Origins of Fluorine NMR Chemical Shifts in Fluorine-Containing Proteins†

E. Y. Lau and J. T. Gerig*

Contribution from the Department of Chemistry, University of California, Santa Barbara, Santa Barbara, California 93106

Received June 21, 1999

Abstract: Fluorine NMR chemical shifts from proteins containing fluorinated amino acids are usually dispersed over a wide range when the protein is in its native conformation. The shift dispersion essentially disappears when the protein is unfolded. Origins of the large protein structure-induced shielding effects are not clear, although they have been ascribed to electric field effects, short-range electron−electron interactions ("van der Waals" effects), and various local magnetic anisotropies. The present work explores the relative contributions of electric fields and short-range electronic interactions to fluorine shielding of 6-fluorotryptophan residues contained in the enzyme dihydrofolate reductase *E. coli* (DHFR), in binary complexes of this enzyme with NADPH and with methotrexate, and in a ternary complex with NADPH and methotrexate. Comparison of computed shielding effects to experimental data suggest an important role for short-range electronic interactions in determining fluorine shielding changes in these proteins and a lesser, but nonnegligible, contribution of electric fields and other anisotropies to observed shielding effects. However, the methods employed for calculation of fluorine shielding effects do not have great predictive power for this enzyme, contrary to what has been possible in other systems. The failure to obtain a clean diagnosis of shielding in this system may be a consequence of high conformational mobility.

Introduction

Many fluorinated analogues of amino acids have been incorporated into proteins and peptides.[1] These materials find utility in applications ranging from pharmacology to magnetic resonance imaging, and can often be profitably studied by fluorine NMR spectroscopy.[2,3] For such proteins, a fluorine chemical shift difference perhaps as large as 8 ppm upfield or downfield may be observed for a particular fluorinated residue upon denaturation of the protein. The reasons for the large shift effects induced by the tertiary structures of proteins have not been convincingly elucidated. The large range of shifts observed suggest that local magnetic anisotropies such as those arising from aromatic ring currents and chemical bonds are not major contributors, since these effects likely can add only about ±1 ppm to the total effect observed.[4]

It has been argued that electric fields arising in the interior of a protein are the primary source of the protein structure-induced fluorine shielding effects.[5−7] This conclusion appears to have been generally accepted[8] even though there are systems where electric field effects alone are insufficient to explain observed shieldings.[9]

Short-range electronic interactions between atoms produce NMR shielding effects.[10] Although often referred to as a "van der Waals" effect, Hartree−Fock calculations by Jameson and de Dios[11] indicate that this contribution is better understood as arising from electronic overlap compression and exchange effects, even though the shielding effects can be approximately described by the classical equations for the van der Waals effect on shielding. Short-range or van der Waals interactions have been put forward as the major source of the protein structure-induced fluorine shielding effect[12] or, at the least, as a nonneglible contributor to the effect.[13]

It was the purpose of the present work to explore the relative importance of electric fields and short-range interactions to protein structure-induced fluorine shielding. The contributions of both considerations to shielding in a protein structure depend critically on interatomic separations and a calculation of either effect from a single, static structure is likely to produce misleading conclusions. Thus, we have extended computer simulations of the molecular dynamics of the enzyme dihydrofolate reductase (*Escherichia coli*) which were reported previ-

* To whom correspondence should be addressed. Tel: 805-893-2113; FAX: 805-893-4120; e-mail: gerig@nmr.ucsb.edu.

† Abbreviations used: DHFR, dihydrofolate reductase (*E. coli*); 6F-Trp, 6-fluorotryptophan; F-DHFR, the form of DHFR in which all tryptophan residues have been replaced by 6-fluorotryptophan; MTX, methotrexate; FOL, folate; GBP, galactose binding protein (*E. coli*).

(1) Hortin, G.; Boime, I. *Methods Enzymol.* **1983** *96,* 777-784.
(2) Gerig, J. T. *Prog. NMR Spectrosc.* **1994** *26,* 293−370.
(3) Danielson, M. A.; Falke, J. J. *Annu. Rev. Biophys. Biomol. Struct.* **1996** *25,* 163−195.
(4) Gregory, D. H.; Gerig, J. T. *Biopolymers* **1991** *31,* 845−858.
(5) Augspurger, J.; Pearson, J.; Oldfield, E.; Dykstra, C. E.; Park, K. D.; Schwartz, D. *J. Magn. Reson.* **1992** *100,* 342−357.
(6) de Dios, A. C.; Pearson, J. G.; Oldfield, E. *Science* **1993** *260,* 1491−1496.
(7) Pearson, J. G.; Oldfield, E.; Lee, F. S.; Warshel, A. *J. Am. Chem. Soc.* **1993** *115,* 6851−6862.
(8) de Dios, A. C. *Prog. NMR Spectrosc.* **1996** *29,* 229−278.
(9) Feeney, J.; McCormick, J. E.; Bauer, C. J.; Birdsall, B.; Moody, C. M.; Starkmann, B. A.; Young, D. W.; Francis, P.; Havlin, R. H.; Arnold, W. D.; Oldfield, E. *J. Am. Chem. Soc.* **1996**, *118,* 8700−8706.
(10) Rummens, F. H. A. *van der Waals Forces and Shielding Effects., NMR Basics Principles and Progress Vol. 10.,* Springer-Verlag: New York, 1975.
(11) Jameson, C. J.; de Dios, A. C. *J. Chem. Phys.* **1992** *97,* 417−434.
(12) Mirzadegan, T.; Humblet, C.; Ripka, W. C.; Colmenares, L. U.; Liu, R. S. H. *Photochem. Photobiol.* **1992** *56,* 883−893.
(13) Chambers, S. E.; Lau, E. Y.; Gerig, J. T. *J. Am. Chem. Soc.* **1994**, *116,* 3603−3604.

10.1021/ja992107w CCC: $19.00
Published on Web 04/19/2000

ously[14] and have used the results to explore potential fluorine chemical shielding effects in an analogue of this protein in which the five tryptophan residues have been replaced by 6-fluorotryptophan. Hoeltzli and Frieden have prepared this enzyme and determined that it has activity and stability toward denaturation that are very similar to those of the native enzyme.[15] They determined the fluorine shift for each fluorinated residue in the apo-protein, two binary complexes (with methotrexate and with NADPH), and a ternary complex formed with methotrexate and NADPH, and thereby have provided a large suite of experimental data for comparison to predictions. Our results suggest that both short-range interactions and electric fields influence the protein structure-induced fluorine shielding effects in this protein and its complexes, but indicate that a clear diagnosis of the relative importance of these effects at any given site within the protein is made difficult by the local mobility of the protein.

Experimental Methods

All molecular dynamics simulations were performed using CHARMM (Version 22.5[16]) following basically the same procedures that have been described previously,[14] and we only summarize the protocols used. Every simulation used an all-atom representation for the protein, ligand, and solvent. The available crystal structures for DHFR in the Brookhaven Protein Data Bank were used as the initial coordinates for the protein and ligands. All five tryptophans in DHFR (residues 22, 30, 47, 74, and 133) were replaced with 6-fluorotryptophan. The partial atomic charge of fluorine was taken to be −0.25 au, and the charge of the aromatic carbon bonded to fluorine was adjusted to 0.25 au. These charges are consistent with charges used in previous MD simulations.[4,7] The fluorinated DHFR was placed in the center of a droplet of TIP3P water molecules 35 Å in radius.[17] The water droplet was sufficient in size to provide at least 15 Å of solvent between all protein atoms and the solvent−vacuum interface. Any water molecule within 2.8 Å of a protein atom or a crystallographic water (included in the simulation) was deleted. Typically, the system (protein, ligand(s), and solvent) consisted of approximately 17000 atoms, and the droplet included approximately 4700 water molecules. The potential energy of the system was minimized using a combination of steepest descent and adopted basis Newton−Raphson methods. Dynamics was started when minimization of the potential energy was completed. The system was heated to 300 K in 5 picoseconds (ps) and equilibrated for 20 ps. Production dynamics was performed for 100 ps. The SHAKE algorithm[18,19] was not used because the number of bonds containing hydrogen exceeded the default array size in the software for this algorithm. Since SHAKE was not used, the angle bending constant of water was increased to 250 kcal/mol·radian^2 to prevent deformation of the molecule. The HOH angle of water fluctuated by ±4.5° during simulations. All nonbonded interactions were cutoff with smoothing functions to conserve the total energy.[20,21] A switching function was used between 10 and 11 Å on the Lennard−Jones potential, and the Coulombic potential was cutoff at 12 Å using a shifting function. A harmonic constraint (k = 50 kcal mol^{-1} Å^{-2}) was placed on the oxygen atom of all water molecules 34 Å beyond the origin (center of the spherical droplet) to prevent evaporation from the solvent surface. A dielectric constant of 1 was used in all simulations. Force field parameters other than the partial atomic charges were obtained from the file PARM.PRM, which was included with the program QUANTA 3.2 (Molecular Simulations Inc., Burlington, MA.). The 1−4 nonbonded interactions were scaled by one-half.[22]

Root-mean-squared deviations (RMSD) of atom positions were examined for all simulations to ensure that trajectories were stable. The ranges of observed RMSDs observed for all systems were similar regardless of whether ligands were present. Typical changes in RMSD over the course of a simulation were similar to those previously reported,[14] with the RMSD changing somewhat over the first 40 ps of production dynamics.

All systems simulated were electroneutral and, in most cases, the side chains of amino acid residues were not ionized. It was found earlier that the charges of side chains do not significantly affect the dynamics of F-DHFR complexes.[14] The ligands MTX and NADPH were modeled as charged moieties. The partial atomic charges for MTX and NADPH were assigned using the Gasteiger method,[23] as implemented in QUANTA (Version 4.0). MTX was modeled with an overall −2 au charge. NADPH was modeled with an overall −3 au charge. The Gasteiger charges of the nicotinamide of NADPH were modified to those obtained from ab initio calculations.[24] The phosphate attached to C2′ of the adenosine ribose was modeled with a −1 au charge. The overall charge of NADPH is consistent with the charge used for NADPH in previous free energy calculations of the DHFR binary complex.[25] The differences between the partial atomic charges for the ribose and adenosine of NADPH used and the charges on the corresponding atoms in similar molecules in the residue topology files supplied with QUANTA were usually less than ±0.1 au.

If crystallographic waters were included in a simulation, they were modeled as TIP3P waters. Weak NOE constraints were placed on water molecules at the corresponding positions in apo-F-DHFR and the F-DHFR complexes if the B-factor for the oxygen of a crystallographic water in published structures of the corresponding nonfluorinated systems was below 30 Å^2. The force constant for all NOE constraints was 1 kcal mol^{-1} Å^{-2} with the values for r_{min} and r_{max} at 2.7 and 3.0 Å, respectively. It was shown previously that omission of NOE constraints on the crystallographic waters in the MD simulations did not affect the dynamics of the DHFR.[14] However, these waters could contribute to the electric field experienced by the fluorines and were therefore included. Details of MD simulations of specific systems follow.

Apo-F-DHFR. Due to high disorder within loop 1 (residues 9−23) of the crystal structure (PDB entry 5DFR.ENT), residues 16−20 did not have coordinates in the database. Coordinates for these residues were obtained by attaching residues 16−20 from the PDB file 7DFR.ENT (DHFR−folate−NADP$^+$ ternary complex) to residues 15 and 21 of apo-DHFR. The potential energy of residues 16−20 was minimized while holding the rest of the protein rigid. After minimization of the potential energy of loop 1, the protein was solvated and the MD calculations carried out as described for the general case above.

Three different simulations for apo-F-DHFR were performed. One (apo-F-DHFR-A) used no constraints other than the harmonic constraints placed on the waters at the water−vacuum interface; the crystallographic waters were deleted. Two replicate simulations (apo-F-DHFR−B and apo-F-DHFR−C), done using different initial velocities, included all 120 crystallographic waters. NOE constraints were used in these simulations and included those placed between the hydrogens of water 329 and the hydroxyl oxygen of Thr113 and the carbonyl oxygen of Tyr111.

F-DHFR−MTX Complex. Simulations of this complex and its nonfluorinated form have been described in detail previously.[14]

F-DHFR−NADPH Complex. Coordinates for the F-DHFR−NADPH complex were obtained from the crystal structure of DHFR complexed with NADP$^+$ (PDB entry 6DFR.ENT). Loop 1 for this complex is also disordered, and there are no coordinates for residues 16−20. The same method used to obtain coordinates for residues 16−20 of apo-F-DHFR was used for this binary complex. The coordinates

(14) Lau, E. Y.; Gerig, J. T. *Biophys. J.* **1997** *73*, 1579−1592.
(15) Hoeltzi, S. D.; Frieden, C. *Biochemistry* **1994**, *33*, 5502−5509.
(16) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminthan, S.; Karplus, M. *J. Comput. Chem.* **1983** *4*, 187−217.
(17) Jorgensen, W. L.; Chandraseklar, J.; Madura, J. D., Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983** *79*, 926−935
(18) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1997** *23*, 327−341.
(19) van Gunsteren, W. F.; Berendsen, H. J. C. *Mol. Phys.* **1977** *34*, 1311−1327.
(20) Tasaki, K.; McDonald, S.; Brady, J. W. *J. Comput. Chem.* **1993** *15*, 667−683.
(21) Steinbach, P. J.; Brooks, B. R. *J. Comput. Chem.* **1994** *15*, 667−683.
(22) Momany, F. A.; Rone, R. *J. Comput. Chem.* **1992** *13*, 888−900.
(23) Gasteiger, J.; Marsili, M. *Tetrahedron* **1980** *36*, 3219−3228.
(24) Cummins, P. L.; Gready, J. E. *J. Mol. Struct.* **1989** *183*, 161−174.
(25) Cummins, P. L.; Ramnarayan, K.; Singh, U. C.; Gready, J. E. *J. Am. Chem. Soc.* **1991** *113*, 8247−8256.

for NADPH were generated by overlaying NADPH on to $NADP^+$. All 72 crystallographic waters were included in the simulations. To preserve electroneutrality in the system, residues Arg44 and His45 were modeled as protonated species, and a single sodium atom was placed in the proximity of the diphosphate bonded to the 5′ carbon of the adenosine ribose. The sodium atom was constrained about the diphosphate group using a r^6 potential well, with a Na^+ force constant of 1 kcal mol^{-1} $Å^{-6}$. This method of charge neutralization was used for free energy simulations of DHFR−NADPH.[25] NOE constraints were used between a single crystallographic water molecule (water 405) and the protein. Replicate simulations (F-DHFR−NADPH-A and F-DHFR−NADPH-B) differed only in their initial velocities.

F-DHFR Complexed with MTX and NADPH. The coordinates for the ternary complex (NADPH and MTX) were obtained from the crystal structure of DHFR complexed with $NADP^+$ and folate (7DFR.ENT). The coordinates for MTX in the complex were obtained by rotating the pteridine ring of folate about the C_6-C_9 bond 180° from its original position in the crystal structure and then overlaying MTX on to the repositioned folate.[26,27] Coordinates for NADPH were generated by overlaying the molecule on to the $NADP^+$ coordinates. In one simulation of the ternary complex (F-DHFR−MTX−NADPH-A) the net charge of the ligands was neutralized by modeling residues Arg44, Arg52, and Arg57 of the enzyme with protonated side chains while two sodium atoms were placed in the proximity of the diphosphate group. The constraints used on the sodium counterions were the same as in the F-DHFR−NADPH simulations. Crystallographic waters were not included in this simulation. In a second calculation (F-DHFR-MTX-NADPH−B), the charges of the ligands were neutralized by modeling residues Arg44, His45, Arg52, and Arg57 as protonated residues along with a single sodium atom proximate to the diphosphate group of NADPH. A r^6 potential was used to constrain the sodium atom. All 55 crystallographic waters were included in this simulation. The corresponding waters in the binary complex (MTX) and ternary complex (MTX and NADPH) of DHFR from *Lactobacillus casei* have measured NMR lifetimes in the nanosecond range.[28−30] The long lifetimes indicate that these two waters are not transient entities but part of a stable complex. Thus, NOE constraints were placed between the two water molecules in the folate binding site (waters 403 and 405) and the protein and MTX; a total of 11 NOE constraints were used in this simulation

Gly-6F-Trp-Gly tripeptide. To provide a model of a 6F-Trp residue of F-DHFR in a denatured protein, two simulations (tripeptide-A, tripeptide-B) were performed with the tripeptide Gly-6-F-Trp-Gly solvated in a sphere of water 17 Å in radius (682 water molecules). To minimize endgroup effects, the N-terminus was acetylated and the C-terminal carboxylate was replaced by an amide group. A harmonic constraint was placed on the β-carbon of the fluorotryptophan with a force constant of 250 kcal mol^{-1} $Å^{-2}$ to prevent significant excursions from the center of the solvent sphere. The SHAKE algorithm was used. A spherical potential function was placed on all the water molecules to prevent "evaporation" from the solvent surface. The simulations differed only in the initial velocities used.

Debye−Waller (B) factors and order parameters (S^2) were calculated from the enzyme trajectories of simulations as previously described.[14] Order parameters where calculated from the data obtained between 70 and 80 ps while Debye−Waller factors were calculated using the entire 100 ps of production dynamics.

The contribution of electric fields to fluorine shielding in F-DHFR was calculated using the multipole shielding polarizability method employed by Pearson, et al.,[7] and followed the same procedures as used in the previous study of fluorobenzene.[31] Assuming that the fluorinated tryptophan ring moves isotropically relative to the magnetic field, the observed (isotropic) shielding parameter σ_E produced by the time-varying electric field is given by

$$\sigma_E = \bar{A}_x\langle V_x\rangle + \bar{A}_{xx}\langle V_{xx}\rangle + \bar{A}_{yy}\langle V_{yy}\rangle + \bar{A}_{zz}\langle V_{zz}\rangle + ... \quad (1)$$

where the brackets indicate the average of an electric field component (V_n) or field gradient component (V_{nm}). Values for the coefficients $\bar{A}_{mn}$ are not available for the fluorinated indole ring so those given by Pearson et al. for fluorobenzene (for a series expansion centered on the fluorine nucleus) were used. We have assumed that the electric field at fluorine is due to the partial charges of the surrounding atoms and can be written

$$\vec{\mathbf{F}} = \frac{C}{\epsilon}\sum_j\left(\frac{q_j}{r^2_j}\right)\frac{\vec{\mathbf{r}}}{r_j} \quad (2)$$

where ϵ is the dielectric constant (assumed to be 1), q_j is the partial charge centered on atom j, $\vec{\mathbf{r}}$ is the vector of length r_j from the fluorine nucleus to solvent atom j, and C is a constant that depends on the units for the other quantities in the equation. The coordinate system for calculating the electric field was defined with fluorine as the origin, the x-direction was along the fluorine−carbon bond, the phenyl portion of the indole ring of 6F-Trp in the xy plane, and the z-axis perpendicular to the aromatic ring. If any atom of a residue or a water molecule was within 15 Å of the fluorine of interest, then all atoms in that molecule were used in the field calculation. Field gradients were calculated from derivatives of eq 2.

To take into account the electric fields generated by molecules beyond the 15 Å cutoff, a reaction field $\vec{\mathbf{R}}$ was calculated using eq 3. The volume (V) used was that of the 15 Å radius sphere centered on the fluorine, ϵ is the bulk dielectric constant, and the summation represents the collective dipole moment of the atoms within the sphere. All molecules in the sphere were used for the calculation. The dielectric constant used for the bulk medium in the reaction field calculations were 2 for protein atoms and 80 for water atoms.

$$\vec{\mathbf{R}} = \frac{8\pi(\epsilon-1)}{3V(2\epsilon-1)}\sum_j q_j\vec{\mathbf{r}}_j(3) \quad (3)$$

The effects of short-range interactions σ_{SR} between a fluorine atom and atoms that surround it were estimated using eq 4. It was assumed that these effects are additive and can be calculated by considering only pairwise interactions. Here U_F and U_i are the first ionization

$$\sigma_{SR} = \frac{-3}{2}B_{SR}\frac{U_F U_i}{(U_F+U_i)}\frac{\alpha_i}{r^6_{F,i}} \quad (4)$$

potentials of the fluorine and interacting atom i, respectively, α_i is the static polarizability of the interacting atom, $r_{F,i}$ is the distance between the two atoms, and B_{SR} is a parameter which incorporates the polarizability of aromatic fluorine. A value for B_{SR} of 79 ppm $Å^3$ eV^{-1} was indicated by earlier work with fluorobenzene[31] and was used in the present work. As a result of the previous studies, all contributions to shielding from fluorine−hydrogen interactions were calculated with a $r^{-6.5}$ distance dependence. Atom polarizabilities were taken to be the values given by Dreisbach,[32] and the first ionization potentials for the atoms were taken from the compilation by Levin and Lee.[33] The atomic polarizabilities for the atoms of water were taken from Applequist, Carl, and Fung.[34] All atoms within 5 Å of a given fluorine were used to calculate the shielding effects from short-range interactions.

Fluorine shift effects by both mechanisms were computed using 100 coordinate sets collected at 1 ps intervals from the MD trajectories.

(26) Bolin, D. J.; Filman, D. J.; Matthews, D. A.; Kraut, J. *J. Biol. Chem.* **1982** *257*, 13663−13672.
(27) Reyes, V. M.; Sawaya, M. R.; Brown, K. A.; Kraut, J. *Biochemistry* **1995** *34*, 2710−2723.
(28) Gerothanassis, I. P.;, Birdsall, B.; Bauer, C. J.; Frenkiel, T. A.; Feeney, J. *J. Mol Biol.* **1992** *226*, 549−554.
(29) Meiering, E. M.; Li, H.; Delcamp, T. J.; Freisheim, J. H.; Wagner, G. *J. Mol. Biol.* **1995** *247*, 309−325.
(30) Meiering, E. M.; Wagner, G. *J. Mol. Biol.* **1995** *247*, 294−308.
(31) Lau, E. Y.; Gerig, J. T. *J. Am. Chem. Soc.* **1996** *118*, 1194−1200.
(32) Dreisback, R. R. *Physical Properties of Organic Compounds II*, *Advances in Chemistry, Series 22*, American Chemical Society: Washington D. C., 1959; pp 1−128.
(33) Levin, R. D.; Lea, S. G. *Ionization Potential and Appearance Potential Measurements (1971−1981)*, NSRDS−NBS-71, GPO, Washington D. C., 1982.
(34) Applequist, J.; Carl, J. R.; Fung, K. K. *J. Am. Chem. Soc.* **1972**, *94*, 2952−2960.

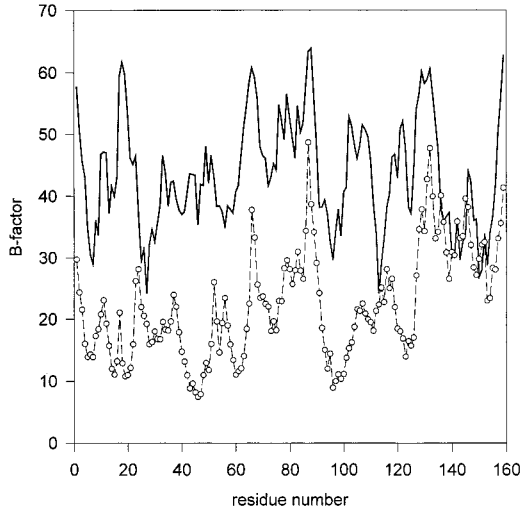

Figure 1. Plot of experimental and calculated Debye−Waller factors for the ternary F-DHFR−MTX−NADPH complex. The solid line is data from a MD calculation (F-DHFR−MTX−NADPH-A) while the dotted line represent experimental points from the crystallographic study of Bystroff, Oatley, and Kraut.[38]

Values for the calculated shielding due to short-range interactions (eq 4) tended to be consistent for each fluorinated residue in F-DHFR, varying by less than 2 ppm between replicate simulations of the same system. The calculated contributions of electric fields to fluorine shielding were much more variable; the calculated effects (eq 1) could differ by as much as 5 ppm for a particular fluorotryptophan in replicate simulations. The used of NOE constraints to keep certain water molecules near crystallographic positions did not affect the calculated fluorine shielding for any residue, the differences in calculated shielding contributions being much smaller than the shift differences between replicate simulations.

Results

Dynamics Calculations. B-Factors. The Debye−Waller (B) factor obtained in crystallographic studies can provide an indication of the mobility of a heavy atom of a protein. An experimental B-factor includes contributions from thermal fluctuations, longer time frame events such as aromatic ring flips, and lattice disorder due to imperfections in the crystal. Removal of the last (often dominant) contribution from experimental B-factors is not trivial.[35−37] On the time scale available to MD simulations only thermal fluctuations are well-sampled.

We have previously reported that B-factors for methotrexate and folate complexes of DHFR (fluorinated and native) calculated from MD simulations carried out using the procedures and assumptions indicated above were in good agreement with those from experiment.[14] Considering only B-factors for the amide nitrogens of the protein backbone, we found in the present work that calculated B-factors for the fluorinated analogues of the ternary complex, the DHFR−NADPH complex, and apoenzyme were all lower than experimental B-factors for the corresponding native (nonfluorinated) systems. Nonetheless, features present in the experimental data indicating structural regions of high and low mobility were often qualitatively reproduced by the simulations (Figures 1 and 2). A major source of the disagreement between the experimental and calculated

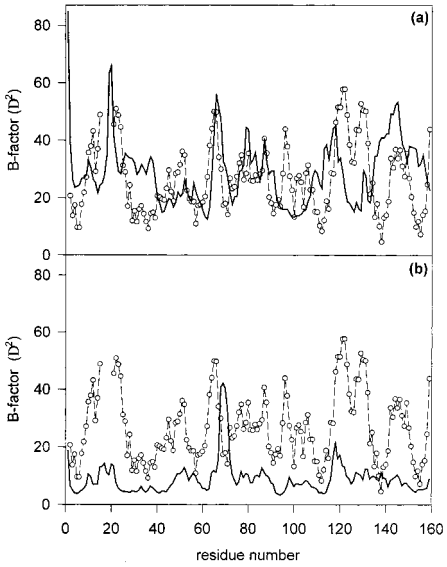

Figure 2. Plot of experimental and calculated Debye−Waller factors for apo-F-DHFR. The solid lines are derived from the apo-F-DHFR-A and apo-F-DHFR-B simulations while the dashed line represents the experimental results of Bystroff and Kraut.[62]

values for the B-factors of these systems is likely lattice disorder since the crystal structures of apo-DHFR, DHFR−$NADP^+$, and DHFR−folate−$NADP^+$ were all solved at a lower resolution (greater than 2 Å, *R* greater than 19%) than the structures of either of the binary complexes DHFR−MTX or DHFR−folate.

The greatest disparity between experimental and calculated B-factors was for the ternary complex (Figure 1). Many of the calculated B-factors of F-DHFR−MTX−NADPH were less than half the experimental values. However, the experimental B-factors for the ternary complex (DHFR−folate−$NADP^+$) used for comparison are known to be increased due to lattice disorder.[38] Hinge residue 88 has the highest experimental B-factor in the ternary complex, and this is reflected in the simulations. Residues 66−69 in turn E are highly mobile in this complex, as was observed previously from simulations of the binary DHFR complexes, and confirmed experimentally for the DHFR-folate complex.

Calculated B-factors from both F-DHFR−NADPH simulations were smaller than the crystallographic B-factors for DHFR−$NADP^+$, although not to the degree observed for the ternary complex.[38] In this case, the disparities are probably not due entirely to lattice disorder, and it may be that the simulation has not adequately sampled all the available configurations for the F-DHFR−NADPH complex. Experiment (DHFR−$NADP^+$) and simulation (F-DHFR−NADPH) both show that the spatial fluctuations for this binary complex are greatly reduced in the solvent exposed loops relative to the fluctuations seen in the crystal structures of DHFR−MTX or DHFR−folate. Differences in spatial fluctuations between F-DHFR−NADPH and F-DHFR−MTX are especially apparent in turn E (residues 63−73) where the calculated B-factors for F-DHFR−NADPH are no greater than at any other position within the protein.[14]

The apo-F-DHFR simulations were the least successful in reproducing features present in experimental data (Figure 2). Experiment suggests that there is significant spatial fluctuation in the three solvent-exposed loops (loop 1, residues 9−23; turn E; the βF-βG loop, residues 117−132). Two of three apo-F-DHFR simulations (apo-F-DHFR−B and apo-F-DHFR−C)

(35) Petsko, G. A.; Ringe, D. *Annu. Rev. Biophys. Bioeng.* **1984**, *13*, 331−371.

(36) Brooks, C. L.; Karplus, M.; Pettitt, B. M. *Proteins: A Theoretical Perspective of Dynamics, Structure and Thermodynamics*; Wiley: New York, 1987.

(37) McCammon, J. A.; Harvey, S. C. *Dynamics of Proteins and Nucleic Acids*; Cambridge University Press: Cambridge, 1987.

(38) Bystroff, C.; Oatley, S. J.; Kraut, J. *Biochemistry* **1990**, *29*, 3263−3277.

Table 1. Mean N−H Order Parameters of F-DHFR

residue range		apo-F-DHFR[a]	MTX complex[a]	NADPH complex[a]	MTX−NADPH complex[a]	experimental[b]
1−8	αA	0.83 (0.09)	0.85 (0.07)	0.86 (0.08)	0.83 (0.10)	0.82 (0.10)
9−23	loop 1	0.78 (0.21)	0.73 (0.28)	0.67 (0.32)	0.78 (0.24)	0.75 (0.07)
24−35	αB	0.87 (0.05)	0.87 (0.07)	0.85 (0.10)	0.85 (0.12)	0.82 (0.04)
35−38	turn C	0.79 (0.13)	0.78 (0.15)	0.81 (0.06)	0.70 (0.15)	0.77 (0.11)
39−43	βB	0.81 (0.05)	0.77 (0.09)	0.85 (0.04)	0.83 (0.07)	0.80 (0.01)
43−50	αC	0.87 (0.04)	0.82 (0.09)	0.88 (0.04)	0.87 (0.04)	0.81 (0.07)
50−58	loop	0.72 (0.34)	0.76 (0.23)	0.72 (0.35)	0.81 (0.14)	0.79 (0.05)
58−63	βC	0.82 (0.06)	0.78 (0.11)	0.84 (0.09)	0.82 (0.06)	0.81 (0.06)
63−73	turn E	0.73 (0.15)	0.64 (0.27)	0.83 (0.04)	0.74 (0.15)	0.70 (0.20)
77−86	αE	0.84 (0.14)	0.85 (0.09)	0.84 (0.06)	0.84 (0.21)	0.85 (0.03)
86−91	turn F,G	0.76 (0.15)	0.74 (0.19)	0.83 (0.04)	0.80 (0.11)	0.73 (0.15)
91−95	βE	0.84 (0.06)	0.81 (0.11)	0.84 (0.06)	0.84 (0.07)	0.80 (0.06)
96−104	αF	0.85 (0.09)	0.86 (0.06)	0.89 (0.02)	0.89 (0.04)	0.83 (0.03)
108−116	βF	0.80 (0.14)	0.78 (0.13)	0.77 (0.21)	0.79 (0.16)	0.82 (0.05)
117−132	loop	0.75 (0.26)	0.70 (0.20)	0.73 (0.20)	0.76 (0.13)	0.72 (0.10)
132−141	βG	0.69 (0.36)	0.76 (0.18)	0.76 (0.20)	0.75 (0.15)	0.82 (0.05)
141−150	turn L	0.82 (0.11)	0.75 (0.19)	0.79 (0.14)	0.83 (0.09)	0.82 (0.04)
150−159	βH	0.82 (0.16)	0.81 (0.16)	0.86 (0.05)	0.84 (0.07)	0.78 (0.06)
α-helices		0.86 (0.10)	0.85 (0.09)	0.86 (0.15)	0.86 (0.14)	0.83 (0.05)
β-sheets		0.79 (0.21)	0.79 (0.15)	0.82 (0.16)	0.80 (0.12)	0.81 (0.07)
turns, loops		0.76 (0.23)	0.73 (0.26)	0.74 (0.25)	0.78 (0.17)	0.75 (0.12)
overall		0.80 (0.22)	0.79 (0.21)	0.79 (0.22)	0.81 (0.16)	0.79 (0.08)

[a] Values given are the average for the region calculated from all simulations performed for that systems. Parenthetic values are standard deviations. [b] Experimental values for the DHFR−folate complex obtained by Epstein, Benkovic, and Wright.[40]

showed only limited mobility in these loops, with the most flexible portion of apo-F-DHFR in all cases being turn E. Only one simulation (apo-F-DHFR-A) produced B-factors similar to the B-factors for the crystal structure of native apo-DHFR (Figure 2).

Backbone N−H Order Parameters. Experimental values for backbone N−H order parameters (S^2) of dihydrofolate reductase are available only for the enzyme−folate complex.[40] These data indicate that many regions in the enzyme undergo dynamical processes on a time scale ranging from 10^{-12} to 10^{-9} s. Previous MD simulations of this complex and its analogue containing 6-fluorotryptophan residues led to calculated N−H order parameters in good agreement with the experimental results.[14] Table 1 presents averages of the calculated order parameters produced by the new simulations of apo-F-DHFR, the binary complexes, and the ternary F-DHFR−MTX−NADPH complex. The calculations indicate that variations in the experimental order parameters seen throughout the structure of the DHFR−folate complex are generally present in these other complexes. The higher mobility indicated for the loop and turn regions of the structure, particularly the solvent-exposed residues 63 to 73, 86 to 91, and 117 to 132, is observed and is quantitatively in good agreement with the experimental values for the folate complex.

Interestingly, the mean N−H S^2 values for individual regions of F-DHFR are similar in the apo-F-DHFR and ternary complex simulations. The only significant differences in the computed order parameters are at β-sheet G and the loop at residues 50−58, which differ by 0.06 and 0.11, respectively, for these two systems.

The most mobile region in the F-DHFR−NADPH simulations is loop 1 (residues 9 to 23). The mean N−H S^2 for this region (0.67) was the lowest of the systems examined. The flexibility observed for loop 1 of this complex is consonant with indications that this part of the DHFR−$NADP^+$crystal is disordered[38] and the results of recent MD simulations.[39] There is significant peak broadening observed for the fluorine NMR signal from F-Trp22 in the F-DHFR−NADPH, a result consistent with exchange broadening arising from motions of loop 1.[15]

We noted in simulations of the DHFR−folate and DHFR−MTX complexes that there appears to be correlated motion of residues in turn E and in the βF-βG loop.[14] This has not been observed in any of the additional simulations reported here.

Experimental N−H order parameters for the native DHFR−folate complex indicate that the backbone N−H bond vectors of the residues comprising the folate binding site and the NADPH binding site are as restricted with regard to angular reorientation as the backbone N−H bond vectors of the entire enzyme, a feature reproduced by the MD simulations of this system.[14] The present work with the apo-F-DHFR and F-DHFR−NADPH systems show that the mean S^2 for N−H bond vectors of residues in the empty folate binding site is comparable to the mean of values for the occupied site. The mean N−H S^2 for the residues of the folate binding site was 0.86 (obtained from the three apo-F-DHFR simulations) while the mean N−H S^2 for this site from the F-DHFR−NADPH simulations was 0.85. Similar, the mobility of the residues which define the NADPH binding site do not differ significantly whether the site is empty or filled. The mean N−H S^2 values for the NADPH binding site are 0.84 and 0.84 from the apo-F-DHFR and F-DHFR−NADPH simulations, respectively. There also appears to be no additional contribution to protein rigidity (in the sense of altering local angular reorientations) by occupation of either the MTX or NADPH binding sites.

Order Parameters for $N_{\epsilon 1}$−$H_{\epsilon 1}$ Bond of 6-Fluorotryptophan. The calculated order parameters for the $N_{\epsilon 1}$−$H_{\epsilon 1}$ bond vector of 6-F-Trp were similar to experimental values found for the DHFR−folate complex (Table 2). One exception observed was for F-Trp22 in one of the F-DHFR−NADPH simulations. In that case, the S^2 for F-Trp22 was 0.27, but this was not due to spatial fluctuations of the indole ring but is rather the result a dihedral transition in the protein backbone that reorients the indole ring of F-Trp22 almost 180°from its position at the beginning of production dynamics. Otherwise, there was a tendency for calculated S^2 values for the $N_{\epsilon 1}$−$H_{\epsilon 1}$ bond vectors to be slightly less than those observed experimentally. There were no indications in the simulations that indole ring

(39) Radkiewicz, J. L.; Brooks, C. L. III. *J. Am. Chem. Soc.* **2000**, *122*, 225−231.

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

Table 2. Order Parameters for $N_{E1}-H_{E1}$ Bond Vectors in F-DHFR

residue	apo-A	apo-B	apo-C	NADPH-A	NADPH-B	MTX-A	MTX-B	MTX−NADPH-A	MTX−NADPH-B	exp[40]
F-Trp-22	0.88	0.90	0.83	0.89	0.27	0.91	0.90	0.92	0.88	0.91
F-Trp-30	0.83	0.89	0.86	0.91	0.89	0.90	0.89	0.83	0.90	0.97
F-Trp-47	0.87	0.89	0.82	0.90	0.92	0.91	0.86	0.89	0.90	0.85
F-Trp-74	0.91	0.91	0.92	0.93	0.91	0.86	0.88	0.80	0.89	0.95
F-Trp-133	0.86	0.86	0.75	0.92	0.74	0.89	0.91	0.92	0.87	0.90

Table 3. Ribose Order Parameters for DHFR Complexes

bond vector	NADPH-A	NADPBBB	ternary-A	ternary-B
		Adenosine Ribose		
H−C1′	0.95	0.92	0.95	0.94
H−C2′	0.93	0.92	0.91	0.94
H−C3′	0.91	0.84	0.92	0.92
H−C4′	0.89	0.85	0.94	0.94
		Nicotinamide Ribose		
H−C1′	0.86	0.46	0.70	0.80
H−C2′	0.80	0.37	0.52	0.40
H−C3′	0.84	0.67	0.84	0.46
H−C4′	0.88	0.74	0.82	0.01

dynamics anywhere in the protein are dependent on the presence or absence of ligand or on the nature of the ligand.

Order Parameters for Ligands. Calculated order parameters for the N−H and C−H bond vectors of MTX from the F-DHFR-MTX have been discussed.[14] The S^2 values for N−H and C−H bond vectors for MTX in the ternary F-DHFR−MTX−NADPH complex were found to be very similar to those of the binary complex. However, the dynamics of the two riboses of NADPH were found to be very different in the F-DHFR−NADPH and ternary complex simulations (Table 3). High flexibility of the ribose ring attached to the nicotinamide is indicated by the order parameters for the C−H bonds of the ring. (Some of the C−H order parameters reported for this ribose are likely incorrect, since the autocorrelation functions of the C−H bond vectors were still decaying when the calculation of S^2 was performed.) In contrast, the adenosine ribose is much more rigid, having order parameters approaching those of the most rigid parts of the enzyme structure. The nicotinamide ribose bridges the major (residues 1−37 and 89−159) and minor domains (residues 38−88) of the protein. There are no strong interactions between this ribose and the protein. The reason for the extreme rigidity of the adenosine furanose ring is likely due to an extensive network of hydrogen bonds that links the sugar and the protein, including hydrogen bonds between the phosphate attached to O2′ of the adenosine ribose and Arg44 (a charged residue). The backbone amides of Arg44 and His45 form hydrogen bonds with O3′ and O4′ of this ribose stabilizing the furanose ring.

MD simulations of DHFR from *Lactobacillus casei* suggest the presence of correlated motions of ligands in the DHFR−MTX−NADPH ternary complex.[41] We have not examined our data to determine if this effect is present in the *E. coli* enzyme.

Order Parameters for C_α−H Bonds. Order parameters for the C_α−H bond vectors were computed for the simulations described and are available in the Ph.D. thesis of E. Y. Lau (University of California, Santa Barbara, CA, 1997). Most MD simulations have shown that the angular reorientation of these bonds is more restricted than reorientation of the corresponding N−H bond (for examples, see refs 42−44), and that observation has been made in our work as well. There are no experimental data for comparison to these calculated quantities.

In summary, the agreement of predictions made from our simulations of the various F-DHFR systems studies with observed Debye−Waller factors and with order parameters from NMR relaxation analysis is reasonably good. To the extent that these comparisons are valid, it appears that the assumptions and protocols used for the MD simulations we have used in this work give reliable indications of the (short-term) dynamics of these systems.

Fluorine Chemical Shifts. The starting point for any discussion of fluorine shielding should be the shielding parameter (σ_{gas}) observed for the fluorinated group in the gas phase at low pressure. There is a significant reduction in the shielding parameter (typically 5−12 ppm) when covalent fluorine is transferred to a condensed phase from the gas phase. Various intermolecular interactions contribute to this shielding change.[45,46] After correction for the effects of changes in bulk magnetic susceptibility, it is usually assumed, as an approximation, that these effects can be represented by a collection of additive terms, shown in eq 5, where σ_A represents the effects of local magnetic anisotropies, σ_{SR} arises from short-range ("van der Waals") interactions between fluorine and surrounding atoms, σ_E is produced by electric fields at the fluorine nucleus, and σ_H results from specific interactions such as the formation of hydrogen bonds.[45] Equation 5 has been used in consideration of shielding effects in systems as diverse as fluorobenzene in a variety of organic solvents[47] to fluorocarbons dissolved in supercritical CO_2.[48] A similar formulation has been used in discussion structure-induced proton shielding effects in proteins.[46]

$$\delta_{obs} = (\sigma_{condensed} - \sigma_{gas}) = \sigma_A + \sigma_{SR} + \sigma_E + \sigma_H \quad (5)$$

In this context, the fluorine shielding difference that is observed between a protein containing a fluorinated amino acid in its native and denatured states, or when a fluorinated small molecule enters a protein binding site, represents a change of environment. There is a gas-to-solution shift change when a fluorinated group enters the (largely) aqueous environment; the protein structure-induced shift arises when the fluorinated group becomes surrounded by the different, possibly much less aqueous, environment provided by the protein. We presume that the gas-to-solution shift and the gas-to-protein shifts can both be treated by means of eq 5, and that there is no need to postulate new kinds of interactions for the protein environment.

The various terms in eq 5 are time-dependent because interatomic interactions between solute and solvent change continually as a result of molecular motions. Thus, the various terms of the equation must somehow be averaged over all system configurations; this is the basis for our concern that the MD

(41) Verma, C. S.; Caves, L. S. D.; Hubbard, R. E.; Roberts, G. C. K. *J. Mol. Biol*. **1997**, *266*, 776−796.
(42) Palmer, A. G.; Case, D. A. *J. Am. Chem. Soc.* **1992**, *114*, 9059−9067.
(43) Smith, P. E.; van Schaik, R. C.; Szyperski, T.; Wuthrich, K.; van Gunsteren, W. F. *J. Mol. Biol.* **1995**, *246*, 356−365.
(44) Philippopulos, M.; Lim, C. *J. Mol. Biol.* **1995**, *254*, 771−792.
(45) Raynes, W. T.; Buckingham, A. D.; Bernstein, H. J. *J. Chem. Phys*. **1962**, *36*, 3481−3488.
(46) Sitkoff, D., Case, D. A. *Prog. NMR Spectrosc.* **1998**, *32*, 165−190.
(47) Suntioien, S.; Laatikainen, R. *Magn. Reson. Chem.* **1991**, *29*, 433−439.
(48) Dardin, A.; DeSimone, J. M.; Samulski, E. T. *J. Phys. Chem. B.* **1998**, *102*, 1775−1780.

methods used for the present work represent in a reliable way the motions available to protein and solvent molecules.

The work of Hoeltzli and Frieden provides values for fluorine shifts for five 6F-Trp residues in each of four forms of F-DHFR.[15] The shift of the tripeptide Gly-6-F-Trp-Gly was assumed to be the same as that of the denatured protein under the conditions of their experiments and provided another value for comparison to a calculated shift. It soon became apparent that fluorine shifts for these systems due to short-range interactions or to electric fields estimated by application of eqs 1 and 4 had little correlation with experiment. It is recognized that there are a number of parameters in both of these equations that were transferred from other systems and therefore may not be appropriate for use with the 6-fluorotryptophan groups in the DHFR systems examined. In the case of eq 4, the value of B_{SR} used was derived from a study of solvent effects on the shielding parameter of fluorobenzene,[31] and the expansion coefficients in eq 1 were derived for fluorobenzene as well.[7]

As a means of exploring the lack of agreement between calculated (δ_{calc}) and observed shieldings (δ_{exp}) in the DHFR systems, we carried out a number of fitting exercises. In the first of these we sought values for the coefficients (R, P_{SR}) in eq 6 which provided the optimum agreement between δ_{calc} and δ_{exp} by the least-squares criterion.

$$\delta_{calc} = R + P_{SR}\sigma_{SR} \tag{6}$$

Both δ_{calc} and δ_{exp} are measured relative to the same arbitrary reference signal. The parameter R in the equation corresponds to the fluorine shift, relative to the 4-fluorophenylalanine reference used by Hoeltzli and Frieden, for a system (peptide or enzyme) in which all protein structural shielding effects are absent. A weak correlation ($r_c = 0.37$) was obtained when comparing calculated shifts for all residues of the enzyme to experimental values in this way. Most of the calculated shifts for residue F-Trp-74 were 2 to 3 ppm different from the corresponding experimental shifts and, if data for F-Trp-74 were excluded from the calculation, r_c improved to 0.64, with the value of the weighting coefficient $P_{SR} = 0.78$, indicating that the contribution of short-range interactions to shielding of the remaining residues was nonnegligible. Correlations done by omitting other fluorinated residues one at a time from consideration did not lead to improved correlations between σ_{SR} and experiment, with such omissions generally producing poorer correlations that the one obtained when only data for residue F-Trp-74 are omitted.

Next the observed shieldings were compared to calculated shielding effects predicted to result from electric fields by fitting eqs 7 to the data. Use of all available data for the enzyme and

$$\delta_{calc} = R + P_{E}\sigma_{E} \tag{7}$$

its complexes produced a weak correlation ($r_c = 0.45$), with a value for P_E of -0.3. Omitting data for F-Trp-74 had virtually no effect on the correlation ($r_c = 0.57$, $P_E = -0.35$). Use of a separate weighting coefficient for each term in Equation 1 did not improve the correlation between computed electric field effect on shielding and the experimental shifts, either when the full data sets were considered or when data for F-Trp-74 in each set was omitted. In either case, elimination of data for other residues did not lead to improved correlations between observed and calculated shielding effects.

Correlations were tried in which the shifts for only one complex were compared to shifts calculated for either mechanism, but in no case did data for a single complex exhibit better agreement between observed and calculated shifts than was apparent for the combined data for all four systems examined by Hoeltzli and Frieden.

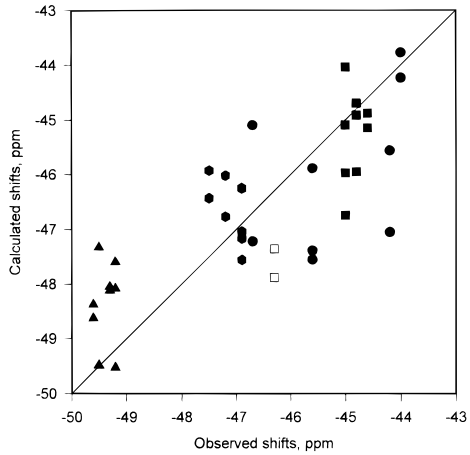

Figure 3. Comparison of observed fluorine shifts for 6F-Trp-containing DHFR and shifts predicted by eq 8 using values of R, P_{SR}, and P_E given in the text. Data for F-Trp74 are not shown but cluster in the lower right corner of the plot. The open squares represent data for the tripeptide while the filled circles, squares, triangles, and hexagons represent data for F-Trp residues 22, 30, 47, and 133, respectively.

Finally, correlations between calculated shifts obtained using eq 8 and experiment were sought in which the R, P_{SR} and P_E were all adjusted. In this case, data for F-Trp-74 were still outliers. However, data for the remaining 38 calculated shifts (from four 6F-Trp residues in nine simulations and two tripeptide simulations) were moderately well reproduced ($r_c = 0.78$) by eq 8 in which $P_{SR} = 0.71$, $P_E = -0.27$, and R was -40.4 ppm. Correlations were also explored in which separate weighting coefficients were used for each term in the expression for the shielding contributions from electric fields with and without the data for residue 74, but the results of these were scarcely different from those obtained with this three-parameter fit.

$$\delta_{calc} = R + P_{SR}\sigma_{SR} + P_{E}\sigma_{E} \tag{8}$$

The plot in Figure 3 compares the observed and calculated shifts obtained with eq 8 and the parameters mentioned for all simulations in which the protein side chains were neutral. The gas-to-solution shift calculated for the tripeptide model of the denatured enzyme is -7.2 ± 0.3 ppm, in moderately good agreement with the shift experienced by fluorobenzene when it is transferred from the vapor to aqueous solution (-6.6 ppm[31]). However, the average error in protein structure-induced shifts predicted by eq 8 was 0.87 ppm, with about 25% of the predicted shifts even having the incorrect sign.

Previous workers have assumed that the charges associated with the amino acid side chains on the surface of a protein can be neglected when estimating the effects of electric fields on fluorine shielding.[7] To check this assumption, simulations of the MTX and MTX-NADPH complexes (F-DHFR−MTX-C, F-DHFR−MTX−NADPH-B) were run with systems in which the fluorinated protein had charged amino acid side chains. Fitting these data to eq 8 showed a modest correlation ($r_c = 0.62$) with the parameters $R = -41.7$, $P_{SR} = 0.52$, and $P_E = -0.20$. The correlation coefficient increases somewhat to 0.74 if F-Trp74 is again excluded from the least-squares fit. Given the scatter in data for systems with neutral side chains and those systems with charged side chains, it appears that inclusion of

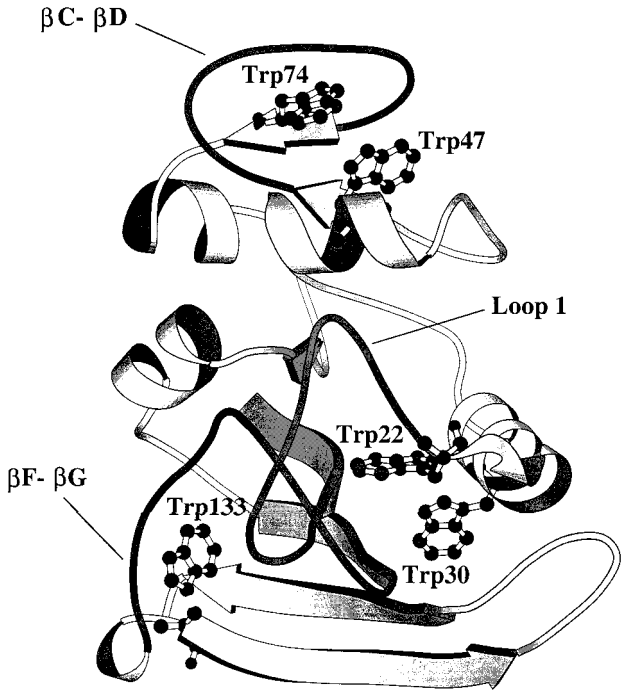

Figure 4. A schematic rendering of DHFR (*E. coli*) showing the relative positions of the five tryptophans in this enzyme. The drawing was generated by MOLSCRIPT.[63]

side chain charges does not have a large effect on computed fluorine shieldings, thus supporting the assumptions made previously.

Discussion

It has been demonstrated that consideration of internal electric fields gives a good accounting of the protein structure-induced fluorine chemical shielding effects for 5-fluorotryptophan residues within galactose binding protein (GBP),[7] although there are indications that shielding in this system can equally be explained by short-range interactions.[13] Gas-to-solution fluorine shifts of fluorobenzene for a number of organic solvents are primarily defined by short-range interactions with smaller but nonnegligible contributions from electric fields in the more polar of these solvents.[31] Given the rather successful treatment of medium-derived shielding effects in these systems, the poor results in the present study are troubling. Certainly, the present work provides no reliable indication that shielding effects of electric fields *or* short-range interactions within DHFR are *dominantly* important for shielding changes produced by the tertiary structure of this protein, although it is probably a safe conclusion that both effects are present.

The need to exclude data for the protein structure-induced shifts of the 6F-Trp74 residue to obtain even the modest correlations observed suggests an unusual environment for the fluorine of this residue. F-Trp47 is proximate to F-Trp74 in the tertiary structure of the protein (Figure 4), and the shift data from the former residue fits the correlation as well as any other. The fluorine of F-Trp47 is directed toward Pro53 and is in van der Waals contact with several of its atoms; the calculated shielding contributions for F-Trp47 were fairly consistent between replicate simulations. In contrast, the fluorine in F-Trp74 is directed toward the exterior of the protein (Figure 4) and is close to the polar side chains of Arg44 and Glu48. The side chain of Arg44 periodically comes within van der Waals contact with this fluorine during all of the F-DHFR simulations. It may be necessary to perform simulations longer than 100 ps before all interactions between the fluorine of F-Trp74 and the components of its environment are well-sampled. As indicated below, ring current effects may also contribute to some of the disparity observed between the calculated and experimental shifts for F-Trp74 since the fluorine of F-Trp74 is in contact with the five-membered ring of the indole of F-Trp47.

There are a number of possible reasons for the generally poor predictions of fluorine shieldings that arise from consideration of electrostatic and short range effects in the F-DHFR systems. These can be grouped into (1) defects in the way electric fields and their effects on shielding are reckoned, (2) defects in the way shielding effects form short-range interactions are estimated, (3) incomplete or misleading representations of the dynamics of DHFR and its complexes, or (4) neglect of other shielding contributions. We briefly consider each of these potential sources of poor predictive performance.

Electric Fields. Coulomb's law was used to calculate the electrostatic field in these simulations. Electronic polarization, which in effect would decrease the electric fields at fluorine, was not taken into account, and a single dielectric constant was used in calculating the electric field experienced by a fluorinated residue. It has been suggested that the dielectric constant of a protein varies throughout the structure, from 2 to 4 near the center of the protein to approximately 20 at the surface.[49] Using any of these values would reduce the shielding effect from electric fields calculated in this work. Explicit treatment of atomic polarizability in the force field in effect reduces the partial atomic charges on atoms by 10−12% relative to static point charges;[50,51] using lower charges could affect the magnitude of the electrostatic field experienced by atoms in the system.

Simple scaling of the calculated shielding contribution predicted by eq 1 appeared to be all that was required to estimate the electric field experienced by the fluorine of fluorobenzene in various solvents, but this approach (eq 6 or 8) appears to offer no benefit in the case of DHFR. Possibly use of the Poisson−Boltzmann equation[52] or a multipole expansion[53] would provide a more accurate estimate of the electric fields, although Augsberger et al. report that a finite difference Poisson−Boltzmann algorithm produced electric fields in their proteins that were in good agreement with the fields computed by the simple Coulombic model used here.[5] Unfortunately, beyond the additional computing costs, obtaining electric field gradients from the Poisson−Boltzmann equation is not trivial and application of the Poisson−Boltzmann method still requires selection of a dielectric constant for the protein.

The dipole shielding polarizabilities used in eq 1 for 6-fluorotryptophan may be significantly different from those of fluorobenzene. The indole ring of 6F-Trp is not symmetric, and off-diagonal elements of the electric field gradients may make significant contributions to fluorine shielding. If the negative weighting parameter (P_E) for the electrostatic field contribution to fluorine shielding for 6-fluorotryptophan in F-DHFR (eq 8) is taken seriously, it may imply that significant revision of the coefficients of eq 1 is needed.

(49) Simonson, T.; Brooks, C. L. *J. Am. Chem. Soc.* **1996**, *118*, 8452−8458.
(50) Caldwell, J. W.; Kollman, P. A. *J. Phys. Chem.***1995**, *99*, 6208−6219.
(51) Cornell, W. D.; Cieplak, P.;, Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(52) Sharp, K. A.; Honig, B. *Annu. Rev. Biophys. Chem.* **1990**, *19*, 301−332.
(53) Lee, F. S.; Warshel, A. *J. Chem. Phys.* **1992**, *97*, 3100−3107.

The calculated electric field contributions to shielding were much more highly variable between replicate simulations than was the case for calculated short range effects. These variations are possibly due to insufficient sampling of electrical environment over the relatively short time span of the simulations.

Short-Range Interactions. Short-range electronic effects appear to have a significant role in defining protein structure-induced fluorine shielding effects in proteins containing the CH_2F group.[9] Our results seem to point to a significant role for short-range interactions in defining the protein structure-induced shielding effects in fluoroaromatic-containing DHFR but do not account for the full range of these effects. The initial B_{SR} value used to calculate the effects on shielding from short-range interactions was that found for fluorobenzene (79 ppm Å^3 eV^{-1}). Although it is chemically reasonable that B_{SR} for 6-fluorotryptophan is similar to that of fluorobenzene, our results ($P_{SR} \sim 0.78$) suggest that a somewhat smaller value may be more appropriate for the 6F-Trp system. Application of eq 4 assumed that short range shielding effects are isotropic and the result of simple atom−atom interactions, a model that is likely not appropriate for a moiety as electronically anisotropic as a fluorinated aromatic ring. Short-range shielding effects in fluorinated aromatic rings need additional investigation by computational methods, including consideration of the anisotropy of these effects. Equation 4 can be taken as an empirical summarization of the results of electronic structure calculations and may eventually better fill this role for the fluorinated indole ring if the polarizability and ionization parameters are considered to be anisotropic.

Dynamics. It is clear that protein flexibility is a key aspect of the mechanism of action of DHFR.[39,54] Overall, our dynamics simulations of F-DHFR are consonant with known properties of the nonfluorinated enzyme and give predicted B-factors and order parameters that qualitatively agree with those of the native protein. However, the dynamics simulations can only provide information about structural motions that take place on a short time scale, from picosecond to nanosecond. NMR studies of the DHFR−MTX complex shows that this system interconverts between two isomers.[55] The apo form of DHFR also exists as two interconverting forms in solution, estimated to be present in a 1:1 ratio.[55−57] Interconversion of these enzyme forms takes place on a time scale much longer that the one sampled by our MD simulations and may involve conformations in which a given fluorine experiences significantly different shielding effects. The rate of interconversion of conformations may nevertheless be fast enough to produce an averaged signal in an experiment, one whose position is reflective of two or more sets of electric and short-range interactions. Hoeltzli and Frieden have noted what is likely exchange broadening of the signal for F-Trp22 in some of their systems, consistent with this possibility.[15] We believe that such conformational mobility is the primary source of the poor correlation between observed and calculated fluorine shielding effects in the F-DHFR systems.

Other Shielding Contributions. A possible explanation for some of the disparity between the calculated and experimental shielding differences, particularly for F-Trp74, maybe due to shielding contributions from local electronic anisotropies. The fluorine of F-Trp74 is in van der Waals contact with the five-membered ring of the indole of F-Trp47. Depending on the orientation of that ring relative to the fluorine, the shielding contribution from aromatic ring currents could be substantial. Ring current shielding contributions were estimated for all 6F-Trp residues in our simulations using the model of Johnson and Bovey.[58] The calculated ring current effects for F-Trp74 vary from approximately −0.40 to 0.47 ppm between simulations while the ring current effects for the other fluorotryptophans were much smaller, less than ±0.15 ppm in all cases. Ring currents shielding effects are very sensitive to the orientation of the aromatic ring relative to the atom of interest, and it may be that the orientations of F-Trp47 or F-Trp74 are not sufficiently well reproduced by the MD simulations. The collective effects of other electronic anisotropies (chemical bonds) of the protein were not estimated but are expected to be small.

The weak agreement between calculated and observed shifts in the various forms of F-DHFR that we have examined stands in contrast to the apparently good predictions of protein structure-induced fluorine chemical shifts obtained for a 5-fluorotryptophan-substituted galactose binding protein (GBP) from *E. coli*.[7] We note that the crystal structures for GBP and DHFR−folate complex have been solved at approximately the same accuracy (resolution ∼ 1.9 Å, $R \sim 15\%$).[27,59] GBP differs from DHFR in that it undergoes only small structural fluctuations, as indicated by its low *B*-factors, compared to the DHFR complexes which have *B*-values 2−3 times larger than those of GPB is certain regions of the structure. Only 20 ps of production dynamics were used for the calculations of Pearson et al., although the rigidity of GBP could mean that a small number of coordinate sets from a MD simulation are adequate to sample the configuration space of this protein.

It may be that electric fields in the interior of a protein are estimated more reliably than the fields that arise from (fluctuating) solvent molecules. GBP is much larger than DHFR, and four of the five fluorotryptophans of the GBP system are located in its interior, while the remaining fluorotryptophan is solvent exposed. Electric field effects were least successful in explaining the observed shift of the solvent-exposed residue. Except for this residue, the relative contribution of solvent to the electric fields experienced by the fluorines in GBP may be less important in GBP than in F-DHFR, leading to a better prediction of shielding when intraprotein fields are correspondingly more important.

The molecular dynamics program used for the GBP study was ENZYMIX,[60] and all residues were modeled as neutral species. The partial atomic charges used for neutral arginine and lysine for the ENZYMIX program are higher than in CHARMM.[16] In ENZYMIX, the neutral carboxylic acid groups of aspartic acid and glutamic acid are modeled with a charge on one oxygen atom and no charge on the other oxygen. Thus, electric fields calculated in the GBP work may not necessarily be comparable to those obtained in the present work.

Surface residues were modeled as neutral amino acids in the GBP simulations since it was observed that surface charges appear not to have a significant effect on the fluorine chemical shifts of 4-fluorotryptophan residues within lysozyme.[61] This assumption may only be valid for 4-fluorotryptophan. The

(54) Sawaya, M.; Kraut, J. *Biochemistry* **1997**, *36*, 586−603.
(55) Falzone, C. J.; Wright, P. E.; Benkovic, S. J. *Biochemistry* **1991**, *30*, 2184−2191.
(56) Falzone, C. J.; Wright, P. E.; Benkovic, S. J. *Biochemistry* **1994**, *33*, 439−442.
(57) Fierke, C. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1987**, *26*, 4085−4094.
(58) Johnson C. E.; Bovey, F. A. *J. Chem. Phys.* **1958**, *29*, 1012−1014.
(59) Vyas, N. K.; Vyas, M. N.; Quiocho, F. A. *Science* **1988**, *242*, 1290−1295.
(60) Lee, F. S.; Chu, Z. T.; Warshel, A. *J. Comput. Chem.* **1993**, *14*, 161−185.
(61) Lian, C.; Le, H.; Montez, B.; Patterson, J.; Harrell, S;, Laws, D.; Matsumura, I.; Pearson, J., Oldfield, E. *Biochemistry* **1994**, *33*, 5238−5245.
(62) Bystroff, C., Kraut, J. *Biochemistry* **1991**, *30*, 2227−2239.
(63) Kraulis, P. *J. Appl. Crystallogr.* **1991**, *24*, 946−950.

fluorine at position 4 of tryptophan will generally be directed toward the protein backbone. Shielding contributions from short-range interactions may be dominant for fluorine in this position because the fluorine tends to remain proximate to the atoms of the protein backbone during dynamics. In this position the electrostatic field contribution to fluorine shielding would be primarily from nearby carbonyl and amide groups. Short-range effects may be less for fluorine at the 5 or 6 positions of the tryptophan indole ring, making nuclei at these positions more sensitive to longer range electric field effects arising from surface charges of the protein.

It is clear that a number of prerequisites must be in place before a quantitative understanding of fluorine shielding effects in proteins will be available. Although it will be useful to obtain more electronic structure calculations related to electric field and short-range shielding effects in the fluoroindole ring system, the ability to model the dynamics of these systems over much longer time frames is probably the most pressing need before much additional progress in this regard can be expected.

Acknowledgment. We thank Prof. M. Karplus for supplying copies of the CHARMM program. This work was supported in part by the U. S. Public Health Service through NIH grant GM-25975 and the University of California, Santa Barbara Committee on Research.

JA992107W