Journal of Computer-Aided Molecular Design, **12:** 325–333, 1998.
KLUWER/ESCOM

Identification and energetic ranking of possible docking sites for pterin on dihydrofolate reductase

Andrey A. Bliznyuk & Jill E. Gready
Computational Molecular Biology and Drug Design Group, John Curtin School of Medical Research, The Australian National University, PO Box 334, Canberra, ACT 2601, Australia

Received 9 March 1998; Accepted 1 December 1997

Key words: dihydrofolate reductase, molecular docking, pterin, solvation calculation

Summary

The reliability of new methodology for detecting sites for ligand binding on the surfaces of proteins has been tested using a range of dihydrofolate reductase (DHFR) crystal structures. Docking of the pterin molecule to ten such DHFR structures has been examined. Initial docking sites were selected using the VDW-FFT method we have developed recently. This procedure was followed by rigid geometry optimization and solvation energy calculations using our parametrized reaction field multipoles (PRFM) method and the finite difference solution of the Poisson equation (FDPB) method. Two different sets of MM parameters, from the OPLS and Amber94 force fields, have been used. In eight cases the energy of the complexes with pterin bound at the active site was the lowest with the recent Amber94 parameters. In one case the spurious first-ranked site was only 1.8 kcal/mol lower in energy compared with the active site. The other 'failure' of the method may, in fact, represent a valid initial binding site. The calculations with the old OPLS parameters gave slightly worse results.

Introduction

The availability of high-resolution structures of enzymes has spurred interest in predicting the position and strength of binding of ligands, such as substrates and inhibitors. The general solution to this problem (molecular docking problem) requires a combination of methods for locating the binding site (or sites) and evaluating the energy of binding.

The difficulty of solving the docking problem arises because the binding energy is hard to calculate. Ideally, evaluation of the free energy is necessary to discriminate between possible binding sites and orientations of a ligand [1]. Unfortunately, such an approach is very time consuming and can be used only when the location of the active site is known. A faster method, based on parametrization of molecular dynamics energies, has been suggested recently for evaluation of binding energies [2–5]. While this may be an attractive alternative to the traditional free energy perturbation (FEP) calculations, it is still much too slow for use in locating possible binding sites on the surface of an enzyme. Recently, the finite difference solution of the Poisson-Boltzmann equation (FDPB) method (see [6–8] for reviews) in combination with molecular mechanics (MM) energy evaluation has been used successfully for evaluation of binding energies [9]. This method is relatively fast and can be used to discriminate among possible binding sites at reasonable compute-time expense, provided that the number of candidate sites can be restricted to approximately a dozen. Hence, the development of a method for locating the appropriate candidate sites assumes major importance.

In addition to the above general methods for estimating the free energy of binding, various parametric functions have been proposed for evaluation of binding constants. Several reviews of these approaches have been published recently, see for example refs. 10 and 11. These methods commonly use available experimental data to obtain parameters for some relatively simple functionals that allow for fast estimation of the binding energy. These functionals have been applied successfully to the task of selecting promis-

ing ligands that can bind to the known active site. The parameters of these functions are based on data for active-site binding only and they do not have well defined physical foundation. Consequently, their usefulness for estimating the binding energies for regions on protein surfaces that do not resemble active sites is doubtful. Hence, in this article, which addresses the more general problem of finding possible docking sites on protein surfaces, we have employed the FDPB method of solvation energy estimation in combination with MM energy evaluation in order to avoid such possible methodological complications.

Several approaches based on the idea of geometric complementarity between a ligand and an enzyme binding site have been proposed. The main concept behind them is to simplify drastically the interaction energy calculations so as to permit a much more sophisticated scan of the protein surface [12–17]. All such methods appear to be able to select several dozen possible binding sites, usually including the actual binding site. However, it should be noted that the methods [12–14,16–17] have not been tested extensively for complexes of proteins with small ligands. For example, only 3 such complexes were included in the test set in ref. 12 and only 2 in refs. 14 and 16. Identifying possible docking sites of small ligands is a quite different problem from searching for possible aggregation sites between two large enzyme molecules. Recently, a new approach to the problem of detecting possible binding sites was described [15]. The so-called 'pocket finder' method produced excellent results on a test set of 10 enzyme-ligand complexes [15]. This method is based on the evaluation of a scoring function, which was optimised for reproducing experimental binding energies of ligands in the active site. As the authors of the ref. 15 pointed out, it is not clear whether the method can be applied successfully to identifying binding sites on the exterior of a protein.

For a test set of 15 complexes of proteins with small ligands we have shown recently that a modification of the approach described in refs. 16–18 allows us to reduce successfully the number of candidate binding sites to a very few (Bliznyuk and Gready, in preparation).

In this article we use the recently developed VDW-FFT procedure (Bliznyuk and Gready, in preparation) in conjunction with the FDPB and parametrized reaction field multipoles (PRFM) [19] methods for solvation energy calculations to find possible binding sites on the surface of the enzyme dihydrofolate reductase (DHFR).

Method

Selection of candidate sites

The calculations begin with a projection of an enzyme onto two 3-D grids of size N×N×N. Points on one grid are assigned to a modified van der Waals (VDW) repulsion energy, defined in Equation 1.

$$V(i, j, k) = \sum_{l} \sqrt{A_{ll}} \,/\, (R + d \,/\, 2)^{12} \qquad (1)$$

Here, (i,j,k) is a grid point index, R is the distance between the grid point and an atom l, d is the grid spacing, and A_{ll} is the standard VDW repulsion parameter. Summation is over all atoms within the cutoff distance from point (i,j,k). We used N equals 64 and a cutoff of 8 Å for the grid calculations. This gives a value of d in the range 0.8–1.0 Å. To avoid large values of V(i,j,k) which may occur if (R+d/2) is small, we use 1 Å whenever (R+d/2) is less than 1 Å. Points on the other grid were assigned to the VDW attraction energy, defined in Equation 2.

$$W(i, j, k) = -\sum_{l} \sqrt{B_{ll}} \,/\, R^{6} \qquad (2)$$

Here, B_{ll} is the standard VDW attraction parameter. The ligand's VDW parameters were projected onto the nearest eight points of grids with the same dimensions as the enzyme grids using Equations (3) and (4).

$$V_L(i, j, k) = \sum_{l} (1 - a \,/\, d)(1 - b \,/\, d)(1 - c \,/\, d)\sqrt{A_{ll}} \qquad (3)$$

$$W_L(i, j, k) = \sum_{l} (1 - a \,/\, d)(1 - b \,/\, d)(1 - c \,/\, d)\sqrt{B_{ll}} \qquad (4)$$

Here, the values of a, b, and c are distances along the x, y, and z coordinates from atom l to the grid point (i,j,k). We found that this simple linear projection worked well and did not experiment with others.

It is easy to calculate the modified VDW energy using the grids. For example, the energy of interaction of the enzyme with the ligand molecule shifted over the distance (p*d,m*d,n*d) is a simple multiplication and summation over grid points:

$$E(p, m, n) = \sum_{i}^{N} \sum_{j}^{N} \sum_{k}^{N} \{V(i, j, k)^{*} V_L(i + p, j + m, k + n) + W(i, j, k)^{*} W_L(i + p, j + m, k + n)\} \qquad (5)$$

Such a form of the energy function allows us to use fast Fourier transform (FFT) to find the minima of E(p,m,n) (for details see refs. 16, 18). The reason for the use of parameter d in Equation 1 is clear from Equation 5. It makes the repulsive grid smoother, thus reducing errors due to the discrete nature of the grid. We find that the use of d in Equation 1 is necessary to produce meaningful results with small ligands and a moderate grid size (N=64).

A similar type of the VDW energy projection onto the grids has been used in ref. 18 for the purpose of calculating the exact MM energies. In order to obtain an accurate estimation of the interaction energies, a very fine grid was utilized and parameter d/2 was not used in Equation 1. The goal of our approach is to select possible binding sites, using the idea of geometrical complementarity between the receptor site and the ligand. Hence, accurate evaluation of the VDW energy is not needed. On the other hand, the use of a simple geometrical projection, as was proposed in ref. 16, did not allow us to select relevant candidate sites (Bliznyuk and Gready, in preparation). Hence, we decided to employ the modified form of the VDW energy, as described in Equations (1–4).

In order to scan all possible orientations of the ligand with respect to the enzyme, we vary three Euler angles which define the orientation by fixed steps of 20°, similarly to the original method [16]. This results in the generation of 360/20 * 360/20 * 180/20 = 2916 possible rotations, only 2628 of which correspond to dissimilar orientations. Hence, the procedures of projecting the ligand onto the grids and the FFT-based search of minima have to be repeated 2628 times. Fortunately, the FFT procedure is fast and the whole operation requires approximately 40 minutes on the SGI Power Indigo^2 (R8000) workstation.

We saved the best 100 positions and orientations of the ligand produced by the FFT procedure. Then we performed MM energy optimization using the simplex and the BFGS methods [20] with the saved ligand positions as starting points. As a result, 200 different candidate sites were obtained. These sites have been grouped, assuming that two sites are different if their RMS difference is larger than 0.2 Å. This procedure usually reduces the number of sites to 50–60. Of these, the ten sites with lowest MM energies have been selected for further investigation. Finally, the PRFM and FDPB calculations of solvation energies were performed on the selected ten complexes. PRFM calculations were done as described in ref. 19. The FDPB calculations were performed using the FD solution of the Poisson equation as given by the DelPhi program [21]. A dielectric constant for water equal to 80, a grid size of 0.4 Å, a probe radius of 1.4 Å and zero ionic strength have been used in the DelPhi calculations. All PRFM and DelPhi calculations were done using the PRFM radii [19] of the atoms.

This procedure was repeated for all ten enzymes investigated. Also, two different sets of protein MM parameters have been used: OPLS [22] and Amber94 [23].

Test set

Most of the tests for various docking methods consist of selecting enzyme-ligand complexes from the Brookhaven Protein data bank (PDB) [24] and then docking the ligand to the enzyme. We decided to choose a more elaborate test. We selected one enzyme, DHFR, for which many different complexes have been deposited in the PDB. These entries differ in a number of ways: the enzyme was crystallised in different space groups; the structure was obtained using the NMR rather than X-Ray method; the complexes were with different ligands; the enzymes come from different species. With more than 30 different complexes now available in the PDB, there is a unique opportunity to investigate how differences in atomic coordinates influence the performance of a docking method. A summary of the selected DHFR complexes is given in Table 1.

Entries 1DHF and 1DRF are complexes of human DHFR with folate but the coordinates were determined in different space groups, P1 and R3, respectively. Files 1DIS and 3DFR are for *L. casei* DHFR. The coordinates of the former were determined using the NMR method, while the latter were obtained from high resolution X-Ray data. *E. coli* protein structures have been solved in many different complexes. Examples used here were: 1DRH, a complex with cofactor NADPH, 4DFR, a complex with inhibitor methotrexate, and 7DFR in a folate ternary complex. For the chicken enzyme ternary complex, 1DR1, two conformations of Tyr31 have been reported [29]. Both have been used in our calculations. They are denoted 1DR1 A and 1DR1 B in the text. Finally, we included in the test set the recently-reported structure of the ternary complex of *P. carinii* DHFR. Unfortunately, we could not incorporate in the test set coordinates of *E. coli apo*-enzyme [33], because residues 16–20, which are near the active side, are disordered and their coordinates are missing in the PDB entry 5DFR.

Table 1. Summary of DHFR enzyme complexes used in the calculations.

Name	Species	Resolution, Å	Point group	Ligands	Conserved acidic residue in the active site	Reference
1DHF	Human	2.3	P1	Folate	Glu30	[25]
1DRF	Human	2.0	R3	Folate	Glu30	[26]
1DIS	*L. casei*	NMR		BDM[a]	Asp26	[27]
3DFR	*L. casei*	1.7	P61	NADPH[b], MTX[c]	Asp26	[28]
1DR1	Chicken	2.2	C2	NADP[d], Biopterin	Glu30	[29]
1DRH	*E. coli*	2.3	P322	NADPH	Asp27	[30]
4DFR	*E. coli*	1.7	P61	MTX[c]	Asp27	[28]
7DFR	*E. coli*	2.5	P322	NADP[d], Folate	Asp27	[31]
1DYR	*P. carinii*	1.9	P21	NADPH[b], TMP[e]	Glu32	[32]

[a] Brodimoprim-4,6-dicarboxylate
[b] Nicotinamide adenine dinucleotide phosphate, reduced form.
[c] Methotrexate
[d] Nicotinamide adenine dinucleotide phosphate, oxidised form.
[e] Trimethoprim

Hydrogen atoms have been added using the InsightII program [32]. Missing side chains in 1DR1 (Arg2, Lys80, Glu104, Lys178), 1DRF (Glu12), 1DRH (Glu17, Leu28, Lys32, Arg52, Arg98, Glu120, Arg159), 3DFR (Lys51), 4DFR (Glu129, Asp131) and 7DFR (Glu17, Arg44, Glu48, Arg52, Arg98, Lys106, Asp116, Glu120, Glu129, Asp131) have been modelled using the 'replace' command, as implemented in the Biopolymer module of the InsightII program [34]. All of these side chains are pointing towards the solvent.

As a suitable simple ligand, we chose pterin, a fragment of the substrate folate (Figure 1). Pterins are known to bind in the active site of DHFRs in an orientation similar to folate [29, 35]. Pterin is a small molecule and does not bind very strongly. Hence, there may be several sites on the surface of DHFRs where pterin would be able to bind. This makes the accurate location of the active site a more demanding task than the binding of methotrexate, where only one strong binding site was located (Bliznyuk and Gready, in preparation). Coordinates for pterin were obtained from an SCF *ab-initio* calculation with geometry optimization and a 6-31G* basis set. Atomic charges for the MM computations were generated by fitting the electrostatic potential using the Merz-Kollman procedure [36–37] as implemented in the Gaussian94 program [38]. VDW parameters were assigned using OPLS [39] and Amber94 [23] data for nucleic acid bases.

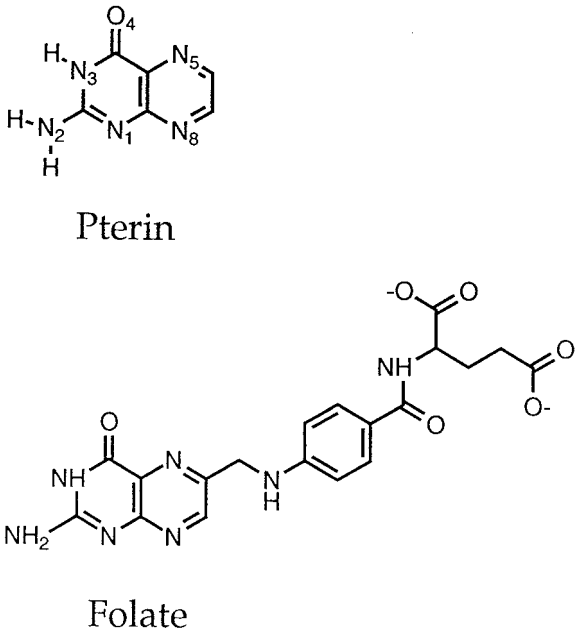

Figure 1. Schematic representation of the DHFR substrate ligand folate and its fragment pterin.

Results and Discussion

The results obtained from the docking calculations are summarised in Table 2. We assumed that the pterin molecule is in the correct active-site position if it binds in the orientation similar to that of the folate [25, 26, 31]. Specifically, that it forms hydrogen bonds (H-bonds) $N_3H\ldots O$ and $N_2H\ldots O$ with the conserved

Table 2. Rank from the docking calculations of the correct positioning of pterin in the active site of the DHFR structures given in Table 1.

Enzyme	OPLS			Amber94		
	MM	PRFM	DelPhi	MM	PRFM	DelPhi
1DHF	2	1	1	1	1	1
1DRF	4	1	1	5	1	1
1DIS	3	1	1	1	1	1
3DFR	2	2	1	5	3	1
1DR1 A	1	1	2	2	1	1
1DR1 B	1	1	1	2	1	1
1DRH	2	2	2	2	1	1
4DFR	1	4	3	1	1	3
7DFR	2	1	1	2	1	1
1DYR	4	2	2	3	3	2

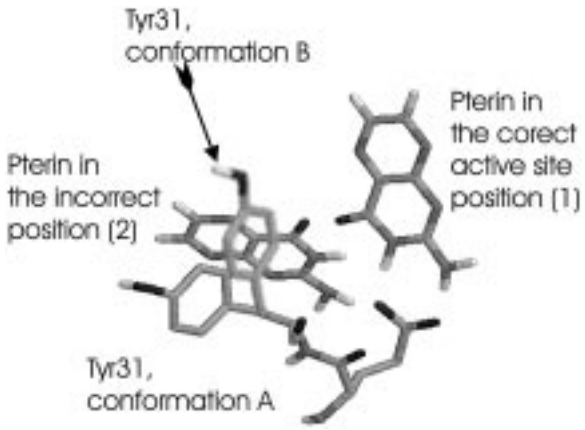

Figure 2. Orientations of pterin relative to the Glu30 and Tyr31 residues in the active site of DHFR in the 1DR1 complex. The B conformation of Tyr31 is shown as a blur.

carboxylate side chain of the Asp or Glu residue (see Table 1 for the numbering of the residue) in the active site and it is placed approximately in the same plane as the carboxylate group.

The combination of the FFT-VDW docking procedure with MM optimization worked quite well. In all cases, the proper active site has been located and ranked among five best candidate sites. Addition of solvation energy usually improves the ranking, predicting in most cases the correct result. Our new PRFM model for solvation energy calculation [17] worked as well as the much more sophisticated and time consuming FDPB method implemented in DelPhi. Good results have been obtained using both OPLS and Amber94 MM parameters, the latter being slightly better.

One of the parameters for the solvation energy calculations for both the FDPB and PRFM methods is the dielectric constant of the protein interior. Although the protein cannot be described rigorously by a single dielectric constant [40], we treat this value as a parameter of the method for the solvation energy calculations. Such a point of view is similar to the approach taken by the developers of the DelPhi program [7]. Our preliminary calculations with a value for the internal dielectric constant of 1 did not give acceptable results. However, with a value of 4 for the internal dielectric constant good ranking was achieved. Therefore, only these latter data are considered below.

In the following discussion we refer mostly to the results obtained using DelPhi which is, in principle, a more accurate method than the PRFM method [17].

While overall the results are quite good, there are some inconsistencies. Probably the most unusual one is the difference between the results for the 1DR1 A and 1DR1 B complexes when using the OPLS parameters. The enzyme coordinates are identical except for those of the Tyr31 residue side chain. Figure 2 shows the correct (1) and incorrect (2) positions of the pterin molecule in the active site with Tyr31 in conformation A.

Pterin in the assumed correct position (1) forms strong hydrogen bonds with Glu30. Pterin in position (2) also interacts with Glu30, but the MM OPLS energy is approximately 2.6 kcal/mol higher. Inclusion of solvation energy within DelPhi makes this position approximately 0.1 kcal/mol more favourable. While the absolute value of this energy difference is not significant and is well within the error threshold of the calculations employed, the direction of the tendency indicated is important. The Amber94 result predicts the MM energy difference of 8.5 kcal/mol for these two positions. Inclusion of the solvation energy reduces this difference but much less than needed for position (2) to become more favourable. This leads to a conclusion of possible deficiency of the OPLS parameters. When the Tyr31 residue is in conformation B the residue is closer to the position (2) of pterin (see Figure 2) thus making this position less favourable due to VDW repulsion.

From the results for the 1DYR enzyme complex both OPLS and Amber94 parameters ranked the active site position as second. The best site is located approximately 20 Å away and is formed by residues 5–8, 118–120, 139–140, and 180 (see Figure 3).

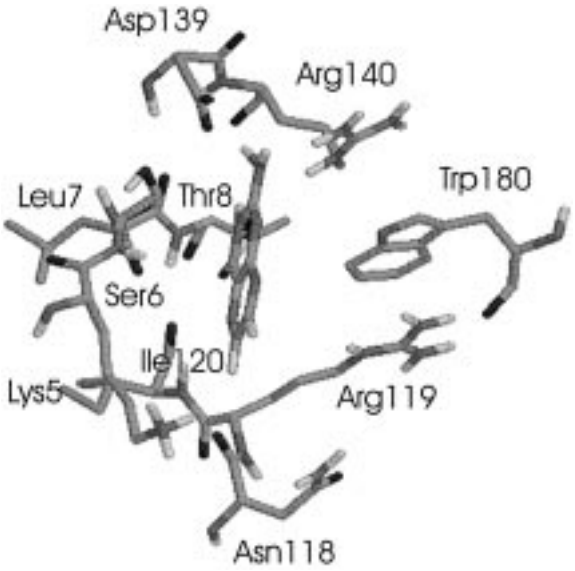

Figure 3. The spurious pterin binding site for the 1DYR DHFR enzyme complex located ∼20 Å from the active site.

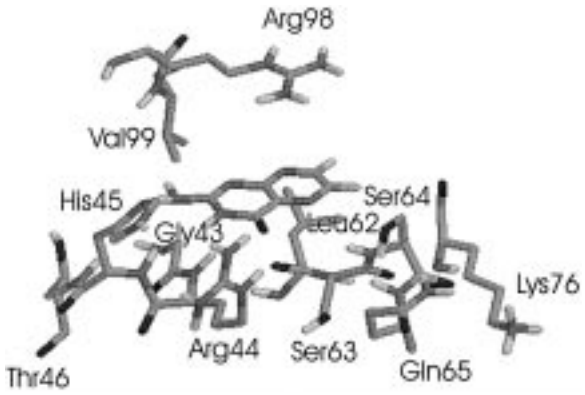

Figure 4. The spurious pterin binding site in the 1DRH DHFR enzyme complex as obtained in the OPLS calculations.

The site is shallow and atoms N_1 and N_8 of pterin (facing out of the plane of the paper) are accessible to the solvent. The pterin NH_2 group forms hydrogen bonds with Asp139, and the 4-oxo group of pterin interacts strongly with Thr8. This site is approximately 7 kcal/mol more stable than the active site according to the calculations with the OPLS parameters. The difference reduces to 1.8 kcal/mol when the Amber94 parameters are used. Such a combination of residues that form this site seems to be unique to DHFR from *P. carinii.* While an interesting result, we believe it is more likely to be an artefact of the MM parameters used than a proper binding site.

Very interesting results were obtained for the *E. coli* enzyme complexes. The active site position in the OPLS calculations for 1DRH was ranked second. This lowest energy site is located approximately 20 Å from the Asp27 residue in a niche formed by Leu62 and Ser63 residues at the bottom and Arg98 at the top (see Figure 4).

The position of the pterin molecule is stabilised by a strong H-bond interaction between the O_4 atom of pterin and the side chain of the Arg44 residue. In the OPLS calculations this site is approximately 1 kcal/mol more stable than the active site. The Amber94 calculation predicts this position to be second, placing the active site position approximately 2 kcal/mol lower in energy. The MM-only energy of this improper site is noticeably higher than the energy of the active site. The difference is approximately 10 and 6 kcal/mol for OPLS and Amber94 calculations, respectively. In both cases, the energy difference is reduced by inclusion of the solvation energy, the OPLS result now leading to the incorrect prediction. Here, as in the case of the 1DR1 calculations, the Amber94 parameters seem to be superior to those for OPLS. It is worth noting that the position of the Arg44 residue side chain is different in the 4DFR and 7DFR enzymes, making the formation of an H-bond with the O_4 atom in pterin impossible. Hence, the analogous site is ranked well below the active site for these enzyme complexes. Here we have a very clear illustration of how a slight difference in the crystal-structure coordinates can have a large effect on the results of the calculations.

Another example of the same character can be seen in the results obtained for docking of pterin into the 4DFR enzyme coordinates. The active site for this structure has been ranked third by both the OPLS and Amber94 calculations. As the two lowest energy sites have less than 0.3 Å RMS difference in the pterin coordinates, we restrict our discussion to the most stable position only. The ligand in this site is located in the direct vicinity of the active site and interacts strongly with the side chain of the Asp27 residue, but is turned approximately 90 degrees compared with the pterin position in the active site (see Figure 5).

This incorrect position is more stable by approximately 3 and 1 kcal/mol for the OPLS and Amber94 calculations, respectively. There is no pterin molecule in this site in the docking of the other *E. coli* enzyme complexes because the Met20 residue has a different conformation and the site is occupied by its thiomethyl group. The position of the flexible ‘Met20

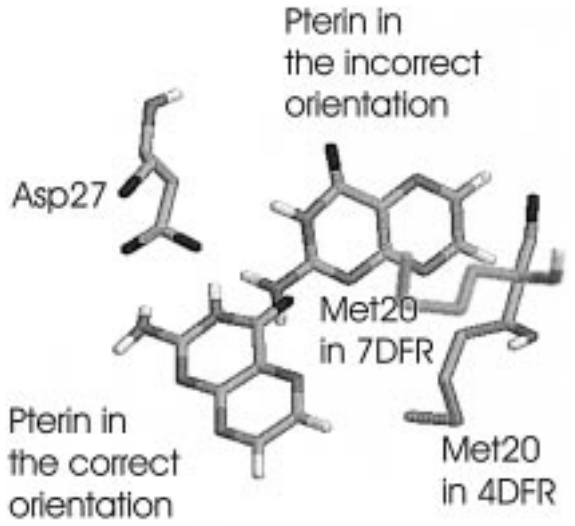

Figure 5. Two different positions of pterin in the active site of the 4DFR DHFR enzyme complex. The position of the Met20 residue in the 7DFR and 1DRH structures are similar; the location of the Met20 residue in the former is shown as a blur.

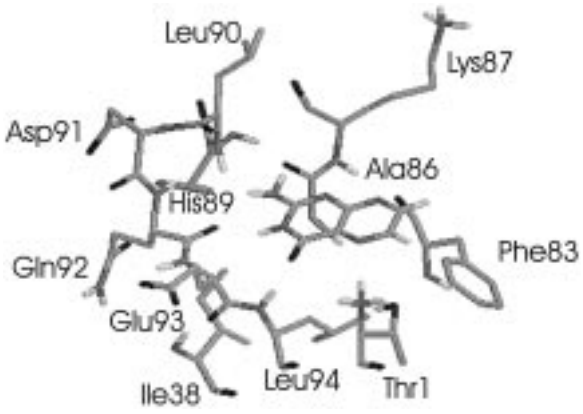

Figure 6. One of the spurious pterin binding sites in the 3DFR DHFR enzyme complex.

loop' (which contains Met20) is known to be affected strongly by the space group in which molecule of the DHFR was crystallised [30,41]. The location of this loop also depends upon whether the enzyme has been crystallised as an *apo*, binary or ternary complex [30,41]. Hence, this position of the ligand may be an initial docking site formed when the Met20 loop is open. The closure of the Met20 loop would facilitate movement of the ligand to the actual active site. We are not aware of any experimental data which could confirm or refute this hypothesis, but it is an interesting idea bearing on the unknown dynamics of ligand binding and complex formation.

In all others cases the calculations with both OPLS and Amber94 parameters were able to predict correctly the active site as the lowest in energy. However, it should be noted that in many cases the difference between the energy of the active-site position and some spurious site is only 1 or 2 kcal/mol. Such a small difference is within the accuracy of much more precise free energy perturbation calculations [42]. In all cases, the spurious sites have a very reasonable geometry and form interactions which are, at least, partly similar to the interactions known experimentally. For example, the ligand position ranked second in the *L. casei* 3DFR OPLS calculations has only 2 kcal/mol higher total energy than the active site. It is located more than 20 Å away on the other side of the protein. Residues 1, 38, 83, 86-87, 89–94 form a shallow pocket in which the pterin ring sits on the top of the β-carbon of Ala86 and forms hydrogen bonds with backbone atoms (see Figure 6). The strongest interactions are N_3H...O Gln92, O_4. . .HN Leu94 and N_5. . .HN Thr1. Such an interaction pattern with the pterin ring is unusual for DHFR enzymes, where strong interaction with the conserved Asp or Glu side chain occurs in the active site. However, molecules containing the pterin ring are known to bind at the active site of glycinamide ribonuclease transformylase [43] and plasmid encoded R67 DHFR [44] forming their main interactions with backbone atoms. The Amber94 calculations for the 3DFR enzyme - pterin complex predict this site to be 2.6 kcal/mol less favourable compared with the active site. The analogous site for the 1DIS *L. casei* DHFR coordinates was ranked third in the Amber94 calculations (7 kcal/mol higher than the active site) and second in the OPLS calculations (5 kcal/mol higher than the active site).

There is a substantial difference in MM energy ranking between these two *L. casei* DHFR coordinate sets (see Table 2). For example, the active site was ranked fifth in Amber94 MM calculations for 1DIS complex and first for the 3DFR complex. According to the NMR data [27] all the His residues are charged in 1DIS giving a total charge on the enzyme equal to $+2$. For the 3DFR protein we assumed all His residues to be neutral and, hence, the total charge was -3. This explains the difference in MM ranking. However, the correct ranking of the active site was obtained with solvation in both cases.

Conclusions

The docking calculations of pterin with DHFR enzyme were performed using the sum of the molecular mechanics energy and solvation energy as the total energy function. The solvation energy was calculated using the FD solution of the Poisson equation. To our knowledge, this is the first attempt to investigate thoroughly possible binding sites of a ligand with an enzyme using such an elaborate energy function. Coordinates for ten different DHFR enzyme complexes have been used as a test set. In eight cases the energy of the complexes with pterin bound at the active site was the lowest when the recent Amber94 parameters [23] were used. In one case (1DYR) the spurious first-ranked site was only 1.8 kcal/mol lower in energy compared with the active site. The other 'failure' of the method (4DFR) may, in fact, represent a valid initial binding site. The calculations with the old OPLS parameters [22] gave slightly worse results.

In all cases, the difference in total energy between the active site and low energy spurious sites were in the range of a few kcal/mol. Hence, the results may be affected by the lack of optimization of the enzyme coordinates. Another possible source of error may result from the position of the ligand being optimized using only the MM type of energy function instead of the total energy. Unfortunately, use of the sophisticated FDPB approach for the solvation energy calculations makes the optimization of the enzyme coordinates and/or the ligand position impractical. Neglect of optimization of the ligand's position was compensated for partially by the method of selecting the probe sites. The method usually generates several orientations of a ligand in any particular site. Also, our test calculations using the quicker PRFM approach for optimization of the ligand positions indicated that the geometry change is slight, usually not more than of 0.1 Å RMS.

Overall, the combination of the VDW FFT (Bliznyuk and Gready, in preparation) method for selection of probe sites with high-level total energy evaluation for accurate discrimination between them was highly successful. The program for performing the VDW FFT calculations can be obtained freely from our web page: http://www.biocomp.anu.edu.au/∼aab.

Acknowledgements

JEG gratefully acknowledges funding from an ANU Strategic Initiative Development Grant.

References

1. Kollman, P.A., Curr. Opin. Struct. Biol., 4(1994) 240.
2. Aqvist, J., Medina, C. and Samuelsson, J.-E., Protein Eng., 7 (1994) 385.
3. Aqvist, J., J. Comput. Chem. 17 (1996) 1587.
4. Carlson, H.A. and Jorgensen, W.L., J. Phys. Chem., 99 (1995) 10667.
5. Gorse, A-D. and Gready, J.E., Prot. Eng., 10 (1997) 23.
6. Sharp, K.A. and Honig, B., Ann. Rev. Biophys. Biophys. Chem., 19 (1990) 301.
7. Honig, B., Sharp, K. and Yang, A., J. Phys. Chem., 97 (1993) 1101.
8. Sharp, K., Curr. Opin. Struct. Biol., 4 (1994) 234.
9. Shen, J. and Quiocho F.A., J. Comput. Chem., 16 (1995) 445.
10. Bamborough, P. and Cohen F.E., Curr. Opin. Struct. Biol., 6 (1996) 236.
11. Lengauer, T. and Rarey, M., Curr. Opin. Struct. Biol., 6 (1996) 402.
12. Lin, S.L. and Nussinov R., J. Mol. Graph., 14 (1996) 78.
13. Fischer, D., Lin, S.L., Wolfson, H.L. and Nussinov, R., J. Mol. Biol., 248 (1995) 459.
14. Kasinos, N., Lilley, G.A., Subbarao, N. and Haneel, I., Prot. Eng., 5 (1992) 69.
15. Ruppert, J., Welch, W. and Jain, A.N., Protein Science, 6 (1997), 524.
16. Katchalski-Katzir, E., Shariv, I., Eisenstein, M., Friesem, A.A., Aflalo, C. and Vakser, I.A., Proc. Natl. Acad. Sci. USA, 89 (1992) 2195.
17. Vakser, I.A. and Aflalo, C., Proteins, 20 (1994) 320.
18. Harrison, R.W., Kourinov, I.V. and Andrews. L.C., Prot. Eng., 7 (1994) 359.
19. Bliznyuk, A.A. and Gready, J.E., J. Phys. Chem., 99 (1995) 14506.
20. Dennis, J.E.,Jr. and Schnabel, R.B., Numerical methods for unconstrained optimization and nonlinear equations, Prentice Hall Inc., New York, 1983.
21. DelPhi program, Biosym Technologies version 2.5., 1993.
22. Jorgensen, W.L. and Tirado-Rives, J., J. Am. Chem. Soc., 110 (1988) 1664.
23. Cornell, W.D., Cieplak, P., Bayly, C.I., Gould, I.R., Merz., K.M. Jr., Ferguson, D.M., Spellmeyer, D.C., Fox, T., Calwell, J.W. and Kollman, P.A., J. Am. Chem. Soc., 117 (1995) 5179.
24. Bernstein, F.C., Koetzle, T.F., Williams, G.J.B., Meyer, E.F. Jr., Brice, M.D., Rodgers, J.D., Kennard, O., Shimanouchi, T. and Tasumi, M., J. Mol. Biol., 112 (1977) 535.
25. Davies II, J.F., Delcamp, T.J., Prendergast, N.J., Ashford, V.A., Freisheim, J.H. and Kraut, J., Biochemistry, 29 (1990) 9467.
26. Oefner, C., D'Arcy, A. and Winkler, F.K., Eur.J.Biochem., 174 (1988) 377.
27. Morgan, W.D., Birdsall, B., Polshakov, V.I., Sali, D., Kompis, I. and Feeney, J., Biochemistry 34 (1995) 11690.
28. Bolin, J.T., Filman, D.J., Matthews, D.A., Hamlin, R.C. and Kraut, J., J. Biol. Chem., 257 (1982) 13650.
29. McTigue, M.A., Davies II, J.F., Kaufman, B.T., Xuong, N-H. and Kraut, J., Biochemistry, 31 (1992) 7264.
30. Reyes, V.M., Sawaya, M.R., Brown, K.A. and Kraut, J., Biochemistry, 34 (1995) 2710.
31. Bystroff, C., Oatley, S.J. and Kraut, J., Biochemistry, 29 (1990) 3263.
32. Champness, J.N., Achari, A., Ballantine, S.P., Bryant, P.K., Delves, C.D. and Stammers, D.K., Structure, 2 (1994) 915.
33. Bystroff, C. and Kraut, J., Biochemistry, 30 (1991) 2227.

34. Insight II, version 95, October 1995. San Diego: Biosym/MSI, 1995.
35. Ivery, M.T.G. and Gready, J.E., Biochemistry, 34 (1995) 3724.
36. Singh, U.C. and Kollman, P.A., J. Comput. Chem., 5 (1984) 129.
37. Besler, B.H., Merz, K.M., Jr. and Kollman, P.A., J. Comput. Chem., 11 (1990) 431.
38. Gaussian 94 (Revision C.2), Frisch, M.J., Trucks, G.W., Schlegel, H.B., Gill, P.M.W., Johnson, B.G., Robb, M.A., Cheeseman, J.R., Keith, T.A., Peterson, G.A., Montgomery, J.A., Raghavachari, K., Al-Laham, M.A., Zakrzewski, V.G., Ortiz, J.V., Foresman, J.B., Cioslowski, J., Stefanov, B.B., Nanayakkara, A., Challacombe, M., Peng, C.Y., Ayala, P.Y., Chen, W., Wong, M.W., Andres, J.L., Replogle, E.S., Gomperts, R., Martin, R.L., Fox, D.J., Binkley, J.S., Defrees, D.J., Baker, J., Stewart, J.P., Head-Gordon, M., Gonzalez, C. and Pople, J.A., Gaussian, Inc., Pittsburgh P.A., 1995.
39. Pranto, J., Wiershke, S.G. and Jorgensen, W., J. Am. Chem. Soc., 113 (1991) 2810.
40. Warshel, A. Computer modelling of chemical reactions in enzymes and solutions, John Wiley and Sons, Inc., New York, 1991.
41. Sawaya, M.R. and Kraut, J., Biochemistry, 36 (1997) 586.
42. Straatsma, T.P, Zacharias, M. and McCammon, J.A., In: van Gunsteren, W.F., Weiner, P.K. and Wilkinson, A.J. (Eds.) Computer Simulation of Biomolecular Systems, Vol. 2, ESCOM, Leiden, 1993, pp. 349–367.
43. Klein, C., Chen, P., Arevalo, J.H., Stura, E.A., Marolewski, A., Warren, M.S., Benkovic, S.J. and Wilson, I.A., J. Mol. Biol., 249 (1995) 153.
44. Narayana, N., Matthews, D.A., Howell, E.E. and Xuong, N., Nature Struct. Biol., 2 (1995) 1018.