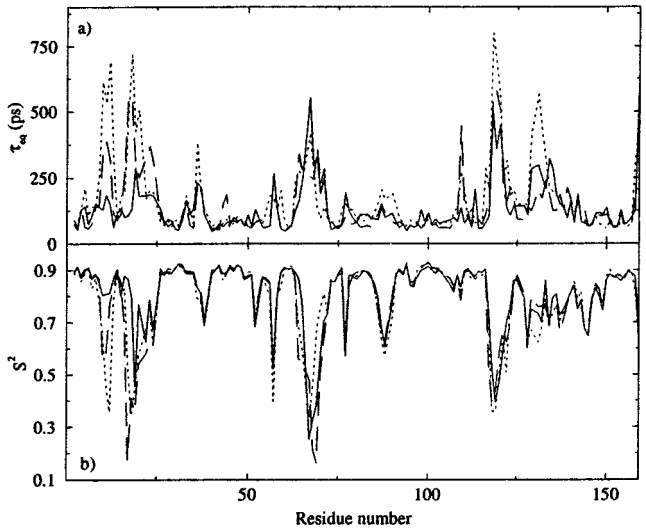



**Figure 4.** (a) The effective internal correlation times per residue. (b) The order parameters per residue. DHFR/DHF/NADPH (solid); DHFR/THF/NADP$^+$ (dashed); DHFR/THF/NADPH (dotted).

**Characterization of Motions.** Comparison of the RMSDs from the MD average structure for various loops illustrates that differences between the trajectories exist. By first orienting the X-ray structure and average MD structure using only the nonloop Cα atoms, RMSDs for the M20, CD, and FG loops can be calculated (Table 1). These regions generally have larger RMSDs than the nonloop Cα atoms. The RMSDs vary between the three complexes, and the loop with the largest deviations for each simulation also differs. Obviously, the behavior of the loop regions is different for each complex. This variation is explored by an investigation of order parameters and relaxation times and by a more in-depth examination of the M20 loop conformational changes.

**(a) Order Parameters, Relaxation Times, and B-Factors.** The flexibility of the backbone can be assessed by determining the N−H bond vector order parameters. Small fluctuations in the angular orientation of the N−H bond result in a high order parameter and indicate low backbone mobility.[51] The results from the three simulations are very similar, although slight differences in the extent of motion exist (Figure 4b). In particular, the M20 loop is less flexible in the **DH** complex. The calculated order parameters from our simulations are in general agreement with those determined by NMR for DHFR/FOL.[52] The major peaks (indicating the most significant motions) occur in the same regions, although the NMR order parameters are much larger (less motion indicated) for these regions. Differences between the three simulations are more apparent in the effective internal correlation times (Figure 4a).[53] The relaxation time of the N−H bond angular correlation function indicates the time scale of backbone motions. The M20, CD, and FG loops all show considerable differences in relaxation times between the three simulations, demonstrating that the ligands affect the dynamics of the protein.

B-factors from the X-ray structure were also compared to those calculated from the MD simulations. Correlations of 0.70 for **DH**, 0.65 for **TP**, and 0.72 for **TH** were observed in

(51) Cavanagh, J.; Fairbrother, W. J.; Palmer, A. G., III; Skelton, N. J. *Protein NMR Spectroscopy: Principles and Practice*; Academic Press: San Diego, 1996.

(52) Epstein, D. M.; Benkovic, S. J.; Wright, P. E. *Biochemistry* **1995**, *34*, 11037−11048.

(53) Palmer, A. G., III; Case, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 9059−9067.

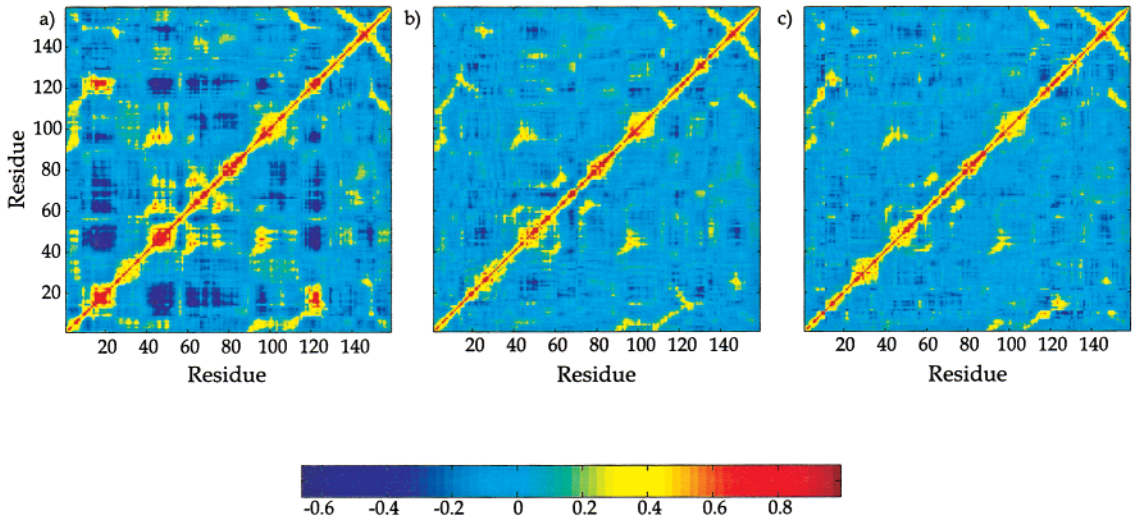



**Figure 5.** Residue−residue based map of correlated motions. Red and yellow indicate regions of positive correlation, and dark blue indicates regions of anti-correlation. (a) DHFR/DHF/NADPH, (b) DHFR/THF/$NADP^+$, (c) DHFR/THF/NADPH.

comparison with the residue-averaged X-ray B-factors for the reference DHFR/FOL/$NADPH^+$ complex. The B-factors are in general agreement with the order parameters: the largest fluctuations occurred in the same regions, and secondary structural elements were less flexible. Like the order parameters, the B-factors were very similar between the three simulations although slight differences exist.

**(b) M20 Loop Conformational Changes.** The extent of the M20 loop conformational changes, which occur in all of the simulations, varies between the three complexes. The M20 loop undergoes a stepwise conversion to a new conformation during the **DH** trajectory. A small change results in a new conformation that fluctuates about an average structure for an extended period of time before another conformational change occurs. Whether or not the M20 loop has completed its conversion by the end of the 10 ns simulation is unclear. In the **TP** simulation, the M20 loop samples several conformations for extended periods of time. The conformations observed include closed, open, a new conformation, and a branching point structure, which lies between all three of these conformations. For **TH**, the M20 loop remains mostly in the closed conformation with short excursions to other conformations: more closed, the branching point, and the new conformation.

In all three simulations, very similar conformations are sampled that do not occur in any X-ray crystal structures. This suggests that a minimum on the force field potential energy surface exists in this region of loop conformations. The existence of this conformation in solution is not precluded by its absence in X-ray structures. However, whether this structure is an aberration of the force field or is truly a solution structure remains to be seen.

Our results demonstrate that the frequency and type of M20 loop conformational changes depend on the ligands bound to the enzyme. Most likely, these differences are necessary to the completion of the catalytic cycle and for the differentiation between reactants and products. In addition, it is interesting to note that small alterations in the ligands can have profound effects on the protein fluctuations.

**Coupled Motions and Catalysis. (a) Correlated Motions.** One can question whether the fluctuations of one residue are related to the fluctuations of another distant residue. Generally, movements of one residue with respect to another appear to be random. However, in some cases the two residues move in concert with each other, either in the same direction (correlated) or in opposite directions (anti-correlated). The extent of the correlation (positive or negative) can be quantified by calculating the covariance between the fluctuations of two residues as shown in Figure 5.[54] Positive correlation (yellow and red regions in the figure) indicates that the two residues generally move in the same direction, while a negative correlation (dark blue regions) indicates that they move in opposite directions.

It is also useful to calculate the correlated motions for parts of the trajectory to determine if the correlation coefficients have converged. Correlated motions have been shown not to be converged after 800 ps in simulations of BPTI, and very different correlation graphs are obtained for lysozyme when sampling different parts of relatively short trajectories.[55] To explore this issue, cross correlations were calculated for each half of the 10 ns trajectory. For the product complexes (TP and TH) agreement is found between the correlations computed from the first half of the trajectory and the entire trajectory. These results are converged. However, for the first 5 ns segment of the reactant trajectory (**DH**) an anti-correlation pattern is observed which is different from that observed overall. The change in the correlation pattern observed for the reactant occurs when the M20 loop begins to change conformation. Correlations during 4−8 ns also agree with the results from the entire 10 ns, suggesting that a consistent pictures of the correlated motions emerge during the last 5 ns.

Analysis of the data presented in Figure 5 shows, as one would expect, that the areas of positive correlation correspond to coupled motions between neighboring $\beta$-strands and neighboring loops. However, the pattern of strong anti-correlation between the M20 loop (residues 14−24) and residues ∼40−80 and between the FG loop (residues 116−125) and residues ∼40−80 in **DH** was not anticipated. These correlations correspond to coupled fluctuations between the structures of the subdomains of the protein that were defined from comparison of X-ray crystal structures[21] and MD simulation studies of the *L. casei* protein.[26] However, what is so surprising is that this coupling disappears in **TP** and **TH**; instead, no significant correlation patterns are observed. The disappearance of this coupling in the two product complexes, even though they also undergo M20 loop conformational changes, shows that the anti-

(54) Ichiye, T.; Karplus, M. *Proteins* **1991**, *11*, 205−217.
(55) Hunenberger, P. H.; Mark, A. E.; Gunsteren, W. F. v. *J. Mol. Biol.* **1995**, *252*, 492−503.

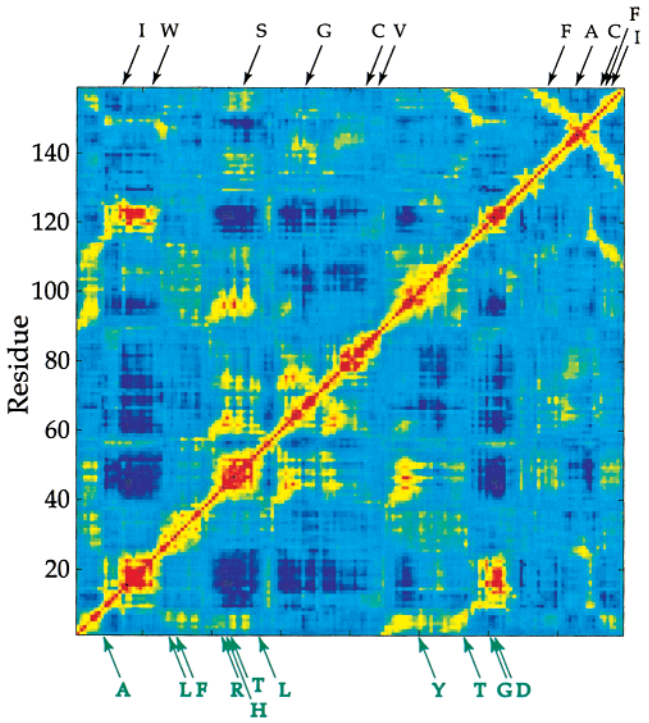



**Figure 6.** Correlations between correlated motions of DHFR/DHF/NADPH and the location of debilitating mutants. Those along the top of the diagram do not affect the first three steps of the catalytic cycle. Those below affect at least one of the first three steps of the catalytic cycle.

correlated motions are caused by some other factor. We conjecture that such coupling may be evidence of a link, at this point undiscovered, between dynamics and catalysis in this system.

**(b) The Effect of Mutations on Dynamics.** Further support for this conjecture comes from mutagenic data. A majority of the mutants that affect the first half of the catalytic cycle occur in regions of the **DH** complex with strongly coupled anti-correlated motion, while those that do not affect catalysis occur mainly in regions of no correlation (Figure 6). The Asp 27 mutant is not included in this analysis, since this residue is involved in the chemical reaction.[56] Mutants included in our analysis must modify catalytic rates through binding or dynamic effects.

There are eleven mutants that affect the rate of at least one of the first three steps of the catalytic cycle by a factor of 6 or more. Seven of these occur in regions that are strongly coupled with distant residues through anti-correlated motions: Ala 9,[57] Arg 44,[58] His 45,[58] Thr 46,[59] Leu 54,[60] Gly 121,[27,28] Asp 122.[29] Perhaps these mutations alter the dynamics of the protein, and the variations in the rates are not just the result of modifications to the binding of the transition state. Interestingly, the four mutants which do not occur in regions with significant coupling to distant residues, Leu 28,[33] Phe 31,[61] Tyr 100,[62] and Thr 113,[63] line the active site. In this case, alterations in the binding geometry probably determine the rate of catalysis.

There are also eleven mutants that do not affect the first three steps of the catalytic cycle. Seven of these occur in regions where there are no anti-correlated motions: Cys 85,[64] Val 88,[65] Phe 137,[66] Ala 145,[67] Cys 152,[64] Phe 153,[68] Ile 155.[68] The Ser 49 mutant[62] occurs in a region with significant coupling and is surrounded by mutants that do affect catalysis. Not every residue in such regions will necessarily influence the kinetics, but the majority should if there is a relationship between dynamics and catalysis. The Gly 67 mutant[57] may be in a situation similar to that of Ser 49. However, no surrounding residues have been mutated to see if this region can effect kinetics through catalysis. Two of the residues, Ile 14[62] and Trp 22,[59] occur in the M20 loop, a region involved in anti-correlated motions. These two residues may have no effect because they are not involved in the conformational changes of the M20 loop.

For the majority of the mutated residues, the kinetic effect of the mutation correlates with the degree of motional coupling in that region: strong coupling through anti-correlated motions occurs with rate modifications, and regions without such correlations show no affect. This, together with the information that these coupled motions disappear in the product complexes, is highly suggestive of a role for protein dynamics in effecting catalysis. The exception are those residues which are in the active site. They cause changes in the rates presumably due to their juxtaposition to the binding pocket or active site residues. This underlies the role that binding plays in determining the rate of catalysis, as was previously understood. Our results reinforce the conjecture that protein dynamics should be considered when exploring how enzymes catalyze reactions, particularly when evaluating the effects of residues far from the active site.

**(c) Hydride Transfer Geometry.** The appearance of motional coupling through the observation of patterns of correlated motions only in the Michaelis complex suggests that these dynamic processes are coupled to the chemical steps in catalysis for this protein system. However, this would be true only if the DHF and NADPH are poised for reaction. The geometry necessary for reaction can be estimated from hydride transfer ab initio calculations reported in the literature.[69−72] These systems, although different in detail, can act as approximate models for the DHFR reaction. An examination of this reaction, *cis* hydride transfer between a pteridine ring (the reactive part of DHF) and (*N*-methyl)nicotinamide (the reactive part of NADPH), at the AM1 and PM3 levels of semiempirical quantum chemistry was also performed.[71] However, examination of the DHFR binding pocket indicates that a *cis* transition state is not possible. Thus, an accurate quantum chemical calculation directly representative of this reaction does not exist. Consensus findings from the earlier calculations were employed to guide our considerations of the "catalytic competence" of the reactive elements from our simulations. From the earlier studies we find

(56) Howell, E. E.; Villafranca, J. E.; Warren, M. S.; Oatley, S. J.; Kraut, J. *Science* **1986**, *231*, 1123−1128.
(57) Miller, P. Ph.D. Thesis, Penn State University, 1998.
(58) Adams, J.; Johnson, K.; Matthews, R.; Benkovic, S. J. *Biochemistry* **1989**, *28*, 6611−6618.
(59) Farnum, M. F.; Magde, D.; Howell, E. E.; Hirai, J. T.; Warren, M. S.; Grimsley, J. K.; Kraut, J. *Biochemistry* **1991**, *30*, 11567−11579.
(60) Murphy, D. J.; Benkovic, S. J. *Biochemistry* **1989**, *28*, 3025−3031.
(61) Chen, J.-T.; Taira, K.; Tu, C.-P. D.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 4093−4100.
(62) Adams, J. A.; Fierke, C. A.; Benkovic, S. J. *Biochemistry* **1991**, *30*, 11046−11054.
(63) Fierke, C. A.; Benkovic, S. J. *Biochemistry* **1989**, *28*, 478−486.
(64) Iwakura, M.; Jones, B. E.; Luo, J.; Matthews, C. R. *J. Biochem.* **1995**, *117*, 480−488.
(65) Ahrweiler, P. M.; Frieden, C. *Biochemistry* **1991**, *30*, 7801−7809.
(66) Dion-Schultz, A.; Howell, E. E. *Protein Eng.* **1997**, *10*, 263−272.
(67) Ohmae, E.; Ishimura, K.; Iwakura, M.; Gekko, K. *J. Biochem.* **1998**, *123*, 839−846.
(68) Dion, A.; Linn, C. E.; Bradrick, T. D.; Georghiou, S.; Howell, E. E. *Biochemistry* **1993**, *32*, 3479−3487.
(69) Wu, Y.-D.; Houk, K. N. *J. Am. Chem. Soc.* **1987**, *109*, 2226−2227.
(70) Wu, Y.-D.; Lai, D. K. W.; Houk, K. N. *J. Am. Chem. Soc.* **1995**, *117*, 4100−4108.
(71) Andres, J.; Moliner, V.; Safont, V. S.; Domingo, L. R.; Picher, M. T.; Krechl, J. *Bioorg. Chem.* **1996**, *24*, 10−18.
(72) Schiott, B.; Zheng, Y.-J.; Bruice, T. C. *J. Am. Chem. Soc.* **1998**, *120*, 7192−7200.

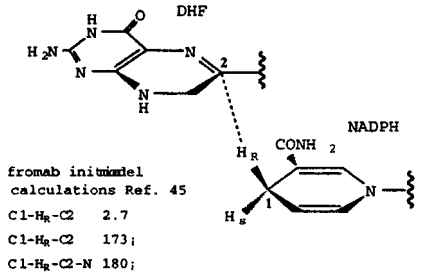



**Figure 7.** Depiction of the reactive complex for hydride transfer.

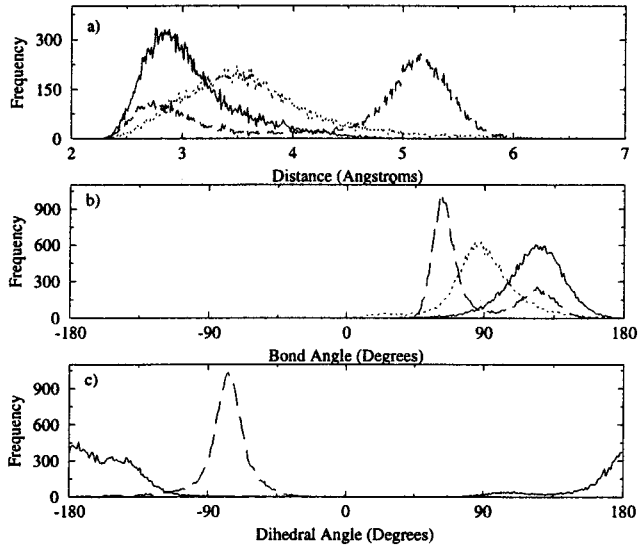



**Figure 8.** (a) Distribution of the distance between transferring hydrogen and the recipient carbon: DHFR/DHF/NADPH C**2**−$H_R$ (solid); DHFR/DHF/NADPH C**2**−$H_S$ (dotted); DHFR/THF/$NADP^+$ C**1**−H (dashed); b) Distributions of bond angle for C**1**−H−C**2**: DHFR/DHF/NADPH *pro-R* H (solid); DHFR/DHF/NADPH *pro-S* H (dotted); DHFR/THF/$NADP^+$ (dashed). (c) Distribution of dihedral angle C**1**−H−C**2**−N: DHFR/DHF/NADPH *pro-R* H (solid); DHFR/THF/$NADP^+$ (dashed).

that: (i) the range of distances between the two carbons which transfer the hydride is 2.6−2.9 Å (average ∼2.75 Å); (ii) the angle between the two carbons and the hydride (∠C−H−C) ranges from 155 to 180° (average ∼168°); (iii) the dihedral angle between the two atoms, hydride and nitrogen, of the DHF model (∠C−H−C−N) is only available from one calculation and is 180°.[69]

The distance, angle, and dihedral angle involved in the hydride transfer were collected and compared to the average ab initio transition state values (Figure 7). The distances between the hydrogen atoms (*pro-R* and *pro-S*) of NADPH and the reactive carbon of DHF were measured. The *pro-R* H−C distance is generally shorter with a peak at a distance of 3 Å (Figure 8). This distance is within reason for the initiation of hydride transfer and agrees with the experimental evidence that the *pro-R* hydrogen is transferred.[73] The average ab initio TS value for the angle (168°) compares well with the simulation distribution of this angle, which peaks at ∼130° (The *pro-S* H angle peak is at ∼90°.). The dihedral angle between the C−(*pro-R*)H of NADPH and the C−N of DHF peaks at −175° in agreement with the ab initio value of 180°. The frequency with which both the bond distance and angle are aligned for hydride transfer was also calculated. The *pro-R* hydrogen of DHF is in a reactive geometry 43% of the time, while the *pro-S* hdyrogen is reaction competent only 7% of the time.

Further support that the reactant complex helps to align the ligands for hydride transfer comes from the absence of this alignment in the product complex. The same distance, angle, and dihedral angle were calculated for the **TP** complex. However, the hydrogen is now on the THF, and the distance being measured is between this H and the reactive carbon of $NADP^+$. The distance and angle in the product complex are seldom close to values conducive to hydride transfer with a large peak around 5.2 Å and ∼60° respectively. In addition, the dihedral angle has a peak at −80° and never approaches the optimal value of 180° during the simulation. The bond distance and angle are aligned for hydride transfer only 16% of the time.

Clearly, in the reactant complex the ligands are aligned for hydride transfer. The *pro-R* hydrogen has the preferred geometry, and the position of DHF will result in the formation of the correct stereoisomer of THF.[74] This alignment disappears in the product complex.

## Conclusions

We have successfully performed three molecular dynamics simulations of DHFR complexes relevant to catalysis to 10 ns. Analysis of our simulation data indicates that our models are highly representative of the complexes studied. The protein maintains its secondary structure as well as the tertiary arrangement of these elements while several flexible loops undergo conformational changes. Surprisingly, the loop changes are very sensitive to which ligands are bound to the protein. Small alterations, such as the addition of one to two hydrogen atoms, cause dramatic differences in the behavior of the protein dynamics.

Strikingly, the Michaelis complex, **DH**, was the only one to exhibit extensive coupling between distant regions of the structure as indicated by the presence of correlated motions. Since all three complexes occur along the kinetic pathway, these differences in coupled motions are probably necessary to the completion of the catalytic cycle, although we have as yet been unable to identify the specific coupled motions that influence catalysis. In addition, the mutagenic data shows a high correspondence between the position of debilitating mutations and the regions of distant residues involved in correlated motions indicating that the mutants could be effecting catalysis by altering the protein dynamics.

Taken together, our results suggest that a relationship between protein dynamics and the rate of catalysis may exist. Despite the fact that the full extent of motions sampled in our calculations of the Michaelis complex may not be complete, our findings are highly suggestive and provide some promise for use of simulation to explore such coupling. Additional evidence is needed to determine the extent of this relationship and its generality. Nevertheless, the effect of dynamics should be considered when investigating other proteins and when designing enzymes.

**Acknowledgment.** We thank the NIH for financial support (GM19276) and (GM56879) and the Pittsburgh Supercomputing Center for its generous allocation of computer resources. We also thank Professor Steve Benkovic for helpful comments and discussions.

JA9913838

(73) Charlton, P. A.; Young, D. W.; Birdsall, B.; Feeny, J.; Roberts, G. C. K. *J. Chem. Soc. Chem. Commun.* **1979**, 922−924.

(74) Fontecilla-Camps, J. C.; Bugg, C. E.; Temple, C.; Rose, J. D.; Montgomery, J. A.; Kisliuk, R. L. *J. Am. Chem. Soc.* **1979**, *101*, 6114.