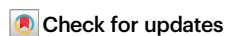


Cryo-EM structure of 1-deoxy-D-xylulose 5-phosphate synthase DXPS from *Plasmodium falciparum* reveals a distinct N-terminal domain

Received: 18 December 2023

Accepted: 17 July 2024

Published online: 05 August 2024

Victor O. Gawriljuk¹, Andre S. Godoy², Rick Oerlemans¹, Luise A. T. Welker¹, Anna K. H. Hirsch^{3,4} & Matthew R. Groves¹✉

Plasmodium falciparum is the main causative agent of malaria, a deadly disease that mainly affects children under five years old. Artemisinin-based combination therapies have been pivotal in controlling the disease, but resistance has arisen in various regions, increasing the risk of treatment failure. The non-mevalonate pathway is essential for the isoprenoid synthesis in *Plasmodium* and provides several under-explored targets to be used in the discovery of new antimalarials. 1-deoxy-D-xylulose-5-phosphate synthase (DXPS) is the first and rate-limiting enzyme of the pathway. Despite its importance, there are no structures available for any *Plasmodium* spp., due to the complex sequence which contains large regions of high disorder, making crystallisation a difficult task. In this manuscript, we use cryo-electron microscopy to solve the *P. falciparum* DXPS structure at a final resolution of 2.42 Å. Overall, the structure resembles other DXPS enzymes but includes a distinct N-terminal domain exclusive to the *Plasmodium* genus. Mutational studies show that destabilization of the cap domain interface negatively impacts protein stability and activity. Additionally, a density for the co-factor thiamine diphosphate is found in the active site. Our work highlights the potential of cryo-EM to obtain structures of *P. falciparum* proteins that are unfeasible by means of crystallography.

Plasmodium falciparum is one of the causative agents of malaria, a life-threatening disease that consistently affects more than 200 million people worldwide each year¹. In 2021, 619,000 people died of malaria, with the majority of cases in the African region (96%) among children under five years old¹. The 21st century has been marked by great efforts to control and eliminate malaria, with the mortality rate and deaths decreasing by approximately 50% and

36%, respectively, between the years 2000 and 2015². However, new challenges and concerns have started to arise since the emergence of artemisinin resistance *P. falciparum* in Pailin, Western Cambodia, in 2009³. Concerns are also growing about the impact of climate change on the global incidence of malaria, as a result of both changes in transmission vectors and vector population growth due to increased flooding⁴.

¹Chemical and Pharmaceutical Biology, Groningen Research Institute of Pharmacy, University of Groningen, Antonius Deusinglaan 1, 9713 AV Groningen, The Netherlands. ²Sao Carlos Institute of Physics, University of Sao Paulo, Av. Joao Dagnone, 1100 - Jardim Santa Angelina, Sao Carlos 13563-120, Brazil.

³Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) - Helmholtz Centre for Infection Research (HZI), Campus Building E8.1, 66123

Saarbrücken, Germany. ⁴Saarland University, Department of Pharmacy, Campus Building E8.1, 66123 Saarbrücken, Germany. ✉e-mail: m.r.groves@rug.nl

Artemisinin-based combination therapies (ACTs) are the first-line treatment for uncomplicated *P. falciparum* malaria worldwide. Due to their high efficacy against most stages of the parasite, ACTs have been one of the main weapons in controlling the disease, making artemisinin resistance a serious challenge^{5–7}. Since its discovery in 2009, artemisinin resistance has emerged in the Greater Mekong Subregion (GMS) of Southeast Asia^{8,9}, Papua New Guinea¹⁰, South America¹¹ and recently in Eastern Africa^{12,13}. In addition, resistance to all known antimalarials has been documented¹⁴, increasing the likelihood of ACT treatment failures due to resistance to both artemisinin and the partner drug, as has already been seen in GMS^{15,16}. Malaria vaccines such as RTS,S^{17,18} or the newly recommended R21/Matrix-M^{19,20} are important new tools that are safe and effective in preventing malaria in children. However, the need for the discovery and development of new antimalarials is still of great importance, both to tackle parasite resistance and to treat individuals for whom vaccination does not provide protection.

The *Plasmodium* genus is part of the *Apicomplexan* phylum, characterized by the presence of a non-photosynthetic plastid known as the apicoplast²¹. This organelle holds several functions important for *P. falciparum* viability, including isoprenoid precursor biosynthesis^{22,23}. Isoprenoids, such as ubiquinone and sterols, are essential molecules for cell homeostasis, making their synthesis crucial for parasite survival²⁴. In contrast to humans, which use the mevalonate pathway for the synthesis of isoprenoid precursors, *Plasmodium* spp. relies solely on the 2-C-methyl-D-erythritol 4-phosphate pathway (MEP pathway) for their synthesis²⁵. The absence of this pathway in humans and its essential nature for *P. falciparum* make it an important source of targets for the development of new antimalarials²⁶.

1-deoxy-D-xylulose-5-phosphate synthase (DXPS) is the first and rate-limiting step of the MEP pathway²⁷. The enzyme catalyses the conversion of glyceraldehyde-3-phosphate and pyruvate to 1-deoxy-D-xylulose-5-phosphate (DXP) using thiamine diphosphate (ThDP) as a co-factor²⁸. Despite being an attractive target, most research on MEP pathway antimalarials is focused on the second enzyme DXR, with fosmidomycin being a clear example of the potential in targeting this pathway²⁹. One of the reasons for the lack of studies targeting *Plasmodium* spp. DXPS is the challenge in its recombinant expression^{30,31}, together with the lack of structural information, making small molecule screening and structure-based drug discovery less attractive when compared to other MEP pathway targets.

The recent discovery of a truncation construct proved to be useful for the crystallisation of several bacterial DXPS enzymes^{32–35}. However, this approach is less helpful for the *P. falciparum* DXPS (*PfDXPS*), due to the presence of low-complexity regions with predicted regions of high disorder, which reduces the success rate of protein crystallisation³¹.

In this work, we use single-particle cryo-EM to elucidate the structure of *PfDXPS* to a final resolution of 2.42 Å, circumventing the disadvantages of crystallisation for this target. The structure is similar to other DXPS, with the addition of an N-terminal cap domain exclusively present in the *Plasmodium* genus. We further explore the N-terminal cap domain with mutational studies targeting the N-terminal cap:Domain I interface, showing that destabilization of this interaction decreases protein stability and activity. While the origins of this domain are still unknown, our study shows that the cap domain is an important feature of *Plasmodium* spp. DXPS. Finally, our results demonstrate the potential of cryo-EM for structure-based drug design projects targeting plasmodial proteins that are unsuitable for protein crystallisation.

Results

Cryo-EM structure of *PfDXPS*

Despite its potential as an antimalarial target, there are no structures available for *PfDXPS*. To avoid the disadvantages of protein crystallisation for this target, we used single-particle cryo-EM to solve the

structure of *PfDXPS* at 2.42 Å. The dataset analysis and final statistics are shown in Supplementary Fig. 1 and Supplementary Table 1.

The *PfDXPS* construct used in this work consisted of a truncation in which the signal and transit peptides were removed, increasing protein expression with no impact on protein activity³⁰. The final model showed a map-model correlation coefficient (CC) of 0.73 and no C β or Ramachandran outliers, with all rotamer outliers being justified by the map. The map is clear for most of the residues, with some areas of lower quality, such as the first 20 residues and segments of the C-terminal domain that are close to the surface and proximal to a low-complexity region. Due to the position of these regions within the overall structure, the lower quality of their maps may be explained by their higher flexibility. Five regions, corresponding to low-complexity regions, are not visible on the map (residues 554–617, 684–704, 754–819, 999–1029, 1049–1066).

Like other DXPS homologues, the protein forms a homodimer, with each monomer consisting of three distinct domains^{33–35}. Domain I (419–753) and Domain II (820–998) together compose the ThDP binding domain and active site; whereas Domain III (1030–1205) mainly makes dimerization contacts. Distinct from all other DXPS structures, *PfDXPS* has an additional α -helix domain of 106 residues at the N-terminus (311–418). The N-terminal cap domain is surface-exposed and makes contacts exclusively with Domain I (Fig. 1). This domain is further discussed in the following sections.

The cryo-EM structure of *PfDXPS* is of a protist enzyme. Although the sequence identity between other homologues is low, the structure has an average root mean square deviation (RMSD) of 1.44 ± 0.14 Å against the available structures of different DXPS (Supplementary Fig. 2). The structural alignment algorithm superposes structures based on their C α -atom positions, optimizing the overall RMSD value by excluding atoms if their individual value is higher than 2. After alignments with all homologues, residues excluded from the calculation comprised of the N-terminal cap domain and some insertions found in *PfDXPS*.

The insertions found either extend secondary structures or connections between them. Most of them are in surface-exposed areas of the third domain where the addition or deletion of residues is also common in other homologues (Supplementary Fig. 2A).

Low-complexity regions

One of the remarkable characteristics of the *P. falciparum* genome is its high AT-content, recombination rate, and the large presence of low-complexity regions (LCR), with almost 90% of *P. falciparum* proteins having at least one region³⁶. LCR can vary in size, with lengths ranging from 20 to 250 amino acids or more, and can be divided into heterogeneous, PolyN, or HighGC regions³⁷. Their composition is mainly hydrophilic and consists of amino acids such as asparagine and lysine. This is one of the main reasons why most *P. falciparum* proteins are larger than their respective orthologues. For example, *PfDXPS* has a molecular mass of 103 kDa compared to other DXPS, which have an average mass of 60 kDa.

The *PfDXPS* structure has five regions with no density visible in the cryo-EM map obtained. We performed a multiple-sequence alignment (MSA) of DXPS from different *Plasmodium* spp. and found that all regions have some degree of low-complexity and low identity within them (Supplementary Table 2). Their lengths were not shared among the species, with some of them having fewer or more residues for the same region (Supplementary Table 3).

The first and longest one is 68 amino acids long and located within residues 554–617 (Fig. 2a). The region is surface-exposed and connects an α -helix to a β -strand in Domain I. The second is located directly after the metal-binding site (684–704) and is known to be poorly conserved and highly disordered in other DXPS³².

The third one lies within the region that is also known as ‘spoon-fork’ motif (754–819), which is not visible in most of the available DXPS

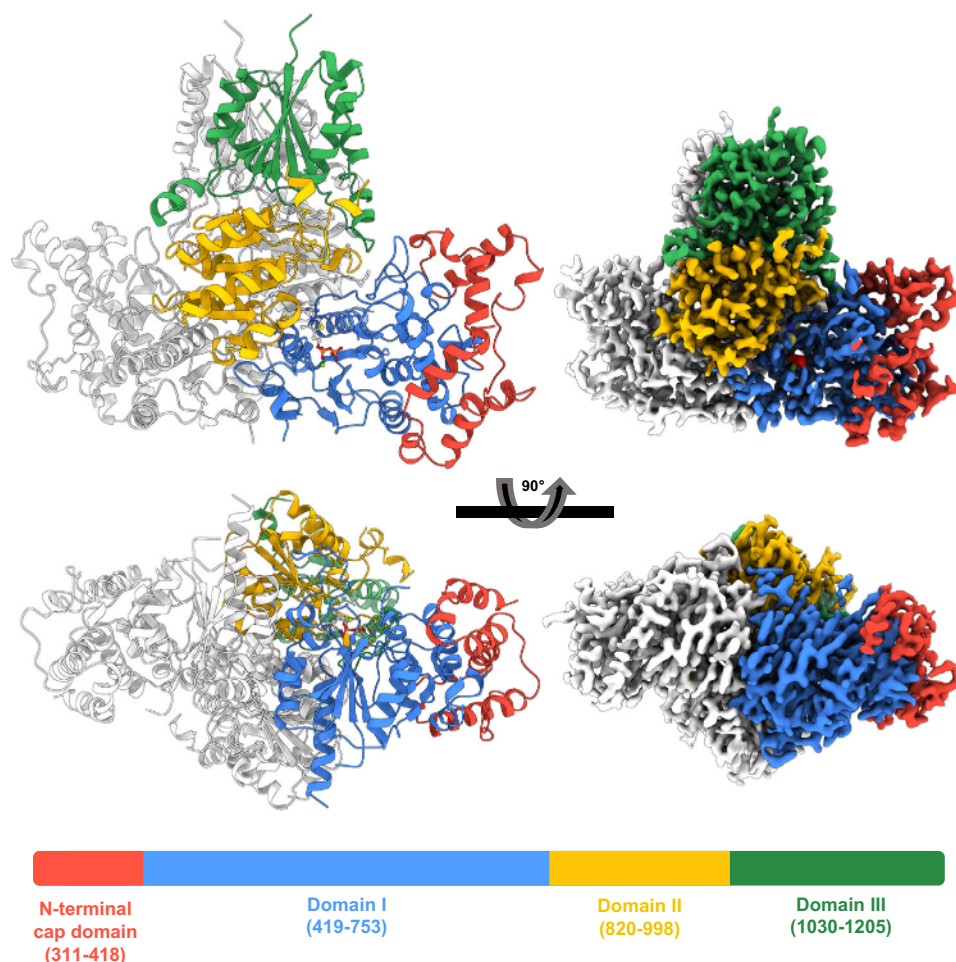


Fig. 1 | Cryo-EM structure of *PfDXPS*. The protein arrangement is a homodimer with each monomer composed of four domains. The N-terminal cap domain is shown in red (311–418), Domain I (419–753) in blue, Domain II (820–998) in yellow,

and Domain III (1030–1205) in green. The cryo-EM map coloured by the domain and contoured at 3.4σ is shown on the right. A *PfDXPS* sequence schematic showing the domain positions is displayed at the bottom.

structures, due to its high flexibility during the open and closed state transitions of the enzyme^{38,39}. However, in *PfDXPS*, this region has additional residues with low-complexity that might increase the flexibility and disorder of this motif. Interestingly, *PfDXPS* and its homolog from *Plasmodium vivax* (*PvDXPS*) have a k_{cat}/K_m for pyruvate approximately ten times lower than that of other DXPS, suggesting that this LCR might affect the stabilization of the pyruvate intermediate through the opening and closing state transition of the spoon-fork motif³⁰. The last two regions are located in close proximity and are only separated by a small conserved segment of 19 residues. The first (999–1029) connects domains II and III, while the second one (1049–1066) is a connection between two β-strands in Domain III (Fig. 2a).

When compared with other apicomplexan organisms, the only LCR that is not shared is LCR-I, which is unique in the *Plasmodium* genus. We performed a truncation of 43 residues in this region (spanning residues 568–611) to evaluate the effect of its presence on *PfDXPS*. The remaining 25 residues were kept to guarantee enough space for linking the two secondary-structure segments in Domain I without hampering the protein fold with a shorter linker.

PfDXPS-ΔLCR-I could still be expressed and purified (Supplementary Fig. 3a) and showed no difference in activity when compared with the wild-type enzyme (Supplementary Fig. 3b), suggesting that LCR-I does not impact protein activity. Interestingly, the truncation in LCR-I resulted in an increase of almost 6 degrees in the protein's melting temperature when compared with the wild-type enzyme (Fig. 2b).

The similar activity with or without one of the LCRs may explain why insertions in these specific regions were propagated, as they show a neutral effect on protein activity. While the specific function of these LCRs in *PfDXPS* or other *Plasmodium* spp. proteins are out of the scope of this study, different theories about their retention and activity exist –such as playing a role in modulating host immune response^{40,41}, serving as cryptic introns for enhanced RNA folding⁴², tRNA ‘sponges’ to compensate for low levels of tRNAs⁴³, or simply an artefact of neutral evolution³⁶. However, the lack of structures for these LCRs may have implications on overall protein stability, as shown here for the LCR-I, requiring the development of additional structural elements to ensure correct physiological function, particularly in enzyme systems known to possess high structural flexibility.

The N-terminal cap domain is exclusive to the *Plasmodium* genus

Unlike other DXPS structures, *PfDXPS* has an additional α-helix domain at its N-terminus. The domain consists of nine α-helices that form a cap-like arrangement enclosing the N-terminal region of Domain I (Fig. 3a, Supplementary Fig. 4a).

The first group of helices encloses the N-terminal α-helix of Domain I, decreasing its solvent-exposed area. Three tyrosine residues (Y329, Y333, Y345) form hydrogen bond interactions with the helix, while the remaining interactions are hydrophobic contacts between buried residues lying in the domain's interface (Fig. 3b). The group of helices is connected to the second one through a linker of 27 residues

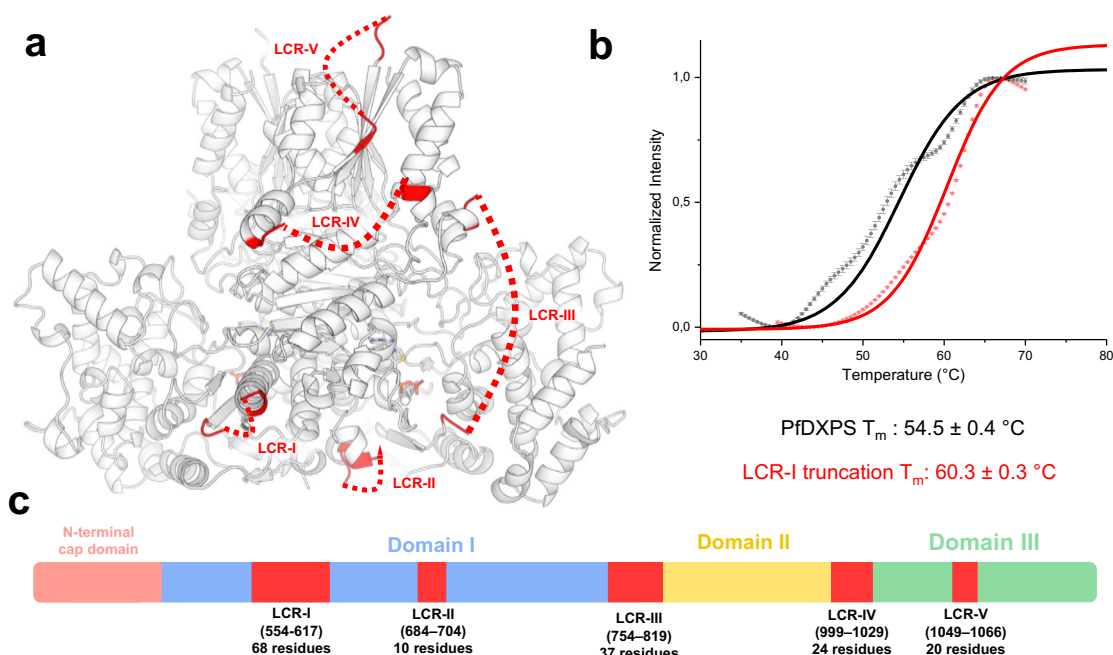


Fig. 2 | Low-complexity regions (LCR) found in *PfDXPS*. **a** *PfDXPS* is represented as a white cartoon and the location of each region is highlighted as red dashed lines in the structures. **b** Differential scanning fluorimetry curves for the LCR-I truncated, shown in red, and wild-type enzyme, shown in black. Raw normalized data are

presented as mean $\pm$ SD and shown as scatter points. Boltzmann-fitted curves are shown as lines. Graphs were plotted from technical quintuplicates ($n = 5$). **c** Sequence schematic with length and position of each LCR region in the *PfDXPS*. Source data are provided as a Source Data file.

with two short helices on both ends. The only interaction with Domain I is a salt bridge between E361 and K431.

The last group of two helices is connected through a linker of 14 residues and folded to hold the loop region below the N-terminal α -helix of Domain I (Fig. 3c). The linker between them forms an extensive network of hydrogen bonds between the main chain and side chains of N396, Y401, and S522. After the hydrogen bond network, the last helix encloses the rest of the N-terminal α -helix of Domain I, completely burying it from the solvent. Similar to most surface-exposed α -helices, the outer part of the cap domain is mainly composed of charged residues, whereas the inner parts that form the interface between the two domains are mostly hydrophobic, contributing to most of the domain's interactions (Supplementary Fig. 4b).

An MSA of the DXPS from all *Plasmodium* spp. available in UniProt shows that the domain is present and conserved across all members of the *Plasmodium* genus, highlighting its importance in these species (Fig. 3d). To evaluate the origins and evolution of this domain, a second MSA was performed, including all DXPS sequences from organisms in the *Apicomplexa* phylum available in UniProt (Supplementary Table 2).

Interestingly, the domain was not found in any other organism, suggesting that it is unique to the *Plasmodium* genus, or that its evolutionary origin is still uncharacterized. A protein structure search was also performed using only with the N-terminal cap domain with the web servers Foldseek⁴⁴, CO-FACTOR⁴⁵ and Dali⁴⁶. The only proteins that provided matches, however, were the DXPS of other *Plasmodium* spp. available in the AlphaFoldDB, suggesting that the domain fold is also unique to these species.

N-terminal cap domain stability is required for optimal enzymatic activity

To evaluate the function of the cap domain, we performed a series of mutational analyses in this region. First, a truncated version of the *PfDXPS* without the cap domain was generated. Without the domain, however, the expression yield of soluble protein was drastically reduced, with the protein present in the insoluble fraction

(Supplementary Fig. 5a). This contrasts with the construct used in the structure determination, which was expressed predominantly in the soluble fraction under identical conditions (Supplementary Fig. 5b).

The surface of Domain I is largely hydrophobic in nature and buried at the interface with the cap domain. Removing the cap domain would expose the hydrophobic surface to the solvent, which could explain the decrease in protein solubility during expression (Supplementary Fig. 4). Enzymatic assays were performed with the eluted fractions from the affinity chromatography but showed no measurable activity (Supplementary Fig. 5d).

Therefore, we conducted a mutational study to further assess the role of the domain. Three mutants were generated to evaluate the effect of disrupting the interface between the cap domain and Domain I. The mutants were selected by identifying hydrophobic residues with a large buried surface area at the interface and replacing these residues with charged ones such as glutamate or aspartate.

The webserver PISA was used to evaluate the residues with the highest buried surface area at the interface⁴⁷. The most buried residues were then evaluated for predicted destabilization after mutation using the webserver DDMut⁴⁸. DDMut predicts changes in Gibbs Free Energy ($\Delta\Delta G$) of the protein after mutation to determine whether the mutation would stabilize or destabilize the protein. The most destabilizing mutations were then checked for sequence conservation across different *Plasmodium* spp. After assessing these criteria, three mutations were chosen (I320D, I341D and F387E), their locations are shown in Fig. 4a.

The mutants were then expressed and purified using a protocol similar to the wild-type enzyme. For both I320D and I341D, the expected band of 103 kDa was not observed. Instead, I320D showed two bands of around 55 kDa in the elution from the immobilized metal affinity chromatography (IMAC), while I341D showed a similar but single band (Supplementary Fig. 6a). Since an identical protocol was followed for all mutants, the presence of different molecular weight bands could indicate protein degradation of *PfDXPS* due to the destabilization of the N-terminal cap domain interaction with Domain I. Activity assays were performed with the IMAC-purified

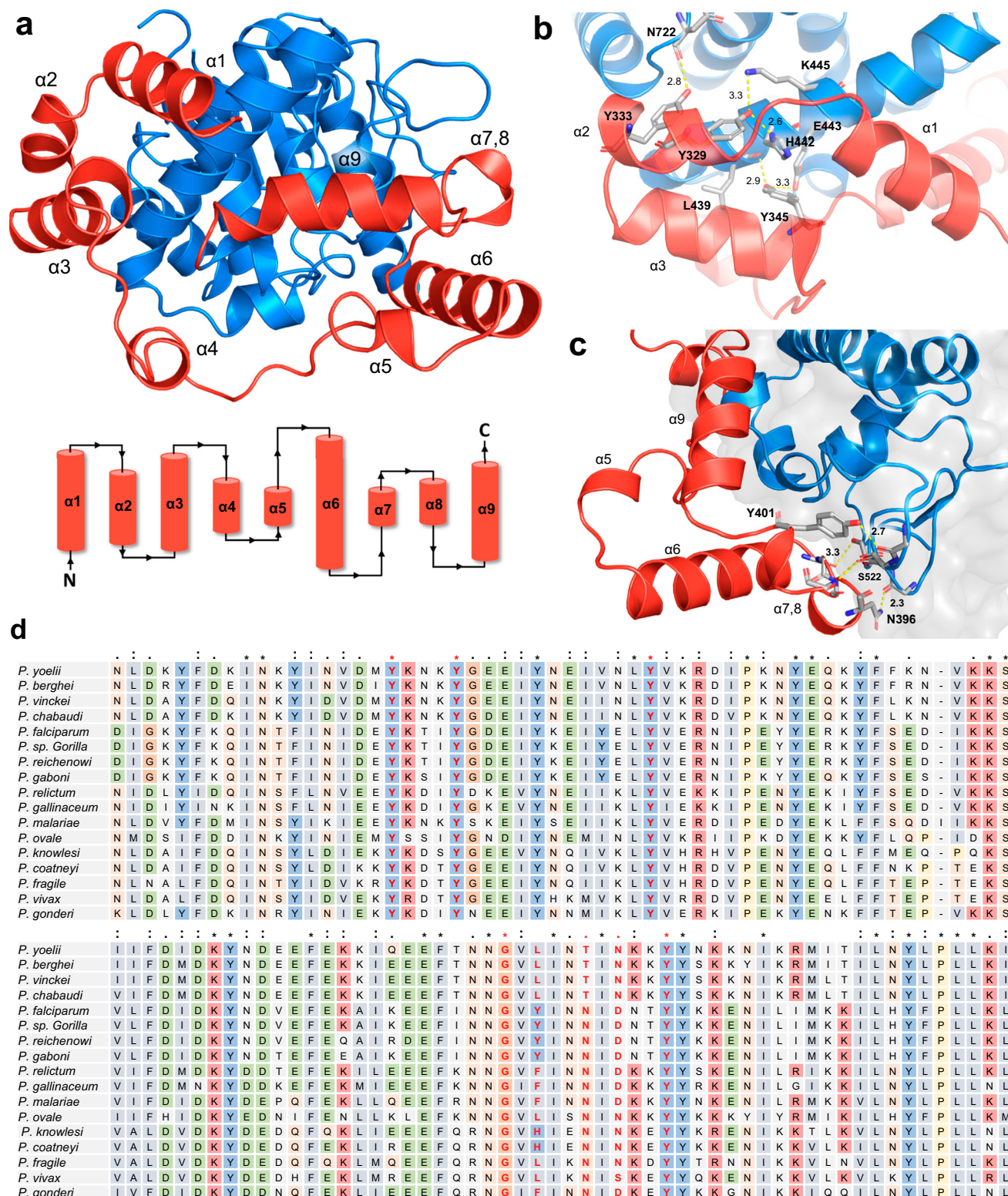


Fig. 3 | Structural analysis of the N-terminal cap domain. **a** Top: the N-terminal cap domain (red) makes only interactions with Domain I (blue); Bottom: 2D-topology of the cap domain. **b** Cap domain interactions with the N-terminal α -helix in Domain I. Hydrogen bonds are depicted as yellow dashes with distances in Ångström. **c** Further interactions between the cap domain and Domain I. Hydrogen

bonds are indicated by yellow dashes with distances in Ångström. **d** Multiple-sequence alignment of DXPS from different *Plasmodium* spp., with the alignment of the cap domain highlighted. The residues making interactions with Domain I in **b** and **c** are shown in red.

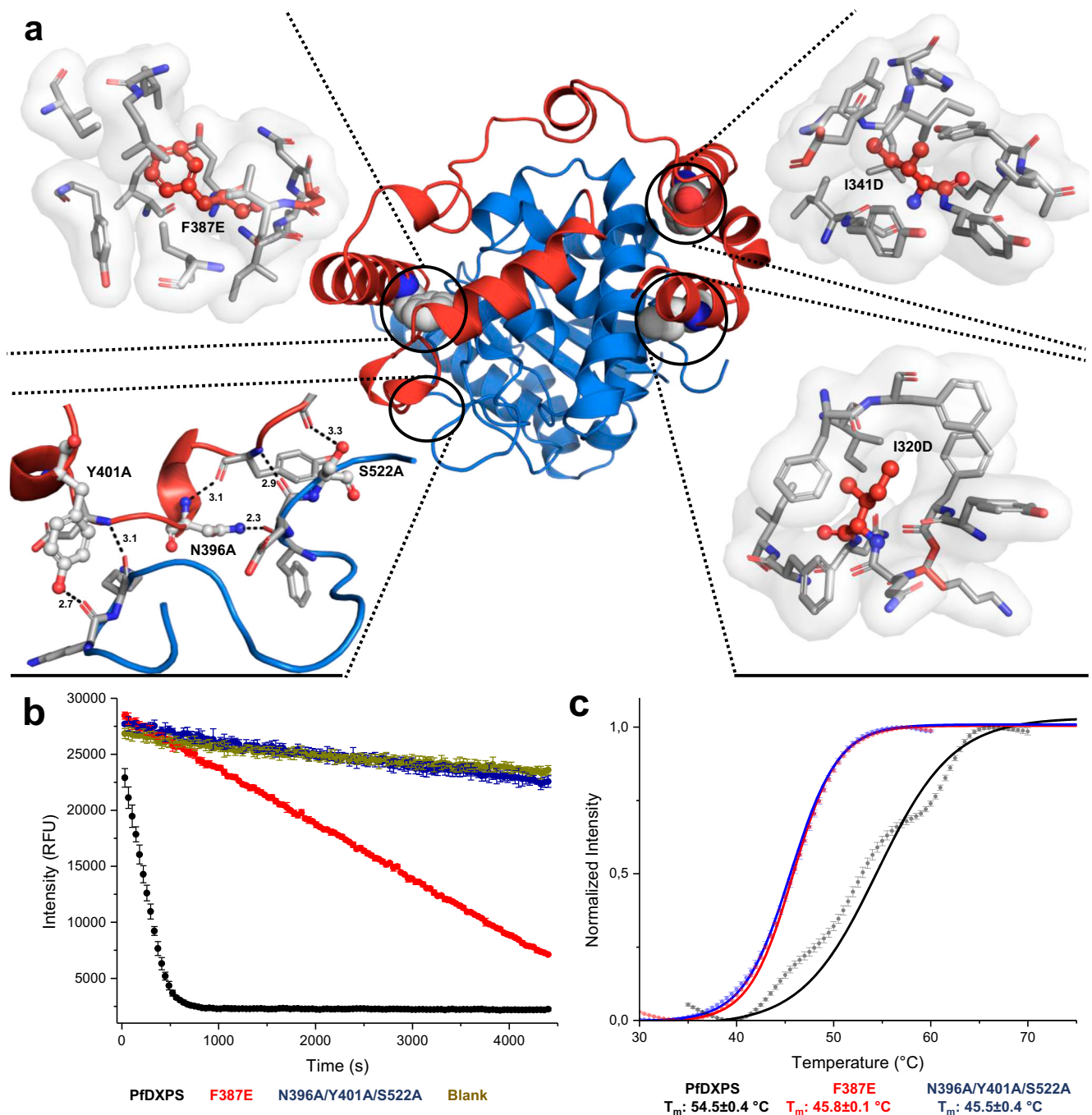


Fig. 4 | Mutation studies of the N-terminal cap domain. a Position of the four mutants for destabilization of the interface between cap domain (depicted as a red cartoon) and Domain I (depicted as a blue cartoon). **b** Graph of intensity versus time for the enzymatic reaction at a fixed protein concentration of 500 nM. Wild-type enzyme is shown in black, the F387E mutant is shown in red, the N396A/Y401A/S522A triple mutant is shown in blue, and blank is shown in dark yellow. Data are presented as scatter points with mean values $\pm$ SD. Graphs were plotted from

technical triplicates ($n = 3$). **c** Differential scanning fluorimetry curves for the mutants and wild-type enzyme. Wild-type enzyme is shown in black, the F387E mutant is shown in red, and the N396A/Y401A/S522A triple mutant is shown in blue. Raw normalized data are presented as scatter points with mean values $\pm$ SD. Boltzmann-fitted curves are shown as lines. Graphs were plotted from technical quintuplicates ($n = 5$). Source data are provided as a Source Data file.

fractions of the mutants, but no activity was observed (Supplementary Fig. 6b).

Considering that the LCRs are potential hotspots for proteolytic degradation due to their highly disordered nature, a degradation product of PfdXPS containing a histidine tag of around 53 to 60 kDa could be obtained if, for instance, LCR-III was cleaved. Since both mutations are closer to the N-terminal portion of the enzyme, the strong destabilization caused by the mutations could negatively

impact the protein fold, potentially leading to protein degradation in our expression system.

On the other hand, F387E resulted in an elution profile for IMAC similar to that of the wild-type enzyme (Supplementary Fig. 6a). Activity assays were performed to evaluate its effect on protein activity. Interestingly, the mutation had a drastic effect, with its activity being reduced to ~10% of the wild-type enzyme (Fig. 4b). Protein stability was also assessed through differential scanning fluorimetry,

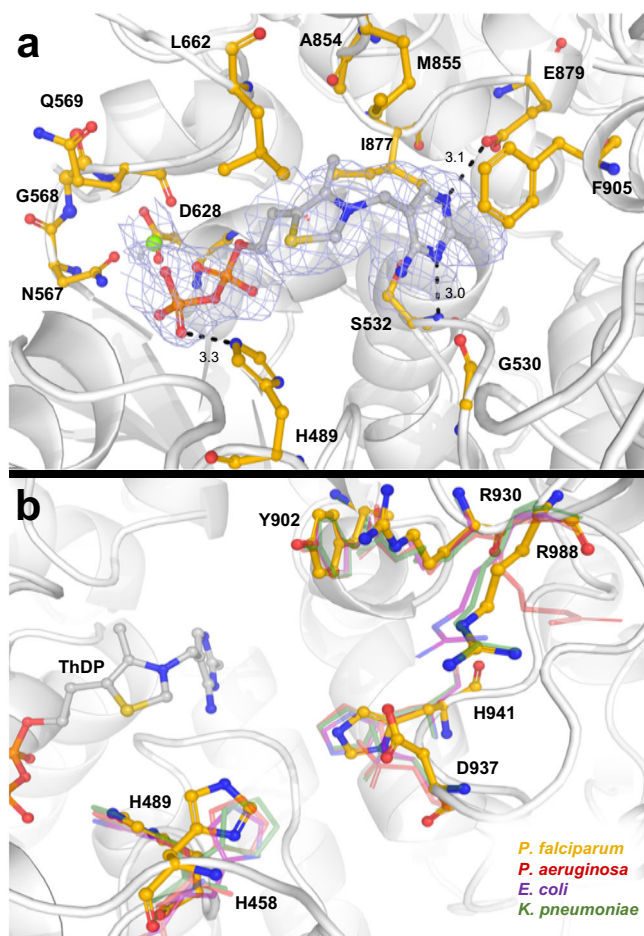


Fig. 5 | ThDP binding site and active site of *PfDXPS*. **a** ThDP shown as grey sticks with residues that make interactions with the co-factor highlighted as yellow sticks. The cryo-EM map density for the co-factor is highlighted as a mesh contoured at 3.4 σ . Hydrogens bonds are depicted as black dashes with distances in Ångström. **b** ThDP is shown as grey sticks, residues that are known to be important for catalysis are highlighted as sticks coloured for different organisms (Yellow for *P. falciparum*, purple for *E. coli*, red for *Pseudomonas aeruginosa* and green for *K. pneumoniae*).

which showed that the mutation reduced the melting temperature by 9 degrees when compared with the wild-type enzyme (Fig. 4c). Therefore, destabilization in the cap domain and Domain I interface has severe effects on enzyme activity and overall stability.

In a parallel approach, we sought to investigate the effects of destabilizing the domain interface by mutating residues that form a network of side-chain hydrogen bond interactions between the two domains. The network is an interaction of the cap domain with a loop in Domain I that becomes completely unstructured in the X-ray apo structures of *Escherichia coli* DXPS and *Klebsiella pneumoniae* DXPS (Fig. 4a)^{34,35}. We performed a triple mutation (Y401A, N397A, S522A) to evaluate the effect of losing this hydrogen bond network in *PfDXPS*. Similarly to F387E, the triple mutant led to a significant decrease of 9 degrees in the measured melting temperature (Fig. 4c). However, the protein activity was now approximately 1% of the wild-type enzyme (Fig. 4b). While the enzyme is still active in this unstable state, it requires a concentration at least 64 times higher than the wild-type to exhibit a similar profile (Supplementary Fig. 7).

Although the cap domain seems to be unique to the *Plasmodium* genus, its presence is necessary to guarantee proper solubility and stability. Destabilization of the interface between Domain I can affect the enzyme in various ways, hindering the purification of a full-length and soluble protein, or resulting in a soluble but unstable and less

active enzyme. Unfortunately, we were unable to obtain the truncated *PfDXPS* in large amounts to assess its enzymatic activity at higher concentrations. Nevertheless, our destabilization studies suggest that while the cap domain may not be required for protein activity, its presence significantly enhances it. Our results are consistent with truncation studies conducted by Handa et al. in 2013 on *P. vivax* DXPS (*PvDXPS*), where the deletion of this domain significantly reduced the expression yield and resulted in a notable loss of activity in the soluble fractions of the protein³¹.

ThDP complex and active site

A clear density assigned to ThDP was found in the active site of the protein. The pyrimidine ring makes a π - π stacking interaction with the conserved F905 and hydrogen bonds with S532. E879 is close to the nitrogen atom in the 1'-position of the pyrimidine ring, allowing the 4'-amino group to act as a proton acceptor for the C2 carbon, making the atom a nucleophile for the DXPS reaction. Finally, the diphosphate moiety coordinates the magnesium cation together with N657, Q659, and D628. The co-factor also makes several hydrophobic interactions with highly conserved residues (I877, M855, A854, L662) (Fig. 5, Supplementary Fig. 8).

The active site of DXPS is highly conserved. All residues shown in mutational studies to be important for catalysis (D937, H489) or to the GAP binding (R930, Y902, R988, H458) have clear density in the map and occupy similar positions to that seen in other DXPS, opening space for the development of broad-spectrum anti-infectives (Fig. 5b)^{35,49–51}. *PfDXPS* and all other DXPS have regions of high flexibility that are usually not seen in X-ray or cryo-EM maps. One of these regions was previously described as the 'spoon-fork' motif, which can adopt an open or closed conformation depending on the step within the enzymatic reaction^{38,39}. Although some density can be seen in areas close to this region, the amino acids of this chain could not be modelled with confidence, suggesting that this model of *PfDXPS* is in an open conformation. Consequently, no density could be modelled for H766, mutation of which has shown to be detrimental for protein activity⁴⁹. The presence of bound ThDP in the cryo-EM maps demonstrates that the currently presented cryo-EM structure provides a resolution suitable for inhibitor binding studies.

Discussion

PfDXPS is an essential enzyme for the isoprenoid biosynthesis in *P. falciparum* and an attractive target for the development of anti-malarials with different modes of action⁵². Development of small molecules targeting *PfDXPS* has been hampered by the lack of structural information on this unusual DXPS, which contains an essential structural element specific to the *Plasmodium* genus. The overall structure of *PfDXPS* is similar to other DXPS, with a conserved active site sharing the same residues and position as many other homologues, emphasizing the potential of a broad-spectrum DXPS inhibitor.

The discovery of the N-terminal cap domain separates the DXPS from the *Plasmodium* genus into a distinct group within these enzymes. Although the cap domain may play no direct role in the catalytic mechanism, destabilization of the interface with Domain I causes a significant decrease in protein activity, opening space for the development of inhibitors targeting this interface.

Important questions still remain, such as where this domain came from and what pressures led to its existence. The decrease in protein stability when compared with the truncated LCR-I, the largest LCR which is unique in the *Plasmodium* genus, could be indicative of one of these changes, but unfortunately, lack of sequence information for intermediate species does not allow us to study the N-terminal cap domain evolution.

Inherent DXPS flexibility together with the presence of several long and disordered LCRs, makes *PfDXPS* harder to express and purify^{30,31}. Nevertheless, our results reveal the potential of cryo-EM for

solving *PfDXPS* structures at resolutions compatible with structure-based drug design approaches, showing the advantage of this technique for plasmodial targets that were previously not feasible by crystallography⁵³.

In summary, the results herein obtained shed light on structural differences of DXPS family members, but most importantly, open new roads through cryo-EM for the development of antimalarials using *PfDXPS*, an important but under-explored target.

Methods

Protein expression and purification

The pET-28a vector containing the *PfDXPS* coding sequence with deletion of the first 310 residues (*PfDXPS*-pET-28a) was kindly provided by Dr. David Merkler (University of South Florida). Details of cloning and construct design can be found elsewhere³⁰.

Escherichia coli BL21 Star™ (DE3) (cat number: C601003, Thermo Fisher Scientific) cells transformed with the *PfDXPS*-pET-28a plasmid were grown in Terrific Broth media at 37 °C at 180 RPM and induced with 1 mM of IPTG after reaching an OD₆₀₀ of 1.0. The cells were further grown for 24 h at 10 °C, 180 RPM and harvested by centrifugation. The cell pellet was resuspended in lysis buffer: 50 mM Tris pH 7.5, 500 mM NaCl, 20 mM imidazole, 5% glycerol, 5 mM β -mercaptoethanol, 1 mM MgCl₂, 50 μ g/mL lysozyme (cat number: HR7-110, Hampton Research) and 10 μ g/mL DNaseI (cat number: 10104159001, Roche) 1 tablet of cComplete™, Mini, EDTA-free protease inhibitor cocktail (cat number: 11836170001, Roche), lysed by sonication and centrifuged at 41000 $\times$ g to obtain the clarified lysate.

The supernatant was loaded onto a HisTrap High-Performance 5 mL column (Cytiva) equilibrated with buffer A (50 mM Tris 7.5, 500 mM NaCl, 5% glycerol). The column was washed with 10 column volumes (CV) of washing buffer (50 mM Tris 7.5, 500 mM NaCl, 50 mM imidazole, 5% glycerol), and the his-tagged protein eluted with 3 CV of buffer B (50 mM Tris 7.5, 500 mM NaCl, 250 mM imidazole, 5% glycerol). The eluted protein was diluted 5 times with buffer A and supplemented with 5 mM MgCl₂ and 1 mM thiamine diphosphate (ThDP). The protein was further purified with a size-exclusion Superdex 200 16/60 HiLoad column (Cytiva) pre-equilibrated with 20 mM HEPES pH 7.5, 150 mM NaCl, 5 mM MgCl₂, 5 mM β -mercaptoethanol. The protein-containing fractions were pooled, concentrated to 9 mg/mL, flash-frozen in liquid nitrogen and stored at -80 °C for further use. Sample purity was assessed by 8% SDS-PAGE.

For activity and stability assays, the proteins were expressed and purified following a similar protocol, excluding the size-exclusion step. After nickel affinity chromatography, the eluted protein was dialyzed overnight in 50 mM Tris pH 7.5, 500 mM NaCl, 5% glycerol, 5 mM β -mercaptoethanol, 2 mM MgCl₂, 30 μ M ThDP. Subsequently, the protein was concentrated flash-frozen in liquid nitrogen and stored at -80 °C for further use. Sample purity was assessed by 10% SDS-PAGE.

Site-directed mutagenesis and cloning

A construct of *PfDXPS* with the N-terminal cap domain removed was produced from the original plasmid *PfDXPS*-pET-28a using the Q5® Site-Directed Mutagenesis Kit (cat number: E0554S, New England Biolabs). The primers used for domain deletion were TTCCGCTGCTGAAAC, CATATGGCTGCCGC with an annealing temperature (T_a) of 61 °C. The final plasmid Δ *PfDXPS*-pET-28a had the sequence confirmed by Sanger sequencing, with the *PfDXPS* sequence from residues 417–1205. For point and triple mutations, the same protocol was used, however, with different primers and annealing temperatures. F387E (CAAAGAAGAAATCAACAACGGCG, ATGG CTTTTCAAATTCAAC, T_a : 57 °C), I320D (CTTCAACAGGATAACACCTTCATCAAC TATTTGCCGATATCGTAC, T_a : 57 °C), I341D (CTA-CAAAGAAGATTACGAAGCTGTACGTGG, ATTCATCACCGTAGATC, T_a : 57 °C), Y401A (CGATAACACCGCGTACAAGAAAGAAAACATCC, ATGTT GTTGATATAAACGC, T_a : 56 °C), N396A (CGTTTATATCGCGAA

CATCGATAACACC, CCGTTGTTGATAAATTCTTC, T_a : 56 °C), S522A (CATCTTCGAAGCGATCTACGATAAATTCGG, TTCAGGAAACCCGAA ATG, T_a : 56 °C).

A synthetic codon-optimized gene of *PfDXPS* (UniProtID: O96694) with a truncation of the first 310 residues and the 43 residues in the low-complexity region I (spanning residues 568–611) was purchased from Eurofins Genomics. The synthetic gene was cloned into the pET-M11 vector using the NcoI and HindIII restriction sites. The final plasmid *PfDXPS* Δ LCR-I-pET-M11 had the sequence confirmed by Sanger sequencing.

Grid preparation

The protein was diluted with size-exclusion buffer to a final concentration of 0.9 mg/mL and applied onto copper Quantifoil 1.2/1.3 grids with a mesh size of 300 at 100% humidity and temperature of 4 °C. Afterwards the grid was plunge frozen using Vitrobot (Thermo Fisher Scientific) with the following parameters: blot force of 0, blotting time of 6 sec, a wait time of 10 sec.

Cryo-EM data collection

Data were collected at the Astbury Biostructure Laboratory EM facility at the University of Leeds using a Titan Krios operated at 300 kV with slit width of 10 eV. Images were collected with a Falcon4i detector operating in counting mode, at a magnification of 165,000 corresponding to a pixel size of 0.74 Å. The dose rate was set to 6.6 e⁻ per pixel per second, and movies were recorded over 3.69 sec fractionated into 44 frames, resulting in a total dose of 44.47 e⁻ per Å². Data were collected using EPU software (Thermo Fisher Scientific) with a defocus range -0.8 μ m to -2.4 μ m. A total of 2070 micrographs were collected in one session.

Cryo-EM data processing

For high-resolution structure determination, the 2070 movies collected were aligned using MotionCor2. Then, micrographs were imported into cryoSPARC v4.1.2⁵⁴, where CTF was corrected using Patch CTF Estimation. An initial set of particles was obtained using Blob picker, extracted with 256 pixels boxes binned 4 times and went to multiple rounds of 2D classification. Then, 73,756 filtered particles were chosen to train a model in DeepPicker, which was used for picking 352,935 particles. These particles were extracted with a 256 pixels box two times binned and were submitted to two rounds of 2D classification with $N = 100$, resulting in a selection of 197,265 particles. These were submitted to one round of heterogeneous classification, and two classes containing a total of 97,205 particles were selected for final refinement. These particles were re-extracted with 400-pixel boxes and submitted to the Homogeneous Refinement algorithm with 1 extra final pass, and C2 applied symmetry, resulting in a refined map with gold-standard Fourier shell correlation (gsFSC) of 2.68 Å. The resulting map was used for the non-uniform refinement with 1 extra final pass, 7 Å initial lowpass resolution, C2 applied symmetry, per-particle defocus optimization, and per-group CTF optimization, resulting in a refined map with gsFSC of 2.52 Å. The resulting map was then refined using Local Refinement algorithm with C2 imposed symmetry, resulting in a map with gsFSC of 2.42 Å. Map resolution was calculated using cryoSPARC Local Resolution. Finally, deepEMhancer v0.14⁵⁵ was used to sharpen the map in the post-processing. Statistics of data collection, processing and model refinement are available in Supplementary Table 1. Schematic of the data processing steps is available in Supplementary Fig. 1.

Modelling, refinement, and analysis

An AlphaFold model of the *PfDXPS* monomer (UniProt: O96694) was used as the starting point for the modelling and refinement. Initially the model was manually treated by deleting the first 310 amino acids as well as the long low-complexity regions with no ordered structure. The

model was then docked into the map using the Dock in map program in the Phenix 1.20.1-4487 suite, searching for 2 subunits with symmetry set to C2. The docked model was further refined with Coot v0.9.6⁵⁶ and Phenix real space refine⁵⁷.

Enzymatic assays

*Pf*DXPS activity was measured through a DXPS-DXR coupled assay previously described, using *E. coli* DXR (*Ec*DXR) as the coupling enzyme⁵⁸. *Ec*DXR expression and purification methods are described in the Supplementary Methods. The reaction was prepared with a final concentration of 300 μ M DL-glyceraldehyde-3-phosphate, 500 μ M pyruvate, 300 μ M thiamine diphosphate, 2 μ M *Ec*DXR and 100 μ M NADPH, in an assay buffer composed of 50 mM HEPES pH 7.5, 100 mM NaCl, 1.5 mM MnCl₂. A reaction with all substrates and coupled enzymes but lacking *Pf*DXPS was used as a blank. Enzyme concentrations for each experiment are shown in the associated Supplementary information and figure legends.

Activity was monitored through the decrease in fluorescence of NADPH oxidation using a plate reader SYNERGY H1 (BioTEK) at wavelengths 340/450 nm (excitation/emission) for 1 hour at 25 °C. Initial velocity was calculated through the slope of the linear region using Microsoft Office Excel 2016 Professional Plus. Activity of the mutants were compared with the wild-type enzyme by calculating the relative activity through Eq. (1).

$$\text{Relative activity} = 100 \times \left(\frac{\text{Mutant Initial velocity}}{\text{Wild type initial velocity}} \right) \quad (1)$$

All experiments were performed in triplicate, final results are shown as the average with the related sample standard deviation calculated through Microsoft Office Excel 2016 Professional Plus.

Thermal shift assays

Thermal shift analysis was performed to measure changes in the thermal denaturation temperature of *Pf*DXPS and mutants. The final concentrations in the reaction were 5 μ M *Pf*DXPS, 5x SYPRO Orange in a buffer composed of 50 mM HEPES pH 7.5, 100 mM NaCl, 1.5 mM MnCl₂, 300 μ M ThDP. A 96-well PCR plate was used, and protein denaturation was achieved by using a continuous heating rate of 1 °C/min from 20 °C to 95 °C using an CFX96 Real-time system (BIO-RAD). Changes in fluorescence of the SYPRO Orange dye were measured every 30 seconds with the HEX filter. All experiments were done in quintuplicate.

The intensity data for each reaction was normalized and averaged, and the deviation was calculated through sample standard deviation using Microsoft Office Excel 2016 Professional Plus. The normalized data were fitted onto a Boltzmann model using OriginPro 9, with the standard deviation as instrumental errors. The melting temperature was obtained through the inflexion point coefficient of the fitted curve.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The collected micrographs generated in this study are available at the Electron Microscopy Public Image Archive (EMPIAR) under accession code [11763](#). The cryo-EM maps and structural models generated in this study are available at the Protein Data Bank (PDB) under accession code [8R2H](#) and at the Electron microscopy Data Bank (EMDB) under accession code [EMD-18842](#). Source data are provided in this paper.

References

1. Organisation, W. H. *World Malaria Report 2021*. World Health Organization (2021).

2. Cibulskis, R. E. et al. Malaria: global progress 2000 – 2015 and future challenges. *Infect. Dis. Poverty* **5**, 61 (2016).
3. Dondorp, A. M. et al. Artemisinin resistance in *Plasmodium falciparum* malaria. *N. Engl. J. Med.* **361**, 455–467 (2009).
4. Samarasekera, U. Climate change and malaria: predictions becoming reality. *Lancet* **402**, 361–362 (2023).
5. Fairhurst, R. M. & Dondorp, A. M. Artemisinin-resistant *Plasmodium falciparum* malaria. *Emerg. Infect.* **10**, 409–429 (2016).
6. Alaithan, H., Kumar, N., Islam, M. Z., Liappis, A. P. & Nava, V. E. Novel therapeutics for malaria. *Pharmaceutics* **15**, 1800 (2023).
7. Hanboonkunupakarn, B. & White, N. J. Advances and roadblocks in the treatment of malaria. *Br. J. Clin. Pharmacol.* **88**, 374–382 (2022).
8. Takala-Harrison, S. et al. Independent emergence of artemisinin resistance mutations among *Plasmodium falciparum* in Southeast Asia. *J. Infect. Dis.* **211**, 670–679 (2015).
9. Boullé, M. et al. Artemisinin-Resistant *Plasmodium falciparum* K13 Mutant Alleles, Thailand-Myanmar Border. *Emerg. Infect. Dis.* **22**, 1503–1505 (2016).
10. Miotto, O. et al. Emergence of artemisinin-resistant *Plasmodium falciparum* with kelch13 C580Y mutations on the island of New Guinea. *PLoS Pathog.* **16**, e1009133 (2020).
11. Mathieu, L. C. et al. Local emergence in Amazonia of *Plasmodium falciparum* k13 C580Y mutants associated with in vitro artemisinin resistance. *Elife* **9**, e51015 (2020).
12. Uwimana, A. et al. Emergence and clonal expansion of in vitro artemisinin-resistant *Plasmodium falciparum* kelch13 R561H mutant parasites in Rwanda. *Nat. Med.* **26**, 1602–1608 (2020).
13. Balikagala, B. et al. Evidence of artemisinin-resistant malaria in Africa. *N. Engl. J. Med.* **385**, 1163–1171 (2021).
14. Haldar, K., Bhattacharjee, S. & Safeukui, I. Drug resistance in *Plasmodium*. *Nat. Rev. Microbiol.* **16**, 156–170 (2018).
15. van der Pluijm, R. W. et al. Determinants of dihydroartemisinin-piperaquine treatment failure in *Plasmodium falciparum* malaria in Cambodia, Thailand, and Vietnam: a prospective clinical, pharmacological, and genetic study. *Lancet Infect. Dis.* **19**, 952–961 (2019).
16. Na-Bangchang, K., Ruengweeraayut, R., Mahamad, P., Ruengweeraayut, K. & Chaijaroenkul, W. Declining in efficacy of a three-day combination regimen of mefloquine-artesunate in a multi-drug resistance area along the Thai-Myanmar border. *Malar. J.* **9**, 273 (2010).
17. El-Moamly, A. A. & El-Sweify, M. A. Malaria vaccines: the 60-year journey of hope and final success-lessons learned and future prospects. *Trop. Med. Health* **51**, 29 (2023).
18. RTS, S Clinical Trials Partnership. Efficacy and safety of RTS, S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: final results of a phase 3, individually randomised, controlled trial. *Lancet (London, England)* **386**, 31–45 (2015).
19. Datto, M. S. et al. Efficacy and immunogenicity of R21/Matrix-M vaccine against clinical malaria after 2 years' follow-up in children in Burkina Faso: a phase 1/2b randomised controlled trial. *Lancet Infect. Dis.* **22**, 1728–1736 (2022).
20. Organisation, W. H. WHO recommends R21/Matrix-M vaccine for malaria prevention in updated advice on immunization. <https://www.who.int/news/item/02-10-2023-who-recommends-r21-matrix-m-vaccine-for-malaria-prevention-in-updated-advice-on-immunization> (2023).
21. Lim, L. & McFadden, G. I. The evolution, metabolism and functions of the apicoplast. *Philos. Trans. R. Soc. B Biol. Sci.* **365**, 749–763 (2010).
22. Sato, S. The apicomplexan plastid and its evolution. *Cell. Mol. Life Sci.* **68**, 1285–1296 (2011).
23. Ralph, S. A. et al. Metabolic maps and functions of the *Plasmodium falciparum* apicoplast. *Nat. Rev. Microbiol.* **2**, 203–216 (2004).

24. Guggisberg, A. M., Amthor, R. E. & Odom, A. R. Isoprenoid biosynthesis in *Plasmodium falciparum*. *Eukaryot. Cell* **13**, 1348–1359 (2014).
25. Jomaa, H. et al. Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. *Science* **285**, 1573–1576 (1999).
26. Cassera, M. B. et al. The methylerythritol phosphate pathway is functionally active in all intraerythrocytic stages of *Plasmodium falciparum*. *J. Biol. Chem.* **279**, 51749–51759 (2004).
27. Estévez, J. M., Cantero, A., Reindl, A., Reichler, S. & León, P. 1-deoxy-D-xylulose-5-phosphate synthase, a limiting enzyme for plastidic isoprenoid biosynthesis in plants. *J. Biol. Chem.* **276**, 22901–22909 (2001).
28. Lois, L. M. et al. Cloning and characterization of a gene from *Escherichia coli* encoding a transketolase-like enzyme that catalyzes the synthesis of D-1-deoxyxylulose 5-phosphate, a common precursor for isoprenoid, thiamin, and pyridoxol biosynthesis. *Proc. Natl Acad. Sci. USA* **95**, 2105–2110 (1998).
29. Masini, T. & Hirsch, A. K. H. Development of inhibitors of the 2C-methyl-D-erythritol 4-phosphate (MEP) pathway enzymes as potential anti-infective agents. *J. Med. Chem.* **57**, 9740–9763 (2014).
30. Battistini, M. R., Shoji, C., Handa, S., Breydo, L. & Merkler, D. J. Mechanistic binding insights for 1-deoxy-D-Xylulose-5-Phosphate synthase, the enzyme catalyzing the first reaction of isoprenoid biosynthesis in the malaria-causing protists, *Plasmodium falciparum* and *Plasmodium vivax*. *Protein Expr. Purif.* **120**, 16–27 (2016).
31. Handa, S. et al. Production of recombinant 1-deoxy-d-xylulose-5-phosphate synthase from *Plasmodium vivax* in *Escherichia coli*. *FEBS Open Bio.* **3**, 124–129 (2013).
32. Gierse, R. M. et al. Identification of a 1-deoxy-D-xylulose-5-phosphate synthase (DXS) mutant with improved crystallographic properties. *Biochem. Biophys. Res. Commun.* **539**, 42–47 (2021).
33. Gierse, R. M. et al. First crystal structures of 1-deoxy-D-xylulose 5-phosphate synthase (DXPS) from *Mycobacterium tuberculosis* indicate a distinct mechanism of intermediate stabilization. *Sci. Rep.* **12**, 7221 (2022).
34. Hamid, R. et al. 1-deoxy-D-xylulose-5-phosphate synthase from *Pseudomonas aeruginosa* and *Klebsiella pneumoniae* reveals conformational changes upon cofactor binding. *J. Biol. Chem.* **299**, 105152 (2023).
35. Xiang, S., Usunow, G., Lange, G., Busch, M. & Tong, L. Crystal structure of 1-deoxy-D-xylulose 5-phosphate synthase, a crucial enzyme for isoprenoids biosynthesis. *J. Biol. Chem.* **282**, 2676–2682 (2007).
36. DePristo, M. A., Zilversmit, M. M. & Hartl, D. L. On the abundance, amino acid composition, and evolutionary dynamics of low-complexity regions in proteins. *Gene* **378**, 19–30 (2006).
37. Zilversmit, M. M. et al. Low-complexity regions in *plasmodium falciparum*: missing links in the evolution of an extreme genome. *Mol. Biol. Evol.* **27**, 2198–2209 (2010).
38. Chen, P. Y. T., DeColli, A. A., Freel Meyers, C. L. & Drennan, C. L. X-ray crystallography-based structural elucidation of enzyme-bound intermediates along the 1-deoxy-D-xylulose 5-phosphate synthase reaction coordinate. *J. Biol. Chem.* **294**, 12405–12414 (2019).
39. Zhou, J. et al. Conformational dynamics of 1-deoxy-d-xylulose 5-phosphate synthase on ligand binding revealed by H/D exchange MS. *Proc. Natl Acad. Sci.* **114**, 9355–9360 (2017).
40. Kemp, D. J., Coppel, R. L. & Anders, R. F. Repetitive proteins and genes of malaria. *Annu. Rev. Microbiol.* **41**, 181–208 (1987).
41. Hughes, A. L. The evolution of amino acid repeat arrays in *Plasmodium* and other organisms. *J. Mol. Evol.* **59**, 528–535 (2004).
42. Xue, H. Y. & Forsdyke, D. R. Low-complexity segments in *Plasmodium falciparum* proteins are primarily nucleic acid level adaptations. *Mol. Biochem. Parasitol.* **128**, 21–32 (2003).
43. Frugier, M. et al. Low complexity regions behave as tRNA sponges to help co-translational folding of plasmodial proteins. *FEBS Lett.* **584**, 448–454 (2010).
44. van Kempen, M. et al. Fast and accurate protein structure search with Foldseek. *Nat. Biotechnol.* <https://doi.org/10.1038/s41587-023-01773-0> (2023).
45. Zhang, C., Freddolino, P. L. & Zhang, Y. COFACTOR: improved protein function prediction by combining structure, sequence and protein-protein interaction information. *Nucleic Acids Res.* **45**, W291–W299 (2017).
46. Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480 (1995).
47. Krissinel, E. & Henrick, K. Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* **372**, 774–797 (2007).
48. Zhou, Y., Pan, Q., Pires, D. E. V., Rodrigues, C. H. M. & Ascher, D. B. DDMut: predicting effects of mutations on protein stability using deep learning. *Nucleic Acids Res.* **51**, W122–W128 (2023).
49. Handa, S. et al. Mechanistic studies of 1-deoxy-D-xylulose-5-phosphate synthase from *deinococcus radiodurans*. *Biochem. Mol. Biol. J.* **04**, 1–11 (2018).
50. Brammer Basta, L. A., Patel, H., Kakalis, L., Jordan, F. & Freel Meyers, C. L. Defining critical residues for substrate binding to 1-deoxy- d -xylulose 5-phosphate synthase - active site substitutions stabilize the predecarboxylation intermediate C2a-lactylthiamin diphosphate. *FEBS J.* **281**, 2820–2837 (2014).
51. Decolli, A. A., Zhang, X., Heflin, K. L., Jordan, F. & Freel Meyers, C. L. Active site histidines link conformational dynamics with catalysis on anti-infective target 1-deoxy- d -xylulose 5-phosphate synthase. *Biochemistry* **58**, 4970–4982 (2019).
52. Gierse, R. M. et al. DXS as a target for structure-based drug design. *Future Med. Chem.* **7**, 2131–2141 (2017).
53. Anton, L., Cobb, D. W. & Ho, C.-M. Structural parasitology of the malaria parasite *Plasmodium falciparum*. *Trends Biochem. Sci.* **47**, 149–159 (2022).
54. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
55. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
56. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
57. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D. Struct. Biol.* **74**, 531–544 (2018).
58. Humnabadkar, V., Jha, R. K., Ghatnekar, N. & De Sousa, S. M. A high-throughput screening assay for simultaneous selection of inhibitors of *Mycobacterium tuberculosis* 1-deoxy-D-xylulose-5-phosphate synthase (Dxs) or 1-deoxy-D-xylulose 5-phosphate reductoisomerase (Dxr). *J. Biomol. Screen.* **16**, 303–312 (2011).

Acknowledgements

We would like to thank Dr. David Merkler (University of South Florida) for kindly providing the *PfDXPS* plasmid. The cryo-EM data were collected at the Astbury Biostructure Laboratory cryo-EM facilities during the cryo-EM training course, which was supported by a Wellcome/MRC funding program (218785/Z/19/Z). V.O.G. was funded by the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement no. 860816.

Author contributions

V.O.G. and L.A.T.W. performed the cloning, expression, and protein purification. V.O.G. and A.S.G. performed the cryo-EM data collection and processing. V.O.G., R.O., and A.S.G. built the atomic model. V.O.G. performed the mutation studies. V.O.G. and R.O. analysed the data and interpreted the results. A.K.H.H. and M.R.G. supervised the work. All authors jointly wrote the manuscript and approved its final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-024-50671-9>.

Correspondence and requests for materials should be addressed to Matthew R. Groves.

Peer review information *Nature Communications* thanks Babatunde Ekundayo and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024