



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Genetic manipulation of histone deacetylases SAP18 and SIR2A in *Plasmodium falciparum*

by

Tayla Rabie

Submitted in partial fulfilment of the requirements of the degree

Magister Scientiae

(Specialisation in Biochemistry)

In the Faculty of Natural and Agricultural Sciences

University of Pretoria

South Africa

September 2021



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

SUBMISSION DECLARATION

I, Tayla Anne Rabie, declare that the dissertation, which I hereby submit for the degree *Magister Scientiae* in the Department of Biochemistry, Genetics and Microbiology at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE:

T. Rabie

DATE

23/09/2021



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

DECLARATION OF ORIGINALITY

UNIVERSITY OF PRETORIA
FACULTY OF NATURAL AND AGRICULTURAL SCIENCES
DEPARTMENT OF BIOCHEMISTRY, GENETICS AND MICROBIOLOGY
DIVISION OF BIOCHEMISTRY

Full names of student: Tayla Anne Rabie

Student number: 15090885

Declaration:

1. I understand what plagiarism is and am aware of the University's policy in this regard.
2. I declare that this dissertation is my own original work. Where other people's work has been used (either from a printed source, Internet or any other source), this has been properly acknowledged and referenced in accordance with departmental requirements.
3. I have not used work previously produced by another student or any other person to hand in as my own.
4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

SIGNATURE: _____

T. Rabie

DATE: _____ 23/09/2021 _____

Acknowledgments

I would like to acknowledge the following people without whom I would not have been able to complete my MSc degree.

Firstly, I would like to acknowledge my supervisor, Prof. Lyn-Marié Birkholtz for all her guidance and support over the duration of my degree. The knowledge and experience I have received from her has helped me achieve my best. I would also like to thank Dr Jandeli Niemand, for her assistance and expertise. I am also so thankful for the Malaria Parasite Molecular Laboratory (M2PL) team for their willingness to assist whenever needed, especially Dr Elisha Mugo for giving advice and input where needed.

I would like to express my appreciation to my parents, Dave and Deborah, for their support and continuous encouragement during my years of study. I would like to acknowledge all my MSc colleagues whom have grown to become great friends. Thank you for all the coffees, chats and laughs to pass some of the stressful and daunting situations.

Lastly, I would like to acknowledge the National Research Foundation for funding this degree through the NRF Master's Innovation Scholarship.

Summary

Malaria parasites are increasingly developing resistance to current drug treatments, emphasizing the continued need for new drugs targeting the disease-causing parasite, *Plasmodium falciparum* and the investigation of targetable biological processes. This includes the unique and unusual epigenome of the parasite, where epigenetic mechanisms regulate both asexual proliferation and sexual differentiation. Here, we investigate the putative histone deacetylase I complex protein, PF3D7_0711400 (SAP18), and a member of the histone deacetylase III family, PF3D7_1328800 (SIR2A). These proteins are thought to regulate histone acetylation, thereby allowing repression of genes in a very specific manner. These genes also show a unique expression profile during cell cycle arrest and re-entry. This phenomenon indicates an important role of both SAP18 and SIR2A in cell cycle progression.

This study aimed to manipulate *sap18* and *sir2A* for the investigation of essentiality and localisation in *P. falciparum* parasites. *Sap18* and *sir2A* were targeted for gene disruption to prove their functional essentiality. The DNA sequence encoding the N-terminus of *sap18* and *sir2A* was manipulated into a specialised gene disruption plasmid (pSLI-TGD) that allows antibiotic-driven selection of integration of truncated *sap18* and *sir2A* genes into the parasite genome. Additionally, these genes were tagged with green fluorescent protein (GFP) to allow for localisation of the protein during asexual proliferation by manipulating the C-terminus of *sap18* and *sir2A*.

The pSLI-Sandwich/pLyn-FRB-nmd3-BSD-*mCherry* plasmids, also allowed for mislocalisation of the protein of interest. This was used to achieve knock sideways of SAP18, however, no effect on asexual proliferation was observed. GFP fluorescence showed that SAP18 localises to the nucleus. Disruption of *sap18* using pSLI-TGD does not produce a growth defect, further supporting the conclusion that SAP18 is not essential in asexual stages. The disruption of SAP18 was evaluated for its effect in cell cycle arrest and re-entry. These parasites arrest in early trophozoite stage and re-enter the cell cycle after the addition of putrescine in the same manner as wild type parasites. These data suggest that SAP18 may not be a major regulator of HDAC1 activity, which differentiates *P. falciparum* from other organisms.

This study has allowed for evaluation of the functional importance of SAP18 to parasite survival and has the potential to provide novel tools for analysis of the epigenetic regulation of the cell cycle in *P. falciparum* parasites.

Table of contents

Acknowledgments.....	iv
Summary	v
List of Figures	viii
List of Tables	x
Chapter 1: Literature review	1
1.1. Malaria burden.....	1
1.2. Control strategies.....	1
1.2.1. <i>Anopheles</i> vector control	1
1.2.2. Parasite control.....	2
1.3. <i>P. falciparum</i> asexual life cycle development.....	2
1.4. <i>P. falciparum</i> sexual development	4
1.5. <i>P. falciparum</i> cell cycle	4
1.5.1. Cell cycle regulation in <i>P. falciparum</i>	5
1.6. Transcriptional regulation of <i>P. falciparum</i> life cycle.....	6
1.7. Epigenetic gene regulation in eukaryotes.....	7
1.8. Epigenetic gene regulation in <i>P. falciparum</i>	8
1.9. Histone deacetylase complexes in eukaryotes.....	11
1.9.1. Class I HDACs.....	11
1.9.2. Class III HDACs.....	12
1.10. SAP18: a SIN3/HDAC or ASAP complex subunit.....	14
1.11. SIR2A: a class III HDAC	15
1.12. Histone deacetylation in life cycle regulation in <i>P. falciparum</i>	16
1.13. Genetic modification strategies in <i>P. falciparum</i>	19
Chapter 2: Materials and Methods	23
2.1. Ethics statement	23
2.2. <i>In silico</i> analysis of <i>P. falciparum</i> <i>sap18</i> and <i>sir2A</i> genes	23
2.2.1. Analysis of SAP18 and SIR2A functional annotations	23
2.2.2. Protein-protein interactions of SAP18 and SIR2A	24
2.3. <i>In vitro</i> culturing of intraerythrocytic parasites	24
2.4. Sorbitol synchronisation.....	25
2.5. <i>In vitro</i> gametocyte induction and maintenance of <i>P. falciparum</i> cultures.....	25
2.6. Cloning into pSLI-plasmids	25
2.6.1. Genomic DNA isolation from <i>P. falciparum</i> NF54 parasites	27
2.6.2. Amplification of gene regions for cloning into SLI plasmids	27
2.6.3. Cloning gene fragments into pSLI plasmids	28
2.6.4. Screening of recombinant plasmids	30
2.7. Large scale plasmid isolation for transfection.....	31
2.8. Transfection of recombinant SLI plasmids	32

2.8.1. Drug selection and screening of transgenic parasites	32
2.9. Confocal microscopy.....	36
2.10. Confirmation of integration via western blot	37
2.11. Knock sideways of NF54- <i>Sap18</i> -2xFKBP-GFP-2xFKBP parasites	38
2.11.1. Transfection and drug selection of pLyn-FRB- <i>mCherry</i> -nmd3-BSD plasmid into NF54- <i>Sap18</i> -2xFKBP-GFP-2xFKBP parasites	38
2.11.2. Functional analysis of NF54- <i>Sap18</i> -2xFKBP-GFP-2xFKBP/FRB- <i>mCherry</i> by knock sideways.....	39
2.12. Functional analysis of disrupted NF54- Δ <i>Sap18</i> -GFP parasites	39
2.13. Confirming the IC ₅₀ of DFMO on NF54 wild type and NF54- Δ <i>Sap18</i> -GFP parasites using a SYBR green I - based fluorescence assay	40
2.14. Assessing the effect of DFMO treatment on NF54 wild type and NF54- Δ <i>Sap18</i> -GFP parasites	40
Chapter 3: Results	41
3.1. <i>In silico</i> analysis and characterisation of <i>sap18</i> and <i>sir2A</i> genes	41
3.1.1. SAP18: a member of the histone deacetylase subunit	41
3.1.2. Protein-protein interactions of SAP18	46
3.1.3. SIR2A: A member of the sirtuin family with NAD ⁺ -binding activity	47
3.1.4. Protein-protein interactions of SIR2A	50
3.2. <i>In vitro</i> cultivation of <i>P. falciparum</i> parasites	52
3.3. PCR amplification	52
3.4. Cloning <i>sap18</i> and <i>sir2A</i> gene fragments into pSLI plasmids.....	53
3.5. Generation of transgenic lines	57
3.5.1. Parasite transfection and selection for episomal uptake.....	57
3.5.2. Selection for integration	60
3.6. Validation of transgenic <i>sap18</i> lines.....	65
3.6.1. Confocal microscopy to assess GFP expression	65
3.6.2. Western blot analysis to detect GFP protein expression	66
3.7. Functional analysis of SAP18 protein with knock sideways.....	67
3.8. Functional and morphological analysis of NF54- Δ <i>Sap18</i> -GFP parasites.....	69
3.9. Involvement of SAP18 in cell cycle control.....	73
3.9.1. DFMO inhibition of <i>P. falciparum</i> parasites	73
Chapter 4: Discussion	78
Chapter 5: Conclusion	82
References	83
Supplementary information	90

List of Figures

Figure 1.1:	The <i>P. falciparum</i> parasite life cycle.....	3
Figure 1.2:	Cell cycle in <i>P. falciparum</i> parasites.....	5
Figure 1.3:	Gene regulation in <i>P. falciparum</i> by histone post-translational modifications	10
Figure 1.4:	Schematic representation of class I and II HDACs complexes in higher eukaryotes.....	12
Figure 1.5:	Transcriptional silencing by SIR2 in <i>S. cerevisiae</i>	13
Figure 1.6:	Structure of the ASAP complex and ubiquitin domain of SAP18 in higher eukaryotes.....	15
Figure 1.7:	Structural organisation of PfSIR2A and its homologue hSIRT5.....	16
Figure 1.8:	Transcriptome profile of <i>sap18</i> and <i>sir2A</i> during cell cycle arrest and re-entry in <i>P. falciparum</i>	18
Figure 1.9:	Transcriptome profile of <i>sap18</i> and <i>sir2A</i> during commitment and development of gametocytogenesis in <i>P. falciparum</i>	19
Figure 2.1:	Cloning strategy to produce recombinant SLI-plasmids for transfection.....	27
Figure 2.2:	Drug selection and screening of transgenic parasites to confirm episomal uptake and integration of SLI-TGD and SLI-Sandwich recombinant plasmids.....	33
Figure 2.3:	Schematic representation of the Selection Linked Integration systems.....	35
Figure 2.4:	Schematic representation of the pLyn-FRB- <i>mCherry</i> -nmd3-BSD plasmid....	39
Figure 3.1:	Phylogenetic tree of SAP18 proteins in different <i>Plasmodium</i> species, other apicomplexans and non-parasitic relatives	41
Figure 3.2:	Schematic representation of PF3D7_0711400 (SAP18) architecture.....	42
Figure 3.3:	Sequence alignment of <i>P. falciparum</i> (residues 608-723) and HMM SAP18 domain consensus.....	43
Figure 3.4:	Sequence alignment of <i>P. falciparum</i> (residues 608-723) and <i>M. musculus</i> (residues 1-143) SAP18.....	43
Figure 3.5:	Predicted structure of PF3D7_0711400 modelled with SWISS model.....	44
Figure 3.6:	Predicted structure of PF3D7_0711400 using AlphaFold database.....	45
Figure 3.7:	A schematic model of the protein-protein interaction complex of PF3D7_0711400 and its predicted functional partners.....	46
Figure 3.8:	Phylogenetic tree of SIR2A proteins in different <i>Plasmodium</i> species, other apicomplexans and non-parasitic relatives.....	48
Figure 3.9:	Schematic representation of PF3D7_1328800 (SIR2A) architecture.....	49
Figure 3.10:	Sequence alignment of <i>P. falciparum</i> (residues 36-213) and HMM SIR2 domain consensus.....	49
Figure 3.11:	X-ray crystal structure of SIR2A	50
Figure 3.12:	A schematic model of the protein-protein interaction complex of PF3D7_1328800 and its predicted functional partners.....	51
Figure 3.13:	Stage morphology of <i>P. falciparum</i> NF54 parasites.....	52
Figure 3.14:	PCR amplification of gene specific homology regions	53
Figure 3.15:	Restriction enzyme digestion of pSLI-TGD and pSLI-Sandwich plasmids....	54
Figure 3.16:	Validation of pSLI-TGD- <i>Sap18</i> and pSLI-SW- <i>Sap18</i> recombinant plasmids.	55
Figure 3.17:	Validation of pSLI-TGD- <i>Sir2A</i> and pSLI-SW- <i>Sir2A</i> recombinant plasmids....	56
Figure 3.18:	Selection and recovery of recombinant <i>P. falciparum</i> parasites after transfection with pSLI-TGD- <i>Sap18</i> , pSLI-SW- <i>Sap18</i> , pSLI-TGD- <i>Sir2A</i> and pSLI-SW- <i>Sir2A</i>	57

Figure 3.19: PCR analysis of pSLI-TGD- <i>Sap18</i> and pSLI-SW- <i>Sap18</i> for episomal uptake into NF54 <i>P. falciparum</i> parasites.....	58
Figure 3.20: PCR analysis of pSLI-TGD- <i>Sir2A</i> and pSLI-SW- <i>Sir2A</i> for episomal uptake into NF54 <i>P. falciparum</i> parasites	59
Figure 3.21: Selection recovery of recombinant <i>P. falciparum</i> parasites after episomal uptake of pSLI-TGD- <i>Sap18</i> , pSLI-SW- <i>Sap18</i> , pSLI-TGD- <i>Sir2A</i> and pSLI-SW- <i>Sir2A</i>	60
Figure 3.22: PCR analysis of pSLI-TGD- <i>Sap18</i> integration into the <i>sap18</i> locus of <i>P. falciparum</i> parasites.	61
Figure 3.23: PCR analysis of pSLI-TGD- <i>Sir2A</i> integration into the <i>sir2A</i> locus of <i>P. falciparum</i> parasites	62
Figure 3.24: PCR analysis of pSLI-SW- <i>Sap18</i> into the <i>sap18</i> locus of <i>P. falciparum</i> parasites.....	63
Figure 3.25: PCR analysis of pSLI-SW- <i>Sir2A</i> integration into the <i>sir2A</i> locus of <i>P. falciparum</i> parasites.....	64
Figure 3.26: Localization of SAP18 in NF54- <i>Sap18</i> -2xFKBP-GFP-2xFKBP and NF54- Δ <i>sap18</i> -GFP parasites.....	65
Figure 3.27: Western blot analysis of GFP expression of NF54- Δ <i>Sap18</i> -GFP parasites compared to wild type NF54 parasites probed with anti-GFP antibodies.....	66
Figure 3.28: Western blot analysis of FKBP-12 expression of NF54- <i>Sap18</i> -2xFKBP-GFP-2xFKBP parasites compared to wild type NF54 parasites probed with anti-FKBP antibodies.....	67
Figure 3.29: PCR analysis of pLyn-FRB- <i>mCherry</i> plasmid for episomal uptake into the NF54- <i>Sap18</i> -2xFKBP-GFP-2xFKBP parasite line.....	68
Figure 3.30: Knock-sideways of NF54- <i>Sap18</i> -2xFKBP-2xFKBP/FRB- <i>mCherry</i> parasite lines assessed via confocal microscopy.....	68
Figure 3.31: Growth rate of untreated and rapalog treated NF54- <i>Sap18</i> -2xFKBP-GFP-2xFKBP/FRB- <i>mCherry</i> asexual stage parasites.....	69
Figure 3.32: Morphological evaluation of NF54- Δ <i>Sap18</i> -GFP and NF54 wild type parasites in the asexual stages.....	70
Figure 3.33: Growth rate of NF54- Δ <i>Sap18</i> -GFP and NF54 wild type asexual stage parasites.....	71
Figure 3.34: Number of merozoites per schizont in NF54- Δ <i>Sap18</i> -GFP and NF54 wild type parasites.....	71
Figure 3.35: Morphological evaluation of NF54- Δ <i>Sap18</i> -GFP parasites in sexual stages.	72
Figure 3.36: Dose-response evaluation of DFMO inhibition of NF54- Δ <i>Sap18</i> -GFP and NF54 wild type parasites.....	74
Figure 3.37: DFMO arrest of NF54- Δ <i>Sap18</i> -GFP and NF54 wild type parasites.....	74
Figure 3.38: DFMO arrest and reversal of NF54- Δ <i>Sap18</i> -GFP and NF54 wild type parasites.....	75
Figure 3.39: Morphological evaluation of DFMO arrest and re-entry of NF54- Δ <i>Sap18</i> -GFP and NF54 wild type asexual stage parasites.....	77

List of Tables

Table 2.1:	Primer sequences of <i>sap18</i> and <i>sir2A</i> gene amplicons.....	28
Table 2.2:	Primer sequences used for plasmid backbone amplification and sequencing.....	30
Table 2.3:	Recombinant plasmids and their functionality and selection markers.....	32
Table 2.4:	PCR primers for the 5', 3' and loci integration check of NF54- Δ <i>Sap18</i> -GFP, NF54- <i>Sap18</i> -2xFKBP-GFP-2xFKBP, NF54- Δ <i>Sir2A</i> -GFP and NF54- <i>Sir2A</i> -2xFKBP-GFP-2xFKBP parasites.....	36
Table 2.5:	Primer sequences used for amplification of pLyn-FRB- <i>mCherry</i> -nmd3 plasmid.....	39
Table 3.1:	Concentration and purity of digested and purified PCR products.....	54
Table 3.2:	Analysis of gametocyte production and conversion in NF54- Δ <i>Sap18</i> -GFP and NF54 wild type sexual stages.....	73

List of abbreviations

ApiAP2	Apicomplexan apatela 2
APS	Ammonium persulfate
ASAP	Apoptosis-and splicing-associated protein
DFMO	DL- α -difluoromethylornithine
DHFR	Dihydrofolate reductase
FKBP12	FK506 binding protein 12
FRB	FKBP12-rapamycin-binding
GFP	Green fluorescent protein
HATs	Histone acetyltransferases
HDACs	Histone deacetylases
HDMs	Histone demethylases
HEPES	2-hydroxyethyl-1-piperazineethanesulfonic acid
HMM	Hidden Markov model
HMTs	Histone methyltransferases
HP1	heterochromatin protein 1
Hpi	Hours post-invasion
IDC	Intraerythrocytic cycle
IRS	Indoor residual spraying
ITNs	Insecticide treated nets
IVM	Integrated vector management
NAG	N-acetylglucosamine
PBS	Phosphate buffered saline
PTMs	Post-translational modifications
RbAps	Retinoblastoma-associated proteins
R-point	Restriction point
SAP18	SIN3-associated polypeptide p18
SDS	Sodium dodecyl sulphate
SIR2	Silent information regulator 2
SLI	Selection-Linked Integration
WHO	World Health Organisation

Chapter 1: Literature review

1.1. Malaria burden

Malaria remains a global health burden and is caused by the parasitic protozoan *Plasmodium* species. In humans, malaria is currently known to be caused by five *Plasmodium* species, *Plasmodium falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and *P. knowlesi* [1]. Among these species, *P. falciparum* and *P. vivax* are the most widespread, and contribute the highest number of infections and mortalities worldwide. *P. vivax* also contributes to the parasite infectious reservoir by forming dormant hypnozoites in the liver that cause recurring malaria. *P. falciparum* still globally accounts for the greatest number of infections and mortalities. In 2019, World Health Organisation (WHO) estimated about 229 million cases of malaria occurred in 87 malaria endemic countries worldwide, with the WHO African Region accounting for about 94 % of cases [1]. An estimated 409 000 deaths occurred in 2019 [1]. Studies have shown that *Plasmodium* infection is increased by lack of education, poorly constructed housing structure, and low income among people in sub-Saharan Africa [1].

1.2. Control strategies

Malaria is a difficult disease to control due to the complex and highly adaptable natures of the vector and parasite. In order to achieve sustainable malaria control, a combination of approaches and tools need to be used.

1.2.1. *Anopheles* vector control

There are over 400 *Anopheles* mosquito species, 40 of which are important in malaria transmission. Malaria vector control uses four types of insecticides, namely organophosphates, organochlorides (specifically dichlorodiphenyltrichloroethane), carbamates and pyrethroids. Insecticide-treated nets (ITNs) and indoor residual spraying (IRS) are recommended for people at risk of malaria in most epidemiological and environmental settings [1]. ITNs have been largely substituted with long-lasting ITNs. These nets are used as an intervention strategy to provide a barrier between the mosquito and the host [1]. IRS involves coating surfaces of a house with a residual insecticide which kills the mosquito when it lands on the surface after a blood meal. New strategies such as integrated vector management (IVM) have been designed to achieve maximum disease-control benefit while minimising the cost. Rather than relying on a single method of vector control, IVM emphasises the importance of first understanding the local vector environment and patterns of disease transmission, followed by utilisation of suitable vector control tools [1].

1.2.2. Parasite control

Parasite control is an important means for both control (chemotherapy) and prevention (chemoprophylaxis). Chemoprophylaxis is the use of medicines to temporarily protect individuals typically travelling to malaria-endemic areas by preventing onset of disease. Chemotherapy is used to treat patients infected with the parasite. The current most widely-used recommended treatment is artemisinin-based combination therapies (artemisinin and its derivatives in combination with drugs such as atovaquone, quinolines and antifolates) but the rapid development of resistance towards therapies is a concern. Quinolines mostly target the intraerythrocytic (IDC) stages and form a complex with haem in the parasite's food vacuole [1]. This results in a build-up of haem to toxic levels, thereby killing the parasite. Chloroquine is the preferred chemoprophylactic agent in malaria regions without resistant *falciparum* malaria [1]. Antifolates work by inhibiting dihydrofolate reductase (DHFR) within dividing schizont stage parasites [2]. Artemisinins have wide activity against various parasite stages. Their overarching mechanism of action is through the generation of free radicals created by ferrous haem or exogenous free iron [2]. Atoquavone works by targeting the cytochrome bc₁ complex in the electron transport system [3].

The complex nature of the parasite and the lack of understanding of pathogenesis has hampered vaccine development. Malaria parasites produce thousands of potential but variant antigens and exposure to malaria parasites does not result in lifelong protection. The most advanced vaccine currently being explored is the sporozoite subunit vaccine RTS,S/AS01, that employs the *P. falciparum* circumsporozoite protein to elicit an immune response. RTS,S/AS01 is the first vaccine to be tested in phase III clinical trials and the first to be evaluated in routine vaccination projects in high risk malaria areas [1]. The RTS,S vaccine has minimised childhood deaths and severe malaria by about 39 % in five to 17 month old children over four years who received a four-dose vaccine series [1]. A new vaccine, R21/MM, a modified version of the RTS,S malaria vaccine, has recently shown promising results. It works similarly to the RTS,S vaccine but employs a different adjuvant to boost immune response [4]. A phase II b trial, which recruited 450 children aged between five and 17 months, showed a vaccine efficacy of 74-77 % over the course of a year [5]. A phase III trial is underway, spanning four African countries and enrolling close to 5 000 children [5].

Malaria elimination requires a complex strategy, including drugs, vaccines, vector control and public health measures. Several factors, such as the parasites highly adaptable nature and the overuse or inadequate use of anti-malarial drugs, has resulted in the spread of resistance [6]. Therefore, there is a continued need to develop new drugs that target both asexual stage parasites to prevent disease, and the sexual stage parasites to prevent transmission.

1.3. *P. falciparum* asexual life cycle development

The *P. falciparum* life cycle is initiated when a female *Anopheles* mosquito feeds on an uninfected human host, thereby releasing sporozoites into the human host (Figure 1.1). The exo-erythrocytic cycle

begins when parasites enter the circulatory system and infect liver cells. The migration of sporozoites from the skin to the liver takes place in the first 15 minutes to a few hours after infection [7]. Sporozoites invade hepatocytes, within which they form a parasitophorous vacuole, and develop and replicate [8]. Sporozoites undergo multiple rounds of nuclear divisions within this vacuole producing 10 000-30 000 merozoites through the process of hepatic schizogony. Hepatic merozoites are released into the bloodstream to subsequently invade erythrocytes to initiate the intraerythrocytic cycle (Figure 1.1) [8].

Once inside the erythrocyte, the merozoite forms into a ring-stage parasite, after which these parasites undergo cell growth to develop into trophozoites. Trophozoite stages extensively change the host erythrocyte environment, consuming the host cytoplasm, decreasing the deformability of the cell membrane, making the membrane more permeable to small solutes, creating new permeability pathways and breaking down haemoglobin into amino acids [9, 10]. Once the parasite initiates asexual intraerythrocytic schizogony, multiple rounds of asynchronous DNA synthesis and endomitotic nuclear divisions result in a multi-nucleated schizont. After cytokinesis, schizonts develop that contain 16-32 merozoites [11]. These schizonts rupture to release merozoites, which then go on to re-infect new erythrocytes. This cycle repeats every 48 h, continuously rupturing, releasing and re-infecting erythrocytes [12].

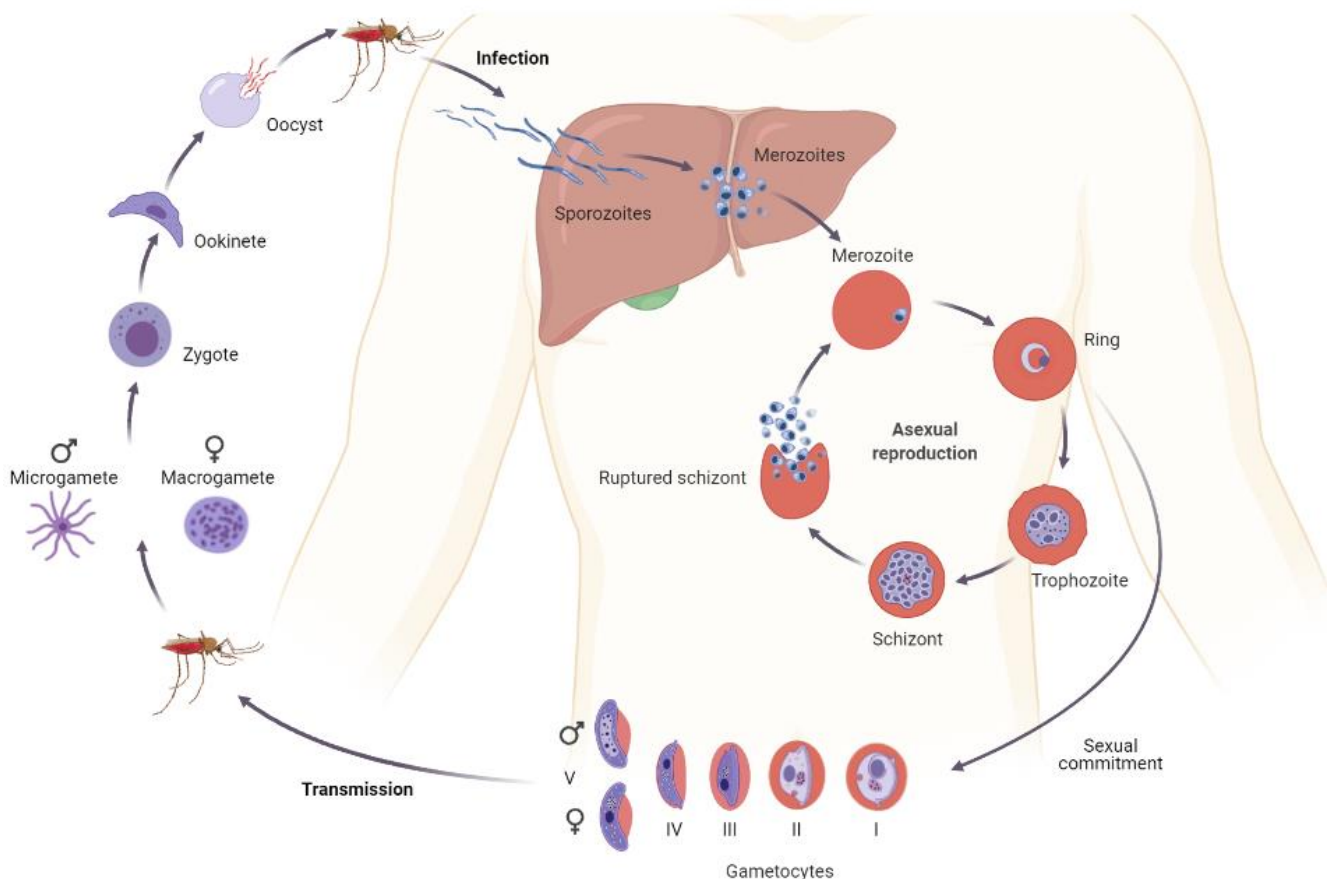


Figure 1.1: The *P. falciparum* parasite life cycle. The parasite completes its life cycle through three separate stages traversing both the *Anopheles* mosquito vector and human host. In the human host, the parasite undergoes an exo-erythrocytic stage inside hepatocytes before invading erythrocytes. The parasite can then either complete intraerythrocytic schizogony or differentiate into five stages of gametocytogenesis. The mature gametocytes are then transmitted to a feeding *Anopheles* mosquito. Image created using graphic components from BioRender (<https://app.biorender.com>).

1.4. *P. falciparum* sexual development

A small proportion (<10 %) of intraerythrocytic parasites can commit to sexual development, forming mature gametocytes that are capable of transmission to a feeding *Anopheles* mosquito (Figure 1.1). Sexually committed merozoites leave the intraerythrocytic cycle to develop into male and female gametocytes over a period of 10-14 days for *P. falciparum* [13, 14]. Commitment to sexual development occurs one asexual cycle before gametocyte manifestation involving all merozoites from a single committed schizont [15, 16]. Additionally, the merozoites from a single schizont become either all male or all female gametocytes, with a female-biased ratio of 3-5:1 [15].

The developing gametocytes are morphologically and functionally different from asexual parasites [17, 18]. Gametocytogenesis proceeds through five morphologically distinct developmental stages, exclusively observed for *P. falciparum* [17]. Stage I gametocytes are found primarily in the bone marrow during infection and resembles an asexual trophozoite that sequesters in the microvasculature. Stage II gametocytes are characterised by a mono-nucleic D-shape [19]. Stage III gametocytes adopt an elongated D-shape with rounded edges. Stage IV gametocytes are developed after additional elongation with sharpened edges. Gametocytes mature to produce rounded edge stage V female (macro-) and male (micro-) gametocytes [20]. Mature stage V gametocytes are released into the blood stream where they circulate for two to three days before they are taken up by the female mosquitoes [21].

Following uptake into the mosquito, the different micro-environment, including an increased pH, decreased temperature and the presence of xanthurenic acid stimulates the activation of male and female gametocytes to differentiate into micro- and macro- gametes, respectively [22]. Differentiation into micro-gametocytes involves three rounds of rapid mitosis, followed by micro-gamete ex-flagellation, resulting in eight flagellated haploid microgametes [23]. The female gametocyte, however, matures into one large macrogamete [24]. Micro- and macro-gametes from different hosts fuse to form a zygote in the mosquito midgut lumen, before the zygote develops into a motile ookinete [25]. The ookinete moves through the mosquito midgut before developing into an oocyst [25]. The oocyst undergoes sporogony which results in the production of multiple sporozoites. When the mature oocyst ruptures, it releases sporozoites which move to the salivary glands, to infect a new host [26] (Figure 1.1).

Apart from the difference in life cycle progression observed during the proliferative asexual and differentiating sexual intraerythrocytic parasite stages, the asexual proliferative cycle is associated with a unique cell cycle.

1.5. *P. falciparum* cell cycle

Eukaryotes rely on precise cell cycle progression and regulation for DNA synthesis and subsequent cell division into daughter cells [27]. The eukaryotic cell cycle includes interphase where the cell synthesises mRNA and proteins, synthesis (S) phase where DNA replication occurs, and mitosis (M) phase where the cell divides. The cell cycle is typically regulated by internal and external

cellular signals at certain checkpoints. These checkpoints ensure that cells either commence progression to the next phase, or enter a quiescent-like state until the cell cycle can recommence [28]. *Plasmodium* species differ notably from this standard model, to allow each cell to produce many new daughter cells during a single replicative cycle [28] (Figure 1.2). Cell cycle progression is thought to be a tightly regulated process as indicated by the stage-specific transcriptional profile during the IDC [28].

The *P. falciparum* cell cycle is uniquely different to that of the classical eukaryotic model. Invading merozoites are generally accepted to be in G_0/G_1 with a $1N$ DNA content [29]. Chromatin decondensation occurs from rings to early trophozoites as the parasites undergo G_1A-G_1B transition [29]. In early trophozoites, an increase in RNA and protein content is seen up to 24 h post invasion [29]. DNA synthesis occurs in late trophozoites after which, endocyclic schizogony occurs described by continuous DNA synthesis without cytokinesis. Rapid nuclear divisions result in multinucleated schizonts ($>2N$). Multiple rounds of S- and M-phase occur followed by a last round of synchronous nuclear division and a single cytokinesis event, and haploid daughter merozoites are released [30] (Figure 1.2). No evidence of the presence of a G_2 phase has been found in *P. falciparum*.

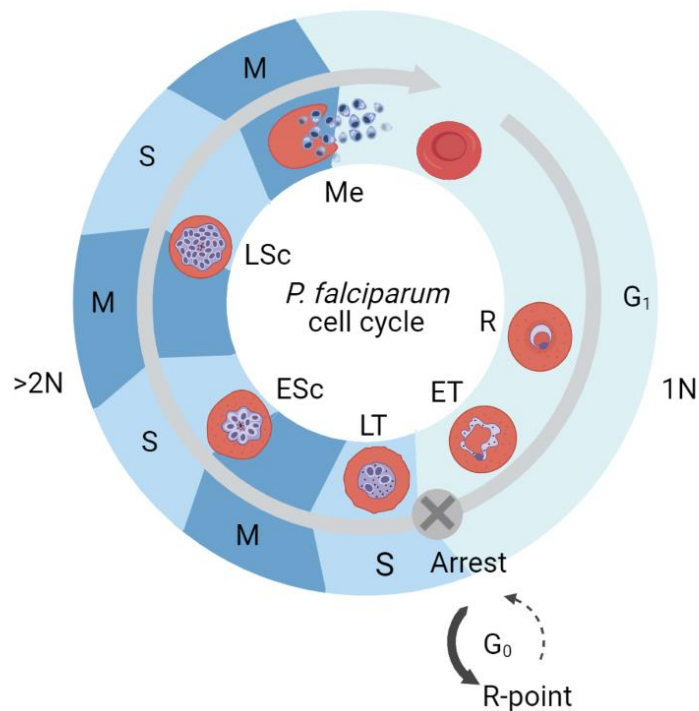


Figure 1.2: Cell cycle in *P. falciparum* parasites. Merozoites progress from ring-stage to early trophozoite stage during G_1 -phase ($1N$). The restriction point (R-point) exists where parasites can either arrest or progress to the next phase producing late trophozoites. Alternating S- and M- phases occur to produce a multinucleated schizont ($>2N$) which ruptures releasing multiple merozoites. Image adapted from [31]. Image created using graphic components from BioRender (<https://app.biorender.com>). R= rings, ET= early trophozoites, ESc = early schizont, LSc = late schizonts, Me = merozoites. $1N$ = single nuclei, $>2N$ = multiple nuclei.

1.5.1. Cell cycle regulation in *P. falciparum*

Cell cycle progression in eukaryotes is characterised by certain checkpoints that enables the cell to undertake to either advance to the next phase or exit the cell cycle and enter G_0 arrest. In *Plasmodium*, little is known about the existence of cell cycle checkpoints. However, induced cell cycle arrest has

suggested a growth factor-dependent restriction point (R-point) that resembles a G₁/S checkpoint [32, 33]. This checkpoint is the commitment point for eukaryotic cells to either progress through the G₁ phase to enter S-phase or enter G₀ arrest. The first line of evidence of an R-point is the entry of the parasite into a quiescent state of G₀ arrest, preventing the entry into a proliferative state it would not be able to complete [32, 33]. The fully reversible nature of the R-point, allowing maintenance of the parasite in an arrested state or re-initiation of the cell cycle, also supports a cell cycle control point at the G₁/S transition. Finally, the parasite progresses rapidly into S-phase after the block in proliferation is lifted without inconsequential irreversible damage to the parasite. Passing this R-point seems to be necessary before entering the mitotic cycles [30].

The above R-point is influenced by the availability of mitogens and particularly, a model to induce such a cell cycle arrest has been developed and involves the removal of polyamines [33]. Depleting polyamines in *P. falciparum* parasites, as in other cells [34, 35], induced cytostasis, concurrent with an arrest in the cell cycle. Polyamines are polycations which are needed for macromolecular synthesis required for S-phase [36]. Polyamine biosynthesis is effectively prevented by the addition of DL- α -difluoromethylornithine (DMFO), a suicide inhibitor of the synthesis enzyme ornithine decarboxylase [37], resulting in induction of the R-point, with reversal thereof possible through provision of exogenous mitogens (particularly polyamines) to allow the cell to re-enter the cell cycle. These results indicate that *P. falciparum* parasites can respond to external stimuli and undertake a quiescence-proliferation decision-making process in the initial phases of its cell cycle.

Cell cycle progression in the parasite is thought to be a tightly regulated process, due to the evidence of the “just-in time” transcriptional profile observed during the IDC [38], the dynamic regulation of its epigenome [39] and the incorporation of signalling events involved in cell cycle progression [33]. The parasite needs to be able to transition between different morphological stages in a human host. To accomplish this, the parasite utilises several types of regulation, including transcriptional and post-transcriptional gene expression, translational repression and post-translational protein modifications. Life cycle progression and gene expression cascades are thought to be driven by transcriptional regulation [40].

1.6. Transcriptional regulation of *P. falciparum* life cycle

The parasite's genome is 23 Mb in size and comprises of approximately 5500 genes distributed over 14 chromosomes, including 57 non-coding RNAs that have not been functionally characterised and about 2000 gene products with unknown function [41, 42]. The regulation of gene expression itself has unique features in the parasite and is controlled by a number of factors, such as transcriptional, epigenetic, post-transcriptional and post-translational control mechanisms [43].

P. falciparum parasites utilise a “just-in-time” transcriptional profile that indicates particular transcripts are only expressed in specific life cycle stages. Approximately 80 % of *P. falciparum* genes are expressed during the IDC [38]. More than 75 % of transcripts are expressed at a single stage of development [44, 45]. In gametocytes, transcriptional control is evident from expression of gametocyte-

specific and sex-specific genes [46]. Approximately 35 % of genes are known to be differentially expressed during gametocytogenesis. Of these, 246 genes show tightly-regulated, specific gametocyte expression profiles that may play a role in transcriptional regulation during gametocyte development [18].

It is therefore clear that the defined developmental stages of the parasite's life cycle are characterised by fluctuations in gene expression [44]. Transcriptional control is a key regulator of eukaryotic gene expression [47] however, a limited number of transcription factors have been found in the parasite which belong to the Apicomplexan apatela 2 (ApiAP2) family of proteins [48, 49]. During the IDC, 21 of the 27 ApiAP2 transcription factors are expressed, including some with roles in specific morphological stages, contributing to the strict control and overall transcriptional pattern observed during asexual replication. This paucity of transcription factors suggests that other mechanisms, such as post-transcriptional, translational, and post-translational regulation, as well as changes in the chromatin landscape, might contribute to regulation of parasite development through differential gene expression [43, 50]. Beyond transcriptional and translational control, there is substantial evidence that the parasite relies on epigenetic control to modulate gene expression throughout the life cycle.

1.7. Epigenetic gene regulation in eukaryotes

Epigenetic mechanisms consist of heritable changes to the eukaryotic genome that alters gene expression without altering the DNA sequence [51]. Epigenetic gene regulation is a highly controlled and reversible process. The epigenome, or total chemical modifications in the genome, controls the expression of certain genes needed for various cellular activities [51].

Epigenetic gene regulation can occur at the level of DNA, nucleosome, chromatin and higher-order nuclear organisation. Epigenetic modifications can occur at the DNA level by the addition of methyl groups to cytosine bases by DNA methyltransferases, producing 5-methylcytosine, usually found in CpG-rich regions [52]. CpG-rich islands are commonly associated with hypermethylation and are associated with heterochromatin formation, with resulting transcriptional repression [53]. Epigenetic modifications at the nucleosome level involve changes in chromatin structure, preventing access of DNA binding factors and resulting in changes in gene expression. Histone post-translational modifications (PTMs) and histone variants are responsible for the overall chromatin structure and composition shaping the cell's epigenetic landscape. The nucleosome core comprises a 147 bp DNA strand coiled around the histone octamer which consists of an H3/H4 tetramer and two H2A/H2B dimers [54]. Nucleosomes are coiled and folded to form a compact and complex chromatin structure. Euchromatic regions are open to transcription factors and are marked by activating PTMs. Heterochromatic regions are not accessible to DNA binding proteins and are marked by repressive PTMs. In addition, variant histones, differing by only a few amino acids, alter nucleosome stability and create novel characteristics to the nucleosome [55].

There are various histone PTMs identified within higher eukaryotes, including acetylation, methylation, phosphorylation, ubiquitination and SUMOylation, with acetylation and methylation being observed

most often. Epigenetic writers and erasers introduce or remove chemical alterations onto either histone tails or DNA bases. Epigenetic readers recognise the modification to mediate a functional result [56].

Histone can be mono-, di- or tri-methylated at lysine residues, while arginine residues can be mono- or di-methylated and the number of methyl groups can have significantly different effects on transcriptional activity [57]. Methyl groups can be deposited on either lysine or arginine residues by histone lysine methyl transferases or histone arginine methyl transferases and removed by the respective demethylases [57].

Contrary to methylation, histones can be acetylated at lysine residues. Histone acetyltransferases (HATs) and histone deacetylases (HDACs) control the deposition and removal of acetylated lysine residues. Acetylation reduces the histone-DNA interaction by countering the positive charge on histone tails, which results in loosening of the chromatin packaging to a transcriptionally active, euchromatic state [58]. Contrastingly, deacetylation results in heterochromatin formation. HAT- and/or HDAC-containing complexes are recruited to promoter regions to produce changes in chromatin status of various genes resulting in changes in the transcriptional activity of these genes [59]. Mammals have 18 HDAC enzymes that can be classed into four types: class -I (HDAC1, 2, 3 and 8), class -II (HDAC4, 5, 6, 7, 9, 10), class -III (sirtuins 1–7) and class- IV (HDAC11) [60]. All HDACs function in combination with other proteins to regulate acetylation levels. Histone acetylation has been shown to be important in progression through the cell cycle [61]. Within *Saccharomyces*, acetylation of various histones is important for the progression through the cell cycle, while chemical inhibitors of HDACs have been reported to have growth effects in mammals [61-63].

1.8. Epigenetic gene regulation in *P. falciparum*

Research has shown that *Plasmodium* parasites have atypical epigenetic regulation machinery. In *P. falciparum*, the core histone complex is retained from other eukaryotes, along with several variants such as H2A.Z, H2Bv, H3.3 and H3Cen, but H1 is absent in these parasites [64]. N-terminal PTMs are characterised by a majority of histone lysine acetylation and methylation marks (Figure 1.3 A). This decreases the overall compaction of chromatin structure in the parasite and results in a mostly transcriptionally active, euchromatic chromatin state during asexual development [64]. This is aided by the unusually AT-rich (81 %) genome and shorter nucleosome spacing [41].

It is thought that rapid chromatin changes contribute to the dynamic gene activation and repression profile observed during asexual replication of *P. falciparum* [38]. During schizogony, nucleosome numbers reach peak occupancy, while nucleosomes at telomeric regions sustain a cluster of genes in a repressed state characterised by heightened nucleosome occupancy, histone deacetylation, and heterochromatin protein 1 (PfHP1) [65]. Mass spectrometry has identified 232 different histone PTMs during the *P. falciparum* asexual life cycle [65]. Repressive histone PTMs are less frequent in *P. falciparum* and are found in condensed chromatin regions [65-67]. Only 10 % of the *P. falciparum* genome carries silencing histone marks, compared to the large proportion of activating histone marks

[65]. These observations coincide with the abundant euchromatin observed during the IDC of *P. falciparum* (Figure 1.3 B).

Epigenetic regulation is known to play an important role in commitment to gametocytogenesis, and a distinct epigenomic PTM landscape characterises gametocyte development [39, 68]. The highly transcriptionally active state present in asexual stages, differs from the general repressive state found during gametocytogenesis. This repressive environment is associated with an abundance of di- and tri-methylation marks particularly associated with H3, which are indicative of a heterochromatic state [69]. After commitment, progression of gametocyte development also requires transcriptional repression, as supported by increased repressive epigenetic marks and HP1 occupancy [39, 69]. These observations reveal a significant gametocyte-specific profile associated with distinct phases of transcription during gametocyte development.

As observed, the histone PTM landscape is dynamic during parasite development showing a stage-specific phase-like abundance profile [39]. Multiple histone PTMs act together, known as the histone code, to result in changes in gene regulation [55]. Histone acetylation and deacetylation has shown to play an important role in gene regulation within higher eukaryotes. However, it will be important to uncover the functional role of HDACs and their interacting partners in *P. falciparum*.

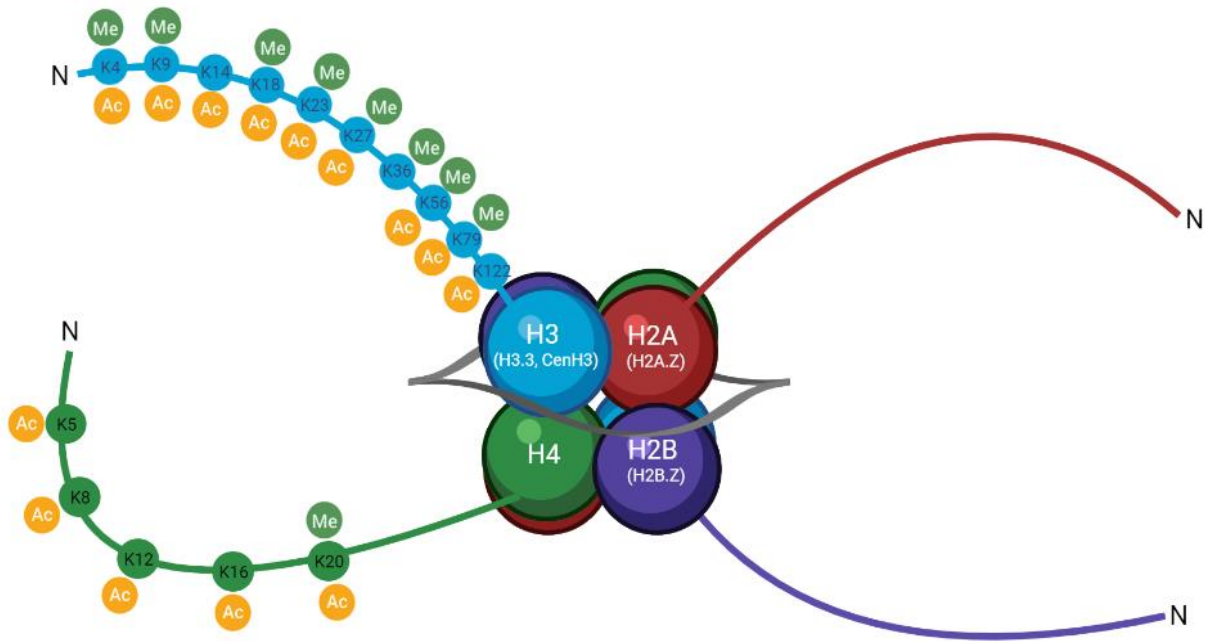
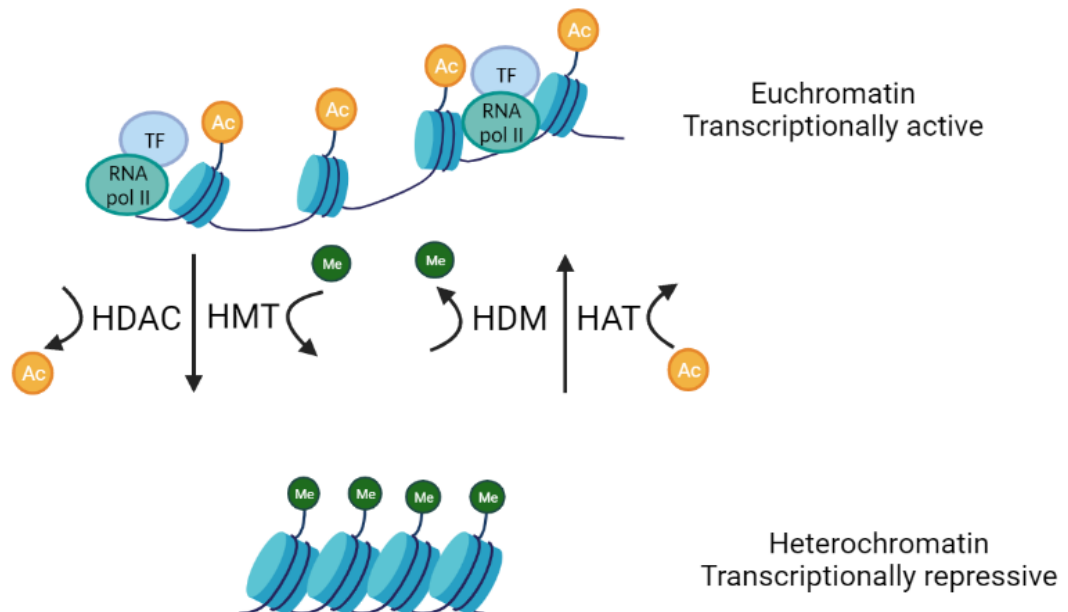
A**B**

Figure 1.3: Gene regulation in *P. falciparum* by histone post-translational modifications. (A) DNA (grey loops) wrapped around histones contain PTMs on their N-terminal tails, shown in their octamer nucleosomal organisation. These histones can be swapped with histone variants (H2A.Z, H2B.Z, H3.3, CenH3), which differ in amino acid composition and can carry distinct PTMs. Several histone PTMs were identified in all stages over life cycle development. For simplicity purposes, stage-specific histone PTMs are not shown. **(B)** The *P. falciparum* genome is found in a euchromatic, transcriptionally active state through histone demethylases (HDMs) or histone acetyltransferases (HATs), but certain epigenetic erasers and writers can produce a heterochromatic state responsible for repression of transcription such as histone deacetylases (HDACs) and histone methyltransferases (HMTs). Histone PTMs affect chromatin landscape through tightening or loosening of nucleosomes, thereby controlling the availability of promoters to RNA polymerases (RNA pol II) and transcription factors (TF). Image created using graphic components from BioRender (<https://app.biorender.com>). Image compiled based on information from [39, 70].

1.9. Histone deacetylase complexes in eukaryotes

HDACs are found as multiprotein complexes and involve interplay with other HDACs or regulator proteins for maximal functionality. HDACs do not contain a DNA-binding motif but consist of multiple interaction domains to recruit HDAC-interacting proteins to allow interaction with their chromatin targets [71]. Class I and III HDACs are the most widely studied.

1.9.1. Class I HDACs

A heterodimer is produced from HDAC1 and HDAC2 which comprises the catalytic core of the SIN3/HDAC complex [60]. The SIN3/HDAC complex is identified as a global transcriptional co-repressor, and remains highly conserved from yeast to humans. The SIN3A protein contains four paired amphipathic α -helix motifs to allow for multiple protein interactions [72]. Several previous studies have identified this complex in a multi-subunit protein complex containing eight proteins, SIN3A, HDAC1, HDAC2, SAP30, SAP18, SDS3 and retinoblastoma-associated proteins (RbAps) RbAP46 and RbAP48 [73, 74] (Figure 1.4 A). Little is known about the structure of the SIN3/HDAC complex and how its subunits are arranged to enable HDAC-mediated deacetylation. A refined model has been created to characterise the spatial arrangement of SIN3A subunits (Figure 1.4 B) [75]. This model provides the crosslinking hotspots of SIN3A centred on the HDAC interaction domain. SDS3, SAP30/30L, SAP130, HDAC2 and HDAC1 can attach to the platform of SIN3A. SAP30L directly contacts HDAC1 to help with the positioning of HDAC1 or alter HDAC1 folding [75].

In addition to the SIN3/HDAC complex, HDAC1 and HDAC2 are present in the NuRD (Figure 1.4 C) and Co-REST (Figure 1.4 D) complexes. The SIN3/HDAC and the NuRD complexes have wide activity in gene transcription [76] and bind various co-factors such as SAP proteins, Mi2, MTA2, and MBD proteins [71, 77]. The NuRD and Co-REST complexes consist of the catalytic core of HDAC1/2 as well as various co-factors. The Co-REST complex (Figure 1.4 D) plays a more precise role in transcriptional repression of neural genes [78]. A class II HDAC, HDAC3 is a major component within the NCoR/SMRT complex (Figure 1.4 E) and a negative regulator of angiogenesis, the formation of new blood vessels [71].

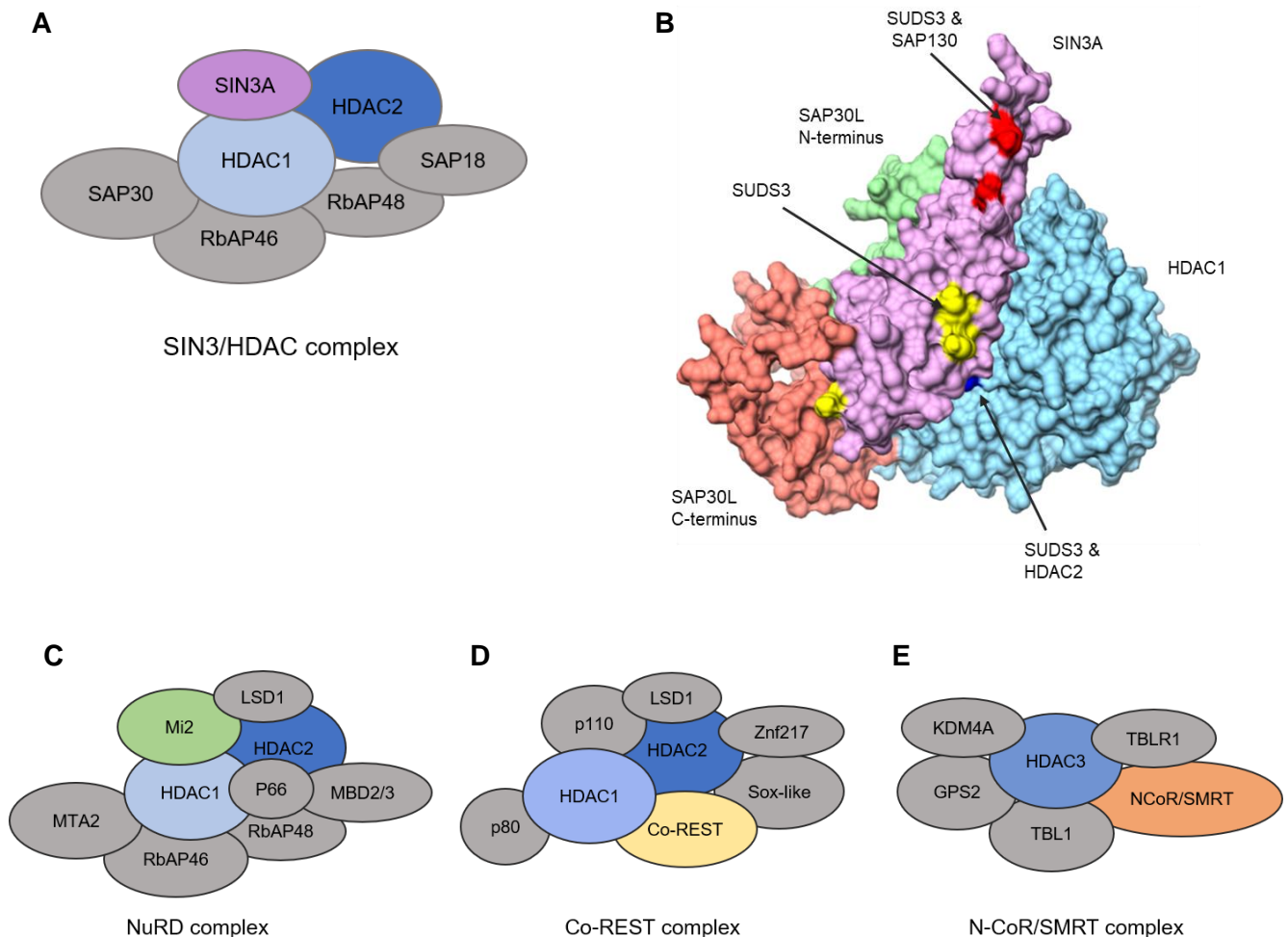


Figure 1.4: Schematic representation of class I and II HDACs complexes in higher eukaryotes. (A) Schematic depiction of the SIN3/HDAC complex which contains HDAC1, HDAC2, SIN3, RbAp46, RbAp48, SAP18 and SAP30. **(B)** Model of the SAP30L/SIN3A/HDAC1 complex generated using SWISS model indicating the SIN3A (purple) residues crosslinked with SUDS3 (yellow), SAP130 (red) and HDAC2 (dark blue). HDAC1 (light blue) interacts with SAP30L (orange and green). Image was obtained from PDB database (PDBDEV_0000043) adapted from [75]. **(C)** The NuRD complex contains HDAC1, HDAC2, RbAP46, RbAP48, chromatin remodeler Mi2, MTA2 (a member of metastasis associated family), methyl CpG binding domain protein MBD2/3, lysine-specific demethylase LSD1 and transcriptional repressor p66. **(D)** The Co-REST complex contains the co-repressor Co-REST, LSD1, Kruppel-like zinc-finger protein Zn217, p80 and Sox-like protein. **(E)** The N-CoR/SMRT complex is the only complex known to contain HDAC3 as well as N-CoR (Nuclear receptor CoRepressor), SMRT (Silencing Mediator for Retinoid and Thyroid receptor), TBL1 (transducin b-like 1), TBLR1 (TBL related 1), GPS2 (G-protein pathway suppressor 2) and the lysine-specific demethylase KDM4A. Image compiled based on information from [71, 79].

1.9.2. Class III HDACs

Silent information regulator 2 (SIR2), a class III HDAC, was originally identified in *Saccharomyces cerevisiae* for its function in silencing at the cryptic mating type loci, *HML* and *HMR* [80]. Sirtuins have a variety of cellular activities within higher eukaryotes including their involvement in ageing, life-span and metabolism [81]. Two complexes exist with SIR2, the telomeric SIR complex (Figure 1.5 A), and the nucleolar RENT complex (Figure 1.5 B). SIR2 functions in the holocomplex which consists of SIR2, SIR3 and SIR4 to maintain a heterochromatin state by acting as a histone deacetylase [82] (Figure 1.5 A). The silencer elements contain binding sites for the origin recognition complex (ORC), repressor activator protein 1 (Rap1) and autonomous replicating sequence-binding factor 1 (Abf1), which together

act as a scaffold for the SIR proteins [83]. Telomere silencing requires the recruitment of SIR4 and SIR2 to silencer elements which triggers histone deacetylation and spreading of heterochromatin (Figure 1.5 A). The RENT complex contains Net1 and Cdc14, which interacts with SIR2 and is required for the interaction to the rDNA locus and hence rDNA silencing (Figure 1.5 B) [84]. Net1 and Pol I also recruit SIR2 to the rDNA to silence transcription [85].

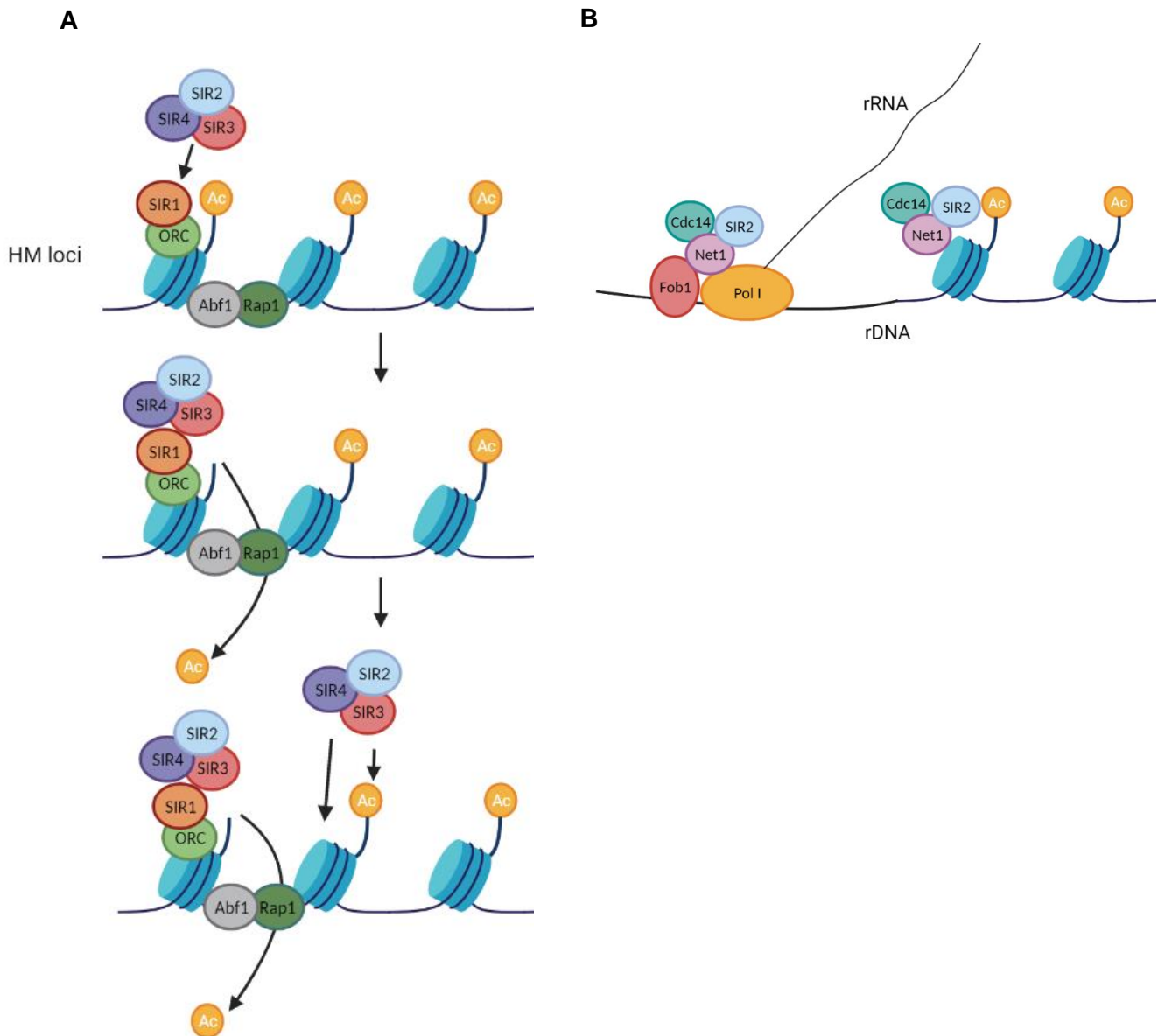


Figure 1.5: Transcriptional silencing by SIR2 in *S. cerevisiae*. (A) SIR2 with SIR3 and SIR4 is recruited to the HM loci by origin recognition complex (ORC), autonomous replicating sequence-binding factor 1 (Abf1) and repressor activator protein 1 (Rap1), which deacetylates N-terminal histone tails. The SIR complex spreads additionally deacetylating neighbouring histones. (B) The SIR2 and regulator of nucleolar silencing and telophase (RENT) complex composed of nucleolar protein 1 (Net1) and a tyrosine-protein phosphatase (Cdc14); interacts with RNA polymerase I (Pol I) and fork-block protein (Fob1) which silences downstream transcription. Image created using graphic components from BioRender (<https://app.biorender.com>). Image compiled based on information from [83, 86].

1.10. SAP18: a SIN3/HDAC or ASAP complex subunit

The HDAC1/2 complexes present in mammals, and RPD3 complex present in *Saccharomyces* and *Drosophila*, are recruited to chromatin where they deacetylate histones to result in heterochromatin formation [87]. The conservation of these SIN3A complex proteins suggests that the function of these complexes and associated co-factors may be conserved as well. In *Homo sapiens*, the SIN3-associated p18 protein, SAP18, is an 18 kDa protein and is known to be a transcriptional co-repressor that is functionally associated with HDAC1 [88]. SAP18 is thought to directly interact with SIN3A to enhance SIN3A-mediated transcriptional repression [88].

A similar SIN3/HDAC complex was identified in *Arabidopsis thaliana*. *Arabidopsis* contains a circadian clock which synchronises with the organisms' activities in their environment to maximise plant fitness. Parasites such as *P. falciparum* exhibit daily circadian rhythms within the host due to clock-control of the hosts behaviour [89]. The *Arabidopsis* genome contains six SIN3A homologues, SIN3-LIKE 1-6 (SNL1-6), four RPD3 homologues (HDA19, HDA9, HDA7, and HDA6), one SAP18 homologue, and two SAP30 homologues (SAP30 function related 1 (AFR1) and AFR2)) [90]. To provide information on the importance of the SIN3/HDAC complex in circadian oscillation, Lee *et al.* (2019) mutated SAP18 and analysed the resulting circadian oscillation pattern [88]. The SAP18 mutant revealed a different expression of circadian clock-associated 1 and cold circadian rhythm RNA binding 2 proteins [91], suggesting that the *Arabidopsis* SIN3/HDAC complex may be involved in circadian oscillation as well as the possible importance of SAP18 to SIN3/HDAC complex functionality. In contrast, within *Drosophila melanogaster*, SAP18 was found not be a component of the SIN3/HDAC complex as it does not co-immunoprecipitate with SIN3 [87]. RNA interference knockdown of SAP18 had no effect on cell progression, hence the function of the SIN3/RPD3 complex appears to be independent of SAP18 [92]. This leads to the question of the functional importance of SAP18 in the SIN3/HDAC complex.

In addition to its function within the SIN3/HDAC complex, SAP18 may be involved in other complexes, such as the ASAP (apoptosis-and splicing-associated protein) complex (Figure 1.6 A). This complex has been identified in several species such as *H. sapiens*, *D. melanogaster* and *Mus musculus* to contain a SAP18 subunit, along with Acinus (beta-hairpin motif) and RNPS1 (RRM domain), that plays a role in transcriptional control and pre-mRNA splicing [93]. The ASAP complex is involved in the exon-junction complex, an mRNP complex associated with post-transcriptional control [94]. RNPS1, an RNA-binding protein with a serine-rich domain, was identified within this complex, and may provide functionality in the regulation of mRNA export, splicing and nonsense-mediated mRNA decay [94]. Acinus, a chromatin condensation inducer, is cleaved during programmed cell death upon activation by caspase-3 [94]. The Acinus–RNPS1 interaction domain recruits the ubiquitin-like domain of SAP18 (Figure 1.6 B) to produce a complex with both RNA- and protein-binding functions [93, 94]. This leads to the possibility of an additional function of SAP18 within a similar ASAP complex in *P. falciparum*.

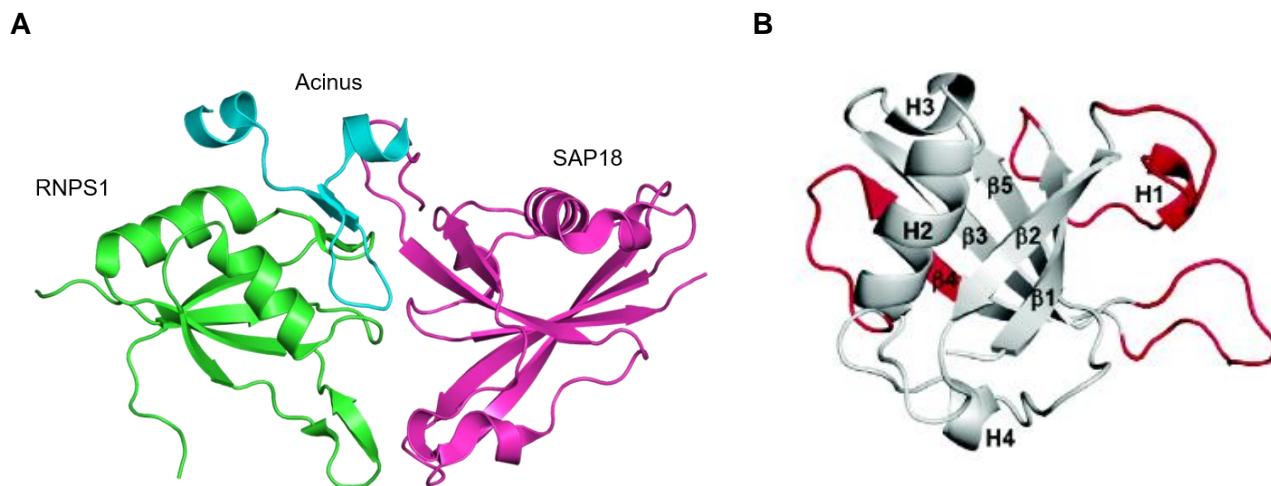


Figure 1.6: Structure of the ASAP complex and ubiquitin domain of SAP18 in higher eukaryotes. (A) RNA-binding protein (RNPS1, green), apoptotic chromatin condensation inducer (Acinus, blue) and SIN3-associated protein (SAP18, pink). Image was obtained from PDB database (accession code 4A8X) adapted from [93]. **(B)** SAP18 contains a ubiquitin-like domain (residues 20-143) including β -sheets ($\beta 1$ -5) and α -helices (H1-4). Image adapted from [95].

1.11. SIR2A: a class III HDAC

Sir2A encodes for a 30.3 kDa protein in *P. falciparum* and is the first epigenetic regulator shown to play a role in transcriptional regulation of antigenic variation through its histone deacetylase activity. PfSIR2A has been shown to control expression of surface antigen-coding *var* genes to evade detection by host immune surveillance [96]. Its major function is the removal of acetyl groups from N-terminal tails of H3 and H4 [97]. SIR2A is found both in the nucleolus and the nuclear periphery near telomeric clusters [98], as was previously identified in *S. cerevisiae* [99]. SIR2A can regulate the expression of *var* genes in *P. falciparum* by encouraging heterochromatic silencing. A recent crystal structure of PfSIR2A (PDB 3JWP) shows that the structure closely resembles that of human class III SIRT5 (PDB ID 2B4Y; Antoschenko *et al.*, unpublished) with a 31 % protein identity (Figure 1.7). PfSIR2A contains a catalytic NAD⁺-binding domain within the Rossmann fold and a variable domain that consists of Zn²⁺-coordinating cysteine residues. Zinc binding is important for the structural integrity of PfSIR2A and its deacetylase function [100]. These findings may indicate unique epigenetic mechanisms that may control parasite proliferation and virulence.

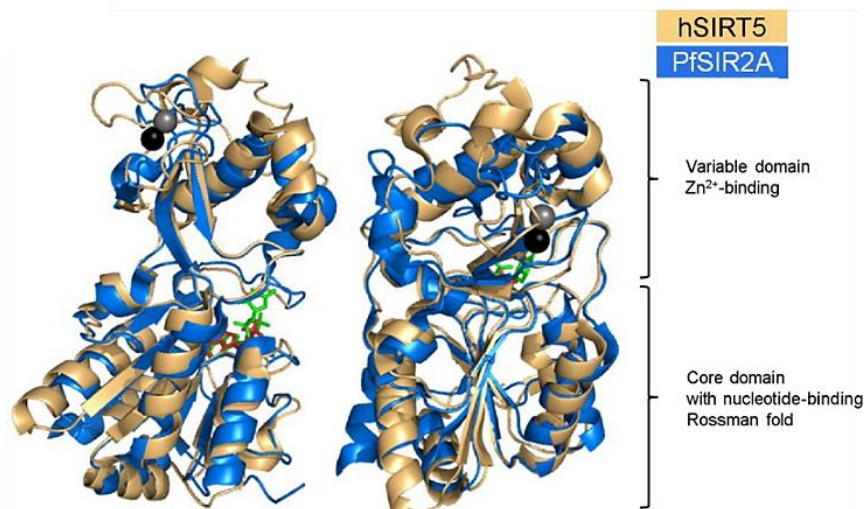


Figure 1.7: Structural organisation of PfSIR2A and its homologue hSIRT5. 3D structural alignment of hSIRT5 and PfSIR2A. PfSIR2A contains adenosine monophosphate (red) present in the NAD⁺-binding region in the core domain. hSIRT5 contains adenosine-5-diphosphoribose (green) present in the NAD⁺-binding region in the core domain. Zn²⁺ is represented as grey (PfSIR2A) and black (hSIRT5) spheres. Image taken from Figure 2 [86].

HDACs and their associated regulatory proteins are important for proper cell functioning and proliferation and their role in *P. falciparum* is an intriguing avenue of investigation

1.12. Histone deacetylation in life cycle regulation in *P. falciparum*

The *P. falciparum* genome contains five known HDAC homologues. These include class I, II and IV, which are zinc-dependent hydrolases, as well as class III HDACs, SIR2. Putative PfHDAC2/ PfHDA1, a class II HDAC (PF3D7_1472200), is transcribed throughout the asexual and gametocyte stages and is essential during asexual stages in *P. falciparum*. Expression of PfHDA1 is reduced upon increased temperature conditions, suggesting it is important in the up-regulation of genes involved in stress response and gametocytogenesis [101]. PfHDA2/PfHDAC3 (PF3D7_1008000), the second class II HDAC, is localised near the nuclear periphery that has been linked to the regulation of clonally variant genes. PfHDA2 silences virulence gene expression and regulates gametocyte conversion [102]. When PfHDA2 is knocked down, it leads to changes in gene expression of both virulence-associated *var* genes and PfApiAP2, a transcription factor associated with sexual conversion [102]. An increase in gametocyte conversion was observed, possibly due to dysregulation of *pfap2-g*. The return of PfHDA2 expression resulted in the of silencing of *var* genes [102].

Class III HDACs, also known as sirtuins, vary from other HDAC classes as they depend on NAD⁺ as their co-factor instead of zinc. Apicomplexans contain two sirtuins, SIR2A and SIR2B [103]. PfSIR2A and PfSIR2B, are localised to telomeric clusters at chromosomal ends and regulate transcriptional repression of surface antigen-coding *var* genes [104]. PfSIR2A is better characterised, having been localised to the nucleolus and shown to play an important role in epigenetic regulation as a histone deacetylase to produce heterochromatin at subtelomeres and rDNA [96, 98]. Neither SIR2A nor SIR2B

are essential in asexual stage parasites, as demonstrated using gene knockout studies [96, 104] however, they may exhibit redundancy, and a double knock-out has not been attempted.

PfHDAC1 (PF3D7_0925700), a class I HDAC, is localised to the nucleus [105] with expression throughout the asexual IDC, gametocytes and the exo-erythrocytic stages [44]. Drug inhibition of PfHDAC1 results in the down-regulation of essential genes, resulting in parasite death [106]. Although HDACs from higher eukaryotes are known to exist as multi-protein complexes, this has only been discovered via *in silico* analyses in *P. falciparum*. Only PfHDAC1 has been functionally validated *in vitro* however, little is known about its associated co-factor proteins. Identifying PfHDAC1 associated proteins could give us insight into the function of this protein as well as possible co-factors essential for PfHDAC1 function. A previous study identified members of PfHDAC1 complex through immunoprecipitation, several of these included homologues from higher eukaryotic species [107].

A previous study has identified multiple genes as differentially expressed during induced cell cycle arrest during the IDC of the parasite [33]. Of these, SAP18 and SIR2A were selected for further investigation to establish whether these epigenetic regulators are important during arrest and re-entry of the parasite into the cell cycle [33].

Several genes encode proteins involved in epigenetic control, and these genes are differentially expressed within the different cell cycle stages of the IDC. Particularly evident is the presence of 14 genes encoding suspected regulatory molecules associated with immediate events allowing re-entry into the cell cycle after arrest [33]. Ring-stage *P. falciparum* parasites arrested with DFMO, and sampled at multiple timepoints post-DFMO treatment (Figure 1.8 A), exhibited significant down-regulation of an epigenetic modulator, *sap18*, upon immediate re-entry into the cell cycle induced through putrescine administration (Figure 1.8 B). This trend continues throughout the parasite cell cycle. By contrast, *sap18* shows increased expression during periods of cell cycle arrest. Decreased *sap18* transcript abundance at cell cycle re-entry may be associated with an increased expression of genes associated with DNA replication and transcription [33]. This phenomenon may indicate an important role of SAP18's influence on cell cycle progression. It is postulated that when *sap18* is down-regulated, such as upon cell cycle re-entry, it is unable to enhance SIN3A activity. Thus, HDAC1 is not able to deacetylate chromatin, maintaining the euchromatic state, leading to the activation of genes associated with proliferation and, ultimately allowing progression into asexual development.

In addition, *sir2A* possesses the opposite profile to that of *sap18*, where it is significantly down-regulated during arrest while significantly up-regulated during stages of re-entry (Figure 1.8 B). It is postulated that *sir2A* is recruited to *var* genes where it deacetylates *var* gene promoter histones, allowing SET10 to bind to H3 and catalyse methylation, marking *var* genes as poised for reactivation in the next cycle [108]. *Sir2A* is down-regulated during cell cycle arrest, which is thought to allow antigenic switching of active *var* genes as well as ribosomal RNA synthesis.

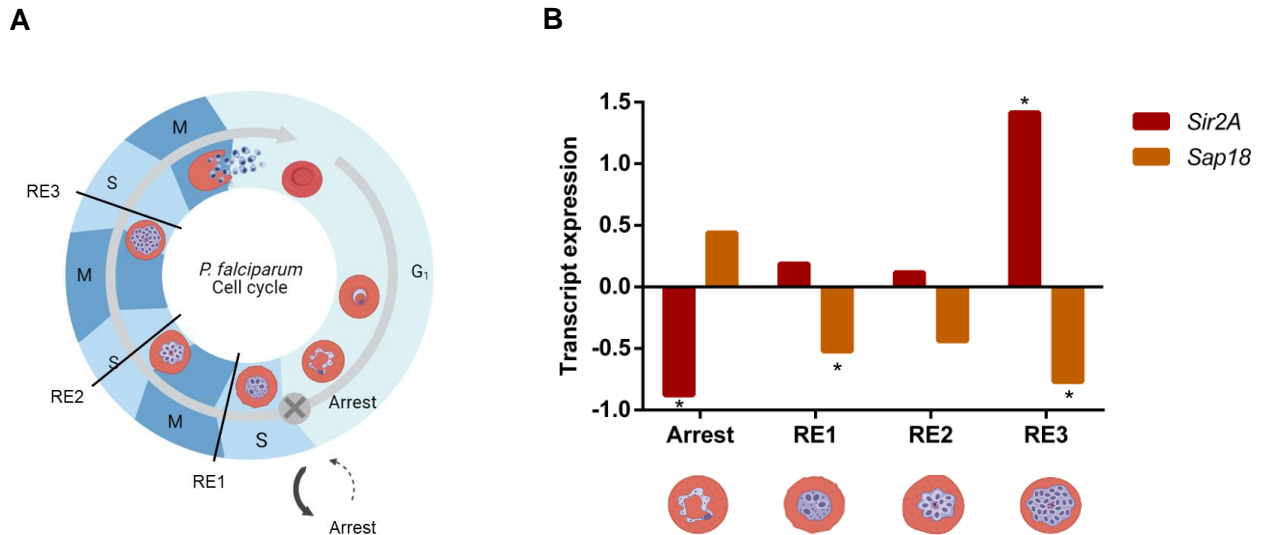


Figure 1.8: Transcriptome profile of *sap18* and *sir2A* during cell cycle arrest and re-entry in *P. falciparum*. (A) Ring-stage *P. falciparum* parasites were arrested with DFMO and sampled after 24 h (Arrest) post treatment. Cell cycle arrest was reversed with putrescine and parasites were sampled after 27 h (RE1), 30 h (RE2) or 36 h (RE3) post DFMO treatment. (B) The transcriptional profile of *sap18* (orange) and *sir2A* (red) of cell cycle arrested *P. falciparum* parasites (Arrest) as well as parasites that re-entered their cell cycles at specific time points (RE1-3). Genes differentially expressed during arrest and/or re-entry are indicated with an asterisk (*). Representative illustration of parasite stages shown below the graph. Image created using graphic components from BioRender (<https://app.biorender.com>). Figure adapted from Figure 3 A [33].

In addition to the involvement in the cell cycle during asexual proliferation, HDACs have also been implicated as important role players during particular stages during gametocytogenesis. *P. falciparum* parasites were sampled at multiple timepoints during commitment and development during gametocytogenesis (Figure 1.9), to reveal a set of 1075 genes identified as an “on switch” that describes gametocyte commitment. Within this, 680 genes (63 %) have increased expression during gametocytogenesis [109]. These included epigenetic regulators associated with cell cycle regulation such as SAP18 and SIR2A. These are thought to play a role by reducing DNA synthesis and creating a block in proliferation required for the parasite to differentiate. After commitment, progression of gametocytogenesis also requires transcriptional repression (an “off switch”). This is supported by increased HP1 occupancy and epigenetic marks associated with heterochromatin formation of early-stage gametocytes [110].

Sir2A expression is up-regulated during gametocytogenesis, resulting in the deacetylation of chromatin with consequent heterochromatin formation (Figure 1.9). *Sap18* is shown to be up-regulated during sexual commitment and maintained throughout gametocytogenesis. This is suggested to promote activity of HDAC1 to ultimately result in heterochromatin formation, supporting previous data that gametocyte development may require repression of proliferative markers to encourage gametocyte development. These epigenetic regulators have not been predicted to play a direct role in commitment to gametocytogenesis, but may contribute to the occupancy of particular histone marks during parasite differentiation.

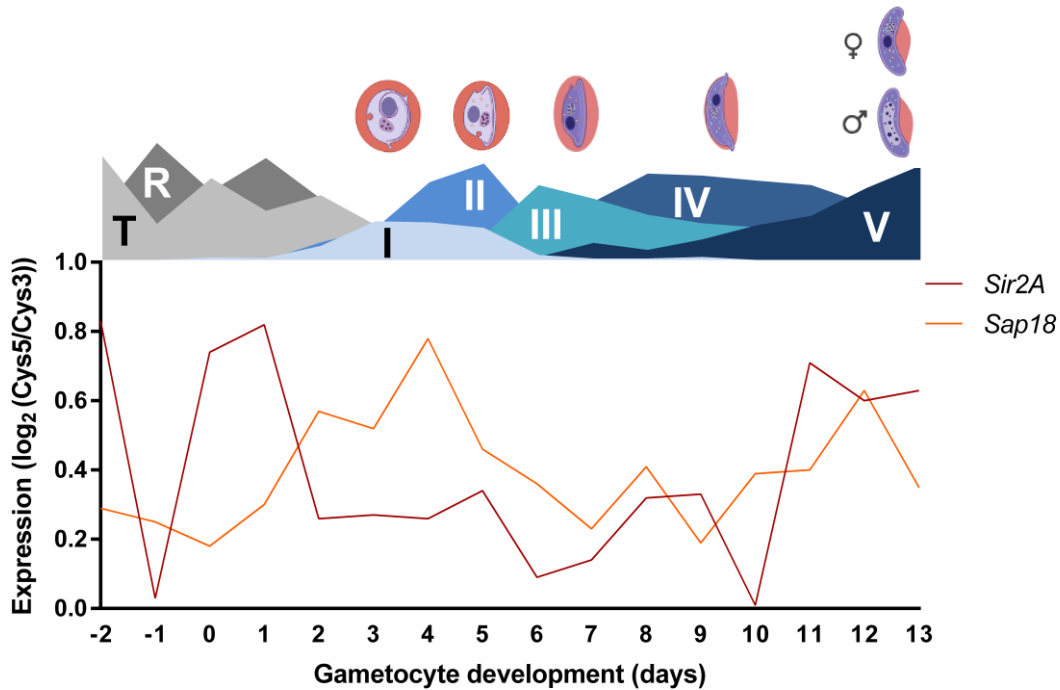


Figure 1.9: Transcriptome profile of *sap18* and *sir2A* during commitment and development of gametocytogenesis in *P. falciparum*. The transcriptome of *sap18* and *sir2A* are both increased in expression during commitment and development during gametocytogenesis. R = rings, T = trophozoites, I-V = stages of gametocytogenesis. Area plot above the graph indicates the timing of appearance and abundance of specific stages during development. Representative illustration of parasite stages shown above the graph. Image created using graphic components from BioRender (<https://app.biorender.com>). Figure adapted from Figure S 5 [109].

Investigation into the importance of SAP18 and SIR2A during asexual proliferation and gametocyte differentiation requires a genetic manipulation approach. This allows manipulation of the genes of interest to discover their functionality within the parasite.

1.13. Genetic modification strategies in *P. falciparum*

Several epigenetic regulator proteins have been shown to be essential for parasite proliferation in *P. falciparum*. This makes it difficult to target these essential proteins with a conventional knockout system. Other approaches that allow us to knock-down or disrupt gene or protein expression in an inducible manner, may provide additional insight to study essential protein function. Previous studies have identified PfHDAC1 [111], PfHDA2 [102] and PfHDA1 [111] to be essential to asexual proliferation using single-crossover recombination or piggyBac mutagenesis, while SIR2A and SIR2B were dispensable to cell growth [112], possibly indicating a redundant function of these two proteins. Other co-factors within the SIN3/HDAC complex, such as the *H. sapiens* homologue of RbAp46/48, known as chromatin assembly factor 1 protein WD40 (PfCAF1) was also shown to be essential with piggyBac mutagenesis [111]. Other methods, such as inhibitors, have been used to study protein function. One such example is JX21108, a potential target drug of PfHDAC1 that inhibits the growth of asexual parasites by reducing the expression of essential genes involved in metabolic processes as well as block parasite transmission [106].

Genome editing tools have become progressively more robust and available. However, genome editing within *P. falciparum* has proven challenging due to the lack of selectable markers, its ability to take up foreign DNA and the parasite's AT-rich genome. To study the biology of *P. falciparum* parasites, multiple tools have been developed in the field to attenuate gene expression at a gene, RNA or protein level. There are various techniques used in *P. falciparum* such as piggyBac transposon mutagenesis, diCre conditional knockout system, FKBP destabilisation domain, DOZI-TetR aptamer system, *glmS* ribozyme knockdown system and CRISPR-Cas9.

The DiCre conditional knock-out system uses FK506 binding protein 12 (FKBP12) and FKBP12-rapamycin-binding (FRB) proteins to dimerise two Cre recombinase proteins [113]. Each Cre polypeptide is fused to either the FKBP12 or FRB proteins. The recombinase identifies short sequences, known as *loxP* sites. DNA flanked by *loxP* sites are created in such a way that as well as the removal of DNA sequences or domains, point mutations can also be generated [113, 114].

Conditional knockdown systems targeting mRNA levels have been used to study essential proteins. One such example is the *glmS* knockdown system, which uses a *glmS* ribozyme sequence attached to the gene of interest, to produce a ribozyme in the mRNA product [115]. Glucosamine-6-phosphate is produced by the parasite upon addition of glucosamine, which leads to cleavage of the mRNA, resulting in transcript instability and degradation. This system also uses a *glmS* ribozyme sequence, which is inactivated by a mutation of the 9th nucleotide of the mRNA, which is thus termed M9. This serves as a control to the active *glmS* system. The mutated *glmS* control system allows confirmation that observed changes are due to knockdown of the mRNA transcripts and not as an effect of the genetic manipulation itself.

The DOZI-TetR system works by inserting several aptamer copies into the 3' untranslated regions of the gene of interest. Once transcribed, these are recognised by the TetR protein, which is bound to DOZI (development of zygote inhibited), a protein that represses translation by sequestering mRNA into P-bodies [116, 117]. Binding of DOZI-TetR to the mRNA leads to localisation to mRNA sequestration sites thereby repressing translation. Repression of translation can be reversed by the addition of tetracycline. One system used to regulate protein levels in *P. falciparum* is the FK506 binding protein destabilisation domain [118]. This domain contains mutations that prompt its misfolding in the absence of the stabilising ligand, Shld1. The addition of Shld1 stabilises the domain and the protein is not degraded [118].

An additional method, Selection-Linked Integration (SLI), is based on homologous recombination of the parasite's genomic DNA. The advantage of SLI is that parasites that contain the gene of interest episomally and parasites that have integrated the gene of interest into the genome can be selected for separately, thus increasing the chance of obtaining recombinant parasites and decreasing the time to integration [78]. The SLI system uses two different drug resistance genes, namely neomycin resistance

marker and human dihydrofolate reductase (hDHFR). The hDHFR gene is on the plasmid backbone and allows selection for parasites which have taken up the plasmid episomally. Episomes are extrachromosomal elements such circular plasmid DNA that are replicated autonomously in the cell. The hDHFR resistance gene allows drug selection with WR99210. This drug disables the *Plasmodium* DHFR enzyme thereby stopping nucleic acid synthesis. Therefore, only parasites that contain the SLI plasmid will survive since the expression of hDHFR takes over the function of the *Plasmodium* DHFR. The gene targeting region on the plasmid is attached via a skip peptide to an additional selection marker, known as the neomycin resistance marker [79, 80]. Neomycin resistance marker can only be expressed after a single crossover event into the parasite's genome which allows parasites carrying the integration to be selected.

This study involves the genetic manipulation of SAP18 and SIR2A to further understand parasite biology and epigenetic mechanisms of transcriptional control in *P. falciparum*.

Hypothesis

SAP18 and SIR2A are important for asexual replication and gametocytogenesis in *P. falciparum*

Aim

To develop and characterise genetically modified *P. falciparum* parasites to evaluate the essentiality, through targeted gene deletion, and functionality, through knock sideways gene function disruption, of SAP18 and SIR2A

Objectives

- 1) Using *in silico* analysis to characterise *sap18* and *sir2A* genes to obtain predictive functional descriptions
- 2) Generate recombinant plasmids for targeted gene deletion (pSLI-TGD) and knock sideways-mediated gene function disruption (pSLI-Sandwich) for both *sap18* and *sir2A* gene fragments
- 3) Generate four genetically modified *P. falciparum* parasite lines (NF54- Δ Sap18-GFP, NF54- Δ Sir2A-GFP NF54-Sap18-2xFKBP-GFP-2xFKBP, NF54-Sir2A-2xFKBP-GFP-2xFKBP)
- 4) Characterise the functionality and essentiality of these lines during asexual proliferation and sexual differentiation

Research outputs

Rabie T., Mugo E., Niemand J. and Birkholtz L-M. SAP18: A novel regulator of epigenetic mediated expression of genes associated with proliferation in *Plasmodium falciparum*. 5th MRC Office of Malaria Research Conference. Poster presentation. Pretoria, South Africa, 2019.

Rabie T., Mugo E., Niemand J. and Birkholtz L-M. Genetic manipulation of histone deacetylases SAP18 and SIR2A in *Plasmodium falciparum*. 6th MRC Office of Malaria Research Conference. Oral presentation. Pretoria, South Africa, 2021

Chapter 2: Materials and Methods

2.1. Ethics statement

All experiments of this study were performed at the Malaria Parasite Molecular Laboratory (M₂PL) with certified Biosafety level P2 facilities for *P. falciparum* parasite cultivation. The University of Pretoria Research Ethics Committee, Health Sciences Faculty has granted ethical clearance (506/2018) for the use of human erythrocytes. The *in vitro* cultivation of human malaria parasites is covered by approved umbrella ethics for the SARChI program under Prof. Birkholtz (NAS ethics approval no 180000094).

2.2. *In silico* analysis of *P. falciparum* *sap18* and *sir2A* genes

Genome modification approaches has provided phenotypic data describing if a gene is essential to *P. falciparum* or *P. berghei*. The essentiality and mRNA expression of *sap18* and *sir2A* genes during the asexual and gametocyte life cycle was analysed, using PlasmogEM (<http://plasmogem.sanger.ac.uk/>, accessed February 2021) and PhenoPlasm (<http://phenoplasm.org/>, accessed February 2021) databases. Both databases were used to establish the expected functionality of these genes.

2.2.1. Analysis of SAP18 and SIR2A functional annotations

The amino acid sequence of SAP18 and SIR2A proteins were analysed using NCBI Protein Blast Search (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, accessed January 2021) which provided results of most similar species within *Plasmodia* as well as outside of this lineage. A multiple sequence alignment was performed among *Plasmodium* species, other species within apicomplexan and more distant free-living relatives using MUSCLE (Multiple Sequence Comparison by Log-Expectation) (<https://www.ebi.ac.uk/Tools/msa/muscle/>, accessed July 2021). A phylogenetic tree was constructed using the maximum likelihood method (Jones-Taylor-Thornton (JTT) or Le Gascuel (LG) model) using 100 bootstrap replications.

InterPro (<https://www.ebi.ac.uk/interpro/>, accessed January 2021) and Pfam (<http://pfam.xfam.org/>, accessed June 2021) were used to predict conserved domains to provide information of the type of superfamily/family these proteins belong to as well as regions in the protein that are highly conserved. The protein architecture such as structural motifs and low complexity regions were identified with SMART (<http://smart.embl-heidelberg.de/>, accessed June 2021). These regions can have various organisations such as repetitive single amino acids and irregular motifs, and are found within the coding and non-coding areas of a genome [119].

A homology model was built for SAP18 and SIR2A using SWISS Model (<https://swissmodel.expasy.org/>, accessed February 2021) to visualise the quaternary structures of the protein using various templates identified by the server to best match the protein. In addition, a

Ramachandran plot was constructed with MolProbity (<http://molprobity.biochem.duke.edu/>, accessed August 2021) to evaluate the model quality based on acceptable angles. When phi (ϕ) (x-axis) and psi (ψ) (y-axis) angles are plotted against each other, a prediction of β sheets and α helices can be made. These can be represented on a Ramachandran plot and this provides information on acceptable rotations around ϕ and ψ angles to assess the quality of the model.

The 3D protein structure generated in SWISS model and the AlphaFold protein structure database (<https://alphafold.ebi.ac.uk/>, accessed July 2021) were compared. This database is based on machine learning predictions by using physical and biological knowledge about protein structure and multi-sequence alignments to create a deep learning algorithm. The Matchmaker tool in UCSF Chimera (version 1.15) was used to superimpose the protein structures by firstly producing pairwise sequence alignments, then matching the aligned residue pairs. The root-mean-square deviation (RMSD) value was used to determine the average deviation between the corresponding atoms of two proteins. A smaller RMSD value the more similar the two structures.

2.2.2. Protein-protein interactions of SAP18 and SIR2A

SAP18 and SIR2A are postulated to work with other epigenetic proteins to create an epigenetic response. The protein-protein interactions were identified using STRING (version 11.5) network algorithm (<https://string-db.org/>, accessed March 2021). This summarises the network of predicted associations for a particular protein [120]. The lines represent the predicted functional associations. These interactions are based on i) proteins with similar functions in different species, ii) proteins with similar metabolic functions and, iii) co-expression of genes [120]. The STRING network was imported into Cytoscape (version 3.8.2) for further analysis. The STRING network parameters were set to include protein-protein interactions with a minimum required interaction score of medium confidence (0.4), with a maximum of 20 proteins interacting.

The yeast-two-hybrid interaction network was also used to predict the interactors of SAP18 and SIR2A. *P. falciparum* protein-protein interactions were identified using a yeast-two-hybrid assay which overcomes the difficulties to express *P. falciparum* proteins in *S. cerevisiae* [121]. For two-hybrid screening, two hybrids are created between the protein of interest and either the DNA binding domain or the activation domain of the transcription factor. These domains can function near each other without direct binding to produce a functional transcriptional factor upstream of the reporter gene [122].

2.3. In vitro culturing of intraerythrocytic parasites

Human blood (various blood types) was collected in a blood bag containing citrate phosphate adenine anticoagulant (Adcock Ingram, South Africa). Serum and buffy coat was removed from whole blood by washing with 1x phosphate buffered saline (PBS) (137 mM NaCl (Sigma-Aldrich, USA), 2.7 mM KCl (Merck, Germany), 4.3 mM Na₂HPO₄ (Merck, Germany), 1.47 mM KH₂PO₄ (Merck, Germany), pH 7.4) followed by centrifugation at 3500 xg for 10 min. Erythrocytes were stored at 4 °C in incomplete culture

medium (RPMI–1640 (Sigma-Aldrich, USA) supplemented with 23.81 mM NaHCO₃ (Sigma-Aldrich, USA), 25 mM 2-hydroxyethyl-1-piperazineethanesulfonic acid (HEPES) (Sigma-Aldrich, USA), 0.024 mg/mL gentamycin (HyClone, USA), 0.2 mM hypoxanthine (Sigma-Aldrich, USA) and 0.2 % (w/v) glucose (Merck, Germany) for no more than two weeks at 50 % haematocrit. Routine culture maintenance of a NF54 *P. falciparum* strain (MRA-1000, a drug sensitive strain from BEI Resources/MR4; <https://www.beiresources.org>), included daily media changes with complete culture media (as above for incomplete but with the addition of 5 g/L Albumax II (Thermo Fisher Scientific, USA). Cultures were gassed with 5 % CO₂, 5 % O₂, and 90 % N₂ (Afrox, South Africa) for about 15-30 s, after which cultures were incubated at 37 °C in a shaking incubator at 60 rpm. Parasitaemia was determined by Giemsa staining under a light microscope (Nikon, Japan) at 1000x magnification. Parasitaemia is represented as a percentage of parasite-infected erythrocytes per number of uninfected erythrocytes, counted in approximately 5 fields of 100 erythrocytes each.

2.4. Sorbitol synchronisation

P. falciparum parasite cultures with a parasitaemia >2 % were synchronised to ~80 % ring-stage parasite population with 5 % (w/v) D-sorbitol (Sigma-Aldrich, USA) treatment at 37 °C for 15 min to allow for iso-osmotic lysis of trophozoite and schizont stage infected erythrocytes. This was followed by three washes in incomplete culture media. Synchronised cultures were maintained in complete culture medium at 5 % haematocrit.

2.5. *In vitro* gametocyte induction and maintenance of *P. falciparum* cultures

A *P. falciparum* NF54 strain (drug sensitive) was used since it is optimal for gametocyte production. Gametocytogenesis was induced by nutrient deprivation and by decreasing the haematocrit. This causes the parasite to stress [123]. Synchronised asexual parasites (>90 % ring-stage parasites) (day -3) were cultured as mentioned above, except they were cultured in complete RPMI media without glucose supplementation (gametocyte induction medium) to obtain a 0.5 % parasitaemia and 6 % haematocrit. After 72 h (day 0) the haematocrit was reduced to 4 %. The cultures were maintained at 37 °C without shaking and daily gametocyte induction media changes [123]. Gametocytogenesis was monitored daily with Giemsa-stained blood smears. Once mostly stage II gametocytes were observed, the asexual parasites were removed with 50 mM N-acetylglucosamine (NAG) (Sigma-Aldrich, USA) treatment (days 6-9) in the presence of complete RPMI media with 0.2 % (w/v) glucose supplementation, to obtain pure gametocyte cultures. NAG is toxic to asexual parasites and prevents proliferation by inhibiting erythrocyte invasion by merozoites [124]. From day 10 onwards, glucose enrichment was maintained until mainly stage V gametocytes were observed [123].

2.6. Cloning into pSLI-plasmids

The SLI system allows for the study and functionality of gene products within *Plasmodium*. The plasmids used for cloning, pSLI-TGD, pSLI-Sandwich, were kindly provided by Dr Tobias Spielmann's lab

(Bernhard Nocht institute for tropical medicine, Hamburg, Germany) [125]. Targeted gene disruption (TGD) can be used to target a shortened sequence in the 5' region of the gene of interest [125]. Once integrated into the genome, this will render the protein non-functional to allow for essentiality studies. pSLI-Sandwich is used for knock sideways and localisation studies where the protein may be essential to parasite proliferation. The 3' end of the gene of interest is fused with an FKBP12 domain which allows induction of the protein from its site of action through the dimerization of the FRB domain [125]. FRB is expressed separately in an additional plasmid, pLyn-FRB-*mCherry*-nmd3-BSD through episomal selection with blasticidin. This plasmid allows for mislocalisation of the protein of interest to the plasma membrane with the dimerization of FRB and FKBP.

A cloning strategy was adopted to generate recombinant plasmids for downstream functional analysis. The pSLI-TGD plasmid contains an N-terminal encoded gene fragment region; in this case *sap18* (426 bp), from 7-432 bp, or *sir2A* (193 bp), from 4-196 bp; which is ligated into pSLI-TGD through directional cloning using *NotI* and *MluI* restriction sites (Figure 2.1). This will produce pSLI-TGD-*Sap18* and pSLI-TGD-*Sir2A* recombinant plasmids. The pSLI-Sandwich plasmid contains a C-terminal encoded gene fragment region of either *sap18* (936 bp), from 1249-2184 bp (excluding the stop codon), or *sir2A* (816 bp), from 4-819 bp (excluding the stop codon); cloned with the restriction sites *NotI* and *AvrII* (Figure 2.1). This will produce pSLI-SW-*Sap18* and pSLI-SW-*Sir2A* recombinant plasmids.

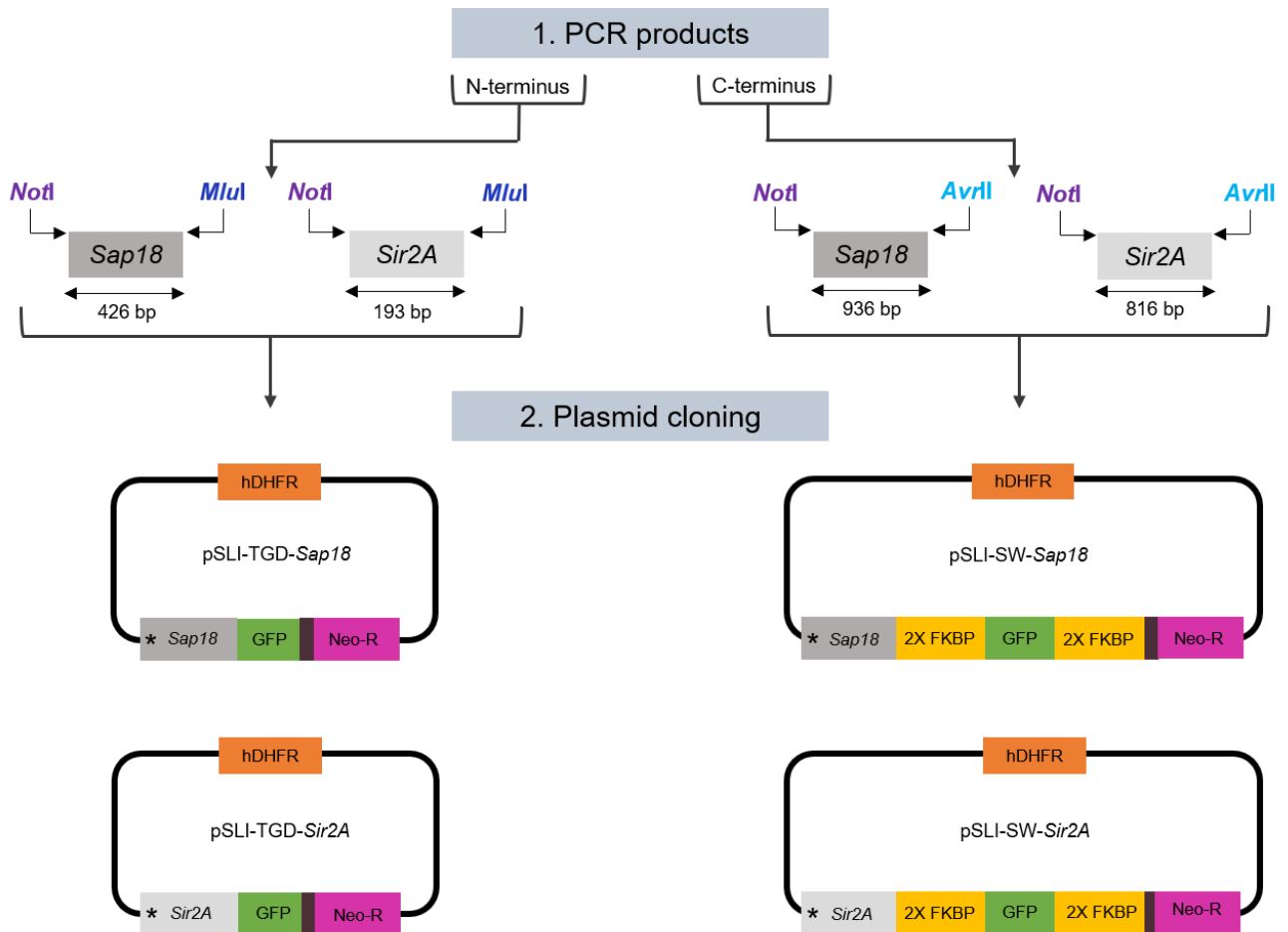


Figure 2.1: Cloning strategy to produce recombinant SLI-plasmids for transfection. The N-terminal encoded gene fragments of *sap18* and *sir2A* were PCR amplified and cloned into pSLI-TGD plasmids, using *NotI* and *MluI*, to produce pSLI-TGD-*Sap18* and pSLI-TGD-*Sir2A* recombinant plasmids. C-terminal encoded gene fragments of *sap18* and *sir2A* were PCR amplified and cloned into pSLI-Sandwich plasmids, using *NotI* and *AvrII*, to produce pSLI-SW-*Sap18* and pSLI-SW-*Sir2A* recombinant plasmids. Asterisks denote stop codon: TAA.

2.6.1. Genomic DNA isolation from *P. falciparum* NF54 parasites

Genomic DNA from intraerythrocytic *P. falciparum* NF54 parasites (>5 % parasitaemia and >80 % trophozoite stage parasites) was isolated using a Quick-gDNA Blood Miniprep kit (Zymo Research, USA) as per the manufacturer's instructions for use in downstream PCR amplification.

2.6.2. Amplification of gene regions for cloning into SLI plasmids

To obtain gene fragments for cloning into SLI plasmids, forward and reverse primers were designed using Benchling software (USA) (<https://benchling.com/>) to amplify N- or C-terminal homology regions of PF3D7_0711400 (*sap18*) and PF3D7_1328800 (*sir2A*) for insertion into pSLI-TGD and pSLI-Sandwich (Table 2.1). Primer design was based on GC content, Gibbs free energy (ΔG) and melting temperature (T_m) values. Primers were designed with flanking restriction enzyme sites, enabling the digestion with *NotI*-HF (5' GCGGCCGC 3') and *MluI*-HF (5' ACGCGT 3') for cloning into pSLI-TGD; and *NotI*-HF and *AvrII* (5' CCTAGG 3') for cloning into pSLI-Sandwich. Four extra nucleotide bases were added onto the 5' ends of the restriction enzyme sites to increase restriction digestion efficiency to allow cloning into the desired locations of the plasmids. The gene-specific fragments of all forward

primers were preceded by a stop codon, following the restriction enzyme region to prevent expression of gene fragments before genomic integration.

N- and C-terminal regions of the *sap18* and *sir2A* genes were PCR amplified for insertion into their respective plasmids using a 2720 Thermal Cycler (Applied Biosystems, Foster City, USA). PCR amplification was performed with 1x KAPA Taq Ready Mix (0.4 U KAPA Taq polymerase, 1.5 mM MgCl₂, 0.2 mM of each ddNTPs, KAPA Biosystems, Wilmington, USA), 5 pmol of each primer (IDT, San Jose, USA) and 50 ng of *P. falciparum* NF54 genomic DNA. Optimised PCR conditions included denaturation at 95 °C for 3 min, followed by 30 cycles of denaturation at 95 °C for 30 s with annealing at 55°C/58°C for 30 s and extension at 68 °C for 90 min, and a final extension at 68 °C for 5 min. DNA bands were visualised for the amplified gene regions on 1.5 % (w/v) agarose (Lonza, Basel, Switzerland) / TrisAcetate-EDTA (TAE; 20 mM Tris, 20 mM acetic acid and 1 mM EDTA, pH 8.0) gel that was stained with ethidium bromide (0.5 µg/mL in 1x TAE) for ~10 min and visualised with a Gel Doc XR+ Imaging system (Bio-Rad, USA). A 1 kb or 100 bp DNA ladder (Promega, USA) was used to determine DNA fragment sizes on the gel.

Table 2.1: Primer sequences of *sap18* and *sir2A* gene amplicons

Gene ID	Gene region amplified		Sequence (5'- 3' orientation) *	Ta (°C)
PF3D7_0711400	N-terminal	F	GCAT <u>GCGGCCGC</u> TAAAGAAGTGAAGTTTGAGTTTAAG	55
		R	CGAT <u>ACGCGT</u> GTTAGAGCTTTCTTTCTTGTCGC	
PF3D7_0711400	C-terminal	F	GCAT <u>GCGGCCGC</u> TAAAGTAGAAATAAAAGTAGAATAAG	55
		R	CGAT <u>CCTAGG</u> TTTTCTGTCCAACATAATAGATAGTAG	
PF3D7_1328800	N-terminal	F	GCAT <u>GCGGCCGC</u> TAAGGTAATTTAATGATTTCTTTTTG	56
		R	CGAT <u>ACGCGT</u> TCCATATATTCTTGGGTCATACTT	
PF3D7_1328800	C-terminal	F	GCAT <u>GCGGCCGC</u> TAAGGTAATTTAATGATTTCTTTTTG	58
		R	CGAT <u>CCTAGG</u> CATTATTTCTTATTTTTTCACTTGACCC	

*Restriction sites with four extra nucleotides were added to primer sequences: *NotI* (GCGGCCGC), *MluI* (ACGCGT), *AvrII* (CCTAGG), as underlined. F= forward, R= reverse.

PCR amplified fragments were purified from the PCR reactions using a Quick-DNA Miniprep kit (Zymo Research, USA), following manufacturer's instructions. Eluted DNA concentration and purity was determined using a Nanodrop ND-1000 spectrophotometer (Thermo Fisher Scientific, USA) by measuring nucleic acid absorbance at 260 nm as well as purity of the sample by evaluating contaminating proteins (A_{280}), salt, carbohydrate and phenol contamination (A_{230}). Eluted DNA purity was determined by the A_{260}/A_{280} and A_{260}/A_{230} ratios, with a ratio of 1.8-2.0 regarded as pure DNA. Purified DNA was stored at -20 °C.

2.6.3. Cloning gene fragments into pSLI plasmids

E. coli DH5 α glycerol stocks (Gibco BRL Life Technologies, USA) containing pSLI-TGD and pSLI-Sandwich plasmids were inoculated into LB-broth (1 % (w/v) tryptone, 1 % (w/v) NaCl and 0.5 % (w/v)

yeast extract, pH 7.5) containing 50 µg/mL ampicillin. Cultures were incubated with agitation at 150-200 rpm at 37 °C for 16 h. pSLI-TGD and pSLI-Sandwich plasmid DNA was isolated from saturated cultures using a NucleoSpin Plasmid (NoLid) Kit (Macherey-Nagel, Germany), according to manufacturer's instructions. Plasmid yield was quantified spectroscopically using the Nanodrop spectrophotometer and plasmids were stored at -20 °C until further use.

The PCR products and SLI-plasmids were digested with their respective enzymes to produce sticky ends for ligation. SLI plasmids and PCR products were prepared for cloning by digesting ~4 µg of plasmid and ~1 µg of PCR product with either 1 U of *NotI*-HF restriction enzyme (New England Biolabs, UK) and *MluI*-HF for pSLI-TGD or, *NotI*-HF and *AvrII* for pSLI-Sandwich in 1x CutSmart buffer (New England Biolabs, UK) at 37 °C for 3 h. The digestion reaction was incubated at 37 °C for 60 min after which it was heat inactivated at 65 °C for 5 min. The PCR products were purified using a NucleoSpin Gel and PCR Clean up kit (Macherey-Nagel, Germany), according to manufacturer's instructions. The plasmid backbones were separated on an 1 % (w/v) agarose gel and subsequently excised from the gel. The digested plasmid backbones were purified using a NucleoSpin Gel and PCR Clean up kit (Macherey-Nagel, Germany), according to the manufacturer's instructions.

The digested PCR products were ligated into their respective plasmid backbones with T4 DNA ligase (New England Biolabs, UK) in an optimal insert: vector molar ratio of 3:1 (as shown in the equation below) and incubated overnight at 4 °C.

$$\frac{\text{ng of vector} \times \text{kb size of insert}}{\text{kb size of vector}} \times \text{insert: vector molar ratio} = \text{ng of insert}$$

Each 20 µL ligation reaction consisted of 1x Rapid Ligation Buffer (60 mM Tris-HCl pH 7.8, 20 mM MgCl₂, 20 mM DTT, 2 mM ATP, 10 % polyethylene glycol), digested plasmid backbone, digested PCR insert and 1 U of T4 DNA ligase. The ligation reactions were incubated at 4 °C overnight before transformation into competent cells.

E. coli DH5α cells were made chemically competent to incorporate plasmid DNA via a CaCl₂ method. *E. coli* DH5α cells (-80 °C) were inoculated into 5 mL of LB-broth (1 % (w/v) tryptone, 1 % (w/v) NaCl and 0.5 % (w/v) yeast extract, pH 7.5) and incubated at 37 °C overnight with shaking at 150-200 rpm. The overnight culture was diluted 1:50 in LB-broth and grown until a mid-logarithmic growth phase, corresponding to an OD₆₀₀ of 0.4. The cells were incubated on ice for 10-20 min, followed by centrifugation at 1865 xg for 30 min at 4 °C. The supernatant was removed, and the cells were washed with 25 mL of 0.1 M CaCl₂ and centrifuged again as above. The supernatant was removed and 2.5 mL CaCl₂ and 375 µL 100 % glycerol was added. The cells were left on ice for 1 h and 100 µL aliquots were frozen at -80 °C.

Competent *E. coli* DH5α cells were thawed from -80 °C on ice. The ligation mixture consisting of plasmid and its corresponding gene fragment insert, was immediately added to competent cells and incubated on ice for 30 min. The cells were heat-shocked for 90 s at 42 °C, allowing plasmid DNA uptake by competent *E. coli* DH5α cells. The cells were immediately placed on ice for 2 min after which 900 µL of

pre-warmed LB-glucose (LB-broth and 20 mM glucose) was added and the cells were incubated at 37 °C for 1 h shaking at 150-200 rpm. The transformation reaction (pellet and/or 100 µL) was plated onto LB-agar plates (1 % (w/v) agar in LB-broth supplemented with 100 µg/mL ampicillin) and incubated overnight at 37 °C in a stationary incubator.

2.6.4. Screening of recombinant plasmids

To confirm the presence of the correct insert within their respective plasmids, colony screening PCR was performed on transformed bacterial colonies using primers listed in Table 2.2. Single colonies were picked and individually inoculated into 100 µL of LB-broth, supplemented with 50 µg/mL ampicillin, for 3 h at 37 °C with shaking at 150-200 rpm. The PCR reaction mixture was set up as follows: 1x KAPA-Taq polymerase, 5 pmol of forward primer and 5 pmol reverse primer (Table 2.1 or 2.2) and 1 µL of bacterial culture. Thermocycler conditions were set up as follows: initial denaturation of 95 °C for 3 min and cycling 30x at 95 °C for 30 s, 52 °C for 30 s, 68 °C for 1 min followed by a final extension at 68 °C for 5 min. PCR reactions were analysed on a 1.5 % (w/v) agarose gel and positive clones were identified based on the presence of a DNA band on the gel. These positive clones were inoculated into 10 mL of LB-broth (containing 50 µg/mL of ampicillin) and grown overnight at 37 °C in a shaking incubator at 150-200 rpm. Plasmid DNA was isolated from positive *E. coli* transformants using a NucleoSpin plasmid (NoLid) Kit (Macherey-Nagel, Germany), according to manufacturer's instructions. Eluted plasmid DNA was assessed for concentration and purity using a Nanodrop ND-1000 spectrophotometer and stored at -20 °C.

Table 2.2: Primer sequences used for plasmid backbone amplification and sequencing

Plasmid source	Primer number		Sequence (5' – 3' orientation)	T _a (°C)
pSLI-TGD	9	F	AGCGGATAACAATTTACACAGGA	52
	10	R	ACAAGAATTGGGACAACCTCCAGTGA	
pSLI-Sandwich	9	F	AGCGGATAACAATTTACACAGGA	52
	11	R	TCTCTGCAGAGCAGCTCTAGCA	

To further validate the positive clones obtained from colony PCR, restriction enzyme digestion was carried out on the isolated plasmids. Approximately 300 ng of recombinant plasmid DNA was digested with 1 U/µg of each of its respective enzymes (*NotI*-HF/*MluI*-HF/*AvrII*) in 1x CutSmart buffer for 3 h at 37 °C and visualised on a 1 % (w/v) agarose gel.

The gene fragments and their flanking regions were also sequenced to confirm the correct sequences were cloned. Sanger sequencing is based on the incorporation of chain-terminating dideoxynucleotides that are fluorescently labelled at the 3' end. The resulting fluorescence can be detected by a CCD camera and translated into a chromatogram. Sanger sequencing was performed to validate the insert of the recombinant plasmid. A 20 µL reaction was set up with 100 ng per 1000 bp of template DNA, 5 pmol of plasmid specific forward or reverse primers (Table 2.2), 4 µL 2x Big dye buffer (Applied

Biosystems, Foster City, USA), 2 µL 1x Big dye reaction mix (Applied Biosystems, Foster City, USA). Thermo-cycling conditions consisted of an initial template denaturation step at 96 °C for 1 min, followed by 25 cycles of denaturation at 60 °C for 10 s, primer annealing at 50 °C for 5 s and elongation at 60 °C for 4 min. The reaction was kept at 4 °C until ethanol precipitation.

Ethanol precipitation was used to purify the DNA before sequencing. A concentration of 3 M NaOAc was added in a 1:10 ratio to the PCR reaction mixture followed by three times the reaction volume of 100 % ice-cold ethanol. The mixture was centrifuged (5415R Eppendorf centrifuge, Hamburg, Germany) at 16 000 *xg* at 4 °C for 30 min. The supernatant was removed, and 250 µL ice-cold 70 % ethanol was added. The tubes were centrifuged again at 16 000 *xg* at 4 °C for 10 min to precipitate the DNA. The supernatant was removed, and DNA was allowed to air dry until all the ethanol had evaporated. Tubes were kept at 4 °C until analysis could be completed. Sequencing analysis was performed at the University of Pretoria AGCT Sequencing facility using an ABI PRISM® Genetic analyser (Applied Biosystems, Foster City, USA). The sequenced results were analysed using Benchling software. The software was used to analyse the chromatograms and the sequence of DNA.

2.7. Large scale plasmid isolation for transfection

Final recombinant plasmid constructs were isolated using a NucleoBond Xtra Midi purification kit (Macherey-Nagel, Germany), according to manufacturer's instructions from saturated *E. coli* DH5 α cultures ($OD_{600} = 4$), to obtain 50-100 µg of plasmid DNA.

To further increase the purity of the eluent, isopropanol precipitation was done. Room temperature isopropanol (3.5 mL) was added to the eluent and vortexed briefly before centrifuging at 15 000 *xg* for 30 min at 4 °C. The supernatant was removed and the pellet was washed with 2 mL of ice cold 70 % ethanol. It was further centrifuged for 5 min at 15 000 *xg* at 4 °C and the supernatant removed. The pellet was left to air dry at room temperature for 10 min before resuspending in dddH₂O.

The eluted plasmid DNA was ethanol precipitated as previously described (section 2.6.4.), including an additional wash step with 1 mL of ice cold 96 % ethanol for 10 min at 15 000 *xg* at 4 °C. The pellet was allowed to dry at room temperature inside a flow hood. The DNA was reconstituted in Cytomix (120 mM KCl, 0.15 mM CaCl₂, 5 mM MgCl₂, 10 mM K₂PO₄, 2 mM EGTA 25 mM HEPES, pH 7.6) to a final concentration of 400 ng/µL and stored at 4 °C until transfection. The naming convention for the recombinant plasmid constructs and their functionality in *P. falciparum* parasites is listed below (Table 2.3).

Table 2.3: Recombinant plasmids, their functionality and selection markers

Recombinant plasmid	Description	Selection markers	Homology region and size	Restriction site	Plasmid backbone
pSLI-TGD- <i>Sap18</i>	Disruption	hDHFR Neomycin	N-terminal 426 bp	<i>NotI</i>	pSLI-TGD
				<i>MluI</i>	
pSLI-SW- <i>Sap18</i>	Knock sideways Localisation	hDHFR Neomycin	C-terminal 936 bp	<i>NotI</i>	pSLI-Sandwich
				<i>AvrII</i>	
pSLI-TGD- <i>Sir2A</i>	Disruption	hDHFR Neomycin	N-terminal 193 bp	<i>NotI</i>	pSLI-TGD
				<i>MluI</i>	
pSLI-SW- <i>Sir2A</i>	Knock sideways Localisation	hDHFR Neomycin	C-terminal 816 bp	<i>NotI</i>	pSLI-Sandwich
				<i>AvrII</i>	

2.8. Transfection of recombinant SLI plasmids

The SLI-system was used to genetically integrate modified genes into *P. falciparum* NF54 parasite lines. A parasitaemia above 5 % (>90 % ring-stage *P. falciparum* NF54 parasite culture, 5 % haematocrit) was cultured as previously described (section 2.3.) and placed back into a shaking incubator at 37 °C in 5 % O₂, 5 % CO₂ and 90 % N₂ atmosphere. After 3 h the culture was harvested by centrifugation (3500 *xg* for 2 min) and the media was aspirated and the parasite-infected erythrocyte pellet was washed once in an equal volume of ice-cold Cytomix (120 mM KCl, 0.15 mM CaCl₂, 2 mM EGTA, 5 mM MgCl₂, 10 mM K₂HPO₄, 25 mM HEPES, pH 7.6). A 200 µL parasite-infected erythrocyte pellet (100 % haematocrit) was subsequently resuspended into the reconstituted plasmid (~100 µg) to a final volume of 450 µL. The mixture was added to a pre-chilled electroporation cuvette and electroporated using a Bio-Rad gene pulser and pulse controller unit set at maximum resistance, 0.31 kV and 950 µF to ensure time constants between 10-20 ms. Electroporated samples were immediately combined with 5 mL of complete media containing 500 µL erythrocytes (50 % haematocrit). The parasite suspension was placed into a 25 cm² canted-neck culture flask and incubated at 37 °C with 5 % O₂, 5 % CO₂ and 90 % N₂ gas supplementation. The parasites were allowed to recover for 2 h after which, they were centrifuged at 3500 *xg* for 2 min and lysed erythrocyte debris was removed. A volume of 5 mL pre-warmed culture media was added to parasites and transferred back into a stationary incubator at 37 °C in 5 % O₂, 5 % CO₂ and 90 % N₂ atmosphere.

2.8.1. Drug selection and screening of transgenic parasites

The transfected parasites were selected for both episomal uptake of the different recombinant plasmids and integration of the plasmid into the parasite's genome. Several checkpoints were used to confirm episomal uptake and integration of the recombinant plasmids into *P. falciparum* NF54 parasites. Figure 2.2 indicates the selection process, recovery and checkpoints for recombinant pSLI-TGD and pSLI-Sandwich plasmids.

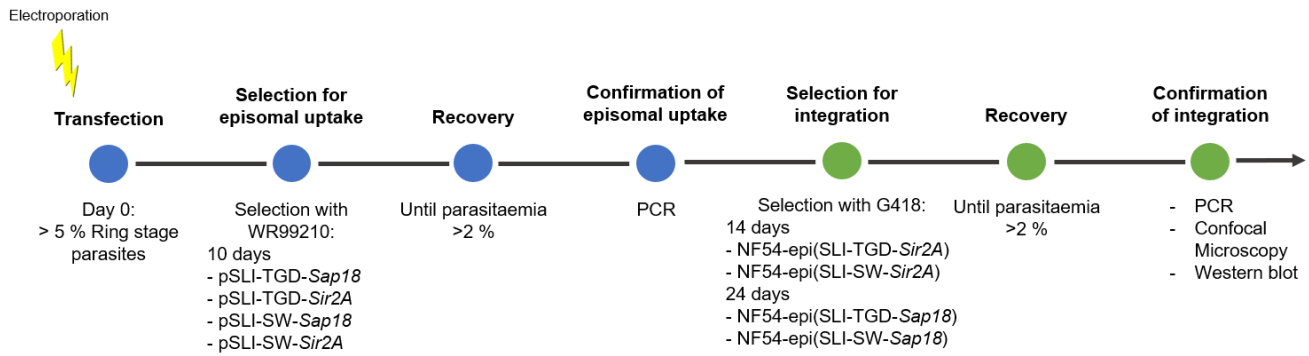


Figure 2.2: Drug selection and screening of transgenic parasites to confirm episomal uptake and integration of SLI-TGD and SLI-Sandwich recombinant plasmids. 5% ring-stage parasites transfected with pSLI-TGD-*Sap18*, pSLI-TGD-*Sir2A*, pSLI-SW-*Sap18* or pSLI-SW-*Sir2A* by electroporation. Parasites were selected for episomal uptake with WR99210 for 10 days and confirmed for episomal uptake via PCR. NF54-epi(SLI-TGD-*Sir2A*) and NF54-epi(SLI-TGD-*Sap18*) parasites were split into 3 technical replicates (performed in two biological replicates) for selection for integration with G418 for 14 (NF54-epi(SLI-TGD-*Sir2A*) or 24 (NF54-epi(SLI-TGD-*Sap18*) days and confirmed via PCR, confocal microscopy or western blot for further functional analysis. NF54-epi(SLI-SW-*Sap18*) and NF54-epi(SLI-SW-*Sir2A*) parasites were selected for integration with G418 for 14 (pSLI-SW-*Sir2A*) or 24 (pSLI-SW-*Sap18*) days and confirmed via PCR, confocal microscopy or western blot for further functional analysis.

Twenty-four hours after recovery, transfected *P. falciparum* NF54 parasites were maintained under selective drug pressure in complete culture media with 4 nM WR99210 (Jacobus Pharmaceutical Company, USA) (Figure 2.2). Media was replaced every 24 h for the first 10 days followed by the removal of drug pressure, and culture media was replaced every second day until parasites recovered. The antifolate drug, WR99210, was used to select for the resistance marker, hDHFR. The parasitaemia was calculated daily for the duration of drug selection to monitor parasite death and proliferation. The parasitaemia was maintained at less than 2 %. Fresh erythrocytes were added every 7 days to replace old lysing erythrocytes and maintain a 5 % haematocrit. After approximately two - three weeks without drug pressure parasites started to appear due to proliferation over the detection threshold.

Following an increase in parasitaemia above 2 %, after the first round of drug selection with WR99210, the recombinant cultures were increased into a 20 mL culture by addition of culture media and fresh erythrocytes. Plasmid DNA was extracted from trophozoite stage parasite cultures with parasitaemia ranging from 3 - 5 % using the Quick-DNA Miniprep Kit (Zymo Research, USA). To determine whether the plasmid was taken up episomally by the parasite, amplification of the flanking regions of the plasmids were done using the primers listed in Table 2.2. The PCR reaction mixture was set up as follows: 1x KAPA-Taq polymerase, 5 pmol of each primer (Table 2.2) and 50 ng of plasmid DNA. Thermocycler conditions were set up as follows: initial denaturation at 95 °C for 3 min, and cycling 30x at 95 °C for 30 s, 52 °C for 30 s, 68 °C for 1 min followed by a final extension at 68 °C for 5 min

After confirmation of episomal uptake via PCR, parasite cultures were frozen away in liquid nitrogen containing a 1:1 ratio of 100 % packed erythrocytes (>5 % ring-stage parasites) to freezing media (28 % (v/v) glycerol, 3 % (w/v) D-sorbitol, 0.65 % (w/v) NaCl). Following episomal uptake, recombinant parasite lines are referred to as NF54-epi(pSLI-TGD-*Sap18*), NF54-epi(pSLI-TGD-*Sir2A*), NF54-epi(pSLI-SW-*Sap18*) and NF54-epi(pSLI-SW-*Sir2A*).

Episomal parasite lines were reduced to 1 % parasitaemia in preparation for selection. NF54-epi(pSLI-TGD-*Sap18*) and NF54-epi(pSLI-TGD-*Sir2A*) parasites were split into 3 flasks to select with 400 µg/mL of G418 (geneticin selective antibiotic) (50 mg/mL in culture media, Gibco, Thermo Fisher Scientific) in performed in two biological replicates. Complete culture media containing G418 was changed every day for the first 10 days then every second day for 14 days (NF54-epi(pSLI-TGD-*Sap18*)) or every second day for four days (NF54-epi(pSLI-TGD-*Sir2A*)) after which the drug pressure was removed and cultured twice weekly until parasites appeared (Figure 2.2). NF54-epi(pSLI-SW-*Sap18*) and NF54-epi(pSLI-SW-*Sir2A*) parasites had undergone a second round of drug selection with 400 µg/mL G418. The G418 containing media was changed every day for the first 10 days then every second day for 14 days (NF54-epi(pSLI-SW-*Sap18*)) or four days (NF54-epi(pSLI-SW-*Sir2A*)) after which the drug pressure was removed and cultured twice weekly until parasites appeared (Figure 2.2).

Following the second round of drug selection with G418, genomic DNA from each of the transgenic parasite cultures was isolated using the Quick-DNA Miniprep Kit (Zymo Research, USA). A PCR across the integration junctions and one verifying absence of the unmodified locus was performed on each of these parasite lines to assess correct integration. Figure 2.3 outlines the location of primers that were designed to assess successful episomal uptake and integration of these lines. The transgenic parasite cultures were assessed for successful integration into the genome by amplification of 5' and 3' flanking genomic regions of the gene of interest (Table 2.4). Wild type parasites were identified by positive amplification of the locus.

The PCR reaction mixture was set up as follows: 1x KAPA-Taq polymerase, 5 pmol of forward primer and 5 pmol reverse primer (Table 2.4) and 50 ng of genomic DNA. PCR amplification conditions were set up as follows: initial denaturation at 95 °C for 3 min, and cycling 30x with initial denaturation of 95 °C for 30 s, annealing for 30 s using temperatures listed in Table 2.4, extension at 68 °C for 1 min followed by a final extension at 68 °C for 5 min. PCR results were analysed on a 1.5 % agarose gel. A 1 kb or 100 bp DNA ladder (Promega, USA) was used to determine DNA fragment sizes on a gel.

After confirmation of integration via PCR, parasite cultures were frozen away in liquid nitrogen containing a 1:1 ratio of 100 % packed erythrocytes (>5 % ring-stage parasites) to freezing media (28 % (v/v) glycerol, 3 % (w/v) D-sorbitol, 0.65 % (w/v) NaCl), as described previously. Following integration, recombinant parasite lines are referred to as NF54-Δ*Sap18*-GFP, NF54-*Sap18*-2xFKBP-GFP-2xFKBP, NF54-Δ*Sir2A*-GFP and NF54-*Sir2A*-2xFKBP-GFP-2xFKBP.

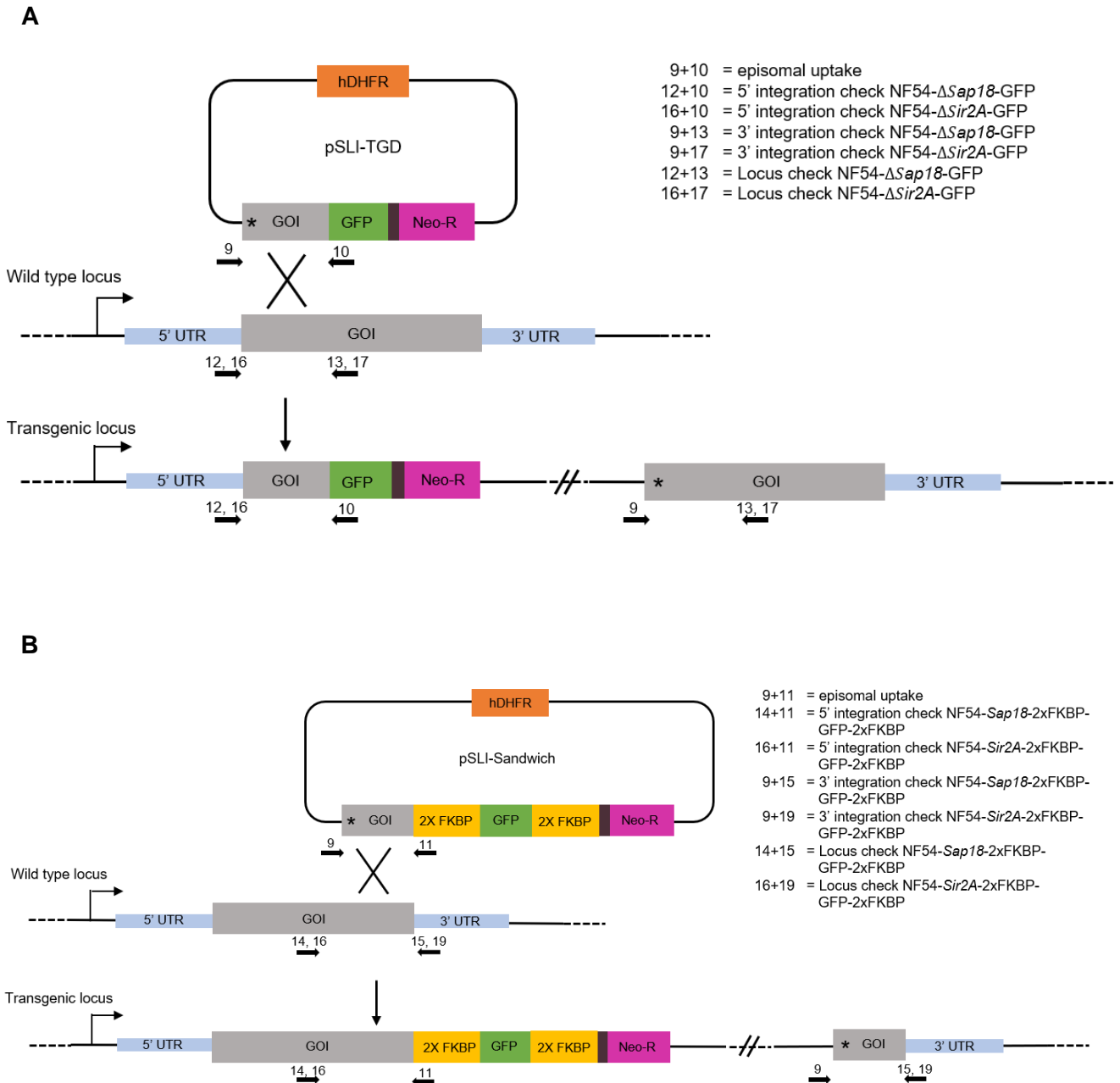


Figure 2.3: Schematic representation of the Selection Linked Integration systems. (A) Schematic of the pSLI-TGD plasmid that undergoes homologous recombination into the wild type locus. The recombinant plasmid includes a 200-500 bp of the 5' end of the gene of interest, followed by a GFP tag, a T2A skip peptide and a neomycin resistance marker cassette. **(B)** Schematic of the pSLI-Sandwich plasmid that undergoes homologous recombination into the wild type locus. The recombinant plasmid includes a 500-1000 bp of the 3' end of the gene of interest (without the stop codon), followed by FKBP-12, a GFP tag, a second FKBP-12, a T2A skip peptide and a neomycin resistance cassette. Asterisks denote a stop codon. Arrows indicate location of primers for episomal uptake and integration validation.

Table 2.4: PCR primers for the 5', 3' and loci integration check of recombinant parasite lines NF54- Δ Sap18-GFP, NF54-Sap18-2xFKBP-GFP-2xFKBP, NF54- Δ Sir2A-GFP and NF54-Sir2A-2xFKBP-GFP-2xFKBP.

Parasite line	Region amplified and size	Primer number		Primer sequence (5' to 3' orientation)	Ta (°C)
NF54- Δ Sap18-GFP	5' 795 bp	12	F	CCCATAATATATAGGTGTATATTTTAATTTTAAAAAAG	56
		10	R	ACAAGAATTGGGACAACCTCCAGTGA	
	3' 731 bp	9	F	AGCGGATAACAATTTACACAGGA	56
		13	R	CGTCTTTATAATTATCATAATAACTGATACTAGAG	
	Locus 919 bp	12	F	CCCATAATATATAGGTGTATATTTTAATTTTAAAAAAG	54
		13	R	CGTCTTTATAATTATCATAATAACTGATACTAGAG	
NF54-Sap18-2xFKBP-GFP-2xFKBP	5' 1237 bp	14	F	GGACAGTGATTATTATGATGAGGAGGG	56
		11	R	TCTCTGCAGAGCAGCTCTAGCA	
	3' 1167 bp	9	F	AGCGGATAACAATTTACACAGGA	56
		15	R	CGGACAACCTTAAATTGACCAAATG	
	Locus 1313 bp	14	F	GGACAGTGATTATTATGATGAGGAGGG	54
		15	R	CGGACAACCTTAAATTGACCAAATG	
NF54- Δ Sir2A-GFP	5' 441 bp	16	F	GCTATTATTTAATAATCCTAGAAAGTACAATCAATGC	52
		10	R	ACAAGAATTGGGACAACCTCCAGTGA	
	3' 379 bp	9	F	AGCGGATAACAATTTACACAGGA	52
		17	R	CGTCTTTATAATTATCATAATAACTGATACTAGAG	
	Locus 441 bp	16	F	GCTATTATTTAATAATCCTAGAAAGTACAATCAATGC	52
		17	R	CGTCTTTATAATTATCATAATAACTGATACTAGAG	
NF54-Sir2A-2xFKBP-GFP-2xFKBP	5' 1039 bp	16	F	GCTATTATTTAATAATCCTAGAAAGTACAATCAATGC	52
		11	R	TCTCTGCAGAGCAGCTCTAGCA	
	3' 1131 bp	9	F	AGCGGATAACAATTTACACAGGA	52
		19	R	GAACAGAATGTGTCCACATTAGGAACA	
	Locus 1193 bp	16	F	GCTATTATTTAATAATCCTAGAAAGTACAATCAATGC	52
		19	R	GAACAGAATGTGTCCACATTAGGAACA	

2.9. Confocal microscopy

For parasite morphology analysis, imaging was performed on a Zeiss LSM 880 Inverted Confocal Laser Scanning Microscope (LSM) with Airyscan detector (Zeiss, Germany). The Airyscan detector improves the signal-to-noise ratio to achieve a higher resolution. Appropriate channels were used based on fluorophore emission and excitation wavelengths. GFP has an excitation wavelength of 488 nm and emission at 500 nm which could be detected through the green channel. Hoechst has an excitation wavelength of 350 nm and emission at 460-490 nm therefore, the 420LP filter was used and could be detected through the blue channel. *mCherry* has an excitation wavelength of 587 nm and an emission of 610 nm which is detected through the red channel. A 100x oil immersion objective was used for

imaging parasites. Zeiss ZEN lite blue edition software (version 3.89) (Zeiss, Germany) was used for initial image processing and ImageJ (version 1.53) (NIH, USA) was used for final fluorescence images.

Firstly, 1 mm coverslips (Menzel-Gläser, Thermo Fisher Scientific, USA) were incubated in 1 M HCl for 24 h and rinsed three times with dddH₂O followed by 95 % ethanol. The coverslips were submerged in 0.1 mg/mL poly-L-Lysine (Sigma-Aldrich, USA) for 5 min and air-dried in the flow hood. These were placed into a 6-well tissue culture plates until use. Parasite cultures at late trophozoite/schizont stage were washed three times in 1x PBS at 300 *xg* for 3 min. The cells were resuspended in 2 mL 1x PBS to 1 % haematocrit and 2 % parasitaemia. The cells were added to the coverslips and centrifuged at 300 *xg* for 3 min. Non-adherent cells were washed off by adding 1 mL of 1x PBS down the slide of the well and the PBS was aspirated. The cells were fixed by adding 2 mL of fixative (4 % (v/v) formaldehyde, 0.025 % (v/v) glutaraldehyde in 1x PBS) onto the coverslips and incubated at room temperature for 1 h. The fixative was aspirated and the coverslips were washed with 2 mL of 1x PBS-T (1x PBS with 0.1 % (v/v) Tween 20). To stain parasite DNA, 2 mL of Hoechst (Thermo Fisher Scientific, USA) (1:100 dilution of 10 mg/mL in 1x PBS) was added to the fixed cells on the coverslip and incubated at room temperature for 30 min. To remove excess Hoechst, the fixed cells on the coverslips were rinsed three times with 1x PBS and prepared for microscopy by placing the cell-coated coverslip face down on a microscope slide with Fluoroshield (Sigma-Aldrich, USA).

2.10. Confirmation of integration via western blot

In addition to confirming the integration at the correct locus via PCR and confocal microscopy, a western blot was performed to confirm the size of the protein-GFP cassette. The parasite culture was pelleted by centrifuging at 4000 *xg* for 3 min and culture media was aspirated. Ice-cold saponin (1.5x vol) (0.15 % (w/v) saponin) was added to the pellet and incubated on ice for 10 min. Ice cold 1x PBS was added to 40 mL, the pellet re-suspended, and centrifuged at 4000 *xg* for 3 min and the supernatant was aspirated. The pellet was washed with 1x PBS until the supernatant was clear of visible haemoglobin.

Isolated parasites were lysed with 50 - 200 μ L of protein lysis buffer (4 % (w/v) sodium dodecyl sulphate (SDS), 0.5 % (v/v) Triton X-100, 0.5x PBS with protease inhibitors) by grinding parasites. Insoluble material was pelleted by centrifugation at 16 000 *xg* for 1 min at room temperature and the supernatant was resuspended in 4x Laemmli buffer (4 % SDS, 120 mM Tris-Cl pH 6.8, 20 % (v/v) glycerol, 0.02 % (v/v) bromophenol blue (Bio-Rad, SA)) with 10 % β -mercaptoethanol (Sigma-Aldrich, USA). Samples were boiled for 5 min at 95 °C. A Pierce[®] BCA Protein Assay Kit (Thermo Fisher Scientific, USA) was used to determine the concentration of the isolated proteins. Bovine serum albumin (BSA) standards were used as the known protein concentrations for the standard curve and the absorbance values at 595 nm were read using a Paradigm multiplate reader (SpectraMax, Molecular Devices, USA).

SDS-detergent denatures the proteins by disrupting non-covalent bonds and imparts a negative charge on the proteins. This linearises proteins allowing a proper separation based on size not shape. The samples were separated on a polyacrylamide gel (10 % separation gel: 30 % (v/v) acrylamide, 10 % (w/v) SDS, TEMED (Bio-Rad), 10 % (w/v) ammonium persulfate (APS), 1.5 M Tris-HCl, pH 8.8; and 4

% stacking gel: 30 % (w/v) acrylamide, 10 % (w/v) SDS, TEMED (Bio-Rad), 10 % (w/v) APS, 1 M Tris-HCl, pH 6.8) at 80 V for the stacking gel and 100 V for the separation gel with constant amps. The gel was transferred to a polyvinylidene fluoride membrane (Bio-Rad, SA) using the wet tank transfer system (Hoefer TE22, USA) for 2 h at 80 V and constant amps. Membranes were blocked in 1x TBS-T (50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1 % (v/v) Tween-20) containing 5 % skim milk powder (Sigma-Aldrich, USA) for 1 h. Dilutions for primary antibodies were as follows: rabbit polyclonal anti-GFP (Abcam, ab290) 1:1000, anti-histone H3 (Abcam, ab18521) 1:10 000 and anti-FKBP12 (Abcam, ab2918) 1:1000. These were incubated with the membrane overnight at 4 °C, followed by washing three times with 1x TBS-T. Secondary antibodies, horseradish peroxidase-conjugated goat anti-rabbit antibodies (Abcam, ab6702) were diluted 1:1000 for anti-GFP and anti-FKBP12, or 1:10 000 for anti-histone H3, and incubated with the membrane for 1 h at room temperature. Visualisation was done using SuperSignal West Pico chemiluminescent substrate (Thermo Fisher Scientific, USA) using a Chemi-Doc imaging system (Bio-Rad, SA).

2.11. Knock sideways of NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites

2.11.1. Transfection and drug selection of pLyn-FRB-*mCherry*-nmd3-BSD plasmid into NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites

In addition to transfection of the *Sap18*-2xFKBP-GFP-2xFKBP plasmid into NF54 wild type parasites, an additional plasmid, pLyn-FRB-*mCherry*-nmd3-BSD, is also required by this system to be expressed episomally. As described previously, knock sideways is based on the ligand-induced dimerization of the proteins FRB and FKBP. The protein of interest is fused with FKBP (present in the initial plasmid – SLI-Sandwich), whereas FRB is separately expressed by the nuclear mislocaliser plasmid (pLyn-FRB-*mCherry*-nmd3-BSD) (Figure 2.4). With the addition of rapalog, FRB and FKBP dimerise and the target protein is removed from its site of action by the mislocaliser, which allows for functional analysis in the absence of protein at its original site of action.

Transfection of pLyn-FRB-*mCherry*-nmd3-BSD was performed using NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites as described previously in section 2.8.2. However, episomal selection was under the control of blasticidin drug selection. Parasites were selected with 2 µg/mL of blasticidin S (50 mg/mL in culture media, Thermo Fisher Scientific, USA) for 10 days and allowed to recover. Recombinant cultures (parasitaemia >2 %) were increased to a 20 mL culture by addition of culture media and fresh erythrocytes. Genomic parasite DNA was extracted from trophozoite parasite cultures with parasitaemia's ranging from 3 - 5 % using the Quick-DNA Miniprep Kit (Zymo Research, USA). To confirm whether the plasmid was taken up episomally by the parasite, PCR amplification of the Lyn, FRB and *mCherry* fragments was done using the primers listed in Table 2.5. Cyrostocks were prepared as mentioned previously (section 2.8.1.). The expression of the mislocaliser plasmid was assessed under a confocal microscope as performed in section 2.10.

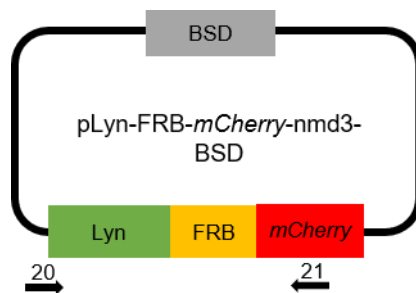


Figure 2.4: Schematic representation of the pLyn-FRB-*mCherry*-nmd3-BSD plasmid. This plasmid includes a lyn proto-oncogene fragment, FRB (FKBP12-rapamycin binding domain) and *mCherry* cassette. Arrows indicate location of primers used for episomal uptake validation.

Table 2.5: Primer sequences used for amplification of pLyn-FRB-*mCherry*-nmd3 plasmid

Primer number	Plasmid source	Primer sequence (5' – 3' orientation)		Ta (°C)
20	pLyn-FRB- <i>mCherry</i> -nmd3-BSD	F	CGAGATGGGATGTATAAAATCAAAGGGA	52
21		R	CCTGGTCACCTTCAGCTTGG	

2.11.2. Functional analysis of NF54-*Sap18*-2xFKBP-GFP-2xFKBP/FRB-*mCherry* by knock sideways

NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites containing the episomally expressed pLyn-FRB-*mCherry*-nmd3-BSD plasmid, were synchronised twice with 5 % D-sorbitol as described previously (section 2.4.). A 5 % ring-stage parasite culture was prepared and split into two culture flasks. One culture flask contained vehicle (control), while the other flask contained 250 nM of rapalog (25 µM in culture media, AP-21967, Clontech, Takara Bio, Japan). Mislocalisation was assessed with the Zeiss Confocal Microscope after 24 h of rapalog treatment. Parasitaemia increase was monitored over 96 h with a starting parasitaemia of 0.5 % (>90 % ring-stage parasites). Every 24 h, the parasitaemia of rapalog treated NF54-*Sap18*-2xFKBP-GFP-2xFKBP/FRB-*mCherry* asexual stage parasites was compared to that of untreated NF54-*Sap18*-2xFKBP-GFP-2xFKBP/FRB-*mCherry* parasites by counting 1000 erythrocytes for each sample. Rapalog was added every 24 h for NF54-*Sap18*-2xFKBP-GFP-2xFKBP/FRB-*mCherry* parasites.

2.12. Functional analysis of disrupted NF54- Δ *Sap18*-GFP parasites

NF54- Δ *Sap18*-GFP and NF54 wild type control parasites were synchronised twice and reduced to a starting parasitaemia of 1 % (>90 % ring-stage parasites). These parasites were visualised by Giemsa-stained light microscopy over 96 h to evaluate morphology and fold changes over the intraerythrocytic cycle. Every 24 h, the parasitaemia of NF54- Δ *Sap18*-GFP asexual stage parasites was compared to that of NF54 wild type parasites by counting 1000 erythrocytes for each slide. In addition, number of merozoites per schizont was determined in comparison to NF54 wild type parasites by counting 40 schizont stage parasites. The new lines' ability to undergo gametocytogenesis was also determined and compared to that of NF54 wild type parasites. Gametocyte stage development was evaluated on

NF54- Δ Sap18-GFP parasites over the 14-day developmental period. Routine culturing and gametocyte induction were performed as previously described.

2.13. Confirming the IC₅₀ of DFMO on NF54 wild type and NF54- Δ Sap18-GFP parasites using a SYBR green I - based fluorescence assay

SYBR Green I-based fluorescence assay was used to assess the IC₅₀ of *in vitro* DFMO on NF54 wild type and NF54- Δ Sap18-GFP parasites. The assay is based on the detection of nucleic acid content. SYBR Green I will bind to double stranded DNA found within parasitic cells. An >80 % synchronised ring-stage intraerythrocytic *P. falciparum* parasite culture (1 % parasitaemia, 2 % haematocrit) was used. A positive control, known as chloroquine disulphate (1 μ M) was used for inhibition and an untreated negative control was used to assess parasite growth. DFMO was dissolved in 1x PBS in a 96-well plate serially diluted two-fold leading to an eighteen-point assay. The parasites were placed in a gas chamber, gassed (5 % CO₂, 5 % O₂, and 90 % N₂) for 1 min and incubated at 37 °C in a stationary incubator for 96 h. An equal volume of parasite suspension was resuspended in SYBR Green lysis buffer ((0.002 % (v/v) SYBR Green I dye (Invitrogen, USA), 20 mM Tris, 0.008 % (w/v) saponin (Sigma-Aldrich, USA), 5 mM EDTA (ethylene dinitrotetraacetic acid, Merck, Germany), 0.08 % (v/v) Triton X-100 (Sigma-Aldrich, USA)). SYBR Green I fluorescence was measured using a Fluoroskan Ascent FL microplate fluorometer (Thermo Fisher Scientific, USA) at excitation wavelength of 485 nm and emission wavelength of 538 nm. The parasite proliferation was normalised to the negative control by subtracting background fluorescence of the positive control and expressing it as a percentage. Data was analysed and a non-linear regression curve was generated using GraphPad 6.01 software to calculate the half-maximal inhibitory concentration (IC₅₀) for DFMO.

2.14. Assessing the effect of DFMO treatment on NF54 wild type and NF54- Δ Sap18-GFP parasites

The effect of DFMO treatment on NF54 wild type parasites and NF54- Δ Sap18-GFP parasites was determined by morphological evaluation and parasitaemia changes over 96 h. Intraerythrocytic *P. falciparum* cultures (1 % parasitaemia, 5 % haematocrit) were grown and synchronised with 5 % (w/v) D-sorbitol twice to produce a tightly synchronised (>90 %) ring-stage population. Parasites were treated with 2xIC₅₀ (2 mM) DFMO at 0-12 hours post-invasion (hpi). After 24 h of DFMO treatment, DFMO-arrested parasites were reversed with 2 mM putrescine (Sigma-Aldrich, USA) and parasite recovery was monitored based on parasitaemia differences and morphology every 24 h up to 96 h post DFMO treatment.

Chapter 3: Results

3.1. *In silico* analysis and characterisation of *sap18* and *sir2A* genes

In silico analysis was performed for both *sap18* and *sir2A* to obtain functional predictions of these genes in *P. falciparum*.

3.1.1. SAP18: a member of the histone deacetylase subunit

In order to confirm and expand on the automated annotation of PF3D7_0711400 as a gene encoding a protein with SAP18 activity, we performed several complementary bioinformatics analyses of this gene. In the first instance, we analysed the conserved nature of this gene both within *Plasmodia* as well as in other species using BLASTP and a Maximum Likelihood phylogenetic tree was constructed using MEGA X (Figure 3.1).

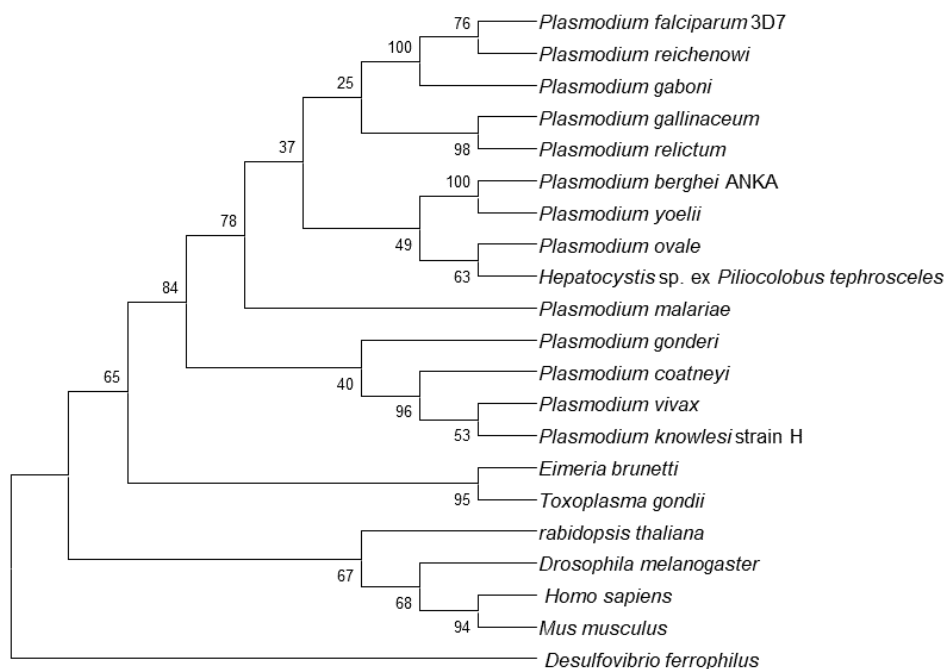


Figure 3.1: Phylogenetic tree of SAP18 proteins in different *Plasmodium* species, other apicomplexans and non-parasitic relatives. Protein sequences were aligned using the MUSCLE algorithm and the phylogenetic tree was created using the Maximum Likelihood method and JTT + G + I matrix-based model. Nodal values represent bootstrap probabilities based on 100 replicates and are shown next to the branches. Evolutionary analysis was performed in MEGA X.

The phylogenetic tree of various SAP18 proteins bears reliable bootstrap percentages of above 70 % for the majority of the nodes, however some nodes may not be accurate predictions of evolutionary relationships due to their low bootstrap values (Figure 3.1). It is clear that PF3D7_0711400 is conserved among all *Plasmodia*, including human, murine and primate species (Figure 3.1). Interestingly, the highest sequence similarity was seen between *P. falciparum* and *P. reichenowi*, while the lowest sequence similarity was seen between *P. falciparum* and *P. ovale*. *P. malariae* showed distinct

separation from *P. falciparum* and its neighbouring species. In addition, *P. malariae* showed the closest similarity to *Hepatocystis*, a member of the apicomplexan family. Non-parasitic relatives, such as *H. sapiens* and *M. musculus*, show a distant evolutionary relationship, suggesting that PF3D7_0711400 functionality may differ in *P. falciparum*. A low sequence similarity was found between *P. falciparum* and *A. thaliana*, which leads back to the question of SAP18 as a regulator in the circadian rhythm. *Desulfovibrio ferrophilus*, a prokaryote, was used as an outgroup as it bears low similarity to *P. falciparum* SAP18.

To obtain the functional annotation of PF3D7_0711400, we first performed protein domain predictions with InterPro and SMART. This evaluated the identity and functionality of PF3D7_0711400 based on similar domains. PF3D7_0711400 contains a SIN3-associated polypeptide p18 (SAP18) homologous superfamily located at the C-terminus at residues 593-728 (InterPro: IPR042534) as well as a SAP18 protein family (InterPro: IPR010516) located at residues 562-728 (Figure 3.2). An ASAP complex subunit, SAP18 (CATH-Gene3D:G3DSA:3.10.20.550) domain was also identified, which leads back to the possibility of PF3D7_0711400 as a component in the ASAP complex. The Pfam database identified a SAP18 domain at residues 608-723. Furthermore, PF3D7_0711400 also contains several low complexity regions (Figure 3.2) and a coiled-coil region (Figure 3.2). These low complexity regions were not identified within the SAP18 domain. These results contribute to the functional annotation of PF3D7_0711400 as a member of the SIN3-associated polypeptide p18 family.

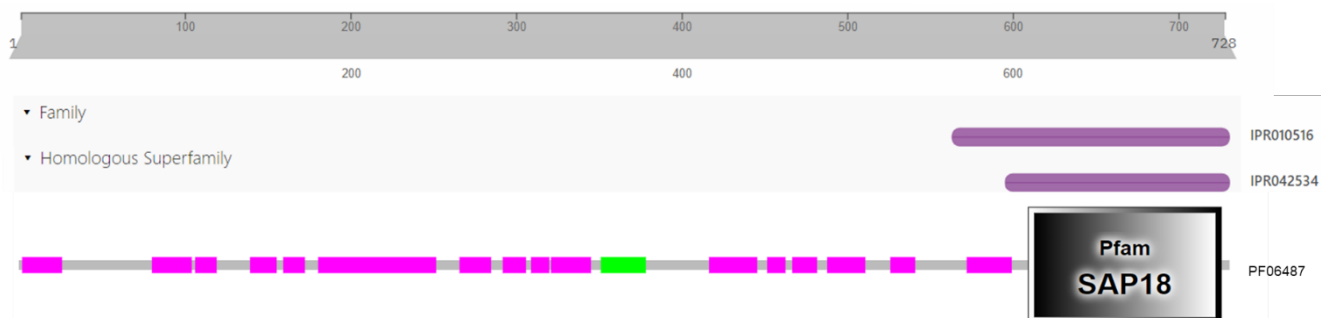


Figure 3.2: Schematic representation of PF3D7_0711400 (SAP18) architecture. A SIN3-associated polypeptide p18 family (residues 562-728, IPR010516, purple) and superfamily (residues 593-728, IPR042534, purple) are located at the C-terminus of PF3D7_0711400. A SIN3-associated polypeptide p18 domain located at residues 608-723 (PF06487, grey). Low complexity regions are shown in pink. Coiled-coil region is shown in green. Information obtained from SMART (<http://smart.embl-heidelberg.de/>) and InterPro (<https://www.ebi.ac.uk/interpro/>).

To evaluate the prediction of PF3D7_0711400 with a SAP18 domain, we used an expect (E)-value parameter and sequence identity to confirm this domain. A SAP18 domain identified at residues 608-723 (Figure 3.2) had an E-value of 1.0×10^{-19} . The E-value is a variable that represents the number of hits due to chance when searching a database of a certain size. A good E-value which gives a confident prediction is much lower than 1. Thus, a lower E-value indicates a more similar sequence identity to the consensus domain sequence. A Hidden Markov model (HMM) consensus SAP18 domain sequence was aligned to the predicted domain sequence of PF3D7_0711400 (Figure 3.3). A consensus sequence is defined from multiple sequence alignments. We observed a sequence identity of 34 % using ClustalO.

Thus, using both E-value parameter and sequence identity, we confirmed that the domain predicted by Pfam, on a sequence level, resembles a SAP18 protein-like subunit.

HMM consensus	IDREKTPFLLRFLFYKTGSFHRLDEFSS-SGKLPSEDELQIYTWKDCTLRELAQLLKSALP	59
<i>P. falciparum</i>	INREKTCPFLLRFLFYKVDKEYNVDDMDILTNDNSNELQIYAWIDITMREIVTLVKDFYK	60
	*:**** *****... :*::: : . *:*****:* * *:***: *::*	
HMM consensus	SLLPSPAVGTRLSFRLVYPDTRRGRYASKDLGGSVSGKKGSEAEDDSEEEEEKERSEEEED	119
<i>P. falciparum</i>	DSRQR---NAQWVFKVFSYEKKKLTFLSKVHST-----IYNYKE	96
	. ::: *::: :::: : ** : :	
HMM consensus	ADKTLADLRFVIGDYLDVAI	139
<i>P. falciparum</i>	DNKTLLSLNYEIGDIILLSI	116
	:*** .*: *** : ::*	

Figure 3.3: Sequence alignment of *P. falciparum* (residues 608-723) and HMM SAP18 domain consensus. Protein sequences were obtained from PlasmoDB (<https://plasmodb.org/>) and Pfam (<http://pfam.xfam.org/>) and aligned with ClustalO (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

Secondly, we evaluated the functional annotation of PF3D7_0711400 based on structure. A homology model was built using the crystal structure of *M. musculus* SAP18 (residues 1-143, PDB 4a90.1) as a template. The SWISS model database (<https://swissmodel.expasy.org/>) uses homology-based modelling to construct a predicted model of the target protein from its amino acid sequence using a 3D structure of a similar homologous protein as a template [126]. A sequence identity of 36 % was identified between *M. musculus* SAP18 template (residues 20-137) to that of the *P. falciparum* SAP18 protein sequence (residues 608-723), however, the sequence coverage was only 16 % over the whole protein sequence, particularly showing coverage over the *P. falciparum* SAP18 domain (residues 608-723) (Figure 3.4). This is to be expected due to the abundance of low complexity regions within the protein (Figure 3.2) as well as the residue length differences between *P. falciparum* SAP18 (728 residues) and *M. musculus* (153 residues).

<i>M. musculus</i>	IDREKTCPLLLRVFTTNGRHH--RMDEFSRGNVPSELQIYTWMDATLKELTSLVKEVY	58
<i>P. falciparum</i>	INREKTCPFLLRFLFYKVDKEYNVDDMD-ILTKDNNSELQIYAWIDITMREIVTLVKDFY	59
	*:*****:***:* : : : ** : : *:*****:* * *:***: *::*	
<i>M. musculus</i>	PEARKKGTHFNFAIVFMDLKRPGYRVEIGSTMSGRKGTDDSMTLQSQKFQIGDYLDIAI	118
<i>P. falciparum</i>	KDSRQRNAQWVFKVFSYEKKKLTFLSKV-HSTIYN--YKEDNKTLLSLNYEIGDIILLSI	116
	::*::: : * : : * : * ** : . :*: ** * : :*** : :*	

Figure 3.4: Sequence alignment of *P. falciparum* (residues 608-723) and *M. musculus* (residues 20-137) SAP18. Protein sequences were obtained from PlasmoDB (<https://plasmodb.org/>) and UniProt (<https://uniprot.org/>) and aligned with ClustalO (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

The SAP18 homology model (residues 608-723) was aligned to residues 20-137 of the *M. musculus* template using MatchMaker utility in UCSF Chimera (version 1.15) (Figure 3.5 A). A root-mean-square deviation (RMSD) value between 103 pruned atom pairs of 0.239 Å, was identified. This shows high structural similarity between the two core structural regions of the SAP18 domains, as a RMSD value closer to 0 indicates more similar 3D structures [127]. However, over all 111 common atom pairs, a

RMSD value of 1.494 Å, was identified. These RMSD value differences may be due to slight structural differences identified at various atom pairs between the model and template. In addition, the *P. falciparum* SAP18 homology model resembles an ubiquitin-like domain found within the SAP18 *H. sapiens* homologue (residues 20-143), which includes β -sheets (β 1-5) and α -helices (H1-4) (Figure 1.5) [95]. Therefore, the *P. falciparum* SAP18 domain identified within PF3D7_0711400, closely resembles the *M. musculus* SAP18 domain, on a structural level, which could potentially provide confidence as to a similar functionality.

A Ramachandran plot was also generated to show the grouping of β -sheets and right-handed α -helices based on phi (ϕ) and psi (ψ) angles. A Ramachandran plot is used to represent allowed regions of a polypeptide backbone based on psi and phi torsion angles of amino acid residues present in a protein structure. It is used to further understand the structure of a protein and the conformation of the amino acids. The Ramachandran plot shows that 92 % of the residues consist of acceptable phi and psi angles of α -helices and β -sheets within the peptide, while 12 out of 1361 residues fall outside the phi and psi allowed torsional angles (Figure 3.5 B).

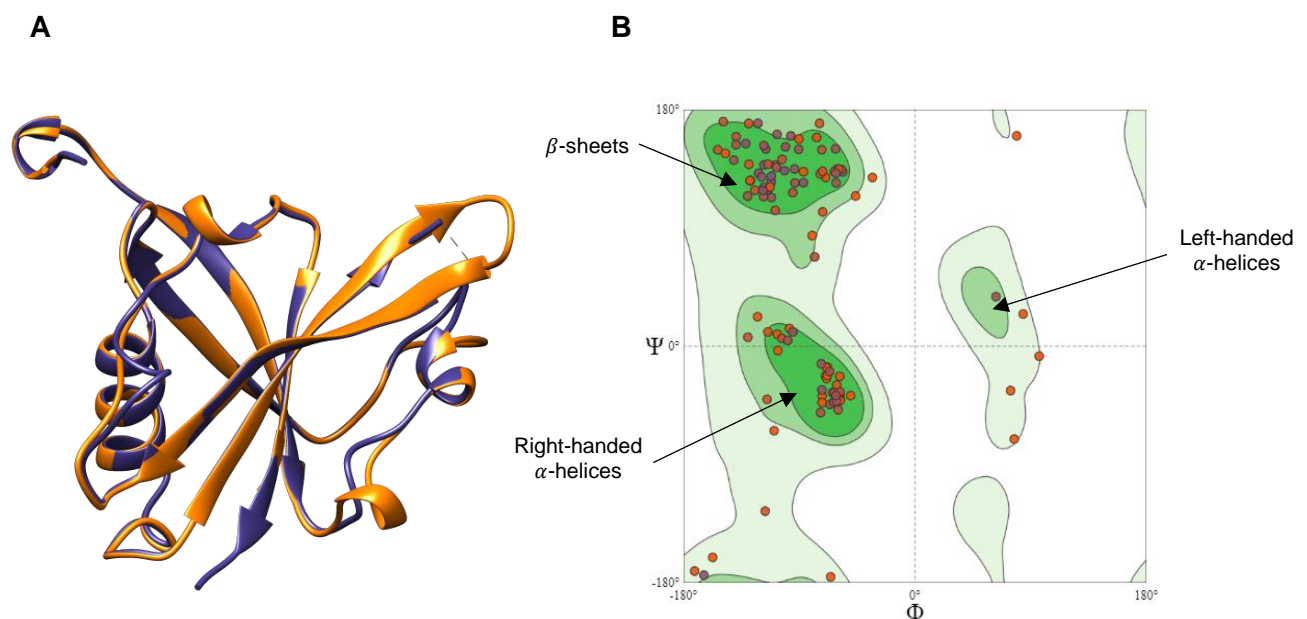


Figure 3.5: Predicted structure of PF3D7_0711400 modelled with SWISS model. (A) Secondary structure alignment of *P. falciparum* SAP18 (608-723, orange) and *M. musculus* SAP18 (PDB 4a90.1) (residues 20-137, purple). The modelled structure represents the SIN3-associated polypeptide p18 domain (residues 608-723). **(B)** Ramachandran plot of PF3D7_0711400 domain model indicating 92 % of residues with allowed phi and psi angles confirming α -helix and β -sheet formation with 12 out of 1361 residues as outliers.

Since homology-based modelling only modelled the structure of the SAP18 domain (residues 608-723), we also looked at the AlphaFold protein structure database (<https://alphafold.ebi.ac.uk/>) which predicts the proteins 3D structure based on deep learning predictions which incorporates both the physical and biological information about protein structure [128]. The AlphaFold database predicted the whole protein structure of SAP18 (residues 1-728), with regions of very low to low confidence scores, which indicate a poor model accuracy (Figure 3.6 A). However, high confidence scores were observed spanning the

SAP18 domain, further contributing to the annotation of PF3D7_0711400 with a SAP18 domain, with possible HDAC1 complex function.

We aligned the AlphaFold predicted *P. falciparum* SAP18 protein structure (<https://alphafold.ebi.ac.uk/entry/Q8IBX6>) (Figure 3.6 A), and the model generated from SWISS model (Figure 3.5 A) using MatchMaker utility in UCSF Chimera (version 1.15) (Figure 3.6 B). We used RMSD to determine the structural similarity between AlphaFold SAP18 protein structure compared to the homology model generated. A RMSD value of 0.71 Å between 76 pruned atom pairs, was obtained however, across all 116 common atom pairs the RMSD value was 4.246 Å. This value is a lot larger when compared to the alignment with the *M. musculus* SAP18 domain, indicating larger structural differences of the homology model to the predicted AlphaFold sequence. Although the proteins sequences of the homology model to AlphaFold template are identical, their secondary structures differ largely at particular outer residues contributing to the large structural differences. However, the RMSD value obtained between the AlphaFold structure and the homology modelled SAP18 structure indicates high structural similarity over the core region (76-103 residues) of the SAP18 domain.

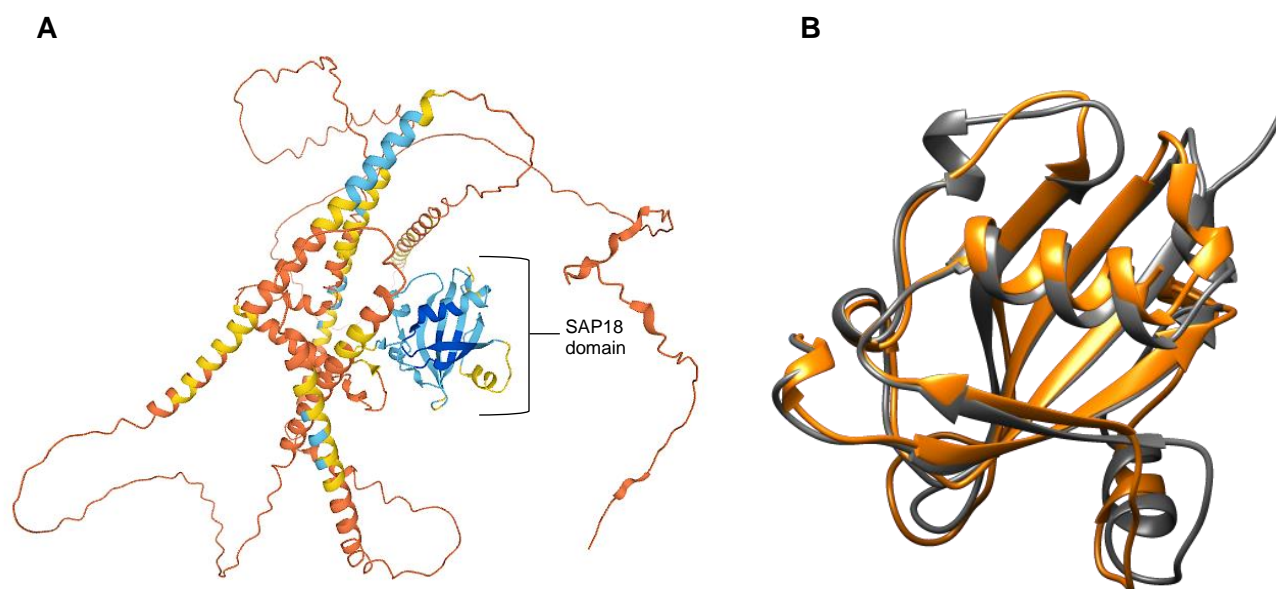


Figure 3.6: Predicted structure of PF3D7_0711400 using AlphaFold database. (A) Predicted *P. falciparum* PF3D7_0711400 structure in AlphaFold (<https://alphafold.ebi.ac.uk/entry/Q8IBX6>). Enclosed region indicates the SIN3-associated polypeptide p18 domain. A per-residue confidence score (pLDDT) is given between 0-100. Dark blue residues indicate very high confidence (pLDDT > 90), light blue residues indicate confidence (90 > pLDDT > 70), yellow residues indicate low confidence (70 > pLDDT > 50), orange residues indicate very low confidence (pLDDT < 50). **(B)** Secondary structure alignment of PF3D7_0711400 modelled with SWISS model (residues 608-723, orange) and the trimmed AlphaFold structure (residues 608-723, grey).

Using both homology-based modelling as well as machine learning predictions, we are able to confirm that the domain predicted by Pfam, on a structural level, resembles a SAP18 protein-like subunit.

3.1.2. Protein-protein interactions of SAP18

Since SAP18 is predicted to form part of the HDAC1 complex, a STRING network (<https://string-db.org/>) was generated based on information obtained from neighbourhood, experiments, databases, co-occurrence and co-expression. (Figure 3.7). Data from STRING network database and information from yeast-two-hybrid study were imported into Cytoscape (version 3.8.2).

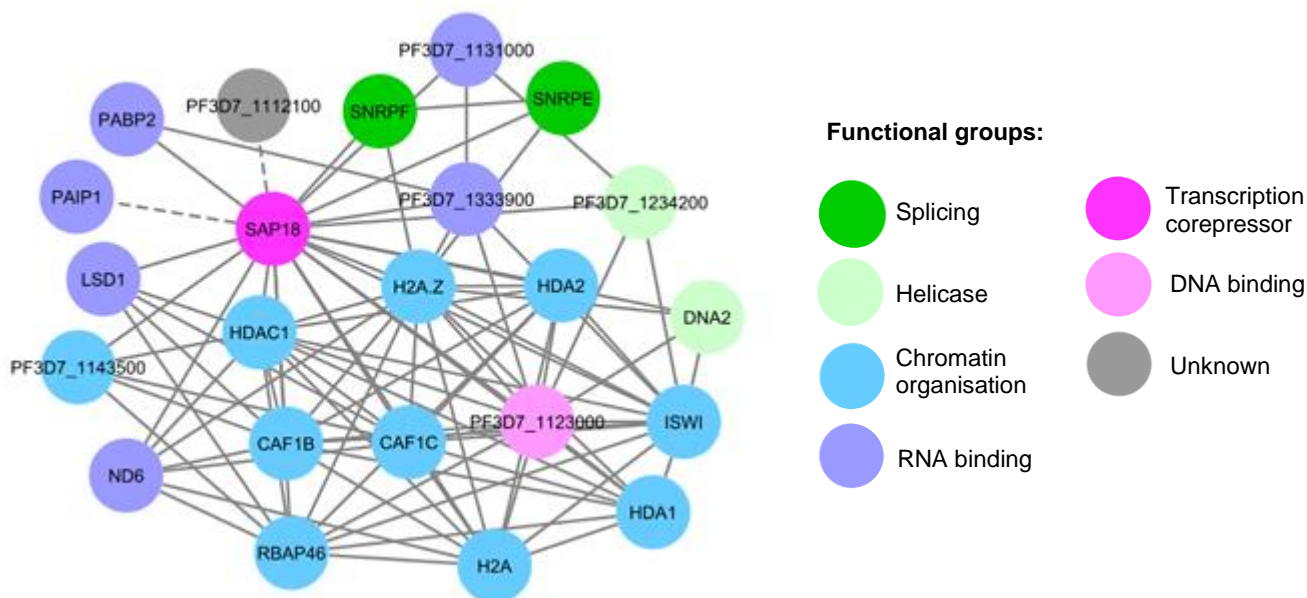


Figure 3.7: A schematic model of the protein-protein interaction complex of PF3D7_0711400 and its predicted functional partners. This network was manually curated and placed in a network based on evidence from STRING interactions and previous yeast-two-hybrid protein interactions [129]. Functional predictions are based on gene ontology obtained from PlasmoDB. Shown in dark pink is SIN3-associated polypeptide p18 SAP18 (PF3D7_0711400) involved in transcriptional corepressor activity. Shown in purple blocks are proteins associated with RNA binding: PAIP1 polyadenylate-binding protein-interacting protein 1, putative PAIP1 (PF3D7_1107300), conserved *Plasmodium* protein LSD1 (PF3D7_1220300), uncharacterised protein (PF3D7_1333900), ND6 protein (PF3D7_0508500), RNA-binding protein s1, putative (PF3D7_1131000), polyadenylate-binding protein 2, putative PABP2 (PF3D7_0923900). Shown in blue blocks are proteins associated with chromatin organisation: conserved *Plasmodium* protein (PF3D7_1143500), uncharacterized protein chromatin assembly factor 1 subunit CAF1B (PF3D7_1329300), histone H2A (PF3D7_0617800), chromatin-remodeling complex ATPase ISWI (PF3D7_0624600), histone H2A.Z (PF3D7_0320900), uncharacterised protein chromatin assembly factor 1 P55 subunit RBAP46 (PF3D7_1433300), chromatin assembly factor 1 protein WD40 domain CAF1C (PF3D7_0110700), histone deacetylase type 1 subfamily HDAC1 (PF3D7_0925700), histone deacetylase, putative HDA2 (PF3D7_1008000), Histone deacetylase, putative HDA1 (PF3D7_1472200). Shown in light pink blocks are proteins associated with DNA binding: uncharacterised protein Myb-like DNA-binding domain, putative (PF3D7_1123000). Shown in light green are proteins associated with helicase activity: DNA replication ATP-dependent helicase/nuclease DNA2 (PF3D7_1106700), conserved *Plasmodium* protein (PF3D7_1234200). Shown in dark green are proteins associated with RNA splicing: small nuclear ribonucleoprotein E putative SNRPE (PF3D7_1350200), small nuclear ribonucleoprotein F putative SNRPF (PF3D7_1126900). Shown in grey blocks are proteins associated with unknown function: conserved *Plasmodium* protein (PF3D7_1112100). Solid lines represent data from STRING (<https://string-db.org/>) and dotted lines represent data from a yeast-two-hybrid study [129].

Several interacting partners of *P. falciparum* SAP18 were identified within the STRING network (Figure 3.7), one of which was HDAC1, a component of the eukaryotic SIN3/HDAC complex (Figure 1.4 A). HDAC1 (total score of 0.87), HDA1 (0.68) and HDA2 (0.68) were identified, as expected, leading to the hypothesis of SAP18 as an interacting partner within the HDAC complex of *P. falciparum*. These data strongly predict SAP18's involvement in the HDAC family of proteins. However, we did not identify any

proteins with SIN3 activity, as would have been expected, as *P. falciparum* SAP18 was identified on both structural and sequence level to contain a SIN3-associated polypeptide p18 domain. This could possibly indicate that no such protein exists in *P. falciparum* or SAP18 has developed an alternative method to enhancing HDAC activity.

By contrast, the ASAP complex found within eukaryotes (Figure 1.8 A), identified three proteins SAP18, Acinus and RNPS1 [94]. Proteins with similar properties were identified within the STRING network of *P. falciparum* (Figure 3.7). A polyadenylate-binding protein-interacting protein 2 (PABP2, PF3D7_0923900) contains an RNA-recognition motif (RRM), known as the ribonucleoprotein (RNP) domain. The yeast-two-hybrid model identified a protein interacting with SAP18 with similar function containing an armadillo (ARM) repeat with an RNA binding function (PAIP1 (PF3D7_1107300)). This provides insight into an additional function of *sap18* as an adapter protein in RNA binding and mRNA processing. Two proteins involved in RNA-splicing via the spliceosome complex were identified in the STRING network; namely small nuclear ribonucleoprotein E (SNRPE) and small nuclear ribonucleoprotein F (SNRPF). These proteins are found within small nuclear ribonucleoproteins (snRNPs), which function in mature mRNA formation, and are a major component of the eukaryotic spliceosome [130]. This coincides with previous data within higher eukaryotes, where Acinus, RNPS1 and SAP18 have been identified in purified spliceosomes or in purified mRNPs [131]. A DNA replication ATP-dependent helicase/nuclease, DNA2 (PF3D7_1106700) and a conserved helicase protein (PF3D7_1234200) were identified in the STRING network which coincides with *sap18*'s additional predicted function in the regulation of transcription.

Evidence of these protein-protein interactions within the SIN3/HDAC and ASAP complex as well as its unique transcription profile at the G1/S transition point leads to the suggestion that *sap18* may indeed be essential to asexual parasite proliferation.

3.1.3. SIR2A: A member of the sirtuin family with NAD⁺-binding activity

To confirm and expand on the automated annotation of PF3D7_1328800 as gene encoding a protein with SIR2 activity, we performed several complementary bioinformatics analyses of this gene. We analysed the conserved nature of this gene both within *Plasmodia* as well as in other species using BLASTP and a phylogenetic tree was constructed using MEGA X (Figure 3.8). Sequence comparison between different *Plasmodium* species revealed a high sequence conservation (Figure 3.8). Apicomplexans show distinct clustering from their non-parasitic relatives although SIR2A appears

highly conserved in species outside of the *Plasmodium* lineage. Possibly indicating a similar evolutionary conserved function in PF3D7_1328800 as a NAD⁺-dependent histone deacetylase.

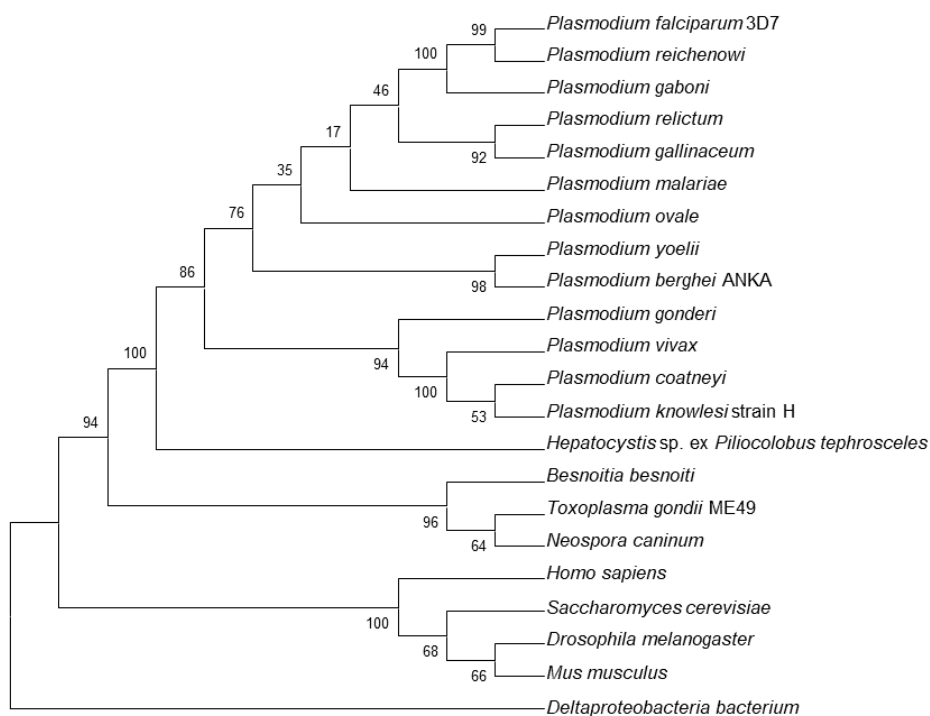


Figure 3.8: Phylogenetic tree of SIR2A proteins in different *Plasmodium* species, other apicomplexans and non-parasitic relatives. Protein sequences were aligned using the MUSCLE algorithm and the phylogenetic tree was created using the Maximum Likelihood method and LG + G matrix-based model. Nodal values represent bootstrap probabilities based on 100 replicates and are shown next to the branches. Evolutionary analysis was performed in MEGA X software (<https://www.megasoftware.net/>).

To obtain the functional annotation of PF3D7_1328800, we first performed protein domain predictions with InterPro. This evaluated the identity and functionality based on similar domains. SIR2A belongs to a sirtuin family of proteins (InterPro: IPR003000) (residues 36-213), which is further divided into five classes (I-IV and U) based on phylogenetic analysis. SIR2A was identified to belong to a SIR2 class III family located at residues 29 to 264 (InterPro: IPR027546) (Figure 3.9). This family consists of enzymes which catalyse NAD⁺-dependent reactions as well as histone deacetylases. Sirtuins typically consist of two highly variable N- and C-terminal domains and a conserved catalytic core domain (InterPro: IPR026590, residues 19-273). SIR2A contains a DHS-like NAD/FAD-binding domain, also known as a thiamine pyrophosphate (TPP) enzyme central domain, located between residues 11-255 (InterPro: IPR029035). The TPP enzyme domain, requires TPP (vitamin B1) as a co-factor and contains a 2-fold Rossmann fold, similar to the FAD/NAD-binding Rossmann fold domain (Figure 1.7). Various zinc, substrate and NAD⁺-binding sites are also present along SIR2A (Figure 3.9), which provide structural integrity to SIR2A and allow it to complete its function.

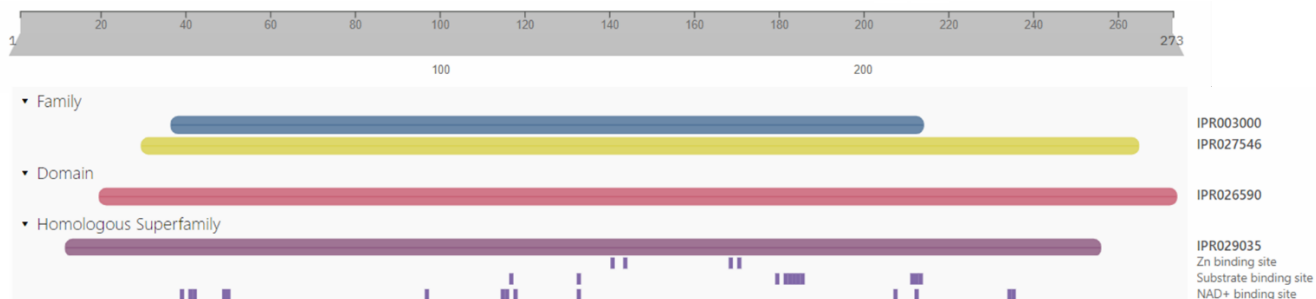


Figure 3.9 Schematic representation of PF3D7_132880 (SIR2A) architecture. A SIR2 catalytic domain is located between amino acid residues 19-273 (IPR026590, pink). SIR2A is part of the sirtuin family of proteins located at residues 36-213 (IPR003000, blue) and belongs to the class III SIR2 family located at residues 29-264 (IPR027546, yellow). SIR2A contains a DHS-like NAD/FAD-binding domain superfamily located at residues 11-255 (IPR029035, purple). NAD⁺-, substrate- and Zn-binding sites are shown. Information obtained from InterPro (<https://www.ebi.ac.uk/interpro/>).

To evaluate the prediction of PF3D7_132880 with a SIR2 domain, we used an E-value parameter and sequence identity to confirm this domain. A SIR2 domain identified at residues 36-213 (Figure 3.9) had an E-value of 4.3×10^{-37} . This value is much lower than 1, making it a confident domain prediction. An HMM consensus SIR2 domain sequence was aligned to the predicted domain sequence of PF3D7_132880 (Figure 3.10). We observed a sequence identity of 35 % using ClustalO. Thus, using both E-value parameters and sequence identity, we confirmed that the domain predicted by InterPro, on a sequence level, resembles a SIR2 protein.

HMM consensus	GAGISTESGIPDFRSKEGL -YAKLAKEELASPEALF SKKLLVEEVFYNIAREL - LAEEAE	58
<i>P. falciparum</i>	GSGETSAESNIPFRGSSNSIWSKYDPRIYGTIWFVK ---YPE -KIWEVIRDISSDYEIE	56
	: *:*:*:*:*:*:*..... ::* . .: .:.. * :*: *:* * *	
HMM consensus	PNAAHKALKKLEKKGKLLRLITQNIIDGLEKKGAGSKKVVVEIHGSLAKAKCVKCEKKIDKEE	118
<i>P. falciparum</i>	INNGHVALSTLESGLYKSVVTQNVNVDGLHEASGNTKVISLHGNVFEAVCCTCNKIVKLNK	116
	* . * *:*:*:* * * :*:*:*:*:*: :*:*:*:*:*:*: * * .:* * .: .:	
HMM consensus	LQEE ---IEAEKVPECEKCGKI1KPDVWVFFGEA1PDK1-KKIIIEKVEEAD111VIGTS1K	174
<i>P. falciparum</i>	IMLQKTSHFHQLPPECEPCGGIFKPNIIILFGEVSSDLLKEAEEIACDLLLVIGTSST	176
	: : .:* ** *:*:*:*:*:*: .:* * :* :*:*:*:*:* *	
HMM consensus	VA 176	
<i>P. falciparum</i>	VS 178	
	*:	

Figure 3.10: Sequence alignment of *P. falciparum* (residues 36-213) and HMM SIR2 domain consensus. Protein sequences were obtained from PlasmoDB (<https://plasmodb.org/>) and Pfam (<http://pfam.xfam.org/>) and, aligned with ClustalO (<https://www.ebi.ac.uk/Tools/msa/clustalo/>).

Secondly, we evaluated the functional annotation of SIR2A based on structure. Figure 3.11 indicates the 3D X-ray crystal structure of SIR2A covering residues 1-273 (PDB 3u31.1) [132]. The crystal structure consists of the SIR2A protein (residues 10-273) and a histone 3 myristoyl lysine 9 peptide (residues 1-9). SIR2A binds to NAD⁺, glycerol and a Zn²⁺ atom to provide its functionality (Figure 3.11). The resolution of the X-ray diffraction structure was 2.2 Å. Thus, we are able to confirm that the domain

predicted by InterPro, on a structural level, resembles a SIR2A protein with histone deacetylase and NAD⁺-binding activity.

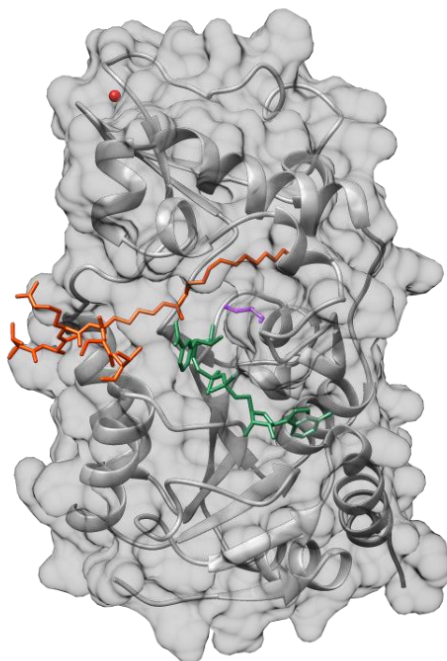


Figure 3.11: X-ray crystal structure of SIR2A. SIR2A crystal structure contains a transcriptional regulatory protein SIR2 homologue (residues 10-273, grey) (PBD 3u31.1A) and a histone 3 myristoyl lysine 9 peptide (residues 1-9, orange) (PBD 3u31.1B). NAD (green), glycerol (purple) and zinc atom (red) residues are shown at their respective binding sites.

3.1.4. Protein-protein interactions of SIR2A

To obtain an indication of the interacting partners of SIR2A in *P. falciparum*, a STRING network (<https://string-db.org/>) was generated based on information obtained from neighbourhood, experiments, databases, co-occurrence and co-expression (Figure 3.12). There is no yeast-two-hybrid data available for this protein. Data from STRING network database was imported into Cytoscape (version 3.8.2).

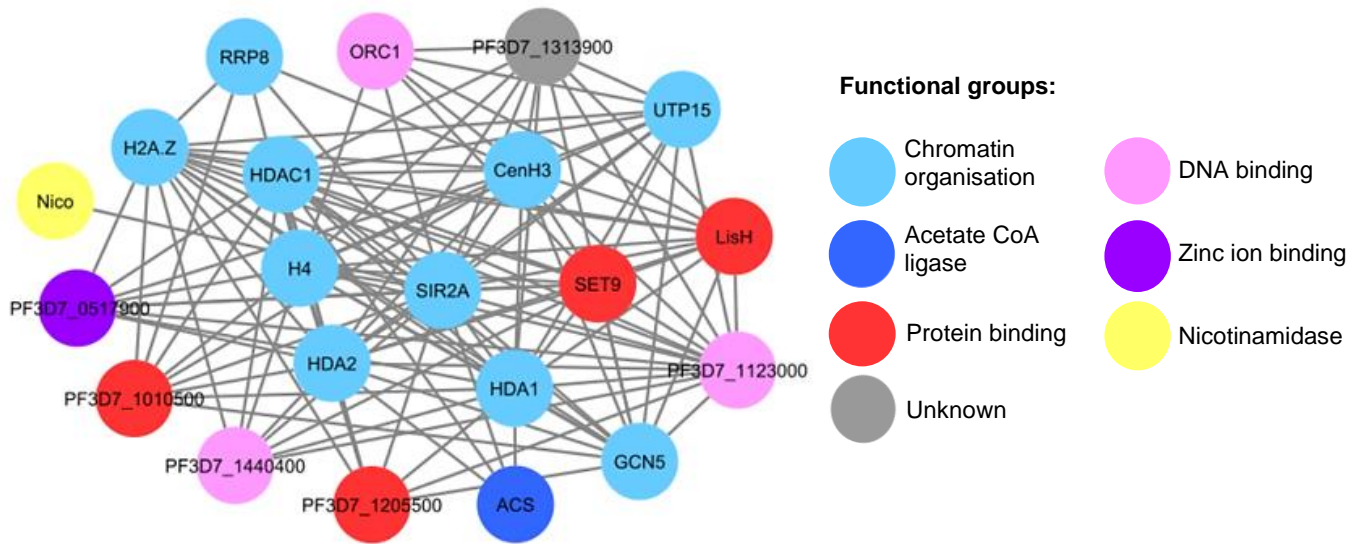


Figure 3.12: A schematic model of the protein-protein interaction complex of PF3D7_1328800 and its predicted functional partners. This network was manually curated and placed in a network based on evidence from STRING interactions. Functional predictions are based on gene ontology obtained from PlasmoDB. Shown in light pink blocks are proteins associated with DNA binding: Uncharacterised protein Myb-like DNA-binding domain, putative (PF3D7_1123000), origin recognition complex subunit 1 ORC1 (PF3D7_1203000), zinc finger protein, putative (PF3D7_1440400). Shown in light blue blocks are proteins associated with chromatin organisation: U3 small nucleolar RNA-associated protein 15 putative UTP15 (PF3D7_1352200), histone deacetylase 2 HDA2 (PF3D7_1008000), histone H4 (PF3D7_1105000), histone H3 CenH3 (PF3D7_1333700), histone deacetylase putative HDA1 (PF3D7_1472200), histone H2A.Z (PF3D7_0320900), histone deacetylase 1 HDAC1 (PF3D7_0925700), transcriptional regulatory protein SIR2A (PF3D7_1328800), histone acetyltransferase GCN5 (PF3D7_0823300), ribosomal RNA-processing protein 8, putative RRP8 (PF3D7_0925200). Shown in grey blocks are proteins with unknown function: ankyrin-repeat protein, putative (PF3D7_1313900). Shown in yellow blocks are proteins associated with Nicotinamidase activity: nicotinamidase putative (PF3D7_0320500). Shown in red blocks are proteins associated with protein binding: ankyrin-repeat protein, putative (PF3D7_1010500), LisH domain-containing protein, putative (PF3D7_1303400), SET domain protein, putative (PF3D7_0508100). Shown in dark blue are proteins associated with acetate CoA ligase activity: acetyl-CoA synthetase, putative ACS (PF3D7_0627800). Solid lines represent data from STRING (<https://string-db.org/>).

The STRING network predicted several functional partners of SIR2A (PF3D7_1328800) including histone H4, HDAC1 and histone H3, providing further evidence into SIR2A's role in chromatin organisation and deacetylase activity (Figure 3.12). SIR2A was also identified in zinc ion binding, as expected, as previous studies have indicated that zinc binding is important for the structural integrity and deacetylase activity of SIR2A [100]. SIR2A was shown to interact with histone acetyltransferase, GCN5 which is a component of the bromodomain epigenetic reader complex [133]. This may lead to further associations of SIR2A and epigenetic reader complexes.

SIR2A is known to catalyse NAD⁺-dependent hydrolysis of medium and long chain fatty acyl groups from lysine residues [132]. Acetyl-co-enzyme A ligase, identified in the STRING network, is an enzyme central to metabolism in prokaryotes and eukaryotes. Previous data suggests the function of NAD⁺-dependent SIR2A in antigenic variation may be accomplished by adding or removing acetylation marks to acetyl-co-enzyme A ligase to regulate its function [132]. This may lead to the possibility of SIR2A as a possible drug target.

3.2. *In vitro* cultivation of *P. falciparum* parasites

Asexual *P. falciparum* parasites were cultivated *in vitro* for the isolation of genomic DNA and for use in transfection (Figure 3.13). Ring-stage parasites are visible around 10-16 hpi. Early trophozoites start the formation of the hemozoin crystal as the parasite increases in size by late trophozoite stage. Schizont stage parasites consist of multiple nuclei where these burst releasing merozoites starting the cycle again.


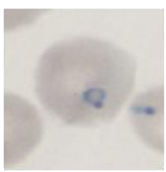
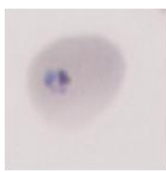
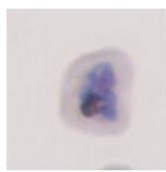
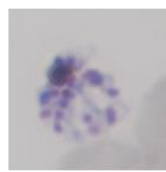
Stage	Merozoite	Ring	Early trophozoite	Late trophozoite	Schizont
Morphology					
hpi	0	10-13	26-29	34-37	42-46

Figure 3.13: Stage morphology of *P. falciparum* NF54 parasites. Merozoite, ring, early trophozoite, late trophozoite and schizont stages are shown with their respective hours post invasion (hpi). The parasites were visualised using Giemsa stain at 1000x magnification.

3.3. PCR amplification

The gene-specific regions of both PF3D7_0711400 (*sap18*) and PF3D7_1328800 (*sir2A*) genes were amplified for cloning into the SLI plasmids. Both the N- and C-terminal gene fragments of each gene were amplified between 220-1000 bp (Figure 2.1). Genomic DNA was isolated from NF54 *P. falciparum* parasites (>80 % trophozoite stage, >5 % parasitaemia) and used as a template for fragment amplification. Genomic DNA was extracted from *P. falciparum* parasites at a concentration of 58 ng/ μ L using a Nanodrop spectrophotometer. The purity of the genomic DNA extracted was spectrophotometrically determined with A_{260}/A_{280} of 1.99 and A_{260}/A_{230} of 1.96.

Single bands corresponding to the expected N- or C-terminal gene fragments were obtained for all four gene fragments: PF3D7_0711400_N-terminal (~450 bp), PF3D7_0711400_C-terminal (~960 bp), PF3D7_1328800_N-terminal (~220 bp) and PF3D7_1328800_C-terminal (~820 bp) (Figure 3.14 A and B). Conditions were optimised based on denaturation, annealing temperatures and duration times to achieve maximal DNA template quantity. PF3D7_0711400 N- and C-terminal gene fragments were best amplified at 55 °C as expected due to the melting temperature of these primer combinations. PF3D7_1328800 N- and C-terminal gene fragments were amplified at 56 °C and 58 °C, respectively. PCR DNA bands of the correct size for each gene fragment were subsequently purified, yielding 138-321 ng/ μ L of DNA with A_{260}/A_{280} and A_{260}/A_{230} ratios ranging from 1.8-2.4.

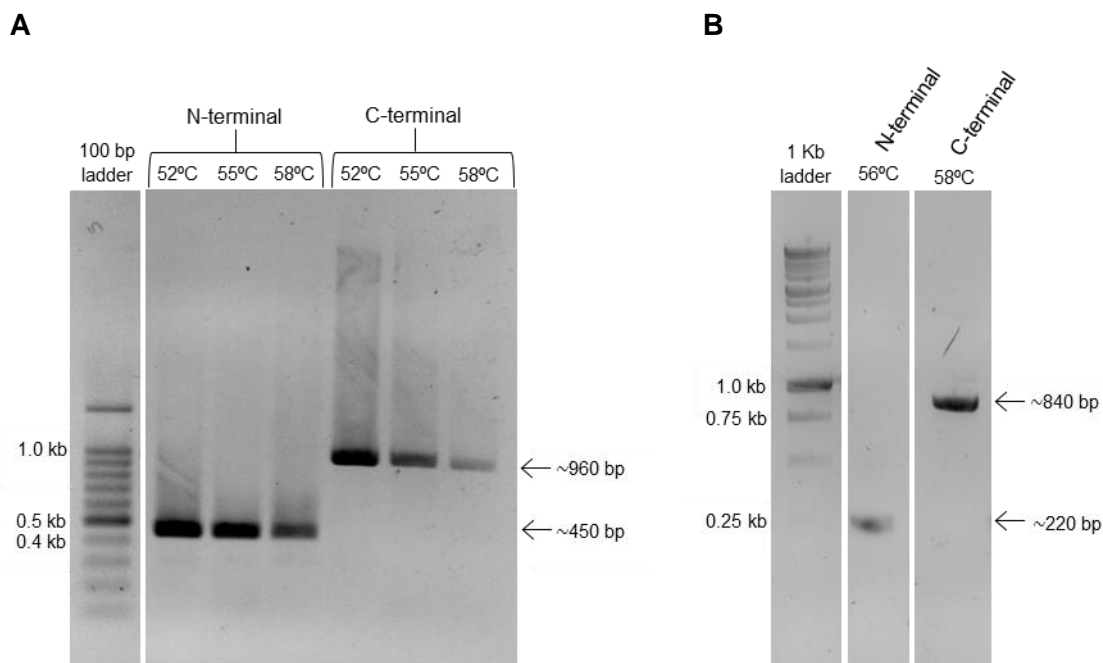


Figure 3.14: PCR amplification of gene specific homology regions. (A) Homology region amplification of N- and C- terminal PF3D7_0711400 at 52 °C, 55 °C and 58 °C using primers shown in Table 2.1. **(B)** Homology region amplification of N- and C- terminal PF3D7_1328800 at 56 °C and 58 °C using primers shown in Table 2.1. The expected amplicon band sizes were obtained: 451 bp (PF3D7_0711400, N-terminal), 961 bp (PF3D7_0711400, C-terminal), 217 bp (PF3D7_1328800, C-terminal), 841 bp (PF3D7_1328800, C-terminal). DNA fragments and 1 kb or 100 bp DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light. For complete gel picture of Figure 3.14 A and B see supplementary information Figure S1 A and B.

3.4. Cloning *sap18* and *sir2A* gene fragments into pSLI plasmids

For the first step of the cloning strategy, pSLI-TGD and pSLI-Sandwich plasmids were isolated from *E. coli* DH5 α cells and purified for downstream ligation reactions. Uncut purified plasmids pSLI-TGD (7260 bp) and pSLI-Sandwich (9050 bp) are represented as circular and supercoiled as shown by the two band sizes in each lane (Figure 3.15 A). pSLI-TGD and pSLI-Sandwich yielded concentrations of 560 ng/ μ L and 592 ng/ μ L, respectively with A_{260}/A_{280} and A_{260}/A_{230} ratios ranging from 1.9-2.4. Uncut plasmids were digested with either *NotI* and *MluI* for pSLI-TGD or *NotI* and *AvrII* for pSLI-Sandwich to isolate the plasmid backbone (Figure 3.15 B). DNA bands correspond to linearised base pair sizes of \sim 6760 bp for SLI-TGD and \sim 8220 bp for SLI-Sandwich plasmid backbones. These bands were gel extracted and purified to produce concentrations of 73 ng/ μ L for SLI-TGD and 76.5 ng/ μ L for SLI-Sandwich. The A_{260}/A_{280} ratios were within expected range of 1.9 to 2, however A_{260}/A_{230} ratios were significantly lower possibly due to the contamination of EDTA, guanidine salts, and oligosaccharides.

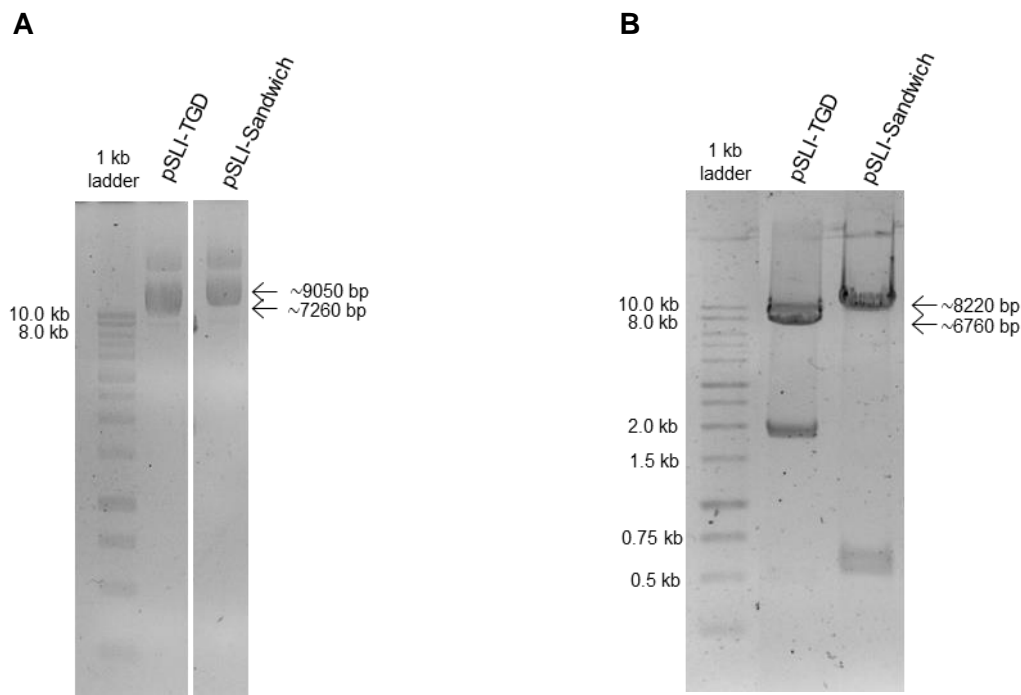


Figure 3.15: Restriction enzyme digestion of pSLI-TGD and pSLI-Sandwich plasmids. (A) Uncut SLI-TGD and SLI-Sandwich plasmids isolated from *E.coli* DH5 α to produce 9050 bp for SLI-Sandwich and a 7260 bp for SLI-TGD. **(B)** Restriction enzyme digestion of the pSLI-TGD plasmid with *NotI* and *MluI* to produce a 6760 bp plasmid backbone. pSLI-Sandwich plasmid digested with *NotI* and *AvrII* to produce a 8220 bp backbone. Plasmids were visualised with 1 % (w/v) agarose/TAE gel with ethidium bromide under UV light. For complete gel picture of Figure 3.15 A see supplementary information Figure S2.

The PF3D7_0711400 and PF3D7_1328800 N-terminal gene fragments were digested with *NotI* and *MluI*, while C-terminal gene fragments were digested with *NotI* and *AvrII*. These digested PCR fragments were subsequently purified to yield concentrations ranging from 98-195 ng/ μ L with A_{260}/A_{280} and A_{260}/A_{230} ratios ranging from 1.9-2.6 (Table 3.1).

Table 3.1: Concentration and purity of digested and purified PCR products

Digested and purified PCR products	Concentration (ng/ μ L)	A_{260}/A_{280}	A_{260}/A_{230}
PF3D7_0711400_N-terminal	121	1.91	1.87
PF3D7_0711400_C-terminal	195	1.88	2.56
PF3D7_1328800_N-terminal	120	1.95	2.14
PF3D7_1328800_C-terminal	98	1.9	2.28

Digested *sap18* PCR products were subsequently ligated to their corresponding digested and purified pSLI-plasmid backbones. After transformation into *E. coli* DH5 α cells, bacterial colonies were screened for the presence of a recombinant plasmid by performing colony screening PCR. Figure 3.16 A shows that pSLI-TGD-*Sap18* and pSLI-SW-*Sap18* recombinant plasmids were confirmed to contain the correct size insert at ~450 bp and ~ 1100 bp, respectively.

Colonies shown to contain the correct insert were selected, inoculated in LB-broth overnight and isolated and purified for restriction enzyme digestion mapping. This resulted in cutting of the gene fragment in each instance, confirming the identity of the plasmids, but also allowing subsequent purification of these plasmids for sequencing. Both pSLI-TGD-*Sap18* and pSLI-SW-*Sap18* were shown to have the correct size insert at ~440 bp and ~ 950 bp, respectively (Figure 3.16 B). In order to determine the correct construction of the above plasmids, the gene insert regions and flanking plasmid backbone, of the recombinant plasmids were sequenced. Chromatogram data shown below the consensus sequence represents the accuracy of the sequencing data (Figure 3.16 C). Combining both forward and reverse sequencing results we were able to correctly identify the restriction sites, N- and C-terminal *sap18* gene fragments as well as GFP regions visible (Figure 3.16 C). These regions were also in-frame to ensure that GFP will be successfully expressed once integrated into the genome.

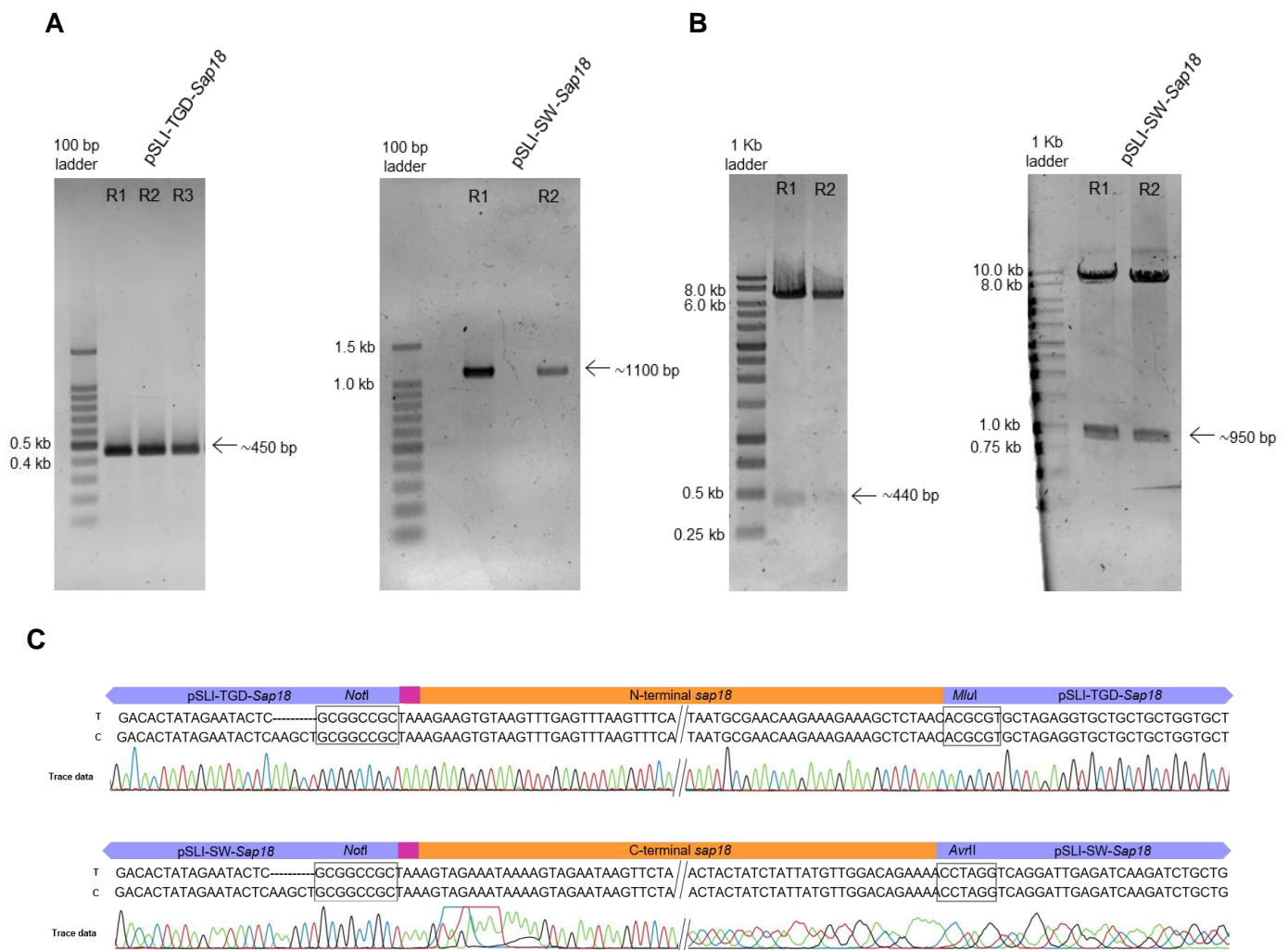


Figure 3.16: Validation of pSLI-TGD-*Sap18* and pSLI-SW-*Sap18* recombinant plasmids. (A) Colony screening PCR of *sap18* N- and C- terminal regions cloned into their respective plasmids. R indicates bacterial clone replicates. The expected region band size was obtained 436 bp (pSLI-TGD-*Sap18*) and 1097 bp (pSLI-SW-*Sap18*). (B) Restriction enzyme mapping of *sap18* N- and C- terminal regions cloned into their respective plasmids using *NotI* and *MluI*/*AvrII* restriction enzymes. The expected region band size was obtained 436 bp (pSLI-TGD-*Sap18*) and 946 bp (pSLI-SW-*Sap18*). (C) Sequence alignments confirming the identity of the *sap18* gene inserts in the respective plasmids. DNA fragments in schemes not drawn to scale. A representative forward or reverse chromatogram (trace data) is shown below the consensus sequence. T: template sequence. C: consensus sequence. DNA fragments and 1 kb or 100 bp DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light.

Digested *sir2A* PCR products were subsequently ligated to their corresponding digested and purified pSLI-plasmid backbones. After transformation into *E. coli* DH5 α cells, bacterial colonies were screened for the presence of a recombinant plasmid by performing colony screening PCR. Figure 3.17 A shows that pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* recombinant plasmids were confirmed to contain the correct size insert at ~380 bp and ~ 980 bp, respectively.

Colonies shown to contain the correct insert were selected, inoculated in LB-broth overnight and isolated for restriction enzyme digestion mapping. Both pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* were shown to have the correct size insert at ~200 bp and ~ 830 bp, respectively (Figure 3.17 B). To further confirm the identity of the recombinant plasmids, Sanger sequencing was performed (Figure 3.17 C). A consensus sequence allowed us to correctly identify the restriction sites, N- and C-terminal *sir2A* gene fragments and GFP regions visible (Figure 3.17 C).

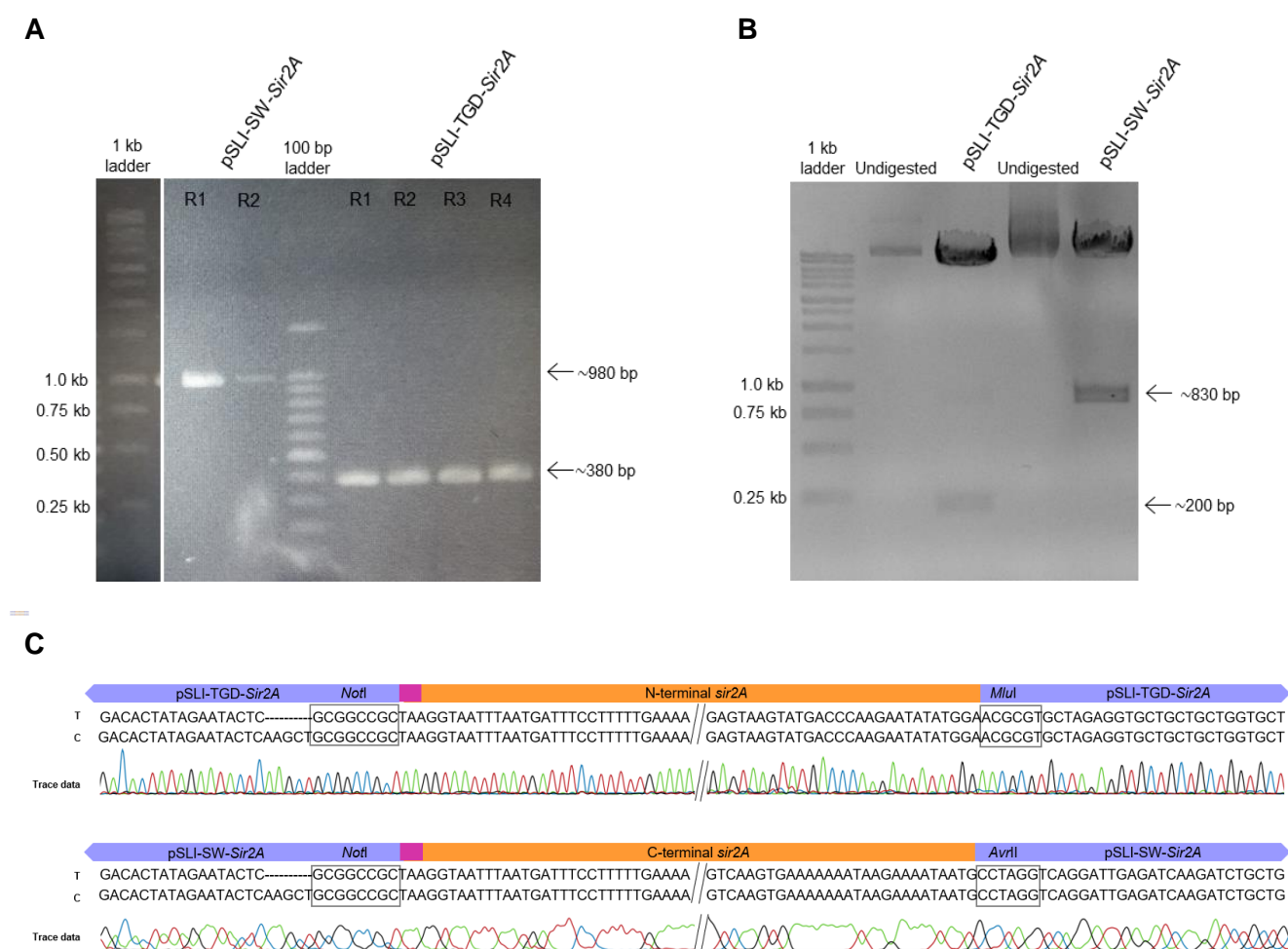


Figure 3.17: Validation of pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* recombinant plasmids. (A) Colony screening PCR of *sir2A* N- and C- terminal regions cloned into their respective plasmids. R indicates bacterial clone replicates. The expected region band size was obtained 379 bp (pSLI-TGD-*Sir2A*) and 977 bp (pSLI-SW-*Sir2A*). **(B)** Restriction enzyme mapping of *sir2A* N- and C- terminal regions cloned into their respective plasmids using *NotI* and *MluI/AvrII* restriction enzymes. The expected region band size was obtained 202 bp (pSLI-TGD-*Sir2A*) and 826 bp (pSLI-SW-*Sir2A*). **(C)** Sequence alignments confirming the identity of the *sir2A* gene inserts. DNA fragments in schemes not drawn to scale. A representative forward or reverse chromatogram (trace data) is shown below the consensus sequence. T: template sequence. C: consensus sequence. DNA fragments and 1 kb or 100 bp DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light. For complete gel picture of Figure 3.17 A see supplementary information Figure S3.

Thus, successful cloning and construction of pSLI-TGD-*Sap18*, pSLI-SW-*Sap18*, pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* plasmids was therefore achieved for transfection.

3.5. Generation of transgenic lines

3.5.1. Parasite transfection and selection for episomal uptake

The SLI system was used to produce NF54 transgenic parasite lines containing non-functional or modified *sap18* and *sir2A* loci. Parasites were monitored over the course of drug selection and recovery. Electroporation of pSLI-TGD-*Sap18* and pSLI-SW-*Sap18* into NF54 wild type parasites resulted in a sudden, but expected, drop in parasitaemia, from 5 % on day 0 to ~3 % on day 1 due to erythrocyte lysis (Figure 3.18 A). Subsequent WR99210 selection for 10 days resulted in a further decrease in parasitaemia to <0.1 %. Following selection, parasites increased to a parasitaemia of ~3.5 % indicating correct episomal uptake of both pSLI-TGD-*Sap18* and pSLI-SW-*Sap18* into wild type NF54 parasites.

Electroporation of pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* into NF54 wild type parasites resulted in a smaller drop in parasitaemia, from 5 % on day 0 to only ~4 % on day 1, possibly due to fewer erythrocytes lysing after electroporation (Figure 3.18 B). Subsequent selection with WR99210 for 10 days also resulted in a further decrease in parasitaemia to <0.1 %. Following selection, parasites increased to a parasitaemia around 2 % indicating correct episomal uptake of pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* into wild type NF54 parasites.

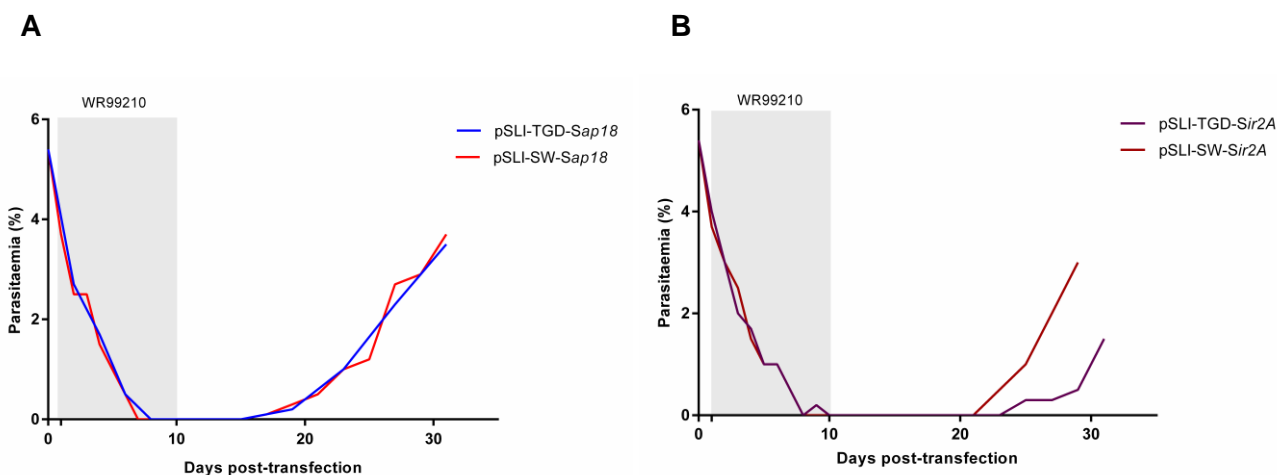


Figure 3.18: Selection and recovery of recombinant *P. falciparum* parasites after transfection with pSLI-TGD-*Sap18*, pSLI-SW-*Sap18*, pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A*. (A) Wild type *P. falciparum* NF54 parasites (>5 % ring-stage parasites) were transfected with pSLI-TGD-*Sap18* and pSLI-SW-*Sap18* (day 0). Parasites were selected with WR99210 for 10 days to select for episomal uptake and parasites were allowed to recover above 2 % parasitaemia. (B) Wild type *P. falciparum* NF54 parasites (>5 % ring-stage parasites) were transfected with pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* (day 0). Parasites were selected with WR99210 for 10 days to select for episomal uptake and parasites were allowed to recover above 2 % parasitaemia. Parasite viability was evaluated by Giemsa-stained blood smears at 1000x magnification.

To test whether transgenic parasites contain the plasmid of interest episomally, parasite lines were expanded to a 5 % parasitaemia after they had recovered from WR99210 drug selection. Plasmid DNA was isolated from NF54-epi(pSLI-TGD-*Sap18*) and NF54-epi(pSLI-SW-*Sap18*) parasites, and the regions flanking the gene of interest on the plasmid backbone were PCR amplified. This confirmed the successful uptake of the correct sized recombinant plasmids within the parasite. A single ~610 bp band was observed in NF54-epi(pSLI-TGD-*Sap18*) parasites, confirming the presence of pSLI-TGD-*Sap18* (Figure 3.19 B). A single ~1100 bp band was observed in NF54-epi(pSLI-SW-*Sap18*) parasites, confirming the presence of pSLI-SW-*Sap18* (Figure 3.19 D).

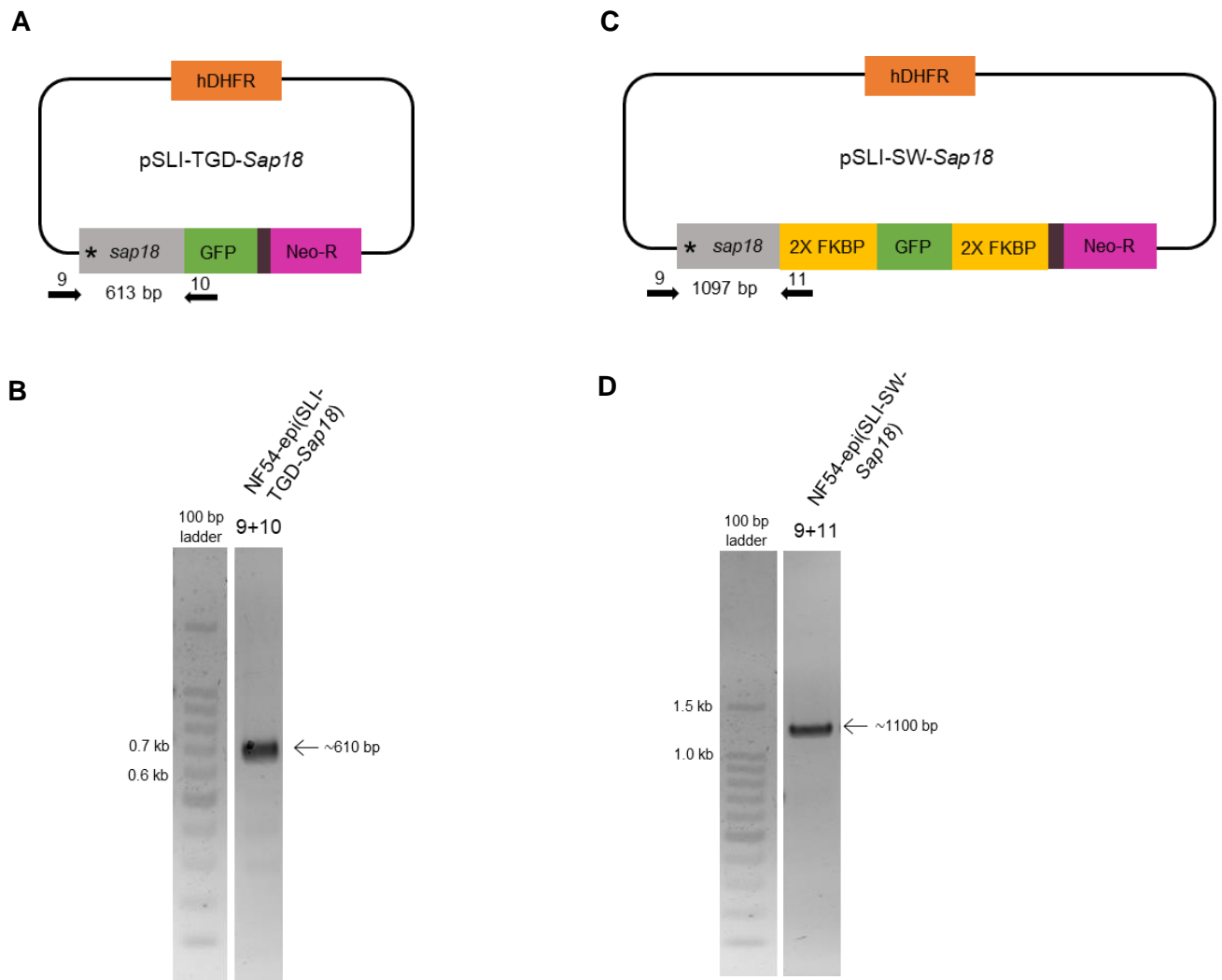


Figure 3.19: PCR analysis of pSLI-TGD-*Sap18* and pSLI-SW-*Sap18* for episomal uptake into NF54 *P. falciparum* parasites. (A) Schematic representation of pSLI-TGD-*Sap18* showing backbone primer locations and expected size. (B) PCR amplification of SLI-TGD-*Sap18* plasmid backbone (613 bp) showing episomal presence in NF54 parasites. (C) Schematic representation of pSLI-SW-*Sap18* showing backbone primer locations and expected size. (D) PCR amplification of pSLI-SW-*Sap18* plasmid backbone (1097 bp) showing episomal presence in NF54 parasites. DNA fragments and 100 bp or 1 kb DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light. For complete gel picture of Figure 3.20 B and D see supplementary information Figure S3.

Similarly, for *sir2A*, we confirmed episomal uptake of recombinant plasmids. Plasmid DNA was isolated from NF54-epi(pSLI-TGD-*Sir2A*) and NF54-epi(SLI-SW-*Sir2A*) parasites and the regions flanking the

gene of interest on the plasmid backbone were PCR amplified. pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* plasmids were present episomally as shown by the band sizes at ~380 and ~980 bp, respectively (Figure 3.20 B and D).

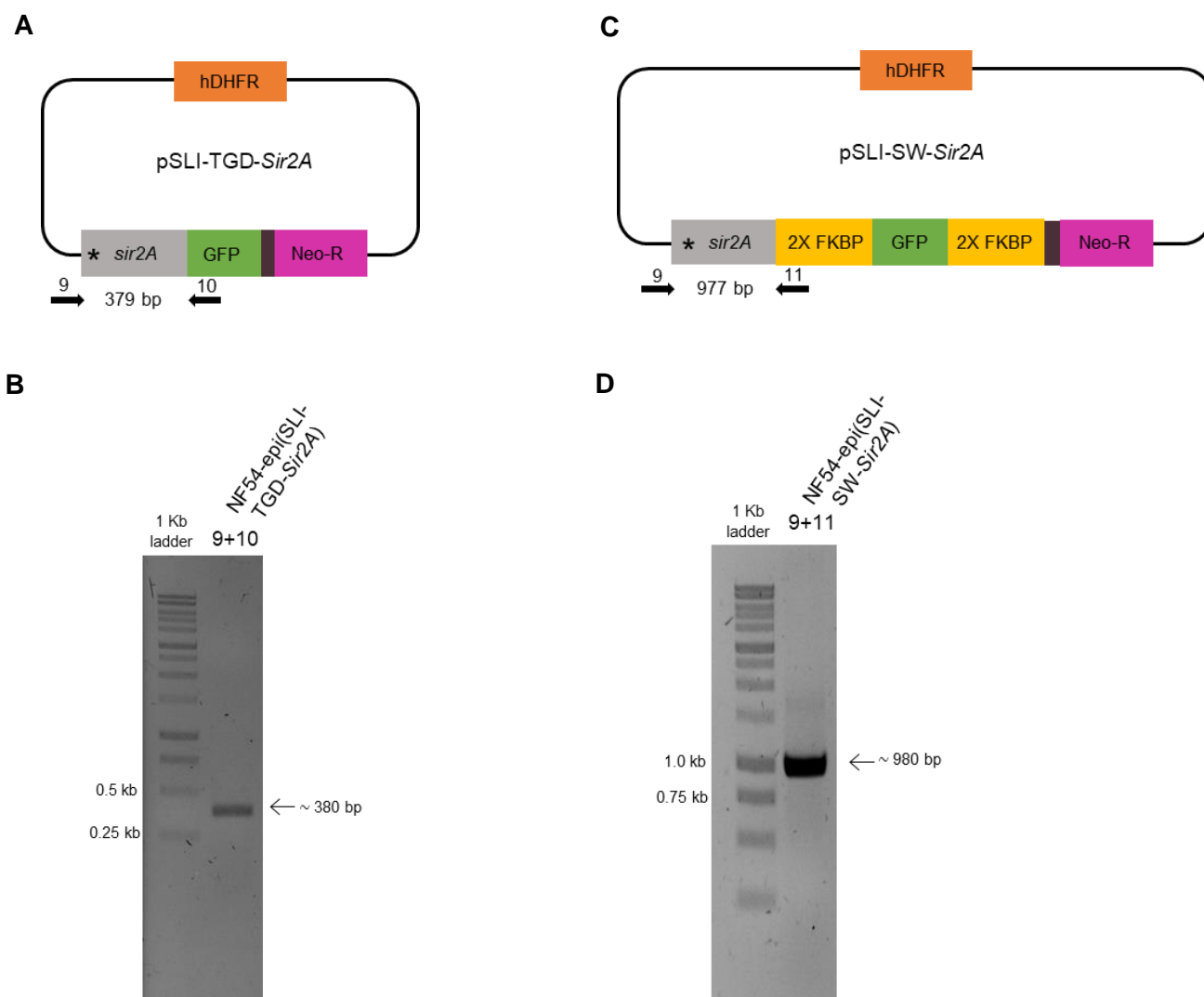


Figure 3.20: PCR analysis of pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* for episomal uptake into NF54 *P. falciparum* parasites. (A) Schematic representation of pSLI-TGD-*Sir2A* showing backbone primer locations and expected size. (B) PCR amplification of pSLI-TGD-*Sir2A* plasmid backbone (379 bp) showing episomal presence in NF54 parasites. (C) Schematic representation of pSLI-SW-*Sir2A* showing backbone primer locations and expected size. (D) PCR amplification of pSLI-SW-*Sir2A* plasmid backbone (977 bp) showing episomal presence in NF54 parasites. DNA fragments and 100 bp or 1 Kb DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light.

Positive episomal uptake of pSLI-TGD-*Sap18*, pSLI-SW-*Sap18*, pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* plasmids was achieved. Subsequently NF54-epi(SLI-TGD-*Sap18*), NF54-epi(SLI-SW-*Sap18*), NF54-epi(SLI-TGD-*Sir2A*) and NF54-epi(SLI-SW-*Sir2A*) parasites could be selected for integration using drug pressure.

3.5.2. Selection for integration

NF54-epi(SLI-TGD-*Sap18*) and NF54-epi(SLI-SW-*Sap18*) parasites were selected for integration for 24 days, where parasites started to die around 5 days after initial G418 drug pressure due to lack of the neomycin resistance marker (Figure 3.21 A). Parasites that had possibly undergone integration, which would contain the neomycin resistance gene, were able to survive the drug pressure and these parasites increased in parasitaemia around 20 days after initial G418 drug selection (Figure 3.21 A). Similarly for *sir2A* integration, NF54-epi(SLI-TGD-*Sir2A*) and NF54-epi(SLI-SW-*Sir2A*) parasites were selected for integration with G418 for only 14 days (Figure 3.21 B) as previous attempts to select for integration for longer periods in these lines resulted in all parasites dying. Parasites that had possibly undergone integration to become neomycin resistant, were able to survive the drug pressure but these parasites increased in parasitaemia only around 50 days after initial G418 drug selection (Figure 3.21 B). Recovery of these parasite lines took longer as the parasites were more stressed, as indicated by a number of gametocytes present.

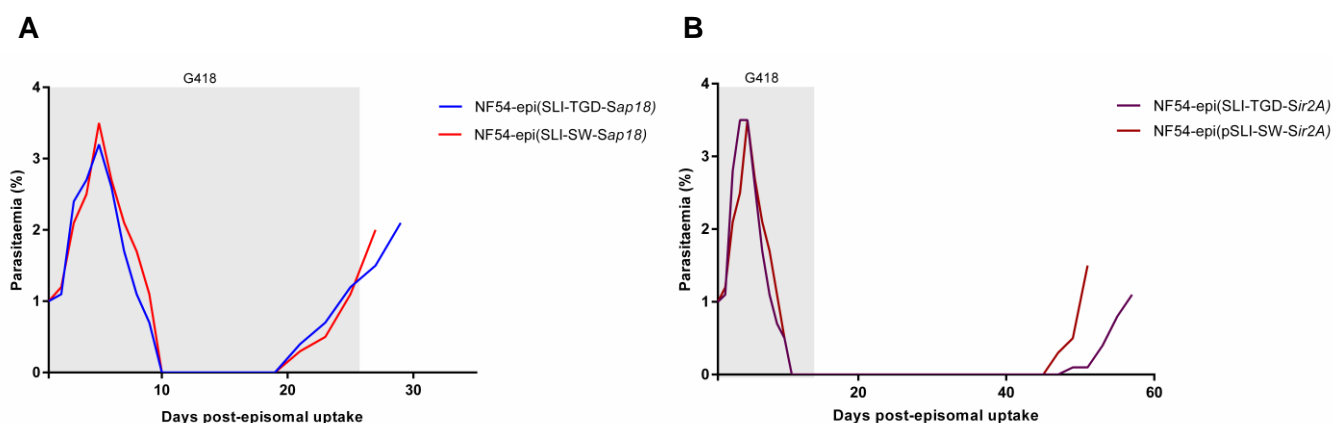


Figure 3.21: Selection recovery of recombinant *P. falciparum* parasites after episomal uptake of pSLI-TGD-*Sap18*, pSLI-SW-*Sap18*, pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A*. (A) NF54-epi(SLI-TGD-*Sap18*) and NF54-epi(SLI-SW-*Sap18*) parasites were selected with G418 for 24 days to select for integration. Parasites were allowed to recover above 2 % parasitaemia. (B) NF54-epi(SLI-TGD-*Sir2A*) and NF54-epi(SLI-SW-*Sir2A*) parasites were selected with G418 for 14 days to select for integration. Parasites were allowed to recover above 2 % parasitaemia. Data represent 1 biological repeat. Parasite viability was evaluated by Giemsa-stained blood smears at 1000x magnification.

Plasmids that had undergone homologous recombination into the parasite's genome were assessed for successful integration by PCR. Genomic DNA was isolated and the flanking 5' and 3' loci regions of both *sap18* and *sir2A* were PCR amplified. Figure 3.22 B indicates 5' (~800 bp) and 3' (~730 bp) integration of NF54- Δ *Sap18*-GFP parasites at the *sap18* locus. The absence of wild type locus parasites within the integrated line indicates a fully transgenic line, now known as NF54- Δ *Sap18*-GFP. Non-specific bands were present within the 5' region of the integrated line which was not to be expected and may require further optimization of PCR conditions.

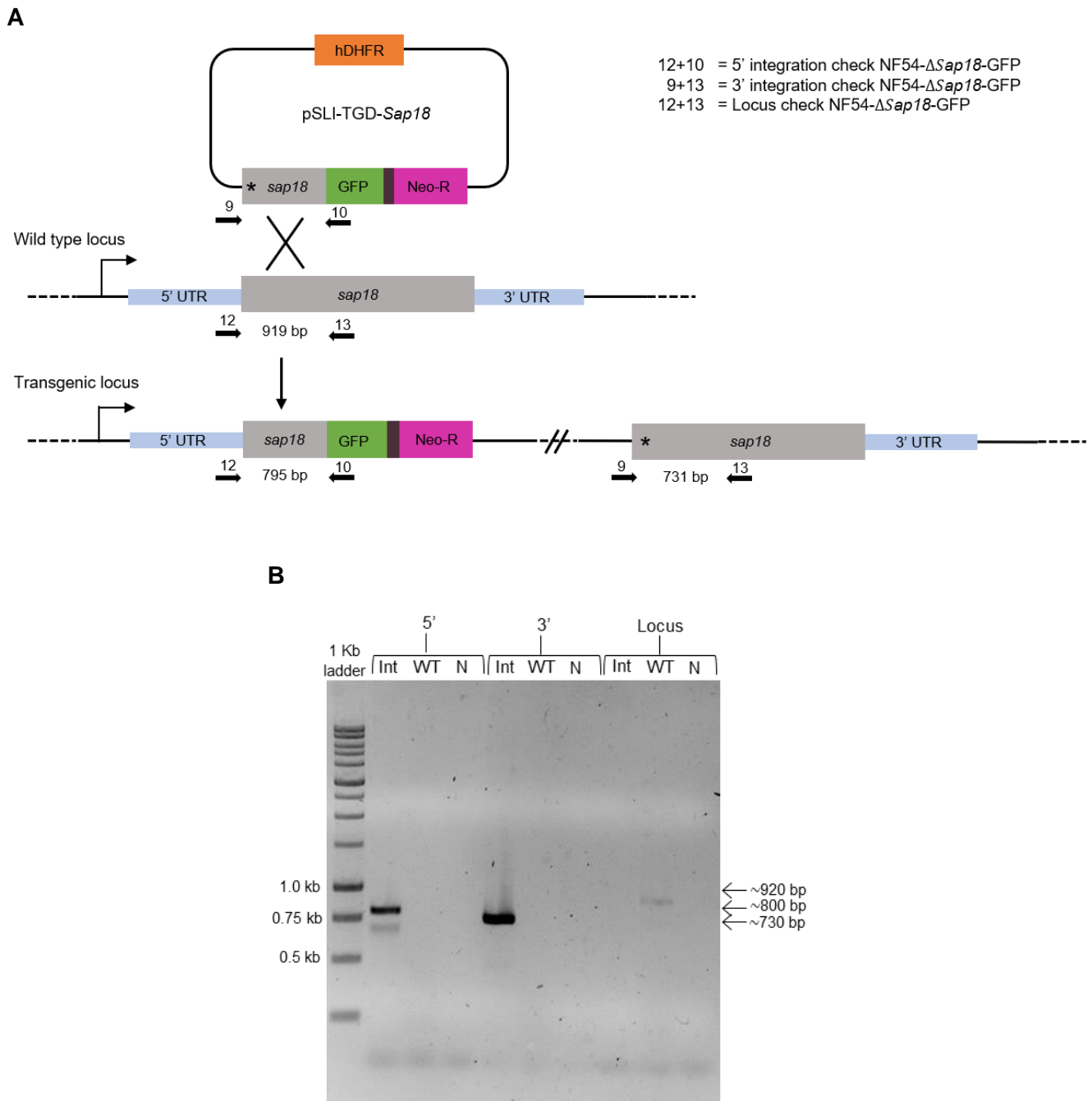


Figure 3.22: PCR analysis of pSLI-TGD-*Sap18* integration into the *sap18* locus of *P. falciparum* parasites. (A) Schematic representation of integration PCR for wild type and transgenic loci showing primer pairs and expected sizes for 5', 3' and loci integration checks. (B) NF54-Δ*Sap18*-GFP parasites were PCR screened for the presence of 5' and 3' integration. The expected DNA band sizes were obtained on the gel (5': 795 bp, 3': 731 bp). The absence of the original locus denotes no wild type parasites were present. NF54 wild type parasites (control) indicated the presence of the original locus (919 bp). PCR products and 1 kb DNA ladder were visualised with 1.5 % (w/v) agarose/TAE gel stained with ethidium bromide under UV light. Int=integrated parasite line, WT= NF54 wild type control, N= negative control.

NF54-Δ*Sir2A*-GFP parasites were assessed for their integration via PCR. After two biological repeats, no integration into the *sir2A* locus was observed in NF54-epi(pSLI-TGD-*Sir2A*) parasites. This was determined by the absence of both 5' or 3' integration within the *sir2A* locus (Figure 3.23 B). This was also confirmed by the presence of the wild type locus in NF54-epi(pSLI-TGD-*Sir2A*) parasites. Lack of integration may be due to the short truncated fragment of *sir2A* (~210 bp) for cloning into the TGD

plasmid which increases the time required to integration for these shorter targeting regions [125]. Additionally, the lack of integration may be due to the lengthy recovery time after selection observed in this parasite line (Figure 3.21 B). Thus, more time may be required to integrate this parasite line. PCR amplification of previous attempts of integration of the NF54-epi(pSLI-TGD-*Sir2A*) parasites are shown in supplementary information Figure S5.

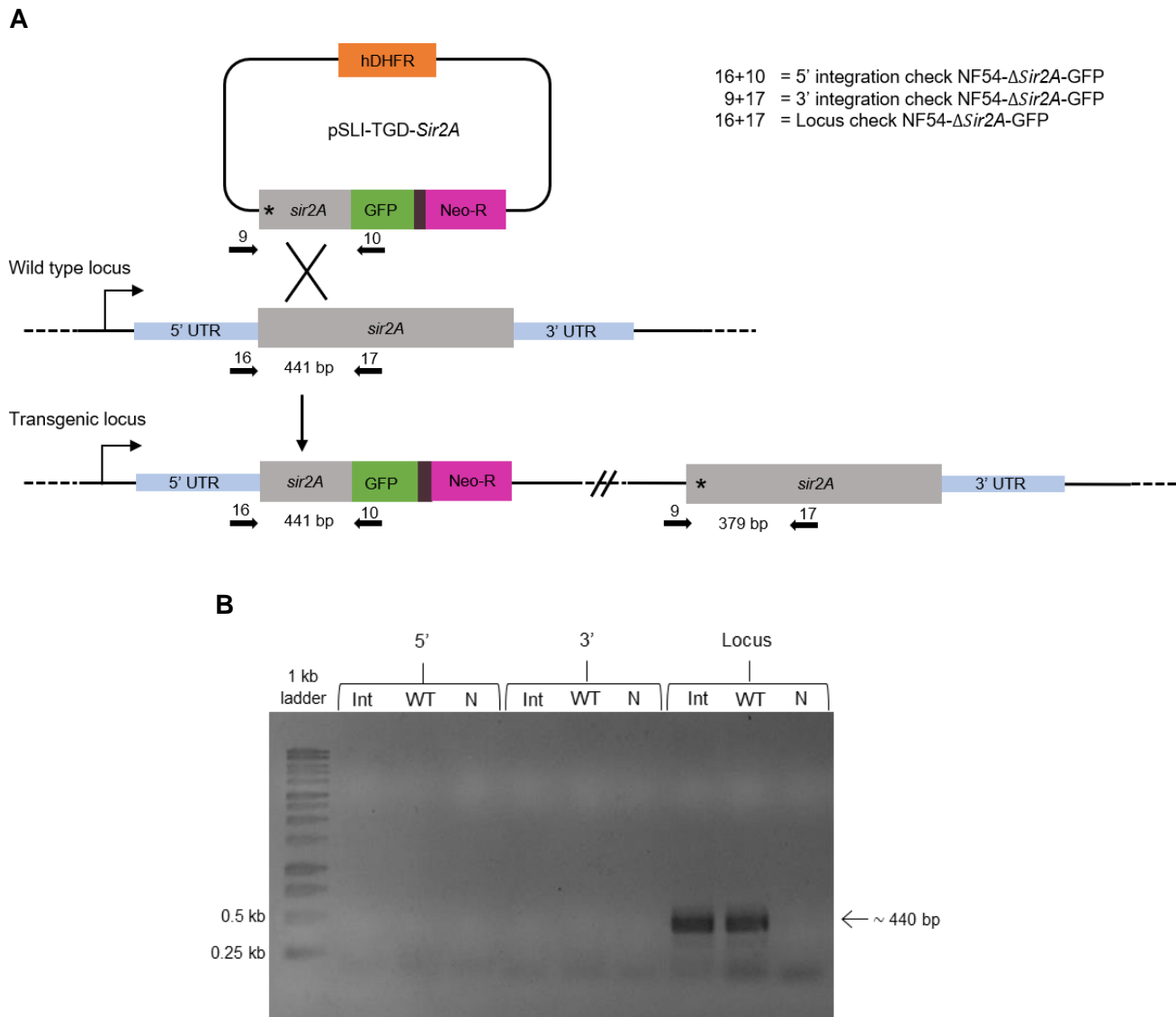


Figure 3.23: PCR analysis of pSLI-TGD-*Sir2A* integration into the *sir2A* locus of *P. falciparum* parasites. (A) Schematic representation of integration PCR for wild type and transgenic loci showing primer pairs and expected sizes for 5', 3' and loci integration checks. (B) NF54- Δ *Sir2A*-GFP parasites were PCR screened for the presence of 5' and 3' integration. The absence of 5' and 3' integration bands indicate no integration. The presence of the original locus denotes wild type parasites were present (441 bp). NF54 wild type parasites (control) indicated the presence of the original locus (441 bp). PCR products and 1 Kb DNA ladder were visualised with 1.5 % (w/v) agarose/TAE gel stained with ethidium bromide under UV light. Int=integrated parasite line, WT= NF54 wild type control, N= negative control.

NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites were assessed for their integration via PCR. NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites showed integration at both the 5' and 3' end of the *sap18* locus as indicated by the presence of DNA bands (Figure 3.24 B). The *sap18* locus was also amplified to identify any unintegrated/wild type parasites present, but no wild type parasites were amplified

indicating full integration of these lines. Non-specific bands were present within the 3' region of the integrated line as well as the wild type locus, which was not expected but may require further optimization of the PCR conditions.

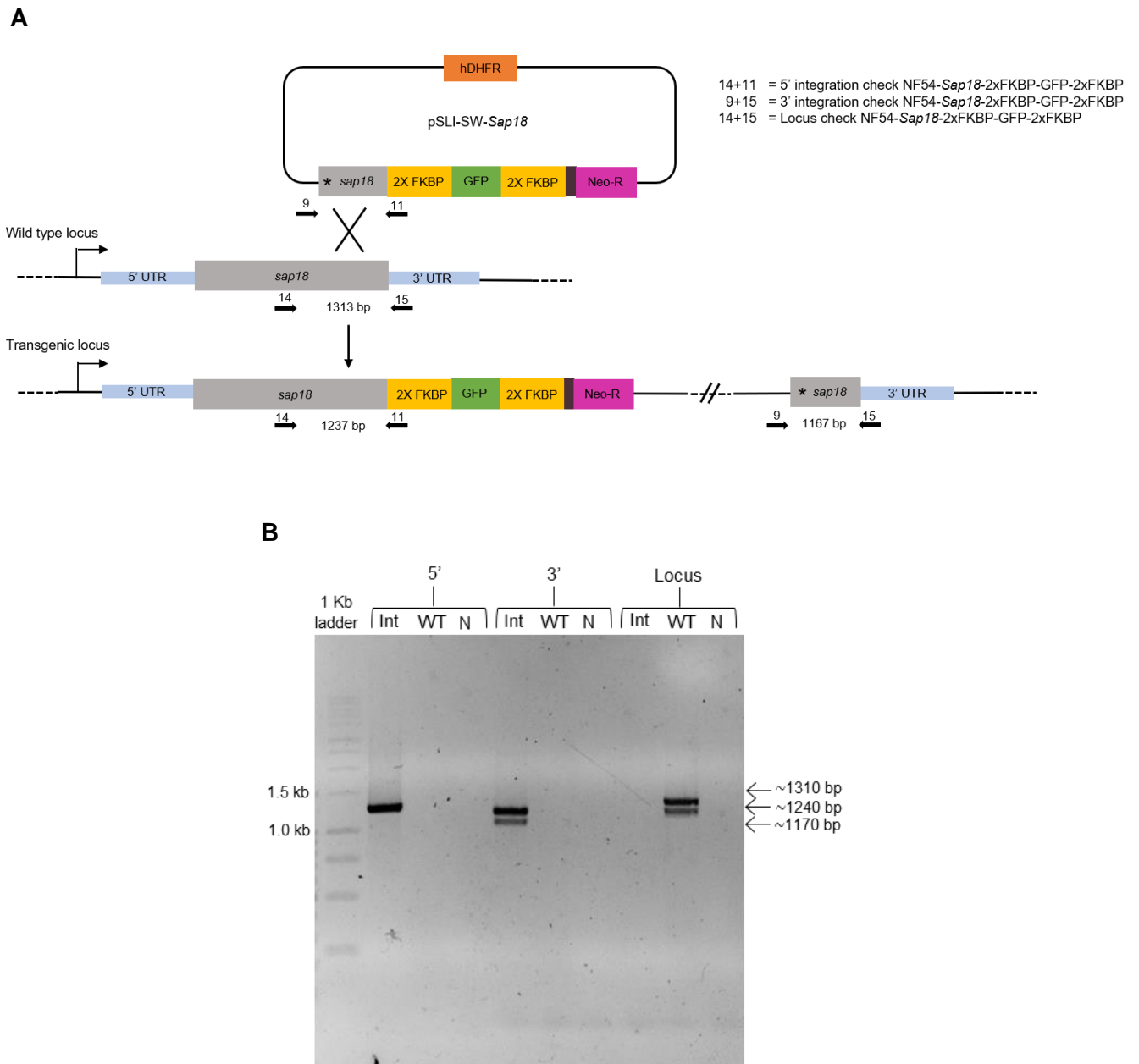
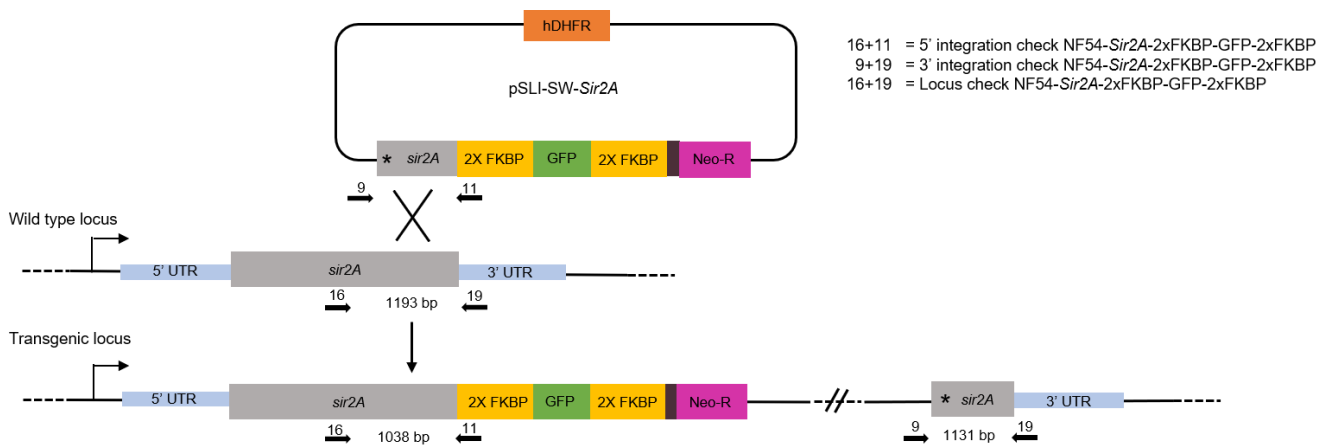


Figure 3.24: PCR analysis of pSLI-SW-*Sap18* into the *sap18* locus of *P. falciparum* parasites. (A) Schematic representation of integration PCR for wild type and transgenic loci showing 5', 3' and loci integration checks. (B) NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites were PCR screened for the presence of 5' and 3' integration. The expected DNA band sizes were obtained on the gel (5': 1237 bp, 3': 1167 bp). The absence of the original locus denotes no wild type parasites were present. NF54 wild type parasites (control) indicated the presence of the original locus (1313 bp). PCR products and 1 kb DNA ladder were visualised with 1.5 % (w/v) agarose/TAE gel stained with ethidium bromide under UV light. Int=integrated parasite line, WT= NF54 wild type control, N= negative control.

NF54-*Sir2A*-2xFKBP-GFP-2xFKBP parasites showed both 5' and 3' integration as well as the lack of wild type parasites at the original *sir2A* locus (Figure 3.25 B). However, manipulation of the *sir2A* locus

proved challenging as integration of NF54-*Sir2A*-2xFKBP-GFP-2xFKBP parasites was only successful after the third biological repeat (Supplementary Figure S6).

A



B

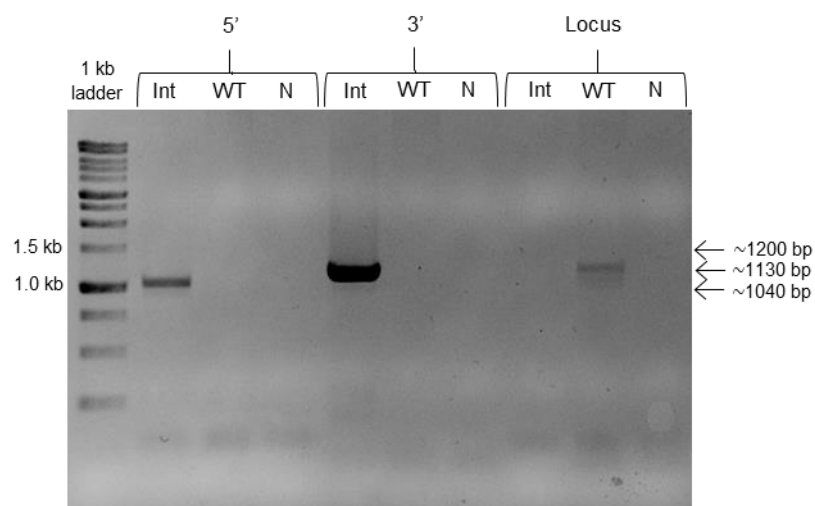


Figure 3.25: PCR analysis of pSLI-SW-*Sir2A* integration into the *sir2A* locus of *P. falciparum* parasites. (A) Schematic representation of integration PCR for wild type and transgenic loci showing 5', 3' and loci integration checks. **(B)** NF54-*Sir2A*-2xFKBP-GFP-2xFKBP parasites were PCR screened for the presence of 5' and 3' integration. The expected DNA band sizes were obtained on the gel (5': 1038 bp, 3': 1131 bp). The absence of the original locus denotes no wild type parasites were present. NF54 wild type parasites (control) indicated the presence of the original locus (1193 bp). PCR products and 1 kb DNA ladder were visualised with 1.5 % (w/v) agarose/TAE gel stained with ethidium bromide under UV light. Int=integrated parasite line, WT= NF54 wild type control, N= negative control.

Thus, we proceeded onto the analysis of NF54-*Sap18*-2xFKBP-GFP-2xFKBP and NF54- Δ *Sap18*-GFP parasite lines.

3.6. Validation of transgenic *sap18* lines

3.6.1. Confocal microscopy to assess GFP expression

For both NF54-*Sap18*-2xFKBP-GFP-2xFKBP and NF54- Δ *Sap18*-GFP parasite lines, SAP18 is now associated and tagged with GFP to allow for the localisation of the protein. Late trophozoite stage populations were selected for visualisation under a confocal microscope. This parasite stage was chosen based on transcriptomic data, where *sap18* is most highly expressed in trophozoites (27-35 hpi) (100-172 TPM) [134, 135], but the abundance of most parasite protein levels peak significantly later (median 11 h) than peak mRNA abundance levels [45], hence 38-46 hpi.

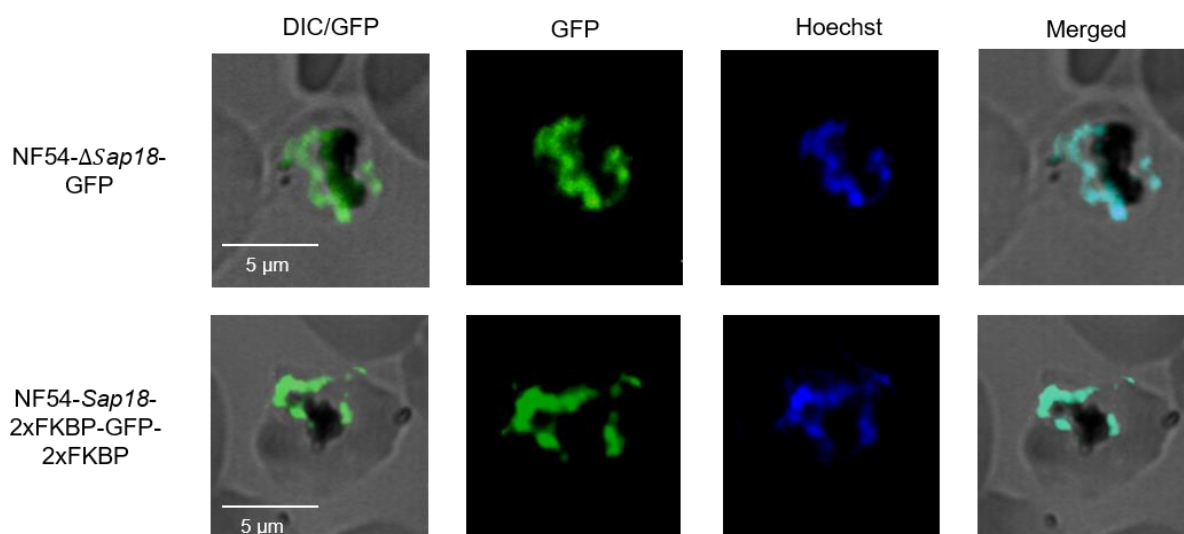


Figure 3.26: Localisation of SAP18 in NF54-*Sap18*-2xFKBP-GFP-2xFKBP and NF54- Δ *sap18*-GFP parasite lines via confocal microscopy. DIC, differential interference contrast; Hoechst, parasite nuclei. Scale bars, 5 μ m.

Figure 3.26 shows GFP-tagged expression of SAP18 in both NF54-*Sap18*-2xFKBP-GFP-2xFKBP and NF54- Δ *Sap18*-GFP parasite lines. This expression was colocalised to the DNA Hoechst stain found in the nucleus of the parasite. This was to be expected as SAP18 is thought to be a subunit of the HDAC1 complex, which has been previously localised to the nucleus [105]. NF54- Δ *Sap18*-GFP parasites express GFP in the same manner and location as that of NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites, indicating truncated SAP18 may still be functional within the nucleus of NF54- Δ *Sap18*-GFP parasites. However, the SAP18 domain is found at the C-terminus and therefore not present in NF54- Δ *Sap18*-GFP parasites, leading to the prediction of no functionality without the presence of the SAP18 domain.

To further validate NF54- Δ *Sap18*-GFP and NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasite lines, we performed western blot to validate GFP protein expression and the size of each cassette.

3.6.2. Western blot analysis to detect GFP protein expression

To confirm expression of SAP18 in late trophozoite/schizont stages, western blot was used to confirm expression of the protein. SAP18-GFP tagged protein (61 kDa) expressed in NF54- Δ Sap18-GFP was confirmed to express GFP by an anti-GFP antibody (Figure 3.27). Histone H3 core antibody was used as a loading control across the samples. Western blot analysis indicated no protein expression at the expected size as compared to NF54 wild type parasites (Figure 3.27), possibly due to the low transcript level of this gene controlling GFP production. There was also high background signal with the anti-GFP antibody as seen by the multiple bands indicating non-specific binding of the antibody.

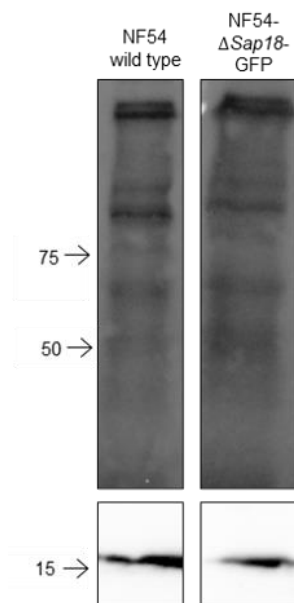


Figure 3.27: Western blot analysis of GFP expression of NF54- Δ Sap18-GFP parasites compared to NF54 wild type parasites probed with anti-GFP antibody. Expected protein size of ~61 kDa. Histone H3 was used as the control for equal protein loading. Precision Plus Protein Dual Colour ladder was used to compare protein size. For complete western blot of Figure 3.27 see in supplementary information Figure S7.

Previous reports have addressed the difficulties to detect GFP expression in pSLI-Sandwich lines [125] thus, we used the FKBP-12 antibody to detect the SAP18-GFP-FKBP-tagged protein (174 kDa) in NF54-Sap18-2xFKBP-GFP-2xFKBP parasites (Figure 3.28). SAP18-GFP-FKBP-tagged protein is shown at ~170 kDa, corresponding to the expected size of 174 kDa. This concludes that SAP18-GFP-FKBP-tagged protein was detected in NF54-Sap18-2xFKBP-GFP-2xFKBP parasites.

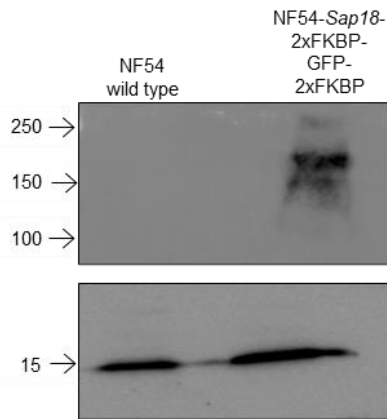


Figure 3.28: Western blot analysis of FKBP-12 expression of NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites compared to NF54 wild type parasites probed with anti-FKBP12 antibody. Expected protein size of ~170 kDa. Histone H3 was used as the control for equal protein loading. Precision Plus Protein Dual Colour ladder was used to compare protein size.

In conclusion, PCR data confirms that the plasmid integration of NF54- Δ *Sap18*-GFP parasites occurred at the correct locus as identified by the correct size of these cassettes. SAP18 may not be essential to asexual stage parasites as parasites recovered after successful integration. In addition, integration of NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites was successful as the correct size cassette was identified via both PCR and western blot. In addition, SAP18 was also localised to the nucleus. The latter could allow nuclear mislocalisation of SAP18 with the use of a knock sideways system.

3.7. Functional analysis of SAP18 protein with knock sideways

A conditional knock sideways experiment was performed on SAP18 to identify its function within the parasite. To perform knock sideways of the protein of interest, this system requires the additional transfection of the pLyn-FRB-*mCherry*-nmd3-BSD plasmid as well as drug pressure with blasticidin to select for episomal presence of this plasmid. A PCR was performed to test for the episomal presence of the plasmid. Figure 3.29 shows that the pLyn-FRB-*mCherry*-nmd3-BSD plasmid is present episomally in NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites, as seen by the expected band at ~650 bp.

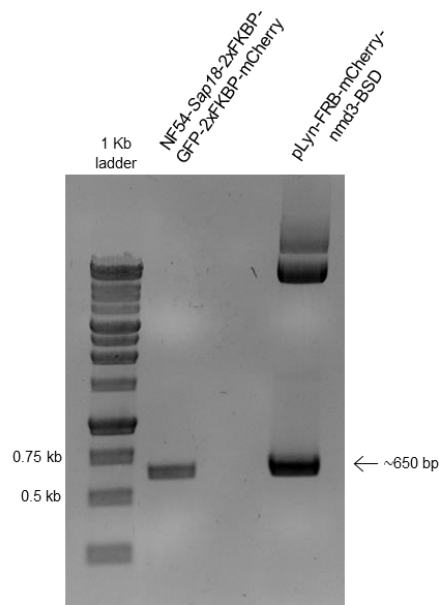


Figure 3.29: PCR analysis of pLyn-FRB-*mCherry*-nmd3-BSD plasmid for episomal uptake into the NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasite line. PCR amplification of the pLyn-FRB-*mCherry*-nmd3-BSD plasmid (~650 bp) showing episomal presence in NF54-*Sap18*-2xFKBP-GFP-2xFKBP parasites. The pLyn-FRB-*mCherry*-nmd3-BSD plasmid is shown as a positive control. DNA fragments and 1 kb DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light.

Subsequently, knock sideways was performed on NF54-*Sap18*-2xFKBP-2xFKBP/FRB-*mCherry* parasites to determine the functional importance of SAP18 to asexual parasites. Late schizont stage parasites were treated with rapalog for 24 h to induce SAP18 mislocalisation to the plasma membrane and parasites were visualised under the Zeiss Confocal Microscope with Airyscan detector (Figure 3.30).

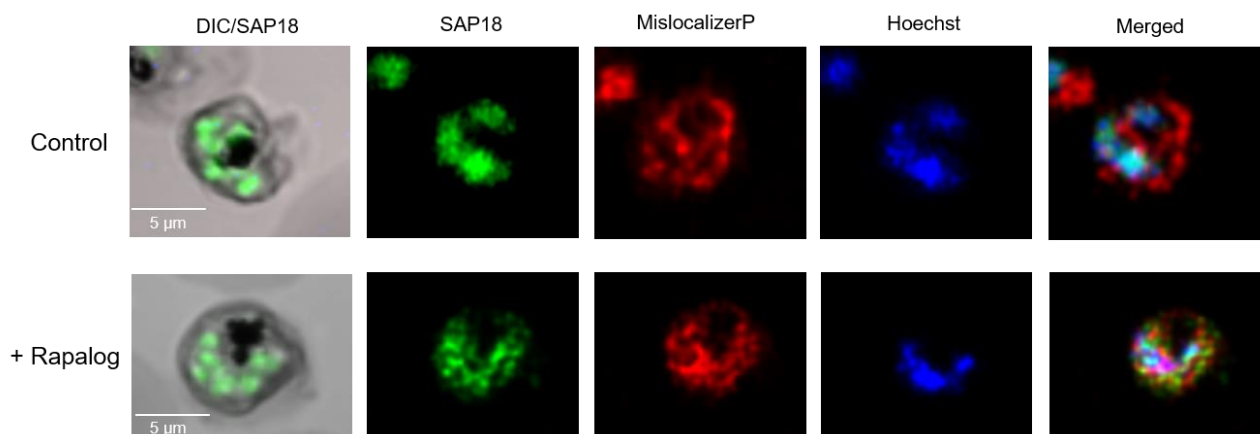


Figure 3.30: Knock sideways of NF54-*Sap18*-2xFKBP-2xFKBP/FRB-*mCherry* parasite lines assessed via confocal microscopy. Functional analysis of C-terminally GFP tagged *sap18*. DIC: differential interference contrast; Hoechst: parasite nuclei; merged: merged red, green and blue signal; MislocaliserP: Plasma membrane mislocaliser. Scale bars, 5 μ m.

Untreated (control) NF54-*Sap18*-2xFKBP-2xFKBP/FRB-*mCherry* parasites showed colocalisation of the SAP18-GFP signal to the Hoechst DNA stain in the nucleus (Figure 3.30). The mislocaliser-*mCherry* signal is present in the plasma membrane however, this cannot be distinguished from surrounding

cytoplasmic signal. A similar localisation profile of the *mCherry* signal was observed by Birnbaum *et al.* (2017) although, Figure 3.30 represents a less homogenous signal, possibly due to differences in the imaging microscopes. Rapalog treated parasites showed *mCherry* signals present in the plasma membrane as expected [125]. Partial mislocalisation was achieved as GFP signal was found to be colocalised to the plasma membrane as well as the nucleus (Figure 3.30). The movement of GFP-tagged SAP18 to the plasma membrane should render it non-functional from its site of action in the nucleus. However, partial mislocalisation was identified by the localisation of GFP still present in the nucleus. This concludes that not all the protein was mislocalised to the plasma membrane. Incomplete mislocalisation may be attributed to the copy number variation of the pLyn-FRB-*mCherry*-nmd3-BSD plasmid within each parasite influencing the strength of the mislocaliser and hence, the degree of mislocalisation. The stage of each parasite may differ slightly thereby affecting the expression of the mislocaliser as later stage replicating parasites produce a higher expression of mislocaliser.

Untreated and rapalog treated NF54-*Sap18*-2xFKBP-GFP-2xFKBP/FRB-*mCherry* parasites were also monitored for 96 h during asexual stage parasite proliferation. This allowed an evaluation of the functional importance of SAP18 to parasite survival upon mislocalisation from its site of action. No observable difference in parasitaemia was identified using the knock sideways system (Figure 3.31), indicating no growth difference between rapalog treated and untreated parasites. This may possibly be due to the incomplete mislocalisation of SAP18 as some proteins may be refractory to mislocalisation. Therefore, it was decided to constitutively disrupt *sap18* using the SLI-TGD plasmid and observe its effects.

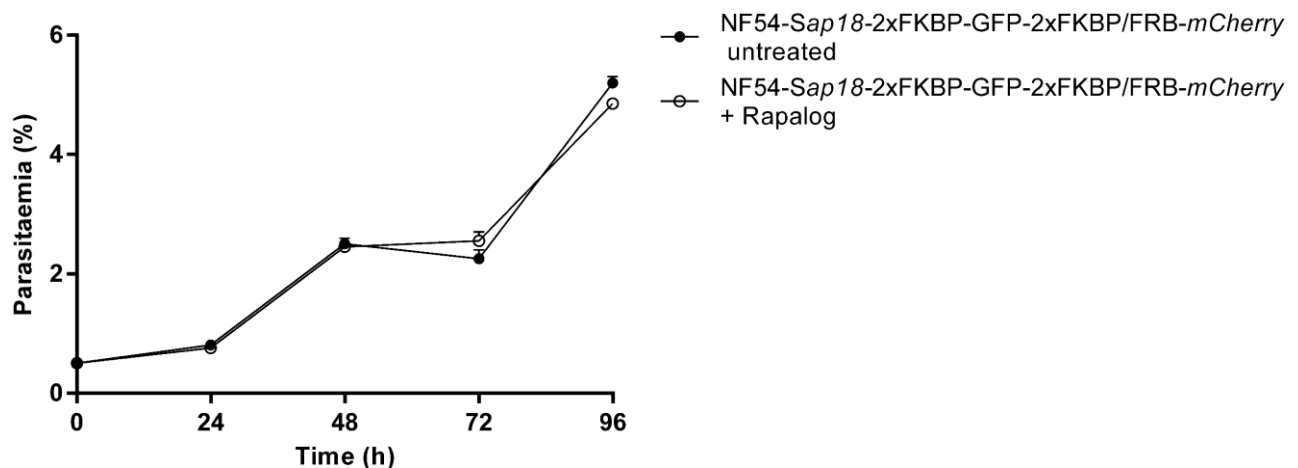


Figure 3.31: Growth rate of untreated and rapalog treated NF54-*Sap18*-2xFKBP-GFP-2xFKBP/FRB-*mCherry* asexual stage parasites. Synchronised parasites were monitored from ring-stage (0-6 hpi) and parasitaemia was determined every 24 h for 96 h. Parasitaemia was determined by using thin smear Giemsa-stained microscope slides with ≥ 1000 erythrocytes counted for each sample. Data are from 2 biological replicates, performed in technical duplicate, mean \pm SEM.

3.8. Functional and morphological analysis of NF54- Δ *Sap18*-GFP parasites

The NF54- Δ *Sap18*-GFP transgenic lines were compared to the NF54 wild type control with regards to growth rate and morphology of the parasite during the asexual life cycle. Parasites were visualised by Giemsa-stained light microscopy to evaluate morphology for the intraerythrocytic asexual stages

(Figure 3.32). The morphology of NF54- Δ *Sap18*-GFP parasites was similar to NF54 wild type parasites during the asexual intraerythrocytic developmental cycle and stayed within the 48 h life cycle timeframe indicating no morphological observed effects in NF54- Δ *Sap18*-GFP parasites (Figure 3.32).

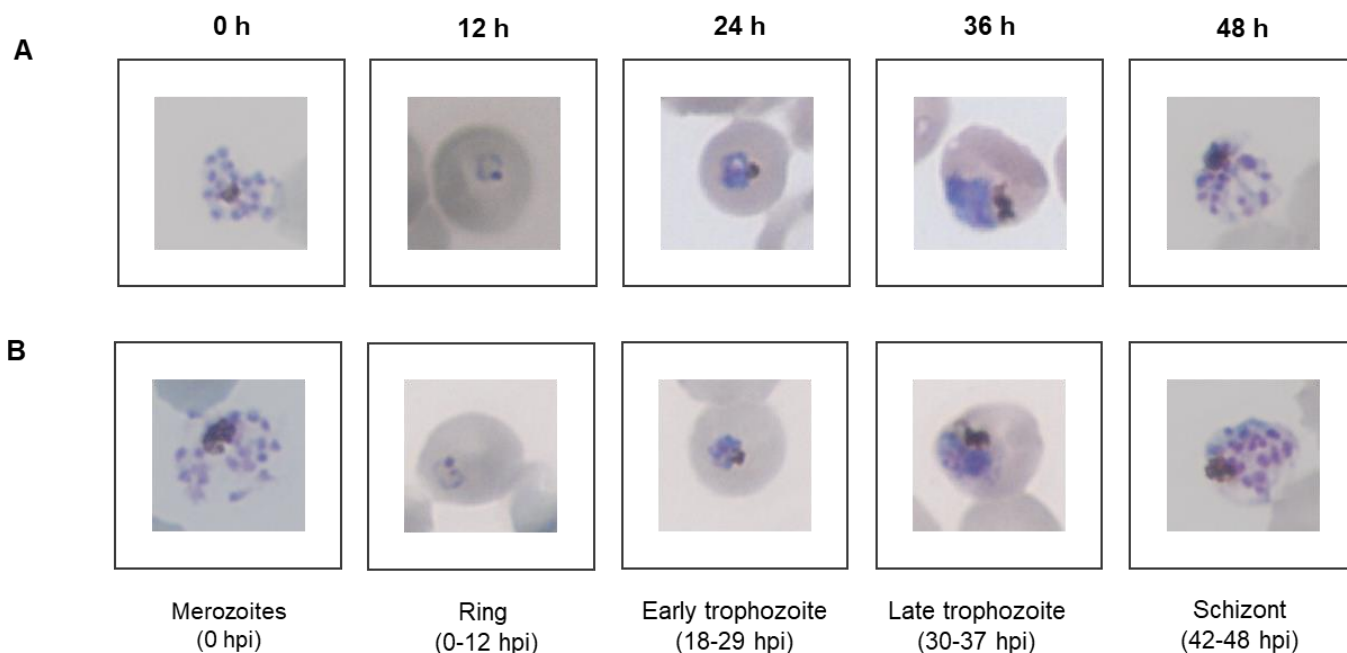


Figure 3.32: Morphological evaluation of NF54- Δ *Sap18*-GFP and NF54 wild type parasites in the asexual stages. Intraerythrocytic development of (A) NF54 wild type (control) and (B) NF54- Δ *Sap18*-GFP asexual stage parasites monitored every 12 h for 48 h visualised by Giemsa-staining under a light microscope at 1000x magnification.

Parasitaemia increase was also monitored over a 96 h period by using 2x synchronised ring-stage parasites (Figure 3.33). Results showed similar fold change increases after one and two asexual replication cycles between the NF54 wild type parasites and NF54- Δ *Sap18*-GFP parasites (Figure 3.33). No difference in mean parasitaemia between these two parasite populations was observed. This confirms that the constitutive disruption of *sap18* does not have an observable effect on parasite replication.

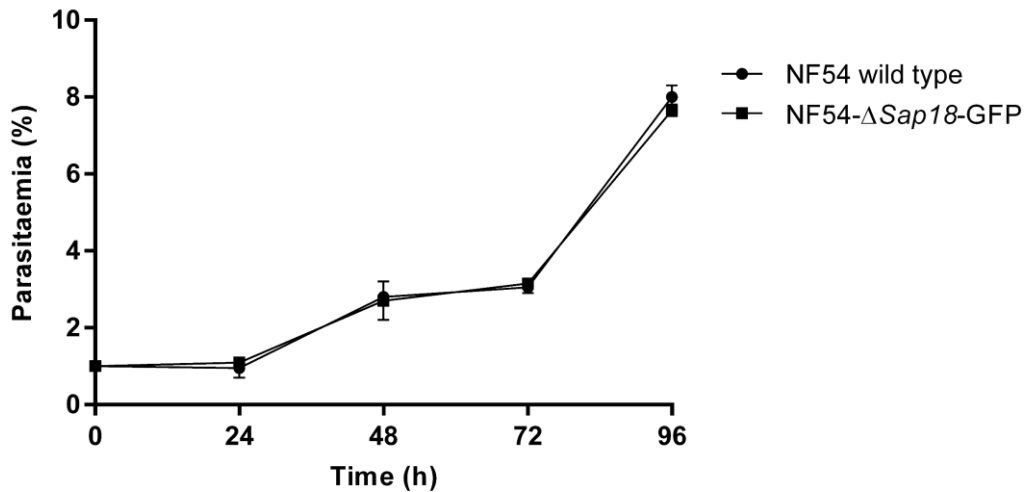


Figure 3.33: Growth rate of NF54-ΔSap18-GFP and NF54 wild type asexual stage parasites. Synchronised parasites were monitored from ring-stage (0-6 hpi) and parasitaemia was determined every 24 h for 96 h. Parasitaemia was determined by using thin smear Giemsa-stained microscope slides with ≥ 1000 erythrocytes counted for each sample. Data are from 2 biological replicates, performed in technical duplicate, mean \pm SEM.

The number of merozoites per schizont were also evaluated between NF54-ΔSap18-GFP and wild type parasites (Figure 3.34). Figure 3.34 shows no significant difference ($P > 0.05$) between the two parasite populations, indicating that *sap18* may not have a functional affect in schizont maturation.

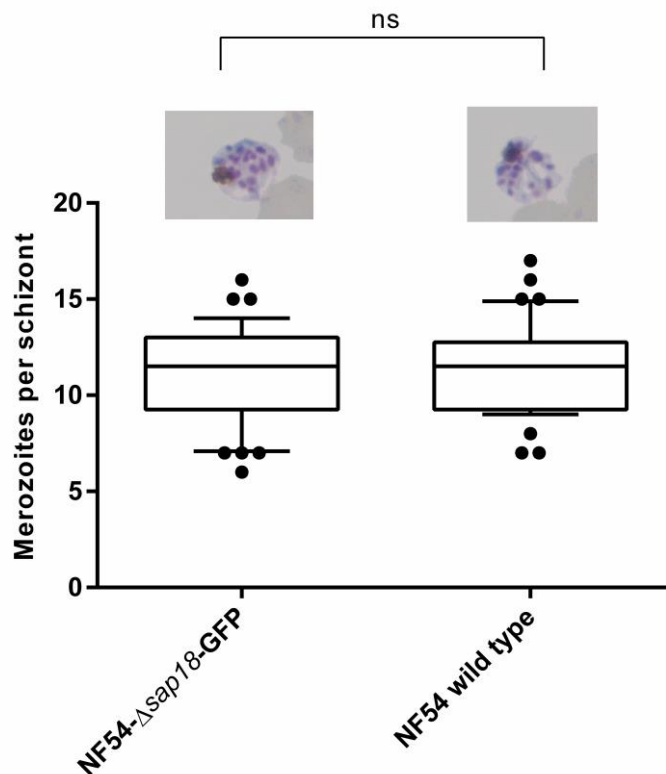


Figure 3.34: Number of merozoites per schizont in NF54-ΔSap18-GFP and NF54 wild type parasites. The number of merozoites was observed in 40 mature, segmented schizonts representative schizonts are shown above box and whiskers). Boxes represent median, 25th percentile and 75th percentile, and whiskers represent 10th and 90th percentiles. The mean number of merozoites per schizont between NF54-ΔSap18-GFP and NF54 wild type parasites was not significantly different ($P > 0.05$) using a two-tailed unpaired students t-test (ns= non-significant).

Next, NF54- Δ Sap18-GFP parasites were induced to undergo gametocytogenesis. These were evaluated based on morphology as well as their ability to commit to gametocytes (Figure 3.35). The development of stage I to V gametocytes for NF54- Δ Sap18-GFP, visualised by Giemsa-stained microscopy, corresponded to NF54 wild type parasites based on morphology (Figure 3.35) and timeframe of development.

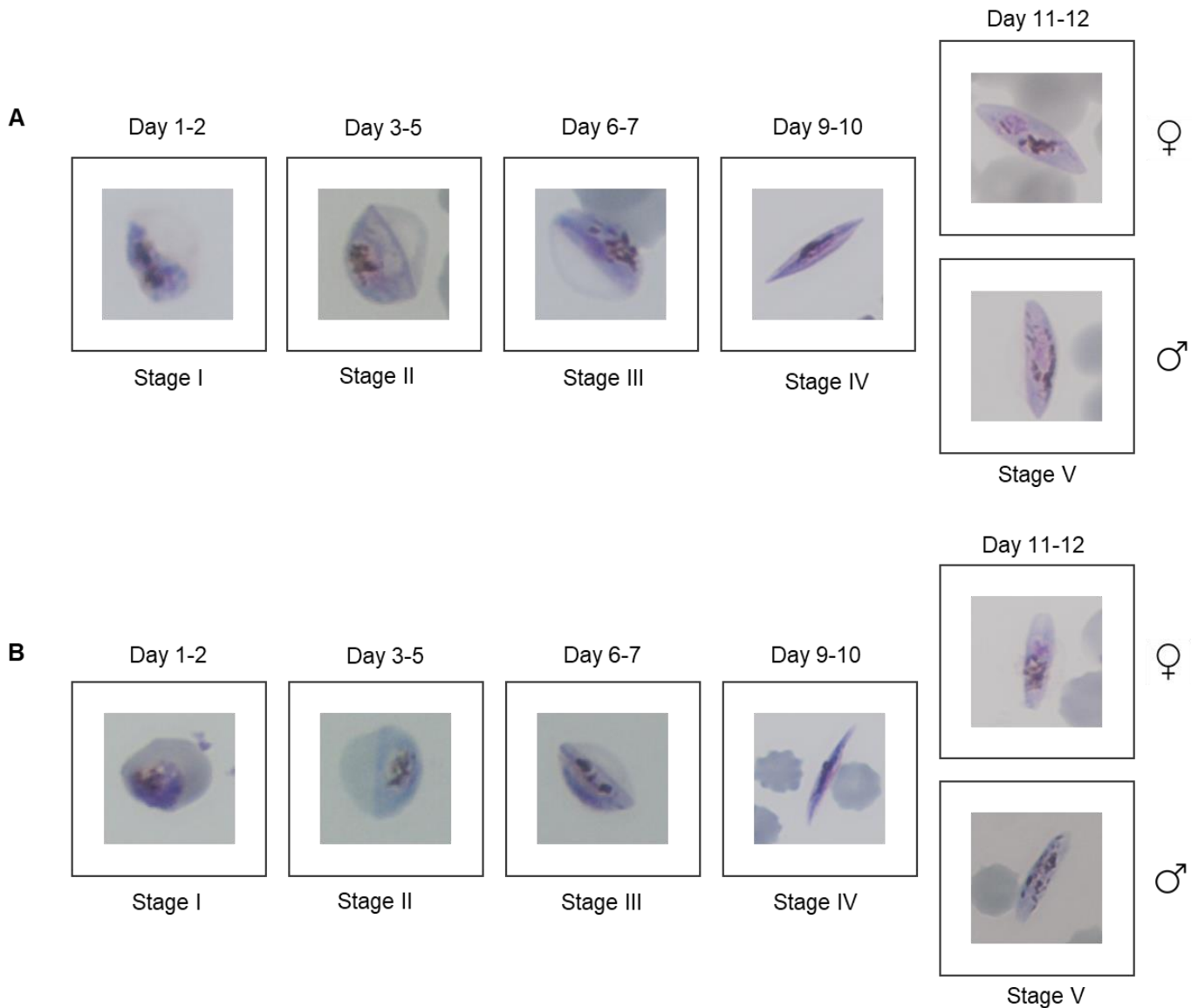


Figure 3.35: Morphological evaluation of NF54- Δ Sap18-GFP and NF54 wild type parasites in sexual stages. Gametocytogenesis of (A) NF54 wild type (control) parasites and (B) NF54- Δ Sap18-GFP parasites. Parasites were visualised using thin smear Giemsa-stained microscope slides.

These parasites were further analysed for the number of gametocytes produced and the conversion rate (Table 3.2), which is the number of asexual stage parasites that commit to gametocytogenesis. No difference in gametocyte production or conversion rate was seen between NF54- Δ Sap18-GFP and NF54 wild type control parasites. This therefore indicates that SAP18 may not be an important regulator for gametocyte development.

Table 3.2: Analysis of gametocyte production and conversion of NF54- Δ Sap18-GFP and NF54 wild type sexual stages.

Parasite line	Average gametocytemia (%) ^a	Average conversion rate (%) ^b
NF54 wild type (control)	1.25 \pm 0.13 (n=2)	29.45 \pm 1.45
NF54- Δ Sap18-GFP	1.06 \pm 0.20 (n=2)	25.93 \pm 0.23

Data are representative of (n) biological replicates, \pm SD.

^a Day 10 after gametocyte induction

^b Conversion factor = $\frac{\text{number of stage 2 gametocytes on day 5}}{\text{number of rings on day 2}}$

3.9. Involvement of SAP18 in cell cycle control

3.9.1. DFMO inhibition of *P. falciparum* parasites

SAP18 is thought to act as an early cell cycle regulator with an increased mRNA abundance during cell cycle arrest and a decrease in abundance during re-entry. Thus, it was hypothesised that *sap18* may function as an important regulator at this cell cycle checkpoint either by affecting the parasites' re-entry or arrest during the cell cycle. We investigated the effect of the disruption of *sap18* on DFMO treatment during cell cycle arrest and re-entry.

The IC₅₀ of DFMO was confirmed against wild type intraerythrocytic NF54 *P. falciparum* parasites as well as transgenic NF54- Δ Sap18-GFP parasites through *in vitro* SYBR Green I fluorescence assay with no significant difference ($P > 0.05$, n=3) observed (Figure 3.36). The IC₅₀ of DFMO indicates its potency as an antiproliferative agent and refers to the concentration at which it is lethal to 50 % of the cells *in vitro*. The IC₅₀ value of DFMO in NF54 wild type parasites was determined as 906.4 \pm 89.3 μ M and corresponds to previous reports of an IC₅₀ 1.25 \pm 0.42 mM for a 3D7 strain [136]. NF54- Δ Sap18-GFP parasites showed no significant difference ($P > 0.05$) to the NF54 wild type parasites, with an IC₅₀ of 646.5 \pm 14.3 μ M (Figure 3.36). For DFMO-induced cell cycle arrest studies, parasites were treated with DFMO at 2xIC₅₀ concentrations.

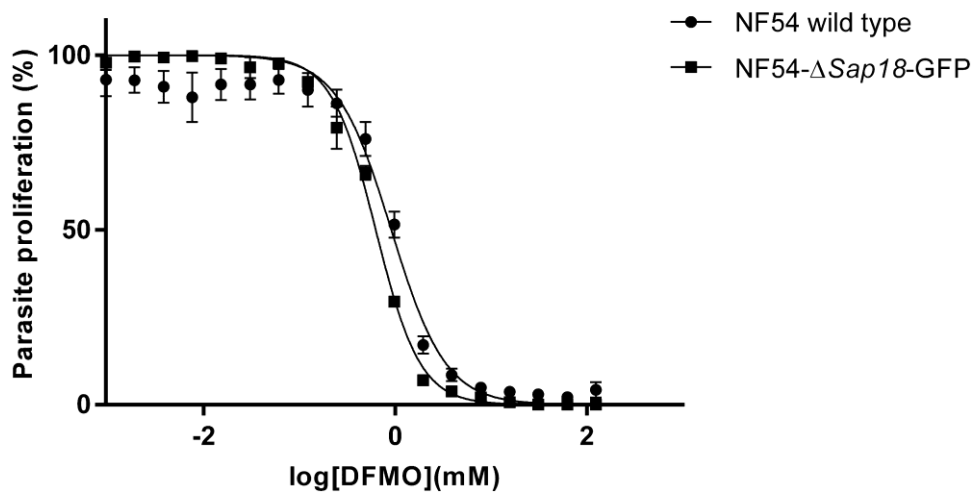


Figure 3.36: Dose-response evaluation of DFMO inhibition of NF54- Δ Sap18-GFP and NF54 wild type parasites. DFMO dose-response against NF54- Δ Sap18-GFP and NF54 wild type parasites were determined using SYBR Green I-based assay for 96 h drug pressure on ring-stage parasite cultures at 1 % haematocrit and 1 % parasitaemia. Results are representative of 3 biological replicates ($n=3 \pm$ SEM), performed in technical triplicate. Non-linear regression curves were generated using GraphPad Prism version 6.01 software. Error bars represent standard deviation. For some points, error bars fall within the symbols. The mean IC_{50} between NF54 wild type and NF54- Δ Sap18-GFP asexual stage parasites was not significantly different using a two-tailed multiple student's t-test ($P>0.05$).

Subsequently, the effect of cell cycle arrest was evaluated on NF54 wild type parasites and NF54- Δ Sap18-GFP parasites. For initial evaluation and quantification of the effect of DFMO induced cell cycle arrest, parasites were evaluated over 96 h with parasitaemia (Figure 3.37).

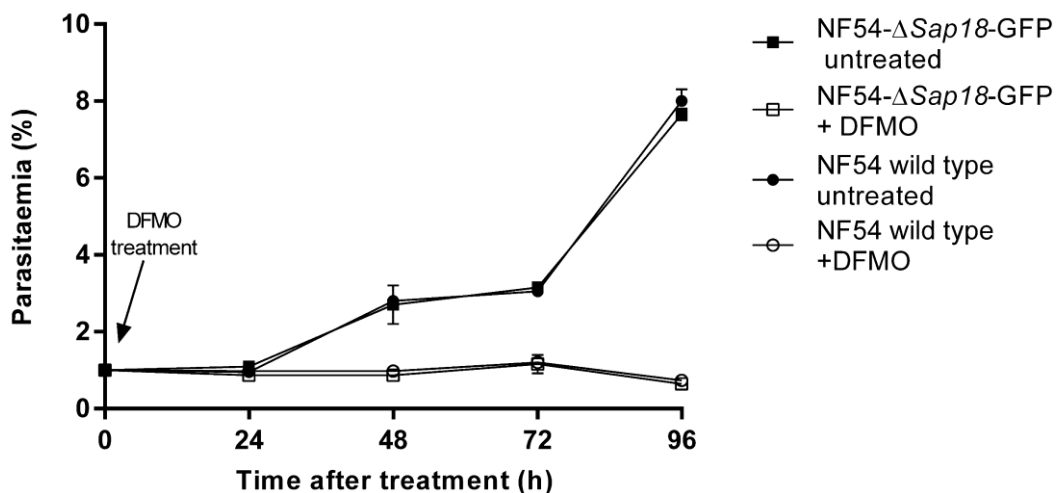


Figure 3.37: DFMO arrest of NF54- Δ Sap18-GFP and NF54 wild type parasites. Proliferation of untreated *P. falciparum* NF54 wild type (\bullet) and NF54- Δ Sap18-GFP (\blacksquare) parasites was monitored over 96 h. NF54 wild type DFMO-arrested parasites (\circ) (1 % parasitaemia, 5 % haematocrit) and NF54- Δ Sap18-GFP DFMO-arrested parasites were monitored over 96 h (\square) (1 % parasitaemia, 5 % haematocrit). Giemsa-stained microscope slides with ≥ 1000 events counted for each sample. Data are from 2 biological replicates, performed in technical duplicate, mean \pm SEM.

Untreated NF54- $\Delta Sap18$ -GFP and NF54 wild type parasites progressed through the life cycle showing no difference in parasitaemia over 96 h, confirming previous data of *sap18* not being essential during parasite proliferation (Figure 3.37).

DFMO treatment maintained both NF54- $\Delta Sap18$ -GFP and wild type parasites in an arrested state at a constant parasitaemia (1-2 %) for 96 h after treatment. After 24 h of DFMO treatment, there was no difference in parasitaemia between both untreated and DFMO-arrested parasite populations (Figure 3.37). A three- or eight-fold difference in parasitaemia between untreated NF54 wild type parasites and DFMO-arrested NF54 wild type parasites was detected at 48 h, 72 h and 96 h. NF54- $\Delta Sap18$ -GFP parasites showed a similar difference in parasitaemia between untreated and DFMO-arrested NF54- $\Delta Sap18$ -GFP parasites at 48 h, 72 h and 96 h. These data indicate NF54 wild type and NF54- $\Delta Sap18$ -GFP parasites are able to arrest upon addition of DFMO treatment, compared to their untreated controls. Subsequently, arrested *P. falciparum* NF54 parasite lines were reversed after 24 h of DFMO treatment through putrescine supplementation (Figure 3.38).

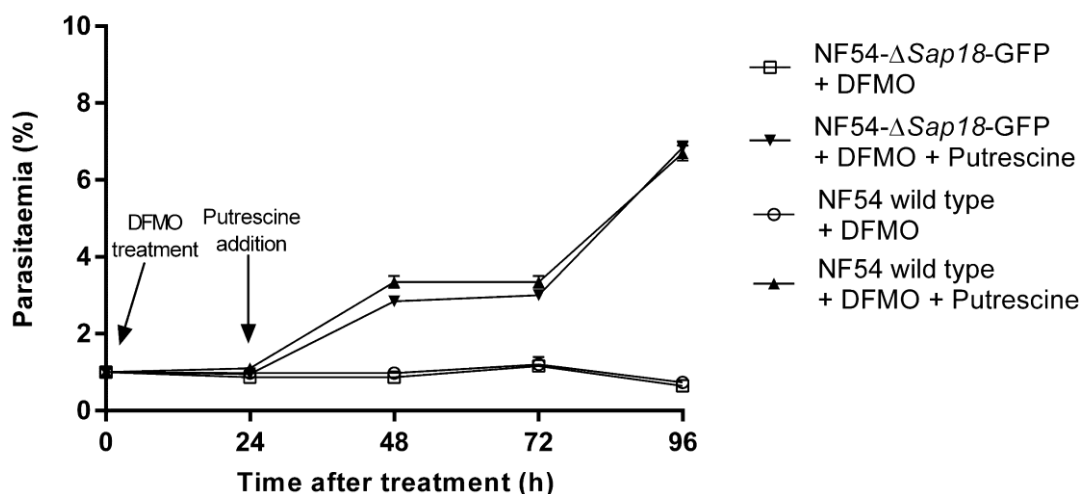


Figure 3.38: DFMO arrest and reversal of NF54- $\Delta Sap18$ -GFP and NF54 wild type parasites. NF54 wild type DFMO-arrested parasites (\circ) (1 % parasitaemia, 5 % haematocrit) were monitored over 96 h or reversed after 24 h of DFMO treatment through 2 mM putrescine supplementation (\blacktriangle). NF54- $\Delta Sap18$ -GFP DFMO-treated parasites (\square) (1 % parasitaemia, 5 % haematocrit) were monitored over 96 or reversed after 24 h of DFMO treatment through 2 mM putrescine supplementation (\blacktriangledown). Giemsa-stained microscope slides with ≥ 1000 events counted for each sample. Data are from 2 biological replicates, performed in technical duplicate, mean \pm SEM.

In the presence of putrescine, both parasite populations progressed through the life cycle showing an increase in parasitaemia (Figure 3.38). Putrescine-rescued and DFMO-arrested NF54 wild type parasites showed a three- or seven-fold difference in parasitaemia at 48 h, 72 h and 96 h, due to parasites re-entering the cell cycle. A similar profile was seen for putrescine-rescued and DFMO-arrested NF54- $\Delta Sap18$ -GFP parasites at 48 h, 72 h and 96 h (Figure 3.38). These data indicated that NF54 wild type and NF54- $\Delta Sap18$ -GFP parasites are able to undergo cell cycle arrest for 24 h and re-enter the cell cycle upon addition of putrescine.

To validate life cycle arrest and re-entry, parasites were aged morphologically over 96 h in both NF54 wild type (Figure 3.39 A) and NF54- $\Delta Sap18$ -GFP (Figure 3.39 B) parasites. After 24 h of DFMO

treatment, there was no visible difference in morphology between both parasite populations of untreated and DFMO-arrested parasites, with the majority of parasites classified at an age between 24-29 hpi. DFMO treatment of both parasite populations produced metabolically and structurally uncompromised arrested parasites and the arrested parasites were able to be reversed upon putrescine supplementation, 24 h after initial treatment. After 48 h, the observed life cycle halt was more evident, where both untreated wild type and NF54- Δ *Sap18*-GFP parasite populations had progressed through schizogony to become rings, while DFMO treated parasites were still halted at early trophozoite stage. In the presence of putrescine, both these parasite populations had re-entered and progressed through the cell cycle however, these parasite populations showed a slight lag in development as they were only schizonts at 48 h after initial treatment, compared to their untreated controls which had progressed through schizogony (Figure 3.39 A and B).

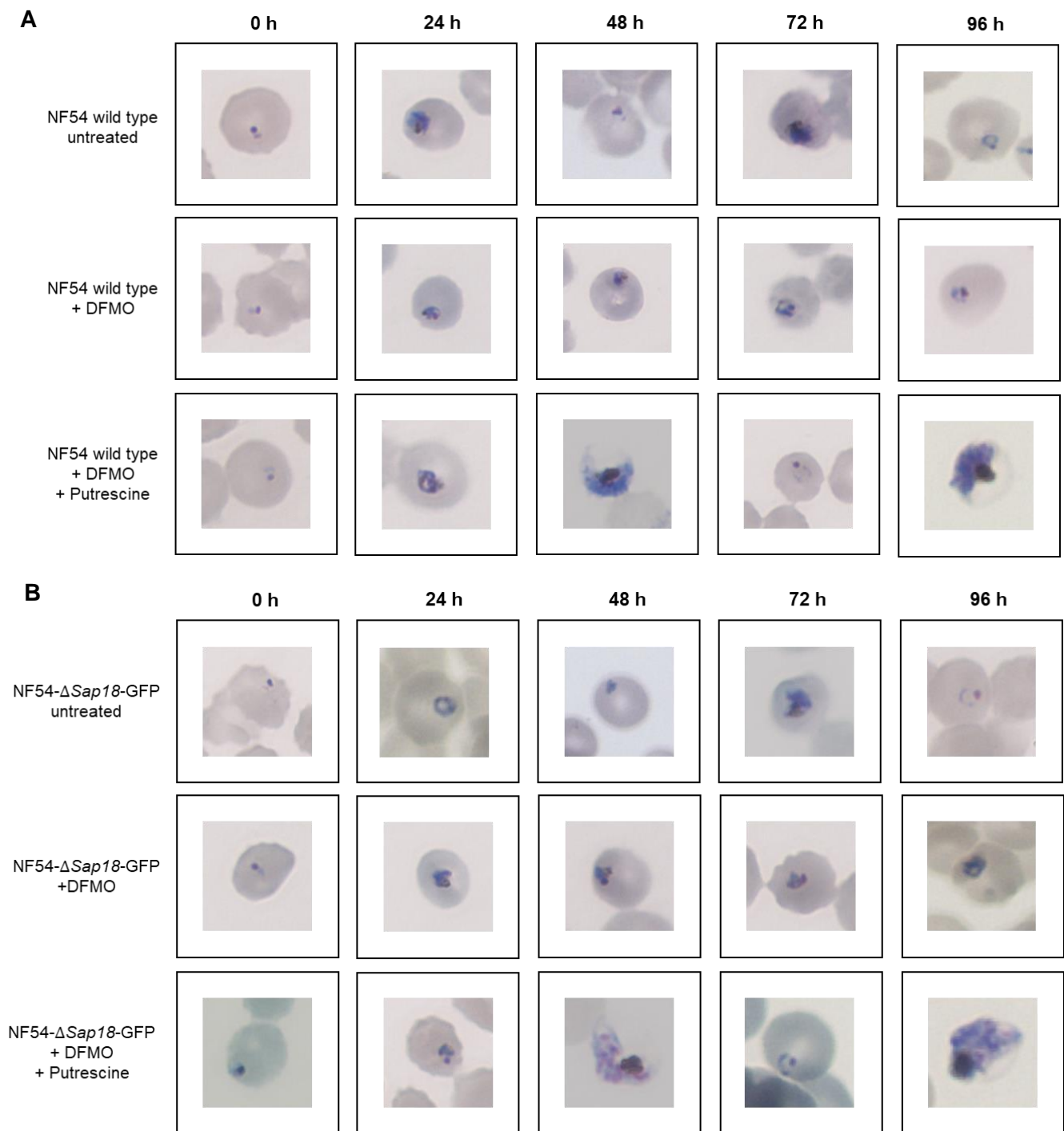


Figure 3.39: Morphological evaluation of DFMO arrest and re-entry of NF54-Δ*Sap18*-GFP and NF54 wild type asexual stage parasites. (A) NF54 wild type parasites monitored over 96 h in the presence of DFMO or addition of putrescine 24 h after arrest visualised by Giemsa staining at 1000x magnification. (B) NF54-Δ*Sap18*-GFP parasites monitored over 96 h in the presence of DFMO or addition of putrescine 24 h after arrest visualised by Giemsa staining at 1000x magnification.

Together, this data shows that both NF54 wild type and NF54-Δ*Sap18*-GFP parasites arrested and re-entered the cell cycle as seen by the increase in parasitaemia and morphology stage progression. Since NF54-Δ*Sap18*-GFP parasites arrested and re-entered the cell cycle similar to that of NF54 wild type parasites, it leads to the assumption that although SAP18 may be thought of as an important cell cycle regulator, it may not influence the parasite's ability to enter cell cycle arrest or re-enter the cell cycle upon supplementation with DFMO or putrescine, respectively

Chapter 4: Discussion

The *P. falciparum* parasite has such a complex life cycle which occurs across a human host and *Anopheles* mosquito vector. It is the ability to progress through different stages and different biological processes such as proliferation and differentiation that implies extraordinary control of these stage transitions to allow the parasite to survive and complete its life cycle. The transcriptome of the parasite is in fact so finely tuned that genes are expressed in a “just-in-time” fashion, only in the particular life cycle stage where they are needed. This implies added levels of control, such as epigenetic regulation of gene expression to regulate life cycle progression. In this study we aimed to develop genetically manipulated lines of the putative histone deacetylase I complex protein SAP18 and a member of histone deacetylase III SIR2A for functional analyses. These genes were chosen based on their association as potential early responders or epigenetic repressors as well as their unique gene expression responses during induced cell cycle arrest and re-entry events. To achieve this aim, we cloned these gene fragments into SLI-TGD or SLI-Sandwich plasmids for transfection into *P. falciparum* NF54 parasites to generate transgenic lines. Several of these developed lines were evaluated for their essentiality and functionality in asexual stage proliferation and gametocyte differentiation.

SAP18 was first identified in higher eukaryotes, as a regulatory component of the SIN3/HDAC complex and it was shown to enhance transcriptional repression by interacting with transcription factors [93]. We hypothesised that SAP18 acts as a SIN3-repressor complex protein within *P. falciparum*. Truncation of SAP18 at the N-terminus using the SLI-TGD plasmid is expected to make the protein non-functional due to the presence of the SIN3-associated polypeptide p18 domain within the C-terminus region of the protein which was deleted in the TGD construct. This study produced a NF54- $\Delta sap18$ -GFP transgenic line that allows constitutive disruption of the gene of interest by producing a truncated N-terminus protein fused to GFP. The transgenic line showed comparable proliferation of asexual stages to the parental NF54 wild type parasite line. A previous knock-out study within *P. berghei* showed only a slight reduction in parasite proliferation, supporting the data from this study that SAP18 is not essential in asexual stage parasites. Gametocyte differentiation did not show differences in gametocytaemia or conversion rate compared to NF54 wild type parasites. Furthermore, the slight reduction in gametocyte conversion may be due to the continuous long-term *in vitro* culturing of transgenic asexual stage parasites.

It may be possible that the truncated SAP18 protein product is still able to perform its function and may be why we do not see a growth defect within the parasite. On the other hand, if the truncated protein is indeed no longer functional, the observed lack of growth defect may indicate that this protein does not have a vital role. Functional annotation of SAP18 predicted a SIN3-associated polypeptide p18 domain with possible HDAC regulatory function. However, disruption of SAP18 showed it was not essential to cell cycle regulation. This may indicate that SAP18 may not have the same regulatory function in *P. falciparum* as in other species or the SIN3/HDAC complex may not be an important cell cycle regulator in *P. falciparum* compared to other species. STRING network data failed to identify any known

interaction of SAP18 with SIN3A or a related protein. It did, however, identify HDAC1 as a possible interaction partner, confirming previous data from other species. The low sequence homology between *P. falciparum* and higher eukaryotes makes modelling of SAP18 difficult. However, we could confirm the presence of a SAP18 domain, due to the high structural similarity to *M. musculus* SAP18 and the AlphaFold predicted structure. Thus, using homology-based modelling, machine learning predictions as well as domain predictions, we are able to confirm that on a structural and sequence level, PF3D7_0711400 resembles a SAP18 protein-like subunit with possible HDAC function. It will therefore be important to identify how SAP18 interacts with HDAC1 to mediate epigenetic repression as well as the importance of SAP18 to HDAC1 activity. It will be interesting to identify the location of SAP18's protein-protein interaction domain to identify the binding interaction of SAP18 and HDAC1.

An additional domain was identified in SAP18 with possible ASAP complex functionality. The *P. falciparum* SAP18 homology model also resembles an ubiquitin-like domain found within the SAP18 *H. sapiens* homologue, which is thought to be required for efficient splicing regulatory activity [95]. A similar structural domain was identified within *P. falciparum* SAP18, possibly indicating a similar functionality in splicing within an ASAP complex within *P. falciparum*. However, no such complex has been identified within *P. falciparum* to date. This study may also provide insight into SAP18 as a possible drug target however, this may be questionable due to the fact that SAP18 appears to be non-essential for asexual proliferation and gametocyte viability.

Since it is a proposed regulator of HDAC1, the result of SAP18 disruption may be greater than what was observed in this study, particularly at the level of consequent changes to downstream gene expression resulting from disruption of HDAC1 activity or the associated histone code. Future studies may look into the downstream effect of SAP18 disruption by analysing the transcriptional changes and histone post-translational modification alteration consequent to SAP18 disruption. Though we expect SAP18 to regulate HDAC1 like its homologues, it may not play the same role in *Plasmodium*. However, it may also be that redundancy occurs between HDAC complexes, and that another HDAC complex assumes the role of HDAC1 upon SAP18 disruption or a protein with similar function to SAP18 taking over its function, though no such protein has been observed in *P. falciparum* to date.

Various genetic manipulation approaches have been developed to study gene function in malaria parasites. As it is difficult to knock-out essential genes, due to the haploid genome, a knock sideways system has been developed. We used the SLI-Sandwich system which allows for proteins of interest to be tagged with GFP at their C-terminus in order to study localisation and protein function. Confocal microscopy of NF54- Δ Sap18-GFP and NF54-Sap18-2xFKBP-GFP-2xFKBP parasite lines showed that SAP18 localises to the nucleus. These data correspond to previous data within mammals where SAP18 is localised to the nucleus and nuclear speckle [131, 137]. Although mislocalisation allowed the GFP-tagged SAP18 fusion protein to mislocalise to the plasma membrane, complete mislocalisation was not achieved as some protein was observed to remain within the nucleus of the parasite. This remaining protein may still be sufficient to carry out its function within the parasite, and may explain the unnoticeable growth defect.

SAP18 is thought to function as regulatory subunit of the HDAC complex, removing acetylation marks which leads to decreased levels of gene expression in affected genes. These biological features of the parasite have yet to be extensively studied in SAP18 and SIR2A genetically modified parasites, and hold great potential in revealing functional aspects of parasite life cycle progression, particularly pertaining to transmission. It remains questionable, however, whether these early responders and/or epigenetic repressors play a role in cell cycle progression or whether these genes simply act as transcriptional repressors to regulate RNA synthesis. As part of our study, SAP18 was evaluated for its effect relating to DFMO treatment. NF54- Δ *Sap18*-GFP and NF54 wild type parasites were arrested with DFMO for 24 hours and allowed to re-enter the cell cycle upon addition of putrescine. Parasites containing a non-functional SAP18 product showed a similar arrest profile compared to control parasites. These parasites also re-entered the cell cycle and progressed through their life cycle, as did the control parasites. This indicates that SAP18 may not be an important early responder necessary for the parasite's arrest and re-entry into the cell cycle. However, further in-depth analysis into the function of SAP18 as a possible downstream effector during these stages of arrest and re-entry is necessary to fully characterise the role of SAP18.

SIR2A was the first known epigenetic factor shown to play a role in transcriptional control of antigenic variation in *P. falciparum*. SIR2A was functionally annotated to contain a SIR2 catalytic domain with NAD⁺-dependent binding. In this study, the SIR2A STRING network identified histones H2A, H3 and H4 as possible interacting partners, which confirms SIR2As function as it binds to these histone tails to deacetylate them. Interestingly, an interaction with GCN5, a histone acetyltransferase was also identified. GCN5 also forms part of the Bromo-domain reader complex within *P. falciparum* [133].

In this study, a NF54-*Sir2A*-2xFKBP-GFP-2xFKBP transgenic parasite line was produced to analyse the functionality and localisation of SIR2A. We were unable to develop a disrupted SIR2A transgenic line using the SLI-TGD system. [125]. SIR2A contains a catalytic domain at residues 19-273, which spans the majority of the protein. Hence, using a small homologous region of the *sir2A* gene for cloning into SLI-TGD for essentiality studies, allowed us to reduce the likelihood of the production of a functional SIR2A protein once integrated into the *sir2A* locus. However, this also brought about challenges to integrate small gene fragments into the locus as it may increase the chance of off-target integration of the recombinant plasmid, aggravated by the biased AT-rich genome in *Plasmodium* [138]. We also observed large differences in the length of time to possible integration of both manipulated *sir2A* parasite lines, compared to the recombinant *sap18* parasite lines. This may explain the lack of integration of the smaller N-terminal *sir2A* homologous fragment, as previous studies suggested there was a correlation of time to integration and length of targeting region [125]. In addition, these differences in the number of days to integration may be due to the differences in the transcript level of *sir2A* and *sap18* as SLI depends on the expression of the neomycin resistance marker, which is under the control of the promoter of the target gene. However, Birnbaum *et al.* (2017) did not find any correlation between expression levels and time to integration.

Another possibility pertaining to the lack of successful integration of N-terminal recombinant *sir2A* parasites, may be due to the possibility of SIR2A being essential to parasite proliferation. This would mean that upon homologous recombination of the truncated *sir2A* into the *sir2A* locus, the parasite would not be able to proliferate and hence, no parasites would be viable for functional studies. However, previous data suggests SIR2A is not essential to asexual parasites, but rather knock-out of SIR2A increases the number of merozoites produced on each round of parasite multiplication and hyperacetylates ribosomal RNA genes [112].

This study may provide a step to aid in the further understanding of the complex nature of these epigenetic regulators, SAP18 and SIR2A and their functionality within the cell cycle.

Chapter 5: Conclusion

This research aimed to develop genetically modified *P. falciparum* SAP18 and SIR2A parasite lines as well as to functionally characterise several of these lines. A cloning strategy was followed to transfect the recombinant SLI plasmids into *P. falciparum* NF54 parasites. Three transgenic parasite lines were successfully created, namely NF54- Δ *Sap18*-GFP, NF54-*Sap18*-2xFKBP-GFP-2xFKBP and NF54-*Sir2A*-2xFKBP-GFP-2xFKBP. However, integration of pSLI-TGD-*Sir2A* was not successful.

A *sir2A* knock sideways transgenic parasite line was generated for future functional studies into the importance of SIR2A during its unique expression during gametocytogenesis as well as its function in the regulation of acetylation of histones. Functional validation of these epigenetic regulators in *P. falciparum* will be important to provide attractive antimalarial drug targets.

Additionally, SAP18 was predicted to contain a SIN3-associated polypeptide p18 domain with possible HDAC regulatory function, due its structural and sequence similarities to higher eukaryotic organisms. Although SAP18 was shown not to be essential within asexual proliferation or gametocyte differentiation, it may play a key role in future studies investigating its importance in HDAC activity, or its functional effect on the resulting histone code within the parasite. SAP18 disruption was shown not to influence the ability of the parasite to arrest and re-enter the cell cycle, which may indicate a possible higher unknown functionality of SAP18 as a proposed cell cycle regulator.

References

- 1 WHO. (2020) World malaria report 2020: 20 years of global progress and challenges. pp. 1-299
- 2 Tilley, L., Straimer, J., Gnädig, N. F., Ralph, S. A. and Fidock, D. A. (2016) Artemisinin action and resistance in *Plasmodium falciparum*. Trends in Parasitology. **32**, 682-696
- 3 Nixon, G. L., Moss, D. M., Shone, A. E., Lalloo, D. G., Fisher, N., O'Neill, P. M., Ward, S. A. and Biagini, G. A. (2013) Antimalarial pharmacology and therapeutics of atovaquone. Journal of Antimicrobial Chemotherapy. **68**, 977-985
- 4 Ays, A., C. K., Snaith, R., Cottingham, M. G., Gilbert, S. C. and Hill. (2017) Enhancing protective immunity to malaria with a highly immunogenic virus-like particle vaccine. Scientific Reports. **7**, 46621
- 5 Dato, M. S., Magloire Natama, H., Somé, A., Traoré, O., Rouamba, T., Bellamy, D., Yameogo, P., Valia, D., Tegneri, M. and Ouedraogo, F. (2021) High efficacy of a low dose candidate malaria vaccine, R21 in 1 adjuvant Matrix-M, with seasonal administration to children in Burkina Faso. The Lancet. **397**, 1809-1818
- 6 Hyde, J. E. (2007) Drug-resistant malaria– an insight. The FEBS Journal. **274**, 4688-4698
- 7 Yamauchi, L. M., Coppi, A., Snounou, G. and Sinnis, P. (2007) *Plasmodium* sporozoites trickle out of the injection site. Cellular Microbiology. **9**, 1215-1222
- 8 Mota, M. M., Pradel, G., Vanderberg, J. P., Hafalla, J. C., Frevert, U., Nussenzweig, R. S., Nussenzweig, V. and Rodríguez, A. (2001) Migration of *Plasmodium* sporozoites through cells before infection. Science. **291**, 141-144
- 9 Tilley, L., Dixon, M. W. A. and Kirk, K. (2011) The *Plasmodium falciparum*-infected red blood cell. The International Journal of Biochemistry & Cell Biology. **43**, 839-842
- 10 Baumeister, S., Winterberg, M., Przyborski, J. M. and Lingelbach, K. (2010) The malaria parasite *Plasmodium falciparum*: cell biological peculiarities and nutritional consequences. Protoplasma. **240**, 3-12
- 11 Bannister, L. and Mitchell, G. (2003) The ins, outs and roundabouts of malaria. Trends in Parasitology. **19**, 209-213
- 12 Biamonte, M. A., Wanner, J. and Le Roch, K. G. (2013) Recent advances in malaria drug discovery. Bioorganic & Medicinal Chemistry Letters. **23**, 2829-2843
- 13 Kaushal, D. C., Carter, R., Miller, L. H. and Krishna, G. (1980) Gametocytogenesis by malaria parasites in continuous culture. Nature. **286**, 490-492
- 14 Brancucci, N. M. B., Gerdt, J. P., Wang, C. Q., De Niz, M., Philip, N., Adapa, S. R., Zhang, M., Hitz, E., Niederwieser, I., Boltryk, S. D., Laffitte, M. C., Clark, M. A., Grüning, C., Ravel, D., Blancke Soares, A., Demas, A., Bopp, S., Rubio-Ruiz, B., Conejo-Garcia, A., Wirth, D. F., Gendaszewska-Darmach, E., Duraisingh, M. T., Adams, J. H., Voss, T. S., Waters, A. P., Jiang, R. H. Y., Clardy, J. and Marti, M. (2017) Lysophosphatidylcholine regulates sexual stage differentiation in the human malaria parasite *Plasmodium falciparum*. Cell. **171**, 1532-1544
- 15 Smith, T. G., Lourenco, P., Carter, R. and Walliker, D. (2000) Commitment to sexual differentiation in the human malaria parasite, *Plasmodium falciparum*. Parasitology. **121**, 127-133
- 16 Bruce, M. C., Alano, P., Duthie, S. and Carter, R. (1990) Commitment of the malaria parasite *Plasmodium falciparum* to sexual and asexual development. Parasitology. **100**, 191-200
- 17 Hawking, F., Wilson, M. E. and Gammage, K. (1971) Evidence for cyclic development and short-lived maturity in the gametocytes of *Plasmodium falciparum*. Transactions of the Royal Society of Tropical Medicine and Hygiene. **65**, 549-559
- 18 Young, J. A., Fivelman, Q. L., Blair, P. L., De La Vega, P., Le Roch, K. G., Zhou, Y., Carucci, D. J., Baker, D. A., Winzeler, E. A. and Winzeler, E. A. (2005) The *Plasmodium falciparum* sexual development transcriptome: A microarray analysis using ontology-based pattern identification. Molecular & Biochemical Parasitology. **143**, 67-79
- 19 Sinden, R. E. (1982) Gametocytogenesis of *Plasmodium falciparum* *in vitro*: an electron microscopic study. Parasitology. **84**, 1-11
- 20 Hanssen, E., Knoechel, C., Dearnley, M., Dixon, M. W. A., Le Gros, M., Larabell, C. and Tilley, L. (2012) Soft X-ray microscopy analysis of cell volume and hemoglobin content in erythrocytes infected with asexual and sexual stages of *Plasmodium falciparum*. Journal of Structural Biology. **177**, 224-232
- 21 Smalley, M. E. and Sinden, R. E. (1977) *Plasmodium falciparum* gametocytes: their longevity and infectivity. Parasitology. **74**, 1-8

- 22 Billker, O., Shaw, M. K., Margos, G. and Sinden, R. E. (1997) The roles of temperature, pH and mosquito factors as triggers of male and female gametogenesis of *Plasmodium berghei* *in vitro*. *Parasitology*. **114**, 1-7
- 23 Vlachou, D., Schlegelmilch, T., Runn, E., Mendes, A., Kafatos, F. C., Firtel, R. and Zernicka-Goetz, M. (2006) The developmental migration of *Plasmodium* in mosquitoes. *Current Opinion in Genetics & Development*. **16**, 384-391
- 24 Guerreiro, A., Deligianni, E., Santos, J. M., Silva, P. A. G. C., Louis, C., Pain, A., Janse, C. J., Franke-Fayard, B., Carret, C. K., Siden-Kiamos, I. and Mair, G. R. (2014) Genome-wide RIP-Chip analysis of translational repressor-bound mRNAs in the *Plasmodium* gametocyte. *Genome Biology*. **15**, 1-16
- 25 Vlachou, D., Zimmermann, T., Cantera, R., Janse, C. J., Waters, A. P. and Kafatos, F. C. (2004) Real-time, *in vivo* analysis of malaria ookinete locomotion and mosquito midgut invasion. *Cellular Microbiology*. **6**, 671-685
- 26 Akaki, M. and Dvorak, J. A. (2005) A chemotactic response facilitates mosquito salivary gland infection by malaria sporozoites. *The Journal of Experimental Biology*. **208**, 3211-3218
- 27 Haase, S. B. and Wittenberg, C. (2014) Topology and control of the cell-cycle-regulated transcriptional circuitry. *Genetics*. **196**, 65-90
- 28 Spencer, Sabrina L., Cappell, Steven D., Tsai, F.-C., Overton, K. W., Wang, Clifford L. and Meyer, T. (2013) The proliferation-quiescence decision is controlled by a bifurcation in CDK2 activity at mitotic exit. *Cell*. **155**, 369-383
- 29 Gritzmacher, C. A. and Reese, R. T. (1984) Protein and nucleic acid synthesis during synchronized growth of *Plasmodium falciparum*. *Journal of Bacteriology*. **160**, 1165-1167
- 30 Arnot, D. E., Ronander, E. and Bengtsson, D. C. (2011) The progression of the intra-erythrocytic cell cycle of *Plasmodium falciparum* and the role of the centriolar plaques in asynchronous mitotic division during schizogony. *International Journal for Parasitology*. **41**, 71-80
- 31 Matthews, H., Duffy, C. W. and Merrick, C. J. (2018) Checks and balances? DNA replication and the cell cycle in *Plasmodium*. *Parasites & Vectors*. **11**, 216-216
- 32 Pardee, A. (1974) A restriction point for control of normal animal cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America*. **71**, 1286-1290
- 33 van Biljon, R., Niemand, J., van Wyk, R., Clark, K., Verlinden, B., Abrie, C., von Grüning, H., Smidt, W., Smit, A., Reader, J., Painter, H., Llinás, M., Doerig, C. and Birkholtz, L.-M. (2018) Inducing controlled cell cycle arrest and re-entry during asexual proliferation of *Plasmodium falciparum* malaria parasites. *Scientific Reports*. **8**, 16581
- 34 Koomoa, D.-L. T., Yco, L., Borsics, T., Wallick, C. J. and Bachmann, A. S. (2008) Ornithine decarboxylase inhibition by DFMO activates opposing signaling pathways via phosphorylation of both Akt/PKB and p27Kip1 in neuroblastoma. *Cancer Research*. **68**, 9825
- 35 Meyskens, F. L. and Gerner, E. W. (1999) Development of difluoromethylornithine (DFMO) as a chemoprevention agent. *Clinical Cancer Research*. **5**, 945-951
- 36 Abraham, A. K. and Pihl, A. (1981) Role of polyamines in macromolecular synthesis. *Trends in Biochemical Sciences*. **6**, 106-107
- 37 Assaraf, Y., Golenser, J., Spira, D. and Bachrach, U. (1984) Polyamine levels and the activity of their biosynthetic enzymes in human erythrocytes infected with the malarial parasite, *Plasmodium falciparum*. *Biochemical Journal*. **222**, 815-819
- 38 Bozdech, Z., Llinás, M., Pulliam, B. L., Wong, E. D., Zhu, J. and DeRisi, J. L. (2003) The transcriptome of the intraerythrocytic developmental cycle of *Plasmodium falciparum*. *PLoS Biology*. **1**, 85-100
- 39 Coetzee, N., Sidoli, S., van Biljon, R., Painter, H., Llinás, M., Garcia, B. A. and Birkholtz, L.-M. (2017) Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual *Plasmodium falciparum* parasites. *Scientific Reports*. **7**, 607
- 40 Gupta, A. P., Chin, W. H., Zhu, L., Mok, S., Luah, Y.-H., Lim, E.-H. and Bozdech, Z. (2013) Dynamic epigenetic regulation of gene expression during the life cycle of malaria parasite *Plasmodium falciparum*. *PLoS Pathogens*. **9**, e1003170
- 41 Gardner, M. J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R. W., Carlton, J. M., Pain, A., Nelson, K. E., Bowman, S., Paulsen, I. T., James, K., Eisen, J. A., Rutherford, K., Salzberg, S. L., Craig, A., Kyes, S., Chan, M.-S., Nene, V., Shallom, S. J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M. W., Vaidya, A. B., Martin, D. M. A., Fairlamb, A. H., Fraunholz, M. J., Roos, D. S., Ralph, S. A., McFadden, G. I., Cummings, L. M., Subramanian, G. M., Mungall, C., Venter, J. C., Carucci, D. J., Hoffman, S. L., Newbold, C., Davis, R. W., Fraser, C. M. and

- Barrell, B. (2002) Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature*. **419**, 498-511
- 42 Aurrecochea, C., Brestelli, J., Brunk, B. P., Dommer, J., Fischer, S., Gajria, B., Gao, X., Gingle, A., Grant, G., Harb, O. S., Heiges, M., Innamorato, F., Iodice, J., Kissinger, J. C., Kraemer, E., Li, W., Miller, J. A., Nayak, V., Pennington, C., Pinney, D. F., Roos, D. S., Ross, C., Stoeckert, C. J., Treatman, C. and Wang, H. (2009) PlasmoDB: a functional genomic database for malaria parasites. *Nucleic Acids Research*. **37**, 539-543
- 43 Lliná, M., Deitsch, K. W. and Voss, T. S. (2008) *Plasmodium* gene regulation: far more to factor in. *Trends in Parasitology*. **24**, 551-556
- 44 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. and Winzeler, E. A. (2003) Discovery of gene function by expression profiling of the malaria parasite life cycle. *Science*. **301**, 1503-1508
- 45 Javier Foth, B., Zhang, N., Kaur Chaal, B., Kwan Sze, S., Rainer Preiser, P. and Bozdech, Z. (2011) Quantitative time-course profiling of parasite and host cell proteins in the human malaria parasite *Plasmodium falciparum*. *Molecular Cell Proteomics*. **10**, 110.006411
- 46 Lasonder, E., Rijpma, S. R., van Schaijk, B. C. L., Hoeijmakers, W. A. M., Kensche, P. R., Gresnigt, M. S., Italiaander, A., Vos, M. W., Woestenenk, R., Bousema, T., Mair, G. R., Khan, S. M., Janse, C. J., Bártfai, R. and Sauerwein, R. W. (2016) Integrated transcriptomic and proteomic analyses of *P. falciparum* gametocytes: molecular insight into sex-specific processes and translational repression. *Nucleic Acids Research*. **44**, 6087-6101
- 47 Consortium, E. P. (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature*. **489**, 57-74
- 48 Balaji, S., Babu, M. 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 Research*. **33**, 3994-4006
- 49 Coulson, R. M. R., Hall, N. and Ouzounis, C. A. (2004) Comparative genomics of transcriptional control in the human malaria parasite *Plasmodium falciparum*. *Genome Research*. **14**, 1548-1554
- 50 Foth, B. J., Zhang, N., Mok, S., Preiser, P. R. and Bozdech, Z. (2008) Quantitative protein expression profiling reveals extensive post-transcriptional regulation and post-translational modifications in schizont-stage malaria parasites. *Genome Biology*. **9**, 177
- 51 Holliday, R. (1987) The inheritance of epigenetic defects. *Science*. **238**, 163-170
- 52 Ooi, S. K. T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C. D., Cheng, X. and Bestor, T. H. (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to *de novo* methylation of DNA. *Nature*. **448**, 714-717
- 53 Cedar, H. and Bergman, Y. (2009) Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics*. **10**, 295-304
- 54 Cooper, G. M. (2000) *The cell: A molecular approach*. Sinauer Associates, Sunderland
- 55 Strahl, B. D. and Allis, C. D. (2000) The language of covalent histone modifications. *Nature*. **403**, 41-45
- 56 Gillette, T. G. and Hill, J. A. (2015) Readers, writers, and erasers: chromatin as the whiteboard of heart disease. *Circulation research*. **116**, 1245-1253
- 57 Bannister, A. J. and Kouzarides, T. (2011) Regulation of chromatin by histone modifications. *Cell Research*. **21**, 381-395
- 58 Eberharter, A. (2002) Histone acetylation: a switch between repressive and permissive chromatin: Second in review series on chromatin dynamics. *EMBO Reports*. **3**, 224-229
- 59 Li, J., Lin, Q., Wang, W., Wade, P. and Wong, J. (2002) Specific targeting and constitutive association of histone deacetylase complexes during transcriptional repression. *Genes & Development*. **16**, 687-692
- 60 Adams, G. E., Chandru, A. and Cowley, S. M. (2018) Co-repressor, co-activator and general transcription factor: the many faces of the Sin3 histone deacetylase (HDAC) complex. *Biochemical Journal*. **475**, 3921-3932
- 61 Magnaghi-Jaulin, L., Ait-Si-Ali, S. and Harel-Bellan, A. (2000) Histone acetylation and the control of the cell cycle. *Progress in Cell Cycle Research*. **4**, 41-47
- 62 Kim, Y. B., Ki, S. W., Yosnida, M. and Horinouchi, S. (2000) Mechanism of cell cycle arrest caused by histone deacetylase inhibitors in human carcinoma cells. *The Journal of Antibiotics*. **53**, 1191-1200
- 63 Sambucetti, L. C., Fischer, D. D., Zabludoff, S., Kwon, P. O., Chamberlin, H., Trogani, N., Xu, H. and Cohen, D. (1999) Histone deacetylase inhibition selectively alters the activity and expression of

cell cycle proteins leading to specific chromatin acetylation and antiproliferative effects. *Journal of Biological Chemistry*. **274**, 34940-34947

64 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

65 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*. *Journal of Proteome Research*. **8**, 3439-3450

66 Karmodiya, K., Pradhan, S. J., Joshi, B., Jangid, R., Reddy, P. C. and Galande, S. (2015) A comprehensive epigenome map of *Plasmodium falciparum* reveals unique mechanisms of transcriptional regulation and identifies H3K36me2 as a global mark of gene suppression. *Epigenetics & Chromatin*. **8**, 1-18

67 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

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

69 Fraschka, S. A., Filarsky, M., Hoo, R., Niederwieser, I., Yam, X. Y., Brancucci, N. M. B., Mohring, F., Mushunje, A. T., Huang, X., Christensen, P. R., Nosten, F., Bozdech, Z., Russell, B., Moon, R. W., Marti, M., Preiser, P. R., Bártfai, R. and Voss, T. S. (2018) Comparative heterochromatin profiling reveals conserved and unique epigenome signatures linked to adaptation and development of malaria parasites. *Cell Host & Microbe*. **23**, 407-420

70 Bechtsi, D. P. and Waters, A. P. (2017) Genomics and epigenetics of sexual commitment in *Plasmodium*. *International Journal for Parasitology*. **47**, 425-434

71 Reichert, N., Choukrallah, M.-A. and Matthias, P. (2012) Multiple roles of class I HDACs in proliferation, differentiation, and development. *Cellular and Molecular Life Sciences*. **69**, 2173-2187

72 Silverstein, R. A. and Ekwall, K. (2005) Sin3: a flexible regulator of global gene expression and genome stability. *Current Genetics*. **47**, 1-17

73 Hassig, C. A., Fleischer, T. C., Billin, A. N., Schreiber, S. L. and Ayer, D. E. (1997) Histone deacetylase activity is required for full transcriptional repression by mSin3A. *Cell*. **89**, 341-347

74 Laherty, C. D., Yang, W.-M., Sun, J.-M., Davie, J. R., Seto, E. and Eisenman, R. N. (1997) Histone deacetylases associated with the mSin3 corepressor mediate mad transcriptional repression. *Cell*. **89**, 349-356

75 Banks, C. A., Zhang, Y., Miah, S., Hao, Y., Adams, M. K., Wen, Z., Thornton, J. L., Florens, L. and Washburn, M. P. (2020) Integrative modeling of a Sin3/HDAC complex sub-structure. *Cell Reports*. **31**, 107516

76 Brunmeir, R., Lagger, S. and Seiser, C. (2009) Histone deacetylase 1 and 2-controlled embryonic development and cell differentiation. *International Journal of Developmental Biology*. **53**, 275-289

77 Lagger, G., O'Carroll, D., Rembold, M., Khier, H., Tischler, J., Weitzer, G., Schuettengruber, B., Hauser, C., Brunmeir, R. and Jenuwein, T. (2002) Essential function of histone deacetylase 1 in proliferation control and CDK inhibitor repression. *The EMBO Journal*. **21**, 2672-2681

78 Lakowski, B., Roelens, I. and Jacob, S. (2006) CoREST-like complexes regulate chromatin modification and neuronal gene expression. *Journal of Molecular Neuroscience*. **29**, 227-239

79 Delcuve, G. P., Khan, D. H. and Davie, J. R. (2012) Roles of histone deacetylases in epigenetic regulation: emerging paradigms from studies with inhibitors. *Clinical Epigenetics*. **4**, 1-13

80 Haber, J. E. and George, J. P. (1979) A mutation that permits the expression of normally silent copies of mating-type information in *Saccharomyces cerevisiae*. *Genetics*. **93**, 13-35

81 Haigis, M. C. and Sinclair, D. A. (2010) Mammalian sirtuins: biological insights and disease relevance. *Annual Review of Pathology: Mechanisms of Disease*. **5**, 253-295

82 Wierman, M. B. and Smith, J. S. (2014) Yeast sirtuins and the regulation of aging. *FEMS Yeast Research*. **14**, 73-88

83 Buck, S. W., Gallo, C. M. and Smith, J. S. (2004) Diversity in the Sir2 family of protein deacetylases. *Journal of Leukocyte Biology*. **75**, 939-950

84 Straight, A. F., Shou, W., Dowd, G. J., Turck, C. W., Deshaies, R. J., Johnson, A. D. and Moazed, D. (1999) Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. *Cell*. **97**, 245-256

- 85 Sandmeier, J. J., French, S., Osheim, Y., Cheung, W. L., Gallo, C. M., Beyer, A. L. and Smith, J. S. (2002) RPD3 is required for the inactivation of yeast ribosomal DNA genes in stationary phase. *The EMBO Journal*. **21**, 4959-4968
- 86 Religa, A. A. and Waters, A. P. (2012) Sirtuins of parasitic protozoa: in search of function (s). *Molecular and Biochemical Parasitology*. **185**, 71-88
- 87 Spain, M. M., Caruso, J. A., Swaminathan, A. and Pile, L. A. (2010) Drosophila SIN3 isoforms interact with distinct proteins and have unique biological functions. *Journal of Biological Chemistry*. **285**, 27457-27467
- 88 Zhang, Y., Iratni, R., Erdjument-Bromage, H., Tempst, P. and Reinberg, D. (1997) Histone deacetylases and SAP18, a novel polypeptide, are components of a human sin3 complex. *Cell*. **89**, 357-364
- 89 O'Donnell, A. J., Prior, K. F. and Reece, S. E. (2020) Host circadian clocks do not set the schedule for the within-host replication of malaria parasites. *Proceedings of the Royal Society B: Biological Sciences*. **287**, 20200347
- 90 Gu, X., Wang, Y. and He, Y. (2013) Photoperiodic regulation of flowering time through periodic histone deacetylation of the florigen gene FT. *PLoS Biology*. **11**, 1001649
- 91 Lee, H. G., Hong, C. and Seo, P. J. (2019) The *Arabidopsis* sin3-hdac complex facilitates temporal histone deacetylation at the Cca1 and Prr9 loci for robust circadian oscillation. *Frontiers in Plant Science*. **10**, 171
- 92 Pile, L. A., Schlag, E. M. and Wassarman, D. A. (2002) The SIN3/RPD3 deacetylase complex is essential for G2 phase cell cycle progression and regulation of SMRTER corepressor levels. *Molecular and Cellular Biology*. **22**, 4965-4976
- 93 Murachelli, A. G., Ebert, J., Basquin, C., Le Hir, H. and Conti, E. (2012) The structure of the ASAP core complex reveals the existence of a Pinin-containing PSAP complex. *Nature Structural Molecular Biology*. **19**, 378-386
- 94 Schwerk, C., Prasad, J., Degenhardt, K., Erdjument-Bromage, H., White, E., Tempst, P., Kidd, V. J., Manley, J. L., Lahti, J. M. and Reinberg, D. (2003) ASAP, a novel protein complex involved in RNA processing and apoptosis. *Molecular and cellular biology*. **23**, 2981-2990
- 95 McCallum, S. A., Bazan, J. F., Merchant, M., Yin, J., Pan, B., de Sauvage, F. J. and Fairbrother, W. J. (2006) Structure of SAP18: a ubiquitin fold in histone deacetylase complex assembly. *Biochemistry*. **45**, 11974-11982
- 96 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
- 97 Merrick, C. J. and Duraisingh, M. T. (2007) *Plasmodium falciparum* Sir2: An unusual sirtuin with dual histone deacetylase and ADP-ribosyltransferase activity. *Eukaryotic Cell*. **6**, 2081-2091
- 98 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. and Scherf, A. (2005) Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell*. **121**, 25-36
- 99 Gotta, M., Laroche, T., Formenton, A., Maillet, L., Scherthan, H. and Gasser, S. M. (1996) The clustering of telomeres and colocalization with Rap1, Sir3, and Sir4 proteins in wild-type *Saccharomyces cerevisiae*. *Journal of Cell Biology*. **134**, 1349-1363
- 100 Chakrabarty, S. P. and Balaram, H. (2010) Reversible binding of zinc in *Plasmodium falciparum* Sir2: structure and activity of the apoenzyme. *Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics*. **1804**, 1743-1750
- 101 Rawat, M., Srivastava, A., Gupta, I. and Karmodiya, K. (2021) Single cell RNA-sequencing reveals cellular heterogeneity, stage transition and antigenic variation during stress adaptation in synchronized *Plasmodium falciparum*. *Microbiology Spectrum*. **9**, 0000821
- 102 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., Llinás, M., Buckee, C. O. and Duraisingh, M. T. (2014) A *Plasmodium falciparum* histone deacetylase regulates antigenic variation and gametocyte conversion. *Cell Host & Microbe*. **16**, 177-186
- 103 North, B. J. and Verdin, E. (2004) Sirtuins: Sir2-related NAD-dependent protein deacetylases. *Genome Biology*. **5**, 1-12
- 104 Tonkin, C. J., Carret, C. K., Duraisingh, M. T., Voss, T. S., Ralph, S. A., Hommel, M., Duffy, M. F., Silva, L. M., Scherf, A., Ivens, A., Speed, T. P., Beeson, J. G. and Cowman, A. F. (2009) Sir2 paralogues cooperate to regulate virulence genes and antigenic variation in *Plasmodium falciparum*. *PLoS Biology*. **7**, 1000084

- 105 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*. *Molecular and Biochemical Parasitology*. **99**, 11-19
- 106 Huang, Z., Li, R., Tang, T., Ling, D., Wang, M., Xu, D., Sun, M., Zheng, L., Zhu, F. and Min, H. (2020) A novel multistage antiplasmodial inhibitor targeting *Plasmodium falciparum* histone deacetylase 1. *Cell Discovery*. **6**, 1-15
- 107 Engel, J. A., Norris, E. L., Gilson, P., Przyborski, J., Shonhai, A., Blatch, G. L., Skinner-Adams, T. S., Gorman, J., Headlam, M. and Andrews, K. T. (2019) Proteomic analysis of *Plasmodium falciparum* histone deacetylase 1 complex proteins. *Experimental Parasitology*. **198**, 7-16
- 108 Duffy, M. F., Selvarajah, S. A., Josling, G. A. and Petter, M. (2012) The role of chromatin in *Plasmodium* gene expression. *Cellular Microbiology*. **14**, 819-828
- 109 Van Biljon, R., Van Wyk, R., Painter, H. J., Orchard, L., Reader, J., Niemand, J., Llinás, M. and Birkholtz, L.-M. (2019) Hierarchical transcriptional control regulates *Plasmodium falciparum* sexual differentiation. *BMC Genomics*. **20**, 1-16
- 110 Coetzee, N., Sidoli, S., van Biljon, R., Painter, H., Llinás, M., Garcia, B. A. and Birkholtz, L.-M. (2017) Quantitative chromatin proteomics reveals a dynamic histone post-translational modification landscape that defines asexual and sexual *Plasmodium falciparum* parasites. *Scientific Reports*. **7**, 607-607
- 111 Zhang, M., Wang, C., Otto, T. D., Oberstaller, J., Liao, X., Adapa, S. R., Udenze, K., Bronner, I. F., Casandra, D. and Mayho, M. (2018) Uncovering the essential genes of the human malaria parasite *Plasmodium falciparum* by saturation mutagenesis. *Science*. **360**, 7847
- 112 Mancio-Silva, L., Lopez-Rubio, J. J., Claes, A. and Scherf, A. (2013) Sir2a regulates rDNA transcription and multiplication rate in the human malaria parasite *Plasmodium falciparum*. *Nature Communications*. **4**, 1-6
- 113 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. *Molecular Microbiology*. **88**, 687-701
- 114 Knuepfer, E., Napiorkowska, M., Van Ooij, C. and Holder, A. A. (2017) Generating conditional gene knockouts in *Plasmodium*—a toolkit to produce stable DiCre recombinase-expressing parasite lines using CRISPR/Cas9. *Scientific Reports*. **7**, 1-12
- 115 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**, 73783
- 116 Ganesan, S. M., Falla, A., Goldfless, S. J., Nasamu, A. S. and Niles, J. C. (2016) Synthetic RNA–protein modules integrated with native translation mechanisms to control gene expression in malaria parasites. *Nature Communications*. **7**, 1-10
- 117 Belmont, B. J. and Niles, J. C. (2010) Engineering a direct and inducible protein– RNA interaction to regulate RNA biology. *ACS Chemical Biology*. **5**, 851-861
- 118 Armstrong, C. M. and Goldberg, D. E. (2007) An FKBP destabilization domain modulates protein levels in *Plasmodium falciparum*. *Nature Methods*. **4**, 1007-1009
- 119 Chaudhry, S., Lwin, N., Phelan, D., Escalante, A. and Battistuzzi, F. (2018) Comparative analysis of low complexity regions in *Plasmodia*. *Scientific Reports*. **8**, 1-9
- 120 Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J. and von Mering, C. (2019) STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*. **47**, 607–613
- 121 Arumugam, J. H. W., Mohammad, A., Mei Fang, S., Hong May, S., Chandra, S. V., David, P. L. and Prakash. (2017) A yeast two-hybrid system for the screening and characterization of small-molecule inhibitors of protein–protein interactions identifies a novel putative Mdm2-binding site in p53. *BMC Biology*. **15**, 1-17
- 122 Clark, D. P., Pazdernik, N. J. and McGehee, M. R. (2019) Chapter 15 - Proteomics: The Global Analysis of Proteins. In *Molecular Biology* (Clark, D. P., Pazdernik, N. J. and McGehee, M. R., eds.). pp. 484-520, Academic Cell
- 123 Reader, J., Botha, M., Theron, A., Lauterbach, S. B., Rossouw, C., Engelbrecht, D., Wepener, M., Smit, A., Leroy, D., Mancama, D., Coetzer, T. L. and Birkholtz, L.-M. (2015) Nowhere to hide: interrogating different metabolic parameters of *Plasmodium falciparum* gametocytes in a transmission blocking drug discovery pipeline towards malaria elimination. *Malaria Journal*. **14**, 1-17

- 124 Gupta, S. K., Schulman, S. and Vanderberg, J. P. (1985) Stage-dependent toxicity of N-acetylglucosamine to *Plasmodium falciparum*. The Journal of Protozoology. **32**, 91-95
- 125 Birnbaum, J., Flemming, S., Reichard, N., Soares, A. B., Mesén-Ramírez, P., Jonscher, E., Bergmann, B. and Spielmann, T. (2017) A genetic system to study *Plasmodium falciparum* protein function. Nature Methods. **14**, 450-456
- 126 Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J. and Schwede, T. (2009) Protein structure homology modeling using SWISS-MODEL workspace. Nature Protocols. **4**, 1-13
- 127 Kufareva, I. and Abagyan, R. (2012) Methods of protein structure comparison. Methods Molecular Biology. **857**, 231-257
- 128 Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A. and Potapenko, A. (2021) Highly accurate protein structure prediction with AlphaFold. Nature, 583–589
- 129 LaCount, D. J., Vignali, M., Chettier, R., Phansalkar, A., Bell, R., Hesselberth, J. R., Schoenfeld, L. W., Ota, I., Sahasrabudhe, S. and Kurschner, C. (2005) A protein interaction network of the malaria parasite *Plasmodium falciparum*. Nature. **438**, 103-107
- 130 Sorber, K., Dimon, M. T. and DeRisi, J. L. (2011) RNA-Seq analysis of splicing in *Plasmodium falciparum* uncovers new splice junctions, alternative splicing and splicing of antisense transcripts. Nucleic Acids Research. **39**, 3820-3835
- 131 Tange, T., Shibuya, T., Jurica, M. S. and Moore, M. J. (2005) Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. RNA. **11**, 1869-1883
- 132 Zhu, A. Y., Zhou, Y., Khan, S., Deitsch, K. W., Hao, Q. and Lin, H. (2012) *Plasmodium falciparum* Sir2A preferentially hydrolyzes medium and long chain fatty acyl lysine. ACS Chemical Biology. **7**, 155-159
- 133 Hoeijmakers, W. A. M., Miao, J., Schmidt, S., Toenhake, C. G., Shrestha, S., Venhuizen, J., Henderson, R., Birnbaum, J., Ghidelli-Disse, S. and Drewes, G. (2019) Epigenetic reader complexes of the human malaria parasite, *Plasmodium falciparum*. Nucleic acids research. **47**, 11574-11588
- 134 D., O. T., Wilinski, D., Assefa, S., Keane, T. M., Sarry, L. R., Böhme, U., Lemieux, J., Barrell, B., Pain, A., Berriman, M., Newbold, C. and Llinás. (2010) New insights into the blood-stage transcriptome of *Plasmodium falciparum* using RNA-Seq. Molecular Microbiology. **76**, 12-24
- 135 Hoeijmakers, W. A., Bártfai, R. and Stunnenberg, H. G. (2012) Transcriptome analysis using RNA-Seq. In Malaria. pp. 221-239, Springer
- 136 Das Gupta, R., Krause-Ihle, T., Bergmann, B. r., Müller, I. B., Khomutov, A. R., Müller, S., Walter, R. D. and Lüersen, K. (2005) 3-Aminoxy-1-aminopropane and derivatives have an antiproliferative effect on cultured *Plasmodium falciparum* by decreasing intracellular polyamine concentrations. Antimicrobial Agents and Chemotherapy. **49**, 2857-2864
- 137 Singh, K. K., Erkelenz, S., Rattay, S., Dehof, A. K., Hildebrandt, A., Schulze-Osthoff, K., Schaal, H. and Schwerk, C. (2010) Human SAP18 mediates assembly of a splicing regulatory multiprotein complex via its ubiquitin-like fold. RNA. **16**, 2442-2454
- 138 Gomes, A. R., Bushell, E., Schwach, F., Girling, G., Anar, B., Quail, M. A., Herd, C., Pfander, C., Modrzynska, K. and Rayner, J. C. (2015) A genome-scale vector resource enables high-throughput reverse genetic screening in a malaria parasite. Cell Host & Microbe. **17**, 404-413

Supplementary information

Figure S1 shows the entire agarose gel of Figure 3.14 A and B in the main text.

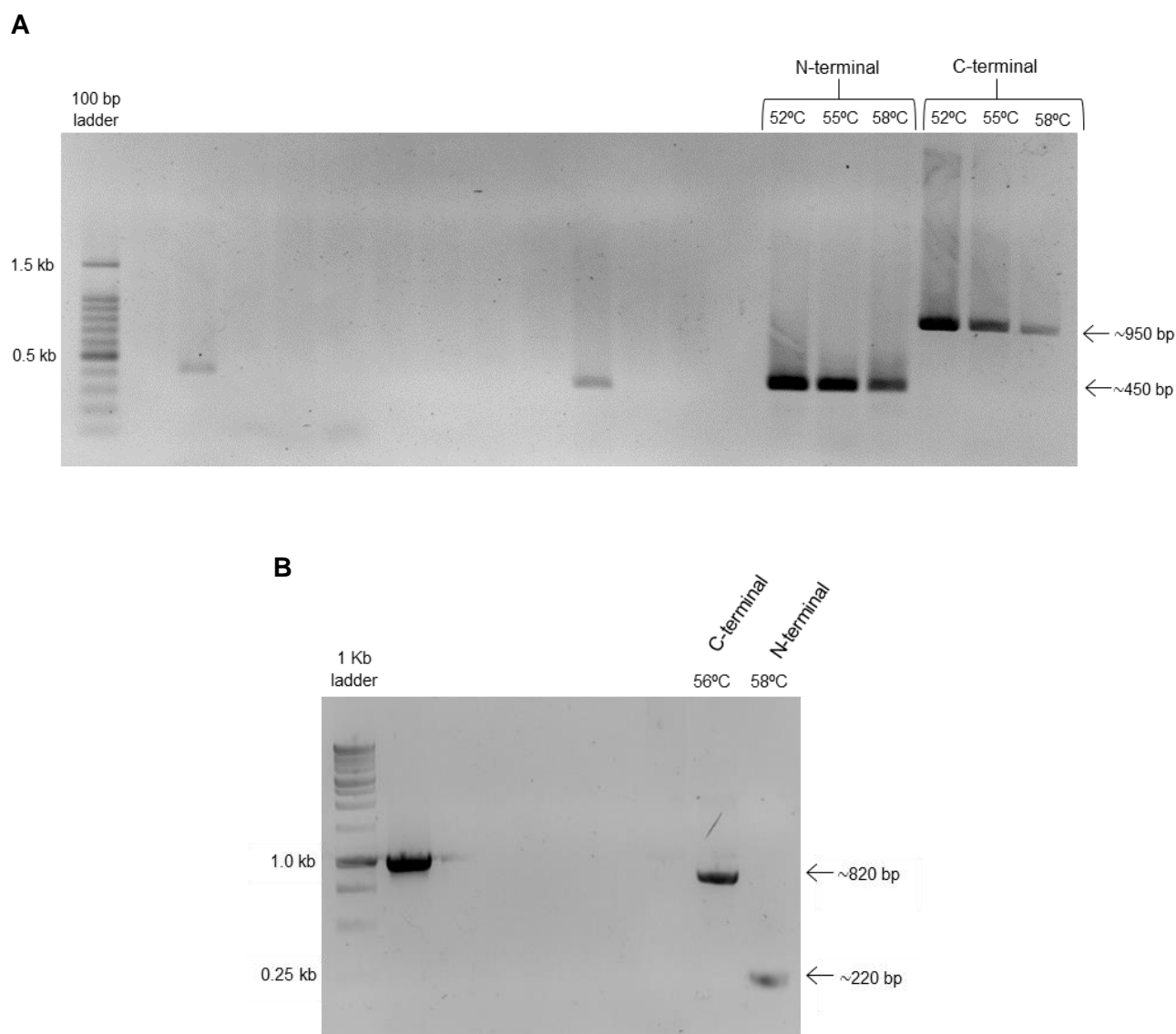


Figure S1: PCR amplification of gene specific homology regions. Homology region amplification of N- and C-terminal PF3D7_0711400 and PF3D7_1328800 gene fragments from NF54 parasites. PCR conditions had an annealing temperature of **(A)** 52°C, 55°C and 58°C or **(B)** 56°C, 58°C. The expected amplicon band sizes were obtained: 451 bp (PF3D7_0711400, N-terminal), 954 bp (PF3D7_0711400, C-terminal), 217 bp (PF3D7_1328800, C-terminal), 822 bp (PF3D7_1328800, C-terminal). DNA fragments, 1 kb and 100 bp DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light. Other bands were not used for this work.

Figure S2 shows the entire agarose gel of Figure 3.15 A in the main text.

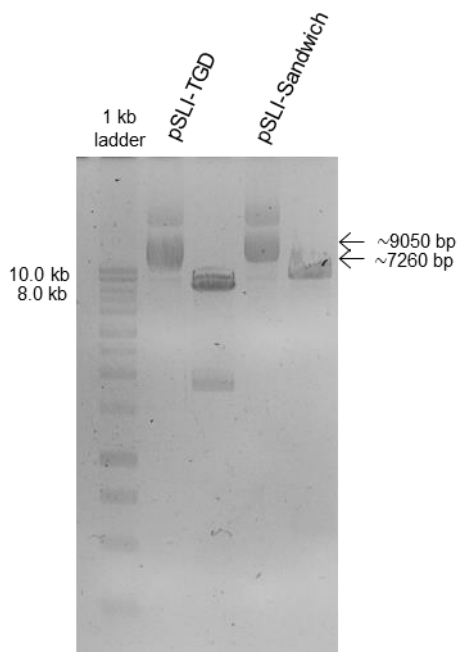


Figure S2: Restriction enzyme digestion of pSLI-TGD and pSLI-Sandwich plasmids. (A) Uncut SLI-TGD and SLI-Sandwich plasmids isolated from *E.coli* DH5 α to produce 9050 bp for SLI-Sandwich and a 7260 bp for SLI-TGD. Plasmids were visualised with 1 % (w/v) agarose/TAE gel with ethidium bromide under UV light. Other bands were not used for this work.

Figure S3 shows the entire agarose gel of Figure 3.17 A in the main text.

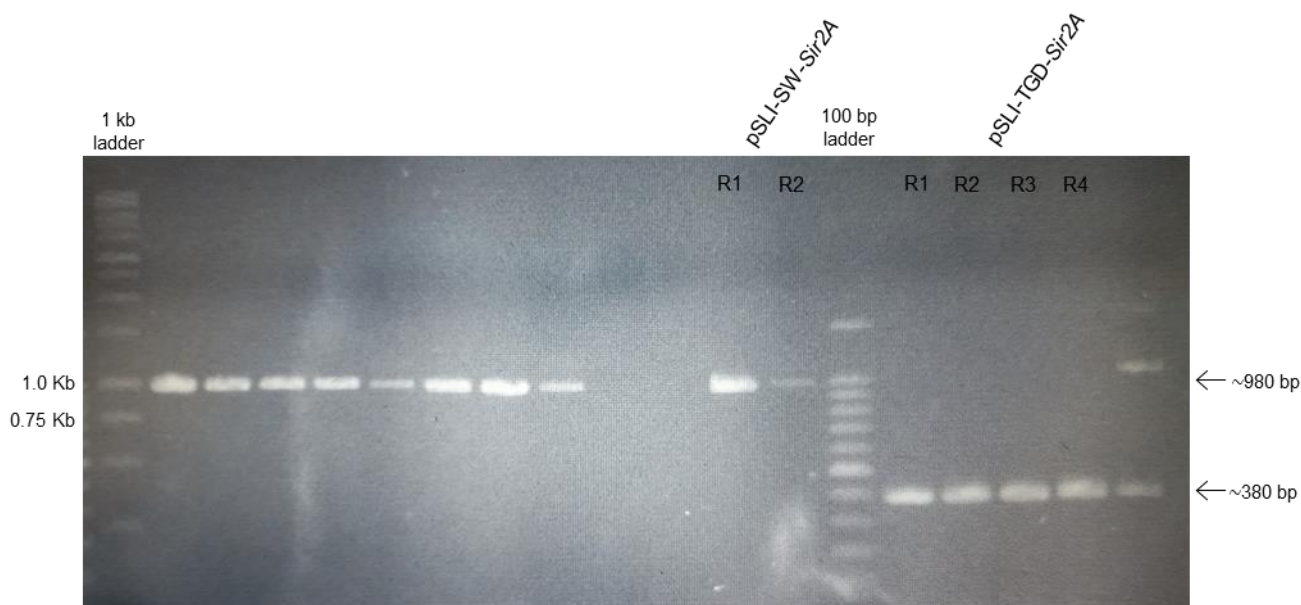


Figure S3: Validation of of pSLI-TGD-*Sir2A* and pSLI-SW-*Sir2A* recombinant plasmids. Colony screening PCR of *sir2A* N- and C- terminal regions cloned into their respective plasmids using primers shown in Table 2.1. (R indicates bacterial clone replicates). The expected region band size was obtained 430 bp (pSLI-SW-*Sir2A*) and 977 bp (pSLI-SW-*Sir2A*). DNA fragments, 1 kb and 100 bp DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light. Other bands were not used for this work.

Figure S4 shows the entire gel of Figure 3.19 B and D in the main text

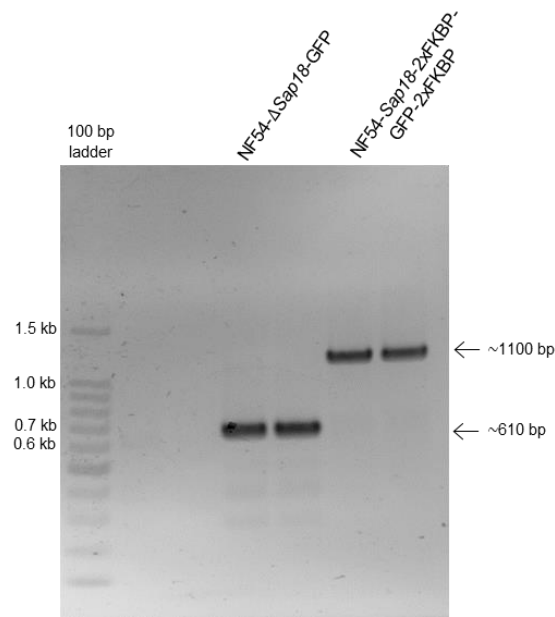


Figure S4: PCR analysis of pSLI-TGD-*Sap18* and pSLI-SW-*Sap18* for episomal uptake into NF54 *P. falciparum* parasites. PCR amplification of pSLI-TGD-*Sap18* (613 bp) using and pSLI-SW-*Sap18* plasmid backbones (1097 bp) using primers listed in Table 2.2. DNA fragments and Promega 100 bp DNA ladder were visualised with a 1.5 % (w/v) agarose/TAE gel with ethidium bromide under UV light.

Figure S5 shows integration attempts of NF54-*Sir2A*-2xFKBP-GFP-2xFKBP parasites

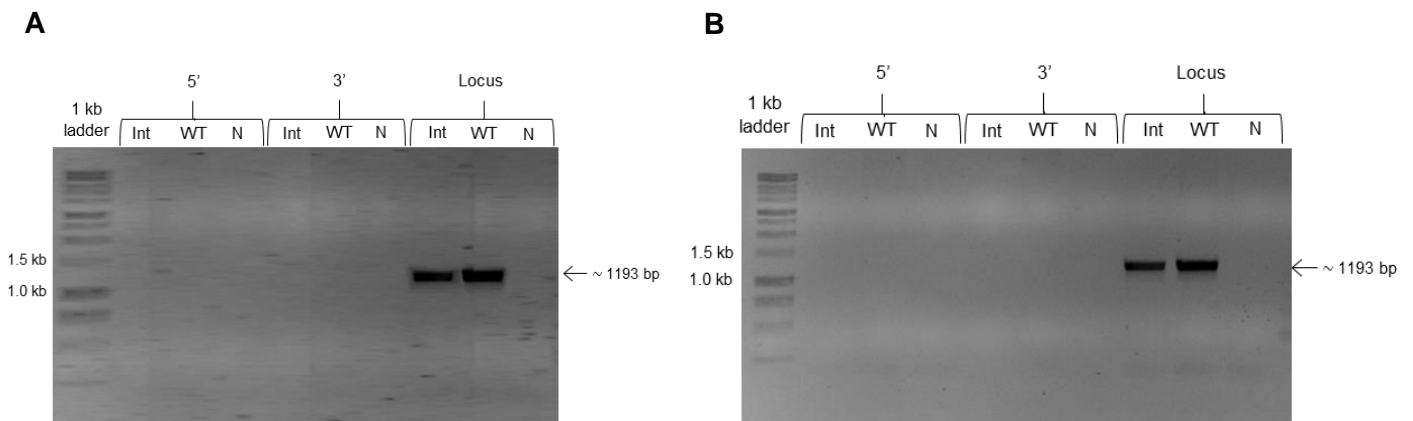


Figure S5: PCR analysis attempts of SLI-Sandwich integration into the *sir2A* locus of *P. falciparum* parasites. (A) and (B) NF54-*Sir2A*-2xFKBP-GFP-2xFKBP parasites were PCR screened for the presence of 5' and 3' integration using primers shown in Table 2.4. The absence of 5' and 3' integration bands indicate no integration. The presence of the original locus in recombinant parasites denotes wild type parasites were present (1193 bp). NF54 wild type parasites (control) indicated the presence of the original locus (1193 bp). PCR products were visualised with 1.5 % (w/v) agarose/TAE gel stained with ethidium bromide under UV light. Int=integrated parasite line, WT= NF54 wild type control, N= negative control.

Figure S6 shows the entire western blot for Figure 3.27 in the main text

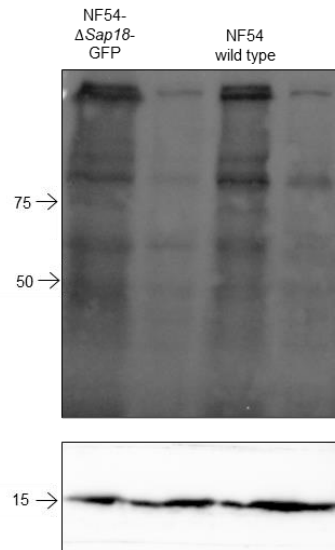


Figure S6: Western blot analysis of GFP expression of NF54- Δ Sap18-GFP parasites compared to NF54 wild type parasites probed with anti-GFP antibody. Expected protein size of ~61 kDa. Histone H3 was used as the control for equal protein loading. Precision Plus Protein Dual Colour ladder was used to compare protein size. Other bands were not used for this work.