Syst Synth Biol (2015) 9:S27–S37
DOI 10.1007/s11693-015-9168-8

RESEARCH ARTICLE

Structural insights into a key carotenogenesis related enzyme phytoene synthase of *P. falciparum*: a novel drug target for malaria

Shalini Agarwal[1,4] · Vijeta Sharma[2] · Swastik Phulera[3] · M. Z. Abdin[4] · R. Ayana[2] · Shailja Singh[2]

Received: 3 May 2014 / Accepted: 10 March 2015 / Published online: 9 April 2015

Abstract Carotenoids represent a diverse group of pigments derived from the common isoprenoid precursors and fulfill a variety of critical functions in plants and animals. Phytoene synthase (PSY), a transferase enzyme that catalyzes the first specific step in carotenoid biosynthesis plays a central role in the regulation of a number of essential functions mediated via carotenoids. PSYs have been deeply investigated in plants, bacteria and algae however in apicomplexans it is poorly studied. In an effort to characterize PSY in apicomplexans especially the malaria parasite *Plasmodium falciparum* (*P. falciparum*), a detailed bioinformatics analysis is undertaken. We have analysed the Phylogenetic relationship of PSY also referred to as octaprenyl pyrophosphate synthase (OPPS) in *P. falciparum* with other taxonomic groups. Further, we in silico characterized the secondary and tertiary structures of *P. falciparum* PSY/OPPS and compared the tertiary structures with crystal structure of *Thermotoga maritima* (*T. maritima*) OPPS. Our results evidenced the resemblance of *P. falciparum* PSY with the active site of *T. maritima* OPPS. Interestingly, the comparative structural analysis revealed an unconserved unique loop in *P. falciparum* OPPS/PSY. Such structural insights might contribute novel accessory functions to the protein thus, offering potential drug targets.

Keywords Apicomplexa · Carotenoid biosynthesis · Phytoene · Isoprenoid · *Plasmodium falciparum* · OPPS · PSY

Introduction

Carotenoids are a wide family of tetraterpenoid pigments synthesized by all plants, fungi, bacteria, and the protozoan parasites of phylum apicomplexa (Goodwin 1980; Lu et al. 2008; Tonhosolo et al. 2009) and fulfill a variety of critical functions in plants. In addition the important health benefits of these carotenoids to animals and humans make the research on carotenoid metabolism exceptionally important. All carotenoids possess a polyisoprenoid structure and are synthesized from the five carbon units' isopentenyl diphosphate (IPP) and its double-bond isomer dimethylallyl diphosphate (DMAP). Unlike higher eukaryotes and many bacteria, plants and apicomplexan protozoa such as malaria parasites produce their isoprenoids (IPP and DMAP) using the non-mevalonate pathway, which takes place in their plastids (Lichtenthaler et al. 1997; Clastre et al. 2007). IPP and DMAP are common precursors for production of geranylgeranyl pyrophosphate (GGPP) required for carotenoid biosynthesis in all organisms. The condensation of two molecules of GGPP to form phytoene, the initial C_{40} carotenoid skeleton represents the first commited step in carotenoid biosynthesis pathway (Cunningham and Gantt 1998). This first specific step in

Shalini Agarwal and Vijeta Sharma have contributed equally to the work.

✉ Shailja Singh
shailja.singh@snu.edu.in

1 International Centre for Genetic Engineering and Biotechnology, New Delhi 110067, India

2 Shiv Nadar University, Gautam Buddha Nagar 201314, UP, India

3 National Centre for Cell Science, NCCS Complex, Pune 411007, India

4 Department of Bitechnology, Jamia Hamdard, Hamdard Nagar, New Delhi 110062, India

Table 1 Table showing the protein ID of enzymes for studying the homology and phylogentic relationships among different taxa

Organism	Entry ID	Protein
Plasmodium falciparum 3D7	XP_001349541.1	Octaprenyl pyrophosphate synthase/phytoene synthase
Babesia bovis T2Bo	XP 001610655.1	Polyprenyl synthetase superfamily protein
Toxoplasma gondii GT1	EPR64929.1	Polyprenyl synthetase superfamily protein
Arabidopsis thaliana	AED92400.1	Phytoene synthase
Pantoea ananatis AJ13355	YP 005941087.1	Phytoene synthase crtb
Rhodobacter capsulatus	WP 023921666.1	Putative phytoene synthase
Staphylococcus aureus subsp. aureus DSM20231	ELP27047.1	Phytoene synthase
Neurospora crassa	AAA19428.1	Phytoene synthase
Synechococcus elongatus PCC 7942	CAA45350.1	Phytoene synthase
Thermus thermophilus	WP_014630564.1	Phytoene synthase
Cucumis melo	CAA85775.1	Phytoene synthase
Dunaliella salina	AAB51287.1	Phytoene synthase
Thermotoga Maritima	pdb|1V4E	Chain a, crystal structure of octaprenyl pyrophosphate synthase
Chlamydia pneumoniae B21	ETR79494.1	Octaprenyl diphosphate synthase
Brassica napus	AHE79749.1	Phytoene synthase
Rubrivivax gelatinosus	WP_014429000.1	Phytoene synthase

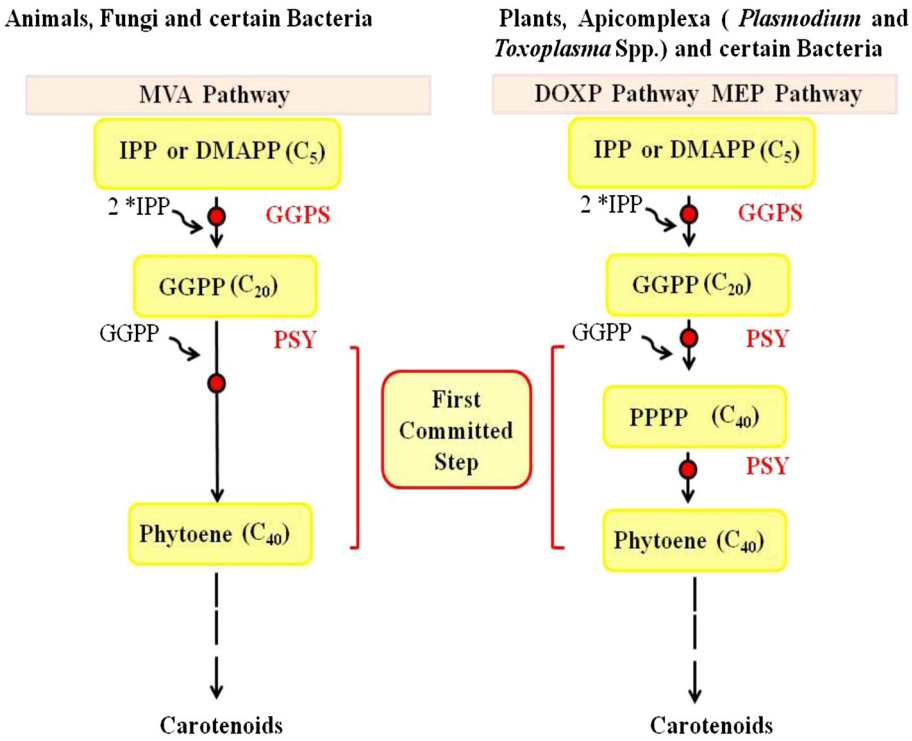


Fig. 1 Schematic of isoprenoid to carotenoid biosynthesis pathway in various organisms. The isoprenoid precursors are synthesized via the 2-Cmethyl-D-erythritol 4-phosphate (MEP) pathway in plastids of higher plants, apicomplexa and many proteobacteria and via mevalonic acid (MVA) pathway in animals, fungi and certain bacteria. Though different at many steps, the enzyme Phytoene synthase (PSY), catalyzing the first commited step of isoprenoid to carotenoid biosynthesis pathway is common in most organisms. PSY acts at the condensation of two molecules of GGPP to form the phytoene, the first commited step of the carotenoid pathway. Names of compounds and pathways are highlighted in black and of enzymes in red. *Abbreviations: IPP* isopentenyl diphosphate, *DMAP* dimethylallyl diphosphate, *GGPP* geranylgeranyl-pyrophosphate, *GGPS* geranyl-geranyl diphosphate synthase, *PPPP* prephytoene diphosphate, *PSY* phytoene synthase. (Color figure online)

carotenogenesis is catalyzed by a transferase enzyme, Phytoene synthase (PSY) (Cunningham and Gantt 1998). PSY, a rate limiting enzyme thus, occupies the driving control between multiple pathways and irreversibly fluxes GGPP away from isoprenoid production towards carotenoid synthesis. PSY hence, plays a central role in the regulation of a number of essential functions mediated via carotenoids.

Significant progress has been made towards the understanding of isoprenoid and carotenoid biosynthesis and catabolism in plants, algae and their endosymbiont apicomplexans. Apicomplexans have a conserved DOXP-MEP pathway for synthesis of isoprenoids, derived from their plastids (Ralph et al. 2004; Jordão et al. 2011) Additionally, biochemical findings demonstrates the presence of carotenoid biosynthesis in the intraerythrocytic stages of the *P. falciparum*, a causative agent of human malaria (Tonhosolo et al. 2005, 2009). Given that, isoprenoids and carotenoids are synthesized in parasites, it is possible that the enzymes driving their synthesis such as PSY are also retained in the parasite metabolism. In apicomplexans PSY is poorly investigated in contrast to studies in higher plants and algae (Salvini et al. 2005; Yan et al. 2005). However, recent studies in *P. falciparum* show PSY as a bifunctional enzyme (PfB0130w), exhibiting octaprenyl pyrophosphate synthase (OPPS) as well as PSY activity, analogous to some bacterial OPPS enzymes (Tonhosolo et al. 2005, 2009). *Toxoplasma gondii* (*T. gondii*), a related apicomplexan parasite, is also found to have weak orthologues of PSY based on searches using genes from higher plants. Bioinformatic analysis and structural comparison of this important enzyme, PSY in diverse group of organisms such as bacteria, higher plants, and apicomplexa could thus, be helpful to provide further insights into its role in carotenoid biosynthesis.

In this study, we have analysed the phylogenetic relationship of *P. falciparum* PSY/OPPS with other taxonomic groups. Further, we have in silico characterized the secondary and tertiary structures and found the conserved motifs which might influence the properties of this enzyme. The structural regions of the modeled *P. falciparum* PSY/OPPS were compared with the crystal structure of a hyper thermophilic bacterium, *Thermotoga maritima (T. maritima*) OPPS with which it is evolutionary related as depicted in this study. Interestingly, our results evidenced the resemblance of *P. falciparum* OPPS/PSY active site with the active site of *T. maritima* OPPS. Also, after comparative structural analysis, we found an unconserved unique loop in *P. falciparum* OPPS/PSY that might contribute novel accessory functions to the enzyme related to carotenoid metabolism in parasite. Therefore, from the perspective of drug designing this study offers a very promising pharmacological target.

Materials and methods

Protein retrieval and sequence analysis

The protein sequences were retrieved from protein database of NCBI (http://www.ncbi.nlm.nih.gov/protein/). The primary accession numbers of enzymes of members of bacteria, fungi, plants and apicomplexans are listed in Table 1. The protein sequence of *P. falciparum* OPPS/PSY (XP_001349541.1) was used as query for PSIBLAST *(blast.ncbi.nlm.nih.gov)* analysis for identification of homologs. The best blast hits included polyprenyl pyrophosphate synthetase superfamily protein (PPS) of *T. gondii* and OPPS of *T. maritima.*

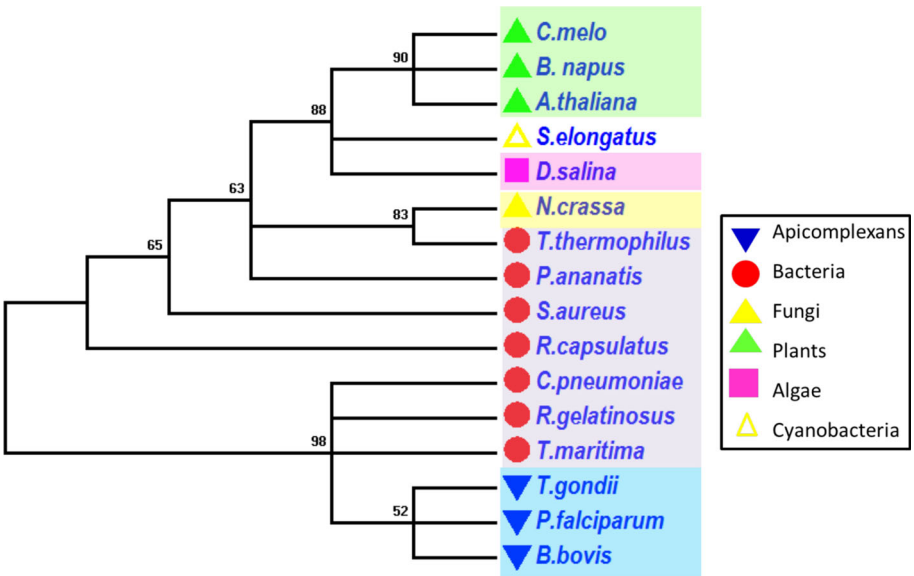


Fig. 2 Phylogenetic relationship of the *P. falciaprum* OPPS/PSY with its counterpart in other various species. The sequences of 16 proteins from different taxonomic groups were compared using CLUSTAL W and an unrooted phylogenetic tree was generated by Neighbor-Joining method using MEGA6. The proteins used for phylogenetic tree construction are the amino acid sequences used for multiple sequence alignment as mentioned in the Table 1. The evolutionary distance is mention by the numbers along the branches

Fig. 3 Prediction of secondary and tertiary structures. **a** Secondary structures displayed under the amino acid sequence of *P. falciparum* OPPS/PSY (*left*) and *T. gondii* PPS (*right*) was predicted using PSIPRED V2.6 (http://bioinf.cs.ucl.ac.uk/psipred/). The *cylinders* represent alpha-helices, *empty arrows* indicate beta-sheets, and *heavy lines* are coiled regions as indicated in the legend. **b** Ribbon representations of models of *P. falciparum* OPPS/PSY (*left*); and *T. gondii* PPS (*right*). The models were generated by SWISS-MODEL. **c** Ramachandran *plot* showing the acceptability of models. Ramachandran *plots* showing the acceptability of models of *P. falciparum* (*left*), and *T. gondii* (*right*). The *plots* were generated by PROCHECK for assessing the stereochemical quality of the protein structure models

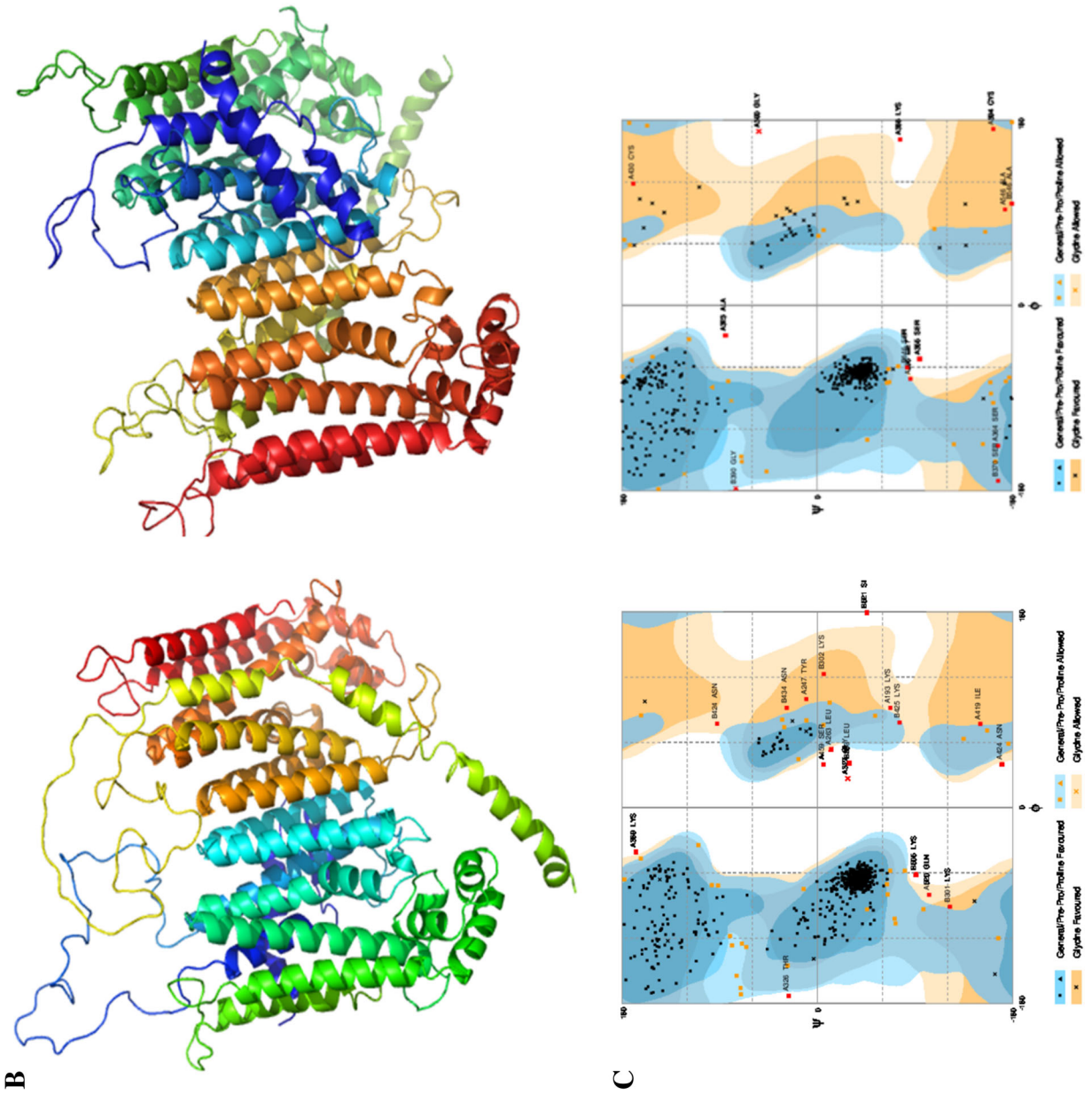


Fig. 3 continued

Phylogenetic analysis

Phylogenetic tree was generated by maximum likelihood method (MLM). Multiple sequence alignment of the amino acid sequences was carried out by using Clustal W (http://www.ebi.ac.uk/Tools/msa/clustalw2/) to construct phylogenetic tree (Larkin et al. 2007). The evolutionary history was inferred using the Neighbor-Joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 14.67214052 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the JTT matrix-based method (Jones et al. 1992) and are in the units of the number of amino acid substitutions per site. The analysis involved total of 16 protein sequences including *P. falciparum* OPPS/PSY. All positions containing gaps and missing data were eliminated. There were a total of 143 positions in the final dataset. Evolutionary analyses were conducted in MEGA6 (Tamura et al. 2013).

Structure prediction and homology modeling

Putative secondary structures were predicted for *P. falciparum* OPPS/PSY and *T. gondii* PPS using PSIPRED V2.6 software (http://bioinf.cs.ucl.ac.uk/psipred/). The 3D models of *P. falciparum* OPPS/PSY, *T. gondii* PPS were generated using SWISS MODEL (Arnold et al. 2006) with parameters of energy minimization values. For visualization and editing of the generated PDB models the PyMOL Molecular Graphics System, Version 1.5.0.4 Schrödinger, LLC (www.pymol.org,) was used.

Model quality and reliability validation

The modeled structures were validated by PROCHECK in SAVES server (http://nihserver.mbi.ucla.edu/SAVES/) and best model was selected. The selected best model was further checked with verify3D (Bowie et al. 1991; Lüthy et al. 1992; Benkert et al. 2011) and Ramachandran plot at PROCHECK (Laskowski et al. 1993; Morris 1992). Ramachandran plot constructed by the PROCHECK explained the regions of amino acids in the plot and the overall quality of the model based on the phi-psi angles.

Motifs prediction

For motifs prediction, the amino acid sequence of *P. faciaprum* OPPS/PSY was subjected to blast search in swissprot database, limiting the results to top 250 matches. The results of the blast search were aligned using MUSCLE (Edgar and Robert 2004) and the alignment was improved empirically by removing sequences with major gaps or sequences not aligning well. A pattern/motif search was then performed using the programme PRATT (http://www.ebi.ac.uk/pratt/) with the initial parameters: C % (minimum percentage of sequences to match) = 50; PL (maximum pattern length) = 50; PN (maximum number of pattern symbols) = 50; PX (maximum number of consecutive Xs) = 5; FN (maximum number of flexible spacers) = 2; FL (maximum flexibility) = 2; FP (maximum flexibility product) = 10; E (search greediness) = 3 (McWilliam et al. 2013).

Fig. 4 Comparison of amino acid sequences of the *T. maritima* OPPS (299aa), *P. falciparum* OPPS/PSY (538aa) and *T. gondii* (676aa) PPS. The sequences are compared by multiple sequence alignment done by using CLUSTAL W. *Colored outlines* indicate identical and similar amino acid residues, respectively. The *encircled* parts in the shown *P. falciaprum* model (*blue*) and *T. gondii* model (*magenta*) correspond to the gaps as shown in the alignment contributing to the loops which is absent in the 1V4E model (*green*). (Color figure online)

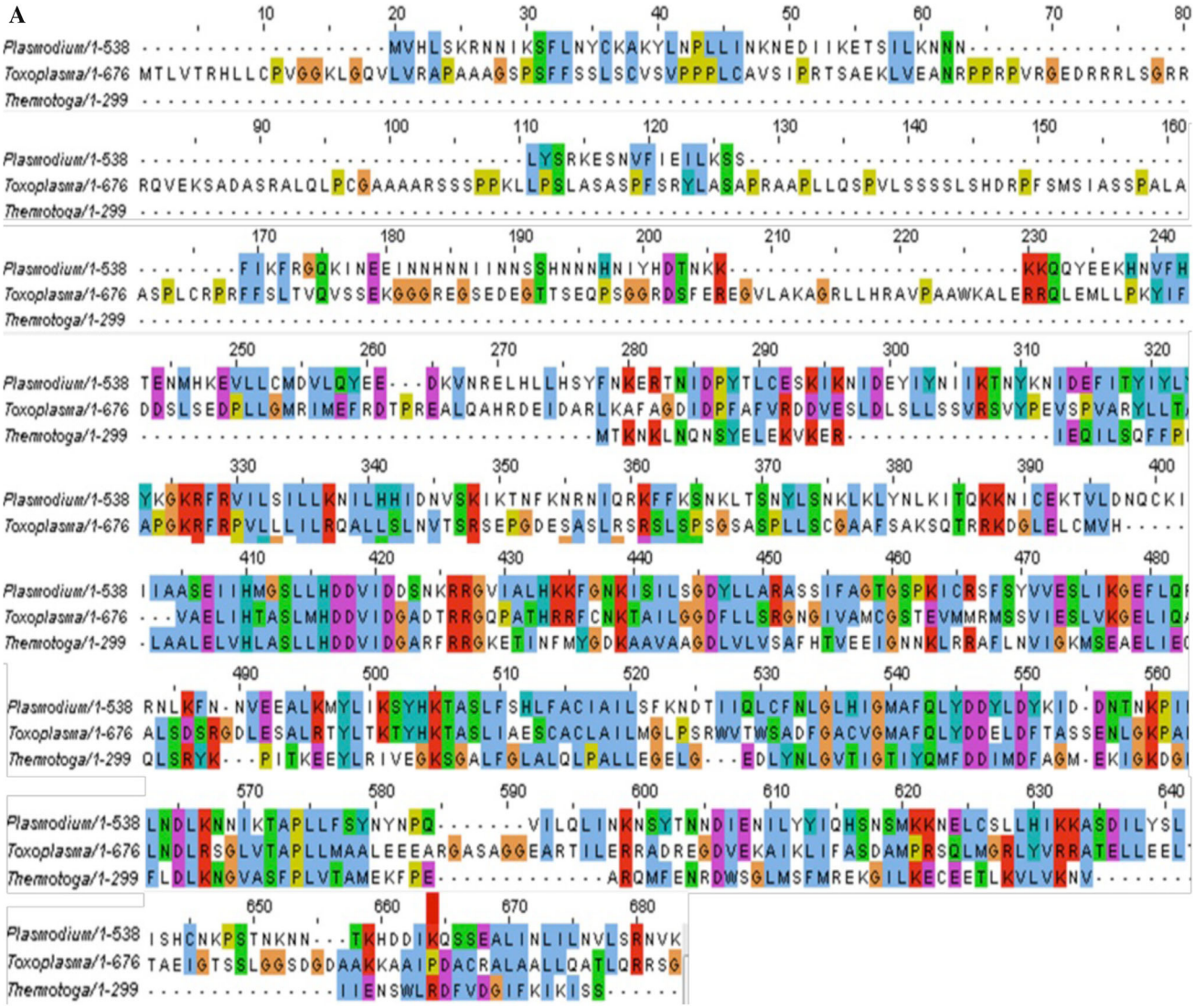
Structural alignment

To compare the structural regions of the modeled protein of *P. falciaprum* OPPS/PSY with the crystal structure (PDB ID 1V4E) of *T. maritima* OPPS, the superposition of the models was assessed. Superposition was done using Superpose program from the CCP4 suite (http://wishart.biology.ualberta.ca/SuperPose/) (Maiti et al. 2004).

Results and discussions

Several studies have investigated isoprenoid and carotenoid biosynthesis and catabolism in plants, algae and their endosymbiont apicomplexans. Apicomplexans have diverse pathways for synthesis of isoprenoids and carotenoids. Though different at many steps, the enzyme PSY, catalyzing the first commited step of isoprenoid to carotenoid biosynthesis pathway is also present in apicomplexans parasites including *P. falciparum*. Schematic for isoprenoid to carotenoid biosynthesis pathway in various organisms, shows that isoprenoid precursors which are further catalysed by PSY, are synthesized via the 2-Cmethyl-D-erythritol 4-phosphate (MEP) pathway in plastids of higher plants, apicomplexa and some bacteria whereas via mevalonic acid (MVA) pathway in animals, fungi and certain bacteria (Fig. 1). As depicted in figure PSY acts at the condensation of two molecules of GGPP to form the phytoene (C_{40} carotenoid), the first commited step of the carotenoid pathway (Fig. 1).

In an attempt to in silico characterize PSY/OPPS in apicomplexans esp. the malaria parasite *P. falciparum*; a detailed bioinformatics analysis of OPPS is undertaken. To identify the


A
Plasmodium/1-538
Toxoplasma/1-676
Thermotoga/1-299
B
Thermotoga maritima
Toxoplasma gondii
Plasmodium falciparum

OPPS protein (accession no. XP_001349541.1), we searched *P. falciparum* OPPS/PSY as query sequence in BLAST. We selected 16 protein members of various taxonomic groups including bacteria, plants, fungi and algae. The selection was based on blast hits and search in protein database (NCBI) of organisms in which the carotenoid biosynthesis pathway is most studied. A complete list of selected sixteen proteins from different taxonomic groups is summarized in Table 1.

The relationship between the proteins of apicomplexans with other taxonomic groups was investigated in detail. We generated multiple sequence alignment by CLUSTAL W2 of selected protein sequences and further generated phylogenetic tree by using MEGA 6 as discussed in "Materials and methods" (Fig. 2). Phylogenetic tree analysis revealed that the sequence for OPPS/PSY of *P. falciaprum* and other apicomplexan parasites along with some bacteria such as *T. maritima* and

Table 2 PRATT results showing motifs in OPPS/PSY of *P. falciparum*. The sequences were analysed using PRATT with parameters as mentioned in materials and methods and the results are summarized as shown. (Color table online)

	Fitness	Hits(Seqs)	Pattern
1:	37.5305	223(223)	E-x(2)-H-x(3)-L-x-H-D-D-x(2)-D-x(4)-R-R
2:	33.3604	223(223)	H-x(3)-L-x-H-D-D-x(2)-D-x(4)-R-R
3:	29.1904	223(223)	L-x-H-D-D-x(2)-D-x(4)-R-R
4:	28.6904	379(223)	L-x(0,1)-H-D-D-x(2)-D-x(4)-R-R
5:	25.0203	223(223)	H-D-D-x(2)-D-x(4)-R-R
6:	24.0203	223(223)	G-x(3)-G-x(3)-Q-x(2)-D-D-x(1,3)-L
7:	20.3503	237(223)	D-x(0,1)-D-x(2)-D-x(4)-R-R
8:	19.8503	247(223)	L-x(2,3)-D-D-x(1,2)-D-x(4)-R
9:	19.8503	223(223)	G-x(3)-Q-x(2)-D-D-x(1,3)-L
10:	16.6802	283(223)	D-D-x-L-D
11:	16.6802	247(223)	D-x(2)-D-x(4)-R-R
12:	16.1802	247(223)	D-D-x(1,2)-D-x(4)-R
13:	16.1802	253(223)	L-x(3,4)-D-x(2)-D-x(4)-R
14:	16.1802	575(223)	L-x(1,2)-D-D-x(2)-D
15:	15.6802	223(223)	Q-x(2)-D-D-x(1,3)-L
16:	15.6802	243(223)	I-x(0,1)-H-x(3)-L-x(2,3)-D
17:	15.1802	256(223)	L-x(3)-D-x(0,2)-D-x(3,4)-R
18:	15.1802	305(223)	D-D-x(1,2)-D-x(2,4)-R
19:	15.1802	625(223)	L-x(0,2)-D-D-x(1,2)-D
20:	15.1802	223(223)	E-x(0,2)-I-x(4,5)-L-x-H
21:	12.5102	288(223)	D-x-L-D
22:	12.5102	263(223)	K-x(3)-K-x(3)-L
23:	12.0102	301(223)	D-x(1,2)-D-x(4)-R
24:	12.0102	512(223)	A-x(4)-D-x(1,2)-L
25:	12.0102	583(223)	T-A-x(1,2)-L
26:	12.0102	278(223)	T-x(1,2)-P-x-L
27:	12.0102	586(223)	A-x(0,1)-S-L
28:	12.0102	438(223)	T-x(1,2)-S-L
29:	12.0102	625(223)	A-x(1,2)-L-x(2)-D
30:	12.0102	464(223)	A-x(4,5)-D-x(2)-L
31:	12.0102	315(223)	A-x(5)-D-x(1,2)-L
32:	12.0102	348(223)	D-x(3,4)-R-R
33:	11.5102	507(223)	K-x(0,1)-A-x(0,1)-L
34:	11.5102	305(223)	I-x(2,4)-L-x(3)-D
35:	11.5102	688(223)	A-x(0,2)-S-L
36:	11.5102	561(223)	A-x(3,4)-D-x(1,2)-L
37:	11.5102	267(223)	G-x(3)-G-x(1,3)-Q
38:	11.5102	678(223)	T-x(0,1)-A-x(1,2)-L
39:	11.5102	720(223)	A-x(1,2)-L-x(1,2)-D
40:	11.5102	444(223)	G-x(1,3)-A-x(2)-L
41:	11.5102	375(223)	K-x(2,3)-S-x(0,1)-L
42:	11.5102	495(223)	T-x(1,3)-S-L
43:	11.5102	489(223)	D-x(1,2)-L-x(1,2)-A
44:	11.5102	659(223)	L-x(1,2)-D-x(1,2)-L
45:	11.5102	330(223)	E-x(1,2)-I-x(4,5)-L
46:	11.5102	300(223)	E-x(0,1)-I-x(4,5)-L
47:	11.5102	310(223)	E-x(1,3)-I-x(4)-L
48:	11.5102	653(223)	T-A-x(0,2)-L
49:	11.0102	761(223)	L-x(1,2)-D-x(0,2)-L
50:	11.0102	835(223)	A-x(0,2)-L-x(1,2)-D

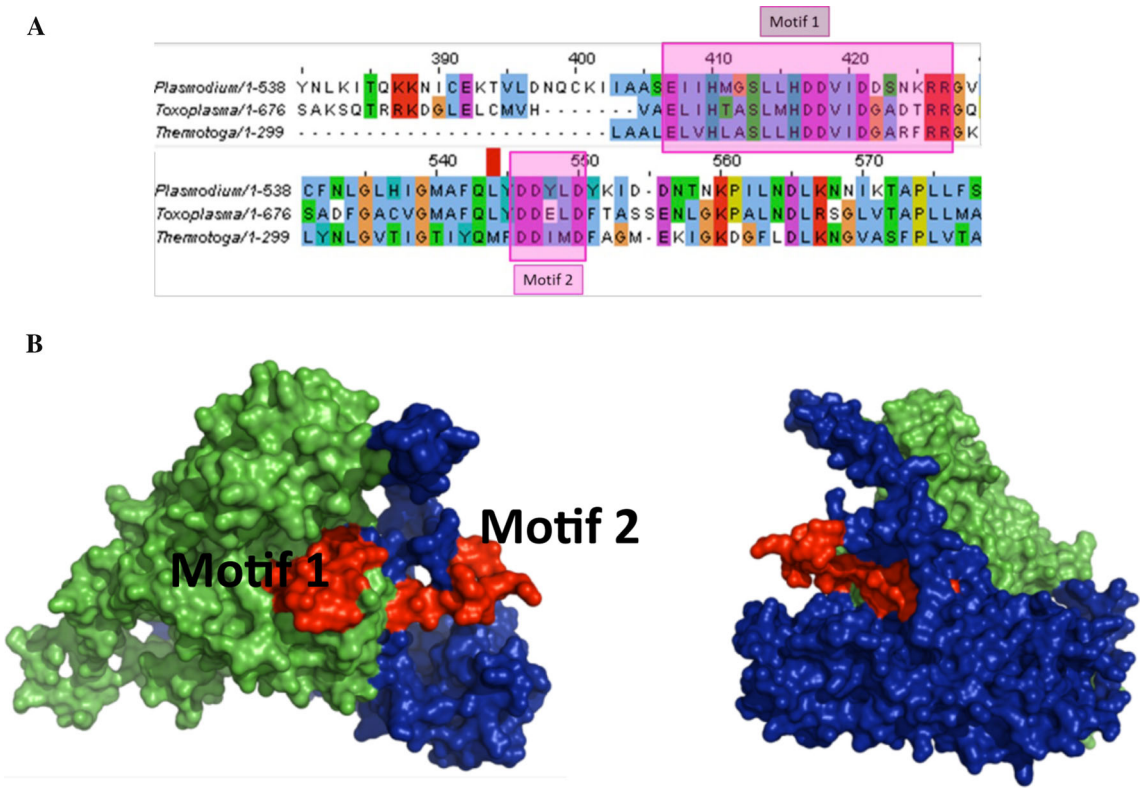


Fig. 5 Motifs in OPPS/PSY of *P. falciparum* and Surface representation. **a** Sequence alignment of *P. falciparum* OPPS, *T. gondii* PPS and *T. maritima* OPPS displaying the conserved motif 1 [E-x(2)-H-x(3)-L-x-H-D-D-x(2)-D-x(4)-R-R] and motif 2[D-D-X(2)-D] shaded as *pink*. **b** The surface representations of the two subunits of *P. falciparum* OPPS, coloured as *green* (chain A) and *blue* (chain B). *Red* indicates the region from residue number 198 to 206 indicated as motifs by Pratt. (Color figure online)

Rubrivivax gelatinosus occupy one cluster (Fig. 2) suggesting that they share some similar biological functions.

To have profound insights about the structure of *P. falciparum* OPPS/PSY secondary and tertiary structure predictions were done. Similarly the secondary and tertiary structure predictions were also performed for *T. gondii* PPS. The secondary structure predicted by using PSIPRED showed alpha-helices, beta-sheets, and coiled regions (Fig. 3a). The tertiary structures of *P. falciparum* OPPS/PSY, *T. gondii* PPS were predicted using SWISS MODEL with parameters of energy minimization values (Fig. 3b). The best modeled structures were validated by PROCHECK in SAVES server. Ramachandran plot showed the acceptability of the best models with 646 (88 %)amino acids arranged in most favored regions and 5 (0.7 %) amino acids in disallowed regions in case of *P. falciparum* and 634 (91.6 %) amino acids arranged in most favored regions and 4(0.6 %) of the amino acids in disallowed regions in case of *T. gondii* respectively (Fig. 3c).

Though having significant sequence homology, several gaps were found in the sequence alignment of P. falciparum OPPS/PSY, T. gondii PPS and T. maritima OPPS. We tried to model the biggest of this gap i.e. 206–261 in case of *P. falciaprum* OPPS/PSY, but no appropriate template could be found for it. The region from residue number 206 to 261 thus is unmodellable and might probably represent the region with some unknown characteristic and might contribute novel accessory functions to this protein (Fig. 4). Nonetheless the possiblity that this unique region would actually form a folded domain like structure cannot be ruled out. Further studies are necessary at this point to be able to comment more about this region of the enzyme as this loop might be deemed to have some physiological relevance and could offer potential drug target.

With the purpose to further explore the properties of the enzyme which might influence its biological activity, we looked for the patterns/motifs that the amino acids may contribute using PRATT. The fifty motifs found on PRATT analysis are listed (Table 2). It is evident from motifs found, that the [E-x(2)-H-x(3)-L-x-H-D-D-x(2)-D-x(4)-R-R] and [D-D-X(2)-D] motifs are conserved in *P. faciaprum* and *T. gondii* and interestingly they are juxtapositioned around the active site of trans OPPS of *T. maritima*. (Guo et al. 2004) (Fig. 5a, Table 2). We further visualized these two conserved motifs of *P. faciaprum* OPPS by PyMol

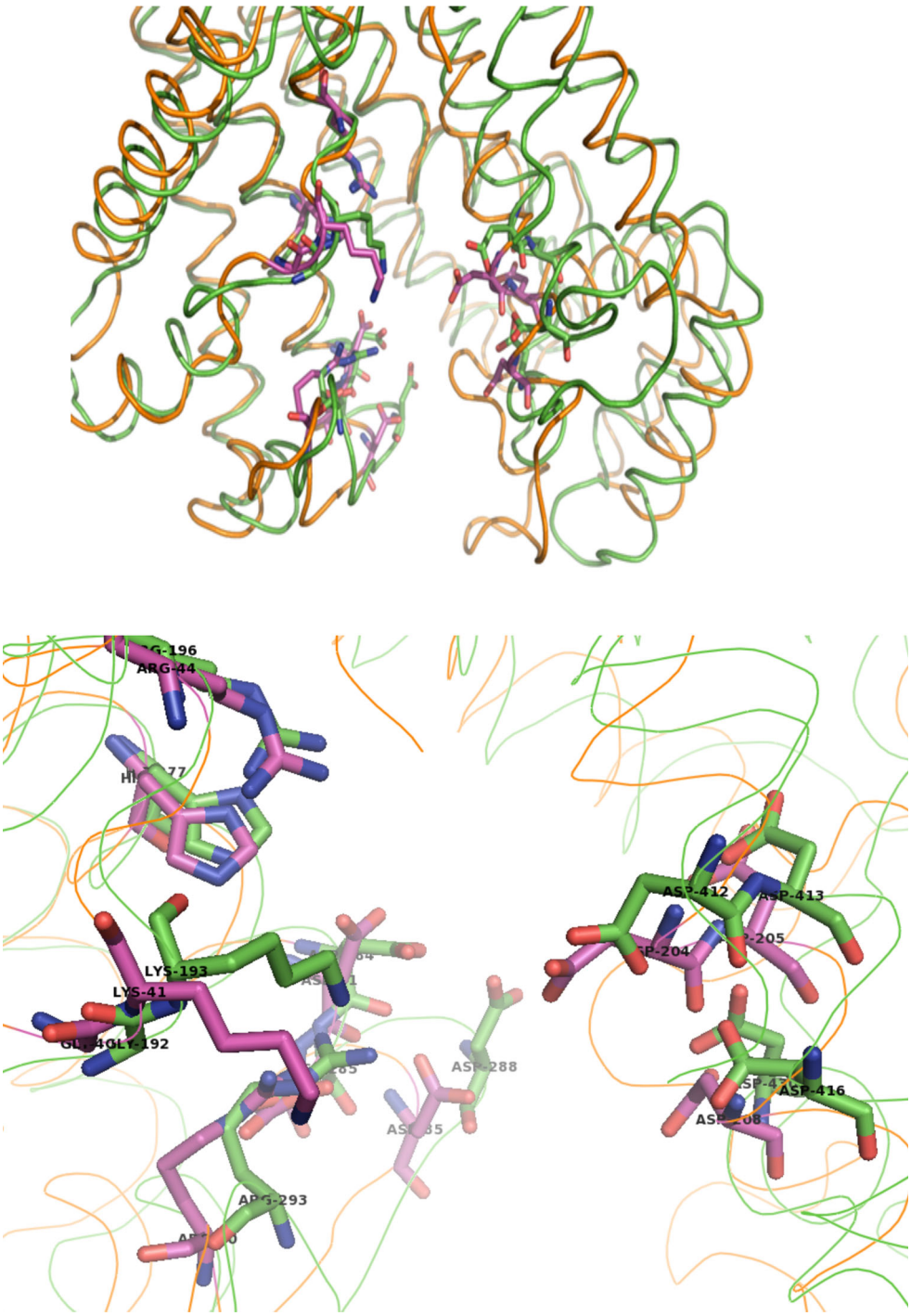


Fig. 6 The predicted substrate binding structure of *P. falciparum* OPPS/PSY. Comparison of substrate binding site residues in superposed structure of *P. falciparum* model over 1V4E. The superposed structuers are shown in ribbon (*P. falciparum* OPPS as *Green* and 1V4E as *saffron*). Active site residues Lys-41, Arg-44, His-74, Asp-81, Asp-82, Asp-85, Arg-90, Arg-91, Asp- 204, Asp-205, and Asp-208 of 1V4E are shown in *sticks* colored by atom (Carbon *green*, nitrogen *blue*, oxygen *red*). While the corresponding residue Arg-192, Phe-193,Iso-196, His-277, Asp-284, Asp-285, Asp-288, Arg-293, Asp-412, Asp-413, Asp-416, Asp-430 of *P. falciparum* model in sticks coloured by atom (Carbon *pink*, nitrogen *blue*, oxygen *red*). (Color figure online)

(Fig. 5b). The surface representations of the model depicts the motifs colored as red (Fig. 5b).

To predict the substrate binding pocket and location of these two conserved motifs superposition of structural regions of OPPS/PSY of *P. falciaprum* was done with the crystal structure of *T. maritima* OPPS using Superpose program from the CCP4 suite. The results of structural alignment show that the amino acid residues corresponding to the substrate binding site in *T. maritima* OPPS are well superposed with the *P. falciparum* residues (Fig. 6). Hence, this structurally conserved region can be considered as substrate binding site in *P. falciparum.*

Such comparison of binding sites and generation of hypothetical protein models of *P. falciparum* OPPS/PSY followed by structural alignment can provide deep insights for the design of novel drugs. Further, an unconserved

unique loop depicted in *P. falciparum* OPPS/PSY based on structural alignment could be exploited for structure based drug designing against malaria parasite.

Acknowledgments Authors would like to thank Shiv Nadar University, Uttar Pradesh for providing the resources to conduct these studies. S.S. is a recipient of the IYBA Award from Department of Biotechnology (DBT). S.A. is supported by senior research fellowship from DBT. V.S. is supported by doctoral fellowship by Shiv Nadar University.

Conflict of interest The authors declare no conflict of interest.

References

Arnold K et al (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics 22:195–201

Benkert P et al (2011) Toward the estimation of the absolute quality of individual protein structure models. Bioinformatics 27: 343–350

Bowie JU et al (1991) A method to identify protein sequences that fold into a known three-dimensional structure. Science 253(5016):164–170

Clastre M et al (2007) The methylerythritol phosphate pathway for isoprenoid biosynthesis in coccidia: presence and sensitivity to fosmidomycin. Exp Parasitol 116:375–384

Cunningham FX Jr, Gantt E (1998) Genes and enzymes of carotenoid biosynthesis in plants. Annu Rev Plant Physiol Plant Mol Biol 49:557–583

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–1797

Goodwin TW (1980) *The biochemistry of carotenoids*. Plants vol I Chapman&Hall. London, p. 98

Guo RT, et al. (2004) Crystal structure of octaprenyl pyrophosphate synthase from hyperthermophilic Thermotoga maritima and mechanism of product chain length determination. J Biol Chem 279:4903–4912. doi: 10.1074/jbc.M310161200

Jones DT et al (1992) The rapid generation of mutation data matrices from protein sequences. Comput Appl Biosci 8:275–282

Jordão FM et al (2011) Isoprenoid biosynthesis in the erythrocytic stages of *Plasmodium falciparum*. Mem Inst Oswaldo Cruz 106(Suppl 1):134–141

Larkin MA et al (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23(21):2947–2948

Laskowski RA et al (1993) PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283–291

Lichtenthaler HK et al (1997) Two independent biochemical pathways for isopentenyl diphosphate and isoprenoid biosynthesis in higher plants. Physiol Plant 101:643–652

Lu S et al (2008) Carotenoid metabolism: biosynthesis, regulation and beyond. J Integr Plant Biol 50:778–785

Lüthy R et al (1992) Assessment of protein models with three-dimensional profiles. Nature 356(6364):83–85

Maiti R, et al. (2004) SuperPose: a simple server for sophisticated structural superposition. Nucleic Acids Res 32(Web Server issue):W590–4

McWilliam H, et al. (2013) Analysis tool web services from the EMBL-EBI. Nucleic Acids Res, (Web Server issue):W597–600. doi: 10.1093/nar/gkt376

Morris AL et al (1992) Stereochemical quality of protein structure coordinates. Proteins 12:345–364

Ralph SA et al (2004) Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. Nature 2:203–216

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425

Salvini M et al (2005) cDNA cloning and expression of the phytoene synthase gene in sunflower. J Plant Phys 162:479–484

Tamura K et al (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol 30:2725–2729

Tonhosolo R et al (2005) Identification, molecular cloning and functional characterization of an octaprenyl pyrophosphate synthase in intra-erythrocytic stages of *Plasmodium falciparum*. Biochem J 392:117–126

Tonhosolo R et al (2009) Carotenoid biosynthesis in intraerythrocytic stages of *Plasmodium falciparum*. J Biol Chem 284:9974–9985

Yan Y et al (2005) Cloning and sequence analysis of the phytoene synthase gene from a unicellular chlorophyte, Dunaliella sanlia. J Agric Food Chem 53:1466–1469