Available online at www.sciencedirect.com
SCIENCE DIRECT
Bioorganic & Medicinal Chemistry 14 (2006) 4433–4443
Bioorganic & Medicinal Chemistry

Elucidation of sulfadoxine resistance with structural models of the bifunctional *Plasmodium falciparum* dihydropterin pyrophosphokinase–dihydropteroate synthase

Tjaart A. P. de Beer, Abraham I. Louw* and Fourie Joubert

Bioinformatics and Computational Biology Unit, Department of Biochemistry, Faculty of Natural and Agricultural Sciences, University of Pretoria, Pretoria 0002, South Africa

Received 18 January 2006; revised 16 February 2006; accepted 17 February 2006
Available online 6 March 2006

Abstract—Resistance of the most virulent human malaria parasite, *Plasmodium falciparum*, to antifolates is spreading with increasing speed, especially in Africa. Antifolate resistance is mainly caused by point mutations in the *P. falciparum* dihydropteroate synthase (DHPS) and dihydrofolate reductase (DHFR) target proteins. Homology models of the bifunctional *P. falciparum* dihydropterin pyrophosphokinase–dihydropteroate synthase (PPPK–DHPS) enzyme as well as the separate domains complete with bound substrates were constructed using the crystal structures of *Saccharomyces cerevisiae* (PPPK–DHPS), *Mycobacterium tuberculosis* (DHPS), *Bacillus anthracis* (DHPS), and *Escherichia coli* (PPPK) as templates. The resulting structures were subsequently solvated and refined using molecular dynamics. The active site residues of DHPS are highly conserved in *S. cerevisiae*, *M. tuberculosis*, *E. coli*, *S. aureus*, and *B. anthracis*, an attribute also shared by *P. falciparum* DHPS. Sulfadoxine was superimposed into the equivalent position of the *p*-aminobenzoic acid substrate and its binding parameters were refined using minimization and molecular dynamics. Sulfadoxine appears to interact mainly with *P. falciparum* DHPS mainly through hydrophobic interactions. Rational explanations are provided by the model for the sulfadoxine resistance-causing effects of four of the five known mutations in *P. falciparum* DHPS. A possible structure for the bifunctional PPPK–DHPS was derived from the structure from the *S. cerevisiae* bifunctional enzyme. The active site residues of *P. falciparum* PPPK are also conserved when compared to *S. cerevisiae*, *Haemophilus influenzae*, and *E. coli*. The informative nature of these models opens up avenues for structure-based drug design approaches toward the development of alternative and more effective inhibitors of *P. falciparum* PPPK–DHPS.

1. Introduction

Malaria infections cause up to 500 million clinical cases and nearly three million fatalities annually, mostly in sub-Saharan Africa. *Plasmodium falciparum* is the most virulent form of the four species infecting humankind, and the global prevalence of parasite resistance to chloroquine and the antifolate combination of pyrimethamine–sulfadoxine are a matter of grave concern.[1–3] Whereas it took several decades before clinical resistance developed against chloroquine, the pace for resistance selection to pyrimethamine–sulfadoxine has been exceptionally fast, which has limited the useful therapeutic life of this drug combination to about five years.[4] Pyrimethamine is a competitive inhibitor of *P. falciparum* dihydrofolate reductase (DHFR; EC 1.5.1.3), which occurs as the N-terminal domain of a bifunctional enzyme with thymidylate synthase (TS; EC 2.1.1.45) as partner. Sulfadoxine (Fig. 1) is a competitive inhibitor of dihydropteroate synthase (DHPS; EC 2.5.1.15), another bifunctional enzyme together with dihydropterin pyrophosphokinase (PPPK; EC 2.7.6.3) at its N-terminal end.[5–8] PPPK catalyzes the transfer of a pyrophosphate group from ATP to 6-hydroxymethyl-7,8-dihydropterin

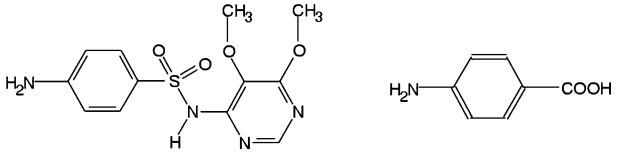

Figure 1. The structures of sulfadoxine (left), a *p*ABA analogue, and *p*ABA (right).

Keywords: Malaria; *Falciparum*; Dihydropteroate; DHPS; PPPK; Sulfadoxine; Homology model.
* Corresponding author. Tel.: +27 12 420 2480; fax: +27 12 362 5302; e-mail: braam.louw@bioagric.up.ac.za

0968-0896/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2006.02.035

to form dihydropterin pyrophosphate (DHP). DHPS condenses DHP and the substrate *p*-aminobenzoic acid (*p*ABA, Fig. 1) to form dihydropteroate as a substrate for folate synthesis. The combination of pyrimethamine and sulfadoxine acts synergistically and when either component is compromised, the effectiveness of inhibiting *P. falciparum* is dramatically reduced. Five point mutations in DHPS[1,7–10] have been linked with resistance to sulfonamides, being Ala 437 Gly, Ser 436 Phe/Ala, Lys 540 Glu, Ala 581 Gly, and Ala 613 Thr/Ser.[1,7–10] Recent studies have shown that two mutations in DHFR (Asn 108 and Arg 59) combined with one in DHPS (Glu 540) are indicative of pyrimethamine–sulfadoxine therapeutic failure in cases of uncomplicated *P. falciparum* malaria.[11]

Molecular modeling has proved to be very useful in elucidating structural interactions between *P. falciparum* DHFR and multiple antimalarial antifolates, and in explaining the poor binding affinities of antifolate antimalarials caused by specific steric constraints imposed by the mutated residues in resistant parasites.[12] Many of the features revealed by modeling efforts were recently confirmed by the crystal structures of wild-type and antifolate-resistant *P. falciparum* DHFR-TS.[5] In the absence of a crystal structure of *P. falciparum* PPPK–DHPS, the detailed mechanism behind sulfadoxine resistance, and that of other sulfonamide drugs caused by the observed amino acid substitutions, has not been elucidated.

Korsinczky et al. recently constructed homology models of *Plasmodium vivax* and *P. falciparum* DHPS, which showed that the replacement of Ala 613 in *P. falciparum* DHPS by Val 585 in *P. vivax* DHPS is most likely responsible for the innate sulfadoxine resistance of *P. vivax*.[13] Ten sulfadoxine contact residues each were identified for *P. falciparum* and *P. vivax* DHPS but only three of the five amino acid residues could be associated with resistance traits. Ser 436, Ala 437, and Ala 613 were shown to be involved in direct contact with sulfadoxine. Both inserts in *P. falciparum* DHPS were however excluded from their model as well as Mg^{2+} and substrates, the latter of which have been shown to influence the orientation of sulfadoxine and other sulfa-drugs in the active site.[14–16] Here, we present an alternative homology model of *P. falciparum* DHPS in which one insert as well as Mg^{2+} and the substrate DHP were included during docking of sulfadoxine in the active site as well as a model of *P. falciparum* PPPK, which includes the substrates. All the models were refined with molecular dynamics (MD) simulations. A positive correlation between the stability of sulfadoxine-binding in the active sites of the wild-type *P. falciparum* DHPS, and mutants with varying levels of resistance, could be demonstrated. A direct comparison with the model of Korsinczky has however not been possible since the coordinates of the latter model are not yet available. A possible structure of the bifunctional enzyme and linker region between *P. falciparum* PPPK and DHPS could be derived from the crystal structure of the *Saccharomyces cerevisiae* bifunctional enzyme.[17] A better understanding of the molecular mechanism by which these mutations confer antifolate resistance without seriously compromising the catalytic activity of *P. falciparum* DHPS would enable the development of new drug leads, as was shown for *P. falciparum* DHFR.[12,18–20]

2. Results and discussion

2.1. DHPS domain structure

Individual models were built to investigate the different combinations of mutations in *P. falciparum* DHPS. The crystal structure of *Mycobacterium tuberculosis* DHPS was selected as one of the templates for homology modeling of the DHPS domain based on a sequence identity of 31% with *P. falciparum* and the conserved nature of the active site residues between *P. falciparum* and *M. tuberculosis* DHPS.[21] The *Bacillus anthracis* structure was also selected due to the inclusion of the product analogue, pteroic acid, in the crystal structure.[22] The active site residues in contact with DHP appear to be highly conserved between the enzymes from the different organisms as deduced from the sequence alignment (Fig. 2). Several studies have indicated that parasite-specific inserts are important for the functioning of *P. falciparum* bifunctional enzymes.[19,20] Therefore, the PPPK–DHPS sequences of *P. vivax*, *Plasmodium yoelii yoelii*, *Plasmodium berghei*, and *Plasmodium chabaudi* were also included to assist in the delineation of the inserts. The five *Plasmodial* sequences have two DHPS parasite-specific inserts when compared to the *M. tuberculosis* and *B. anthracis* sequences, a feature that has been reported for many other *Plasmodial* proteins. *S. cerevisiae* also has species-specific inserts in the DHPS domain when compared to other species. There is a high degree of conservation in the sequences of the *Plasmodial* DHPS domains except for the lengths of the parasite-specific inserts. DHPS parasite-specific insert 1 is conserved between *P. berghei*, *P. yoelii yoelii*, and *P. chabaudi* but not in *P. vivax* and *P. falciparum*, either in length or sequence. *P. chabaudi*, however, has an extra eight-residue insert within the conserved sequence of DHPS insert 1. DHPS insert 2 also appears to be highly conserved between the five *Plasmodial* species, especially at the C-terminal end. However, *P. vivax* and *P. falciparum* have an extra 58 and five residues, respectively, inserted at the N-terminal side of insert 2, whereas *P. yoelii yoelii* has an extra 18 residue insert interrupting the predicted α-helix (Fig. 1) shared between the different *Plasmodial* species. Interestingly, the unique insert in *P. vivax* consists of a 7-amino acid motif ([G,A,S]-K-L-T-N-G-[D,E]), which is repeated nine times.

From the model it was apparent that *P. falciparum* DHPS is part of the triosephosphate isomerase family (TIM barrel) of enzymes and a search of SCOP confirmed this (Figs. 3 and 4). The wild-type model with either pteroic acid or with both substrates, and models with both substrates and with different combinations of mutations, were built. All eight residues in the model, which are predicted to be in contact with the substrate, correspond to those of the known crystal structure contacts (Table 1). The pteroic acid in the model was

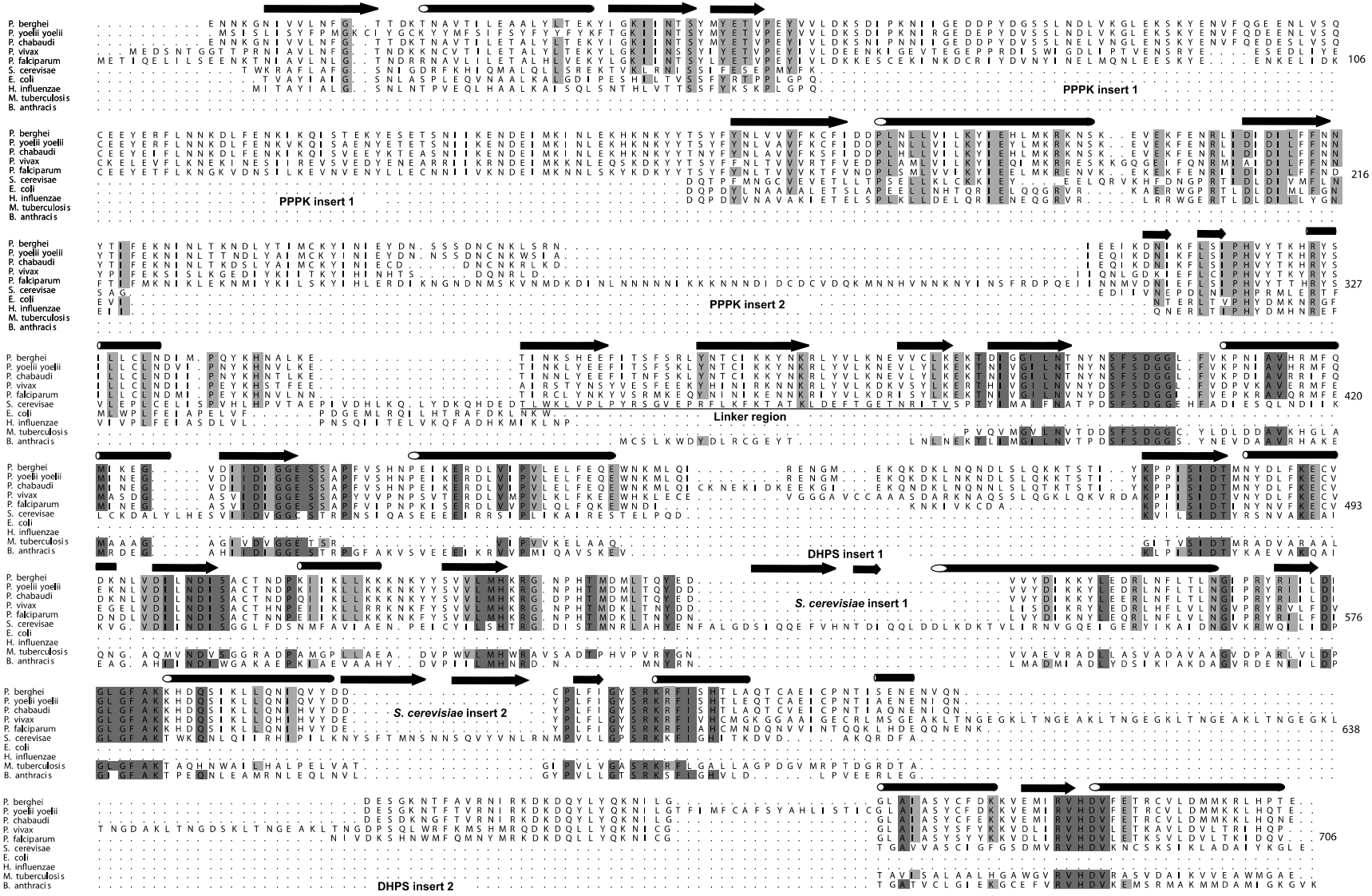

Figure 2. The alignment of the complete bifunctional PPPK–DHPS sequences from *Plasmodium berghei*, *Plasmodium yoelii yoelii*, *Plasmodium chabaudi*, *Plasmodium vivax*, *Plasmodium falciparum*, and *Saccharomyces cerevisiae*. The sequences of *Escherichia coli* and *Haemophilus influenzae* were aligned with the PPPK domain and the sequences of *Mycobacterium tuberculosis* and *Bacillus anthracis* were aligned with the DHPS domain. Residues conserved throughout the sequences are shaded. Only the residue numbers for the full PPPK–DHPS from *P. falciparum* are given. Species-specific inserts are labeled as well as the linker region between the two domains. Arrows indicate β-plates and cylinders indicate α-helices.

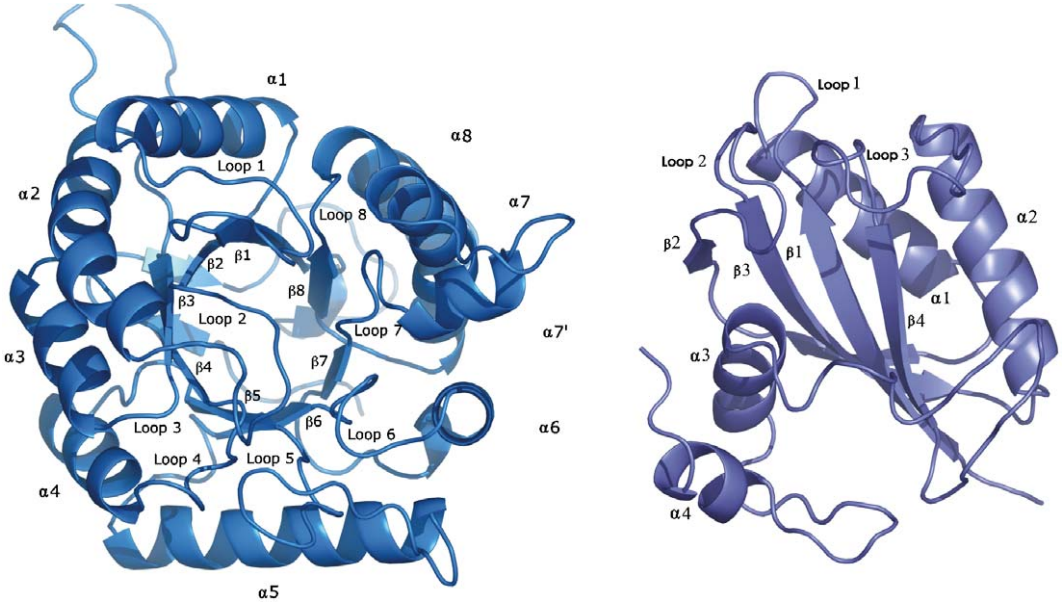

Figure 3. A steric view of each domain of the *Plasmodium falciparum* PPPK–DHPS model. DHPS is colored blue, PPPK is colored green and parasite-specific inserts and locations are indicated in yellow. The substrates are shown in the active site with the metal ion colored gray.

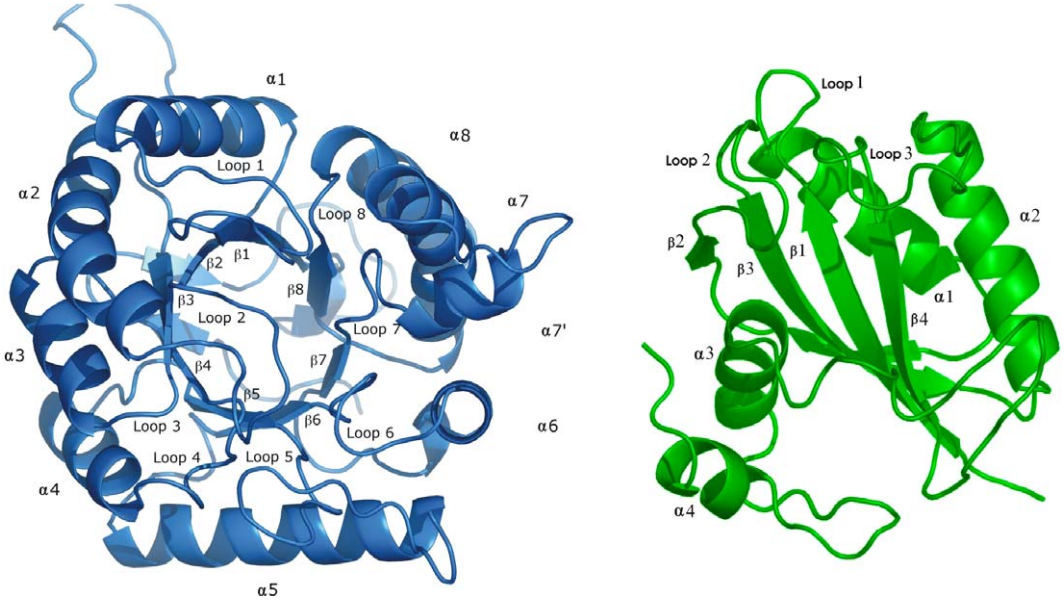

Figure 4. An annotated view of the *P. falciparum* PPPK and DHPS domain models. DHPS is on the left and PPPK on the right.

Table 1. The contacts between DHPS and DHP in the various known crystal structures compared to the contacts in the PfDHPS model

M. tuberculosis	*S. aureus*	*E. coli*	*B. anthracis*	*P. falciparum* model
Asn 105	Asn 103	Asn 115	Asn 120	Asn 502
Met 130	Met 128	Met 129	Met 145	Met 529
Asp 177	Asp 167	Asp 185	Asp 184	Asp 575
Phe 182	Phe 172	Phe 188	Phe 189	Phe 580
Ser 211	Ser 201	Ser 219	Ser 218	Ser 607
Lys 213	Lys 203	Lys 221	Lys 220	Lys 609
Arg 253	Arg 239	Arg 255	Arg 254	Arg 686
His 255	His 241	His 257	His 256	His 688

modified to the structure of the natural substrates using InsightII (Accelrys). PROCHECK confirmed that all the parameters were within normal ranges. Secondary structure predictions indicated that the central part of DHPS insert 2 forms an α-helix. Given the parasite-specific nature of the inserts, this helix is likely to be

involved in domain–domain interactions between the PPPK and DHPS domains as shown for DHFR-TS[20] (see below).

2.2. DHPS ligand-binding site

Since complex formation occurs between sulfonamide and DHP, the simultaneous presence of DHP and magnesium is implied when sulfonamides bind in the active site.[23,24] Therefore, DHP and magnesium were included in the active site during subsequent minimization and MD. The coordinates used were based on the *M. tuberculosis* and *B. anthracis* crystal structures in complex with magnesium and pteroic acid, respectively. To optimize the position of *p*ABA after construction, the complex was subjected to minimization. The *p*ABA part of pteroic acid in the *B. anthracis* crystal structure was shown to associate with Lys 220 (Lys 609 in *P. falciparum*; Table 1). This orientation of *p*ABA was used as a guide to judge the orientation of *p*ABA in *P. falciparum* DHPS during and after minimization and MD. Following minimization of the complete *P. falciparum* DHPS structural model (including DHP, *p*ABA, and magnesium), the proximity of *p*ABA was close enough (<3 Å) to DHP to allow formation of dihydropteroate (Table 1).

2.3. Sulfadoxine resistance in DHPS

Sulfadoxine resistance in DHPS has been attributed to five mutations (Ala 437 Gly, Ser 436 Phe/Ala, Lys 540 Glu, Ala 581 Gly, and Ala 613Thr/Ser[1,7–10]). All these mutations, except for Ala 518, occur in loops. Sulfadoxine is a structural analogue of *p*ABA. To investigate the binding of sulfadoxine, it was partially superimposed onto *p*ABA in the DHPS active site. After minimization of the DHPS–sulfadoxine complex in the presence of water, sulfadoxine showed hydrophobic interactions with various amino acids (Fig. 5). Sulfadoxine may have long-range interactions with the hydroxyl group of Ser 436 in the wild-type enzyme. This was evident during the various simulations when a water molecule was anchored for an extended length of time between sulfadoxine and the hydroxyl group on Ser 436. This linking of sulfadoxine and Ser 436 through a water molecule may assist in stabilizing sulfadoxine in the active site. Thus, when Ser 436 mutates to Ala, this long-range interaction between sulfadoxine and Ser 436 is lost. The molecular dynamics simulation also indicated that the Ser 436 Phe mutation causes sterical clashes between sulfadoxine and Phe 436. The long-range interaction with sulfadoxine is lost when Ser 436 is mutated to Ala. The Ala 437 Gly mutation seems to result in a loss of hydrophobic interaction with sulfadoxine (results not shown). The wild-type simulation showed that the methoxy groups on sulfadoxine undergo hydrophobic interactions with the side chain of Lys 540. The Lys 540 Glu mutation resulted in a loss of this interaction due to a shortening of the hydrophobic side chain as well as an increase in the hydrophilicity of residue 540 (Fig. 5). This increase in hydrophilicity caused the loop on which Glu 540 is located, to open up more than in the wild-type and thus exposed the active site to solvent (results not shown). No rational explanation for the Ala 613Thr/Ser mutation was apparent from the model. Our model shows that Ala 613 and Arg 608 do not interact with sulfadoxine as in the model by Korsinczky. The side chains of Arg 608 and Ala 613 are pointing away from the active site of our model as a result of being located on a relatively flexible part of the protein. This flexibility also increases the distance between sulfadoxine and Arg 608 and Ala 613. However, the orientation of sulfadoxine in the active site is slightly different from that of Korsinczky as the model presented here includes the pterin substrate.

Ala 581 forms part of a hydrophobic cavity in which the methoxy groups of sulfadoxine associate. The Ala 581 Gly mutation also resulted in a loss of hydrophobic interaction between sulfadoxine and *P. falciparum* DHPS. The side chain of Lys 540 also forms part of this cavity. These results partially explain the observations by Triglia et al.[25] as well as those by Berglez et al.[1] regarding the effect each mutation has on sulfadoxine binding. They observed similar trends in increasing resistance with the increase in number of mutations. Table 2 compares the simulation results with the results obtained by Berglez et al.[1]

2.4. Loop movement in DHPS

MD simulations of other TIM barrels[26] and of *P. falciparum* DHPS showed that some loop movements occur (most noticeably loop 2 which contains Ala 437 and Ser 436). Flexibility of loop 2 was strongly suggested in the *M. tuberculosis* crystal structure due to its absence from the crystal structures.[21] These loops are involved in active site closure as shown for other TIM barrels.[9,26,27] The loop movement in *P. falciparum* DHPS indicates a similar mechanism of active site closure.

The DHPS domain model presented here differs in several respects from the model of Korsinczky et al.:[13] (1) these authors did not include any of the insertions with the assumption that this omission is unlikely to affect the active site or predictions made from the resultant model. However, deletion mutagenesis studies of inserts in *P. falciparum* DHFR-TS and AdoMetDC/ODC as well as information obtained by the crystal structure of the former enzyme provided clear evidence for functional roles of parasite-specific inserts in the activities and domain interactions within bifunctional complexes.[15,16,28] Insert 1 was included in the model during MD simulations of DHPS but results showed that its inclusion did not result in any significant influence on the stability of the molecule. However, the effect of insert 2 is still uncertain, since it was excluded from this study due to the unavailability of a template. (2) The model presented here includes the natural substrates (DHP and *p*ABA) as well as the magnesium co-factor. The model also demonstrates the interaction of sulfadoxine with DHP, which has been shown to be present in the active site during sulfadoxine binding.[23,24] (3) Lastly, MD simulations were applied on the proposed structure of *P. falciparum* DHPS after docking procedures, since Perola and Charifson showed that more than 60% of

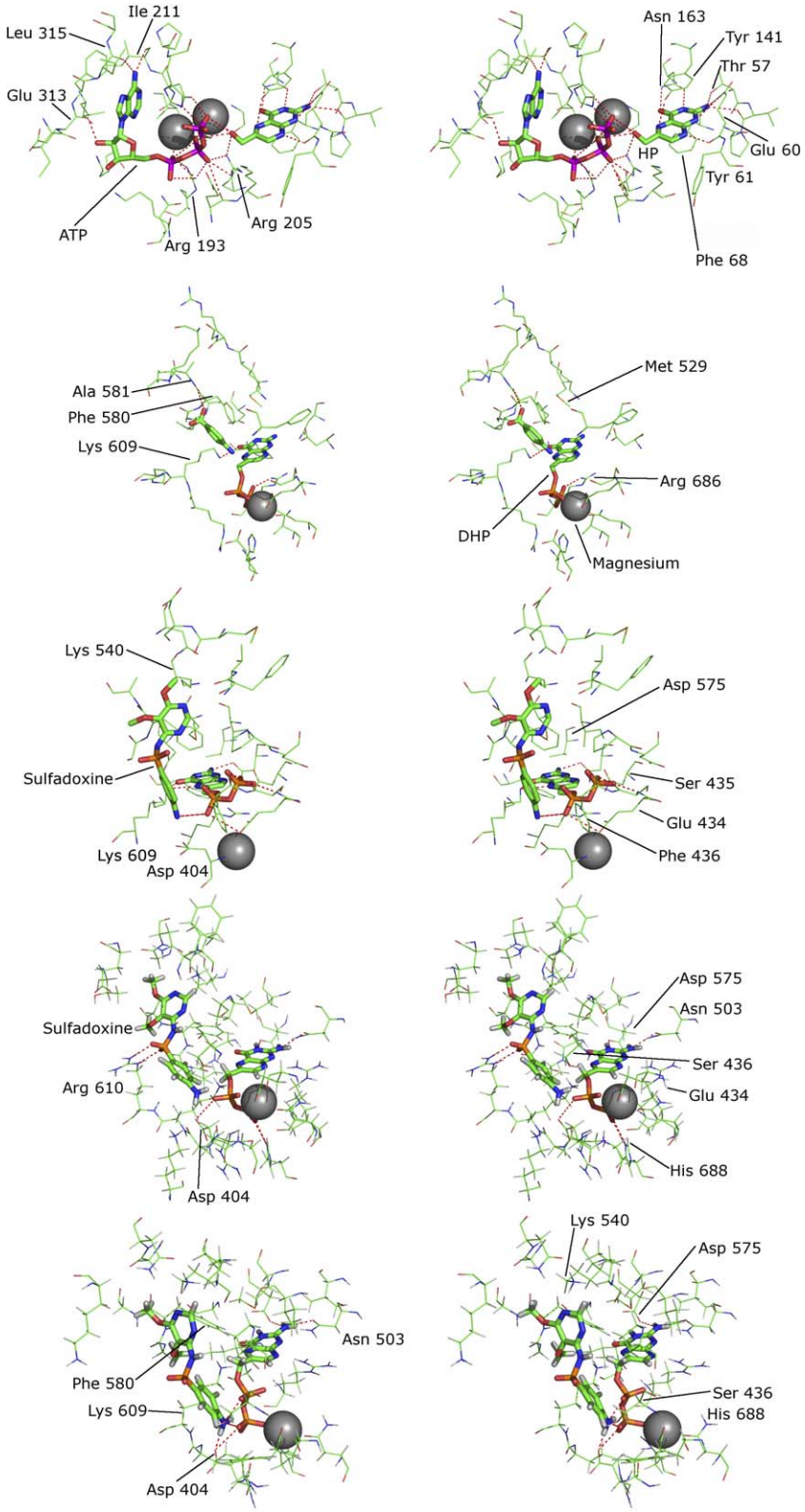

Figure 5. Steric views of some of the active sites of PPPK and DHPS as well as mutations and their effect on substrate binding. From top to bottom: PPPK active site and substrate analogues, DHPS active site and substrates, Ser 436 Phe mutation, Ala 437 Gly mutations, and Lys 540 Glu mutation.

ligands do not bind in their minimum energy conformation.[29] The average RMSD of the protein backbone in protein–sulfadoxine complex during simulation is 1.53 Å, while the average RMSD for the amino acids involved in resistance is 1.14 Å. Using MD accounted for the flexibility of the ligands as well as the protein, thus resulting in a more accurate representation of ligand binding. A more comprehensive understanding is thus provided of the action of sulfadoxine on *P. falciparum* DHPS, and the role of point mutations in sulfadoxine

Table 2. The resistance levels of the various DHPS mutants to sulfadoxine

Mutants	Experimental[1]	Model	Explanation
A437G	Low (3× increase in resistance)	Low	A437G allows more rotational freedom to S436 thus breaking the interaction between S436 and sulfadoxine (interaction lasts less than 5 ps)
A437G + K540E	Medium (6× increase in resistance)	Medium	A437G allows more rotational freedom to S436 thus breaking the interaction between S436 and sulfadoxine (interaction last less than 3 ps). The change in charge K540E from +1 to −1 prevents hydrogen bonding to occur between residue 540 and sulfadoxine
A437G + S436A	Medium[6]	Medium	A437G allows more rotational freedom to S436 thus breaking the interaction between S436 and sulfadoxine. S436A prevents the formation of the interaction (interaction last less than 3 ps) with electron lone pair on the oxygen atom in S436 as it is replaced with a neutral CH_3 group
A437G + A581G	High (11× increase in resistance)	Medium	A437G allows more rotational freedom to S436 thus breaking the interaction between S436 and sulfadoxine (interaction last less than 3 ps). A581 has a long-range stabilising effect on the interaction between K540 and sulfadoxine. By removing this interaction K540 achieves more rotational freedom of movement and the hydrogen bonding between sulfadoxine and K540 is weaker
S436F + A437G + A613S	High (>14× increase in resistance)	High	S436F causes a partial steric blocking of the active site thus preventing the larger sulfadoxine from binding but still allowing *p*ABA to bind. A613S has a long-range effect as there occurs a charge reversal with the mutation
Wild-type	Sensitive	None	The interaction between sulfadoxine and S436 and K540 last more than 10 ps

Experimental resistance levels are adapted from the work by Berglez et al.[1] The resistance levels for the model are derived from the stability of DHPS–sulfadoxine active site interactions: the longer sulfadoxine stays in the active site, the less resistance the mutant has.

resistance. However, the model also supports some aspects of the model published by Korsinczky et al.: (1) both models indicate hydrophobic interactions between sulfadoxine and Ala 437, and (2) both models show that the Ser 436 Phe mutation results in steric clashes between sulfadoxine and *P. falciparum* DHPS.

2.5. PPPK domain structure and ligand-binding site

The PPPK domain shows high sequence conservation between the *Plasmodial* species including insert 1 (Fig. 2). The N-terminal of PPPK insert 2 shows low sequence conservation between the *Plasmodial* species. A further prominent feature of insert 2 is the low complexity of the sequence toward the end of the *P. falciparum*-specific insert. The alignments also suggested that *P. berghei*, *P. chabaudi*, and *P. yoelii yoelii* are more closely related to each other than to *P. falciparum* or *P. vivax*. This suggestion is supported by phylogenetic studies of the *Plasmodium* adenylosuccinate lyase gene.[30]

For modeling of the *P. falciparum* PPPK domain the sequence of *E. coli*[31] was used. The PPPK domain exhibited a ferrodoxin-like fold, which consists of five α-helices and four β-sheets (a central anti-parallel α-sheet flanked on one side by three α-helices and the other side by two α-helices; Figs. 3 and 4) but lacks a helix. The active site of the model compared well with the known crystal structures (Table 3), and retains the ring stacking required to anchor 6-hydroxymethyl-7,8-dihydropterin (HP) in the active site (*E. coli*: Tyr 53 and Phe 123, *P. falciparum*: Phe 68 and Tyr 141). Most of the 23 residues identified by Blaszczyk et al.[31] to be in the PPPK active site are conserved in *P. falciparum* as well as all of the interactions with the substrate (Table 3). The hydrogen bond forming capacity of HP is saturated in *E. coli* and the same residues and interactions are present in the *P. falciparum* model (results not shown).

2.6. Loop movement in PPPK

The loops in *E. coli* PPPK were shown to be very mobile,[31] which supports the molecular dynamics results on the *P. falciparum* PPPK domain model. During the simulation of the PPPK domain model it tended to dissociate, which probably indicates the necessity of the DHPS domain, inserts, and the linker region for stabilization of the enzyme. The substrates (ATP and HP) and the two Mg^{2+} atoms remained in the active site, suggesting that the active site residues anchoring the substrates are positioned correctly. During the first phase of the simulation, the substrates moved to within 3 Å of each other in the orientation needed for catalysis to take place.

2.7. Influence of magnesium in the PPPK and DHPS active sites

Kasekarn et al.[32] showed that *P. falciparum* DHPS activity is dependent on the presence of magnesium. Magnesium was therefore included as a co-factor during modeling of both domains. These models show that magnesium anchors the pyrophosphate group of ATP (PPPK) and DHP (DHPS) thus ensuring that the substrate maintains the correct orientation in the active

Table 3. Active site residue comparison between the PPPK crystal structures of *Escherichia coli*, *Haemophilus influenzae*, and *Plasmodium falciparum* model (Ec: *E. coli*, Hi: *H. influenzae*)

E. coli	*H. influenzae*	*P. falciparum*	Interaction with PPPK
Gly 8	Gly 9	Gly 23	Interacts with HP, highly conserved
Thr 42	Ser 43	Thr 57	Forms hydrogen bonds with primary amine on HP
Pro 43	Lys 44	Val 58	Use peptide backbone O for HP interaction
Leu 45	Leu 46	Glu 60	Glu 60 uses εO to form hydrogen bond with HP, Leu 45
Tyr 53	Tyr 54	Phe 68	Side-chain associates with HP, substrate sandwiched between Tyr (Ec 53, Hi 54) and Phe (Ec 123, Hi 123)
Asn 55	Asn 56	Asn 163	Interacts with the 4-oxo group on ATP
Gln 74	Gln 76	Lys 185	Interacts with ATP through the terminal NH group
Glu 77	Glu 78	Glu 188	Interacts with Mg^{2+} and with ATP
Arg 82	Arg 83	Arg 193	Interacts with the oxygens on the α-phosphate and β-phosphate groups of ATP
Arg 84	Arg 85	Asn 195	Interacts with the oxygens on the α-phosphate group of ATP, helps with loop stabilization
Arg 88	Arg 88	Lys 200	Interacts with ATP through the ε-amino group
Trp 89	Trp 89	Lys 201	Interacts with the substrate and γ-phosphate oxygen of ATP
Arg 92	Arg 92	Arg 205	Helps with loop stabilization and interacts with oxygens on the α-phosphate and β-phosphate groups of ATP
Asn 95	Asn 95	Asp 210	Interacts with HP and Mg^{2+} through oxygens
Asn 97	Asn 97	Asp 210	Interacts with Mg^{2+} through oxygens
Ile 98	Ile 98	Ile 211	Residues interact with ATP though their amide and/or carbonyl groups
Arg 110	Arg 110	Glu 313	Arg 110 uses backbone O for hydrogen bond with ribose part of ATP, Glu uses terminal O
Thr 112	Thr 112	Leu 315	Both residues use peptide backbone O for interaction
Pro 114	Pro 114	Ile 317	Forms cavity for ATP, interacts with ATP core
His 115	His 115	Pro 318	Forms cavity for ATP, interacts with ATP core
Tyr 116	Tyr 116	His 319	Forms cavity for ATP, interacts with ATP core
Arg 121	Arg 121	His 324	Forms cavity for ATP, interacts with ATP core
Phe 123	Phe 123	Tyr 326	Side-chain associates with HP, substrate sandwiched between Tyr 53 and Phe 123

site and hence assisting in the positioning of DHP, *p*ABA or sulfadoxine in DHPS, and ATP and HP in PPPK.

2.8. Bifunctional domain organization

The *P. falciparum* PPPK–DHPS bifunctional model was constructed based on the crystal structure of the bifunctional *S. cerevisiae* PPPK–DHPS. Several short, structured amino acid residue sequences could be identified in the crystal structure of *P. falciparum* DHFR-TS that are involved in interdomain interactions including two α-helices, respectively, localized in insert 1 of DHFR and in the junction region connecting it to the TS domain.[16] *P. falciparum* AdoMetDC/ODC has extra inserted sequences in addition to the hinge region. Deletion of individual inserts in the bifunctional AdoMetDC/ODC was shown to diminish the corresponding enzyme activity, and in some instances also to decrease the activity of the neighboring, non-mutated domain.[15] The *S. cerevisiae* structure shows that species-specific inserts also contribute to interdomain interactions (β-sheet 5′, 5″, 7′ and 7″). The short length of *P. falciparum* DHPS insert 1 (12 residues) and the apparent absence of low-complexity regions usually associated with parasite-specific inserts[16] warranted the inclusion of insert 1 during the modeling procedure.

The *S. cerevisiae* bifunctional structure revealed for the first time the structure of the linker region between PPPK and DHPS domains. Using this structure as a template for the *P. falciparum* model, a putative structure for the linker region between PPPK–DHPS is proposed (Fig. 5). This putative structure assumes the same three-stranded, anti-parallel β-sheet conformation as seen in *S. cerevisiae* PPPK–DHPS. *S. cerevisiae* PPPK–DHPS also has a second area, which forms a three-stranded, anti-parallel β-sheet to assist in the stabilization of the bifunctional complex. This feature is located in the two *S. cerevisiae*-specific inserts in the DHPS domain (one insert forms one sheet and the second the other two sheets). The *P. falciparum* PPPK in the absence of insert 1 in the bifunctional model did not dissociate during MD, as was observed for the domain alone. Therefore, associations between the regions (α1, α3, α4, and β2; Figs. 4 and 6) in the *P. falciparum* PPPK domain and regions on the DHPS domain (α4) and the linker region (β2), which are similar to those shown for the *S. cerevisiae* bifunctional enzyme, are proposed to also stabilize the *P. falciparum* bifunctional complex (Fig. 6). In the case of *S. cerevisiae*, there is an additional interaction between α1 and α3 of PPPK and the yeast-specific DHPS insert. The specific structure of the yeast insert was superimposed in the equivalent position on the DHPS domain of the bifunctional *P. falciparum* PPPK–DHPS. Given its position in the bifunctional enzyme it is apparent that insert 1 (~50 amino acids) of *P. falciparum* PPPK (not modeled) could further stabilize the *P. falciparum* bifunctional enzyme.

3. Conclusion

The bifunctional model presented here provides the first glimpse of the structure of *P. falciparum* PPPK–DHPS. Resistance against sulfonamides is on the increase and the homology model of the *P. falciparum* DHPS domain provides some clarification of the mechanism behind sulfadoxine resistance. The model suggests that sulfadoxine forms hydrophobic interactions with Ala 437,

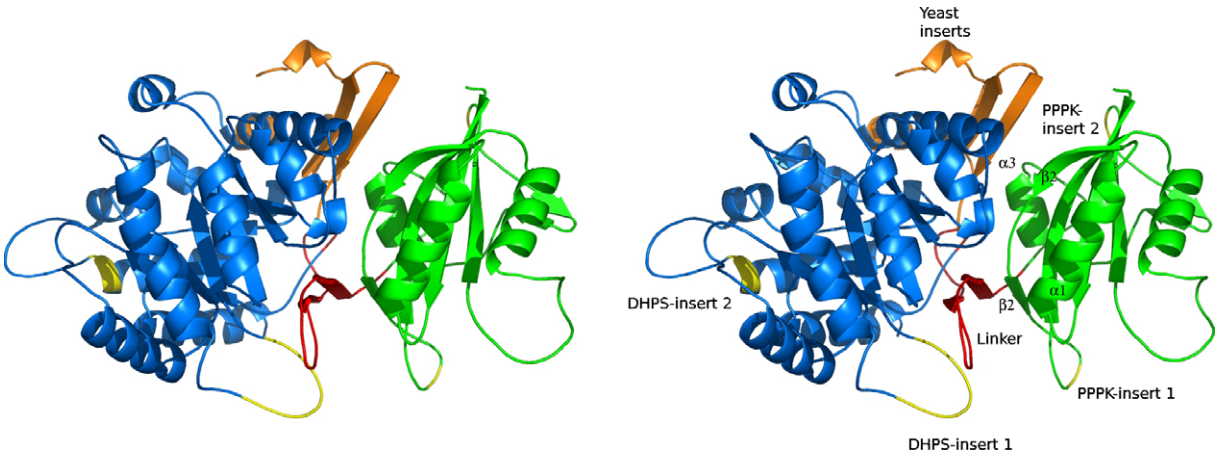

Figure 6. A steric view of the bifunctional *Plasmodium falciparum* PPPK–DHPS model. DHPS is colored blue, PPPK is colored green, parasite-specific inserts and locations are indicated in yellow, yeast specific inserts are colored orange, and the linker region is colored red. The substrates are shown in the active site with the metal ion colored gray. The yeast-specific insert was superimposed on the *P. falciparum* model to indicate its relative position.

the Lys 540 side chain, and Ala 581 but polar interactions with Ser 436. The Ala 437 Gly, Lys 540 Glu and Ala 581 Gly mutations result in a lower affinity of sulfadoxine for the active site and thus less efficiency as an inhibitor. The normal functioning of the protein does not appear to be affected by these mutations since *p*ABA associates mostly with Lys 609 (Lys 220 in *B. anthracis* and Lys 807 in *S. cerevisiae*) in *P. falciparum* PPPK–DHPS. No clarity is provided by the model regarding the role of the Ala 613 Ser mutation in resistance, which has been shown to occur less often.[1–3,7–10] The range of highly specific interactions demonstrated in the active site of *P. falciparum* DHPS model opens up avenues to explore in the design of novel lead compounds. The structure of the PPPK domain emphasizes the similarity between the substrate backbone structure of the PPPK and DHPS domains, which can be exploited by rational drug design strategies against both domains.

4. Methods

4.1. Model building

The bifunctional *P. falciparum* PPPK–DHPS sequence (TREMBL Accession No.: Q27865) was aligned with the sequences of *S. cerevisiae* (PPPK–DHPS), *H. influenzae* (Swissprot id: P43777), *E. coli* (PPPK[31] and DHPS,[33]) *B. anthracis* (DHPS),[22] and *M. tuberculosis* (DHPS).[21] The translated sequences of *P. chabaudi* PPPK–DHPS (TREMBL Accession No.: Q9BLN5), *P. berghei* PPPK–DHPS (TREMBL Accession No.: Q9BIX9), *P. vivax* PPPK–DHPS (Genbank Accession No.: AY186730) and *Plasmodium yoelii yoelii* PPPK–DHPS (Genbank acc. no: EAA21661) were included to assist with the delineation of parasite-specific inserted sequences. The alignments were prepared using ClustalX[34] followed by manual refinement. For the bifunctional model, the *S. cerevisiae* crystal structure (2BMB) was used. The *M. tuberculosis* crystal structure (1EYE) as well as the *B. anthracis* crystal structure (1TX0) were selected as templates for the modeling of the DHPS domain based on the higher identity between *M. tuberculosis* (31%) and *P. falciparum*, and the availability of a product analogue in the *B. anthracis*[21,22] structure. Residue 383 in the alignment served as the start of the sequence modeled due to an absence of template. The modeling of the PPPK domain was based on the crystal structure of *E. coli* (1DY3). The modeling started at residue 12 as the first residues had no available template.

P. falciparum DHPS insert 1 was retained during modeling due to its relatively short length. The alignment of *P. falciparum* with *P. chabaudi*, *P. berghei*, *P. vivax*, and *P. yoelii yoelii* revealed a conserved motif at the end of insert 2, which was predicted with HMMSTR[35] to be an α-helix. In the case of *P. yoelii yoelii*, this helix is interrupted with an 18-residue sequence (shown as an insert in DHPS insert 2) of undefined structure. No low-complexity regions were evident in *P. falciparum* DHPS using the SEG algorithm,[36] but PPPK insert 2 was identified as a low-complexity region.

MODELLER 6v3[37–39] was used for homology modeling. The *Plasmodium*-specific inserts were omitted (with the exception of DHPS insert 1) due to a lack of template structure. None of the ligands were included in the construction of the bifunctional model but were included in the construction of the separate domains. In order to investigate the properties of each of the domains in more detail, separate models based on the same alignment were constructed. The homology model of the wild-type DHPS domain was constructed with the inclusion of pteroic acid (product analogue from *B. anthracis* crystal structure) and Mg^{2+} ion (based on the *M. tuberculosis* crystal structure) during the modeling process. The metal ion was included to retain conserved interactions between the active site amino acids, metal ion, and the substrate.[21] A total of 15 models were generated and since the models were all super-imposable with differences only in the loop sections, the model with the best MODELLER objective score was used for further refinement. PROCHECK was used to verify that all the protein parameters such as bond lengths and bond

angles were within acceptable limits.[40] Following model building the product analogue (pteroic acid) was converted to DHP and *p*-aminobenzoic acid using the Builder module in InsightII (Accelrys), followed by optimization in Discover 3 (Accelrys; 500 steps steepest descent minimization, ESFF[41] forcefield). The ESFF forcefield was selected as it is parameterized for magnesium atoms. This reduced the need to parameterize the forcefield for this specific atom. Certain license restrictions prevented the use of other forcefields, which also contained parameters for magnesium.

In the PPPK model, the substrates in the active site (two Mg^{2+} ions, ATP, and 6-hydroxymethyl-7,8-dihydropterin (HP)) were included in the modeling process. Fifteen models were generated and minor differences only occurred in the loop sections. The model characteristics were evaluated using PROCHECK. HP was converted to the natural substrate using the Builder module in InsightII (Accelrys), followed by optimization in Discover 3 (Accelrys; 500 steps steepest descent minimization, ESFF forcefield).

4.2. Molecular dynamics (MD)

The bifunctional PPPK–DHPS model was subjected to molecular dynamics using Yasara[42] to investigate the stability of the two domains. The simulation was run in water at a pH 7.0 for 2.985 ns at a temperature of 323 K. Yasara was chosen due to certain technical limitations, which prevented the simulation of the full bifunctional protein being run on the local cluster using CHARMM 29b2.[43]

The homology models of each individual domain were subjected to molecular dynamics using CHARMM 29b2 to investigate loop movement in the model, as well as to relax energy clashes. The model was energy-minimized for 2500 steps and then solvated using TIP3 water molecules. The model was then subjected to 500 steps of steepest descent minimization followed by 200 steps of adopted basis Newton–Raphson minimization. It was then heated to 300 K in steps of 3 K over a time period of 10 picoseconds (ps), left to equilibrate for 3 ps, and MD calculations were subsequently performed for 1 nanoseconds (ns). Subsequently, for the DHPS domain, sulfadoxine was partially superimposed on *p*ABA in the minimized *P. falciparum* DHPS structure as an initial position, and then underwent 500 steps of steepest descent minimization, followed by solvation of the protein and then 100 ps of MD. Hydrogen bonds were calculated using InsightII (Accelrys). During the MD simulations both the ligand and the protein were flexible, thus allowing for a much more accurate representation of the interaction between protein, substrate, sulfadoxine, and water. Models of the known resistance-causing mutations in *P. falciparum* DHPS domain were also made by mutating the desired residues and minimizing the resultant model for 200 steps. The mutation and minimization were done using the builder module of InsightII (Accelrys). The mutated models were subjected to the same methods as the wild-type models to investigate the effect of the mutations on sulfadoxine binding.

Acknowledgments

This manuscript is based on the work supported by the National Research Foundation (NRF) of South Africa, the University of Pretoria, and the National Bioinformatics Network. We also thank Prof. D. Litthauer for assistance with the Yasara simulations, G Wells for critical comments, and Prof. M. Lawrence for providing the crystal structure coordinates of *S. cerevisiae* PPPK–DHPS (2BMB) prior to public release. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and therefore, the NRF does not accept any liability in regard thereto.

References and notes

1. Berglez, J.; Iliades, P.; Sirawaraporn, W.; Coloe, P.; Macreadie, I. *Int. J. Parasitol.* **2004**, *34*, 95.
2. Bwijo, B.; Kaneko, A.; Takechi, A.; Zungu, I. L.; Moriyama, Y.; Lum, J. K.; Tshukahara, T.; Mita, T.; Takahashi, N.; Bergqvist, Y.; Björkman, A.; Kobayakawa, T. *Acta Trop.* **2003**, *85*, 363.
3. Eriksen, J.; Mwankusye, S.; Mduma, S.; Kitua, A.; Swedberg, G.; Tomson, G.; Gustaffson, L. L.; Warsame, M. *Trans. R. Soc. Trop. Med. Hyg.* **2004**, *98*, 347.
4. Sibley, C. H.; Hyde, J. E.; Sims, P. F.; Plowe, C. V.; Kublin, J. G.; Mberu, E. K.; Cowman, A. F.; Winstanley, P. A.; Watkins, W. M.; Nzila, A. M. *Trends Parasitol.* **2001**, *17*, 582.
5. Yuvaniyama, J.; Chitnumsub, P.; Kamchonwongpaisan, S.; Vanichtanankul, J.; Sirawaraporn, W.; Taylor, P.; Walkinshaw, M. D.; Yuthavong, Y. *Nat. Struct. Biol.* **2003**, *5*, 345.
6. Wattanarangsan, J.; Chusacultanachai, S.; Yuvaniyama, J.; Kamchonwongpaisan, S.; Yuthavong, Y. *Mol. Biochem. Parasitol.* **2003**, *126*, 97.
7. Triglia, T.; Menting, J. G. T.; Wilson, C.; Cowman, F. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 13944.
8. Triglia, T.; Menting, J. G. T.; Wilson, C.; Cowman, F. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7149.
9. Nzila, A. M.; Mberu, E. K.; Sulo, J.; Dayo, H.; Winstanley, P. A.; Sibley, C. H.; Watkins, W. M. *Antimicrob. Agents Chemother.* **2000**, *44*, 991.
10. Basco, L. K.; Tahar, R.; Ringwald, P. *Antimicrob. Agents Chemother.* **1998**, *42*, 1811.
11. Talisuna, A. O.; Nalunkuma-Kazibwe, A.; Langi, P.; Mutabingwa, T. K.; Watkins, W. W.; Marck, E. V.; Egwang, T. G.; D'Alessandro, U. *Infect. Genet. Evol.* **2004**, *4*, 321.
12. Rastelli, G.; Sirawaraporn, W.; Sompornpisut, P.; Vilaivan, T.; Kamchonwongpaisan, S.; Quarrell, R.; Lowe, G.; Thebtaranonth, Y.; Yuthavong, Y. *Bioorg. Med. Chem.* **2000**, *8*, 1117.
13. Korsinczky, M.; Fishcer, K.; Chen, N.; Baker, J.; Rieckmann, K.; Cheng, Q. *Antimicrob. Agents Chemother.* **2004**, *48*, 2214.
14. Pizzi, E.; Frontali, C. *Genome. Res.* **2001**, *11*, 218.
15. Birkholtz, L. M.; Wrenger, C.; Joubert, F.; Wells, G. A.; Walter, R. D.; Louw, A. I. *Biochem. J.* **2004**, *377*, 439.
16. Yuvaniyama, J.; Chitnumsub, P.; Kamchonwongpaisan, S.; Vanichtanankul, J.; Sirawaraporn, W.; Taylor, P.; Walkinshaw, M. D.; Yuthavong, Y. *Nat. Struct. Biol.* **2003**, *10*, 357.
17. Lawrence, M. C.; Iliades, P.; Fernley, R. T.; Berglez, J.; Pilling, P. A.; Macreadie, I. G. *J. Mol. Biol.* **2005**, *348*, 3652.

18. Toyoda, T.; Brobey, R. K. B.; Sano, G.-I.; Horii, T. *Biochem. Biophys. Res. Comm.* **1997**, *235*, 515.
19. McKie, J. H.; Douglas, K. T.; Chan, C.; Roser, S. A.; Yates, R.; Read, M.; Hyde, J. E.; Dascombe, M. J.; Yuthavong, Y.; Sirawaraporn, W. J. *J. Med. Chem.* **1998**, *41*, 1367.
20. Warhurst, D. C. *Drug Discovery Today* **1998**, *3*, 538.
21. Baca, A. M.; Sirawaraporn, R.; Turley, S.; Sirawaraporn, W.; Hol, W. G. J. *J. Mol. Biol.* **2000**, *302*, 1193.
22. Babaoglu, K.; Qi, J.; Lee, R. E.; White, S. W. *Structure* **2004**, *12*, 1705.
23. Swedberg, G.; Castensson, S.; Skold, O. *J. Bacteriol.* **1979**, *137*, 129.
24. Roland, S.; Ferone, R.; Harvey, R. J.; Styles, V;L.; Morrison, R. W. *J. Biol. Chem.* **1979**, *254*, 10337.
25. Triglia, T.; Wang, P.; Sims, P. F. G.; Hyde, J. E.; Cowman, A. F. *EMBO J.* **1998**, *17*, 3807.
26. Derreumaux, P.; Schlick, T. *Biophys. J.* **1998**, *74*, 72.
27. Rozovsky, S.; Jogl, G.; Tong, L.; McDermott, A. E. *J. Mol. Biol.* **2001**, *310*, 271.
28. Shallom, S.; Zhang, K.; Jiang, L.; Rathod, P. K. *J. Biol. Chem.* **1999**, *273*, 37781.
29. Perola, E.; Charifson, P. S. *J. Med. Chem.* **2004**, *47*, 2499.
30. Kedzierski, L.; Escalante, A. A.; Isea, R.; Black, C. G.; Barnwell, J. W.; Coppel, R. L. *Infect. Genet. Evol.* **2002**, *1*, 297.
31. Blaszczyk, J.; Shi, G.; Yan, H.; Ji, X. *Structure Fold. Des.* **2000**, *8*, 1049.
32. Kasekarn, W.; Sirawaraporn, R.; Chahomchuen, T.; Cowman, A. F.; Sirawaraporn, W. *Mol. Biochem. Parasitol.* **2004**, *137*, 43.
33. Achari, A.; Somers, D. O.; Champness, J. N.; Bryant, P. K.; Rosemond, J.; Stammers, D. K. *Nat. Struct. Biol.* **1997**, *490*, 4.
34. Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucleic Acids Res.* **1994**, *22*, 4673.
35. Bystroff, C.; Thorsson, V.; Baker, D. *J. Mol. Biol.* **2000**, *301*, 173.
36. Wootton, J. C.; Federhen, S. *J. Comput. Chem.* **1993**, *17*, 149.
37. Marti-Renom, M. A.; Stuart, A.; Fiser, A.; Sánchez, R.; Melo, F.; Sali, A. *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291.
38. Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779.
39. Fiser, A.; Do, R. K.; Sali, A. *Protein Sci.* **2000**, 1753.
40. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283.
41. Shi, S.; Yan, L.; Yang, Y.; Fisher-Shaulsky, J.; Thacher, T. *J. Comput. Chem.* **2003**, *24*, 2069.
42. Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.
43. Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; Karplus, M. *J. Comput. Chem.* **1983**, *4*, 187.