Structural Insights into the Inactive Subunit of the Apicoplast-localized Caseinolytic Protease Complex of Plasmodium falciparum*

Majida El Bakkouri,† Sumit Rathore,‡ Charles Calmettes, Amy K. Wernimont, Kaiyun Liu,§ Dipto Sinha,¶ Mohd Asad,¶ Patrick Jung, Raymond Hui, Asif Mohmmed,§ and Walid A. Houry

From the †Department of Biochemistry, University of Toronto, Toronto, Ontario M5S 1A8, Canada, the ‡Malaria Group, International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi 110067, India, and the §Structural Genomics Consortium, University of Toronto, Toronto, Ontario MSG 1L7, Canada

The ATP-dependent caseinolytic protease, ClpP, is highly conserved in bacteria and in the organelles of different organisms. In cyanobacteria, plant plastids, and the apicoplast of the genus Plasmodium, a noncatalytic paralog of ClpP, termed ClpR, has been identified. ClpRs are found to form heterocomplexes with ClpP resulting in a ClpRP tetradecameric cylinder having less than 14 catalytic triads. The exact role of ClpR in such a complex remains enigmatic. Here we describe the x-ray crystal structure of ClpR protein heptamer from Plasmodium falciparum (PfClpR). This is the first structure of a ClpR protein.

The structure shows that the PfClpR monomer adopts a fold similar to that of ClpP, but has a unique motif, which we named the R-motif, forming a β turn located near the inactive catalytic triad in a three-dimensional space. The PfClpR heptamer exhibits a more open and flat ring than a ClpP heptamer. PfClpR was localized in the P. falciparum apicoplast as is the case of PfClpP. However, biochemical and structural data suggest that, contrary to what has been observed in other organisms, PfClpP and PfClpR do not form a stable heterocomplex in the apicoplast of P. falciparum.

The ATP-dependent caseinolytic proteases (Clp) have essential roles in protein quality control by removing misfolded, damaged, and regulatory proteins. These proteases are highly conserved in prokaryotes, in the mitochondria of many eukaryotes, and in the chloroplast of plants (1–4). The proteases consist of a proteolytic component for protein degradation and a chaperone forming a AAA+ superfamily (5, 6) that uses ATP for substrate recognition, unfolding, and then transfer to the proteolytic component. Clp ATPase chaperones typically form hexameric complexes and contain, in addition to other domains, one (class II) or two (class I) nucleotide-binding AAA+ domains that have characteristic Walker A and Walker B nucleotide binding and recognition motifs (5, 7–10).

The proteolytic component, termed ClpP, forms two heptameric rings resulting in a cylindrical structure (3, 11). ClpP is a serine protease containing the canonical Ser-His-Asp catalytic triad. Substrates enter the proteolytic chamber through the two axial pores located on opposite sides of the cylinder; the generated peptides of about 7 to 8 residues (12, 13) then exit the chamber through equatorial side pores that we proposed to form due to the dynamics in the structure of the ClpP complex (14–20). The AAA+ chaperones catalyze the translocation of substrates through the ClpP axial pores and might also modulate the dynamics of the ClpP cylinder for substrate degradation and the exit of the generated peptides.

In organisms such as cyanobacteria and many plants, multiple isoforms of ClpP were found and a noncatalytic paralog of the protease, termed ClpR, was also identified (1, 3, 21, 22). ClpR has high sequence similarity to ClpP but is noncatalytic because it lacks the catalytic Ser residue, however, it was shown to be a core component of the Clp protease complex in cyanobacteria (23, 24) and plants (25). In these organisms, the Clp chaperone-protease complex typically consists of several paralogs of ClpP/ClpR and multiple ATPase chaperones. For exam-

* This work was supported in part by a program support grant from Department of Biotechnology, Government of India (to A. M.) and Canadian Institutes of Health Research. The atomic coordinates and structure factors (codes 4GM2 and 4HNK) have been deposited in the Protein Data Bank (http://wwwpdb.org/).

† Recipient of a fellowship from the Canadian Institutes of Health Research (CIHR) Strategic Training Program in Protein Folding and Interaction Dynamics: Principles and Diseases Grant TGF-53910.

‡ Supported by research fellowships from Indian Council of Medical Research, Government of India.

§ Recipient of a fellowship from the Canadian Institutes of Health Research (CIHR) Strategic Training Program in Protein Folding and Interaction Dynamics: Principles and Diseases Grant TGF-53910.

¶ Recipient of a fellowship from the Canadian Institutes of Health Research (CIHR) Strategic Training Program in Protein Folding and Interaction Dynamics: Principles and Diseases Grant TGF-53910.

© 2013 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.
ple, the unicellular cyanobacterium *Synechococcus elongatus* has three ClpP paralogs and one ClpR. Three proteolytic core complexes were identified in *S. elongatus*: SeClpP3R, SeClpP1P2, and SeClpP1R that interact with different AAA+ chaperones (23). Additional studies showed that the SeClpP3R cylinder consists of two identical heptameric rings of three ClpP3 and four ClpR arranged in a specific layout (24).

The Clp core complex in the plastid of *Arabidopsis thaliana* consists of five different proteolytic ClpP proteins (ClpP1, ClpP3–ClpP6) and four different ClpR proteins (ClpR1–ClpR4) (25). Recently, the subunit stoichiometry of the chloroplast AtClpRP tetradecameric complex was characterized by mass spectrometry and was found to have one heptameric ring consisting of ClpP3, -4, -5, -6 in a 1:2:3:1 ratio and the other heptameric ring consisting of three ClpP1s and one of each of ClpR1, -2, -3, -4 (26). Analyses of various ClpP and ClpR mutants in plants showed that the proteolytic and nonproteolytic subunits have differential contributions to the function of the complex but all of them, except for ClpR1, are essential for embryo or seedling development (21, 22).

Recent studies from our group (27, 28) demonstrated that the *Plasmodium falciparum* protozoan parasite that causes malaria in humans contains four Clp ATPases, which we termed *Pf* ClpB1, *Pf* ClpB2, *Pf* ClpC, and *Pf* ClpM, one *Pf* ClpP (gene ID PF3D7_0307400; old gene ID was PFC0310c), and one *Pf* ClpR (gene ID PF3D7_1436800; old gene ID was PF14_0348). Expression of all ClpP chaperones and proteases was confirmed in blood-stage parasites. *Pf* ClpB1, *Pf* ClpC, *Pf* ClpM, and *Pf* ClpP were experimentally found to be localized in the apicoplast. *Pf* ClpB2 (also known as Hsp101) was found in the parasitophorous vacuole, whereas *Pf* ClpR was only predicted to be targeted to the apicoplast but not experimentally confirmed. Unlike the plant plastids, the apicoplast is a nonphotosynthetic organelle that does, however, accommodate several important metabolic pathways essential for parasite survival. Hence, proteins in the apicoplast are potential targets for the development of novel antimalarial drugs. Indeed, inhibiting *Pf* ClpP using β-lactones showed that the protease plays an important role in development of functional apicoplasts (28). We earlier showed that both recombinant *Pf* ClpP and *Pf* ClpR can form separate homoeptamer rings as observed by size exclusion chromatography, analytical ultracentrifugation, and electron microscopy (27, 28). However, the x-ray structure of *Pf* ClpP solved in our study showed that the protein oligomerizes as a compacted tetradecamer at high concentrations required for crystallography. No stable complexes were observed between *Pf* ClpP and *Pf* ClpR (27).

In this work, we describe the x-ray crystal structure of the *Pf* ClpP protein. This is the first structure of a ClpR protein. The structure shows that the *Pf* ClpR monomer adopts a fold similar to that of other ClpPs. However, arrangement of the *Pf* ClpR heptamer has some unique features including a more open and flat ring, as well as, the presence of a unique motif surrounding the inactive catalytic site. The biochemical and structural data indicate that both *Pf* ClpP and *Pf* ClpR are localized in the parasite apicoplast, but that they form either an unstable complex or separate complexes.

## EXPERIMENTAL PROCEDURES

### Parasite Culture, Plasmid Construct, and Parasite Transfection—*Plasmodium falciparum* strain 3D7 was cultured with human erythrocytes (4% hematocrit) in RPMI media (Invitrogen) supplemented with 10% O+ human serum using a protocol described previously (29). Parasite cultures were synchronized by repeated sorbitol treatment following the protocol of Lambros and Vanderberg (30). To generate a transfection vector construct, the full-length *Pf* clpR gene (732 bp) was amplified from *P. falciparum* 3D7 genomic DNA. The amplified PCR product was then cloned in-frame to the N terminus of GFP into the vector pHH2 (31, 32). The *Pf* ClpP–GFP fragment was subcloned into the Xhol site of transfection vector pARL1.1*α* to yield construct pARL-ClpP–GFP. Synchronized *P. falciparum* 3D7 ring-stage parasites were transfected with 100 μg of purified plasmid DNA (Plasmid Maxi Kit, Qiagen) by electroporation (33), and the transfected parasites were selected over 2.5 nm WR99210 drug.

### Isolation of Total DNA and RNA, cDNA Synthesis, and Quantitative Real-time PCR—The genomic DNA was isolated from an *in vitro* culture of *P. falciparum* following a standard protocol (34). Total RNAs were isolated from synchronized *P. falciparum* 3D7 parasite cultures using a mini RNA isolation kit (Qiagen). An aliquot of 50 ng of total RNA was used to synthesize cDNA using cDNA synthesis kit (Invitrogen) following the manufacturer’s protocol. Gene-specific primers were designed using Beacon Designer 4.0 software, for *Pf* clpP, *Pf* clpR, eba175, and falcipain-2 (35), and 18S rRNA control primers were used following Blair et al. (36). Quantitative real-time PCR was carried out in triplicate using the iCycler version 3.0 (Bio-Rad). Each reaction contained equal amounts of cDNA, 100 ng, of both the gene-specific primers, and 1 μl SYBR Green PCR mix (Bio-Rad). Threshold cycle (*C*<sub>T</sub>) values were calculated using iCycler software. Standard curves for each gene were obtained by using different dilutions of wild-type genomic DNA (100 to 1 ng) as template, and these standard curves were used to determine genome equivalents of *C*<sub>T</sub> values for the respective gene and 18S rRNA in each RNA sample (36). Genome equivalents of each gene were normalized using that of 18S rRNA for all the RNA samples.

### Fractionation of Parasite Lysate—Parasites were isolated from tightly synchronized *P. falciparum* 3D7 cultures at different developmental stages by lyses of infected erythrocyte with 0.15% saponin. Cell-free parasite lysate was prepared from the parasite pellet and fractionated as described earlier (28). Briefly, the parasite pellet was suspended in 1× PBS, lysed by three freeze-thaw cycles, and the lysate was clarified by centrifugation at 20,000 × g for 30 min at 4 °C. The cell-free extract (~5 mg protein) was fractionated on a Superose 6 HR 10/30 column (GE Healthcare) equilibrated with the lysis buffer. Eluted fractions were analyzed by Western blot using the respective antibodies. Antibodies used include: rabbit anti-*Pf* ClpP (1:1,000), rabbit anti-*Pf* ClpR (1:1,000), mice anti-GFP (1:1,000); rabbit anti-histidine-rich protein II (HRPII) (1:2,000), secondary antibody (anti-rabbit or anti-mouse, 1:2,000) conjugated to horseradish peroxidase. Bands were visualized using the ECL detection kit (GE Healthcare).
The Structure of ClpR from *P. falciparum*

**Fluorescence Microscopy and Indirect Immunofluorescence Assay—** *P. falciparum* culture transfected with pARL-ClpR plasmid was synchronized by two sorbitol treatments. Parasites at different developmental stages were collected from the culture for fluorescence microscopy and stained with DAPI at a final concentration of 2 μg ml⁻¹ and MitoTracker Red CMXRos (Invitrogen) at a final concentration of 20 nM in 1× PBS for 15 min at 37 °C. The parasites were viewed using a Nikon TE 2000-U fluorescence microscope.

**Cryo-immunoelectron Microscopy—** Immunoelectron microscopy was carried out on transgenic *P. falciparum* parasites expressing PfClpR-GFP at trophozoite stages as described earlier (28, 37). Parasites were fixed in 4% paraformaldehyde, 0.04% glutaraldehyde in 1× PBS at 4 °C for 1 h and, subsequently, embedded in gelatin and infiltrated with a cryo-preservation and plasticizer (2.3 M sucrose, 20% polyvinyl pyrrolidone). After freezing in liquid nitrogen, samples were sectioned with a Leica Ultracut UCT cryo-ultramicrotome (Leica Microsystems, Barrington, IL) at −260 °C. Ultra-thin sections were blocked with 5% fetal bovine serum and 5% normal goat serum in 1× PBS for 30 min and, subsequently, stained with rabbit anti-GFP antibody (Abcam, 1:500 dilution in blocking buffer), washed thoroughly, and then incubated with 18-nm colloidal gold-conjugated anti-rabbit IgG for 1 h. Sections were stained with 0.3% uranyl acetate, 1.7% methyl cellulose, and visualized under a JOEL 1200EX transmission electron microscope (JEOL USA, Peabody, MA). All labeling experiments were conducted in parallel with controls omitting the primary antibody or using pre-immune sera as primary antibodies.

**Protein Expression and Purification—** PfClpR constructs with N-terminal His₆ tag followed by a tobacco etch virus cut site were expressed in *Escherichia coli* SG1146, which lacks the gene coding for the endogenous *E. coli* ClpP. Cells were grown in Terrific Broth media containing 100 μg/ml of ampicillin and 34 μg/ml of chloramphenicol at 37 °C. Once the A₆₀₀ reached 0.6, protein expression was induced by the addition of 1 mM isopropyl 1-thio-D-galactopyranoside. Cells were incubated overnight at 18 °C and then harvested by centrifugation and re-suspended in buffer A (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 10 mM imidazole, 10% (v/v) glycerol, and 1 mM β-mercaptoethanol). Cells were sonicated or subjected to French press and the protein was purified using a column of Ni²⁺-nitrilotriacetic acid-agarose beads (Qiagen). After loading the sample, the column was washed in three steps using buffer A with 20, 50, and 100 mM imidazole consecutively. The protein was eluted with 500 mM imidazole. The purest fractions were pooled and dialyzed in buffer B (50 mM Tris-HCl, pH 7.5, 25 mM NaCl, 10% (v/v) glycerol, and 1 mM DTT), and subjected to ion exchange chromatography using a MonoQ column. The purest fractions were pooled once more and subjected to size exclusion chromatography using a Superdex 200 HR 10/300 GL (Amersham Biosciences) column in buffer C (10 mM HEPES, pH 7.5, and 300 mM NaCl). Fractions were pooled, concentrated, and quantified using absorbance with the following extinction coefficients: 16,515 M⁻¹ cm⁻¹ for PfClpR(P61-E244) and 19,495 M⁻¹ cm⁻¹ for PfClpR(S49-E244) obtained from the ProtParam Tools on the ExPASy Server (38).

**Crystallization, Data Collection, and Processing—** The hanging-drop vapor-diffusion method of crystallization was used to obtain PfClpR crystals. Initially, commercial screening kits, such as Hampton Index screen and Nextal JCSG+ Suite, were used to find a starting condition for crystallization. Single crystals were obtained for two different constructs of PfClpR: PfClpR(P61-E244) and PfClpR(S49-E244). Crystals of the first construct were obtained in crystallization conditions containing 0.1 M Tris-HCl, pH 8.5, and 20% (v/v) ethanol. Crystals of the second construct were obtained in a crystallization condition containing 0.25 M calcium acetate, pH 8.6, 7.5% PEG 20K, and 7.5% PEG MME500. 1 μl of crystallization solution was added to 1 μl of ~14 mg/ml of PfClpR protein. The reservoir contained 700 μl of crystallization solution. Crystals typically grew in 2 days at room temperature. We were not able to grow crystals in cryoprotectant condition, hence, single crystals were transferred gradually to new drops with higher concentrations of glycerol or ethylene glycol. Several such crystals were tested for x-ray diffraction. The best crystals were flash-frozen in liquid nitrogen in the presence of 25% ethylene glycol for PfClpR(P61-E244) and 10% glycerol for PfClpR(S49-E244). Diffraction data at 2.8-Å resolution were collected for crystals of PfClpR(P61-E244) at the beamline 19-ID at the Advanced Photon Source, Argonne National Laboratory. Diffraction data at 2.9-Å resolution were collected for crystals of PfClpR(S49-E244) at the CHESS A1 beamline. Data were processed and scaled with HKL2000 (39). Phaser program (53) from the CCP4 crystallographic program suite was used to search for 7 copies within the asymmetric unit. The program found 3 copies and the symmetry between two subunits was used to generate the heptamer. PfClpR structure was visualized and built using COOT (40) and refinement was done with REFMAC (41) and Phenix (42). Crystallographic details and refinement statistics are summarized in Table 1. The coordinates have been deposited in the protein structure database under the accession numbers 4GM2 for PfClpR(P61-E244) and 4HNK for PfClpR(S49-E244).

**RESULTS**

**Localization of PfClpR in Transgenic Parasites—** Initially, we carried out experiments to confirm the apicoplast localization of PfClpR. According to PlasmoDB (43), the first 49 residues of PfClpR are predicted to harbor the hydrophobic N-terminal signal sequence and the apicoplast-targeting transit peptide sequence (Fig. 1A). A GFP-targeting approach was employed to study the localization of PfClpR in the parasite by expressing plasmid-borne full-length PfClpR protein with a C terminally fused GFP in transgenic parasites (Fig. 1A). Immunoblot analysis using GFP-specific antibodies confirmed the presence of the fusion protein in trophozoite-stage transgenic parasites and not in the wild type parasite lines (Fig. 1B). Western blot against HRPII was used as loading control (Fig. 1B). Fluorescent microscopic images of the transgenic parasites at early and late trophozoite stages showed that the PfClpR-GFP fusion protein was localized in an apicoplast-like cellular organelle (Fig. 1C, green). To ascertain that the GFP fusion protein is not in the parasite mitochondria, the transgenic parasites were co-stained with MitoTracker dye, the mitochondria-specific live stain. The
The Structure of ClpR from P. falciparum

GFP fluorescence pattern (Fig. 1C, green) was observed in close association but distinct from mitochondrial staining (Fig. 1C, red). The parasite nuclei were stained with DAPI (blue) and slides were visualized by fluorescence microscope. Scale bar represents 2 μm. D, localization of PfClpR by immunoelectron microscopy. Ultra-thin sections of transgenic P. falciparum parasites expressing PfClpR-GFP were labeled with anti-GFP antibody and gold-labeled secondary antibody. Labeling was observed in the apicoplast, which has the characteristic four membranes. Scale bar represents 250 nm.

To further ascertain the localization of PfClpR-GFP fusion protein in the apicoplast, immunoelectron microscopic studies were carried out. Labeling with anti-GFP antibody followed by gold-labeled secondary antibody demonstrated the localization of the fusion protein in the lumen of the apicoplast, which was identified as a four-membrane structure in the parasite (Fig. 1D). From these experiments, we concluded that PfClpR is indeed localized to the apicoplast.

Analysis of the Transcription and Translation of PfClpR in the Asexual Blood-stage Parasites—The transcriptional pattern of the PfclpR gene in different developmental stages of the asexual blood-stage parasites was assessed by quantitative real-time PCR using total RNA extracted from tightly synchronized parasite cultures at early ring (ER), late ring (LR), trophozoite (T), late trophozoite (LT), early schizont (ES), and late schizont (LS) stages (8, 16, 24, 32, 40, and 48 h after invasion, respectively). Stage-specific expression of PfClpP, eba-175, and falcipain-2 were also analyzed. B, Western blot analyses of highly synchronized parasites at ring (lane 1), trophozoite (lane 2), and schizont (lane 3) stages with anti-PfClpR and anti-PfClpP antibodies. A parallel blot was probed with anti-HRP II antibodies to show equal loading. C, molecular mass of the native PfClpR and PfClpP protease complexes in the parasite. The parasite lysate was fractionated over a Superose 6 column and eluates were analyzed by Western blot using anti-PfClpR and anti-PfClpP antibodies. The elution patterns of the molecular mass standards are indicated: thyroglobulin (660 kDa), ferritin (450 kDa), and aldolase (158 kDa).

Western blot analysis of total parasite lysates detected a band of ~23 kDa in the trophozoite and schizont-stage parasites using anti-ClpR antibodies (Fig. 2B). This band corresponds to the calculated molecular mass of the processed PfClpR corresponding to residues ~49–244 (27) (Fig. 1A). No band was stage parasites (Fig. 2A). Maximum transcription of PfclpP was also observed in the same parasite stages. As control, the same set of samples were also analyzed by quantitative PCR for two other P. falciparum genes, erythrocyte binding antigen-175 (eba-175) and the cysteine protease falcipain-2. As expected, eba-175 transcript levels were found to be maximum in samples from late schizont-stage parasites, whereas falcipain-2 showed maximum transcript levels in trophozoite-stage parasites (Fig. 2A).
detected using preimmune sera; in addition, the anti-PfClpR antibodies did not react with the lysate of uninfected RBCs (data not shown). Western blot against HRP II was used as loading control (Fig. 2B). As expected anti-PfClpP antibodies detected a band in the trophozoite- and schizont-stage parasites. Both the transcription and translation analyses results suggest that PfClpR is expressed in blood-stage parasites at trophozoite and schizont stages, which is similar to that of PfClpP (28).

**PfClpR Forms a Multisubunit Complex**—To determine the size of the native protein complex of PfClpR, the parasite lysate was subjected to size exclusion chromatography using a Sepharose-6 column. The eluted samples were analyzed by Western blot using anti-PfClpR, as well as anti-PfClpP antibodies. The native PfClpR was detected in fractions of molecular mass of ~440 kDa (Fig. 2C). On the other hand, PfClpP eluted as a much larger complex of about 660 kDa. We had found earlier that the mature, processed PfClpP and PfClpR have similar masses of about 25 kDa (27, 28). Hence, PfClpP seems to be part of a larger complex that, unexpectedly, does not include PfClpR. This suggests that PfClpP and PfClpR do not form a stable protein complex in the parasite or that they form separate complexes.

**Crystallization and Structure Determination of PfClpR**—Because no structure has been reported for ClpR from any organism and to better understand the function of this inactive protease, we carried out experiments to solve the x-ray structure of PfClpR. We had shown earlier that the mature PfClpR consists approximately of residues 49–244 (Fig. 3A). Based on the crystal structure of ClpP from different species, residues from Ser49 to Ile68 of PfClpR might form flexible axial loops, which would present a challenge to crystallization. Therefore, attempts were made to crystallize both PfClpR(P61-E244) as well as PfClpR(P61-E244). The constructs contained a His6 tag and a cut site for the tobacco etch virus protease (MGSSHSS-GRENLYFQG) prior to Ser49 or Pro61; these residues were not removed in the crystallization trials. Single crystals were obtained and diffracted at 2.8-Å resolution for PfClpR(P61-E244) and 2.9 Å for PfClpR(S49-E244) (Table 1). The structure was solved by molecular replacement using PfClpP protomer as template. When PfClpP heptamer was used as a template, the molecular replacement program failed to find a solution because PfClpP and PfClpR heptamers have different conformations (see below).

**Overall Structure of PfClpR**—Using gel filtration, analytical ultracentrifugation, and electron microscopy, we previously confirmed that PfClpR exists mainly as a heptamer in solution (27). For both constructs, PfClpR was also crystallized as a 7-fold symmetric single ring of ~106 Å in diameter and a central pore of ~30 Å (Fig. 3, B–D). The PfClpR(S49-E244) asymmetric unit included two 7-fold heptamers (chains A to N) related by a translational symmetry (not a double ring structure). The asymmetric unit of PfClpR(P61-E244) includes a 7-fold heptamer (chains A to G) with the same packing as for PfClpR(S49-E244).

Overall, PfClpR(P61-E244) and PfClpR(S49-E244) structures are similar (Fig. 4) showing that the presence of ethanol in the crystallization condition of PfClpR(P61-E244) does not drastically affect the PfClpR structure. The root mean square deviation for the protomers is 0.66 Å and for the heptamers is 1.4 Å. Below, we will mainly refer to PfClpR(P61-E244).

This first atomic structure of a ClpR is similar to other ClpP structures with a root mean square deviation for PfClpR(P61-E244) of 1.5 Å to PfClpP and 2 Å to Escherichia coli ClpP (EcClpP). Each PfClpR subunit has an α/β-fold made up of six

### TABLE 1
Crystallographic data collection and refinement statistics

<table>
<thead>
<tr>
<th></th>
<th>PfClpR(P61-E244)</th>
<th>PfClpR(S49-E244)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>C2</td>
<td>C2</td>
</tr>
<tr>
<td>Unit cell a, b, c (Å)</td>
<td>182.89, 102.16, 96.41</td>
<td>178.79, 105.85, 186.4</td>
</tr>
<tr>
<td>α, β, γ (°)</td>
<td>90.00, 114.53, 90.00</td>
<td>90.00, 118.10, 90.00</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.979</td>
<td>0.979</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>50.2-5 (2.85-2.80)</td>
<td>50.2-2.9 (2.95-2.90)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.8 (3.8)</td>
<td>3.3 (3.0)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8 (100)</td>
<td>96.7 (87.8)</td>
</tr>
<tr>
<td>NCS redundancy</td>
<td>7-fold</td>
<td>2x7-fold</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.051 (0.755)</td>
<td>0.075 (0.764)</td>
</tr>
<tr>
<td>Rwork</td>
<td>27.46 (1.55)</td>
<td>20.07 (1.75)</td>
</tr>
<tr>
<td>Number of reflections used for refinement</td>
<td>39733</td>
<td>121188</td>
</tr>
<tr>
<td>Number of residues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amino acid residues</td>
<td>1202</td>
<td>2353</td>
</tr>
<tr>
<td>Water molecules</td>
<td>69</td>
<td>70</td>
</tr>
<tr>
<td>Glyceraldehyde</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>Rmerge</td>
<td>0.199</td>
<td>0.200</td>
</tr>
<tr>
<td>Rwork</td>
<td>0.229</td>
<td>0.244</td>
</tr>
<tr>
<td>Average B factors (Å2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>102.3</td>
<td>89.2</td>
</tr>
<tr>
<td>Solvent</td>
<td>80.4</td>
<td>67.0</td>
</tr>
<tr>
<td>Ligand (Glyceraldehyde)</td>
<td>-</td>
<td>94.7</td>
</tr>
<tr>
<td>Wilson B factors (Å2)</td>
<td>83.5</td>
<td>71.9</td>
</tr>
<tr>
<td>Ramachandran plot (MolProbity)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Favorable regions</td>
<td>96.6%</td>
<td>97.3%</td>
</tr>
<tr>
<td>Additional Allowed</td>
<td>3.1%</td>
<td>2.7%</td>
</tr>
<tr>
<td>Outliers</td>
<td>0.3%</td>
<td>0.04%</td>
</tr>
<tr>
<td>RMSD bond lengths/angles</td>
<td>0.0048Å/0.915°</td>
<td>0.0068Å/1.126°</td>
</tr>
<tr>
<td>PDB ID</td>
<td>4GM2</td>
<td>4HNK</td>
</tr>
</tbody>
</table>

Values in parentheses refer to the highest resolution shell. The various crystallographic parameters are defined as follows: Rmerge = Σ||Fcalc|| - ||Fo||/ΣFo, where Fo is the intensity of the i-th observation, <I> is the mean intensity of the reflection, and the summation extends over all data. Rmerge = Σ||Fobs|| - ||Fc||/ΣFc, where Fobs and Fcalc represent, respectively, the observed and calculated structure factors. Rmerge was calculated using 5% of the observed reflections excluded from refinement. Excluded data were randomly selected.

**FIGURE 3. Domain organization of PfClpR.** A, sequence alignment EcClpP, PfClpP, and PfClpR using ClustalW2 (49) and visualized with ESPript (50). Residues of the Ser-His-Asp catalytic triad are indicated by a red asterisk. The R-motif is highlighted in purple. The handle region and axial loop are boxed in green and red, respectively. The arrows indicate the start site of the crystallized PfClpR proteins (following the His-tag tobacco etch virus tag). PfClpP and EcClpP secondary structures are shown on the top and bottom of the sequence alignment, respectively. Residues that are identical in the three sequences are highlighted in blue, whereas those that are highly similar are in red. B, ribbon diagram of PfClpR(P61-E244) protomer. The head domain and handle region are indicated. The R-motif is colored purple and the residues in the inactive catalytic triad (Gly-Asn-Asp) are shown in green and boxed. C, ribbon diagram of PfClpR(P61-E244) heptamer. In the top two panels, each subunit is colored differently. In the bottom panel, the subunits are all in gray, whereas the R-motif and the Gly-Asn-Asp sequence are colored as described in B, D, shown are overlays of the seven PfClpR(P61-E244) protomers of the heptameric ring in two different orientations. Arrows point to the handle region (blue arrow) and the R-motif (red arrow).
The Structure of ClpR from P. falciparum

repeats of the α/β unit (αA/β1, αB/β2, αC/β3, αD/β4/β5, αE/β7, αF/β8) forming the head domain and an additional protruding α/β unit (αE/β6) forming the handle region (Fig. 3, A and B). Residues Pro61, Glu174 and Lys211-E244 comprise the head domain, whereas residues Ser175-Glu210 form the handle region. As is typically found in other ClpPs, the head domain comprises 8 β strands arranged into two layers of β sheets. On one side, the β sheets pack against a layer of α helices from the same subunit whereas, on the other side, they pack against a layer of α helices of the neighboring subunit within the heptamer (Fig. 3C).

We were not able to build the axial loop (Ser49-Ile60) of PfClpR(S49-E244) because of weak electron density. Also, the handle region is either absent or is missing some residues in all the subunits of this construct. For PfClpR(P61-E244), the final models of subunits A and E include all residues, Pro61-Lys244. Residues Ser175-Lys192, Gln174-Ile189, Gln174-Glu193, Gln174-Asn186, and Tyr176-Ile187 in the handle region are missing in subunits B, C, D, F, and G, respectively. As shown in Figs. 3D and 5, the two handle regions that we could trace in PfClpR(P61-E244) have different orientations due to crystal packing.

The Ser-His-Asp catalytic triad is missing in PfClpR and is replaced by residues Gly148-Asn173-Asp223, which we will call the inactive catalytic triad, which are located in a cleft between the head domain and handle region similar to the location of the catalytic triad in ClpPs (Figs. 3, B and C, and 6A). The head domain of PfClpR includes an insertion formed by the 10 residues Asp111-Gly120 not present in ClpPs (Fig. 3A); we name Asp109-Gly120 as the R-motif. The equivalent segments in A. thaliana and cyanobacterium Synechococcus have been modeled as an extension of helix C followed by a loop (24, 25). In our PfClpR structure, the R-motif adopts a structured β turn that is largely equivalent in all 7 subunits (Fig. 3D) and protrudes from the head domain without affecting the α/β-fold (Fig. 3D). It also includes an extension of helix C (residues NG). The R-motif is adjacent to the presumed active site and faces the internal chamber of the heptamer (Fig. 6A).

In ClpP, each subunit has a hydrophobic groove forming a continuous substrate-binding surface leading to the catalytic triad (44, 45). An examination of surface topology and surface potential of PfClpR reveals that this groove is present in the PfClpR structure and is formed by residues Ile121, Ile149, Ile152, Leu172, Asn173, Ser175, Val201, Ile205, and Leu220 (Fig. 6, A and B). The presence of the R-motif results in the formation of a deep cleft without significantly restricting accessibility to the groove from the internal chamber of the PfClpR heptamer (Fig. 6B). This is clearly observed if the PfClpR structure is overlaid with the structure of the Helicobacter pylori ClpP protease.
that is in complex with a heptapeptide NVLGFTQ (PDB code 2ZL2). The heptapeptide is found to nicely fit into that deep cleft of PfClpR (Fig. 6B). The amino acid sequence of the R-motif is highly conserved in the genus Plasmodium with a consensus sequence of: D(L/I)(D/E)NNKI(I/V)NLNG. This suggests that the motif may play an essential role related to the function of this protein.

The Oligomerization of PfClpR—The PfClpR(P61-E244) heptamer is stabilized by extensive interactions among the head domains with about 30% of the total surface area making contacts within the heptamer. Compared with EcClpP and PfClpP structures, the PfClpR(P61-E244) heptamer is compressed by about 10 Å along the z axis and is wider by about 14 Å along the x axis (Fig. 7A). This is the result of a twist of each subunit in the PfClpR(P61-E244) heptamer by about 15° outwards leading to a more open structure of the PfClpR(P61-E244) heptamer (Fig. 7, B and C). As a consequence of the open shape of PfClpR(P61-E244), the surface area of the heptameric complex is larger than that of PfClpP or EcClpP: ~56,630 Å² for PfClpR(P61-E244), ~46430 Å² for PfClpP, and ~50,970 Å² for EcClpP. This observation leads us to conclude that a flat ring of PfClpR is probably not favorable for an interaction with the PfClpP ring, unless a conformational change in PfClpR is induced. This might explain our inability to observe a stable PfClpRP complex using the purified proteins (27).

DISCUSSION

In eukaryotes, the 26 S proteasome plays an essential role in controlling the levels of key regulatory proteins and in degrading abnormal polypeptides; it also plays a key role in cell cycle
progression and the regulation of numerous transcription factors (46). In prokaryotes, mitochondria, and plant chloroplasts, these tasks are carried out by the Clp chaperone-protease system (47). We had earlier identified and characterized the Clp chaperones and proteases in *P. falciparum* (27, 28, 48). In this study, we have experimentally confirmed that, like *Pf*ClpP, *Pf*ClpR is localized in the apicoplast of the parasite (Fig. 1, C and D). Furthermore, both proteins were found to be expressed in blood-stage parasites at trophozoite and schizont stages (Fig. 2, A and B, and Ref. 27 and 28) suggesting that they might function together. However, it was surprising to find that the two proteins do not co-elute when a *P. falciparum* lysate is fractionated (Fig. 2C).

We also determined the crystal structure of *Pf*ClpR protein, which exists as a heptamer in solution (27). Although the structure of *Pf*ClpR monomer is similar to that of *Pf*ClpP (Figs. 3D and 7C), the dimensions and shapes of *Pf*ClpP and *Pf*ClpR heptameric rings are different (Figs. 4 and 7, A and B). The structure we obtained for *Pf*ClpR showed a flattened heptameric complex compared with that of ClpP (Fig. 7A) that is unlikely to be able to interact stably with an equivalent heptameric *Pf*ClpP by the intercalation of the handle regions as typically observed in ClpP tetradecamers. These results correlate with the presence of separate *Pf*ClpR and *Pf*ClpP complexes in the parasite as observed in the size exclusion fractionation experiment (Fig. 2C). However, at this stage, we cannot exclude the possibility that a AAA+ chaperone or some other cofactor might induce a conformational change in *Pf*ClpR or *Pf*ClpP to allow for the formation of a transient tetradecameric complex consisting of two homo-heptameric rings *Pf*ClpR, *Pf*ClpP. The presence of a substrate might also stabilize a *Pf*ClpR, *Pf*ClpP complex. Alternatively, a functional hetero-heptameric *Pf*ClpRP complex in the apicoplast may be present at very low levels to be detected by the fractionation experiment of Fig. 2C; however, we have not yet been able to form such a complex in vitro. It should be noted that *Pf*ClpP formed a tetradecamer in the crystal but a heptamer in solution (27), indicating a concentration-dependent oligomerization.

If indeed *Pf*ClpR and *Pf*ClpP do not form a complex in the apicoplast of *Plasmodium*, then this suggests the presence of a separate inactive *Pf*ClpR complex with distinct activities from *Pf*ClpP. *Pf*ClpR might indirectly regulate the activity of *Pf*ClpP by competing for interactions with cognate AAA+ proteins. Alternatively, the *Pf*ClpR ring might have a chaperone-like activity. This is suggested by the presence of the deep hydrophobic clefts formed by the R-motif and the substrate-binding surface leading to the inactive catalytic triad (Fig. 6). Exposed hydrophobic segments in unfolded or misfolded proteins can bind in these clefts. However, it should be pointed out that these possible alternative functions for *Pf*ClpR would be specific for apicomplexa and would not be consistent with the fact that hetero-oligomeric complexes have been observed in plant plastids and cyanobacteria.

It should be noted that, whereas the R-motif seems to be present in most of the cyanobacterial and plant ClpRs, the amino acid sequence of the motif is not conserved across species. Furthermore, some ClpRs have other insertions as well. For example, SeClpR has an insertion between B1 and aB, which is not present in plant or *Plasmodium* ClpRs (24, 25, 27). The role of such insertions is not known; these insertions might be required to further optimize or modify the function of the protein in a species-specific manner.

In conclusion, our study provides the first structure of a ClpR protein. The structure helps to further our understanding of the role that ClpR plays in different organisms. However, the function of *Pf*ClpR in *Plasmodium* and the possible presence of a *Pf*ClpRP complex remains enigmatic and requires additional studies.

Acknowledgments—We are grateful to Alan Cowman (Walter and Eliza Hall Institute of Medical Research, Australia) for pRALI vector, and Guy Schiehser and David Jacobus (Jacobus Pharmaceuticals, NJ, USA) for the drug WR99210. We thank Wandy Beatty for help with the immunoelectron microscopic studies and the Rotary Blood Bank, New Delhi, for providing the human RBCs.

REFERENCES


