



**HAL**  
open science

# Characterization of *P.falciparum* histone methyltransferases : biological role and possible targets for new intervention strategies

Shuai Ding

► **To cite this version:**

Shuai Ding. Characterization of *P.falciparum* histone methyltransferases : biological role and possible targets for new intervention strategies. Parasitology. Université Pierre et Marie Curie - Paris VI, 2016. English. NNT : 2016PA066578 . tel-01956977

**HAL Id: tel-01956977**

**<https://theses.hal.science/tel-01956977v1>**

Submitted on 17 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Universit  Pierre et Marie Curie

Ecole doctorale Complexit  du Vivant

*Biologie des Interactions H te-Parasite*

**Characterization of *P. falciparum* histone  
methyltransferases: biological role and possible targets  
for new intervention strategies**

Shuai DING

Th se de doctorat de Parasitologie

Dirig e par Artur SCHERF & Dominique MAZIER

Pr sent e et soutenue publiquement le 15 D cembre 2016

Devant un jury compos  de :

Pr. Vincent MARECHAL	Pr�sident du jury
Pr. Nicolai SIEGEL	Rapporteur
Pr. Luisa Miranda FIGUEIREDO	Rapporteur
Pr. Agathe SUBTIL	Examineur
Dr. Beno�t GAMAIN	Examineur
Dr. Arnaud CHENE	Examineur
Pr. Dominique MAZIER	Co-Directeur de th�se
Pr. Artur SCHERF	Directeur de th�se



当我不遗忘 也不想曾经

尝欢

洒泪

问道

寻真



If you seek Truth, you will not seek to gain a victory by every possible means;  
and when you have found Truth, you need not fear being defeated.

*Epictetus*



# SUMMARY

---

<b>ACKNOWLEDGEMENTS</b>	<b>i</b>
<b>TABLE OF CONTENTS</b>	<b>iii</b>
<b>LIST OF FIGURES</b>	<b>v</b>
<b>LIST OF TABLES</b>	<b>vi</b>
<b>ABBREVIATIONS</b>	<b>vii</b>
<b>INTRODUCTION</b>	<b>01</b>
<b>RESULTS</b>	<b>39</b>
<b>DISCUSSION</b>	<b>130</b>
<b>BIBLIOGRAPHY</b>	<b>138</b>





# ACKNOWLEDGEMENTS

This thesis represents not only my work at the bench and at the computer, it is also an unforgettable experience of my life work at Institut Pasteur and specifically within the **Biologie des Interactions Hôte-Parasite** Laboratory.

First and foremost, I wish to express my heartfelt gratitude to my supervisor, professor **Artur Scherf** for his guidance, patience, and tough love. By his seniority, creativity and rigor, he has set a standard by which I always found hard to live up to. At the same time, he served as a strong, ever present support when I became lost in my experiments or neglected scientific thinking, and when results were partial, doubtful or confusing. Thank you for believing in me, and that I am capable of overcoming my faults and achieving my goals. I would also particularly like to thank **Dominique Mazier** for her continuous support and encouragement.

My greatest appreciation is extended to **Nicolas A. Malmquist**. His outstanding work on enzymatic biochemistry has been indispensable in forming the basis of my projects. He gave me meticulous guidance and warm encouragement in the first two years of my PhD. In a similar vein, I would like to acknowledge **Shruthi S Vembar**, who helped me to shape my thoughts and this thesis, and guided me over this last year of development. She has supported me academically and emotionally through road right up to the end of this thesis.

This work was made possible by the continuous assistance, guidance and support from each member of BIHP, and to whom I am greatly indebted. Special thanks go to **Patty Chen**, without whom the histone methyltransferase projects described here would not be so exciting. Thanks also for the care and kindness she's shown me, as well as those delicious homemade Chinese, French and innovative dishes she has made. I have greatly benefited from **Julien Guizetti**, who has single-handedly taught me microscopy technique, and gave insightful comments and suggestions. I owe my deepest gratitude to **Anne Cozanet** for her patience, efforts and readiness to help, otherwise my life in France would have been difficult and frustrating. I am fortunate to have **Claudia Demarta-Gatsi**, **Gigliola Zanghì**, **Anna Barcons Simon** by my side, these lovely girls gave me valuable moral support, especially towards the end of the finishing line. I am particularly indebted to **Salah Mechieri** and **Roger Peronet** for providing the antibodies, which were the most important tools for my projects. Generous support given by **Denise Mattei**, **Rosaura Hernandez Rivas**, **Sebastian Baumgarten**, **Dorothea Droll** and **Jessica Bryant** have been a great help in this dissertation. My sincere thanks to **Aurélie Claes** and **Christine Scheidig-Benatar** for sharing experiences and

discussing technical details. I also received enormous help from the former members of my lab, **Rafael Miyazawa Martins**, **Cameron Macpherson**, **Jose-Juan Lopez-Rubio** and **Mehdi Ghorbal**.

I would like to acknowledge the members of my Thesis Committee, **Vincent Marechal** for his indispensable contribution to assess and evaluate my PhD projects in the past three years. I am deeply grateful to **Nicolai Siegel** and **Luisa Miranda Figueiredo** for being my rapporteur and providing insightful questions and comments. My sincere thanks go to **Agathe Subtil**, **Benoît Gamain** and **Arnaud Chene** for being part of my committee. I would also like to thank **Anavaj Sakuntabhai** as my tutor in Pasteur and UPMC.

Finally, my heartfelt thanks and appreciation to all of my family and friends in helping me to move forward during last three years. It would be impossible to mention all of them here, but without their emotional support, accomplishing this work would be difficult if not impossible. I'll love you forever.

# TABLE OF CONTENTS

---

## INTRODUCTION

<b>1. General Introduction to Malaria</b>	<b>01</b>
<b>1.1 Malaria Epidemiology</b>	<b>01</b>
<b>1.2 Plasmodium Biology</b>	<b>02</b>
1.2.1 Life Cycle	02
1.2.2 Chronic <i>P. falciparum</i> Malaria Infections	03
<b>1.3 Current situation and Challenges of Malaria Control</b>	<b>04</b>
1.3.1 Antimalarial Drugs and Drug Resistance	04
1.3.2 Vaccine Candidate in Development	06
<b>2. General Epigenetic Regulation in <i>P. falciparum</i></b>	<b>08</b>
<b>2.1 Epigenetic Mechanisms of Gene Regulation</b>	<b>08</b>
2.1.1 Chromatin Structure	08
2.1.2 Histone Modification	09
2.1.3 Histone Variants	10
2.1.4 Nuclear Organization	11
2.1.5 DNA Methylation	13
<b>2.2 Histone Modification Machinery</b>	<b>14</b>
2.2.1 “Writers” and “Erasers” of the Histone Code	14
2.2.2 “Readers” of the Histone Code	17
<b>2.3 Histone Modification in eu- and heterochromatin in <i>P. falciparum</i></b>	<b>20</b>
<b>3. The Roles of Histone Methylation and Histone Methylation Modifiers</b>	<b>21</b>
<b>3.1 Just-in-Time Transcription during Asexual Development</b>	<b>21</b>
<b>3.2 Histone Lysine Methylation Regulates Virulence Gene Expression</b>	<b>23</b>
3.2.1 Monoallelic <i>var</i> Expression	23
3.2.3 Histone H3 Marks and <i>var</i> Gene Activation	25
3.2.3 The Histone Reader PfHP1 is involved in <i>var</i> gene Silencing	26

<b>3.3 The Interplay of Histone Methylation and Deacetylation Regulates Sexual Commitment</b>	<b>27</b>
<b>3.4 Histone Lysine Methylation may Contribute to Hypnozoite Quiescence</b>	<b>29</b>
<b>3.5 Predicted HKMTs in <i>P. falciparum</i></b>	<b>33</b>
<b>3.6 PfSET Proteins Associated with <i>var</i> gene Expression</b>	<b>33</b>
3.6.1 PfSET10 Maintains the Active <i>var</i> gene in Poised State	33
3.6.2 PfSETvs Represses <i>var</i> gene Expression	34
<b>3.7 Development of HKMT inhibitors as antimalarials</b>	<b>34</b>
<b>3.8 Emerging Roles of HKMT-mediated Lysine Methylation on Histone and Non-Histone Proteins</b>	<b>36</b>
<b>4. Scope of the Thesis</b>	<b>38</b>
 <b>RESULTS</b>	
<b>5. Results</b>	<b>38</b>
5.1 Article I	38
5.2 Article II	78
5.3 Article III	104
5.4 Result IV	126
 <b>DISCUSSION</b>	
<b>6 Discussion</b>	<b>130</b>
<b>7 Perspective</b>	<b>135</b>
<b>8 Conclusion</b>	<b>137</b>
<b>9 Bibliography</b>	<b>138</b>

## LIST OF FIGURES

---

Figure 1.1. <b>Malaria deaths per 100,000 populations in 2013</b> .....	1
Figure 1.2. <b>The life cycle of <i>Plasmodium spp.</i> indicating the identification of post-translational protein modifications (PTM) in essentially all stages of the life cycle</b> .....	2
Figure 1.3. <b>Timeline for the introduction of different antimalarials, with the lower panel indicating the time of emergence of parasite resistance</b> .....	5
Figure 2.1. <b>Histone H3 post-translational modifications in <i>Plasmodium spp.</i> and domains that recognize the modified histone tails</b> .....	10
Figure 2.2. <b>Nuclear organization spatially regulates transcription of select genes</b> .....	12
Figure 3.1. <b>Overview of the transcriptome of <i>P. falciparum</i> intraerythrocytic development at 1 h resolution</b> .....	22
Figure 3.2. <b>Monoallelic expression of <i>var</i> genes</b> .....	24
Figure 3.3. <b>Monoallelic expression of <i>var</i> genes regulated by histone PTMs</b> .....	25
Figure 3.4. <b>Distribution of select histone PTMs along <i>P. falciparum</i> chromosomes</b> .....	26
Figure 3.5. <b>Proposed model for the regulation of sexual commitment in <i>P. falciparum</i></b> ..	29
Figure 3.6. <b>Domain organisations of predicted SET-domain HKMTs in <i>P. falciparum</i></b> ..	31
Figure 3.7. <b>Phylogenetic analyses of SET and post-SET domains from Apicomplexan lysine methyltransferases identify several subfamilies of protein lysine methyltransferases</b> .....	32

## LIST OF TABLES

---

Table 2.1. <b>Predicted and verified histone PTM writers and erasers in <i>P. falciparum</i></b> ...	16
Table 2.2. <b>Predicted and verified histone mark readers in <i>P. falciparum</i></b> .....	19
Table 3.1. <b>Predicted HKMTs in <i>P. falciparum</i></b> .....	33

# ABBREVIATIONS

---

ACTs	Artemisinin Combination Therapies
CD	Chromo Domain
ChIP	Chromatin immunoprecipitation
CQ	Chloroquine
CSD	Chromo Shadow Domain
CSP	Circumsporozoite Protein
EMP1	Erythrocyte Membrane Protein 1
GAPs	Genetically Attenuated Parasites
HATs	Histone Acetyltransferases
HDACs	Histone Deacetylases
HKDMs	Histone lysine Demethylases
HKMTs	Histone lysine Methyltransferases
HP1	Heterochromatin Protein 1
HRMTs	Histone arginine Methyltransferases
HRF	Histamine-Releasing Factor
hpi	hours post-erythrocyte invasion
IDC	Intraerythrocytic Development Cycle
JHDMs	Jumonji C-domain containing demethylases
MT	microtubules
NESs	Nuclear Export Signals
NLSs	Nuclear Localization Signals
PHDs	Plant Homeodomains
PPM	Parasite Plasma Membrane
PTM	Post-Translational Modifications
PVM	Parasitophorous Vacuole Membrane
IRS	Indoor Residual Spraying
LLINs	Long Lasting Insecticide-treated Nets
LSD1	Lysine-Specific Demethylase1
MYND	Myeloid-Nervy-DEAF1
SAM	S-Adenosyl Methionine
SET	Su(var)3-9-Enhancer of zeste-Trithorax
Sir2	Sirtuin 2 protein
SP	Sulfadoxine-Pyrimethamine
TSSs	Transcription Start Sites
ups	Upstream promoter sequence





# INTRODUCTION

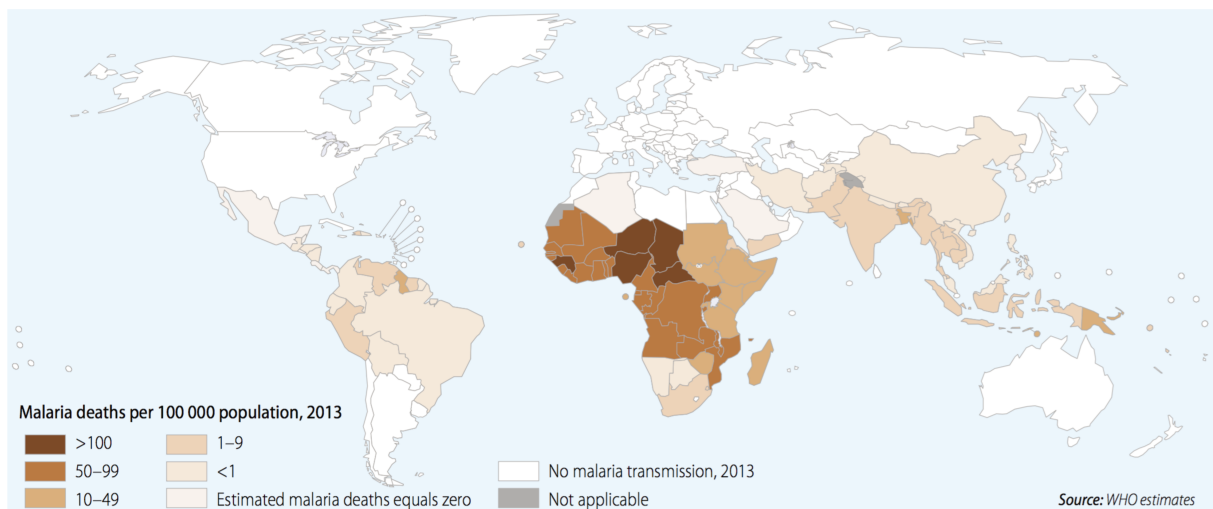




# 1. A General Introduction To Malaria

## 1.1 Malaria Epidemiology

Human malaria is caused by protozoan parasites of the genus *Plasmodium*, with *Plasmodium falciparum* being responsible for the most severe form of malaria among the five species, *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and *P. knowlesi*, that infect humans. According to the World Malaria Report 2015, an estimated 3.3 billion people in 97 countries and territories are at risk of being infected with malaria and developing disease, and 1.2 billion are at high risk (>1 in 1000 chance of getting malaria in a year). In 2014, an estimated 214 million cases of malaria and 438,000 malaria deaths occurred worldwide (WHO, 2015). The burden is heaviest in the African Region, where an estimated 88% of malaria cases and 90% of malaria-related deaths occur, especially in children under five years of age, who account for 78% of all deaths (Figure 1.1) (WHO, 2014).



**Figure 1.1.** Malaria deaths per 100,000 population in 2013. Taken from the World Malaria Report 2014 (WHO, 2014).



Transmission begins when the *Anopheles* vector takes a blood meal that contains mature male and female gametocytes (Figure 1.2). Within the mosquito midgut, the male gametocyte undergoes rapid nuclear division, producing eight flagellated microgametes that fertilize the female macrogamete. The resulting ookinete traverses the wall of the mosquito gut and encysts on the exterior of the gut wall as an oocyst. The oocysts matures over several days (13-14 days in the case of *P. falciparum*) producing thousands of sporozoites, which when released into the mosquito midgut, travel to the vector's salivary glands and enter through the injected saliva the next human host.

### **1.2.2 Chronic *P. falciparum* Malaria Infections**

Accumulated data over the last 50 years show unequivocally that *P. falciparum* can survive for a long period of time in the peripheral blood of the human host, with the longest duration reported to date being 9 years (Howden et al., 2005). In individuals who have developed a protective immune response, it is still unclear how the parasite evades the host immune mechanism to maintain a low-grade parasite load over months and years. But it is widely accepted that the malaria parasite establishes a chronic infection by escaping the host immune responses directed against a series of antigens that expose variable surface epitopes and/or by selecting parasite populations with distinct polymorphic antigens (Druilhe and Pérignon, 1997).

The most important molecule implicated in *P. falciparum* immune evasion is erythrocyte membrane protein 1 (PfEMP1), the major erythrocyte surface antigen mediating parasite sequestration in the microvasculature. The parasite varies the expression of ~60 PfEMP1-encoding *var* genes such that a single *var* gene is expressed at any given time, with the remaining 59 genes maintained in a transcriptionally inactive state. The mechanisms of *var* gene regulation are discussed in detail in Sections 2 and 3. In addition, expression of particular PfEMP1 subtypes, which bind to alternative host endothelial receptors, may result in parasite enrichment in different tissues, causing a range of symptoms and disease severity (Merrick et al., 2012; Scherf et al., 2008a).

## 1.3 Current Situation and Challenges of Malaria Control

The past decade witnessed great efforts to control malaria, especially given the increased availability of new and diverse strategies and tools against both the parasite and the vector. However, malaria still represents a major health burden, particularly in Africa. Key challenges such as the rise of drug and insecticide resistance, and the unavailability of a licensed, highly efficient malaria vaccine, are discussed below.

### 1.3.1 Antimalarial Drugs and Drug Resistance

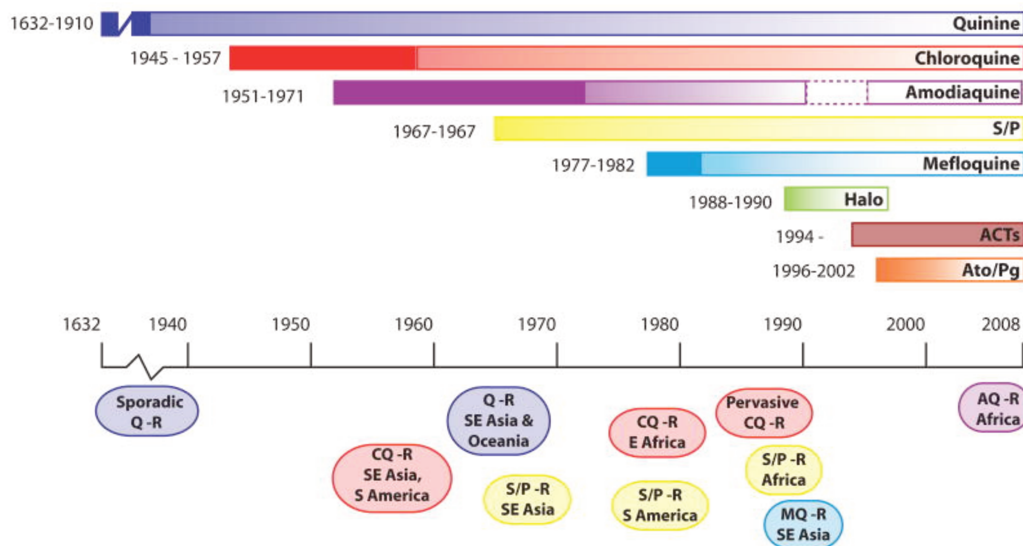
Currently available antimalarials can be divided into three broad categories according to the chemical structure and mode of action: the aryl aminoalcohol compounds (quinine, quinidine, chloroquine, amodiaquine, mefloquine, halofantrine, lumefantrine, piperazine, tafenoquine), the antifolate compounds (pyrimethamine, proguanil, chlorproguanil, trimethoprim), and the artemisinin compounds (artemisinin, dihydroartemisinin, artemether, artesunate). Of these, the artemisinin drugs have the broadest time window of action on different stages of the parasite and produce the most rapid therapeutic response.

Since the introduction of the first drug to tackle malaria, the rapid emergence of drug resistant pathogens has been observed (Figure 1.3). Chloroquine (CQ) was introduced in 1945 with resistant parasites cropping up in the early 1960s (Blount, 1967). Following widespread resistance of *P. falciparum* to CQ, sulfadoxine-pyrimethamine (SP) came to the scene in the 1960s as a cheap and well-tolerated alternative. However, SP-resistance was reported in the same year (Figure 1.3) and widespread resistance was documented over a relatively short period of time in most endemic countries (Ekland and Fidock, 2008).

Artemisinin and its derivatives are currently the most active antimalarial drugs available and have been introduced around the world as an integral part of both prophylactic and first-line therapies. They are often used in combination with other antimalarials, the so-called Artemisinin Combination Therapies or ACTs. The 5 ACTs currently recommended for use by the WHO are artemether plus lumefantrine, artesunate plus amodiaquine, artesunate plus mefloquine, artesunate plus sulfadoxine-pyrimethamine, and dihydroartemisinin plus piperazine (WHO, 2006). The choice of therapy is based on the efficacy of the combination in the endemic country or area of intended use.

In 2009, researchers reported concerns that artemisinin was taking longer to clear *P. falciparum* parasites from malaria patients in the Thailand-Cambodia border (Dondorp et al., 2009). In successive years, further investigation in the Greater Mekong sub-region reported increased incidence of treatment failure among patients with ACT ranges from 2008 to 2015 (Ashley et al., 2014; World Health Organization, 2015). An international collaborative study subsequently identified a molecular marker in the kelch13 (K13) gene (located on *P. falciparum* chromosome 13) that is associated with the slow clearance phenotype (Ariey et al., 2013). This finding has spurred researchers to monitor the prevalence of K13 mutations in Cambodia, Thailand and other Greater Mekong sub-regions, thus determining the extent of artemisinin resistance in south-east Asia (Straimer et al., 2014, Mbengue et al., 2015). These studies highlights that the pace at which artemisinin resistance is spreading is rapid, presenting a major challenge to elimination efforts.

Other preventive measures against malaria recommended by the WHO include anti-vector procedures such as indoor residual spraying (IRS), the use of long lasting insecticide-treated nets (LLINs), and destruction of larval sites. However, the effectiveness of these tools is being threatened by the emergence and spread of insecticide resistance. These issues emphasize the need for novel transmission blocking strategies that can work toward malaria eradication, directed both against the parasite and vector.



**Figure 1.3.** Timeline for the introduction of different antimalarials, with the lower panel indicating the time of emergence of parasite resistance. Q=Quinine; CQ=Chloroquine; S/P=Sulfadoxine and Pyrimethamine; MQ=Mefloquine; AQ=Amodiaquine; ACTs= Artemisinin Combination Therapies; Ato/Pg= atovaquone/proguanil; Taken from (Ekland and Fidock, 2008)



### 1.3.2 Vaccine Candidates in Development

Vaccines are generally believed to be important in the elimination of *P. falciparum* and other *Plasmodium spp.*, because their efficiency has been proven in the global eradication of smallpox and polio from the Western Hemisphere. As compared to chemotherapy, vaccines are safer and more convenient for young infants and children, pregnant women, and travelers. Nonetheless, the prospects of achieving malaria elimination are diminished by the lack of a licensed malaria vaccine. So far, most of efforts have been directed towards single antigens, and the candidate antigens can be classified into three groups based on the target life cycle stage: pre-erythrocytic, erythrocytic and transmission-blocking (Barry and Arnott, 2014).

To prevent pre-erythrocytic infection, a vaccine should block sporozoites from entering and multiplying in the liver. The best characterized pre-erythrocytic vaccine is RTS,S/AS01, which was engineered using immunogenic regions of the circumsporozoite protein (PfCSP) and a viral envelope protein of hepatitis B virus (HBsAg), to which was added a chemical adjuvant AS01 (Partnership et al., 2012). RTS,S/AS01 is the first malaria vaccine to have completed phase 3 clinical testing, and shows partial effectiveness against clinical disease in young African children up to 4 years after vaccination (Greenwood, 2015).

Since all of the symptoms of malaria occur during the erythrocytic stages, the majority of vaccine candidates are designed to prevent parasite asexual reproduction and development. One approach is to target merozoite antigens to prevent red blood cell invasion. Another approach is to target the major surface protein PfEMP1 to block the process of erythrocyte adherence to host cells, which is accountable for many clinical symptoms associated with malarial infection (such as placental adhesion during pregnancy malaria and others (see section 1.2.2)) (Hviid, 2010). However, because there are hundreds to thousands of distinct alleles for PfEMP1 in field isolates, the development of PfEMP1-based vaccines has lagged behind (Barry & Arnott, 2014).

The last group consists of transmission-blocking vaccines that target antigens expressed during life cycle stages in the mosquito host. Vaccine candidates are divided into two classes based on the target antigen: the first class includes surface proteins of the extra-cellular male and female gametes (Figure.1.2), such as Pfs48/45, while the second class includes surface proteins expressed on the zygote and ookinete stages of the parasites (Figure.1.2), such as

Pfs25 (Carter, 2001). These vaccines would greatly assist elimination efforts to prevent onward transmission.

In the past decade, a different and potentially improved vaccination strategy involves the use of attenuated parasites, either through irradiation or genetic ablation, to generate sterile protection against malaria. The best studied amongst them are Radiation Attenuated Sporozoites, and Genetically Attenuated Parasites (GAPs) that arrest either during liver stage or blood stage development (Nganou-Makamdop and Sauerwein, 2013). For instance, Mécheri and colleagues studied the role of the parasite's histamine-releasing factor (HRF) in the development of host immune response and found that mice infected with blood stage *P. berghei* HRF-deficient parasites (*Pb* $\Delta$ *hrf*) cleared parasites rapidly leading to a long-lasting, wide-ranging, cross-stage and cross-species protective effect (Demarta-Gatsi et al., 2016). This indicated that *P. falciparum*  $\Delta$ *hrf* parasites could be developed as a blood stage GAP vaccine. The other GAP-based vaccine, *Pf* $\Delta$ *slarp* $\Delta$ *b9* and its equivalent rodent mutant that arrest in pre-erythrocytic liver stages, have also been evaluated for their efficacy and safety, and are being developed for clinical testing (Annoura et al., 2014; van Schaijk et al., 2014).

## 2. General Epigenetic Regulation in *P. falciparum*

### 2.1 Epigenetic Mechanisms of Gene Regulation

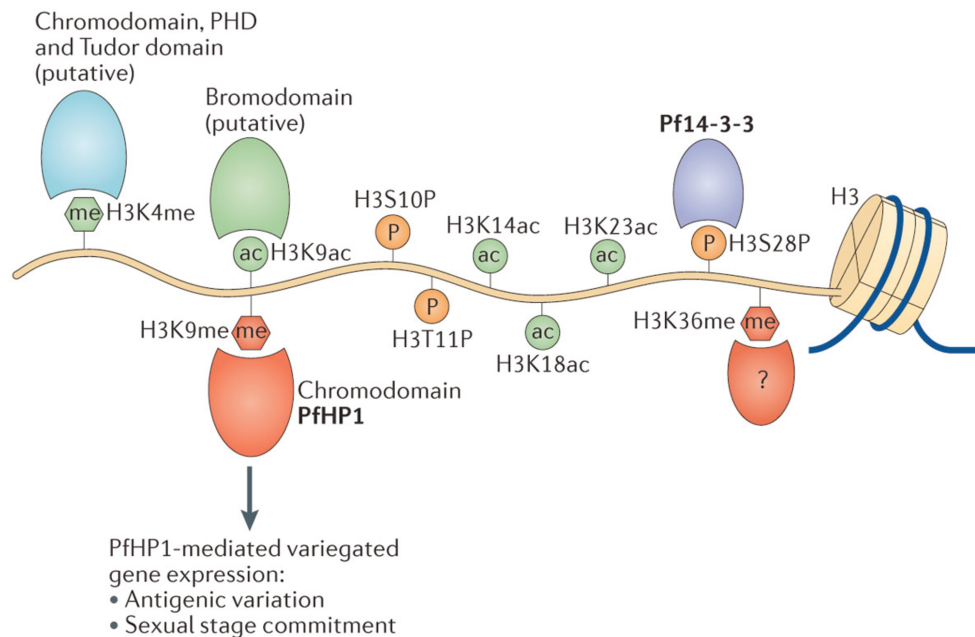
#### 2.1.1 Chromatin Structure

Genomic DNA in eukaryotic cells is packaged with special proteins termed histones to form protein/DNA complexes called chromatin. The nucleosome is the basic unit of chromatin structure and consists of an octamer of 2 copies each of the core histones, H2A, H2B, H3 and H4, around which is wrapped about 147bp of DNA. The *P. falciparum* genome encodes four main types of core histones (H2A, H2B, H3 and H4) and four histone variants H2A.Z, H2Bv, H3.3, and the centromere-specific H3, CenH3 (Miao et al., 2006). In most eukaryotes, nucleosome occupancy decreases upstream of transcriptionally active genes and increases in the regulatory regions of repressed genes (Huebert et al., 2012). In contrast, several studies in *P. falciparum* asexual blood stages suggested lower nucleosome occupancy in intergenic regions of all genes relative to the gene body and for the majority of genes, gene expression did not correlate with changes in nucleosome enrichment (Ponts et al., 2010; Westenberger et al., 2009). This was confirmed by a third study in which genes expressed during the asexual stage showed a nucleosome-free region at the transcription start sites (TSSs) and predicted upstream promoters (Ponts N, Harris EY, Lonardi S, 2011). The most recently published large-scale map of nucleosome positions provides a structural and regulatory framework to the transcriptional unit by demarcating landmark sites. Bártfai and colleagues observed transcription-coupled eviction of nucleosomes on strong transcription start sites during intraerythrocytic development, and demonstrated that nucleosome positioning and dynamics can be predictive for functional DNA elements (Kensche et al., 2015). Finally, nucleosome dynamics is a point of contention in subtelomeric regions: Ponts et al. reported an over-enrichment of nucleosomes for telomeric and subtelomeric regions in comparison with the rest of the genome, while Westenberger et al. observed that subtelomeric regions showed the highest fluctuations in nucleosome occupancy at different intra-erythrocyte stages of the parasite (Ponts et al., 2010; Westenberger et al., 2009). Taken together, how higher-order structures of chromatin are formed, regulated, and their effects on genomic activity and function in *P. falciparum* continues to remain elusive.

### 2.1.2 Histone Modifications

Histone post-translational modifications (PTMs) are covalent modifications of critical residues, typically lysine, arginine, serine or tryosine, in the N-terminal tails of nucleosomal histones. Histone PTMs can impact gene expression by altering chromatin structure or recruiting other modifiers. In general, histone modifications can affect many DNA-related processes, including transcription, recombination, DNA repair and replication, and chromosomal organization (Birney et al., 2007; Groth et al., 2007; Ruthenburg et al., 2007).

Mass spectrometric analyses have shown that *P. falciparum* histones carry more than 60 PTMs, including acetylation (ac), methylation (me), phosphorylation (phos), lipidation, etc (Dastidar et al., 2012; Treeck et al., 2011; Trelle et al., 2009). Very recently, a more systematic mass spectrometry study revealed 232 histone PTMs throughout the intraerythrocytic development cycle of *P. falciparum*, of which 160 had never been detected in *Plasmodium* and 88 had never been identified in any other species (Saraf et al., 2016). Acetylation and methylation of N-terminal histone lysines are the two most common PTMs with distinct distributions along both euchromatin and heterochromatin in *P. falciparum*, and can be broadly grouped into three categories: those that show positional changes during the intraerythrocytic development cycle (IDC), those that show variable abundance through the IDC, and those that occurs at a specific position in the genome and are maintained throughout the IDC (Bozdech et al., 2003; Preiser, 2013). The best-studied marks in *P. falciparum* are histone H3 lysine 9 acetylation (H3K9ac, a predominantly euchromatic PTM), histone H3 lysine 9 mono-, di-, or tri- methylation (H3K9me1,2 or 3, respectively, predominantly heterochromatic marks) and lysine 4 methylation (H3K4me3, primarily euchromatic marks) (Figure 2.1). These are discussed in greater detail in section 2.2. Of note, H3K9me3 is enriched in subtelomeric regions of all chromosomes and chromosome-internal regions of chromosomes 4, 7, 8 and 12 (Figure 2.2A).



Nature Reviews | Microbiology

**Figure 2.1.** Histone H3 post-translational modifications in *Plasmodium spp.* and domains that recognize the modified histone tails. Taken from (Doerig et al., 2015).

Malaria parasites have a large repertoire of enzymes catalyzing PTMs and proteins with PTM-binding modules (Figure 2.1; discussed in more detail in Sections 2.2 and 2.3). In some instances, these modifications have been implicated in regulating the transitions between developmental stages (discussed in more detail in Sections 2.2 and 2.3). Nevertheless, most studies in *P. falciparum* have focused on the correlation between a subset of histone modifications and gene expression of the *var* gene family, as discussed in sections 3.2 and 3.6.

### 2.1.3 Histone Variants

Histone variants, which differ from their canonical counterparts, are used to create specialized nucleosomes, generating functionally unique chromatin domains (Talbert and Henikoff, 2010). The specialization imparted by the histone H2A variant H2A.Z has been studied extensively in *P. falciparum*. Initial studies indicated that PfH2A.Z localizes to euchromatic intergenic regions along with H3K4me3 and H3K9ac throughout the IDC, and may be associated with promoter strength but not temporal activity (Bártfai et al., 2010). In addition, *P. falciparum* H2B.Z has been shown to dimerize with H2A.Z. These double variant nucleosomes show genome-wide enrichment in the promoter regions of most genes, including active *var* genes, suggesting that transcriptional regulation is highly dependent on the incorporation of variant histones (Petter et al., 2011).

Genome-wide mapping found that PfCenH3 demarcates centromeric regions of all *P. falciparum* chromosomes, which are also enriched in PfH2A.Z but devoid of typical pericentric heterochromatin marks such as H3K9ac and H4K20me3 (Hoeijmakers et al., 2011). A recent study found that the other variant form of histone H3, PfH3.3, demarcates GC-rich euchromatic coding region and subtelomeric repetitive sequences, and during trophozoite stages predominantly occupies AT-rich intergenic regions. Moreover, PfH3.3 specifically marks the promoter region of the active and poised *var* gene, but not silenced *var* genes (Fraschka et al., 2016). This distribution pattern is similar to that of the PfH2A.Z/PfH2B.Z double-variant nucleosomes around the transcription start site of the active *var* gene during the ring stage (Petter et al., 2013). Notably, PfH3.3 is retained at the AT-rich *var* promoter throughout the poised state (i.e., during parasite division in the trophozoite and schizont stages; discussed in Sections 3.2.1 and 3.2.2), thereby potentially contributing to epigenetic memory of *var* gene expression.

Taken together, the aforementioned studies suggest that the amino acid sequence variability of histones can, by itself or in a conjunction with histone modifications, regulate chromatin dynamics, which is important for gene expression and demarcating functional chromosome regions.

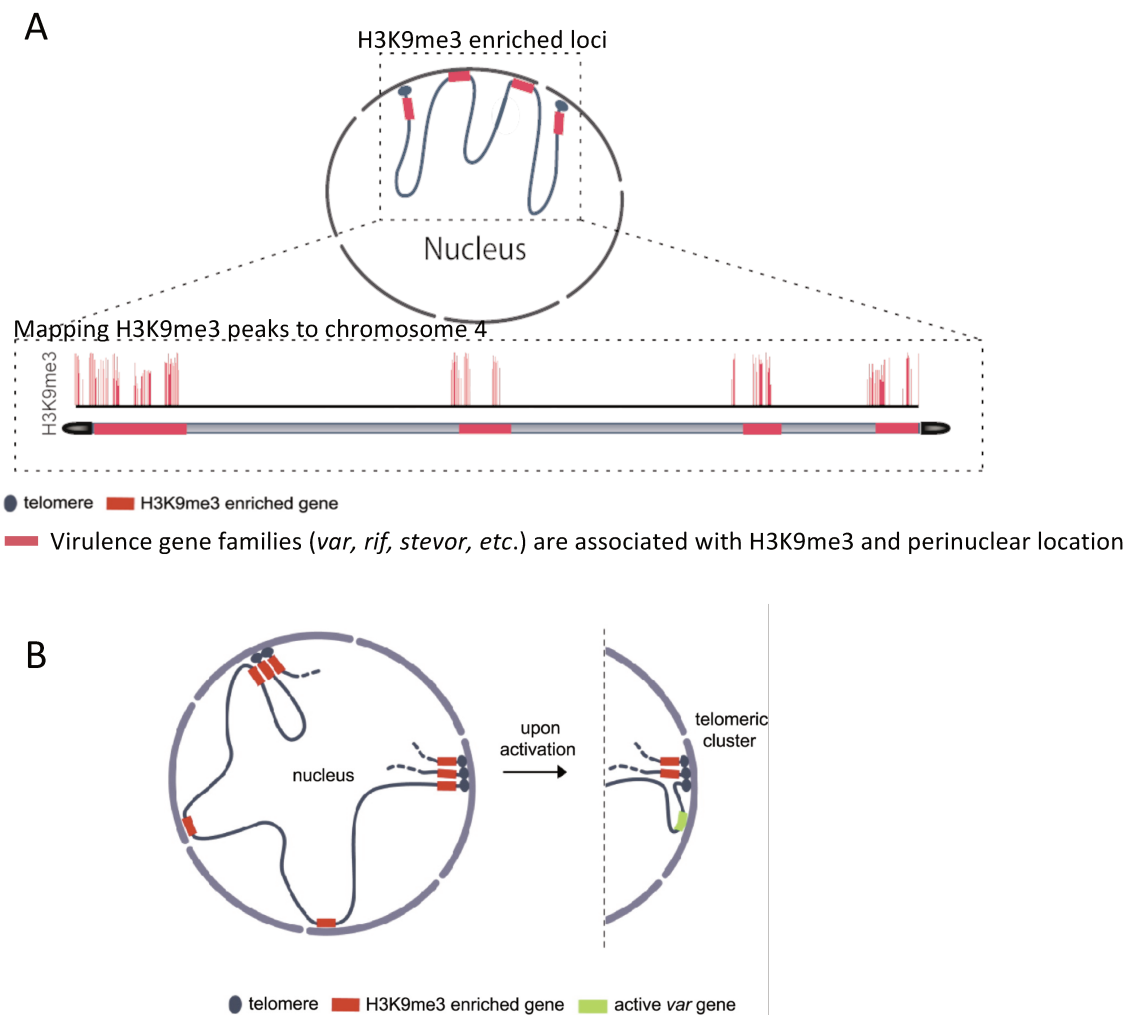
#### **2.1.4 Nuclear Organization**

The organization of the genome in the eukaryotic nucleus is complex and dynamic. Various features of nuclear architecture, including compartmentalization of molecular machines and the spatial arrangement of certain genetic loci, contributes to the nuclear gene regulatory process (Schneider and Grosschedl, 2007). However, increasing evidences has indicated that virulence gene activation or silencing in *P. falciparum* is associated with repositioning of the locus relative to other nuclear compartments and genomic loci.

Within the haploid *P. falciparum* genome, the approximately 60 copies of *var* genes, 180 copies of *rif* genes, and 24 copies of *stevor* genes are primarily located in subtelomeric regions, with a few clusters located in internal chromosomal regions (Gardner et al., 2002). More details about genome organization of *var* genes and how a single *var* gene is maintained in a transcriptionally active state will follow shortly in section 3.2.1. The subtelomeric as well as internal clonally variant genes localize to the nuclear periphery as part of 4-7 telomeric clusters (Freitas-Junior et al., 2000; Ralph et al., 2005). Lopez-Rubio et al. further

demonstrated that a mark considered specific for repressed *var* genes, H3K9me3, is associated with the perinuclear location of these telomeric clusters (Figure 2.2A) (Lopez-Rubio et al., 2009).

Notably, the single active *var* gene segregates away from these repressive centres and moves into a permissive site containing required transcription factors required for mRNA production (Figure 2.2B) (Freitas et al., 2005; Kyes et al., 2007; Lopez-Rubio et al., 2009; Ralph et al., 2005). The transition from the active to the poised state has been associated with the histone methyltransferase PfSET10 and H3K4me2 marks (Volz et al., 2012a) (See Section 3.6.1 for more details).



**Figure 2.2. Nuclear organization spatially regulates transcription of select genes.** A. Model for nuclear spatial organization of H3K9me3-enriched domains in a single chromosome. B. The relocation of an active *var* gene to a distinct perinuclear expression site. Adapted from (Lopez-Rubio et al., 2009; Ralph et al., 2005)

Besides, experimental data support that recombination efficiency of *var* genes is enhanced by the physical tethering of homologous region (Freitas-Junior et al., 2000; Taylor et al., 2000), suggesting nuclear architecture may determine the hot spots of recombination in malaria parasites. Future progress in our understanding of *Plasmodium* nuclear substructure and dynamics is essential to elucidate antigenic variation.

### 2.1.5 DNA Methylation

DNA modifications such as methylation of cytosine at the C5 position, 5mC, and methylation of adenine at the N6 position, m6A, are well established as regulators of gene expression in eukaryotic and prokaryotic systems. More recently, 5-hydroxymethylcytosine or 5hmC, an oxidation derivative of 5mC, has emerged as a key regulator of a number of cellular processes in eukaryotes, with the crosstalk between 5mC and 5hmC levels modulating gene suppression and heterochromatin remodeling. Active conversion of 5mC to 5hmC by Ten-eleven-translocases (Tet) may facilitate the maintenance of DNA methylation patterns during cell division, in coordination with DNA methyltransferases (Tahiliani et al., 2009). Moreover, 5hmC is not recognized by methyl-CpG-binding proteins that typically recognize 5mC, thus preventing the recruitment of histone deacetylases (discussed in section 2.2) and leading to the formation of transcriptionally competent chromatin (Clouaire and Stancheva, 2008; Jin et al., 2010).

In *P. falciparum*, based on bisulfite conversion and high throughput sequencing (BS-seq), a genome-wide map of DNA cytosine methylation was generated during the intraerythrocytic developmental cycle (Ponts et al., 2013). This study revealed that the malaria genome displays core promoter hypermethylation patterns and may share common features with undifferentiated plant and mammalian cells. Notably, intra-exonic methylation inversely correlated with gene transcription, and sharp transitions of methylation occurred at nucleosome and exon-intron boundaries. However, this study has several caveats: 1) The researchers analyzed a bulk parasite culture, which was not time-synchronized. Given the cyclic pattern of gene expression during the intraerythrocytic developmental cycle (discussed in section 3.1), a direct correlation cannot be made between DNA methylation and gene expression; 2) Because *var* genes are expressed only at 8-16 hours post erythrocyte-invasion, this study did not provide insights into whether 5mC levels contribute to *var* expression; and 3) Importantly, BS-seq cannot differentiate between 5mC and 5hmC.



Recently, Scherf and colleagues determined the existence of 5hmC in the *P. falciparum* genome and quantified the genome-wide distribution of 5hmC using BS-seq and oxidative bisulfite sequencing (oxBS-seq). They observed dynamic changes in 5hmC genomic enrichment in different stages of parasite development (ring, trophozoite and schizont). Particularly, 5hmC was differentially enriched in genes encoding select erythrocyte invasion proteins and was absent from the promoter of the single active *var* gene in comparison to the 59 silent *var* genes (Vembar and Scherf et al, manuscript in preparation). In the future, research could focus on establishing the correlation between DNA methylation and gene expression in *P. falciparum*, and could address the following questions: 1) How vital is DNA methylation in *P. falciparum*; 2) How it plays a role in modulating gene expression or even parasite development; and 3) The molecular components of the *P. falciparum* DNA methylation and demethylation machinery.

## **2.2 Histone Modification Machinery**

### **2.2.1 “Writers” and “Erasers” of the Histone Code**

Histone PTMs are controlled by the opposing actions of various enzymes for their addition and removal. Epigenetic writers such as histone acetyltransferases (HATs), histone lysine methyltransferases (HKMTs) and arginine methyltransferases (HRMTs) lay down epigenetic marks on histone N-terminal tails (Table 2.1). In contrast, erasers such as histone deacetylases (HDACs) and lysine demethylases (HKDMs) catalyze the removal of marks (Jenuwein and Allis, 2001) (Table 2.1).

HATs and HDACs function antagonistically to control histone acetylation states. HATs transfer the acetyl group from acetyl-CoA to the lysine residues in histones tails and are currently categorized into 5 families: GNATs (GCN5 *N*-acetyltransferases), MYSTs (MOZ, Ybf1/Sas3, Sas2, and Tip60), p300/CBP (CREB-binding protein), general transcription factor HATs, and nuclear hormone-related HATs (Lee and Workman, 2007). Ten predicted HATs belonging to the GNAT or MYST families are distributed in the malaria parasite genomes, as shown in Table 2.1. Among them, PfGCN5 preferentially acetylates H3K9 and K14 *in vitro* (Fan et al., 2004), and the perturbation of H3K9ac levels by attenuation of PfGCN5 activity leads to parasite growth inhibition (Cui et al., 2007). Another HAT protein PfMYST, which preferentially acetylates H4K5, H4K8, H4K12 and H4K16, is

reported to be involved in intraerythrocytic cell cycle regulation and DNA repair (Miao et al., 2010).

In *P. falciparum*, there are five HDAC homologues/orthologues (Table 2.1). PfHDAC1 is a class I enzyme homologous to yeast Rpd3 and is a nuclear protein (Joshi et al., 1999). PfHDAC2 and PfHDAC3 are provisionally assigned to class II and have not been characterized. The class III HDACs, also referred to as Sirtuins, use NAD<sup>+</sup> as a reactant to deacetylate acetyl-lysine residues of protein substrates; in *P. falciparum*, two paralogs PfSir2A and PfSir2B of the well-characterized *Saccharomyces cerevisiae* HDACIII, Sir2, have been annotated (Duraisingh et al., 2005; Tonkin et al., 2009). Although PfSir2A and PfSir2B are not essential to *P. falciparum* survival *in vitro*, gene knockouts abolished silencing and mutually exclusive expression of virulence genes and induced leaky transcription from numerous genes (Duraisingh et al., 2005; Tonkin et al., 2009). Substrates of PfSir2A include acetylated K9 and K14 of histone H3 and acetylated K16 of histone H4 (French et al., 2008). Moreover, PfSir2A plays a role in propagating silent heterochromatin from the telomere to chromosome internal regions, thus silencing telomere-proximal genes (discussed in more detail in section 3.2.1) (Duraisingh et al., 2005; Mancio-Silva et al., 2008). In addition, a novel negative regulatory function has been reported for PfSir2A in fine-tuning the transcription of ribosomal RNA genes which results in different levels of merozoite production (Mancio-Silva et al., 2013).

The methylation of lysine or arginine residues can occur in several states with complex readouts (Zhang and Reinberg, 2001). The lysine residues can house either mono-, di- or trimethyl moieties on their amine group, whereas arginine residues can carry mono- or dimethyl groups on their guanidinyll group. The *P. falciparum* genome encodes ten HKMTs and three conserved protein HRMTs (also referred to as PRMTs, for protein arginine methyltransferase) (Table 2.1). HKMTs typically contain a catalytic Su(var)3-9-Enhancer of zeste-Trithorax (SET) domain: since the focus of this thesis is PfHKMTs, they are discussed in greater detail in section 3.5. Of three putative HRMTs, PfPRMT1 has been experimentally confirmed to have methyltransferase activity *in vitro* and produces the activating histone mark H4R3 (Fan et al., 2009). PfPRMT1 contains an extended N-terminal region that is essential for the catalytic activity of this enzyme, and is present in both the cytoplasm and nucleus of asexual stage parasites (Fan et al., 2009). However, it is unknown whether the enzyme can affect transcription *in vivo*. The other two putative PfPRMTs were annotated as PfPRMT5

and PfCARM1 owing to sequence homology, although PfCARM1 is quite diverged from other known CARM1 homologues, including that found in *Toxoplasma gondii* (Saksouk et al., 2005). Both PfPRMT5 and PfCARM1 contain stretches of asparagine residues, which are encoded by AT-rich codons, and low complexity amino acid sequences.

The three histone lysine demethylases that have been bioinformatically identified belong to two families, lysine-specific demethylase1 (LSD1) and Jumonji C-domain containing demethylases (JHDMS). In contrast to HDACs, which are being considered as targets to develop parasite-specific small molecule inhibitors, the HKDMS are not well characterized. At present only one study indicated that PfLSD1 is involved in controlling variant expression of *var* genes and maintaining telomere repeat length (Comeaux, 2009).

**Table 2.1** Predicted and verified histone PTM writers and erasers in *P. falciparum*.

<b>Modification</b>	<b>Domain Family</b>	<b>Gene Name</b>	<b>Gene ID</b>	<b>Validated Activity</b>
<b>Histone acetylation</b>	GNAT	PfGCN5	PF3D7_0823300	(Fan et al., 2004)
			PF3D7_0109500	
			PF3D7_1437000	
			PF3D7_1003300	
			PF3D7_1323300	
			PF3D7_0805400	
	GNAT-related		PF3D7_1227800	
	MYST	PfMYST	PF3D7_1118600	
			PF3D7_0809500	
HAT1	PfHAT1	PF3D7_0416400	(Patel et al. 2009)	
<b>Histone deacetylation</b>	Class I	PfHDAC1	PF3D7_0925700	
	Class II	PfHDAC2	PF3D7_1472200	(Merrick & Duraisingh, 2007)
		PfHDAC3	PF3D7_1008000	
	Class III	Sir2A	PF3D7_1328800	
		Sir2B	PF3D7_1451400	
<b>Histone lysine</b>	SET	PfSET1	PF3D7_0629700	

		PfSET2	PF3D7_1322100	(Cui et al., 2008, Jiang et al., 2013)
		PfSET3	PF3D7_0827800	
		PfSET4	PF3D7_0910000	
		PfSET5	PF3D7_1214200	
		PfSET6	PF3D7_1355300	
		PfSET7	PF3D7_1115200	(Chen&Ding et al., 2016)
		PfSET8	PF3D7_0403900	(Cui et al., 2008)
		PfSET9	PF3D7_0508100	
		PfSET10	PF3D7_1221000	(Volz et al., 2012)
<b>Histone arginine methylation</b>	Class I	PfRMT1		
	Putative Class I	PfRMT4/PfCAR M1	PF3D7_0811500	
	Class II	PfRMT5	PF3D7_1361000	
<b>Histone demethylation</b>	LSD	PfLSD1	PF3D7_1211600	
	JMJC	PfJmJC1	PF3D7_0809900	
		PfJmJC2	PF3D7_0602800	

### 2.2.2 “Readers” of the Histone Code

The epigenetic “writers” are complemented by a wide range of “readers”, proteins that contain binding domains to recognize modified amino acid side chains (Figure 2.1 and Table 2.2). The best-studied histone reader in *P. falciparum* is the heterochromatin protein PfHP1 (Flueck et al., 2009; Pérez-Toledo et al., 2009). According to protein sequence analysis, PfHP1 is composed of two characteristic domains, a chromodomain (CD) that recognizes methylated lysines, and a chromo shadow domain (CSD) that enables oligomerization of HP1 protein leading to nucleosome aggregation and hence the formation of densely packed heterochromatin. PfHP1 binds to H3K9me3 *in vitro* and its genome-wide enrichment highly correlates with H3K9me3 distribution (Flueck et al., 2009; Pérez-Toledo et al., 2009).

Bioinformatic analyses predict a number of genes encoding other methylation or acetylation recognition domains including bromodomains, chromodomains, Tudor domains and plant homeodomains (PHDs) (Figure 2.1 and Table 2.2). A bromodomain is an approximately 110 amino acid protein domain that belongs to an extensive family of evolutionarily conserved protein modules which form a hydrophobic pocket that recognizes the acetylated lysine (Owen et al., 2000; Zeng and Zhou, 2002). Bromodomains recruit proteins such as HATs to specific chromosomal sites and contribute to chromatin remodeling (Zeng & Zhou, 2002). *P.*

*falciparum* has seven predicted bromodomain proteins that could potentially bind acetylated histones. One of them is the conserved histone acetyltransferase GCN5, which has been described above. The other characterized bromodomain protein, PfBDP1, has emerged as a regulator of several genes, including invasion-related genes, through binding to acetylated histones present in nucleosomes at the corresponding regulatory sites (Josling et al., 2015).

A chromodomain (chromatin organization modifier) contains ~ 50 amino acids that bind to methylated histone lysines (mono-, di- and tri-) and is well conserved across different eukaryotes. In mammals, chromodomain-containing proteins regulate chromatin remodeling and the formation of heterchromatin regions (Jones et al., 2000). There are 4 chromodomain-containing proteins in *P.falciparum*, PfMYST and PfHP1 contain a single chromodomain, while the chromodomain-helicase-DNA-binding protein 1 homologue contains two chromodomains (Volz et al., 2010).

The Tudor domain was originally identified as a region of 50 amino acids found in the *Drosophila* Tudor protein and was subsequently shown to bind to methylated lysine or arginine residues. Structural studies of human Tudor indicated that it adopts a strongly bent anti-parallel  $\beta$ -sheet structure consisting of five  $\beta$ -strands with a barrel-like fold and recognizes symmetrically dimethylated histones (Sprangers et al., 2003). Only one Tudor domain-containing protein PfTSN has been reported to be essential in the parasite's life cycle. Biochemical analysis showed that PfTSN possesses nuclease activity and that the Tudor domain binds to RNA (Hossain et al., 2008). The role of Tudor domain-containing protein in epigenetic regulation of gene expression in *P. falciparum* has not been studied as yet.

The 65 amino acid residue PHD finger was first discovered in homeodomain proteins of *Arabidopsis thaliana*, and resembles the metal binding RING domain and FYVE domain. It occurs as a single finger, but more frequently co-exists with adjacent PHD fingers and other histone PTM-binding modules (Ruthenburg et al., 2007). The PHD domain binds to specific acetyl and methyl marks on histone tails, recruiting transcription factors and nucleosome-associated complexes. Based on the specificity of PHD modules toward PTMs, PHD domains can be divided into those that binds unmodified histone H3 tails, those that are capable of binding H3K4me3, those that exhibit preference for H3K9me3 and H3K36me3, and those that are specific for histone H3 or H4 acetylated at various lysine residues (H3/H4ac) (Shi et al., 2007). In *P. falciparum*, although the complete gene has undergone a major expansion,

the PHD region is conserved between *Arabidopsis* Histone lysine N-methyltransferase gene and *Plasmodium spp.* (Kishore et al., 2013). Interestingly, bioinformatics tools have identified that one or several PHDs are located at the proximal and upstream regions of the SET domain in multiple *P. falciparum* HKMTs (e.g. PfSET10 harbors a PHD domain) (Cui et al., 2008).

Another histone mark reader protein in *P. falciparum* has been identified as Pf14-3-3I, which recognizes phosphorylated histones (Dastidar et al., 2013). Three putative 14-3-3 proteins are predicted in *P. falciparum*, of which Pf14-3-3I and Pf14-3-3II are expressed at higher levels in the asexual parasite stages, and share 70% and 25% sequence similarity with that of human, *Nicotiana tobaccum* and *Cryptosporidium parvum* 14-3-3 proteins. *In vitro* assays demonstrated that Pf14-3-3I binds to H3S10ph and H3S28ph, with a preference for H3S28ph over H3S10ph (Dastidar et al., 2013). In a number of eukaryotic organisms, the 14-3-3 protein family plays critical roles in cell signaling events that control progression through the cell cycle and transcriptional alterations. General mechanisms of 14-3-3 action include changes in activity of bound ligands, altered association of bound ligands with other cellular components, and changes in intracellular localization of 14-3-3-bound cargo (Yaffe, 2002). The *in vivo* function of 14-3-3 proteins in *P. falciparum* still awaits characterization.

**Table 2.2.** Predicted and verified histone mark readers in *P. falciparum*.

<b>Modification</b>	<b>Domain Family</b>	<b>Gene Name</b>	<b>Gene ID</b>	<b>Validated Binding?</b>
<b>Histone acetylation</b>	Bromodomain	PfGCN5	PF3D7_0823300	
		PfSET1	PF3D7_0629700	
			PF3D7_0110500	
			PF3D7_1033700	
			PF3D7_1212900	
			PF3D7_1234100	
			PF3D7_1475600	
<b>Histone methylation</b>	Chromodomain	PfHP1	PF3D7_1220900	(Perez-Toledo et al., 2009; Flueck et al., 2009)

		PfCHD1	PF3D7_1023900		
			PF3D7_1140700		
	Chromodomain-like	PfMYST	PF3D7_1118600		
	PhD domain	PfSET1	PfSET1		
		PfSET2	PfSET2		
		PfSET10	PfSET10		
		PfLSD1	PfLSD1		
		PfISWI	PfISWI		
				PF3D7_1433400	
				PF3D7_1008100	
				PF3D7_1310300	
				PF3D7_1475400	
				PF3D7_0310200	
				PF3D7_1141800	
				PF3D7_1360700	
			PF3D7_1460100		
<b>Histone phosphorylation</b>	14-3-3	PF14-3-3I	PF3D7_0818200	(Dastidar et al., 2012)	
		PF14-3-3II	PF3D7_1362100		

### 2.3 Histone Modifications in eu- and heterochromatin in *P. falciparum*

Several studies published over the past decade have identified various histone PTMs in different *P. falciparum* strains and at different stages of parasite intraerythrocytic development (Dastidar et al., 2012; Saraf et al., 2016; Treeck et al., 2011; Trelle et al., 2009). Moreover, the genome-wide distribution of select marks has also been described (Grewal and Jia, 2007; Lopez-Rubio et al., 2009)

The overarching observation is that the euchromatic marks H3K9ac and H3K4me3 are widely distributed across the genome, including in intergenic regions, and are enriched at promoters of actively transcribed genes (Lopez-Rubio et al., 2007, 2009). In contrast, the heterochromatin mark H3K9me3 is usually seen in broad domains over silenced subtelomeric and chromosome-internal regions (Chookajorn et al., 2007, Lopez-Rubio et al., 2007; Lopez-Rubio, Mancio-Silva, & Scherf, 2009). Intriguingly, unlike other organisms (Grewal and Jia, 2007), enrichment of H3K9me3 in *P. falciparum* does not spread into the centromere core. In subtelomeric regions, H3K9me3 is associated with gene families such as *var*, *rifin* and *stevor*, which are characterized by high variability between strains and contribute to phenotypic plasticity of blood stage parasites (Lopez-Rubio, Mancio-Silva, & Scherf, 2009).

### **3. The Roles of Histone Methylation and Histone Methyltransferases in *P. falciparum***

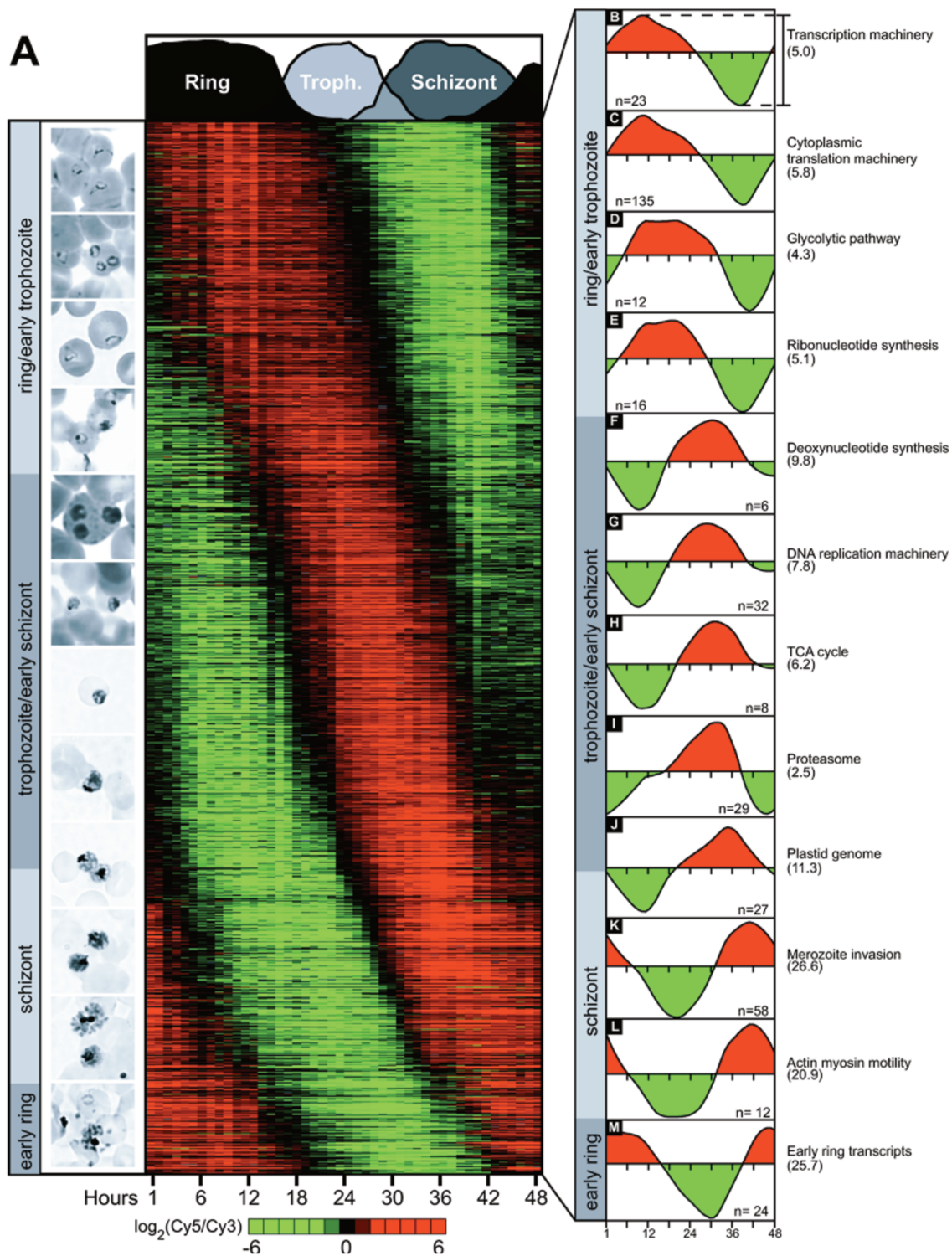
#### **3.1 Just-in-time transcription during asexual development**

The sequencing of the *P. falciparum* genome in 2002 revealed that the 23.3 Mb genome is partitioned amongst 14 chromosomes and is highly AT-rich throughout the genome (Gardner et al., 2002). The current gene annotation identifies ~5700 open reading frames, and nearly 60% of these lack homologs in other organisms, in spite of being conserved in other *Plasmodium* species. In the past decade, numerous studies have illuminated our understanding of transcriptional, post-transcriptional, and post-translational gene regulatory processes associated with *Plasmodium* life cycle progression (Doerig et al., 2015; Vembar et al., 2016).

Transcriptomic studies showed an unusual, continuous cascade of gene expression in *P. falciparum* asexual blood stages, as shown in the left panel of Figure 3.1, with functionally related genes being transcribed at the same time, thus determining the distinct morphology and physiology of each developmental stage (Bozdech et al., 2003; Foth et al., 2011). The right panel of Figure 3.1 illustrates that most genes associated with basic metabolic and cellular functions are expressed in the ring and trophozoite stages, while genes encoding surface molecules involved in host-parasite interactions are expressed in the late stages. This led researchers to postulate that transcription of most asexual stage genes occurs just at the time when protein function is needed, the so-called “just-in-time” transcription model.

Intriguingly, genome-wide analysis of histone modifications and transcriptomic data from *P. falciparum* asexual stages provided a direct link between histone methylation and acetylation and gene activation (Salcedo-Amaya et al., 2009). In addition, genome-wide nucleosome occupancy studies indicated a link between the nucleosome positioning and transcriptional activity (Ponts N, Harris EY, Lonardi S, 2011). However, none of these studies provide a comprehensive explanation for “just-in-time” transcription. As discussed in Section 2.1.5, unpublished data from Scherf and colleagues links 5hmC levels in promoters of stage-specific genes to activation, providing additional regulatory information to an already complex process. Additionally, post-transcriptional regulation of mRNA outcomes by RNA-binding and –regulatory proteins further fine-tunes the steady state transcriptome of any given parasite stage (Vembar et al., 2016).





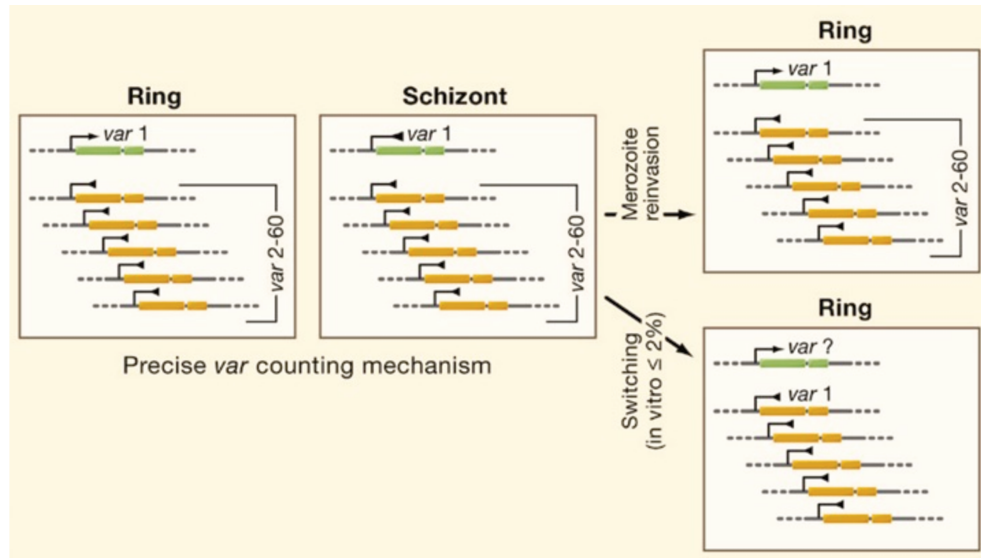
**Figure 3.1.** Overview of the transcriptome of *P. falciparum* intraerythrocytic development at 1 h resolution. The morphological characteristics of the intraerythrocytic parasites are shown on the left, aligned with transcriptional profiles by phase of sequential gene expression. The temporal ordering of biological processes and functions are shown on the right. Each graph corresponds to the average expression profile for the genes in each set and the mean peak-to-trough amplitude is shown in parentheses. Taken from (Bozdech et al., 2003b).

## 3.2 Histone lysine methylation regulates virulence gene expression

### 3.2.1 Monoallelic *var* expression

*P. falciparum* employs a strategy of antigenic variation to evade destruction by the host immune system (Scherf et al., 2008a). A wide range of clonally variant gene families have been identified in *P. falciparum*, the most important being the *var* gene family, comprising of approximately 60 genes encoding the surface molecule PfEMP1 (discussed in section 1.2.2) (Smith et al., 1995; Su et al., 1995). Approximately 60% of *var* genes are located within polymorphic subtelomeric domains of *P. falciparum* chromosomes, adjacent to regions that are composed of a mosaic of non-coding repetitive elements called telomere-associated repeat elements or TAREs. On average, one to three *var* genes exist close to telomeres pointing in different directions, either in a tail-to-tail, head-to-tail, or head-to-head orientation (Kraemer et al., 2007). In addition, some chromosomes (4,7, 8 and 12) harbor central *var* genes that are organized as tandem head-to-tail arrays (Gardner et al., 2002). Bioinformatic analysis of *var* genes has led to their classification into the *ups A, B, C, or E* sub-types based on the sequence of the 5' flanking regions (Lavstsen et al., 2003).

Each parasite expresses a single *var* gene at any given time, maintaining the remaining 59 members of the family in a transcriptionally silent state using a precise counting mechanism (Scherf et al., 2008a, 2008b). The greatest level of *var* gene transcription occurs in early ring stage parasites (8-16 hours post-erythrocyte invasion (hpi)), while the levels decrease dramatically in the trophozoite and schizont stages (Kyes et al., 2003); in contrast, the peak level of PfEMP1 protein expression appears at 25 hpi (mid-stage trophozoites). Regulation occurs at the level of transcriptional initiation in a monoallelic manner (Scherf et al., 2008a). The same *var* gene is expressed in subsequent generations, although a minority of parasites may switch the expression to a new *var* gene to avoid immune clearance and prolong infection. Fast and slow switching from one *var* gene to the next has been observed in parasites *in vitro* (Horrocks et al., 2004). Figure 3.2 is a schematic diagram depicting monoallelic *var* expression.



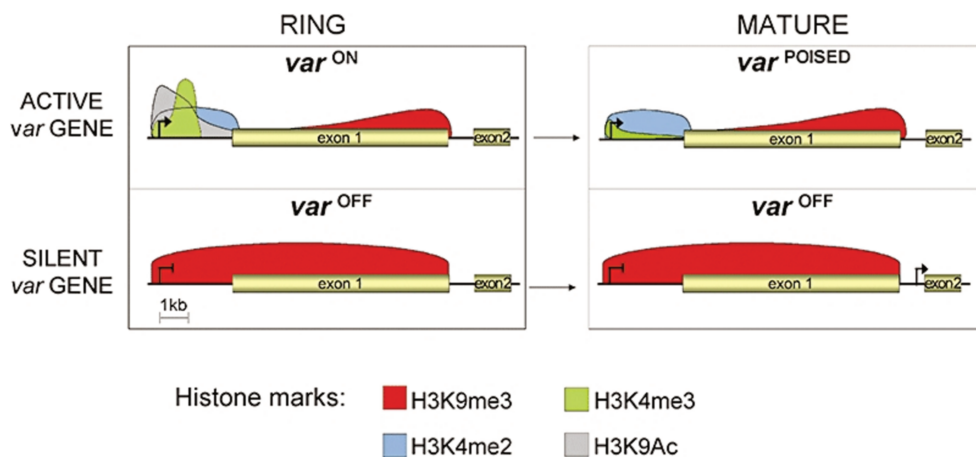
**Figure 3.2** Monoallelic expression of *var* genes. Taken from (Scherf et al., 2008b).

The study of antigenic variation motivated the initial investigation of epigenetic regulation in malaria parasites (Scherf et al., 2008a). The histone deacetylase PfSir2A and histone acetylation were identified as the first epigenetic factors involved in monoallelic expression of *var* genes (Duraisingh et al., 2005; Freitas-Junior et al., 2005). This was followed by several studies showing that factors such as spatial nuclear localization, noncoding RNAs from the *var* intron and histone variants also contributed to the silencing and counting process (Scherf et al., 2008a, 2008b). Moreover, several *cis*-acting DNA sequence elements of the *var* gene locus have been postulated to be key to monoallelic expression (Guizetti and Scherf, 2013). More recently, a trans-acting GC-rich noncoding RNA was identified as a regulator of the gene counting mechanism (Guizetti et al., 2016).

Histone modifications are the best-studied components of the epigenetic regulation machinery. They are essential for monoallelic *var* gene expression as well as general gene activation. In the next sub-sections, we focus on the mechanism by which histone lysine methylation and acetylation regulate monoallelic expression in a dynamic manner. In particular, we discuss *var* gene activation marks (H3K4me2, H3K4me3 and H3K9ac), repression marks (H3K9me3 and H3K27me3) and marks correlated with epigenetic memory (H3K4me2).

### 3.2.2 Histone H3 marks and *var* gene activation

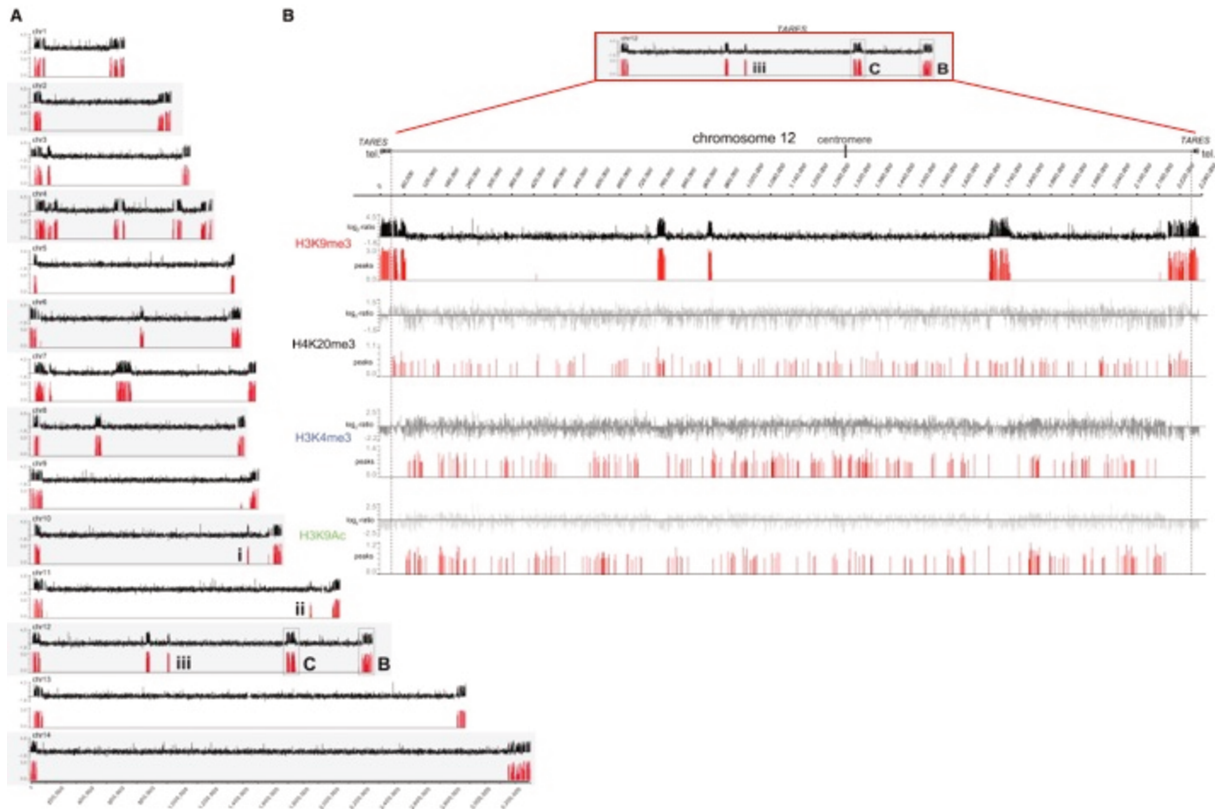
Chromatin immunoprecipitation (ChIP) analysis demonstrated the highly dynamic change in histone H3 methylation status. As shown in Figure 3.3, silent *var* genes are enriched in H3K9me3 at the 5'UTR and gene body. In contrast, *var* gene activation is linked to the replacement of H3K9me3 with H3K9ac and H3K4 methylation in promoter regions (Figure 3.3) (Lopez-Rubio et al., 2007, 2009). Furthermore, when *var* gene transcription ceases in the trophozoite and shizont stage, H3K4me2 is the major PTM in the promoter region of the so-called poised *var* gene, maintaining epigenetic memory for its reactivation in the next replication cycle (Lopez-Rubio et al., 2007; Volz et al., 2012a).



**Figure 3.3** Monoallelic expression of *var* genes regulated by histone PTMs. Taken from (Scherf et al., 2008b).

Lopez-Rubio *et al* analyzed the genome-wide distribution of different histone modifications and observed that the typical mark of facultative heterochromatin H3K9me3 was enriched at the loci of silenced *var* genes both at subtelomeric and internal chromosomal regions (Lopez-Rubio et al., 2009). The most striking feature when comparing the H3K9me3 distribution in *P. falciparum* to other eukaryotes is that H3K9me3 is restricted to subtelomeric and select chromosome internal regions in *P. falciparum* (Figure 3.4). While in many eukaryotes, the centromeres and pericentromeric regions are embedded in heterochromatin, which contributes to faithful chromosome segregation and genome stability, *P. falciparum* H3K9me3 is enriched in genes that show a variegated expression profile such as the clonally variant *var* genes at pericentromeric heterochromatin (Grewal and Jia, 2007; Lopez-Rubio et al., 2009).

This pivotal study was followed by other investigations that characterized the whole genome profiles of other histone methylation and acetylation marks, to obtain a complete picture of how these chromatin modifications contribute to the tight regulation of *var* gene expression (Bartfai et al., 2010; Salcedo-Amaya et al., 2009). Figure 3.4 shows the dynamics of histone modifications in select *var* gene loci.



**Figure 3.4.** Distribution of select histone PTMs along *P. falciparum* chromosomes. A. Genome-wide identification of the H3K9me3 landscape in the 14 chromosomes of *P. falciparum* using ChIP-on-chip. B. The distribution profiles of different histone modifications including the repressive mark H3K9me3, another repressive mark (H4K20me3) and activation marks (H3K4me3 and H3K9ac) on chromosome 12. Taken from (Lopez-Rubio et al., 2009).

### 3.2.3 The histone code reader PfHP1 is involved in *var* gene silencing

The chromodomain of PfHP1 binds to H3K9me3 *in vitro* and PfHP1 is enriched at H3K9me3 marks throughout the genome (Flueck et al., 2009; Pérez-Toledo et al., 2009). Genome-wide analysis indicated that PfHP1 occupancy is restricted to 425 genes mostly associated with variegated gene expression and includes gene families coding for virulence factors required for erythrocyte invasion and host-parasite interactions. PfHP1 is localized mainly at the nuclear periphery during asexual blood stage development, co-localizing with telomeric

clusters where its target genes are located (Flueck et al., 2009). The conditional depletion of PfHP1 de-repressed 52 of 60 *var* genes, confirming an essential role for PfHP1 in *var* gene silencing (Brancucci et al., 2014). Notably, PfHP1 is absent from centromeric regions of *P. falciparum*, which is different from the human host, where HP1 localizes to the centromere and plays a broader role in gene expression (Ayyanathan et al., 2003).

### **3.3 The interplay of histone methylation and acetylation regulates sexual commitment**

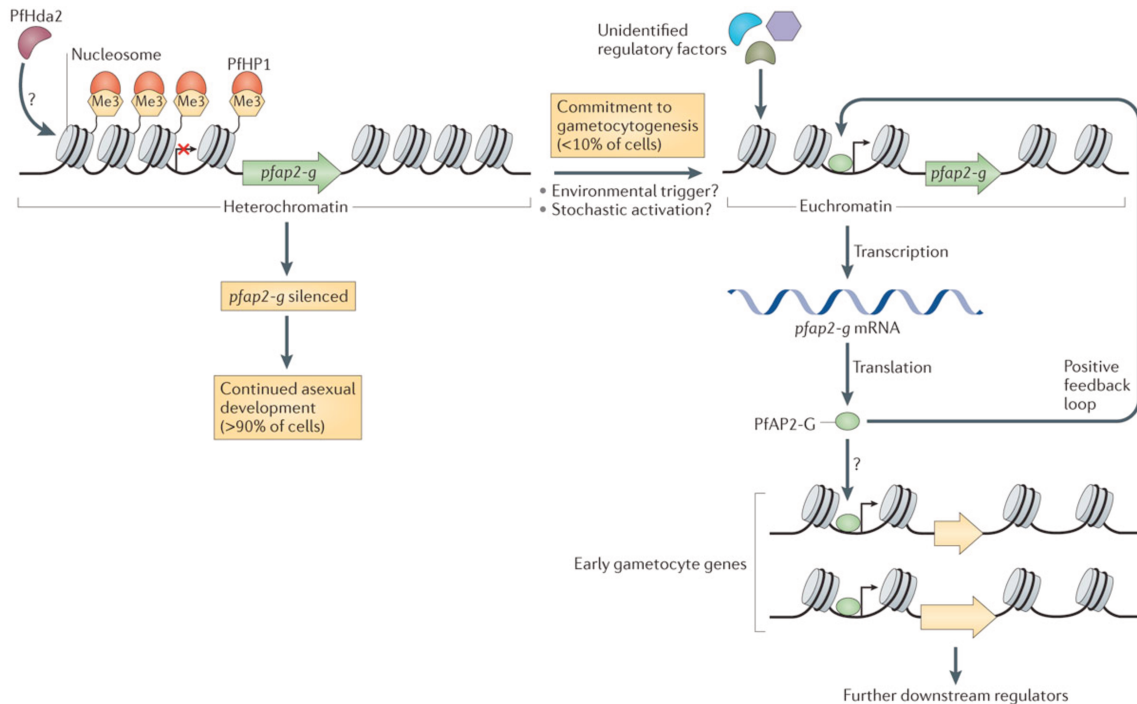
Although the distinct crescent shape of *P. falciparum* gametocytes was described as early as 1880, the molecular mechanisms involved in commitment to sexual development have remained a mystery. Recent studies have shown that a coordinated interplay between histone modifications and specialized transcription factors regulates the sexual development of malaria parasites (Kafsack et al., 2014; Sinha et al., 2014). Recent studies indicate that facultative chromatin leads to variegated expression of a transcription factor that initiates sexual development in malaria parasites, in a mechanism similar to *var* gene regulation. This transcription factor belongs to the Apicomplexan AP2 (ApiAP2) family of DNA-binding proteins, which are homologous to the Arabidopsis/Ethylene Response Factor (AP2/ERF) family of transcription factors found in plants (De Silva et al., 2008). The *P. falciparum* genome encodes 27 members of the ApiAP2 family; however, only some of them have been studied in detail in different stages of *Plasmodium* development (Balaji, Madan Babu, Iyer, & Aravind, 2005; Flueck et al., 2010; Painter, Campbell, & Llinás, 2012).

A single member of the PfAP2 family, PfAP2-G, is silenced by facultative heterochromatin similar to *var* genes, and was suggested to play a role in parasite development (Lopez-Rubio et al., 2009). Recently two independent groups have proved that AP2-G is a master regulator of gametocytogenesis, as a transcriptional switch controlling a divergent decision in malaria parasites (Kafsack et al., 2014; Sinha et al., 2014). In *P. falciparum*, PfAP2-G regulates a handful of early gametocyte genes, and appears to be necessary for repression of the silent loci prone to spontaneous activation (Kafsack et al., 2014). In the murine malaria parasite *P. berghei*, in addition to PbAP2-G as a key regulator of gametocyte commitment, Sinha *et al* identified a second ApiAP2 member PbAP2-G2, the deletion of which significantly modulated but did not abolish gametocytogenesis, indicating that a cascade of ApiAP2 proteins may be involved in sexual commitment in malaria parasites (Sinha et al., 2014).

While studying heterochromatin-mediated gene silencing in the asexual stages of blood development, Brancucci *et al* observed that the conditional knock-down of PfHP1 caused not only de-repression of *var* genes, but also up-regulation of AP2-G and thereby increased gametocyte production (up to a 10-25 fold increase) (Brancucci et al., 2014). This demonstrated that PfHP1 depletion renders heterochromatin at the PfAP2-G locus highly unstable, leading to activation of PfAP2-G (Figure 3.5). Besides, PfHP1 facilitates heterochromatin compaction by recruiting other methyltransferases or bridging neighboring nucleosomes; this might lead to up-regulation of known early gametocyte genes that are not directly associated with PfHP1 (Flueck and Baker, 2014).

A similar phenotype was observed when PfHda2, a component of the histone deacetylation machinery, was inactivated (Figure 3.5). Besides de-repressing *var* genes, a switch to sexual differentiation was observed. It was suggested that PfHda2 depletion leads to the hyper-acetylation of H3K9 residues in nucleosomes along heterochromatic gene loci, marking an active promoter state (Coleman et al., 2014).

Thus, integrated epigenetic and transcriptional mechanisms regulate *Plasmodium* developmental stage commitment. Any factor that is involved in establishing histone methylation-dependent chromatin and maintaining facultative heterochromatin can disturb the fine balance between asexual and sexual stage development.



**Figure 3.5.** Proposed model for the regulation of sexual commitment in *P. falciparum*. In a small population of parasites, silencing regulator HP1 and repressive histone code H3K9me3 are removed, resulting in *pfap2-g* is transcription and expression. A rise in PfAP2-G levels probably activates the expression of early gametocyte genes, thereby initiating gametocytogenesis. PfHda2 is also required for silencing of the *pfap2-g* locus. Taken from (Josling and Llinás, 2015).

### 3.4 Histone lysine methylation may contribute to hypnozoite quiescence

Hypnozoites represent non-dividing forms in the infected liver, and are associated with latency and relapse in some types of human malaria infections in the absence of fresh infectious mosquito bites. *P. ovale* and *P. vivax* produce hypnozoites as well as the monkey malaria parasite *P. cynomolgi*. Relapse can occur after symptoms have been absent for weeks, months, or even years, making the study of hypnozoites of great biological and medical significance. Currently, the identification of hypnozoites within a hepatocyte is performed by immunostaining with antibodies against Hsp70 and the manual counting of small (hypnozoite) and large (schizont) forms (Dembele et al., 2011). However, hypnozoites remain poorly characterized due to the inherent problem to produce sufficient numbers for molecular and biochemical analysis.

One hypothesis is that epigenetic mechanisms may control hypnozoite commitment and/or activation (Doerig et al., 2015). A key study supports this hypothesis: Dembélé *et al* established a long-term *in vitro* cultivation system of *Plasmodium* liver stages and were able

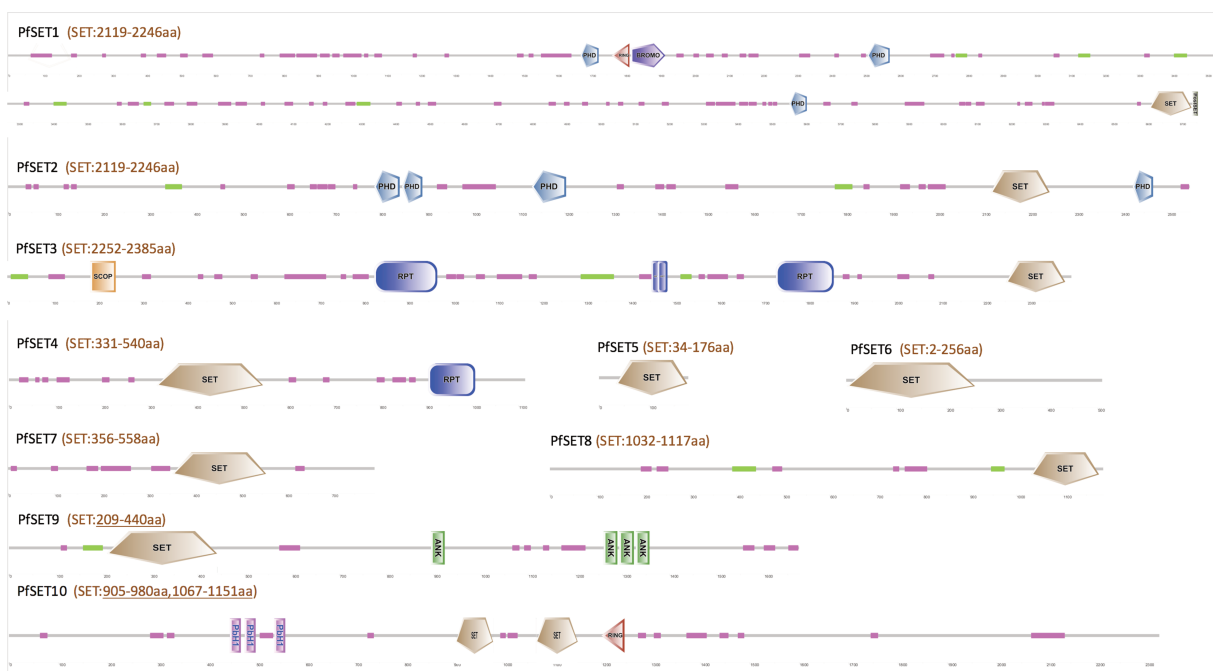


to detect hypnozoites of the monkey malaria parasite *P. cynomolgi* (Dembélé et al., 2014). This pioneering work opened multiple avenues to study the quiescence mechanism at the cellular and molecular level for the first time and to screen small molecule inhibitors that potentially inhibit HKMTs. Strikingly, in the presence of lower concentrations of the *Plasmodium*-specific HKMT inhibitor TM2-115, they observed a significantly higher proportion of dividing hepatic forms relative to hypnozoites. Furthermore, when liver stage cultures that had been treated with atovaquone to clear dividing forms were subsequently treated with TM2-115, there was a re-emergence of dividing forms, indicative of hypnozoite activation. Taken together, this suggested that histone methylation-dependent modifications are involved in the regulation of hypnozoite quiescence, which is similar to sexual stage commitment where alternative developmental decisions need to be made. More importantly, this study also unveiled a novel strategy of eliminating hypnozoites: to treat hypnozoites with HKMT inhibitors to awaken them from quiescence and then kill them with a second drug that targets mature dividing forms.

### **3.5 Predicted HKMTs in *P. falciparum***

In 2000, the Su(var)3-9-Enhancer of zeste-Trithorax (SET) domain of the mouse/human protein Suv39h1/KMT1a was demonstrated to be the catalytic domain for protein lysine methylation (Rea et al., 2000). Since then, many SET domain-containing molecules have been reported to be histone lysine methyltransferases (HKMTs) in different species (Pontivianne et al., 2010; Qian and Zhou, 2006). In addition to the canonical SET domain-containing enzymes, *Saccharomyces cerevisiae* Dot1 and its mammalian homolog DOT1L (DOT1-Like) also possess histone methyltransferase activity toward histone H3 lysine 79 (Nguyen and Zhang, 2011). Specific HKMTs can mono- or di-methylate, symmetrically or asymmetrically, specific histone lysine residues through the following mechanism: S-Adenosyl methionine (SAM; the methyl donor) and the lysine residue of the substrate histone tail are bound and properly oriented in the catalytic pocket of the SET domain. Next, a nearby tyrosine residue deprotonates the  $\epsilon$ -amino group of the lysine residue, after which the lysine chain makes a nucleophilic attack on the methyl group on the sulfur atom of SAM, transferring the methyl group to the lysine side chain, and generating methylated histone and S-Adenosyl homocysteine (SAH) as a by-product.

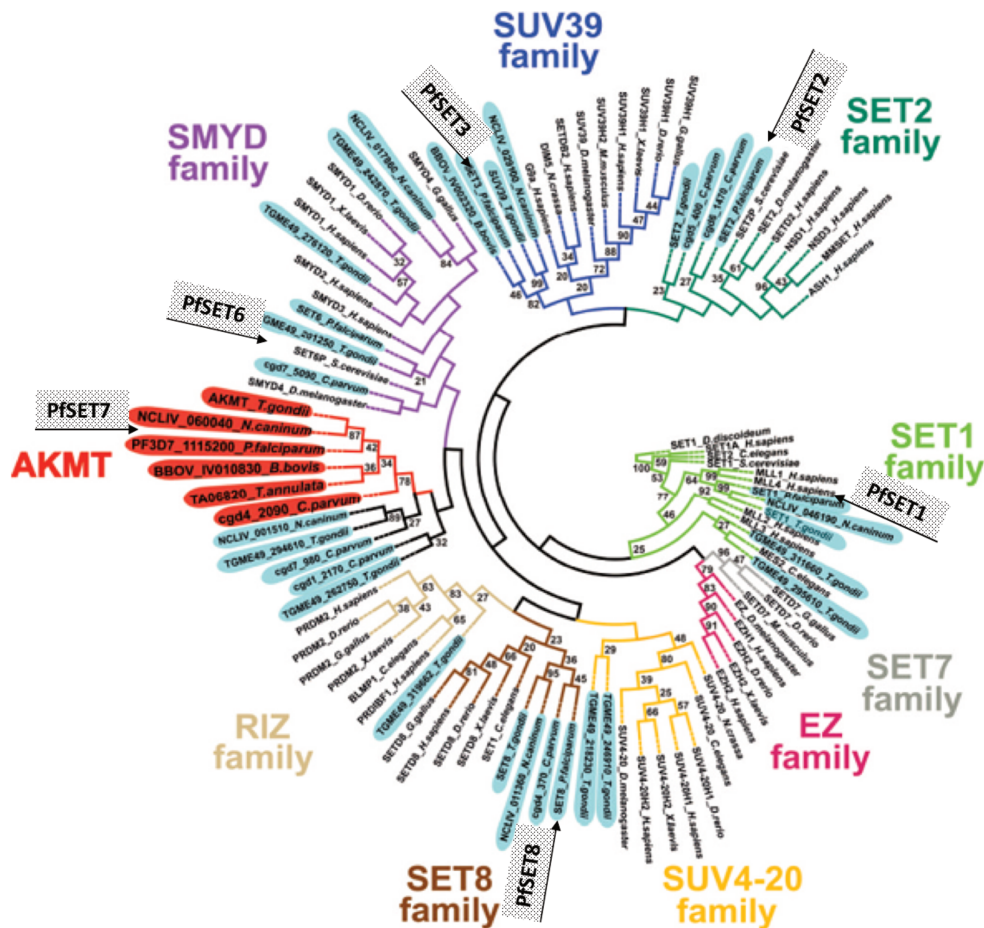
In *P. falciparum*, while no *Dot1* ortholog is evident, ten predicted SET domain-containing HKMTs have been found, based on the SET consensus sequence. The genes encoding SET-domain containing proteins are distributed on 9 chromosomes and their sizes range from 178 to 6,753 aa (Table 3.1). PfSET1, PfSET2 and PfSET8 are named based on the sequence similarities with *S. cerevisiae* SET1, *S. cerevisiae* SET2 and human SET8/PrSET7 proteins, respectively (Cui et al., 2008). Analysis of the domain organization of the PfSETs revealed the presence of additional domains such as the Myeloid-Nervy-DEAF1 (MYND) domain, the bromodomain, plant homeodomain (PHD) motif and HDAC-interacting domain (Figure 3.6). PfSET1 and PfSET2, for example, contain PHD domains (Figure 3.6), domains that have been shown to bind to all nucleosomal histones and associate with chromatin-mediated transcriptional regulation (Eberharther et al., 2004). PfSET4 and PfSET6 belong to the SMYD3 family, where the conserved SET domain is split into two parts by a MYND domain (Figure 3.6). The MYND domain is a zinc finger motif that primarily functions as a protein–protein interaction module (Spadaccini et al., 2008; Spellmon et al., 2015).



**Figure 3.6** Domain organisations of predicted SET-domain HKMTs in *P. falciparum*.

Phylogenetic analysis separates the PfSET proteins into six clusters, based on their methylation specificity (Figure 3.7 and Table 3.1) (Cui et al., 2008; Sivagurunathan et al., 2013). Consistent with this subdivision, methylation marks have been found on amino acid residues K4, K9 and K36 of histone H3 and lysine K20 of histone H4 by specific anti-methyl histone lysine antibodies and by proteomic studies of asexual blood stages (also see section

2.3). However, because bacterial expression of recombinant PfSET proteins with detectable HKMT activity was not successful, the substrate specificity and resulting methylation patterns for the PfSETs remain undefined (Cui et al., 2008).



**Figure 3.7.** Phylogenetic analyses of SET and post-SET domains from Apicomplexan lysine methyltransferases identify several subfamilies of protein lysine methyltransferases. Adapted from (Sivagurunathan et al., 2013).

To provide insight into the potential functions of the PfSETs, Jiang *et al* attempted to knockout the predicted PfSET genes (Jiang et al., 2013). Six of the ten PfSET genes were refractory to knockout attempts via double crossover homologous recombination (Table 3.1), suggesting essentiality of these genes in asexual blood stage parasites. Independently, to visualize their location within the parasite, Volz *et al* expressed epitope-tagged versions of select PfSET proteins that were identified as components of the nuclear proteome of asexual blood stage parasites (Volz et al., 2010). For example, the authors found that the putative H3K9me3 methyltransferase PfSET3 clusters around the nuclear periphery, similar to PfSir2

and telomeric clusters containing silenced *var* genes. Below we discuss the two PfSET proteins that have been functionally characterized in *P. falciparum* blood stages.

**Table 3.1** Predicted HKMTs in *P. falciparum*

Gene Name	Domain Family	Length (AA)	Predicted Site-Specificity	KnockOut
PfSET1	SET1	6,753	H3K4me1-3	NO
PfSET2	SET2	2,548	H3K36me2,3	YES
PfSET3	Suv39	2,399	H3K9me2,3	NO
PfSET4	SMYD3	1,114	H3K4	YES
PfSET5	Unknown	178	Unknown	YES
PfSET6	SMYD3	509	H3K4	NO
PfSET7	AKMT	810	Unknown	NO
PfSET8	SET8/Pr-SET7	1,186	H3K20me1-3	YES
PfSET9	Unknown	1,674	Unknown	NO
PfSET10	SET	2,329	H3K4me1-3	NO

Adapted from (Cui et al., 2008; Jiang et al., 2013)

### 3.6 PfSET proteins associated with *var* gene expression

#### 3.6.1 PfSET10 maintains the active *var* gene in a poised state

PfSET10, a putative H3K4 methyltransferase, encodes a 271 kDa protein with little sequence similarity to other known HKMT SET domains in *P. falciparum*, *S. cerevisiae*, *Drosophila melanogaster* and humans. In addition to the SET domain, PfSET10 contains a PHD domain (Figure 3.6), which shares a number of amino acid residues with zinc finger domains that preferentially bind non-methylated histone H3. PfSET10 localizes to a single spot within the nuclear periphery of late ring, trophozoite and schizont stage parasites, which is coincident with the perinuclear *var* expression site within an apparent euchromatic region (Volz et al., 2010, 2012b). Given that the PfSET10 nuclear signal is detected only after *var* transcription, researchers have suggested that PfSET10 actually associates to the *var* expression site to maintain the active *var* gene in a poised state; this also correlates with H3K4me2, the predicted outcome of PfSET10 methylation and a histone mark associated with the poised state and epigenetic memory (see section 3.2.2). Additionally, PfActin-1, which is involved in *var* gene positioning and spatial chromosome organization (Zhang et al., 2011), was detected as one of PfSET10's interacting factors. Taken together, Volz et al. suggested that PfSET10

maintains a permissive chromatin environment at the active *var* promoter, and transmits this to daughter generations via epigenetic memory.

Moreover, it is likely that PfSET10 “marks” a transcription site within the nucleus that is utilized by not only *var* genes but also other potentially essential genes. This could explain the essential character of the PfSET10 gene during blood stage development (Volz and Cowman, 2012).

### **3.6.2 PfSETvs represses *var* gene expression**

PfSET2, also called PfSETvs, is an orthologue of the *D. melanogaster* ASH1 protein, which is responsible for H3K36-specific methylation. Disruption of PfSETvs causes loss of H3K36me3 at the transcription start site of all silent *var* genes and leads to their de-repression. Intriguingly, H3K9me3 depletion was also observed at the promoter of the de-repressed *var* genes in PfSETvs knockout parasites, indicating that the deposition of both histone modification patterns may be inter-dependent (Jiang et al., 2013). Notably, the authors demonstrated that multiple *var* transcripts are made in a single nucleus, which are subsequently translated and transported to the erythrocyte membrane allowing the functional clustering of multiple PfMEP1s in a single parasite. This provides a new strategy to develop antimalarial vaccines by generating transgenic strains that express most or all PfEMP1 proteins (see section 1.2.2).

A second independent study supported the role of PfSETvs in regulating *var* gene expression, and also described the recruitment of PfSETvs to *var* genes through direct interactions with the unphosphorylated C-terminal domain of RNA polymerase II (Ukaegbu et al., 2014). The implications of this to monoallelic *var* expression is unclear.

## **3.7 Development of HKMT inhibitors as antimalarials**

Efforts to control malaria are hampered by the development of drug resistance in parasites, insecticide resistance in mosquitoes, and the lack of an effective vaccine (as described in section 1.3). Here we focus on effective malaria drug discovery, which should include both the development of improved antimalarials from existing compounds and the discovery of new drug targets and novel small molecule inhibitors.

As described in the previous sections, virulence gene expression in *P. falciparum* is governed by the interplay of different histone methyl marks in coordinating transcriptional activity (Scherf et al., 2008a). Given the essentiality of several histone PTM modifiers to parasite survival and the wide-spread impact of epigenetic machineries on parasite development, there are obvious implications for rational drug design against histone methylation and acetylation that could potentially interfere with all life cycle stages. Moreover, the principle of using drugs to alter protein PTMs is quite different from currently known antimalarial mechanisms, which target heme polymerization, induce oxidative stress or interfere with other metabolic processes, presenting as an optimal/ attractive combination therapy for treatment of malaria.

The HKMT inhibitor BIX-01294, which is a diaminoquinazoline that targets human G9a and the closely related H3K9 methyltransferase G9a-like protein, and one of its analogues TM2-115, were found to inhibit the growth of drug-resistant and sensitive *P. falciparum* parasites *in vitro* with IC<sub>50</sub> values  $\leq 100$  nM; the inhibitory effect was rapid and effective against all stages of the intraerythrocytic life cycle (Malmquist et al., 2012). BIX-01294 or TM2-115 treatment significantly reduced parasite histone H3K4me3 levels in a concentration-dependent and exposure time-dependent manner (Malmquist et al., 2012). Given these promising effects, BIX-01294 and TM2-115 were further characterized for drug efficacy and pharmacokinetic properties: 1) Activities against *ex vivo* clinical isolates of both *P. falciparum* and *P. vivax* were similar, with potencies of 300 to 400 nM; 2) Gametocyte inhibition occurred at micromolar levels; 3) Both compounds displayed oral efficacy in *in vivo* mouse models of *P. berghei* and *P. falciparum* infection (Malmquist et al., 2015). Together with studies that assessed derivatives of BIX-01294 and TM2-115 for antimalarial activity, it is strongly evident that the transcriptional activation mark H3K4me3 is critical for all intraerythrocytic life cycle stages such that its disruption for a short amount of time results in irreversible commitment to cell death. Moreover, these inhibitor-based growth studies have motivated increased efforts into active recombinant PfHKMT production to both biological and enzymatic characterization (Chen and Scherf, unpublished) and target-based discovery of novel or specific inhibitors for the malarial enzyme (Scherf and Fuchter, unpublished).

In conclusion, the modulation of epigenetic factors by small molecular weight chemical compounds presents as an attractive research direction in antimalarial drug discovery. Moreover, given that particular histone states are linked to important biological processes of *P.*

*falciparum*, specific inhibitors could be potentially used to explore the fascinating biology of this parasite.

### **3.8 Emerging roles of HKMT-mediated lysine methylation on histone and non-histone proteins**

Although reported as early as 1964, the origin and the function of histone lysine methylation remained a mystery for a long time (Allfrey et al., 1964). It was the epigenetic field that rekindled the interest in histone lysine methylation, following the observation that H3K9me3 led to the recruitment of HP1 to chromatin and consequently promoted heterochromatin formation and spreading (Lachner et al., 2001). From then on, more and more studies have been focusing on the widespread modification of histone proteins by methylation and its effects on gene expression. It is believed that the high versatility and density of histone methylation modifications have evolved to sustain certain as yet unknown functions that may be essential for life.

Increasing evidence has demonstrated that lysine methylation also occurs on various non-histone proteins in the nucleus, especially on transcription and chromatin regulators (Zhang et al., 2015). For example, the HKMT SETD7 methylates both transcription factors and epigenetic regulators, which can lead to either gene activation or repression. SETD7-mediated methylation of the tumor suppressor protein p53 and estrogen receptor  $\alpha$  (ER $\alpha$ ) stabilizes these proteins and is required for their activation as transcription activators (Chuikov et al., 2004; Subramanian et al., 2008). In contrast, SETD7-mediated methylation of another tumor suppressor retinoblastoma protein (Rb) at K873 promotes its interaction with the heterochromatin protein HP1, in turn impeding the binding of cyclin-dependent kinase (Munro et al., 2010). Moreover, SETD7 methylates the histone-modifying enzyme SUV39H1 at residues K105 and K123, and which inhibits the H3K9 methyltransferase activity of SUV39H1, leading to heterochromatin relaxation and genome instability (Wang et al., 2013).

Notably, the enzymes responsible for histone lysine methylation also target non-histone substrates in the cytoplasm (Hamamoto et al., 2015). Take the largely cytoplasmic KMT SMYD3 for instance. SMYD3 is overexpressed in numerous human tumors, and is regarded as a transcriptional potentiator of multiple cancer-promoting genes. Recently Mazur et al. identified MAP3K2 kinase as a cytoplasmic target of SMYD3, and moreover, they elucidated a key role for SMYD3-mediated lysine methylation in integrating cytoplasmic kinase-

signalling cascades and driving tumour cell proliferation (Mazur et al., 2014). In addition to the above evidence that a KMT regulates the function of a cytoplasmic protein, lysine methylation and other modifications on non-histone proteins constitute a more general protein PTM code, which is widely involved in the regulation of protein-protein interaction, protein stability, and in some cases, protein localization (Zhang et al., 2012).

Thus far, the proteome-wide studies in *P. falciparum* have focus on histone modification and its association with specific cellular functions (Saraf et al., 2016). Future efforts involving the high-throughput identification of protein modifications and better understanding of the writers, erasers and readers of non-histone methylation and its regulatory networks, will provide novel functional roles of lysine methylation in *P. falciparum*.



## 4. Scope of the thesis

Epigenetic control via reversible histone methylation regulates transcriptional activation throughout the malaria parasite genome, controls the repression of multi-copy virulence gene families, determines sexual stage commitment, and most likely regulates the development of quiescent liver stage parasites. Six of the ten predicted PfHKMTs belonging to the SET domain superfamily appear to be essential for parasite asexual blood stage development. However, except for PfSET10 and PfSETvs, their biological function remains completely unknown. Our laboratory is focusing on these 5 HKMTs (PfSET1, PfSET3, PfSET6, PfSET7 and PfSET9), and has expressed and purified several enzymatically active, recombinant versions of these methyltransferases. Nevertheless, a major part of the work presented in this thesis is centered around the biological role of PfSET7 and PfSET6, both of which are refractory to classical genetic disruption in blood stages.

The main objectives of this thesis work are to use molecular and cellular biology, biochemistry, and reverse genetics tools to:

- Characterize the cellular localization of PfSET7 and PfSET6 during blood stage development (both asexual and sexual);
- Perform inducible knockout/knockdown of PfSET7 and PfSET6 and to characterize its effect on parasite development and gene expression, and thereby validate these SET proteins as novel targets for inhibitor screening;
- Identify the genome-wide chromatin occupancy and/or protein interacting partners of PfSET7 and PfSET6;

A minor objective is to use a CRISPR/Cas9-based genome editing approach to create a parasite line that expresses different individual *var* genes with distinct fluorescent tags. The transgenic parasites will serve as a reporter strain to screen hit compounds by flow cytometry and fluorescence microscopy, thus identifying compounds that mediate derepression of *var* genes, by likely targeting HKMTs enzymes or the epigenetic reader HP1. Furthermore, this strategy aims to generate an unbiased approach to screen for parasites that simultaneously express more than one *var* genes. Genome sequencing may identify novel factors contributing to *var* monoallelic expression.

# RESULTS





## 5.1 Article I

*Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites

Chen PB and Ding S *et al*, Scientific Reports, 2016

### Manuscript Highlight:


- We report the first large-scale production of an active, full-length recombinant PfHKMT using a baculovirus-based insect cell expression system.
- We performed detailed *in vitro* enzymatic characterization using SAM-dependent methyltransferase assays and concluded that PfSET7 methylates H3K4 and H3K9, targeting the latter particularly in the presence of a pre-existing H3K14Ac mark.
- We observed that PfSET7 localizes to distinct cytoplasmic foci adjacent to the nucleus in erythrocytic and liver stage parasites, and throughout the cytoplasm in salivary gland sporozoites.
- Taken together, our data suggest that PfSET7 as being a cytosolic protein methyltransferase.

### Individual Contribution:

- I generated transgenic parasite lines for visualizing cellular localization of PfSET7 and analyzing growth and gene expression phenotypes upon PfSET7 knock-down
- I performed western blotting and immunofluorescence microscopy to elucidate the localization of PfSET7 in blood stage parasites.
- I participated in the manuscript writing.



# SCIENTIFIC REPORTS



OPEN

## *Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites

Received: 09 December 2015

Accepted: 01 February 2016

Published: 23 February 2016

Patty B. Chen<sup>1,2,3,\*</sup>, Shuai Ding<sup>1,2,3,\*</sup>, Gigliola Zanghi<sup>4</sup>, Valérie Soulard<sup>4</sup>, Peter A. DiMaggio<sup>6</sup>, Matthew J. Fuchter<sup>7</sup>, Salah Mecheri<sup>1,2,3</sup>, Dominique Mazier<sup>4,5</sup>, Artur Scherf<sup>1,2,3</sup> & Nicholas A. Malmquist<sup>1,2,3</sup>

Epigenetic control via reversible histone methylation regulates transcriptional activation throughout the malaria parasite genome, controls the repression of multi-copy virulence gene families and determines sexual stage commitment. *Plasmodium falciparum* encodes ten predicted SET domain-containing protein methyltransferases, six of which have been shown to be refractory to knock-out in blood stage parasites. We have expressed and purified the first recombinant malaria methyltransferase in sufficient quantities to perform a full enzymatic characterization and reveal the ill-defined PfSET7 is an AdoMet-dependent histone H3 lysine methyltransferase with highest activity towards lysines 4 and 9. Steady-state kinetics of the PfSET7 enzyme are similar to previously characterized histone methyltransferase enzymes from other organisms, however, PfSET7 displays specific protein substrate preference towards nucleosomes with pre-existing histone H3 lysine 14 acetylation. Interestingly, PfSET7 localizes to distinct cytoplasmic foci adjacent to the nucleus in erythrocytic and liver stage parasites, and throughout the cytoplasm in salivary gland sporozoites. Characterized recombinant PfSET7 now allows for target based inhibitor discovery. Specific PfSET7 inhibitors can aid in further investigating the biological role of this specific methyltransferase in transmission, hepatic and blood stage parasites, and may ultimately lead to the development of suitable antimalarial drug candidates against this novel class of essential parasite enzymes.

While global mortality due to malaria has decreased since the beginning of this century, this parasitic disease continues to claim approximately 0.6 million lives per year, particularly in the vulnerable populations of children under five years of age and pregnant women<sup>1</sup>. Malaria eradication efforts have been hindered by the incredible ability of the parasite to develop resistance to existing antimalarials, prompting the search for novel essential factors to serve as potential drug targets. The complex life cycle of human malaria parasites involves an insect vector phase, a liver stage at the onset of infection, an asexual blood stage responsible for disease pathogenesis

<sup>1</sup>Unité Biologie des Interactions Hôte-Parasite, Département de Parasites et Insectes Vecteurs, Institut Pasteur, Paris 75015, France. <sup>2</sup>CNRS, ERL 9195, Paris 75015, France. <sup>3</sup>INSERM, UMR 1201, Paris 75015, France. <sup>4</sup>Sorbonne Universités, UPMC Univ Paris 06, INSERM U1135, CNRS ERL 8255, Centre d'Immunologie et des Maladies Infectieuses (CIMI-Paris), 91 Bd de l'hôpital, 75013, Paris, France. <sup>5</sup>AP HP, Centre Hospitalo-Universitaire Pitié-Salpêtrière, 75013 Paris, France. <sup>6</sup>Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom. <sup>7</sup>Department of Chemistry, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to N.A.M. (email: nicholas.malmquist@pasteur.fr)

and a sexual stage permitting disease transmission. Transition through the various stages of the complex parasite life cycle is a highly controlled process regulated at the level of transcriptional gene activation. Indeed, in the experimentally tractable asexual blood stage, a clear stage-specific transcriptional cascade as been observed<sup>2</sup>. Aside from the recent discovery of a plant-derived family of transcription factors<sup>3</sup>, no canonical gene regulatory elements have been identified in *P. falciparum*. However, a variety of distinct epigenetic mechanisms have been discovered which contribute to the transcriptional control of single copy genes as well as the orchestration of clonally variant virulence gene families<sup>4–8</sup>.

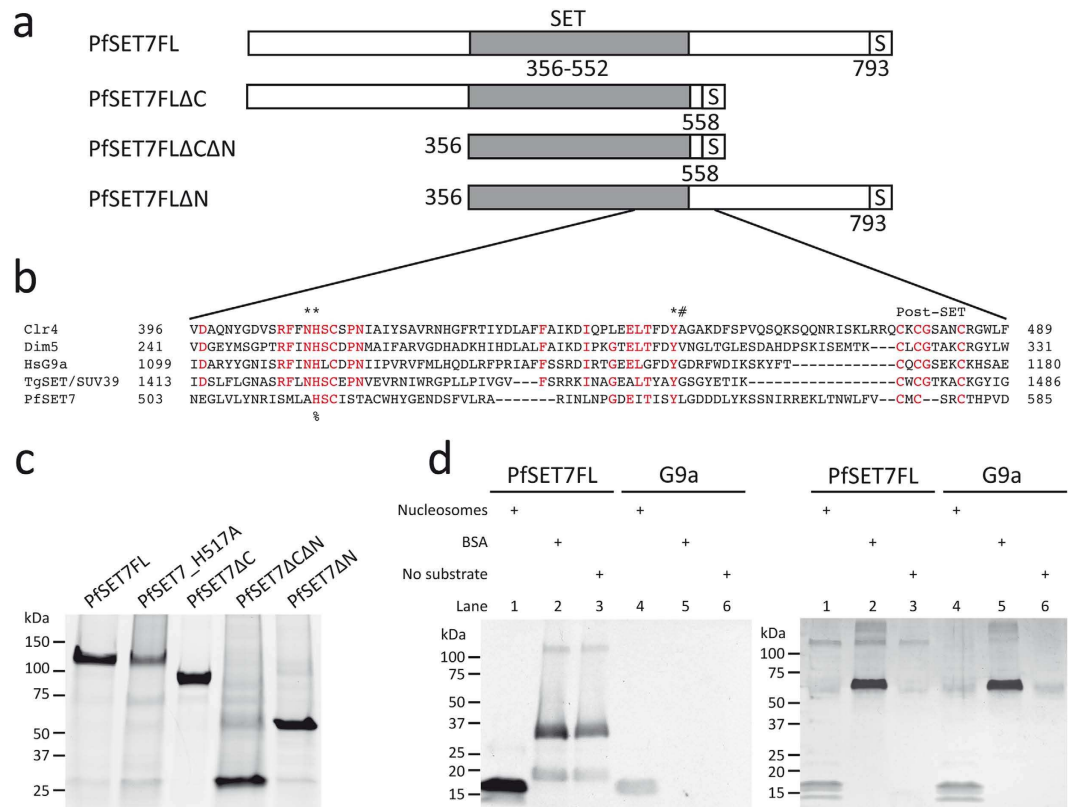
Histone post-translational modifications (PTMs) play a particularly important role in the developmental progression of blood stage malaria parasites. The role of histone PTMs was first demonstrated in the monoallelic expression of a virulence gene family known as the *var* genes, which are involved in antigenic variation and pathogenesis<sup>4</sup>. Subsequent studies have shown that transcriptional activation and silencing of virtually all genes is associated with histone methylation or acetylation<sup>5,6</sup>. The tri-methylation of histone H3 lysine 4 (H3K4me3) and acetylation of histone H3 lysine 9 (H3K9ac), associated with transcriptional activation in the conserved histone code, is indeed associated with actively transcribed genes, including the single expressed *var* gene, in *P. falciparum*<sup>5,6</sup>. Tri-methylation of histone H3 lysine 9 (H3K9me3), a repressive mark in the conserved histone code, stands out in malaria parasites for its association with the variegated gene expression of clonally variant gene families and genes encoding a parasite-induced erythrocyte permeation pathway and with regulating the commitment to transmission stage parasites<sup>5,6,9,10</sup>. Importantly, H3K9me3 is not associated with general transcriptional repression across the genome<sup>5,6</sup>. Tri-methylation of histone H3 lysine 36 (H3K36), an activation mark in the conserved histone code, plays a dual function in *P. falciparum*, as H3K36me3 is associated with transcriptionally active genes throughout the genome but also is associated with silenced members of *var* genes and other clonally variant gene families<sup>11</sup>. Additional methylation PTMs exist on *P. falciparum* histones, some of which are conserved throughout eukaryotes while others are unique to *Plasmodium*<sup>12</sup>. These methylation marks exist in combination with other PTMs such as acetylation and phosphorylation<sup>13</sup>, but their functions, either singly or in combination, remain to be elucidated.

Protein methyltransferase enzymes (PMTs) catalyze the mono-, di- or tri-methylation of lysine residues (PKMTs) or the mono- or di-methylation of arginine residues (PRMTs). PKMTs and PRMTs were initially thought to be specific for the methylation of a single protein substrate, and while some members of this class of enzyme do exhibit a high degree of protein substrate specificity, it is widely recognized that certain individual PMTs are able to methylate multiple protein substrates, including both histones in the nucleus and non-histone proteins in the cytosol<sup>14</sup>. Since both PKMTs and PRMTs have been associated with a variety of diseases including cancer, neurodegenerative and inflammatory diseases, PMT enzymes have emerged as a target class for drug discovery against human disease<sup>15</sup>.

Recently, we have expanded our fundamental research on the role of histone PTMs in malaria parasite gene regulation to include translational research into discovering PKMT inhibitors for both the further study of parasite biology using a chemical biology approach and for the potential development of novel classes of future antimalarials. In the absence of recombinant target parasite enzymes we initiated an approach based on targeting malaria parasite histone methylation using a known inhibitor of a human PKMT, BIX-01294<sup>16–18</sup>. These studies established histone methylation to be a viable target for antimalarial drug discovery, as BIX-01294 and its derivatives cause rapid and specific parasite death with the concomitant reduction in parasite histone methylation levels<sup>16</sup>. These inhibitor-based results have motivated increased efforts into recombinant PfPKMT enzyme production to both fully characterize these enzymes and to enable a target-based approach to discovering specific malaria parasite enzyme inhibitors.

Computational analysis predicts ten PKMTs and three PRMTs in the *P. falciparum* genome. All ten identified PfPKMT genes contain a catalytic methyltransferase SET domain, named after the *Drosophila* chromatin-modifying enzymes *Su(var)3–9*, *Enhancer of zeste*, and *Trithorax*<sup>19</sup>. SET domain containing proteins were initially studied in the context of specifically methylating lysine residues of histones, but subsequent work has revealed numerous non-histone protein substrates for SET domain proteins<sup>20</sup>. Knock-out studies have indicated a subset of PfPKMTs are essential in blood stage parasites<sup>11</sup>. Despite significant efforts, the molecular characterization of PfPKMT enzyme activity has been constrained by the inability to produce sufficient quantities of recombinant proteins. Cui *et al.* were able to express four PfPKMTs (PfSET1, PfSET2, PfSET3, PfSET8) using a wheat germ expression system, but only assign histone H4 lysine 20 (H4K20) methyltransferase activity to PfSET8 and histone H3 methyltransferase activity to PfSET2 through Western blots analysis and autoradiography of enzyme reactions containing nucleosomes as protein substrates<sup>21</sup>. PfSET2, since re-named PfSETvs, was confirmed to have H3K36 methyltransferase activity through the observed reduction of this mark along *var* genes in PfSETvs knock-out parasites<sup>11</sup>. Volz, *et al.* were able to detect low-level H3K4 methyltransferase activity for an affinity-tagged version of endogenous PfSET10 immunoprecipitated from transgenic parasites<sup>22</sup>. To date, no isolated PfPKMT enzymes have undergone biochemical characterization to determine substrate specificity and enzyme kinetic parameters.

In this report we describe the first large-scale production and enzymatic characterization of a recombinant *P. falciparum* PKMT, the poorly understood PfSET7, purified from a baculovirus expression system. Recombinant PfSET7 displays comparable kinetics to other characterized PKMTs from human and mouse with regards to enzyme turnover and AdoMet methyl-donor utilization. Nucleosome labeling experiments reveal that PfSET7 extensively methylates H3K4 and H3K9, but modifies the latter particularly in the presence of pre-existing acetylated histone H3 lysine 14 (H3K14ac). Immunofluorescence imaging of blood stage parasites, however, reveals PfSET7 to localize to distinct foci outside of the parasite nucleus. PfSET7 was also identified by immunofluorescence to be present in motile salivary gland sporozoite stage parasites and in liver stage parasite forms. Since PfSET7 was refractory to genetic knock-out in blood-stage parasites<sup>11</sup>, this enzyme is presumed to be essential in at least that stage, and therefore represents a promising antimalarial drug target candidate. With a



**Figure 1. Protein production and purification.** (a) Schematic of the PfSET7 full-length protein and truncated proteins, all expressed as recombinant affinity-tagged proteins. SET, SET domain; S, 2 × strep-tag. (b) Alignment of the C-terminal part of the SET domain and post-SET domains of H3K9 methyltransferases from *Schizosaccharomyces pombe* Clr4 (O60016), *Neurospora crassa* Dim5 (Q8X225), *Homo sapiens* HsG9a (Q96KQ7), *Toxoplasma gondii* SET/SUV39 (TGME49\_255970), *Plasmodium falciparum* PfSET7 (PF3D7\_1115200). Conserved residues are highlighted in red. (\*) Residues involved in catalysis. (#) Final amino acid of the SET domain. (%) Indicates residue mutated in the catalytic mutant PfSET7\_H517A. Post-SET, post-SET domain. (c) SDS-PAGE of recombinant PfSET7FL, the catalytic mutant PfSET7\_H517A, PfSET7ΔC, PfSET7ΔCΔN and PfSET7ΔN. (d) Autoradiograph (left panel) of enzyme reactions with PfSET7FL (lanes 1, 2 and 3) and MmG9a (lanes 4, 5 and 6) with nucleosomes, BSA and enzyme alone. Film was exposed 72h. Duplicate reactions were resolved by SDS-PAGE and silver stained (right panel).

biochemically characterized enzyme amenable to large-scale expression and purification, target based inhibitor discovery can proceed. Specific PfSET7 inhibitors will be useful tools to further explore the biological function of PfSET7 and have potential to be developed into a novel class of future antimalarials.

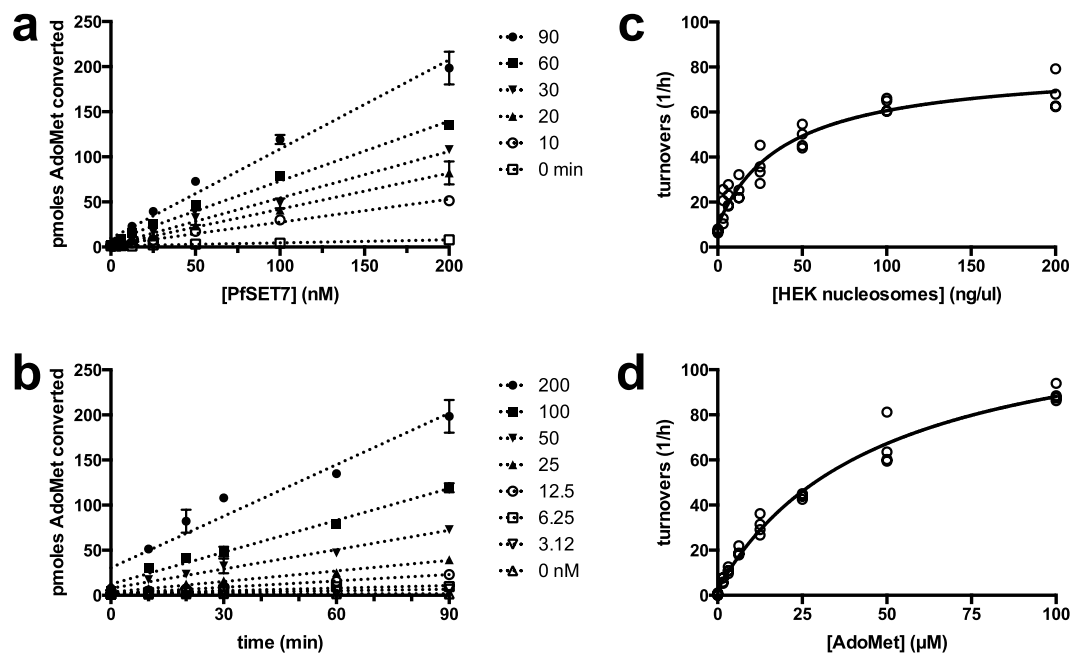
## Results

**Expression and purification of recombinant wild-type and mutant PfSET7.** The 793 amino acid PfSET7 protein contains a central catalytic SET domain followed by a short post-SET domain, which is flanked by a 355 amino acid N-terminal arm and 240 amino acid C-terminal arm (Fig. 1a). The SET domain is identifiable due to its homology with PKMTs from other species<sup>21</sup>. The C-terminal region of the PfSET7 SET domain (Fig. 1b), which contains the catalytic residues, shows higher homology to other organisms compared to the N-terminal region<sup>21</sup>. As with most of the *Plasmodium* genome, the PfSET7 gene contains a high percentage (72%) of AT nucleotides. To reduce the high AT content for heterologous recombinant protein expression we obtained a codon optimized version of wild-type full-length PfSET7 (PfSET7FL). In anticipation of potential recombinant protein production in secretory systems, two potential N-glycosylation sites were mutated. Recombinant protein expression attempts in mammalian cells and wheat germ were unsuccessful. Ultimately, an active recombinant PfSET7 enzyme fused to a C-terminal 2x-strep tag was successfully produced in a baculovirus-based Sf9 insect cell expression system.

To study the impact of the N- and C-termini on methyltransferase activity, truncated mutants of PfSET7 (Fig. 1a) were made by removing the N-terminus or the C-terminus or both termini, leaving only the SET domain. Finally, to confirm that our purified methyltransferase activity is specific to PfSET7, a catalytically inactive mutant, PfSET7\_H517A, was produced by the mutation of a single conserved catalytic residue.

Purified PfSET7FL and PfSET7\_H517A are expected to migrate at 98 kDa, though a single protein band is consistently observed at approximately 120 kDa (Fig. 1c). PfSET7ΔC is expected at 70 kDa but migrates at





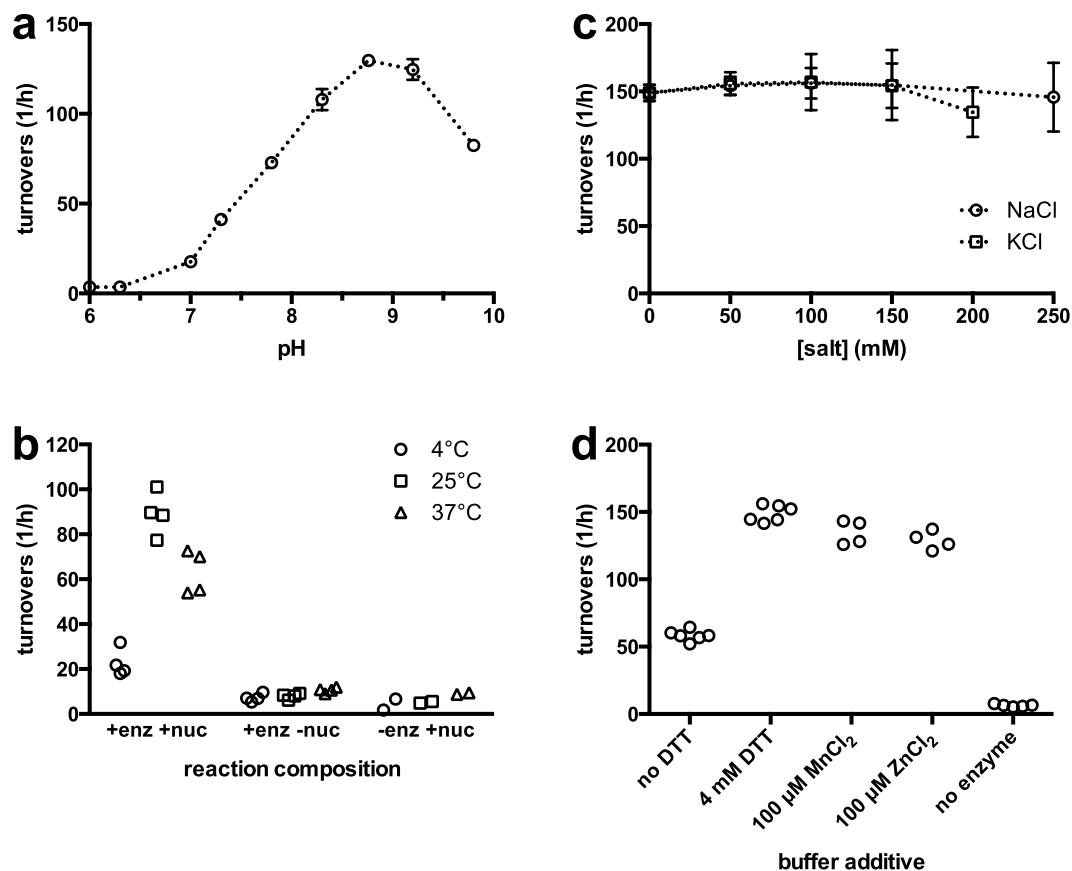
**Figure 2. Steady state enzyme kinetics of PfSET7FL.** (a) Enzyme concentration and (b) time dependent enzyme activity. (c) Saturation kinetics for HEK nucleosomes yield a  $K_m$  value of  $40 \pm 9$  ng/ul. (d) Saturation kinetics for methyl-donor AdoMet yield a  $K_m$  value of  $48 \pm 7$   $\mu$ M. Data in (a,b) are mean  $\pm$  SD of one representative experiment. Data in (c,d) are replicate values from duplicate experiments and were fitted to the Michaelis-Menton equation using GraphPad Prism software.

85 kDa. PfSET7 $\Delta$ C $\Delta$ N and PfSET7 $\Delta$ N are both expressed at an expected size of 27 and 55 kDa, respectively. Recombinant PfSET7 proteins nonetheless were purified to homogeneity.

**Steady state enzyme kinetics.** To investigate the methyltransferase activity of PfSET7FL, the purified enzyme was tested in a PKMT assay using methyl- $^3$ H-AdoMet as the methyl donor. Mouse G9a, a known histone H3K9 PKMT was used as a positive control. Both enzymes were tested for methyltransferase activity in the presence of nucleosomes as a protein substrate, with BSA as a general protein substrate, or without a protein substrate. Enzyme reaction contents were migrated on SDS-PAGE and exposed to film (Fig. 1d left panel). A second gel loaded with identical reactions using unlabelled AdoMet was silver stained to show total protein content (Fig. 1d right panel). Methyltransferase activity of PfSET7FL was confirmed by the presence of bands on the autoradiograph. In reactions with PfSET7FL and G9a with nucleosomes (Lanes 1 & 4), a band at 17 kDa corresponding to histone H3 is apparent. PfSET7FL produces a second band below H3, corresponding to histone H4. There was no methyl transfer to BSA by either enzyme (lanes 2 & 5). However, in the absence of nucleosomes or in the presence of BSA, PfSET7FL transfers methyl groups to proteins that migrate at 19 kDa, 32 kDa and 125 kDa (Lanes 2 & 3). The band at 125 kDa co-migrates with PfSET7FL and could represent automethylation of the enzyme occurring in the absence of nucleosomes, as these bands are absent when nucleosomes are present<sup>23</sup>. The two lower bands could be methylation of degradation fragments of the full-length enzyme, as there are no other common proteins added to these reactions. These results suggest PfSET7FL is primarily a histone H3 methyltransferase, with possible additional activity toward histone H4.

To characterize the enzymatic properties of PfSET7FL, time and enzyme concentration dependent activity of PfSET7FL was examined using radiometric assays (Fig. 2a,b). Enzyme activity was linear from 12.5–200 nM for up to 90 minutes of reaction time. Subsequent experiments therefore used 25 nM enzyme in 60 minutes reactions to remain in this linear range. Nucleosome-dependent activity was tested at nucleosome concentrations of 0–200 ng/ul, revealing a  $K_m$   $40 \pm 9$  ng/ul (Fig. 2c). AdoMet-dependent activity was examined at AdoMet concentrations of 0–100  $\mu$ M, yielding a  $K_m$  for AdoMet of  $48 \pm 7$   $\mu$ M (Fig. 2d). Subsequent experiments used nucleosome and AdoMet at saturating concentrations of 200 ng/ul and 100  $\mu$ M, respectively. PfSET7FL revealed a turnover of between 80 to 150  $h^{-1}$ , which is comparable to other characterized PKMTs such as mouse G9a at 88  $h^{-1}$ <sup>24</sup> and Dim5 at 180  $h^{-1}$ <sup>25</sup>.

Experiments to test pH dependence showed PfSET7FL to display peak activity near pH 8.8 (Fig. 3a), with activity decreasing above and below this pH value. Performing enzyme reactions at three different temperatures of 4  $^{\circ}$ C, 25  $^{\circ}$ C and 37  $^{\circ}$ C showed the highest activity at 25  $^{\circ}$ C (Fig. 3b). The presence of monovalent salts has been reported to be detrimental to PKMT activity for certain recombinant enzymes. However, up to 250 mM NaCl or 200 mM KCl (Fig. 3c) does not appear to affect PfSET7FL enzyme activity. Various PKMT characterization buffers reported in the literature contain divalent magnesium cations. To investigate any role for magnesium on PfSET7FL,  $MgCl_2$  was added at 100  $\mu$ M but did not appear to alter PfSET7FL enzyme activity. SET domain



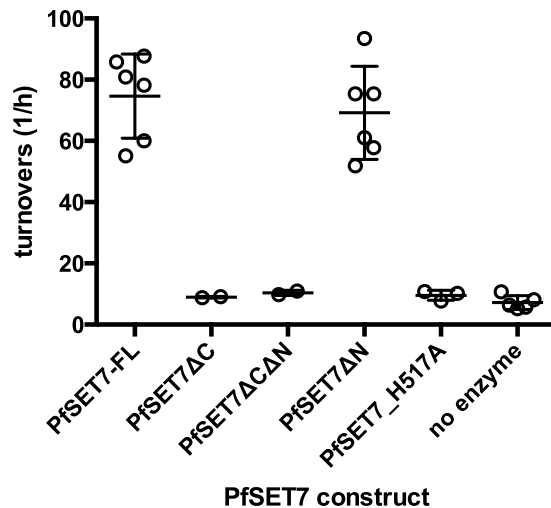
**Figure 3. Characterization of PfSET7FL enzyme reaction conditions.** (a) PKMT reactions were performed to determine pH dependent enzyme activity on a pH range of 6–10. (b) Temperature dependent enzyme activity where reactions were carried out at 4°C (circles), 25°C (squares), or 37°C (triangles). Control reactions without nucleosomes or without enzyme were performed in parallel. (c) The effect of two common salts NaCl (circles) and KCl (squares) were tested up to 250 mM or 200 mM, respectively. (d) The effect of removing DTT or adding divalent metal ions on enzyme activity. All reactions, except where noted, contain 4 mM DTT. Reactions contained 100 μM MnCl<sub>2</sub> or ZnCl<sub>2</sub> where noted. Data from at least two independent experiments are represented in (a,c) as mean ± SD. Individual data from two independent experiments are presented in (b,d).

proteins are known to contain zinc binding sites<sup>26</sup>, thus the effect of zinc ions on PfSET7FL enzyme activity was tested. The data show supplementing reactions with 100 μM zinc has little effect on apparent enzyme activity. Finally, the presence of the reducing agent dithiothreitol (DTT) was tested, and the data reveal 4 mM DTT results in increased apparent PfSET7 enzyme catalysis relative to no reducing agent present (Fig. 3d). These data define the optimal assay conditions for PfSET7FL for the present analyses for any future *in vitro* PfSET7FL studies, including further enzyme characterization or subsequent inhibitor discovery.

**PfSET7 mutational analysis.** To identify functionally important regions of PfSET7, truncation mutants were examined for catalytic activity (Fig. 4). PfSET7ΔN demonstrated PKMT activity similar to that of the full-length enzyme, indicating that the N-terminus of PfSET7 is not essential for enzyme activity.

While some PKMT enzymes contain only a catalytic SET domain<sup>27</sup>, others also possess a cysteine rich post-SET domain that is implicated in protein folding and formation of the active site<sup>28</sup>. In proteins containing a post-SET domain, this domain has been shown to be necessary for enzyme activity<sup>29</sup>. PfSET7 contains a cysteine rich putative post-SET domain similar to post-SET domains from other PKMTs, though the canonical CxCxxx motif as seen in Dim5 and G9a is a CxCxxC motif in PfSET7 (Fig. 1b). To test if the putative post-SET domain of PfSET7 is essential for enzyme activity, two mutants lacking the post-SET domain, PfSET7ΔC and PfSET7ΔCΔN, were produced. These mutants display no enzyme activity, indicating that the putative post-SET domain of PfSET7 may indeed be necessary for catalysis as seen in other PKMT enzymes containing post-SET domains.

Wild-type PfSET7 has two conserved catalytic residues in the SET domain (Fig. 1b), H517 and Y551<sup>21</sup>. A third catalytic residue, N516, is present in PKMTs such as Clr4, Dim5 and HsG9a, but absent in PfSET7, which contains an arginine at this location. Mutation of the conserved H517 has been shown to abolish PKMT activity<sup>30,31</sup>. Accordingly, the catalytic mutant PfSET7\_H517A shows no HKMT activity compared to PfSET7FL (Fig. 4). This confirms that the methyltransferase activity we observe is specific to PfSET7 and that mutation of a single conserved residue is sufficient to abolish the activity of this purified recombinant enzyme.



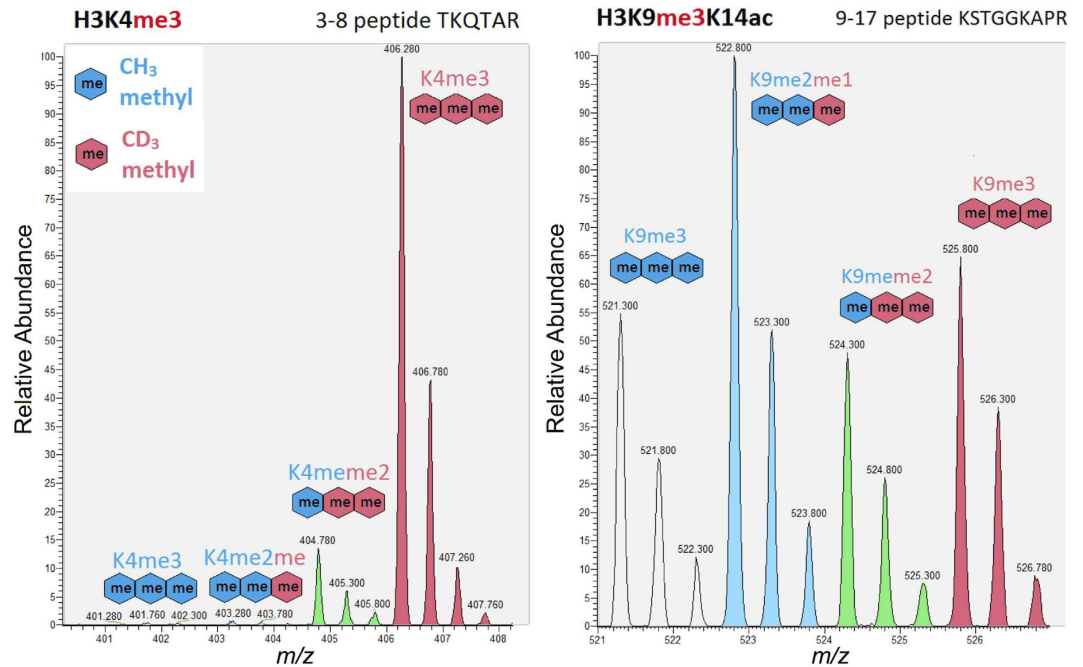
**Figure 4. PfSET7 mutational analysis.** Enzyme activity of full-length, truncated and catalytic dead versions of PfSET7. Data from two independent experiments are represented as mean  $\pm$  SD.

**Mass spectrometry identification of PfSET7 histone lysine methylation targets.** To identify the target amino acid specificity of PfSET7, we employed deuterium-labeled AdoMet (CD3-AdoMet) and isolated human nucleosomes in a mass spectrometry based enzyme assay. These experiments have the advantage of providing a protein substrate containing a range of pre-existing histone PTMs and can identify an enzyme preference for specific target residues and any influence of modifications to neighboring residues. Recombinant PfSET7 enzyme (1  $\mu$ M) was incubated with human nucleosomes (0.1 mg/mL, isolated from HEK293 nuclei) and CD3-AdoMet (1 mM) for 3 hours. Histones were prepared for mass spectrometry analysis (as described in the Methods section) to identify the lysine targets that were methylated by PfSET7, which are distinguishable from existing methylations by a CD3 mass shift (+3 Daltons) in the  $m/z$  of the modified peptide. A mass shift of 3 Daltons is necessary to sufficiently distinguish new methyl groups added by the recombinant HKMT from previously existing histone methylation, as naturally occurring isotopes (e.g.  $^{13}\text{C}$ ,  $^{15}\text{N}$ , etc) and their combinations result in abundant isotopic peaks that are +1 and +2 Daltons after the monoisotopic peak (i.e. the mass based on the most abundant isotopes:  $^{12}\text{C}$ ,  $^{14}\text{N}$ ,  $^{16}\text{O}$ , etc). The same *in vitro* labeling experiment was performed using recombinant mouse G9a as a reference control as it is known to methylate H3K9. For instance, the right panels of Figures S1 and S2 reveal significant levels of mouse G9a methylation for H3K9me3K14ac and H3K9me2K14un, respectively, where the isotopic peak cluster for every methylation event has been colour-coded based on the total number of CD3 methyl groups (blue, green and red isotopic peaks represent the signal for 1, 2 and 3 CD3 methyl groups, respectively).

Extensive labeling of H3K4 di- and tri-methylation (H3K4me2, H3K4me3) was observed for PfSET7 (left panel of Fig. 5 for H3K4me3; H3K4me2 in Fig. S1 and corresponding tandem mass spectra in Figures S3–S7), but not for mouse G9a. Interestingly, PfSET7 appears to modify H3K4 to its highest methyl occupancy, as observed by the peak intensities in the mass spectrum (in the left panel of Fig. 5, the highest intensity peak corresponds to H3K4me3 with three new CD3 methyl groups; similar trend observed in Fig. S1 for H3K4me2). PfSET7 was also observed to methylate H3K9 in the presence of existing K14 acetylation on nucleosome substrates (right panel of Fig. 5 for H3K9me3K14ac; corresponding tandem mass spectra in Figures S8–S13) to the same extent as mouse G9a (compare Figs. 5 to S1). However, whereas mouse G9a exhibited a similar degree of H3K9 methylation activity for unmodified H3K14 substrates (right panels of Figs S1 and S2), the extent of PfSET7-mediated K9 methylation was orders of magnitude lower in the absence of H3K14 acetylation, as shown in Fig. S2 for H3K9me2K14un (see also tandem MS in Figures S14–S17). This implies PfSET7 preferentially recognizes nucleosome substrates containing acetylation on H3K14.

To a lesser extent PfSET7 was also observed to methylate H3K36, particularly when H3K27 is unmodified (see Figures S18 and S19 for H3K36me1 and H3K36me3, respectively). A tandem mass spectrum was identified to support evidence of PfSET7 methylation of H3K27me2 substrates to H3K27me3 (Fig. S20). However, it should be noted that methylation activity for H3K27me2/me3 is also observed for mouse G9a, which can occur under *in vitro* conditions since the amino acid sequences around H3K9 and H3K27 are homologous (ARKST and ARKSA, respectively). Methylation of other common lysine targets (i.e. H3K18, H3K23, H3K79 and H4K20) was not observed for PfSET7.

**Cellular localization of PfSET7 in blood stage parasites.** To identify the cellular location of PfSET7 in parasites, a Pf3D7-SET7-HA-glmS integrant transfected line was generated through markerless insertion at the PfSET7 genomic locus of a  $3 \times$  HA affinity tag and the glmS ribozyme using the CRISPR/Cas9 genome editing tool. Correct genomic integration of this transgenic parasite line was confirmed by PCR and endogenous tagged protein expression was confirmed by Western Blot (Fig. S21). This glmS construct was designed to produce an inducible knock-down parasite line, but was found to be inefficient at PfSET7 mRNA and protein reduction in the presence of glucosamine inducer, and induced parasites displayed no obvious growth phenotype (Fig. S22).

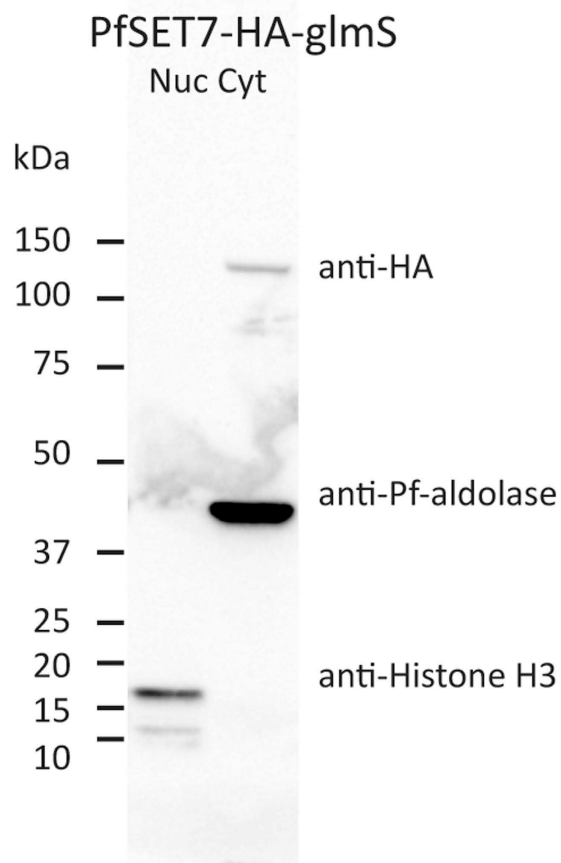


**Figure 5. PfSET7 extensively methylates H3K4 and H3K9 in the presence of H3K14ac.** Recombinant PfSET7 enzyme was incubated with human nucleosomes and CD3-labeled S-adenosyl-methionine to identify the histone lysine targets of PfSET7, which are distinguishable from existing methylations (blue colored ‘me’ symbols) by a CD3 mass shift (+3 Daltons; red colored ‘me’ symbols) in the  $m/z$  of the modified peptide. Significant labeling of H3K4me2 and H3K4me3 by PfSET7 was observed (shown for H3K4me3 in the left panel), where the major product is the fully methylated H3K4 substrate. Extensive labeling of H3K9me2 and H3K9me3 by PfSET7 was also observed, but particularly in the presence of existing H3K14 acetylation on nucleosome substrates (shown for H3K9me3K14ac in the right panel), implying this methyltransferase exhibits some degree of nucleosome specificity (compare extent of labeling of H3K9me2K14un substrate by PfSET7 and mG9a in Supplementary Fig. S2). Annotated tandem mass spectra for all observed labeled histone peptides are provided in the Supplementary Figures.

To determine the subcellular distribution of PfSET7, whole cell extracts of the integrant Pf3D7-SET7-HA-glmS line were biochemically separated into nuclear and cytoplasmic fractions, which were examined by Western Blot for PfSET7-HA via the HA-tag and for aldolase or histone H3 for cytoplasmic or nuclear protein controls, respectively. The results show PfSET7 is present in the cytoplasmic fraction but not visible in the nuclear fraction (Fig. 6). These results suggest PfSET7 may have a role outside of the nucleus, which is unexpected for an apparent histone methyltransferase.

**PfSET7 localizes to a distinct cytoplasmic compartment in blood stages.** To assess the intracellular localization of PfSET7, *P. falciparum* parasites were analyzed by immunofluorescence microscopy (IF) using monoclonal anti-HA or polyclonal anti-PfSET7 antibodies in PfSET7-HA-glmS or wild-type parasites, respectively. Staining patterns using either antibody are virtually identical, suggesting both antibodies are equivalent in detecting PfSET7, and that affinity tagging the endogenous protein does not lead to aberrant expression or localization. (Fig. S23). Through the three major asexual blood stages, PfSET7 signal exhibited a punctate pattern adjacent to the nucleus, with a concomitant increase in staining during progression through mature trophozoites and schizonts (Fig. S24). To determine whether PfSET7 is localized at the nuclear periphery or outside of the nucleus, PfSET7 staining was compared to that of PfHP1, which localizes to discrete foci within the nucleus but at the nuclear periphery<sup>32</sup>. Consistent with the biochemical fractionation experiments, PfSET7 staining is distinct from the nuclear periphery marker PfHP1 (Fig. 7a), further supporting PfSET7 cytoplasmic localization. The punctate fluorescent signal suggests PfSET7 could be localized to a cytoplasmic organelle rather than evenly distributed throughout the cytoplasm.

The discrete IF staining of PfSET7 led us to explore whether the protein localizes to other DNA-containing organelles in malaria parasites, namely the single parasite mitochondrion or the apicoplast, an organelle specific to the phylum apicomplexa. A previous study reported PfSET5 localizes to the mitochondria, but did not examine PfSET7<sup>33</sup>. To determine if PfSET7 is in the parasite mitochondrion, we combined IF staining for PfSET7 and MitoTracker Deep Red FM staining for the mitochondrion. The data show PfSET7 does not co-localize with mitochondrial staining and is therefore not located within this organelle (Fig. 7a). To determine if PfSET7 is in the apicoplast we used a parasite line expressing HA-tagged triosephosphate transporter (PfoTPT-HA), a polytopic membrane protein of the apicoplast<sup>34</sup>. The IF images show no signal overlap between PfSET7 and PfoTPT, indicating PfSET7 is not located in the parasite apicoplast (Fig. 7a). Other markers for organelles whose ontogeny begins in early stage were used to evaluate their proximity to PfSET7, including anti-Bip for the endoplasmic



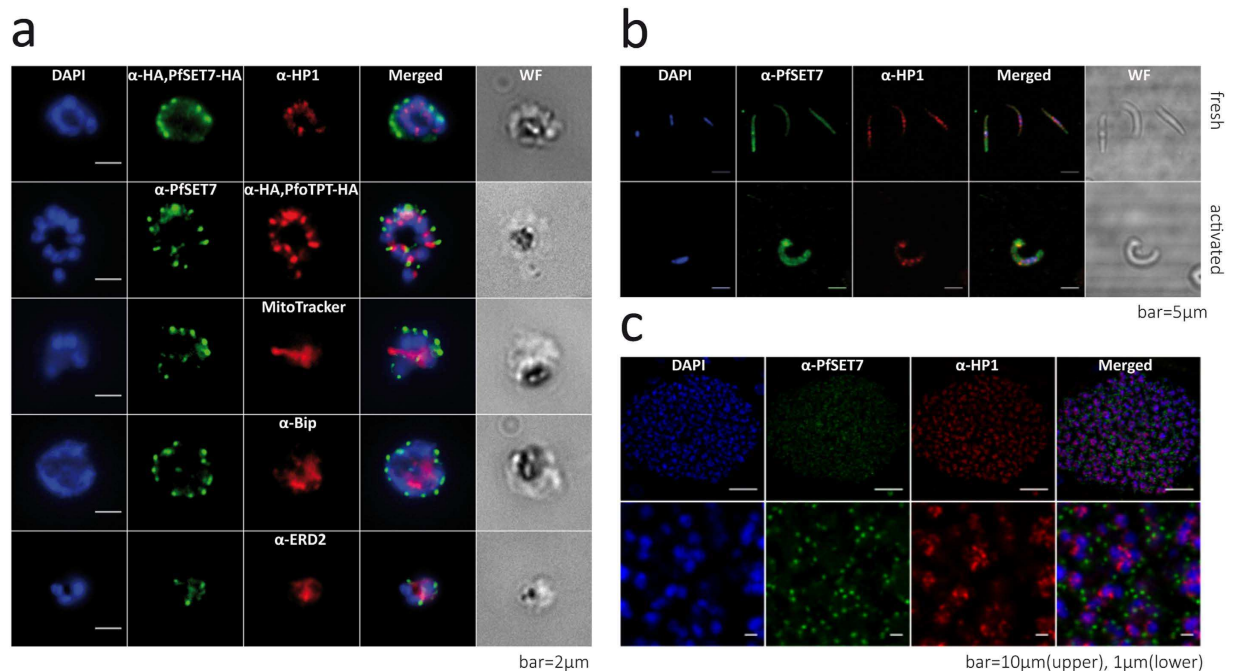
**Figure 6. PfSET7 expression and cytoplasmic localization.** Western blot of nuclear and cytoplasmic fractions of asynchronous parasites using anti-HA antibodies against PfSET7-HA-glmS. The two fractions were monitored for purity using control antibodies anti-Pf-aldolase (predicted MW: 29 kDa) for the cytoplasm and anti-histone H3 (predicted MW: 15 kDa) for the nucleus. Anti-HA recognized the PfSET7-HA band of >100 kDa (predicted MW: 98 kDa) in the cytoplasm.

reticulum and anti-EDR2 for the cis-golgi apparatus (Fig. 7a). Results from IF studies utilizing markers for these two organelles reveal that PfSET7 does not co-localize with either the endoplasmic reticulum or the cis-golgi. These combined IF results suggest PfSET7 is localized to a distinct compartment in the cytoplasm.

**PfSET7 expression in *P. falciparum* sporozoites and in liver stage schizonts.** To characterize PfSET7 expression in pre-erythrocytic parasite stages, immunofluorescence imaging was performed on *P. falciparum* mosquito salivary gland stage sporozoites and on hepatic forms. As motile sporozoites productively invade hepatocytes, these forms undergo a radical morphological remodeling, a process first discernible by a bulbous expansion that progressively enlarges to become the early spherical hepatic stage<sup>35</sup>. PfSET7 staining of freshly isolated sporozoites shows an uneven and clustered localization in the cytoplasm (Fig. 7b). Upon sporozoite transformation into early hepatic forms under axenic conditions, a similar cytoplasmic staining is observed with a more intense focal signal at the growing bulb (Fig. 7b). Nearly fully mature hepatic schizonts (just prior to merozoite egress) were imaged in liver sections from humanized mice seven days post-sporozoite inoculation (Fig. 7c). PfSET7 is found in these hepatic schizonts in a punctate pattern as distinct dots observed near the nucleus of each forming merozoite. This cytoplasmic and nucleus-associated fluorescence pattern is similar to that observed in asexual blood stage parasites.

## Discussion

Epigenetic gene regulation plays a critical role in malaria parasites in the control of general transcriptional regulation, monoallelic expression of virulence genes and in the commitment to transmission stage development. A major mechanism of epigenetic regulation in *P. falciparum* is mediated by histone post-translational modifications, which have been linked to transcriptional activation throughout the *Plasmodium* genome and to transcriptional repression of *P. falciparum* multi-copy gene families<sup>36</sup>. Building upon our fundamental research into parasite gene regulation, we have taken a chemical biology approach to target malaria methyltransferase enzymes in an effort to identify the biological role of individual PfPKMTs, ideally through the development of specific enzyme inhibitors. We have recently reported the identification and development of one PKMT inhibitor series entirely through phenotypic readouts<sup>16–18</sup>, but the inability to produce functional recombinant PfPKMTs has



**Figure 7. Immunofluorescence localization of PfSET7 at different stages of the *Plasmodium falciparum* life cycle.** (a) Asexual blood stage parasites. Markers used were: DAPI as a DNA marker, anti-HA in PfSET7-HA parasites, anti-PfSET7 in wild-type parasites, anti-HP1 as a nuclear periphery marker, anti-HA in PfoTPT-HA parasites as an apicoplast marker, MitoTracker Deep Red FM staining for the mitochondrion, anti-Bip as an ER marker and anti-ERD2 as a cis-golgi marker. The scale bars correspond to 2  $\mu$ m. (b) Mosquito salivary gland stage sporozoites. The PfSET7 is expressed in fresh (upper panel) and activated sporozoites (lower panel) characterized by the transformation bulb. Sporozoites are visualized by DAPI staining, anti-PfSET7 and anti-HP1 immunostaining. The scale bars correspond to 5  $\mu$ m. (c) Liver stage parasites. Upper panel: liver stage schizonts from TK-NOG humanized mice were immunostained with anti-PfSET7 and anti-HP1 at day 7 post-infection. The scale bar corresponds to 10  $\mu$ m. Lower panel: higher magnification of the image above. The scale bar corresponds to 1  $\mu$ m.

greatly limited our ability to take a target-based approach to PfPKMT inhibitor discovery. Indeed, the difficulty with producing recombinant PfPKMT enzymes by us and other researchers has greatly limited the functional understanding of this entire class of parasite enzymes. Notably, we tested numerous bacterial, eukaryotic and *in vitro* protein expression systems before successfully producing sufficient quantities of active PfSET7 enzyme using a baculovirus-based insect cell expression system.

Here we report the first large-scale production of an active full-length recombinant PfPKMT, allowing a detailed *in vitro* enzyme kinetics investigation of this important enzyme class in malaria parasites. PfSET7 exhibits kinetic characteristics similar to other PKMTs from other organisms in terms of steady-state AdoMet cofactor utilization, peak activity at pH ~9 and a dependence upon a catalytic histidine residue<sup>24,27,37</sup>. Mutational analysis of truncated PfSET7 enzymes revealed at least the post-SET domain of the C-terminal protein beyond the SET domain is required for PKMT activity, as truncated enzymes lacking this segment of the protein were inactive. PfSET7 does not exhibit decreased activity at increased salt concentrations, as has been reported for some PKMT enzymes when using peptide substrates but not nucleosome protein substrates, a result that has previously been suggested to support the interpretation that nucleosomes are the physiological substrate of a given PKMT<sup>38</sup>. Methyl transfer to isolated nucleosome protein substrates is saturable, further suggesting histones are a legitimate protein substrate for PfSET7. As additional support that PfSET7 is a histone methyltransferase, nucleosome labeling studies using heavy-labeled AdoMet reveals that PfSET7 extensively methylates H3K4 and H3K9, the latter target particularly in the presence of a pre-existing H3K14ac mark. Despite the specific methylation of histones, which are overwhelmingly present in the parasite nucleus, Western blot analysis and immunofluorescence localization reveal PfSET7 is detected primarily in the cytosol in blood stage parasites. Additional immunofluorescence localization in salivary gland sporozoites and liver stage parasites reveals PfSET7 to be localized throughout the cell in sporozoites and, in liver stage parasites, located at distinct foci in the cytosol adjacent to parasite nuclei similar to blood stage parasites. The specific role of PKMT enzymes in sporozoites has not been studied. PKMT enzymes in liver stage hypnozoite forms of *P. cynomolgi* malaria parasites have been linked to the maintenance of quiescence<sup>39</sup>, though no specific PfPKMT has been implicated in this process.

The functional significance of H3K4, H3K9 and to a lesser extent H3K36 methylation by PfSET7 is not presently understood. The methylation of H3K9 in the presence of existing H3K14 acetylation has been reported previously in mass spectrometry-based analyses of *P. falciparum* histones<sup>12</sup>, indicating this is indeed a bona fide histone PTM combination present in malaria parasites. By ChIP analysis, the H3K9me3 mark in *P. falciparum*

has been closely linked to the monoallelic expression of genes encoding variant surface antigens<sup>5,6</sup>. Importantly, these ChIP studies used commercially available antibodies specific for the single H3K9me3 histone modification, which would presumably not discriminate between the presence or absence of H3K14ac with regards to H3K9me3. Further investigation could therefore determine whether the H3K9me3K14ac mark is responsible for regulating a subset of the genes associated with H3K9me3, and thus implicate PfSET7 in an even finer level of epigenetic control over multi-copy gene families. This finding also highlights the strengths of using mass spectrometry to detect heavy methyl labeling of modified nucleosome substrates, as assays employing unmodified histone substrates (peptides, histone proteins or nucleosomes) would fail to detect such activity. While methylation of H3K4 and H3K9 (in the presence of K14ac) methylation by PfSET7 were observed to be most abundant, further studies will be required to elucidate the significance of the H3K27 and H3K36 methylations observed in lower abundance (Figs S18–S20). Notably, the control experiment using mouse G9a enzyme reveals similar methylation of H3K27me3, perhaps resulting from the sequence homology between H3K9 and H3K27, and thus could represent a consequence of the *in vitro* assay conditions. However, the PfSET7 mediated methylation of H3K36 in the context of unmethylated H3K27 substrates is not observed in the mouse G9a enzyme control, and thus may represent an additional interesting target to investigate further.

The primarily cytosolic localization of PfSET7 in asexual blood stage parasites was unanticipated for a putative histone methyltransferase. However, PKMT enzymes which are localized in both the nucleus and the cytosol have been reported, including the human histone H3K9 specific PKMT SetDB1<sup>40</sup>. As well, newly synthesized histones have been reported in other organisms to be methylated at H3K9 and acetylated at H3K14 within histone-chaperone protein complexes in the cytoplasm<sup>41,42</sup>. This opens the possibility that PfSET7 is a cytosolic histone methyltransferase acting on newly synthesized histones in a PfSET7-defined subcellular cytosolic compartment. Future pull-down studies could confirm whether similar histone-chaperone protein complexes exist in *Plasmodium*. Alternatively, PfSET7 may methylate recycled, pre-modified histones. Very little is known about the order of addition of histone PTMs or histone recycling in *Plasmodium*, and additional investigation of PfSET7 through genetic or chemical attenuation may be able to address the cytosolic role of PfSET7 on the methylation of newly synthesized or recycled histones. Alternatively, many PKMT enzymes which were originally reported to be histone methyltransferases were subsequently found to possess one or more additional non-histone protein substrates<sup>15</sup>. As such, it is entirely possible that PfSET7 has additional non-histone protein substrates in the parasite cytosol, though further studies would be required to address this hypothesis.

Altogether, the production and enzymatic characterization of recombinant PfSET7 enzyme now allows for small molecule inhibitor discovery against this histone methyltransferase previously reported to be essential in blood stage parasites<sup>11</sup>. Inhibitors identified through target-based discovery can be used as tools to dissect the biological role of PfSET7 with regards to histone and potential non-histone protein methylation, and the role of the H3K9me3K14ac histone PTM combination for which PfSET7 appears to be specific. Chemical biology experiments using PfSET7 specific inhibitors can be corroborated using, for example, inducible knock-down of PfSET7 in blood stage parasites, the stage in which various genetic techniques have been established and continue to be developed. However, chemical genetics investigations using specific PfSET7 inhibitors may prove much more powerful in determining the biological role of PfSET7 in sporozoite and liver stage parasites, which are much less amenable to genetic manipulation, but present fascinating biology as the obligate initial stage of malaria infection and are responsible for malaria relapses in *P. vivax* and *P. ovale* infections. Ultimately, successfully developed small molecule inhibitors of PfSET7 may enter the antimalarial drug discovery pipeline as a much needed new chemical entity against this novel target class.

## Methods

**Protein expression constructs.** Codon optimized full-length PfSET7 (PfSET7FL) gene (PlasmoDB Gene ID PF3D7\_1115200) was synthesized (Genscript), fused to a C-terminal double Strep-tag and ligated into the transfer vector pVL1393. Two potential glycosylation sites in PfSET7FL, S122A and T784A, were mutated to allow expression in secretory systems. The SET domain of PfSET7 was identified using Prosite (prosite.expasy.org). Truncated constructs of PfSET7 were generated by PCR (Pfu Ultra II, Agilent Technologies) using specific primers (Table S1). PfSET7 $\Delta$ C (residues 1–558), PfSET7 $\Delta$ C $\Delta$ N (residues 356–558), PfSET7 $\Delta$ N (residues 356–793) and the double strep-tag, were amplified and inserted into pVL1393 via Gibson assembly (NEB). The PfSET7\_H517A catalytic mutant was generated by site directed mutagenesis using PfSET7FL as template (primers in Table S1). Clones were verified to be error free and in frame by sequencing (GATC).

**Protein expression and purification.** Recombinant proteins were expressed using a baculoviral expression system following manufacturer's instructions. Briefly, 5  $\mu$ g of plasmid, 500 ng linearised baculovirus DNA (BestBac 2.0, Expression Systems) was transfected with Cellfectin (Life Technologies) into Sf9 insect cells grown in media (Insect Xpress, Lonza) supplemented with 5% fetal bovine serum and 50  $\mu$ g/ml gentamycin. Positive baculoviral clones were selected and viral stocks for protein production were made by several cycles of viral amplification. Recombinant proteins were purified using StrepTrap HP columns (GE Healthcare). Briefly, pellets from infected Sf9 insect cells were resuspended in lysis buffer (10 mM TRIS, pH 7.5, 300 mM NaCl, 1% Triton X-100, 10% Glycerol and protease inhibitors), lysed by sonication using 6 bursts of 15 seconds (Vibra-Cell, Sonics and Materials) with a 1 minute pause on ice between bursts. The lysate was clarified by centrifugation (20 000  $\times$  g, Beckman Coulter) for 1 h at 4  $^{\circ}$ C. The cleared lysate was filtered through 0.45  $\mu$ m filters (Durapore, Millipore), loaded onto the column, washed with wash buffer (10 mM TRIS, pH 8, 150 mM NaCl, 1 mM EDTA with protease inhibitors), then eluted (10 mM TRIS, pH 8, 150 mM NaCl, 1 mM EDTA and 2.5 mM desthiobiotin). Protein concentration was determined by Bradford Assay (Bio-Rad Protein Assay, Bio-Rad). Purified recombinant proteins were aliquoted, flash frozen in liquid nitrogen and stored at  $-80^{\circ}$ C until use. All proteins were judged to be >90% pure by SDS-PAGE.

**Nucleosome extraction.** To isolate nuclei, HEK cell pellets were resuspended in lysis buffer (20 mM HEPES, pH 7.5, 250 mM sucrose, 4 mM MgCl<sub>2</sub>, 0.5% NP-40, 0.4 mM PMSF and 5 mM TCEP), homogenized with a dounce (Wheaton), then centrifuged at 2000 × *g* for 15 minutes at 4 °C. The nuclear pellet was suspended in wash buffer (20 mM HEPES, pH 7.5, 250 mM sucrose, 4 mM MgCl<sub>2</sub>, 0.4 mM PMSF and 5 mM TCEP), centrifuged and nuclei were resuspended in wash buffer. To 0.5 ml of suspended nuclei, 20 ml of 0.65 M NaCl-sodium phosphate buffer (0.65 M NaCl, 50 mM phosphate, pH 6.8, 0.4 mM PMSF and 5 mM TCEP) was added dropwise with constant stirring, then vortexed for 1 hour at 4 °C. The nuclear lysate was passed through a column loaded with pre-wet Hydroxylapatite Fast Flow (Calbiochem) then washed with 0.65 M NaCl-sodium phosphate buffer. Nucleosomes were isolated with elution buffer (2.5 M NaCl, 50 mM phosphate, pH 6.8, 0.4 mM PMSF and 5 mM TCEP). Fractions containing nucleosomes were pooled and buffer-exchanged (HiPrep Desalting, GE Healthcare) into 50 mM sodium phosphate, pH 6.8. Pooled fractions were concentrated to 1 mg/ml by centrifugation (Amicon Ultra, Millipore) and stored at −20 °C.

***In vitro* methyltransferase assay.** Standard reactions were performed in a total volume of 25 μl in KMT buffer (50 mM TRIS, pH 8.8, 5 mM MgCl<sub>2</sub>, 4 mM dithiothreitol), with HEK nucleosomes (0.2 mg/ml final) as substrate, unless otherwise stated. For autoradiography, 2.5 μl of S-[<sup>3</sup>H-methyl]-adenosyl-L-methionine (<sup>3</sup>H-AdoMet) was added (3.7 μM final, 15 Ci/mmol, Perkin-Elmer). For kinetic characterization, <sup>3</sup>H-AdoMet was diluted with unlabeled AdoMet (Sigma) to give a final concentration of 100 μM at 0.025 Ci/mmol. PfSET7 enzyme was used at 25 nM and reactions were incubated at room temperature for 1 hour unless otherwise stated. HKMT reactions at various pH values were performed using citrate-HEPES-CHES buffer (50 mM Tri-Sodium Citrate, 50 mM HEPES, 50 mM CHES, 5 mM MgCl<sub>2</sub> and 4 mM DTT)<sup>43</sup>. 10 μl of each reaction was spotted in duplicate onto P81 Filters (Unifilter 96 well plate, Whatman), washed with 200 mM ammonium bicarbonate, dried, overlaid with scintillation fluid (Betaplate Scint, Perkin-Elmer) and read in a scintillation counter (MicroBeta2, Perkin-Elmer) and counts were converted to enzyme turnover. Analysis by fluorography was performed by resolving 20 μl of standard HKMT reactions (25 nM of PfSET7FL or G9a, 20 μg/ml of either BSA or nucleosomes, 2.5 μl of <sup>3</sup>H-AdoMet at 15 Ci/mmol, in a total volume of 25 μl in KMT buffer) on 4–15% SDS-PAGE, the gel was rinsed in EN<sup>3</sup>HANCE (Perkin-Elmer) according to manufacturer's instructions, dried and exposed to film (BioMax MR Film, Carestream). Duplicate reactions were performed with unlabeled AdoMet, resolved on 4–15% SDS-PAGE and silver stained (Dodeca, Silver Stain Kit, Bio-Rad). Mutational analysis experiments tested enzymes at 25 nM and 100 nM.

**Histone extraction and LC-MS/MS.** Deuterium labeled S-[<sup>2</sup>H<sub>3</sub>-methyl]-adenosyl-L-methionine (CD3-AdoMet) was synthesized using iodomethane-d<sub>3</sub>, AgOTf and HCOOH, according to the method of Stecher<sup>44</sup>. Nano-liquid chromatography tandem mass spectrometry (nanoLC-MS/MS) was utilized to identify sites of PfSET7 methylation on histone variants H3 and H4 based on the observation of mass shifts resulting from the incorporation of CD3 from CD3-AdoMet. Histones were acid extracted from the *in vitro* reaction buffer and processed for nanoLC-MS/MS analysis as previously described<sup>45</sup>. Briefly, the histone proteins were propionylated to label unmodified and monomethylated lysines with a propionyl group (+56 Da), which will result in a trypsin digest cleaving only at arginine residues. This produces the same histone peptide sequences regardless of any post-translational modifications present on lysine residues, and thus enables relative quantitation between different modifications on the same peptide based on signal intensity in the mass spectrometer. Subsequent propionylation of the N-termini of the corresponding histone peptides results in enhanced chromatographic resolution of isobaric (i.e. same mass) peptides (e.g. trimethylation [42.0469 Da] and acetylation [42.0106 Da] of lysine) based on their degree of hydrophobicity, thereby eliminating issues relating to signal interference and increasing confidence in PTM identification<sup>46</sup>. Approximately 1 μg of histone peptides were chromatographically resolved using an Ultimate 3000 RS-LC-nano System (Dionex), with an Acclaim PepMap100, C18 stationary phase, 2 μm particle size, 100 Å pore size, 75 μm internal diameter x 15 cm length column (Thermo Fisher). The nanoLC gradient started at 1% B (5% H<sub>2</sub>O, 95% MeCN, and 0.1% formic acid) and 99% A (0.1% formic acid and 100% H<sub>2</sub>O) and increased to 30% B over 35 min followed by an increase to 95% B over 30 mins. Real-time tandem mass spectra were acquired on an LTQ Velos Pro linear ion trap (Thermo Scientific) with a 110 min acquisition time. Targeted zoom scans were performed for *m/z* values corresponding to the modified histone peptides to mitigate dynamic range issues and tandem mass spectra corresponding to methylated histone peptides were acquired using a priority queue.

**Transgenic parasite plasmid constructs.** The CRISPR/Cas9 genome editing approach was employed to generate HA-tagged PfSET7 transgenic parasite line as described previously<sup>47</sup>. We used Protospacer Workbench to identify sgRNA sequences for PfSET7 and computationally predict off-target sites<sup>48</sup>. A 20-nt gRNA target sequence (5'-CACTGATGCTCCTCAAATTG-3') directing Cas9 cleavage near the C-terminal end of PfSET7 was cloned into the sgRNA-expression cassette in the pL6 plasmid. Two homology regions were inserted in separate cloning steps using Gibson assembly to generate the pL6-Pfset7 plasmid. The primers used are listed in Table S1. A second plasmid, pUF1-Cas9, containing an engineered *S. pyogenes* endonuclease Cas9 cassette and regulatory elements of *Pfalciparum* U6 snRNA was reported previously<sup>47</sup>.

**Parasite culture and transfections.** Asexual blood-stage parasites were cultivated using standard methods, and synchronous cultures were obtained by sorbitol treatment and plasmon enrichment. Parasites were transfected using the pre-loaded red blood cells method<sup>49</sup>. Positive selection using 2.66 nM WR99210 and 1.5 μM DSM1 and negative selection using 40 μM 5-fluorocytosine were applied sequentially in culture until the integrant transfected line was established. Genomic integration was verified by PCR.



**PfSET7 antibody production.** C57BL/6 and BALB/C mice were immunized 4 times at 14-day intervals with 20 µg doses of purified recombinant PfSET7-FL in Freund's Complete Adjuvant for the initial injection and Freund's Incomplete Adjuvant for booster injections to generate a polyclonal anti-PfSET7 antibody. Specificity of immune sera was determined by Western blot (Fig. S25).

**Subcellular fractionation and western blotting.** Subcellular extract preparations were prepared as described previously<sup>50</sup>. Equal amounts of protein were loaded and separated on 4–12% SDS-PAGE gel (Invitrogen), and analyzed by western blot using antibodies anti-PfSET7 (mouse), anti-HA (mouse, Roche 12CA5 or rat, Roche 3F10), anti-Pf-aldolase (Abcam ab38905) or anti-Histone H3 (mouse, Abcam ab1791), all at 1:2000 dilution. Secondary antibodies were anti-rabbit IgG-HRP (GE NA934V, Lot 9568295) and anti-mouse IgG-HRP (GE NA931V, Lot 9648752), both at 1:4000. Blots were developed using ECL (Thermo Scientific).

**Immunofluorescence microscopy.** Immunofluorescence microscopy on extracellular and intracellular parasites was performed as described previously<sup>51</sup>. Antibody dilutions are as follows: anti-PfHP1 (rabbit, Genscript) 1:4000, anti-PfSET7 (mouse) 1:2000, anti-HA (mouse or rat, Roche) 1:2000, anti-Bip (rabbit, MR4) 1:1000, anti-ERD2 (rat, MR4) 1:1000, and secondary goat anti-rat/rabbit/mouse Alexa-Fluor 488 or 568 conjugates (Invitrogen), all at 1:2500. Mitochondria were stained with 100 nM MitoTracker-DeepRedFM for 20 minutes at 37 °C in the dark before fixation and permeabilization. Images were captured using a Nikon Eclipse 80i microscope with a CoolSnapHQ2 camera (Photometrics). NIS Elements version 3.0 software was used for image acquisition, and Fiji software was used for analysis.

*P. falciparum* (NF54) sporozoites were isolated by aseptic dissection of the salivary glands of infected *Anopheles stephensi* obtained from the Department of Medical Microbiology, University Medical Centre, St Radboud, Nijmegen, the Netherlands.  $5 \times 10^4$  sporozoites placed on poly-L-lysine coated coverslips were either fixed or activated for 1 h at 37 °C and then fixed in 4% PFA for 10 min at room temperature. Samples were blocked in normal goat serum diluted 1:500 for 2 h at 37 °C, washed twice in PBS and incubated with anti-PfSET7 antibody at 1:1000 and anti-PfHP1 at 1:2000 for 1 h, washed twice in PBS, then incubated with secondary goat anti-mouse IgG coupled to Alexa-Fluor 488 or 594 diluted 1:500 with DAPI 1 µg/ml in the dark for 1 h.

Fifty micron thick serial sections of frozen liver from humanized TK-NOG mice infected with *P. falciparum* were prepared as previously described<sup>52</sup>. Sections were incubated overnight with anti-PfSET7 antibody diluted 1:1000 and anti-PfHP1 antibody diluted 1:2000, washed twice in PBS, then incubated with secondary antibodies and DAPI as above, at 37 °C. All antibodies were diluted in PBS containing 1% BSA and 0.2% Triton X-100. All sections were washed twice in PBS before being mounted in anti-fading medium and stored at 4 °C before analysis. Immunostained sporozoites and liver sections were examined under a confocal microscope (Olympus FV-1000, Plateforme d'Imagerie Cellulaire PICPS, La Pitié-Salpêtrière, Paris) and images were analysed using ImageJ software.

**Ethics Statement.** All animal care and experiments involving mice were conducted at the Institut Pasteur, approved by the 'Direction Départementale des Services Vétérinaires' de Paris, France (Permit Number N° 75-066 issued on September 14, 2009) and performed in compliance with institutional guidelines and European regulations ([http://ec.europa.eu/environment/chemical?s/lab\\_animals/home\\_en.htm](http://ec.europa.eu/environment/chemical?s/lab_animals/home_en.htm)). A statement of compliance with the French Government's ethical and animal experiment regulations was issued by the Ministère de l'Enseignement Supérieur et de la Recherche under the number 00218.01.

## References

- World Health Organization. *World Malaria Report 2014*. WHO (2014). at <[http://www.who.int/malaria/publications/world\\_malaria\\_report\\_2014/report/en/](http://www.who.int/malaria/publications/world_malaria_report_2014/report/en/)> Date of access: 15/01/2016.
- Bozdech, Z. *et al.* The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* **1**, E5 (2003).
- De Silva, E. K. *et al.* Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc. Natl. Acad. Sci. USA* **105**, 8393–8 (2008).
- Freitas-Junior, L. H. *et al.* Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* **121**, 25–36 (2005).
- Lopez-Rubio, J. J., Mancio-Silva, L. & Scherf, A. Genome-wide Analysis of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear Repressive Centers in Malaria Parasites. *Cell Host Microbe* **5**, 179–190 (2009).
- Salcedo-Amaya, A. M. *et al.* Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. USA* **106**, 9655–60 (2009).
- Shock, J. L., Fischer, K. F. & DeRisi, J. L. Whole-genome analysis of mRNA decay in *Plasmodium falciparum* reveals a global lengthening of mRNA half-life during the intra-erythrocytic development cycle. *Genome Biol.* **8**, R134 (2007).
- Gunasekera, A. M. *et al.* Widespread distribution of antisense transcripts in the *Plasmodium falciparum* genome. *Mol. Biochem. Parasitol.* **136**, 35–42 (2004).
- Cortés, A. *et al.* Epigenetic silencing of *Plasmodium falciparum* genes linked to erythrocyte invasion. *PLoS Pathog.* **3**, e107 (2007).
- Kafsack, B. F. C. *et al.* A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* **507**, 248–52 (2014).
- Jiang, L. *et al.* PfSETvs methylation of histone H3K36 represses virulence genes in *Plasmodium falciparum*. *Nature* **499**, 223–7 (2013).
- Trelle, M. B., Salcedo-Amaya, A. M., Cohen, A. M., Stunnenberg, H. G. & Jensen, O. N. Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite *Plasmodium falciparum*. *J. Proteome Res.* **8**, 3439–50 (2009).
- Dastidar, E. G. *et al.* Comprehensive Histone Phosphorylation Analysis and Identification of Pf14-3-3 Protein as a Histone H3 Phosphorylation Reader in Malaria Parasites. *PLoS One* **8**, e53179 (2013).
- Hamamoto, R., Saloura, V. & Nakamura, Y. Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat. Rev. Cancer* **15**, 110–24 (2015).
- Copeland, R. A., Solomon, M. E. & Richon, V. M. Protein methyltransferases as a target class for drug discovery. *Nat. Rev. Drug Discov.* **8**, 724–732 (2009).

16. Malmquist, N. A., Moss, T. A., Mecheri, S., Scherf, A. & Fuchter, M. J. Small-molecule histone methyltransferase inhibitors display rapid antimalarial activity against all blood stage forms in *Plasmodium falciparum*. *Proc. Natl. Acad. Sci.* **109**, 16708–16713 (2012).
17. Sundriyal, S. *et al.* Development of Diaminoquinazoline Histone Lysine Methyltransferase Inhibitors as Potent Blood-Stage Antimalarial Compounds. *Chem. Med. Chem.* **9**, 2360–2373 (2014).
18. Malmquist, N. A. *et al.* Histone Methyltransferase Inhibitors Are Orally Bioavailable, Fast-Acting Molecules with Activity against Different Species Causing Malaria in Humans. *Antimicrob. Agents Chemother.* **59**, 950–959 (2015).
19. Jenuwein, T., Laible, G., Dorn, R. & Reuter, G. SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell. Mol. Life Sci.* **54**, 80–93 (1998).
20. Biggar, K. K. & Li, S. S.-C. Non-histone protein methylation as a regulator of cellular signalling and function. *Nat. Rev. Mol. Cell Biol.* **16**, 5–17 (2015).
21. Cui, L. L., Fan, Q., Cui, L. L. & Miao, J. Histone lysine methyltransferases and demethylases in *Plasmodium falciparum*. *Int. J. Parasitol.* **38**, 1083–1097 (2008).
22. Volz, J. C. *et al.* PfSET10, a *Plasmodium falciparum* Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division. *Cell Host Microbe* **11**, 7–18 (2012).
23. Chin, H. G. *et al.* Automethylation of G9a and its implication in wider substrate specificity and HP1 binding. *Nucleic Acids Res.* **35**, 7313–7323 (2007).
24. Patnaik, D. *et al.* Substrate specificity and kinetic mechanism of mammalian G9a histone H3 methyltransferase. *J. Biol. Chem.* **279**, 53248–58 (2004).
25. Rathert, P., Cheng, X. & Jeltsch, A. Continuous enzymatic assay for histone lysine methyltransferases. *Biotechniques* **43**, 602–608 (2007).
26. Xiao, B., Wilson, J. R. & Gamblin, S. J. SET domains and histone methylation. *Curr. Opin. Struct. Biol.* **13**, 699–705 (2003).
27. Fang, J. *et al.* Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr. Biol.* **12**, 1086–1099 (2002).
28. Min, J., Zhang, X., Cheng, X., Grewal, S. I. S. S. & Xu, R.-M. Structure of the SET domain histone lysine methyltransferase Clr4. *Nat. Struct. Biol.* **9**, 828–832 (2002).
29. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–599 (2000).
30. Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. & Jenuwein, T. Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* **410**, 116–120 (2001).
31. Gerace, E. L., Halic, M. & Moazed, D. The methyltransferase activity of Clr4Suv39h triggers RNAi independently of histone H3K9 methylation. *Mol. Cell* **39**, 360–72 (2010).
32. Flueck, C. *et al.* *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog.* **5**, e1000569 (2009).
33. Volz, J. *et al.* Potential epigenetic regulatory proteins localise to distinct nuclear sub-compartments in *Plasmodium falciparum*. *Int. J. Parasitol.* **40**, 109–21 (2010).
34. Botté, C. Y. *et al.* Atypical lipid composition in the purified relict plastid (apicoplast) of malaria parasites. *Proc. Natl. Acad. Sci. USA* **110**, 7506–11 (2013).
35. Kaiser, K., Camargo, N. & Kappe, S. H. I. Transformation of sporozoites into early exoerythrocytic malaria parasites does not require host cells. *J. Exp. Med.* **197**, 1045–50 (2003).
36. Doerig, C., Rayner, J. C., Scherf, A. & Tobin, A. B. Post-translational protein modifications in malaria parasites. *Nat. Rev. Microbiol.* **13**, 160–172 (2015).
37. Trievel, R. C., Beach, B. M., Dirk, L. M. a., Houtz, R. L. & Hurley, J. H. Structure and catalytic mechanism of a SET domain protein methyltransferase. *Cell* **111**, 91–103 (2002).
38. Xiao, B. *et al.* Specificity and mechanism of the histone methyltransferase Pr-Set7. *Genes Dev.* **19**, 1444–54 (2005).
39. Dembélé, L. *et al.* Persistence and activation of malaria hypnozoites in long-term primary hepatocyte cultures. *Nat. Med.* **20**, 307–312 (2014).
40. Loyola, A. *et al.* The HP1 $\alpha$ -CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. *EMBO Rep.* **10**, 769–775 (2009).
41. Alvarez, F. *et al.* Sequential establishment of marks on soluble histones H3 and H4. *J. Biol. Chem.* **286**, 17714–17721 (2011).
42. Loyola, A., Bonaldi, T., Roche, D., Imhof, A. & Almouzni, G. PTMs on H3 Variants before Chromatin Assembly Potentiate Their Final Epigenetic State. *Mol. Cell* **24**, 309–316 (2006).
43. Newman, J. Novel buffer systems for macromolecular crystallization. *Acta Crystallogr. D. Biol. Crystallogr.* **60**, 610–2 (2004).
44. Stecher, H. *et al.* Biocatalytic Friedel-Crafts Alkylation Using Non-natural Cofactors. *Angew. Chemie Int. Ed.* **48**, 9546–9548 (2009).
45. Leroy, G. *et al.* A quantitative atlas of histone modification signatures from human cancer cells. *Epigenetics Chromatin* **6**, 20 (2013).
46. Bartke, T., Borgel, J. & DiMaggio, P. a. Proteomics in epigenetics: new perspectives for cancer research. *Brief. Funct. Genomics* **12**, 205–18 (2013).
47. Ghorbal, M., Gorman, M., Macpherson, C. R., Martins, R. M. & Scherf, A. Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nat. Biotechnol.* **2**, (2014).
48. MacPherson, C. R. & Scherf, A. Flexible guide-RNA design for CRISPR applications using Protospacer Workbench. *Nat. Biotechnol.* **33**, 805–6 (2015).
49. Deitsch, K. W. Transformation of malaria parasites by the spontaneous uptake and expression of DNA from human erythrocytes. *Nucleic Acids Res.* **29**, 850–853 (2001).
50. Lanzer, M., de Bruin, D. & Ravetch, J. V. A sequence element associated with the *Plasmodium falciparum* KAHRP gene is the site of developmentally regulated protein-DNA interactions. *Nucleic Acids Res.* **20**, 3051–6 (1992).
51. Guizetti, J., Martins, R. M., Guadagnini, S., Claes, A. & Scherf, A. Nuclear pores and perinuclear expression sites of var and ribosomal DNA genes correspond to physically distinct regions in *Plasmodium falciparum*. *Eukaryot. Cell* **12**, 697–702 (2013).
52. Soulard, V. *et al.* *Plasmodium falciparum* full life cycle and *Plasmodium ovale* liver stages in humanized mice. *Nat. Commun.* **6**, 7690 (2015).

## Acknowledgements

We thank Stéphane Petres and the Recombinant Proteins in Eukaryotes platform, Institut Pasteur, Paris, for providing us with the pVL1393 plasmid, for making the PfSET7FL enzyme and for valuable advice. We thank Cyrille Y. Botté and Geoff McFadden for providing us with the parasite line expressing PfoTPT-HA. We gratefully acknowledge Katie Addison and Nicole Trainor for contributing to the synthesis of heavy-labeled AdoMet. The expression plasmid for mouse G9a enzyme was kindly provided by Thomas Jenuwein, Max Planck Institute, Freiburg. This work was funded by European Research Council Advanced Grants (PlasmoEscape 250320 and PlasmoSilencing 670301), the French Parasitology Consortium ParaFrap (ANR-11-LABX0024) and the Agence Nationale de Recherche (HypEpiC ANR-14 CE 160013 02).

### Author Contributions

P.B.C. expressed and purified enzymes and performed enzyme characterization and kinetics experiments. S.D. performed cellular localization experiments in blood stage parasites. G.Z. performed cellular localization experiments in sporozoite and liver stage parasites. V.S. provided frozen mouse liver sections infected with *P. falciparum*. P.A.D. designed, performed, analyzed and interpreted mass spectrometry experiments. P.A.D. and M.J.F. supervised synthesis of heavy-labeled AdoMet. S.M. supervised anti-PfSET7 antibody production. D.M. and A.S. supervised the project and interpreted the results. N.A.M. designed, performed, analyzed and interpreted enzyme kinetics experiments, designed and performed nucleosome labeling experiments, interpreted and synthesized all results and wrote the manuscript with input from all authors. All authors contributed to the writing of and reviewed the manuscript.

### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Chen, P. B. *et al.* *Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites. *Sci. Rep.* **6**, 21802; doi: 10.1038/srep21802 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

## Title

*Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites

## Authors

Patty B. Chen<sup>1,2,3,§</sup>, Shuai Ding<sup>1,2,3,§</sup>, Gigliola Zanghi<sup>4</sup>, Valérie Soulard<sup>4</sup>, Peter A. DiMaggio Jr<sup>6</sup>, Matthew J. Fuchter<sup>7</sup>, Salah Mecheri<sup>1,2,3</sup>, Dominique Mazier<sup>4,5</sup>, Artur Scherf<sup>1,2,3</sup>, Nicholas A. Malmquist<sup>1,2,3,\*</sup>

<sup>1</sup> Unité Biologie des Interactions Hôte-Parasite, Département de Parasites et Insectes Vecteurs, Institut Pasteur, Paris 75015, France. <sup>2</sup> CNRS, ERL 9195, Paris 75015, France. <sup>3</sup> INSERM, UMR 1201, Paris 75015, France. <sup>4</sup> Sorbonne Universités, UPMC Univ Paris 06, INSERM U1135, CNRS ERL 8255, Centre d'Immunologie et des Maladies Infectieuses (CIMI-Paris), 91 Bd de l'hôpital, 75013, Paris, France. <sup>5</sup> AP HP, Centre Hospitalo-Universitaire Pitié-Salpêtrière, 75013 Paris, France. <sup>6</sup> Department of Chemical Engineering, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom. <sup>7</sup> Department of Chemistry, Imperial College London, South Kensington Campus, London SW7 2AZ, United Kingdom.

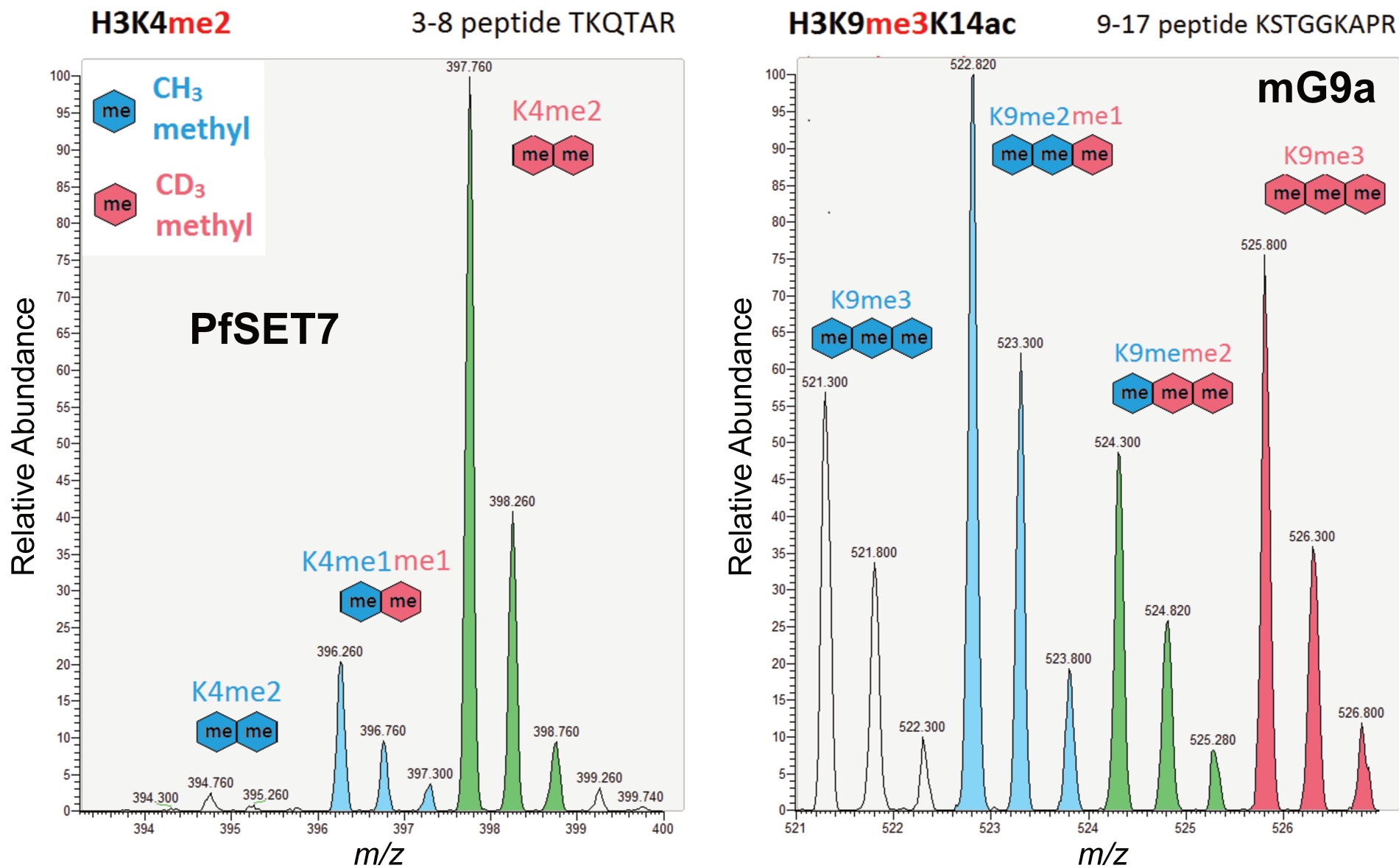
\*Correspondence and requests for materials should be addressed to N.A.M (e-mail: [nicholas.malmquist@pasteur.fr](mailto:nicholas.malmquist@pasteur.fr))

§These authors contributed equally to this work.

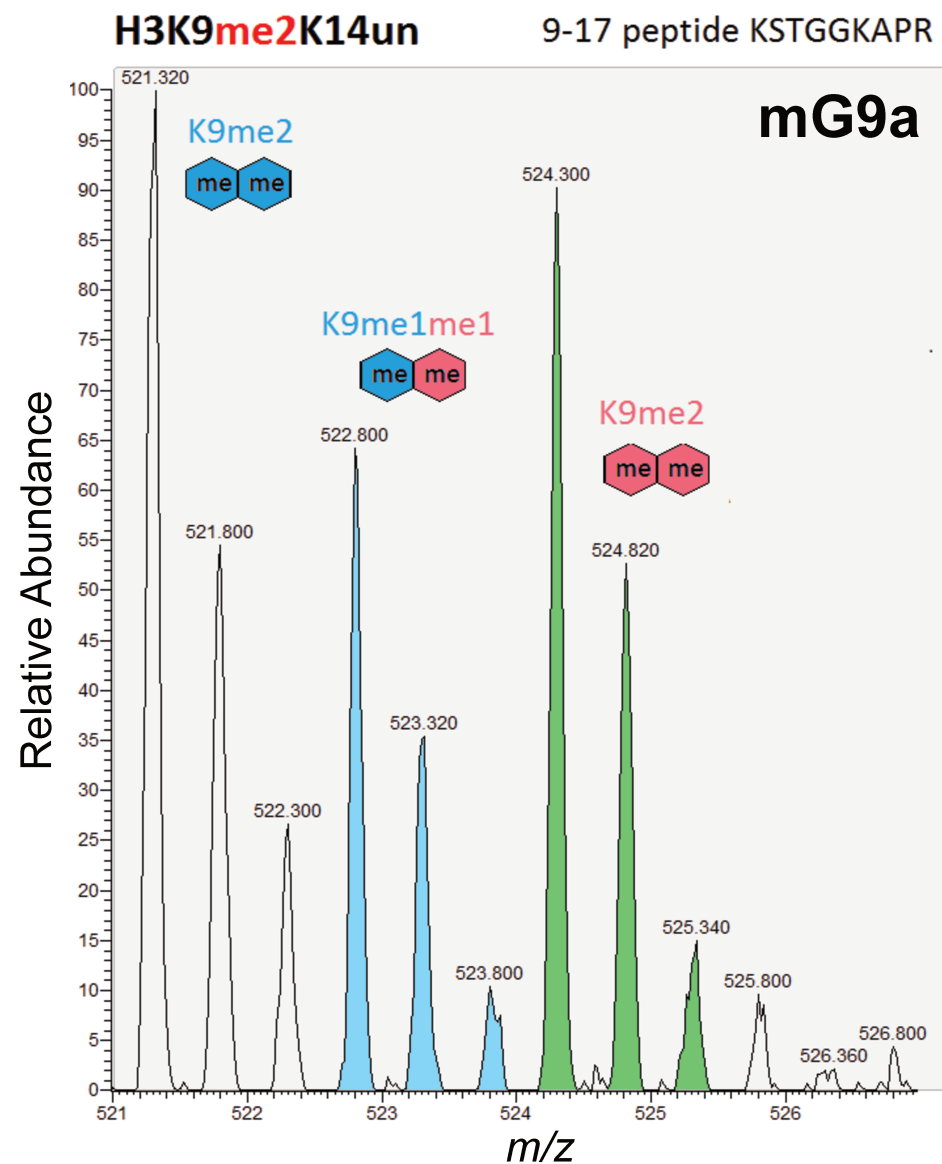
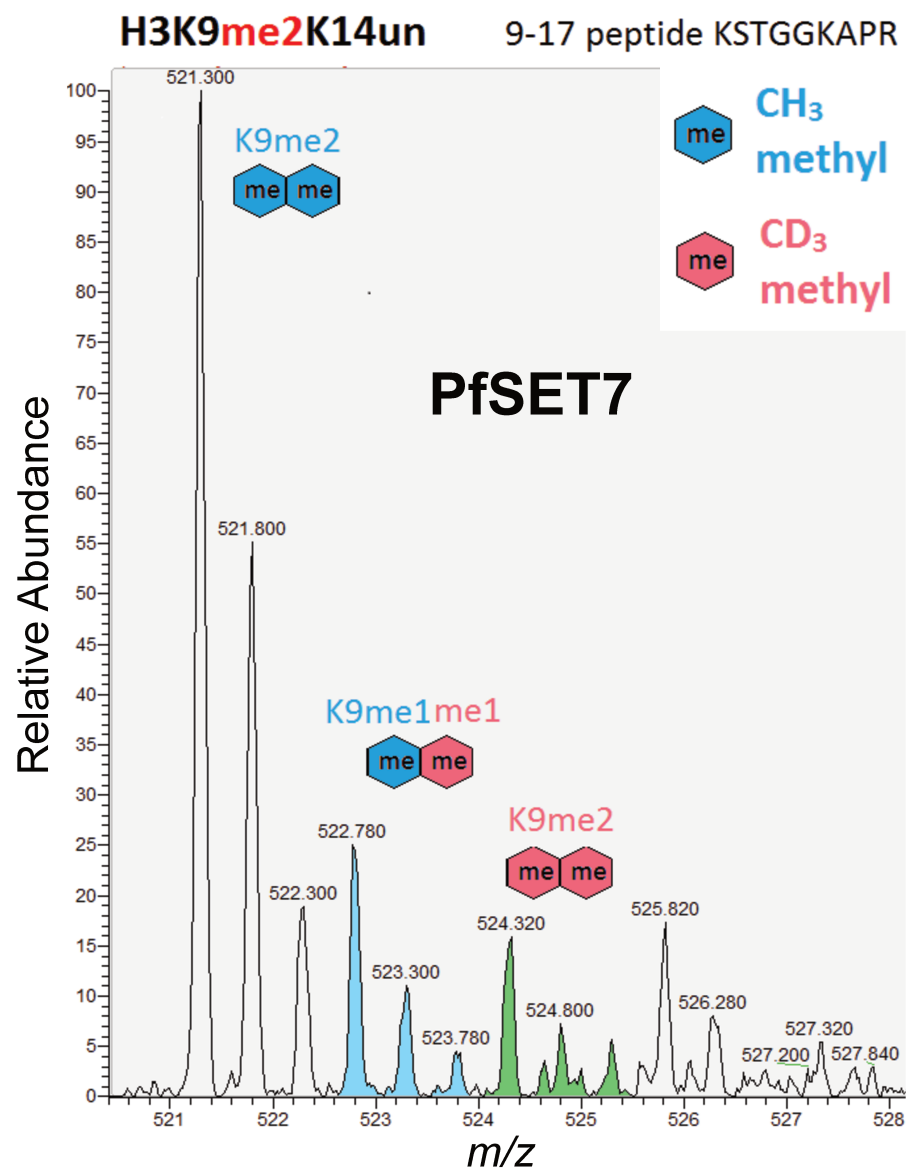
## Supplementary Data

Product	Primer	Sequence
PfSET7ΔC	Primer 1F Primer 3R	5'-CATCGGGCGCGGATCCATGGAACTATTTTTAGGAACTC-3' 5'-ATAGAGGTCATCGTCGCCAGGTA-3'
PfSET7ΔCΔN	Primer 2F Primer 3R	5'-CATCGGGCGCGGATCCATGAACGACGTGGAAATCTTCAATGTCA-3' 5'-ATAGAGGTCATCGTCGCCAGGTA-3'
PfSET7ΔN	Primer 2F Primer 5R	5'-CATCGGGCGCGGATCCATGAACGACGTGGAAATCTTCAATGTCA-3' 5'-AGATCTGCAGCGGCCGGATCAGCGGGTTTAACTCACTTCTC-3'
Strep-Tag	Primer 4F Primer 5R	5'-GACGATGACCTCTATGACGATGACGATAAGGCCGGTTG-3' 5'-AGATCTGCAGCGGCCGGATCAGCGGGTTTAACTCACTTCTC-3'
PfSET7_H517A	Primer 6F Primer 7R	5'-CGCATCTCTATGCTGGCAGCCTCATGCATTAGTACAGCCTG-3' 5'-CAGGCTGTACTAATGCATGAGGCTGCCAGCATAGAGATGCG-3'
pL6-Pfset7 plasmid Homology Region 1	Primer 8F Primer 9R	5'-TTCCGCGGGGAGGACTAGTCCAATTATTTCCAAGTGAATGT-3' 5'-TATCCACCGCCTGGCGGCCCTCTTCGATCTGAGGAGCAT-3'
pL6-Pfset7 plasmid Homology Region 2	Primer 10F Primer 11R	5'-TGTAGGAGGGGACGGCGCCATATATTTAGCAATAACTTC-3' 5'-AATTTTTTTTACAAAATGCTTAAGTACCTTAGAACAAATAGG-3'
Integration confirmation	Primer 12F Primer 13R Primer 14F Primer 15R	5'-GTATGTATGTGTAGTAGGTGTACTCATCC-3' 5'-GCATAGTCTGGTACGTCATAGGGATACG-3' 5'-CATGGGGACAACAATGGGGAGAAAGAGG-3' 5'-GCTTAGTTGACGAGGATGGAGGTTATCG-3'
PfSET7 qPCR	Primer qF Primer qR	5'-GATTTGTCTGATGCCTTAGC-3' 5'-TTTCCCATTGAGATGTATCC-3'

Table S1 | Primers for generating PfSET7 mutants, pL6-Pfset7 constructs and quantitative PCR.



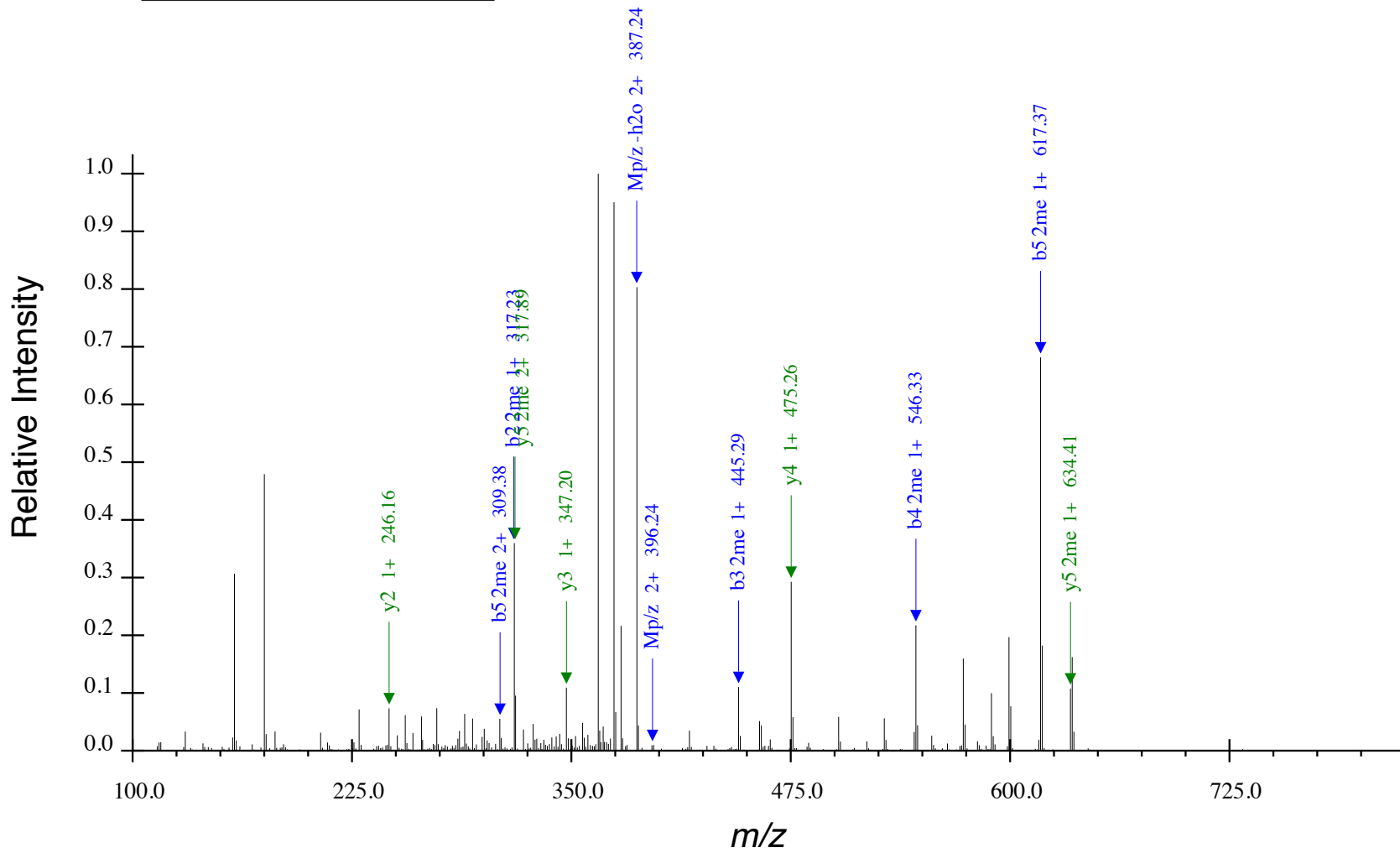
**Figure S1** | Left Panel: PfSET7 methylation of H3K4me2, where the most abundant product is H3K4 with two new CD3 methyl groups (green series of isotopic peaks). Right Panel: heavy methyl labeling observed for mouse G9a-mediated methylation of H3K9me3K14ac.



**Figure S2** | Left Panel: PfSET7 methylation of H3K9me2K14un, which is relatively low in abundance compared to substrates containing H3K14ac (compare to Fig. 5). Right Panel: heavy methyl labeling observed for mouse G9a-mediated methylation of H3K9me2K14un.

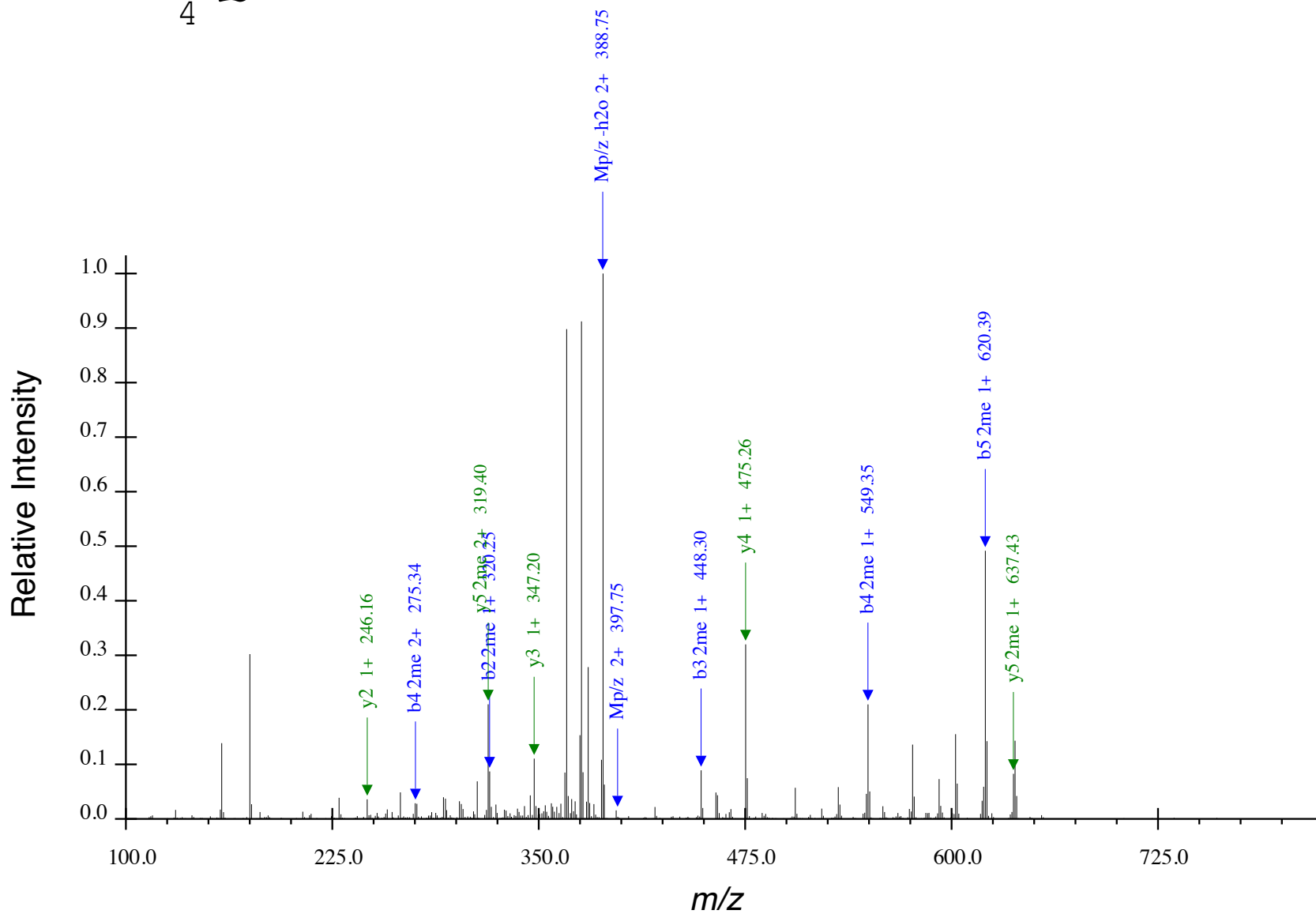
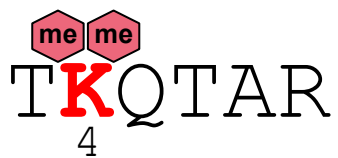


x2

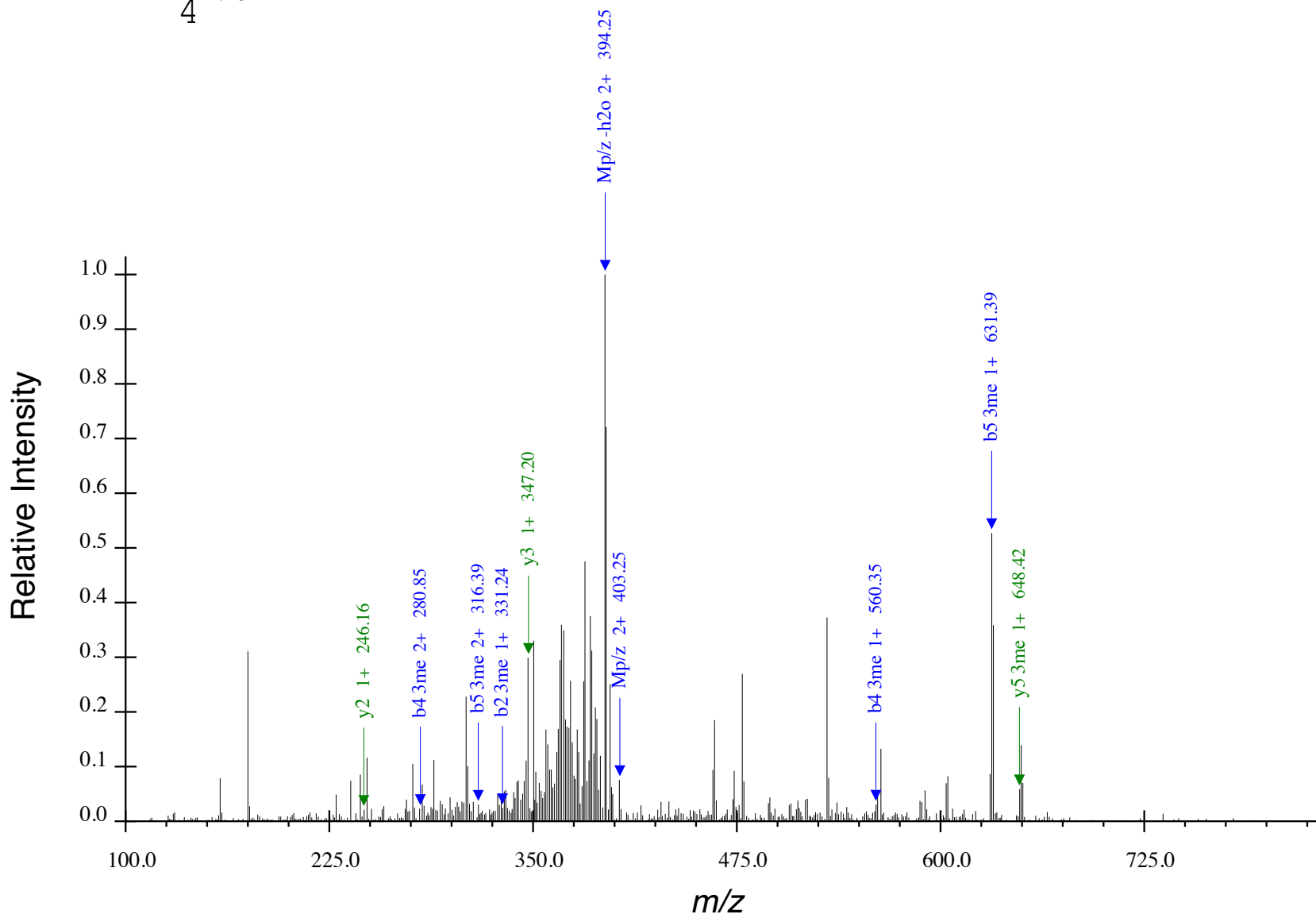


**Figure S3** | Annotated CID tandem mass spectrum for H3K4me2(CH<sub>3</sub>, CD<sub>3</sub>), z = 2, Mp/z = 396.24. N-terminus has been propionylated (+56.0627 Da).

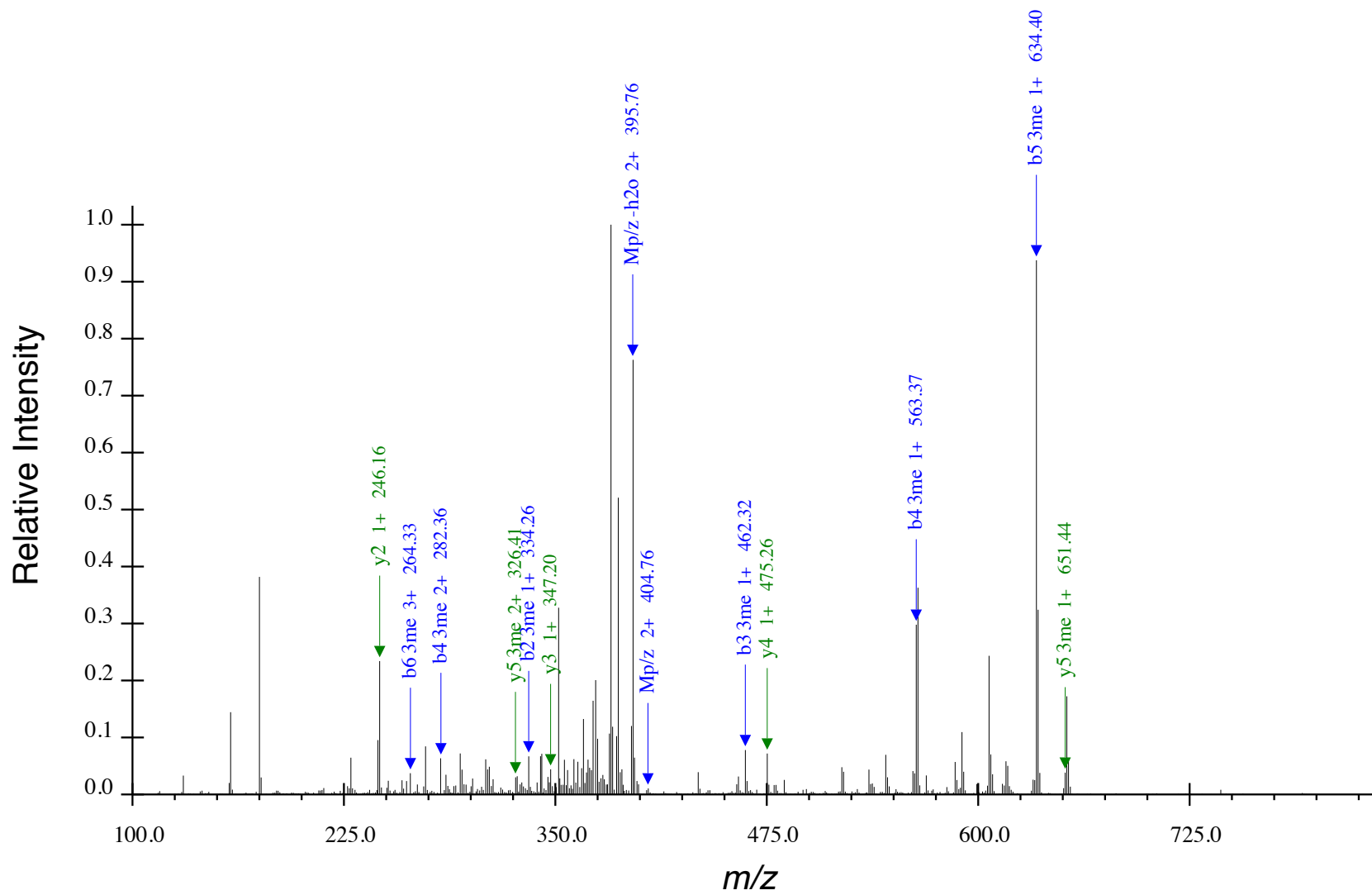




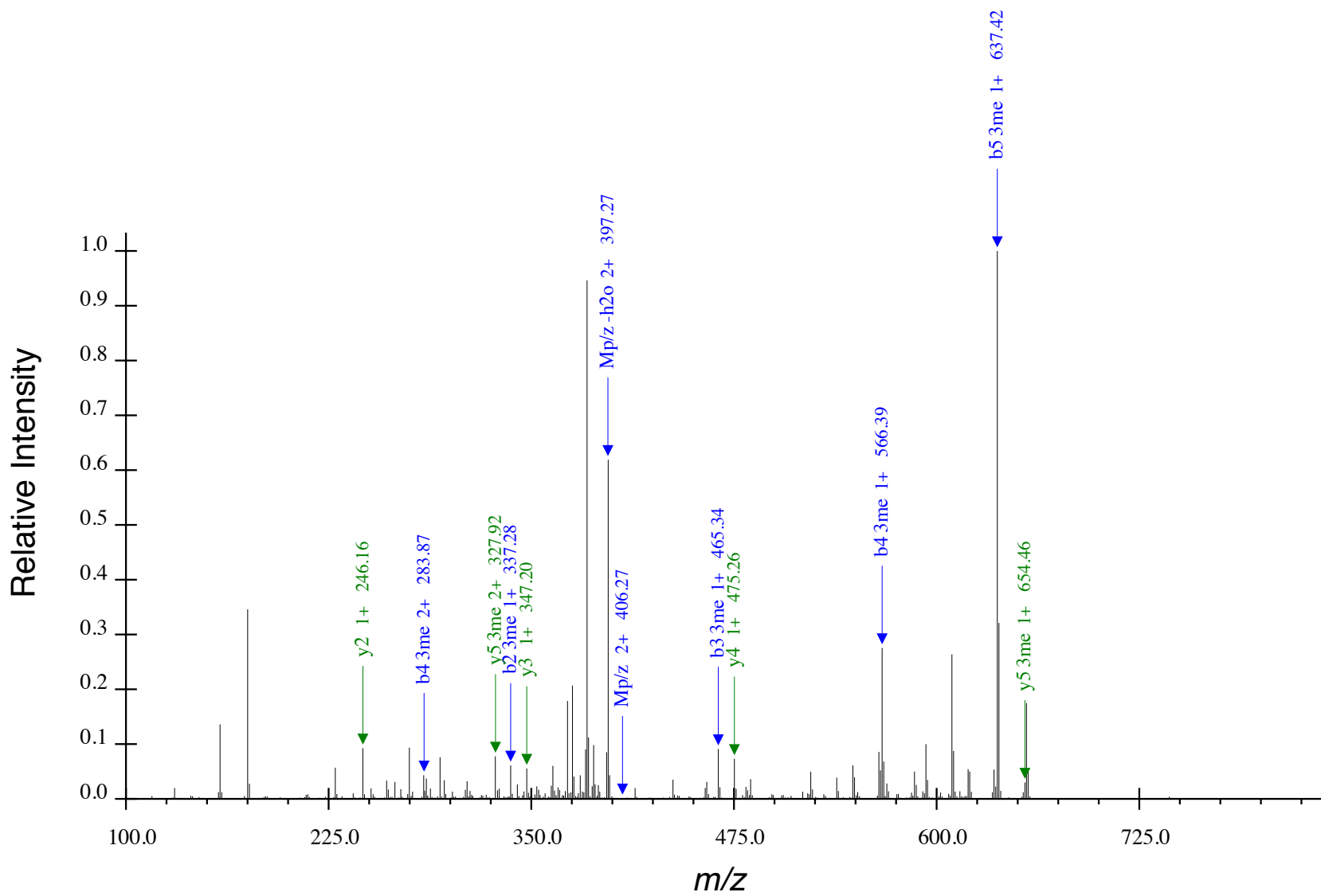
**Figure S4** | Annotated CID tandem mass spectrum for H3K4me2(2CD<sub>3</sub>), z = 2, Mp/z = 397.75. N-terminus has been propionylated (+56.0627 Da).



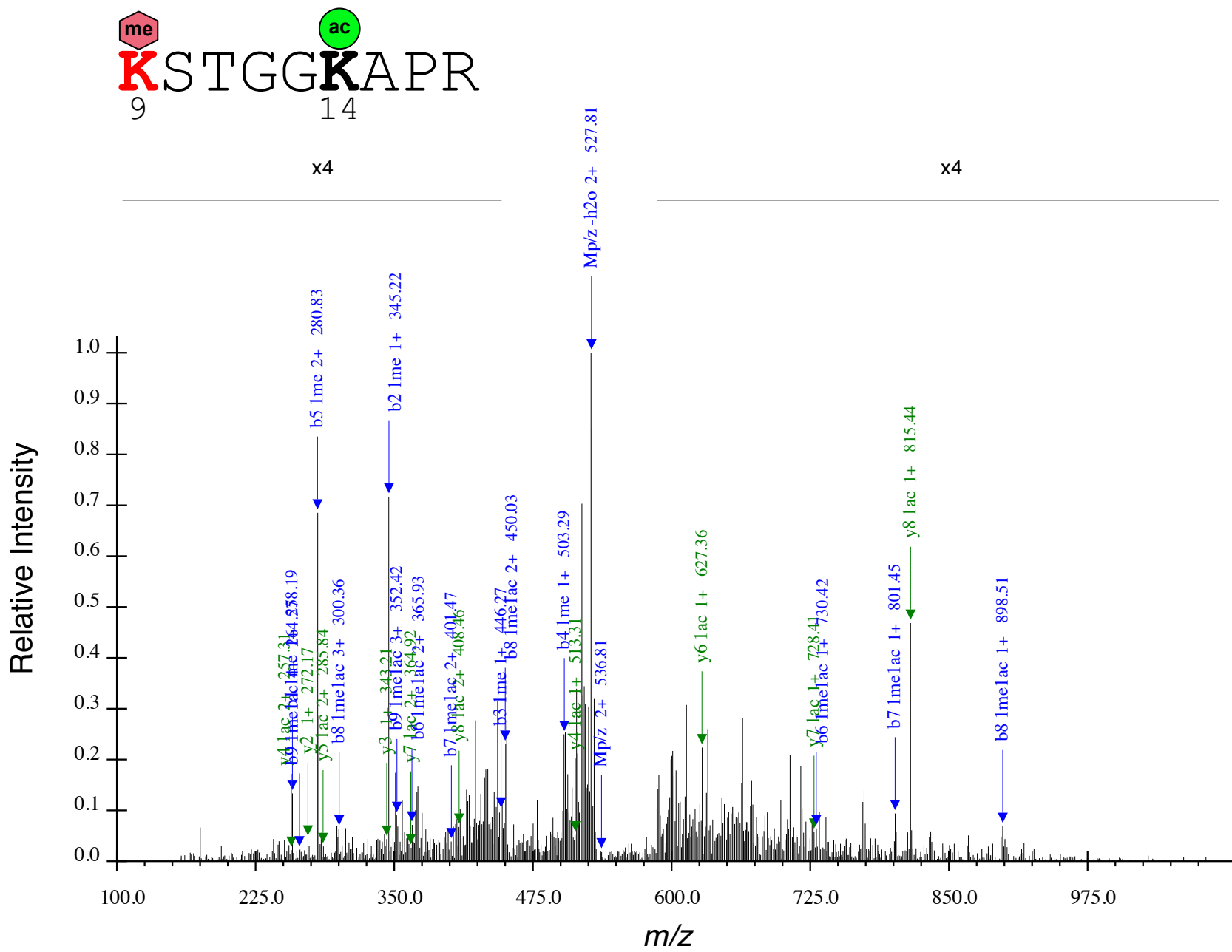
**Figure S5** | Annotated CID tandem mass spectrum for H3K4me3(2CH<sub>3</sub>, CD<sub>3</sub>), z = 2, Mp/z = 403.25. N-terminus has been propionylated (+56.0627 Da).



**Figure S6** | Annotated CID tandem mass spectrum for H3K4me3(CH<sub>3</sub>,2CD<sub>3</sub>), z = 2, Mp/z = 404.76. N-terminus has been propionylated (+56.0627 Da).



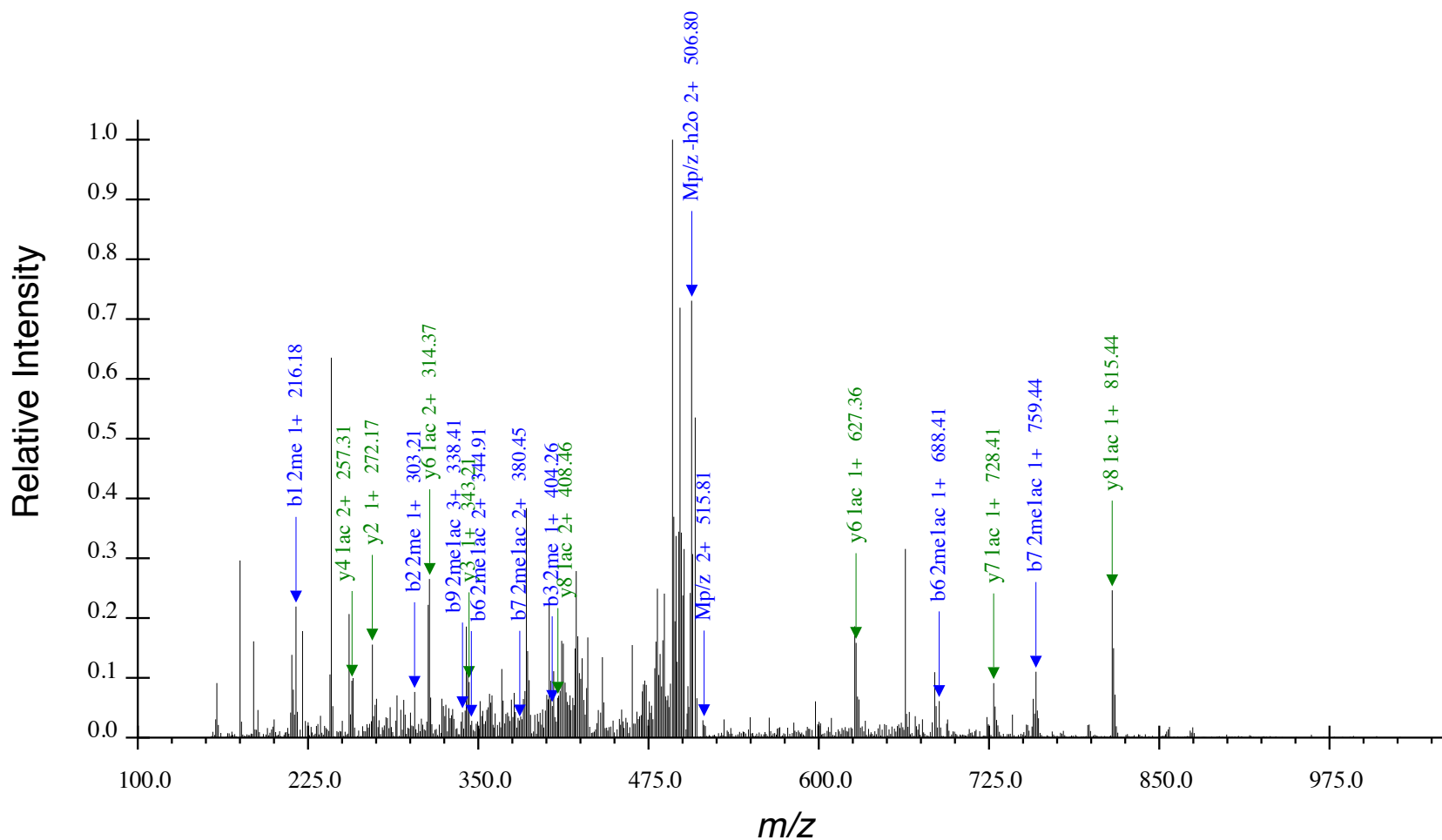
**Figure S7** | Annotated CID tandem mass spectrum for H3K4me3(3CD<sub>3</sub>), z = 2, Mp/z = 406.27. N-terminus has been propionylated (+56.0627 Da).



**Figure S8** | Annotated CID tandem mass spectrum for H3K9me1(CD<sub>3</sub>)K14ac, z = 2, Mp/z = 536.81. N-terminus and K9 have been propionylated (+56.0627 Da).



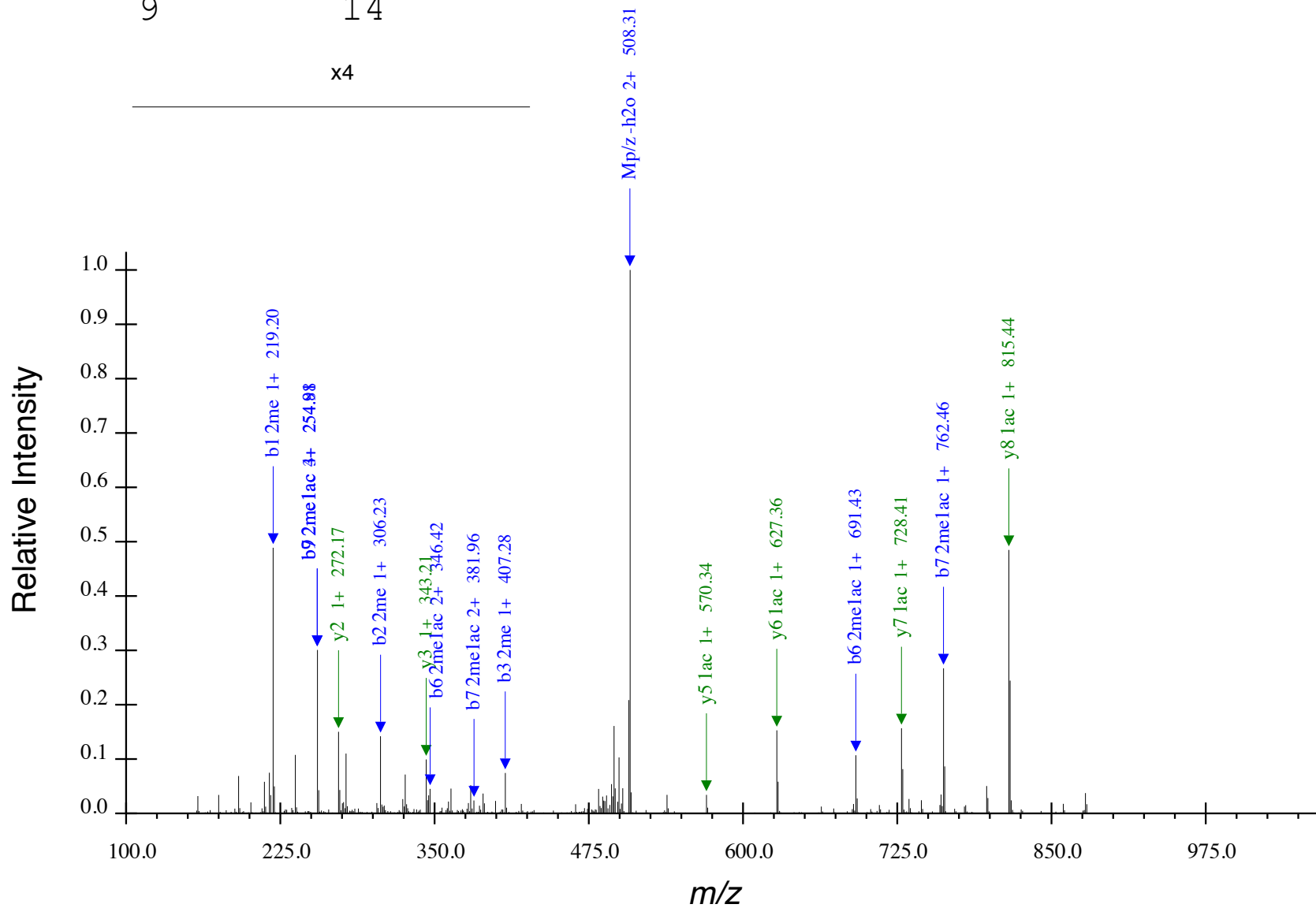
x4



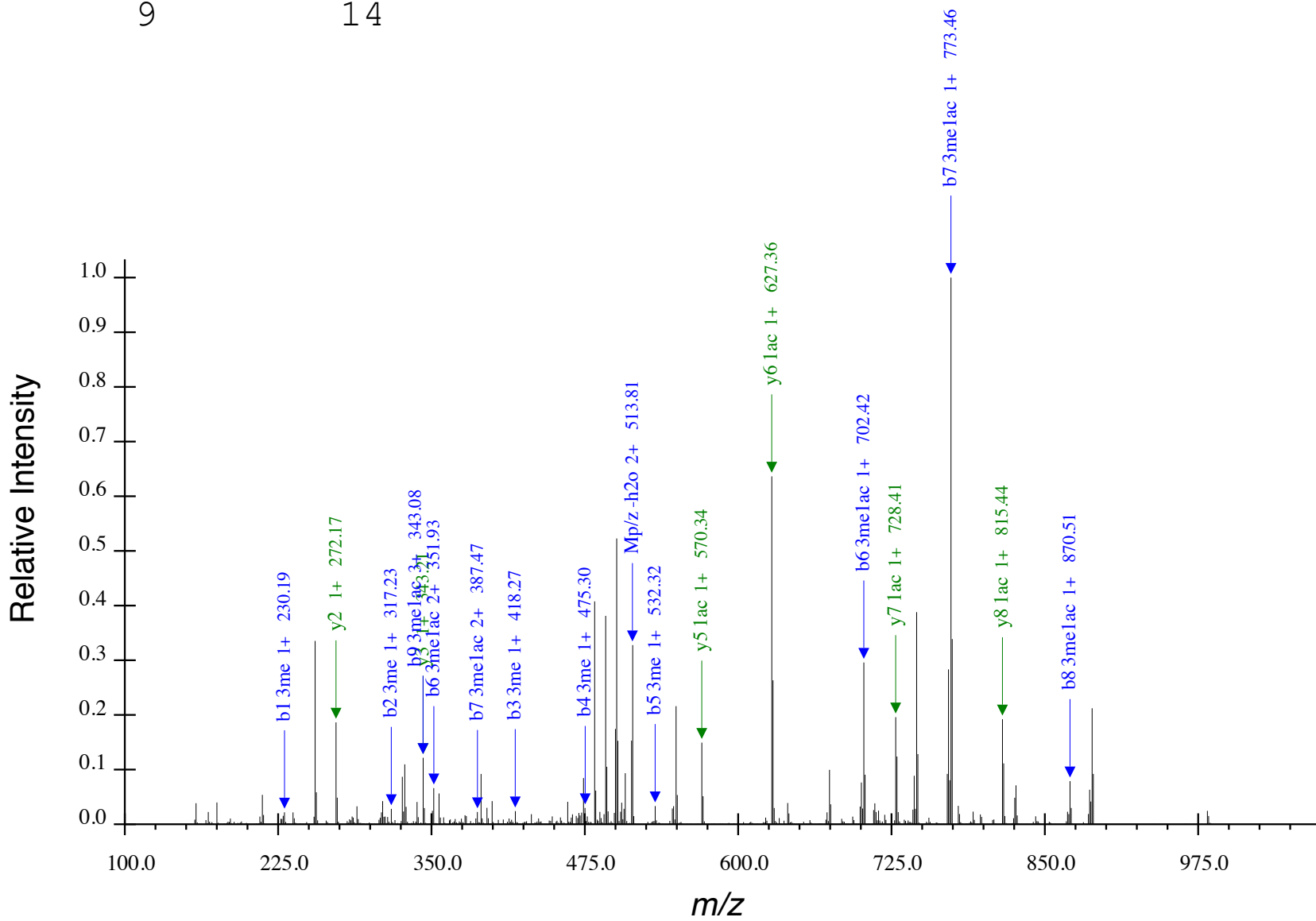
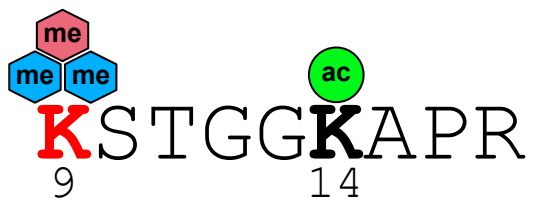
**Figure S9** | Annotated CID tandem mass spectrum for H3K9me2(CH<sub>3</sub>,CD<sub>3</sub>)K14ac, z = 2, Mp/z = 515.81. N-terminus has been propionylated (+56.0627 Da).



x4

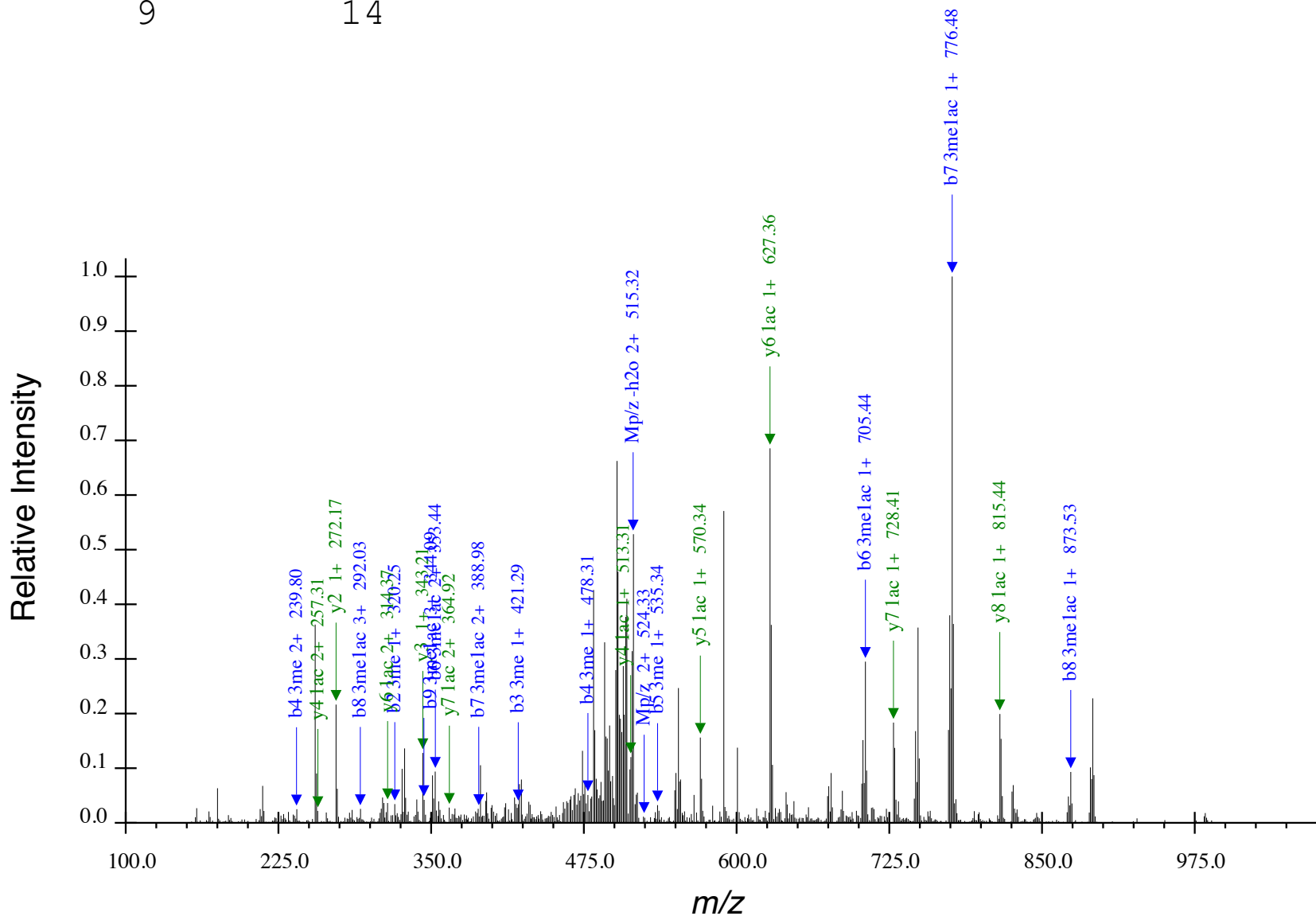
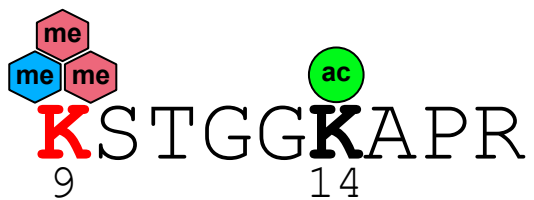


**Figure S10** | Annotated CID tandem mass spectrum for H3K9me2(2CD<sub>3</sub>)K14ac, z = 2, Mp/z = 517.31. N-terminus has been propionylated (+56.0627 Da).

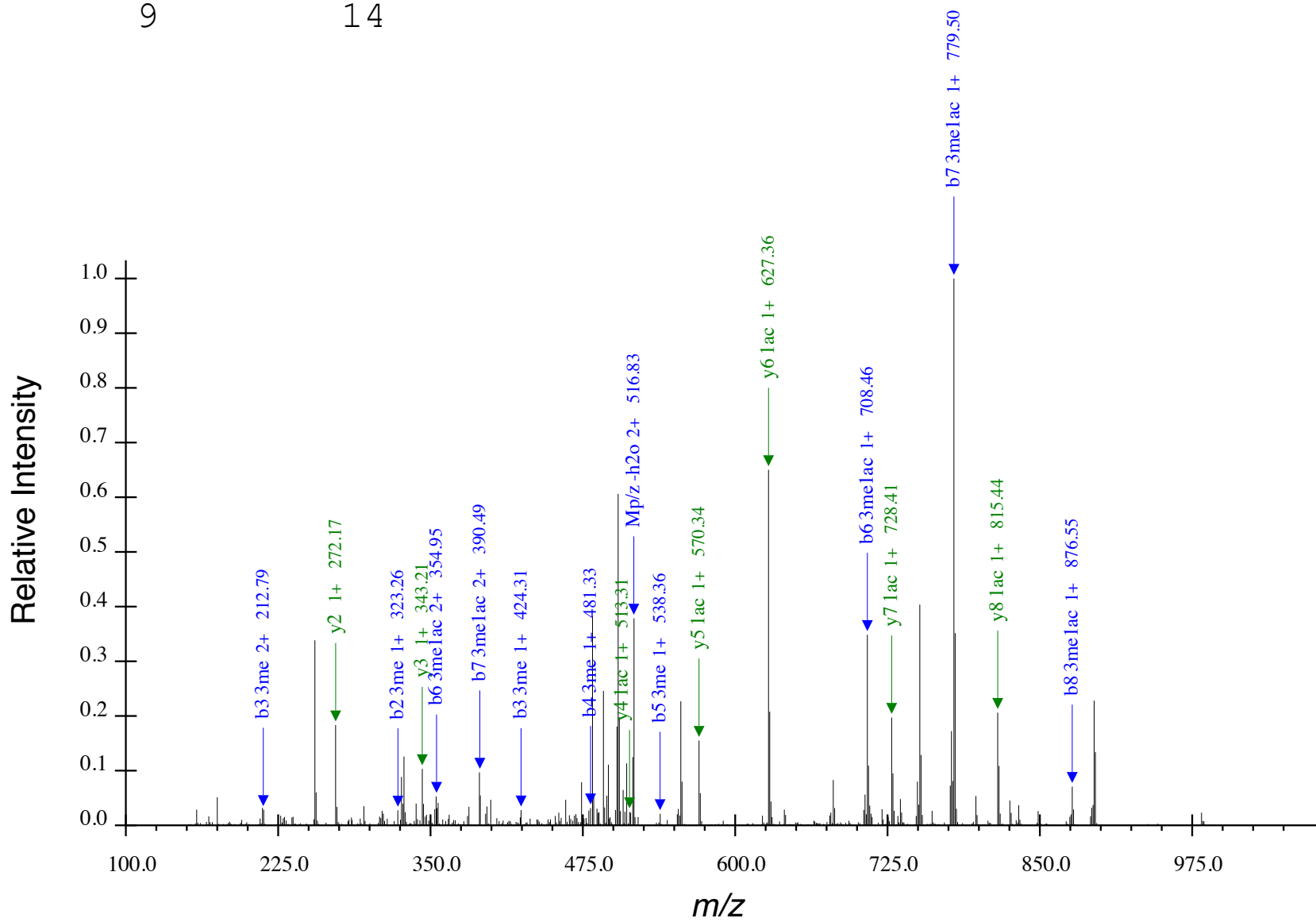
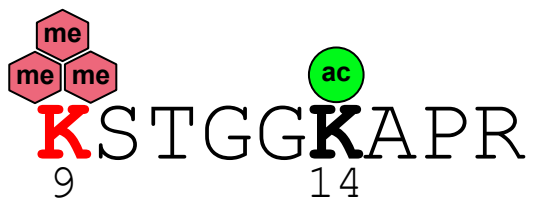


**Figure S11** | Annotated CID tandem mass spectrum for H3K9me3(2CH<sub>3</sub>,CD<sub>3</sub>)K14ac, z = 2, Mp/z = 522.82. N-terminus has been propionylated (+56.0627 Da).

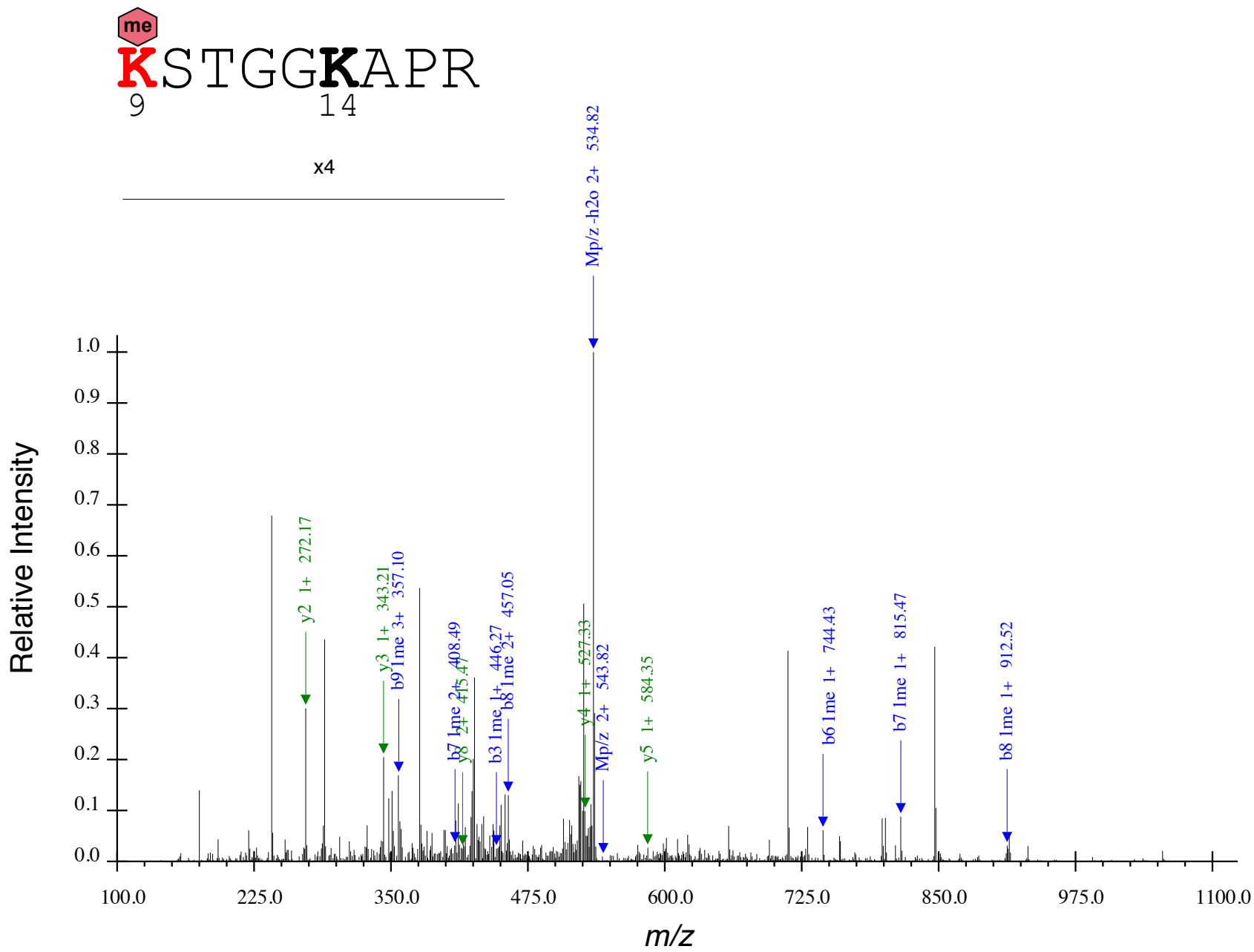




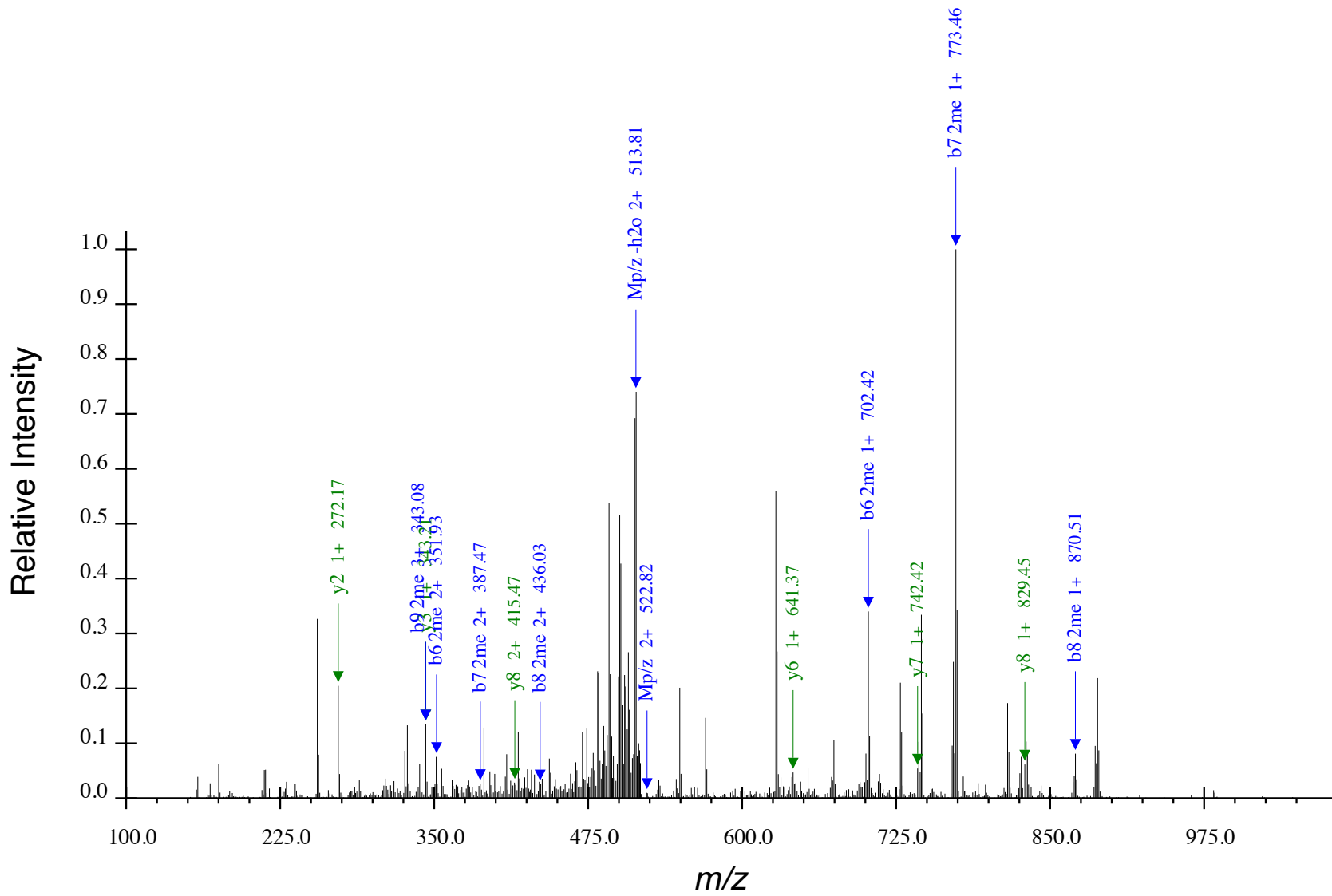
**Figure S12** | Annotated CID tandem mass spectrum for H3K9me3(CH<sub>3</sub>,2CD<sub>3</sub>)K14ac,  $z = 2$ ,  $Mp/z = 524.33$ . N-terminus has been propionylated (+56.0627 Da).



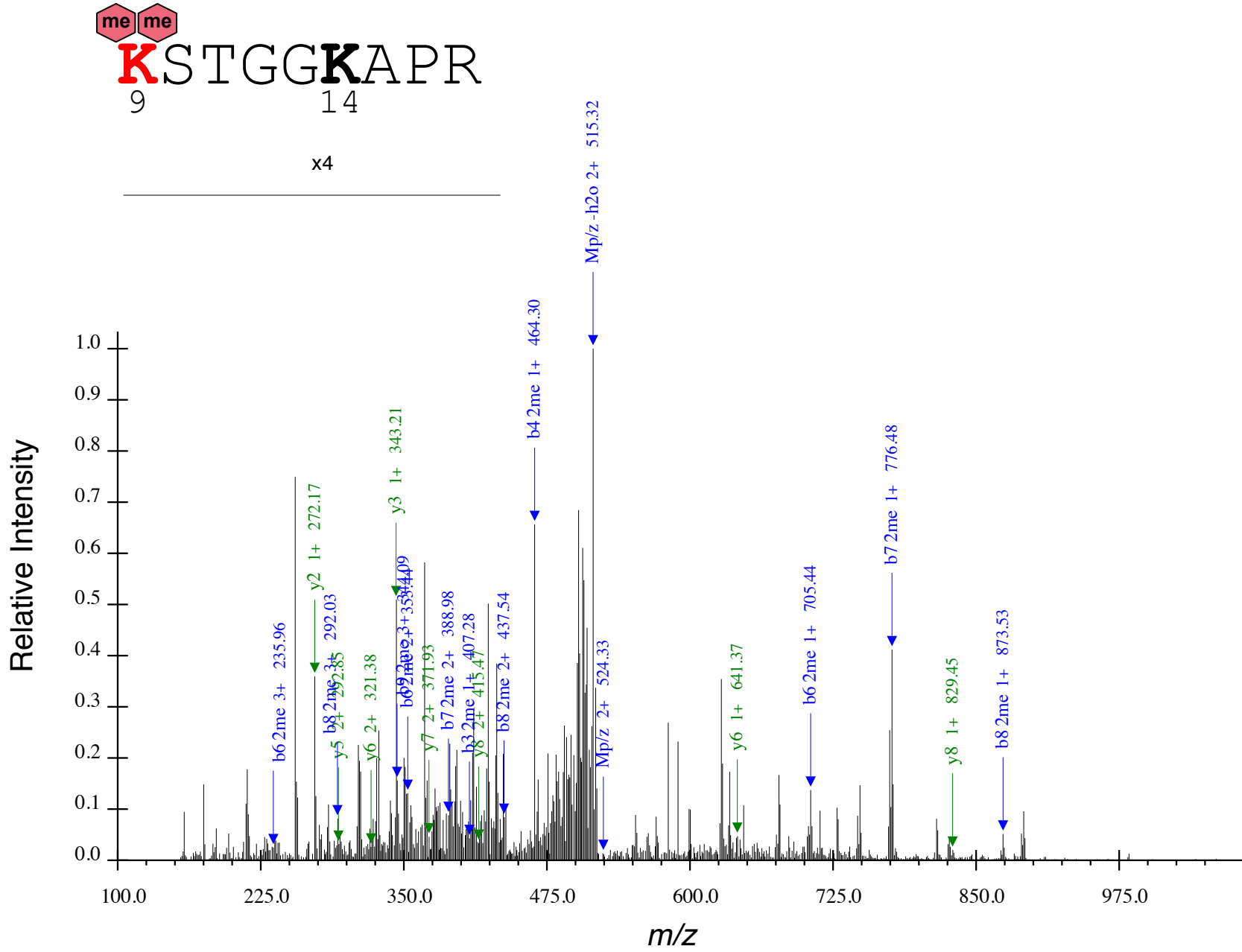
**Figure S13** | Annotated CID tandem mass spectrum for H3K9me3(3CD<sub>3</sub>)K14ac, z = 2, Mp/z = 525.84. N-terminus has been propionylated (+56.0627 Da).



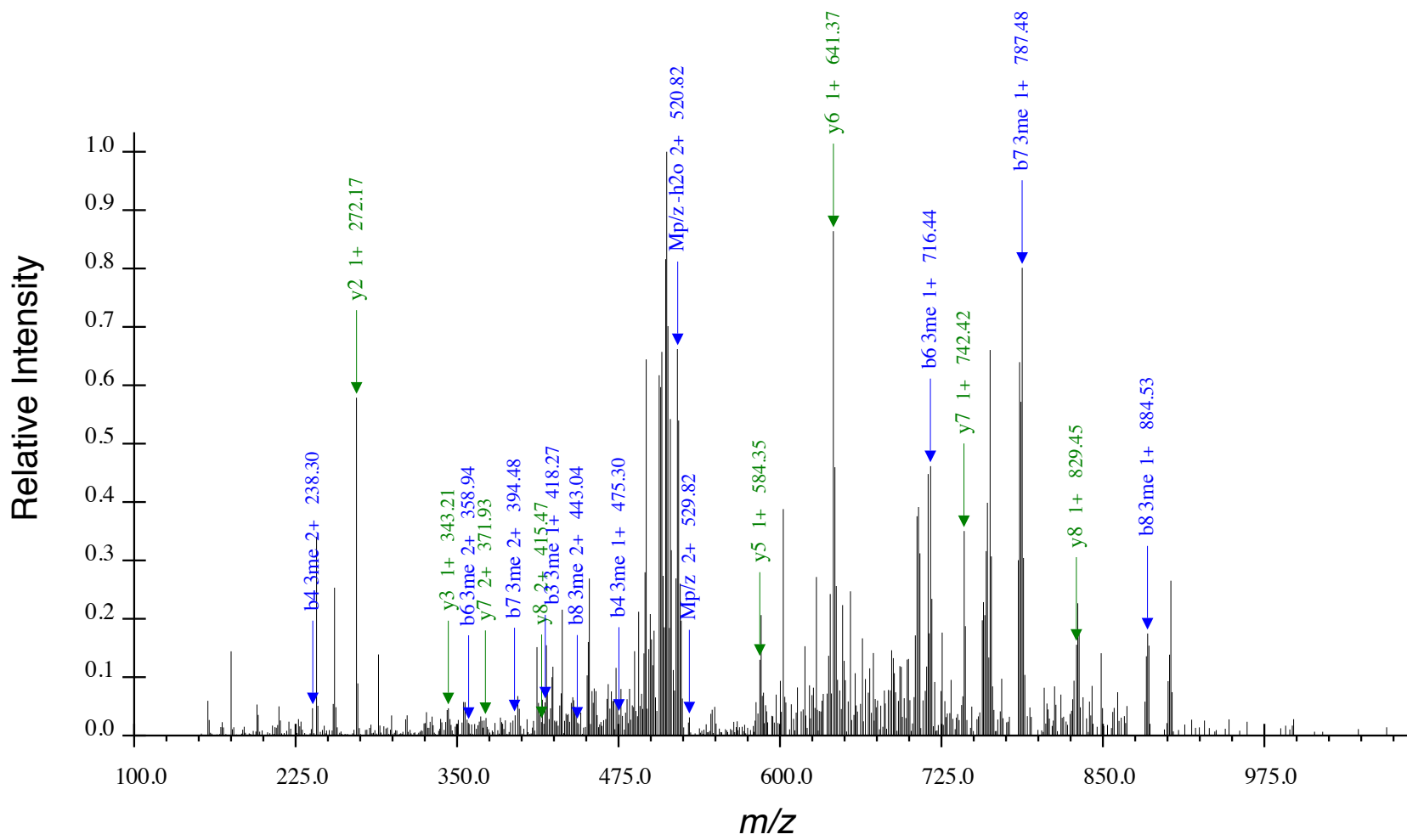
**Figure S14** | Annotated CID tandem mass spectrum for H3K9me1(CD<sub>3</sub>)K14un,  $z = 2$ ,  $Mp/z = 543.82$ . N-terminus, K9 and K14 have been propionylated (+56.0627 Da).



**Figure S15** | Annotated CID tandem mass spectrum for H3K9me2(CH<sub>3</sub>,CD<sub>3</sub>)K14un, z = 2, Mp/z = 522.82. N-terminus and K14 have been propionylated (+56.0627 Da).

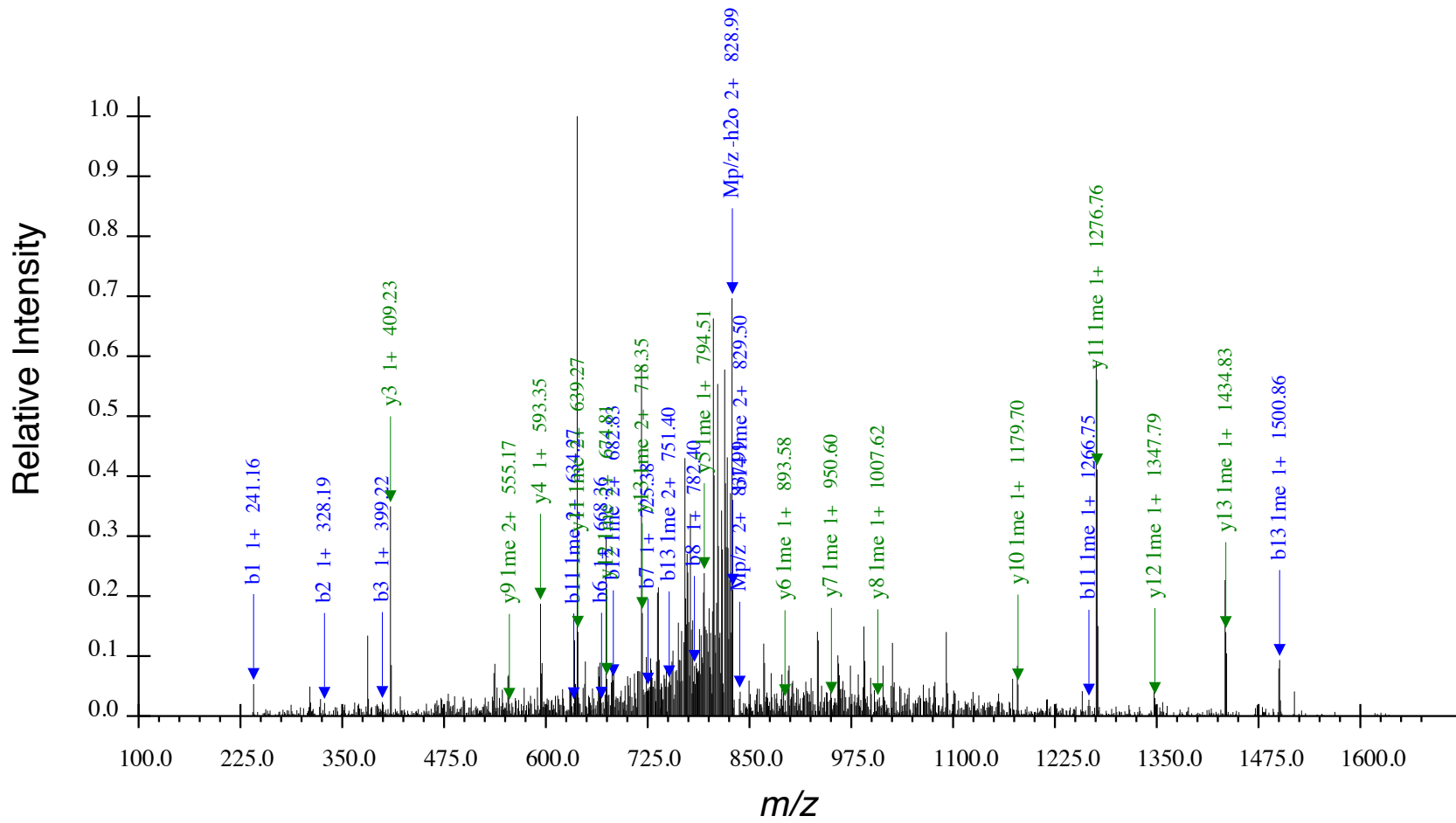


**Figure S16** | Annotated CID tandem mass spectrum for H3K9me2(2CD<sub>3</sub>)K14un, z = 2, Mp/z = 524.33. N-terminus and K14 have been propionylated (+56.0627 Da).



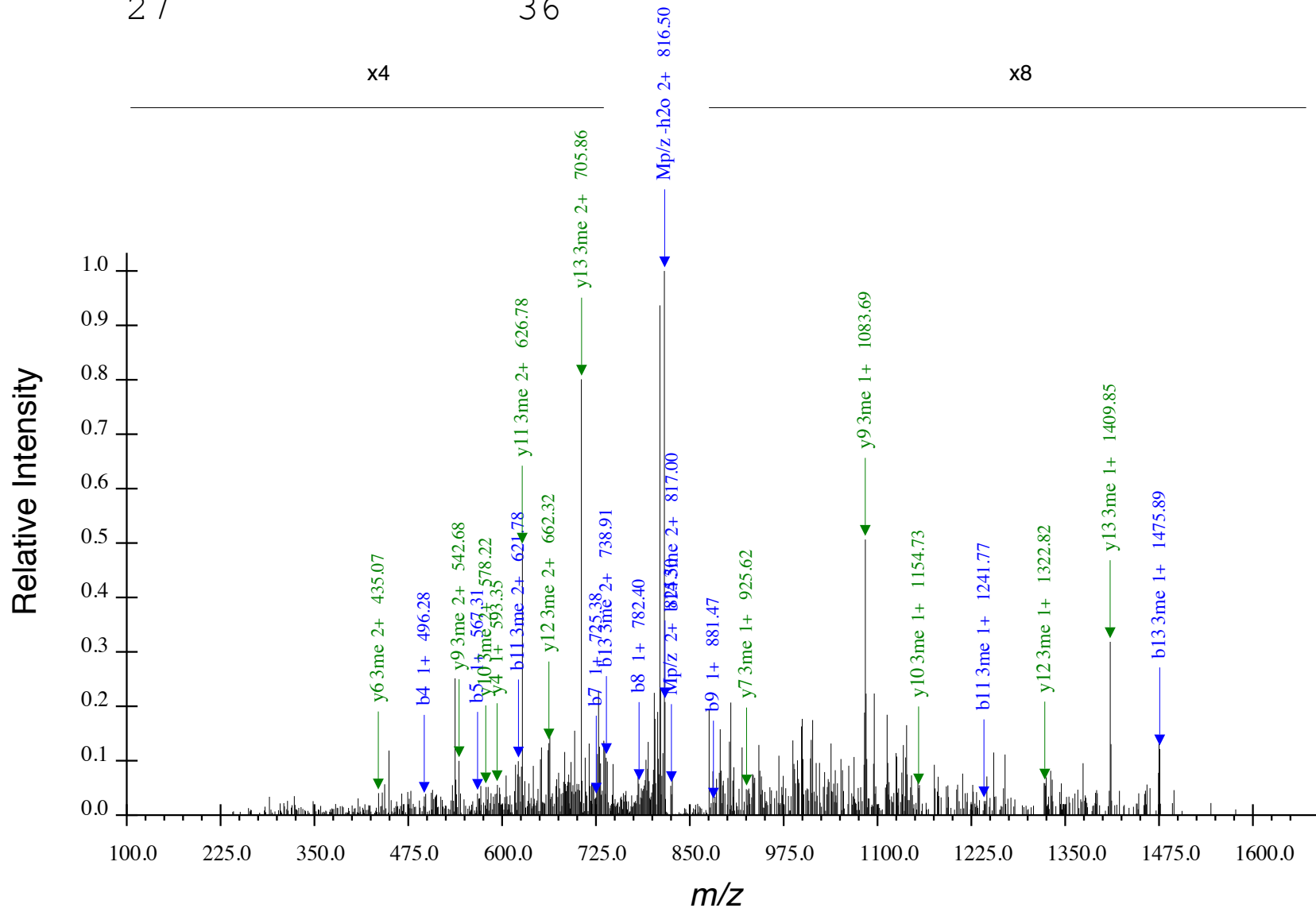
**Figure S17** | Annotated CID tandem mass spectrum for H3K9me3(2CH<sub>3</sub>,CD<sub>3</sub>)K14un, z = 2, Mp/z = 529.82. N-terminus and K14 have been propionylated (+56.0627 Da).

KSAPATGGVKKPHR  
 27 me 36



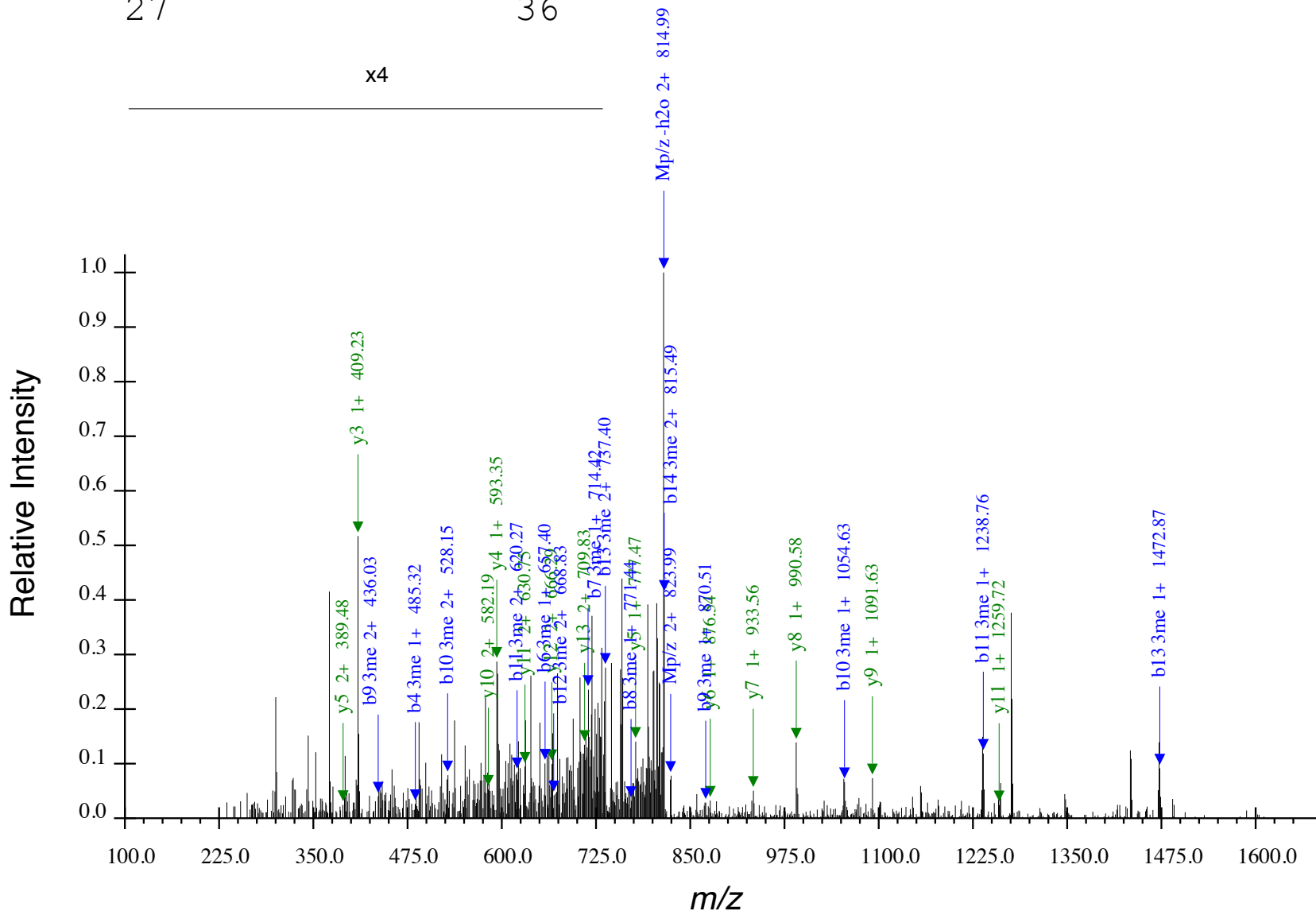
**Figure S18** | Annotated CID tandem mass spectrum for H3K27unK36me1(CD<sub>3</sub>), z = 2, Mp/z = 837.99. N-terminus, K27, K36 and K37 have been propionylated (+56.0627 Da).

**K**SAPATGGV**K**KPHR
   
 27 36



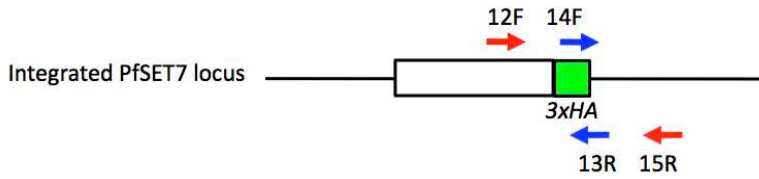
**Figure S19** | Annotated CID tandem mass spectrum for H3K27unK36me3(CH<sub>3</sub>,2CD<sub>3</sub>), z = 2, Mp/z = 825.50. N-terminus, K27 and K37 have been propionylated (+56.0627 Da).



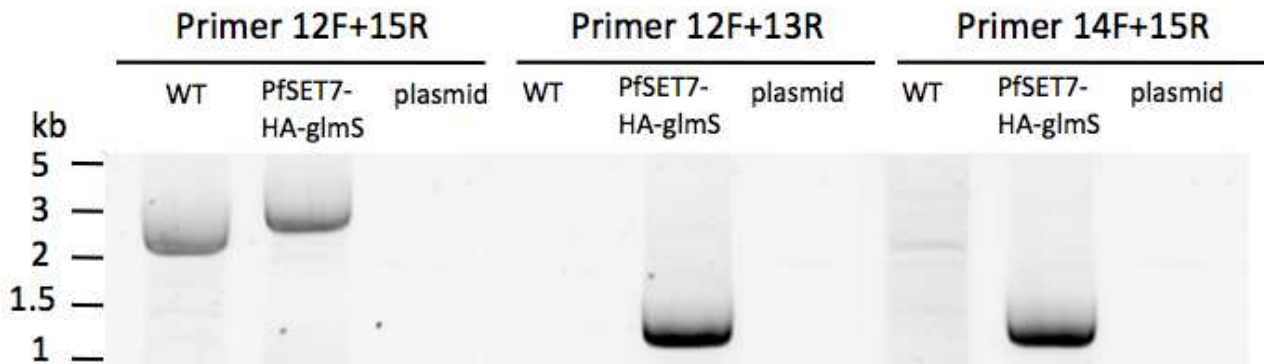


**Figure S20** | Annotated CID tandem mass spectrum for H3K27me3(2CH<sub>3</sub>,CD<sub>3</sub>)K36un, z = 2, Mp/z = 823.99. N-terminus, K36 and K37 have been propionylated (+56.0627 Da).

**a**



**b**



**c**

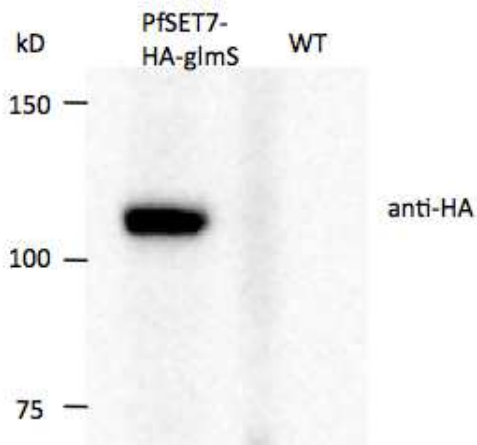
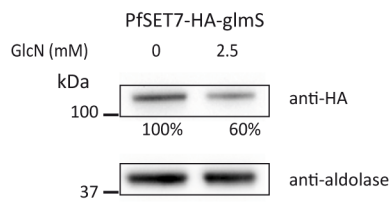


Figure S21 | **Confirmation of PfSET7-HA-glmS integration.** (a) Schematic of the PCR verification of PfSET7-HA-glmS integration. White rectangles represent the wild PfSET7 locus and green box indicate the 3xHA tag. Red arrows represent the primers used to identify the upstream and downstream of homologous region. Blue arrows represent the primers located in HA tag sequence. (b) Agarose gel electrophoresis of the upstream, downstream and insertion confirmation PCRs. Left panel: Either wild-type (2248 bp) or the integrated allele (2615 bp) are visualized after amplification with 12F+15R. Middle and right panel: note the presence of integration after amplification with 12F+13R (1230 bp) and 14F+15R (1247 bp). (c) Western blot of PfSET7 in transgenic PfSET7-HA-glmS line and wild-type parasite lines. The anti-HA antibody recognizes a band of >100 kDa (PfSET7-HA predicted MW: 98 kDa).

a



b

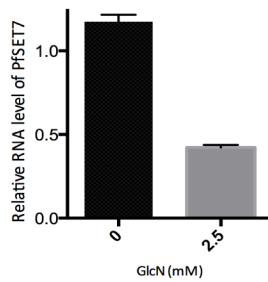


Figure S22 | **Generation of a PfSET7 knock-down line after 2 generations of glucosamine treatment.** (a) Western blot analysis of treated PfSET7-HA-glmS parasites shows 40% decrease in protein levels compared to the control. The PfSET7 expression was normalized to the protein aldolase. The blot image presented has been cropped to the area used for densitometry. (b) Quantitative PCR analysis showing 60% reduction of PfSET7 mRNA levels in treated parasites. Seryl-tRNA synthetase (PF07\_0073) was used as an endogenous control. The  $\Delta\text{CT}$  for the PfSET7 gene was determined by subtracting the PfSET7 gene CT value from the seryl-tRNA synthetase gene CT value. Copy number was then converted using the formula (copy number =  $2^{-\Delta\text{Ct}}$ ). Data are presented as mean  $\pm$  SD of two independent experiments.

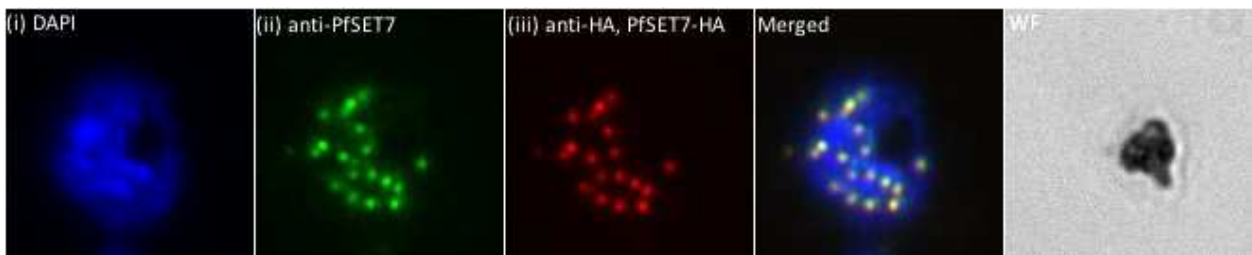


Figure S23 | **Comparison of anti-PfSET7 and anti-HA antibodies in the PfSET7-HA-glmS parasite line.** Both anti-PfSET7 and anti-HA produced a strong punctate staining dots that are consistent with each other. They are equivalent in detecting PfSET7.

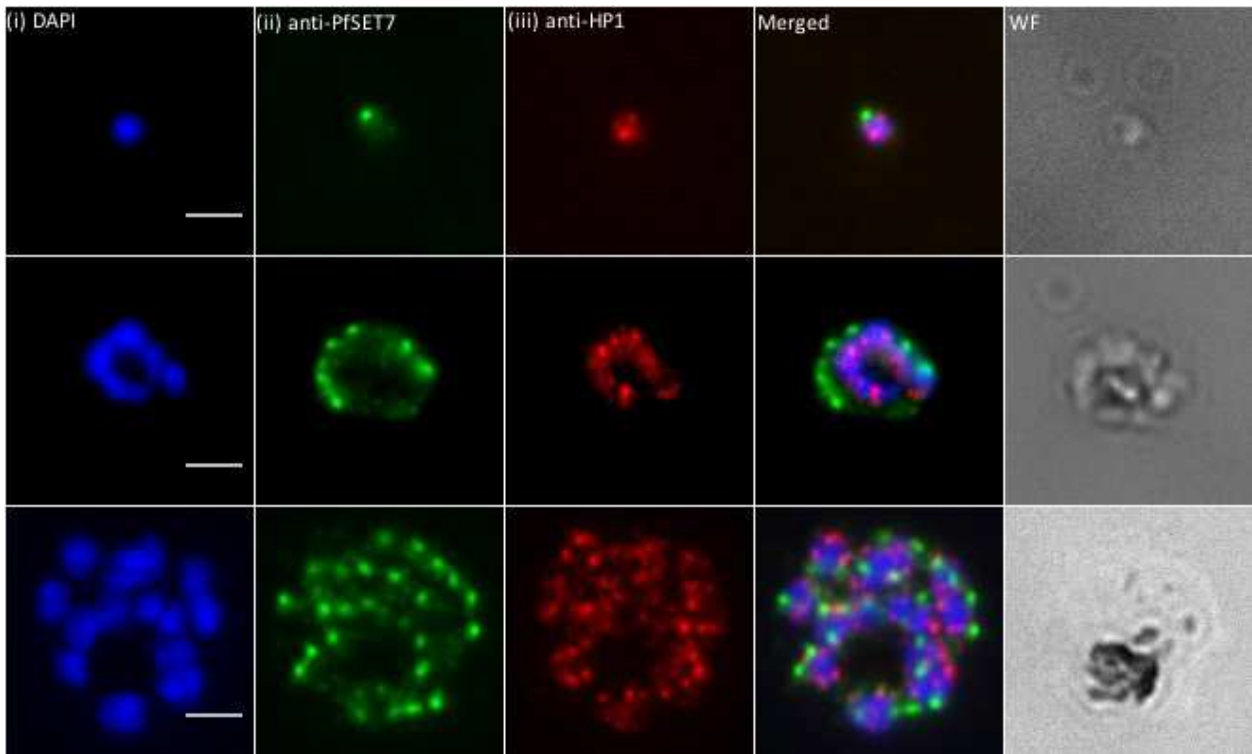


Figure S24 | **Cellular localization of PfSET7 at different blood cell stages.** IF of *P. falciparum* erythrocytic stage rings, trophozoites and schizonts using anti-PfSET7 combined with the nuclear periphery marker PfHP1. Throughout the three asexual blood stages, PfSET7 shows a punctate pattern adjacent to nucleus and distant from nuclear periphery. Scale bar is 2  $\mu$ m.

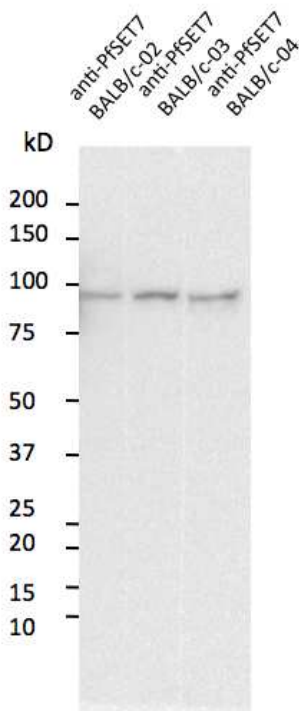


Figure S25 | **Western blot of parasite lysates using the native PfSET7 antibody.**

The anti-PfSET7 antibody from three separate animals all recognize a single band representing the 94 kDa native PfSET7 protein. Identical lanes from a single membrane were cut, probed with the three antibodies individually and then imaged together as oriented on the original membrane.

## 5.2 Article II

A SMYD-type methyltransferase PfSET6 associates with a subset of gene loci of *Plasmodium falciparum* and forms multiple cytoplasmic foci during asexual, sexual blood stage and sporozoite stage

Ding S *et al*, in preparation

### Manuscript Highlight:

- We expressed the recombinant full length PfSET6 and showed that it possesses methyltransferase activity to histone H3 *in vitro*, albeit to lower extent than PfSET7.
- We observed dynamic cellular localization of PfSET6 during parasite growth: nuclear localization in ring stages, a punctuated pattern in the cytoplasm of mature, blood stage parasites and sexual blood stage gametocytes, and discrete nuclear and cytoplasmic foci in salivary gland sporozoites.
- Using chromatin immunoprecipitation and next generation sequencing (ChIP-seq) analysis, we found that PfSET6 associates with a subset of gene loci in ring stages in a manner that correlates with transcriptional repression.

### Individual Contribution:

- I performed western blotting and immunofluorescence microscopy to elucidate the localization of PfSET6 in blood stage parasites.
- I performed ChIP-seq experiments (with anti-PfSET6, anti-PfHP1, anti-histone H3K9me3 and anti-histone H3K4me3 antibodies) and data analysis to identify genome-wide occupancy of PfSET6 in ring stage parasites and interpreting its potential role in epigenetic regulation.
- I generated transgenic parasite lines for establishing conditional gene excision of PfSET6.
- I wrote the manuscript.

**A SMYD-type methyltransferase PfSET6 associates with a subset of gene loci of *Plasmodium falciparum* and forms multiple cytoplasmic foci during asexual, sexual blood stage and sporozoite stage**

Ding Shuai<sup>1,2,3</sup>, Chen Patty B.<sup>1,2,3</sup>, Baumgarten Sebastian<sup>1,2,3</sup>, Vembar Shruthi<sup>1,2,3</sup>, Zhanghi Gigliola<sup>4,5,6</sup>, Mecheri Salah<sup>1,2,3</sup>, Mazier Dominique<sup>4,5,6</sup> and Scherf Artur<sup>1,2,3</sup>\*

<sup>1</sup>*Unité Biologie des Interactions Hôte-Parasite, Département de Parasites et Insectes Vecteurs, Institut Pasteur, Paris 75015, France;*

<sup>2</sup>*CNRS, ERL 9195, Paris 75015, France;*

<sup>3</sup>*INSERM, Unit U1201, Paris 75015, France;*

<sup>4</sup>*Sorbonne Universités, Université Pierre et Marie Curie-Paris 6, UMR S945 Paris, F-75013 France*

<sup>5</sup>*Institut National de la Santé et de la Recherche Médicale U945, F-75013 Paris, France*

<sup>6</sup>*AP-HP, Groupe Hospitalier Pitié-Salpêtrière, Service Parasitologie-Mycologie, Paris, F-75013, France*

- *Corresponding author: Scherf Artur email: artur.scherf@pasteur.fr*

## Abstract

Epigenetic control via reversible histone methylation regulates transcriptional activation of *Plasmodium falciparum* blood stage development activation including sexual stage commitment and default silencing of clonally variant virulence gene families. *P. falciparum* encodes ten predicted histone lysine methyltransferases (HKMTs) with a predicted catalytic SET (Su(var)3-9 and 'Enhancer of zeste') protein domain. HKMTs represent attractive drug targets but only PfSET2, PfSET7 and PfSET10 have been characterized to a limited extent in malaria parasites. PfSET6 was shown refractory to genetic disruption in blood stages, indicating an essential role in parasite development. Here we have expressed the recombinant full length PfSET6 in insect cells. Purified PfSET6 showed methyltransferase activity to histone H3. Immunofluorescence (IF) assays using antibodies raised against recombinant PfSET6 revealed a nuclear localization pattern in ring stage and a punctuated pattern in the cytoplasm of mature stage parasites. This cellular distribution is supported by cellular fractionation and western blot analysis, which showed that anti-PfSET6 detect a band of the expected size (60 kDa) in nuclear and cytoplasmic fractions of parasite cell lysates. In addition, PfSET6 is enriched within foci-like structures in the cytoplasm of sexual stage gametocytes and forms discrete foci in the cytoplasm and nucleus of salivary gland sporozoites. Genome-wide chromatin immunoprecipitation assays of ring stage parasites, when PfSET6 is nuclear, showed an enrichment of PfSET6 within the promoter and gene body of a subset of genes including select clonally variant *var* genes in subtelomeric and chromosome-internal regions. In summary, our work suggests that PfSET6 may have a cytoplasmic role in addition to the predicted nuclear function during the life cycle. Given that PfSET6 apparently targets multiple substrates during the parasite life cycle, this could explain its essentiality during blood stage growth.

## INTRODUCTION

Despite major efforts, malaria remains the most significant human parasitic disease of humans and claims approximately 0.6 million lives per year<sup>1</sup>. Global efforts to control malaria have been hampered by the development of drug resistance in parasites and insecticide resistance in mosquitoes, and the lack of an effective vaccine. The human malaria parasites, such as *Plasmodium falciparum* and *P. vivax*, undergo a complex life cycle within the mosquito vector and the human host. Development within the human host involves a pre-erythrocytic liver stage, asexual intra-erythrocytic stages responsible for all disease pathology, and sexual stages permitting disease transmission. Stage transitions in the multistage life cycle are highly regulated and orchestrated by distinct programs of gene activation. Of the transcriptional waves in *P. falciparum*, the intra-erythrocytic developmental cycle (IDC) is the best studied. It revealed that the ring, trophozoite and schizont stages of the IDC have distinct transcriptomic profiles resulting in a cascade of gene expression, with every mRNA reaching peak levels at only one timepoint of this 48 hour cycle<sup>2,3</sup>.

A variety of epigenetic and post-transcriptional mechanisms have been found to contribute to orchestrate gene activation or repression in *P. falciparum*<sup>4-6</sup>. Genome-wide analysis of histone post-translational modifications (PTMs) provide the evidence linking histone methylation and acetylation to a gene's transcriptional status<sup>7</sup>, particularly for the well-known clonally variant gene family *var*, which is responsible for antigenic variation in *P. falciparum* blood stage pathogenesis. Acetylation and methylation of N-terminal histone lysines are the two most common PTMs with distinct distributions along both euchromatin and heterochromatin in *P. falciparum*<sup>8,9</sup>. The overarching observation is that the euchromatic marks H3K9ac (acetylation of histone H3 lysine 9) and H3K4me3 (tri-methylation of histone H3 lysine 4) are widely distributed across the genome, including in intergenic regions, and are enriched at promoters of actively transcribed genes<sup>7,10,11</sup>. In contrast, the heterochromatin marker H3K9me3 (tri-methylation of histone H3 lysine 9) is distributed in broad domains over silenced regions<sup>10-12</sup>. Specifically, H3K9me3 is associated with multigene families such as *var*, *rifin* and *stevor*, and genes encoding a parasite-induced erythrocyte permeation pathway and with regulating the commitment to sexual differentiation<sup>11</sup>. The therapeutic implication from the above mentioned findings is that the rational drug design against histone modification machineries could potentially interfere with all life cycle stage. Preliminary studies indicated that a known anti-cancer drug BIX-01294, which is a diaminoquinazoline that targets the human SET domain



protein G9a, and one of its analogues TM2-115 caused rapid and specific parasite death with a concomitant reduction in parasite histone methylation levels<sup>13-15</sup>. Of interest, the same drugs also affected the quiescent state of *Plasmodium* hypnozoites by waking them thus allowing their develop into mature dividing forms<sup>16</sup>. These studies established histone methylation can be an attractive direction in antimalarial drug discovery. Moreover, given that particular histone methylation states are linked to important biological processes of *P. falciparum*, specific inhibitors could be potentially used to explore the fascinating biology of this parasite.

As the most essential components of histone methylation mark writers, protein methyl transferase (PKMTs) catalyze the mono-, di- or tri-methylation of lysine residues or the mono- or di-methylation of arginine residues (PRMTs). The Su(var)3-9-Enhancer of zeste-Trithorax (SET) domain was demonstrated to be the catalytic domain for lysine methylation, consisting of binding sites for histone substrate and methyl donor molecule S-(5'-Adenosyl)-L-methionine (SAM). In *P. falciparum*, ten SET domain-containing HKMTs have been bioinformatically predicted<sup>17</sup>. Six of the ten PfSET genes were refractory to knock-out attempts via double crossover homologous recombination, suggesting that these genes play a role in development of asexual blood stage parasites<sup>18</sup>. Until now, three PfSET proteins have been characterized in *P. falciparum* to a certain limit. PfSET10, a putative H3K4 methyltransferase implicated in clonally variant gene expression, localizes to a single spot within the nuclear periphery of all asexual blood stages, coinciding with the perinuclear poised *var* expression site<sup>19,20</sup>. PfSET2, also called PfSETvs, is predicted to mediate H3K36-specific methylation. Disruption of PfSETvs causes loss of H3K36me3 at the transcription start site of all silent *var* genes and leads to their de-repression, indicating it plays a pivotal role in broadly silencing *var* genes<sup>18</sup>. PfSET7, the only *P. falciparum* SET-domain protein that has been purified and enzymatically characterized, shows robust methyltransferase activity *in vitro* and localizes to apical organelles in schizont stages of the IDC<sup>21</sup>. Moreover, PfSET7 methylates histone H3K4 and H3K9 in the presence of pre-existing acetylated histone H3 lysine 14 (H3K14ac). However, because gene knockout and mRNA knockdown of PfSET7 was not achieved, its *in vivo* function remains unknown.

In this report we produced recombinant full length PfSET6 and investigated the biological function of PfSET6 in different life cycle stages. We showed that the cellular localization of PfSET6 changes during the IDC and partially overlaps with *var* repression centres, putative mRNA-containing granules and apical organelles. This points to non-histone PfSET6 targets

outside the nucleus. Moreover, genome-wide occupancy analyses indicated that PfSET6 associates with particular genome regions in a manner that correlates with transcriptional repression.

## RESULTS

### Recombinant PfSET6 possesses histone methyltransferase activity

The PfSET6 open reading frame of 1,530 nucleotides (PF3D7\_1355300) encodes a 509 amino acid protein containing a SET domain (codons 1-250) and Myeloid-Nervy-DEAF1 (MYND) domain (codons 62-100) at the amino terminus (Figure 1a), suggesting that PfSET6 belongs to the SMYD3 family of methyltransferases. Notably, the conserved SET domain is split into two parts by the MYND domain (Figure 1a). Phylogenetic analysis of PfSET6 orthologues in Apicomplexans using EggNOG indicated that while *Plasmodium* SET6 proteins are well conserved with 85.7% identity, they are genetically divergent from *Toxoplasma* and *Cryptosporidium* SET proteins (Supplementary Figure 1).

Next, to determine whether PfSET6 possesses methyltransferase activity *in vitro*, we expressed recombinant full-length PfSET6 fused to a C-terminal 2x-Streptavidin tag using a baculovirus-based Sf9 insect cell expression system (Figure 1a). Due to the high AT-content of the set6 gene (76.1%), for heterologous protein expression, we utilized a codon-optimized version of set6. The resulting protein was purified using StrepTrap HP columns and resolved using denaturing gel electrophoresis. When visualized by Bio-Rad Stain-Free detection, the gel revealed two migrating bands, one of approximately 60kDa, and the other band which was consistently observed at 50kDa and has been validated as a recombinant PfSET6 by western blotting using anti-Streptavidin antibodies (Figure 1b).

Biochemical characterization of methyltransferase activity of recombinant PfSET6 was performed using a radioactive assay described earlier<sup>21</sup>. We used purified histones to measure the incorporation of radiolabeled methyl groups from a S-adenosyl methionine (SAM) donor over time. We observed that recombinant PfSET6 exhibited histone methyltransferase activity that was 3-fold higher than the control (Figure 1c). However, this was significantly lower than PfSET7 or human G9a methyltransferase activity<sup>21</sup>. It has previously been shown that SMYD3 proteins optimally function in the presence of co-factors such as chaperone hsp90. Therefore, we repeated the methyltransferase assay in the presence of total parasite lysates.

However, we did not observe an increase in PfSET6 activity over time indicating that such interaction probably has less reciprocal effect on methyltransferase activity and nuclear chaperone activity compared with *in vivo* (data not shown).

To explore the enzyme specificity, we performed a second independent analysis of PfSET6 activity. The methyltransferase reaction was resolved by denaturing gel electrophoresis and radiolabeled proteins detected using autoradiography. As shown in Figure 1d, a single methylated band at 17kDa corresponding to histone H3 was apparent in reactions containing both PfSET6 and nucleosomes (lane 1). In the absence of nucleosomes or in the presence of an unrelated substrate BSA, there was no methyl transfer, suggesting that PfSET6 is indeed a histone H3 methyltransferase.

### **PfSET6 localizes to the nucleus and cytoplasm of asexual blood stage parasites in a dynamic and punctate pattern**

Anti-PfSET6 antibodies generated in mice against recombinant full-length of PfSET6 were used to determine cellular localization of PfSET6 during blood stage growth. First, the specificity of the anti-PfSET6 antibodies was assessed using western blotting of either parasite lysates (Supplementary Figure 2, left lane) or purified PfSET6 (right lane). These results showed that the anti-PfSET6 antibodies recognize a single band representing the 60kDa native PfSET6 protein in parasite lysates.

Second, to determine the sub-cellular distribution of PfSET6, whole parasites extracts were biochemically separated into nuclear and cytoplasmic fractions, and examined by western blotting with anti-PfSET6 antibodies. Antibodies against PfAldolase or histone H3 were used as controls for cytoplasmic and nuclear fraction, respectively. Figure 2a shows that PfSET6 is present in both cytoplasmic and nuclear fractions of extracts prepared from an asynchronous parasite culture, suggesting that PfSET6 may methylates non-histone substrates in the cytoplasm and have unique roles outside of the nucleus.

To assess the intracellular localization of PfSET6, *P. falciparum* asexual erythrocytic stage parasites were analyzed by immunofluorescence microscopy (IF). As shown in Figure 2b, PfSET6 exhibits a predominantly perinuclear distribution in ring stages, with this pattern changing in trophozoites and schizonts, where the PfSET6 signal is more diffuse throughout the parasite cytoplasm, albeit with distinct foci of enrichment. Notably, nuclear PfSET6

staining is maintained in trophozoites and schizonts. Taken together, these results suggested that PfSET6 cellular localization is dynamic during the IDC.

To determine whether the punctate pattern of PfSET6 in ring stages coincided with the well-characterized epigenetic regulator and marker of facultative heterochromatin, PfHP1, co-IF assays were performed with anti-PfSET6 and anti-PfHP1 antibodies. In rings and schizonts, PfSET6 staining partially overlapped with PfHP1 foci at the nuclear periphery (Figure 3a), implying a role for PfSET6 in gene regulation. PfSET6 staining in schizonts was also compared to other known markers of cytoplasmic organelles: BiP for the endoplasmic reticulum, ERD2 for the *cis*-golgi, CLAG3.2 for rhoptries and EBA175 for micronemes. As shown in Figure 3b, PfSET6 staining partially overlapped with BiP and EBA175, but not the other marker analyzed. This indicated that PfSET6 involved in some critical cellular processes, such as protein synthesis, and some important biological process like invasion. Finally, because the punctate cytoplasmic pattern of PfSET6 was reminiscent of mRNA-containing ribonucleoprotein (mRNA) granules, we combined PfSET6 staining with mRNP markers such as members of the DNA/RNA-binding PfAlba family. As shown in Figure 3c, PfSET6 signals partially overlapped with PfAlba1, 2 or 4 staining, suggesting the role of protein methyltransferases in post-transcriptional gene regulation or mRNA regulation or translational repression has been described previously<sup>22,23</sup>.

### **PfSET6 localizes to distinct foci in different sub-compartments of *P. falciparum* transmission stages**

To determine if PfSET6 continues to be expressed after the parasite differentiates into sexual erythrocytic stages, we performed co-IF assays with anti-Pfg377 antibodies. The reference marker Pfg377 is a gametocyte antigen associated with osmiophilic bodies, membrane-bound vesicles that are found in female gametocytes and become progressively more abundant as gametocyte reaches full maturity; that are believed to be involved in gamete egress from the host cell<sup>24</sup>. Figure 4a shows that in stage IV and V gametocytes, PfSET6 predominantly localized at the parasite membrane, with a low dispersed signal in the cytoplasm, which partially overlapped with Pfg377. Additionally, we observed that PfSET6 localizes in discrete cytoplasmic foci and co-localizes with the nucleus of salivary gland sporozoites (Figure 4b). These combined IF suggest that PfSET6 localizes to different cellular sub-compartments during parasite development in a dynamic manner, indicating that it may methylate both histone and non-histone substrates to achieve its biological function.

## Identification of PfSET6 genome-wide occupancy

Given that PfSET6 is predicted to be a HKMT and that it shows a nuclear localization pattern that partially coincides with an epigenetic regulator, we next asked whether PfSET6 could play a role in transcriptional regulation. We analyzed the genome-wide distribution of PfSET6 in ring stages by performing chromatin-immunoprecipitation with anti-PfSET6 antibodies (Supplementary Figure S3) and next generation sequencing (ChIP-seq). As shown in the circular genome-wide map of Figure 5a, we observed that PfSET6 localized to all chromosomes, with no apparent enrichment at chromosomal ends. Moreover, when we performed peak enrichment analysis using MACS2, we found that PfSET6 occupied hundreds genomic positions, near promoters, within gene bodies and within intergenic regions. These corresponded to a total of 94 genes, reflecting less than 2% of the parasite's coding genome (supplementary Table1).

Specifically, PfSET6 bound to select subtelomeric *var* genes on chromosome 1, 2, 3, 5, 6, 7, 9 and 13, as well as chromosome-central *var* gene clusters in chromosome 4. Notably, PfSET6 occupancy was not restricted to *var* loci, but also included other clonally variant gene families such as *rifin* and *stevor*. Other PfSET6-bound genes encode exported proteins (of the *epf1*, *epf3*, *epf4* families), liver stage antigen (*lsa1*), the gametocyte-specific protein (*Pf11-1*), histone phosphosite binding protein (*14-3-3I*), heterochromatin protein (*hp1*), DNA/RNA-binding protein (*alba2*), ribosomal RNA loci and hypothetical conserved proteins.

Next, we asked whether PfSET6 chromosomal occupancy coincided with known epigenetic regulators (i.e., PfHP1) and histone post-translational modifications (i.e., H3K9me3, a repressive mark, and H3K4me3, the predicted modification of PfSET6). On the gross genomic level, we did not observe a significant overlap between PfSET6 distribution and the distribution of any of these marks, as shown in Figure 6a. However, in select cases, PfSET6 occupancy coincided with the repressive chromatin markers PfHP1 and H3K9me3 (Figure 6b), validating the IF observations of Figure 3a. Notably, there was no correlation between PfSET6 and H3K4me3 distribution indicating that they are functionally unrelated.

## DISCUSSION

Histone methylation has been shown to play an important role in the control of general transcriptional regulation, monoallelic expression of virulence genes, and in the commitment to transmission stage development in *P. falciparum*<sup>6</sup>. These important biological processes are still ill-defined. SET domain proteins have been associated to histone methyltransferase activity in malaria parasites, however the biological role of most *plasmodial* SET domain encoding genes remains unknown. Here we report the expression, purification and biological characterization of PfSET6, a protein lysine methyltransferase which was refractory to classical genetic disruption in asexual blood stages. PfSET6 has an unusual domain structure belonging to the SET and MYND domain-containing proteins (SMYD), that are a special class of protein lysine methyltransferases involved in methylation of histones and non-histone targets in other eukaryotes. It has been linked to cardiomyocyte differentiation and cardiac morphogenesis and proliferation of cancer cells<sup>25,26</sup>.

PfSET6 enzymatic activity showed specific histone H3 methylation using radiolabeled SAM as a methyl donor and histones as substrate. Many efforts have been made to produce an KMT active PfSET proteins, and the only example of a recombinant enzyme that shows a very high KMT activity on histones was reported recently for PfSET7<sup>21</sup>. Although we used the same expression and purification protocol, the full length PfSET6 showed weak activity when compared to PfSET7. One explanation for this maybe that the assay conditions used in this study such as PH, temperature, substrate and enzyme concentration are not optimal for PfSET6 function. Changing these conditions could improve the rate of PfSET6 catalysis *in vitro*. Alternatively, in human cancer research, it was shown that SMYD3-mediated di- and trimethylation of histone H3K4 was enhanced by interaction with heat shock protein 90 $\alpha$  (HSP90 $\alpha$ )<sup>27</sup>. This suggests that plasmodial SET6 may require interacting partners such as chaperon for full activity. Pull-down experiments using anti-PfSET6 antibodies may reveal such partners and help to improve its KMT activity, a feature needed to high-throughput small inhibitor screening assay development.

We observed the changes in the localization of PfSET6 during asexual blood stage development: from nuclear in rings to cytoplasmic in trophozoites and schizonts (Figure 2); this is consistent with the dual nuclear and cytoplasmic localization pattern of SMYD3<sup>28</sup>. In the nucleus, PfSET6 staining partially co-localized with the heterochromatin marker PfHP1 (Figure 3a), suggesting

a potential role in transcriptional repression. In the cytoplasm, PfSET6 distribution was not restricted to any specific sub-cellular compartment but was weakly diffused with partial enrichment in the ER, microneme and mRNA granules (Figure 3b, c). This points to a possible role in targeting RNA complexes, that have been shown to be involved in translational delay and repression of invasion genes<sup>22,23</sup>. The IFA staining show that most substrates of PfSET6 reside in the cytoplasm, suggesting that PfSET6 catalytic efficiency may be orders of magnitude higher on non-histone targets as compared to histones. In *P. falciparum*, although considerable research has been devoted to explore the role of histone PTMs, very little is known about more general protein PTMs and how they are regulated. Further pull down studies could reveal PfSET6 interacting partners and shed light on substrates that are methylated.

The PfSET6 ChIP-Seq data in ring stage parasites was unanticipated for a putative H3K4 methyltransferases, but there was no correlation between PfSET6 and H3K4me3 distribution. By contrast, some genes associated with PfSET6 were also enriched in the transcriptional repressor HP1, together with H3K9me3. Notably, the majority of rest PfSET6-demarcated gene code for proteins which are exported into the host erythrocyte or assisted the exportation process and closely related to the pathogenesis. Moreover, the association of PfSET6 with liver stage antigen *lsa1* and gametocyte-specific protein Pf11-1 provide information about its potential involvement in parasites developmental decision-making. It is noteworthy that 14-3-3i coding for the histone phosphor-site binding protein was also shown in the list of gene set enrichment. In malaria parasites how the hierarchy of multiple modifications is regulated and how distinct combinatorial modification patterns are established or maintained remains to be explored. One possible explanation for the silencing-associated occupancies of PfSET6 could be the various mechanisms of SMYD3 mediated. SMYD3 contains a conserved MYND-type domain, which contains a zinc finger motif that facilitating protein–protein interaction, and other DNA-binding consensus motifs. Due to its localization within the catalytic SET domain, the DNA-interaction mediated by MYND could prevent the methyltransferase enzyme catalyzing modifications to histone lysine residue properly<sup>28</sup>.

Altogether, the production of full length recombinant PfSET6 in insect cells demonstrated histone methyltransferase activity. Although several attempts to generate a knock-down transgenic line failed, the rapid and inducible knock-down of PfSET6 are still hoping to be established in the near future, providing answers to questions regarding to its essentiality for parasite survival, impact on genome-wide transcripome and proteome-wide methylation, and

even the mechanisms of action for PfSET6 in parasite development. Given that PfSET6 is expressed in all life cycle stages so far analyzed and may target distinct biology pathways, this makes this protein a very good novel target for drug development against multiple parasite stages.

## REFERENCE

1. WHO. World Malaria Report 2015. *Who*. 2015.
2. Bozdech Z, Llinás M, Pulliam BL, Wong ED, Zhu J, DeRisi JL. The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol*. 2003;1(1):85-100. doi:10.1371/journal.pbio.0000005.
3. Foth BJ, Zhang N, Chaal BK, Sze SK, Preiser PR, Bozdech Z. Quantitative time-course profiling of parasite and host cell proteins in the human malaria parasite *Plasmodium falciparum*. *Mol Cell Proteomics*. 2011;10(8):M110.006411. doi:10.1074/mcp.M110.006411.
4. Doerig C, Rayner JC, Scherf A, Tobin AB. Post-translational protein modifications in malaria parasites. *Nat Rev Microbiol*. 2015;13(3):160-172. doi:10.1038/nrmicro3402.
5. Guizetti J, Scherf A. Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell Microbiol*. 2013. doi:10.1111/cmi.12115.
6. Scherf A, Lopez-Rubio JJ, Riviere L. Antigenic variation in *Plasmodium falciparum*. *Annu Rev Microbiol*. 2008;62(2):445-470. doi:10.1146/annurev.micro.61.080706.093134.
7. Salcedo-Amaya AM, van Driel MA, Alako BT, et al. Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A*. 2009;106(24):9655-9660. doi:10.1073/pnas.0902515106.
8. Trelle MB, Salcedo-Amaya AM, Cohen AM, Stunnenberg HG, Jensen ON. Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite *Plasmodium falciparum*. *J Proteome Res*. 2009;8(7):3439-3450. doi:10.1021/pr9000898.
9. Saraf A, Cervantes S, Bunnik EM, et al. Dynamic and combinatorial landscape of histone modifications during the intra-erythrocytic developmental cycle of the malaria parasite. *J Proteome Res*. 2016;acs.jproteome.6b00366. doi:10.1021/acs.jproteome.6b00366.
10. Lopez-Rubio JJ, Gontijo AM, Nunes MC, Issar N, Hernandez Rivas R, Scherf A. 5' Flanking Region of Var Genes Nucleate Histone Modification Patterns Linked To Phenotypic Inheritance of Virulence Traits in Malaria Parasites. *Mol Microbiol*. 2007;66(6):1296-1305. doi:10.1111/j.1365-2958.2007.06009.x.
11. Lopez-Rubio JJ, Mancio-Silva L, Scherf A. Genome-wide Analysis of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear Repressive Centers in Malaria Parasites. *Cell Host Microbe*. 2009;5(2):179-190. doi:10.1016/j.chom.2008.12.012.
12. Chookajorn T, Dzikowski R, Frank M, et al. Epigenetic memory at malaria virulence genes. *Proc Natl Acad Sci U S A*. 2007;104(3):899-902. doi:10.1073/pnas.0609084103.
13. Malmquist NA, Moss TA, Mecheri S, Scherf A, Fuchter MJ. Small-molecule histone methyltransferase inhibitors display rapid antimalarial activity against all blood stage forms in *Plasmodium falciparum*. *Proc Natl Acad Sci*. 2012;109(41):16708-16713. doi:10.1073/pnas.1205414109.



14. Sundriyal S, Malmquist N a, Caron J, et al. Development of Diaminoquinazoline Histone Lysine Methyltransferase Inhibitors as Potent Blood-Stage Antimalarial Compounds. *ChemMedChem*. 2014;9(10):2360-2373. doi:10.1002/cmdc.201402098.
15. Malmquist NA, Sundriyal S, Caron J, et al. Histone methyltransferase inhibitors are orally bioavailable, fast- Acting molecules with activity against different species causing malaria in humans. *Antimicrob Agents Chemother*. 2015. doi:10.1128/AAC.04419-14.
16. Demb L, Franetich ois, Lorthiois A, et al. Persistence and activation of malaria hypnozoites in long-term primary hepatocyte cultures. 2014. doi:10.1038/nm.3461.
17. Cui LL, Fan Q, Cui LL, Miao J. Histone lysine methyltransferases and demethylases in *Plasmodium falciparum*. *Int J Parasitol*. 2008;38(10):1083-1097. doi:10.1016/j.ijpara.2008.01.002.
18. Jiang L, Mu J, Zhang Q, et al. PfSETvs methylation of histone H3K36 represses virulence genes in *Plasmodium falciparum*. *Nature*. 2013;499. doi:10.1038/nature12361.
19. Volz J, Carvalho TG, Ralph SA, et al. Potential epigenetic regulatory proteins localise to distinct nuclear sub-compartments in *Plasmodium falciparum*. *Int J Parasitol*. 2010. doi:10.1016/j.ijpara.2009.09.002.
20. Volz JC, Bártfai R, Petter M, et al. PfSET10, a *Plasmodium falciparum* Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division. *Cell Host Microbe*. 2012. doi:10.1016/j.chom.2011.11.011.
21. Chen PB, Ding S, Zanghi G, et al. *Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites. *Sci Rep*. 2016;6(February):21802. doi:10.1038/srep21802.
22. Dastidar EG, Dayer G, Holland ZM, et al. Involvement of *Plasmodium falciparum* protein kinase CK2 in the chromatin assembly pathway. *BMC Biol*. 2012;10(1):5. doi:10.1186/1741-7007-10-5.
23. Vembar SS, Macpherson CR, Sismeiro O, Coppée J-Y, Scherf A. The PfAlba1 RNA-binding protein is an important regulator of translational timing in *Plasmodium falciparum* blood stages. *Genome Biol*. 2015;16(1):212. doi:10.1186/s13059-015-0771-5.
24. Sannella AR, Olivieri A, Bertuccini L, et al. Specific tagging of the egress-related osmiophilic bodies in the gametocytes of *Plasmodium falciparum*. *Malar J*. 2012;11(1):88. doi:10.1186/1475-2875-11-88.
25. Gottlieb PD, Pierce S a, Sims RJ, et al. Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. *Nat Genet*. 2002;31(may):25-32. doi:10.1038/ng866.
26. Hamamoto R, Furukawa Y, Morita M, et al. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat Cell Biol*. 2004;6(8):731-740. doi:10.1038/ncb1151.
27. Silva FP, Hamamoto R, Kunizaki M, Tsuge M, Nakamura Y, Furukawa Y. Enhanced methyltransferase activity of SMYD3 by the cleavage of its N-terminal region in human cancer cells. *Oncogene*. 2008;27(19):2686-2692. doi:10.1038/sj.onc.1210929.
28. Mazur PK, Gozani O, Sage J, Reynoird N. Novel insights into the oncogenic function of the SMYD3 lysine methyltransferase. *Transl Cancer Res*. 2016;5(3):330-333. doi:10.21037/tcr.2016.06.26.

## MATERIALS AND METHODS

**Protein expression construct.** Codon optimized full length PfSET6 (PlasmoDB Gene ID PF3D7\_1355300) was synthesized (Genscript), fused to a C-terminal double Strep-tag and ligated into the transfer vector pVL1393. A single potential glycosylation site S325A was removed for possible expression in secretory systems. The SET and SMYD domains of PfSET6 were identified using Prosite (prosite.expasy.org). Clones were verified to be error free and in frame by sequencing (GATC).

**Protein expression and purification.** To identify construct that provided sufficient amounts of Strep-tagged proteins, small-scale expression and purification was performed as described previously<sup>21</sup>. Briefly, recombinant PfSET6 was expressed using a baculoviral expression system following manufacturer's instructions. Proteins were then purified using StrepTrap HP columns (GE Healthcare). Final protein concentration was determined by Bradford Assay (Bio-Rad Protein Assay, Bio-Rad), and samples were aliquoted, flash frozen in liquid nitrogen and stored at -80°C until use.

***In vitro* methyltransferase assay.** Standard reactions were performed in a total volume of 25  $\mu$ l in KMT buffer (50 mM TRIS, pH 8.8, 5 mM MgCl<sub>2</sub>, 4 mM dithiothreitol), with HEK nucleosomes (0.2 mg/ml final) as substrate and 2.5  $\mu$ l of S-[<sup>3</sup>H-methyl]-adenosyl-L-methionine (3H-AdoMet) was added (3.7  $\mu$ M final, 15 Ci/mmol, Perkin-Elmer). PfSET6 enzyme was used at 1  $\mu$ M and reactions were incubated at room temperature for 1 hour. Analysis by fluorography was performed by resolving 20  $\mu$ l of standard HKMT reactions (2  $\mu$ M of PfSET6, 20  $\mu$ g/ml of either BSA or nucleosomes, 2.5  $\mu$ l of methyl<sup>3</sup>H-AdoMet at 15Ci/mmol, in a total volume of 25  $\mu$ l in KMT buffer) and incubated at room temperature for 2 hours. Reactions were separated on 4-15% SDS-PAGE gel, after migration the gel was rinsed in EN3HANCE (Perkin-Elmer) according to manufacturer's instructions, dried and exposed to film (BioMax MR Film, Carestream). 20  $\mu$ l of each reaction was spotted in duplicate onto P81 Filters (Unifilter 96 well plate, Whatman), washed with 200 mM ammonium bicarbonate, dried, overlaid with scintillation fluid (Betaplate Scint, Perkin-Elmer) and read in a scintillation counter (MicroBeta2, Perkin-Elmer).

**Transgenic parasite plasmid constructs.** The CRISPR/Cas9 genome editing approach was employed to generate different PfSET6 transgenic parasite line as described previously. A 20-nt gRNA target sequence directing Cas9 cleavage near the C-terminal end and in the SET

domain were cloned into the sgRNA-expression cassette in the pL6 plasmid, respectively. For each constructs, two homology regions of approximately 300-400bp were inserted in separate cloning steps using Gibson assembly. We generated a series of plasmid to knockdown PfSET6 that include: pL6-SET6-HA-glms, pL6-SET6-glms, pL6-SET6-HA-glms-hsp86, pL6-SET6-DiCre. A second plasmid, pUF1-Cas9, contained an engineered *S. pyogenes* endonuclease Cas9 cassette and regulatory elements of *P. falciparum* U6 snRNA.

**Parasite culture and transfections.** Asexual blood-stage parasites were cultivated using standard methods, and synchronous cultures were obtained by sorbitol treatment and plasmion enrichment. Parasites were transfected using the pre-loaded red blood cells method. Positive selection using 2.66 nM WR99210 and 1.5  $\mu$ M DSM1 and negative selection using 40  $\mu$ M 5-fluorocytosine were applied sequentially in culture until the integrant transfected line was established. To verify the integration and integration breakpoints, PCR was performed on genomic DNA from transgenic line or wild type strain.

**PfSET6 antibody production.** C57BL/6 and BALB/C mice were immunized 4 times every 14 days with 20ug doses of purified recombinant PfSET6 in Freund's Complete Adjuvant for the initial injection and Freund's Incomplete Adjuvant for booster injections to generate a polyclonal anti-PfSET6 antibody. Specificity of immune sera was determined by western blot.

**Subcellular fractionation and western blotting.** Subcellular extracts preparation was prepared as described previously. Equal amounts of protein were loaded and separated on 4-12% SDS-PAGE gel (Invitrogen), and analyzed by western blot using antibodies anti-PfSET6 (mouse) at 1:500 dilutions, anti-Pf-Aldolase (Abcam ab38905) or anti-Histone H3 (mouse, Abcam ab1791) at 1:2000 dilution. Secondary antibodies were anti-rabbit IgG-HRP (GE NA934V, Lot 9568295) and anti-mouse IgG-HRP (GE NA931V, Lot 9648752), both at 1:4000. Blots were developed using ECL (Thermo Scientific).

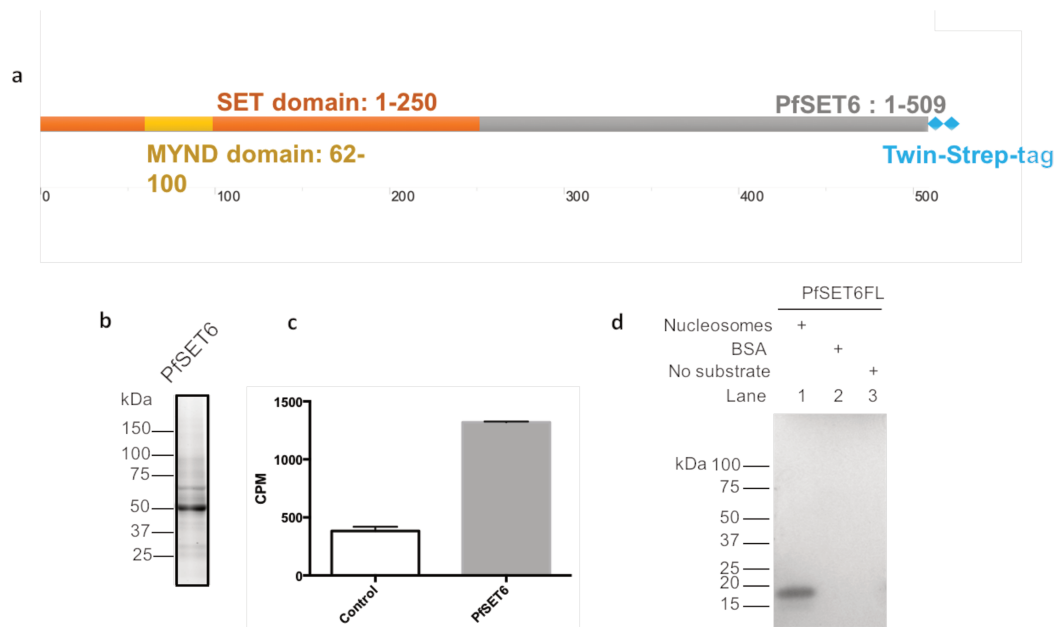
**Immunofluorescence microscopy.** Immunofluorescence microscopy were performed as described previously. Antibody dilutions are as follows: anti-PfSET6 (mouse) 1:1000, anti-PfHP1 (rabbit, Genscript) 1:4000, anti-BiP (rabbit, MR4) 1:1000, anti-ERD2 (rat, MR4) 1:1000, anti-PfEBA175 (rabbit, MR4) 1:1000, anti-PfCLAG3.2 (rabbit, MR4) 1:1000, anti-PfAlba1 (rabbit, peptide antibody) 1:200, anti-PfAlba2 (rabbit, peptide antibody) 1:200, anti-PfAlba4 (rabbit, peptide antibody) 1:200, anti-Pfg377 (rabbit, MR4) 1:1000, and secondary goat anti-rat/rabbit/mouse Alexa-Fluor 488 or 568 conjugates (Invitrogen), all at 1:2500.

Images were captured using a Nikon Eclipse 80i microscope with a CoolSnapHQ2 camera (Photometrics). NIS Elements version 3.0 software was used for image acquisition, and Fiji software was used for analysis.

**ChIP-Seq and data analysis.** Synchronous cultures of ring stage wild-type 3D7 parasites clone G10 were used. Cross-linked chromatin was prepared by adding formaldehyde to the culture followed by addition of glycine. Nuclei were isolated by homogenization suspended in SDS buffer. Chromatin was sheared by sonication in a Bioruptor® Pico to a size of 300bp. The commercially available antibodies to H3K9me3, H3K4me3, monoclonal anti-HP1, homemade polyclonal anti-PfSET6 were added and incubated at 4°C overnight. The antibody-protein-DNA complex is affinity purified using Magna ChIPTM ProteinG Magnetic Beads (Millipore, 16-662). As controls, DNA corresponding to the ChIP input or DNA immunoprecipitated using rabbit IgG were processed. Reverse cross-linking was followed by RNase treatment and Proteinase K digestion, DNA purification using the DNA Clean column. The MicroPlex Library Preparation kit (Diagenode, c05010012) is used for preparing libraries for sequencing. Pooled, multiplexed libraries were sequenced on an Illumina NextSeq® 500/550 system as a 150 nucleotide single-end run.

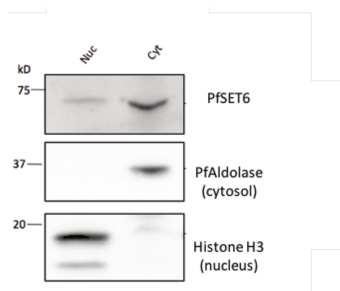
The resulting data were demultiplexed using bcl2fastq2 (Illumina) to obtain fastq files for downstream analysis. A minimum of two biological replicates was analyzed for each antibody and each parasite stage. Quality control of fastq files was performed using the FastQC software. Sequencing reads were mapped to the *P. falciparum* genome (v.3, GeneDB) with the Burrows-Wheeler Alignment tool (BWA MEM) using default parameters. After unique read mapping and deduplication, the aligned reads were considered for peak-calling analysis with the MACS2 software 5 after controlling for a false discovery rate of 5% (default setting of MACS2) using Benjamini and Hochberg's correction method. For the genome-wide representation of PfSET6 from ring stage parasites, the coverages were calculated as average RPKM (reads per kilobase of genome sequence per one million mapped reads) over bins of 1000bp width using bamCoverage from the package deepTools2 (v2.2.4, PMID: 27079975). Coverage differences between the PfSET6 ChIP experiment, Mouse IgG mock and the input control (Fig. X) were calculated using bamCompare from the same package. Circular and linear coverage plots were generated using circos (v0.69, PMID: 19541911) and the Integrated Genomics Viewer (v2.3.72, PMID: 22517427).

## FIGURE LEGENDS

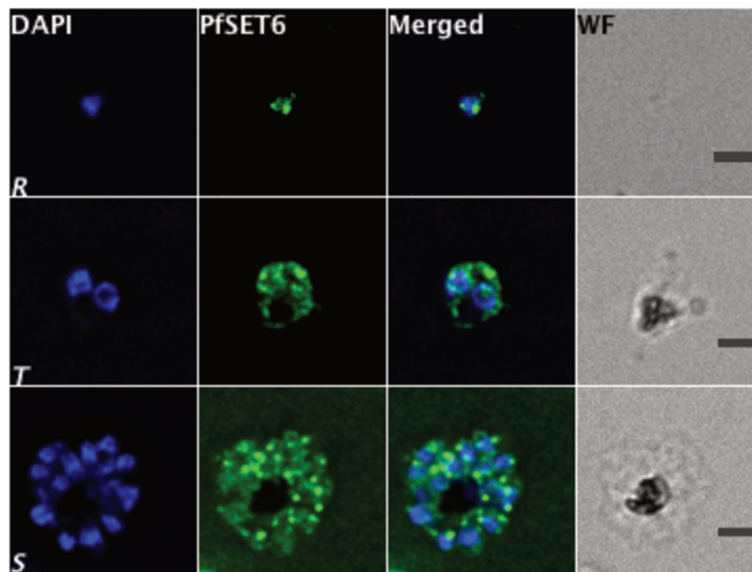


**Figure 1 PfSET6 is a bona fide histone methyltransferase.** a. Schematic representation of the domain organization of PfSET6 protein. The location of the 2x-Streptavidin tag for recombinant protein expression is indicated. b. Purified full-length recombinant PfSET6 was visualized by Bio-Rad Stain-Free detection (left) and was validated by western blotting using anti-Strep (right). c. Enzymatic assays with with radiolabeled S-adenosylmethionine as a methyl donor and histones as substrate and quantified as counts per minute (cpm). d. Autoradiograph of enzyme reactions with PfSET6 with nucleosomes, BSA and enzyme alone (lanes 1,2 and 3). Film was exposed 72h.

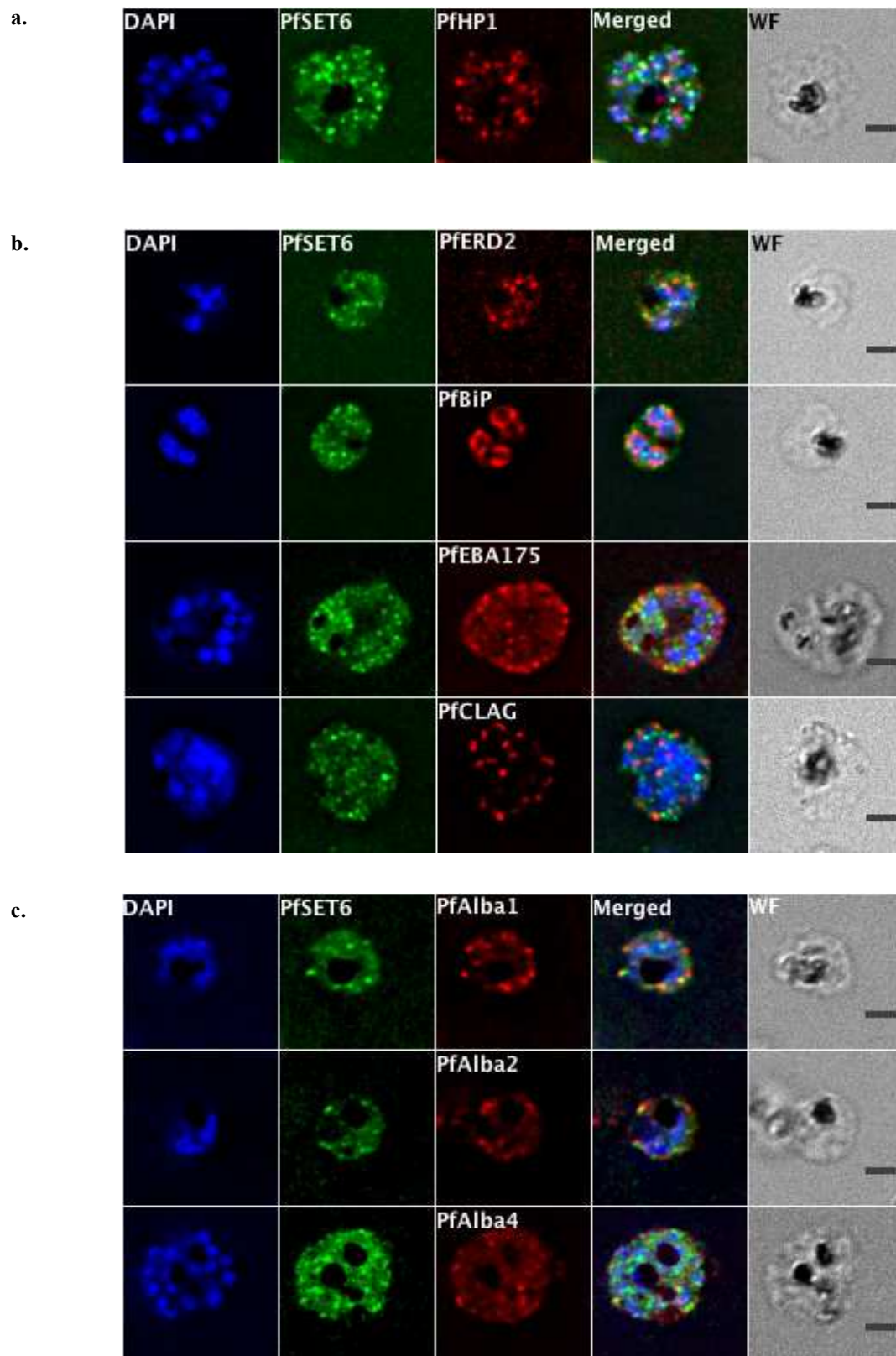
a.



b.

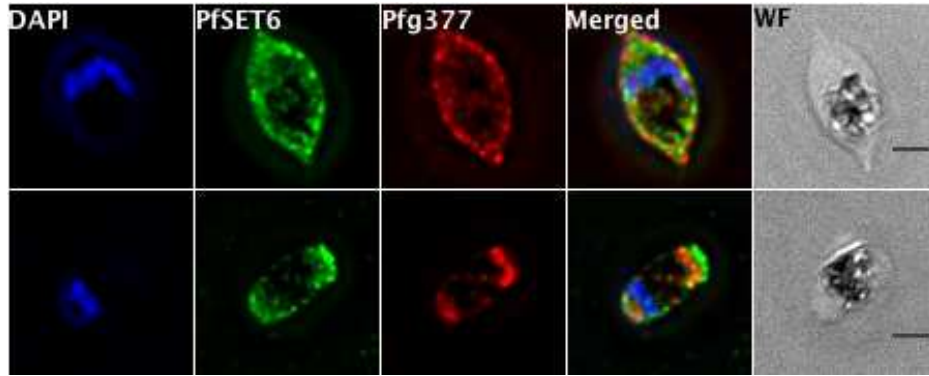


**Figure 2 PfSET6 localizes to the nucleus and cytoplasm of asexual blood stage parasites.** a. Western blotting of parasite nuclear and cytoplasmic fractions of asynchronized parasites using anti-PfSET6 antibodies. b. IF localization of PfSET6 in rings, trophozoites and schizonts. The scale bars correspond to 2 μm.

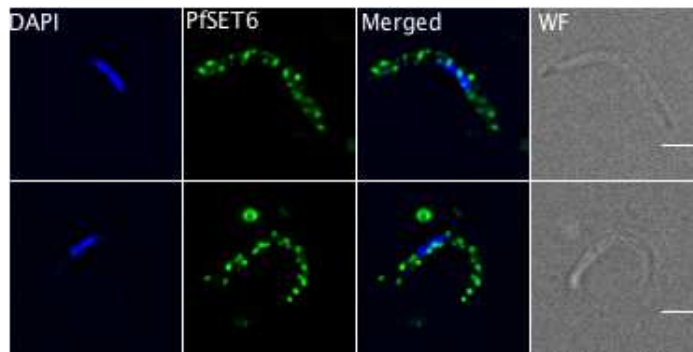


**Figure 3 PfSET6 localization is dynamic during asexual blood stage.** a. co-IF of schizonts using anti-PfSET6 combined with the nuclear periphery marker PfHP1. b. co-IF of late stage parasites using anti-PfSET6 combined with different markers of cytoplasmic organelles, anti-Bip as an ER marker, anti-ERD2 as a cis-golgi marker, anti-EBA175 as microneme, anti-CLAG as rhoptries. c. co-IF of late stage parasites using anti-PfSET6 combined with markers of mRNP granules including PfAlba1, 2 and 4. The scale bars correspond to 2µm.

a.



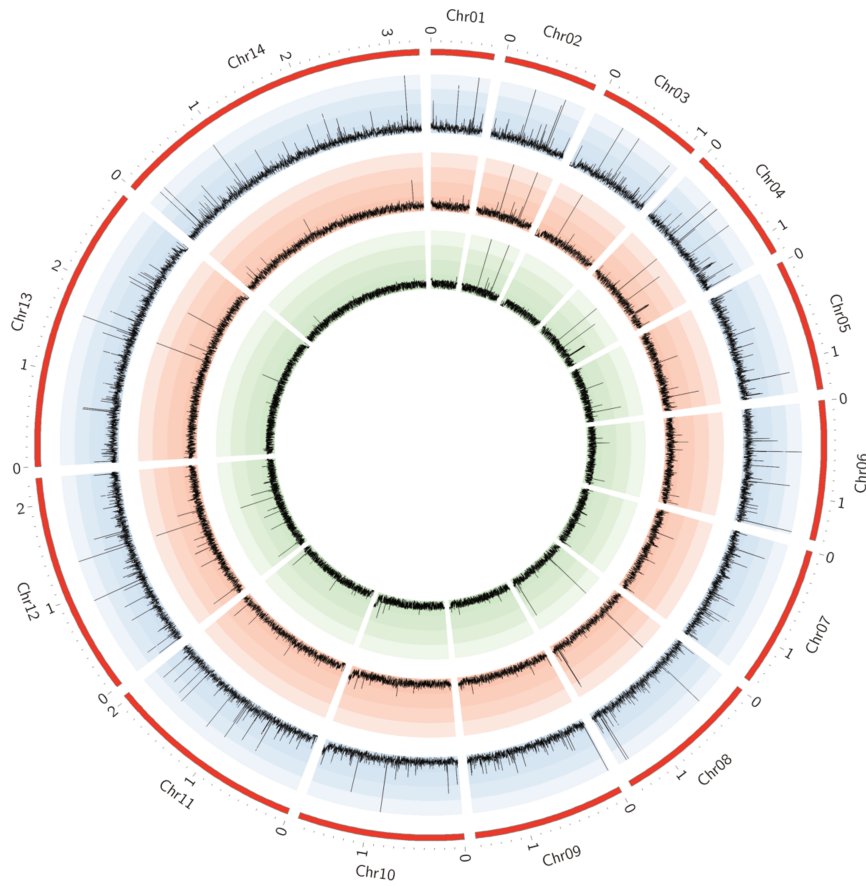
b.



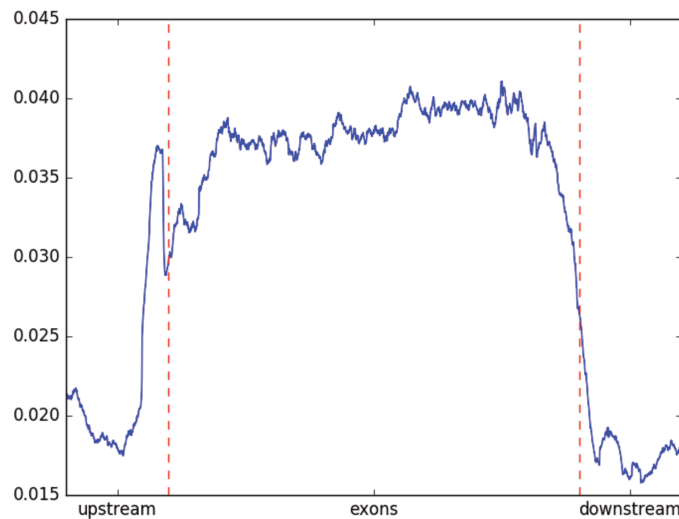
**Figure 4 PfSET6 localizes to discrete foci in the cytoplasm of *P. falciparum* transmission stages.** a. IF of stage IV and V gametocytes using anti-PfSET6 combined with female gametocyte-specific marker anti-Pfg377. The scale bars correspond to 2 $\mu$ m. b. IF localization of PfSET6 in salivary gland sporozoites. The scale bar corresponds to 2  $\mu$ m.



a.

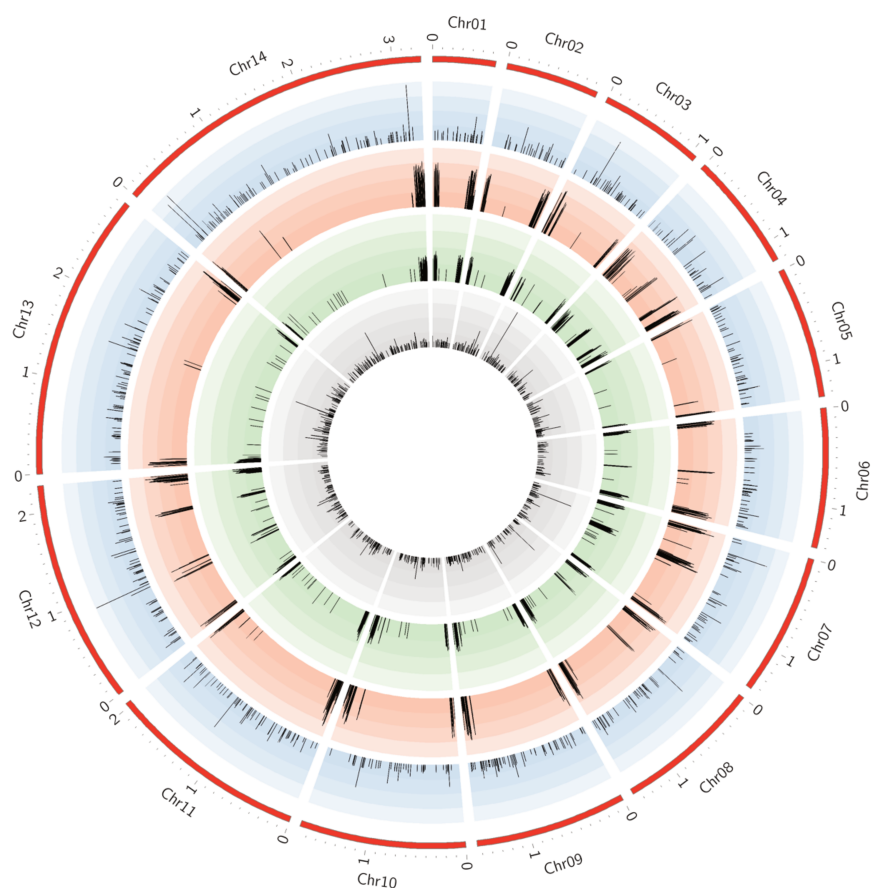


b.



**Figure 5 ChIP-Seq analysis of PfSET6 in ring stage parasites.** a. A circular *P. falciparum* genome with coverage depth plots. Blue circle background represents the signal of PfSET6-IP, red circle background represents mouse IgG mock, and green circle background represents input. The chromosome sizes are indicated in megabases around the outside of the plot in gray. The position within the genome sequence is indicated at the top. The scale for all three is 0 to 700 RPKM values over bins of 1000 bp/bin. b. Plot of the mean coverage over distance.

a.

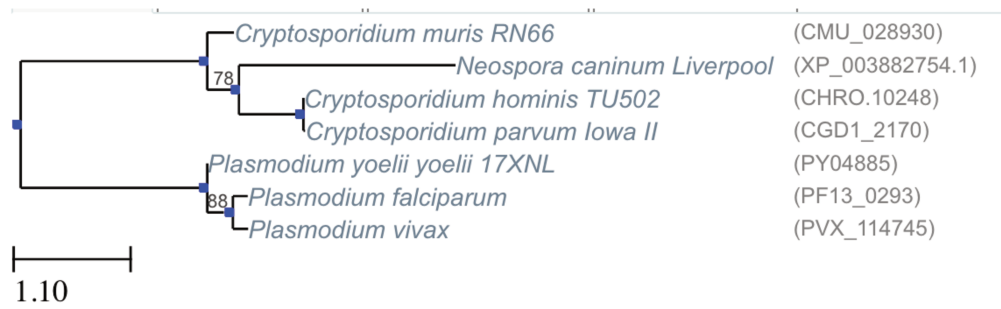


b.

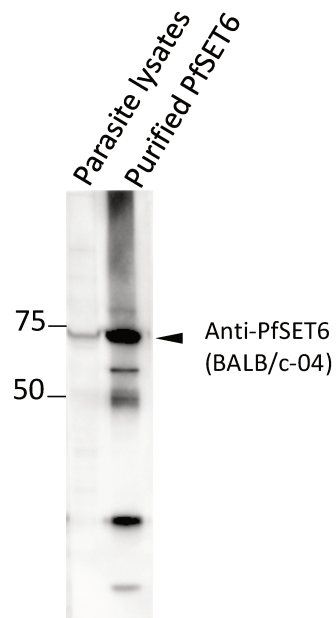
Correlation coefficients	PfSET6-IP
PfHP1-IP	0.4780
H3K9me3-IP	0.5431
H3K4me3-IP	0.0707

**Figure 6 PfSET6 genome-wide distribution correlates with the repressive mark histone H3K4me3 and its reader, PfHP1.** a. Circular representation of normalized peak location in PfSET6 (blue), PfHP1 (red), H3K9me3 (green) and H3K4me3 (grey). The height of bar represents the fold-enrichment of the represented state over the respective input data. b. The correlation coefficients between various CHIP datasets.

## SUPPLEMENTARY



**Figure 1** Phylogenetic tree of PfSET6 and Apicomplexa orthologous group.



**Figure 2** Western blotting of parasite lysates and purified PfSET6 protein using the native anti-PfSET6 antibodies.

**Table 1 List of genes with PfSET6 occupancy as determined by MACS2 peak calling analysis.**

[Gene ID]	[Genomic Location]	[Product Description]
PF3D7_0101100	Pf3D7_01_v3: 69,304 - 70,417 (-)	exported protein family 4 (EPF4)
PF3D7_0101200	Pf3D7_01_v3: 71,624 - 72,426 (+)	exported protein family 3 (EPF3)
PF3D7_0101400	Pf3D7_01_v3: 75,982 - 76,809 (-)	exported protein family 1, pseudogene (EPF1)
PF3D7_0101500	Pf3D7_01_v3: 78,241 - 79,891 (+)	erythrocyte membrane protein 1 (PfEMP1), exon 2, pseudogene (VAR)
PF3D7_0101700	Pf3D7_01_v3: 84,791 - 86,152 (-)	erythrocyte membrane protein 1 (PfEMP1), exon 2, pseudogene (VAR)
PF3D7_0112700	Pf3D7_01_v3: 477,279 - 481,382 (+)	28S ribosomal RNA
PF3D7_0113000	Pf3D7_01_v3: 487,892 - 490,127 (-)	glutamic acid-rich protein (GARP)
PF3D7_0114000	Pf3D7_01_v3: 542,480 - 546,983 (+)	exported protein family 1 (EPF1)
PF3D7_0213800	Pf3D7_02_v3: 558,585 - 560,594 (+)	conserved Plasmodium protein, unknown function
PF3D7_0217500	Pf3D7_02_v3: 720,437 - 722,661 (+)	calcium-dependent protein kinase 1 (CDPK1)
PF3D7_0222000	Pf3D7_02_v3: 876,984 - 877,832 (+)	exported protein family 1 (EPF1)
PF3D7_0222200	Pf3D7_02_v3: 881,469 - 882,267 (-)	exported protein family 3 (EPF3)
PF3D7_0222300	Pf3D7_02_v3: 883,452 - 884,312 (+)	exported protein family 4, pseudogene (EPF4)
PF3D7_0223300	Pf3D7_02_v3: 909,350 - 911,054 (+)	erythrocyte membrane protein 1 (PfEMP1), exon 2
PF3D7_0300300	Pf3D7_03_v3: 49,772 - 51,152 (-)	erythrocyte membrane protein 1 (PfEMP1), exon 2, pseudogene (VAR)
PF3D7_0300600	Pf3D7_03_v3: 58,307 - 58,519 (-)	Plasmodium exported protein, unknown function, fragment
PF3D7_0310200	Pf3D7_03_v3: 427,290 - 440,942 (+)	phd finger protein, putative
PF3D7_0324000	Pf3D7_03_v3: 1,005,952 - 1,006,800 (+)	exported protein family 1 (EPF1)
PF3D7_0324200	Pf3D7_03_v3: 1,010,342 - 1,011,140 (-)	exported protein family 3 (EPF3)
PF3D7_0324300	Pf3D7_03_v3: 1,012,355 - 1,013,464 (+)	exported protein family 4 (EPF4)
PF3D7_0324700	Pf3D7_03_v3: 1,024,030 - 1,025,411 (+)	erythrocyte membrane protein 1 (PfEMP1), exon 2, pseudogene (VAR)
PF3D7_0411800	Pf3D7_04_v3: 519,832 - 525,618 (-)	conserved Plasmodium protein, unknown function
PF3D7_0413100	Pf3D7_04_v3: 591,949 - 599,849 (-)	erythrocyte membrane protein 1, PfEMP1 (VAR)
PF3D7_0420700	Pf3D7_04_v3: 935,031 - 941,875 (-)	erythrocyte membrane protein 1, PfEMP1 (VAR)
PF3D7_0420900	Pf3D7_04_v3: 946,169 - 953,773 (-)	erythrocyte membrane protein 1, PfEMP1 (VAR)
PF3D7_0421000	Pf3D7_04_v3: 956,849 - 956,949 (+)	unspecified product
PF3D7_0421100	Pf3D7_04_v3: 958,067 - 965,611 (-)	erythrocyte membrane protein 1, PfEMP1 (VAR)
PF3D7_0423400	Pf3D7_04_v3: 1,055,665 - 1,056,318 (+)	asparagine-rich protein (AARP)
PF3D7_0500200	Pf3D7_05_v3: 29,233 - 29,304 (-)	erythrocyte membrane protein 1 (PfEMP1), pseudogene
PF3D7_0503400	Pf3D7_05_v3: 140,710 - 141,471 (+)	actin-depolymerizing factor 1 (ADF1)
PF3D7_0504700	Pf3D7_05_v3: 180,421 - 187,470 (+)	centrosomal protein CEP120, putative (CEP120)
PF3D7_0532000	Pf3D7_05_v3: 1,292,410 - 1,296,199 (+)	28S ribosomal RNA
PF3D7_0600400	Pf3D7_06_v3: 18,586 - 22,721 (-)	erythrocyte membrane protein 1, PfEMP1 (VAR)
PF3D7_0600500	Pf3D7_06_v3: 26,557 - 27,830 (+)	rifin (RIF)
PF3D7_0601000	Pf3D7_06_v3: 42,097 - 43,218 (-)	exported protein family 4 (EPF4)
PF3D7_0601100	Pf3D7_06_v3: 44,450 - 45,216 (+)	exported protein family 3 (EPF3)
PF3D7_0601300	Pf3D7_06_v3: 48,874 - 49,722 (-)	exported protein family 1 (EPF1)
PF3D7_0615900	Pf3D7_06_v3: 662,435 - 667,855 (-)	conserved Plasmodium protein, unknown function
PF3D7_0631200	Pf3D7_06_v3: 1,313,708 - 1,315,253 (-)	erythrocyte membrane protein 1 (PfEMP1), pseudogene (VAR)

PF3D7_0631300	Pf3D7_06_v3: 1,316,873 - 1,317,721 (+)	exported protein family 1 (EPF1)
PF3D7_0631500	Pf3D7_06_v3: 1,321,320 - 1,322,094 (-)	exported protein family 3 (EPF3)
PF3D7_0631600	Pf3D7_06_v3: 1,323,294 - 1,324,435 (+)	exported protein family 4 (EPF4)
PF3D7_0631700	Pf3D7_06_v3: 1,326,737 - 1,327,918 (+)	rifin, pseudogene
PF3D7_0701000	Pf3D7_07_v3: 52,086 - 53,636 (+)	erythrocyte membrane protein 1 (PfEMP1), pseudogene
PF3D7_0701300	Pf3D7_07_v3: 62,220 - 63,401 (-)	rifin, pseudogene (RIF)
PF3D7_0701400	Pf3D7_07_v3: 65,718 - 66,854 (-)	exported protein family 4, pseudogene (EPF4)
PF3D7_0701500	Pf3D7_07_v3: 68,066 - 68,840 (+)	exported protein family 3 (EPF3)
PF3D7_0701700	Pf3D7_07_v3: 72,477 - 73,325 (-)	exported protein family 1 (EPF1)
PF3D7_0704600	Pf3D7_07_v3: 216,024 - 229,072 (-)	E3 ubiquitin-protein ligase (UT)
PF3D7_0726000	Pf3D7_07_v3: 1,086,570 - 1,090,357 (+)	28S ribosomal RNA
PF3D7_0732800	Pf3D7_07_v3: 1,410,435 - 1,412,231 (+)	erythrocyte membrane protein 1 (PfEMP1), exon 2
PF3D7_0814200	Pf3D7_08_v3: 687,340 - 688,086 (+)	DNA/RNA-binding protein Alba 1 (ALBA1)
PF3D7_0818200	Pf3D7_08_v3: 830,006 - 831,454 (+)	14-3-3 protein (14-3-3I)
PF3D7_0819700	Pf3D7_08_v3: 888,126 - 890,471 (-)	conserved Plasmodium protein, unknown function
PF3D7_0831800	Pf3D7_08_v3: 1,374,236 - 1,375,299 (-)	histidine-rich protein II (HRPII)
PF3D7_0833200	Pf3D7_08_v3: 1,426,884 - 1,428,135 (+)	rifin (RIF)
PF3D7_0900800	Pf3D7_09_v3: 52,648 - 54,020 (-)	erythrocyte membrane protein 1 (PfEMP1), exon 2, pseudogene (VAR)
PF3D7_0924500	Pf3D7_09_v3: 993,245 - 997,675 (-)	conserved Plasmodium membrane protein, unknown function
PF3D7_1027000	Pf3D7_10_v3: 1,126,523 - 1,131,958 (-)	conserved Plasmodium protein, unknown function
PF3D7_1035200	Pf3D7_10_v3: 1,394,839 - 1,396,596 (+)	S-antigen
PF3D7_1035800	Pf3D7_10_v3: 1,420,533 - 1,422,671 (+)	probable protein, unknown function (M712)
PF3D7_1036400	Pf3D7_10_v3: 1,436,316 - 1,439,804 (+)	liver stage antigen 1 (LSA1)
PF3D7_1038400	Pf3D7_10_v3: 1,519,021 - 1,547,825 (+)	gametocyte-specific protein (Pf11-1)
PF3D7_1039200	Pf3D7_10_v3: 1,576,441 - 1,576,991 (-)	Plasmodium exported protein, unknown function, pseudogene
PF3D7_1039600	Pf3D7_10_v3: 1,590,146 - 1,590,973 (+)	exported protein family 1, pseudogene (EPF1)
PF3D7_1039800	Pf3D7_10_v3: 1,594,503 - 1,595,277 (-)	exported protein family 3 (EPF3)
PF3D7_1039900	Pf3D7_10_v3: 1,596,487 - 1,597,605 (+)	exported protein family 4 (EPF4)
PF3D7_1100900	Pf3D7_11_v3: 60,928 - 61,769 (-)	RESA-like protein, pseudogene
PF3D7_1101400	Pf3D7_11_v3: 76,820 - 78,000 (-)	rifin, pseudogene (RIF)
PF3D7_1101500	Pf3D7_11_v3: 80,306 - 81,424 (-)	exported protein family 4 (EPF4)
PF3D7_1101600	Pf3D7_11_v3: 82,633 - 83,415 (+)	exported protein family 3 (EPF3)
PF3D7_1101800	Pf3D7_11_v3: 87,041 - 87,889 (-)	exported protein family 1 (EPF1)
PF3D7_1102100	Pf3D7_11_v3: 104,300 - 104,852 (+)	Plasmodium exported protein, unknown function, pseudogene
PF3D7_1108100	Pf3D7_11_v3: 356,083 - 360,434 (+)	conserved Plasmodium protein, unknown function
PF3D7_1124000	Pf3D7_11_v3: 948,591 - 950,707 (+)	endoplasmic reticulum oxidoreductin, putative (ERO1)
PF3D7_1148600	Pf3D7_11_v3: 1,925,942 - 1,933,150 (+)	18S ribosomal RNA
PF3D7_1220900	Pf3D7_12_v3: 831,252 - 832,052 (-)	heterochromatin protein 1 (HP1)
PF3D7_1224000	Pf3D7_12_v3: 974,372 - 975,541 (+)	GTP cyclohydrolase I (GCHI)
PF3D7_1228400	Pf3D7_12_v3: 1,153,352 - 1,157,557 (+)	conserved Plasmodium protein, unknown function
PF3D7_1248700	Pf3D7_12_v3: 1,994,775 - 2,000,744 (+)	conserved Plasmodium protein, unknown function
PF3D7_1249700	Pf3D7_12_v3: 2,023,455 - 2,026,181 (-)	conserved Plasmodium protein, unknown function

PF3D7_1254200	Pf3D7_12_v3: 2,207,830 - 2,209,115 (+)	rifin (RIF)
PF3D7_1254300	Pf3D7_12_v3: 2,211,543 - 2,212,554 (+)	stevor
PF3D7_1300400	Pf3D7_13_v3: 47,583 - 48,769 (+)	rifin (RIF)
PF3D7_1300800	Pf3D7_13_v3: 60,097 - 61,172 (-)	erythrocyte membrane protein 1-like (VAR-like)
PF3D7_1302000	Pf3D7_13_v3: 112,792 - 113,815 (-)	EMP1-trafficking protein (PTP6)
PF3D7_1318300	Pf3D7_13_v3: 753,910 - 759,432 (+)	conserved Plasmodium protein, unknown function
PF3D7_1327000	Pf3D7_13_v3: 1,134,070 - 1,138,040 (-)	conserved Plasmodium protein, unknown function
PF3D7_1342400	Pf3D7_13_v3: 1,666,832 - 1,668,153 (-)	casein kinase II beta chain (CK2beta2)
PF3D7_1346300	Pf3D7_13_v3: 1,849,528 - 1,850,517 (+)	DNA/RNA-binding protein Alba 2 (ALBA2)
PF3D7_1371000	Pf3D7_13_v3: 2,800,004 - 2,802,154 (+)	18S ribosomal RNA
PF3D7_1371300	Pf3D7_13_v3: 2,802,945 - 2,807,159 (+)	28S ribosomal RNA
PF3D7_1372900	Pf3D7_13_v3: 2,865,893 - 2,867,226 (+)	erythrocyte membrane protein 1 (PfEMP1), exon 2, pseudogene (VAR)
PF3D7_1400100	Pf3D7_14_v3: 1,393 - 5,343 (+)	erythrocyte membrane protein 1 (PfEMP1), truncated, pseudogene

### 5.3 Article III

Analysis of epigenetic landscape of *P. falciparum* reveals expression of a clonally variant PfEMP1 member on the surface of sporozoites

Zanghi G, Vembar SS, Baumgarten S, Ding S *et al*, under review

#### Manuscript Highlight:

- We mapped, for the first time, the genome-wide distribution of eu- and heterochromatic features in sporozoites, including the transcriptionally repressive histone mark H3K9me3 and its reader protein PfHP1, and the transcriptionally activating histone mark histone H3K9ac.
- We found that the sporozoite genome is organized into heterochromatin-containing gene islands, in particular at chromosome ends, which localize to perinuclear clusters.
- We also identified a single *var* gene (*var<sup>sporo</sup>*) as being depleted of the repressive marks and further confirmed its expression in sporozoites. Strikingly, the encoded PfEMP1 protein was efficiently trafficked to the sporozoite surface and expressed, even in the lack of knob-like structures. This indicated that the epigenetic regulation of *var* gene transcription is maintained during the formation of sporozoites in the mosquito.
- Overall, this work demonstrates that epigenetically controlled variegated gene expression is important in mosquito stages. Moreover, the identification of a highly polymorphic sporozoite surface antigen is highly relevant for vaccine development.

#### Individual Contribution:

- I performed ChIP-seq experiments of the H3K9me3-binding protein PfHP1 in ring stage parasites and compared its enrichment to that observed in sporozoite stage *P. falciparum* parasites.

## **Analysis of epigenetic landscape of *P. falciparum* reveals expression of a clonally variant PfEMP1 member on the surface of sporozoites**

Gigliola Zanghi<sup>1</sup>, Shruthi S. Vembar<sup>2,3,4</sup>, Sebastian Baumgarten<sup>2,3,4</sup>, Shuai Ding<sup>2,3,4</sup>, Julien Guizetti<sup>2,3,4</sup>, Jessica M. Bryant<sup>2,3,4</sup>, Denise Mattei<sup>2,3,4</sup>, Anja T.R. Jensen<sup>5</sup>, Jean-François Franetich<sup>1</sup>, Robert Sauerwein<sup>7</sup>, Mallaury Bordessoulles<sup>1</sup>, Olivier Silvie<sup>1</sup>, Olivier Scatton<sup>6</sup>, Dominique Mazier<sup>1\*</sup> & Artur Scherf<sup>2,3,4\*</sup>

<sup>1</sup> Sorbonne Universités, Université Pierre et Marie Curie-Paris 6, UMR S945, Paris, F-75003; Institut National de la Santé et de la Recherche Médicale, U945, F-75013 Paris,; AP-HP, Groupe Hospitalier Pitié-Salpêtrière, Service Parasitologie-Mycologie, Paris, F-75013, France;

<sup>2</sup> Unité Biologie des Interactions Hôte-Parasite, Département de Parasites et Insectes Vecteurs, Institut Pasteur, 75724 Paris 75015, France;

<sup>3</sup> CNRS, ERL 9195, 75724 Paris 75015, France;

<sup>4</sup> INSERM, Unit U1201, 75724 Paris 75015, France;

<sup>5</sup> Centre for Medical Parasitology, Department of Immunology & Microbiology, Faculty of Health and Medical Sciences, University of Copenhagen and Department of Infectious Diseases, Copenhagen University Hospital (Rigshospitalet), Copenhagen, Denmark;

<sup>6</sup> Service de Chirurgie Digestive, Hépatobilio-Pancréatique et Transplantation Hépatique, Hôpital Pitié-Salpêtrière, Paris, France.

<sup>7</sup> Department of Medical Microbiology and Radboud Center for Infectious Diseases, Radboud university medical center, PO Box 9101, 6500 HB, Nijmegen, The Netherlands.

\*To whom correspondence should be addressed: Artur Scherf, Email: [artur.scherf@pasteur.fr](mailto:artur.scherf@pasteur.fr).  
or Dominique Mazier, Email: [dominique.mazier@upmc.fr](mailto:dominique.mazier@upmc.fr)



## Summary:

Malaria parasites have developed sophisticated strategies during asexual blood stage development that enable chronic infection, pathogenesis and transmission via epigenetically controlled gene expression<sup>1</sup>. In the human parasite *Plasmodium falciparum*, specific epigenetic signatures dictate the transcriptional activation or silencing of immune evasion genes such as the *var* multigene family, which encodes the major virulence factor Erythrocyte Membrane Protein 1 (PfEMP1). However, the epigenetic landscape in other life cycle stages and its impact on host-parasite interaction have not been investigated to date. Here, we map the genome-wide distribution of eu- and heterochromatic features in sporozoites, which migrate through different mosquito and human host cells before infecting a human hepatocyte. Using chromatin immunoprecipitation (ChIP) analysis of Heterochromatin Protein 1 (PfHP1) occupancy and single cell DNA fluorescence *in situ* hybridization (FISH), we found that the genome of sporozoites is organized into heterochromatin-containing gene islands, in particular at chromosome ends, which form 2-4 perinuclear clusters. Importantly, when comparing sporozoites to blood stage parasites, we observed a significant spreading of heterochromatin into gene regions that encode erythrocyte exported proteins. The observed sporozoite PfHP1 signature predicted the expression of a single *var* gene, (*PF3D7\_0809100*), which we confirmed by immunolabeling of the surface of live sporozoites with a specific antibody against the respective PfEMP1 protein. Overall, our work demonstrates that epigenetically controlled variegated gene expression is important in mosquito stages of the *P. falciparum* life cycle. Moreover, the identification of a highly polymorphic sporozoite surface antigen is highly relevant for vaccine development based on attenuated sporozoites.

The most devastating form of human malaria is caused by the protozoan parasite *P. falciparum*, with more than 200 million people infected annually and an estimated 438,000 deaths in 2015<sup>2</sup>. Malaria is transmitted by the bite of an infected *Anopheles* mosquito, which harbors sporozoites in its salivary glands. From the point of injection into the skin, sporozoites migrate via blood vessels to the liver, cross the sinusoidal cell layer separating the blood and the liver, and finally invade hepatocytes where asexual reproduction leads to the release of thousands of merozoites into the blood stream<sup>3</sup>. These infected red blood cells produce new infective merozoites that initiate a new infective cycle. The persistence and pathogenesis of the human malaria parasite *P. falciparum* during blood stage proliferation relies on on the successive expression of variant adhesion molecules termed PfEMP1 encoded by approximately 60 *var*

genes<sup>4</sup>. This immune evasion mechanism called antigenic variation is orchestrated by different layers of epigenetic factors<sup>1</sup>. Monoallelic expression of a single *var* member relies on the default silencing of the rest of the *var* gene family members via facultative heterochromatin. Variegated gene expression is used by the parasite beyond host immune evasion to create phenotypic plasticity in genetically identical parasites during blood stage including sexual commitment of a subset of parasites<sup>5-9</sup>. Although variegated gene expression appears to have evolved as a survival strategy to promote prolonged blood stage infection in humans, it is unknown if it is used by sporozoites, a stage that has been successfully used to provide immune protection to human volunteers<sup>10</sup>.

To fill this important knowledge gap, we developed a robust, low-cell input ChIP assay followed by high throughput sequencing (ChIP-seq) protocol to work with limiting sporozoite numbers. We analyzed the genome-wide distribution of the transcriptionally repressive histone post-translational modification mark (PTM) histone H3-lysine 9-trimethyl (H3K9me3), its reader protein PfHP1, and the transcriptionally activating histone PTMmark H3K9ac. In asexual blood stages, PfHP1 and H3K9me3 regulate transcription of sub-telomeric, clonally variant virulence gene families such as *var* by establishing facultative heterochromatin at promoter regions and gene bodies<sup>5,11,12</sup>. These heterochromatic regions span about hundred kilobases (kb), forming so-called heterochromatin islands. In sporozoites, we observed an enrichment of both H3K9me3 and PfHP1 in sub-telomeric regions of all 14 chromosomes as well as in central chromosome regions of chromosomes 4, 6, 7, 8 and 12 (Fig. 1a and 2a). In contrast, the activating mark H3K9ac is virtually absent from these heterochromatic loci and is dispersed along chromosome regions that contain primarily housekeeping genes (Fig. 1a and 2a). When comparing the genome-wide PfHP1 this pattern to that in asexual blood stage parasites (Fig. 1b and 2a), we observed a general similarity<sup>5</sup>. We noted, however, a significant sporozoite-specific spreading of subtelomeric heterochromatin from nearly all chromosome ends towards central chromosomal regions (Fig. 2a and Fig. 2b). These extended heterochromatic regions of about 50 kb are highly enriched in genes encoding parasite proteins that are typically exported to the host erythrocyte during blood stage development remodeling the erythrocyte to accommodate the parasite's needs<sup>13</sup> (Fig. 2b). Our data suggest for the first time that sporozoites may activate a facultative heterochromatin boundaries remodelling pathway controlled in a stage specific manner and that their extension may function to silence genes that are not expressed in sporozoites to silence genes encoding exported proteins traslocated during intra-erythrocytic development mostly on the surface of the infected erythrocytes (IE).

In blood stage parasites the maintenance of heterochromatin islands are physically linked tethered into several foci at of these regions to the nuclear periphery<sup>5</sup>. To determine if a similar spatial chromosome arrangement is involved/exists in sporozoites, we performed DNA-FISH using a probe corresponding to the conserved subtelomeric repeat TARE6 (Telomere-Associated Repeat 6)<sup>5,9</sup> (Fig. 3a) and indirect immunofluorescence assays (IFAs) with antibodies against PfHP1 (Fig 3b). We observed an average of 2-4 TARE6-containing foci per nucleus, which localized to the nuclear periphery and partially or completely overlapped with PfHP1-containing foci (Fig. 3c). These data indicate a conserved heterochromatin nuclear spatial organization in the nucleus between asexual blood stage parasites and migratory sporozoites.

In asexual stages, only one of the 60 *var* gene members is transcribed at a time to ensure expression of a single erythrocyte surface adhesion protein, implicated in immune evasion and pathogenesis<sup>14,15</sup>. To determine if heterochromatin may control the default silent state of the *var* virulence multigene family in sporozoites as in blood stage parasites we performed bioinformatic analysis of three independent H3K9me3 and PfHP1 ChIP-seq experiments. This analysis revealed that a single member, *Pf3D7\_0809100*, of the *var* gene family showed strongly reduced H3K9me3 and PfHP1 signal in its promoter region and gene body (Fig. 4a), suggesting that *Pf3D7\_0809100* may be expressed in sporozoites. *Pf3D7\_0809100* (here called *var*<sup>sporo</sup>) is located in the central region of chromosome 8, adjacent to two other *var* genes, which retain higher levels heterochromatin marks (Fig. 4a). Notably, other clonally variant gene families such as *rif*, *Pfmc2TM*, *PfACS* and *clag* families, which are located proximal to subtelomeric *var* genes on 13 of the 14 chromosomes, appear to maintain blood stage-like H3K9me3 and PfHP1 patterns, with only two of 180 *rif* genes, one of 13X *Pfmc2TM* genes and five of 13 *PfACS* genes depleted for these marks in Table 1. In addition, the previously identified master regulator of sexual commitment, AP2-G<sup>7,8</sup> shows a PfHP1 enrichment in sporozoites that is similar to that seen in blood stage parasites (Supplementary Fig S. 2).

As we observed the same pattern of PfHP1 depletion for the single *Pf3D7\_0809100* *var* gene in three independent sporozoite preparations, we investigated the possibility that this epigenetic signature was established during the asexual blood stage of the NF54 strain used to generate the sporozoites analyzed in this study. RT-qPCR analysis of all *var* genes in a synchronized bulk culture of NF54 ring stage parasites showed that five different *var* gene transcripts are present at relatively high levels, whereas most other members, including *var*<sup>sporo</sup>, are present at very low levels (Fig. 4b). These data strongly suggest that an epigenetic reset of *var* gene transcription occurs during sexual development, with the specific transcriptional upregulation

of *var*<sup>sporo</sup> during the formation of sporozoites in the mosquito.

To determine if *var*<sup>sporo</sup> is indeed expressed in sporozoites, as predicted by H3K9me3 and PfHP1 distribution (Fig. 4a), we performed IFAs on fixed or live sporozoites using either an antibody against the conserved intracellular C-terminal ATS (Acid Terminal Segment) of that reacts with all PfEMP1 proteins<sup>16</sup> or a specific antibody against the extracellular variable CIDR domain (Cysteine-Rich Inter domain region) of *var*<sup>sporo</sup> PfEMP1. Both antibodies reacted with fixed sporozoites and showed a surface membrane-like staining pattern that partially overlapped with Circumsporozoite protein (CSP), the major surface antigen of sporozoites (Fig. 4c top and middle rows). In contrast, no surface staining was observed with antibodies against the extracellular domain of the PfEMP1 encoded by *var2CSA* (*Pf3D7\_1200600*) and a second *var* gene (*Pf3D\_70412700*) (data not shown) that retain H3K9me3/PfHP1 marks, which is predicted to be repressed by the PfHP1 profile (Fig. 4c bottom row and Supplementary S3). When IFA was repeated with live sporozoites which have intact cell membranes, the anti-*var*<sup>sporo</sup> PfEMP1 antibody, but not the anti-ATS antibody, stained the surface of sporozoites (Fig. 4d), again in a pattern that partially overlapped with CSP (Fig. 4c and 4d). Trypsin treatment of live sporozoites (50 µg/ml) abolished antibody surface reactivity confirming that the CIDR domain of *var*<sup>sporo</sup> PfEMP1 is extracellular (data not shown). Together, these results provide the first evidence of a member of the highly polymorphic PfEMP1 protein being expressed on the surface of the sporozoites, with the ATS domain most likely at the cytoplasmic side of the parasite membrane and the CIDR domain exposed to the extracellular milieu. This manner of PfEMP1 expression contrasts with what is observed in blood stage parasites. In intra-erythrocytic parasites, PfEMP1 proteins are exported across the parasite and the parasitophorous vacuolar membrane to the surface of the erythrocyte, where they serve as adhesion molecules that can bind to a variety of endothelial receptors such as CD36, ICAM, etc., causing malaria pathogenesis by obstructing blood vessels in critical organs such as the brain<sup>4,15</sup>. Given the potential of *var*<sup>sporo</sup> to interact with hepatocytes to contribute to hepatocytes infection, we pre-incubated freshly isolated sporozoites with either the anti-*var*<sup>sporo</sup> antibody directed against the CIDR domain or pre-immune serum before adding them to a primary human hepatocyte culture. The *var*<sup>sporo</sup> antiserum did not reduce infection significantly at 1/50 serum dilution and showed similar level of infection as the corresponding preimmune serum in three independent experiments (Supplementary S4). Anti-CSP antibody confirmed the experimental procedure showing a potent inhibitory effect. Our hepatocyte invasion assay does not support a role for *var*<sup>sporo</sup> PfEMP1 in hepatocyte infection, although better antibodies

targeting different regions of this very large protein are needed to further explore its potential interactions with host cell receptors.

This is, to our knowledge, the first genome-wide analysis of important epigenetic signatures in *P. falciparum* sporozoites. Adaptation of ChIP-seq to this stage has uncovered several features that will greatly inform our view of the biology of malaria transmission to humans. Importantly, the transcription of a single *var* gene was predicted by PfHP1 enrichment patterns, and we confirm the corresponding PfEMP1 expression by surface IFA of sporozoites in three independent mosquito infection experiments. Our data strongly suggest that a specific member of the *var* multigene family (~60 members) is selected for expression in sporozoites. This novel concept is further supported by our finding that NF54 blood stage parasites primarily transcribe *var* genes that are distinct from the identified *var*<sup>sporo</sup>. Thus, it is unlikely that *var*<sup>sporo</sup> transcriptional activation is simply maintained in the transition from human blood to mosquito stage. It is currently unclear why a single *var* gene would be specifically expressed in mosquito stage sporozoites. In the future, *var*<sup>sporo</sup> gene knock out experiments are needed to shed light on the biology of *var*<sup>sporo</sup> PfEMP1 during sporozoite development in the mosquito and migration to human hepatocytes. Using *var*<sup>sporo</sup> mutants to infect mosquitos could also reveal if the mechanisms of *var* gene switching are maintained in sporozoites.

In addition, our work adds a new dimension to vaccine development based on live attenuated sporozoites. Successful protection of human volunteers has been reported after challenge with homologous *P. falciparum* parasites. The discovery of a highly polymorphic strain-specific sporozoite surface antigen may indicate that protective immune response targets variant specific antigens. This emphasizes the need to challenge volunteers with heterologous parasite strains to evaluate the degree of protection due to polymorphic antigens. Future work on *var*<sup>sporo</sup> may shed light on the homing mechanisms of sporozoites to hepatocytes.

## **Acknowledgements**

We thank Benoit Gamain for providing anti var2CSA antibodies. This work was supported by a European Research Council Advanced Grant (PlasmoSilencing 670301), ANR grant HypEpiC (ANR-14-CE16-0013) and the French Parasitology consortium ParaFrap (ANR-11-LABX0024). G.Z. and S.D. are supported by a ParaFrap PhD fellowship, S.S.V. was supported by a fellowship Carnot-Pasteur-Maladies Infectieuses and J.B. by an EMBO fellowship (ALTF 180-2015).

## Authors' contributions

G.Z., S.S.V., D.M. and A.S. conceived and designed the experiments; G.Z., S.S.V., S.D. performed ChIP experiments G.Z., S.S.V. and S.B. analyzed the ChIP data; J.B. performed qRT-PCR analysis, D.M. and A.T.R.J. contributed reagents/materials for *var*<sup>sporo</sup> localization experiments, G.Z. and J.G. performed imaging experiments, J.F., M.B., O.S., R.S. and G.Z. contributed to sporozoite preparation and O.S. to primary human hepatocyte infection studies. G.Z. and A.S. wrote the manuscript. All authors read and approved the final manuscript.

## Competing financial interests:

The authors declare that they have no competing interests.

## References

1. Guizetti, J. & Scherf, A. Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell Microbiol* 15, 718-726 (2013).
2. WHO. Fact Sheet: World Malaria Report 2015. <http://www.who.int/malaria/publications/world-malaria-report-2015/report/en/> (2015).
3. Prudêncio, M., Rodriguez, A. & Mota, M. M. The silent path to thousands of merozoites: the *Plasmodium* liver stage. *Nat Rev Microbiol* 4, 849-856 (2006).
4. Smith, J. D. The role of PfEMP1 adhesion domain classification in *Plasmodium falciparum* pathogenesis research. *Mol Biochem Parasitol* 195, 82-87 (2014).
5. Lopez-Rubio, J. J., Mancio-Silva, L. & Scherf, A. Genome-wide analysis of heterochromatin associates clonally variant gene regulation with perinuclear repressive centers in malaria parasites. *Cell Host Microbe* 5, 179-190 (2009).
6. Rovira-Graells, N. et al. Transcriptional variation in the malaria parasite *Plasmodium falciparum*. *Genome Res* 22, 925-938 (2012).
7. Sinha, A. et al. A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. *Nature* 507, 253-257 (2014).
8. Kafsack, B. F. et al. A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* 507, 248-252 (2014).
9. Brancucci, N. M. et al. Heterochromatin protein 1 secures survival and transmission of malaria parasites. *Cell Host Microbe* 16, 165-176 (2014).
10. Richie, T. L. et al. Progress with *Plasmodium falciparum* sporozoite (PfSPZ)-based malaria vaccines. *Vaccine* 33, 7452-7461 (2015).

11. Salcedo-Amaya, A. M. et al. Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 106, 9655-9660 (2009).
12. Flueck, C. et al. *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS pathogens* 5, e1000569 (2009).
13. de Koning-Ward, T. F., Dixon, M. W., Tilley, L. & Gilson, P. R. *Plasmodium* species: master renovators of their host cells. *Nat Rev Microbiol* 14, 494-507 (2016).
14. Scherf, A., Lopez-Rubio, J. J. & Riviere, L. Antigenic variation in *Plasmodium falciparum*. *Annu Rev Microbiol* 62, 445-470 (2008).
15. Miller, L. H., Ackerman, H. C., Su, X. Z. & Wellems, T. E. Malaria biology and disease pathogenesis: insights for new treatments. *Nat Med* 19, 156-167 (2013).
16. Nacer, A. et al. Clag9 is not essential for PfEMP1 surface expression in non-cytoadherent *Plasmodium falciparum* parasites with a chromosome 9 deletion. *PLoS One* 6, e29039 (2011).

## Methods section

### Isolation and purification of *P. falciparum* sporozoites

*P. falciparum* (NF54) sporozoites were obtained from the Department of Medical Microbiology, University Medical Centre, St Radboud, Nijmegen, the Netherlands. Adult *Anopheles stephensi* females were infected with NF54 strain of *P. falciparum*<sup>1</sup>. After 14-21 days from an infective blood meal, the salivary glands were aseptically dissected and purified on 17% Accudenz gradient as previously described<sup>2</sup>.

### Chromatin Immunoprecipitation and Next Generation Sequencing

ChIP was performed as previously described (Lopez-Rubio et al. 2013), using as starting material of  $5 \times 10^6$  parasites for *P. falciparum* sporozoites and  $1 \times 10^9$  parasites of the laboratory-adapted strain 3D7, clone G7, for *P. falciparum* asexual ring stages, using 0.5 ug of anti-H3K9me3, anti-H3K9ac (Millipore) and anti-PfHP1 rabbit polyclonal antibodies<sup>3</sup>. To generate Illumina-compatible sequencing libraries, the immunoprecipitated DNA was processed using the MicroPlex Library Preparation Kit (Diagenode) according to manufacturer's instructions. As controls, DNA corresponding to the ChIP input or DNA immunoprecipitated using rabbit IgG were processed. Pooled, multiplexed libraries were sequenced on an Illumina NextSeq® 500/550 system as a 150 nucleotide single-end run. The resulting data were demultiplexed using bcl2fastq2 (Illumina) to

obtain fastq files for downstream analysis. A minimum of two biological replicates was analyzed for each antibody and each parasite stage.

### **ChIP-seq Data Analysis**

Quality control of fastq files was performed using the FastQC software. Sequencing reads were mapped to the *P. falciparum* genome (v.3, GeneDB) with the Burrows-Wheeler Alignment tool (BWA MEM) using default parameters <sup>4</sup>. Uniquely mapped reads of quality Q20 were considered for peak-calling analysis with the MACS2 software <sup>5</sup> after controlling for a false discovery rate of 5% (default setting of MACS2) using Benjamini and Hochberg's correction method <sup>6</sup>. For the genome-wide representation of HP1, H3K9ac and H3K9m3 from both sporozoites (Fig 1,3 and Fig S1, S2) and blood stage parasites (Fig. 1a,b), the coverages of the respective chromatin marks were calculated as average RPKM (reads per kilobase of genome sequence per one million mapped reads) over bins of 1000bp width using bamCoverage from the package deepTools2 (v2.2.4, PMID: 27079975). Coverage differences between the HP1 ChIP experiment and the input control (Fig. 3a) were calculated using bamCompare from the same package. Circular and linear coverage plots were generated using circos (v0.69, PMID: 19541911) and the Integrated Genomics Viewer (v2.3.72, PMID: 22517427).

### **Immunofluorescence imaging of sporozoites**

For *var* protein staining  $5 \times 10^4$  sporozoites were placed on poly-L-lysine coated coverslips either live or fixed in 4% PFA for 10 min at room temperature were blocked with 4%BSA/PBS for 30 min, followed by incubation with primary antibody for 1h, three washes with PBS, and detection with Alexa Fluor 488- or 568-conjugated anti-rat IgG, anti-guineapig and anti-rabbit respectively (Life technologies) diluted 1:1000 in PBS/3%BSA. After three final washes in PBS cells were mounted in Vectashield containing DAPI for nuclear staining. Images were captured using a Deltavision Elite imaging system (GE healthcare) with a 60x 1.42 NA objective and a DV Elite CMOS camera. Fiji (<http://fiji.sc>) was used for image analysis of at least two independent replicas.

### **Protein expression and antisera**

The cysteine-rich interdomain region (CIDR)-alpha of PF3D7\_0809100 was cloned into the *Baculovirus* transfer vector pAcGP67-A (BD Biosciences). Recombinant *Baculovirus* were generated by cotransfection of the pAcGP67-A-CIDR construct gene and Bsu36I-linearized Bakpak6 *Baculovirus* DNA (BD Biosciences) into insect Sf9 cells. Recombinant CIDR product was



expressed by infection of insect High Five cells with recombinant *Baculovirus*. CIDR protein was purified from culture supernatants on Co<sup>2+</sup> metal chelate agarose columns <sup>7</sup>. Antibodies to *Baculo*-expressed CIDR were raised in Wistar rats by subcutaneous injection as previously described <sup>8</sup>. All experimental animal procedures were approved by The Danish Animal Procedures Committee ("Dyreforsøgstilsynet") as described in permit no. 2008/561-1498 and according to the guidelines described in Danish act LBK 1306 (23/11/2007) and BEK 1273 (12/12/2005). Rabbit anti var2CSA was a gift of Benoit Gamain <sup>9</sup>. The guinea pig antiserum against recombinant PfEMP1 ATS (Acid Terminal Sequence) region was published earlier <sup>10</sup>. Mouse Mab C6 directed against the *P. falciparum* CS protein has been described previously <sup>11</sup>.

### **DNA-FISH and immunofluorescence imaging**

Sporozoites were fixed in suspension with 4% paraformaldehyde in PBS over night at 4°C. Parasites were then deposited on polylysine coated #1.5 cover glasses, permeabilized 15 min with 0.1% Triton-X-100, and subjected to same DNA-FISH labeling conditions as described previously for blood stage parasites (Mancio-Silva et al. 2008). For sequential immunofluorescence and DNA-FISH parasites were labeled first with rabbit anti-HP1 at 1:2000 in PBS/3%BSA followed by detection with Alexa-Fluor 488-conjugated anti-rabbit IgG (Life technologies). After an additional fixation step we carried out DNA-FISH using a biotinylated LNA probe (Exiqon) labeling all TARE6 telomeric repeat regions 5'-Biotin-+AC+T+AACA+TA+GG+T+CT+T+A-Biotin-3' detected with Streptavidin-Alexa568 (Life Technologies) at 1:2000 in PBS/3%BSA.

### **Trypsin treatment**

Living  $1 \times 10^6$  sporozoites in 200  $\mu$ l Leibovitz's L-15 medium were treated with Trypsin (Gibco) with a final concentration of 50 g/ml for 40 minutes at 25°C to avoid sporozoites activation. Then sporozoites were washed 3 times with Leibovitz's L-15 medium containing serum. The staining and imaging were performed as previously described.

### **Real-Time PCR analysis**

Total RNA was isolated from highly synchronized parasite cultures by sapoin lysis followed by purification with miRNeasy kit (Qiagen). RNA was DNase (Qiagen) treated and reverse transcription was performed using random hexamer primers and SuperScript VILO reverse transcriptase (ThermoFisher Scientific). Quantitative PCR was performed with the resultant cDNA in triplicate on a CFX Real-Time PCR System (BioRad) using Power SYBR Green (Life

Technologies) and primers specific for almost all *var* genes, designed in a previous study<sup>12</sup>. Transcript levels were determined using the quantity mean of each triplicate as calculated from a standard curve based on a serial dilution of genomic DNA. *var* gene transcription was normalized to that of a housekeeping gene, seryl tRNA synthetase.

### **Liver stage development inhibition assay**

Primary human hepatocytes were isolated from human liver fragments collected during unrelated surgery after informed consent was obtained from patients undergoing a partial hepatectomy as part of their medical treatment at Service de Chirurgie Digestive, Hepato-Bilio-Pancréatique et Transplantation Hépatique, Hopital Pitie Salpêtrière, Paris, France. Collection and use of this material for the purposes of the study presented here were undertaken in accordance with French national ethical guidelines under Article L, 1121-1 of the 'Code de la Santé Publique', with approval from the ethics committee of the Centre Hospitalo-Universitaire Pitié-Salpêtrière, Assistance Publique-Hôpitaux de Paris, France. Primary human hepatocyte cultures were prepared from freshly isolated liver fragments as previously described<sup>13</sup>. Cells were seeded at a density of  $8 \times 10^4$  cells per well in 96-well culture plates coated with rat tail collagen I (Becton Dickinson) and cultured at 37 °C, 5% CO<sub>2</sub> in William's Medium E (Life Technologies) supplemented with 10% FCS (Perbio),  $5 \times 10^{-5}$  M hydrocortisone hemisuccinate (Upjohn Laboratories SERB, France), 5 µg per ml insulin (Sigma), 2 mM l-glutamine, 1% penicillin–streptomycin solution (100X, stock solution, GIBCO). After cell adherence (12–24 h), 2% DMSO was added to the culture medium until infection. The different heat-inactivated rat sera of var-sporo and the correspondent pre-immune rat sera were added to the *P. falciparum* sporozoites at a final dilution of 1/50 dilution. A total of  $3 \times 10^4$  sporozoites from the *P. falciparum* NF54 strain in 50 µl of fresh medium were added next to the cells. Plates were centrifuged at 2000 rpm for 10 min at room temperature to facilitate parasite settling onto the cells. Medium changes were carried out at 3, and 48h post addition of sporozoites. Cultures were stopped at day 3 days post infection by fixing with cold methanol. Parasite numbers within hepatocytes were assessed via immunofluorescence using anti-PfHSP70 antibodies as previously described<sup>14</sup>. Quantification was performed by microscopy or using the arrayscan XTi imaging system (Thermofisher).

## References

1. Ponnudurai, T. et al. Infectivity of cultured *Plasmodium falciparum* gametocytes to mosquitoes. *Parasitology* **98 Pt 2**, 165-173 (1989).
2. Kennedy, M. et al. A rapid and scalable density gradient purification method for *Plasmodium* sporozoites. *Malar J* **11**, 421 (2012).
3. Chen, P. B. et al. *Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites. *Sci Rep* **6**, 21802 (2016).
4. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754-1760 (2009).
5. Zhang, Y. et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol* **9**, R137 (2008).
6. Y, B. & Y, H. Controlling false discovery rate: A practical and powerful approach to multiple testing. *Royal statistical society* **57**, 289-300 (1995).
7. Jensen, A. T. et al. *Plasmodium falciparum* associated with severe childhood malaria preferentially expresses PfEMP1 encoded by group A var genes. *J Exp Med* **199**, 1179-1190 (2004).
8. Bengtsson, A. et al. A novel domain cassette identifies *Plasmodium falciparum* PfEMP1 proteins binding ICAM-1 and is a target of cross-reactive, adhesion-inhibitory antibodies. *J Immunol* **190**, 240-249 (2013).
9. Gangnard, S. et al. Structure of the DBL3X-DBL4ε region of the VAR2CSA placental malaria vaccine candidate: insight into DBL domain interactions. *Sci Rep* **5**, 14868 (2015).
10. Nacer, A. et al. Clag9 is not essential for PfEMP1 surface expression in non-cytoadherent *Plasmodium falciparum* parasites with a chromosome 9 deletion. *PLoS One* **6**, e29039 (2011).
11. Okitsu, S. L. et al. Structure-activity-based design of a synthetic malaria peptide eliciting sporozoite inhibitory antibodies in a virosomal formulation. *Chem Biol* **14**, 577-587 (2007).
12. Salanti, A. et al. Selective upregulation of a single distinctly structured var gene in chondroitin sulphate A-adhering *Plasmodium falciparum* involved in pregnancy-associated malaria. *Mol Microbiol* **49**, 179-191 (2003).
13. Silvie, O. et al. A role for apical membrane antigen 1 during invasion of hepatocytes by *Plasmodium falciparum* sporozoites. *J Biol Chem* **279**, 9490-9496 (2004).
14. Rénia, L. et al. A malaria heat-shock-like determinant expressed on the infected hepatocyte surface is the target of antibody-dependent cell-mediated cytotoxic mechanisms by nonparenchymal liver cells. *Eur J Immunol* **20**, 1445-1449 (1990).

## Figure Legends

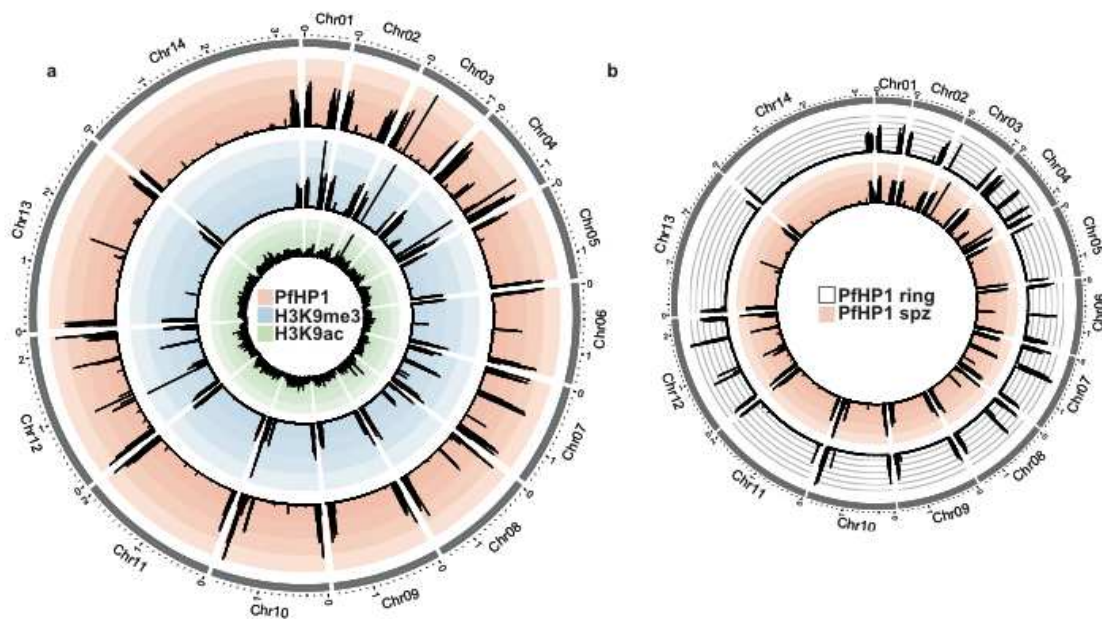


Figure 1

### Figure 1

**ChIP-seq analysis of transcriptionally repressive and activating chromatin features in *P. falciparum* sporozoites.** Circular representation of ChIP-seq data for all 14 chromosomes. (a) Chromatin enrichment (black) of PfHP1 (red), H3K9me3 (blue) and H3K9ac (green) in sporozoites. (b) Comparison of PfHP1 ChIP data obtained for sporozoites (red) and ring blood stage parasites (black). For both images the chromosome sizes are indicated in megabases around the outside of the plot in gray. The position within the genome sequence is indicated at the top. In all cases, coverage plots are represented as average RPKM values over bins of 1000 bp/bin.

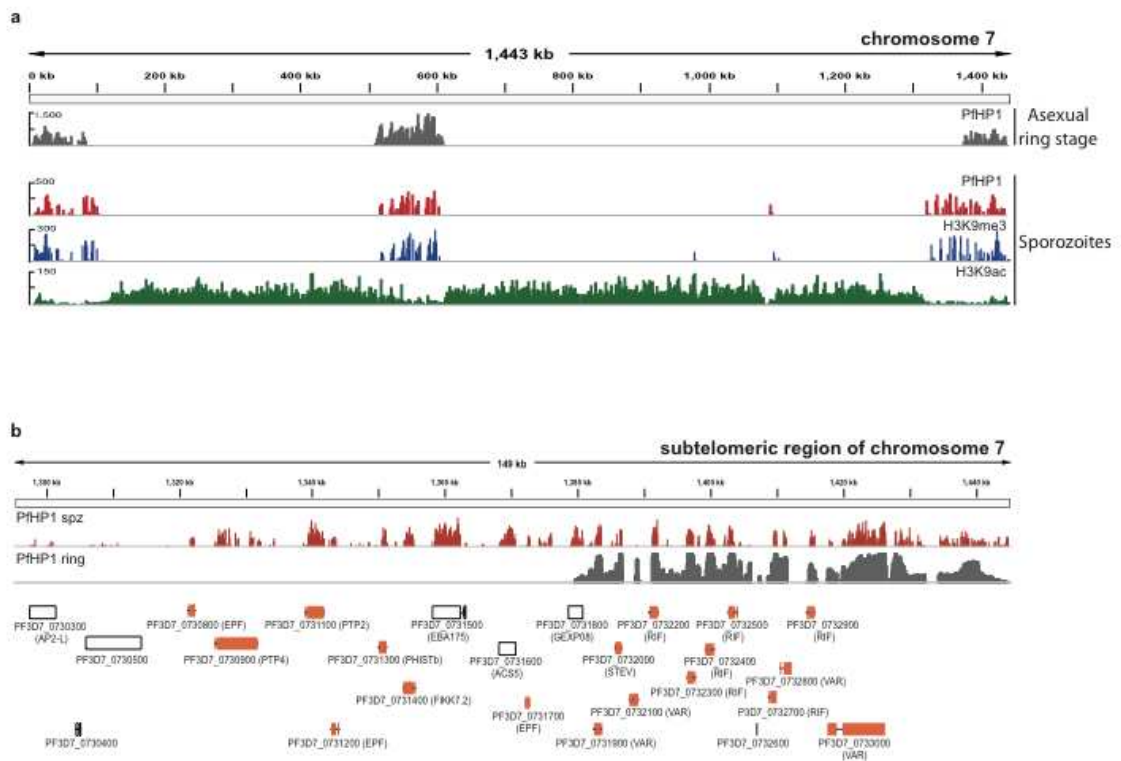


Figure 2

## Figure 2

**Extended heterochromatin islands in subtelomere regions in sporozoites.** (a) Chromosome 7 is shown to illustrate the differences in the PfHP1 distribution at subtelomeric regions between sporozoite (bottom three tracks) and asexual blood stage parasites (top track). (b) Subtelomeric region of chromosome 7 with the telomere at the right, including annotated genes at the bottom. ChIP-seq data show the enrichment of PfHP1 for sporozoites (red) and for the asexual blood stage parasites (black). In sporozoites PfHP1 covers an extra region of 80 kb that is highly enriched in genes encoding exported proteins of infected red blood cells (orange). In (a) coverage plots are represented as average RPKM values over bins of 1000 bp/bin, in (b) coverage plots are represented as an average of 1 single nucleotide/bin. In all cases the position on the genome sequence is indicated at the top.

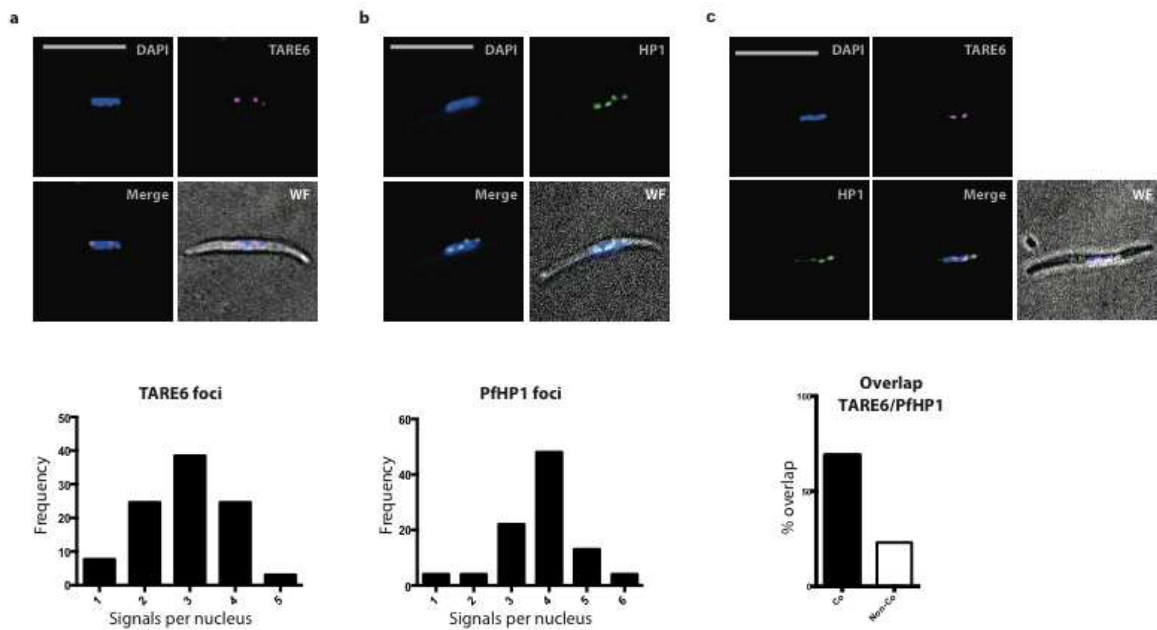


Figure 3

**Figure 3:**

**Frequency and spatial nuclear organization of chromosome ends and heterochromatin islands in sporozoites.** (a) DNA-FISH of telomeres (using a TARE6 probe in magenta) in sporozoites. Frequency distribution of nuclear TARE6 foci with an average of  $3 \pm 1$  foci per nucleus. (b) Immunofluorescence analysis of PfHP1 (green) and its nuclear distribution with an average of  $4 \pm 0.09$  foci per locus. (c) Combined immunofluorescence analysis of PfHP1 (green) and DNA-FISH of telomeres (using a TARE6 probe in magenta) in sporozoites with a percentage of 70% of TARE6 PfHP1 co-localization. DNA was stained with DAPI (blue), indicating the nucleus. Scale bar 5  $\mu$ m. From this experiment, frequency distribution and overlap of TARE6 (telomere clusters) and nuclear PfHP1 foci was established for 50 nuclei for three biological replicates.

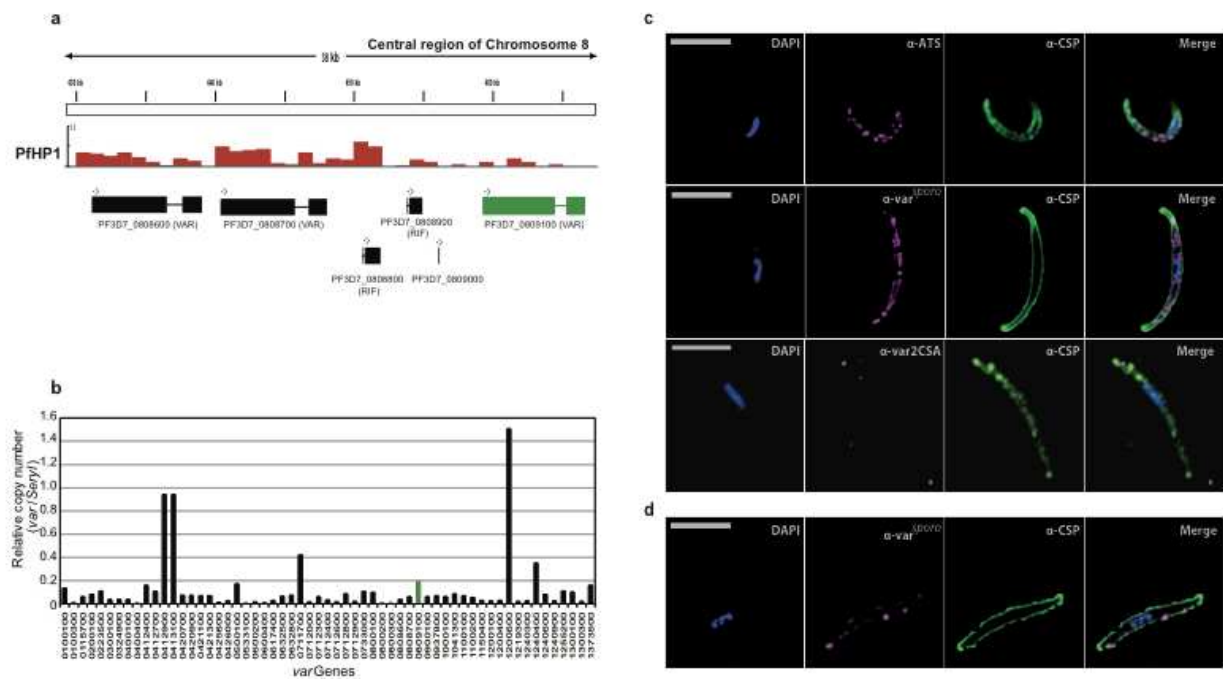


Figure 4

## Figure 4

### Analysis of PfHP1 occupancy reveals sporozoite surface expression of a *var* gene member.

(a) Bioinformatic analysis of ChIP-seq data revealed a single *var* gene - *Pf3D7\_0809100* (*var*<sup>sporo</sup>, highlighted in green) – with extremely low PfHP1 enrichment (indicated on the y-axis) and high enrichment for the neighboring *var* genes (in black) on chromosome 8 (indicated on the x-axis). The position within the genome sequence is indicated at the top. (b) RT-qPCR analysis of all *var* gene RNA transcripts from highly synchronized ring stage parasites 12 hours post infection. *var* gene cDNA levels are normalized to those of seryl tRNA synthetase. The RNA transcript level for *var*<sup>sporo</sup> is highlighted in green. (c) Immunofluorescence analysis of fixed and permeabilized sporozoites using anti-ATS (top row, magenta), anti-*var*<sup>sporo</sup> PfEMP1 (middle row, magenta), anti-*var*2CSA (bottom row, magenta), and anti-CSP (all rows, green) antibodies. (d) Surface immunofluorescence of live sporozoites using anti-CSP (green) and anti-*var*<sup>sporo</sup> (magenta) antibodies. For (c) and (d) DNA was stained with DAPI (blue), indicating the nucleus, and the scale bar is 5  $\mu$ m.

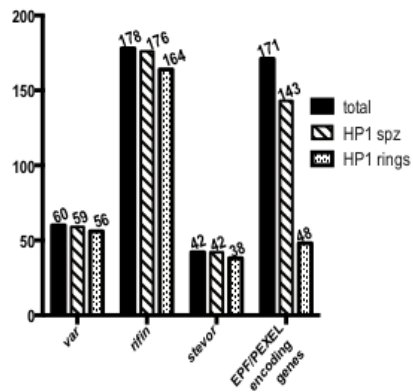


Figure S1

### Supplementary Fig. 1

#### Comparison of PfHP1 enrichment in genes of sporozoites and asexual ring stage parasites.

Bar graph shows clonally variant gene families *var*, *rifin* and *stevor* and genes predicted to be exported into the erythrocyte cytoplasm. Sporozoite show significantly higher PfHP1 coverage for exported proteins. Numbers on top of each bar indicate gene numbers.



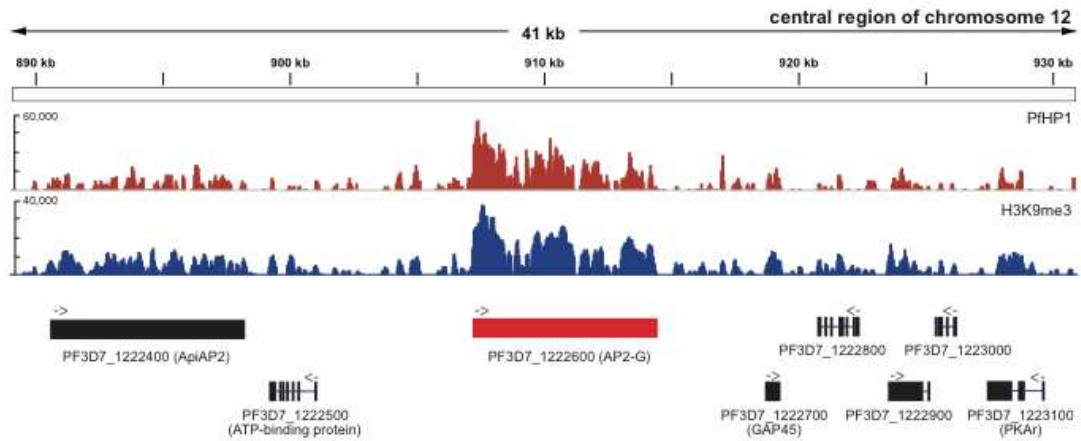


Figure S2

### Supplementary Fig. 2

**H3K9me3 and PfHP1 are enriched at the AP2 gene on chromosome 12 in sporozoites.** ChIP-seq analysis shows the enrichment and colocalization of PfHP1 (red) and H3K9me3 (blue) at the AP2G locus (shown at the bottom). The position on the genome sequence is indicated at the top and coverage plots are represented as an average of 1 single nucleotide/bin.

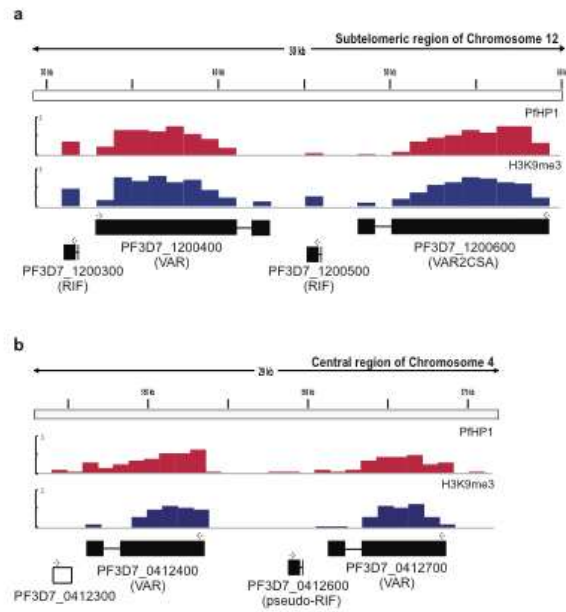


Figure S3

### Supplementary Fig. 3

**Enrichment levels of PfHP1 and H3K9me3 for silent *var* genes used in this study.** (a) ChIP-seq data showing enrichment and colocalization of PfHP1 (red) and H3K9me3 (blue) at the locus of *var2CSA*. (b) PfHP1 (red) and H3K9me3 (blue) enrichment at *var PF3D7\_0412700* locus. In all cases, coverage plots are represented as average of 1000 bp/bin. The position on the genome sequence is indicated at the top.

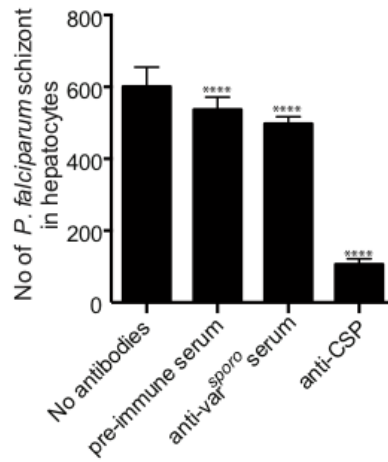


Figure S4

#### Supplementary Fig. 4

**Sporozoite invasion inhibition assay in primary human hepatocytes.** Columns in the graph show the mean of intracellular parasites of three independent biological experiments. Error bars in graph indicate standard deviation. p-value < 0.0001.

**Table1 : PfHP1 enrichment in clonally variant gene families**

<b>Name</b>	<b>PfHP1 Enriched</b>	<b>PfHP1 devoid Gene ID</b>	<b>Comments</b>
<i>var</i>	59/60	<i>PF3D7_0809100</i>	Cytoadherence, immune escape, IE surface
<i>rifin</i>	176/178	<i>PF3D7_1219200</i> <i>PF3D7_1240700</i>	IE surface
<i>stevor</i>	42/42	-	IE surface
<i>PfMC-2TM</i>	12/13	<i>PF3D7_0101300</i>	IE surface
<i>PfACS</i>	8/13	<i>PF3D7_0215000</i> <i>PF3D7_0215300</i> <i>PF3D7_0619500</i> <i>PF3D7_1238800</i> <i>PF3D7_0525100</i>	Metabolic proteins
<i>clag</i>			Rhoptry proteins, new IE permeation pathway
<i>clag3.1/ clag3.2</i>	2/2	-	
<i>clag2/clag8/clag9</i>	3/3	-	

IE, infected erythrocyte

## 5.4 Results IV (Supplementary)

### Generation of Transgenic Dual-labeled Parasite Lines

#### Scope of the Project:

The goals of this project were to create tools to drive forward epigenetic drug screening, and to identify novel factors contributing to *var* monoallelic expression.

#### Context of the Project:

Using the CRISPR/Cas9 marker-free genome editing approach, we are now in a position to create a transgenic parasite line that expresses different PfEMP1 proteins with distinct fluorescent tags such as GFP and RFP. Based on the principle of *var* monoallelic expression, a wild-type parasite will at any given time express either PfEMP1-1-GFP or PfEMP1-2-RFP, but not both. However, if monoallelic expression is disrupted, then more than one *var* gene will be transcriptionally active, resulting in double-fluorescent parasites. Once the proof of concept has been established, the dual-labeled parasite line will be used to screen small molecule inhibitors that mediate derepression of *var* genes, which likely target HKMT enzymes or the epigenetic reader HP1, using as readout flow cytometry and fluorescence microscopy. Furthermore, the parasite line will be chemically mutagenized and analyzed for *var* derepression: novel factors contributing to *var* monoallelic expression will be identified using deep sequencing of genomic DNA.

#### Specific Aims:

- Design and generation of GFP-tagged PfEMP1 in the *P. falciparum* 3D7 strain, G7 clone with a single predominantly active *var* gene *PF3D7\_0412700*; Determination of the fluorescent properties of PfEMP1-0412700-GFP.
- Design and generation of RFP-modified rarely activated *var* gene locus *PF3D7\_1100200* in the *P. falciparum* 3D7 strain; Determination of the fluorescent properties of predominantly expressed PfEMP1-1100200-RFP.
- Generate the dual-labeled transgenic parasite line.
- Confirm the applicability of the dual-labeled reporter line for drug screening by knocking out the gene encoding PfSETvs.

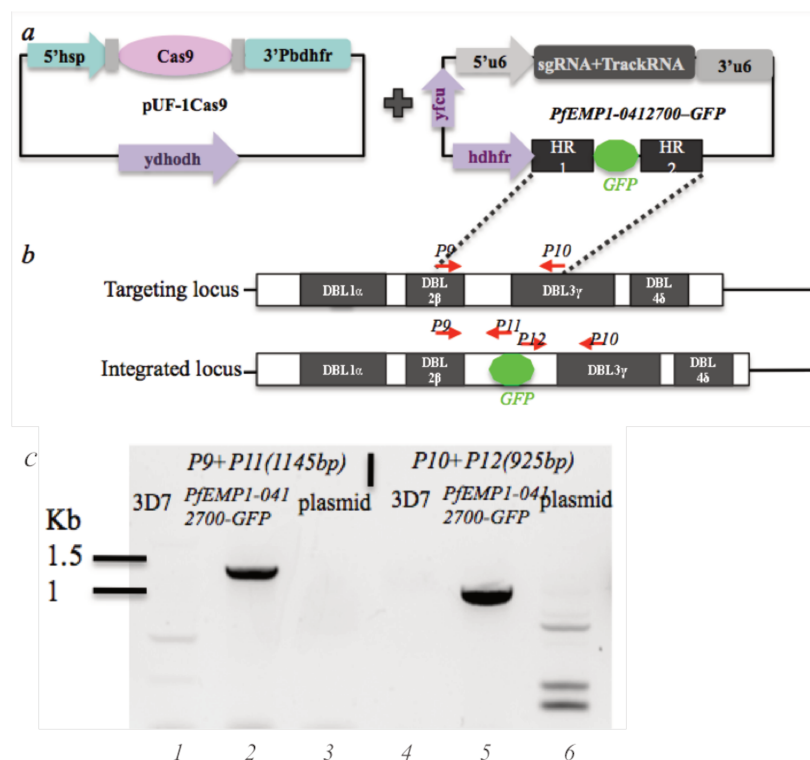
## Current Progress:

- Obtained PfEMP1-1-GFP transgenic parasites and confirmed expression of the tagged protein by live cell imaging.
- Obtained PfEMP1-2-RFP transgenic parasites.

## Results:

### Tagging the endogenous locus of the *PF3D7\_0412700 var* gene with GFP

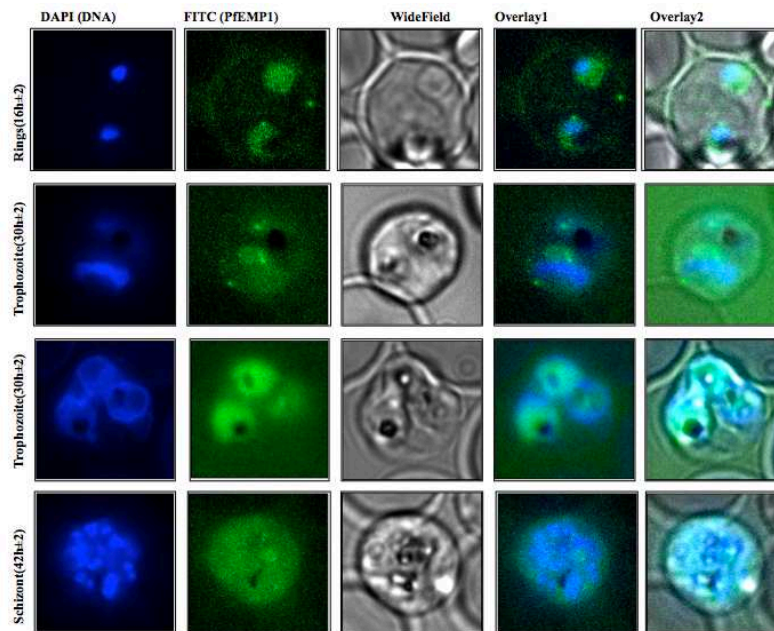
The first step involved the design of constructs to insert the GFP cassette between two DBL domains of *PF3D7\_0412700 var* open reading frame using CRISPR/Cas9-based genome editing (Figures 1a and b). After transfection of the constructs into the G7-D4 sub-clone of *P. falciparum* 3D7, transgenic parasites (*PfEMP1-0412700-GFP*) were confirmed using PCR amplification of the modified locus (Figure 1c) and sequencing of the PCR product.



**Figure 1. Integration of GFP into the *PF3D7\_0412700 var* locus.** a. Schematic of the CRISPR/Cas9-based strategy for GFP-tagging of the *var* gene locus. The Cas9 endonuclease bearing nuclear localization signals is expressed from the pUF1-Cas9 episome. The pL6-based-episome, pPfEMP1-0412700-GFP, contains the guide RNA sequence for Cas9 and homology regions for double crossover homologous recombination. b. The insertion of the GFP cassette should occur between the DBL2 $\beta$  and DBL3 $\gamma$  domains of PfEMP1-0412700. The

primers used for PCR in part C are indicated as p9-p12. c. PCR-based confirmation of GFP insertion into the *PF3D7\_0412700 var* locus. Lanes 1 and 4 correspond to genomic DNA prepared from the parental line 3D7, Lanes 2 and 5 correspond to genomic DNA prepared from the transgenic *PfEMP1-0412700-GFP* line and Lanes 3 and 6 correspond to the p*PfEMP1-0412700-GFP* plasmid. Primers P9 and P11 amplify a 1145 bp band, while primers, P10 and P12 amplify a 925 bp band.

Having confirmed integration of GFP into the *PF3D7\_0412700* locus in-frame, we next performed live cell imaging of the transgenic parasite to determine if *PfEMP1-0412700-GFP* is produced and efficiently trafficked to the parasite surface. As shown in Figure 2, in trophozoite stages (30 h), when *var* mRNA translation is maximal, *PfEMP1-0412700-GFP* signal can be visualized as a punctuated staining pattern in the RBC cytoplasm and RBC membrane. However, in later stages, the GFP signal is predominantly in the parasite cytoplasm or in the parasitophorous vacuole.

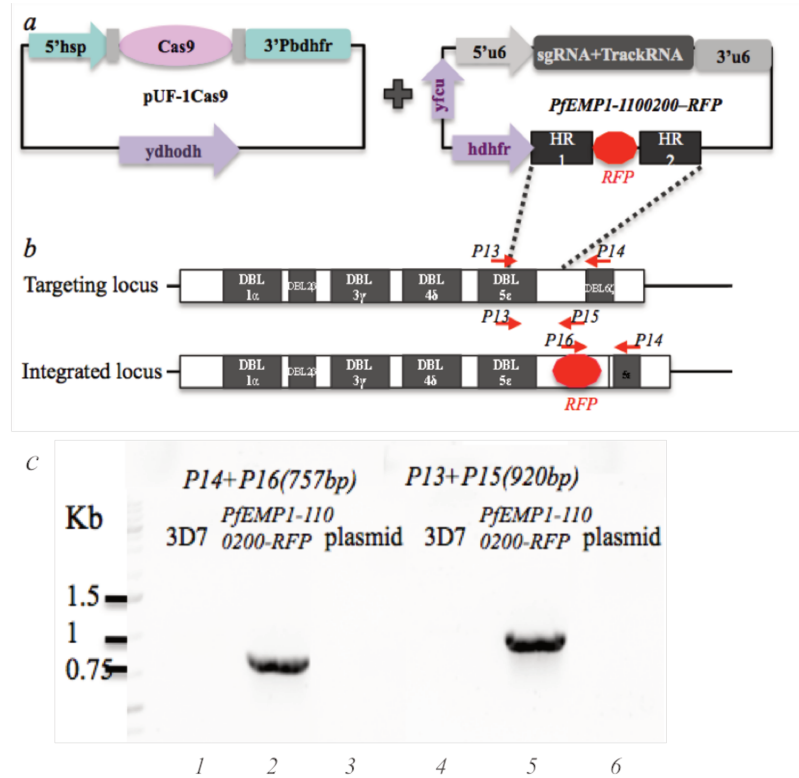


**Figure 2** *PfEMP1-0412700-GFP* distribution pattern in live parasites. The stage of parasite growth is indicated on the left. DAPI specifically stains nuclei. The last two panels are overlays of either DAPI/GFP or DAPI/GFP/bright-field images.

### Tagging the endogenous locus of the *PF3D7\_1100200 var* gene with RFP

To tag the second endogenous *PF3D7\_1100200 var* gene locus with, a similar strategy to the one described for *PfEMP1-0412700-GFP* was used. The RFP tag was inserted between amino acids X and Y of the *PF3D7\_1100200 var* open reading frame using CRISPR/Cas9-based genome editing (Figures 3a and b); *PF3D7\_1100200 var* is rarely activated in culture

presenting as an excellent candidate to understand *var* deregulation. Genomic integration of RFP to generate the *PfEMP1-1100200-RFP* transgenic parasite line was confirmed by PCR amplification and sequencing as shown in Figure 3c.



**Figure 3. Integration of RFP into the *PF3D7\_1100200 var* locus.** a. Schematic of the CRISPR/Cas9-based strategy for RFP-tagging of the *var* gene locus. The Cas9 endonuclease bearing nuclear localization signals is expressed from the pUF1-Cas9 episome. The pL6-based-episome, pPfEMP1-1100200-RFP, contains the guide RNA sequence for Cas9 and homology regions for double crossover homologous recombination. b. The insertion of the GFP cassette should occur between the DBL5 $\epsilon$  and DBL6 $\zeta$  domains of PfEMP1-1100200. The primers used for PCR in part C are indicated as p13-p15. c. PCR-based confirmation of GFP insertion into the *PF3D7\_1100200 var* locus. Lanes 1 and 4 correspond to genomic DNA prepared from the parental line 3D7, Lanes 2 and 5 correspond to genomic DNA prepared from the transgenic *PfEMP1-1100200-RFP* line, and Lanes 3 and 6 correspond to the pPfEMP1-0412700-GFP plasmid. Primers P13 and P15 amplify a 920 bp band, while primers, P14 and P16 amplify a 757 bp band.





# DISCUSSION





## 6. Discussion

In *P. falciparum*, PTMs have been shown to play an important role in the control of general transcriptional regulation, monoallelic expression of virulence genes, and in the commitment to transmission stage development. In particular, histone methylation has been linked to transcriptional activation throughout the *Plasmodium* life cycle and to transcriptional repression of *P. falciparum* multi-copy gene families via its reader protein PfHP1. Ten SET domain-containing histone methyltransferases or HKMTs have been bioinformatically predicted in *P. falciparum*; six of them appear to be essential for parasite asexual blood stage development based on traditional gene knockout experiments (non-inducible homologous gene replacement strategies) (Cui et al., 2008; Jiang et al., 2013). Until now, only two PfSET members have been characterized to some degree, revealing a role in virulence gene expression. The predicted H3K4-specific methyltransferase PfSET10 was suggested to associate with the poised *var* gene expression site in late blood stage parasites, while the H3K36-specific methylase PfSETvs (PfSET2) is required to keep monoallelic expression of *var* genes operational in blood stage parasites (Jiang et al., 2013; Volz et al., 2012). Based on potential importance in various biological aspects of SET domain containing genes in malaria parasites, a major part of my dissertation project focuses on PfSET7 and PfSET6, which have not been previously characterized.

### **PfSET7-mediated methylation may be an extension of the histone code to other proteins**

The 793-amino acid PfSET7 protein contains a central catalytic SET domain followed by a short post-SET domain at its C-terminal end. When HKMTs were first described in *Plasmodium* in 2008 by Cui *et al*, phylogenetic analysis classified PfSET7 as a putative histone H3K4-specific methyltransferase belonging to a clade that was separate from but related to SMYD3, a H3K4-specific methylase (Cui et al., 2008). In accordance with this, our *in vitro* enzyme kinetics investigation revealed that PfSET7 extensively methylates H3K4 and H3K9. In particular, H3K9 is specifically targeted in the presence of a pre-existing H3K14ac mark. Next, *in vivo* analysis revealed that PfSET7 is expressed during the entire asexual blood stage and detectable only in the cytosol as distinct foci (Chen and Ding et al., 2016). This was unanticipated for a putative histone methyltransferase. However, there are several implications of these observations.

First, many HKMT enzymes which had been originally reported as histone methyltransferases were subsequently found to possess non-histone protein substrates (Rea et al., 2000). Our data also point to PfSET7 as being a protein methyltransferase with substrates in the parasite cytosol. Future studies could focus on identifying the substrates targeted by PfSET7 for methylation using techniques such as immuno-affinity capture of methylation substrates and bio-orthogonal click chemistry, in which chemical reporters enable the installation of affinity tags appended to existing methylation using enzymes or selective chemical reactions (Luo, 2011; Weller and Rajsiki, 2005).

Second, in other organisms, newly synthesized histones have been reported to be methylated at H3K9 and acetylated at H3K14 within histone-chaperone protein complexes prior to their deposition into nucleosomes (Alvarez et al., 2011; Loyola et al., 2006). Though the mechanism by which newly synthesized histones are imported into the nucleus and deposited onto replicating chromatin is poorly understood, several HDACs and HKMTs that act on newly synthesized histones have been identified and play a role in chromatin maturation and DNA replication (Campos et al., 2010; Shinkai and Tachibana, 2011): PfSET7 may be one such protein. To define whether post-translationally modified, pre-deposition histones are imported into the nucleus in *P. falciparum*, histone-containing complexes from parasite cytoplasmic fractions could be biochemically purified and then characterized for their PTM profiles.

Third, our study primarily focused on parasite blood stages (asexual and sexual), with localization analysis of PfSET7 in pre-erythrocytic sporozoite and liver stages. However, it is possible that PfSET7 may localize to the nucleus of parasite mosquito stages. Although the presence of a nuclear localization signal (NLS) is not evident from PfSET7 sequence, its import/export from the nucleus could be mediated by its co-factors in a stage-dependent manner, which requires in-depth study.

As mentioned above, immunofluorescence staining revealed a punctate cytoplasmic pattern of PfSET7 in asexual blood stages. Particularly, in the late-schizont stage, the anti-PfSET7 antibodies always stained the apical tip, which was reminiscent of the apical localization of proteins relevant to parasite motility and red blood cell egress and invasion. Co-localization of different apical markers and high-resolution immune electron microscopy need to be performed to obtain a better view of PfSET7's role in this process. Indeed, a more recent phylogenetic analysis of *Plasmodium* and other Apicomplexan SET-domain proteins showed

that PfSET7 actually belongs to the AKMT clade (for Apical complex lysine methyltransferase) (Sivagurunathan et al., 2013). AKMT orthologs and the closely related SMYD subfamily have a distinctive structural difference: the post-SET cysteine cluster in AKMTs contains 2 more cysteine residues (CXCX2CX11CX2C) and is significantly longer than that (CXCX2C) in SMYDs (Sivagurunathan et al., 2013). Furthermore, studies in *Toxoplasma gondii*, another Apicomplexan parasite, reported that AKMTs modulate key steps required for motility, invasion and egress, suggesting that protein methylation by PfSET7 may be a conserved process in Apicomplexans (Heaslip et al., 2011; Sivagurunathan et al., 2013).

It is worth mentioning that recent work revealed that covalent modifications of actin regulate important cellular processes. For example, HDAC6 specifically acetylates actin and participates in actin reorganization (Yildirim et al., 2008). In addition, His-73 methylation regulates actin's interdomain flexibility and stability (Terman and Kashina, 2014). Based on PfSET7's cellular localization in schizonts, it is tempting to speculate that PfSET7 may play a role in actin methylation to regulate host erythrocyte invasion, a process that is dependent on actin-myosin motor activity at the apical tip of the merozoite. This could also explain the redistribution of PfSET7 that occurs in activated sporozoites that are preparing to invade hepatocytes.

A striking observation from our immunofluorescence assays is the enrichment of PfSET7 at the parasite membrane in gametocytes, suggesting that PfSET7 undergoes redistribution in sexual stages and might facilitate the remarkable morphological transformation as parasites prepare itself for transmission. Dixon et al. rendered a three-dimensional model of the inner membrane complex of gametocytes, composed of RBC membrane, parasitophorous vacuole membrane (PVM), parasite plasma membrane (PPM), the sub-pellicular membranes and a layer of sub-pellicular microtubules (MT) from outside to inside. Future immune electron microscopy and pull-down studies would be required to address the co-localization of PfSET7 with the gametocyte's inner membrane complex. Overall, that PfSET7 is continuously expressed throughout the life cycle (only ookinete and oocyst have not been studied as yet) and undergoes cellular redistribution to associate with distinct cellular components is a strong indicator that it might have different biological roles in different life cycle stages.

## **PfSET6 partially associates with transcriptional repression pathways in the nucleus and post-transcriptional regulators in the cytoplasm**

This 509-amino acid protein belongs to the SET and MYND domain-containing 3 (SMYD3) family that is a special class of protein lysine methyltransferases involved in methylation of histones and non-histone targets in other eukaryotes. In cancer research, SMYD3 has been found overexpressed in different types of tumors (Hamamoto et al., 2004). SMYD3 oncogenic activity involves H3K4 methylation-elicited transcriptional activation, and functional interactions with non-histone proteins in the cytoplasm that regulate cancer cell proliferation (Peserico et al., 2016). In *P. falciparum*, immunofluorescence staining showed that PfSET6 localizes to nuclear and cytoplasmic compartments during blood stage growth. This indicated that, while in ring stages, its substrates may be localized to either cellular compartment, in mature stages, most PfSET6 substrates reside in the cytoplasm and are likely to be non-histones. *Bioinformatics* analysis of PfSET6 proteins to *identified* both nuclear localization signals (NLSs; residues 160-192) and nuclear export signals (NESs; at the C-terminal), suggesting that PfSET6 might shuttle between between the nucleus and the cytoplasm through coordination of select transport pathways.

Notably, the partial enrichment of PfSET6 in cytoplasmic foci containing microneme components or Alba-defined mRNA granules suggested that PfSET6 could methylate either erythrocyte invasion components or mRNA-protein complexes under post-transcriptional control. In fact, proteomic studies have revealed that some of these proteins, for example the RNA-binding Alba proteins are methylated (Ghosh-Dastidar and Scherf, unpublished). PfSET6 could be the effector of this methylation, which can be confirmed by performing the *in vitro* methyltransferase assays with recombinant Albas as substrates.

Given PfSET6's expression in the nuclear fraction of ring stages and its partial co-localization with the epigenetic regulator PfHP1 in immunofluorescence assays, we performed ChIP-seq assays to identify its genome-wide occupancy. Different from predicted H3K4 methyltransferases that are responsible for transcriptional activation (Alvarez-Venegas and Avramova, 2002; Santos-Rosa et al., 2002) PfSET6 associates with particular genomic regions in a manner that correlates mostly with transcriptional repression. In select cases, PfSET6 occupancy coincided with the repressive chromatin marks PfHP1 and H3K9me3, supporting its localization pattern to repression foci in ring stage parasites. However, it is important to note that our formaldehyde-based crosslinking

ChIP-seq might not have fully captured PfSET6-chromatin interactions. Indeed, while the PfSET6 occupied sites revealed by our analysis are predictive of the types of genes that PfSET6 targets, we cannot rule out other short-lived interactions. To reveal the full extent of PfSET6 occupancy, alternate methods can be applied such as treating cells with bifunctional membrane-permeating cross-linkers to preserve the protein-protein association before the standard formaldehyde crosslinking, which has been used to analyze the genome-wide distribution of HATs and HDACs in human CD4<sup>+</sup> T cells (Wang et al., 2009).

Although the PfSET6 ortholog SMYD3 was initially described as a histone H3K4-specific di- and tri-methyltransferase *in vitro* (see above), recent studies showed that SMYD3 methylates histone H4K5, and this site serves as a far more effective substrate than H3, particularly in a nucleosome context (van Aller et al., 2012). Another cardiogenesis study showed that m-Bop, a MYND and SET domains-containing muscle-restricted protein, facilitates the development of ventricular myocytes by recruitment of HDACs and functioning as a transcriptional repressor (Gottlieb et al., 2002). Moreover, it has previously been shown that SMYD3 proteins optimally function in the presence of co-factors, such as chaperone hsp90 in cancerous cells (Silva et al., 2008). This suggests that *Plasmodium* SET6 may require to interact with co-factors for full activity. Taken together with our observations that PfSET6 weakly methylates histone H3 *in vitro* and the nuclear and cytoplasmic localization of PfSET6, future work with the recombinant protein could determine its preferential methylation site in histones and non-histone substrates, both in the presence and absence of co-factors. Considering that the nature of the enzyme-substrate interactions is transient and most enzymes have a slew of target proteins, co-immunoprecipitation with specific anti-PfSET6 antibodies (Free et al., 2009) or the so-called BioID method, in which the enzyme is tagged with biotin ligase domain of birA and associated proteins identified by proximity-dependent biotinylation (Roux et al., 2013; Schweingruber et al., 2016), could be used to identify the interacting partners of PfSET6.



## 7. Perspective

Our work raises a number of new and important questions: How are SET domain-containing proteins recruited to different cellular sites and what is their precise molecular mechanism of action. What are the substrates at these cellular sites? Are there specific co-factors that regulate their function *in vivo*? When and how is lysine methyltransferase activity is switched on in multiple life cycle stages? More generally, can we gain further insights into the complex modes of epigenetic regulation that are utilized by this devastating parasites?

To study this in more detail, we aimed to conditional knockdown PfSET7 and PfSET6 using the *glmS* ribozyme system, which is an inducible mRNA degradation system. In principle, the ribozyme sequence is inserted into the 3'-UTR of the coding region of a gene of interest. In response to sugar treatment, ribozyme self-cleavage occurs, resulting in degradation of the corresponding mRNA and knock-down of protein expression. Shaw et al. have demonstrated that the *glmS* ribozyme is active in *P. falciparum* and can be used as a tool to modulate target gene expression (Prommana et al., 2013). In the case of PfSET7, we successfully inserted an HA-tag followed by the *glmS* ribozyme using a CRISPR/Cas9-based genome editing approach (Ghorbal et al., 2014; MacPherson and Scherf, 2015) and detected tagged protein expression. Unfortunately, after several attempts with varying concentrations of the sugar and incubation times, we did not obtain efficient PfSET7 mRNA and protein reduction and the treated parasites displayed no obvious growth phenotype. We also tried to insert different-sized fragments of a heterologous 3'UTR sequence that allowed optimal *glmS* activity in an episomal context into the 3'UTR of *set7*, but were unable to generate such transgenic parasites. Lastly, when applied to PfSET6, we never obtained transgenic parasites modified at the 3' end, indication that the 3' end sequence may be crucial for function of PfSET6 in blood stage.

Recently, we applied a new strategy based on the inducible DiCre/lox system to excise the SET domain for PfSET7 and PfSET6, which has been proven efficient to study essential genes in *P. falciparum* by Treeck and colleagues (Collins et al., 2013; Jones et al., 2016). Briefly, this strategy involves the placement of loxP sites on either side of a region of interest: in most cases, in the context of a heterologous intron at the 5' end and within the 3'UTR, downstream of the STOP codon, at the 3' end. Upon induction of the DiCre recombinase, site-specific removal of the gene sequence between the two loxP sites would generate a gene knockout and allow rapid functional analysis. I have generated a loxP-intron-modified *set6*

transgenic lines and am currently characterizing its response to DiCre induction; I have transfected parasite lines with the SET7 constructs and await the results. I anticipate that in the near future, these strains will serve as tools to study the impact of PfSET6 and PfSET7 on regulating the blood stage transcriptome and proteome-wide methylation, or even shed the light on the mechanisms of action for PfSET7 and PfSET6 in parasite survival and development.

In addition, having produced enzymatically active full-length PfSET7, we are now in a position to perform target-based screening to identify specific small molecule inhibitors for PfSET7. These can help to determine the biological role of PfSET7 in sporozoite and liver stage parasites, which are much less amenable to genetic manipulation. This includes the fascinating biology of malaria relapses in *P. vivax* and *P. ovale* infections.

During my PhD period, we generated a new, highly specific antibody against PfHP1, and validated it in different functional assays including western blotting, immunofluorescence and ChIP-seq. Several of these studies were the basis for a collaborative project with Gigliola Zanghi in Dominique Mazier's laboratory, where we analyzed the role of PfHP1 in sporozoite epigenetic gene regulation (see Article III, page 104). This work opened new avenues revealing clonally variant expression of *var* genes in sporozoites and exposure of PfEMP1 on the surface of sporozoites, possibly as a sensor of the extracellular milieu. This novel aspect has important implications not only for the biology of sporozoites, but also for vaccine development.

Since the CRISPR-Cas9 system was developed in Scherf's laboratory (Ghorbal et al., 2014; MacPherson and Scherf, 2015), a new idea to develop a reporter system to screen for defects in *var* monoallelic expression came to us. As part of my PhD work, I exploited the CRISPR-Cas9 technology to generate a PfEMP1-dual-labeled transgenic parasite line, which could potentially be induced by drug treatment or physical and chemical mutagenesis (see Result IV, page 126). First, I inserted GFP into the endogenous coding region of exon I of a *var* gene, in-frame, and successfully detected PfEMP1-GFP expression on the surface of asexual stage parasites, establishing the proof of principle for the project. Then I designed constructs to insert RFP into a second *var* gene and obtained PfEMP1-RFP transgenic parasites. The next step would be the transfection of the plasmid driving PfEMP1-RFP expression into PfEMP1-0412700-GFP transgenic parasites, and the resultant parasites would exhibit a double fluorescent profile only when monoallelic expression is disrupted. This part has not

been realized as yet due to time constraints to finish the PfSET7 and PfSET6 projects. However, once these PfEMP1-dual-labeled parasites are successfully generated, they could be used to explore empirically for novel factors that are involved in monoallelic expression.

## **8. Conclusion**

This PhD work underlines the concept that computational prediction of protein function needs experimental validation to explore the full extent of SET protein cellular targets and biological function in Apicomplexan parasites. It remains to be seen how many of the 10 predicted SET domain proteins are specific for histone methylation only. Clearly, our study reveals that non-histone methylation is much more important in *P. falciparum* than previously anticipated. Systematic mass spectrometry to detect methylated proteins in combination with inducible gene knockout mutant parasites may help to identify targets of SET domain containing proteins. The fact that at least PfSET7 and PfSET6 are expressed in different life cycle stages, makes them as novel targets for drug development that could against severe disease and to block pathogen transmission.

## 9. Bibliography

- Aller, G.S., Reynoird, N., Barbash, O., Huddleston, M., Liu, S., Zmoos, A.F., McDevitt, P., Sinnamon, R., Le, B., Mas, G., et al. (2012). Smyd3 regulates cancer cell phenotypes and catalyzes histone H4 lysine 5 methylation. *Epigenetics* 7, 340–343.
- Alvarez, F., Muñoz, F., Schilcher, P., Imhof, A., Almouzni, G., and Loyola, A. (2011). Sequential establishment of marks on soluble histones H3 and H4. *J. Biol. Chem.* 286, 17714–17721.
- Alvarez-Venegas, R., and Avramova, Z. (2002). SET-domain proteins of the Su(var)3-9, E(z) and Trithorax families. *Gene* 285, 25–37.
- Annoura, T., Van Schaijk, B.C.L., Ploemen, I.H.J., Sajid, M., Lin, J.W., Vos, M.W., Dinmohamed, A.G., Inaoka, D.K., Rijpma, S.R., Van Gemert, G.J., et al. (2014). Two Plasmodium 6-Cys family-related proteins have distinct and critical roles in liver-stage development. *FASEB J.* 28, 2158–2170.
- Ariey, F., Witkowski, B., Amaratunga, C., Beghain, J., Langlois, A.-C., Khim, N., Kim, S., Duru, V., Bouchier, C., Ma, L., et al. (2013). A molecular marker of artemisinin-resistant Plasmodium falciparum malaria. *Nature* 505, 50–55.
- Ashley, E. a., Dhorda, M., Fairhurst, R.M., Amaratunga, C., Lim, P., Suon, S., Sreng, S., Anderson, J.M., Mao, S., Sam, B., et al. (2014). Spread of Artemisinin Resistance in Plasmodium falciparum Malaria. *N. Engl. J. Med.* 371, 411–423.
- Ayyanathan, K., Ayyanathan, K., Lechner, M.S., Lechner, M.S., Bell, P., Bell, P., Maul, G.G., Maul, G.G., Schultz, D.C., and Schultz, D.C. (2003). Heterochromatin protein 1 (HP1) is a key component of constitutive heterochromatin in. *Genes Dev.* 1855–1869.
- Bártfai, R., Hoeijmakers, W.A.M., Salcedo-Amaya, A.M., Smits, A.H., Janssen-Megens, E., Kaan, A., Treeck, M., Gilberger, T.W., Francoijs, K.J., and Stunnenberg, H.G. (2010). H2A.Z demarcates intergenic regions of the Plasmodium falciparum epigenome that are dynamically marked by H3K9ac and H3K4me3. *PLoS Pathog.* 6.
- Balaji, S., Madan Babu, M., Iyer, L.M., and Aravind, L. (2005). Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.* 33, 3994–4006.
- Barry, A.E., and Arnott, A. (2014). Strategies for designing and monitoring malaria vaccines targeting diverse antigens. *Front. Immunol.* 5, 1–16.
- Birney, E., Stamatoyannopoulos, J. a, Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447, 799–816.
- Blount, R.E. (1967). Management of chloroquine resistant falciparum malaria. *Trans Am Clin*

Clim. Assoc 78, 196–204.

Bozdech, Z., Llinás, M., Pulliam, B.L., Wong, E.D., Zhu, J., and DeRisi, J.L. (2003a). The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biol.* 1, 1–16.

Bozdech, Zbynek; Preiser, P. (2013). No Title. In *Malaria Parasites: Comparative Genomics, Evolution and Molecular Biology*, S.L.D.K.W. Carlton, Jane M; Perkins, ed. p. 96.

Brancucci, N.M.B., Bertschi, N.L., Zhu, L., Niederwieser, I., Chin, W.H., Wampfler, R., Freymond, C., Rottmann, M., Felger, I., Bozdech, Z., et al. (2014). Heterochromatin protein 1 secures survival and transmission of malaria parasites. *Cell Host Microbe* 16, 165–176.

Campos, E.I., Fillingham, J., Li, G., Zheng, H., Voigt, P., Kuo, W.-H.W., Seepany, H., Gao, Z., Day, L. a, Greenblatt, J.F., et al. (2010). The program for processing newly synthesized histones H3.1 and H4. *Nat. Struct. Mol. Biol.* 17, 1343–1351.

Carter, R. (2001). Transmission blocking malaria vaccines. *Vaccine* 19, 2309–2314.

Chen, P.B., Ding, S., Zanghi, G., Soulard, V., DiMaggio, P.A., Fuchter, M.J., Mecheri, S., Mazier, D., Scherf, A., and Malmquist, N.A. (2016). *Plasmodium falciparum* PfSET7: enzymatic characterization and cellular localization of a novel protein methyltransferase in sporozoite, liver and erythrocytic stage parasites. *Sci. Rep.* 6, 21802.

Chookajorn, T., Dzikowski, R., Frank, M., Li, F., Jiwani, A.Z., Hartl, D.L., and Deitsch, K.W. (2007). Epigenetic memory at malaria virulence genes. *Proc. Natl. Acad. Sci. U. S. A.* 104, 899–902.

Chuikov, S., Kurash, J.K., Wilson, J.R., Xiao, B., Justin, N., Ivanov, G.S., Mckinney, K., Tempst, P., Prives, C., Gamblin, S.J., et al. (2004). Regulation of p53 activity through lysine methylation. *Nature* 432, 353–360.

Clouaire, T., and Stancheva, I. (2008). Methyl-CpG binding proteins: Specialized transcriptional repressors or structural components of chromatin? *Cell. Mol. Life Sci.* 65, 1509–1522.

Coleman, B.I., Skillman, K.M., Jiang, R.H.Y., Childs, L.M., Altenhofen, L.M., Ganter, M., Leung, Y., Goldowitz, I., Kafsack, B.F.C., Marti, M., et al. (2014). A *Plasmodium falciparum* Histone Deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host Microbe* 16, 177–186.

Collins, C.R., Das, S., Wong, E.H., Andenmatten, N., Stallmach, R., Hackett, F., Herman, J.P., Müller, S., Meissner, M., and Blackman, M.J. (2013). Robust inducible Cre recombinase activity in the human malaria parasite *Plasmodium falciparum* enables efficient gene deletion within a single asexual erythrocytic growth cycle. *Mol. Microbiol.*

Cui, L., Miao, J., and Cui, L. (2007). Cytotoxic effect of curcumin on malaria parasite *Plasmodium falciparum*: Inhibition of histone acetylation and generation of reactive oxygen species. *Antimicrob. Agents Chemother.* 51, 488–494.

- Cui, L.L., Fan, Q., Cui, L.L., and Miao, J. (2008). Histone lysine methyltransferases and demethylases in *Plasmodium falciparum*. *Int. J. Parasitol.* 38, 1083–1097.
- Dastidar, E.G., Dayer, G., Holland, Z.M., Dorin-Semblat, D., Claes, A., Chene, A., Sharma, A., Hamelin, R., Moniatte, M., Lopez-Rubio, J.-J., et al. (2012). Involvement of *Plasmodium falciparum* protein kinase CK2 in the chromatin assembly pathway. *BMC Biol.* 10, 5.
- Dastidar, E.G., Dzek, K., Krijgsveld, J., Malmquist, N. a, Doerig, C., Scherf, A., and Lopez-Rubio, J.J. (2013). Comprehensive Histone Phosphorylation Analysis and Identification of Pf14-3-3 Protein as a Histone H3 Phosphorylation Reader in Malaria Parasites. *PLoS One* 8, e53179.
- Demarta-Gatsi, C., Smith, L., Thiberge, S., Peronet, R., Commere, P.-H., Matondo, M., Apetoh, L., Bruhns, P., Ménard, R., and Mécheri, S. (2016). Protection against malaria in mice is induced by blood stage–arresting histamine-releasing factor ( HRF )–deficient parasites. *J. Exp. Med.* jem.20151976.
- Dembele, L., Gego, A., Zeeman, A.M., Franetich, J.F., Silvie, O., Rametti, A., Le Grand, R., Dereuddre-Bosquet, N., Sauerwein, R., van Gemert, G.J., et al. (2011). Towards an in vitro model of plasmodium hypnozoites suitable for drug discovery. *PLoS One* 6, 1–7.
- Dembélé, L., Franetich, J.-F., Lorthiois, A., Gego, A., Zeeman, A.-M., Kocken, C.H.M., Le Grand, R., Dereuddre-Bosquet, N., van Gemert, G.-J., Sauerwein, R., et al. (2014). Persistence and activation of malaria hypnozoites in long-term primary hepatocyte cultures. *Nat. Med.* 20, 307–312.
- Doerig, C., Rayner, J.C., Scherf, A., and Tobin, A.B. (2015). Post-translational protein modifications in malaria parasites. *Nat. Rev. Microbiol.* 13, 160–172.
- Dondorp, A.M., Nosten, F., Yi, P., Das, D., Phyo, A.P., Tarning, J., Ph, D., Lwin, K.M., Arie, F., Hanpithakpong, W., et al. (2009). Artemisinin Resistance in. *Drug Ther. (NY)*. 361, 455–467.
- Druilhe, P., and Pérignon, J.L. (1997). A hypothesis about the chronicity of malaria infection. *Parasitol. Today* 13, 353–357.
- Duraisingh, M.T., Voss, T.S., Marty, A.J., Duffy, M.F., Good, R.T., Thompson, J.K., Freitas, L.H., Scherf, A., Crabb, B.S., and Cowman, A.F. (2005). Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. *Cell* 121, 13–24.
- Eberharter, A., Vetter, I., Ferreira, R., and Becker, P.B. (2004). ACF1 improves the effectiveness of nucleosome mobilization by ISWI through PHD-histone contacts. *EMBO J.* 23, 4029–4039.
- Ekland, E.H., and Fidock, D.A. (2008). In vitro evaluations of antimalarial drugs and their relevance to clinical outcomes. *Int. J. Parasitol.* 38, 743–747.
- Fan, Q., An, L., and Cui, L. (2004). *Plasmodium falciparum*. *Eukaryot. Cell* 3, 264–276.

Fan, Q., Miao, J., Cui, L., and Cui, L. (2009). Characterization of PRMT1 from *Plasmodium falciparum*. *Biochem. J.* 421, 107–118.

Flueck, C., and Baker, D.A. (2014). Malaria parasite epigenetics: When virulence and romance collide. *Cell Host Microbe*.

Flueck, C., Bartfai, R., Volz, J., Niederwieser, I., Salcedo-Amaya, A.M., Alako, B.T.F., Ehlgren, F., Ralph, S.A., Cowman, A.F., Bozdech, Z., et al. (2009). *Plasmodium falciparum* heterochromatin protein 1 marks genomic loci linked to phenotypic variation of exported virulence factors. *PLoS Pathog.* 5, e1000569.

Foth, B.J., Zhang, N., Chahal, B.K., Sze, S.K., Preiser, P.R., and Bozdech, Z. (2011). Quantitative time-course profiling of parasite and host cell proteins in the human malaria parasite *Plasmodium falciparum*. *Mol. Cell. Proteomics* 10, M110.006411.

Fraschka, S.A.-K., Henderson, R.W.M., and Bartfai, R. (2016). H3.3 demarcates GC-rich coding and subtelomeric regions and serves as potential memory mark for virulence gene expression in *Plasmodium falciparum*. *Sci. Rep.* 6, 31965.

Freitas, L.H., Hernandez-Rivas, R., Ralph, S.A., Montiel-Condado, D., Ruvalcaba-Salazar, O.K., Rojas-Meza, A.P., Mâncio-Silva, L., Leal-Silvestre, R.J., Gontijo, A.M., Shorte, S., et al. (2005). Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* 121, 25–36.

Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* 407, 1018–1022.

Freitas-Junior, L.H., Hernandez-Rivas, R., Ralph, S.A., Montiel-Condado, D., Ruvalcaba-Salazar, O.K., Rojas-Meza, A.P., Mâncio-Silva, L., Leal-Silvestre, R.J., Gontijo, A.M., Shorte, S., et al. (2005). Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* 121, 25–36.

French, J.B., Cen, Y., and Sauve, A.A. (2008). *Plasmodium falciparum* Sir2 is an NAD<sup>+</sup>-dependent deacetylase and an acetyllysine-dependent and acetyllysine-independent NAD<sup>+</sup> glycohydrolase. *Biochemistry* 47, 10227–10239.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., et al. (2002). Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* 419, 498–511.

Ghorbal, M., Gorman, M., Macpherson, C.R., Martins, R.M., Scherf, A., and Lopez-Rubio, J.J. (2014). Genome editing in the human malaria parasite *Plasmodium falciparum* using the CRISPR-Cas9 system. *Nat Biotechnol* 32, 819–821.

Gottlieb, P.D., Pierce, S. a, Sims, R.J., Yamagishi, H., Weihe, E.K., Harriss, J. V, Maika, S.D., Kuziel, W. a, King, H.L., Olson, E.N., et al. (2002). Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and

morphogenesis. *Nat. Genet.* 31, 25–32.

Greenwood, B.M. (2015). Efficacy and safety of RTS,S/AS01 malaria vaccine with or without a booster dose in infants and children in Africa: Final results of a phase 3, individually randomised, controlled trial. *Lancet* 386, 31–45.

Grewal, S.I.S., and Jia, S. (2007). Heterochromatin revisited. *Nat. Rev. Genet.* 8, 35–46.

Groth, A., Rocha, W., Verreault, A., and Almouzni, G. (2007). Chromatin Challenges during DNA Replication and Repair. *Cell* 128, 721–733.

Guizetti, J., and Scherf, A. (2013). Silence, activate, poise and switch! Mechanisms of antigenic variation in *Plasmodium falciparum*. *Cell. Microbiol.*

Guizetti, J., Barcons-Simon, A., and Scherf, A. (2016). Trans-acting GC-rich non-coding RNA at var expression site modulates gene counting in malaria parasite. *Nucleic Acids Res.* in press, 1–9.

Hamamoto, R., Furukawa, Y., Morita, M., Iimura, Y., Silva, F.P., Li, M., Yagyu, R., and Nakamura, Y. (2004). SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat. Cell Biol.* 6, 731–740.

Hamamoto, R., Saloura, V., and Nakamura, Y. (2015). Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat. Publ. Gr.* 15.

Heaslip, A.T., Nishi, M., Stein, B., and Hu, K. (2011). The motility of a human parasite, *Toxoplasma gondii*, is regulated by a novel lysine methyltransferase. *PLoS Pathog.*

Horrocks, P., Pinches, R., Christodoulou, Z., Kyes, S.A., and Newbold, C.I. (2004). Variable var transition rates underlie antigenic variation in malaria. *Proc. Natl. Acad. Sci. U. S. A.* 101, 11129–11134.

Hossain, M.J., Korde, R., Singh, S., Mohammed, A., Dasaradhi, P.V.N., Chauhan, V.S., and Malhotra, P. (2008). Tudor domain proteins in protozoan parasites and characterization of *Plasmodium falciparum* tudor staphylococcal nuclease. *Int. J. Parasitol.* 38, 513–526.

Howden, B.P., Vaddadi, G., Manitta, J., and Grayson, M.L. (2005). Chronic *falciparum* malaria causing massive splenomegaly 9 years after leaving an endemic area. *Med. J. Aust.* 182, 186–188.

Huebert, D.J., Kuan, P.-F., Keleş, S., and Gasch, A.P. (2012). Dynamic changes in nucleosome occupancy are not predictive of gene expression dynamics but are linked to transcription and chromatin regulators. *Mol. Cell. Biol.* 32, 1645–1653.

Hviid, L. (2010). The role of *Plasmodium falciparum* variant surface antigens in protective immunity and vaccine development. *Hum. Vaccin.* 6, 84–89.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* (80-. ). 293, 1074–1080.



Jiang, L., Mu, J., Zhang, Q., Ni, T., Srinivasan, P., Rayavara, K., Yang, W., Turner, L., Lavstsen, T., Theander, T.G., et al. (2013). PfSETvs methylation of histone H3K36 represses virulence genes in *Plasmodium falciparum*. *Nature* 499.

Jin, S.G., Kadam, S., and Pfeifer, G.P. (2010). Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res.* 38, 1–7.

Jones, D.O., Cowell, I.G., and Singh, P.B. (2000). Mammalian chromodomain proteins: Their role in genome organisation and expression. *BioEssays* 22, 124–137.

Jones, M.L., Das, S., Belda, H., Collins, C.R., Blackman, M.J., and Treeck, M. (2016). A versatile strategy for rapid conditional genome engineering using loxP sites in a small synthetic intron in *Plasmodium falciparum*. *Nat. Sci. Reports* 6, 21800.

Joshi, M.B., Lin, D.T., Chiang, P.H., Goldman, N.D., Fujioka, H., Aikawa, M., and Syin, C. (1999). Molecular cloning and nuclear localization of a histone deacetylase homologue in *Plasmodium falciparum*. *Mol. Biochem. Parasitol.* 99, 11–19.

Josling, G. a, and Llinás, M. (2015). Sexual development in *Plasmodium* parasites: knowing when it's time to commit. *Nat. Rev. Microbiol.* 13, 573–587.

Josling, G.A., Petter, M., Oehring, S.C., Gupta, A.P., Dietz, O., Wilson, D.W., Schubert, T., Längst, G., Gilson, P.R., Crabb, B.S., et al. (2015). A *Plasmodium Falciparum* Bromodomain Protein Regulates Invasion Gene Expression. *Cell Host Microbe* 17, 741–751.

Kafsack, B.F.C., Rovira-Graells, N., Clark, T.G., Bancells, C., Crowley, V.M., Campino, S.G., Williams, A.E., Drought, L.G., Kwiatkowski, D.P., Baker, D.A., et al. (2014). A transcriptional switch underlies commitment to sexual development in malaria parasites. *Nature* 507, 248–252.

Kensche, P.R., Hoeijmakers, W.A.M., Toenhake, C.G., Bras, M., Chappell, L., Berriman, M., and B?rtfai, R. (2015). The nucleosome landscape of *Plasmodium falciparum* reveals chromatin architecture and dynamics of regulatory sequences. *Nucleic Acids Res.* 44, 2110–2124.

Kishore, S.P., Stiller, J.W., and Deitsch, K.W. (2013). Horizontal gene transfer of epigenetic machinery and evolution of parasitism in the malaria parasite *Plasmodium falciparum* and other apicomplexans. *BMC Evol. Biol.* 13, 37.

Kraemer, S.M., Kyes, S. a, Aggarwal, G., Springer, A.L., Nelson, S.O., Christodoulou, Z., Smith, L.M., Wang, W., Levin, E., Newbold, C.I., et al. (2007). Patterns of gene recombination shape var gene repertoires in *Plasmodium falciparum*: comparisons of geographically diverse isolates. *BMC Genomics* 8, 45.

Kyes, S., Christodoulou, Z., Pinches, R., Kriek, N., Horrocks, P., and Newbold, C. (2007). *Plasmodium falciparum* var gene expression is developmentally controlled at the level of RNA polymerase II-mediated transcription initiation. *Mol. Microbiol.* 63, 1237–1247.

Kyes, S.A., Christodoulou, Z., Raza, A., Horrocks, P., Pinches, R., Rowe, J.A., and Newbold,

- C.I. (2003). A well-conserved *Plasmodium falciparum* var gene shows an unusual stage-specific transcript pattern. *Mol. Microbiol.* 48, 1339–1348.
- Lachner, M., O’Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116–120.
- Lanouette, S., Mongeon, V., Figeys, D., and Couture, J.F. (2014). The functional diversity of protein lysine methylation. *Mol. Syst. Biol.* 10, 1–26.
- Lavstsen, T., Salanti, A., Jensen, A.T.R., Arnot, D.E., and Theander, T.G. (2003). Subgrouping of *Plasmodium falciparum* 3D7 var genes based on sequence analysis of coding and non-coding regions. *Malar. J.* 2, 27.
- Lee, K.K., and Workman, J.L. (2007). Histone acetyltransferase complexes: one size doesn’t fit all. *Nat. Rev. Mol. Cell Biol.* 8, 284–295.
- Lopez-Rubio, J.J., Gontijo, A.M., Nunes, M.C., Issar, N., Hernandez Rivas, R., and Scherf, A. (2007). 5' Flanking Region of Var Genes Nucleate Histone Modification Patterns Linked To Phenotypic Inheritance of Virulence Traits in Malaria Parasites. *Mol. Microbiol.* 66, 1296–1305.
- Lopez-Rubio, J.J., Mancio-Silva, L., and Scherf, A. (2009). Genome-wide Analysis of Heterochromatin Associates Clonally Variant Gene Regulation with Perinuclear Repressive Centers in Malaria Parasites. *Cell Host Microbe* 5, 179–190.
- Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 Variants before Chromatin Assembly Potentiate Their Final Epigenetic State. *Mol. Cell* 24, 309–316.
- MacPherson, C.R., and Scherf, A. (2015). Flexible guide-RNA design for CRISPR applications using protospacer workbench. *Nat. Biotechnol.* 33, 1–2.
- Malmquist, N.A., Moss, T.A., Mecheri, S., Scherf, A., and Fuchter, M.J. (2012). Small-molecule histone methyltransferase inhibitors display rapid antimalarial activity against all blood stage forms in *Plasmodium falciparum*. *Proc. Natl. Acad. Sci.* 109, 16708–16713.
- Malmquist, N.A., Sundriyal, S., Caron, J., Chen, P., Witkowski, B., Menard, D., Suwanarusk, R., Renia, L., Nosten, F., Jiménez-Díaz, M.B., et al. (2015). Histone methyltransferase inhibitors are orally bioavailable, fast- Acting molecules with activity against different species causing malaria in humans. *Antimicrob. Agents Chemother.*
- Mancio-Silva, L., Rojas-Meza, A.P., Vargas, M., Scherf, A., and Hernandez-Rivas, R. (2008). Differential association of Orc1 and Sir2 proteins to telomeric domains in *Plasmodium falciparum*. *J Cell Sci* 121, 2046–2053.
- Mancio-Silva, L., Lopez-Rubio, J.J., Claes, A., and Scherf, A. (2013). ARTICLE Sir2a regulates rDNA transcription and multiplication rate in the human malaria parasite *Plasmodium falciparum*.
- Mazur, P.K., Reynoird, N., Khatri, P., Jansen, P.W.T.C., Wilkinson, A.W., Liu, S., Barbash,

- O., Van Aller, G.S., Huddleston, M., Dhanak, D., et al. (2014). SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature* 510, 283–287.
- Merrick, C.J., Huttenhower, C., Buckee, C., Amambua-Ngwa, A., Gomez-Escobar, N., Walther, M., Conway, D.J., and Duraisingh, M.T. (2012). Epigenetic dysregulation of virulence gene expression in severe plasmodium falciparum malaria. *J. Infect. Dis.* 205, 1593–1600.
- Miao, J., Fan, Q., Cui, L., Li, J., Li, J., and Cui, L. (2006). The malaria parasite Plasmodium falciparum histones: Organization, expression, and acetylation. *Gene* 369, 53–65.
- Miao, J., Fan, Q., Cui, L., Li, X., and Wang, H. (2010). The MYST family histone acetyltransferase regulates gene expression and cell cycle in malaria parasite Plasmodium falciparum. *Mol. Microbiol.* 78, 883–902.
- Munro, S., Khaire, N., Inche, A., Carr, S., and La Thangue, N.B. (2010). Lysine methylation regulates the pRb tumour suppressor protein. *Oncogene* 29, 2357–2367.
- Nganou-Makamdop, K., and Sauerwein, R.W. (2013). Liver or blood-stage arrest during malaria sporozoite immunization: The later the better? *Trends Parasitol.* 29, 304–310.
- Nguyen, A.T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. [*Genes Dev.* 2011] - PubMed - NCBI. *Genes Dev.* 3, 1345–1358.
- Organization, W.H. (2006). Guidelines for the treatment of malaria. *Who/Htm/Mal/2006.1108* Isbn 978 92 4 154694 2 266 p.
- Owen, D.J., Ornaghi, P., Yang, J.C., Lowe, N., Evans, P.R., Ballario, P., Neuhaus, D., Filetici, P., and Travers, a (2000). The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J.* 19, 6141–6149.
- Partnership, S.C.T., Agnandji, S.T., Lell, B., Fernandes, J.F., Abossolo, B.P., Methogo, B.G.N.O., Kabwende, A.L., Adegnik, A.A., Mordmüller, B., Issifou, S., et al. (2012). A phase 3 trial of RTS,S/AS01 malaria vaccine in African infants. *N. Engl. J. Med.* 367, 2284–2295.
- Pérez-Toledo, K., Rojas-Meza, A.P., Mancio-Silva, L., Hernández-Cuevas, N.A., Delgadillo, D.M., Vargas, M., Martínez-Calvillo, S., Scherf, A., and Hernandez-Rivas, R. (2009). Plasmodium falciparum heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. *Nucleic Acids Res.* 37, 2596–2606.
- Peserico, A., Germani, A., Sanese, P., Barbosa, A.J., Di, V., Fittipaldi, R., Fabini, E., Bertucci, C., Varchi, G., Pat, M., et al. (2016). A SMYD3 Small-Molecule Inhibitor Impairing Cancer Cell Growth. *230*, 2447–2460.
- Petter, M., Lee, C.C., Byrne, T.J., Boysen, K.E., Volz, J., Ralph, S.A., Cowman, A.F., Brown, G. V., and Duffy, M.F. (2011). Expression of P. falciparum var genes involves exchange of the histone variant H2A.Z at the promoter. *PLoS Pathog.*

- Petter, M., Selvarajah, S.A., Lee, C.C., Chin, W.H., Gupta, A.P., Bozdech, Z., Brown, G. V., and Duffy, M.F. (2013). H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy var gene promoters in the malaria parasite *Plasmodium falciparum*. *Mol. Microbiol.* 87, 1167–1182.
- Pontivianne, F., Blevins, T., and Pikaard, C.S. (2010). *Arabidopsis* Histone Lysine Methyltransferases. *Adv. Bot. Res.* 2296, 1–18.
- Ponts, N., Harris, E., and Prudhomme, J. (2010). Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Res.* 20, 228–238.
- Ponts, N., Fu, L., Harris, E.Y., Zhang, J., Chung, D.-W.D., Cervantes, M.C., Prudhomme, J., Atanasova-Penichon, V., Zehraoui, E., Bunnik, E.M., et al. (2013). Genome-wide mapping of DNA methylation in the human malaria parasite *Plasmodium falciparum*. *Cell Host Microbe* 14, 696–706.
- Ponts N1, Harris EY, Lonardi S, L.R.K. (2011). Nucleosome occupancy at transcription start sites in the human malaria parasite: a hard-wired evolution of virulence? *Infect Genet Evol.* 2011, 716–724.
- Prommana, P., Uthaipibull, C., Wongsombat, C., Kamchonwongpaisan, S., Yuthavong, Y., Knuepfer, E., Holder, A.A., and Shaw, P.J. (2013). Inducible Knockdown of *Plasmodium* Gene Expression Using the glmS Ribozyme. *PLoS One* 8.
- Qian, C., and Zhou, M.M. (2006). SET domain protein lysine methyltransferases: Structure, specificity and catalysis. *Cell. Mol. Life Sci.* 63, 2755–2763.
- Ralph, S.A., Scheidig-Benatar, C., and Scherf, A. (2005). Antigenic variation in *Plasmodium falciparum* is associated with movement of var loci between subnuclear locations. *Proc. Natl. Acad. Sci. U. S. A.* 102, 5414–5419.
- Rea, S., Eisenhaber, F., O’Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599.
- Le Roch, K.G., Zhou, Y., Blair, P.L., Grainger, M., Moch, J.K., Haynes, J.D., De La Vega, P., Holder, A.A., Batalov, S., Carucci, D.J., et al. Discovery of Gene Function by Expression Profiling of the Malaria Parasite Life Cycle.
- Luo, M. (2011). Bioorthogonal Profiling of Protein Methylation Using Azido Derivative of S-Adenosyl-L-methionine. *J. Am. Chem. Soc.* 134, 5909–5915.
- Rea, S., Eisenhaber, F., O’Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D., et al. (2000). Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* 406, 593–599.
- Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *Nat. Rev. Mol. Cell Biol.* 8, 983–994.
- Saksouk, N., Bhatti, M.M., Kieffer, S., Aaron, T., Musset, K., Garin, J., Jr, W.J.S., Hakimi, M., and Smith, A.T. (2005). Histone-Modifying Complexes Regulate Gene Expression

Pertinent to the Differentiation of the Protozoan Parasite *Toxoplasma gondii* Histone-Modifying Complexes Regulate Gene Expression Pertinent to the Differentiation of the Protozoan Parasite *Toxoplasma go.* *Mol. Cell. Biol.* 25, 10301–10314.

Salcedo-Amaya, A.M., van Driel, M.A., Alako, B.T., Trelle, M.B., van den Elzen, A.M.G., Cohen, A.M., Janssen-Megens, E.M., van de Vegte-Bolmer, M., Selzer, R.R., Iniguez, A.L., et al. (2009). Dynamic histone H3 epigenome marking during the intraerythrocytic cycle of *Plasmodium falciparum*. *Proc. Natl. Acad. Sci. U. S. A.* 106, 9655–9660.

Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C.T., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature* 419, 407–411.

Saraf, A., Cervantes, S., Bunnik, E.M., Ponts, N.P., Sardi, M.E., Chung, D.-W.D., Prudhomme, J., Varberg, J., Wen, Z., Washburn, M.P., et al. (2016). Dynamic and combinatorial landscape of histone modifications during the intra-erythrocytic developmental cycle of the malaria parasite. *J. Proteome Res.* [acs.jproteome.6b00366](https://doi.org/10.1021/acs.jproteome.6b00366).

van Schaijk, B.C.L., Ploemen, I.H.J., Annoura, T., Vos, M.W., Lander, F., van Gemert, G.-J., Chevalley-Maurel, S., van de Vegte-Bolmer, M., Sajid, M., Franetich, J.-F., et al. (2014). A genetically attenuated malaria vaccine candidate based on *P. falciparum* b9/slarp gene-deficient sporozoites. *Elife* 3, 1–18.

Scherf, A., Lopez-Rubio, J.J., and Riviere, L. (2008a). Antigenic variation in *Plasmodium falciparum*. *Annu. Rev. Microbiol.* 62, 445–470.

Scherf, A., Rivière, L., and Lopez-Rubio, J.J. (2008b). SnapShot: var Gene Expression in the Malaria Parasite. *Cell* 134, 10–11.

Schneider, R., and Grosschedl, R. (2007). Dynamics and interplay of nuclear architecture, genome organization, and gene expression. *Genes Dev.* 21, 3027–3043.

Shi, X., Kachirskaja, I., Walter, K.L., Kuo, J.H.A., Lake, A., Davrazou, F., Chan, S.M., Martin, D.G.E., Fingerman, I.M., Briggs, S.D., et al. (2007). Proteome-wide analysis in *Saccharomyces cerevisiae* identifies several PHD fingers as novel direct and selective binding modules of histone H3 methylated at either lysine 4 or lysine 36. *J. Biol. Chem.* 282, 2450–2455.

Shinkai, Y., and Tachibana, M. (2011). H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev.*

De Silva, E.K., Gehrke, A.R., Olszewski, K., León, I., Chahal, J.S., Bulyk, M.L., and Llinás, M. (2008). Specific DNA-binding by apicomplexan AP2 transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* 105, 8393–8398.

Sinha, A., Hughes, K.R., Modrzynska, K.K., Otto, T.D., Pfander, C., Dickens, N.J., Religa, A. a, Bushell, E., Graham, A.L., Cameron, R., et al. (2014). A cascade of DNA-binding proteins for sexual commitment and development in *Plasmodium*. *Nature* 507, 253–257.

Sivagurunathan, S., Heaslip, A., Liu, J., and Hu, K. (2013). Identification of functional

modules of AKMT, a novel lysine methyltransferase regulating the motility of *Toxoplasma gondii*. *Mol. Biochem. Parasitol.*

Smith, J.D., Chitnis, C.E., Craig, A.G., Roberts, D.J., Hudson-Taylor, D.E., Peterson, D.S., Pinches, R., Newbold, C.I., and Miller, L.H. (1995). Switches in expression of *Plasmodium falciparum* var genes correlate with changes in antigenic and cytoadherent phenotypes of infected erythrocytes. *Cell* 82, 101–110.

Spadaccini, R., Perrin, H., Bottomley, M.J., Ansieau, S., and Sattler, M. (2008). Retraction notice to “Structure and functional analysis of the MYND domain” [*J. Mol. Biol.* (2006) 358, 498-508] (DOI:10.1016/j.jmb.2006.01.087). *J. Mol. Biol.* 376, 1523.

Spellmon, N., Holcomb, J., Trescott, L., Sirinupong, N., and Yang, Z. (2015). Structure and Function of SET and MYND Domain-Containing Proteins. *Int. J. Mol. Sci.* 16, 1406–1428.

Sprangers, R., Groves, M.R., Sinning, I., and Sattler, M. (2003). High-resolution X-ray and NMR structures of the SMN Tudor domain: Conformational variation in the binding site for symmetrically dimethylated arginine residues. *J. Mol. Biol.* 327, 507–520.

Straimer, J., Gnädig, N.F., Witkowski, B., Amaratunga, C., Duru, V., Ramadani, A.P., Dacheux, M., Khim, N., Zhang, L., Lam, S., et al. (2014). K13-propeller mutations confer artemisinin resistance in *Plasmodium falciparum* clinical isolates. *Science* (80-. ). 2624, 428–431.

Su, X.Z., Heatwole, V.M., Wertheimer, S.P., Guinet, F., Herrfeldt, J. a, Peterson, D.S., Ravetch, J. a, and Wellems, T.E. (1995). The large diverse gene family var encodes proteins involved in cytoadherence and antigenic variation of *Plasmodium falciparum*-infected erythrocytes. *Cell* 82, 89–100.

Subramanian, K., Jia, D., Kapoor-Vazirani, P., Powell, D.R., Collins, R.E., Sharma, D., Peng, J., Cheng, X., and Vertino, P.M. (2008). Regulation of Estrogen Receptor ?? by the SET7 Lysine Methyltransferase. *Mol. Cell* 30, 336–347.

Tahiliani, M., Koh, K.P., Shen, Y., Pastor, W.A., Bandukwala, H., Brudno, Y., Agarwal, S., Iyer, L.M., Liu, D.R., Aravind, L., et al. (2009). Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* 324, 930–935.

Talbert, P.B., and Henikoff, S. (2010). Histone variants--ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* 11, 264–275.

Taylor, H.M., Kyes, S.A., and Newbold, C.I. (2000). Var gene diversity in *Plasmodium falciparum* is generated by frequent recombination events. *Mol. Biochem. Parasitol.* 110, 391–397.

Terman, J.R., and Kashina, A. (2014). Post-translational modification and Regulation of Actin. *25*, 30–38.

Tonkin, C.J., Carret, C.K., Duraisingh, M.T., Voss, T.S., Ralph, S.A., Hommel, M., Duffy, M.F., Da Silva, L.M., Scherf, A., Ivens, A., et al. (2009). Sir2 paralogs cooperate to

regulate virulence genes and antigenic variation in *Plasmodium falciparum*. *PLoS Biol.* 7, 0771–0788.

Treeck, M., Sanders, J.L., Elias, J.E., and Boothroyd, J.C. (2011). The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries. *Cell Host Microbe* 10, 410–419.

Trelle, M.B., Salcedo-Amaya, A.M., Cohen, A.M., Stunnenberg, H.G., and Jensen, O.N. (2009). Global histone analysis by mass spectrometry reveals a high content of acetylated lysine residues in the malaria parasite *Plasmodium falciparum*. *J. Proteome Res.* 8, 3439–3450.

Ukaegbu, U.E., Kishore, S.P., Kwiatkowski, D.L., Pandarinath, C., Dahan-Pasternak, N., Dzikowski, R., and Deitsch, K.W. (2014). Recruitment of PfSET2 by RNA Polymerase II to Variant Antigen Encoding Loci Contributes to Antigenic Variation in *P. falciparum*. *PLoS Pathog.*

Vembar, S.S., Droll, D., and Scherf, A. (2016). Translational regulation in blood stages of the malaria parasite *Plasmodium* spp. : systems-wide studies pave the way. *Wiley Interdiscip. Rev. RNA.*

Volz, J.C., and Cowman, A.F. (2012). Unveiling malaria's "cloak of invisibility." *Virulence* 3, 449–451.

Volz, J., Carvalho, T.G., Ralph, S.A., Gilson, P., Thompson, J., Tonkin, C.J., Langer, C., Crabb, B.S., and Cowman, A.F. (2010). Potential epigenetic regulatory proteins localise to distinct nuclear sub-compartments in *Plasmodium falciparum*. *Int. J. Parasitol.*

Volz, J.C., Bártfai, R., Petter, M., Langer, C., Josling, G.A., Tsuboi, T., Schwach, F., Baum, J., Rayner, J.C., Stunnenberg, H.G., et al. (2012a). PfSET10, a *Plasmodium falciparum* Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division. *Cell Host Microbe* 11, 7–18.

Volz, J.C., Bártfai, R., Petter, M., Langer, C., Josling, G.A., Tsuboi, T., Schwach, F., Baum, J., Rayner, J.C., Stunnenberg, H.G., et al. (2012b). PfSET10, a *Plasmodium falciparum* Methyltransferase, Maintains the Active var Gene in a Poised State during Parasite Division. *Cell Host Microbe*.

Wang, Z., Zang, C., Cui, K., Schones, D.E., Barski, A., Peng, W., and Zhao, K. (2009). Genome-wide Mapping of HATs and HDACs Reveals Distinct Functions in Active and Inactive Genes. *Cell* 138, 1019–1031.

Wang, D., Zhou, J., Liu, X., Lu, D., Shen, C., Du, Y., Wei, F.-Z., Song, B., Lu, X., Yu, Y., et al. (2013). Methylation of SUV39H1 by SET7/9 results in heterochromatin relaxation and genome instability. *Proc. Natl. Acad. Sci. U. S. A.* 110, 5516–5521.

Weller, R.L., and Rajski, S.R. (2005). LETTERS DNA Methyltransferase-Moderated Click Chemistry. 2003–2006.

Westenberger, S.J., Cui, L., Dharia, N., Winzeler, E., and Cui, L. (2009). Genome-wide

nucleosome mapping of *Plasmodium falciparum* reveals histone-rich coding and histone-poor intergenic regions and chromatin remodeling of core and subtelomeric genes. *BMC Genomics* 10, 610.

WHO (2015). World Malaria Report 2015. Who.

World Health Organization (2014). World Malaria Report 2014 (Geneva).

World Health Organization (2015). Status report on artemisinin and ACT resistance September 2015 Key messages Artemisinin resistance Molecular marker of artemisinin resistance. 1–8.

Yaffe, M.B. (2002). How do 14-3-3 proteins work? - Gatekeeper phosphorylation and the molecular anvil hypothesis. *FEBS Lett.* 513, 53–57.

Yildirim, F., Gertz, K., Kronenberg, G., Harms, C., Fink, K.B., Meisel, A., and Endres, M. (2008). Inhibition of histone deacetylation protects wildtype but not gelsolin-deficient mice from ischemic brain injury. *Exp. Neurol.* 210, 531–542.

Zeng, L., and Zhou, M.M. (2002). Bromodomain: An acetyl-lysine binding domain. *FEBS Lett.* 513, 124–128.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: Interplay between different covalent modifications of the core histone tails. *Genes Dev.* 15, 2343–2360.

Zhang, Q., Huang, Y., Zhang, Y., Fang, X., Claes, A., Duchateau, M., Namane, A., Lopez-Rubio, J.-J., Pan, W., and Scherf, A. (2011). A critical role of perinuclear filamentous actin in spatial repositioning and mutually exclusive expression of virulence genes in malaria parasites. *Cell Host Microbe* 10, 451–463.

Zhang, X., Wen, H., and Shi, X. (2012). Lysine methylation : beyond histones Modifying Enzymes for Lysine Methylation. *Acta Biochim Biophys Sin* 44, 14–27.

Zhang, X., Huang, Y., and Shi, X. (2015). Emerging roles of lysine methylation on non-histone proteins. *Cell. Mol. Life Sci.* 72, 4257–4272.