Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Studies of *Toxoplasma gondii* and *Plasmodium falciparum* enoyl acyl carrier protein reductase and implications for the development of antiparasitic agents

Stephen P. Muench,[a] Sean T. Prigge,[b] Rima McLeod,[c] John B. Rafferty,[a] Michael J. Kirisits,[c] Craig W. Roberts,[d] Ernest J. Mui[c] and David W. Rice[a]*

[a]The Krebs Institute for Biomolecular Research, Department of Molecular Biology and Biotechnology, University of Sheffield, Firth Court, Western Bank, Sheffield S10 2TN, England, [b]Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD 21205, USA, [c]Department of Ophthalmology and Visual Sciences, Paediatrics (Infectious Diseases) and Pathology and the Committees on Molecular Medicine, Genetics, Immunology and The College, The University of Chicago, Chicago, IL 60637, USA, and [d]Department of Immunology, University of Strathclyde, Glasgow G4 0NR, Scotland

Correspondence e-mail: d.rice@sheffield.ac.uk

Received 26 October 2006
Accepted 11 December 2006

PDB References: PfENR–NAD^+–triclosan complex, 2o2y, r2o2ysf; TgENR–NAD^+–triclosan complex, 2o2s, r2o2ssf; apo TgENR, 2o50, r2o50sf.

Recent studies have demonstrated that submicromolar concentrations of the biocide triclosan arrest the growth of the apicomplexan parasites *Plasmodium falciparum* and *Toxoplasma gondii* and inhibit the activity of the apicomplexan enoyl acyl carrier protein reductase (ENR). The crystal structures of *T. gondii* and *P. falciparum* ENR in complex with NAD^+ and triclosan and of *T. gondii* ENR in an apo form have been solved to 2.6, 2.2 and 2.8 Å, respectively. The structures of *T. gondii* ENR have revealed that, as in its bacterial and plant homologues, a loop region which flanks the active site becomes ordered upon inhibitor binding, resulting in the slow tight binding of triclosan. In addition, the *T. gondii* ENR–triclosan complex reveals the folding of a hydrophilic insert common to the apicomplexan family that flanks the substrate-binding domain and is disordered in all other reported apicomplexan ENR structures. Structural comparison of the apicomplexan ENR structures with their bacterial and plant counterparts has revealed that although the active sites of the parasite enzymes are broadly similar to those of their bacterial counterparts, there are a number of important differences within the drug-binding pocket that reduce the packing interactions formed with several inhibitors in the apicomplexan ENR enzymes. Together with other significant structural differences, this provides a possible explanation of the lower affinity of the parasite ENR enzyme family for aminopyridine-based inhibitors, suggesting that an effective antiparasitic agent may well be distinct from equivalent antimicrobials.

1. Introduction

The widespread use of antimalarial agents such as quinolines, antifolates and atovaquone/proguanil has provided a powerful selective pressure driving the development of multidrug resistance in several malarial strains. This increase in resistance of *Plasmodium falciparum*, the most dangerous and widespread species of malarial parasite, and the lack of development of novel low-cost antimalarial medicines has contributed to the rise in the incidence of malaria, which currently kills more than two million people annually (Breman, 2001). Furthermore, the closely related apicomplexan parasite *Toxoplasma gondii* is estimated to infect 25% of the world's population (Sibley, 2003). Moreover, it has been reported as being responsible for the deaths of a significant number of European patients suffering from acquired immunodeficiency syndrome (Hill & Dubey, 2002) and is the third most common cause of food-borne deaths in the United States (Mead *et al.*, 1999). *T. gondii* parasites can also be passed from mother to foetus during pregnancy, leading to congenital ophthalmological neurological defects (Boyer & McLeod,

2002). The total healthcare burden of toxoplasmosis is estimated to be up to $5 billion per annum within the USA and there are still no currently available treatments that eliminate the latent form of the parasite.

Recent studies have revealed that it may be possible to control apicomplexan parasite infections by targeting processes that reside in their apicoplast organelle, which contains over 500 different proteins responsible for carrying out a number of key metabolic pathways (Waller *et al.*, 1998; Zuther *et al.*, 1999). This organelle is thought to have arisen through the process of secondary endosymbiosis. Thus, the apicomplexan progenitor endocytosed an ancient alga which contained a cyanobacterial-derived plastid obtained in a previous primary endosymbiotic event. Consistent with this, analysis of the genes encoding the enzymes of apicoplast-located pathways suggests that they are closely related to those found in prokaryotes and the chloroplasts of plants (Zuther *et al.*, 1999; Fast *et al.*, 2001; Kohler *et al.*, 1997). For example, in apicomplexan parasites the fatty-acid biosynthesis pathway resembles the type II fatty-acid synthase that is found in bacteria and plant chloroplasts and in which each catalytic step of the pathway is encoded on a separate polypeptide (Magnuson *et al.*, 1993), rather than the type I FAS found in man (Smith, 2003). Enoyl acyl carrier protein reductase (ENR) carries out one of two reductive steps in the type II FAS pathway and has been shown to be the target of several families of antimicrobial compounds, including the diazaborines (Baldock *et al.*, 1996), aminopyridine-based inhibitors (Payne *et al.*, 2002; Seefeld *et al.*, 2003) and triclosan (Levy *et al.*, 1999; Fig. 1), a biocide which is found in many household formulations such as toothpastes, soaps, mouthwashes and plastics (Bhargava & Leonard, 1996).

Sequence alignment of *T. gondii* and *Plasmodium* ENRs reveals considerable similarity to the enzymes from other species, with the closest resemblance being to those of plant rather than bacterial origin, with for example approximately 50% sequence identity between *Brassica napus* and *T. gondii* ENR (Fig. 2). Compared with other ENRs, a striking feature of the *Plasmodium* ENR enzymes is the presence of a large polar low-complexity insert of variable size which is thought to flank the substrate-binding pocket. In *T. gondii* ENR, a similar but smaller insert consisting of only six residues can be observed. In the bacterial enzymes, only the ENR from *Mycobacterium tuberculosis* (MtENR) shows a significant insert at this position. However, in MtENR this insert has been implicated in allowing the enzyme to accommodate larger substrates such as mycolic fatty acids (Rozwarski *et al.*, 1999) and its glycine-rich hydrophobic nature is in contrast to that found in the apicomplexan ENRs, which contain a predominantly polar hydrophilic insert of unknown function.

Recently, triclosan has been shown to retard the growth of *T. gondii* and *P. falciparum* with an IC_{50} of less than 1 μM and with a K_i of 0.4 nM for the *P. falciparum* ENR enzyme (Kapoor *et al.*, 2001; McLeod *et al.*, 2001; Surolia & Surolia, 2001). However, a major challenge for the development of drugs targeted against the apicomplexan family is the need for the inhibitor to cross the four membranes of the apicoplast in addition to the barriers provided by the host cell and parasite. Further problems in drug delivery arise when targeting the bradyzoite stage of *T. gondii*, in which the parasites reside within a cyst composed of host and parasite constituents. Recent studies have shown that triclosan has been observed to retard the growth of both *P. falciparum* and *T. gondii* trachyzoites (McLeod *et al.*, 2001; Surolia & Surolia, 2001); moreover, the attachment of triclosan to a releasable octa-arginine linker produced a potent antiparasitic agent that could also enter encysted *T. gondii* bradyzoites (Samuel *et al.*, 2003).

In this paper, we report the structures of apo *T. gondii* ENR (TgENR) and of both TgENR and *P. falciparum* ENR (PfENR) in complex with NAD^+ and triclosan solved to 2.9, 2.6 and 2.2 Å, respectively. The structures have allowed comparisons to be made with those of ENRs from bacterial and plant origin in order to identify novel features of the enzyme that could be utilized in a program of rational drug design.

2. Materials and methods

2.1. Structure determination of PfENR

PfENR is a tetramer with a molecular weight of approximately 38 000 Da and its overexpression, purification and crystallization were carried out as reported previously (Muench *et al.*, 2003). Data were collected to 2.2 Å using an ADSC Quantum 4 detector at station 14.1 at the Daresbury Synchrotron Radiation Source (SRS) at 100 K using 20%(*v*/*v*) glycerol as a cryoprotectant. Analysis of the diffraction data using the autoindexing routine in the program *DENZO* (Otwinowski & Minor, 1997) showed that the crystals belong to the primitive monoclinic system, with unit-cell parameters a = 88.2, b = 82.4, c = 94.8 Å, $\alpha = \gamma = 90$, $\beta = 90.8^\circ$ and a tetramer in the asymmetric unit. The data were subsequently processed using the *DENZO*/*SCALEPACK* (Otwinowski & Minor, 1997) package and analysis of the pattern of systematic absences was consistent with the space group being assigned as $P2_1$. Data-collection and processing statistics are given in Table 1.

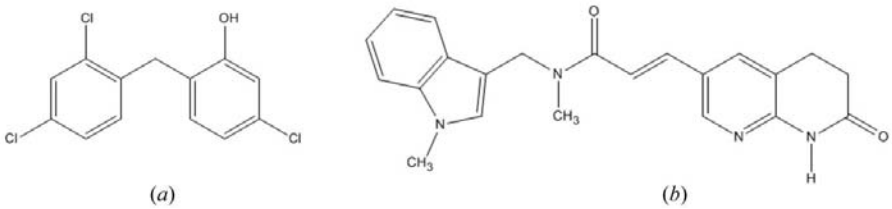

Figure 1
The structural formulae of (*a*) triclosan and (*b*) (*E*)-*N*-methyl-*N*-(1-methyl-1*H*-indol-3-ylmethyl)-3-(7-oxo-5,6,7,8-tetrahydro-1,8-naphthyridin-3-ylacrylamide (compound 29) produced using the program *ISIS/Draw*.

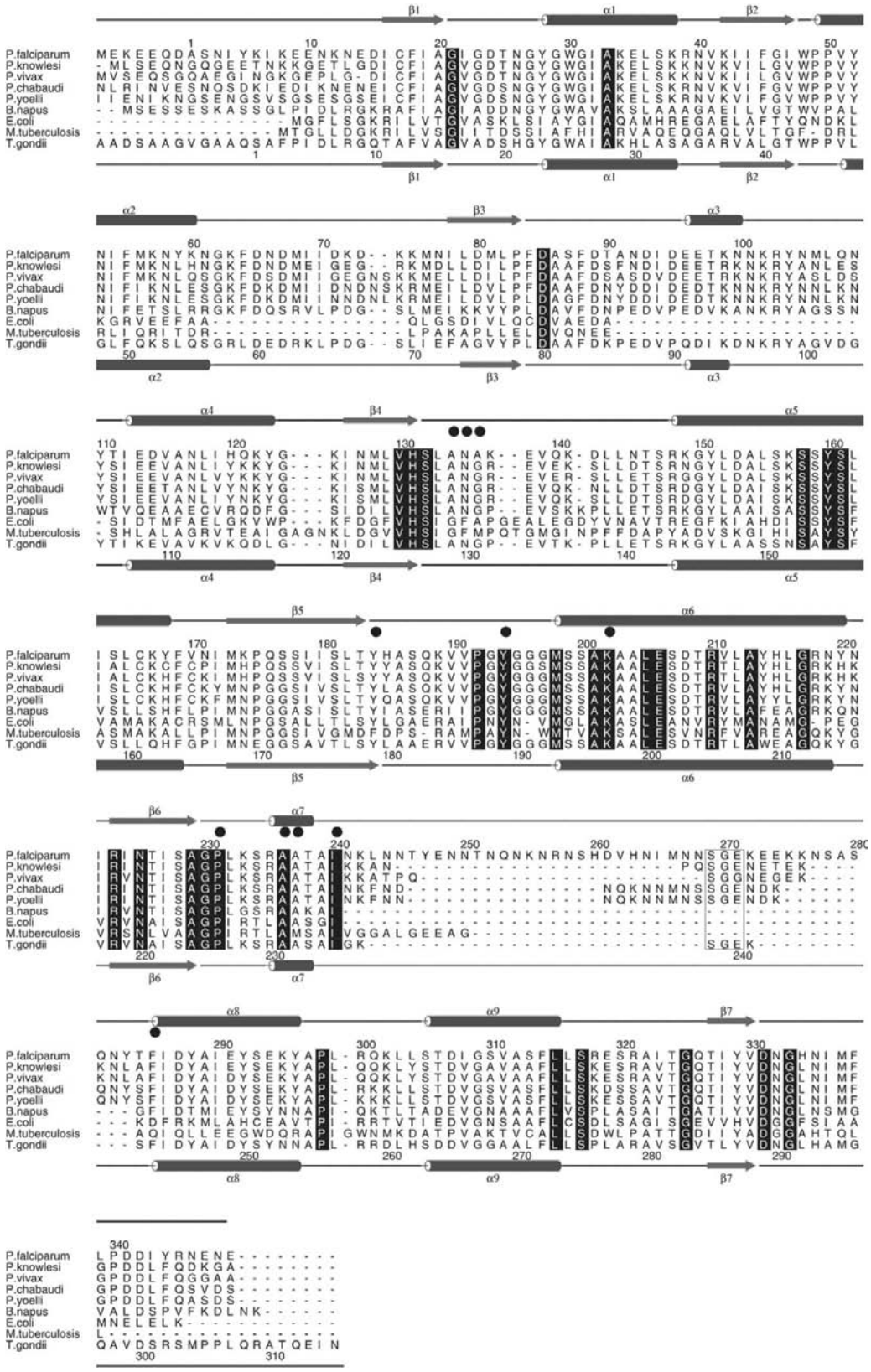

Figure 2
A structure-based sequence alignment of the ENR enzymes from *P. falciparum*, *P. knowlesi*, *P. vivax*, *P. chabaudi*, *P. yoelii*, *B. napus*, *E. coli*, *M. tuberculosis* and *T. gondii*. The elements of secondary structure and the sequence numbering, which is from the start of the mature enzyme for the *P. falciparum* and *T. gondii* enzymes, are shown above and below the alignment, respectively, with cylinders representing α-helices and arrows β-strands. Residues fully conserved in the above sequences are shown in black boxes and residues which are within 4 Å of the inhibitor triclosan are highlighted by a black circle. Those residues which form the conserved SGE motif (residues 238–240 in TgENR) are enclosed by a black box.

Table 1
Data-collection and refinement statistics for apo TgENR, PfENR–NAD$^+$–triclosan and TgENR–NAD$^+$–triclosan complexes.

Values in parentheses are for the highest resolution shell. N/A, not applicable.

	PfENR–NAD$^+$–triclosan	TgENR–NAD$^+$–triclosan	Apo TgENR
Data-collection statistics			
Space group	$P2_1$	$P3_221$	$P3_221$
Wavelength used (Å)	0.9600	1.542	1.542
Resolution range (Å)	50–2.2 (2.26–2.2)	30–2.6 (2.7–2.6)	30.0–2.9 (3.0–2.9)
Unique reflections	77321	21105	13758
Multiplicity	3.1	5.2	10.1
Completeness (%)	97.2 (94.7)	99.9 (99.4)	99.7 (99.8)
$I/\sigma(I) > 3$ (%)	72.6 (49.6)	74.4 (42.0)	—
Average $I/\sigma(I)$	—	—	16.2 (4.5)
$R_{\rm merge}$† (%)	0.095 (0.395)	0.082 (0.48)	0.18 (0.65)
Refinement statistics			
Resolution limits (Å)	30.0–2.2	20.0–2.6	30.0–2.9
$R_{\rm cryst}$‡ (%)	20.0	22.0	21.0
$R_{\rm free}$§ (%)	25.5	28.0	28.0
R.m.s.d. values			
Bond lengths (Å)	0.02	0.015	0.012
Bond angles (°)	1.8	1.7	1.4
Ramachandran plot¶			
Most favoured (%)	90.7	89.0	89.0
Additionally allowed (%)	9.3	11.0	11.0
Generously allowed (%)	0.0	0.0	0.0
Disallowed (%)	0.0	0.0	0.0
Molecules in ASU	4	2	2
Protein atoms	9161	4514	4288
Substrate atoms	244	122	0
Water molecules	493	37	11
Mean B values (Å^2)			
Protein††	22.0 [21.0]	57.0 [56.0]	42.0 [41.0]
Cofactors	23.0	54.0	N/A
Water molecules	23.0	45.0	26.0
Missing residues	Ala1–Glu14 (*A*, *B*, *C*), Ala1–Lys12 (*D*), Lys242–Asn282 (*A*, *B*, *C*), Asn241–Asn282 (*D*), Asn347–Glu340 (*A*, *B*), Glu348–Glu340 (*C*), Arg346–Glu340 (*D*)	Ser1–Phe3 (*A*), Ser1–Ala2 (*B*), Ser238–Gly239 (*B*), Leu307–Asn315 (*A*, *B*)	Ser1–Phe3 (*A*), Ser1–Ala2 (*B*), Leu227–Lys241 (*A*), Leu227–Ser242 (*B*), Leu307–Asn315 (*A*, *B*)

† $R_{\rm merge} = \sum_{hkl}|I_i - I_{\rm m}|/\sum_{hkl} I_{\rm m}$, where I_i and $I_{\rm m}$ are the observed intensity and mean intensity of related reflections, respectively. ‡ $R_{\rm cryst} = \sum_{hkl}(|F_{\rm obs}| - |F_{\rm calc}|)/\sum_{hkl}|F_{\rm obs}|$. § $R_{\rm free}$ was calculated for 5% of the data omitted randomly. ¶ Percentage of residues in regions of the Ramachandran plot according to *PROCHECK* (Laskowski *et al.*, 1993). †† Values in square brackets are the B factor for main-chain atoms only.

Phases for the ternary PfENR–NAD$^+$–triclosan complex were determined by the molecular-replacement method using the program *AMoRe* (Navaza, 1994) and the structure of the *B. napus* ENR–NAD$^+$ complex (PDB code 1eno) as a search model, from which the coordinates for the NAD$^+$ and triclosan were omitted. Following a clear solution of the rotation/translation function, the model was subjected to rigid-body refinement using *REFMAC*5 (Murshudov *et al.*, 1997) using data in the resolution range 10–3.0 Å. The resulting electron-density maps were examined and revealed clear density in each of the four subunits of the tetramer for both NAD$^+$ and triclosan. The model was subsequently subjected to TLS refinement in *REFMAC*5 (Murshudov *et al.*, 1997) and rebuilt in an iterative process using data to 2.2 Å. Following model building, solvent molecules were added using the program *ARP* (Lamzin & Wilson, 1997), but were only refined if they made appropriate contacts to the protein and had B factors below 60 Å^2. Analysis using the program *PROCHECK* (Laskowski *et al.*, 1993) shows that no nonglycine residues lie within disallowed regions of the Ramachandran plot, with 90.7 and 9.3% being in the most favoured and additionally allowed regions, respectively. The final refinement statistics are given in Table 1.

2.2. Structure determination of the TgENR–NAD$^+$–triclosan complex

The overexpression, purification and crystallization of the TgENR–NAD$^+$–triclosan complex were carried out as described previously (Muench *et al.*, 2006). Preliminary X-ray analysis of crystals flash-frozen in 25% glycerol showed that they diffracted well to 2.6 Å resolution. A complete data set was subsequently collected to 2.6 Å resolution at 100 K using a rotation oscillation of 1° over 90° and an exposure time of 10 min on a MAR345 image-plate detector mounted on a Rigaku MM007 generator. Analysis of the diffraction data using the autoindexing routine in the program *DENZO* (Otwinowski & Minor, 1997) showed that the crystals belong to the trigonal point group 32, with unit-cell parameters $a = 78.1$, $b = 78.1$, $c = 188.5$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. Analysis of systematic absences suggested that the space group was either $P3_221$ or $P3_121$. Assuming that the asymmetric unit contains a dimer, the $V_{\rm M}$ value is 2.4 Å^3 Da^{-1} for the TgENR–NAD$^+$–triclosan crystals, which is within the range of $V_{\rm M}$ values observed for protein crystals (Matthews, 1977). The data were subsequently processed using the *DENZO*/*SCALEPACK* (Otwinowski & Minor, 1997) package. Data-collection and processing statistics are given in Table 1.

The structure of the TgENR–NAD$^+$–triclosan complex was determined by molecular replacement using the program *AMoRe* (Navaza, 1994) with the PfENR–NAD$^+$–triclosan complex structure as a search model, from which the coordinates for the NAD$^+$ and triclosan were omitted along with those residues believed to adopt different positions in TgENR and PfENR (63–75, 93–108 and 238–241). The resulting rotation/translation solution gave a clear solution in space group $P3_221$. The model was subjected to rigid-body refinement using *REFMAC*5 (Murshudov *et al.*, 1997) and data in the range 20–4.0 Å. The model was subsequently rebuilt and refined in an iterative process using TLS refinement in

*REFMAC*5 (Murshudov *et al.*, 1997) using data in the range 20–2.6 Å with the addition of water atoms being carried out using the program *ARP* (Lamzin & Wilson, 1997) in the same manner as for the PfENR enzyme. Analysis using the program *PROCHECK* (Laskowski *et al.*, 1993) showed that no non-glycine residues lie in the disallowed regions of the Ramachandran plot, with 89% and 11% being in the most favoured and additionally allowed regions, respectively. The final refinement statistics are given in Table 1.

2.3. Structure determination of apo TgENR

Crystals of apo TgENR were grown in 0.1 *M* Tris pH 8.5 and 25%(*v*/*v*) *t*-butanol, taking 4 days to reach approximate dimensions of 0.15 × 0.1 × 0.1 mm. These crystals were subsequently flash-frozen in 30% glycerol and diffracted well to 2.9 Å resolution at 100 K. A data set was collected using a rotation range of 1° over 200° and an exposure time of 10 min per degree on a MAR345 image-plate detector mounted on a Rigaku MM007 generator. These data were subsequently processed using the programs *MOSFLM* (Leslie, 1992) and *SCALA* (Evans, 1997) and showed that the crystals belonged to the trigonal point group 32, with unit-cell parameters a = 76.04, b = 76.04, c = 187. 47 Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. The structure of apo TgENR was determined by molecular replacement using the program *AMoRe* (Navaza, 1994) with the TgENR–NAD$^+$–triclosan complex structure as a search model, from which the coordinates for the NAD$^+$ and triclosan were omitted along with residues 63–75, 93–108 and 238–241. The resulting rotation/translation solution gave a clear solution in space group $P3_221$. The model was subjected to rigid-body refinement and was subsequently rebuilt and refined in an iterative process by TLS refinement in *REFMAC*5 (Murshudov *et al.*, 1997) using data in the range 20–2.9 Å. An additional 11 water molecules were added in the same manner as for the PfENR enzyme. Analysis using the program *PROCHECK* (Laskowski *et al.*, 1993) showed that no nonglycine residues lie in the disallowed regions of the Ramachandran plot, with 89% and 11% being in the most favoured and additionally allowed regions, respectively. The final refinement statistics are given in Table 1.

The electrostatic surface potentials shown in Figs. 3(*b*) and 5 were calculated using the program *GRASP* (Nicholls *et al.*, 1991) and all figures except Fig. 1 and Fig. 4 were produced in *PyMOL* (DeLano, 2002). Fig. 4 was produced using the graphics program *TURBO-FRODO* (Roussel *et al.*, 1990).

2.4. Sequence identification and alignment

The ENR sequences for the *Plasmodium* species *P. yoelli*, *P. vivax*, *P. knowlesi* and *P. falciparum* and for *T. gondii* were taken from the SWISS-PROT database (accession Nos. Q7RHY7, Q6TEI3, Q6TEI4, Q9BH77 and Q6UCJ9, respectively). The sequence for *P. chabaudi* ENR was obtained by *BLAST* searching the PlasmoDB database (http://www.plasmoDB.org) using the PfENR amino-acid sequence. The sequences were initially aligned using *ClustalW* (Higgins & Thompson, 1994), edited by hand and further manipulated within *ALSCRIPT* (Barton, 1993) to obtain a structural alignment diagram (Fig. 2).

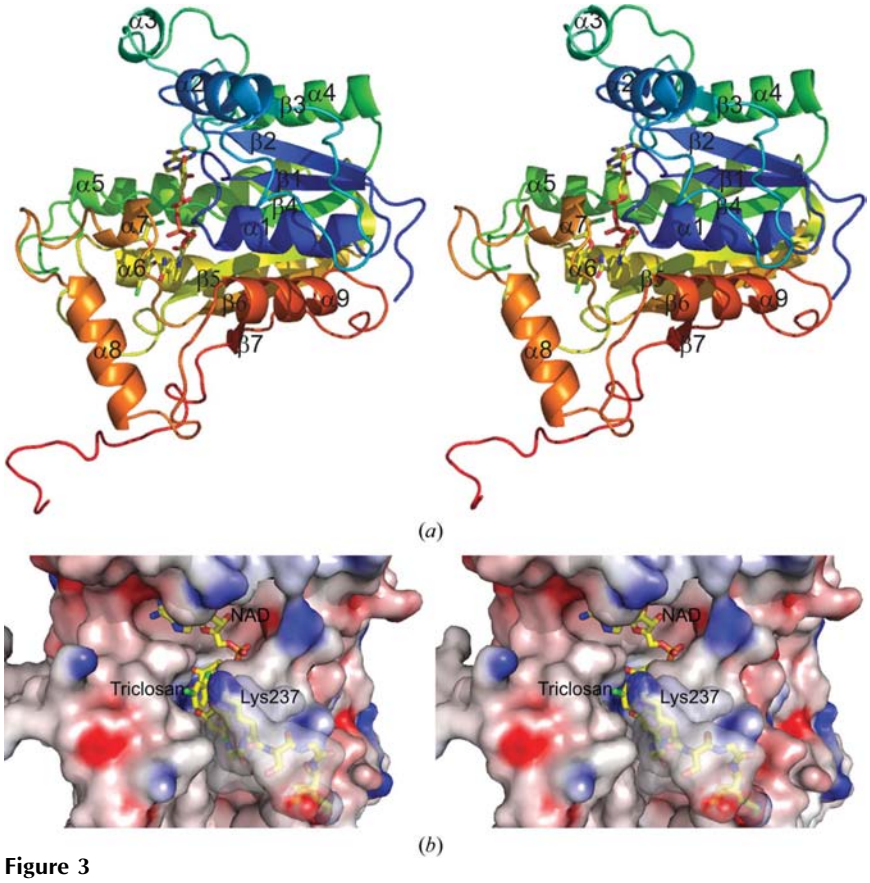

Figure 3
(*a*) Stereoview of the TgENR monomer in complex with NAD$^+$ and triclosan, coloured from blue (N-terminus) to red (C-terminus), with each secondary-structure element labelled. The atom colours for triclosan and the NAD$^+$ cofactor are yellow, blue, red, purple and green for carbon, nitrogen, oxygen, phosphorus and chlorine, respectively. (*b*) Surface representation of the TgENR monomer, showing the distribution of electrostatic charge. The position of the inserted loop, which is a feature of the apicomplexan family, is displayed below the partially transparent depiction of the protein surface as a stick representation and Lys237, which is fully conserved within the loop region of the apicomplexan family, is labelled. The position of the NAD$^+$ cofactor and the triclosan inhibitor which define the active site and the ACP-binding region are shown, demonstrating their proximity to Lys237.

3. Results and discussion

3.1. Overall fold of the complexes of TgENR and PfENR with NAD$^+$ and triclosan

The sequences of PfENR and TgENR share 42% identity and the folds of the enzymes are structurally very similar, with superposition of the core residues giving an r.m.s.d. of the C$^\alpha$ atoms of 0.6 Å. Each

monomer has overall dimensions of 45 × 50 × 60 Å and analysis of the secondary structure (as defined by the program *PROMOTIF*; Hutchinson & Thornton, 1996) shows that TgENR is formed of nine α-helices (α1–α9) comprising 108 residues (∼36%), seven β-strands (β1–β7) formed by 43 residues (∼14%), six 3_{10}-helices formed by 18 residues (∼6%) and a number of loops. The secondary-structure elements form an arrangement reminiscent of the Rossmann fold common to several nucleotide-binding enzymes (Rossmann *et al.*, 1974), with a parallel β-sheet flanked on one side by helices α3, α4, α5 and α6 and on the other by helices α1, α2 and α9, with the α7 and α8 helices lying at the C-terminal end of the β-sheet (Fig. 3*a*). The NAD^+ cofactor binds at the C-terminal end of the β-sheet and makes important interactions with residues at the end of or following each of the β-strands except for β7. Residues from helices α5, α6 and α8 form the inhibitor-binding pocket, with α7 forming a lid over the inhibitor and making extensive van der Waals interactions.

To date, several structures of PfENR have been published in complex with a variety of inhibitors [PDB codes 1nhg, 1nnu, 1nhw (Perozzo *et al.*, 2002), 1zsn, 1zw1, 1zxb, 1zxl (Freundlich *et al.*, 2005) and 1uh5 (Swarnamukhi *et al.*, 2004)], all of which display strong structural similarity, in particular around the NAD^+-binding sites, to the structure reported here. Moreover, in our crystal form and all of the previously solved PfENR structures, the loop region corresponding to the low-complexity insert common to the apicomplexan ENR family between residues Ile240 and Thr283 cannot be seen owing to disorder (Perozzo *et al.*, 2002; Pidugu *et al.*, 2004). In contrast, the smaller equivalent insert in TgENR between helices α7 and α8 (Gly236–Lys241) can be seen in subunit *A*, where it forms a loop that caps the ends of both helices and lies close to the bound inhibitor, but makes no direct contacts. The loop is linked to the flanking residues by a number of hydrogen bonds and an ion-pair interaction between the side chains of Lys241 and Asp249. Inspection of the ENR sequences from the majority of apicomplexan parasites shows an apparent conservation of Lys237 and of an SGE motif (Ser238–Glu240; Fig. 2). However, the difference in insert size makes it difficult to envisage conservation of the position of the SGE motif within the apicomplexan family and the sequence similarity may be only a chance resemblance. Interestingly, examination of an elec-

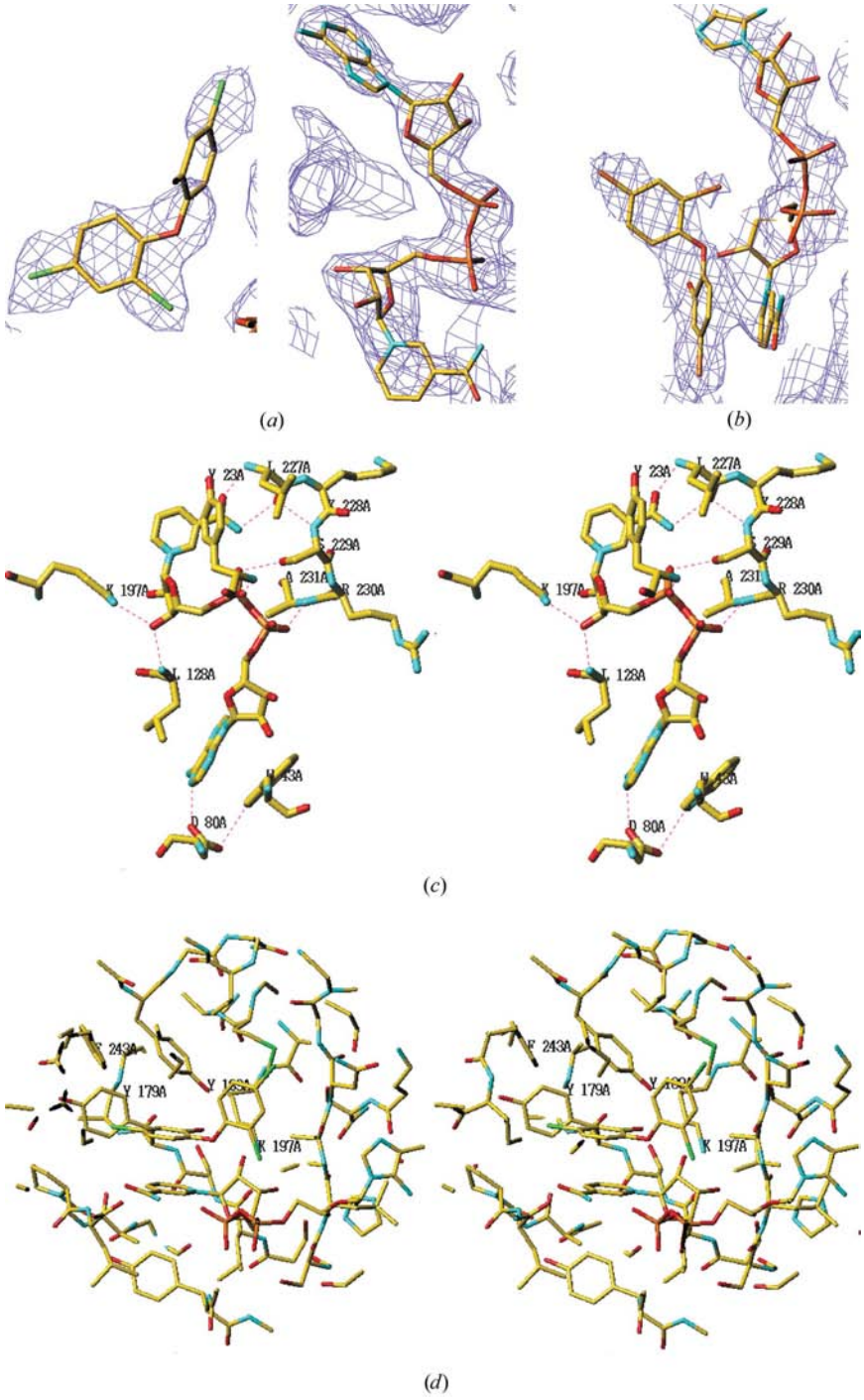

Figure 4
$2F_{obs} - 1F_{calc}$ electron-density maps for (*a*) PfENR and (*b*) TgENR contoured at 0.8σ produced after initial rigid-body refinement in *REFMAC*5 (Murshudov *et al.*, 1997). Both NAD^+ and triclosan were omitted from the model during refinement, but are represented in stick format in order to show their unambiguous position in the initial electron-density maps. (*c*) Stereo diagram of the residues responsible for forming a hydrogen-bonding network to the NAD^+ cofactor in subunit *A* of TgENR. (*d*) Stereo diagram of the triclosan-binding site of subunit *A* of TgENR, with the active-site residues Tyr179, Tyr189, Lys197 and Phe243 labelled. (*c*) and (*d*) use the colour scheme yellow, red, blue, green and orange for carbon, oxygen, nitrogen, chlorine and phosphorus, respectively, and were produced in *TURBO-FRODO* (Roussel *et al.*, 1990).

trostatic charge potential surface of a TgENR monomer shows that the face of the putative fatty-acid binding pocket is lined with several positively charged residues, including Lys237. Moreover, recent studies have strongly implicated helix α8 and its preceding residues (which would include the inserted region in apicomlexan parasites) in forming interactions with the acyl carrier protein, suggesting that Lys237 may play a role in binding the predominantly negatively charged acyl carrier protein (Rafi *et al.*, 2006; Fig. 3*b*).

3.2. The nicotinamide adenine dinucleotide-binding and triclosan-binding sites in PfENR and TgENR

Analysis of the electron-density maps in the region of the active site showed good density for the entire nucleotide cofactor and for the triclosan inhibitor in all subunits of the *P. falciparum* and *T. gondii* ENR enzymes, allowing their unambiguous positioning (Figs. 4*a* and 4*b*). The cofactor is bound in an extended conformation, with both ribose sugar rings adopting a C2′-*endo* conformer and the nicotinamide ring adopting a *syn* conformation. In TgENR, the adenine ring binds within a pocket formed by the main-chain atoms of residues Gly16–Ala18, Leu79–Ala82 and Leu128–Asn130 and by the side chains of residues Asp80, Asn130 and Asn152, with one face of the adenine ring forming π–π stacking interactions with Trp43 and with Leu128 making van der Waals interactions with the other face (Fig. 4*c*). The pyrophosphate moiety makes interactions with the main chain of Gly22, Tyr23 and Ala231, the side-chain atoms of Asp19, Ser229 and Ala231 and the dichloro ring of triclosan. The nicotinamide ribose moiety forms packing interactions with Tyr23, Ser127, Ala129, Leu177 and Lys197 and the bound inhibitor triclosan and forms a hydrogen bond from the nicotinamide ribose to the main-chain N atom of Leu128. The nicotinamide ring forms interactions with the side-chain atoms of Tyr23, Tyr179, Ala224 and Ser229 and with the main chain of Ser178, Ala224, Gly225, Pro226 and Leu227, with additional π–π stacking interactions with the 4-chlorophenoxy ring of triclosan. In addition to an extensive hydrogen-bonding network within the NAD^+-binding site, Leu227 forms two hydrogen bonds from its main-chain N and carbonyl O atoms to the amide O and N atoms of the nicotinamide, allowing the orientation to be unambiguously determined (Fig. 4*c*). The only significant difference in the binding of the NAD^+ cofactor between *T. gondii* and *P. falciparum* ENR is the positioning of Arg235 within PfENR, which in subunit *D* forms additional hydrogen bonds to the ribose moiety of the NAD^+ that are not seen in the other subunits.

The mode of triclosan binding to both PfENR and TgENR is very similar, with the 4-chlorophenoxy ring of the triclosan in TgENR forming van der Waals interactions with the side chains of Tyr179, Tyr189, Pro226, Ala232, Ile235, Phe243 and Ile244. The triclosan 2,4-dichlorophenoxy ring is located within a pocket formed by the peptide backbone of residues Leu128–Ala131, the pyrophosphate and nicotinamide moieties of NAD^+ and the side chains of Val134, Met193, Lys197, Ala231 and Ile235 (Fig. 4*d*).

3.3. Analysis of the apo TgENR structure

The overall fold of the apo TgENR structure is similar to that of the TgENR–NAD^+–triclosan complex, with an r.m.s.d. of the C^α atoms of 0.5 Å, and as such will not be described in detail. The most significant difference is the disorder of residues Lys228–Lys241, which form an ordered loop in the TgENR–NAD^+–triclosan complex, with Ala230–Ser232 forming a helix which packs against the triclosan inhibitor. This ordering of the loop region upon inhibitor binding has been seen within the bacterial and plant ENR family and has been implicated in the slow tight binding behaviour of triclosan (Seefeld *et al.*, 2003; Levy *et al.*, 2001; Qiu *et al.*, 1999; Roujeinikova *et al.*, 1999). Furthermore, PfENR has also been shown to display a two-step inhibition mechanism, with the initial formation of a stable triclosan complex being followed by a conformational change in the loop region, creating a tight inhibitor complex (Kapoor *et al.*, 2004). The only other difference is a slight movement of residues Trp43–Ser69, with Trp43 adopting a position such that it makes stacking interactions with the adenine ring of the NAD^+ cofactor in the TgENR–NAD^+–triclosan complex but is in a different position in the TgENR apo form.

3.4. Comparison of the apicomplexan ENR structures to those of other species

In order to investigate the similarities and differences between members of the ENR family, the structures of the *B. napus*, *Escherichia coli*, *T. gondii* and *P. falciparum* enzymes in complex with NAD^+ and triclosan were compared. As might be expected on the basis of their sequence similarity, the parasite ENRs are more closely related to the plant and cyanobacterial enzymes than those of bacterial origin, with both containing significant inserts both before and after β3 compared with their bacterial counterparts. The second of these inserts, which includes α3, is involved in the formation of a large groove around one of the twofold axes of the ENR tetramer which is absent in the bacterial ENR enzymes (Fig. 5). Furthermore, analysis of the ENR sequences from several plant, *Plasmodium*, *Eimeria* and *Toxoplasma* species reveals that the insert shows strong overall sequence conservation and moreover that all the sequences contain an R/K*x*NKRY motif. The position of this motif is such that the residues directly follow α3 and are all solvent-exposed, with the insert being unique to the ENR of plastid-containing organisms. A further striking feature of the surface-potential calculations is the overall positive charge of PfENR (pI 8.2) when compared with the *T. gondii*, *B. napus* and *E. coli* enzymes, whose pI values are 6.3, 5.4 and 5.6, respectively (Fig. 5). However, a role for the structural conservation of a large groove around one of the twofolds in the plant and apicomplexan ENRs has yet to be determined.

3.5. Conservation of water in the active site

In a proposed mechanism of catalysis for the short-chain dehydrogenase reductase (SDR) family, to which ENR belongs, a key role is played by two water molecules within the

active site which form part of a proton-relay system to the bulk solvent which replenishes the proton donated to the substrate by the catalytic tyrosine (Benach *et al.*, 1998; Filling *et al.*, 2002; Price *et al.*, 2004; Schlieben *et al.*, 2005). Moreover, the superposition of 3β/17β-hydroxysteroid dehydrogenase, a member of the SDR family, onto PfENR reveals the strong spatial conservation of a water molecule next to the conserved catalytic Lys residue, with close spatial conservation of the second water molecule (Fig. 6*a*). These two waters reside in a pocket formed by nine residues (His131, Leu133, Ala134, Ser158, Ser159, Ser181, Leu182, Lys202 and Leu205), all of which are fully conserved within the *T. gondii* and *Plasmodium* ENR family, with the spatial conservation of both water molecules in *E. coli*, *Helicobacter pylori*, *M. tuberculosis* and *B. napus* ENRs (PDB codes 1d8a, 1jvf, 1bvr and 1d70, respectively; Fig. 6*a*).

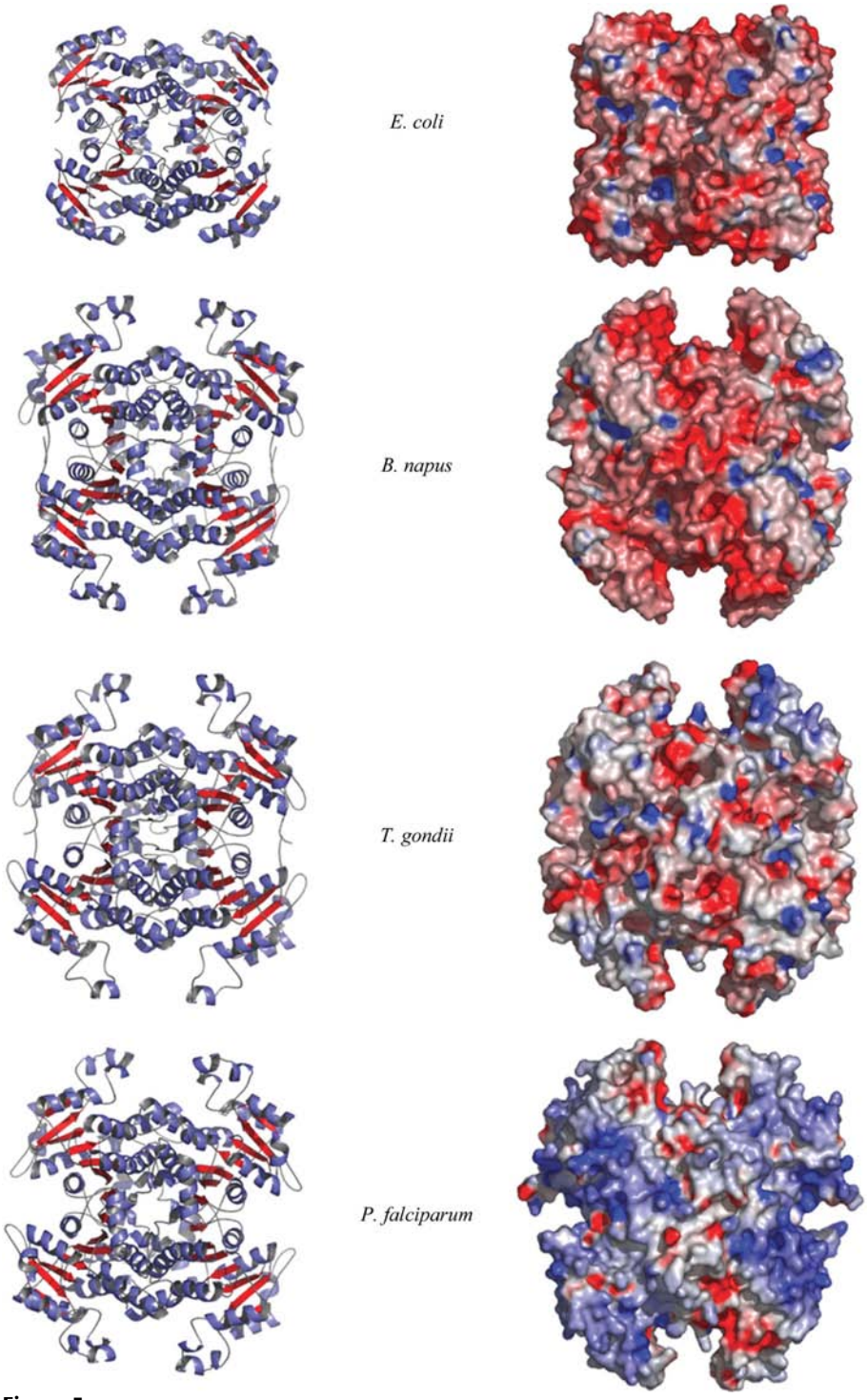

Figure 5
A representation of the surface potential and structure of the *E. coli*, *B. napus*, *T. gondii* and *P. falciparum* ENR tetramers. The large groove present in the parasite and plant ENR enzymes is formed by an insert after helix α3, which is not found within the bacterial ENR family. All four structures are on the same scale and are coloured by electrostatic charge, with red and blue representing negative and positive charges, respectively.

3.6. Inhibitor binding and conformational flexibility in the structure of ENR

The program *ESCET* (Schneider, 2000) was used to compare the position of the C^{α} atoms in the EcENR–NAD^+ complex and the various EcENR–inhibitor complexes that have been solved through co-crystallization, which includes the diazaborines, triclosan and the aminopyridine derivative compound 29 [(*E*)-*N*-methyl-*N*-(1-methyl-1*H*-indol-3-ylmethyl)-3-(7-oxo-5,6,7,8-tetrahydro-1,8-naphthyridin-3-yl)acrylamide] (Seefeld *et al.*, 2003; Levy *et al.*, 1999; Heerding *et al.*, 2001). All other inhibitor complexes [imidazole (Miller *et al.*, 2002) and benzamide-based inhibitors (Payne *et al.*, 2002)] were ignored owing to the observation that for ENR inhibitor soaking and co-crystallization are not equivalent (Heerding *et al.*, 2001). This comparison showed only one region of difference, around the inhibitor-binding site between residues 190–209 in EcENR (229–249 in TgENR), corresponding to α6 and α7 (α7 and α8 in the Tg and Pf enzymes) and the flanking residues (Fig. 6*b*). In the structure of the EcENR–NAD^+ complex and in the TgENR apo enzyme, this region corresponds to a disordered loop. However, on binding of inhibitors belonging to the diazaborine or aminopyridine family or triclosan, residues 190–198 in EcENR adopt a helical structure (Seefeld *et al.*, 2003; Levy *et al.*, 2001; Qiu *et al.*, 1999; Roujeinikova *et al.*, 1999). Furthermore, comparing the structures of the various inhibitor complexes reveals that in the

E. coli enzyme both helices α6 and α7 and their connecting polypeptide chain adopt significantly different positions which appear to be associated with the size and position of the inhibitor, with Gly199, Ile200, Lys201, Phe203 and Met206 showing the greatest movements. In addition, those residues within the *E. coli* enzyme that show a significant change in the side-chain conformation between the structures of the various inhibitor complexes are a subset of those residues that display the greatest main-chain flexibility.

Analysis of the inhibitor-binding pocket for both PfENR and TgENR reveals that both are very similar, with only one sequence difference at position 131 (TgENR numbering), a residue that points away from the pocket, in which Ala is replaced by Gly. Analysis of the sequence conservation across the ENR family of the 11 residues shown to be involved in binding compound 29 in EcENR (Gly93, Phe94, Ala95, Leu101, Tyr146, Tyr156, Met159, Ala196, Ile200, Lys201 and Met206) reveals that three residues are not conserved (Met206, Lys201 and Leu101). In contrast, of the ten residues which make up the triclosan-binding pocket in EcENR (Gly93, Phe94, Ala95, Tyr146, Tyr156, Lys163, Ala196, Ala197, Ile200 and Phe203), all are either fully or strongly conserved.

3.7. Implications of the apicomplexan ENR structure for drug design

Modelling studies on TgENR using the aminopyridine-based inhibitor compound 29 (Seefeld *et al.*, 2003) suggests that if the inhibitor were bound in the same position as

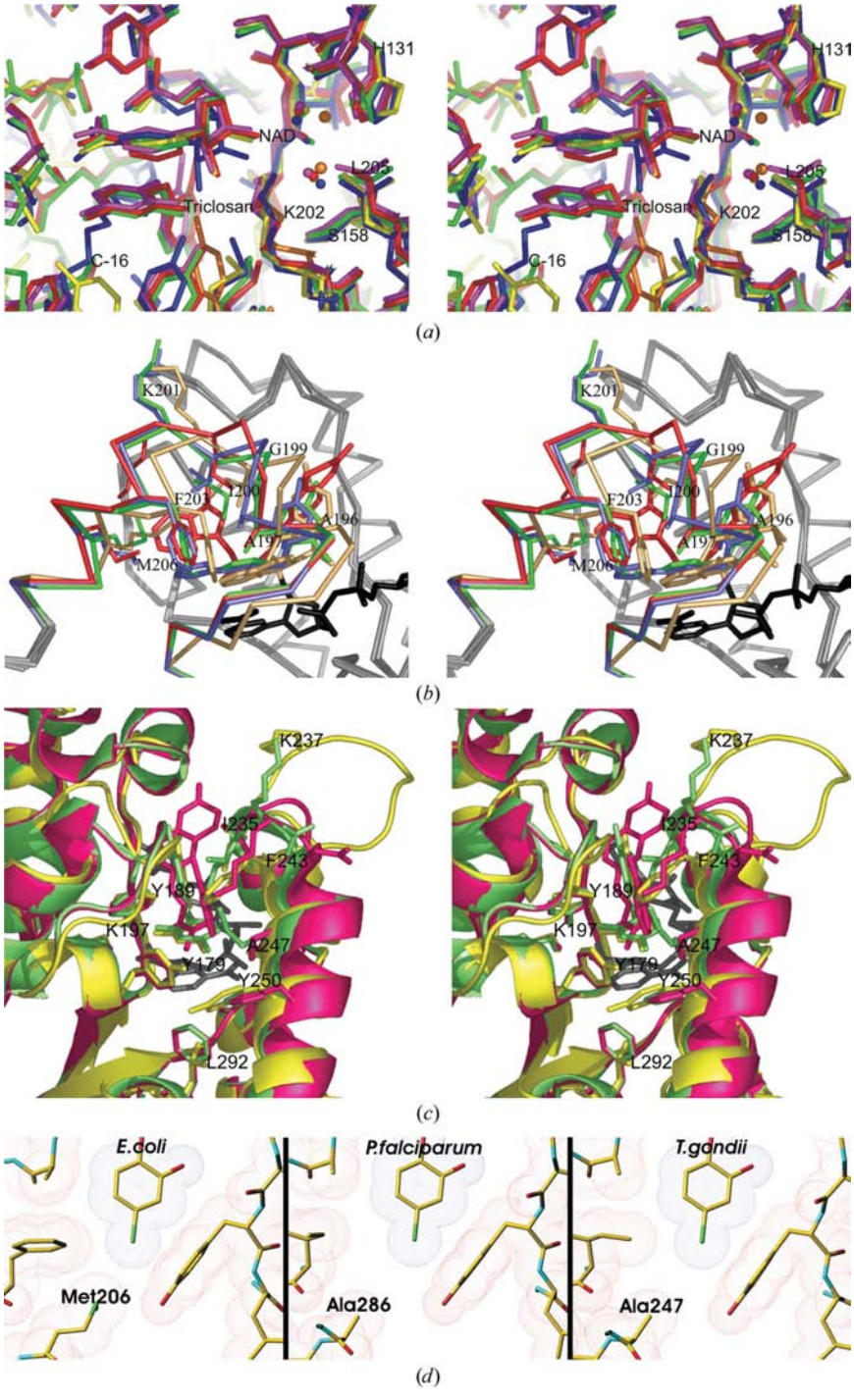

Figure 6
(*a*) Stereoview of the complexes of ENR from *P. falciparum* (red), *B. napus* (magenta) and *E. coli* (green) with NAD^+ and triclosan and of *M. tuberculosis* ENR in complex with NAD^+ (yellow) and with NAD^+ and C-16 fatty-acid substrate, labelled C-16 (blue). In addition to ENR, the catalytic Lys and Tyr residues of the 3β/17β-hydroxysteroid dehydrogenase SDR enzyme are shown (orange). The positions of two apparently conserved water molecules in the active site of each of these enzymes are shown as coloured spheres. The water molecules in ENR form interactions with several active-site residues, including the catalytic lysine and the 3′-OH of the NAD^+ nicotinamide ribose. (*b*) Stereoview of the differences in position for the substrate-binding helix (α6 or α7 in the bacterial and plant/parasite ENRs, respectively) upon the binding of different inhibitors to EcENR. The loop region which moves upon inhibitor binding and the corresponding inhibitor in the EcENR complex is coloured cream, green, red and blue for benzo-diazaborine, triclosan, compound 29 and imidazole, respectively. For clarity, only those residues which make direct interactions with the inhibitors in one or more complexes are shown in stick representation and numbered. (*c*) Close-up of the *E. coli* ENR–NAD^+–compound 29 complex (red), *E. coli* ENR–NAD^+–triclosan complex (green) and *T. gondii* ENR–NAD^+–triclosan complex (yellow) superimposed, with critical residues involved in inhibitor binding shown in stick representation with the TgENR structure numbered. Modelling of compound 29 into the structure of TgENR in complex with NAD^+ (black) and triclosan shows that a severe steric clash might occur owing to the difference in the structure close to Phe242, which in EcENR is replaced by Asp202. (*d*) Representation of the van der Waals surfaces of the protein (coloured red) within the active site of *E. coli*, *T. gondii* and *P. falciparum* ENR close to the phenolic ring of triclosan (blue). Both the *P. falciprum* and *T. gondii* ENRs appear to be less closely packed with triclosan owing to the presence of an alanine residue at the base of the binding pocket in contrast to the bulkier methionine residue in the *E. coli* enzyme.

seen in EcENR there would be likely to be severe clashes with residues Ala231, Ile235, Phe243, Ile244 and Ala247. However, if the structure of TgENR is as flexible as that of EcENR, then this steric hindrance may be relieved to some extent by the movement of α7 and α8. However, even then an adverse contact with the side chain of Phe243 which lies at the start of α8 in TgENR and makes van der Waals contact with the chlorine of the 4-chlorophenoxy ring of triclosan might remain (Fig. 6*c*). In EcENR and other members of the bacterial ENR family, the residue at this position is smaller (Gly, Asn or Asp) and points towards the solvent.

Sequence analysis of the ENR family shows that in the region of the inhibitor-binding pocket the bacterial and plant ENR sequences all contain a bulky hydrophobic residue (Met, Leu or Ile) close to the 4-chlorophenoxy ring of triclosan, whereas the ENRs of apicomplexan species such as *T. gondii* and *Plasmodium* have a fully conserved alanine residue (Fig. 2 and Fig. 6*d*). This sequence change produces an increase in the space at the base of the binding pocket and reduces the van der Waals packing interactions with the triclosan and aminopyridine-based inhibitors in comparison to other members of the ENR family (Fig. 6*d*). Analysis of the structure suggests that the production of a triclosan analogue containing more bulky constituents to replace the Cl atom at the 4-position of the chlorophenoxy ring might allow the van der Waals contacts to be enhanced, improving inhibitor binding. Indeed, recent studies have shown that whilst subtle changes at this position did not improve inhibitor binding, the effects were markedly different between PfENR and ENRs of bacterial origin (Chhibber *et al.*, 2006; Sullivan *et al.*, 2006; Freundlich *et al.*, 2005, 2006). This finding may allow the development of novel ENR inhibitors that have a significantly higher affinity for the apicomplexan ENR family, which have been shown in this study to have a significantly similar binding-site architecture.

4. Conclusion

The structures of both PfENR and TgENR have permitted a detailed comparison to be made between members of the apicomplexan, plant and bacterial ENR families. This has revealed that the apicomplexan enzymes differ in the active-site region when compared with those of bacterial origin. This would suggest that ENR inhibitors developed as part of an antibacterial program might require optimization in order to act as antiparasitic agents. Despite these apparent problems, the prospect for the design of a family of drugs which could target the apicomplexan parasites is good and the current study has provided a toolkit which could be utilized by future drug-discovery programs.

References

Baldock, C., Rafferty, J. B., Sedelnikova, S. E., Baker, P. J., Stuitje, A. R., Slabas, A. R., Hawkes, T. R. & Rice, D. W. (1996). *Science*, **274**, 2107–2110.

Barton, G. J. (1993). *Protein Eng.* **6**, 37–40.

Benach, J., Atrian, S., Gonzalez-Duarte, R. & Ladenstein, R. (1998). *J. Mol. Biol.* **282**, 383–399.

Bhargava, H. N. & Leonard, P. A. (1996). *Am. J. Infect. Control*, **24**, 209–218.

Boyer, K. & McLeod, R. (2002). *Principles and Practice of Pediatric Infectious Diseases*, 2nd ed., edited by S. Long, C. Proeber & L. Pickering, pp. 1303–1322. New York: Churchill Livingstone.

Breman, J. G. (2001). *Am. J. Trop. Med. Hyg.* **64**, 1–11.

Chhibber, M., Kumar, G., Parasuraman, P., Ramya, T. N., Surolia, N. & Surolia, A. (2006). *Bioorg. Med. Chem.* **14**, 8096–8098.

DeLano, W. L. (2002). *PyMOL Molecular Graphics System.* DeLano Scientific LLC, San Carlos, CA, USA. http://www.pymol.org.

Evans, P. R. (1997). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **33**, 22–24.

Fast, N. M., Kissinger, J. C., Roos, D. S. & Keeling, P. J. (2001). *Mol. Biol. Evol.* **18**, 418–426.

Filling, C., Berndt, K. D., Benach, J., Knapp, S., Prozorovski, T., Nordling, E., Ladenstein, R., Jornvall, H. & Oppermann, U. (2002). *J. Biol. Chem.* **277**, 25677–25684.

Freundlich, J. S., Anderson, J. W., Sarantakis, D., Shieh, H. M., Yu, M., Valderramos, J. C., Lucumi, E., Kuo, M., Jacobs, W. R. Jr, Fidock, D. A., Schiehser, G. A., Jacobus, D. P. & Sacchettini, J. C. (2005). *Bioorg. Med. Chem. Lett.* **15**, 5247–5252.

Freundlich, J. S., Yu, M., Lucumi, E., Kuo, M., Tsai, H. C., Valderramos, J. C., Karagyozov, L., Jacobs, W. R. Jr, Schiehser, G. A., Fidock, D. A., Jacobus, D. P. & Sacchettini, J. C. (2006). *Bioorg. Med. Chem. Lett.* **16**, 2163–2169.

Heerding, D. A., Chan, G., DeWolf, W. E., Fosberry, A. P., Janson, C. A., Jaworski, D. D., McManus, E., Miller, W. H., Moore, T. D., Payne, D. J., Qiu, X., Rittenhouse, S. F., Slater-Radosti, C., Smith, W., Takata, D. T., Vaidya, K. S., Yuan, C. C. & Huffman, W. F. (2001). *Bioorg. Med. Chem. Lett.* **11**, 2061–2065.

Higgins, D. & Thompson, J. (1994). *Nucleic Acids Res.* **22**, 4673–4680.

Hill, D. & Dubey, J. P. (2002). *Clin. Microbiol. Infect.* **8**, 634–640.

Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**, 212–220.

Kapoor, M., Dar, M. J., Surolia, A. & Surolia, N. (2001). *Biochem. Biophys. Res. Commun.* **289**, 832–837.

Kapoor, M., Reddy, C. C., Krishnasastry, M. V., Surolia, N. & Surolia, A. (2004). *Biochem. J.* **381**, 719–724.

Kohler, S., Delwiche, C. F., Denny, P. W., Tilney, L. G., Webster, P., Wilson, R. J. M., Palmer, J. D. & Roos, D. S. (1997). *Science*, **275**, 1485–1489.

Lamzin, V. S. & Wilson, K. S. (1997). *Methods Enzymol.* **277**, 269–305.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

Leslie, A. G. W. (1992). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **26**.

Levy, C. W., Baldock, C., Wallace, A. J., Sedelnikova, S., Viner, R. C., Clough, J. M., Stuitje, A. R., Slabas, A. R., Rice, D. W. & Rafferty, J. B. (2001). *J. Mol. Biol.* **309**, 171–180.

Levy, C. W., Roujeinikova, A., Sedelnikova, S., Baker, P. J., Stuitje, A. R., Slabas, A. R., Rice, D. W. & Rafferty, J. B. (1999). *Nature (London)*, **398**, 383–384.

McLeod, R., Muench, S. P., Rafferty, J. B., Kyle, D. E., Mui, E. J., Kirisits, M. J., Mack, D. G., Roberts, C. W., Samuel, B. U., Lyons, R. E., Dorris, M., Milhous, W. K. & Rice, D. W. (2001). *Int. J. Parasitol.* **31**, 109–113.

Magnuson, K., Jackowsk, S., Rock, C. O. & Cronan, J. E. Jr (1993). *Microbiol. Rev.* **57**, 552–642.

Matthews, B. W. (1977). *The Proteins*, Vol. 3, edited by H. Neurath & R. L. Hill, pp. 468–477. New York: Academic Press.

Mead, P. S., Slutsker, L., Dietz, V., McCaig, L. F., Bresee, J. S., Shapiro, C., Griffin, P. M. & Tauxe, R. V. (1999). *Emerg. Infect.* **5**, 607–625.

Miller, W. H., Seefeld, M. A., Newlander, K. A., Uzinskas, I. N., Burgess, W. J., Heerding, D. A., Yuan, C. C., Head, M. S., Payne,

D. J., Rittenhouse, S. F., Moore, T. D., Pearson, S. C., Berry, V., DeWolf, W. E. Jr, Keller, P. M., Polizzi, B. J., Qiu, X., Janson, C. A. & Huffman, W. F. (2002). *J. Med. Chem.* **45**, 3246–3256.
Muench, S. P., Prigge, S. T., Zhu, L., Kirisits, M. J., Roberts, C. W., Wernimont, S., McLeod, R. & Rice, D. W. (2006). *Acta Cryst.* F**62**, 604–606.
Muench, S. P., Rafferty, J. B., McLeod, R., Rice, D. W. & Prigge, S. T. (2003). *Acta Cryst.* D**59**, 1246–1248.
Murshudov, G., Vagin, A. & Dodson, E. (1997). *Acta Cryst.* D**53**, 240–255.
Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.
Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins*, **11**, 281–296.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Payne, D. J. *et al.* (2002). *Antimicrob. Agents Chemother.* **46**, 3118–3124.
Perozzo, R., Kuo, M., Sidhu, A. S., Valiyaveetil, J. T., Bittman, R., Jacobs, W. R. Jr, Fidock, D. A. & Sacchettini, J. C. (2002). *J. Biol. Chem.* **277**, 13106–13114.
Pidugu, L. S., Kapoor, M., Surolia, N., Surolia, A. & Suguna, K. J. (2004). *J. Mol. Biol.* **343**, 147–155.
Price, A. C., Zhang, Y. M., Rock, C. O. & White, S. W. (2004). *Structure*, **12**, 417–428.
Qiu, X., Janson, C. A., Court, R. I., Smyth, M. G., Payne, D. J. & Abdel-Meguid, S. S. (1999). *Protein Sci.* **8**, 2529–2532.
Rafi, S., Novichenok, P., Kolappan, S., Zhang, X., Stratton, C. F., Rawat, R., Kisker, C., Simmerling, C. & Tonge, P. J. (2006). *J. Biol. Chem.* **281**, 39285–39293.
Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). *Nature (London)*, **250**, 194–199.
Roujeinikova, A., Sedelnikova, S., Boer, J. G., Stuitje, A. R., Slabas, A. R., Rafferty, J. & Rice, D. W. (1999). *J. Biol. Chem.* **274**, 30811–30817.
Roussel, A., Fontecilla-Camps, J. C. & Cambillau, C. (1990). *Acta Cryst.* A**46**, C66–C67.
Rozwarski, D. A., Vilchèze, C., Sugantino, M., Bittman, R. & Sacchettini, J. C. (1999). *J. Biol. Chem.* **274**, 15582–15589.
Samuel, B. U., Hearn, B., Mack, D., Wender, P., Rothbard, J., Kirisits, M. J., Mui, E., Wernimont, S., Roberts, C. W., Muench, S. P., Rice, D. W., Prigge, S. T., Law, A. B. & McLeod, R. (2003). *Proc. Natl Acad. Sci. USA*, **100**, 14281–14286.
Schlieben, N. H., Niefind, K., Muller, J., Riebel, B., Hummel, W. & Schomburg, D. (2005). *J. Mol. Biol.* **349**, 801–813.
Schneider, T. R. (2000). *Acta Cryst.* D**56**, 714–721.
Seefeld, M. A. *et al.* (2003). *J. Med. Chem.* **46**, 1627–1635.
Sibley, L. D. (2003). *Traffic*, **4**, 581–586.
Smith, S. (2003). *Prog. Lipid Res.* **42**, 289–317.
Sullivan, T. J., Truglio, J. J., Boyne, M. E., Novichenok, P., Zhang, X., Stratton, C. F., Li, H. J., Kaur, T., Amin, A., Johnson, F., Slayden, R. A., Kisker, C. & Tonge, P. J. (2006). *ACS Chem. Biol.* **1**, 43–53.
Surolia, N. & Surolia, A. (2001). *Nature Med.* **7**, 167–173.
Swarnamukhi, P. L., Kapoor, M., Surolia, N., Surolia, A. & Suguna, K. (2004). *J. Mol. Biol.* **343**, 147–155.
Waller, R. F., Keeling, P. J., Donald, R. G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A. F., Besra, G. S., Roos, D. S. & McFadden, G. I. (1998). *Proc. Natl Acad. Sci. USA*, **95**, 12352–12357.
Zuther, E, Johnson, J. J., Haselkorn, R., Mcleod, R. & Gornicki, P. (1999). *Proc. Natl Acad. Sci. USA*, **96**, 3387–13392.