Antimalarial Drug Discovery: Structural Insights
Lunev, Sergey

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2018

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 5

Crystal structure of truncated aspartate transcarbamoylase from *Plasmodium falciparum*

This chapter has been published

Sergey Lunev, Soraya S. Bosch, Fernando A. Batista, Carsten Wrenger and Matthew R. Groves
Abstract

De novo pyrimidine biosynthesis pathway of Plasmodium falciparum is a promising target for antimalarial drug discovery. The parasite requires a supply of purines and pyrimidines for growth and proliferation and is unable to take up pyrimidines from the host. Direct (or indirect) inhibition of de novo pyrimidine biosynthesis via dihydroorotate dehydrogenase (PfDHODH), the fourth enzyme of the pathway, has already been shown to be lethal for the parasite. In the second step of the plasmodial pyrimidine synthesis pathway, aspartate and carbamoyl phosphate are condensed to N-carbamoyl-L-aspartate and inorganic phosphate by aspartate transcarbamoylase (PfATC). In this paper, a 2.5-Å crystal structure of PfATC is reported. The space group of PfATC crystals was determined to be monoclinic P 2₁, with unit-cell parameters of a = 87.0, b = 103.8, c = 87.1, α = 90.0, β = 117.7, γ = 90.0. The presented PfATC model shares a high degree of homology with the catalytic domain of E. coli ATC. There is as yet no evidence of existence of the regulatory domain in PfATC. Similarly to E. coli, the PfATC was modeled as homo-trimer where each of the three active sites is formed on the oligomeric interface. Each active site comprises the residues from two adjacent subunits in the trimer with high degree of evolitional conservation. Here, we also describe the activity loss due to mutagenesis of the key active site residues.
1. Introduction

According to World Health Organization (WHO) 3.3 billion people worldwide are at risk of being infected with malaria. The disease is considered endemic in more than 100 countries. In 2015 the WHO estimated that there were more than 200 million infections and approximately half a million deaths of malaria, where the majority of death cases occurred amongst children under 5 years old in Africa [1].

Malaria is caused by protozoan parasites of genus *Plasmodium* (*P. falciparum, P. vivax, P. knowlesi, P. malariae* and *P. ovale curtisi*). *P. falciparum* is responsible for the most dangerous and fatal form of the disease. The complex life cycle of the parasite, as well as the spread of current drug-resistant strains, make malaria treatment highly challenging [2]. There is an urgent need of new anti-malarial agents, making the identification of new drug targets very important [3-7].

The intraerythrocytic stage of *P. falciparum* is associated with an extraordinary resource uptake from the host cell. Active proliferation during this stage requires a supply of purines and pyrimidines for parasite growth to support the production of DNA and parasite replication. Malaria parasites lack the *de novo* purine synthesis pathway and take up host cell purines for growth [2, 8]. Inhibition of this pathway was shown to be lethal for *P. falciparum in vitro* [9, 10]. Early biochemical studies on *Plasmodium* parasites *P. berghei* [11-13], *P. knowlesi* [14] and *P. lophurae* [15, 16] demonstrated the inability of *Plasmodium* species to metabolize pyrimidines. The parasites lack a thymidine kinase, the enzyme responsible for salvaging host thymidine, as was confirmed by the completion of the genome sequence [17]. As the parasite lacks the pyrimidine import pathway [18, 19], pyrimidine biosynthesis pathway has been demonstrated to be major target for antimalarial drug research [20-22].

Plasmodial pyrimidine biosynthesis is affected, directly or indirectly, by many of the current antimalarials [23]. For example, the most widely used current antimalarial drug atovaquone [24] is known to inhibit cytochrome *b* in complex III of the respiratory chain and thus collapse the mitochondrial intra-membrane potential. This inhibition causes failure to
provide oxidized ubiquinone to the fourth enzyme in the pyrimidine biosynthetic pathway, dihydroororate dehydrogenase (PfDHODH) [25, 26], which was validated as an essential enzyme for the parasite and a promising drug target [22, 23, 27]. The search for selective PfDHODH inhibitors that might lead to the new antimalarial therapy remains a current topic amongst antimalarial researchers [22, 28-31].

Aspartate transcarbamoylase (ATC, EC 2.1.3.2) catalyzes the second step of pyrimidine biosynthesis, the condensation of aspartate and carbamoyl phosphate to form N-carbamoyl-L-aspartate and inorganic phosphate. The ATC from Escherichia coli was fully characterized by William Lipscomb and colleagues [32]. The E. coli ATC is known to be a highly regulated enzyme: the rate of pyrimidine biosynthesis is stimulated or inhibited in response to cellular levels of the end products of the pathway [33]. In E. coli the ATC holoenzyme is composed of six catalytic and six regulatory subunits: three regulatory pairs coordinate two catalytic trimers.

ATCs are also promising targets for anti-proliferative and anti-tumor drugs as Madani and coworkers showed that while almost undetectable in healthy human brain tissue, significant ATC levels were present in all studied tumor samples [34]. In 2014, ATC from P. falciparum (PfATC) was also reported as a potential target for gametocidal drug development [35].

PfATC is a polypeptide of 375 amino acids with a predicted molecular mass of 43.3 kDa (PlasmoDB ID: PF3D7_1344800; Table 1). BLAST analysis of PfATC sequence shows that the N-terminal region of approximately 36 amino acids does not have known structural homologs [36]. In addition no homologs of the ATC regulatory chain have yet been identified in the plasmodial genome. The amino acids (55-372) can be aligned to the catalytic domain of P. abyssi ATC [37] with an identity of 38%. Based on the absence of structural analogs in the Protein Data Bank (PDB) and predicted presence of apicoplast-targeting signal [38], the N-terminal region of 36 amino acids was removed and a truncated construct PfATC-Met3 was used in this study. Attempts to obtain the full-length PfATC have failed due to spontaneous proteolysis near the N-terminus (data not shown).
The truncated PfATC (PfATC-Met3) consists of 349 amino acids including 339 of wild type PfATC (37-375) and SAWSHPQFEK sequence of Strep-tag (IBA3) at the C-terminus with approximate molecular mass of 40.3 kDa. Recent attempts to characterize PfATC were published [39, 40], but a crystal structure is necessary to study the molecular basis of the catalytic mechanism, support validation of PfATC as a drug target and potentially support structure based drug design.

In this manuscript, the crystal structure and preliminary characterization of a truncated PfATC as well as its localization within the parasite is reported. Structural information was used to design a mutant of PfATC with significantly reduced catalytic activity; the activity profiles of the mutant and wild type enzymes are reported. These data are essential in understanding the role of plasmodial aspartate metabolism and might lead to a novel antimalarial therapy.

2. Materials and Methods

2.1. Cloning

A full length gene encoding for PfATC was amplified via reverse transcriptase PCR using *P. falciparum* total RNA from *P. falciparum* as a template. In the first step, the cDNA was synthesized from RNA using Maxima First Strand cDNA Synthesis Kit for RT-qPCR (Thermo Scientific). In the second step, full-length PfATC gene was amplified from the cDNA template using sequence-specific sense (5´-GCGCGCGGTCTCAATTGATT-GAAATATTGGCCTGAC-3´) and antisense (5´-GCGCGCGGTCTCGCGCTGCTAGATTGAGAATA-3´) primers. The PCR reaction was performed using SuperMix HiFi polymerase mix (Invitrogen). The generated PCR product was digested with *BsaI* (restriction sites underlined) and cloned into pASK-IBA3 vector (www.iba-lifesciences.com) previously digested with the same enzyme. The final expression plasmid pASK-IBA3-PfATC-full encoded the full-length version of PfATC with the additional amino acids SAWSHPQFEK (Strep-tag) at the C-terminus (Table S1). Cloning of the truncated PfATC (PfATC-Met3) was performed via
PCR amplification using pASK-IBA3-PfATC-full plasmid as a template, sense (5′-GCGCGCGGTCTCCATTTTATATCAATAGCAAG-3′) and antisense (5′-GCGCGCGGTCTCCCGCTGCTGCTAGTTGATGAAAAAATGAG-3′) sequence-specific primers.

The PCR fragment was similarly cloned into pASK-IBA3 vector previously digested with BsaI. The final expression plasmid pASK-IBA3-PfATC-Met3 encoded the truncated version of PfATC (residues 37-375) with the additional amino acids SAWSHPQFEK (Strep-tag) at the C-terminus (Table 1). All plasmid samples were verified by sequencing.

2.2. Expression

PfATC-Met3 was recombinantly expressed using E. coli Rosetta 2 (DE3) pLysS (Nalgene) competent cells transformed with pASK-IBA3-PfATC-Met3 expression plasmid. The optimal cell line and inductor concentration were chosen based on preliminary small-scale expression trials (data not shown). The culture was propagated in 1 L of selective LB media supplemented with 50 μg ml⁻¹ ampicillin, 35 μg ml⁻¹ chloramphenicol and 4 mM MgCl₂ at 310 K in 2 L baffled Erlenmeyer flasks (Nalgene) and induced with 60 ng ml⁻¹ of Anhydrotetracycline (AHT) according to the expression trial results. The temperature of the culture was lowered to 291 K after induction and cells were harvested by centrifugation after overnight expression. Centrifugation was performed using SLA-3000 rotor (Thermo Scientific) at 10,000 x g for 30 min.

2.3. Purification

Purification of recombinant PfATC-Met3 was performed via Strep-Tag chromatography according to the manufacturer’s recommendations (IBA Lifesciences). The bacterial pellet from 1 L culture was resuspended in 40 mL Lysis buffer [50 mM Tris (Tris(hydroxymethyl)aminomethan) pH 8.0, 300 mM NaCl, 5% (v/v) glycerol, 3 mM β-mercaptopethanol (BME)].
<table>
<thead>
<tr>
<th>Gene</th>
<th>Truncated Pf/ATC (wild type)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source organism</td>
<td><em>Plasmodium falciparum</em> strain 3D7</td>
</tr>
<tr>
<td>DNA source</td>
<td>pASK-IBA3-Pf/ATC-full plasmid</td>
</tr>
<tr>
<td>Forward primer (<em>BsaI</em>)</td>
<td>5'-GGCGCGCGGTCTCCATGTTTTATATCAATAGCAAG-3'</td>
</tr>
<tr>
<td>Reverse primer (<em>BsaI</em>)</td>
<td>5'-GGCGCGCGGTCTCCCGCGCTGCTAGTGTGATGAAAAAT-GAG-3'</td>
</tr>
<tr>
<td>Expression vector</td>
<td>pASK-IBA3</td>
</tr>
<tr>
<td>Expression host</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>Complete amino acid sequence of the construct produced</td>
<td>MFYINSKYKIDLDKMKMNKSVINIDDVEELLALYTSKQE KILKNENEDSKYELFVCFSVFPSTRTRCSRDFAAILKGLSKVLI TDMMNSTSFSYKGETVEDAFLKILSTYVDIYRDPSSKMNVDIAVSSSS KPIINANGTGEHPTQSLLLLFYTHNYPPFFILDRNINLNLIAVFG DLKNGRTVHSLSKLSRYNVFSNFVCKSLNPDKDIVNTITYNLKK NNFYSDSICKYFDFNLEEGLEDVHIYMMTRIKERFTDVEYNQYKN AFILSNKTLENRDDTKILHPLPRVNEIKVEVDSNPKSVYFTQAEN GLYVRMALLYLIFSSTS <strong>SANSHPQFEK</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Gene</th>
<th>Pf/ATC-R109A/K138A double mutant (Pf/ATC-RK)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA source</td>
<td>pASK-IBA3-Pf/ATC-Met3 plasmid</td>
</tr>
<tr>
<td>Forward primer (R109A)</td>
<td>5'-GTTCCTTGACAAACAGCTACAGCAACAAGATGTTCTTTTGAT-GC-3'</td>
</tr>
<tr>
<td>Reverse primer (R109A)</td>
<td>5'-GCATCAAAGAACAATCTTTGTGCTGTACTTGGTTCAAG-GAAG-3'</td>
</tr>
<tr>
<td>Forward primer (K138A)</td>
<td>5'-CTGATATGATCAATCTTTCTTCTTTATCGCGGAGAAGCTTT-GAAGATGCC-3'</td>
</tr>
<tr>
<td>Reverse primer (K138A)</td>
<td>5'-GGCATCTCTTACACGTTCTCTCCCAGATAAAAGAGTGGGTCAAT-TGATACAGGC-3'</td>
</tr>
<tr>
<td>Expression vector</td>
<td>pASK-IBA3</td>
</tr>
<tr>
<td>Expression host</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>Complete amino acid sequence of the construct produced</td>
<td>MFYINSKYKIDLDKMKMNKSVINIDDVEELLALYTSKQE KILKNENEDSKYELFVCFSVFPSTRTRCSRDFAAILKGLSKVLI TDMMNSTSFSYKGETVEDAFLKILSTYVDIYRDPSSKMNVDIAVSSSS KPIINANGTGEHPTQSLLLLFYTHNYPPFFILDRNINLNLIAVFG DLKNGRTVHSLSKLSRYNVFSNFVCKSLNPDKDIVNTITYNLKK NNFYSDSICKYFDFNLEEGLEDVHIYMMTRIKERFTDVEYNQYKN AFILSNKTLENRDDTKILHPLPRVNEIKVEVDSNPKSVYFTQAEN GLYVRMALLYLIFSSTS <strong>SANSHPQFEK</strong></td>
</tr>
</tbody>
</table>

Table 1. The cloning details for truncated version of *Pf*ATC-Met3 and the mutant *Pf*ATC-RK. *BsaI* restriction sites in the primers are underlined. Point mutations are shown in bold and highlighted. Additional Strep-tag residues at the C-terminus are shown in bold italic.
The lysis was performed by sonication on ice. The supernatant containing soluble Strep-tagged PfATC-Met3 was clarified by centrifugation at 45,000 x g (SS-34 rotor, Thermo Scientific) and filtered using 0.45 μm filter membrane (Whatman). The filtered supernatant was incubated with 2.0 mL Strep affinity resin (Strep Tactin Superflow, IBA Lifesciences) for 20 minutes at 277 K and subsequently poured onto a gravity-flow column (Bio Rad) and washed with 100 mL Lysis Buffer. The protein was eluted with 8 mL Elution buffer [50 mM Tris pH 8.0, 300 mM NaCl, 2.5 mM desthiobiotin, 5% (v/v) glycerol, 3 mM BME]. The eluate was pooled and concentrated at 277 K to 2 mg mL⁻¹ using Vivaspin 4 concentration column with a 10 kDa cutoff (Sartorius Stedim Biotech).

In order to select the correct buffer for further purification a Thermal Shift Assay [41, 42] was performed using CFX96 Real-Time system (Bio-Rad). SYPRO Orange dye (5000x stock, Invitrogen) was added to the protein sample (2 mg mL⁻¹) at 1:500 ratio. Each experiment consisted of 5 μL of the protein/dye mixture and 45 μL of the buffer component to be screened. Water was used as a control instead of the buffer sample. Inflection points in graphs of relative fluorescence units (RFU) against temperature were determined manually and used as an indicator of the sample thermal stability in presence of the screened buffer components. Components that exhibited positive thermal shift in comparison to a water control sample were used to select a size exclusion chromatography (SEC) buffer [20 mM Tris pH 8.0, 300 mM NaCl, 10 mM Na-Malonate, 5% (v/v) glycerol and 2 mM BME]. The remaining protein was concentrated to a volume of 1 mL (Vivaspin4, Sartorius Stedim Biotech) and purified via SEC using a HiLoad 16/60 Superdex 75 column (GE Healthcare) pre-equilibrated with SEC buffer using NGC liquid chromatography system (BioRad). The purified protein eluted as a single peak and was pooled and concentrated to 10 mg mL⁻¹ at 277 K (Vivaspin4, Sartorius Stedim Biotech). The final concentration was determined based on the protein theoretical absorbance at 280 nm [Abs 0.1% (1 mg ml⁻¹) = 0.84; http://web.expasy.org/protparam].
The concentrated protein was immediately used in crystallization trials. The estimated protein purity was better than 95% (Figure 1, a) as assessed by Coomassie Brilliant Blue-stained SDS-PAGE [43]. The final yield of purified PfATC-Met3 was 6 mg per liter of culture. Expression and purification of the mutant version of PfATC-Met3 were performed identically.

### 2.4. Crystallization

Screening for crystallization conditions for PfATC-Met3 was performed using a high-throughput crystallization robot (Gryphon, Art Robbins) against commercially available sparse-matrix screening kits (JCSG plus...
and PACT premier; Molecular Dimensions Ltd.). All experiments were performed at 293 K using the sitting drop vapor diffusion technique in 96-well MRC2 plates (Molecular Dimensions Ltd.). Equal volumes (0.1 μL) of protein solution and crystallization reagent were equilibrated against 50 μL of reservoir solution. Medium-size single crystals appeared overnight in various conditions containing PEG 3350/4000. Further optimization was performed using the hanging-drop technique by varying the precipitant concentration, ionic strength, pH and buffer conditions. The optimized conditions consisted of equal amounts (1.5 μL) of 9 mg ml⁻¹ PfATC-Met3 sample and 0.2 M Sodium Sulfate, 5 mM MgSO₄, 15 % (w/v) PEG 3350 in 0.1 M bis-Tris Propane pH 7.5 as the crystallization solution at 293 K. Rhomboid-shaped diffraction-quality crystals (Figure 1b) with maximum dimensions of 200 μm appeared overnight and were subsequently harvested using mounted round LythoLoops (Molecular Dimensions Ltd.), incubated in the Cryo-buffer (below) and flash-cooled in liquid nitrogen prior to shipment to the synchrotron. The Cryo-buffer was chosen based on an estimation from [44] and consisted of 0.2 M Sodium Sulphate, 5 mM MgCL₂, 15 % (w/v) PEG 3350, 20% glycerol in 0.1 M bis-Tris Propane pH 7.5.

2.5. Data Collection

Cryo-cooled PfATC-Met3 crystals were sent to European Synchrotron Radiation Facility (ESRF, Grenoble) using dry-shipping cryo-container (Taylor-Wharton) and a 2.4 Å data set was collected at 100 K in the nitrogen stream at the ID23-2 beamline. Initial characterization of the crystals and the optimization of data collection parameters were performed using BEST software [45, 46]. The space group of PfATC crystals was calculated to be monoclinic P 2₁ with unit-cell parameters of a = 87.0, b = 103.8, c = 87.1, α = 90.0, β = 117.7, γ = 90.0, the solvent content was calculated to be 59.4 %. The data were processed using the XDSapp [47] software. The analysis also showed that the crystal was twinned. The Matthews coefficient [48] predicted a trimeric form of PfATC-Met3 in the asymmetric unit.
unit. The data collection and processing statistics are reported in Table 2.

2.6. Data processing and Refinement

The structure of PfaTC-Met3 was solved by molecular replacement using BALBES [49] software within the CCP4 package [50]. The crystal structure of ATC from Pyrococcus abyssi [37] was used as a search model (38% identity) yielding a solution of three molecules in the asymmetric unit. The model was further optimized via manual rebuilding in COOT [51] and refined with REFMAC5 software [52]. The final refinement steps were carried out with global NCS restraints, TLS restraints calculated via TLSMD web server [53, 54]. These steps also included twin-refinement. The final 2.5-Å model of PfaTC-Met3 consisted of 3 molecules in the asymmetric unit with R factor of 0.18 and free R factor of 0.22 (Table 3) and has been deposited in the Protein Data Bank (PDB) with the accession code 5ILQ. The structure figures in this manuscript were prepared with PyMol [55].

2.7. Mutagenesis

The double mutant PfaTC-Met3-R109A/K138A (hereafter PfaTC-RK) was created via site-directed mutagenesis using the previously described pASK-IBA3-PfaTC-Met3 expression plasmid as a template. The PCR reaction was performed using Phusion Hot Start II DNA polymerase (Thermo Scientific) according to the manufacturer’s recommendations. For mutagenesis details please refer to Table 1. The generated mutant plasmid was treated with DpnI enzyme for 2 hours at 310 K to eliminate the parental template. The mutations were validated by sequencing. Similarly to the wild type the mutant construct also encoded a C-terminal Strep-tag. Expression and purification of PfaTC-RK mutant were performed identically to PfaTC-Met3.
Table 2. Data collection and processing

Values for the outer shell are given in parentheses.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diffraction source</td>
<td>ESRF ID 23-2</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.98</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>100</td>
</tr>
<tr>
<td>Detector</td>
<td>PILATUS 6M</td>
</tr>
<tr>
<td>Crystal-detector distance (mm)</td>
<td>337.55</td>
</tr>
<tr>
<td>Rotation range per image (°)</td>
<td>0.1</td>
</tr>
<tr>
<td>Total rotation range (°)</td>
<td>129</td>
</tr>
<tr>
<td>Exposure time per image (s)</td>
<td>0.04</td>
</tr>
<tr>
<td>Space group</td>
<td>P 2_1</td>
</tr>
<tr>
<td>a, b, c (Å)</td>
<td>86.96, 103.80, 87.11</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.0, 117.6, 90.0</td>
</tr>
<tr>
<td>Mosaicity (°)</td>
<td>0.2</td>
</tr>
<tr>
<td>Resolution range (Å)</td>
<td>45.0 – 2.42 (2.56 – 2.42)</td>
</tr>
<tr>
<td>Total No. of reflections</td>
<td>126651 (20324)</td>
</tr>
<tr>
<td>No. of unique reflections</td>
<td>50423 (8051)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95.9 (95.6)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>2.51 (2.52)</td>
</tr>
<tr>
<td>&lt;I/σ(I)&gt;</td>
<td>10.81 (1.98)</td>
</tr>
<tr>
<td>CC_{1/2}</td>
<td>99.7 (76.6)</td>
</tr>
<tr>
<td>R_{meas}(%)</td>
<td>7.6 (58.0)</td>
</tr>
<tr>
<td>Overall B factor from Wilson plot (Å^2)</td>
<td>56.64</td>
</tr>
</tbody>
</table>

R_{meas} is defined as \( \sum_{hkl} \sum_{i} | I_i(hkl) - <I(hkl)> | / \sum_{hkl} \sum_{i} I_i(hkl) \), where \( I_i(hkl) \) is the \( \text{i} \)th intensity measurement of reflection \( hkl \) and \( <I(hkl)> \) is the average intensity from multiple observations.

The free R-factor was calculated using 5% of randomly selected reflections omitted from the refinement.

2.8. Activity Assays

The kinetic properties of PfATC-Met3 as well as PfATC-RK were investigated according to [56], [57] and [58] with minor modifications. Briefly, the reaction was carried out at room temperature in a total volume of 160
μl of 14 mM L-aspartate and 1 mM carbamoyl phosphate (CP) saturated substrate solution in 200 mM Tris-Acetate buffer pH 8. The reaction was stopped by 80 μl of 25 mM Ammonium molybdate in 4.5 M H₂SO₄. After all the reactions were stopped 160 μl of 0.5 μM malachite green in 0.1 % (w/v) poly(vinyl alcohol) (PVA) were added and incubated for 30 min at room temperature. Subsequently, the absorption of the samples at a wavelength of 620 nm was measured. The analyses were evaluated from at least three independent quadruplicate assays using GraphPad Prism 4 (GraphPad Software, USA).

2.9 Transfection

The open reading frame of the full-length PfATC was cloned in front of the GFP gene of the expression vector pARL1a and subsequently transfected into P. falciparum. Briefly, erythrocytes were washed in cytomix [59] and an aliquot of 450 ml was combined with approximately 100 mg of the construct resuspended in 50 ml of TE buffer. The red blood cells were electroporated using a Gene-Pulser X-Cell total system (BIORAD) at 0.31 kV and 900 mF. After electroporation the cells were transferred into pre-warmed RPMI medium and inoculated with 50 ml of red blood cells infected with 10% schizonts to give a parasitaemia of 1%. Four hours post transfection the culture medium was exchanged. Parasites were grown for 24 h without drug selection before the medium was supplemented with 5 nM of WR 99210. Parasites were first observed after 16–60 days of selection and live parasites were analyzed by fluorescent microscopy using an Axioskop 2 plus microscope (ZEISS). In order to visualize the nucleus, parasites were incubated with Hoechst 33342 dye according to the manufacturer’s recommendation (Invitrogen). For Co-localization experiments to visualize the ER, transfected Plasmodium was supplemented with 2 mM of ER-Tracker™ Red BODIPY-TR (Invitrogen) prior to microscopy.
3. Results

A truncated version of ATC from *Plasmodium falciparum* was cloned, recombinantly expressed, purified and crystallized. Full-length *Pf*ATC is a protein of 375 residues and calculated molecular weight of 43.3 kDa. *BLAST* [36] analysis has revealed no structural homologs of the first 36 N-terminal residues (Figure 2). *Pf*ATC shows 38% identity (84% query cover) with the catalytic subunit of ATC from *Pyrococcus abyssi* (PDB code: 1ML4; [37]) and 39% identity (84% query cover) with the *Escherichia coli* ATC (PDB code: 1D09; [60]). Nucleotide *BLAST* analysis also shows 37% identity (86% query cover) with ATC domain of Human CAD (carbamoyl-phosphate synthetase 2, aspartate transcarbamylase, and dihydroorotase) protein. The crystal structure of this domain is not yet available, although the preliminary data have been reported [62]. Further analysis with *PlasmoAP* tool [38] showed a strong prediction for the N-terminal region to contain an apicoplast targeting sequence with a possible cleavage site between residues 27-28 (IRT-KK) (Table S1). The GFP-labelling experiments *in vivo* showed distinct localisation of *Pf*ATC rather than in ER supporting the hypothesis of apicomplexian nature of the enzyme (Figure 3). Based on these data, the third methionine in the sequence (Met37) was selected as a starting point of our construct.

Attempts to express and purify full-length *Pf*ATC with sufficient purity have failed. SDS-PAGE analysis of the elution fraction of Strep-purification showed three distinct Strep-tagged (C-terminal) bands with sizes around 43 kDa and further size-exclusion purification showed three overlapping peaks that could not be separated with sufficient purity (data not shown). Use of protease-inhibitors cocktail during the lysis and purification did not show any significant improvement. The presence of at least three versions of *Pf*ATC of different lengths in the elution fraction suggests that the full-length protein is proteolyzed near the N-terminus due to the apicoplast-targeting nature of this region. The activity assay performed with *Pf*ATC-Met3 showed that it possesses catalytic activity and further experiments have been performed with this truncated *Pf*ATC construct.
Figure 2. The homology analysis of \( PfATC-Met3 \) was performed using BLAST tool [36] and visualized via Tcoffee [61]. Red, yellow, green and blue colours represent good, average, bad and weak alignment, respectively. Residues essential for the active sites are shown with arrows. Absolutely, strongly and slightly conserved residues are marked with `*`, `:**` and `:` symbols, respectively.
Table 3. Structure solution and refinement

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution range (Å)</td>
<td>20.0 – 2.5</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>95.2</td>
</tr>
<tr>
<td>σ cutoff</td>
<td>1</td>
</tr>
<tr>
<td>No. of reflections, working set</td>
<td>42863</td>
</tr>
<tr>
<td>No. of reflections, test set</td>
<td>2354</td>
</tr>
<tr>
<td>Final $R_{\text{cryst}}$</td>
<td>0.18</td>
</tr>
<tr>
<td>Final $R_{\text{free}}$</td>
<td>0.22</td>
</tr>
<tr>
<td>No. of non-H atoms</td>
<td>7909</td>
</tr>
<tr>
<td>Protein</td>
<td>7801</td>
</tr>
<tr>
<td>Ion</td>
<td>23</td>
</tr>
<tr>
<td>Water</td>
<td>88</td>
</tr>
<tr>
<td>R.m.s. deviations</td>
<td></td>
</tr>
<tr>
<td>Bonds (Å)</td>
<td>0.019</td>
</tr>
<tr>
<td>Angles (°)</td>
<td>1.96</td>
</tr>
<tr>
<td>Average $B$ factors (Å$^2$)</td>
<td>57.39</td>
</tr>
<tr>
<td>Protein</td>
<td>58.04</td>
</tr>
<tr>
<td>Ion</td>
<td>56.76</td>
</tr>
<tr>
<td>Water</td>
<td>51.47</td>
</tr>
<tr>
<td>Ramachandran plot</td>
<td></td>
</tr>
<tr>
<td>Most favoured (%)</td>
<td>93.60</td>
</tr>
<tr>
<td>Allowed (%)</td>
<td>5.98</td>
</tr>
</tbody>
</table>

Free R-factor was calculated using 5% of randomly selected reflections omitted from the refinement. Values in parentheses correspond to the highest resolution shell.

The recombinant PfATC-Met3 was produced with a C-terminal Strep-tag (Iba Lifesciences) and purified by Strep-chromatography according to the manufacturer’s recommendations. In order to achieve higher purity and homogeneity, PfATC-Met3 was further purified via SEC and the final purity was assessed via SDS-PAGE analysis (Figure 1, a). The yield of pure homogeneous protein was approximately 6 mg per liter of culture. In order to increase the post-purification stability of the protein sample and increase crystallization chances, the SEC buffer was chosen based on the results of Differential Scanning Fluorimetry. The fresh purified protein was concentrated to 9 mg mL$^{-1}$ and used for crystallization experiments.

Multiple crystals of PfATC-Met3 appeared overnight in various crys-
tallization screening conditions containing PEG (polyethyleneglycol) 3350/4000. Optimization of the initial conditions yielded single PfATC crystals with maximum dimension of 0.2 mm (Figure 1, b). The final crystallization condition was 0.1 M bis-Tris Propane pH 7.5, 0.2 M Na-Sulfate, 5 mM Mg-Sulfate and 15% (w/v) PEG 3350.

The optimized crystals diffracted X-rays to a maximum resolution of 2.4 Å at ID23-2 beamline (ESRF, Grenoble). The space group of the PfATC-Met3 crystals was determined to be monoclinic P 2₁, with unit-cell parameters of \( a = 87.0, b = 103.8, c = 87.1, \alpha = 90.0, \beta = 117.7, \gamma = 90.0 \). The presence of three PfATC molecules per asymmetric unit was confirmed by molecular replacement and the 2.5-Å structure of PfATC-Met3 has been deposited in the PDB with the accession code 5ILQ.

3.1. Overall structure of PfATC-Met3

PfATC-Met3 is a homo-trimer with three active sites formed at the oligomeric interfaces (Figure 4, a, b). Superposition of the PfATC-Met3 structure and catalytic subunit of *E. coli* ATC (PDB code: 1D09; [60]) showed high level of sequence and secondary structure conservation (Figure 2, Figure 5). BLAST [36] analysis against nine known homologs (structures available) showed that of 375 residues 46 (12.3%) were absolutely conserved, 43 (11.5%) were strongly conserved and 25 (6.7%) slightly conserved. The loops formed by 84-91, 203-212 and 265-275 residues of PfATC-Met3 are slightly longer than their *E. coli* analogues (residues 34-36, 150-155 and 206-208, respectively). In addition, the N-terminal region of PfATC-Met3 model (residues 37-48) is also longer and shows no secondary structure. Poor electron density coverage suggests that this region is highly mobile. For this reason N-terminal residues 37-46 (chain A) and 37-43 (chains B, C) were excluded from the final model as well as the loops formed by the residues 297-311 (all three chains) and C-terminal region of chain A (residues 375-378). The presence of additional electron density near the C-terminus was modelled as the Strep-tag (SAWSHPQFEK) with
the last 2-3 residues (FEK) excluded.

In 1991 Stevens and co-workers reported the essential residues that form the active site of ATC based on the structure of *E. coli* ATC, confirmed by site-specific mutagenesis experiments [65]. In *Pf*ATC-Met3, all the homologous residues are absolutely conserved (*BLAST*) among the ATC’s from other organisms (Figure 2, 6). Each subunit of the trimer hosts an active site comprised with two residues (Ser135 and Lys138) from the adjacent chain.

*PISA* analysis [66] of the *Pf*ATC-Met3 structure showed that the oligomeric contact between each subunit pair (hereafter A and B) is formed by 32 (Buried Surface Area (BSA) of 899 Å²) and 27 residues (BSA of 956 Å²), respectively (Figure 7, a-f). The interface from the subunit A consists of 9 absolutely conserved, 3 strongly conserved and 2 slightly conserved residues. The adjacent interface B has 8 absolutely, 4 strongly and 2 slightly conserved residues, respectively. Each subunit has a total surface area of approximately 14,300 Å² where 1855 Å² (13%) belong to the intraoligomeric interfaces.

**Figure 3.** Fluorescent microscopy imaging of *P. falciparum* transfected with GFP-labeled *Pf*ATC shows distinct localisation of *Pf*ATC (see materials and methods 2.9).

**Figure 4.** The overall model of *Pf*ATC-Met3. (a) *Pf*ATC-Met3 is a homo-trimer with three active sites (shown with stars) formed at the oligomeric interfaces. (b) Schematic view of *Pf*ATC homo-trimer.

**Figure 5.** The secondary structure of *Pf*ATC-Met3 compared to the *E. coli* ATC (a) and *B. subtilis* ATC (b). Coordinates of *Pf*ATC-Met3 (green) structure and catalytic subunits of *E. coli* ATC in liganded R-state (magenta, PDB code: 1D09, [60]) and unliganded T-state *B. subtilis* ATC (blue, PDB code: 3R7D, [63]) were superimposed using COOT [51, 64].
**Figure 6.** Superposition of the key active site residues of PfATC-Met3 (green) and *E. coli* ATC (magenta, PDB code: 1D09) that exhibit the absolute conservation amongst homologous species (see Figure 2).

**Figure 7.** PfATC-Met3 structure analysis reveals that each subunit forms two oligomeric contacts with other subunits of the trimer (a). Residues that form these contacts are shown in blue (b-f). Absolutely (red), strongly (orange) and slightly conserved (green) residues of these interfaces are shown (c,e). Active site residues (magenta) are also shown on the surface (f).
3.2. Mutagenesis studies and activity profile of the truncated version of PfATC

Based on the mutagenic studies on *E. coli* ATC summarized in [32] and overall structure similarity, a mutant PfATC-Met3 version was constructed. In *E. coli*, point mutations of the essential active site residues of the catalytic chain ATC resulted in significant loss of activity. For example, guanidinium group of Arg54 of *E. coli* ATC was shown to be crucial for binding of the substrate carbamoyl-phosphate (CP) and enzymatic condensation of CP and aspartate, and when removed (Arg54Ala mutation) caused 17,000-fold loss in activity and 13-fold reduction in affinity for CP [67, 68]. The amino group of Lys84, that completes the active site of the adjacent subunit, interacts with both Arg54 and aspartate upon binding; Lys84Asn mutation has resulted in 1200-fold activity loss of the *E. coli* enzyme [60]. Both mutations did not affect the folding and the structure of the enzyme as confirmed by X-ray data (data not shown). Comparison between the crystal structures of PfATC-Met3 and the *E. coli* ATC (PDB code: 1D09; [60]) showed that the Arg109 and Lys138 residues of the active site of PfATC-Met3 are homologous to Arg54 and Lys84 of *E. coli* ATC. The mutant version PfATC-RK (R109A/K138A) was designed and cloned using site-specific mutagenesis technique and recombinantly expressed according to the same protocol as the PfATC-Met3 (See Materials and Methods section).

In order to confirm the presence of the catalytic activity of the recombinant PfATC-Met3 a phosphate-detection system based on malachite green [57] was established. The specific *in vitro* activity of PfATC-Met3 was $11.04 \pm 1.04 \, \mu\text{mol min}^{-1} \, \text{mg}^{-1}$. Activity assays with the double mutant PfATC-RK was also performed and showed significantly lower value of $0.85 \pm 0.30 \, \mu\text{mol min}^{-1} \, \text{mg}^{-1}$. Additional native PAGE electrophoresis and DLS (Dynamic light scattering) experiments (data not shown) confirmed the presence of oligomeric PfATC-Met3 assembly in the solution for both wild type and mutant versions. These data have confirmed the initial hypothesis that the mutations of the key active site residues (R109A and
K138A) significantly reduce the catalytic activity while having no adverse effect on the oligomeric assembly.

4. Discussion

Pyrimidine biosynthesis of *P. falciparum* represents highly attractive target for future antimalarial drug development. Determined crystal structure of truncated aspartate transcarbamoylase from *Plasmodium falciparum*, the enzyme that catalyses the second step of the pathway, shows high degree of evolitional conservation amongst its homologs from other organisms. The active site of the enzyme shows the highest degree of conservation, making the structure-based design of specific active site targeting inhibitors of PfATC and its further validation as a drug target highly challenging.
Table S1. Macromolecule production information (Full-length PfATC)

<table>
<thead>
<tr>
<th>Source organism</th>
<th>Plasmodium falciparum strain 3D7</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA source</td>
<td>cDNA generated from total RNA via reverse transcriptase PCR</td>
</tr>
<tr>
<td>Forward primer (BsaI)</td>
<td>5’-GCGCGCGGTCTTCATGATTGAAATATTTGCCACTGC-3’</td>
</tr>
<tr>
<td>Reverse primer (BsaI)</td>
<td>5’-GCGCGCGGTCTCCCGCGCTAGTTGATGAAATGAG-3’</td>
</tr>
<tr>
<td>Expression vector</td>
<td>pASK-IBA3</td>
</tr>
<tr>
<td>Expression host</td>
<td>Escherichia coli</td>
</tr>
</tbody>
</table>

Complete amino acid sequence of the construct produced:

MIEIFCTAIVITILIVFVYMIIIRTRKRLKLDNMFYINSKYKIDLDKIMTKMKNSVINVIDDEELLAILYTSQFEKILKNEDSKYLENKFCSVFLEPSTRTRCSFDAAILKLGSKVLNITDMSF YK
GETVEDAFKILSTYVDGIYRDPSKNVIAVSSSSKPIINAGNTGEHPTQSLDFYTIIIHYFPIFILDRNNKKLNNIAFGDKNRTVHLSKLLSRYNVSFNFVSCSFLNPKDINVITYLNKKNFYSDDSIKYFDNLEEGLEDVHIIYMRIGKERFDVDEYNQYFNAPFILSNKTLENTREDTILHPLPRVNEIKVEVDSNFKSVYFTQAENGLYVRMALLYLIFSSTSSAHSHPQFEK

Table S1. The cloning details for the full version of PfATC. BsaI restriction sites in both primers are underlined. Predicted apicoplast targeting sequence (PlasmoAP, [69]) is also underlined. Predicted cleavage site (IRT-KK) is shown in italic. Additional Strep-tag residues at the C-terminus are shown in bold italic.
References


37. Van Boxtel B, Cunin R, Khan S, Maes D. Aspartate transcar-


