Conserved cysteine residues in malaria chorismate synthase indicate their important role in protein structure and function

Harvinder Kour Khera1,2, Susheel Kumar Singh1, GN Bhat1 & Subhash Singh1,2*

1Pharmacology Division, CSIR-Indian Institute of Integrative Medicine, Canal Road, Jammu-180 001, Jammu & Kashmir, India
2Academy of Scientific & Innovative Research (AcSIR), New Delhi, India

Received 12 July 2016; revised 07 October 2016

Predicted protein sequences of Chorismate synthase (Cs) from different Plasmodium species have shown high number of conserved cysteine residues when compared to the predicted Cs protein sequences from other bacterial, fungal or plant species. To better understand the structure and function of malaria Cs, we have cloned, expressed and purified recombinant Cs from Plasmodium falciparum (rPfCs) in E.coli. The rPfCs exhibited disulfide linkages as indicated by mobility shifts observed for this protein when compared between non-reduced and reduced-alkylated conditions on SDS-PAGE. Antibodies generated against rPfCs also detected similar mobility shift for the native parasite protein from asexual blood stage culture indicating that the conserved cysteine residues in native parasite Cs play an important role in the protein structure and function.

Keywords: Chorismate synthase, Cysteines, Disulfide linkages, Malaria, Plasmodium falciparum

Malaria continues to be a major global health issue. Prevalence of resistance against the existing antimalarial drugs1, there is a need for developing of new effective antimalarials2. Metabolic pathways unique to the pathogen (which do not occur in the host) have obvious potential advantages for specific chemotherapeutic intervention. Shikimate pathway is one such pathway identified in the malarial parasite3, which the parasite likely inherits from its algal origin. Its absence in vertebrate host makes it an attractive target for antimalarial drugs4. Shikimate pathway consists of seven steps, with Chorismate synthase (Cs, EC 4.2.3.5) catalyzing the final step of this pathway viz., conversion of 5-enolpyruvylshikimate-3-phosphate (EPSP) to Chorismate. Cs catalyzes a unique biochemical reaction involving anti-1,4 elimination of 3-phosphate group and the C-(6proR) hydrogen from EPSP. Chorismate is an important metabolite and a common precursor for aromatic amino acid, ubiquinone and folate biosynthesis5.

Shikimate pathway, in general and Cs in particular, plays a major role in maintaining the virulence of several bacterial pathogens. Genetic mutations affecting this pathway such as that of aroC (Cs homolog in bacteria) deletions have resulted in loss of virulence in different bacterial pathogens such as Brucella suis6, Salmonella enterica serovar Typhimurium7 and Yersinia enterocolitica8. The Shikimate pathway in malaria has been shown to be druggable through the use of glyphosate, herbicide, which inhibits the penultimate step of this pathway catalyzed by the enzyme EPSP synthase9. Silencing of the Cs gene through dsRNA in Plasmodium falciparum has also been reported to inhibit the parasite growth10.

Though the Shikimate pathway is a potential target for chemotherapeutic intervention in the treatment of malaria, molecular understanding of this pathway is incomplete11. Apart from Cs, only a few other Shikimate pathway genes in malaria genome have been identified based on sequence, though biochemical activity has been detected for almost all the enzymes of this pathway12. Molecular characterization of parasite Cs has reported

*Correspondence:
Phone : +91-191-2585006/(07-18); Ext: 355
E-mail: subhash@iiim.res.in; subhash@iiim.ac.in

Abbreviations: BCPIP, 5-Bromo-4-chloro-3'-indolylphosphate; Cs, Chorismate synthase; DHQase, 3-Dehydroquinate dehydratase; dsRNA, double stranded RNA; EPSP, 5-enolpyruvyl-shikimate; GST, Glutathione S-transferase; HRP, Horse radish peroxidase; IAA, Iodoacetamide; IPTG, Isopropyl β-D-thiogalactopyranoside; NBT, Nitro-blue tetrazolium; Ni-NTA, Nickel-Nitrilotriacetic acid; NR, Non-reduced; OD, Optical density; PABA, p-Amino benzoic acid; PfCs, Chorismate synthase of Plasmodium falciparum; RA, Reduced-alkylated; ROPT, Reversibly oxidized protein thiol; SDH, Shikimate dehydrogenase; SK, Shikimate kinase; TBST, Tris-based saline Tween-20 (0.05%)
conflicting results, based on which it is not clear whether malaria Cs protein is monofunctional or bifunctional\textsuperscript{13,14}. Even homology based structural modelling of PfCs derived from observed Cs structures in bacterial Cs\textsuperscript{15} or yeast Cs\textsuperscript{16} proteins have not provided clarity in this regard.

In this study, we tried to clone and express recombinant PfCs His-tagged protein, as a tool for better molecular and functional understanding of this enzyme.

Materials and Methods

All reagents used were of analytical grade and from Sigma-Aldrich, India unless otherwise indicated.

Sequence alignment

Sequences retrieved from National Centre for Biotechnology Information (NCBI) database (http://www.ncbi.nlm.nih.gov/protein/) were aligned using ClustalW with default parameters\textsuperscript{17}. The graphical enhancement of the aligned Cs sequences was done using ESPript 2.2 server\textsuperscript{18}, and binding sites/active site residues are as reported by Tapas et al\textsuperscript{15}.

Cloning, expression, and purification of rPfCs

C-terminal His-tagged and codon optimized PfCs gene was synthesized (Geneart, Germany) and cloned in the expression vector pET-28(a) under the control of IPTG inducer and used to transform E. coli BL21 (DE3). Transformants were selected by kanamycin resistance. Expression of stable and soluble rPfCs was obtained upon induction with 100 µM IPTG at OD 0.6 for 22 h at 18°C in the presence of 2% MgCl\textsubscript{2}. The induced bacterial pellets were lysed by sonication in 50 mM Tris and 300 mM NaCl in the presence of protease inhibitor cocktail without EDTA (Roche Diagnostic, Germany). rPfCs purification was done using Ni-NTA (Qiagen, Germany) column with increasing imidazole concentrations.

Generation of anti-rPfCs antibodies

Purified rPfCs was used to generate polyclonal antibodies in New Zealand White rabbit. 100 µg rPfCs was used for each immunization. Primary immunization was done using Freund's complete adjuvant. Four subsequent booster doses were given on days 21, 41, 63 and 90 with Freund's incomplete adjuvant. Serum was collected 2 weeks after the last boost. The pre- and post-immunization sera were checked for antibody production of anti-PfCs antibodies by Western Blot using purified rPfCs. Non-specific reactivity of the sera from PfCs immunized animals was ruled out by confirming lack of the reactivity against BSA, casein, and other His-tag malaria recombinant proteins. Anti-PfCs antibodies were affinity purified against rPfCs using Aminolink Purification Kit (Thermo Fisher Scientific, USA) as per instructions of the manufacturer and used in subsequent experiments.

Preparation of parasite extract

Asexual blood stage of \textit{P. falciparum} (3D7 strain) was cultured as per standard protocol\textsuperscript{19} and harvested by centrifugation at 5% parasitemia consisting of mainly trophozoite stage. Synchronization of culture was achieved using sorbitol treatment\textsuperscript{20}. The harvested culture was lysed using saponin\textsuperscript{21} in lysis buffer containing 50 mM Tris, 0.5% NP-40 and 100 mM NaCl, and homogenised using a Heidolph RZR 2051, USA, a homogenizer for 10 min on ice. The soluble parasite extract was collected after centrifugation at 100000g for 20 min at 4°C and stored at –80°C in the presence of protease inhibitor cocktail (Roche Diagnostic, Germany).

Western blot and Gel shift assay

About 1 µg of purified rPfCs or native parasite extract (from \(3 \times 10^6\) asexual blood stage trophozoites) was loaded per well of a standard Tris-Glycine based 10% SDS-PAGE gel\textsuperscript{22} under NR or RA conditions. For preparation of RA sample, protein samples (rPfCs or parasite extract) were mixed with Dithiothreitol (DTT) containing SDS-loading buffer (2X) and incubated at 95°C for 10 min followed by addition of 1/5 volume of freshly made 1 M Iodoacetamide (IAA). The sample was vortexed and spun briefly. For the preparation of NR sample, protein sample was mixed with SDS-loading buffer (2X) without DTT, and instead of IAA, 1/5 volume of Type-I water (American Society for Testing and Materials) was added.

Western Blot was performed as per standard protocol\textsuperscript{21}. Antibodies used for detection of Western Blots are as follows: primary anti-His-tag mouse antibody (Cell Signalling Technology, USA) at a dilution of 1:2500 in TBST; affinity purified anti-rPfCs primary antibody raised in rabbit, at a dilution of 1:10; secondary anti-mouse AP-linked IgG (Cell Signalling Technology, USA) at a dilution of 1:3500; and secondary anti-rabbit HRP-linked IgG at a
dilution of 1:2000. All antibody dilutions were made in TBST with 1.25% skimmed milk. Blots were either developed using nitro-blue tetrazolium (NBT)/5-bromo-4-chloro-3'-indolyphosphate (BCPIP) (Biorad Laboratories, USA) for detecting AP activity or Immobilon substrate (Millipore, Germany) for detecting HRP activity.

For gel shift assay, rPfCs or parasite extract was resolved on a 10% SDS-PAGE gel under RA and NR conditions separated by at least one empty lane. Mobility shift of PfCs between RA and NR conditions was compared on Western blot using anti-His antibody or affinity purified anti-rPfCs antibody.

Results

Malaria Cs sequences show high number of conserved cysteine residues

Sequence alignment of the predicted amino acid sequences for the Cs protein from different Plasmodium sp. is shown in Fig. 1. The aligned amino-acid sequences have been numbered as per the predicted PfCs sequence from the 3D7 strain (NCBI ref No. AAB63293.1). High level of amino acid sequence homology (> 60% identity) is observed when comparing the PfCs sequence with Cs from other Plasmodium sp. indicating that Cs is likely to play a significant role in the overall parasite metabolism. We found that the Cs sequences from the apicomplexan parasites contain higher number of cysteine residues (12 or more cysteine residues) when compared with the Cs homologs in randomly chosen organisms from different taxa (bacterial Cs: 0-5 cysteines; fungal Cs: 5-8 cysteines and plant Cs: 7-9 cysteines), as shown in Table 1.

Overall, Cs proteins from Plasmodium sp. share <50% identity with the Cs homologs from plant, bacteria, and fungi. Cs protein from Plasmodium sp. when compared to other protozoan species (Toxoplasma gondii, Phaeodactylum tricornutum), bacteria (Escherichia coli), plant (Arabidopsis thaliana) and fungi (Saccharomyces cerevisiae), only show two conserved cysteines residues C23 and C275. Within the Plasmodium sp. the predicted Cs sequences show additional fully conserved cysteines residues at positions C285, C304, C437, C475 and C487; and partial conservation at residues C149, C239, C287, C380, C421 and C486 in at least three or more Plasmodium sp.

Cloning, expression, and purification of rPfCs

The predicted open reading frame of PfCs sequence comprises of a single exon, which was codon optimized to encode a C-terminal hexa-histidine tagged rPfCs protein. Using standard protocol of induction at OD 0.6 with 1 mM IPTG at 37°C for 3 h resulted in expression of insoluble rPfCs mostly associated with inclusion bodies (Fig 2A). Optimization of soluble expression of rPfCs was achieved in the presence of 2% MgCl₂. With 100 µM IPTG for 22 h post induction at 18°C. Presence of MgCl₂ was found to enhance the expression of the soluble rPfCs protein (Fig. 2B). The rPfCs protein was purified using Ni-NTA and increasing concentration of imidazole (Fig. 2C). Confirmation of rPfCs protein was done by Western Blot using anti-His antibody (Fig. 2D). The observed molecular weight of the rPfCs protein (60 kDa) on SDS-PAGE is in good conformation with its predicted polypeptide mass.

rPfCS protein shows presence of intra-molecular disulfide linkages

Intrigued by the presence of 12 cysteine residues in the PfCs sequence we checked for the presence of disulfide linkages in the rPfCs. We performed gel shift assay for comparing mobility-shift of the rPfCs protein on SDS-PAGE under non-reduced and reduced-alkylated conditions. Fig. 3A shows Western blot analysis of rPfCS using anti-His tag antibodies. A shift in the mobility of monomeric rPfCs protein was observed when compared between non-reducing and reduced, alkylated state on SDS-PAGE. rPfCs displayed a higher apparent molecular mass under reduced-alkylated condition due to complete denaturation leading to a complete linearization of the polypeptide chain and addition of alkylating groups (around 57 Da per cysteine residue). The faster mobility of the monomeric rPfCs protein was observed when compared between non-reducing and reduced-alkylated conditions on SDS-PAGE. rPfCs displayed a higher apparent molecular mass under reduced-alkylated condition due to complete denaturation leading to a complete linearization of the polypeptide chain and addition of alkylating groups (around 57 Da per cysteine residue). The faster mobility of the monomeric protein band in non-reducing conditions as compared to the same in reduced, alkylated state, suggest changes in the hydrodynamic radius of the protein. Under non-reducing conditions on SDS-PAGE, rPfCs protein is possibly in a more compact conformation owing to the presence of the disulfide linkages. Thus, results from the gel shift assay indicate the presence of intra-molecular disulfide linkages in rPfCs protein.
Fig. 1 — Multiple sequence alignment of predicted Cs protein sequences from different *Plasmodium* sp. (indicated by red alphabets) as compared to Cs homologs from other organisms (indicated by black alphabets) as retrieved from the NCBI database. [The respective accession numbers are as follows, f: Cs of *P. falciparum* (AAB63293.1), r: Cs of *P. reichenowi* (XP_012762041.1), v: Cs of *P. vivax* (EDL46514.1), k: Cs of *P. knowlesi* strain H (XP_002261447.1), b: Cs of *P. berghei* strain ANKA (XP_678920.1), y: Cs of *P. yoelii* 17XNL (XP_724321.1), c: Cs of *P. chabaudi* (XP_743671.1), t: Cs of *Toxoplasma gondii* (AAB52422.1), p: Cs of *Phaeodactylum tricornutum* CCAP 1055/1 (XP_002177933.1), a: Cs of *Arabidopsis thaliana* (NP_001031158.1), e: Cs of *Escherichia coli* F11 (EDV69285.1), s: Cs of *Saccharomyces cerevisiae* (CAA42745.1). Identical residues and similar amino acids are highlighted with red and yellow colour background, respectively. The three conserved signatory sequences of Cs are shown in boxes. The red triangles under the aligned sequences indicate the FMN interacting active site residue; pink stars indicate the residues involved in the formation of FMN binding pocket, squares in blue colour indicates the residues interacting with EPSP and brown circle indicates the EPSP binding site residue. Cysteines are marked in green, conserved cysteines (when found to be aligned with 3 or more Cs sequences) are marked with green stars above the aligned sequences. Amino acid numbering is with respect to *P. falciparum* sequence]
Native P. falciparum Cs protein also shows presence of intra-molecular disulfide linkages

Anti-PfCs antibodies were affinity purified against rPfCs protein and used for detection of native PfCs protein from asexual blood stage culture in gel shift assay. Western blot analysis of both rPfCs (Fig. 3B) and native parasite protein (Fig. 3C) resolved on SDS-PAGE under non-reducing and reduced-alkylated conditions was performed using purified anti-rPfCs antibodies. As seen in Fig. 3C, native PfCs monomeric protein also displays mobility shift on SDS-PAGE when compared under non-reducing and reduced-alkylated conditions similar to rPfCs (Fig. 3B). This observation strongly indicates the presence of intra-molecular disulfide linkages in the native PfCs protein expressed in asexual
blood stage development of the parasite. However, the exact positions of these disulfide bonds need to be mapped for further confirmation.

Additionally, a faster migrating PfCs band (possibly a partially degraded or cleaved form) is also observed which exhibits similar mobility shift between reduced vs non-reduced states suggesting that this degraded PfCs band also has some disulfide linkages.

**Discussion**

The enzymes and metabolic intermediates of Shikimate pathway appear to be universal for all organisms harbouring this pathway. However, phylogenetic divergences for this pathway have been observed. Shikimate pathway is present in plastid in green plants while it is cytosolic in fungi and bacteriaootnote{24}. In prokaryotes, all seven enzymes of this pathway are monofunctional each encoded by a separate specific gene. In plants, 3-Dehydroquinate dehydratase (DHQase) and Shikimate dehydrogenase (SDH) are fused to form a bifunctional enzyme complex. In fungi and certain apicomplexa (T. gondii) DHQS, 5-Enolpyruvyl-shikimate synthase (EPSPS), Shikimate kinase (SK), SDH and DHQase are fused to form a penta-functional AROM complexootnote{12}. Though a molecular understanding of Shikimate pathway in malaria parasite remains to elucidated, it has been shown to be druggable through inhibitory effects of glyphosate (a herbicide)ootnote{9}. Malaria Cs is the most readily identified gene based on its sequence similarity with homologs in other organismsootnote{11}. It’s functional characterization, reported by different groups, has, however, provided somewhat conflicting viewsootnote{13,14}.

The Cs enzyme shows divergent evolution. In most of the plants and fungi, Cs is known to itself possess NADP-driven flavin reductase activity and thus, termed as “bifunctional”, whereas, Cs from most of the prokaryotes lack this flavin reductase activity and need to use reduced flavin from the cellular environment and thus termed as “monofunctional”ootnote{12}. Cs from *Mycobacterium tuberculosis* seems to be an exception, as it has been demonstrated to be bifunctionalootnote{25}.

There seems to be a similarity in genomic organization of the Shikimate pathway genes between the fungi and apicomplexan parasites (as observed in *Toxoplasma gondii*, which is closely related to malaria parasite) such that both have a penta-functional *AROM* complex. Based on this analogy, Cs from apicomplexan parasites could be expected to be bifunctional. This was confirmed for Cs from *T. gondii* by Ehammer _et al_ using a cellular screen based on a Cs deficient *Saccharomyces cerevisiae* strain auxotrophic for aromatic amino acids where prototrophy was restored through complementation with Cs transgene from different organisms. In this screen, while expression of bifunctional Cs transgenes restored prototrophy on their own (since _S. cerevisiae_ Cs is bifunctional), monofunctional Cs (from bacteria) required additional co-complementation with oxidoreductase. However, in this study PfCs expression (though confirmed at mRNA level) did not produce the active enzyme in *S. cerevisiae* either in presence or absence of co-complementation with oxidoreductase and therefore, this study could not conclude whether PfCs was actually a bifunctional or monofunctional enzyme.

In another study, Fitzpatrick _et al_ reported PfCs to be monofunctional. Their study was based on rPfCs expressed in *E. coli* as an N-terminal GST-fusion protein, which had no detectable enzyme activity. However, thrombin mediated cleavage of the GST-domain yielded an unstable rPfCs protein with very low enzyme activity as compared to the ones of the *E. coli* and *N. crassa* Cs enzymes. Therefore, the observation of Fitzpatrick _et al._ that PfCs is monofunctional needs to be confirmed using a more stable protein preparation.

In our study sequence alignment of the predicted Cs protein sequences from different Plasmodium sp. showed conserved cysteines as a striking feature. The presence of a noticeable high number of conserved cysteines in Cs from apicomplexan parasites indicates their importance in the structural-functional role. Cysteine is often the least abundant amino acid, but when present, cysteines are generally highly conserved and understood to play a critical role at reactivity protein cysteine thiols may broadly be categorised into four types 1) permanent structural disulfide bonds which are typically observed in oxidizing environments, 2) thiols that coordinate with metals, typically iron, zinc or copper 3) those that are permanently in the reduced state and 4) those that are susceptible to reversible and irreversible oxidation. Due to the available multiple oxidation states of sulfur, cysteine thiols are also targets for various post-translational modifications including oxidation, S-nitrosylation, palmitoylation, prenylation, and Michael addition with oxidized lipid species. The oxidation status of the protein cysteine thiols is
largely dependent on its environment. In general, oxidised protein cysteine thiols promoting the formation of disulfide linkages occurs in subcellular compartments with higher oxidizing environments whereas the reducing cytosolic environment promotes either permanently reduced protein thiols or reversibly oxidized protein thiols (ROPTs). ROPTs have been reported to play an important role in enzyme activity and its regulation. It has been observed that cysteines with disulfide bond are generally very well conserved in contrast to unpaired cysteines with a free thiol group. Our study with the purified rPfCs expressed in E. coli seems to suggest that some of these protein cysteine thiols are involved in the formation of disulfide linkages as evidenced by the mobility changes when compared between non-reduced and reduced-alkylated conditions in the gel-shift assay. We have observed that the native parasite PfCs protein also exhibits similar mobility changes in the gel-shift assay, indicating that the putative disulfide linkages in the recombinant PfCs are likely of biological relevance.

Whether all of the protein cysteine thiols in PfCs are involved in disulfide linkages, or some of them function as ROPTs or form metal co-ordinates, remains to be elucidated. This would be largely dependent upon the redox potential of the subcellular compartment where this protein is localized. Based on the available report of PfCs localization, it is likely that PfCs has both cytosolic and another non-cyto-solic subcellular distribution. Dual localization of some Shikimate pathway enzymes has also been observed in plants where proteins are both cytosolic and also trafficked to the chloroplast.

In silico modelling of PfCs structure based on reported structures of Cs proteins from Helicobacter pylori and Saccharomyces cerevisiae have failed to ascribe any role for the noticeably high number of cysteine residues. This is understandable, as Cs proteins from both yeast and bacteria contain fewer cysteine residues than the apicomplexan parasites. Sequence comparison of the malaria Cs proteins with those from bacteria or yeast Cs protein sequences show the presence of long insertions in the sequence in the malaria Cs protein (data not shown). Correlating the large insertions with increased number of cysteines, one may speculate that the cysteines might be involved in disulfide bonding or other postranslational modifications which may play important role in protein structure and function.

In this study, we successfully expressed and purified rPfCs protein with a small C-terminal hexa-histidine tag which possibly does not interfere with the protein structure and function. Using antibodies generated against rPfCs we have shown that both rPfCs, and native parasite Cs show a shift in their mobility on SDS-PAGE when compared between reduced-alkylated and non-reduced conditions indicating a likely role for the conserved cysteine residues in the formation of disulfide bonds. The rPfCs protein is, therefore, an important tool in further elucidation of malaria Cs structure and function as well as for identification of inhibitors of this important parasite target. This study, thus provides new insights into PfCs structure and function.

Acknowledgement

Author HKK received Senior Research Fellowship from the Indian Council of Medical Research, New Delhi. This study received financial support from CSIR-Indian Institute of Integrative Medicine (IIM), Jammu, through MLP6004 and OLP002 budget heads.

References