Functional genetic evaluation of DNA house-cleaning enzymes in the malaria parasite: dUTPase and Ap4AH are essential in Plasmodium berghei but ITPase and NDH are dispensable

Hirdesh Kumar a,b,c, Jessica Kehrer a,b, Mirko Singer a, Miriam Reinig a, Jorge M. Santos d, Gunnar R. Mair e,f and Friedrich Frischknecht e

a Integrative Parasitology, Department of Infectious Diseases, University of Heidelberg Medical School, Heidelberg, Germany; b Institut de Medicina Molecular, Faculdade de Medicina da Universidade de Lisboa, Lisbon, Portugal; c Department of Biomedical Sciences, 2008 College of Veterinary Medicine, Iowa State University, Ames, IA USA

dUTPase; ITPase; NuDiX; enzymes; drug targets; Malaria; house-cleaning enzymes; Plasmodium berghei; dNTP; pyrophosphohydrolase

ABSTRACT

Background: Cellular metabolism generates reactive oxygen species. The oxidation and deamination of the deoxynucleoside triphosphate (dNTP) pool results in the formation of non-canonical, toxic dNTPs that can cause mutations, genome instability, and cell death. House-cleaning or sanitization enzymes that break down and detoxify non-canonical nucleotides play major protective roles in nucleotide metabolism and constitute key drug targets for cancer and various pathogens. We hypothesized that owing to their protective roles in nucleotide metabolism, these house-cleaning enzymes are key drug targets in the malaria parasite.

Methods: Using the rodent malaria parasite Plasmodium berghei we evaluate here, by gene targeting, a group of conserved proteins with a putative function in the detoxification of non-canonical nucleotides as potential antimalarial drug targets: they are inosine triphosphate pyrophosphatase (ITPase), deoxyuridine triphosphate pyrophosphatase (dUTPase) and two NuDiX hydrolases, the diadenosine tetraphosphate (Ap4A) hydrolase and the nucleoside triphosphate hydrolase (NDH).

Results: While all four proteins are expressed constitutively across the intraerythrocytic developmental cycle, neither ITPase nor NDH are required for parasite viability. dutpase and ap4ah null mutants, on the other hand, are not viable suggesting an essential function for these proteins for the malaria parasite.

Conclusions: Plasmodium dUTPase and Ap4A could be drug targets in the malaria parasite.

1. Introduction

Malaria is a devastating parasitic disease killing around 400,000 people every year, mostly young children. The occurrence of drug resistance and the absence of an efficient vaccine demand a better understanding of the parasite’s biology to aid the identification of suitable and novel drug and vaccine targets [1–4]. Plasmodium, the causative agent of the disease, is haploid throughout most of its life cycle. This protozoan uses multiple rounds of endomitosis followed by cellular morphogenesis to produce unicellular progeny in a process called schizogony [5]. This strategy is employed by the parasite to increase the population at specific time points during its life cycle: in the red blood cell a single P. falciparum merozoite produces up to 32 daughter cells [6]; following transmission to the mosquito an oocyst can generate hundreds of sporozoites [7] while in the liver thousands of merozoites emerge from a hepatocyte infected with a single sporozoite [8]. During sexual development, the male gametocyte produces eight motile gametes in a matter of minutes following transmission to the mosquito vector [9].

The haploid nature of the Plasmodium genome and its rapid DNA replication cycles perhaps pose an increased risk for the mis-incorporation of non-canonical nucleotides into the DNA; yet the enzymes preventing such DNA damage in the malaria parasite are largely unknown. Intracellular nucleotide pools are essential DNA building blocks, serve as energy storage molecules, and act as cofactors and regulators in various metabolic and signal transduction pathways. Reactive oxygen species produced during cellular redox reactions can oxidize the DNA precursor pool and produce so-called non-canonical deoxynucleoside triphosphates (dNTPs). For example oxidative deamination of the nitrogenous base produces deoxyinosine triphosphate (dTTP) from deoxyadenosine triphosphate (dATP), deoxyuridine triphosphate (dUTP) from deoxycytosine triphosphate (dCTP) and deoxyxanthosine triphosphate (dXTP) from deoxyguanosine triphosphate (dGTP). Thymine lacks a free amino group and therefore is not oxidatively deaminated. Oxidation of purine bases on the other hand results in nucleotides like 2-oxo-dATP, 8-oxo-dATP, and 8-oxo-dGTP.

Such modified nucleotide analogs can be incorporated into DNA during replication [10,11]. When recognized by the DNA
repair system, single strand breaks are introduced which need repairing. As a consequence, cellular growth can slow down or the cell may die due to the accumulation of deleterious mutations and double strand breaks [11].

To prevent DNA damage, cells have evolved housecleaning or DNA sanitation enzymes to hydrolyze toxic nucleotides into their non-toxic monophosphate forms, which are poor substrates for their respective nucleoside kinases and thus not phosphorylated to their toxic nucleotide forms. Based on structural features, sanitation enzymes belong to four superfamilies:

ITPases (inosine triphosphate pyrophosphatase) are defined by the Ham1 (6-n-HydroxyaAMinopurine sensitive) domain consisting of a long central beta-sheet forming the floor of the active site [12];

dUTPases (deoxy-uridine triphosphatase) which present a trimeric fold, each with an eight-stranded jelly-roll beta barrel [13];

the NuDX (nucleoside diphosphate linked to some other moiety, X) superfamily is defined by the NuDX box domain G-x(S)-E-x(S)-[UA]-x-R-E-x(2)-E-E-x-G-U, where U is a hydrophobic residue and x any residue [14];

and finally, the fourth group contains all-a-helical NTPases which are active against dNTPs as well as dNDPs with the generation of dNMPs as an end product [15].

The four groups share no sequence homology and are specific for their respective substrates. ITPases hydrolyze ITP and XTP as well as their deoxy forms (dTTP, dXTP) into their respective monophosphates [16,17], which prevents the incorporation of dITP and dXTP into DNA. Escherichia coli and yeast mutants that lack ITPase suffer increased sensitivity to the mutagen 6-N-hydroxymaminopurine (HAP) due to the mispairing of dCTP or dITP leading to A:T to G:C and G:C to A:T transitions [18,19].

The human protein NUDT1/MTH1, which catalyzes the same reaction, is involved in maintaining the health of yeast cells. The enzyme MutT hydrolyzes both 8-oxo-GTP and 8-oxo-dGTP and thus performs house-cleaning at both DNA and RNA level. In mutT depleted E. coli a 100 to 10,000 times increase in occurrence of AT to CG transversions was reported [34,35] and this unidirectional mutation causes GC enrichment in chromosomal DNA [36]. The human protein NUDT1/MTH1, which catalyzes the same reactions, is key to the survival of cancer cells [37].

All-a-helical NTPases are pyrimidine-specific house-cleaning enzymes. They include the Mycobacterium tuberculosis (Mtbo) MazG protein that degrades 5-OH-dCTP and thus prevents CG to TA mutations [38] and mouse RS21-C6 which is a mammalian 2-deoxycytidine 5′-triphosphate pyrophosphohydrolase that has a preference for deoxynucleoside triphosphates and cytosine bases [39].

The haploid nature of the Plasmodium genome and its rapid DNA replication cycles pose perhaps an increased risk for DNA damage; yet the enzyme repertoire preventing the incorporation of non-canonical dNTPs into the DNA of the malaria parasite is largely unknown and its requirement for cell viability has not been explored. Here, we bioinformatically identify sanitation enzyme homologs in the rodent malaria model P. berghei and – through reverse genetics – define their importance for parasite life cycle progression in the mouse host and mosquito vector.

2. Materials and methods
2.1. Ethics statement
All animal experiments were performed according to the FELASA and GV-SOLAS standard guidelines. German authorities (Regierungspräsidium Karlsruhe) approved the animal experiments.

2.2. Bioinformatics analyses
Human nucleotide sanitation enzyme sequences were retrieved from the Universal Protein Resource (http://www.uniprot.org/) repository (Table 1) and searched against the Plasmodium database (www.PlasmoDB.org) using the Domain Enhanced Lookup Time Accelerated-Basic Local Alignment Search Tool (DELTA-BLAST) of protein-BLAST search. The default algorithm parameters used in all searches were as follows: algorithm: DELTA-BLAST (Domain Enhanced Lookup Time Accelerated-Basic Local Alignment Search Tool); max target sequence: 500; expected threshold: 10; word size: 3; scoring matrix: BLOSUM62; threshold: 0.05.

ClustalW alignments were performed at http://www.genome.jp/tools-bin/clustalw. Alignment files were formatted with box shade at https://embnet.vital-it.ch/software/BOX_form.html.

Percentage identity matrices were performed at https://www.ebi.ac.uk/Tools/msa/clustalo/. The information for ortholog groups for dUTPase and ITPase was retrieved from the PlasmoDB database (www.PlasmoDB.org). Both ortholog groups were individually searched in the OrthoMCL (http://orthomcl.org/orthomcl/) database to find similar sequences in the genomes of human, yeast and of various Plasmodium species. The
orthologs in *Plasmodium* and human were retrieved in fasta format and submitted online to Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/) to perform multiple sequence alignments. Clustal Omega uses seeded guide trees and hidden Markov model (HMM) profile-profile techniques to perform a global alignment of sequences. The input protein sequences were submitted with the following parameters: dealign input sequences: no, mbed-like clustering guide-tree: yes, mbed-like clustering iteration: yes, number of combined iterations: 0, max guide tree iterations: −1, max hmm iterations: −1, order: input. The percentage identity matrices and globally aligned sequences were downloaded.

### Table 1. Repertoire of house-cleaning enzymes in *Plasmodium berghei*.

| Homo sapiens | Plasmodium berghei |
|--------------|--------------------|
| Gene | uniprot.org | Gene | Protein | plasmodb.org | Pfam domain | E-value |
| DUT | P33316 | deoxyuridine 5’-triphosphate nucleotidohydrolase | dUTPase | PBANKA_0921300 | dUTPase | 1.70e-16 |
| ITPA | Q9BY32 | Ham1-like protein | ITPAse | PBANKA_0618300 | Ham1p_like | 1.10e-38 |
| NUDT1; MTH1 | P36639 | Nucleoside diphosphate hydrolase | NDH | PBANKA_1361900 | NUDIX | 1.00e-06 |
| NUDT2; AP4AH | P50583 | bis(5’-nucleosyl)-tetraphosphatase [asymmetrical] | Ap4AH | PBANKA_1235300 | NUDIX | 5.20e-14 |

![Figure 1. Overview of deoxyribonucleotide metabolism. A schematic of nucleotide metabolism and the enzymes involved. The structures of canonical and non-canonical nucleotides are shown in blue and red, respectively. House-cleaning NTP pyrophosphatases are marked in colored boxes. NDH: NuDiX hydrolases; ITPase: Inosine triphosphate pyrophosphatase; Ap4AH: Diadenosine tetraphosphate hydrolase; and dUTPase: Deoxyuridine-triphosphatase. '?' next to Ap4AH indicates the possible conserved function of 8-oxo-dGTP hydrolysis in *Plasmodium*. Among these house-cleaning enzymes, dUTPase performs the dual house-cleaning and house-keeping functions as follows: dUTPase prevents entry of dUTP to DNA (house-cleaning) and provides dUMP for the synthesis of dTTP (house-keeping). The enzyme missing in *Plasmodium* is shown with broken line. NDPK: Nucleoside-diphosphate kinases; DHFR-TS: Dihydrofolate reductase-thymidylate synthase.](image1)

#### 2.3. GFP-tagging of *Plasmodium berghei* house-cleaning genes

We generated transgenic parasite lines expressing full-length C-terminally GFP-tagged proteins from episomal plasmids (dUTPase and ITPase) or through the integration of linearized vectors into the corresponding endogenous loci (Ap4AH and NDH) as outlined in (Figure 2(d-g)). *P. berghei* WT parasites were transfected as described [40]. Protein expression was monitored by microscopy and Western blotting. For fluorescence microscopy, a drop of tail blood was mixed with Hoechst 33,342 on a slide that was immediately covered with the coverslip and visualized under Zeiss 200 M Axiovert widefield (63x)
microscope. Image processing was performed with ImageJ. For Western blotting blood-stage parasites were released from red blood cells with saponin and resuspended in RIPA buffer (5 M NaCl, 0.5 M EDTA pH 8, 1 M Tris pH 8, 1% NP-40, 10% Na-deoxycholate, 10% SDS, dH₂O) containing protease inhibitor (complete, Mini, EDTA-free, Roche). The mixture was loaded on to a Mini-Protein TGX Gel, 4–15% (BioRad) for electrophoresis and transferred onto a nitrocellulose membrane using the Trans-Blot Turbo Transfer System (BioRad). The membrane was blocked with 5% milk in Tris (pH 7.5) supplemented with 1% Tween 20

Figure 2. Expression analyses of four house-cleaning enzymes in the blood stage. (a). Protein sequence models (predicted using SMART) are drawn to scale with functional domains indicated in gray. Common enzyme names and gene IDs are given above the domains. (b). Sanitation enzyme conservation in Plasmodium spp. (c). Percentage identity matrices of different house-cleaning enzymes in P. berghei ANKA (Pb), P. yoelii yoelii 17X (Py), P. falciparum 3D7 (Pf) and P. vivax Sal-1 (Pv). (d-g). GFP-tagging approach of house-cleaning enzymes using circular plasmids (d, e) or linear DNA (f, g) for genome integration (left panel). Middle panel: live fluorescence expression images in the indicated blood stage parasites depicting merged Hoechst DNA and GFP signal; scale bar = 5 μm. Right panels: western blot analyses of blood stage protein samples. Predicted sizes of GFP-fusion proteins are given below each blot in kDa. The expected size corresponds to individual protein +27 kDa (GFP-tag). We observed no protein degradation in the two essential enzymes i.e. dUTPase and AP4AH. We suspect this degradation (also the smear in ITPase) appeared due to the in vitro preparatory steps for the western-blots. Importantly we find full-length expression for each fusion protein. Please note that control blots (WT parasites and GFP alone) are not shown in figure but were used in all experiments.
(TBST) for 2 h at RT and probed with an anti-GFP antibody (1:5000 dilution, Clontech Laboratories) diluted in 5% milk in TBST overnight at 4°C. The following day, the membrane was washed 3 times with TBST and incubated with anti-mouse (1:10,000 dilution, GE Healthcare) secondary antibody in 5% milk in TBST for 1 h at RT. Proteins were detected using ECL substrate (Pierce).

2.4. RT-PCR analyses

Total RNA was extracted from different stages of the *Plasmodium* life cycle and cDNA was prepared according to manufacturer’s instructions (Invitrogen). Transcripts were detected with the oligonucleotide primers that bind specifically to the gene. All primer details are given in Supplementary Table 1.

2.5. Generation of *Plasmodium berghei* knockout parasites

In order to delete individual *P. berghei* house-cleaning genes (dUTPase: PBANKA_0921300, ITPase: PBANKA_0618300, Ap4AH: PBANKA_1235300, NDH: PBANKA_1361900), 5′UTR and 3′UTR regions were amplified from genomic DNA of mixed blood stages of *P. berghei* strain ANKA by polymerase chain reaction (PCR) using Phusion High-Fidelity DNA Polymerase (New England Biolabs, USA) with primer sets as shown in Table S3; fragments were digested with restriction endonucleases as indicated in Table S3 and ligated either side of the human dihydrofolate reductase or a yeast bifunctional enzyme cytosine deaminase fused to an uridyl phosphoribosyl transferase (hdhfr-yFCU) as selection cassettes [41]. Linearized vectors were transfected into purified schizonts of *P. berghei* ANKA using standard protocols [40]. Transfected parasites were selected with pyrimethamine dissolved in drinking water (conc. 0.07 mg/mL) starting at 24 h post transfection. Parasitemia was monitored from day 7 onwards. Parasites were harvested above 2% parasitemia and cloned by limiting dilution for further genotype and phenotype analysis. The integration of the vector into the genome was confirmed by PCR using primer sets as shown in Supplementary Table 1.

2.6. Analyses of parasite viability

Pyrimethamine-selected mutant parasites were cloned and genotyped as described [40]. To test for blood stage development, an NMRI mouse infected with cloned parasites was sacrificed to collect blood at a parasitemia of 0.2–0.5%. The blood was diluted with PBS and 100 µL of the diluted blood containing 5000 infected red blood cells was injected into a C57BL/6 mouse (in a group of 4). The parasitemia was measured daily by counting Giemsa stained thin blood smears. The experiment was repeated on three separate occasions.

To quantify mosquito stage development, naïve *Anopheles stephensi* mosquitoes were allowed to blood-feed on infected mice. Mosquitoes were dissected on day 16-post infection to examine the midgut oocyst loads. The midguts were stained with 1% mercurochrome and the oocysts were counted per infected mosquitoes. On the same day, mosquitoes were dissected to isolate salivary glands from which in turn sporozoites were isolated [42].

Gliding motility was quantified as follows: freshly isolated salivary gland sporozoites were re-suspended in RPMI medium (PAA Laboratories GmbH, Germany) supplemented with 3% bovine serum albumin (BSA, Roth) and 50 µL of re-suspended sporozoites transferred to glass bottom well plates, and spun at 209 g for 5 min at RT. Next, the sporozoites were imaged with an inverted Axiovert 200M Zeiss microscope using the Axiovision 4.6 software and a 25x Aplanop objective (NA = 0.25) in differential interference contrast illumination. Movies were recorded at an interval of 3 s per frame for the period of 3 min.

To examine liver stage development we conducted the following assays: 1) *In vitro* liver stage development and indirect immunofluorescence assay (IFA): For IFAs, HepG2 cells were infected with salivary gland sporozoites, incubated for 24 and 48 h and fixed with 4% paraformaldehyde. Fixed slides were washed twice with PBS, permeabilized with 0.5% Triton X-100 and blocked with phosphate buffered saline (PBS) containing 3% BSA. Mouse anti-HSP70 antibody (obtained through the MR4 as part of the BEI Resources Repository, NIAID, NIH: *Mus musculus* b5, MRA-662, deposited by MF Wiser) and Alexa-Fluor® 594-conjugated anti-mouse IgG (Molecular Probes) were used as primary and secondary antibodies, respectively. Host cell and parasite nuclei were stained with the DNA stain Hoescht 33,342. 2) *In vivo* liver stage development: female C56BL/6 mice were used in groups of four animals. Freshly isolated salivary gland sporozoites were suspended in PBS and 100 µL solution containing 10,000 sporozoites were injected into the tail vein of each mouse. The parasitemia was monitored daily on Giemsa stained slides from a drop of blood obtained from the tail vein [43].

3. Results

3.1. Bioinformatic identification of the house-cleaning enzyme repertoire of the rodent malaria model *Plasmodium berghei*, the human *P. falciparum* parasite and related apicomplexan lineages

NudX families are key to cell survival and promising drug targets. In order to identify such proteins in the malaria model, *Plasmodium berghei* we performed basic local alignment search tool (BLAST) searches using human DNA sanitation enzymes as query sequences. Our approach returned single homologs for ITIAse and dUTIAse, as well as two NudX domain-containing proteins with similarities to NUDT1/MTH1 and NUDT2/PA44AH (Supplementary figures 1–4); the latter two are annotated as nucleoside diphosphate hydrolase (NDH from here on) and bis(5′-nucleosyl)-tetraphosphatase [asymmetrical] (Ap4AH from here on). Each protein contains a well-supported Pfam protein domain (Table 1 and Figure 2(a)) [44]). A multiple sequence alignment of the orthologs from selected *Plasmodium* species revealed a high degree of conservation across experimental rodent (*P. berghei* and *P. yoelii*) and human malaria parasites (*P. falciparum* and *P. vivax*) (Supplementary Figures 5–8). Among the four enzymes, dUTIAse is the most conserved with a minimum amino acid sequence identity of 81% between *P. vivax* and *P. berghei* dUTIAse. *P. vivax* dUTIAse contains an additional loop region, which is absent in other *Plasmodium* species revealed a high degree of conservation across experimental rodent (*P. berghei* and *P. yoelii*) and human malaria parasites (*P. falciparum* and *P. vivax*) (Supplementary Figures 5–8). Among the four enzymes, dUTIAse is the most conserved with a minimum amino acid sequence identity of 81% between *P. vivax* and *P. berghei* dUTIAse. *P. vivax* dUTIAse contains an additional loop region, which is absent in other *Plasmodium* species.
species, and therefore is thus least similar to other *Plasmodium* dUTPases as well. The other enzymes are slightly less conserved with the lowest sequence identity of 59% occurring between PfITPase and PbiITPase (Figure 2(b,c)).

Based on the homologies with the human sanitation enzymes and experimental data, these conserved proteins likely catalyze the following reactions:

\[
dUTP = \rightarrow (d)UMP + PPI; \\
(d)XTP = \rightarrow (d)XMP + PPI; \\
8\text{-}\text{oxo}\text{-dGTP} + H_2O = \text{8\text{-}oxo\text{-dGMP}} + PPI; \\
\]

and P1,P4-bis(S’-guanosyl) tetraphosphate + H2O ≤ GTP + GMP [31,45–48].

A comparative sequence and structural analyses of these four house-cleaning enzymes with their respective human orthologs revealed that the overall architecture, as well as the key catalytic residues (dUTPase: D190; ITPase: E44; Ap4AH and NDH: NuDiX domain residues), are conserved between the parasite and host proteins suggesting the conserved mechanism of action (Supplementary Figure 9A–B). On the other hand, we also observed residues in the active site cavity that differ between the host and the parasite enzymes in each case (Supplementary Figure 9C). Such differences in active site residues have been explored to design [49] and test [50] targetspecific inhibitors. Therefore, these parasite-specific enzymes may be selectively inhibited. Indeed, *P. falciparum* dUTPase has been explored to design selective parasite-specific inhibitors [33] but whether this or other enzymes are essential for the parasite is unknown. Low sequence identities of the *P. berghei* enzymes to their human orthologs (Supplementary Figure 9D) suggest these enzymes could be selectively inhibited.

*P. berghei* and *P. falciparum* belong to the Apicomplexa phylum which includes parasites of humans and livestock. They include *Toxoplasma gondii* which causes infections in immunocompromised individuals, *Cryptosporidium* (causing cryptosporidiosis), *Theileria* (theileriosis and East Coast fever in cattle) and *Eimeria* (coccidiosis in poultry). Using BLASTP we explored the repertoire of house-cleaning enzymes in these related protozoa using *P. berghei* as query sequences. Nucleotide sanitation enzymes are present in the three apicomplexan lineages Pioplasmodia, Hemosporida, and Coccidia. Specifically dUTPase, ITPase and Ap4AH were conserved across the two *Plasmodium* species, *Toxoplasma gondii*, *Eimeria* spp., *Theileria annulata*, and *Cryptosporidium* spp; NDH on the other hand was only found in *Plasmodium* and *Theileria* (Table 2; Supplementary Figures 10–13).

### 3.2. House-cleaning enzymes are constitutively expressed throughout the *Plasmodium berghei* intraerythrocytic developmental cycle

A recent RNAseq study of blood stage and ookinete *P. berghei* life cycle stages indicated that all four genes are constitutively transcribed [51]; *dutP* several folds higher than the remaining three candidates, which all share similar expression levels (PlasmoDB, [http://plasmodb.org/plasmo/](http://plasmodb.org/plasmo/)). Our independent RT-PCR analyses in blood and mosquito stage parasites confirmed the whole genome data and expression of these genes throughout the life cycle with transcription evident in oocysts and salivary gland sporozoites as well (Supplementary Figure 14A). Additional support for constitutive protein expression throughout the parasite life cycle comes from *P. falciparum* 3D7, *P. vivax* Sal-1 and *P. yoelii* 17X proteomics (PlasmoDB, [http://plasmodb.org/plasmo/](http://plasmodb.org/plasmo/)) (Supplementary Figure 14B). For example, PfITPase is expressed in different erythrocytic stages, in gametocytes (male and female both), as well as in the salivary gland sporozoites. In absence of any liver stage expression data available for *P. falciparum*, we checked the expression level of the PfITPase ortholog in *P. yoelii* (Py), a rodent parasite, which confirmed the liver-stage expression of PfITPase (Supplementary Figure 14B). The details of protein expression of the remaining three enzymes are also given in supplementary Figure 14B. Using data from these large-scale proteomics studies, we confirmed that ITPase, dUTPase, and NDH are expressed in all three stages of life cycle, i.e. blood-, mosquito- and liver-stage. Although these studies report no blood-stage expression for PfAp4AH, a recent publication confirmed that PfAp4AH is indeed expressed during *P. falciparum* blood stage [52].

In order to address expression timing and localization, we generated transgenic *P. berghei* parasite lines expressing each protein fused to green fluorescent protein (GFP) at its C-terminus (Figure 2(d–g)). *ap4ah* and *ndh* were tagged at the endogenous loci, while we expressed dUTPase-GFP and ITPase-GFP episomally under the control of the native promoter using 986 and 865 basepairs of genomic sequence preceding each open reading frame, respectively. Protein expression in transfected parasites was followed by live microscopy and revealed a clear cytoplasmic GFP signal for all four proteins in asexual blood stages (ring, trophozoite, and schizont) as well as gametocytes. Western blot analyses showed GFP-tagged dUTPase and ITPase to run higher than their expected molecular weights, while Ap4AH:GFP and NDH::GFP were detected at their predicted positions; NDH with a clear doublet and a smaller 25 kDa band correlating with the size of the cleaved GFP tag (Figure 2(d–g)).
Taken together, all four house-keeping enzymes show constitutive expression across the intraerythrocytic developmental cycle; they represent thus suitable antimalarial drug targets.

3.3. Functional genetic analyses of the *P. berghei* house-cleaning enzyme repertoire

We next determined which of the four putative enzymes is required for *P. berghei* life cycle progression by attempting the generation of gene deletion mutants and defining each mutant’s fitness during rodent and mosquito infection. Gene knockout plasmids, designed to replace each open reading frame with a gene encoding a pyrimethamine resistance marker (*Toxoplasma gondii* dhfr/ts, or human *dhfr* fused to yeast *fcu*), were transfected into wild type schizonts; mutant parasites were selected with pyrimethamine following standard protocols [40].

Deletion of *dutpase* failed in three independent attempts. This suggests a likely essential role for the enzyme during blood stage development; most likely due to its dual role in detoxification of dUTP as well as being the sole source for dUMP for dTTP synthesis in *Plasmodium*. Together with previous studies in which *P. falciparum* dUTPase has been reported as a potential therapeutic target [32, 53], our result confirms the conserved essential role of this enzyme in different *Plasmodium* species.

Establishing a clonal *ap4ah*(-) mutant also failed although one transfection attempt out of four had resulted in the integration of the plasmid into the genome (Supplementary Figure 15A). The isolation of an isogenic, mutant clone however proved unsuccessful in two attempts (n = 16); limiting dilution cloning produced either non-infected mice or animals infected with wild type parasites only. In a third cloning attempt, we injected 5 mice with 10 instead of the standard 0.9 parasites from the mixed population of wild type and *ap4ah*(-) parasites. All mice became infected, but genotyping by PCR revealed that only wild type had survived (Supplementary Figure 15A). These results suggest that lack of Ap4AH causes too severe a growth defect and thus, like dUTPase, the protein is required for rodent malaria infection.

Unlike *dutpase* and *ap4ah*, we readily established mutant lines lacking *itpase* or *ndh* (Supplementary Figure 15B and C). In mammals both ITPase as well as the NUDIX domain-containing protein NUDT16 take part in ITP metabolism [54]. Therefore, we proceeded to generating an *itpase*(-);*ndh*(-) double mutant in order to determine whether *P. berghei* NDH and ITPase could compensate each other’s function in the single knockout lines. To this end we first removed the yFCU-hDHFR marker by negative 5-fluorocytosine selection from the *itpase*(-) clone and then transfected the *ndh* deletion plasmid into the selection marker-free *itpase*(-) mutant; limiting dilution cloning was used to generate an isogenic *itpase*(-);*ndh*(-) line (Supplementary Fig. 15D). We first monitored blood stage development and parasitemia curves after injection of 5,000 wild type or *itpase*(-);*ndh*(-) infected red blood cells in C57BL/6 mice. (Figure 3(a)). All mice infected with the wild type control showed health deterioration and were therefore euthanized. Although the two genes are not essential for blood-stage development (Figure 3(a)).
development, mutant parasites were less-virulent with prolonged host survival rates. A similar blood-stage characteristic has been reported for cysteine-protease inhibitor, PblCP knockout parasites [55].

Next, we quantified the number of oocysts following a mosquito blood meal. At day 16, the number of cysts per midgut was similar between the two groups (Figure 3(b)) indicating that neither ITPase nor NDH are required during DNA replication and sporozoite formation in the Anopheles mosquito. We did detect a significant reduction in sporozoite motility speed of salivary gland sporozoites in an in vitro assay on glass microscope slides (Figure 3(c)). The double mutant itpase(-)ndh(-) showed a speed of 1.2 ± 0.4 μm/s (n = 78) versus 1.5 ± 0.5 μm/s (n = 107) for wild type sporozoites (p < 0.001, Mann–Whitney U-test). However, this in vitro difference did not translate into a change of in vivo infectivity (Figure 3(d)), which requires larger differences in sporozoite motility [42]. DNA replication is particularly strong during liver development of the parasite, where a single sporozoite produces several thousand daughter progenies that are eventually released into the bloodstream. Imaging of liver cell development in HepG2 cells (Figure 3(d)) nor the initial establishment of blood stage infections identified a loss in fitness in the double mutant. The prepatent period – the time it takes for blood stage parasites to appear in thin blood smears – varied only slightly between 3.75 and 3.25 days (Figure 3(e)) and parasitemias quantified at day six following the infection were also within the normal range. Finally, similar blood stage infections were caused by intravenous injection of 10,000 salivary gland sporozoites (Figure 3(f)) and each of the parasite lines caused lethal blood stage infections with dynamics mirroring the injection of 5000 iRBC [56,57].

Although we did not observe a strong phenotype for itpase(-)ndh(-) parasite as the parasite could complete the entire life cycle, yet we observed a survival of the mice, which often succumb to experimental cerebral malaria around day 7 post-infection. This suggests that the itpase(-)ndh(-) parasite line is less virulent that the wild-type controls. Whether this is a phenotype of the deletion of the two genes or a clonal effect would require the generation of several more clones. We think that the subtlety of the effect does not justify the additional animal experiments.

Taken together our classical reverse genetics approach has uncovered the essentiality of dutpase and ap4ah and reported itpase and ndh as two redundant genes for the rodent malaria model P. berghei allowing life cycle progression between the mammalian host and mosquito vector. Further work needs to be done to study the functional details of these essential enzymes to support their potential role as the therapeutic targets. In addition, the role of these essential enzymes in the other life-stages could be explored via generating the conditional knockouts in different stages.

4. Discussion

House-cleaning enzymes play critical roles in maintaining genomic DNA integrity. Helical structure and histone-based packaging that protect DNA are absent for nucleotide precursors and they are thus prone to chemical modifications. However, unlike artificial nucleotide analogs, which are widely used in cancer therapy and as antiviral agents (abacavir and didanosine against HIV; lamivudine and entecavir against hepatitis B) [58], endogenous modifications of nucleotide pools by house-cleaning enzymes is poorly understood. The characterization of such proteins can provide a fundamental contribution to our understanding of nucleotide pool maintenance and sanitation. In pathogens, determining the requirements for certain enzymes for cell viability is an essential part during the identification and selection of suitable and novel targets for selective inhibitors that can be developed as drugs for parasites of humans as well as livestock.

Here, we identified four proteins in the rodent malaria parasite P. berghei with homology to evolutionarily conserved house-cleaning enzymes. While all proteins are constitutively expressed, reverse-genetics identified an essential nature of P. berghei dutPase and Ap4AH; ITPase and NDH were dispensable for the establishment and maintenance of an infection in the experimental mouse host and the Anopheles mosquito.

In accordance with our in vivo experimental data, ndh had recently been found to be dispensable in a global phenotype screen for P. berghei [59]. No information is yet available for the remaining three P. berghei genes dutPase, ap4ah and itpase (20 May 2018; http://plasmogem.sanger.ac.uk/). Here we compared our functional genetics studies with two global screens addressing the fitness of mutants in the human malaria parasite P. falciparum and the related apicomplexan parasite Toxoplasma gondii (Table 3). Like P. berghei dutPase the Toxoplasma homolog was found essential in a genome-wide loss-of-function CRISPR screen for growth in infected human fibroblasts [60]. P. falciparum dutPase, surprisingly, on the other hand, was identified as a dispensable gene in a genome-wide saturation mutagenesis screen by piggyBac transposon insertion [61]. This is in contrast to our current finding and the previously reported studies on P. falciparum dutPase [32,53]. The T. gondii data again mirrored our findings for ITPase, identifying the protein as a dispensable component

| KO          | P. berghei PlasmoGEM screen | P. falciparum piggyBac screen | Toxoplasma CRISPR screen |
|-------------|------------------------------|--------------------------------|--------------------------|
| dUTPase     | PBANKA_0921300 E             | nd                             | PF3D7_1127100 D (0.85) (0) TGGT1_233140 E (-4.26) |
| ITPase      | PBANCA_0618300 D             | nd                             | PF3D7_0720800 E (0.14) (~2.83) TGGT1_202300 D (0.42) |
| NDH         | PBANKA_1361900 D             | D 0.92                         | PF3D7_1349100 D (0.96) (~0.92) No homolog n/a |
| Ap4AH       | PBANKA_1235300 E             | nd                             | PF3D7_0520600 E (0.13) (~3.06) TGGT1_214780 D (~0.09) |

E essential gene; D dispensable gene; table based on P. berghei work of this paper, P. berghei PlasmoGEM screen (interrogated on 20 May 2018), P. falciparum piggyBac screen P. falciparum data represent (mutagenesis index score) (mutant fitness score); and Toxoplasma gondii CRISPR screen; nd not determined; n/a not applicable.
of the sanitation enzyme repertoire, while *falciparum* ITPase appeared to be essential for cell viability. We suspect that the host environment is a plausible factor that contributes to the different phenotypes in our *in vivo* study and previously reported *in vitro*, genome-wide saturation mutagenesis screen. *ap4ah*-deficient cells were neither viable in *P. falciparum* nor *P. berghei*; in *Toxoplasma* lack of the gene resulted in a minor growth defect. *In vitro* studies have shown that *P. falciparum* dUTPase catalyzes the hydrolysis of dUTP to dUMP through a conserved nucleophilic substation (SN2) mechanism in which a water molecule initiates the in-line nucleophilic attack [31]. The homology between the *berghei* and *falciparum* proteins strongly suggests that dUTPase in *P. berghei* performs the same reaction. That will prevent the incorporation of uracil into genomic DNA but also provide dUMP for dTTP formation in the absence of dCMP to dUMP deamination in *Plasmodium*. Whether one or both of these functions are the true causes for its essential nature in *P. berghei* is, however, not certain. Some cells can tolerate high concentrations of dUTP in genomic DNA. For example Warner et al. [62] reported that *E. coli* lacking functional dUTPase were viable even though 20% of DNA thymine residues had been replaced with uracil. Also, *Drosophila* can tolerate up to 1097 ± 55 uracil per million bases [29]. Secondly, there is also demonstrated thymidine uptake by *P. falciparum* mediated by the essential plasma membrane equilibrative nucleoside transporter 1 (PfENT1) [61,63], a protein that is conserved in *P. berghei*. The protein may thus be required for an as yet unknown function within the cell [24,27].

The *P. falciparum* ortholog of Ap4AH (bis(5'-nucleosyl)tetraphosphatase) hydrolyzes diadenosine 5',5'-P1,P4-tetraphosphatase, which is a by-product catalyzed by tRNA synthetase and ligases, to GTP and GMP [45]. At low concentrations, Ap4A activates microphthalmia transcription factor (MITF) dependent gene expression in mast cells [64] and also regulates ATP sensitive K+ channels in cardiac myocytes which otherwise get inhibited at higher concentration [65]. High level of Ap4A is linked with DNA damage, stress responses, and apoptosis. Enzymes hydrolyzing Ap4A prevent these effects [66]. *E. coli* NDX-4, an enzyme that hydrolyzes Ap4A to ATP has also been reported to hydrolyze toxic 8-oxo-dGTP [67]. Our failed attempts to delete and maintain an *ap4ah*(-) parasite and recent structural and functional studies of the *P. falciparum* homolog Ap4A suggest that Ap4AH is a potentially druggable enzyme [52].

Unlike *E. coli*, yeast and mice, the deletion of *itpase* in *P. berghei* had no deleterious effect on life cycle progression in the scope of a rodent infection. Although the parasite does not possess an obvious homolog to human NUDT16, a functional analogue of ITPase [54,68], we cannot rule out the presence of a compensatory mechanism in *Plasmodium*. It is unlikely that NDH performs this function; deletion of both *itpase* and *ndh* did not affect parasite growth. This may again be due to the compensatory house-cleaning functions of multiple enzymes in the *Plasmodium*. This is interesting to note that the number of NUDIX domain proteins varies from 0 to 30 in different prokaryotes and eukaryotes that hints towards the metabolic complexity and adaptability of different organisms [14].

5. Conclusions

We report here that *Plasmodium* deoxy-uridine triphosphatase and Ap4AH are essential house-cleaning enzymes while inosine triphosphate pyrophosphatase and NDH are redundant for the malaria parasite. Most importantly, our failed attempts to delete dUTPase in *P. berghei* in combination with previous drug-inhibitory assays of *P. falciparum* dUTPase [69] suggest that this enzyme is indeed essential for blood-stage growth and like Ap4AH is a potential malaria drug target.

Authors’ contributions

HK: Conceptualization, Formal analysis, Investigation, Supervision, Methodology, Visualization, Writing – original draft, Writing – review & editing
JK: Formal analysis, Investigation, Supervision, Methodology, Visualization, Writing – review & editing
MS: Investigation, Supervision, Visualization, Writing – review & editing
GM: Conceptualization, Project administration, Formal analysis, Resources, Supervision, Methodology, Writing – original draft, Writing – review & editing
FF: Funding acquisition, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing

Funding

This work was supported by a fellowship from the Deutsche Akademischer Austauschdienst (DAAD; awarded to H Kumar) and grants from the Chica and Heinz Schaller Foundation, the European Research Council (StG 281719), the Deutsche Forschungsgemeinschaft (SFB 1129) and the Heidelberg University cluster of excellence CellNetworks (awarded to F Frischknecht). JM Santos was supported by the Portuguese Fundação para a Ciência e a Tecnologia (SFRH/BD/63849/2009).

Declaration of interest

F Frischknecht was a member of the EU FP7 Network of Excellence EVMaR. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

Reviewer disclosures

Peer reviewers on this manuscript have no relevant financial or other relationships to disclose.

ORCID

Hirdesh Kumar https://orcid.org/0000-0002-8488-3001

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Gamo FJ, Sanz LM, Vidal J, et al. Thousands of chemical starting points for antimalarial lead identification. Nature. 2010 May 20;465(7296):305–310. PubMed PMID: 20485427; eng.
2. Mu J, Awadalla P, Duan J, et al. Genome-wide variation and identification of vaccine targets in the Plasmodium falciparum genome. Nat Genet. 2007 Jan;39(1):126–130. PubMed PMID: 17159981; eng.
A comprehensive review about malaria biology.

A thorough analysis of malaria parasites and their characteristic stages.

**Important study on Nudix hydrolases.**

A detailed study on the function and role of Nudix hydrolases in various biological processes.

**An interesting review about Nudix hydrolases.**

A comprehensive review highlighting the significance and applications of Nudix hydrolases.

**Important study on Plasmodium falciparum dUTPase to design selective nucleoside inhibitors.**

Investigating the potential of constructing selective nucleoside inhibitors targeting dUTPase.

**An interesting review about dUTPase.**

An in-depth examination of the dUTPase enzyme and its diverse functions.

**New abovementioned review about dUTPase.**

Expanding the discussion on dUTPase with recent findings and developments.

**A comprehensive review about dUTPase.**

A broad overview of dUTPase, its structure, and functions.

**A new comprehensive review about dUTPase.**

Updating the understanding of dUTPase with the latest research.

**A detailed study about dUTPase.**

A methodical exploration of the role and importance of dUTPase.

**Recent studies about dUTPase.**

A compilation of recent research on dUTPase, highlighting new insights and advancements.
recombination. Mol Biochem Parasitol. 2006;150(1):118–121. Pubmed PMID: 16901558; eng.
42. Bane KS, Lepper S, Kehrer J, et al. The actin filament-binding protein coronin regulates motility in Plasmodium sporozoites. PLoS Pathog. 2016;12(7):e1005710. Pubmed PMID: 27409081; eng.
43. Kumar H, Sattler JM, Singier M, et al. Protective efficacy and safety of liver stage attenuated malaria parasites. Sci Rep. 2016;6:26824. Pubmed PMID: 27241521; eng.
44. Letunic I, Doerks T, Bork P. SMART: recent updates, new developments and status in 2015. Nucleic Acids Res. 2015 Jan;43(Database issue):D257–D260. Pubmed PMID: 25300481; eng.
45. Osman W, Endo S, Oh-Hashi K, et al. Molecular characterization and mutational analysis of recombinant diadenosine 5′,5′′-P(1),P(4)-tetraphosphate hydrolase from Plasmodium falciparum. Biol Pharm Bull. 2012;35(7):1191–1196. Pubmed PMID: 22791172; eng.
46. Stemnak P, Kursula P, Flodin S, et al. Crystal structure of human inosine triphosphatase. Substrate binding and implication of the inosine triphosphatase deficiency mutation P32T. J Biol Chem. 2007 Feb 2;282(5):3182–3187. Pubmed PMID: 17138556; eng.
47. Sakai Y, Furuichi M, Takahashi M, et al. A molecular basis for the selective recognition of 2-hydroxy-dATP and 8-oxo-dGTP by human MTH1. J Biol Chem. 2002 Mar 8;277(10):8579–8587. Pubmed PMID: 11756418; eng.
48. Ge H, Chen X, Yang W, et al. Crystal structure of wild-type and mutant human Ap4A hydrolase. Biochem Biophys Res Commun. 2013 Mar 1;432(1):16–21. Pubmed PMID: 23384440; eng.
49. Kumar H, Shah A, Sobhia ME. Novel insights into the structural requirements for the design of selective and specific aldose reductase inhibitors. J Mol Model. 2012;18(5):1791–1799. Pubmed PMID: 21833829; eng.
50. Hoepfner D, McNamara CW, Lim CS, et al. Selective and specific inhibitors of the Plasmodium falciparum dUTPase. Bioorg Med Chem. 2012;20(2):145–151. Pubmed PMID: 22794426; eng.
51. Otto TD, Böhme U, Jackson AP, et al. A comprehensive evaluation of rodent malaria parasite genomes and gene expression. BMC Biol. 2014;12(1):86. Pubmed PMID: 25359557; eng.
52. Sharma A, Yogavel M, Sharma A. Structural and functional attributes of malaria parasite diadenosine tetraphosphate hydrolase. Sci Rep. 2016;6:19981. Pubmed PMID: 26829485; eng.
53. Hampton SE, Schipani A, Bosch-Navarrete C, et al. Investigation of acyclic uridine amide and 5′-amido nucleoside analogues as potential inhibitors of the Plasmodium falciparum dUTPase. Bioorg Med Chem. 2013 Sep 15;21(18):5876–5885. Pubmed PMID: 23916149; eng.
54. Abolhassani N, Iyama T, Tsuchimoto D, et al. NUDT16 and ITPA play a dual protective role in maintaining chromosome stability and cell growth by eliminating diDP/dIDP and diTP/iTP from nucleotide pools in mammals. Nucleic Acids Res. 2010;38(9):2981–2903. Pubmed PMID: 20081199; eng.
55. Lehmann C, Heitmann A, Mishra S, et al. A cysteine protease inhibitor of plasmodium berghei is essential for exo-erythrocytic development. PLoS Pathog. 2014 Aug;10(8):e1004336. Pubmed PMID: 25166051; eng.
56. Sa ECC, Nyboer B, Heiss K, et al. Plasmodium berghei EXP-1 interacts with host Apolipoprotein H during Plasmodium liver-stage development. Proc Natl Acad Sci U S A. 2017 Feb 14;114(7):E1138–e1147. Pubmed PMID: 28137845; eng.
57. Zheng H, Tan Z, Zhou T, et al. The TLR2 is activated by sporozoites and suppresses intrahepatic rodent malaria parasite development. Sci Rep. 2015 Dec 15;5:18239. Pubmed PMID: 26667391; eng.
58. Jordheim LP, Durantel D, Zoulim F, et al. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. Nat Rev Drug Discov. 2013;12(6):447–464. Pubmed PMID: 23722347; eng.
59. Gomes AR, Bushell E, Schwach F, et al. A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. Cell Host Microbe. 2015;17(3):404–413. Pubmed PMID: 25732065; eng.
60. Sidik SM, Huet D, Ganesan SM, et al. A genome-wide CRISPR screen in toxoplasma identifies essential apicomanx genes. Cell. 2016 Sep 8;166(6):1433–1435.e12. Pubmed PMID: 27594426; eng.
61. Zhang M, Wang C, Otto TD, et al. Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. Science. 2018 May 4;360(6388). Pubmed PMID: 29724925; eng.
62. Warner H, Duncan B, Garrett C, et al. Synthesis and metabolism of uracil-containing deoxyribonucleic acid in Escherichia coli. J Bacteriol. 1981;145(2):687–695. Pubmed PMID: 6109711; eng.
63. Downie MJ, Saliba IK, Howitt SM, et al. Transport of nucleosides across the Plasmodium falciparum parasite plasma membrane has characteristics of PFENT1. Mol Microbiol. 2006 May;60(3):738–748. Pubmed PMID: 16629674; eng.
64. Lee Y-N, Nechushtan H, Figov N, et al. The function of lysyl-tRNA synthetase and Ap4A as signaling regulators of MITF activity in FcεRI-activated mast cells. Immunity. 2004;20(2):145–151. Pubmed PMID: 14975237; eng.
65. Jovanovic A, Alekseev AE, Terzic A. Intracellular diadenosine polyphosphates: a novel family of inhibitory ligands of the ATP-sensitive K+ channel. Biochem Pharmacol. 1997;54(2):219–225. Pubmed PMID: 9271325; eng.
66. McLennan AG. Dinucleoside polyphosphates – friend or foe? Pharmacol Ther. 2000;87(2):73–89. Pubmed PMID: 11007992; eng.
67. Arczewska KD, Baumieier C, Kassahun H, et al. Caenorhabditis elegans NDX-4 is a MutT-type enzyme that contributes to genomic stability. DNA Repair (Amst). 2011;10(2):176–187. Pubmed PMID: 21116960; eng.
68. Iyama T, Abolhassani N, Tsuchimoto D, et al. NUDT16 is a (deoxy) inosine diphosphatase, and its deficiency induces accumulation of single-strand breaks in nuclear DNA and growth arrest. Nucleic Acids Res. 2010 Aug;38(14):4834–4843. Pubmed PMID: 20385596; eng.
69. Belen Cassera M, Zhang Y, Z Hazleton K, et al. Purine and pyrimidine pathways as targets in Plasmodium falciparum. Curr Top Med Chem. 2011;11(16):2103–2115. Pubmed PMID: 21619511; eng.