Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Structures of dihydrofolate reductase-thymidylate synthase of *Trypanosoma cruzi* in the folate-free state and in complex with two antifolate drugs, trimetrexate and methotrexate

Olga Senkovich,[a]‡ Norbert Schormann[a]‡ and Debasish Chattopadhyay[a,b]*

[a]Center for Biophysical Sciences and Engineering, University of Alabama at Birmingham, Birmingham, AL 35294, USA, and [b]Department of Medicine, University of Alabama at Birmingham, Birmingham, AL 35294, USA

‡ These authors contributed equally.

Correspondence e-mail: debasish@uab.edu

Received 28 January 2008
Accepted 1 April 2009

PDB References: dihydrofolate reductase-thymidylate synthase, folate-free, 2h2q, r2h2qsf; TMQ complex, 3clb, r3clbsf; MTX complex, 3cl9, r3cl9sf.

The flagellate protozoan parasite *Trypanosoma cruzi* is the pathogenic agent of Chagas disease (also called American trypanosomiasis), which causes approximately 50 000 deaths annually. The disease is endemic in South and Central America. The parasite is usually transmitted by a blood-feeding insect vector, but can also be transmitted *via* blood transfusion. In the chronic form, Chagas disease causes severe damage to the heart and other organs. There is no satisfactory treatment for chronic Chagas disease and no vaccine is available. There is an urgent need for the development of chemotherapeutic agents for the treatment of *T. cruzi* infection and therefore for the identification of potential drug targets. The dihydrofolate reductase activity of *T. cruzi*, which is expressed as part of a bifunctional enzyme, dihydrofolate reductase–thymidylate synthase (DHFR-TS), is a potential target for drug development. In order to gain a detailed understanding of the structure–function relationship of *T. cruzi* DHFR, the three-dimensional structure of this protein in complex with various ligands is being studied. Here, the crystal structures of *T. cruzi* DHFR-TS with three different compositions of the DHFR domain are reported: the folate-free state, the complex with the lipophilic antifolate trimetrexate (TMQ) and the complex with the classical antifolate methotrexate (MTX). These structures reveal that the enzyme is a homodimer with substantial interactions between the two TS domains of neighboring subunits. In contrast to the enzymes from *Cryptosporidium hominis* and *Plasmodium falciparum*, the DHFR and TS active sites of *T. cruzi* lie on the same side of the monomer. As in other parasitic DHFR-TS proteins, the N-terminal extension of the *T. cruzi* enzyme is involved in extensive interactions between the two domains. The DHFR active site of the *T. cruzi* enzyme shows subtle differences compared with its human counterpart. These differences may be exploited for the development of antifolate-based therapeutic agents for the treatment of *T. cruzi* infection.

1. Introduction

Chagas disease is a serious public health problem that is endemic to 19 countries in Latin America, where approximately 100 million people remain at risk. The disease is caused by a protozoan parasite, *Trypanosoma cruzi*, which is primarily transmitted by the blood-feeding insects hematophagous triatomines. However, congenital transmission and transmission *via* blood transfusion are also common. About one third of infected humans develop a chronic form of Chagas disease, which causes damage to the heart, esophagus and colon (Croft, 1997). Although in recent years several

Latin American countries have made significant progress in controlling the infection, current estimates of the number of infected individuals (∼8 million) still remain high (World Health Organization, 2007). Despite the enormous global burden of Chagas disease, no drug is effective in its chronic stage and those used for treatment of the acute disease result in toxic side effects (Croft, 1997; Decampo, 2001). Only two drugs, nifurtimox and benznidazole, are able to cure up to 50% of recent infections, but they are not effective against the intracellular stage of the parasite. Although one recent study showed that treatment with benznidazole can cure chronic infections in mice (Bustamante *et al.*, 2008), side effects and efficacy remain major concerns in its use. Without any hope of a vaccine in the foreseeable future, there is an urgent need for effective chemotherapy. Therefore, identifying new drug targets and developing a detailed understanding of the structural and functional aspects of these targets are extremely important.

Since folate is an essential nutrient for living organisms and the reduced form of folate is a precursor of cofactors that are required for the synthesis of thymidylate, purine nucleotides, methionine, serine and glycine necessary for DNA, RNA and protein synthesis, enzymes involved in folate metabolism have been extensively used as drug targets (McGuire, 2003). Among these, dihydrofolate reductase (DHFR), an NADPH-dependent enzyme, which catalyzes the reduction of folate (FA) to dihydrofolate (DHF) and of DHF to tetrahydrofolate (THF), is one of the most thoroughly studied proteins and continues to serve as one of the most frequently used targets for cancer chemotherapy. Moreover, DHFR has diverged greatly through evolution and DHFRs from different organisms show dramatic differences in their inhibition by certain folate analogs. This selectivity forms the basis for the development of specific antifolates for the treatment of bacterial and parasitic infections (Schweitzer *et al.*, 1990; Ferone, 1984).

The dihydrofolate reductase (DHFR) activity of *T. cruzi* is a potential target for rational drug design. Considering that trypanosomatid parasites are folate auxotrophs and that the *dhfr-ts* gene is essential in related parasites such as *Leishmania*, DHFR activity is likely to be an essential function in *T. cruzi* (Scott *et al.*, 1987; Beck & Ullman, 1990). Genetic studies in *Leishmania* showed that mutants lacking the gene require a supply of thymidine for growth and fail to survive in macrophages *in vitro* and in animal infections (Nare *et al.*, 1997). Moreover, we observed that the lipophilic antifolate drug trimetrexate (TMQ), a potent inhibitor of *T. cruzi* DHFR (*Tc*DHFR) function, kills the infective (trypomastigote) and intracellular (amastigote) forms of the parasite at low nanomolar concentrations, with IC_{90} values of 36 and 72 n*M*, respectively (Senkovich *et al.*, 2005). However, although TMQ is an FDA-approved drug, its use is complicated by associated toxicity. Nevertheless, the proven track record of DHFR inhibitors in the treatment of parasitic diseases such as malaria emphasizes the need for further exploration of this target for the development of an antifolate-based therapy for *T. cruzi* infection.

Bacterial and mammalian DHFRs are monomeric proteins of 18–25 kDa, while TS is usually a homodimer of 60–70 kDa.

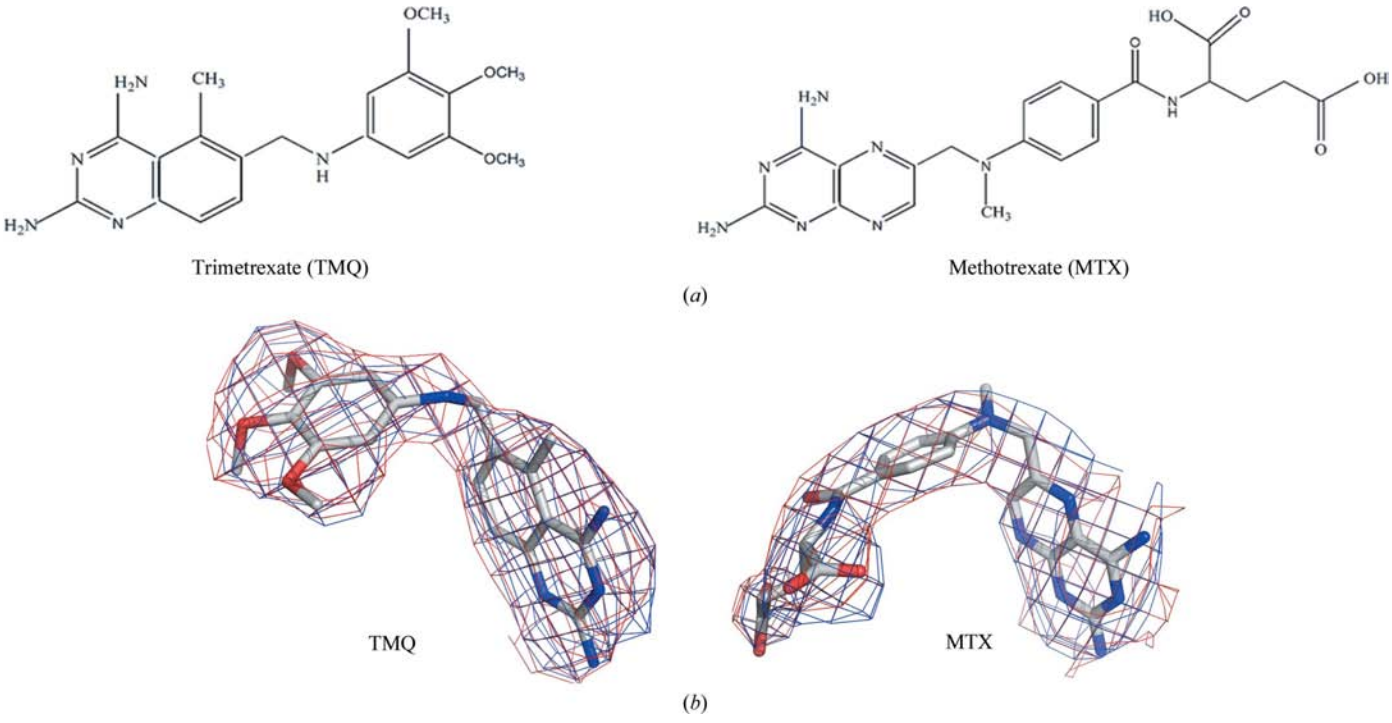

Figure 1
Structures and representative electron-density maps for TMQ and MTX. (*a*) Chemical structures of TMQ and MTX. (*b*) X-ray poses for TMQ and MTX are shown as stick models in element color (white, C; blue, N; red, O) with representative difference electron-density maps superimposed as a mesh in red ($2F_o - F_c$ map; 1σ level) and blue ($F_o - F_c$ OMIT map; 3σ level).

In contrast, in *T. cruzi* and other protozoa DHFR and TS activities are expressed as a bifunctional enzyme, dihydrofolate reductase-thymidylate synthase (DHFR-TS), in which the N-terminal DHFR domain is joined by a linker to the TS domain (Ferone & Roland, 1980; Ivanetich & Santi, 1990; Reche *et al.*, 1994; Knighton *et al.*, 1994). The native bifunctional enzyme is a homodimer with a molecular weight of approximately 110–140 kDa. There is evidence of functional interactions between the domains, presumably *via* conformational changes in the domains of individual subunits of the dimer (Knighton *et al.*, 1994). The *Tc*DHFR domain (residues 21–232) possesses only 36% identity (53 identical amino acids) to its human counterpart and has an additional 20 amino-acid residues at the N-terminus. Moreover, several key residues involved in binding antifolates in human DHFR (hDHFR) are replaced by other amino acids in equivalent positions in the *T. cruzi* protein. However, the potential of *T. cruzi* DHFR-TS (*Tc*DHFR-TS) as a drug target remains largely unexplored (Zuccotto *et al.*, 1999, 2001; Gilbert, 2002).

Here, we describe three crystal structures of the *T. cruzi* DHFR-TS enzyme (*Tc*DHFR-TS). All three structures contain an $NADP^+$ (NAP) molecule bound in the cofactor-binding pocket in the DHFR domain. The first of the described structures (referred to hereafter as 'folate-free') contains a deoxyuridine monophosphate (dUMP) molecule bound in the TS active site, but the DHFR active site is in the ligand-free state. The second structure contains a TMQ molecule (for chemical structure, see Fig. 1*a*) bound in the DHFR active site; in the TS domain a sulfate ion occupies the site where the phosphate group of dUMP binds in the folate-free structure. The third structure contains the classical antifolate methotrexate (MTX; Fig. 1*a*) in the DHFR domain and contains a dUMP molecule in the TS domain. Comparison of the active sites in the *Tc*DHFR domain with those of hDHFR may allow the exploitation of the observed differences for developing specific (selective) inhibitors of the parasitic enzyme using a structure-based strategy.

2. Materials and methods

2.1. Expression and purification of *Tc*DHFR-TS

The expression and purification of *Tc*DHFR-TS has been described previously (Senkovich *et al.*, 2005). Briefly, recombinant *Tc*DHFR-TS with a vector-encoded N-terminal hexahistidine tag was expressed in *Escherichia coli* Rosetta (DE3) pLysS cells and purified from the soluble fraction of the cell lysate by immobilized metal affinity on a Talon column using an imidazole gradient. Pooled fractions were further purified by size-exclusion chromatography on a Superdex 200 column. Purified *Tc*DHFR-TS eluted as a dimeric protein of approximate molecular weight 110 kDa.

2.2. Crystallization and structure determination of *Tc*DHFR-TS complexes

All complexes were formed by incubating *Tc*DHFR-TS (10 mg ml^{-1}) with 1 m*M* $NADP^+$, 1 m*M* dUMP and 1–25 µ*M* inhibitor (for inhibitor complexes) for 3 h on ice. The hanging-drop vapor-diffusion method at room temperature was used for crystallization.

Crystals of the cofactor complex (folate-free) appeared in 2 d and reached maximum dimensions of 0.1 × 0.1 × 0.04 mm within a week. The reservoir contained 16–20% 5-methylpentane-2,3-diol (MPD), 8–12% PEG 4000, 0.1 *M* potassium citrate pH 6.0. Intensity data for this complex were collected on the Advanced Photon Source (APS) synchrotron beamline (SBC 19BM) at 103 K using 30% MPD, 12% PEG 4000, 0.1 *M* potassium citrate pH 6.0 as cryoprotectant. Intensity data were processed using the *HKL*-2000 program package (Otwinowski & Minor, 1997). The crystals of the folate-free complex belonged to the orthorhombic space group $C222_1$, with unit-cell parameters $a = 88.68$, $b = 137.25$, $c = 189.27$ Å. Considering two subunits of *Tc*DHFR-TS to be present in the asymmetric unit, the values of the calculated Matthews coefficient and the solvent content were 2.4 $Å^3$ Da^{-1} and 50%, respectively. The crystal structure was solved by molecular replacement using the program *MOLREP* (Vagin & Teplyakov, 1997) and a search model built from the *Leishmania major* DHFR-TS (*Lm*DHFR-TS) structure (Knighton *et al.*, 1994) by replacing non-identical residues with alanine. The top solution had an *R* factor of 0.38 with a correlation coefficient of 0.62. The two *Tc*DHFR-TS subunits in the asymmetric unit are related by a noncrystallographic twofold axis. Calculation of an electron-density map to 2.6 Å resolution at this stage allowed fitting of most of the residues. For all three structures described here the graphics program package *Coot* (Emsley & Cowtan, 2004) was used for map fitting and *REFMAC*5 (Murshudov *et al.*, 1997) was used for refinement. When applicable, medium NCS restraints were applied during restrained refinement by maximum likelihood using TLS refinement (Winn *et al.*, 2001). NCS restraints were removed in the final rounds of refinement. Cofactors and inhibitor molecules were placed and refined. Finally, based on reasonable hydrogen-bonding criteria, water molecules were placed at 3σ peaks in the final $F_o - F_c$ electron-density map, followed by final rounds of refinement.

The crystallization medium for the TMQ complex contained 1.5–1.7 *M* ammonium sulfate, 0.01 *M* magnesium chloride, 5% ethylene glycol and 0.05 *M* Tris pH 7.5. These crystals appeared in 3 d and grew to maximum dimensions of 0.2 × 0.2 × 0.15 mm in two weeks. A crystal was flash-frozen in mother liquor containing 25% ethylene glycol (EDO) and the intensity data collected on the IMCA 17ID beamline at APS were processed using the *d*TREK* program package (Pflugrath, 1999). These crystals belonged to the tetragonal space group $P4_32_12$, with four molecules in the asymmetric unit (V_M = 4.2 $Å^3$ Da^{-1}; 70% solvent content). The crystal structure of this complex was solved by molecular replacement using the structure of one subunit of ligand-free *Tc*DHFR-TS (after the removal of $NADP^+$ and dUMP) as a search model with the program *Phaser* (McCoy *et al.*, 2007). The top solution provided the location of four DHFR-TS subunits in the asymmetric unit related by noncrystallographic 222 symmetry. Electron-density maps calculated after initial rigid-body

Table 1
Crystal data and data-collection statistics for *Tc*DHFR-TS.

Values in parentheses are for the highest resolution shell.

	Folate-free	TMQ	MTX
Space group	$C222_1$	$P4_32_12$	$I4_122$
Unit-cell parameters (Å)			
a	88.68	176.57	121.05
b	137.25	176.57	121.05
c	189.27	251.85	260.76
No. of subunits per ASU	2	4	1
Solvent content (%)	49.7	70.5	69.7
V_M (Å^3 Da^{-1})	2.44	4.17	4.05
Resolution range (Å)	50.00–2.40 (2.49–2.40)	24.97–3.00 (3.11–3.00)	19.78–3.30 (3.42–3.30)
No. of reflections	222579	520583	74534
Unique reflections	45818 (4372)	77643 (7755)	14555 (1601)
Average redundancy	4.9 (4.6)	6.7 (6.8)	5.1 (5.3)
Completeness (%)	99.5 (95.8)	97.2 (98.3)	97.5 (99.5)
R_{merge} (%)	6.1 (24.0)	15.2 (47.2)	10.3 (43.3)
Reduced χ^2	Not determined	0.99 (1.15)	0.94 (1.15)
$I/\sigma(I)$	8.2 (2.2)	7.9 (3.0)	9.6 (3.2)

Table 2
Refinement statistics for *Tc*DHFR-TS.

Values in parentheses are for the highest resolution shell. NA, not applicable.

	Folate-free	TMQ	MTX
Resolution (Å)	43.15–2.40 (2.45–2.40)	19.94–3.00 (3.08–3.00)	19.78–3.30 (3.38–3.30)
No. of reflections	41407 (2746)	73568 (5613)	13809 (1041)
Completeness (%)	95.2 (88.3)	97.1 (98.3)	97.5 (99.2)
R_{all} (%)	20.8	20.3	20.8
R_{work} (%)	20.6 (23.5)	20.1 (28.6)	20.6 (26.1)
R_{free} (%)	24.1 (28.2)	23.9 (34.3)	24.9 (29.3)
No. of atoms			
Overall	8471	16812	4157
Protein	8028	16134	4015
Ligands	136	300	101
Ions	NA	140	13
Waters	307	190	20
EDO	NA	48	8
Wilson *B* factor (Å^2)	30.0	54.0	87.0
Average *B* factors (Å^2)			
Overall	37.4	38.7	90.8
Protein	37.1	37.5	90.3
Ligands†	56.0‡ [59.4, 52.6]	62.4§ [57.6, 61.4, 59.0, 71.8]	98.4¶
Cofactor (NAP)†	56.4 [54.5, 58.3]	80 [72.4, 75.5, 79.7, 92.4]	127.2
Ions	NA	109.5	114.6
Waters	42.9	27.8	58.1
EDO	NA	60.1	84.0
R.m.s. deviations			
Bonds (Å)	0.007	0.007	0.010
Angles (°)	1.11	1.14	1.32
Coordinate error ESU	0.19	0.25	0.35
Correlation coefficient			
F_oF_c	0.93	0.94	0.94
F_oF_c free	0.92	0.91	0.90

† Values in square brackets are the *B* factors for ligand or cofactor in different subunits in the aymmetric unit. ‡ Ligand is dUMP. § Ligand is TMQ. ¶ Ligands are dUMP and MTX.

refinement allowed us to place NADP$^+$ and TMQ in all four subunits (*A*, *B*, *C* and *D*), although the density for NADP$^+$ and for TMQ was poor in subunit *B*. One sulfate ion (from the crystallization medium) was located at the site normally occupied by the phosphate group of dUMP.

Crystals of the MTX complex were grown from 17% MPD, 10% PEG 4000, 0.1 *M* potassium citrate pH 6.0 and reached maximum dimensions of 0.15 × 0.15 × 0.1 mm in two weeks. Intensity data were collected on the IMCA 17ID beamline at APS using a crystal flash-frozen in mother liquor containing 25% EDO and were processed using the *d*TREK* program package (Pflugrath, 1999). These crystals belonged to the tetragonal space group $I4_122$, with one subunit in the asymmetric unit (V_M = 4.1 Å^3 Da^{-1}, 70% solvent content). The crystal structure of this complex was solved by molecular replacement using the structure of the *Tc*DHFR-TS protein chain (subunit *A*) of the TMQ complex as a search model with the program *Phaser* (McCoy *et al.*, 2007). Electron-density maps calculated after initial rigid-body refinement allowed us to place NADP$^+$ and MTX into the active site of the DHFR domain and dUMP in the TS domain.

Electron-density maps representing the inhibitors TMQ and MTX are shown in Fig. 1(*b*). Data-collection and refinement statistics are presented in Tables 1 and 2. Coordinates for all three structures have been deposited in the PDB (folate-free, 2h2q; TMQ complex, 3hbb; MTX complex, 3cl9).

2.3. Docking outline

For docking, we used the program package *eHiTS* (SimBioSys Inc.) together with the graphical interface *CheVi*. The program package and its use have been described in detail in recent publications (Zsoldos *et al.*, 2006, 2007; Hurt *et al.*, 2006; Deng *et al.*, 2007; Schormann *et al.*, 2008). Briefly, a structure of trimethoprim (TMP) was created by sketching in two dimensions, with subsequent generation of three-dimensional coordinates using the program package *MarvinBeans*/*JChem* (ChemAxon; for details see http://www.chemaxon.com/jchem/intro/index.html). The TMP molecule was then docked into the active sites of *Tc*DHFR and *E. coli* DHFR (*Ec*DHFR). Atomic coordinates for the *Tc*DHFR protein were taken from the structure of the TMQ complex and those for *Ec*DHFR were from the respective MTX complex. The docked conformations used were those that provided the best fit of the 2,4-diaminopyrimidine moiety of TMP with the 2,4-diaminopteridine scaffold of TMQ (*Tc*DHFR) and MTX (*Ec*DHFR).

3. Results and discussion

3.1. Quality of the structures

The final model of the *Tc*DHFR-TS folate-free state was refined to a resolution of 2.4 Å (*R* and R_{free} were 20.6% and 24.1%, respectively) and contains two DHFR-TS subunits (*A*, 514 residues; *B*, 492 residues) with one $NADP^+$ molecule bound to each DHFR domain, one molecule of dUMP in each TS domain and a total of 307 water molecules. Residues 36–42, 61–65, 113–116, 182–188 and the C-terminal residues 516–521 are missing from subunit *B*. A total of 99.5% of the residues are in the allowed regions (90.6% are in the core region) and four residues (Tyr343 in subunits *A* and *B*, and Asn119 and Lys87 in subunit *B*) are in the disallowed region of the Ramachandran plot. Electron density for the Tyr343 side chain was excellent in both subunits, with the aromatic side chain pointing into a hydrophobic pocket and stabilized *via* stacking interactions with Phe299. Lys87 and Asn119 in subunit *B* are in solvent-exposed regions and the electron density for the side chains was poor.

The structure of the TMQ complex was refined to a resolution of 3.0 Å (*R* and R_{free} were 20.1% and 23.9%, respectively). It contains four *Tc*DHFR-TS subunits in the asymmetric unit (*A*, 510 residues; *B*, 508 residues; *C*, 510 residues; *D*, 509 residues). One molecule of $NADP^+$ and one TMQ molecule were bound to each DHFR domain in each subunit (Fig. 1*b*). A sulfate ion (presumably from the precipitant) occupies the phosphate-binding site in each TS domain. Residues 115–120 and 516–521 of each subunit are missing from the final model. Approximately 99.4% of all residues in the final model lie in the allowed regions (88.3% are in the core region) of the Ramachandran plot and seven residues (Leu3 in subunit *A*, Arg32 and Lys65 in subunit *B*, Ser2, Asp114 and Tyr343 in subunit *C* and Ala112 in subunit *D*) are in the disallowed region. Electron density for the Leu3 (*A*), Arg32 (*B*), Lys65 (*B*) and Asp114 (*C*) side chains and Ala112 (*D*) are poor. The Ser2 (*C*) side-chain hydroxyl group forms a hydrogen bond to the main-chain carbonyl O atom of Met1 (*C*). The aromatic side chain of Tyr343 in subunit *C* points into a hydrophobic pocket and the side-chain hydroxyl group makes a hydrogen bond to the main-chain N atom of Cys403.

The structure of the MTX complex was refined to a resolution of 3.3 Å ($R = 20.6\%$, $R_{free} = 24.9\%$). The DHFR domain contains $NADP^+$ and MTX in the active site (see Fig. 1*b*), while the TS domain contains dUMP. Eight residues in a loop region (111–118) and the C-terminal residues 515–521 are missing from the final model. 99.8% of all protein residues are in the allowed regions (84.6% are in the core region) of the Ramachandran plot and only one residue (Phe4) lies in the disallowed region. Electron density for the main chain and side chain of this residue is excellent, but the side-chain phenyl ring lies in a solvent-accessible region.

Data-collection and refinement statistics for the three structures are shown in Tables 1 and 2.

3.2. Overall structure

Although the histidine tag was not cleaved off the protein used for crystallization, these residues were not seen in the electron density. In *Tc*DHFR-TS, the DHFR (residues 1–232) and TS (residues 235–521) domains are connected by a short linker. The electron density for the linker residues was excellent in each molecule of all three structures. The overall structure of *Tc*DHFR-TS is very flexible and the two-domain protein packs in various crystal lattices in different complexes.

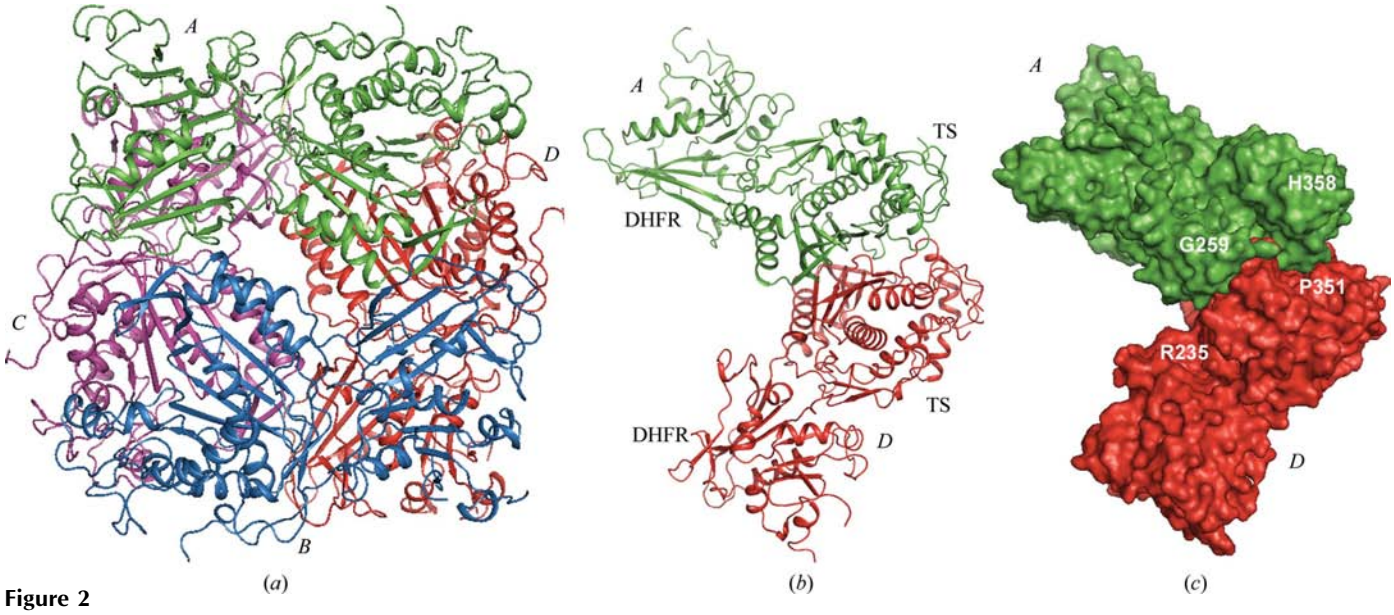

Figure 2
Packing of *Tc*DHFR-TS in the TMQ complex. (*a*) Cartoon representation showing the arrangement of *Tc*DHFR-TS in the crystal structure with TMQ bound in the DHFR active site. The four subunits of *Tc*DHFR-TS are arranged as two sets of dimers (subunits *AD* and *BC*). (*b*) Cartoon diagram showing the interaction between subunits *A* and *D*. Similar interactions exist between subunits *B* and *C*. These interactions involve packing of the five-stranded β-sheet structures of each TS domain against each other. (*c*) The extensive protein–protein interaction involving the TS domain of subunits *A* (green) and *D* (red) in the dimer is shown in surface representation.

Crystal structures suggest that interactions between the TS domains may be responsible for the dimeric nature of the bifunctional *Tc*DHFR-TS enzyme. In the folate-free structure, the presence of noncrystallographic twofold symmetry (NCS) near the loop residues 202–206 places the two TS domains (in subunits *A* and *B*) in the asymmetric unit away from each other. The r.m.s. deviation between these two subunits is 0.7 Å for 486 C^{α} atoms. However, from analyses of the structures of *Tc*DHFR-TS–inhibitor complexes and DHFR-TS enzymes of other organisms, it becomes apparent that the biological dimeric unit in the folate-free structure is formed through interaction between TS domains from subunits related by crystallographic symmetry. On the other hand, in the TMQ complex, which contains two sets of dimers (*AD* and *BC*), subunits within each dimer are NCS-related (Figs. 2*a* and 2*b*). In the MTX complex, which has only one subunit in the asymmetric unit, the homodimer is formed between molecules related by crystallographic symmetry. In all structures, the TS domains from two subunits of the dimer form extensive protein–protein interactions (Fig. 2*c*). For example, in the TMQ complex the dimer interface contains 53 residues from each partner subunit and a total of 21 and 25 hydrogen bonds are involved in forming the dimers *AD* and *BC*, respectively. The calculated interface accessible surface area is 2200 $Å^2$ (*AD*, *BC*), which accounts for ∼9% of the accessible surface area of the DHFR-TS subunits and ∼16% of the accessible surface area of the TS domains (the accessible surface area was calculated using the *PROTORP* protein–protein interaction server; see http://www.bioinformatics.sussex.ac.uk/protorp/). The substrate-binding pockets in both the DHFR and TS domains are partially exposed to solvent (see later). Although the *Tc*DHFR-TS structures do not reveal any direct evidence of channeling of the substrate between the two domains, the possibility of such transport cannot be ruled out at this time (Knighton *et al.*, 1994; Elcock *et al.*, 1996).

As will be discussed later, various degrees of local disorder resulting from the inherent flexibility, especially in the DHFR domain, and the interactions of protein residues with various ligands (cofactor, inhibitor *etc.*) and changes in the local environment are noticed in *Tc*DHFR-TS structures. Moreover, when the asymmetric unit contains multiple subunits there is considerable asymmetry resulting from varying degrees of movement of the DHFR and TS domains with respect to each other and the movement of flexible regions of the protein chains. Although individual domains in different *Tc*DHFR-TS subunits superimpose fairly well, the angle between the DHFR and TS domains varies. For example, the r.m.s. deviation between pairs of DHFR domains from various subunits in the TMQ complex ranges between 0.5 and 0.7 Å, but for the full-length protein the r.m.s. deviation between some subunit pairs increased to 1.4 Å. The movement of the two domains in the *Tc*DHFR-TS molecule is revealed in the rotation angles between different subunits (within the same asymmetric unit or in different structures), which vary in the range 0–13° as calculated by the *DynDom* web server (http://fizz.cmp.uea.ac.uk/dyndom/). The motion of the protein structures in the inhibitor complexes is reflected in a higher average *B* factor of the protein residues compared with that of the water molecules.

The overall structure of *Tc*DHFR-TS is very similar to that of *Lm*DHFR-TS, with which it shares 68.8% sequence identity. The r.m.s. deviation between these structures is 1.4 Å for 396 matched C^{α} atoms. In both structures a short linker (two residues) connects the DHFR and the TS domains, while in DHFR-TS from *Cryptosporidium hominis* (*Ch*DHFR-TS) and *Plasmodium falciparum* (*Pf*DHFR-TS) the linker is much longer (58 and 89 residues, respectively). The longer linker results in a different arrangement of the domains in the homodimer with respect to each other: the protein shows the canonical TS dimer interface with the two DHFR domains residing on the 'shoulders' of the TS monomers. Because the short linker in *Tc*DHFR-TS and *Lm*DHFR-TS restricts the orientation of the DHFR domains relative to the TS domains, the DHFR domains are rotated approximately 180° relative to the *Ch*DHFR and *Pf*DHFR domains. As a result, the DHFR and TS active sites are on the same side of the monomer in *Tc*DHFR-TS and *Lm*DHFR-TS, while they are on opposite sides of the monomer in *Ch*DHFR-TS and *Pf*DHFR-TS.

3.3. DHFR domain

The architecture of the DHFR domain is highly conserved in various organisms. The fold (SCOP database; http://scop.mrc-lmb.cam.ac.uk/scop/) is characterized by a mixed β-sheet of eight strands (order 43251687), with strand 8 being antiparallel to the rest (three layers: $\alpha/\beta/\alpha$). Approximately 83–85% of the solvent-accessible surface of the cofactor is buried in these structures and the solvent accessibility is reduced by ∼3% when an inhibitor is bound (*LPC*/*CSU* server; http://bioportal.weizmann.ac.il/oca-bin/lpccsu). The specificity for the cofactor is determined by several contacts between the adenine phosphate O atoms: AOP1 with Thr102 OG and Ser100 OG (for example, 2.8 and 2.4 Å, respectively, in the folate-free state), AOP2 with Ser101 OG (2.6 Å) and AOP3 with Arg78 NH1 (2.6 Å). The amide N and O atoms of the nicotinamide moiety are anchored *via* hydrogen bonds to the backbone O and N atoms of Ala28 (2.8 and 2.9 Å, respectively). The α-phosphate O atoms AO1 and AO2 form hydrogen bonds to Thr80 OG1 (2.5 Å) and Ser158 OG (2.6 Å), respectively. The phosphate O atom (NO2) of the β-phosphate group forms hydrogen bonds to the backbone N atoms of Gly157 (3.0 Å) and Ser158 (2.7 Å). The adenine ring only makes hydrophobic interactions with the protein. The cofactor in the TMQ and MTX complexes also shows similar interactions with protein residues. The cofactor molecules in different subunits display varying degrees of disorder, as reflected in their average *B* factors.

There are three conserved insertions, residues 1–18, 60–65 and 101–115 (*T. cruzi* numbering), in the DHFR domain of trypanosomal enzymes compared with DHFR domains from other species. The moderate sequence identity (∼36%) between the hDHFR and *Tc*DHFR domains is reflected in the divergence of their three-dimensional structures outside of the core region. While the r.m.s. deviation between C^{α} atoms of

the 146 matching residues (69% match rate) is 1.2 Å, the overall r.m.s. deviation between the two structures is 2.7 Å.

The N-terminal extension (residues 1–20) of the *Tc*DHFR domain folds back on the TS domain of the respective subunit, forming a major interaction surface connecting the two domains and extending over the entire length of the TS domain on one side (Fig. 3*a*). Although the *B* factors of the residues in this loop are higher than the average value for the entire protein, the quality of electron density for this region was very good in all three structures. Similar interactions between the TS domain and the N-terminal tail of the DHFR domain have also been noticed in the bifunctional enzymes from other parasites (Knighton *et al.*, 1994; Yuvaniyama *et al.*, 2003). While the length of this extension varies (22 residues in *Lm*DHFR-TS and six residues in *Pf*DHFR-TS) and its exact function is not known, the influence of this non-active-site segment in the modulation of the catalytic activities of the DHFR and TS domains of the *L. major* and *P. falciparum* enzymes has been shown. Interestingly, while sequential removal of amino acids from the extension resulted in a corresponding loss of DHFR activity of the *P. falciparum* enzyme, deletion of the entire extension from the *L. major* enzyme resulted in a higher DHFR catalytic rate, suggesting that the segment plays different roles in these enzymes (Dasgupta & Anderson, 2008). As shown in Fig. 3(*b*), in *Tc*DHFR-TS the C-terminal end of this N-terminal loop packs against two helices, composed of residues 132–145 and 156–171, in the DHFR domain. The conformation of these helices

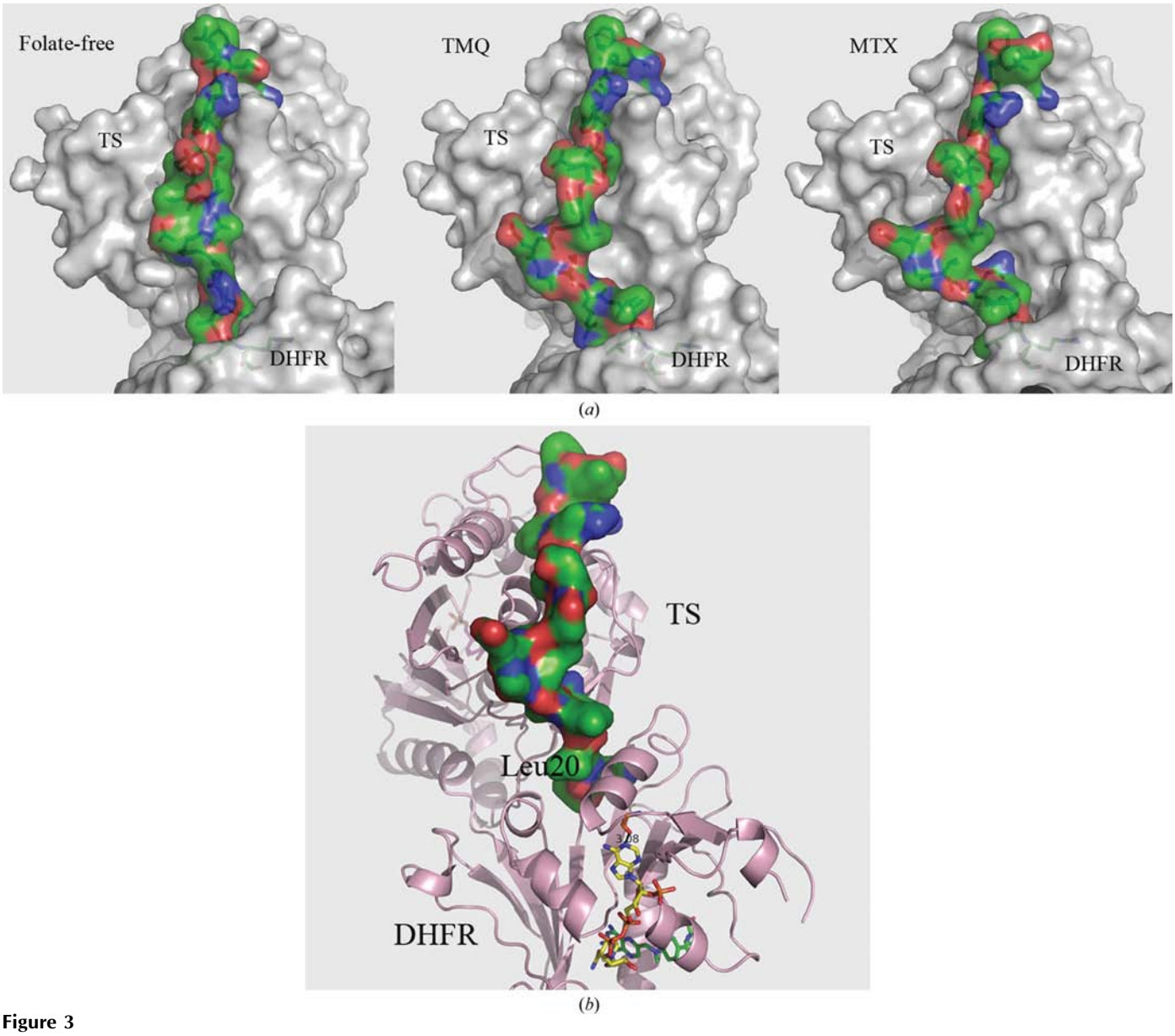

Figure 3
Interaction between domains *via* the N-terminal tail. (*a*) Surface drawing showing the packing of the N-terminal loop (residues 1–20) in the folate-free, TMQ and MTX complexes. The loop is shown with C atoms in green, N atoms in blue and O atoms in red. (*b*) Cartoon drawing showing the interaction of the above loop with both domains. $NADP^+$ (NAP) and MTX are shown in stick representation [blue, N; red, O; yellow, C (NAP); green, C (MTX)]. Gly131 is also shown as a stick model. The peptide N atom of Gly120 is 3.08 Å away from N7 in the adenine ring of NAP.

Table 3
Active-site residues in human (CO4, MTX) and *T. cruzi* (TMQ, MTX) DHFR domains.

The protein residues of hDHFR and the *Tc*DHFR domain that interact with the inhibitors are listed. Identical residues in equivalent positions are shown in bold. Charged residues are marked with asterisks.

(*a*) hDHFR (1pd8; CO4); *Tc*DHFR (3clb; TMQ).

hDHFR	Ile7	**Val8**	**Ala9**	Leu22	Glu30*	Phe31	**Phe34**	Gln35*†	**Thr56**	**Ser59**	**Ile60**	**Pro61**
*Tc*DHFR	Val26	**Val27**	**Ala28**	Ile41	Asp48*	Met49	**Phe52**		**Thr80**	**Ser83**	**Ile84**	**Pro85**
hDHFR	Asn64*	**Leu67**	Val115	**Tyr121**	**Thr136**							
*Tc*DHFR	Phe88	**Leu91**	Ile154	**Tyr160**	**Thr178**							

(*b*) hDHFR (1dls; MTX); *Tc*DHFR (3cl9; MTX).

hDHFR	Ile7	**Val8**	**Ala9**	Tyr22‡	Arg28*§	Glu30*	Phe31	**Phe34**	Gln35*		**Ser59**
*Tc*DHFR	Val26	**Val27**	**Ala28**	Ile41		Asp48*	Met49	**Phe52**	Arg53*	Thr80§	**Ser83**
hDHFR	**Ile60**		Asn64*	**Leu67**		**Arg70***	Val115	**Tyr121**	**Thr136**		
*Tc*DHFR	**Ile84**	Pro85§	Phe88	**Leu91**	Pro92§	**Arg94***	Ile154	**Tyr160**	**Thr178**		

† The equivalent residue in *Tc*DHFR (Arg53) makes no contact with the inhibitor. ‡ Tyr22 is a mutation in hDHFR. § The equivalent residue in hDHFR or *Tc*DHFR does not form contacts with MTX.

may be sensitive to the presence of the bound cofactor. For example, the main-chain N atom of Gly131, the residue immediately preceding the first helix, forms a contact with the adenine-ring N atom N7 of the bound $NADP^+$. Although in the absence of a cofactor-free structure the effects of cofactor binding on the local structure cannot be predicted accurately, it is clear that residues 128–131 are important for formation of the cofactor-binding pocket. Thus, while the mechanism of domain–domain communication is unclear, the results of studies on the *Leishmania* and *Plasmodium* enzymes suggest a possible role for the N-terminal 20 residues of *Tc*DHFR-TS in modulation of the enzymatic function of this enzyme. It remains to be seen whether these mechanisms may be targeted for the development of novel therapeutic agents.

3.4. Increased flexibility in the DHFR domain induced by MTX binding

The structure of the DHFR domain displays variations in the stability of the loop regions that separate the secondary-structural elements. In molecular-dynamics studies the native state of *Ec*DHFR was described to be an ensemble and the region spanning residues 60–90 was found to be the least stable in the entire protein (Pan *et al.*, 2000; McElheny *et al.*, 2005). This region, which contains an α-helix and two β-strands connected by a number of loops, contains several residues that are involved in cofactor or substrate binding. The molecular-dynamics studies confirmed the findings of a series of crystallographic analyses that showed that binding of MTX to DHFR resulted in a decrease in atomic *B* factors throughout the protein and a concomitant increase in the mobility of the loop (residues 63–72) connecting the β-strands. Notably, this loop is 15 Å away from the substrate-binding site. Results of these studies on *Ec*DHFR demonstrate that conformational changes resulting from ligand binding propagate throughout the DHFR structure and contradicts the classical view of stabilization of protein conformations upon ligand binding (or the classical view of 'freezing out' as described by Pan *et al.*, 2000). This notion agrees well with our observations from the structures of *Tc*DHFR-TS in the folate-free and inhibitor-bound states. In *Tc*DHFR the region corresponding to residues 60–90 of *Ec*DHFR is much longer owing to two insertions of seven and 16 amino acids and extends from residues 95 to 150. The insertions are accommodated by lengthening of both β-strands and loops in the *Tc*DHFR structure. The segment of *Tc*DHFR equivalent to the *Ec*DHFR loop residues 63–72 is 25 residues long and contains a 3_{10}-helix (residues 105–111). This region of the *Tc*DHFR structure is more ordered in the folate-free state than in the inhibitor complexes. In fact, in subunit *A* of the folate-free structure the entire region is ordered and in subunit *B* only four residues (113–116) are disordered, while in different subunits of TMQ and MTX complexes the 3_{10}-helix is either short or absent and a number of residues are missing from the loops.

3.5. TS domain

The TS domain is highly homologous in various organisms and the fold of this domain in *Tc*DHFR-TS is similar to those in human TS (hTS) and in parasitic DHFR-TS enzymes. There is a noticeable bend in the β-strands of the *Tc*TS domain (Supplementary Fig. 1[1]); a similar bending is observed in the corresponding region of *Lm*DHFR-TS. The bending is characterized by the presence of small loops that interrupt three of the strands in the five-stranded β-sheet. It should be noted that the packing of the β-sheets from two subunits constitute the major interactions in the assembly of the dimers in monofunctional TS as well as in the bifunctional DHFR-TS enzymes. The TS active site is formed by residues from both subunits in the dimer. In the folate-free structure the side-

[1] Supplementary material has been deposited in the IUCr electronic archive (Reference: HV5125). Services for accessing this material are described at the back of the journal.

chain atoms of Ser424 and Arg423 from the corresponding subunit in the asymmetric unit (*A* or *B*) as well as Arg383 and Arg384 from a symmetry-related molecule are hydrogen bonded to the phosphate O atoms of dUMP. Hydrogen bonds are observed between the hydroxyl group of the deoxyribose group and protein residues His464 and Tyr466, while uracil atoms O2, O4 and N3 display hydrogen bonds to the main-chain atoms of Asp426 and Asn434. In the TMQ complex one sulfate ion instead occupies the phosphate-binding site. In each dimer (*AD* and *BC*), arginine residues (Arg384 and Arg423) from both subunits form hydrogen bonds to the sulfate O atoms. Irrespective of whether the substrate or just the sulfate is bound, the positions of the side chains and main chains of residues forming the dUMP-binding pocket remain unchanged.

3.6. Details of inhibitor binding and comparison of DHFR domains in the folate-free structure and the inhibitor complexes

The TMQ-binding pocket in the *Tc*DHFR domain is predominantly hydrophobic (see Table 3). The only charged active-site residue (Asp48) forms hydrogen bonds with the inhibitor. Both O atoms of the carboxyl side chain of Asp48 are involved in interactions with the amino group and the pyrimidine-ring N atom of the quinazoline ring. Other polar interactions between the protein and TMQ are weak (3–4 Å). The cofactor provides the remaining interactions (66 total contacts between TMQ, protein and cofactor, calculated for a distance cutoff of 2–4 Å). The quinazoline ring of TMQ stacks against the nicotinamide ring of $NADP^+$ on one side and the aromatic side chain of Phe52 on the other. 83% of the solvent-

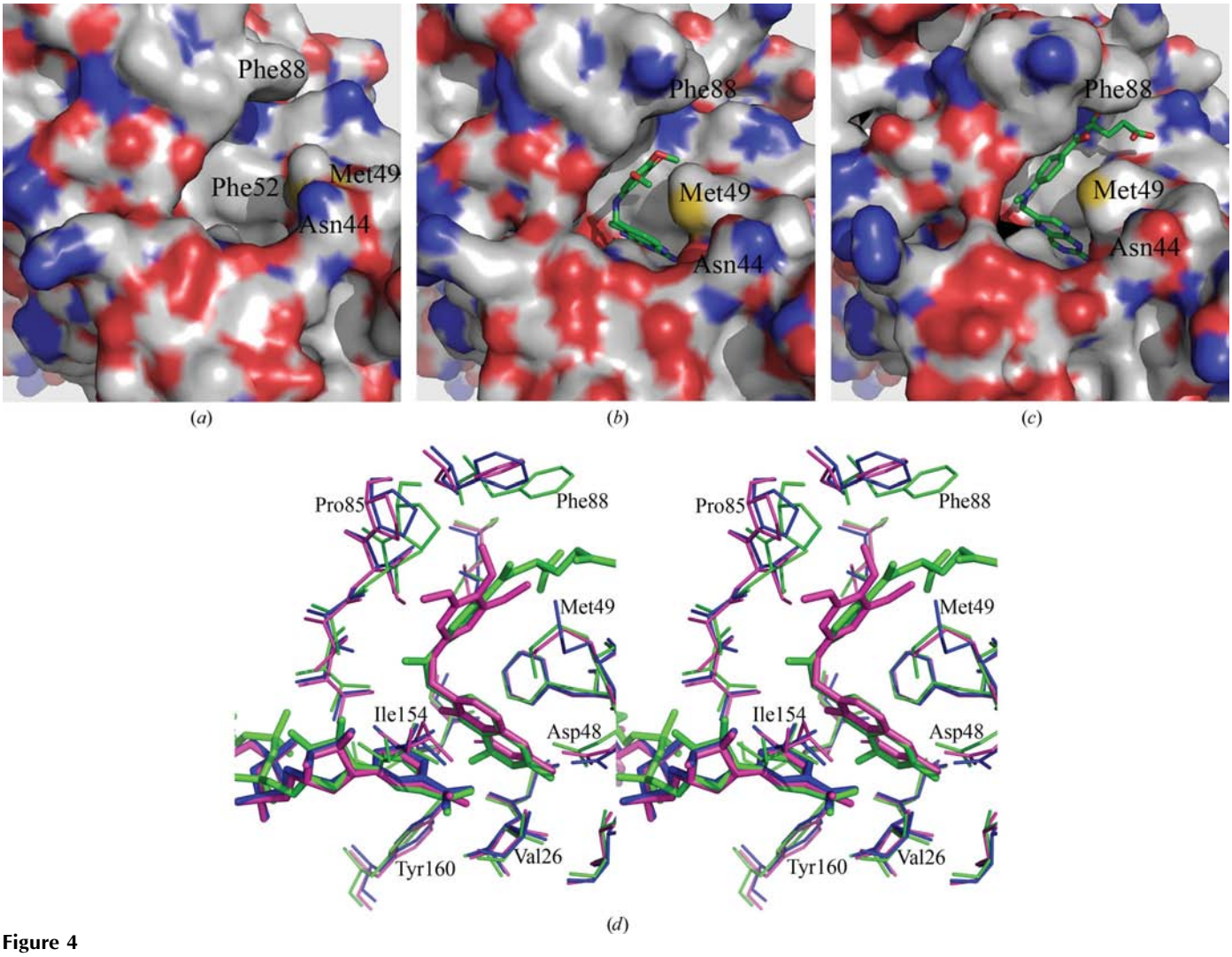

Figure 4
Partial closure of the substrate-binding pocket in the *Tc*DHFR domain upon inhibitor binding. The atomic coordinates of the DHFR domains of the folate-free state and inhibitor complexes of the *Tc*DHFR-TS structures were used for superimposition. (*a*)–(*c*) Surface representations of the residues near the binding pocket are shown in element colors (white, C; blue, N; red, O; yellow, S). (*a*) Folate-free structure. (*b*) TMQ complex. (*c*) MTX complex. The positions of the TMQ and MTX molecules are shown as stick models in element colors (green, C; blue, N; red, O). (*d*) Stereo figure of the superimposition of active-site residues in the DHFR domain for the three structures. The stereoview shows active-site residues, cofactor and inhibitors. Color codes are blue, folate-free structure; magenta, TMQ complex; green, MTX complex. Active-site residues are shown in line representation; the cofactor NAP and inhibitors (TMQ, MTX) are shown as stick models.

accessible surface of TMQ (uncomplexed, 606 Å^2; complexed, 106 Å^2) is buried in the complex (*LPC*/*CSU* server; http://bioportal.weizmann.ac.il/oca-bin/lpccsu). In the MTX complex, the inhibitor is surrounded by 19 active-site residues and forms hydrogen bonds to Asp48 and Arg94. Hydrogen bonds to the side-chain O atoms of Tyr160 and Thr178 and the main-chain O atoms of Val26 and Ile154 are also observed and the total number of contacts between MTX, protein residues and cofactor is 100 (for a distance cutoff of 2–4 Å). 91% of the solvent-accessible surface of MTX (uncomplexed, 700 Å^2; complexed, 60 Å^2) is buried upon binding to the *Tc*DHFR active site (*LPC*/*CSU* server).

Comparison of the *Tc*DHFR domain in the folate-free and inhibitor-bound states demonstrates that upon inhibitor binding movement of residues near the active site results in partial closure of the pocket (Figs. 4*a*, 4*b* and 4*c*), especially owing to the movement of the side chain of Met49. The volume of the entire binding pocket is reduced by ∼20% in the TMQ structure (1240 Å^3) when compared with the folate-free (1540 Å^3) structure (*CASTp* server; see http://sts-fw.bioengr.uic.edu/castp/calculation.php). In addition, several

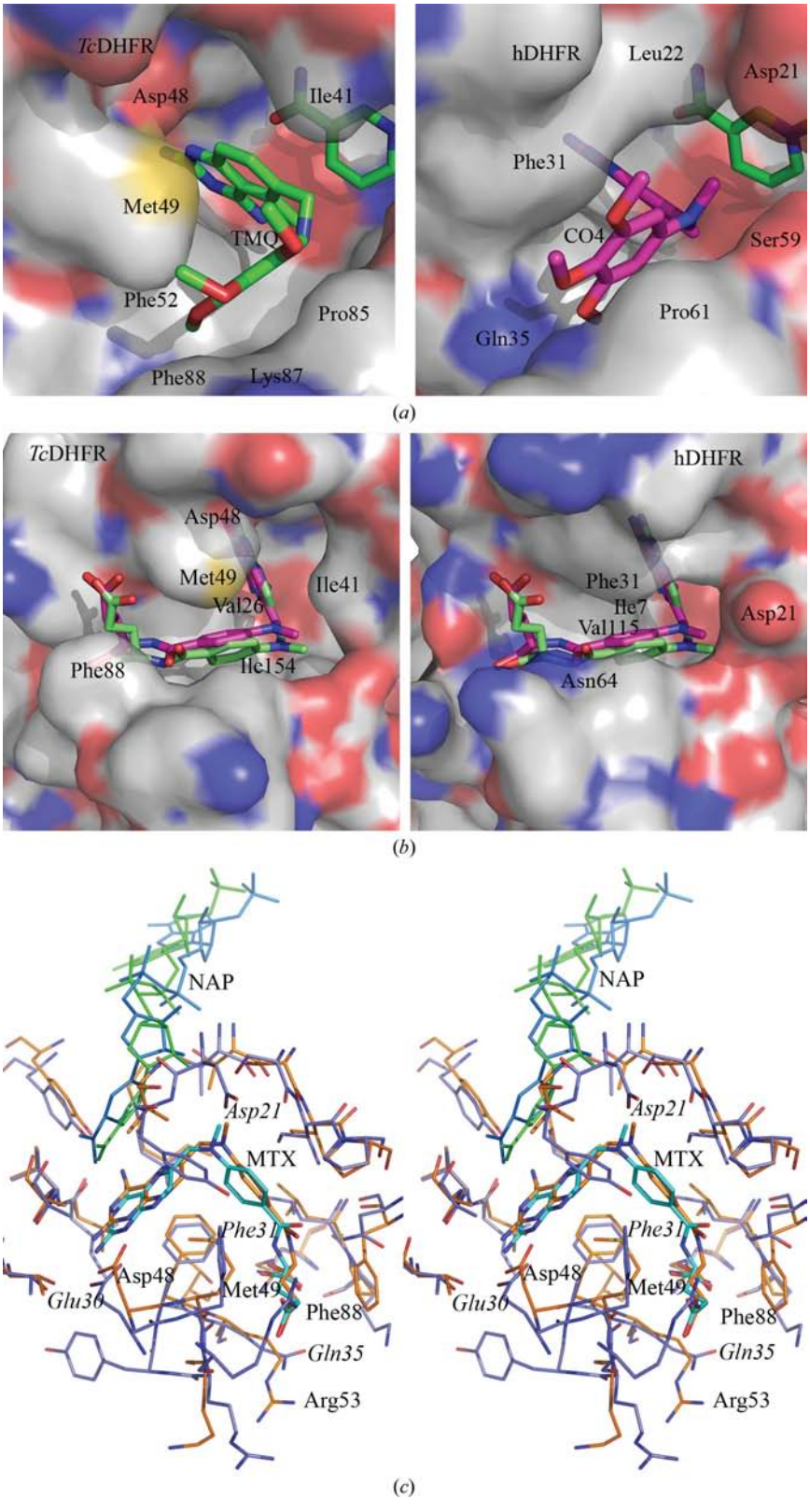

Figure 5
TMQ and MTX binding in the *Tc*DHFR domain and comparison with hDHFR. (*a*) The TMQ-binding pocket in hDHFR and *Tc*DHFR. Atomic coordinates for the *Tc*DHFR domain of the TMQ complex are superimposed on the hDHFR domain bound to the inhibitor CO4. Residues near the inhibitor-binding site are shown in surface representation. Inhibitor molecules are not included in the surface calculation. The position of the cofactor is shown as a stick model. Residues in *Tc*DHFR (left) and hDHFR (right) near the active site are shown. The TMQ and CO4 molecules are shown as stick models in element color [blue, N; red, O; green, C (TMQ); pink, C (CO4)]. (*b*) The MTX-binding pocket in hDHFR and *Tc*DHFR. Atomic coordinates for the *Tc*DHFR domain of the MTX complex are superimposed with those of the hDHFR–MTX complex. Residues near the inhibitor-binding site are shown in surface representation. Inhibitor molecules are not included in the surface calculation. Cofactor molecules are bound but are not shown. Residues in *Tc*DHFR (left) and hDHFR (right) near the MTX-binding site are shown. MTX in *Tc*DHFR is displayed as a stick model in element color (green, C; blue, N; red, O). The position of MTX in hDHFR is also shown as a stick model (rose, C; blue, N; pink, O) for comparison. Some residues are labeled to highlight the differences near the active-site pocket. (*c*) Stereo figure of the superimposition of active-site residues in the MTX complexes of *Tc*DHFR and hDHFR. Active-site residues are shown as line drawings in element color (*Tc*DHFR: slate, C; blue, N; red, O; hDHFR: orange, C; blue, N; red, O). Residues of *Tc*DHFR are labeled in regular font and those of hDHFR in italics. MTX molecules are shown as stick models for both proteins with color codes blue, N; red, O; cyan, C (*Tc*DHFR); yellow, C (hDHFR). The position of the cofactor is shown as a stick model in cyan for *Tc*DHFR and green for hDHFR.

other active-site residues (Asp48, Phe52, Ile84, Phe88 and Leu91) and the nicotinamide moiety of the cofactor shift position (Fig. 4*d*).

3.7. Active-site comparison in *Tc*DHFR and hDHFR inhibitor complexes

In order to develop species-specific inhibitors for parasitic DHFR, it will be necessary to improve the selectivity of an active compound by reducing the binding affinity for the human enzyme while maintaining the potency against the target enzyme. The increase in selectivity is required to reduce the cytotoxicity. Although the active site in the *Tc*DHFR domain appears to be more hydrophobic than the site in hDHFR (see Table 3), suggesting that the former should favor the binding of lipophilic inhibitors, both TMQ and MTX demonstrate a similar binding mode in the active-site pocket of both enzymes and these compounds have a similar inhibitory activity against both enzymes (Senkovich *et al.*, 2005). The structure of a complex of hDHFR with CO4, a structural analog of TMQ, has been reported. Comparison of the active-site pocket in the hDHFR–CO4 complex (PDB code 1pd8) with those in the *Tc*DHFR domain in the TMQ complex is shown in Fig. 5(*a*). Active-site residues that interact with the bound inhibitors in these structures show six differences, of which three are nonconservative (hDHFR: Asp21, Phe31 and Asn64; *Tc*DHFR: Ile41, Met49 and Phe88) and the rest are conservative (hDHFR: Ile7, Glu30 and Val115; *Tc*DHFR: Val26, Asp48 and Ile154) substitutions (Table 3*a*). Fig. 5(*b*) shows the active-site pockets in the MTX complexes of hDHFR (PDB code 1dls) and *Tc*DHFR. Among the active-site residues of these two proteins that interact with MTX (Table 3*b*) six are different, of which three are nonconservative (hDHFR, Phe31, Gln35 and Asn64; *Tc*DHFR, Met49, Arg53 and Phe88) and the remainder are conservative (hDHFR, Ile7, Glu30 and Val115; *Tc*DHFR, Val26, Asp48 and Ile154) substitutions. In hDHFR, amino group NA2 of the pteridine moiety of MTX forms a hydrogen bond to the side chain of Glu30, while amino group NA4 makes hydrogen bonds to the main-chain carbonyl O atom of Ile7 and Val115. Hydrogen bonds are also observed between O atoms of the glutamate moiety in MTX and N atoms of the side chains of Asn64 and Arg70. In *Tc*DHFR, O atom O2 of the glutamate moiety of MTX forms a hydrogen bond to the side chain of Arg94. In this structure,

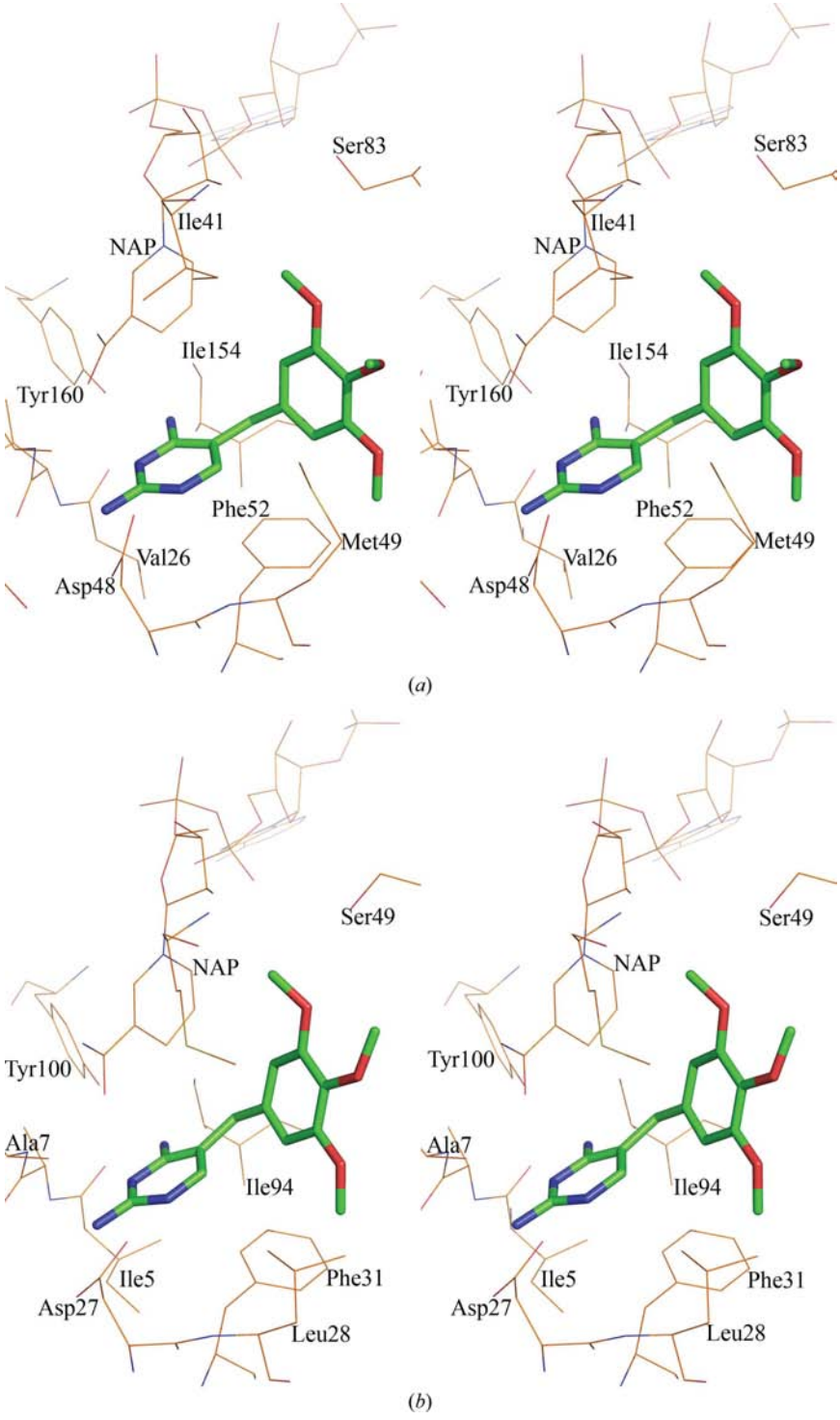

Figure 6
Comparison of the *Tc*DHFR domain, *Ec*DHFR and hDHFR in complex with TMP. Stereo figures showing the active-site residues in the *Tc*DHFR and *Ec*DHFR domain with the docked pose of TMP and in the NMR solution structure of the hDHFR–TMP complex (PDB code 1yho). Protein residues are shown as line drawings. The respective cofactor is labeled NAP. TMP is shown as a stick model. Some residues that interact with the inhibitor (distance ≤ 4 Å) are labeled. (*a*) TMP docked into the *Tc*DHFR active site. (*b*) TMP docked into the *Ec*DHFR active site.

amino group NA2 of the pteridine moiety of MTX forms a hydrogen bond to the side chain of Asp48, while amino group NA4 is hydrogen bonded to the main-chain carbonyl O atom of Val26, the main-chain carbonyl O atom of Ile154 and the side-chain hydroxyl of Tyr160. A stereoview of the superimposition of the residues at and near the MTX-binding sites in *Tc*DHFR and hDHFR is shown in Fig. 5(*c*).

3.8. Comparison of interactions of TMP with *T. cruzi*, *E. coli* and human DHFR

Although the overall protein fold of DHFR is conserved, the primary sequence has diverged considerably among various species. The sequence identity between prokaryotes, eukaryotes and protozoa averages in the range 15–40%. Protozoa exhibit about 30% sequence identity to eukaryotes, while the identity to prokaryotes is about 20%. Among the trypanosomatid parasites (*Leishmania*, *T. cruzi*, *T. brucei*) the sequence identity reaches 46–58%. Prokaryotic DHFR enzymes show the lowest sequence identity to either eukaryotes or protozoans. The differences in the active site have led to the development of species-specific DHFR inhibitors that are highly selective for certain DHFRs.

The 2,4-diaminopyrimidine analog trimethoprim (TMP) is a potent and selective inhibitor of bacterial DHFRs. The experimental IC_{50} values for *Ec*DHFR and *Lactobacillus casei* DHFR (*Lc*DHFR) are 10 n*M* and 0.26 µ*M*, respectively, while the IC_{50} for hDHFR is 230 µ*M* (Gangjee *et al.*, 1998). NMR solution structures of complexes of *L. casei* DHFR with TMP show that the inhibitor is completely surrounded by active-site residues providing a multitude of interactions. In the crystal structure of the *Ec*DHFR–TMP complex, in addition to Asp27, which interacts with the inhibitor, Ile5, Ala7, Leu28, Phe31, Ile50, Leu54, Arg57, Ile94, Tyr100 and Thr113 have been described to be near the inhibitor (Baker *et al.*, 1981; Kuyper *et al.*, 1985). In the solution structure of the complex of hDHFR with TMP (PDB code 1yho), the protonated N1 atom of TMP interacts with Glu30 (Kovalevskaya *et al.*, 2005). The inhibitor interacts with a number of hydrophobic residues, namely Phe31, Phe34, Ile60 and Val115, and with Thr56.

In the absence of a crystal structure of *Tc*DHFR complexed with TMP, we used docking and modeling to predict the protein–inhibitor interactions. The results of this study suggested that the *Tc*DHFR active-site residues Val26 (main chain), Asp48 (side chain), Tyr160 (side chain) and Ile154 (main chain) as well as carbonyl O atom O7N of the nicotinamide group of the cofactor form hydrogen bonds to the 2,4-diaminopyrimidine moiety of TMP. Three residues (Ile41, Phe52 and Ser83) of *Tc*DHFR are engaged in hydrophobic interactions with TMP (Fig. 6*a*).

Since the coordinates for the crystal structure of the *Ec*DHFR complex are not available, we generated the docked pose of TMP in the active site in order to compare the interaction of the inhibitor with the *E. coli*, *T. cruzi* and human enzymes. In addition to several conserved contacts involving equivalent residues in the three enzymes and TMP, in the *Ec*DHFR complex both carboxyl O atoms of Asp27 interact with N4 of the diaminopyrimidine ring and the amino group attached to the C3 atom of the ring. Moreover, the hydroxyl group of Tyr100 forms a contact with N4 of the amino group in the inhibitor (Fig. 6*b*). In the *Tc*DHFR–TMP complex (docked) there is one interaction involving Asp48 and the ring N atom and one relatively strong interaction between the amino substituent (Fig. 6*a*). In hDHFR, on the other hand, one of the carboxyl O atoms of Glu30 shows a strong interaction with the N1 atom of the diaminopyrimidine ring, but there is no electrostatic interaction with the amino group (Fig. 6*c*). The differences in the total number of interactions between protein and inhibitor may qualitatively account for the intermediate level of inhibitory activity of TMP against *T. cruzi* DHFR ($IC_{50} \simeq 2$ µ*M*) compared with those against bacterial and mammalian DHFR enzymes (Gangjee *et al.*, 1998).

The wide variations in observed activity of different inhibitors against bacterial, mammalian and protozoan DHFR enzymes result from diverse and complex interactions of these inhibitors with protein residues as well as cofactor molecules. Only in rare circumstances can the differences in activity be partially and semi-quantitatively explained by visual inspection and superimposition of crystal structures of the enzyme from different species with the same

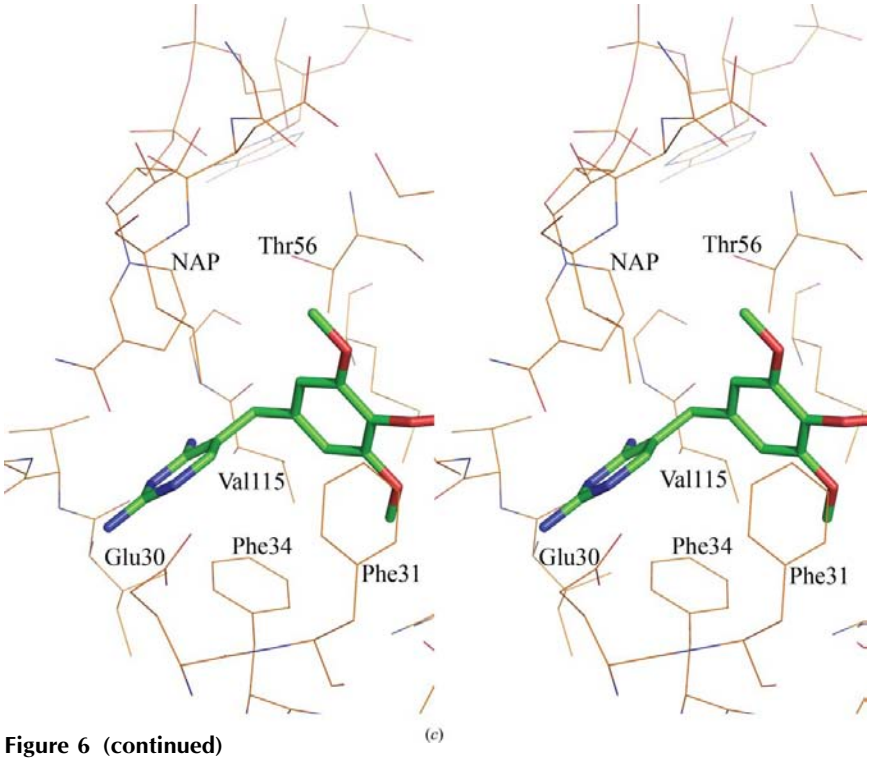

Figure 6 (continued)
(*c*) TMP bound to hDHFR.

inhibitors. Quantitative structure–activity relationship (QSAR) studies may be useful in addressing this problem. The availability of enzyme-inhibitory properties of molecules of diverse chemical structures and structural data on enzyme–inhibitor complexes provides the foundation for developing sophisticated QSAR models for designing selective and potent inhibitors of the parasitic DHFR. Although challenging, the species-specific differences between *Tc*DHFR and hDHFR may be exploited for the development of selective inhibitors of the parasitic enzyme.

We thank the staff at the IMCA and SBC beamlines. Use of the Argonne National Laboratory IMCA and SBC beamlines at the Advanced Photon Source was supported by the US Department of Energy, Office of Energy Research under contract No. W-31-109-ENG-38. We also thank Kelly Bradley and Bessie Lane for help with crystallization. This work was supported in part by NASA cooperative agreement NCC8-246. TMQ (as Neutrexin) was a kind gift from MedImmune Inc. The authors thank Dr Dave Matthews for kindly providing the coordinates of the *L. major* DHFR-TS structure.

References

Baker, D. J., Beddell, C. R., Champness, J. N., Goodford, P. J., Norrington, F. E. A., Smith, D. R. & Stammers, D. K. (1981). *FEBS Lett.* **126**, 49–52.
Beck, J. T. & Ullman, B. (1990). *Mol. Biochem. Parasitol.* **43**, 221–230.
Bustamante, J. M., Bixby, L. M. & Tarleton, R. L. (2008). *Nature Med.* **14**, 542–550.
Croft, S. L. (1997). *Parasitology*, **114**, S3–S15.
Dasgupta, T. & Anderson, K. S. (2008). *Biochemistry*, **47**, 1336–1345.
Decampo, R. (2001). *Curr. Pharm. Des.* **7**, 1157–1164.
Deng, J., Sanchez, T., Al-Mawsawi, L. Q., Dayam, R., Yunes, R. A., Garafalo, A., Bolgerd, M. B. & Neamati, N. (2007). *Bioorg. Med. Chem.* **15**, 4985–5002.
Elcock, A. H., Potter, M. J., Matthews, D. A., Knighton, D. R. & McCammon, J. A. (1996). *J. Mol. Biol.* **262**, 370–374.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Ferone, R. (1984). *Handbook of Experimental Pharmacology. Part II: Antimalarial Drugs*, edited by W. Peters & W. H. G. Richards, pp. 207–221. New York: Springer-Verlag.
Ferone, R. & Roland, S. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 5802–5806.
Gangjee, A., Vidwans, A. P., Vasudevan, A., Queener, S. F., Kisliuk, R. L., Cody, V., Li, R., Galitsky, N., Luft, J. R. & Pangborn, W. (1998). *J. Med. Chem.* **41**, 3426–3434.
Gilbert, H. (2002). *Biochim. Biophys. Acta*, **1587**, 249–257.
Hurt, D. E., Sutton, A. E. & Clardy, J. (2006). *Biorg. Med. Chem. Lett.* **16**, 1610–1615.
Ivanetich, K. M. & Santi, D. V. (1990). *FASEB J.* **4**, 1591–1597.
Knighton, D. R., Kan, C. C., Howland, E., Janson, C. A., Hostomska, Z., Welsh, K. M. & Matthews, D. A. (1994). *Nature Struct. Biol.* **1**, 186–194.
Kovalevskaya, N. V., Smurnyy, Y. D., Polshakov, V. I., Birdsall, B., Bradbury, A. F., Frenkiel, T. & Feeney, J. (2005). *J. Biomol. NMR*, **33**, 69–72.
Kuyper, L. F., Roth, B., Baccanari, D. P., Ferone, R., Beddell, C. R., Champness, J. N., Stammers, D. K., Dann, J. G., Norrington, F. E., Baker, D. J. & Goodford, P. J. (1985). *J. Med. Chem.* **28**, 303–311.
McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. & Read, R. J. (2007). *J. Appl. Cryst.* **40**, 658–674.
McElheny, D., Schnell, J. R., Lansing, J. C., Dyson, H. J. & Wright, P. E. (2005). *Proc. Natl Acad. Sci. USA*, **102**, 5032–5037.
McGuire, J. (2003). *Curr. Pharm. Des.* **9**, 2593–2613.
Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.
Nare, B., Hardy, L. & Beverley, S. M. (1997). *Parasitology*, **114**, 12883–12891.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Pan, H., Lee, J. C. & Hilser, V. J. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 12020–12025.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Reche, P., Arrebola, R., Olmo, A., Santi, D. V., Pecanowska, D. G. & Ruiz-Perez, L. (1994). *Mol. Biochem. Parasitol.* **65**, 247–258.
Schormann, N., Senkovich, O., Walker, K., Wright, D. L., Anderson, A. C., Rosowsky, A., Anathan, S., Shinkre, B., Velu, S. & Chattopadhyay, D. (2008). *Proteins*, **73**, 889–901.
Schweitzer, B. I., Dicker, A. P. & Bertino, J. R. (1990). *FASEB J.* **4**, 2441–2452.
Scott, D. A., Coombs, G. H. & Sanderson, B. E. (1987). *Mol. Biochem. Parasitol.* **23**, 139–149.
Senkovich, O., Bhatia, V., Garg, N. & Chattopadhyay, D. (2005). *Antimicrob. Agents Chemother.* **49**, 3234–3238.
Vagin, A. & Teplyakov, A. (1997). *J. Appl. Cryst.* **30**, 1022–1025.
Winn, M. D., Isupov, M. N. & Murshudov, G. N. (2001). *Acta Cryst.* D**57**, 122–133.
World Health Organization (2007). *Wkly Epidemiol. Rec.* **82**, 259–260.
Yuvaniyama, J., Chitnumsub, P., Kamchonwongpaisan, S., Vanichtanankul, J., Sirawaraporn, W., Taylor, P., Walkinshaw, M. D. & Yuthavong, Y. (2003). *Nature Struct. Biol.* **10**, 357–365.
Zsoldos, Z., Reid, D., Simon, A., Sadjad, B. S. & Johnson, A. P. (2006). *Curr. Protein Pept. Sci.* **7**, 421–435.
Zsoldos, Z., Reid, D., Simon, A., Sadjad, B. S. & Johnson, A. P. (2007). *J. Mol. Graph. Model.* **26**, 198–212.
Zuccotto, F., Brun, R., Pacanowska, D. G., Ruiz Perez, L. M. & Gilbert, I. H. (1999). *Bioorg. Med. Chem. Lett.* **9**, 1463–1468.
Zuccotto, F., Zvelebil, M., Brun, R., Chowdhury, S. F., Di Lucrezia, R., Leal, I., Maes, L., Ruiz-Perez, L. M., Pacanowska, D. G. & Gilbert, I. H. (2001). *Eur. J. Med. Chem.* **36**, 395–405.