Journal of Computer-Aided Molecular Design, **12:** 119–131, 1998.
KLUWER/ESCOM

Computation of affinity and selectivity: Binding of 2,4-diaminopteridine and 2,4-diaminoquinazoline inhibitors to dihydrofolate reductases

John Marelius[a], Malin Graffner-Nordberg[b], Tomas Hansson[a], Anders Hallberg[b] & Johan Åqvist[a,*]
[a]*Department of Molecular Biology, Uppsala University, Biomedical Centre, Box 590, S-751 24 Uppsala, Sweden*
[b]*Department of Organic Pharmaceutical Chemistry, Uppsala University, Biomedical Centre, Box 574, S-751 23 Uppsala, Sweden*

Received 27 June 1997; Accepted 20 October 1997

Key words: binding free energy, drug design, molecular dynamics

Summary

Binding energy calculations for complexes of mutant and wild-type human dihydrofolate reductases with 2,4-diaminopteridine and 2,4-diaminoquinazoline inhibitors are reported. Quantitative insight into binding energetics of these molecules is obtained from calculations based on force field energy evaluation and thermal sampling by molecular dynamics simulations. The calculated affinity of methotrexate for wild-type and mutant enzymes is reasonably well reproduced. Truncation of the methotrexate glutamate tail results in a loss of affinity by several orders of magnitude. No major difference in binding strength is predicted between the pteridines and the quinazolines, while the N-methyl group present in methotrexate appears to confer significantly stronger binding. The recent improvement, which is used here, of our linear interaction energy method for binding affinity prediction, as well as problems with treating charged and flexible ligands are discussed. This approach should be suitable in a drug discovery context for prediction of binding energies of new inhibitors prior to their synthesis, when some information about the binding mode is available.

Introduction

Access to powerful methods for prediction of binding constants for ligands to target receptors prior to organic synthesis is highly desirable in the drug discovery process. Such predictions can be valuable for lead generation and for lead optimisation and serve to highlight areas in which to focus chemical synthesis efforts. A computational approach to estimation of binding constants (or, equivalently, free energies of binding) can be particularly fruitful when the structure of the receptor is known. Predictions of the affinity of a given inhibitor for different enzyme variants and mutants may be equally valuable.

Various computational methods for estimating ligand binding affinity have been devised, each representing a different choice of computational demand versus accuracy [1]. Free energy perturbation (FEP) theory [2–4] combined with conformational sampling by molecular dynamics (MD) or Monte Carlo (MC) simulation can be used to calculate free energy changes upon small modifications of the ligand or receptor [5, 6] However, FEP calculations become quite complicated and computationally expensive for structurally dissimilar inhibitors and for calculation of absolute free energies of binding. In the latter case, precautions must also be taken to ensure a correct standard state when molecules are annihilated or created [7]. Scoring functions applied to single conformations of the docked complex is a more empirical approach to affinity prediction. They are generally based on identifying individual points of intermolecular interaction such as hydrogen bonds, ionic interactions and hydrophobic interactions, as well as entropy estimates, in a given conformation of the receptor-ligand complex and as-

* To whom correspondence should be addressed.

signing a binding energy score to each contributing factor [8–11]. These methods include estimates of the effect of solvation, either implicitly by their parametrisation to fit a set of experimental values or explicitly by measuring the change in the solvent-accessible surface area upon binding. There are also examples of utilising scaled molecular mechanics energies for minimised structures to obtain binding energy estimates [12].

In the present work we use a different computational method for binding affinity prediction based on linear relationships between average molecular mechanics energies and binding free energies as described below. As part of an interdisciplinary drug discovery project we wish to examine whether this methodology is adequate for quantitative analysis of inhibitors of human dihydrofolate reductase (DHFR). Analogues of methotrexate (MTX, **3** in Figure 1, a classic DHFR inhibitor) are particularly challenging from a computer simulation viewpoint since they are flexible and have three ionic sites. We also wish to evaluate the potential of the present method for predicting the effects of enzyme mutations on the affinity for a given inhibitor. Calculations of absolute binding free energies are reported for three non-classical DHFR inhibitors to the wild-type enzyme and for methotrexate to two different models of the wild-type and for two mutant enzymes. To investigate the sensitivity of the method to perturbations of the starting model two different crystal structures of the enzyme are used. The set of inhibitors studied here allows us to assess, e.g., the contribution from the MTX glutamate tail as well as the difference between a pteridine (**4**) and a quinazoline (**5**) in the absence of this glutamate tail and also the difference between an inhibitor with an N-methyl group (**5**) as in MTX and one without it (**6**) as in the enzyme's substrate (**1**).

Dihydrofolate reductase (DHFR, E.C. 1.5.1.3) catalyses the nicotine adenine dinucleotide phosphate (NADPH) dependent reduction of dihydrofolate **1** (Figure 1) to produce the co-factor tetrahydrofolate **2** that is needed for reductive one-carbon transfer reactions. These reactions are essential for the biosynthesis of nucleic acids. Dihydrofolate reductases are well established targets for antibacterial (e.g., trimethoprim), antifungal (e.g., piritrexim, trimetrexate), antiprotozoal (e.g., pyrimethamine) and antineoplastic (e.g., methotrexate, **3**) agents. Many inhibitors of DHFR exploit structural differences between enzymes from different classes of organisms and fair to good selectivity is often achieved. Furthermore, methotrexate has recently been introduced in the treatment of rheumatoid arthritis [13–15].

The LIE method

The linear interaction energy (LIE) method introduced by Åqvist et al. in 1994 [16] for calculating absolute free energies of binding is a semi-empirical method which is computationally less expensive than FEP. It is still similar to FEP in that it is based on thermodynamic cycles and includes thermal conformational sampling. However, the method involves only simulations of the two physical states (bound and free ligand) and requires no transformation processes. It has been described in detail elsewhere [17–20] but we will present an outline here, followed by a description of a recent extension of the method [21].

The main idea is to consider contributions from polar and non-polar (all non-electrostatic) interactions to the total binding energy separately. The polar part can be treated using the electrostatic linear response approximation while the non-polar contribution must be calculated using an empirical formula calibrated against a set of experimental binding data. The approach is schematically depicted in Figure 2 where the two vertical legs of the cycle denote the free energies associated with 'turning on' the electrostatic interactions between the ligand and its surroundings in the bound and free states, respectively. The binding free energy in this cycle corresponds to the difference between these two 'electrostatic legs' plus a non-polar contribution represented by the lower horizontal leg.

The electrostatic free energies (vertical legs) are given in the linear response approximation by the simple equation

$$\Delta G_i^{polar} = \frac{1}{2}(\langle \Delta V_{on-off} \rangle_{off,i} + \langle \Delta V_{on-off} \rangle_{on,i}), \quad (1)$$

where $\langle \Delta V_{on-off} \rangle$, the thermal average of the difference between the potentials for the 'on' and 'off' states, reduces to $\langle V_{l-s}^{el} \rangle$ if the 'electrostatic legs' only involve turning on or off the *inter*molecular ligand-surrounding electrostatic interactions (V_{l-s}^{el}). This yields

$$\Delta G_i^{polar} = \frac{1}{2}(\langle V_{l-s}^{el} \rangle_{off,i} + \langle V_{l-s}^{el} \rangle_{on,i}). \quad (2)$$

In the above equations, the notations $\langle \cdot \rangle_{off,i}$ and $\langle \cdot \rangle_{on,i}$ represent thermal averages sampled with these intermolecular electrostatic interactions off and on,

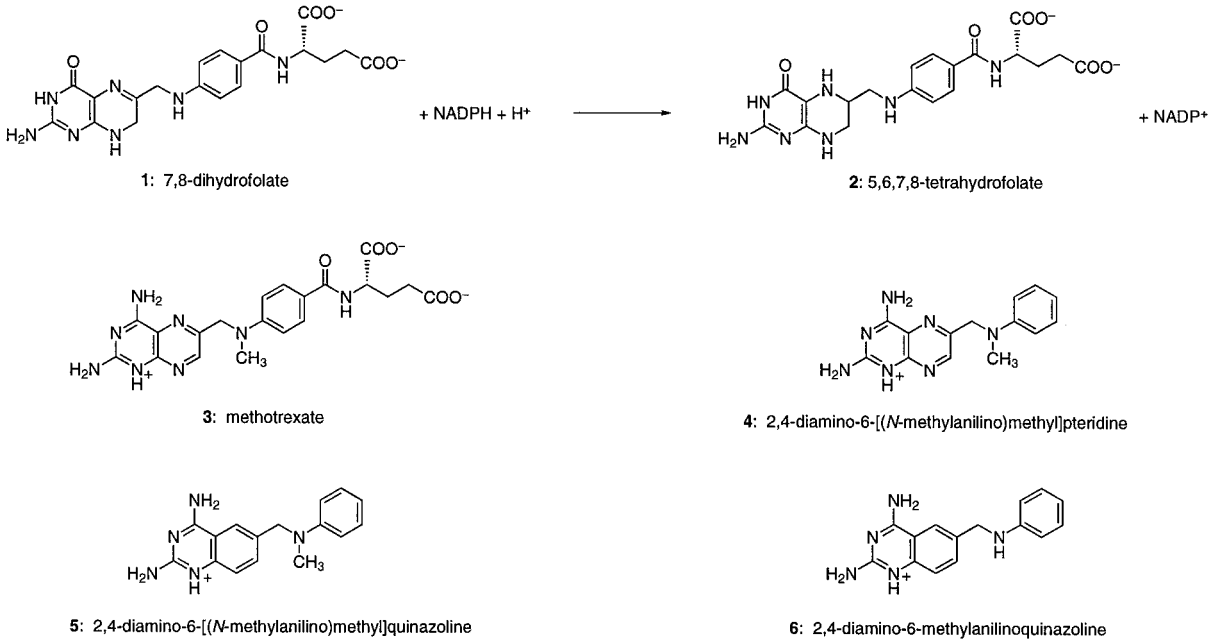

Figure 1. Substrate, product and the studied inhibitors of DHFR. The folates **1** and **2** are depicted in their anticipated protonation state at physiological pH, whereas the inhibitors **3**–**6** are shown in the actual protonation states used in the dynamics simulations.

respectively, and *i* simply denotes which leg in Figure 2 is considered (bound or free). Note that, even if the sampling potential has V^{el}_{l-s} turned off, it is possible to calculate what V^{el}_{l-s} would be for the generated configurations. Not surprisingly, this average turns out to be close to zero in most cases [22] and additionally, the binding cycle of Figure 2 involves the difference of two such averages. Neglecting the first term of Equation 2 and taking the difference between the two electrostatic legs gives the electrostatic binding contribution

$$\Delta\Delta G^{polar} = \beta(\langle V^{el}_{l-s}\rangle_{bound} - \langle V^{el}_{l-s}\rangle_{free}), \qquad (3)$$

where $\beta = \frac{1}{2}$ according to the linear response approximation. Note here that *intra*molecular interactions (both in the ligand and surroundings) are then also assumed to obey linear response with respect to V^{el}_{l-s} and are thus implicitly taken into account.

The simple form of Equations 1 and 2 is the consequence of assuming a quadratic dependence of the free energy on the intermolecular electrostatic interaction strength. That is, the potentials of mean force as a function of the collective polarisation of the system are parabolas with equal curvature for the '*on*' and '*off*' states. The derivative of the potential of mean force is then linear, whence the term linear response. This is the standard assumption in continuum macroscopic electrostatic theory, where polarisation is usually said to be proportional to the applied electric field with a given dielectric constant. For other recent applications of linear response approaches in microscopic simulations, see [23, 24].

In order to estimate the non-polar binding contribution (horizontal leg in Figure 2) it may first be noted that 'non-polar solvation energies', or more precisely solvation energies of non-polar compounds, scale linearly with molecular size measures such as surface area, volume or chain length [25, 26]. Furthermore, it has been found in computer simulations that average intermolecular solute-solvent van der Waals (Lennard-Jones) energies also scale linearly with such size measures [16, 27]. Taken together, these two observations suggest that it should be possible to estimate a non-polar binding contribution in analogy with the electrostatic treatment above:

$$\Delta\Delta G^{non\text{-}polar} = \alpha(\langle V^{vdW}_{l-s}\rangle_{bound} - \langle V^{vdW}_{l-s}\rangle_{free}). \qquad (4)$$

This leads to the final estimate for the total binding energy

$$\Delta G_{bind} = \alpha(\langle V^{vdW}_{l-s}\rangle_{bound} - \langle V^{vdW}_{l-s}\rangle_{free}) + \beta(\langle V^{el}_{l-s}\rangle_{bound} - \langle V^{el}_{l-s}\rangle_{free}). \qquad (5)$$

The implication of Equation 4 is not that van der Waals interactions are the origin of the non-polar

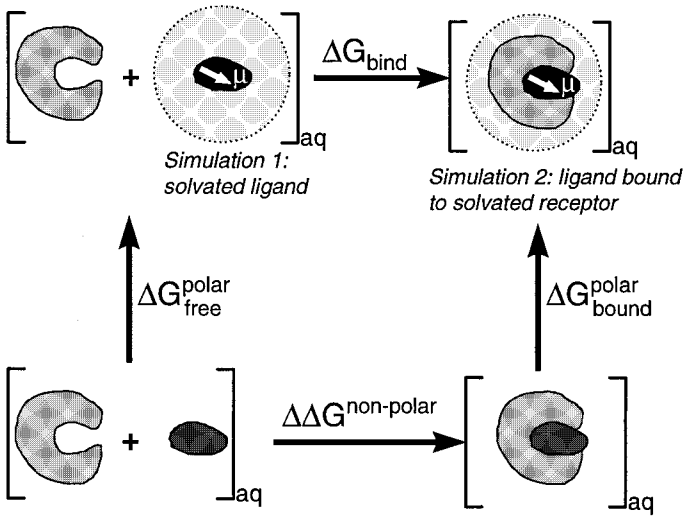

Figure 2. The thermodynamic cycle used in the LIE method. The binding free energy ΔG_{bind} is given by the sum of a non-polar contribution $\Delta\Delta G^{non\text{-}polar}$ and a polar contribution $\Delta\Delta G^{polar} = \Delta G^{polar}_{bound} - \Delta G^{polar}_{free}$. The arrows marked μ indicate that intermolecular electrostatic ligand-surrounding interactions are present. The actual simulation systems are denoted by dotted circles.

binding contribution, but merely reflects their correlation with non-polar solvation energies. Furthermore, because the value of α is determined from calibration against experimental data, it may also implicitly take into account other possible contributions (force field errors, systematic entropy terms, concentration effects, etc.).

Originally the parameter α was calibrated against experimental binding data for a set of endothiapepsin inhibitors, giving $\alpha = 0.16$ utilising the GROMOS force field [16]. With $\beta = \frac{1}{2}$ in Equation 5 the method thus relies on the validity of the linear response approximation. This model has been used in several studies that yielded reasonable binding estimates for different systems [17–20]. A more detailed study of the validity of Equation 2 for a number of compounds in water (altogether 55 different systems) revealed that deviations from electrostatic linear response do occur, particularly for dipolar compounds containing hydroxyl groups [22]. The calculations involved direct comparison of the left- and right-hand sides of Equation 2, utilising FEP simulations. These results suggest that the value $\beta = \frac{1}{2}$, while valid for ionic compounds, should be reduced for neutral dipolar ones. Since these deviations appear to vary systematically with the composition of the solute, a revised model that takes them into account has recently been presented [21]. The revised model still has only one free parameter α, but four different values of β are used according to a simple classification of the solute (compounds with ionic groups: $\beta = \frac{1}{2}$, dipolar without hydroxyl groups: $\beta = 0.43$, dipolar with one hydroxyl group: $\beta = 0.37$, dipolar with more than one hydroxyl group: $\beta = 0.33$). Note that $\beta = \frac{1}{2}$ is used for all compounds with ionic groups (even if they contain hydroxyls) since the electrostatics is then dominated by their strong interaction and linear response is obeyed [22].

With these fixed β's, α was optimised using experimental data for 18 receptor-ligand complexes yielding $\alpha = 0.181$ and an average unsigned binding energy error of 0.6 kcal/mol [21]. In [21] the effects of relaxing the constraint of equal scaling factors (α and β) in the bound and free states as well as the effects of introducing a constant term in the LIE expression were also investigated by optimising the parameters of the more general LIE equation

$$\Delta G_{bind} = \alpha_{prot}\langle V^{vdW}_{l-s}\rangle_{bound} - \alpha_{wat}\langle V^{vdW}_{l-s}\rangle_{free} + \beta_{prot}\langle V^{el}_{l-s}\rangle_{bound} - \beta_{wat}\langle V^{el}_{l-s}\rangle_{free} + \gamma \qquad (6)$$

under various combinations of constraints such as $\alpha_{prot} = \alpha_{wat}$, $\beta_{prot} = \beta_{wat}$ and $\gamma = 0$. The improvement of the fit to experimental data by introducing another degree of freedom in the model (adding a free parameter) was weighed against the increased complexity of the model. The model with four fixed β's and α as the single free parameter was found to give the smallest errors in leave-one-out cross-validation tests. When γ was a free parameter, its optimised value was always very close to zero, indicating that constant con-

tributions to the solvation energies in water and protein tend to cancel.

Hence, the model we employ in the present work is that with four fixed β's, but since all the analysed inhibitors are charged ($\beta = \frac{1}{2}$ according to [21]) the only difference compared to the original model [16] is in the value of α (now 0.181).

Computational models and methods

Our simulation models of human dihydrofolate reductases are based on the crystal structure by Lewis et al. of the ternary complex of the mutant leucine 22 to tyrosine (L22Y) with MTX and NADPH [28, Protein Data Bank entry 1DLS] and on the structure by Oefner et al. of the complex of the wild-type enzyme with folate [29, entry 1DRF]. Using the InsightII software [30], a model of the wild-type derived from the L22Y mutant and a model of the L22F mutant were constructed. With these models we can study both the sensitivity of the binding energy estimates to small differences in the starting model and the precision in predicting the structure and affinity for a given inhibitor of computer-modelled mutant enzymes. The initial geometries of all the inhibitors are derived from the MTX co-ordinates in the mutant structure (1DLS). Docking of NADPH, which is known to be important for the binding affinity of inhibitors [28], and MTX into the other wild-type enzyme model (1DRF) was done manually with InsightII and involved slight modifications in the conformation of some side chains near the binding sites, guided by their conformation in the 1DLS complex. Docking of **4**, **5** and **6** was achieved by simply overlaying them on the portion of MTX they have in common in the 1DLS structure.

Our simulation systems for the bound state are spheres containing the inhibitor placed in the binding site of the solvated enzyme. The simulation of the free state is carried out in a sphere of equal size filled with water molecules and with the inhibitor at the centre.

For finite simulation systems of this type, it is important that constant free energy contributions from the outside cancel between the bound and free states. Particularly, the systems must be equivalent with respect to the free energy associated with polarisation of the medium outside the simulation spheres. This energy is given in the continuum approximation by Born's formula $\Delta G_{\mathrm{Born}} = -k \cdot \frac{Q^2}{2r} \cdot (1 - \frac{1}{\epsilon})$, where Q is the net charge of the system in an interaction sphere of radius *r*, outside of which the dielectric constant is ε and *k* is the constant in Coulomb's law. This means that the two simulated systems should be of equal size and have equal net charge. Either, counter ions must be added to one of the systems (which has been shown earlier to slow down convergence [20]), or excess charge on the protein must be removed [18, 20, 31]. In any case, it is necessary to neutralise charged groups near the boundary of the simulation sphere which will otherwise be unscreened and introduce vacuum-like fields into the system. To fulfil these requirements, a number of charged residues far from the ligand and near the boundary of the simulation sphere were modelled using the alternate set of GROMOS residues with neutral dipolar groups [32] rather than the standard set [33]. For charged ligands, this procedure means that we neglect the Coulomb interactions between the ligand and the neutralised charges of the protein. Since all these groups are far (≥ 15 Å) from the ligand, the energy can be approximated by Coulomb's law (Equation 7) with a high dielectric constant characteristic of long-range electrostatic interactions in solvated proteins [34–37].

$$\Delta\Delta G_{\mathrm{el.corr}} = k \cdot \sum_{\substack{p \in \text{neglected ionic sites,} \\ l \in \text{ligand atoms}}} \frac{q_p \cdot q_l}{\epsilon \cdot r_{p-1}} \qquad (7)$$

Here, q_p is the formal (integer) charge of the neglected ionic group, q_l is the partial charge of the ligand atom, r_{p-l} is the distance between the ligand atom and a central atom in the ionic group (e.g., NZ in lysine), k is the constant in Coulomb's law and ε is the dielectric constant. We calculated this correction with $\epsilon = 80$ using the initial co-ordinates (the sensitivity to this procedure has been addressed earlier [20]). The simulated models of DHFR based on 1DLS were charged at D21, R28, E30, K55, E62, R65, K68, R70, H87 and D145, whereas the model based on 1DRF was charged at D21, E30, K55, E62, K63, R65, K68 and R70. Together with the charge on NADPH (see below), this adds up to zero. With these choices, all charged protein (or co-factor) groups within 10 Å of the charged moieties of the inhibitor were included (with the exception of R32 in 1DLS which is somewhat closer but solvent exposed). The slight difference between the models was introduced to see whether the estimated free energies of binding are sensitive to changes in the modelling of charged groups of the enzyme. All other formal charges were thus neglected in the simulations and instead taken into account by Equation 7.

The inhibitors were modelled as protonated at N1 according to NMR studies of MTX bound to DHFR by Cocco et al. [38, 39] in both the bound and free states.

However, MTX (**3**), which has a pK_a of 5.73 [38], and the second 2,4-diaminopteridine inhibitor (**4**), which is expected to have a similar pK_a, are not protonated at N1 in water at pH 7.65 where the experimental binding constants were measured [28]. The free energies estimated in our calculations are for the reaction from protonated free inhibitor to protonated bound inhibitor and thus need to be corrected by adding the positive free energy for protonation of the free ligand. The correction is given by

$$\Delta\Delta G_{pK} = k_B T \cdot \ln 10 \cdot (pH_{exp} - pK_a) = 1.36\ \text{kcal/mol} \cdot (pH_{exp} - pK_a), \quad (8)$$

where k_B is Boltzmann's constant. This is also consistent with experiments showing stronger inhibition by MTX at pH 6.0 than at pH 7.5 [40, 41]. In applying Equation 8, a pK_a of 5.73 was used for both the 2,4-diaminopteridine inhibitors **3** and **4**. The pK_a values of the 2,4-diaminoquinazolines **5** and **6** were estimated to be in the range of 7 to 8, based on data in [42] and [43]. Since these pK_a's are critical, we decided to measure the pK_a of a substituted 2,4-diaminoquinazoline using potentiometric titration [44, 45]. This was done using a Sirius PCA-101 titrator (Sirius Analytical Instruments Ltd, East Sussex, England) equipped with a Ross type combination glass electrode calibrated for mixed solvent titrations. The mixed solvent approach was employed because of the limited aqueous solubility of the compound across the pH range. Methanol was used as cosolvent and the ionic strength was adjusted with 0.15 M KCl. Three separate titrations were performed with different water/methanol ratios to obtain pK_a's in the presence of methanol (psK_a values). Aqueous pK_a values were calculated by extrapolation to 0% methanol using the Yasuda-Shedlovsky relationship [44, 45]: a linear plot of $psK_a + \log[H_2O]$ versus $1/\epsilon$, where ϵ is the dielectric constant of the water/methanol mixture. This yielded a p*K*$_a$ value of 7.75 and hence no pH correction was applied to the 2,4-diaminoquinazolines.

Hydrogen atoms on the aromatic rings of the inhibitors were included but aliphatic methyl and methylene groups were replaced by 'extended' carbon atoms. The partial charges on the phenyl CH dipoles were set to ± 0.14. These parameters were found to reproduce the free energy of hydration of benzene in FEP calculations better than the standard extended-atom GROMOS phenyl groups. The partial charges for the other portions of the inhibitors were derived from semi-empirical quantum mechanics calculations with the AM1/SM2 Hamiltonian [46]. The values were adjusted to be compatible with the GROMOS force field by scaling the QM charges to match the values for groups defined in the force field library (e.g., the glutamate residue which is part of MTX) and by slight modifications to achieve electroneutral groups of less than about 10 atoms. The partial charges of NADPH were also determined as above. The simulation model carried a charge of -2 on the pyrophosphate moiety, but the adenosine 2′-phosphate was replaced by a corresponding dipolar group since it was far (>17 Å) from the inhibitor and close to the boundary of the simulation sphere.

Molecular dynamics simulations of spherical systems of radius 18.5 Å were carried out with a modified [47, 16] version of the GROMOS force field [33] with a time step of 1 fs using a modified version (Åqvist, unpublished) of the program Enzymix [48]. In the 'free' simulations the inhibitors were surrounded by about 870 SPC [49] water molecules, giving a total of ca. 2600 particles, and held centred in the sphere by harmonic positional restraints on three of the atoms of the pteridine ring. This was done since overall rotation of the ligand in water was found to slow down convergence of the ligand-surrounding interaction energies. Restraining of the ionic groups was also attempted but gave, not surprisingly, significantly different interaction energies. In the 'bound' simulations, comprising about 3100 particles, some 900 protein atoms far away from the active site extended out of the simulation sphere and were held fixed. The position of the centre of the sphere was chosen to minimise the distances to all inhibitor atoms and to charged groups of the enzyme. The carboxylate groups of MTX were initially 5–9 Å from the surface of the sphere. The volume of the sphere not occupied by protein or ligand was filled by about 380 water molecules. In both the 'bound' and 'free' simulations, the water molecules were confined to the sphere by a half-harmonic radial restraining potential and their polarisation restrained according to the SCAAS model [50]. Long-range electrostatic interactions (beyond 10 Å) were calculated using the local reaction field method [51] which entails multipole expansion of the long-range electrostatic potentials. The temperature was maintained at 300 K by coupling ($\tau = 10$ fs) to a temperature bath [52].

The enzyme-inhibitor-water complexes were equilibrated using a procedure of slow heating during which the protein atoms were restrained to their initial co-ordinates by a 5 kcal/(mol·Å^2) harmonic potential. The first 1 ps of dynamics was run at 1 K with very strong coupling ($\tau = 0.2$ fs) to the temperature bath

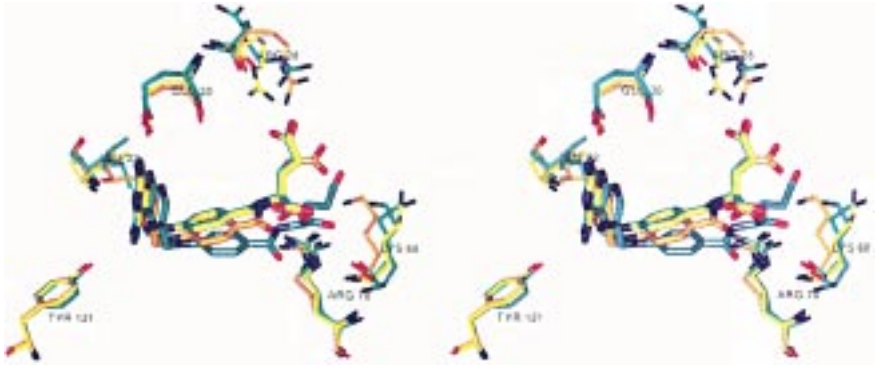

Figure 3. Stereo view of two models of DHFR with bound methotrexate (MTX) before and during MD simulation. MTX and seven interacting residues of the enzyme are shown. Yellow: initial structure of the wild-type enzyme built from the 1DLS structure of the L22Y mutant. Orange: average structure of the same 1DLS-derived complex computed from the trajectory between 475 and 500 ps. Light green: initial structure from the 1DRF structure of the wild-type enzyme into which MTX was docked. Dark green: average structure (as above) of the 1DRF-derived complex. Nitrogen and oxygen atoms have been coloured blue and red, respectively. The 2,4-diaminopteridine ring is held tightly in place by the enzyme during the simulations. The carboxylate groups of MTX interact with several positively charged protein side chains. Q30 interacts with the positively charged pteridine ring. L22, lining the binding pocket, is included to show the location of the studied mutations.

(very similar to steepest descent energy minimisation), followed by 6 ps at 50 K, 2 ps at 150 K and 7 ps at 300 K. After this the restraints were relaxed. Throughout the free state simulations, the inhibitors were held in the centre of the sphere by 5 kcal/(mol·Å^2) harmonic restraints at three points of the aniline group, thus avoiding overall rotation. The initial conformations of the inhibitors were the same in the bound and free simulations.

Energy data from the equilibration phase of the simulations was not included in the interaction energy averages. The start of the production phase (from which energy data were included in averages) was determined after the simulation so that the average ligand-surrounding electrostatic interaction energies in the first and second half of the time series after that point were equal, i.e., the energy had reached a stable long-time average value. If many such points followed by data that could be split in two consecutive blocks of equal length with equal mean could be found, the earliest one was chosen. The standard error of the average interaction energies was calculated according to the method presented by Allen and Tildesley [53] based on the concept of statistical inefficiency. This method is a general scheme to compensate for correlated data in a time series by estimating the fraction of the data that can be considered independent. The procedure involves partitioning the data into consecutive blocks, taking averages for each block and calculating the variance of these averages. The calculation is repeated for progressively larger and larger blocks and finally extrapolated to infinite block length. The statistical inefficiency s is defined by

$$s = \lim_{t_b \to \infty} \frac{t_b \cdot \mathrm{Var}(\bar{E}_b)}{\mathrm{Var}(E)} \tag{9}$$

where E is the series of sampled ligand-surrounding electrostatic interaction energies and $\bar{E}_b$ is the series of block average values. Given the statistical inefficiency s of the data set, it is easy to calculate the standard error as

$$\Delta\langle E\rangle = \sqrt{\frac{\mathrm{Var}(E)}{N^*}} \tag{10}$$

where $N^* = N/s$ is the number of independent data points and N is the total number.

Results

Molecular dynamics simulations of the enzyme and inhibitor models described above were performed and the required ligand-surrounding interaction energies were measured and averaged for LIE binding calculations.

Simulations of methotrexate (MTX) bound to the 1DRF model of the wild-type enzyme, to the 1DLS-derived wild-type model and L22F mutant model, as well as the corresponding simulation of free MTX in water were carried out. Furthermore, simulations of **4**, **5** and **6** bound to the 1DLS-derived wild-type model and free in water were performed. Following

restrained heating of the systems, all these simulations extended for 500 ps, but some were prolonged to a total of 750 ps, 1000 ps or, for MTX in the free state, to 1500 ps (see Table 1). From the resulting trajectories it is possible to extract stable long-term energy averages with reasonable precision, given this challenging situation with charged and flexible molecules. These long trajectories allow studies of the variations of the interaction energies on the hundreds-of-ps time-scale, in order to assess the dependence of the calculated binding energies on the length of the simulated trajectory. Particularly, the prolonged trajectories allow us to gauge the expected necessary trajectory lengths for accurate binding energy predictions in future simulations.

These simulations yielded good agreement between simulated and crystallographic structures of the complexes as shown in Figure 3. In a comparison between the initial structures and average structures computed on the trajectory between 475 and 500 ps, rms co-ordinate deviations were calculated for all heavy (non-hydrogen) atoms within 8 Å of any inhibitor atom, totally 1100 atoms. This yielded a value of 1.01 Å for the complex of MTX and NADPH with the 1DLS-derived wild-type DHFR model and a value of 1.86 Å for the MTX+NADPH+DHFR complex derived from 1DRF. The corresponding rms deviations for only the protein backbone atoms were 0.51 Å and 1.22 Å, respectively. The mutant L22Y complex which was taken directly from the 1DLS PDB entry gave values of 0.93 Å (heavy) and 0.47 Å (backbone). As a reference, the rms deviations between the same atom sets of the initial co-ordinates of the two different wild-type enzyme models are 1.00 Å (heavy) and 0.48 Å (backbone). Thus, the deviations that arise during the long simulations are of the same order as the differences between two crystal structures of the same enzyme but with different ligands (see below). Furthermore, comparing different structures along the simulation trajectories, we find that significant differences between consecutive 25 ps average structures are limited to some solvent-exposed polar residues, while the residues lining the substrate binding pocket, as well as the inhibitor itself, maintain their relatively stable positions throughout the simulation. While large-scale motion of the protein is eliminated by restraining the part of the protein outside the simulation sphere, the good agreement still indicates that our simulations can be expected to reproduce interactions within the simulation sphere in a realistic fashion, i.e., no unnatural distortive forces arise due to the specific assumptions applied in modelling these complexes.

In the enzyme models derived from the 1DLS ternary complex where NADPH is included, NADPH moves only very slightly around its initial position, whereas in the 1DRF-derived model into which NADPH was docked a minor reorganisation of the protein to accommodate the new ligand and a small (1 Å) movement of the nicotine amide portion of the co-factor can be seen. We make an analogous observation near the substrate binding site where the loop formed by residues 55–68 moves slightly together with the inhibitor in the 1DRF model, while in the 1DLS-derived wild-type model this loop remains very close to its initial position.

Table 1 shows the resulting ligand-surrounding interaction energy averages. It is apparent that for all these complexes with potent inhibitors, the interaction energies in the bound state are more negative than the corresponding values for the free state, indicating that both electrostatic and non-polar interactions will contribute to the free energy of binding as calculated according to Equation 5. The part of the trajectory used for averaging was determined according to the procedure described above, so that any trend in the data was eliminated. The standard errors of the average energies were calculated using the method of Allen and Tildesley based on the concept of statistical inefficiency [53], as outlined above. The Coulombic free energy correction (Equation 7 using the starting co-ordinates) for neglected charges in the simulation models of the enzymes and the free energy correction related to the protonation states of the inhibitor in the bound and free state (Equation 8) are also given in Table 1.

The electrostatic interaction energy between methotrexate and its surroundings is very large due to its three ionic groups. Furthermore, its instantaneous value fluctuates within about -500 ± 100 kcal/mol. This is attributable to the great conformational flexibility especially of the glutamic acid moiety which carries two of the three net charges, as illustrated in Figure 4. Therefore, the task of calculating the quantity $\Delta\Delta G^{polar}$ according to Equation 3 to a precision of a few kcal/mol becomes a very challenging one, mainly due to statistical considerations. In contrast, **4**, **5** and **6**, where the net charge is smaller and located on a non-flexible part of the molecules, yield smaller and much more stable interaction energy averages. Disregarding the differences in flexibility, one would expect the standard error of the electrostatic

Table 1. Average ligand-surrounding interaction energies from simulations

Enzyme	Inhibitor		Total simulation time (ps)	Time used for averaging (ps)[a]	$\langle E^{el} \rangle$ (kcal/mol)[b]	$\langle E^{vdw} \rangle$ (kcal/mol)[c]	Electrostatic correction (kcal/mol)[d]	pH correction (kcal/mol)[e]
Human DHFR 1DRF	**3** (MTX)	bound	1000	739	–498.9±4.7	–52.3±0.7	–0.9	2.6
		free	1500	1126	–478.5±3.4	–21.1±0.2		
Human DHFR modelled from 1DLS	**3** (MTX)	bound	750	658	–499.7±5.0	–52.7±0.4	–0.6	2.6
		free	1500	1126	–478.5±3.4	–21.1±0.2		
Human DHFR L22Y mutant 1DLS	**3** (MTX)	bound	500	450	–493.8±4.4	–54.1±0.4	–0.6	2.6
		free	1500	1126	–478.5±3.4	–21.1±0.2		
Human DHFR L22F mutant modelled from 1DLS	**3** (MTX)	bound	750	318	–492.2±2.0	–53.9±0.2	–0.8	2.6
		free	1500	1126	–478.5±3.4	–21.1±0.2		
Human DHFR modelled from 1DLS	**4**	bound	750	424	–108.4±0.4	–47.0±0.3	+0.1	2.3
		free	500	316	–99.2±0.5	–25.1±0.2		
Human DHFR modelled from 1DLS	**5**	bound	500	500	–102.4±0.3	–47.5±0.3	+0.1	0
		free	500	500	–98.6±0.3	–24.7±0.1		
Human DHFR modelled from 1DLS	**6**	bound	500	422	–106.3±0.5	–42.6±0.1	+0.1	0
		free	500	500	–105.7±0.3	–23.4±0.1		

[a] Determined to eliminate trends in the data due to non-equilibrium conditions.
[b] Electrostatic (Coulomb) ligand-surrounding interaction energy.
[c] van der Waals (Lennard-Jones 6–12 potential) ligand-surrounding interaction energy.
[d] Free energy contribution from neglected charges on the protein calculated according to Equation 7.
[e] Free energy of protonation of the inhibitor in water at the pH where the experimental binding constant was determined.

interaction energy to scale linearly with its absolute magnitude. This makes charged ligands problematic in that their electrostatic energies will always be larger than those for ligands with no ionic sites. Furthermore, it turns out that even with very long trajectories, the attainable relative standard errors are larger (around 1%) for MTX than for the smaller inhibitors (around 0.5%), presumably caused by the combined effects of conformational flexibility and net charge.

Free energies of binding were calculated from the energy averages listed in Table 1, according to Equation 5 where the coefficient α was 0.181 and β was $\frac{1}{2}$, and corrected according to Table 1. The resulting values together with the corresponding experimentally determined values are listed in Table 2. The error bars given are the square root of the sum of the squared average errors given in Table 1.

The calculated free energies of binding for MTX to the wild-type enzyme, –14.1 kcal/mol and –14.2 kcal/mol (Table 2) for the 1DRF and 1DLS models, respectively, indicate that in this case the results of LIE binding calculations are not so sensitive to differences in the structure of the initial complex and to slight variations in the modelling of charged groups of the protein. The calculations turn out to reproduce the large absolute free energy of binding, $\Delta G_{expt.} = -15.3$ kcal/mol, fairly accurately. Encouragingly, of the different LIE models discussed in [21] the one found to be the best for the calibration set (the

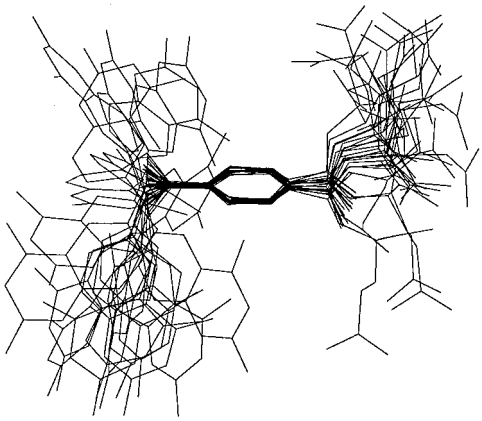

Figure 4. Conformations of methotrexate in water. This flexible molecule explores large regions of conformational space causing significant energy fluctuations during the simulation. The picture shows 20 snapshots from a 1500 ps dynamics trajectory. The central aniline moiety is restrained at three points to eliminate overall rotation and translation.

Table 2. Binding free energies calculated with the LIE method

Enzyme	Inhibitor	Calculated ΔG_{bind} (kcal/mol)[a]	Experimental ΔG_{bind} (kcal/mol)	Ref.
Human DHFR 1DRF	**3** (MTX)	–14.1±2.9	–15.3	28, 29
Human DHFR modelled from 1DLS	**3** (MTX)	–14.2±3.0	–15.3	28
Human DHFR L22Y mutant 1DLS	**3** (MTX)	–11.6±2.8	–11.2	28
Human DHFR L22F mutant modelled from 1DLS	**3** (MTX)	–11.0±2.0	–12.7	28
Human DHFR modelled from 1DLS	**4**	–6.2±0.3	–7.4[b]	55
Human DHFR modelled from 1DLS	**5**	–5.9±0.2	no data	
Human DHFR modelled from 1DLS	**6**	–3.6±0.3	no data	

[a] Calculated from average inhibitor-surrounding interaction energies in Table 1 and corrected for neglected charges of the protein and for protonation state of the inhibitors (see Table 1).
[b] Uncertain value based on IC_{50} measurement on pigeon liver DHFR.

model used herein) also turns out to yield the smallest errors for the DHFR inhibitors.

For methotrexate, electrostatic interactions of both the pteridine ring and the tail carboxylate groups are important for strong binding. In particular, the ion pairs between Q30 and the pteridine ring and R70 and the α-carboxylate are conserved throughout the simulations in both models of the enzyme. These two residues also appear to be the only entirely conserved charged ones in the active site region among different DHFR sequences [54]. K68, although initially not in direct contact with MTX, finds a stable position close to both carboxylates in both models. Initial and average structures of the two complexes are shown in Figure 3.

The effects of the mutations L22Y and L22F on MTX binding are also reasonably well reproduced, although the relative binding free energy between the two mutants is not correctly predicted. The standard error of the calculated relative free energy of binding (2.4 kcal/mol) is determined only from the bound simulations. This error overshadows the binding free energy difference (0.6 kcal/mol) between the mutants, for which the selectivity result in this case is inconclusive. However, the significantly stronger binding to the native enzyme is correctly predicted.

The calculated statistical errors for the smaller inhibitors are very small. At this stage only one experimental value is available, but the results clearly confirm the important effect of truncating the glutamate tail. The 2.3 kcal/mol difference between **5** and **6**, which differ only with respect to the N-methyl group on **5**, is notable. This methyl group fits into a protein cavity and appears to thereby render **5** less flexible. As can be seen in Table 1, the largest difference appears in the electrostatic interaction energies. In contrast, compound **6** is more flexible in the binding site due to the absence of the methyl group. This affects both the quinazoline moiety, which moves nearly 1 Å from its initial position, and the conformation of the phenyl ring. Compounds **4** and **5** differ mainly in their electrostatic interactions with the protein, with **4** giving the most negative average value. The binding free energies are still very similar due to the difference in the pK_a values between the 2,4-diaminopteridine and 2,4-diaminoquinazoline inhibitors. That is, the protonated form of **5** is more stable than that of **4** at physiological pH. For the smaller inhibitors **4**, **5** and **6** that lack carboxylate groups, the binding is attributable by similar amounts to electrostatic and non-polar interactions. All these retain more translational and conformational freedom in the bound state than MTX.

Discussion

In this work, the main objective has been to examine whether the LIE approach is useful for ligand design towards dihydrofolate reductases. Methotrexate-like inhibitors of this enzyme pose some rather difficult problems for force-field based estimation of absolute binding energies. Here, several ionic groups, both in the ligand, enzyme and co-factor, give rise to large electrostatic interaction energies on the order of several hundred kcal/mol. In addition, ligands like methotrexate are flexible and explore a wide range of conformations in solution under normal thermal sampling conditions (Figure 4). Taken together, these

circumstances lead to slow convergence of especially V^{el}_{l-s}. Nevertheless, we have shown that stable results can be attained with reasonable computational efforts. The calculated binding data for the different inhibitors and mutant enzymes in Table 2 are in rather good agreement with available experiments.

It should perhaps be emphasised here that other force-field approaches to binding calculations such as FEP do not appear to have reached this level of accuracy for absolute binding energies for the type of complexes studied here. Furthermore, unlike some other force-field based scoring methods, our procedure has not been specifically parametrised for the system under study but seems fairly general. It is of course always possible to optimise the method for a particular system by reparametrisation on a relevant training set. The method yields results of similar quality with the new parametrisation [21] for systems like endothiapepsin [16], HIV-1 proteinase [20], glucose binding protein (GBP) [17] and trypsin [18]. A recent test on K^+ binding to 18-crown-6 gave a binding energy of -2.0 kcal/mol, also in fair agreement with experiment (-2.7 kcal/mol) [J.Å., J.M., and T.H., unpublished results]. The same type of methodology as that used herein has been successfully applied by other groups as well to deazapterin DHFR inhibitors with aliphatic substituents [56], hydration energies of organic compounds [27], cytochrome P450cam ligands [57] and human thrombin inhibitors [58].

A drawback of the present approach compared to empirical scoring functions is, of course, the larger computational demand. The main reasons for this are, first, complicated energy function evaluations (many interactions) and, second, thermal sampling of many configurations. This means that for exhaustive conformational searches, such as required in *ab initio* docking, our approach cannot compete with empirical scoring methods. However, for tasks like exploration of chemical structures around a lead compound (or similar) when some 3D structural information is available, the method has proven useful. In the latter type of application, at present it is not clear precisely which available scoring method has the best predictive power. The long computation times required for accurate results with our method may be worthwhile since the evaluation of many interactions in many configurations takes some potentially important physical properties into account. For instance, if binding of a certain ligand is significantly influenced by long-range interactions (and not just atom-atom contacts) this might be difficult to elicit by a scoring function that does not take these into account. Furthermore, the binding free energy is not determined by a single conformation but indeed by a thermal average. Thus, if conformational fluctuations are important for the binding energy in a given system, our approach tries to capture this effect. Moreover, the explicit treatment of water molecules, while again time-consuming, may in some cases offer a more realistic picture of the binding mode. For example, in the case of glucose binding to GBP, three water molecules provide very specific and strong interactions with the ligand in its binding site. Methods that have been parametrised without explicit treatment of water molecules will thus erroneously attribute such binding contributions to something else.

The reported accuracy of empirical scoring functions for estimating binding affinity for receptor-ligand complexes of known structure is quite impressive [8–11]. In general, however, the exact conformation of the complex is unknown and can only be estimated by manual or automated docking procedures. The accuracy of empirical scoring functions under these circumstances is more difficult to assess since they are generally parametrised exclusively on crystal structures of tight-binding complexes. It seems particularly difficult to parametrise a scoring function to correctly take unfavourable interactions into account, since there will be few such examples in the calibration set. This problem has been addressed by Jain by construction of a differentiable scoring function [11] that can be used (as part of a docking procedure [59]) to optimise the score by moving the ligand along the score gradient. The LIE approach avoids the bias inherent in a calibration set composed only of minimum energy structures and can thus be expected to be more resilient. Also, even when the initial conformation suffers from minor inaccuracies, the dynamics simulation allows the system to converge to a relaxed equilibrium structure, as the 1DRF simulation illustrates. In this context, is it worth noting that, e.g., the LUDI scoring function [8] gives a 2.8 kcal/mol binding energy difference for MTX between the two starting structures used here.

In the case of MTX, the large electrostatic interaction energies and their fluctuations give rise to a considerable uncertainty in the binding energy predictions. Whether the present method used to estimate the error is the best remains to be seen. In earlier studies we have, e.g., used the difference between the first and second halves of the trajectory as the error measure, which is now by definition zero. Nevertheless, the convergence errors are likely to be the main cause of

the erroneous prediction of the L22Y vs. L22F relative affinity for MTX. For such a small mutation one might expect standard FEP calculations to give reliable results. Regarding the applicability of the LIE procedure to mutations of the receptor rather than the ligand, there is so far little data to judge from (see also [57]). In principle, if the linear response approach with its current parametrisation is valid for differing receptor sites, then studies of protein mutants is just a special case. In particular, the linear response assumption means that all other interactions will always respond linearly to the ligand-surrounding ones and are thus implicitly included. This means that intramolecular interactions in the ligand and in the surroundings as well as intermolecular interactions in the surroundings, e.g., differences in solvation/desolvation of receptor sites, do not enter explicitly. A simple example of this is ion hydration where water-water interactions need not be measured to calculate the free energy according to linear response [22]. However, as said earlier [16] it may well be the case that receptor sites of different nature require different parametrisation both due to deviations from linear response [22] and the 'hydrophobic contribution' embedded in the α term. It is of course conceivable that even a point mutation could cause such a change in the nature of the binding site.

The present calculations show that both the 2,4-diaminopteridine ring and the glutamic acid tail are important for binding of MTX. The compounds lacking the tail are predicted to bind 10^5–10^6-fold weaker, in agreement with experimental data for **4**. The residual binding affinity of the double ring system with only an attached N-methylaniline moiety does not appear particularly sensitive to heteroatom substitution on the second ring, as indicated by the small difference between the 2,4-diaminopteridine and 2,4-diaminoquinazoline compounds **4** and **5**. As noted above, the difference between **5** and **6** appears to be that the N-methyl group of compound **5** contributes to fixing its conformation in a more favourable manner by steric hindrance caused by the positioning of this group in a cavity. A similar steric effect appears to be responsible for the reduced affinity of L22Y and L22F for MTX, where the corresponding phenyl rings of residue 22 are not able to fill the space normally occupied by the leucine side chain. This leucine side chain also makes contact with the pteridine ring.

Conclusions

In summary, we have presented binding affinity calculations for complexes of mutant and wild-type human DHFR with 2,4-diaminopteridine and 2,4-diaminoquinazoline inhibitors. These charged and flexible molecules are difficult to treat in microscopic simulations aimed at obtaining stable energy averages. However, the calculation methodology used here, which is based on force field energy evaluation and MD thermal sampling, has been shown to still yield convergent results with reasonable computational efforts. In combination with an improvement of our linear interaction energy method for binding energy prediction [21] the present approach shows promise for the ability to explore novel compound classes, at least when some information about the binding mode is available. This is of particular interest in a drug discovery context for accurate prediction of binding energies of new inhibitor candidate molecules.

Acknowledgements

We wish to thank one referee for insightful comments regarding R28. We gratefully acknowledge support from the Swedish Research Council for Engineering Sciences (TFR), the Swedish Natural Science Research Council (NFR) and the Swedish Foundation for Strategic Research (SSF).

References

1. Ajay and Murcko, M.A., J. Med. Chem., 38 (1995) 4953.
2. Kollman, P., Chem. Rev., 1993 (1993) 2395.
3. Jorgensen, W.L., Acc. Chem. Res., 22 (1989) 184.
4. Beveridge, D.L. and DiCapua, F.M., Annu. Rev. Biophys. Biophys. Chem., 18 (1989) 439.
5. Singh, U.C., Proc. Natl. Acad. Sci. USA, 85 (1988) 4280.
6. Singh, U.C. and Benkovic, S.J., Proc. Natl. Acad. Sci. USA, 85 (1988) 9519.
7. McCammon, J.A., Gilson, K., Given, J.A. and Bush, B.L., Biophys. J., 72 (1997) 1047.
8. Böhm, H.-J., J. Comput.-Aided Mol. Design, 8 (1994) 243.
9. Wallqvist, A., Jernigan, R.L. and Covell, D.G., Protein Sci., 4 (1995) 1881.
10. Head, R.D., Smythe, M.L., Oprea, T.L., Waller, C.L., Greene, S. and Marshall, G.R., J. Am. Chem. Soc., 118 (1996) 3959.
11. Jain, A.N., J. Comput.-Aided Mol. Design, 10 (1996) 427.
12. Holloway, M.K., Wai, J.M., Halgren, T.A., Fitzgerald, P.M., Vacca, J.P., Dorsey, B.D., Levin, R.B., Thompson, W.J., Chen, L.J., deSolms, S.J., Gaffin, N., Ghosh, A.K., Giuliani, E.A., Graham, S.L., Guare, J.P., Hungate, R.W., Lyle, T.A., Sanders, W.M., Tucker, T.J., Wiggins, M., Wiscount, C.M., Woltersdorf, O.W., Young, S.D., Darke, P.L. and Zugay, J.A., J. Med. Chem., 38 (1995) 305.

13. DeGraw, J.I., Colwall, W.T., Crase, J., Lane Smith, R., Piper, J.R., Waud, W.R. and Sirotnak, F.M., J. Med. Chem., 40 (1997) 370.
14. Piper, J.R., DeGraw, J.I., Colwell, W.T., Johnson, C.A., Lane Smith, R., Waud, W.R. and Sirotnak, F.M., J. Med. Chem., 40 (1997) 377.
15. Matsuoka, H., Ohi, N., Mihara, M., Suzuki, H., Miyamoto, K., Maruyama, N., Tsuji, K., Kato, M., Akimoto, T., Takeda, Y., Yano, K. and Kuroki, T., J. Med. Chem., 40 (1997) 105.
16. Åqvist, J., Medina, C. and Samuelsson, J.E., Protein Eng., 7 (1994) 385.
17. Åqvist, J. and Mowbray, S.L., J. Biol. Chem., 270 (1995) 9978.
18. Åqvist, J., J. Comput. Chem., 17 (1996) 1587.
19. Hultén, J., Bonham, N.M., Nillroth, U., Hansson, T., Zuccarello, G., Bouzide, A., Åqvist, J., Classon, B., Danielsson, H., Karlén, A., Kvarnström, I., Samuelsson, B. and Hallberg, A., J. Med. Chem., 40 (1997) 885.
20. Hansson, T. and Åqvist, J., Protein Eng., 8 (1995) 1137.
21. Hansson, T., Marelius, J. and Åqvist, J., J. Comput.-Aided Mol. Design (1998), in press.
22. Åqvist, J. and Hansson, T., J. Phys. Chem., 100 (1996) 9512.
23. del Buono, G.S., Figueirido, F.E. and Levy, R.M., Proteins, 20 (1994) 85.
24. Lee, F.S., Chu, Z.T., Bolger, M.B. and Warshel, A., Protein Eng., 5 (1992) 215.
25. Abraham, M.H., J. Am. Chem. Soc., 104 (1982) 2085.
26. Ben-Naim, A. and Marcus, Y., J. Chem. Phys., 81 (1984) 2016.
27. Carlson, H.A. and Jorgensen, W.L., J. Phys. Chem., 99 (1995) 10667.
28. Lewis, W.S., Cody, V., Galitsky, N., Luft, J.R., Pangborn, W., Chunduru, S.K., Spencer, H.T., Appleman, J.R. and Blakley, R.L., J. Biol. Chem., 270 (1995) 5057.
29. Oefner, C., d'Arcy, A. and Winkler, F.K., Eur. J. Biochem., 174 (1988) 377.
30. InsightII, Biosym/MSI, San Diego, CA, USA, 1995.
31. Åqvist, J. and Warshel, A., Chem. Rev., 93 (1993) 2523.
32. Åqvist, J., van Gunsteren, W.F., Leijonmarck, M. and Tapia, O., J. Mol. Biol., 183 (1985) 461.
33. van Gunsteren, W.F. and Berendsen, H.J.C., Groningen Molecular Simulation (GROMOS) Library Manual, Biomos, Groningen, The Netherlands, 1987.
34. Warshel, A., Russell, S.T. and Churg, A.K., Proc. Natl. Acad. Sci. USA., 81 (1984) 4785.
35. Sternberg, M.J., Hayes, F.R., Russell, A.J., Thomas, P.G. and Fersht, A.R., Nature 330 (1987), 86.
36. Bashford, D., Karplus, M. and Canters, G.W., J. Mol. Biol., 203 (1988) 507.
37. Åqvist, J., Protein Eng., 5 (1992) 469.
38. Cocco, L., Roth, B., Temple Jr., C., Montgomery, J.A., London, R.E. and Blakley, R.L., Arch. Biochem. Biophys., 226 (1983) 567.
39. Cocco, L., Groff, J.P., Temple Jr., C., Montgomery, J.A., London, R.E., Matwiyoff, N.A. and Blakley, R.L., Biochemistry, 20 (1981) 3972.
40. Jarabak, J. and Bachur, N.R., Arch. Biochem. Biophys., 142 (1971) 417.
41. Greco, W.R. and Hakala, M.T., J. Biol. Chem., 254 (1979) 12104.
42. Perrin, D.D., Dissociation Constants of Organic Bases in Aqueous Solution, Butterworths, London, 1965.
43. Perrin, D.D., Dempsey, B. and Serjeant, E.P., pKa Prediction for Organic Acids and Bases, Chapman and Hall, London, 1981.
44. Yasuda, M., Bull. Chem. Soc. Jpn., 32 (1959) 429.
45. Shedlovsky, T. and Kay, R.L., J. Am. Chem. Soc., 60 (1956) 151.
46. Cramer, C.J. and Truhlar, D.G., Science, 256 (1992) 213.
47. Åqvist, J., Fothergill, M. and Warshel, A., J. Am. Chem. Soc., 115 (1993) 631.
48. Warshel, A. and Creighton, S., In van Gunsteren, W.F. and Weiner, K.P. (Eds), Computer Simulation of Biomolecular Systems: Theoretical and Experimental Approaches, ESCOM, Leiden, The Netherlands, 1989, p. 120.
49. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F. and Hermans, J., In Pullman, B. (Ed.), Intermolecular Forces, Reidel, Dordrecht, The Netherlands, 1981, p. 331.
50. King, G. and Warshel, A., J. Chem. Phys., 91 (1989) 3647.
51. Lee, F.S. and Warshel, A., J. Chem. Phys., 97 (1992) 3100.
52. Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., di Nola, A. and Haak, J.R., J. Chem. Phys., 81 (1984) 3684.
53. Allen, M.P. and Tildesley, D.J., Computer Simulation of Liquids, Oxford University Press, Oxford, U.K., 1987.
54. Bertani, L.E. and Campbell, J.L., Gene, 147 (1994) 131.
55. Montgomery, J.A., Elliot, R.D., Straight, S.L. and Temple Jr, C., Ann. N.Y. Acad. Sci., 186 (1971) 227.
56. Gorse, A.-D. and Gready, J.E., Protein Eng., 10 (1997) 23.
57. Paulsen, M.D. and Ornstein, R.L., Protein Eng., 9 (1996) 567.
58. Jones-Hertzog, D.K. and Jorgensen, W.L., J. Med. Chem., 40 (1997) 1539.
59. Welsh, W., Ruppert, J. and Jain, A.N., Chem. Biol., 3 (1996) 449.