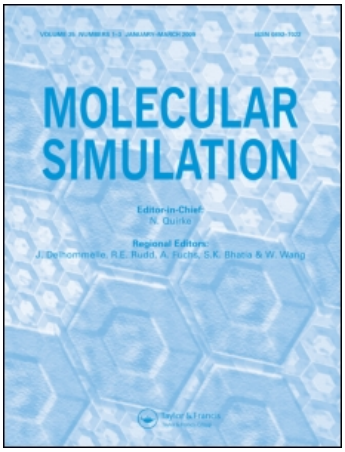

Molecular Simulation
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713644482

The Influence of Starting Coordinates in Free Energy Simulations of Ligand Binding to Dihydrofolate Reductase

Peter L. Cummins[a]; Jill E. Gready[a]
[a] Department of Biochemistry, University of Sydney, Australia

To cite this Article Cummins, Peter L. and Gready, Jill E.(1995) 'The Influence of Starting Coordinates in Free Energy Simulations of Ligand Binding to Dihydrofolate Reductase', Molecular Simulation, 15: 3, 155 — 175
To link to this Article: DOI: 10.1080/08927029508024052
URL: http://dx.doi.org/10.1080/08927029508024052

Molecular Simulation, 1995, Vol. 15, pp. 155–175
Reprints available directly from the publisher
Photocopying permitted by license only

THE INFLUENCE OF STARTING COORDINATES IN FREE ENERGY SIMULATIONS OF LIGAND BINDING TO DIHYDROFOLATE REDUCTASE

PETER L. CUMMINS and JILL E. GREADY

Department of Biochemistry, University of Sydney, N.S.W. 2006, Australia

(Received March 1995, accepted April 1995)

Molecular dynamics (MD) simulation combined with free energy perturbation (FEP) methods have been used to study the key structural differences and relative free energies for the binding of 6-methyl-N5-deazapterin (N8 protonated) and the 8-substituted compound, 6,8-dimethyl-N5-deazapterin (N3 protonated), to dihydrofolate reductase (DHFR). The free energy changes have been calculated using a variety of initial X-ray coordinates derived from bacterial and vertebrate (including human) DHFRs, and both with and without the reduced cofactor nicotinamide adenine dinucleotide (NADPH) bound. Given a sufficiently long simulation time for the FEP calculations (ca. 200 ps), all structures obtained after mutating 6,8-methyl-N5-deazapterin to 6-methyl-N5-deazapterin exhibited hydrogen bond formation between a backbone carbonyl group of DHFR and H(N8) of 6-methyl-N5-deazapterin, analogous to that found in the X-ray crystal structure of N5-deazafolate (N8 protonated) bound to human DHFR. However, both simulation and experiment suggest this additional H-bonding does not greatly enhance thermodynamic stability, with experiment indicating at most a factor of 2 difference in the relative affinities of the two ligand cations for vertebrate DHFR. Moreover, a binding differential of 10 *in favour* of the protonated 8-substituted compound is found experimentally for bacterial DHFR. The MD/FEP calculations suggest that the relative cost of ligand desolvation may largely cancel the lowering of free energy obtained in the active site, resulting in predicted binding differences within the range indicated by the vertebrate and bacterial DHFR experiments. However, the theoretical free energy changes could not be obtained with the accuracy required for the rationalization of the observed species dependence. While sampling difficulties are known to be inherent in MD simulation methodologies, these studies with several initial coordinate sets have demonstrated the contribution of coordinate choice to this problem. The results indicate that for demanding protein-ligand binding problems such as this one, the accuracy of the method may be no better than ± 2 kcal/mol.

KEY WORDS: molecular dynamics, free energy perturbation, dihydrofolate reductase inhibitors, hydrophobic hydration.

INTRODUCTION

The goal of rational drug design is to optimize the physical and chemical properties of a molecule for a highly specific function, usually for strong binding to a particular protein or DNA. Information regarding the nature of specific ligand-protein interactions and the relationship between structure and biological function can be gleaned from X-ray crystallographic studies of complexes formed between substrates or inhibitors and the enzymes. The prediction of relative thermodynamic stabilities, however, requires free energy calculations using molecular dynamics (MD) or Monte Carlo simulation techniques. These so-called free energy perturbation (FEP) or thermodynamic cycle methods may prove to be useful tools in the overall drug-design

strategy [1]. The enzyme dihydrofolate reductase (DHFR) catalyses the nicotinamide adenine dinucleotide phosphate (NADPH) dependent reduction of folate to tetrahydrofolate, and dihydrofolate to tetrahydrofolate, and is a target for various antineoplastic and antibacterial drugs [2,3]. DHFR has become a popular test case for computer-aided drug design and consequently several MD/FEP calculations on the DHFR-binding of drugs such as methotrexate and trimethoprim have been reported in the literature [4–10].

The design rationales we used in the development of mechanism-based 8-substituted-pterin (8-R-pterin) substrates of DHFR [11–15] suggested that 8-R-N5-deazapterins might well be inhibitors of DHFR. We refer to the earlier work [11–15] for details, but simply note here that 8-substitution which produces a more basic compound is central to the pterin-substrate design. Recent experimental work carried out in this laboratory has shown that 8-substituted-N5-deazapterins, including 6,8-dimethyl-N5-deazapterin, are indeed inhibitors of DHFR and display strong binding in ternary complexes with NADPH cofactor [16]. The non-8-R-substituted compound, N5-deazafolate, is also known to be an inhibitor and binds more tightly than folate to DHFR from *E. coli* and chicken liver [17]. Quantum chemical calculations of protonation behaviour and electronic spectra [13, 18] indicate that N5-deazapterins (i.e. analogues of N5-deazafolate) protonate at N8, whereas N3 is the protonation site for the 8-R-substituted compounds, findings also supported by experiment [15, 18, 19]. Structures of the N8 protonated and neutral forms of 6-substituted-N5-deazapterins (e.g. 6-methyl-N5-deazapterin and N5-deazafolate) are shown in Figure 1. The structures of N3 protonated and neutral 6-R'-8-R-N5-deazapterins (e.g. 6,8-dimethy-N5-deazapterin) are shown in Figure 2. Both the 8-substituted (N3 protonated) and non-8-substituted (N8 protonated) cations are stabilized by the resonance-delocalized extended-guanidinium group [13, 20, 21].

One of the key features in the X-ray structure of human DHFR complexed with N5-deazafolate [22] is the hydrogen bond formed between H(N8) of the ligand and the carbonyl oxygen of Ile-7. It has been assumed that this H-bond interaction is responsible for the tighter binding of N5-deazafolate compared with folate which is not

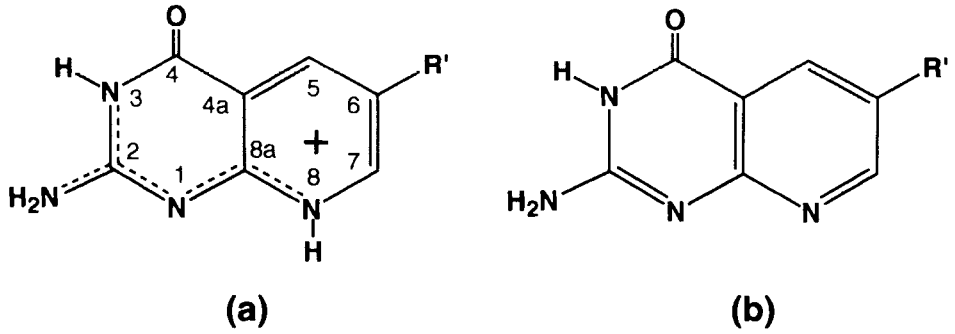

Figure 1 Structures for (a) stable cations showing the extended guanidinium resonance, and (b) neutral forms of 6-R'-N5-deazapterin. In the inhibitor N5-deazafolate, the side chain is R' = methylene(*p*-aminobenzoyl)-L-glutamate.

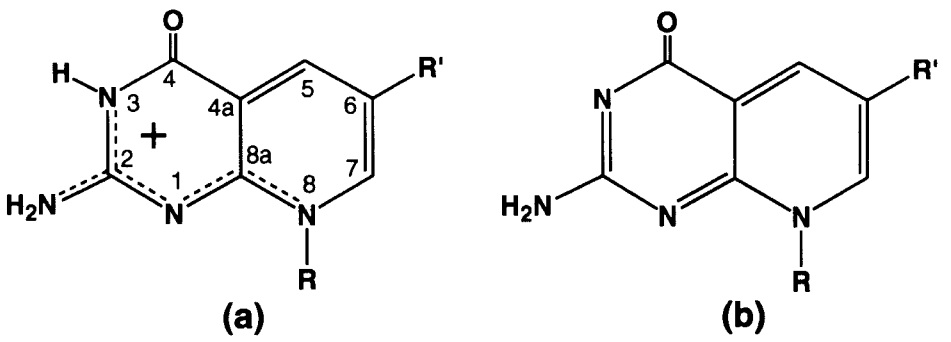

Figure 2 Structures for (a) stable cations showing the extended guanidinium resonance, and (b) neutral forms of 6-R′-8-R-N5-deazapterins.

H-bonded in the X-ray structure and, therefore, presumably is not protonated on N8 [22]. In a recent crystal structure determination of the *E. coli* DHFR.N5-deazafolate complex [23] this H-bonding is absent, suggesting the binding of unprotonated ligand (at pH ~ 7). Clearly, the H-bond interaction will not be present if the N5-deaza-inhibitor is substituted in the 8 position, for example by a methyl group. Whether this 8-substitution translates into reduced or increased affinity of the protonated species for DHFR will depend on the relative free energies of desolvation and ligand binding. The difference in the dissociation constants of the cation form of the two ligands (Figures 1 and 2) from vertebrate DHFRs can be estimated from experiment [14, 16, 24] to be very small (no more than a factor of two). This result seems to suggest that the gain in stability due to possible H-bond formation may be offset by an almost equal gain in solution stability, thereby giving little net gain in binding. However, the experimental results for *E. coli* DHFR [24] suggest that the 8-substituted cation (Figure 2) binds 10 times (1.5 kcal/mol) more strongly. In principle, MD simulation combined with the FEP approach is suited to the study of this type of problem, as it allows the calculation of both the relative solvation free energies for the binding of N8 protonated 6-methyl-N5-deazapterin and N3 protonated 6,8-dimethyl-N5-deazapterin to DHFR.

The practical difficulties inherent in the estimation of free energy differences by molecular simulation are well known and have been discussed in various review articles [25–28]. Many of these difficulties are associated with the choice of initial conditions, subsequent MD simulation conditions and the efficiency of configuration-space sampling. Note that the actual X-ray structure of the complex formed between DHFR and 6,8-dimethyl-N5-deazapterin has not been determined. Consequently we must use starting coordinates for DHFR from the available X-ray studies on complexes where substrates or inhibitors are bound. These substrate or inhibitor molecules are then deleted from the structures and 6,8-dimethyl-N5-deazapterin modeled into the binding site using the known binding-geometry information [12–14]. If the initial X-ray structure turns out to be a poor guess to the structure of the true complex, lengthy equilibration may be required, and *a priori* it is not clear how long the simulation must be run in order to obtain adequate statistics for a free energy determination. A number of free energy determinations should be made by varying the simulation time in order

to test the validity of the result. The success of free energy simulations, therefore, depends critically on the development of protocols which are appropriate for a given problem.

In this work we have examined the convergence of the computed free energies and structures obtained by MD simulation by performing calculations under different conditions, i.e. initial X-ray coordinates, simulation times and force-field parameters. The MD/FEP simulations were carried out on several examples of DHFR from bacterial [29–31] and vertebrate [22, 32–34] sources. Many of these X-ray crystal structures have the cofactor $NADP^+$ or NADPH bound, and all have either a known substrate or inhibitor molecule bound in the active site. For some of the complexes studied, the mutations were carried out in both directions, i.e. the free energy was calculated for the change 6,8-dimethyl-N5-deazapterin → 6-methyl-N5-deazapterin and also for the change 6-methyl-N5-deazapterin → 6,8-dimethyl-N5-deazapterin. As an additional source of error in the calculations derives from the approximations used to obtain potential energy functions in MD simulations, the sensitivity of the computed free energies and structures to the possible H-bond interaction was tested by carrying out simulations using two potential energy functions which differ only in their parameter assignment for the description of interactions between the enzyme and H(N8) of the 6-methyl-N5-deazapterin ligand.

METHODS

Free Energy Perturbation Methods

The free energy differences are obtained by transforming or "mutating" the potential energy parameters of a ligand A (6,8-dimethyl-N5-deazapterin) to those of a second ligand B (6-methyl-N5-deazapterin) during the MD simulation. The relative thermodynamic stability of A and B is given by the free energy difference [1]

$$\Delta\Delta F_{\text{bind}} = \Delta F_{\text{bind}} - \Delta F_{\text{solv}} \tag{1}$$

where ΔF_{solv} is obtained from the simulation in which the mutation A → B is carried out for the free ligand in solution and ΔF_{bind} is from the simulation in which the mutation A → B is carried out for the ligand bound to the protein. The procedure follows closely that used in a previous study on DHFR-binding ligands [35]. The MD/FEP simulations were carried out using the AMBER 3.2 programs [36] and the A → B mutation achieved via λ coupling [37] of the potential energy function (V), where the entire ligand molecule was taken to be the perturbed group of atoms. The coupling parameter λ was divided into discrete values, λ_i, to yield "windows" of uniform width $\Delta\lambda_i = \lambda_{i\pm1} - \lambda_i$. For each window an equilibration simulation was performed followed by a simulation in which all data were collected for the calculation of the perturbation free energy from the formula [1, 37]

$$\Delta F(\pm\Delta\lambda_i) = -\beta^{-1}\ln\langle\exp\{-[V(\lambda_{i\pm1}) - V(\lambda_i)]\beta\}\rangle_{\lambda_i} \tag{2}$$

where $\beta = (kT)^{-1}$ and the ensemble average is taken over the intermediate state defined by λ_i. The total free energy difference for the mutation is obtained by summing over windows and taking the average (and standard error) of the plus and minus perturbation free energies $\Delta F(+\Delta\lambda_i)$ and $\Delta F(-\Delta\lambda_i)$,

$$\Delta F = \frac{1}{2}\sum_i \{\Delta F(+\Delta\lambda_i) - \Delta F(-\Delta\lambda_i)\} \tag{3}$$

In AMBER [36] the potential energy functions used for MD simulations are of the general form

$$V = V_{bad} + V_{ele} + V_{vdw} \tag{4}$$

where V_{bad} includes all bond, angle and dihedral terms, V_{ele} is the electrostatic term arising from nonbonded interactions between atomic partial charges, and V_{vdw} is the nonbonded van der Waals (vdW) interaction term (which may be modified for H-bond interactions). In the MD/FEP calculations the bond energy terms were made redundant (all zero), since the SHAKE algorithm [38] was used to constrain all bond lengths to their specified equilibrium values.

In the present study, the terms in V which correspond to the ligands' intramolecular energy were omitted from the free energy calculation. These terms represent the interactions between atoms of the ligand, i.e. the perturbed group of atoms. These energies have no physical meaning for these ring systems due to the approximate nature of MD force-fields [4, 14, 35]. For example, the ligands' atomic charges are derived to reproduce the electrostatic potential outside the molecules [39], and consequently electrostatic interactions between atoms in the ligand are poorly defined. The use of SHAKE [38] constrains the distances between chemically-bonded atoms and, hence, the ligands' internal vibrations are not properly described. Also, the angle and dihedral terms are only crude approximations taken from the AMBER protein force field [40]. However, accurate vibrational contributions to the free energy for molecules in the gas phase can be estimated readily at the semiempirical AM1 level [41] using standard procedures [42]. For 6,8-dimethyl-N5-deazapterin and 6-methyl-N5-deazapterin we calculate values of -3.3 and -3.2 kcal/mol, respectively. Thus, the free energy difference is negligible. Moreover, these ligands are compact resonance-stabilised [20, 21] ring systems with rigid torsions. Therefore, internal vibrations are only weakly coupled to intermolecular motions in condensed phases, thereby justifying their neglect in MD/FEP calculations.

Clearly the inclusion of the intramolecular terms for these types of ligands may lead to physically unrealistic values for the free energy change. Our experience is that the neglect of the ligands' intramolecular terms yields good agreement with the experimentally-observed trends in dissociation constants for similar DHFR-binding ligands [14, 35]. In this model of the interaction only nonbonded terms between ligand and environment are used in the calculation of the free energy change, as these are physically well-defined when used to describe intermolecular forces.

As shown in Figure 3, the free energy calculations were divided into four stages. Energy minimization (BORN module) was followed by an initial MD equilibration for 7 ps before the mutations were started (GIBBS module). The mutation $A \rightarrow B$ was carried out in two stages via an intermediate state A' (electrostatic decoupling) [43]. The intermediate state A' takes on the electrostatic terms of the final state B, but has the same intramolecular and vdW terms as the initial state A. Thus, the mutation $A \rightarrow A'$ gives the "electrostatic" contribution ΔF_{ele}. The transformation is completed by the mutation $A' \rightarrow B$ to give the "vdW" contribution ΔF_{vdw}, where the vdW and all remaining potential energy terms are changed to those of the final state B. The total free energy is obtained as the sum of ΔF_{ele} and ΔF_{vdw}. There are two reasons for performing the mutation in the stages described. Firstly, as pointed out by Rao and Singh [43], electrostatic decoupling reduces the likelihood of path dependencies that may arise in the calculation of ΔF_{solv}. Secondly, as in the present case the electrostatic mutation represents only a small perturbation on the system and contributes negligibly to the net free energy change [35], the transformation $A \rightarrow A'$ effectively allows additional equilibration time prior to the major structural mutation $A' \rightarrow B$ (i.e. $CH_3 \rightarrow H$). We have found that for the active-site dynamics approach used here, the rms deviations from the initial X-ray coordinates reach a maximum after 40 to 50 ps [44]. Hence, in this work we have economised by following a relatively short initial equilibration (7 ps) with a 42 ps mutation of the electrostatic terms. In previous work [35], we found that a 40 ps simulation time for the calculation of ΔF_{bind} was sufficient to reproduce the experimentally-observed ordering of binding constants for a series of related DHFR-binding ligands. In the present work we have sampled configuration

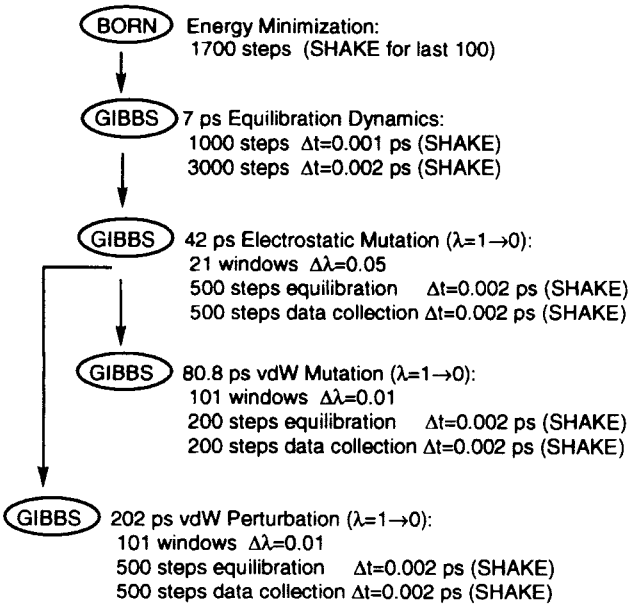

Figure 3 Simulation conditions for the calculation of the solvation free energy differences ΔF_{solv} and binding free energy differences ΔF_{bind}. Δt = integration time step, $\Delta \lambda$ = increment in the coupling parameter.

space more thoroughly by performing the vdW mutation over longer simulation times of ca. 80 and 200 ps. The 200 ps simulation is 100 ps longer than any previously-reported FEP simulation in DHFR [4–10, 14, 35, 45]. For ΔF_{solv}, and ΔF_{bind} for some of the ligand-protein complexes, the calculation of the free energy for the reverse vdW mutation, i.e. $B \rightarrow A'$, was carried out as a further test of the sampling quality.

Potential Energy Parameters

The standard AMBER all-atom force-field parameters of Weiner *et al.* [40] were adopted for the DHFR molecule, while the NADPH cofactor force field parameters that were used are described in an earlier work on cofactor binding to DHFR [45]. Water molecules were assigned the TIP3P force field parameters of Jorgensen *et al.* [46]. For the ligands, 6,8-dimethyl-N5-deazapterin and 6-methyl-N5-deazapterin, structural parameters and atomic partial charges were obtained at the semiempirical AM1 level [41]. The AM1 partial atomic charges (Q) and the vdW parameters (i.e. the Lennard-Jones well depth ε and vdW radii R) for the ligands are given in Table 1. The atomic charges were obtained using methods [39] based on a least-squares fit of the molecular electrostatic potential (MEP) computed within the neglect of diatomic

Table 1 Partial charges Q (a.u.) and van der Waals parameters, R(Å) and ε(kcal/mol), used in the non-bonded potential terms for the 6,8-dimethyl-N5-deazapterin and 6-methyl-N5-deazapterin cations

atom[a]	*6,8-dimethyl-N5-deazapterin*			*6-methyl-N5-deazapterin*		
	Q	R	ε	Q	R	ε
C8a	0.75902	1.85	0.12	0.80869	1.85	0.12
N1	−0.81704	1.75	0.16	−0.86008	1.75	0.16
C2	0.78848	1.85	0.12	0.85078	1.85	0.12
N2	−0.51692	1.75	0.16	−0.52741	1.75	0.16
H2A	0.32454	1.00	0.00	0.32980	1.00	0.00
H2B[b]	0.31536	1.00	0.00	0.31761	1.00	0.00
N3	−0.64056	1.75	0.16	−0.68701	1.75	0.16
H3	0.33181	1.00	0.00	0.34137	1.00	0.00
C4	0.86360	1.85	0.12	0.88708	1.85	0.12
O4	−0.46672	1.60	0.20	−0.46185	1.60	0.20
C4a	−0.62301	1.85	0.12	−0.67329	1.85	0.12
C5	0.21355	1.85	0.12	0.27036	1.85	0.12
H5	0.13016	1.54	0.01	0.11504	1.54	0.01
C6	−0.09604	1.85	0.12	−0.11401	1.85	0.12
C6′	−0.24409	1.80	0.06	−0.23255	1.80	0.06
H6A	0.10732	1.54	0.01	0.10281	1.54	0.01
H6B	0.11343	1.54	0.01	0.11185	1.54	0.01
H6C	0.10878	1.54	0.01	0.10858	1.54	0.01
C7	0.06383	1.85	0.12	0.05536	1.85	0.12
H7	0.15847	1.54	0.01	0.17674	1.54	0.01
N8	−0.11044	1.75	0.16	−0.22605	1.75	0.16
C8/H8	−0.10345	1.80	0.06	0.30503	1.00	0.00 & 0.02
H8A	0.10859	1.54	0.01	0.00000	1.54	0.00
H8B	0.10776	1.54	0.01	0.00000	1.54	0.00
H8C	0.12241	1.54	0.01	0.00000	1.54	0.00

[a] See Figures 1 and 2 for a guide to the atom labels.
[b] The hydrogen closest to N3.

differential overlap (NDDO) approximation. Apart from the 8-methyl group that is being mutated to H, the vdW parameters are exactly the same for both ligands and consequently do not contribute to the perturbation free energy given by Equation (2). The choices for R and ε in Table 1 were derived from appropriate analogs in the all-atom force-field for proteins [40], except for hydrogen atoms attached to nitrogen for reasons we discuss below.

In order to gauge the sensitivity of the free energies to the H-bond interaction of interest, i.e. between H(N8) and the enzyme, the vdW contributions have been computed in two ways corresponding to different choices of the Lennard-Jones 6–12 interaction potential between H(N8) of the ligand and the protein atoms. The vdW term for any pair of atoms i and j may be written in the form

$$E_{vdw} = (\varepsilon_i \varepsilon_j)^{1/2} [(R_i + R_j)^{12}/R_{ij}^{12} - 2(R_i + R_j)^6/R_{ij}^6] \quad (5)$$

where ε_i, ε_j, R_i, and R_j are parameters and R_{ij} is the interatomic distance. We have used the two values $\varepsilon_i = 0.02$ kcal/mol and $\varepsilon_i = 0$ (i.e. $E_{vdw} = 0$), for the case where the atom (i) refers to H(N8) of the ligand. Hence, in Table 1 there are two values given for this parameter. For each value of ε_i the vdW radii were kept at $R_i = 1.0$ Å. The AM1 partial atomic charges also remain the same for each simulation with a different value of ε_i. The value $\varepsilon_i = 0.02$ is from the Weiner *et al.* [40] protein force-field for hydrogens attached to nitrogen, while $\varepsilon_i = 0$ is generally more appropriate for any H-bonded interaction, since $\varepsilon_i > 0$ usually leads to an underestimation of the H-bond strength [47]. For the remaining H atoms attached to N, *i.e.* H(N2) and H(N3), we have chosen $\varepsilon = 0$ for both ligands, since these are known [12, 13, 22] to be involved in H-bonding with a conserved carboxylate side-chain in the enzyme (Asp in bacterial and Glu in vertebrate enzyme).

Molecular Dynamics Simulation Conditions

For the calculation of ΔF_{solv}, the 6,8-dimethyl-N5-deazapterin ligand was solvated with boxes of Monte Carlo water using the EDIT module of AMBER. Only those water molecules within 12 Å (in the x, y and z directions) of any solute atom were retained to form a rectangular simulation box of dimension 33.1 Å × 30.4 Å × 26.0 Å (z axis perpendicular to the plane of the ring system) containing 852 water molecules. The simulations for the ligands in solvent were carried out for the closed, isothermal, isobaric (NTP) ensemble under periodic boundary conditions using a cutoff radius of 8 Å for all nonbonded interactions, a temperature of 300 K and pressure of 1 atm. For the calculation of ΔF_{bind}, the starting coordinates for the DHFR molecule (and $NADP^+$/NADPH where present) were taken from selected X-ray crystal structure determinations: lcDHFR (*Lactobacillus casei* DHFR) [29], ecDHFR (*Escherichia coli* DHFR) [30, 31] clDHFR (chicken liver DHFR) [32, 33] and rhDHFR (recombinant human DHFR) [34]. The DHFR.ligand and DHFR.NADPH.ligand complexes were neutralized by adding Na^+ and Cl^- counterions. It is not computationally feasible at present to solvate the entire protein within a simulation box due to the excessively large numbers of water molecules that would be required. Consequently, in order to reduce the computational cost of the simulations to a reasonable level while minimizing the

loss of accuracy due to neglect of the bulk solvent effects, we have used an active-site dynamics approach [48]. In this approach a molecular dynamics zone is defined by including all atoms or residues within a certain radius of the active site, thus partitioning the system into "non-dynamic" atoms whose positions are frozen at the initial X-ray coordinate values and "dynamic" atoms which are the only ones allowed to move during the various stages of the calculations. An active-site dynamics zone confined to be within a radius of ca. 16 Å from the ligand center of mass appears to be a reasonable choice on the basis of previous studies on ligand-binding to DHFR [4–9, 35]. The region within this radius was solvated using EDIT (crystallographic water molecules were retained in the simulations) in order to fill any gaps in the structure. The dynamic atoms were then defined by including all protein residues, counterions and water molecules within the 16 Å radius from the ligand mass center (yielding 2000–2200 dynamic atoms). Water molecules and counterions were restrained from leaving the 16 Å dynamics region by applying a small half-harmonic (0.6 kcal/mol/Å^2) radial restoring force [48]. The ΔF_{bind} simulations were carried out at constant temperature of 300 K and using an 8 Å cutoff for all nonbonded interactions, consistent with the ΔF_{solv} calculations.

RESULTS AND DISCUSSION

Both free energy changes, ΔF_{bind} and ΔF_{solv}, defined in Equation (1) have been obtained as the sum of electrostatic (ΔF_{ele}) and vdW (ΔF_{vdw}) contributions according to the scheme shown in Figure 3. The various free energy terms are given in Table 2, where the change for a total mutation (in solvent or DHFR) is given by $\Delta F_{total} = \Delta F_{ele} + \Delta F_{vdw}$. The results in parentheses are for the 200 ps vdW mutation. Note that for all systems the electrostatic components are quite small compared with the vdW components. A relatively small free energy change for ΔF_{ele} (i.e. mutation of the coulomb term V_{ele}) is to be expected since both the total ligand charge is conserved and the partial charges (Q) in Table 1 indicate only minor differences in the molecular charge distributions and, hence, MEPs of the two ligands. Thus, our original assumption that the electrostatic mutation should be regarded as essentially an equilibration phase seems well justified (see FEP methods).

Free Energy Changes for 80 ps vdW Mutation

The vdW component depends heavily on the enzyme source of the initial coordinates for the complexes and also the ε_i parameter (where i = H(N8)). With the exception of the lcDHFR.ligand complex, the parameter $\varepsilon_i = 0$ kcal/mol consistently gives lower ΔF_{vdw} values than $\varepsilon_i = 0.02$ kcal/mol. However, no definite trend emerges when comparing ΔF_{vdw} for the different initial structures or when considering the presence or absence of NADPH cofactor. The relative thermodynamic stabilities of ligand binding, $\Delta\Delta F_{bind}$, also vary markedly, giving a range of values depending on initial coordinates and ε_i. The simulation performed on lcDHFR($\varepsilon_i = 0$) yields $\Delta\Delta F_{bind} = 1.37 \pm 0.15$ kcal/mol, i.e. the 6,8-dimethyl-N5-deazapterin complex is a factor of ca. 10 more stable than the 6-methyl-N5-deazapterin complex. In contrast, rhDHFR($\varepsilon_i = 0$)

Table 2 Free energy changes (kcal/mol) calculated using the procedure outlined in Figure 3 for the mutation 6,8-dimethyl-N5-deazapterin → 6-methyl-N5-deazapterin in solution and for the ligands bound to various dihydrofolate reductases with and without NADPH cofactor. Results for 80 ps simulation, with 200 ps simulation in parentheses

System	ε_i [a]	ΔF_{ele} [b]	ΔF_{vdW} [b]	ΔF_{total} [c]	$\Delta\Delta F_{bind}$ [c]
solution	0.02	-0.25 ± 0.01	-3.83 ± 0.07	-4.08 ± 0.08	—
		-0.25 ± 0.01	(-3.48 ± 0.05)	(-3.73 ± 0.06)	—
	0.00	-0.25 ± 0.01	-6.63 ± 0.02	-6.88 ± 0.03	—
		-0.25 ± 0.01	(-6.11 ± 0.06)	(-6.36 ± 0.07)	—
lcDHFR[d]	0.02	-0.63 ± 0.02	-6.05 ± 0.15	-6.68 ± 0.17	-2.60 ± 0.25
	0.00	-0.63 ± 0.02	-4.88 ± 0.10	-5.51 ± 0.12	1.37 ± 0.15
		-0.63 ± 0.02	(-7.42 ± 0.02)	(-8.05 ± 0.04)	(-1.69 ± 0.11)
lcDHFR.NADPH[d]	0.02	-0.43 ± 0.01	-6.53 ± 0.08	-6.96 ± 0.09	-2.88 ± 0.17
	0.00	-0.43 ± 0.01	-9.33 ± 0.05	-9.76 ± 0.06	-2.88 ± 0.09
ecDHFR[e]	0.02	-0.52 ± 0.01	-3.82 ± 0.10	-4.34 ± 0.11	-0.26 ± 0.19
	0.00	-0.52 ± 0.01	-7.90 ± 0.09	-8.42 ± 0.10	-1.54 ± 0.13
		-0.52 ± 0.01	(-7.34 ± 0.01)	(-7.86 ± 0.02)	(-1.50 ± 0.09)
ecDHFR[f]	0.02	-0.69 ± 0.02	-4.40 ± 0.01	-5.09 ± 0.03	-1.01 ± 0.11
	0.00	-0.69 ± 0.02	-5.82 ± 0.03	-6.51 ± 0.05	0.37 ± 0.08
		-0.69 ± 0.02	(-5.57 ± 0.07)	(-6.26 ± 0.09)	(0.10 ± 0.16)
ecDHFR.NADPH[f]	0.02	-0.49 ± 0.01	-3.79 ± 0.02	-4.28 ± 0.03	-0.20 ± 0.11
	0.00	-0.49 ± 0.01	-7.02 ± 0.11	-7.51 ± 0.12	-0.63 ± 0.15
clDHFR[g]	0.02	-0.54 ± 0.01	-2.79 ± 0.03	-3.33 ± 0.04	0.85 ± 0.12
		-0.54 ± 0.01	(-3.62 ± 0.07)	(-4.16 ± 0.08)	(-0.43 ± 0.14)
	0.00	-0.54 ± 0.01	-5.39 ± 0.02	-5.93 ± 0.03	0.95 ± 0.06
clDHFR.NADPH[g]	0.02	-0.43 ± 0.01	-3.60 ± 0.02	-4.03 ± 0.03	0.05 ± 0.11
	0.00	-0.43 ± 0.01	-5.35 ± 0.03	-5.78 ± 0.04	1.10 ± 0.07
		-0.43 ± 0.01	(-4.41 ± 0.01)	(-4.84 ± 0.02)	(1.52 ± 0.09)
rhDHFR[h]	0.02	-0.25 ± 0.01	-6.29 ± 0.07	-6.54 ± 0.08	-2.46 ± 0.16
	0.00	-0.25 ± 0.01	-10.29 ± 0.17	-10.54 ± 0.18	-3.66 ± 0.21
		-0.25 ± 0.01	(-7.34 ± 0.02)	(-7.59 ± 0.03)	(-1.23 ± 0.10)

[a] vdW parameters $\varepsilon_i = 0.0$ and 0.02 kcal/mol for H(N8) of 6-methyl-N5-deazapterin.
[b] mean value [Equation (3)] and standard error of the two free energies, $\Delta F(+\Delta\lambda)$ and $\Delta F(-\Delta\lambda)$.
[c] the total free energy change (i.e. ΔF_{solv} for solvation or ΔF_{bind} for binding) is the sum of the electrostatic (ΔF_{ele}) and vdW (ΔF_{vdw}) components
[d] DHFR.NADPH.methotrexate coordinates (ref. 28)
[e] DHFR.methotrexate coordinates (ref. 29)
[f] DHFR.$NADP^+$.folate coordinates (ref. 30)
[g] DHFR.$NADP^+$.biopterin coordinates (ref. 32)
[h] DHFR.folate coordinates (ref. 33)

yields $\Delta\Delta F_{bind} = -3.66 \pm 0.21$ kcal/mol, i.e. the 6-methyl-N5-deazapterin complex is the more stable by a factor greater than 10^2.

The simulations for rhDHFR($\varepsilon_i = 0$), lcDHFR($\varepsilon_i = 0$), clDHFR.NADPH($\varepsilon_i = 0$), ecDHFR($\varepsilon_i = 0.02$) and in solution ($\varepsilon_i = 0$ and $\varepsilon_i = 0.02$) were carried out in the reverse direction (i.e. $\lambda = 0 \rightarrow 1$) for the vdW mutation. In these reverse-mutation calculations, the final structure obtained from the forward mutation is used as the initial structure for the reverse mutation. The results (note that the signs have been changed for comparison with results in Table 2) are summarized in Table 3, together with the differences between the corresponding forward mutations given in Table 2. The differences

Table 3 The free energy $\Delta F_{reverse} = -\Delta F_{vdw}$ (kcal/mol) for the reverse van der Waals mutation and the corresponding forward-reverse differences ($\Delta F_{forward} - \Delta F_{reverse}$) between vdW terms

system	*time*[a]	ε_i	$\Delta F_{reverse}$	$\Delta F_{forward} - \Delta F_{reverse}$[b]
solvent	80	0.02	-3.38 ± 0.02	-0.45
	80	0.00	-5.56 ± 0.10	-1.07
lcDHFR	80	0.00	-3.90 ± 0.04	-0.98
ecDHFR	80	0.02	-4.29 ± 0.08	0.47
clDHFR.NADPH	80	0.00	-5.63 ± 0.05	0.28
rhDHFR	80	0.00	-8.55 ± 0.01	-1.74
	200	0.00	-9.15 ± 0.05	1.81

[a] Simulation time (ps) for vdW mutation.
[b] $\Delta F_{forward}$ are the ΔF_{vdw} values from Table 2.

between free energies calculated for the forward and reverse mutations vary considerably. In some cases, particularly rhDHFR where the difference is almost 2 kcal/mol, the free energy change appears to be quite strongly coupled to the mutation pathway, while for clDHFR.NADPH the difference is only 0.3 kcal/mol indicating weak path dependence. Also, with regard to sampling in the ΔF_{bind} calculations, it will be noted in Table 2 that simulations were carried out on ecDHFR binary complexes with initial coordinates from two different X-ray determinations. In one simulation the initial coordinates for the DHFR molecule were taken from the ecDHFR.methotrexate structure, while in the second simulation the DHFR coordinates originated from the ecDHFR.$NADP^+$.folate structure. The differences between ΔF_{bind} for these two simulations are calculated from the results in Table 2 to be 0.7 kcal/mol for $\varepsilon_i = 0.02$ and 2 kcal/mol for $\varepsilon_i = 0.0$. Uncertainties of these magnitudes (ca. 2 kcal/mol) are likely to obscure any species-dependent (i.e. bacterial vs vertebrate) or cofactor-dependent trends in inhibitor binding [15–17].

In understanding the origin of differences between the various results for ΔF_{bind} and ΔF_{solv} in Table 2, it is instructive to examine the change in the vdW component of the free energy as a function of the coupling parameter λ (or window number), as shown in Figure 4, and also to examine differences in active-site structures obtained from MD simulation, as depicted in the stereo views in Figures 5 and 6. Note first that the magnitude of the vdW component will depend on the strength of the interaction between ligand and its surroundings. In solution, strong electrostatic interactions cause the polar water molecules to be attracted to the ion, which at equilibrium result in a significant overlap between the electronic charge clouds of the water and ion. In the simulations, this effect is manifested in repulsive vdW interaction terms for the solvent in direct contact with the molecular ion, which may be quite large for the hydrophobic points of contact. Consequently, as the hydrophobic CH_3 group is contracted into H during the forward mutation, the negative free energy change can be seen as a release of vdW energy.

In principle, similar arguments apply to interactions in the active site of an enzyme, although the important difference here is that the active site structure is comparatively rigid and inhomogeneous, rather than fluid and homogeneous as in water. An obvious consequence of this rigidity is that the vdW repulsion may also be influenced by the

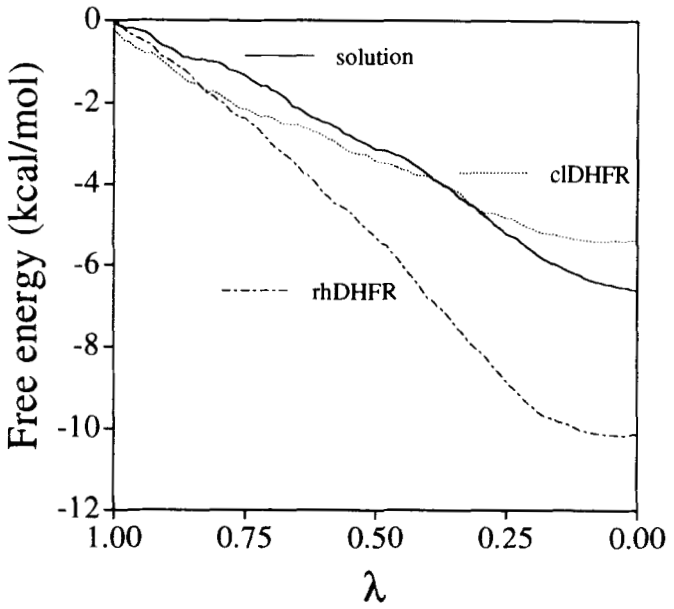

Figure 4 van der Waals components ($\varepsilon_i = 0$ kcal/mol) of the free energy change 6,8-dimethyl-N5-deazapterin → 6-methyl-N5-deazapterin as a function of the coupling parameter λ for ΔF_{bind} (where the ligand is bound to clDHFR and rhDHFR) and for ΔF_{solv}. A "window" corresponds to an increment of 0.01 in λ.

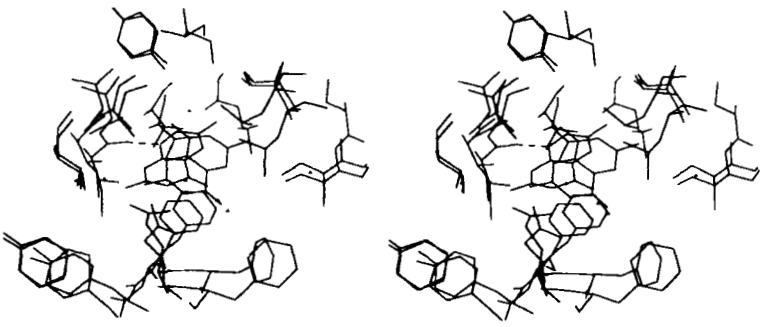

Figure 5 Least-squares superposition of 6,8-dimethyl-N5-deazapterin and 6-methyl-N5-deazapterin complexes of rhDHFR obtained from MD simulation ($\varepsilon_i = 0$ kcal/mol). The H-bond between the carbonyl oxygen of Ile-7 and H(N8) of the ligand (H atoms not shown) is indicated by a broken line. Positions of water molecules are indicated by the small triangles.

packing arrangement of the ligand and enzyme in the active site. An example of the effect this difference may have on free energies is illustrated graphically in Figure 4. During the initial stage of the mutation the vdW components of ligand-binding (ΔF_{bind}) to rhDHFR and clDHFR are quite similar, but after about 30 windows ($\lambda \sim 0.7$) the two rates of free energy change diverge abruptly. In fact, the rate of free energy change in clDHFR falls below that in solution, ΔF_{solv}. As the vdW repulsive

Figure 6 Least-squares superposition of 6,8-dimethyl-N5-deazapterin and 6-methyl-N5-deazapterin complexes of clDHFR obtained from MD simulation ($\varepsilon_i = 0$ kcal/mol). The H-bond between the carbonyl oxygen of Ile-7 and H(N8) of the ligand (H atoms not shown) is indicated by a broken line. Positions of water molecules are indicated by the small triangles.

energy arises from steric interactions, it would appear that the clDHFR simulation takes a different path and begins sampling configurations in which the overlap of CH_3 with the protein is much reduced. It is evident from Figures 5 and 6 that quite large structural changes take place in the active site during the mutation, and consequently significant variations in sampling for the free energy calculation may also be expected. The least-squares superposition of structures in Figure 5 (clDHFR) and Figure 6 (rhDHFR) show that the ligand and solvent-exposed side chains of the protein have orientations that change significantly between the final and initial states. Figure 7 shows the RMS deviations (Å) of the rhDHFR($\varepsilon_i = 0$) C_α coordinates from their initial X-ray values obtained after a least-squares superposition (note that the non-dynamic

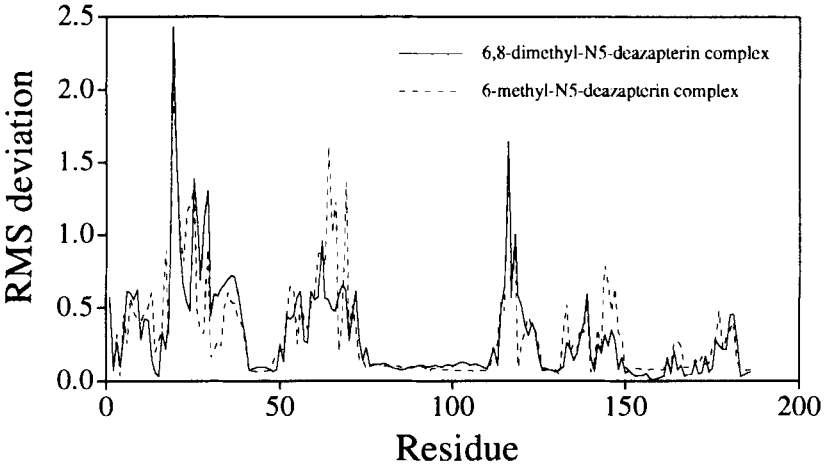

Figure 7 RMS deviations (Å) of the MD simulation C_α (DHFR backbone) coordinates from their initial X-ray values. The MD simulation coordinates were taken from the rhDHFR ($\varepsilon_i = 0$) calculation and the C_α coordinates least-squares fitted to the rhDHFR X-ray coordinates.

residues show effectively zero deviation). Some structural variations are also apparent between the DHFR backbones of the 6-methyl-N5-deazapterin and 6,8-dimethyl-N5-deazapterin complexes.

Relative Thermodynamic Stabilities and Hydrogen Bonding Based on the 80 ps vdW Mutation

In the initial state the 8-methyl compound is bound. During the course of the FEP mutation, the 8-methyl group is transformed into a hydrogen and the distance between the protein backbone and the H(N8)-position of the ligand becomes progressively shorter. The resulting H-bonds between H(N8) of 6-methyl-N5-deazapterin and a carbonyl oxygen of the protein backbone obtained at the end of a simulation can be seen in Figures 5 and 6. The complete H-bonding pattern between the 6-methyl-N5-deazapterin ligand and enzyme is illustrated schematically in Figure 8. The H-bond interaction of H(N3) and H(N2) with the carboxylic acid side chain of Asp (bacterial) or Glu (vertebrate) is conserved for both ligands in all the DHFRs. Various H-bonding geometries between water molecules and the carboxylate-ligand system are also evident in Figures 5 and 6. In the majority of structures obtained by MD simulation (typified by Figures 5 and 6), the carbonyl oxygens of Leu-4 in lcDHFR, Ile-5 in ecDHFR, Ile-7 in clDHFR and rhDHFR are H-bonded to H(N8) of 6-methyl-N5-deazapterin. In order to see the possible effect of the H-bonding at H(N8) on the relative thermodynamic stabilities of 6-methyl-N5-deazapterin and 6,8-dimethyl-N5-deazapterin bound to DHFR, it would be useful to compare $\Delta\Delta F_{bind}$ with the H-bond strength. For the purposes of such a comparison, the H-bond distances may be taken as a suitable indicator of the H-bond strength.

Figure 9 shows the correlation between thermodynamic stability ($\Delta\Delta F_{bind}$ results from Table 2) and the N8-oxygen distance in the 6-methyl-N5-deazapterin complexes. This distance has a value of 2.7 Å in the X-ray structure of the rhDHFR.N5-deazafolate complex [22], from which ligand protonation and H-bonding are inferred, compared

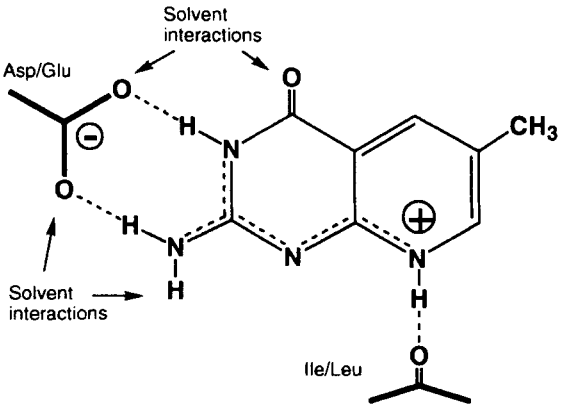

Figure 8 Schematic of the H-bonding between 6-methyl-N5-deazapterin and DHFR obtained in the MD/FEP simulations. Arrowed atoms are also involved in H-bonding with solvent water molecules.

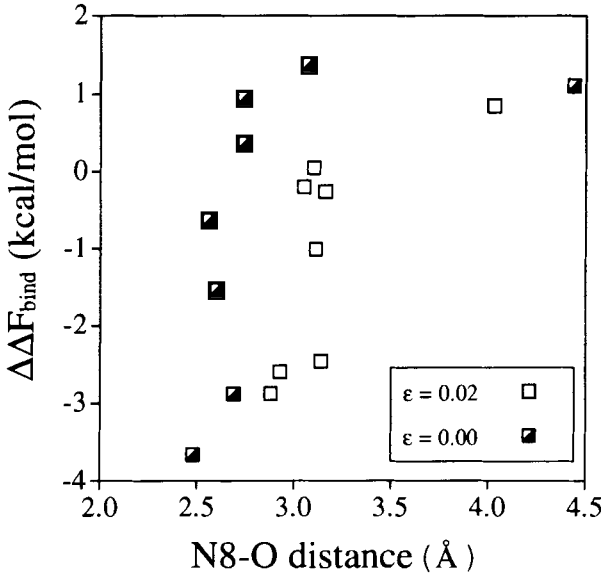

Figure 9 Correlation between the calculated $\Delta\Delta F_{bind}$ values (Table 2) and the N8-O distances (O = carbonyl oxygen of protein backbone, Figure 8) in the 6-methyl-N5-deazapterin MD complexes.

with a distance of 3.4 Å in the ecDHFR.N5-deazafolate X-ray structure [23] suggesting binding of unprotonated ligand. As Figure 9 shows, although a range of distances is obtained from the simulations, the majority lie between 2.5 and 3.2 Å, which may be regarded as the H-bonded range. These distances for $\varepsilon_i = 0$ are within ± 0.2 Å of the experimentally observed distance (i.e. 2.7 Å) in the rhDHFR.N5-deazafolate complex, whereas the parameter $\varepsilon_i = 0.02$ kcal/mol yields H-bonds which are longer by 0.2 to 0.5 Å, consistent with previous experience that $\varepsilon_i > 0$ leads to an underestimation of H-bond strength [47]. The majority of simulations with N8-O distance within the H-bonded range yield $\Delta\Delta F_{bind} < 0$, suggesting that complexes involving H-bonded 6-methyl-N5-deazapterin are the thermodynamically more stable ones. Nevertheless, several of the H-bonded structures at the higher end of the H-bonded range are found to give $\Delta\Delta F_{bind} > 0$, namely lcDHFR ($\varepsilon_i = 0$), ecDHFR ($\varepsilon_i = 0$) and clDHFR ($\varepsilon_i = 0$), while clDHFR.NADPH ($\varepsilon_i = 0.02$) gives $\Delta\Delta F_{bind} \approx 0$. Two of the structures [clDHFR($\varepsilon_i = 0.02$) and clDHFR.NADPH ($\varepsilon_i = 0$)] are well outside the H-bonded range at 4.0 and 4.5 Å, respectively, i.e. distances which are typical of the N8-O separations found in the initial 6,8-dimethyl-N5-deazapterin complexes. Both of these structures correspond to $\Delta\Delta F_{bind} > 0$. It can be seen that the 6-methyl-N5-deazapterin structures in Figures 10 and 11 exhibit a higher degree of overlap with the initial 6,8-dimethyl-N5-deazapterin structures than those shown in Figures 5 and 6 where H-bonding at N8 takes place. Not surprisingly, therefore, H-bonding appears to involve quite a substantial displacement of the binding geometry from the initial non-H-bonded state. This fact is also reflected in the small difference (Table 3) between the free energies for the forward and reverse vdW mutations in the calculations on

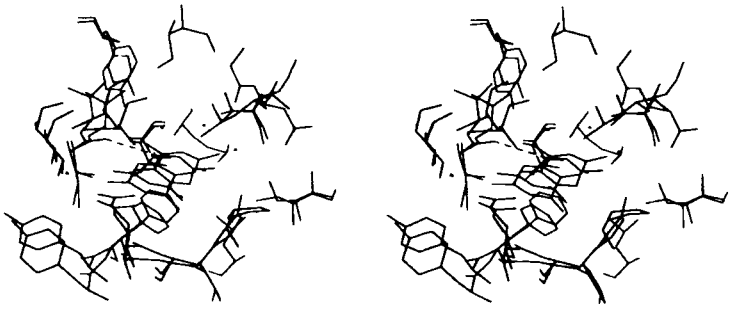

Figure 10 Least-squares superposition of 6,8-dimethyl-N5-deazapterin and 6-methyl-N5-deazapterin complexes of clDHFR obtained from MD simulation ($\varepsilon_i = 0.02$ kcal/mol). The interaction between the carbonyl oxygen of Ile-7 and H(N8) of the ligand (H atoms not shown) is indicated by broken lines. Positions of water molecules are indicated by the small triangles.

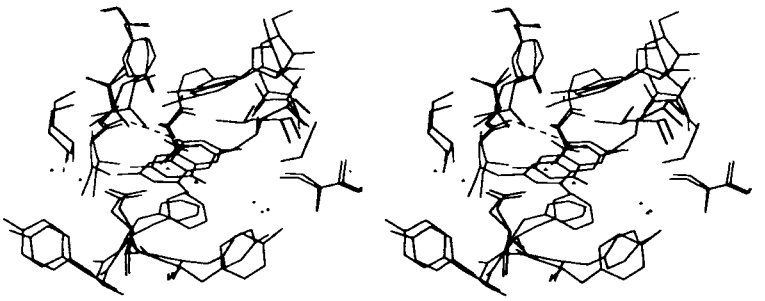

Figure 11 Least-squares superposition of 6,8-dimethyl-N5-deazapterin and 6-methyl-N5-deazapterin complexes of clDHFR.NADPH obtained from MD simulation ($\varepsilon_i = 0$ kcal/mol). The interaction between the carbonyl oxygen of Ile-7 and H(N8) of the ligand (H atoms not shown) is indicated by broken lines. Positions of water molecules are indicated by the small triangles.

clDHFR.NADPH ($\varepsilon_i = 0$) (Figure 11). As no large variations in structure occur during the vdW mutation (in either direction), the free energies are not expected to be very dependent on the mutation path, unlike the H-bonded structures which often give rise to a much larger hysteresis.

Free Energies and H-bonding from the 200 ps vdW Mutation

The vdW mutation was repeated for several complexes using the longer simulation time of 200 ps (500 MD steps equilibration, 500 MD steps data collection). The results are included in Tables 2 and 3 for comparison with the 80 ps simulations. There are only small (< 0.5 kcal/mol) differences between the 80 and 200 ps estimates of $\Delta\Delta F_{bind}$ for the two ecDHFR($\varepsilon = 0$) simulations. Thus, regardless of whether the simulation time for the vdW mutation is 80 ps or 200 ps, values for $\Delta\Delta F_{bind}$ differ by almost 2 kcal/mol for the two initial ecDHFR structures ($\varepsilon = 0$). A small (< 0.5 kcal/mol)

difference between the 80 and 200 ps simulations is also obtained for clDHFR.NADPH($\varepsilon = 0$). By contrast, on increasing the simulation time substantial changes in $\Delta\Delta F_{bind}$ (> 1 kcal/mol) are obtained for lcDHFR($\varepsilon = 0$), clDHFR($\varepsilon = 0.02$) and rhDHFR($\varepsilon = 0$). For lcDHFR the change on going from 80 ps to 200 ps is as high as 3 kcal/mol. The reverse 200 ps mutation for rhDHFR($\varepsilon = 0$) gives a hysteresis of 1.8 kcal/mol which is approximately the same absolute magnitude as for the corresponding 80 ps mutation (Table 3). These results indicate that the free energies are generally still not converged after 200 ps.

For the clDHFR.NADPH($\varepsilon_i = 0$) complex, the 200 ps simulation gave an MD structure with an N8-O distance of 2.75 Å compared with 4.44 Å obtained from the 80 ps simulation. The simulations performed for clDHFR($\varepsilon = 0.02$) yielded 3.21 Å for 200 ps and 4.03 Å for 80 ps. Thus, we observe that the H-bond has formed over a longer simulation time.

Comparison with Experiment

Dissociation constants (K_d) for binary complexes of 6,8-dimethyl-N5-deazapterin and 6-methyl-N5-deazapterin with human, chicken and *E. coli* DHFRs have been measured as a function of pH [16, 24]. These show, respectively, pH minima or limiting values at pH $\sim$ 6.6 and $\sim$ 4.2, which are comparable with the pK_a's of the compounds [19] and consistent with the proposed mechanism for their binding to DHFR in terms of ionization constants of the ligands and the active-site carboxylate group [13, 16]. K_d's for ternary complexes for 6,8-dimethyl-N5-deazapterin also show a minimum at pH $\sim$ 6.6 but with values 15 to 45 times lower: 16 and 1 μM for chicken and 31 and 0.7 μM for human DHFRs [16]. Using the proposed ligand-binding model and the measured binary K_d's, it is possible to estimate pH-independent constants (or the equivalent free energies) for the dissociation of the cation form of ligands from the enzymes [16, 49], i.e. analogous quantities as for the simulations. Unfortunately, K_d's for ternary complexes of 6-methyl-N5-deazapterin at pH 4.2 cannot be measured because NADPH is unstable at this low pH. Consequently, in Table 4 we compare the

Table 4 Experimental and theoretical estimates of the difference in binding free energy (kcal/mol) for the change 6,8-dimethyl-N5-deazapterin → 6-methyl-N5-deazapterin in binary complexes

Enzyme	*Experiment*[a]	*Theoretical*[b]
E. coli DHFR	1.4 ± 0.5	-0.70[c]
Chicken DHFR	0.0 ± 0.5	-0.43
Human DHFR	0.0 ± 0.5	-1.23

[a] Estimated difference based on measured dissociation constants corrected for pH dependence (see text): errors (± 0.5 kcal/mol) are due mostly to uncetainty in the binding of 6-methyl-N5-deazapterin.

[b] $\Delta\Delta F_{bind}$ values from 200 ps simulations in Table 2. See text for discussion of errors.

[c] Mean of 200 ps simulations for the two ecDHFR structures.

estimated experimental and FEP-calculated (selected from the 200 ps simulations in Table 2) differences in binding free energies for the binary complexes only. For 6-methyl-N5-deazapterin and 6,8-dimethyl-N5-deazapterin with human and chicken DHFRs the dissociation constants are all in a very narrow range, and a difference in binding free energy of more than ± 0.5 kcal/mol is unlikely. The FEP result for clDHFR only is within these experimental bounds. The experimental results for *E. coli* DHFR are somewhat different, suggesting that the 6,8-dimethyl-compound binds of the order of 10 times (~ 1.4 kcal/mol) more strongly than the 6-methyl-compound in the binary complex. The FEP result deviates from the experimental ecDHFR value by ~ 2 kcal/mol.

It is easy to see from the results in Table 2, that errors of 10 to 20% in the values of ΔF_{bind} and ΔF_{solv} would lead to errors of up to 2 kcal/mol in the $\Delta\Delta F_{bind}$ values. The values of $\Delta\Delta F_{bind}$ from the MD/FEP calculations (in Table 2) when averaged over all simulation times and species of DHFR yield -0.59 ± 1.59 kcal/mol for $\varepsilon = 0$ and -0.99 ± 1.26 kcal/mol for $\varepsilon = 0.02$ consistent with a binding differential within a factor of 10 (i.e. about 1.4 kcal/mol difference in free energy). In addition to the sampling errors, the average differences between the $\varepsilon = 0$ and $\varepsilon = 0.02$ results indicate that some uncertainty can also be expected to arise from the approximate nature of the force fields.

CONCLUSION

Molecular dynamics (MD) combined with free energy perturbation (FEP) methods have been used to study the relative affinities of 6-methyl-N5-deazapterin (N8 protonated) and 6,8-dimethyl-N5-deazapterin (N3 protonated) ligands for dihydrofolate reductase from bacterial (*L. casei* and *E. coli*), avian (chicken liver) and human sources with and without the presence of NADPH cofactor. Both the computed free energies and final structures from the MD simulations were found to vary considerably, depending on the source of the X-ray coordinates from which starting configurations are obtained, simulation time and description of the H-bond interaction. Even after 200 ps, for many of the complexes the free energies had not converged sufficiently to make a prediction of the relative thermodynamic stabilities accurate to within 2 kcal/mol. Errors of less than 10% in the absolute values of the relative solvation and binding free energies (ΔF_{solv} and ΔF_{bind}) would be required to achieve an uncertainty of less than 1 kcal/mol. As relatively long simulation times have been used compared with analogous work in the literature [4–10, 14, 35, 45], a requirement for even longer simulation times would indicate that, in this particular case at least, conventional simulation techniques based on Boltzmann sampling is an inefficient way to predict precise values of relative thermodynamic stabilities. However, despite these uncertainties, an enzyme-binding differential between 6-methyl-N5-deazapterin and 6,8-dimethyl-N5-deazapterin cations of not more than 10 would be predicted, in broad agreement with estimates from experimental dissociation constants.

The two different sets of initial X-ray coordinates for *E. coli* (ecDHFR) that were used to model the binary complex gave quite different free energy changes. While these differences may be reduced over even longer simulation times, the results bring into

question the reliability of the molecular modelling process whereby the initial protein structure is assumed to be close to a known X-ray structure of a binary or ternary complex in which a structurally-similar ligand is bound. The very extensive structural information now available for DHFR complexes has demonstrated significant protein flexibility, including more recently dependence on crystal form [50]. Even in this atypical modelling case where many sets of coordinates are available it is not possible to predict *a priori* which complex structure might provide the most appropriate starting structure for a new test ligand. Consequently, we strongly recommend calculating free energies using coordinates from more than one X-ray determination, if available, in order to check the validity of the results. The results also show that the nature of the interactions and force-field should be carefully considered in the error assessment. The molecular process studied here, i.e. formation of an H-bond during the mutations, is quite a demanding test of simulation protocols with respect to both sampling and force-fields, and to this extent our experience may be regarded as a "worst case".

Although configuration-space sampling problems in the MD/FEP simulations make the quantitative prediction of free energy differences and the rationalization of observed species dependencies difficult, the structures obtained by MD simulation exhibited H-bond formation between H(N8) of the protonated 6-methyl-N5-deazapterin ligand and the carbonyl oxygen of a conserved Ile or Leu residue in the active site of DHFR. While most of the H-bonds formed over the 80 ps simulations for the vdW mutations, a small number of simulations resulted in only incomplete H-bond formation. However, on increasing the simulation time to 200 ps, these latter structures also exhibited H-bonding. The H-bond formation between H(N8) and a carbonyl oxygen of the protein backbone obtained in the MD simulations is consistent with that observed in the X-ray structure of N5-deazafolate bound to human DHFR in which the ligand is deduced to be protonated on N8.

Acknowledgements

Financial support from the National Health and Medical Research Council (NH&MRC), and equipment grants from the Clive and Vera Ramaciotti and Utah Foundations are gratefully acknowledged. We also acknowledge Australian Supercomputer Technology for generous allocations of time on their FACOM VP2200 supercomputer. We thank M.T.G. Ivery for supplying the experimental dissociation constant data and V. Reyes for a preprint of reference 23.

References

[1] B. L. Tembe and J. A. McCammon, "Ligand-receptor interactions", *Comput. Chem.*, **8**, 281–283 (1984).

[2] R. L. Blakley, "Dihydrofolate reductase", in *Folates and Pterins, Chemistry and Biochemistry of Folates*, Vol. 1, R. L. Blakley and S. J. Benkovic, eds., Wiley, New York, 1984, pp. 191–253.

[3] J. Kraut and D. A. Matthews, "Dihydrofolate reductase", in *Biological Macromolecules and Assemblies*, Vol. 3, F. A. Jurnak and A. McPherson, eds., Wiley, New York, 1987, pp. 1–72.

[4] U. C. Singh, 'Probing the salt bridge in the dihydrofolate reductase-methotrexate complex by using the coordinate-coupled free-energy perturbation method", *Proc. Natl. Acad. Sci. USA*, **85**, 4280–4284 (1988).

[5] U. C. Singh and S. J. Benkovic, "A free energy perturbation study of the binding of methotrexate to dihydrofolate reductase", *Proc. Natl. Acad. Sci. USA*, **85**, 9519–9523 (1988).

[6] C. L. Brooks, "Thermodynamic calculations on biological molecules", *Int. J. Quantum Chem. Quantum Biol. Symp.*, **15**, 221–234 (1988).
[7] C. L. Brooks and S. H. Fleischman, "A theoretical approach to drug design 1. Rêlative solvation thermodynamics for the antibacterial compound trimethoprim and ethyl derivatives substituted at the 3',4' and 5' positions", *J. Am. Chem. Soc.*, **112**, 3307–3312 (1990).
[8] S. H. Fleischman and C. L. Brooks, "Protein-drug interactions: Characterization of inhibitor binding in complexes of DHFR with trimethoprim and related derivatives", *Proteins: Struct. Funct. Genet.*, **7**, 52–61 (1990).
[9] J. J. McDonald and C. L. Brooks, "Theoretical approach to drug design 2. Relative thermodynamics of inhibitor binding by chicken dihydrofolate reductase to ethyl derivatives of trimethoprim substituted at 3',4'- and 5'-positions", *J. Am. Chem. Soc.*, **113**, 2295–2301 (1991).
[10] P. R. Gerber, A. E. Mark and W. F. van Gunsteren, "An approximate but efficient method to calculate free energy trends by computer simulation: Application to dihydrofolate reductase – inhibitor complexes", *J. Comput. -Aided Mol. Design*, **7**, 305–323 (1993).
[11] V. Thibault, M. J. Koen and J. E. Gready, "Enzymic properties of a new mechanism-based substrate for dihydrofolate reductase", *Biochemistry*, **28**, 6042–6049 (1989).
[12] J. E. Gready, "Design of new mechanism-based substrates for dihydrofolate reductase", in *Chemisty and Biology of Pteridines 1989*, H.-Ch. Curtius, S. Ghisla and N. Blau, eds, de Gruyter, Berlin, 1990, pp. 23–30.
[13] J. E. Gready, P. L. Cummins and P. Wormell, "Computer-aided design of mechanism-based pterin analogues and MD/FEP simulations of their binding to dihydrofolate reductase", in *Advances in Experimental Medicine and Biology* Vol. 338 (Chemistry and Biology of Pteridines and Folates), J. E. Ayling, M. G. Nair and C. M. Baugh, eds, Plenum, New York, 1993, pp. 487–492.
[14] P. L. Cummins and J. E. Gready, "Novel mechanism-based substrates of dihydrofolate reductase and the thermodynamics of ligand binding: A comparison of theory and experiment for 8-methylpterin and 6,8-dimethylpterin", *Proteins: Struct. Funct. Genet.*, **15**, 426–435 (1993).
[15] M. T. G. Ivery and J. E. Gready, "Structure-activity relationships for the 8-alkylpterins: a new class of mechanism-based substrates for dihydrofolate reductase (DHFR)", *Biochemistry*, **34**, 3724–3733 (1995).
[16] M. T. G. Ivery and J. E. Gready, "Structure-activity relationships and pH dependence of binding of 8-alkyl-N5-deazapterins to dihydrofolate reductase", *J. Med. Chem.*, **37**, 4211–4221 (1994).
[17] S. R. Stone, J. A. Montgomery and J. F. Morrison, "Inhibition of dihydrofolate reductase from bacterial and vertebrate sources by folate, aminopterin, methotrexate and their 5-deaza analogues", *Biochem. Pharmacol.*, **33**, 175–179 (1984).
[18] P. Wormell and J. E. Gready, "Electronic spectra of some pterins and deazapterins", *Chem. Phys.*, **179**, 55–69 (1994).
[19] M. T. G. Ivery and J. E. Gready, "An improved procedure for the preparation of 2-amino-8-alkylpyrido-[2,3-d]pyrimidin-4(3*H*)-ones (8-alkyl-N5-deazapterins)", *J. Heterocyc. Chem.*, **31**, 1385–1397 (1994).
[20] M. L. Williams and J. E. Gready, "Guanidinium-type resonance stabilization and its biological implications 1. The guanidinium and extended-guanidinium series", *J. Comput. Chem.*, **10**, 35–54 (1989).
[21] M. J. Jordan and J. E. Gready, "Guanidinium-type resonance stabilizatoin and its biological implications 2. The doubly-extended-guanidinium series", *J. Comput. Chem.*, **10**, 186–202 (1989).
[22] J. F. Davies, T. J. Delcamp, N. J. Prendergast, V. A. Ashford, J. H. Freisheim and J. Kraut, "Crystal structures of recombinant dihydrofolate reductase complexed with folate and 5-deazafolate, *Biochemistry*, **29**, 9467–9479 (1990).
[23] V. M. Reyes, M. R. Sawaya and J. Kraut, "Isomorphous crystal structures of *Escherichia coli* dihydrofolate reductase complexed with folate, 5-deazafolate and 5,10-dideazatetrahydrofolate: Mechanistic implications", *Biochemistry*, **34**, 2710–2723 (1995).
[24] M. T. G. Ivery and J. E. Gready, unpublished.
[25] W. F. van Gunsteren, "Methods for calculations of free energies and binding constants: Successes and problems", in *Computer Simulation of Biomolecular Systems*, W. F. van Gunsteren and P. K. Weiner, eds, ESCOM, Leiden, 1989, pp. 27–59.
[26] D. A. Pearlman and P. A. Kollman, "Free energy perturbation calculations: Problems and pitfalls along the gilded road", in *Computer Simulation of Biomolecular Systems*, W. F. van Gunsteren and P. K. Weiner, eds, ESCOM, Leiden, 1989, pp. 101–119.
[27] C. A. Reynolds, P. M. King and W. G. Richards, "Free energy calculations in molecular biophysics", *Mol. Phys.*, **76**, 251–275 (1992).
[28] W. F. van Gunsteren, "Molecular dynamics studies of proteins", *Curr. Opin. Struct. Biol.*, **3**, 277–281 (1993).

[29] T. J. Bolin, D. J. Filman, D. A. Matthews, R. C. Hamlin and J. Kraut, "Crystal structures of *Escherichia coli* and *Lactobacillus casei* refined at 1.7 angstroms resolution. I. General features and binding of methotrexate", *J. Biol. Chem.*, **257**, 13650–13662 (1982).

[30] E. E. Howell, J. E. Villafranca, M. S. Warren, S. J. Oatley and J. Kraut, "Functional role of aspartic acid-27 in dihydrofolate reductase revealed by mutagenesis", *Science*, **231**, 1123–1128 (1986).

[31] C. Bystroff, S. J. Oatley and J. Kraut, "Crystal structure of *Escherichia coli* dihydrofolate reductase: The $NADP^+$ holoenzyme and the folate.$NADP^+$ ternary complex. Substrate binding and a model for the transition state", *Biochemistry*, **29**, 3263–3277 (1990).

[32] D. A. Matthews, T. J. Bolin, J. M. Burridge, D. J. Filman, K. N. Volz, B. T. Kaufman, C. R. Beddell, J. N. Champness, D. K. Stammers and J. Kraut, "Refined crystal structure of *Escherichia coli* and chicken liver dihydrofolate reductase containing bound trimethoprim", *J. Biol. Chem.*, **260**, 381–391 (1985).

[33] M. A. McTigue, J. F. Davies, B. T. Kaufman and J. Kraut, "Crystal structure of chicken liver dihydrofolate reductase complexed with $NADP^+$ and biopterin", *Biochemistry*, **31**, 7264–7273 (1992).

[34] C. Oefner, A. D'Arcy and F. K. Winkler, "Crystal structure of human dihydrofolate reductase complexed with folate", *Eur. J. Biochem.*, **174**, 377–385 (1988).

[35] P. L. Cummins and J. E. Gready, "Computer-aided drug design: A free energy perturbation study on the binding of methyl-substituted pterins and N5-deazapterins to dihydrofolate reductase", *J. Comput.-Aided Mol. Design*, **7**, 535–555 (1993).

[36] U. C. Singh, P. K. Weiner, J. W. Caldwell and P. A. Kollman, AMBER (Version 3.2), Dept. Pharamaceutical Chem., Univ. of California, San Francisco, 1988.

[37] U. C. Singh, F. K. Brown, P. A. Bash and P. A. Kollman, "An approach to the application of free energy perturbation methods using molecular dynamics", *J. Am. Chem. Soc.*, **109**, 1607–1614 (1987).

[38] W. F. van Gunsteren and H. J. C. Berendsen, "Algorithms for macromolecular dynamics and constraint dynamics", *Mol. Phys.*, **34**, 1311–1327 (1977).

[39] P. L. Cummins and J. E. Gready, "The electrostatic potential in the semiempirical molecular orbital approximation" *Chem. Phys. Lett.*, **225**, 11–17 (1994).

[40] S. J. Weiner, P. A. Kollman, D. T. Nguyen and D. A Case, "An all atom force field for simulations of proteins and nucleic acids", *J. Comput. Chem.*, **7**, 230–252 (1986).

[41] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy and J. J. P. Stewart, "AM1: A new general purpose quantum mechanical model", *J. Am. Chem. Soc.*, **107**, 3902–3909 (1985).

[42] M. J. S. Dewar and G. P. Ford, "Ground states of molecules. 44. MINDO/3 calculations of absolute heat capacities and entropies of molecules without internal rotations", *J. Am. Chem. Soc.*, **99**, 7822–7829 (1977).

[43] B. G. Rao and U. C. Singh, "Hydrophobic hydration: A free energy perturbation study", *J. Am. Chem. Soc.*, **111**, 3125–3133 (1989).

[44] P. L. Cummins and J. E. Gready, "Solvent effects in active-site molecular dynamics simulations on the binding of 8-methyl-N5-deazapterin and 8-methyl-pterin to dihydrofolate reductase", submitted to *J. Comput. Chem.*

[45] P. L. Cummins, K. Ramnarayan, U. C. Singh and J. E. Gready, "Molecular dynamics/free energy perturbation study on the relative affinities for the binding of reduced and oxidized NADP to dihydrofolate reductase", *J. Am. Chem. Soc.*, **113**, 8247–8256 (1991).

[46] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, "Comparison of simple potential functions for simulating liquid water", *J. Chem. Phys.*, **79**, 926–935 (1983).

[47] A. T. Hagler, J. R. Maple, T. S. Thacher, G. B. Fitzgerald and U. Dinur, "Potential energy functions for organic and biomolecular systems", in *Computer Simulation of Biomolecular Systems*, W. F. van Gunsteren and P. K. Weiner, eds, ESCOM, Leiden, 1989, pp. 149–167.

[48] C. L. Brooks, A. Brunger and M. Karplus, "Active site dynamics in protein molecules: A stochastic boundary molecular-dynamic approach", *Biopolymers*, **24**, 843–865 (1985).

[49] S. R. Stone and J. F. Morrison, "The interaction of an ionizing ligand with enzymes having a single ionizing group. Implications for the reaction of folate analogues with dihydrofolate reductase", *Biochim. Biophys. Acta*, **745**, 237–246 (1983).

[50] M. R. Sawaya and J. Kraut, personal communication.