Journal of Computer-Aided Molecular Design (2005) 19: 165–187
DOI 10.1007/s10822-005-3693-6

Multiple ligand-binding modes in bacterial R67 dihydrofolate reductase

Hernán Alonso[a], Malcolm B. Gillies[a], Peter L. Cummins[a], Andrey A. Bliznyuk[b] & Jill E. Gready[a,*]
[a]*Computational Proteomics Group, John Curtin School of Medical Research, The Australian National University, P.O. Box 334, 2601, Canberra, ACT, Australia;* [b]*ANU Supercomputer Facility, The Australian National University, 0200, Canberra, ACT, Australia*

Received 13 October 2004; accepted in revised form 11 March 2005

Key words: AutoDock, consensus scoring, docking, FlexX, GROMACS, interligand cooperativity, ligand mobility, molecular dynamics, reactive complex

Summary

R67 dihydrofolate reductase (DHFR), a bacterial plasmid-encoded enzyme associated with resistance to the drug trimethoprim, shows neither sequence nor structural homology with the chromosomal DHFR. It presents a highly symmetrical toroidal structure, where four identical monomers contribute to the unique central active-site pore. Two reactants (dihydrofolate, DHF), two cofactors (NADPH) or one of each (R67•DHF•NADPH) can be found simultaneously within the active site, the last one being the reactive ternary complex. As the positioning of the ligands has proven elusive to empirical determination, we addressed the problem from a theoretical perspective. Several potential structures of the ternary complex were generated using the docking programs AutoDock and FlexX. The variability among the final poses, many of which conformed to experimental data, prompted us to perform a comparative scoring analysis and molecular dynamics simulations to assess the stability of the complexes. Analysis of ligand–ligand and ligand–protein interactions along the 4 ns trajectories of eight different structures allowed us to identify important inter-ligand contacts and key protein residues. Our results, combined with published empirical data, clearly suggest that multipe binding modes of the ligands are possible within R67 DHFR. While the pterin ring of DHF and the nicotinamide ring of NADPH assume a stacked *endo*-conformation at the centre of the pore, probably assisted by V66, Q67 and I68, the tails of the molecules extend towards opposite ends of the cavity, adopting multiple configurations in a solvent rich-environment where hydrogen-bond interactions with K32 and Y69 may play important roles.

Abbreviations: R67 DHFR – R67 dihydrofolate reductase; DHF – dihydrofolate; DHFH+ – N5-protonated dihydrofolate; NADPH – reduced nicotinamide adenine dinucleotide phosphate; NMN – nicotinamide-ribose-phosphate moiety of NADPH; pte – pterin ring of folate; nic – nicotinamide ring of NADPH; pABA-Glu – *para*-aminobenzoyl glutamic acid tail of folate; 2′,5′-ADP – adenosine diphosphate ribose moiety of NADPH

Introduction

Tetrahydrofolate (THF) and its derivatives are essential cofactors derived from the vitamin folic acid. They are involved in the metabolism of the amino acids serine, glycine, and methionine and

*Correspondence to: Prof Jill E. Gready, Computational Proteomics Group, John Curtin School of Medical Research, The Australian National University, P.O. Box 334, Canberra ACT 2601, Australia. Phone: +61 2 6125 8304, Fax: +61 2 6125 0415, Email: jill.gready@anu.edu.au

are also used in the synthesis of purine and thymine nucleotides. Once inside the cell, folate is sequentially reduced to dihydrofolate (DHF) and THF by a nicotinamide adenosine diphosphate (NADPH)-specific enzyme, dihydrofolate reductase (DHFR). Because of its metabolic importance, chromosomal DHFR constitutes an interesting target for drug design. However, some microorganisms have evolved alternative enzymes that can carry out the reduction of folate and, therefore, bypass the otherwise lethal blockage of chromosomal DHFR. *Citobacter* sp. type 2 DHFR (R67), a R-plasmid encoded enzyme which is thought to have evolved in response to trimethoprim antibiotic pressure, is one of them [1, 2]. Understanding of the approach R67 DHFR uses to bind the ligands and promote the reaction is not only of clinical relevance but also important as a basic research question. On the one hand, the presence of this enzyme in broad-range plasmids and integrons [1, 3–5] facilitates the propagation of antibiotic resistance through horizontal transfer. On the other hand, despite catalysing the same reaction as the chromosomal DHFR, it has neither sequence nor structural homology with it.

The chromosomal-encoded DHFR, or type 1 DHFR, is a biologically ubiquitous enzyme that has been extensively studied since the 1950s [6]. There have been major advances in elucidating its kinetic pathways, catalytic mechanism and structure [7, 8]. On the other hand, very little is known about R67 DHFR, a type 2 enzyme isolated only in the mid 1970s [1]. This kind of DHFR has been found only as single-gene cassettes in integrons [9] and does not have any known homologues. Its sequence and 3-D structure are unrelated to those of the conventional DHFR enzyme, and it also presents a substantially different catalytic mechanism. While the type 1 DHFR catalyses the transfer of the *pro-R* hydride from the nicotinamide ring of NADPH to the pterin ring of DHF in an *endo* conformation [10], the type 2 enzyme promotes an *exo* reaction [11]. This implies that the relative orientation of the ligands inside the active site and the potential energy surface for the reaction must be significantly different between the chromosomal DHFR and R67 DHFR enzymes. Moreover, while the active site of the chromosomal enzyme is a narrow binding groove where both reactants DHF and NADPH bind to different specific locations [12], R67 DHFR presents a spacious active-site cavity without a distinctive binding position for each ligand.

R67 DHFR is a homotetrameric enzyme, with a highly symmetrical D2 (point group 222) structure, where four monomers contribute to the assembly of a single central active-site pore that traverses the tetramer (Figure 1). The openings of this pore are ellipsoidal, with a major radius of 24 Å and a minor radius of 18 Å. The pore contracts towards the centre reaching 12 and 9 Å in the middle of the active site. This double funnel-like structure presents four equivalent binding sites which, due to steric constraints, can accommodate up to two ligands at the same time [13]. This includes combinations of two reactants (R67•DHF•DHF), two cofactors (R67•NADPH•NADPH), or one of each (R67•DHF•NADPH), the last one being the productive ternary complex. The reduction catalysed by this type 2 DHFR involves a hydride-ion transfer from the nicotinamide ring of NADPH to the N5-protonated pterin ring of dihydrofolate (DHFH+). As the enzyme itself does not seem to possess a proton donor group in the active site [14], it has been suggested that it may use DHFH+ as substrate [15]; hence, an accurate representation of the reactive complex should include the protonated substrate (R67•DHFH+•NADPH).

Even after much experimental work, including X-ray crystallography, interligand nuclear Overhauser effects (NOEs) and other NMR studies (Table 1 and Figure 2), the structure of a reactant-like complex (i.e. R67•folate•NADPH or R67•DHF•NADPH) has not been determined yet. Interligand NOE experiments suggest that while the pterin and nicotinamide rings of the ligands are located close to each other near the centre of the pore, the rest of the molecules extend towards opposite ends. Particular residues within the active site have been found to be important for the overall catalytic process, including K32, Q67, I68 and Y69. The four copies of these residues, located in equivalent monomers, have been proposed to be involved in the binding of both folate/DHF and NADPH. Despite these advances there is not a clear understanding of the way in which R67 DHFR interacts with the ligands to catalyse the reaction. The empirical information available is not adequate for positioning the ligands within the active site. The main difficulty arises from the symmetry of the enzyme (Figure 1c), which presents two equivalent residues on each side of the

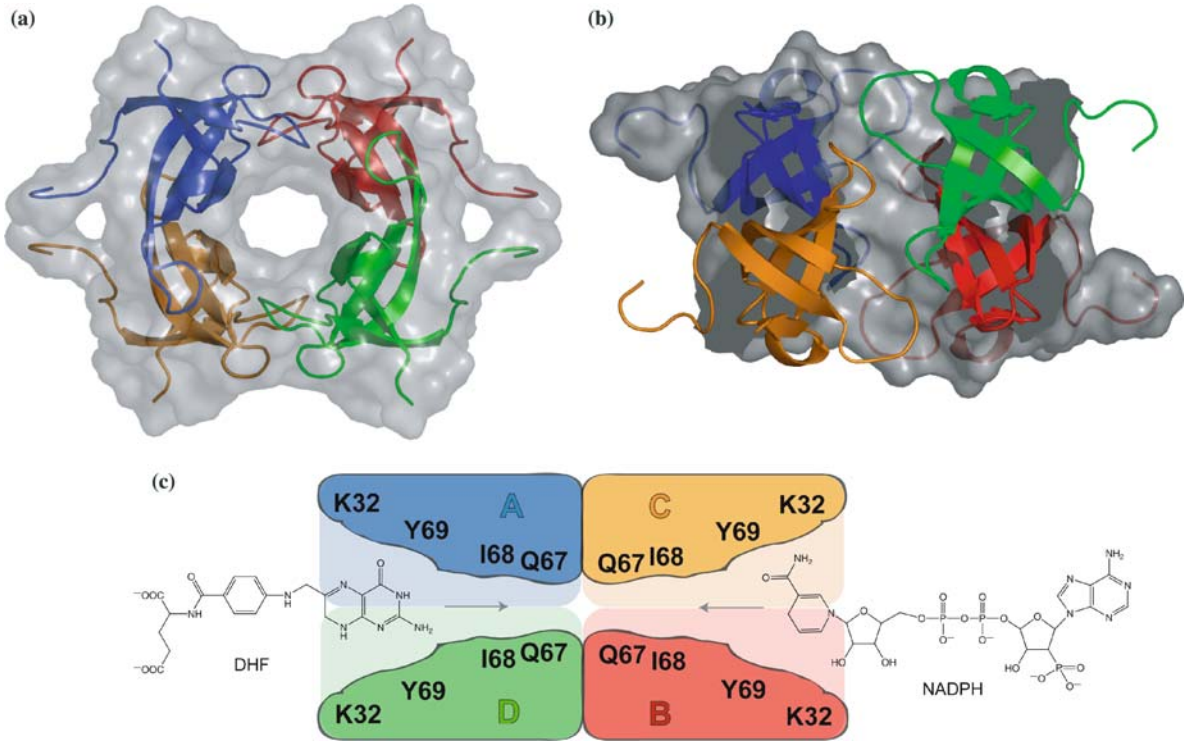

Figure 1. Structure of R67 dihydrofolate reductase (DHFR) [14]. The four monomers that form the active enzyme are depicted as differently coloured cartoons (chain A: blue, chain B: red, chain C: orange and chain D: green), within a surface representation of the tetramer. (a) Top view of the structure: the active-site pore is located at the centre of the complex, where four monomers contribute to the binding of the ligands. (b) Side view of the enzyme: the surface representation of the molecule has been limited to half of the tetrameric structure to facilitate visualisation of the double funnel-like central active-site pore. This presents elliptical openings with a major radius of 24 Å and a minor radius of 18 Å and contracts towards the centre reaching 12 and 9 Å, respectively, in the middle of the active site. (c) Schematic representation of the tetrameric structure of R67 DHFR and the reactants DHF and NADPH. Amino acids suggested by experiment to be important for the binding of the ligands have been roughly positioned along the active-site pore. Each side of the pore, A–D or C–B, presents two sets of equivalent residues and is expected to accommodate one of the two ligands, DHF or NADPH.

pore, each capable of establishing similar interactions with the ligands. The possibility of non-unique binding modes is further increased by the flexibility of the ligands, which can adopt multiple conformations within the central pore without violating the empirical restrictions. Therefore, in order to obtain a reliable structure it is necessary to thoroughly explore the conformational space of the ligands within the enzymatic cavity, and afterwards select the most stable poses.

The first attempt to generate a theoretical model of the complex R67•DHF•NADPH was carried out by Howell et al. [21]. They used the docking programs DOCK [23] and SLIDE [24] to produce structures of a R67•folate•NMN (nicotinamide mononucleotide moiety of NADPH) complex together with a R67•pterin•NADPH complex [21]; a structure for the complete reactive complex was not reported. These partial models were not analysed for stability nor was their reliability tested with other methodologies (e.g. molecular dynamics, MD). Many new experimental studies have been conducted on R67 DHFR since the publication of this initial study [17, 19, 20, 25–27], providing new insights into the mechanism R67 DHFR uses to catalyse the reaction. Some of this new data, such as the interaction between the hydroxyl group of Y69 and NADPH observed in mutational analysis [25], is not in agreement with the previous docking results [21] and, therefore, this model requires re-examination.

In the present study, we used the AutoDock [28] and FlexX [29] docking programs to explore the conformational space of the ligands, protonated DHF (DHFH+) and NADPH, inside the active site of R67 DHFR. The different poses were further

Table 1. Experimental observations pertinent to the conformations adopted by the ligands DHF/folate and NADPH/NADP+ within the active site of R67 DHFR.

Ligand/residue	Observation	Method
Folate/DHF	Only the position of the pterin ring could be determined crystallographically. The O4 atom appears H-bonded to the backbone of I68	X-ray crystallography [14]
	The glutamate moiety does not have a defined binding orientation	Interligand Overhauser Effects [16]
	It is involved in at least one ionic interaction with the enzyme, probably K32, at low salt concentrations	Isothermal titration calorimetry, ionic strength effects [17]
	Important ILOE peaks connecting H9 (folate) with H4 and H5 of NADP+	Interligand Overhauser Effects [16]
	During the reduction the hydrogen atom is transferred to the C6 *si* face of the ring	*In vivo* growth assay of *E. coli* strain C600 [11]
NADPH/NADP+	Adopts an extended conformation when bound to the enzyme	Interligand Overhauser effects [16]
	The ribonicotinamide bond adopts a *syn* conformation while the adenosine moiety adopts an *anti* conformation	NMR, Interligand Overhauser Effects [16, 18]
	The diphosphate and the phosphate groups participate in two ionic interactions with K32 residues	Isothermal titration calorimetry, ionic strength effects, fluorescence quenching [17]
	The diphosphate subunit may interact with Y69	NMR [19]
	Its binding causes large chemical shifts in the signal of the side chain of Q67	NMR [19]
	Its binding to the enzyme is accompanied by minor changes in the backbone	NMR [19]
	The *pro-R* H4 hydrogen of the nicotinamide ring is the one transferred during the reaction	NMR [11]
K32	Interacts with both ligands. May confer a general positive potential to the pore	Mutational analysis, isothermal titration calorimetry, ionic strength effects [17, 20]. Electrostatic potential calculations [21]
Q67	Appears interacting in pairs with a symmetry-related Q67 residue. Undergoes large chemical shifts upon NADPH binding	X-ray crystallography [14], NMR [19]
I68	Its backbone forms H-bond interactions with the carbonyl group of the pterin ring	X-ray crystallography [14]
Y69	Its hydroxyl group may interact with both ligands	Mutational analysis [20, 22]

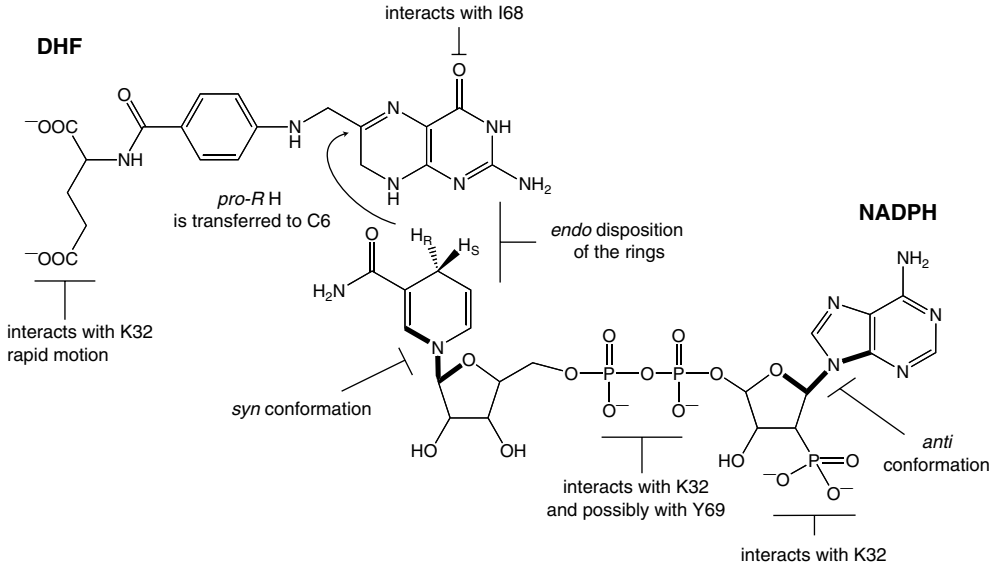

Figure 2. Schematic representation of the reactants DHF and NADPH and the interactions they establish with the enzyme R67 DHFR, as observed experimentally.

ranked using a consensus scoring analysis and three different scoring functions. Finally, the best conformations were submitted to MD simulations, to test their stability over time. The results of the MD simulations provided further insight regarding the behaviour of the ligands within the protein environment, the flexibility of key residues involved in protein–ligand interactions, and the overall strategy used by the enzyme to accommodate the reactants within its spacious active-site pore.

Materials and methods

Docking

R67 DHFR exhibits a unique active-site pore where both reactant and cofactor bind to equivalent sites. The size and symmetry of the cavity and the importance of cooperativity between ligands for the formation of a productive ternary complex [13] makes it necessary to adopt non-conventional docking approaches. Although there have been several recent reports comparing the performance of various docking programs and scoring functions [30–36], the results of these suggest that the performance of different docking tools is very dependent on the target protein and that comparative analysis may lead to different results depending on the properties examined (quality of the top-ranked poses, quality of all poses, predicted binding energies, docking time, etc). As most of these studies focus their attention on the screening of virtual libraries, the settings of the programs are optimised for fast docking rather than docking quality and the systems analysed include mainly small-sized molecules and proteins with simple active sites, excluding difficult cases such as R67 DHFR. We chose the docking programs AutoDock [28], which implements a genetic algorithm, and FlexX [29], with an incremental construction algorithm, to carry out the docking analysis. Both programs have been used successfully to dock several ligands [37–41], including methotrexate, a folate analogue, to the active site of a type 1 DHFR [28, 41, 42]. Moreover, AutoDock has been shown to be adequate for predicting multiple binding modes [43], while FlexX is claimed to be particularly good for studies where the active site is predominantly lipophilic but features polar groups that must form specific interactions with the ligands [32], as is the case for R67 DHFR.

Three different docking strategies were pursued, comprising: a sequential docking, where one molecule/moiety was docked after another; a superligand approach, where both DHFH+ and NADPH were covalently bonded and docked as a single entity; and a manual placement, where the

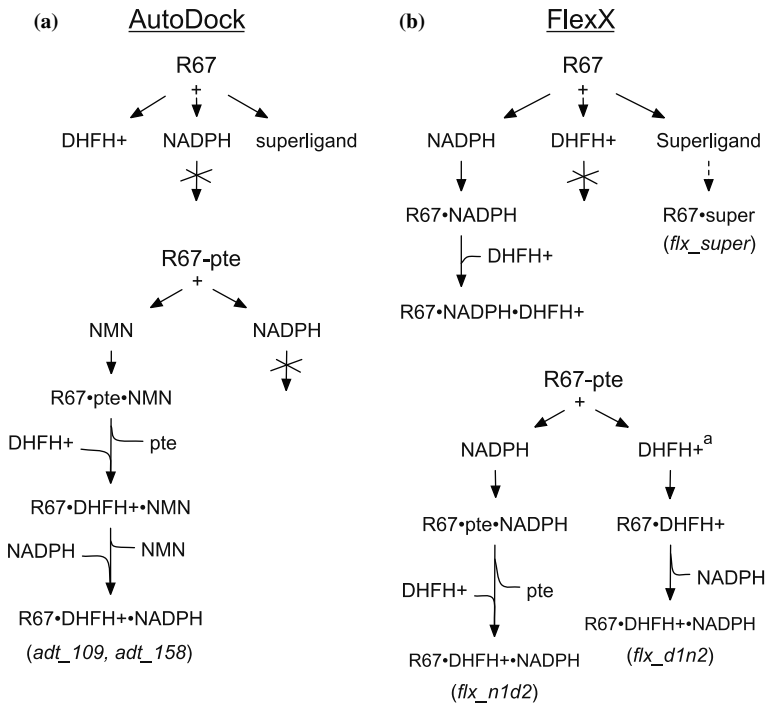

Figure 3. Docking sequence performed with AutoDock [28] (a) and FlexX [29]. (b) Crossed arrows indicate unsuccessful attempts, while dotted arrows denote doubtful results. Those complexes shown in italics correspond to the structures that were further analysed, including complexes adt_109 and adt_158 obtained with AutoDock and complexes flx_d1n2, flx_n1d2 and flx_super generated with FlexX. R67-pte, R67 DHFR including the crystal position of the pterin ring of folate; NMN, nicotinamide-ribose-phosphate moiety of NADPH; pte, pterin ring. [a]This corresponds to a guided docking of DHFH+ into R67 DHFR, taking the crystal position of the pterin ring as reference.

ligands were manually guided within the active-site pore (Figure 3).

Structure preparation

There are only two R67 DHFR crystal structures available, one of the *apo* enzyme (1VIE), and one of a R67•folate complex (1VIF) [14]. A comparison of these structures shows minimal conformational change upon ligand binding (backbone RMSD 0.02 Å, sidechain RMSD 0.05 Å). Therefore, the crystal structure of R67 DHFR complexed with folate (1VIF) was considered an appropriate starting point for the building of the complexes without further modifications. As the 1VIF structure provided two alternative positions of the pterin ring, the one with the *si* face of the C6 centre exposed was selected (R67•pte), based on the stereochemistry of the reaction [11]. Structural coordinates for the ligands, protonated dihydrofolate (DHFH+) and NADPH, were obtained by modification of the crystal-structure coordinates of folate and NADP+ from 1RA2 [8]. A superligand complex was constructed by connecting the H atom to be transferred in the reaction to both C4 of NADPH and C6 of DHFH+.

AutoDock

AutoDock 3.0.5 [28], which uses a stochastic-search based algorithm, was employed for the generation of ternary complexes. Hydrogens and charges were added to the ligands using InsightII [44] (CFF91 force field). Autotors, part of the AutoDock docking package, was used to define the active torsions. Kollman charges were assigned to the enzyme using the AutoDock tools program, and the Addsol module was used to generate the volume parameters. Affinity grids of $60 \times 90 \times 60$

points separated by 0.375 Å were produced around the active-site pore using AutoGrid. The Lennard–Jones and H-bonding potentials supplied with the program were used to calculate the interaction grids.

The Lamarckian genetic algorithm (LGA) was used for the conformational search. Each LGA job consisted of 200 runs and the number of iterations of the pseudo-Solis and Wets local search was set to 300. The initial population was between 200 and 250 structures, while the maximum number of energy evaluations and generations was 10 million; hence, the energy evaluations were the limiting factor. These values were chosen based on the systematic study done by Hetényi and van der Spoel on the influence of different parameters in the docking of peptides to proteins, without prior knowledge of the binding site [45]. The rest of the parameters were set to default values. The final structures were clustered and ranked according to the native AutoDock scoring function.

FlexX

FlexX 1.13.5 [29], which implements an incremental construction algorithm, was the second docking program used. The active-site region was defined by the residues K32, A36, Y46, G64, S65, V66, Q67, I68 and Y69. Both automatic and manual anchor-base selections were performed. The placement of the anchor bases was done using the triangle hashing technique and manual base placement; in the last case the crystal position of the pterin ring, R67•pte, was used as the reference structure. The final poses were clustered and scored using the basic scoring function of FlexX (similar to the Böhm function [46]). Use of the particle positioning option, to model possible water molecules important for the binding process [47], was also employed, but no important differences were found (results not shown).

Manual docking

InsightII [44] was used to manipulate one of the ternary structures obtained with AutoDock in order to change the conformations of DHFH+ and NADPH ligands to those in conformity with all the experimentally observed properties (see below). The final structures were optimised by a short energy minimisation cycle (100 steps of steepest-descent) with GROMACS [48, 49].

Analysis of the structures

The generation of a series of ternary structures through the use of different methods produced a diverse set of complexes sufficient for further studies. In all cases, the final poses were analysed and selected taking into account all the available empirical information. Only the best-ranked structures of each docking approach that fulfilled most of the experimentally observed properties, with particular emphasis on those closely related to the hydride transfer step (e.g. distance between the reacting centres), were further studied (Figure 3). The constraints used for the analysis and selection of the best poses included:

- C6(pte)-C6(nic) distance: only structures with the reacting centres within 4 Å were further analysed.
- C6(pte)-H-C4(nic) angle: guided by the *endo* model for the reactant state proposed by Castillo et al. [11] The angle was required to be larger than 100°.
- C8A(pte)-C6(pte)-C4(nic)-N1(nic) dihedral: the *endo* conformation of the rings required dihedral angles between –90° and 90°.
- *anti* conformation of the adenosine glycosidic angle and *syn* conformation of the ribonicotinamide glycosidic angle of NADPH according to transfer NOE and interligand NOE (ILOE) studies [16, 18].
- H-bond interactions between the phosphate groups of NADPH and K32 [17, 27].

Using the coordinates for some of the published docked complexes for folate and NADPH, kindly provided by Dr E. Howell (University of Tennessee), a ternary structure was generated. The coordinates of folate corresponding to the highest scoring R67•folate•NMN complex generated by DOCK, and those of NADPH, from a R67•pte•NADPH complex consistent with NMR data, were combined into a ternary structure. Coordinates for the final R67•DHFH+•NADPH complex (howH), were generated by simple modifications of the protonation and reduction state of the pterin ring of folate.

Multiple scoring

The selected ternary complexes were ranked using three different scoring functions: AutoDock native scoring function [28], Böhm-like scoring function [46] and DrugScore function [50]. While AutoDock and the Böhm-like scoring functions belong to the so-called empirical group, where the binding free energy is calculated using a summation of terms fit to experimental binding constants, DrugScore is a knowledge-based potential, where distance-dependent interactions obtained from crystallographic data are used to evaluate the poses.

Due to the nature of the complex, two basic approaches were used for the scoring of the structures. A serial score, where the interaction between one of the ligands (DHFH + or NADPH) and the protein (R67 DHFR) was initially evaluated, followed by the scoring of the second ligand binding into the previously evaluated binary complex (R67•DHFH + ↔NADPH or R67•NADPH↔DHFH +), the sum of these two values being the overall score. And an independent score, where the interaction energy of each ligand with the protein was assessed independently (R67↔DHFH + , R67↔NADPH), and the sum of these values was considered as the total energy, i.e. interactions between the ligands were neglected. Therefore, the contribution to the total energy given by the interaction between the ligands was considered to be the difference between the serial and the independent scores.

Consensus scoring

The evaluation of the interaction energy between a docked molecule and its receptor is not precise, as the calculation often involves several approximations and the use of empirically or theoretically parameterised functions. Therefore, there is not a ''best'' scoring function, because different algorithms will behave differently in different situations. The results often depend on the nature of the complex itself, the type of scoring function, the parameters it uses to evaluate different poses, the training data set used during the parameterisation procedure, etc. Therefore, instead of relying on only one function, several algorithms can be used and their results combined into a single consensus score [51, 52].

The results of the three different scoring functions were combined into four different consensus scores. The simplest combination was the addition of the three ranks, giving the sum rank. The worst–best rank involved assigning each pose the second-worst rank value, while the sum of all but the worst score provided the so-called deprecated rank-sum. Finally, we performed a consensus score (cons-score) where the different poses were given one vote in favour if their score was within the best half (top four) or best third (top three) structures out of the eight complexes according to each function. The final cons-score was the combination of points for the best-third and the best-half criteria.

MD simulation

After performing the consensus score, all R67• DHFH + •NADPH complexes were further analysed using MD simulations. The GROMACS simulation package [48, 49] was used to study the evolution of the complexes with time. The simulations were carried out using the OPLS-AA force field [53, 54]. Charges and parameters for the ligands were assigned by fragment comparison to homologous parameterised structures, except for the reacting rings which were calculated at B3LYP/6-31G* level using N5-protonated 6-methyl-dihydropterin and 1-methyl-nicotinamide fragments as models.

The different structures were solvated in a triclinic box of SPC water molecules [55], leaving a minimum distance of 8.5 Å between the complex and the edge of the box. The final charge of the systems was –4; no counter ions were added as they were seen to form strong ionic interactions with the charged ligands during MD simulations (results not shown), perturbing their interactions with the protein. The Particle-Mesh Ewald (PME) method was used to treat long-range electrostatic interactions [56]. The solvated structures were submitted to a short energy minimisation process (100 steps of steepest-descent) and further optimised by performing a 100 ps MD simulation of the water molecules only. Finally, the complete systems were simulated for 4 ns each. All atoms were treated explicitly and the SHAKE algorithm was used to constrain the length of covalent bonds between hydrogen and heavy atoms. The time

step was 0.001 ps and the Coulomb and van der Waals cutoffs were set to 1.2 nm. The simulations were performed at constant volume and temperature. Initial velocities were assigned according to a Maxwell distribution at 300 K. The protein and ligands, and the water molecules were coupled separately to a bath at 300 K using the Nose–Hoover algorithm and a coupling time τ_T of 0.1 ps. The centre of mass motion was removed in every step, while the neighbour list update was done every 10 steps.

An initial 1ns simulation of the *apo* enzyme was carried out in order to test the stability of the system under the chosen conditions. The different trajectories were analysed using several auxiliary programs provided with the GROMACS package. These included, *g_bond* for the calculation of distances, *g_angle* for the measurement of angles and dihedrals and *g_hbond* for the H-bond interactions between ligands and protein. In the last case, all the possible donors and acceptors in both the ligands and the protein were taken into account, and a distance of 3 Å and an angle of 60° were used as the cutoff for the interaction.

Results

Docking

The main challenge of this docking study was to position two different molecules (DHFH + and NADPH) within one common central pore, taking into account that the interactions between the molecules themselves may be critical for the overall binding process [13]. Therefore, two different strategies were developed to dock the ligands DHFH + and NADPH into R67: a step-by-step docking procedure, where one molecule/moiety was docked after another and a superligand approach, where the two molecules were covalently bound and docked as a single unit.

The ternary complexes R67•DHFH + •NADPH were constructed using the docking programs AutoDock and FlexX, as well as a manual approach. Assessment of the final poses was judged against the empirically observed properties (Table 1). Only those structures that presented reasonable interactions between the protein and the ligands, and between the ligands themselves, were studied further.

AutoDock docking

Only a few of the docking approaches tried with AutoDock worked (Figure 3). No sensible structures were obtained when docking DHFH + or NADPH into the *apo* enzyme. It was necessary to initially take into account the crystallographic position of the pterin ring in the enzyme structure (R67•pte) in order to obtain some reasonable structures. As the docking of the complete structure of NADPH into R67•pte did not produce any sensible conformations we proceeded to dock a truncated version of NADPH, a nicotinamide-ribose-Pi moiety (NMN). This approach gave good results: 25% of 200 final structures clustered in the largest group (rmsd smaller than 1 Å) with the lowest docking energy, all of them presenting reasonable conformations with a C6(pte)-C4(nic) distance shorter than 4 Å (Figure 4a). After selecting the best-ranked R67-pte•NMN complex, the pterin ring was removed and DHFH + was docked. Although half of the 200 final structures presented the pterin ring of DHFH + in close proximity to the crystallographic position, the positioning of the *para*-amino benzoyl glutamic acid tail (pABA-Glu) showed significant variability, in agreement with published experimental results [16, 57] and previous docking studies [21] (Figure 4b). The pABA-Glu tail adopted two major conformations, determined by the interaction of the glutamic moiety of DHFH + with one of the two K32 residues present in the half-pore region occupied by the reactant. There was an extended conformation, where both the pterin ring and the Glu moiety were interacting with the same dimer-dimer interface (D/C or B/A); and a bent one, where the pterin ring of DHFH + was interacting with the interface between chains D and C (or B and A) while the glutamic moiety was H bonded to the K32 residue of the other interface, B/A (or D/C). Raman studies of DHF in R67 [57] showed a split signal of the C6 = N5 stretching mode. It was suggested that this was the result of two possible binding modes, in which the substituents of the C6-R bond adopted different orientations. Also, recent asymmetric mutational analysis showed that double mutants where K32 from monomers A and B or A and C were changed to methionine exhibited similar kinetic parameters, leading to the conclusion that both topologies permitted similar interactions between K32 and

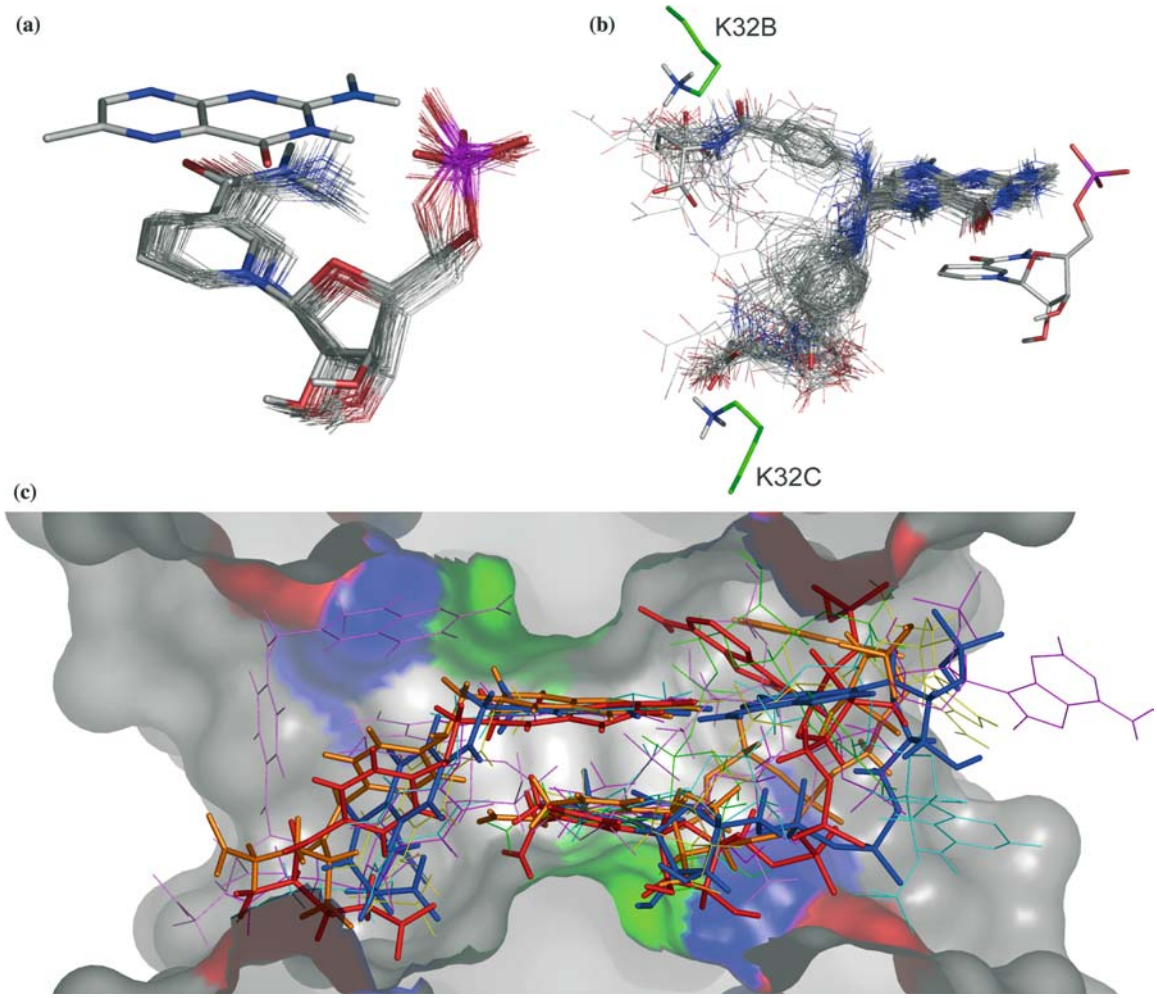

Figure 4. Relative positions of ligands in the docked structures. (a) Cluster of the highest scoring NMN conformations docked into R67•pte using AutoDock. This corresponds to the largest cluster of structures (rmsd smaller than 1 Å), which also exhibit the lowest average energy. The top scoring pose and the pterin ring are drawn in tubular form, whereas the remainder of the nicotinamide-ribose-Pi subunits are represented as wireframe. Only polar hydrogen atoms considered during the docking process are shown. (b) Orientations of different N5-protonated dihydrofolate conformers obtained after the docking of the substrate into R67•NMN using AutoDock. All structures present the pterin ring of dihydrofolate in a position close to that observed in the crystallographic structure. The *para*-amino-benzoyl-glutamic tail, on the other hand, adopts two main conformations depending on the location of the K32 residue with which the glutamic group is interacting (monomers B and C in this case). A representative of each of these two main groups is shown in tubular form, with the remainder in wireframe. This dual positioning of the pABA-Glu tail agrees with Raman studies [57]. Only polar hydrogen atoms are shown. (c) Conformation of eight selected complexes of N5-protonated dihydrofolate and NADPH within the active site of R67 DHFR. Complexes adt_109 (red), flx_n1d2 (blue) and howH (orange) appear in tubular form, with the others in wireframe (adt_158: green, flx_d1n2: yellow, flx_super: violet, man_react: magenta, and man_TS: cyan). The active site of the protein has been included as a trimmed surface, where the position of key amino acids has been highlighted in different colors, K32: red, Q67: green, I68: white and Y69: blue. It can be seen that while the pterin ring of DHFH+ and the nicotinamide ring of NADPH are located in similar positions (except flx_super), the tails extend towards opposite ends adopting several different conformations.

DHF [58]. All these observations are in clear agreement with our results.

Finally, after selecting the best R67•DHFH+•NMN complex, the nicotinamide mononucleotide unit was removed and NADPH was docked. This new placement provided very few structures with a reasonable overall conformation; there were no two conformers with the same positioning of the tail. These results, combined with the failed docking of the superligand molecule,

suggest that the size of the docked molecules could be a limiting factor in the generation of reliable complexes using AutoDock. The ligands, especially NADPH and the superligand, have many torsional angles, 18 and 27, respectively, and, therefore, the conformational space that the program has to explore might be too large to be properly handled by the LGA algorithm. On the other hand, when small molecules were used, NMN (6 active torsions) or DHFH+ (11 active torsions), there was less variability within the results and they were in good agreement with the experimental data. Two of the final R67•DHFH+•NADPH complexes were chosen and selected for further analysis (see below).

FlexX docking

FlexX performed better than AutoDock when docking NADPH into the *apo* enzyme. The use of an incremental growth algorithm may have provided a better treatment of the large number of active torsions of NADPH compared with the LGA approach of AutoDock. However, no poses were obtained for the automatic docking of DHFH+. When the crystallographic position of the pterin ring was used for the docking of NADPH, or for the manual guidance of DHFH+, some reasonable complexes were generated. Selected structures from the guided positioning of DHFH+, 15 in total, and the automatic docking of NADPH, 6, were further used to dock the second ligand. All the final complexes obtained were automatically screened to find those that best fulfilled the experimentally determined characteristics. There was a lot of variability among the tertiary complexes and generally the most sensible poses were poorly ranked and were not part of a big cluster, as in AutoDock. The docking of the superligand did not provide very clear results; very few structures conformed to experimental characteristics. The best poses were selected for further study (Figure 3).

Due to the heterogeneity of structures resulting from the docking studies, we also attempted a manual docking, taking into account all the experimentally observed properties. This variability among final poses was also observed by Howell et al. in their original construction of partial complexes using the docking programs DOCK and SLIDE [21]. Therefore, we considered that the inability of the docking programs to find a unique structure is likely due to the features of the system rather than deficiencies in the docking algorithms themselves.

Two different complexes (man_react and man_TS) were generated, the interactions between the rings were adjusted based on the structures published by Castillo et al. [59] for an *in vacuo* model of the reaction. The main difference between the two structures, apart from that of the reacting rings, was the positioning of the phosphate groups and adenine ring of NADPH.

A superposition of the best structures of each successful docking approach showed important differences in the conformations of the ligands, especially of the 2′,5′-ADP tail of NADPH (Figure 4c), even though all of them conformed to the empirical observations. The variability among the final poses, and the lack of more detailed experimental data made discrimination of the best structure far from trivial. We carried out some comparisons in order to test the stability and feasibility of the eight possible ternary complexes; these comprised the two structures generated with AutoDock (adt_109 and adt_158), three with FlexX (flx_d1n2, flx_n1d2 and flx_super), two manually constructed complexes (man_react and man_TS) and the structure built from published results (howH).

Comparative scoring

Three scoring algorithms were used to rank the structures: the AutoDock native scoring function [28], the Böhm-like scoring implemented in FlexX [41] and DrugScore [41, 50]. As two different molecules were docked inside the same active site, two different strategies were employed to evaluate the poses: a serial scoring and an independent scoring (refer to Materials and methods for details). In the serial scoring of the eight different complexes, the order in which the ligands were scored did not influence the final result (data not shown). Although the energy of the independent scores was higher (less negative) than that of the serial ones, the relative order of the complexes was not altered (Figure 5a). The contribution to the final energy given by the interaction between the two ligands was very uniform among the different structures according to DrugScore, with average values of –14 kJ/mol, while AutoDock and the

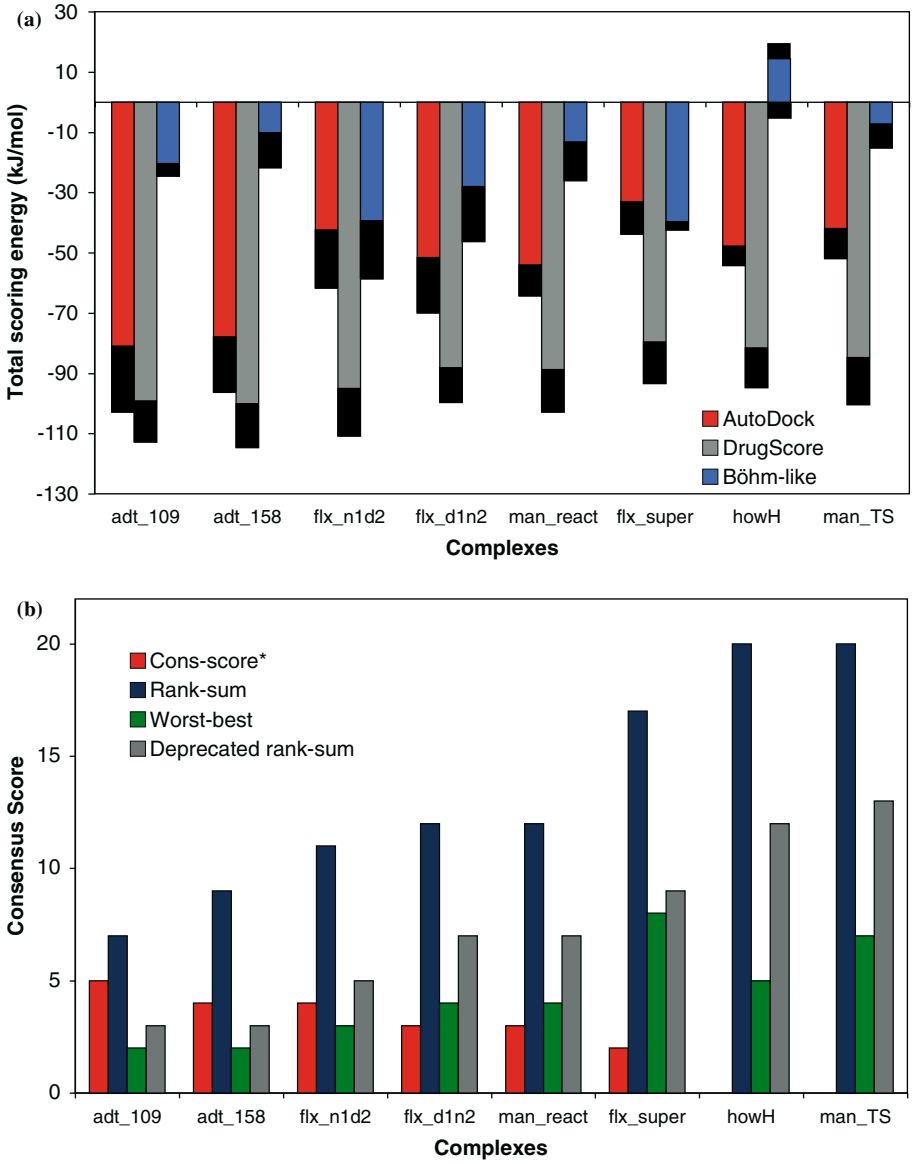

Figure 5. Scoring analysis of docked structures. (a) Final docked energy after serial scoring of different selected poses using three different algorithms: AutoDock native scoring function [28], and DrugScore [50] and Böhm-like [46] algorithms implemented in FlexX. The interaction energy of N5-protonated dihydrofolate with R67 DHFR was first computed, followed by the scoring of the NADPH moiety into the binary complex R67•DHFH+. Shaded regions represent the contribution to the final energy given by the interaction between the two ligands, calculated as the difference between the total energy of the serial and the independent scores. (b) Consensus scoring, the results of the serial scoring procedure were combined into a cons-score, rank-sum, worst-best and deprecated rank-sum. *Cons-score is the only consensus score where a bigger score corresponds to a better structure.

Böhm-like function predicted a greater range of interligand interactions energies. In the case of AutoDock, the values ranged from –6.6 kJ/mol for howH to –22.1 kJ/mol for adt_109, and in the case of the Böhm-like function from –4.3 kJ/mol for adt_109 to –19.5 kJ/mol for flx_n1d2. These values represent around 13% of the total docked energy in the case of DrugScore, about 20% for AutoDock, and up to a 50% in the case of the Böhm-like function. This significant variability among the different functions could be reflecting the fact that they have been optimised to estimate the interaction energies between ligands and protein and not between ligands themselves.

Overall, the three scoring functions gave different rankings (Figure 5a). The best correlation was observed between AutoDock and DrugScore ($r = 0.82$), whereas the results of the Böhm-like function were considerably different. While AutoDock and DrugScore favoured the complexes generated with AutoDock (adt_109 and adt_158), the Böhm-like function identified the FlexX structures as the best ones (flx_n1d2 and flx_d1n2). It is interesting to note that the structure constructed from coordinates of the published partial complexes is one of the lowest scoring poses, and it is the only one to yield a positive energy with the Böhm-like scoring function.

Consensus scoring

In order to resolve the disparities among the scoring functions, and identify the best ternary complex out of the eight selected structures, the individual scores were combined into four different consensus rankings: cons-score, rank-sum, worst-best rank and deprecated rank-sum [51] (Figure 5b). The results of all these combinations agreed on the overall final ranking of the structures, giving the following order of structures from best to worst: adt_109, adt_158, flx_n1d2, flx_d1n2, man_react, flx_super, howH and man_TS.

MD simulations

After the consensus scoring, all the complexes were further analysed by conventional MD simulations. The structures were solvated and simulated under periodic boundary conditions for 4 ns using the GROMACS MD package [49]. In order to evaluate the stability of the complexes, and the likelihood of their leading to a productive transition state, several parameters were monitored during the simulation. These included the interactions between the nicotinamide and pterin rings, such as the distance between the two reactive centres C6(pte) and C4(nic), the angle C6(pte)-H-C4(nic) and the dihedral C8A(pte)-C6(pte)-C4(nic)-N1(nic), and the H-bond interactions between the ligands and the protein.

Movement of the protein and the ligands during the simulations

The *apo* enzyme was initially simulated for a period of 1 ns to test the stability of the system. The structure did not exhibit major deviations from the crystal-structure reaching a global backbone RMSD of 3 Å while the active-site backbone presented a stable RMSD of 1.5 Å; similar values were obtained for all other complexes (see Figure 6 as an example). The conformations of key residues, including K32, V66, Q67, I68 and Y69 (see below) were also analysed along the different trajectories. Most of these presented conformations comparable with those found in the crystal structure; the only residues that showed variable torsional angles were K32 and Q67, which participate in different H-bond interactions with the ligands, as, discussed below. The movement of Q67 upon ligand binding has been observed experimentally (Table 1).

To analyse the movements of the ligands, DHFH+ and NADPH were each divided into three different regions, for which the RMSDs were tracked during the simulations, after superposition of the active-site backbone (Figure 6a, b). It may be seen that while the reacting rings, PTER and NICO, exhibit deviations comparable with those of the backbone, the tails of the molecules display much greater fluctuations. Movement of the ligand regions becomes more evident towards the extremities of the molecules, GLU and RIBADEPHO, which are situated close to the openings of the active-site cavity. The same pattern was observed for all but two simulations, flx_super and flx_n1d2, where the RMSD of the rings was comparable with that of the other regions of the ligands, indicating the instability of these conformations.

Interactions between the reacting pterin and nicotinamide rings

There is enough experimental evidence to support the idea that the disposition of the reacting rings within R67 DHFR is quite different from that found in the chromosomal enzyme. Although the pterin and the nicotinamide rings are known to adopt an *endo* conformation in R67 DHFR there are no studies on the mechanism of the reaction and, therefore, there is not a clear picture of the global arrangement of the reactants. There is only one study where an *endo*-transition state for a model of the reaction *in vacuo* was presented as an alternative to the *exo* conformation found in the ubiquitous type 1 DHFR [16, 59]. According to this model, the reactant-like complex presents a C6(pte)-C4(nic) distance of 4.379/4.974 Å (AM1

and HF/4-31G, respectively), a C6(pte)-H-C4(nic) angle of 128.63/108.82° (AM1 and HF/4-31G) and a C8A(pte)-C6(pte)-C4(nic)-N1(nic) dihedral angle of 117.2/62.9° (AM1 and HF/4-31G). Upon formation of the transition state, the distance between the carbon atoms shortens to about 2.7 Å, the angle increases to approximately 168° and the dihedral angle between the two rings changes to about 350°. Although these calculations were done for an *in vacuo* model and neither AM1 nor HF/4-31G methods can be expected to provide very accurate results, they should be adequate to estimate the range of distances (4.5 ± 1.5 Å), angles (> 100°) and dihedrals (50 ± 60°) that would characterise a productive ternary complex of R67•DHFH + •NADPH during MD simulations (Figure 7). We note that in the reported calculations good agreement was found between the same *in vacuo* model, but with the *exo* conformation of the rings, and a QM/MM model of the reaction within a type 1 DHFR.

C6(pte)-C4(nic) distance

Four out of the eight complexes (adt_158, flx_d1_n2, howH and man_TS) presented a C6(pte)-C4(nic) distance smaller than 5 Å after 4 ns of simulation (Figure 7a). While most structures showed a more or less stable interacting distance during the simulation, in adt_158 the distance decreased remarkably after 2.6 ns of simulation from 5.5 Å to about 3.8 Å. Therefore, all but one structure (flx_n1d2) had reasonable distances between the two reacting carbon centres after 4 ns of simulation.

C6(pte)-H-C4(nic) angle

Most complexes presented angles within the expected range, between 110° and 170° (Figure 7b). As can be seen, the structures that presented the shortest C6(pte)-C4(nic) distances (adt_158, flx_d1n2, howH and man_TS) are the ones with the largest angles, while the structures that exhibit the smallest angles, flx_super and man_react, are among the complexes with the longest carbon–carbon distances.

Dihedral angle C8A(pte)-C6(pte)-C4(nic)-N1(nic)

This angle is one of the most stable parameters during the simulations (Figure 7c). All the complexes exhibited dihedral angles corresponding to

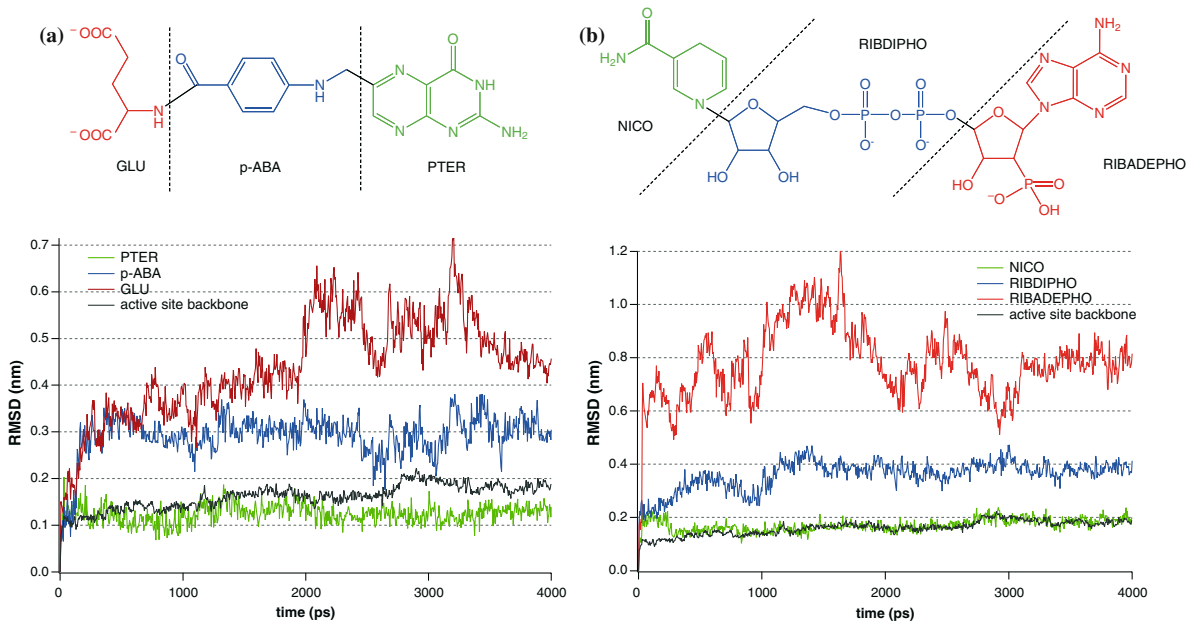

Figure 6. RMS deviations of different regions of the ligands DHFH + (a) and NADPH (b) after superposition of the backbone of residues lining the active-site pore (K32, A36, Y46, G64, S65, V66, Q67, I68 and Y69). These values correspond to the 4 ns MD simulation of the complex howH. All simulations except flx_n1d2 and flx_super presented the same pattern.

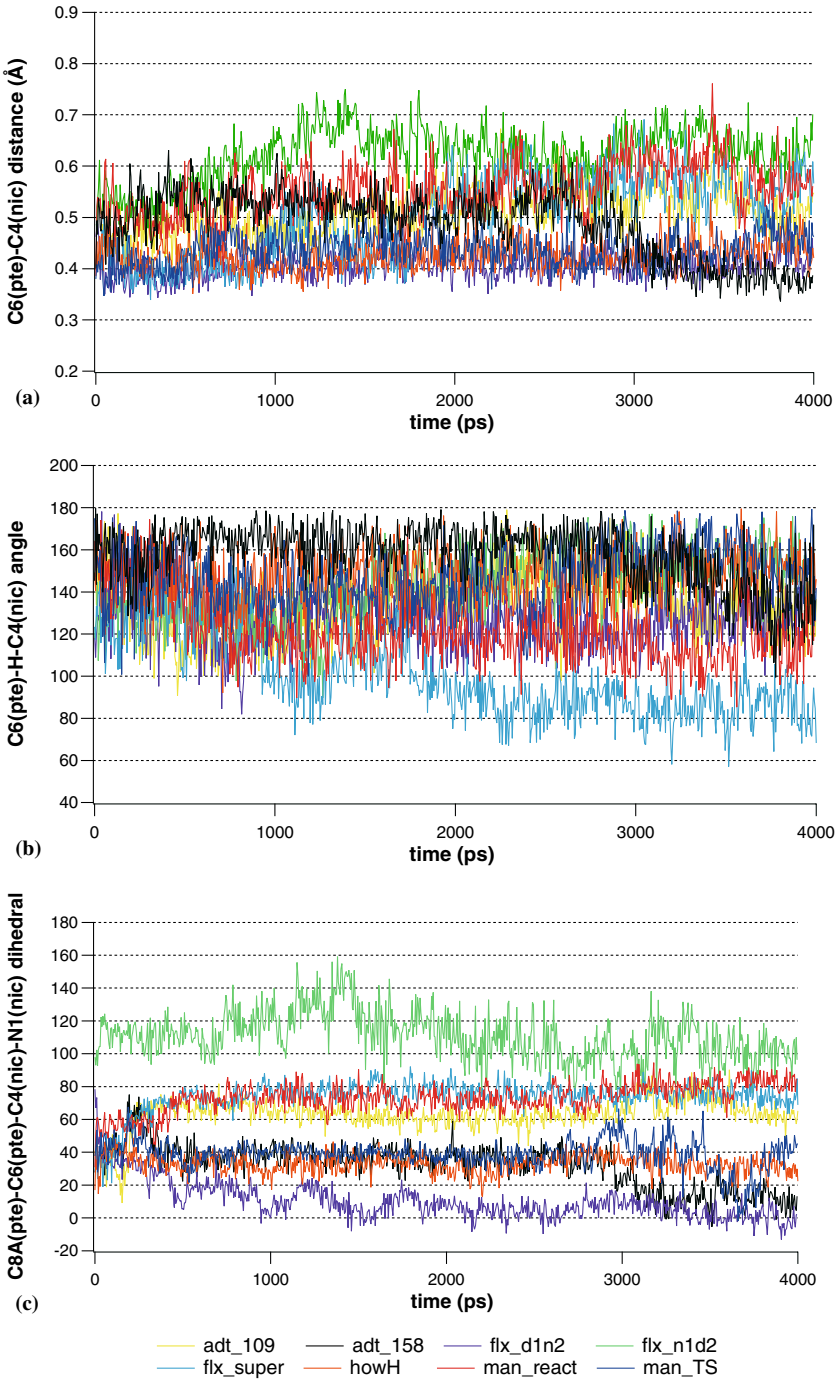

Figure 7. Molecular dynamic simulations of selected complexes. The changes in C6(pterin)-C4(nicotinamide) distance (a), C6(pterin)-H9-C4(nicotinamide) angle (b) and C8A(pterin)-C6(pterin)-C4(nicotinamide)-N1(nicotinamide) dihedral angle (c) in eight selected complexes were monitored during 4 ns simulations. Complexes adt_109 and adt_158 were originally generated using AutoDock; complexes flx_d1n2, flx_n1d2 and flx_super were obtained with FlexX; complexes man_react and man_TS correspond to manually docked structures; howH was constructed using published partial structures [21].

the *endo* conformation of the interacting rings. There is a very good correlation between the C6(pte)-C4(nic) distance and the dihedral of the complexes: those structures with the shortest distances (adt_158, flx_d1n2, howH and man_TS) are the ones that presented the smallest dihedrals, closest to 0°, while the structures with the longest distances (flx_n1d2, man_react, flx_super and adt_109) had the largest dihedrals. Moreover, as the C6(pte)-C4(nic) distance of adt_158 decreases after 2.8 ps of simulation, the inter-ring dihedral angle does so as well. It seems that for a closer interaction to occur, the two rings have to adopt a specific relative orientation, where the atom centres of the two rings are almost on top of one another.

Overall, the values obtained for the interactions between the two reacting rings are within the expected range. Surprisingly, most of the complexes seemed to have an adequate and stable disposition of the rings after 4 ns of simulation. Nevertheless, when the average structures of the last 100 ps were compared important differences were found. Although the reacting rings were located at the centre of the pore, the general conformation of the ligands, particularly that of the tails, was significantly different across the different complexes.

Interactions between the ligands and the protein

After analysis of the interactions between the two reacting rings, the contacts between the ligands and the protein were studied (see Table 1 for experimental evidence). The presence of H bonds between the ligands, DHFH + and NADPH, and the enzyme, R67 DHFR, were monitored during the simulation (Figure 8). The most stable H-bond interactions were identified in all eight different simulations and are described below.

DHFH + - R67 DHFR interactions

Residues K32, Y46, T51, S65, V66, Q67, I68 and Y69 were seen to form stable interactions with DHFH + in at least one of the eight simulations. Those residues participating in interactions present in six or more cases included:

- K32: its interaction was present in all simulations. Its $NH3^+$ terminal group interacts with the glutamic moiety. It was a quite stable interaction during the simulations.
- V66: its interaction was present in all but one simulation (flx_super). The carbonyl group of the backbone appears to be interacting with HN5, HN8 or HN10, probably stabilizing the protonated structure and assisting in the positioning of the ring.
- I68: present in all but flx_n1d2. Interacts through its amide backbone (NH) with O4 of the pterin ring. This interaction was stable over all the simulation period.
- Y69: present in all but two cases (flx_super and howH). The hydroxyl group of the phenyl ring interacts with the glutamic group of DHFH +; it is not as stable as the interaction of K32.

NADPH - R67 DHFR interactions

The cofactor NADPH was found to have H-bond interactions with several residues among the 8 complexes, including K32, A36, G64, S65, V66, Q67, I68, Y69, P70, A72, and A73. Those interactions present in at least five of the simulations included:

- K32: interacts with NADPH in all simulations. Its $NH3^+$ group is interacting with the pyrophosphate group (PPi), the phosphate group (Pi) or both. The interaction is present most of the time, except for adt_158, where the time at which the interaction appears matches the closer approach of the rings.
- V66: appears in all but one simulation (flx_d1n2). In most cases its backbone forms a H bond with the ribose subunit bound to the nicotinamide ring or with the amide group of the nicotinamide ring. This is a stable interaction.
- Q67: present in 5 of the 8 cases, except adt_158, flx_d1n2 and flx_n1d2. The amide of this residue interacts with the adenine ring (adt_109), PPi (flx_super and man_react), and the nicotinamide ring (flx_super, howH and man_TS).
- I68: interacts with NADPH in 5 simulations, excluding flx_super, howH and man_TS. In three cases it has a very stable interaction with the amide group of the nicotinamide ring (adt_109, flx_d1n2 and man_recact), while in

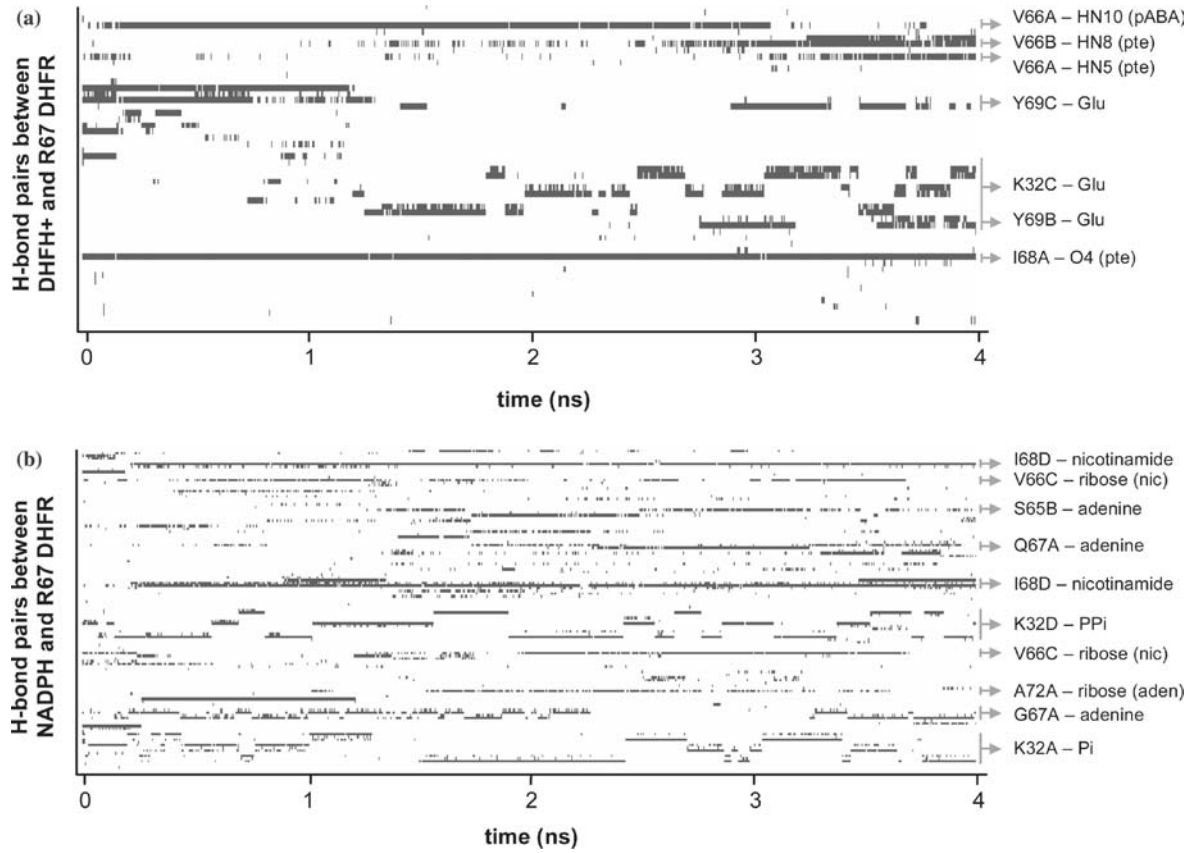

Figure 8. Hydrogen-bond interactions between the ligands, N5-protonated dihydrofolate (a) and NADPH (b), and R67 DHFR during 4 ns of simulation of the complex adt_109. The presence of an H-bond interaction at any time is represented by a black mark. The most stable and long lasting interactions are emphasised with grey arrows and specified by the number and chain of the residue involved and the group or atom of the ligand participating in the interaction. pABA; *para*-amino-benzoyl tail of DHFH+, pte: pterin ring, Glu: glutamic moiety; ribose (aden): ribose moiety bound to the adenine ring ; ribose (nic): ribose moiety bound to the nicotinamide ring; Pi: phosphate group; PPi: pyrophosphate group.

others it also interacts with PPi (adt_158) and the adenine ring (flx_n1d2).

The interactions observed for residues K32, Q67, I68 and Y69 are in clear agreement with experimental results (Table 1). The H-bond contacts between NADPH and R67 DHFR were more dynamic and varied than those observed for DHFH+. The same residues were seen interacting with different regions of NADPH in different complexes, suggesting that multiple configurations of the tail of NADPH were possible. When the average structures of the complexes during the last 100 ps of simulation were compared, it was clear that the characteristics of the active site allowed for a wide distribution of structures. While the conformation of DHFH+ was similar in the different complexes, the tail of NADPH adopted various positions. It is important to note that the same amino acids (K32, Q67, V66, I68 and Y69) were involved in nearly all interactions between both DHFH+ and NADPH and the protein.

Water distribution and water-mediated interactions

The active site of R67 DHFR is a spacious pore filled with water molecules. In order to analyse the changes in water distribution upon ligand binding, we determined the number of water molecules within a radius of 5 Å of those residues lining the active-site cavity (K32, A36, Y46, G64, S65, V66, Q67, I68 and Y69) in the *apo* form, and in the different R67•DHFH+•NADPH complexes

(Figure 9a). While the average distribution of water molecules remains relatively unchanged around those residues positioned close to the openings of the pore, there is an important displacement of waters at the central region of the cavity. While the pterin and nicotinamide rings adopt a stacked conformation at the centre of the active-site pore, where there is little if any water access, the tails of the ligands extend towards opposite ends in a solvent-rich environment (Figure 9b).

Although there are water molecules mediating interactions between the tails of the ligands and the protein, the multiple conformations lead to non-conserved interactions among different simulations. In the case of the reacting rings, only the pterin ring was seen to form a stable water-mediated interaction involving N3 and the residues Q67 or I68.

Discussion

Despite catalysing the reduction of DHFH+ using NADPH as a cofactor, R67 DHFR cannot be considered a typical enzyme. It is not the kind of reaction it promotes which makes this such an unusual enzyme, but its origins and characteristics which set it apart. Although its structure and mechanism of action have been the focus of many studies, details of the way in which it binds both the reactant and the cofactor and promotes the reaction remain unknown. The elusiveness of the system to experimental definition prompted us to perform a computational dissection of R67 and its ligands in order to define how the enzyme accommodates the reactants within the active site.

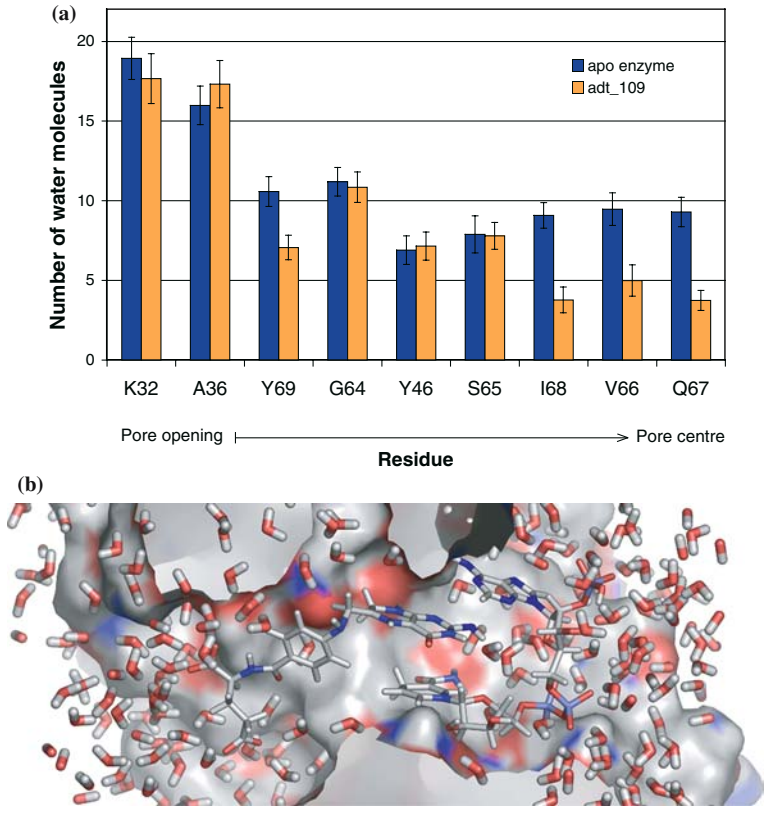

Figure 9. (a) Average water distribution along the active-site cavity for both the *apo* and ligand-bound enzyme. The number of water molecules within a radius of 5 Å from selected active-site residues (K32, A36, Y46, G64, S65, V66, Q67, I68 and Y69) was averaged over 1ns for the simulation of both the *apo* enzyme and the adt_109 complex. The order of the residues in the plot corresponds to their distribution along the pore, with those residues closer to the openings located on the left side and those at the middle of the pore on the right side. (b) Final disposition of the ligands inside R67 DHFR after 4 ns simulation of the complex adt_109. The surface representation of the protein has been clipped to make the central active-site pore visible. Water molecules within 4 Å of the ligands and the active-site pore have been included. Whereas the reacting rings occupy a central position with very little water access, the tails of the ligands are located towards the openings of the pore in a solvent-rich environment.

Interligand interactions

The initial modelling of the ternary complex R67•DHFH + •NADPH using docking techniques was not a trivial process due to the size and symmetry of the active-site pore. It was not possible to position the ligands *de novo* within the *apo* enzyme, but rather it was necessary to consider the presence of a core fragment of the partner, such as the pterin ring or the NMN moiety, to produce sensible structures. The need for the presence of a second molecule within the active site for the proper positioning of the ligands correlates very well with the interligand cooperativity observed experimentally [13]. The interaction between the two reacting molecules, particularly that of the pterin and nicotinamide rings, seems important for adequate binding and positioning of the ligands within the active-site pore.

The 4 ns MD simulations of eight different complexes provided further insights into the behaviour and stability of the ligands within the protein environment and both ligand-ligand and ligand–protein interactions important for establishment of a reactant-like R67•DHFH + •NADPH ternary structure. It was seen that most of the structures analysed presented reasonable distances between the reacting centres after 4 ns of simulation, having the C6(pte) and C4(nic) atoms less than 6 Å apart. The structures with the shortest distances (adt_158, flx_d1n2 and howH) showed larger C4(pte)-H-C6(nic) angles and smaller inter-ring dihedrals than those of the complexes with longer C6(pte)-C4(nic) distances (flx_super, man_react and flx_n1d2). The complexes flx_super and flx_n1d2 presented not only the longest distances but also the most significant fluctuations during the MD simulations, suggesting that these two structures do not correspond to stable conformations. Thus, the MD simulations allowed us to discriminate between structures which despite showing conformity with the empirically observed constraints (see Figure 4c for a comparison between adt_109, howH and flx_n1d2), present dissimilar stabilities.

Other workers have also explored the use of MD simulations to evaluate the relative stability of different complexes identified from a docking protocol. This approach offers a more detailed comparison among several structures as solvation effects and the flexibility of the protein can be taken into account explicitly. Some recent work in which MD simulations were used to validate or refine docked structures include a study on trypanothione reductase and glutathione reductase [60], the binding of the inhibitor propidium to the peripheral anionic site of human acetylcholinesterase [61], a combined docking approach presented by Moitessier et al. [62], and the binding of glucose to insulin [63]. In all these cases the MD simulations provided useful information on the stability of different binding positions. Docking calculations and MD simulations are, thus, complementary, with the first providing a list of more or less reliable starting structures and the second allowing further discrimination on the basis of relative stability after relaxation of the molecular complexes. It is likely this approach will be used more in the future literature.

Protein–ligand interactions

Despite most of the structures showing very similar interactions between the two reacting rings, the general conformation of the tails of the ligands, particularly that of NADPH, were significantly different. Nonetheless, when the H-bond interactions between the ligands and the protein were studied it was seen that in most cases the same amino acids were engaged, supporting the idea of an active site where the same residues can accommodate multiple binding conformations of the tails of the ligands. Moreover, the same residues were found interacting not only with different conformations of a single ligand, but with both DHFH + and NADPH. This dual-binding role played by equivalent amino acids has been previously suggested [22], and is now supported by the results of our MD simulations.

The most important residues involved in the protein–ligand interactions include K32, V66, Q67, I68 and Y79. The interaction of K32 and Y79 with the charged glutamic and phosphate groups of the pABA-Glu and 2′,5′-ADP tails clearly agree with the available experimental data (Table 1) offering, therefore, a better model of the ternary complex than that previously published [21]. Residues V66, Q67 and I68 are located near the centre of the pore and interact with the reacting rings.

Despite the importance of the direct interaction between DHFH + and NADPH for the overall binding process and the correct positioning of the

rings, the roles of residues V66, Q67 and I68 should not be discounted. The peptide carbonyl group of V66 appears H bonded to HN5 of the pterin ring, probably playing a key role in the stabilisation of the reactive protonated structure. It also establishes H-bond interactions with the hydroxyl group of the ribonicotinamide moiety of NADPH, helping in the correct positioning of the cofactor ring. Residues Q67 were not observed to adopt the paired H-bonded conformation described in the crystal structure, but instead the amide groups formed H-bond interactions with the reacting molecules in most cases. This change in conformation of Q67 upon ligand binding agrees with NMR analysis (Table 1). Interactions of the pterin ring and the nicotinamide ring with Q67 were found in five complexes; these H bonds were more varied and less stable than those observed for V66. The NH backbone group of I68 formed one of the most stable interactions seen with O4 of the pterin ring, which may indicate its importance in the correct positioning of the ring. In the case of NADPH, I68 interacts through its backbone with the nicotinamide ring, again, assisting favourable placement of the reacting group. The only persistent water-mediated interaction among different simulations involves the N3 atom of the pterin ring and residues Q67 or I68.

Multiple conformations

Other studies have shown that a particular ligand may bind to an enzyme by adopting flexible conformations. Recent studies have suggested that the increased flexibility that many inhibitors show when bound to the active site [64–66] may be important for the design of new and more potent drugs, which are less susceptible to usual enzyme mutations producing drug resistance. Not only drugs, but also natural substrates of enzymes have been found to adopt multiple binding forms. Many enzymes exhibit quite a wide substrate specificity, and, therefore, have evolved to accommodate a broad variety of reactants [67–69].

The possibility of multiple conformations is associated with the flexibility of ligands, which can adopt different orientations to adjust and maximise the interaction with a given conformation of the active site at a low energy cost. This adaptation, and the variety of probable conformations, will depend also on the properties of the active site and its capacity to accommodate different ligand positions. The arrangement of the monomers in the ternary structure of R67 results in the presence of two symmetry-related equivalent residues within each half of the pore, either of which can establish interactions with DHFH+ or NADPH. This factor, combined with the spaciousness of the active site, can account for the possibility of multiple binding modes. This potential was clearly seen during both docking and MD simulation studies. When docking DHFH+ into R67•NMN (Figure 4b), the glutamic moiety was found to adopt two alternative conformations, interacting with one of two equivalent K32 groups from different monomers; this agrees with Raman studies of DHF in R67 [57]. The docking of NADPH provided a large variety of structures, none of which presented the same conformation of the 2′,5′-ADP moiety. MD simulations of the complexes provided further evidence of their mobility within the active site. While the reacting rings located at the centre of the pore present deviations similar to those of the backbone, the tails of the ligands show significant fluctuations over the whole simulation time. Two structures did not show this pattern, flx_n1d2 and flx_super; both of these showed significant deviations for both reacting rings also. The flx_super complex was originally very different from the rest of the simulated structures, having the rings positioned outside the centre of the pore (Figure 4c); the MD simulation demonstrated this was an inadequate placement, showing it to be a useful tool to discriminate the most stable conformations from a set of docking solutions that were consistent with experimental data but non-unique.

The water distribution analysis clearly showed that upon ligand binding most of the water molecules located at the centre of the cavity are displaced by the reacting rings, while the average water distribution along the rest of the active site remains more or less constant. These changes leave the negatively charged tails within a solvent-rich environment while the reacting rings are positioned in a region with little if any water access.

The positioning and behaviour of the ligands within R67 DHFR is, therefore, very different from that observed in the Type 1 or chromosomal enzyme, which presents a narrow binding groove

where both molecules are positioned in different specific locations [12]. The residues lining these binding sites appear to have been optimised to accommodate each ligand selectively; therefore, the mobility of their tails is substantially hindered compared with that observed within R67 DHFR.

Our results are in good agreement with the available experimental information (Table 1). The mobility of the tail regions has been observed not only in both X-ray crystallography and NMR studies but is also in agreement with mutational analysis. Asymmetric mutations of Y69 indicate that variable positions of the pABA-Glu tail of DHF are tolerated [26], and studies of K32 showed that two mutations on different half pores produce topologies that allow comparable interactions between K32 and the Glu tail of DHF [20]. Furthermore, mutations at the centre of the pore (Q67H) are accompanied by inhibition of the catalytic activity due to non-productive binding [70], while mutations close to the pore surface (K32M and Y69F) are not.

In the present paper we have shown the importance of using several computational techniques when dealing with molecular systems where more than one conformation of the ligands seems possible. Docking approaches alone are not adequate for identifying the correct structure of an enzyme–ligand complex, even when comparative scoring analysis is used to select the best pose. Further studies using MD simulations are necessary to evaluate the evolution of the complexes with time and to test the stability and feasibility of the structures.

Role of R67 DHFR in the catalysis of the reaction

The size and shape of the active site, the position of charged residues, the lack of a direct proton donor or other residue directly involved in the catalytic process, and the multiple interactions that can be established between the protein and ligands support the idea of this being a recently evolved enzyme. Despite R67 DHFR lacking many of the characteristics of more efficient catalysts, R67 is an enzyme that uses a simple yet effective approach to achieve catalysis of the reaction. The double-funnel shape of the active site combined with flexible H-bond interactions between the protein and the ligands help to bring the reactants together and keep them within the pore, thus promoting the encounter of the reacting rings at the centre. V66, Q67 and I68 may contribute to the formation of the reactive complex by stabilizing and confining the stacked rings in the middle of the pore, where they adopt an *endo*-like conformation in an environment with little if any water access, while the charged pABA-Glu and 2′,5′-ADP tails extend towards the opening of the pore adopting multiple conformations in a solvent-rich environment, where K32 and Y69 play important roles (Figure 9b).

Conclusions

Our docking analyses and MD simulations clearly suggest that there is more than one possible conformation of the ligands DHFH+ and NADPH within the active-site pore of R67 DHFH that agrees with experimentally determined properties. The feasibility of these different complexes was further supported by MD simulations, which provided evidence of their stability in spite of the different positioning of the ligand tails. While the reacting rings adopt a stacked *endo* conformation in the vicinity of the centre of the active site, assisted by residues V66, Q67 and I68, the tails extend towards opposite ends of the pore adopting multiple conformations, which are stabilised by H-bond interactions with K32 and Y69 in a solvent-rich environment.

Acknowledgements

We thank Dr E. Howell for providing the atom coordinates of partial complexes of R67 DHFR with folate and NADPH, and sharing with us unpublished experimental results and for helpful comments. We are also grateful to BioSolveIT who provided us with an extended trial version of their docking program FlexX and analysis module DrugScore. We also acknowledge funding from the ANU IAS (Institute of Advanced Studies) block grant and Australian Research Council (ARC) grant DP0346292, and computing time allocations on the Australian Partnership for Advanced Computing (APAC) National Facility. Finally, we thank the anonymous reviewers for their constructive comments.

Note added in proof: A review comparing the different enzymic mechanisms of R67 and chromosomal DHFRs has recently been published (E.E. Howell, Chembiochem., 6 (2005) 590).

References

1. Pattishall, K.H., Acar, J., Burchall, J.J., Goldstein, F.W. and Harvey, R.J., J. Biol. Chem., 252 (1977) 2319.
2. Acar, J.F., Goldstein, F.W., Gerbaud, G.R. and Chabbert, Y.A., Ann. Microbiol. (Paris), 128A (1977) 41.
3. Zolg, J.W. and Hanggi, U.J., Nucl. Acids Res., 9 (1981) 697.
4. Radstrom, P., Skold, O., Swedberg, G., Flensburg, J., Roy, P.H. and Sundstrom, L., J. Bacteriol., 176 (1994) 3257.
5. L'Abee-Lund, T.M. and Sorum, H., Microbiol. Drug Resist, 7 (2001) 263.
6. Huennekens, F.M., Protein Sci., 5 (1996) 1201.
7. Fierke, C.A., Johnson, K.A. and Benkovic, S.J., Biochemistry, 26 (1987) 4085.
8. Sawaya, M.R. and Kraut, J., Biochemistry, 36 (1997) 586.
9. Grape, M., Sundstrom, L. and Kronvall, G., Microbiol. Drug Resist, 9 (2003) 317.
10. Rajagopalan, P.T., Lutz, S. and Benkovic, S.J., Biochemistry, 41 (2002) 12618.
11. Brito, R.M., Reddick, R., Bennett, G.N., Rudolph, F.B. and Rosevear, P.R., Biochemistry, 29 (1990) 9825.
12. Bystroff, C., Oatley, S.J. and Kraut, J., Biochemistry, 29 (1990) 3263.
13. Bradrick, T.D., Beechem, J.M. and Howell, E.E., Biochemistry, 35 (1996) 11414.
14. Narayana, N., Matthews, D.A., Howell, E.E. and Nguyenhuu, X., Nat. Struct. Biol., 2 (1995) 1018.
15. Park, H., Zhuang, P., Nichols, R. and Howell, E.E., J. Biol. Chem., 272 (1997) 2252.
16. Li, D., Levy, L.A., Gabel, S.A., Lebetkin, M.S., DeRose, E.F., Wall, M.J., Howell, E.E. and London, R.E., Biochemistry, 40 (2001) 4242.
17. Hicks, S.N., Smiley, R.D., Hamilton, J.B. and Howell, E.E., Biochemistry, 42 (2003) 10569.
18. Brito, R.M., Rudolph, F.B. and Rosevear, P.R., Biochemistry, 30 (1991) 1461.
19. Pitcher, W.H., 3rd, DeRose, E.F., Mueller, G.A., Howell, E.E. and London, R.E., Biochemistry, 42 (2003) 11150.
20. Hicks, S.N., Smiley, R.D., Stinnett, L.G., Minor, K.H. and Howell, E.E., J. Biol. Chem., 279 (2004) 46995.
21. Howell, E.E., Shukla, U., Hicks, S.N., Smiley, R.D., Kuhn, L.A. and Zavodszky, M.I., J. Comput. Aided Mol. Des., 15 (2001) 1035.
22. Strader, M.B., Smiley, R.D., Stinnett, L.G., VerBerkmoes, N.C. and Howell, E.E., Biochemistry, 40 (2001) 11344.
23. Ewing, T.J.A. and Kuntz, I.D., J. Comput. Chem., 18 (1997) 1175.
24. Schnecke, V., Swanson, C.A., Getzoff, E.D., Tainer, J.A. and Kuhn, L.A., Proteins, 33 (1998) 74.
25. Smiley, R.D., Stinnett, L.G., Saxton, A.M. and Howell, E.E., Biochemistry, 41 (2002) 15664.
26. Stinnett, L.G., Smiley, R.D., Hicks, S.N. and Howell, E.E., J. Biol. Chem., 279 (2004) 47003.
27. Strader, M.B., Chopra, S., Jackson, M., Smiley, R.D., Stinnett, L., Wu, J. and Howell, E.E., Biochemistry, 43 (2004) 7403.
28. Morris, G.M., Goodsell, D.S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K. and Olson, A.J., J. Comput. Chem., 19 (1998) 1639.
29. Kramer, B., Rarey, M. and Lengauer, T., Proteins, 37 (1999) 228.
30. Kontoyianni, M., McClellan, L.M. and Sokol, G.S., J. Med. Chem., 47 (2004) 558.
31. Kellenberger, E., Rodrigo, J., Muller, P. and Rognan, D., Proteins, 57 (2004) 225.
32. Schulz-Gasch, T. and Stahl, M., J. Mol. Model (Online), 9 (2003) 47.
33. Bursulaya, B.D., Totrov, M., Abagyan, R. and Brooks, C.L., 3rd, J. Comput. Aided Mol. Des., 17 (2003) 755.
34. Taylor, R.D., Jewsbury, P.J. and Essex, J.W., J. Comput. Aided Mol. Des., 16 (2002) 151.
35. Wang, R., Lu, Y. and Wang, S., J. Med. Chem., 46 (2003) 2287.
36. Ferrara, P., Gohlke, H., Price, D.J., Klebe, G. and Brooks, C.L., 3rd, J. Med. Chem., 47 (2004) 3032.
37. Cheng, F. and Oldfield, E., J. Med. Chem., 47 (2004) 5149.
38. Rockey, W.M. and Elcock, A.H., Proteins, 48 (2002) 664.
39. Mulakala, C. and Reilly, P.J., Proteins, 49 (2002) 125.
40. Brown, W.M. and Vander Jagt, D.L., J. Chem. Inform. Comput. Sci., 44 (2004) 1412.
41. Kramer, B., Rarey, M. and Lengauer, T., Proteins, 37 (1999) 228.
42. Rarey, M., Kramer, B., Lengauer, T. and Klebe, G., J. Mol. Biol., 261 (1996) 470.
43. Rosenfeld, R.J., Goodsell, D.S., Musah, R.A., Morris, G.M., Goodin, D.B. and Olson, A.J., J. Comput. Aided Mol. Des., 17 (2003) 525.
44. INSIGHT II. 1998 Biosym/MSI: San Diego, CA.
45. Hetenyi, C. and van der Spoel, D., Protein Sci., 11 (2002) 1729.
46. Bohm, H.J., J. Comput. Aided Mol. Des., 8 (1994) 243.
47. Rarey, M., Kramer, B. and Lengauer, T., Proteins, 34 (1999) 17.
48. Berendsen, H.J.C., Vanderspoel, D. and Vandrunen, R., Comput. Phys. Commun., 91 (1995) 43.
49. Lindahl, E., Hess, B. and van der Spoel, D., J. Mol. Mod., 7 (2001) 306.
50. Gohlke, H., Hendlich, M. and Klebe, G., J. Mol. Biol., 295 (2000) 337.
51. Clark, R.D., Strizhev, A., Leonard, J.M., Blake, J.F. and Matthew, J.B., J. Mol. Graph Model, 20 (2002) 281.
52. Charifson, P.S., Corkery, J.J., Murcko, M.A. and Walters, W.P., J. Med. Chem., 42 (1999) 5100.
53. Kaminski, G.A., Friesner, R.A., Tirado-Rives, J. and Jorgensen, W.L., J. Phys. Chem. B, 105 (2001) 6474.
54. Jorgensen, W.L. and Tirado-Rives, J., J. Am. Chem. Soc., 110 (1988) 1657.
55. Berendensen, H.J.C., Postma, J.P.M., van Gunsteren, W.F. and Hermans, J. Interactions Models for Water in Relation to Protein Hydration. D. Reidel Publishing Company, Dordrecht, 1981, pp. 331–342.
56. Kastenholz, M.A. and Hunenberger, P.H., J. Phys. Chem. B, 108 (2004) 774.
57. Deng, H., Callender, R. and Howell, E., J. Biol. Chem., 276 (2001) 48956.
58. Hicks, S.N., Smiley, R.D., Stinnett, L.G., Minor, K.H. and Howell, E.E., J. Biol. Chem. 279 (2004) 46995.
59. Castillo, R., Andres, J. and Moliner, V., J. Am. Chem. Soc., 121 (1999) 12140.

60. Iribarne, F., Paulino, M., Aguilera, S., Murphy, M. and Tapia, O., J. Mol. Model (Online), 8 (2002) 173.
61. Cavalli, A., Bottegoni, G., Raco, C., De Vivo, M. and Recanatini, M., J. Med. Chem., 47 (2004) 3991.
62. Moitessier, N., Henry, C., Maigret, B. and Chapleur, Y., J. Med. Chem., 47 (2004) 4178.
63. Zoete, V., Meuwly, M. and Karplus, M., Proteins, 55 (2004) 568.
64. Choe, J., Suresh, S., Wisedchaisri, G., Kennedy, K.J., Gelb, M.H. and Hol, W.G.J., Chem. Biol., 9 (2002) 1189.
65. Minke, W.E., Pickens, J., Merritt, E.A., Fan, E.K., Verlinde, C.L.M.J. and Hol, W.G.J., Acta Crystallogr. D-Biol. Crystallogr., 56 (2000) 795.
66. Lewi, P.J., de Jonge, M., Daeyaert, F., Koymans, L., Vinkers, M., Heeres, J., Janssen, P.A., Arnold, E., Das, K., Clark, A.D. Jr., Hughes, S.H., Boyer, P.L., de Bethune, M.P., Pauwels, R., Andries, K., Kukla, M., Ludovici, D., De Corte, B., Kavash, R., Ho, C. and Lewis, P.J., J. Comput. Aided Mol. Des., 17 (2003) 129.
67. Bak-Jensen, K.S., Andre, G., Gottschalk, T.E., Paes, G., Tran, V. and Svensson, B., J. Biol. Chem., 279 (2004) 10093.
68. Ko, T.P., Williams, R. and McPherson, A., Acta Crystallogr. D-Biol. Crystallogr., 52 (1996) 160.
69. Thompson, P.R., Schwartzenhauer, J., Hughes, D.W., Berghuis, A.M. and Wright, G.D., J. Biol. Chem., 274 (1999) 30697.
70. Park, H., Bradrick, T.D. and Howell, E.E., Protein Eng., 10 (1997) 1415.