UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Inhibition of the histone deacetylase family as a drug target in the human malaria parasite, *Plasmodium falciparum*

Nabila Ismail
s28264828

Submitted in partial fulfilment of the degree:

Magister Scientiae Biochemistry

In the Faculty of Natural and Agricultural Sciences

Department of Biochemistry
University of Pretoria

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Submission declaration

I declare that the thesis/dissertation which is herewith submitted for the degree *Magister Scientiae* Biochemistry at the University of Pretoria is my own work and has not previously been submitted by them for a degree at this or any other tertiary institution.

Signed: _____

Nabila Ismail

Date: _____

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Plagiarism declaration

University of Pretoria
Faculty of Natural and Agricultural Sciences
Department of Biochemistry

Full name: **Nabila Ismail** Student number: **s28264828**
Title of work: **Inhibition of the histone deacetylase family as a drug target in the human malaria parasite, *Plasmodium falciparum***

Declaration:

1. I understand what plagiarism entails and am aware of the University's policy in this regard.
2. I declare that this thesis is my own, original work. Where someone else's work was used (whether from a printed source, the internet or any other source due acknowledgement was given and reference was made according to departmental requirements.
3. I did not make use of another student's previous work and submit it as my own.
4. I did not allow and will not allow anyone to copy my work with the intention of presenting it as his or her own work.

Signature:_____ Date:_____

Acknowledgements

In the name of God, the Beneficent, the Merciful

First and foremost, I am grateful to the Almighty, for giving me the opportunity to complete this degree. His guidance and grace has helped me to be all that I am today and accomplish that, which I thought was not possible. It is said that one who is not thankful to man, cannot be thankful to God. Therefore, I would like to begin by thanking all those who have made this work possible, either directly or indirectly.

To my family; especially my parents (Talib and Shameema Ismail), for their continued support and encouragement, I am grateful. To my sisters (Zakiya, Sadika and Fadila) and their husbands (Irfaan and Zameer, respectively), I appreciate your faith in me. To my dearest friends, Marna, Bernice, Bianca, Johann and Tasneem who have made this degree seem achievable when I thought it wasn't.

To all those who have contributed to this work and provided the necessary support and expertise, you have been extremely kind to me. Here, I would like to mention Prof. M. Jung (University of Freiburg; for provision of the HDAC substrate) and Dr. Kristina Keller (from the same lab, for enquiries regarding HDAC assays), Dr. P. Burger (for collaboration with the *in silico* work), Dr. S. Stoychev (for queries regarding MS), Dr. S. Smit (for carrying out the MS analysis at the CAF proteomics facility), Dr. J. Niemand (for advice regarding assay development) and Dr. N. October (for chemical synthesis). Also, the members of the malaria research group and the Department of Biochemistry.

Last, but definitely not least, to my mentors, Prof. Birkholtz and Prof Louw, thank you for seeing me as being capable and helping me to carry out this degree to completion. I am grateful for your advice, motivation and your continuous struggle to help me achieve my fullest potential. Your support on this journey has meant a great deal to me.

I pray that God blesses you all for your kindness, support and the valuable time that you have sacrificed.

I am grateful to the NRF for funding of this degree through the NRF Master's Innovation bursary as well as for partial funding to attend the GRC Malaria conference. Here, thanks also go to MVI Path for their additional funding. For attendance to the MIM conference, I am grateful to the UPCSMC, for financial assistance as well as for providing a platform for interaction with other malaria researchers in South Africa.

Summary

The asexual life cycle of *Plasmodium falciparum* parasites takes only 48 hours, allowing for rapid replication. The continuous infection, rupturing and re-infection of erythrocytes results in the pathogenicity of this disease. Schizogony (nuclear division) in *P. falciparum* parasites occurs via alternation between the S and M phases of the cell cycle where DNA synthesis occurs in the mature trophozoite and schizont stages, followed by mitosis to form daughter merozoites. Merozoites then give rise to ring stages after they have infected erythrocytes and the ring stages continue their development to trophozoites. This cyclic development, known as the intra-erythrocytic developmental cycle, has a unique transcriptional regulation, which is closely linked to cell cycle regulation. However, the intricacies that these mechanisms are controlled by are still unidentified. One of the means in which the *P. falciparum* parasite's complex life cycle is controlled is by means of epigenetics. Epigenetics refers to the heritable changes on a phenotypic level, which are independent of changes on a genetic level.

One group of enzymes that participates in the parasite's epigenetic control is the *Plasmodium* histone deacetylases. Inhibition of histone deacetylases (HDACs) results in hyperacetylation, which causes aberrant gene transcription and eventually results in parasite death. Comparative analyses of three histone isolation methods and analysis of *P. falciparum* parasite histones and their post-translational modifications (PTMs) by mass spectrometry techniques identified both epigenetically relevant and novel PTMs in *P. falciparum* parasite histones and led to the discovery of an adapted histone isolation method for investigation of histone PTM landscapes. When this modified method was used for the investigation of histones that were isolated from *P. falciparum* parasites treated with HDAC inhibitors compared to untreated parasites, differences were seen in the PTM landscape.

Subsequent *in silico* screening strategies were used to identify ten compounds from the Medicines for Malaria Venture (MMV) Malaria Box, which target the active site of the zinc-requiring *Pf*HDAC1. From these, eight compounds showed inhibition of proliferation of cultured *P. falciparum* parasites. Ensuing, the adaptation of an HDAC assay to investigate histone deacetylase inhibition was used to validate these compounds as possible PfHDAC1 inhibitors, with at least two of the compounds showing significant inhibition of PfHDAC1 activity, comparable to that of the known HDAC inhibitor, suberoylanilide hydroxamic acid, SAHA. The use of *in silico* screening of a large library of compounds, such as the MMV Malaria Box, successfully narrows down candidates for possible anti-malarials with drug-like properties by identification of their cellular targets. This work is method-based and facilitates the investigation of the epigenetic landscape of histones, and the identification of novel HDAC inhibitors.

Table of Contents

Table of figures

List of Tables

Abbreviations

ABHA: azelaic bishydroxamic acid

ac: acetylation

AMC: 7-amino-4-methylcoumarin

BCA: bicinchoninic acid

CAF: Central Analytical Facility

CDK: cyclin dependent kinase

CQ: chloroquine

HAT: histone acetyl transferase

HDAC: histone deacetylases

H2A: histone H2A

H2A.Z: histone H2A.Z (apicomplexan specific)

H2B: histone H2B

H2B.Z: histone H2B.Z also known as H2Bv

H3: histone H3 also known as H3.1

H3K9Ac: Histone H3, lysine at position 9, acetylation

H3K9me3: Histone H3, lysine at position 9, trimethylation

H3.3: histone H3.3

H3 cen: histone H3 centromeric

H4: histone H4

HMGB: high mobility group box

IDC: intra-erythrocytic developmental cycle

logP: log of the partition coefficient

HMT: histone methyl transferase

me: monomethylation

me2: dimethylation

me3: trimethylation

MMV: Medicines for Malaria Venture

MS: Mass spectrometry

PfHDAC: *P. falciparum* histone deacetylase

PfMYST: *P. falciparum* MYST (MOZ, Ybf1/Sas3, Sas2 and Tip60)

PfSir: *P. falciparum* sirtuin

ph: phosphorylation

PHD: plant homeodomain

PI: protease inhibitor

PTM: post translational modification

SDS-PAGE: polyacrylamide gel electrophoresis

RFU: relative fluorescence unit

Ro5: Lipinski's Rule of 5

SAHA: suberoylanilide hydroxamic acid

SBHA: suberohydroxamic acid

Sir: sirtuin (A class III histone deacetylase)

TCA: trichloroacetic acid

TSA: trichostatin A

Z-MAL: (*S*)-[5-Acetylamino-1-(4-methyl-2-oxo-2H-chromen-7-ylcarbamoyl)-pentyl]-carbamic acid benzyl ester

Chapter 1

Introduction

Malaria is caused by the transmission of a protozoan parasite from a female *Anopheles* mosquito (i.e. the malaria vector) to humans after a blood meal. These protozoan parasites belong to the genus *Plasmodium*, of which five species cause malaria, these are: *P. vivax, P. ovale, P. falciparum, P. malariae* and *P. knowlesi* (zoonotic transfer) [1]. *P. falciparum* causes the most severe form of malaria and is regarded as extremely virulent due to the severity of clinical symptoms it causes including anaemia, kidney and multi-organ failure as well as cerebral malaria [2]. The majority of malaria cases and deaths (95%) occur on the African continent due to tropical and sub-tropical climates [3], which are ideal for survival of the malaria vector. Major efforts have been made by international consortia and partnerships such as the ***M***edicines for ***M***alaria ***V***enture (MMV, www.mmv.org), the ***C***o-ordination, ***R***ationalisation, and ***I***ntegration of anti-***mal***arial ***D***rug ***D***iscovery & development ***I***nitiatives (CRIMALDDI, www.crimalddi.eu) and the ***Mal***aria ***E***radication ***R***esearch ***A***genda (malERA, malera.tropika.net) to reduce the global burden of malaria through prioritising objectives, which need to be accomplished and mobilising resources to help realise them [4].

The elimination of malaria requires an interdisciplinary approach for effective vector and parasite control. Vector control generally comprises of the distribution of insecticide-treated bed nets or wall linings and spraying of insecticides in malaria endemic areas [4]. Parasite control on the other hand requires diagnostics (evaluation of the infected individual and the type of malaria contracted) prior to severe disease onset, followed by treatment with an effective antimalarial [4]. Thus, poor health care facilities, irregular follow up of control treatments, vector resistance to insecticides, climate change and the emergence of parasite resistance to numerous anti-malarials negatively contribute to the burden of malaria despite massive control efforts [5]. Parasite control requires a well-populated anti-malarial pipeline since the ability of the parasite to rapidly mutate, which results in drug resistance, has been seen with numerous drugs such as chloroquine, the anti-folates (sulphadoxine-pyrimethamine) and recently to artemisinin derivatives in combination therapies [6]. The key to finding novel anti-malarials thus lies in investigation of drugs from chemically distinct classes, with limited toxicity and unique modes of action [7]. One of the foci of current anti-malarial drug research has been aimed against the parasite cell cycle and epigenetic aspects in order to perturb gene expression, ultimately causing cell death [8].

1.1 Cell cycle of *P. falciparum* parasites in relation to their life cycle

In eukaryotes, cell division occurs in a concise, ordered way to produce daughter cells that are identical to their parent cells and possess all the biochemical molecules required for growth and proliferation [9]. The four major cell cycle phases that dividing cells undergo are: gap phase 1 (G_1), synthesis phase (S-phase), gap phase 2 (G_2) and mitotic phase (M-phase) [10]. An additional quiescent state (that requires a stimuli for cells to move out of) is known as the G_0 phase [10]. Replicating somatic cells grow in size and synthesise proteins required for DNA synthesis during G_1-phase, after which they enter S-phase where they replicate their chromosomes [10]. This is followed by a second gap phase then M-phase, which is divided into several stages. As unicellular eukaryotes, *P. falciparum* parasites undergo complex cell and developmental cycles (Figure 1.1).

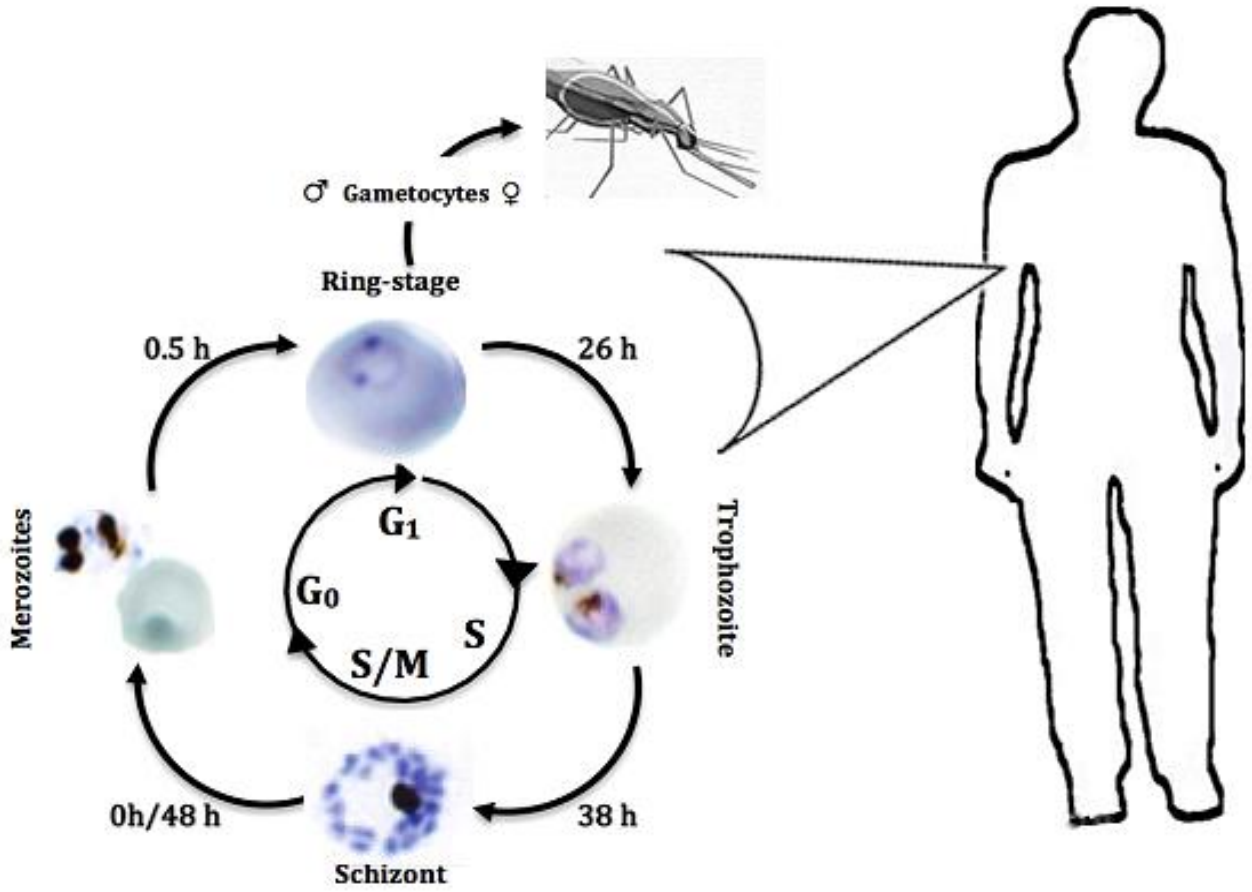


Figure 1.1: *P. falciparum* parasite's life and cell cycles. Intra-erythrocytic development begins when merozoites (G_0) infect erythrocytes and develop into ring stages. These ring stages develop into early trophozoites, with both stages in the G_1 phase (transition from G_{1a} to G_{1b}). Trophozoites mature and DNA-synthesis begins (S-phase). The progression to the schizont stage occurs when the trophozoite nucleus undergoes mitosis. The schizont stage is characterised by multiple asynchronous nuclear divisions to form a multi-nucleated schizont. Rupturing of schizonts releases many merozoites, which can infect new erythrocytes. Alternatively, merozoites, which are programmed for sexual development, can infect erythrocytes and form sexually committed ring stages, which undergo transition to the sexual stages (i.e. male and female gametocytes). A female *Anopheles* mosquito can take up these gametocytes, where progression through the sexual life cycle will give rise to new asexual stages, which can continue the parasite's asexual life cycle. [Figure compiled from [11-13], image of mosquito obtained from National Institute of Allergy and Infectious Diseases (NIAID)].

The life cycle of the parasite begins when a female *Anopheles* mosquito transmits sporozoite forms of the parasite from its salivary glands into the blood stream of a human host during a blood meal [14]. Sporozoites travel to the liver, infect hepatocytes and undergo hepatocytic schizogony (nuclear division) producing up to 40 000 merozoites [11]. These merozoites subsequently infect erythrocytes where they replicate asexually from ring forms to metabolically active trophozoites to multi-nucleated schizonts giving rise to ~32 daughter merozoites. Merozoites released from ruptured schizonts can then infect previously uninfected erythrocytes (Figure 1.1). Intra-erythrocytic development is crucial to the pathogenicity of this parasite [11, 15].

P. falciparum parasites do not undergo classical progression through the cell cycle as is seen with other eukaryotes. Chromatin organisation has been found to alter with different stages in the asexual life cycle. It is understood that infecting merozoites in the G_0-phase with condensed chromatin [11, 16, 17] develop from ring forms to early trophozoites, which is similar to the G_{1a} to G_{1b} transition seen in mammalian cells and is characterised by the de-condensation of chromatin (Figure 1.1) [18]. The early trophozoite stages show large increases in RNA and protein levels [19] followed by DNA synthesis, which begins at the late trophozoite stages. These mature trophozoites are now progressing from the G_1- to S-phase on their developmental path to the schizont stage [11, 15]. Mitosis then occurs as the first step in schizont development and involves division of the trophozoite nucleus [13].

Nuclear divisions in *P. falciparum* parasites are asynchronous; division does not occur simultaneously nor the same number of times in each schizont and mitosis and DNA synthesis are variable [16]. Asynchronous schizogony thus produces merozoites through alternating rounds of DNA synthesis and mitosis (S/M phases) [12, 13], which are released after cytokinesis [20, 21]. Sexually committed merozoites (programmed to enter the sexual stage) infect erythrocytes and form sexually committed ring stages [22] (Figure 1.1) [11, 15, 16]. These ring stages then undergo transition from asexual reproduction to sexual reproduction (Figure 1.1), when they develop into male and female gametocytes that circulate the blood stream and are transferred to another mosquito during a blood meal (Figure 1.1). Once ingested by a mosquito, these haploid gametocytes fuse and mature to form a diploid ookinete, which migrates to the haemocele of the mosquito where it matures into an oocyst [14]. Sporozoites that are formed within the oocyst are released and travel to the salivary glands of the mosquito, allowing for continuation of the cycle.

The progression through both cell and life cycles needs to be regulated, with many forms of cell cycle regulation occurring in the parasite. One of the key regulatory factors are the cyclins, which are synthesised at specific cell cycle phases where they activate cyclin dependent kinases (CDKs), after which these cyclins are degraded prior to the following cell cycle phase [23]. CDKs phosphorylate target proteins, which carry out cell cycle related processes [23]. Through these

signalling pathways, cell cycle control is achieved and prevents deleterious events from occurring due to the presence of certain checkpoints, which must be satisfied [23]. In higher eukaryotes, progression through the M-phase involves a number of mitotic kinase families such as cyclin-dependent kinases, Polo-like kinases, Nima-related kinases and Aurora kinases [24]. Bioinformatic analysis aimed at identifying homologues of cyclins and CDKs in *P. falciparum* parasites found 85-99 genes encoding for kinase-related enzymes [24]. Four genes corresponding to the NIMA-related (***N***ever ***I***n ***M***itosis ***A***) and three genes to the Aurora protein families have been identified. Additionally, a small cluster of genes, which seem to belong to the Polo-like family of proteins were also identified using phylogenetic analysis [24]. Among the several CDK-like proteins identified, the most prominent are Pfmrk, PfPK5 and PfPK6 [25]. An example of one of these, PfPK5, has increased expression and activity around 36 h post invasion and co-localise with DNA. Inhibition of this kinase resulted in reduced DNA synthesis [25]. Although their biochemical characterisation is still incomplete, few putative cyclins have been identified in *P. falciparum* parasites [25].

Another key factor in cell cycle regulation is the regulation of gene expression, as different life stages are dependent on the expression of particular genes for development into subsequent stages [26]. The activation and silencing of certain genes also takes place when parasites differentiate into different morphological and sexual forms [26]. Furthermore, gene expression itself is regulated in different ways including regulation at transcriptional and translational levels [27].

Plasmodium gene expression is thought to occur with a “just-in-time” mechanism, where genes are expressed as they are needed during the different life cycle stages, specifically during asexual development [28, 29], implying strict transcriptional control [30]. This regulation of expression also extends to parasite adaptation to perturbations, and the parasite is capable of naturally selecting transcriptional patterns, which are conducive to its ultimate fitness [31]. The poor correlation between mRNA accumulation and subsequent translation [27] and an apparent paucity of transcription factors compared to other eukaryotes [32, 33] is a possible indication that mechanisms other than transcriptional control alone exist for gene expression control in the parasite. Epigenetics has been found to play an important role in the regulation of *P. falciparum* parasite gene expression [26, 34, 35] and refers to heritable changes in gene expression, which are independent of alterations in DNA sequences [36].

1.2 Epigenetics in *P. falciparum* parasites

A number of epigenetic mechanisms in eukaryotes contribute to the regulation of gene expression. These include DNA methylation, chromatin remodelling (changes from facultative heterochromatin to active euchromatin), transcription factors (*trans* regulatory elements) and non-coding RNA [37]. Epigenetic regulatory mechanisms in the malaria parasite, however, show some deviations from

the accepted eukaryotic mechanisms. An overview of epigenetic regulation and its effect on chromatin structure in *P. falciparum* parasites can be seen in Figure 1.2.

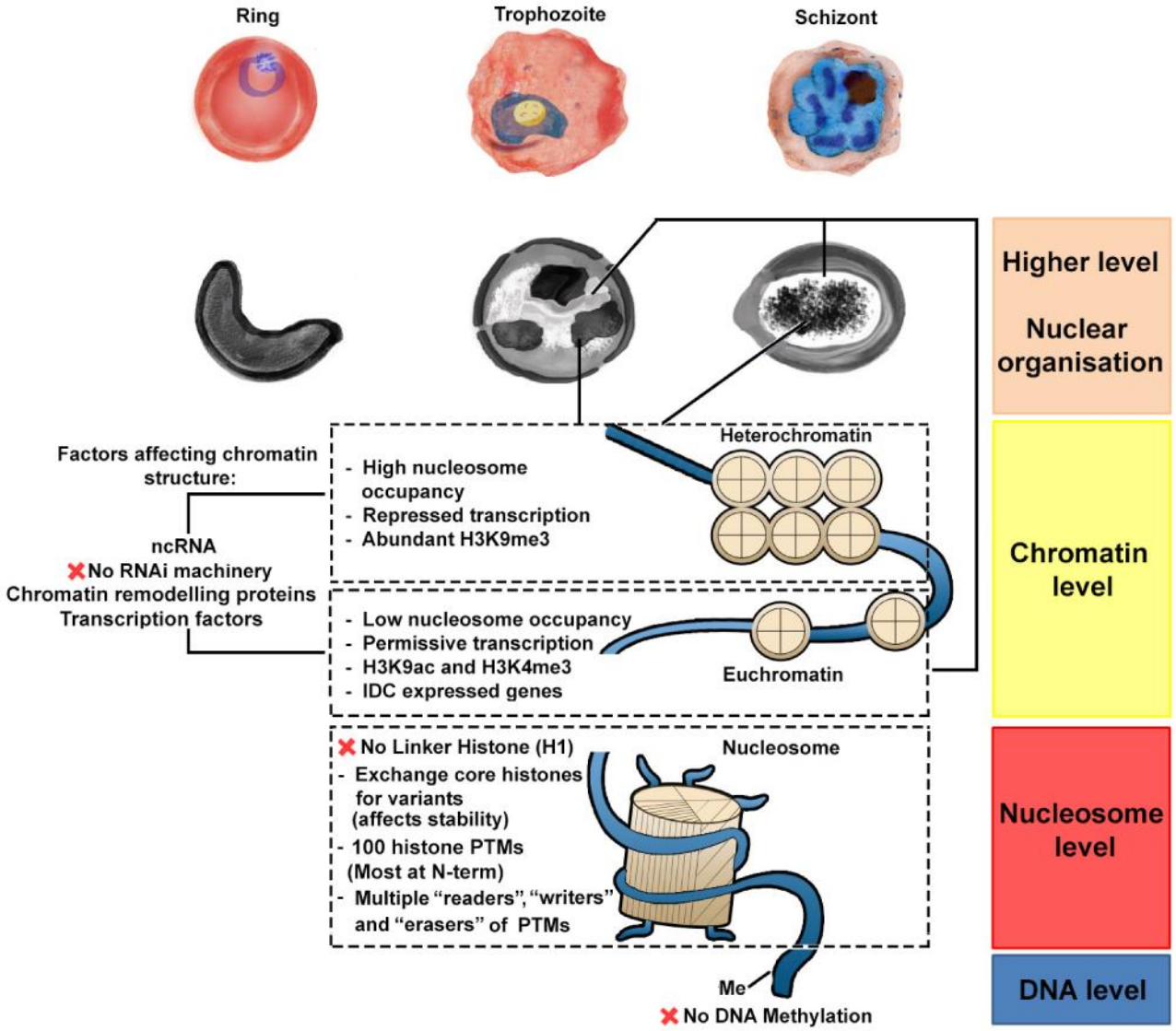


Figure 1.2: An overview of epigenetic regulation in *P. falciparum* parasite asexual stages. The nuclear architecture of three asexual stages is shown. Rings have a cup-shaped nucleus with no clear distinction between heterochromatic and euchromatic regions. Trophozoite stages have scattered heterochromatic islands (black) interspersed with euchromatin. The schizont stage has multiple nuclei (of which only one is displayed) with defined heterochromatin and euchromatin regions. Heterochromatin is the silent form of chromatin, where histone modification marks like H3K9me3 are present. Conversely, euchromatin allows for transcription and is marked by H3K9ac and H3K4me3 marks, amongst others. Unlike other eukaryotes, *P. falciparum* parasites lack RNAi machinery, DNA methylation marks and the linker histone (H1), but have abundant factors affecting chromatin structure, allowing for transition between heterochromatin and euchromatin. The primary level of organisation is the nucleosome, which has ~155 bp of DNA and various factors, such as histone PTMs, which affect nucleosome occupancy. There are abundant PTMs, which mark certain histone residues, and are "read", "written" or "erased" by numerous proteins, to ultimately affect chromatin structure. Figure compiled from [34, 37-39].

Three distinct levels of epigenetic regulation can be described, including a higher level of regulation regarding nuclear organisation, epigenetic regulation at the chromatin level and lastly at the nucleosome level. Unlike other eukaryotes, *P. falciparum* parasites lack RNAi machinery, DNA

methylation marks (DNA level regulation) and the linker histone (H1) [37, 38], but have abundant factors affecting its chromatin structure, allowing for transition between heterochromatin and euchromatin. Approximately 90% of the *Plasmodium* genome occurs in the form of euchromatin particularly during the last half of the intra-erythrocytic developmental cycle (IDC) [38]. Chromatin based epigenetic regulatory mechanisms are important to three biological processes in the parasite: 1) control of life cycle progression; 2) clonally variant gene expression like for the *var* gene family (encoding different erythrocyte membrane proteins allowing immune evasion by the parasite) and 3) mediating adaptive transcriptional responses to external perturbations [40]. The contribution of each level of epigenetic regulation in *P. falciparum* parasites is discussed below.

1.2.1 Epigenetic mechanisms at the higher level of nuclear organisation

The positioning of certain genes plays a role in the regulation of gene expression, where silenced genes move from hetero- to euchromatic regions, to permit access to the transcription machinery as nucleosome occupancy is lowered in euchromatic regions. Clear evidence is provided that these regions move spatially within the nucleus as the parasite progresses through its life cycle [39] (Figure 1.2). Rings have a cup-shaped nucleus with no clear distinction between heterochromatic and euchromatic regions [39] contributing towards our understanding of the low transcriptional activity of rings [41]. Trophozoite stages have scattered heterochromatic islands interspersed with euchromatin [39]. The schizont stage has multiple nuclei with defined heterochromatin and euchromatin regions [39]. The detectable presence of defined euchromatin regions in both trophozoite and schizont stages correlates with the high levels of transcriptional activity seen in these stages [41].

1.2.2 Epigenetic regulation at the chromatin level

Chromatin remodelling refers to the processes that result in a change in chromatin structure from hetero- to euchromatin and the parasite has abundant factors influencing this transition. This includes histone post-translational modifications (PTMs), which indirectly affect chromatin structure through the recruitment of transcription factors or effector proteins [37], remodelling proteins and changes in nucleosome occupancy (Figure 1.2).

1.2.2.1 Changes in nucleosome occupancy

Nucleosome occupancy affects the access of transcription machinery to DNA. Thus, lowered nucleosome occupancy is required for active transcription of genes, which is seen during the S-phase, and increased nucleosome occupancy is seen during the late schizont stages [34]. Eleven SWI2/SNF2 ATPase catalytic domains have been found in *P. falciparum* parasites [32], of which 7 are possible chromatin modifiers [33]. The presence of these ATPases suggests that ATP-

dependent remodelling of chromatin could be occurring in *P. falciparum* parasites. This entails moving of nucleosomes to expose transcription sites, nucleosome displacement or the exchange of core histones for variants. Although the mechanism of action of all the ATPases in *P. falciparum* parasites is not clear, investigations show that nucleosome occupancy in *P. falciparum* parasites changes between the asexual and sexual life stages [42]. Other *P. falciparum* parasite chromatin remodelling proteins include the ***h***igh ***m***obility ***g***roup ***b***ox (HMGB) proteins, which can bind to and bend DNA, affecting nucleosome mobility. HMGB1 is primarily found in asexual stage parasites and HMGB2 in sexual stages [27].

1.2.2.2 Transcription factors (*Trans* regulatory elements)

There are a number of *trans* regulatory elements that impact on chromatin structure and provide either an additional or concerted level of epigenetic regulation. Binding of a single transcription factor or a complex of transcription factors can be facilitated by histone PTMs or by other effector molecules, resulting in various responses. These include the recruitment of additional transcription factors, propagation or formation of heterochromatin, or localisation of chromatin defined sub-nuclear regions. Although a number of enzymes are involved in alteration of chromatin structure (PTMs readers, writers and erasers (discussed below) and chromatin remodellers), transcription factors are unique in that they are characterised by a DNA-binding motif [43]. The Pfmyb family was amongst the first family of transcription factors to be identified [44]. Pfmyb1 is found in trophozoite-stage nuclei and is thought to affect transcription due to putative DNA binding properties [44]. The 27 member ***Api***complexan ***Ap***etela 2 (ApiAP2) transcription factor family possess parasite specific members, making it a valuable antimalarial target to investigate. Certain members of this family have been found to promote heterochromatin formation, assisting in *var* gene silencing [45, 46]. Members of the Alba family of proteins contribute to heterochromatin structure, affect *var* gene expression and can inhibit *in vitro* transcription [34, 47, 48]. Of the four Alba proteins found in *P. falciparum* parasites [46], PfAlba3 and Pflba4 were found to interact with the histone deacetylase PfSir2a [47, 48]. Together, these various transcription factor families expand possible epigenetic drug targets in *P. falciparum* parasites due to their parasite specificity and their impact on *var* gene expression.

Thus, on the chromatin level, the factors discussed above assist in changing the structure of chromatin from hetero- to euchromatin or *vice versa*. These two different chromatin types have varied nucleosome occupancy and a further level of regulation takes place at the nucleosomes.

1.2.3 Nucleosome level epigenetic regulation

In eukaryotes, DNA does not occur as an independent molecule, but rather in association with other nuclear proteins [10]. On the first level of organisation, DNA associates with an octamer of

histones to form nucleosomes, thus resembling “beads on a string”, where a length of DNA wraps around a core of 8 histone proteins (two polypeptides each of H2A, H2B, H3 and H4), and is locked into place by the linker histone H1, followed by a length of linker DNA [10]. This is followed by organisation of the nucleosomes into chromatin fibres, which is the second level of organisation [10]. Epigenetic regulation at the level of nucleosomes is characterised by the replacement of core histones with variant histones as well as histone PTMs.

1.2.3.1 Replacement of core histones with variants

P. falciparum parasites exhibit a similar nucleosome and chromatin organisation as higher eukaryotes, mediated by various nucleosome assembly proteins. *Plasmodium* genes encoding core histones (H2A, H2B, H3.1 and H4) as well as variant histones (H2Av (H2A.Z), H2Bv (H2B.Z), H3v (H3.3) and H3 centromeric), have been identified [49]. However, histone H1 has not been found within the genome sequence or amongst isolated histones from *P. falciparum* parasites [49, 50]. Increased core histone expression coincides with increased DNA synthesis [50] in late trophozoite and schizont stages [19, 51]. The replacement of core histones with variant histones impacts on the stability of chromatin structure and nucleosomes. In higher eukaryotes, an H2A.v replacement is said to result in reduced nucleosome stability and replacement with H3.3 results in increased gene transcription [52]. In *P. falciparum* parasites, H2A.Z (apicomplexan-specific histone variant) localises to the euchromatic regions and is involved in *var* gene activation [53, 54]. The other apicomplexan-specific histone variant, H2B.Z, together with H2A.Z, forms a double-variant nucleosome subtype, which is abundant at AT-rich 3’ intergenic regions on promoters [55, 56]. *Plasmodium* centromeres are delimited by histone H3 centromeric, as the name indicates [57] and this variant histone plays a role in chromosome segregation and kinetochore formation. Although the function of H2B.Z is not known, the acetylation of at least 6 residues on its N-terminal tail [49] is thought to be an indication of its association with euchromatin [34]. Thus, like other eukaryotes, *P. falciparum* parasites exploit amino acid differences of histone variants to facilitate changes in chromatin structure and transcriptional regulation [56].

1.2.3.2 Histone PTMs

PTMs affect intrinsic protein characteristics, such as activity, affinity and hydrophobicity by altering the size, charge and structure of the protein [58]. These changes result in varying functioning of proteins [58]. Proteins can either be modified at single or multiple residues, or can be modified with more than one type of PTM. The N- and C-termini of histones extend beyond the globular nucleosomal structure [10]. These termini are called “histone tails” and have a high content of basic amino acids such as lysine (modified at their Ɛ-amino group) and arginine residues (modified at the nitrogen on their guanidium group).

In addition to the indirect effect that histone PTMs can have on chromatin structure (Section 1.2.2.2), they can also contribute directly to chromatin structure changes by directly influencing interaction of the negatively charged DNA phosphate backbone with the positively charged residues in the histone tails [50, 59, 60]. For example, masking of the positive charge on a lysine residue causes a decreased interaction with negatively charged DNA leading to euchromatin (open or relaxed chromatin) allowing for increased transcription. Conversely, deacetylation results in an increased interaction as seen in heterochromatin (condensed) thereby decreasing transcription [61].

The combination of PTMs that histone tails undergo is hypothesised to be a “histone code”, which is interpreted by certain enzymes to either condense or relax chromatin, thereby regulating transcription of chromatin [62]. The code can be “written”, “erased” and “read” by various types of proteins [62]. Enzymes that carry out PTM of histones “write” the code (e.g. acetylation of histones by ***h***istone ***a***cetyl ***t***ransferases, HATs) and enzymes which remove these modifications, “erase" the code (e.g. deacetylation by ***h***istone ***d***e***ac***etylases, HDACs) [62]. “Reading” of the code is conducted by effector molecules, such as transcription factors and chromatin remodelling enzymes, discussed above [26]. PTMs of histones facilitate binding of these effector molecules, which ultimately affect chromatin remodelling and stability [63]. Chromatin-remodelling enzymes possess domains for binding certain histone PTMs. The bromodomain binds acetylated lysine residues, the chromodomain binds methylated lysines and the plant homeodomain (PHD) binds trimethylated H3K4 [26]. From the variety of modifications that *P. falciparum* parasite histones can undergo, methylation and acetylation are the most common [37].

In *P. falciparum* parasites, a PTM of the Ɛ-amino group of lysine in histones can only be conducted in the absence of other modifications, with acetylation impeding other forms of modification [59]. Thus, HATs and HDACs generally control the possibility of other modifications from occurring [59]. In context of epigenetic regulation, the most common PTMs of histones include acetylation, methylation, phosphorylation, ubiquitination and sumoylation among others [37, 64]. In Figure 1.3, all the PTMs of *P. falciparum* parasite histones that were previously detected in various reports [37, 49, 50, 65, 66] are graphically compiled. This catalogue of PTMs were accumulated using a variety of techniques, including antibody detection and mass spectrometry (MS) analysis of histone PTMs that were either enriched for or sought without any enrichment technique.

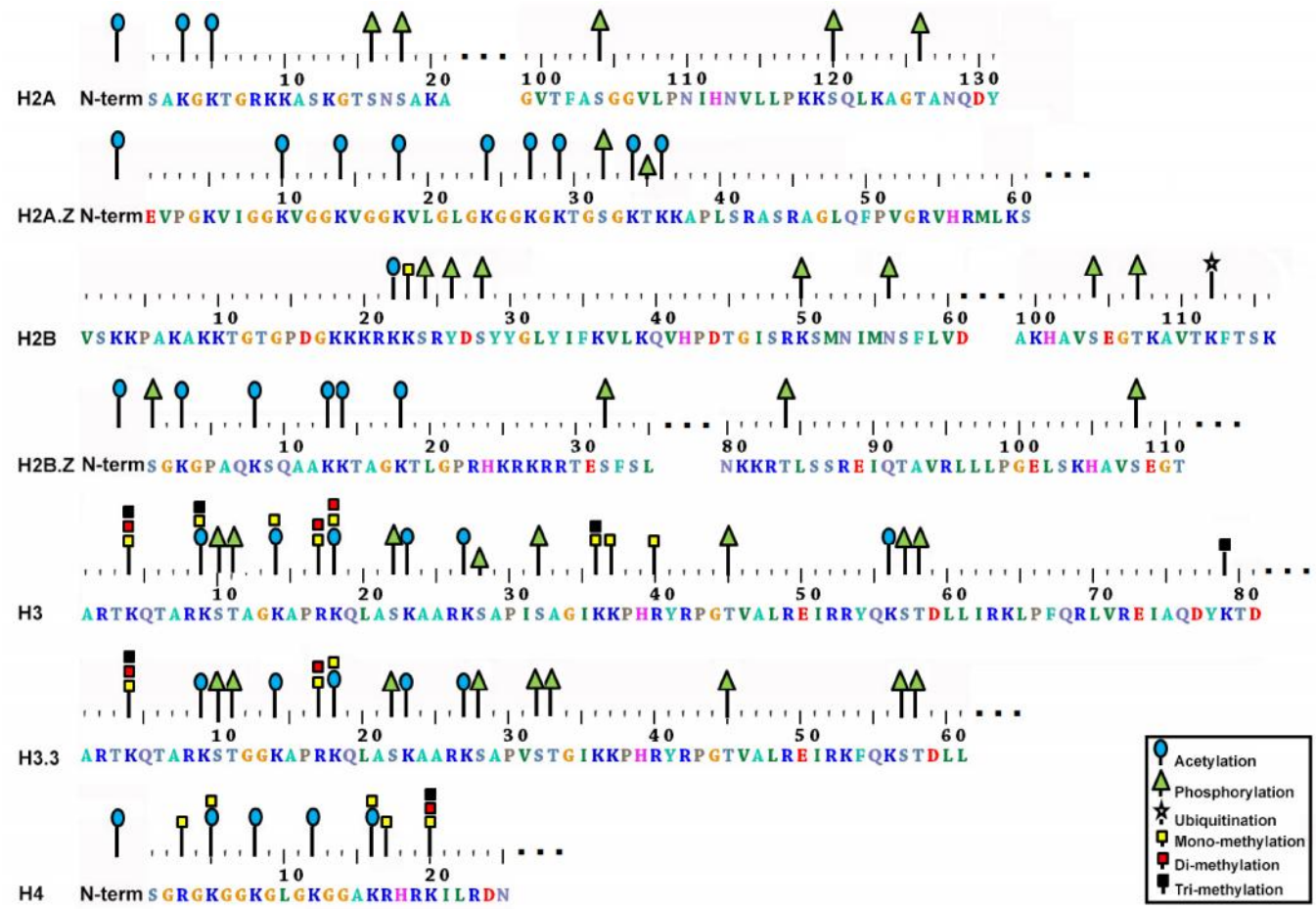


Figure 1.3: Current post-translational modifications identified on P. falciparum parasites' core and variant histones. Graphical depiction of histone-tail modifications on the histone sequences. Ellipses are indicative of sequences where no modifications exist. Sumoylation was identified on H4, but its exact location is unknown. There are 8 histone proteins in P. falciparum parasites, however, no modifications were found on the variant histone, H3 centromeric. Information compiled from [37, 49, 50, 65, 66].

Histone Ubiquitination

Ubiquitination is a reversible process involving the conjugation of one or more ubiquitin peptides to certain lysine residues in proteins to facilitate downstream effects [10]. The ubiquitin pathway involves an ubiquitin-activating enzyme (E1), which donates an ubiquitin-thioester to the second enzyme (E2) involved in the pathway, forming an E2-ubiquitin-thioester, as well as an ubiquitin-ligase (E3), which conjugates ubiquitin to the target protein [10]. Many well-conserved homologues of E1 and E2 have been found in *P. falciparum* parasites, with *P. falciparum* parasite E3 enzymes being the most abundant and diverse [67]. Ubiquitination modifications are regulatory mechanisms, which can be controlled by phosphorylation [68]. While ubiquitinated proteins in general are destined for degradation, it has been found that ubiquitination can be involved in regulation of other cellular processes, such as DNA repair and mitotic progression [67, 68]. Histones have been found to be mono-ubiquitinated only but poly-ubiquitination is known to occur on non-histone proteins. In *Plasmodium*, not much is known about histone ubiquitination, though it has been found in other

higher eukaryotes that H2A ubiquitination causes silencing of genes, while H2B ubiquitination can cause either activation or repression [69]. These effects on transcription are due to the fact that histone ubiquitination serves as a signalling molecule with further downstream effects resulting in transcription factor recruitment or effects on histone methylation [69]. *P. falciparum* parasite histone ubiquitination of lysine 112 (K112) can be seen in Figure 1.3 and to date is the only known ubiquitination of *Plasmodium* histones found.

Histone Sumoylation

Sumoylation refers to the conjugation of lysine residues in proteins to a SUMO (***S***mall ***U***biquitin-like ***Mo***difier) or Nedd8 (another ubiquitin-like protein) [70-72]. Sumoylation of histones has been found to occur at H4 within the parasite, which represses gene expression [37, 70]. Furthermore, sumoylation has been found to act along a similar, but distinct pathway to ubiquitination, although in competition with this process [70, 71]. Sumoylation of histones and other proteins alters their interaction with other proteins and affect parasite proteins in a stage-specific manner [70, 72]. Sumoylated HDACs have altered deacetylase activity, thereby affecting transcription [71].

Histone Phosphorylation

Phosphorylation of histones occurs on serine, threonine and tyrosine residues. There are many determined functional consequences of histone phosphorylation in higher eukaryotes including DNA repair, cell division and replication, apoptosis, and nuclear hormone signalling [73]. Phosphorylation of target residues is catalysed by kinases and 65 kinases have been identified in *P. falciparum* parasites [74], though only CK2 kinase has been found to associate with histones *in vitro* [75]. “Readers” of phosphorylation marks include the Pf-14-3-3 protein, which shows a preference for binding H3S28ph rather than H3S10ph [65]. At least 29 histone phosphorylation marks have been identified in *P. falciparum* parasites with no clear indication of their functional consequence on chromosomal structure yet [65, 66].

Histone Methylation

Methylation states, which occur commonly in *P. falciparum* parasite histones, play an important role in epigenetic regulation within the parasite. Lysine residues of histone tails can be methylated with one, two or three methyl groups, while histone arginine residues can be mono- or dimethylated (Figure 1.3) [64]. Methylation occurs primarily at lysine 4, 9 and 36 on H3 and at lysine 20 on H4 (Figure 1.3) [64, 76]. There are two classes of ***h***istone ***m***ethyl ***t***ransferases (HMTs), which are classified according to the amino acids they methylate, namely K-HMTs, which methylate lysine residues and R-HMTs, which methylate arginine residues [64]. In *P. falciparum* parasites, at least 9 genes encoding K-HMTs, which possess a SET-domain, are found [76].

Methylation of histones can be reversed by demethylases within the parasite. Lysine specific demethylases, containing a Jumonji-C domain, have been found in *P. falciparum* and other apicomplexan parasites [76]. In general, H3K4, H3K36 and H3K79 methylation marks are prominent in transcriptionally active genes [77], while H4K20 and H3K9me3 and H3K27me3 are associated with silent genes (Figure 1.2) [78]. H3K9me3 is an extensively studied modification as it is associated with heterochromatin and silent *var* gene marks [77, 79]. This modification is bound by the chromodomain containing heterochromatin-binding protein (HP1) [80].

Histone Acetylation

Histone acetylation is carried out by HATs, which were found to acetylate histone lysine residues 9 and 14 in *P. falciparum* parasites [30, 37, 81]. Acetylation of histones lowers their affinity for DNA as positively charged histone tails no longer interact strongly with negatively charged DNA [61]. This allows greater access to DNA by transcription factors and polymerases leading to increased levels of gene expression [82, 83]. Thus, hyperacetylation leads to activation of gene expression, while hypoacetylation leads to gene silencing [64]. HATs fall into two groups; Type-A HATs, which acetylate histones directly involved in regulation of chromatin assembly and transcription [64] and Type-B HATs, which acetylate lysine 5 and 12 of H4 [64]. Transcriptionally active genes have been are marked by H3K9ac, carried out by a HAT, which if disrupted, has been shown to have growth-inhibiting effects on the *P. falciparum* parasite [84]. Perturbations in the functioning of other *P. falciparum* HATs such as PfMYST (***M***OZ, ***Y***bf1/Sas3, ***S***as2, and ***T***ip60) results in disrupted schizogony [37]. By contrast, HDACs are responsible for deacetylation of lysine residues, acting in opposition to HATs [85]. Both HDACs and HATs have received marked attention recently due to their essential nature to parasite viability, with HDAC inhibitors particularly showing potential as antimalarial agents. The rest of this discussion will focus specifically on HDAC proteins of *P. falciparum* parasites.

1.3 Classes of HDACs in *P. falciparum* parasites

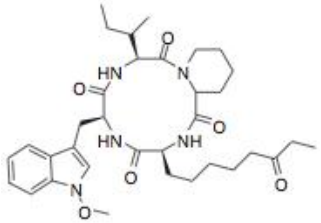
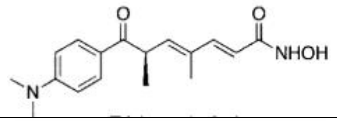
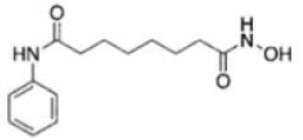
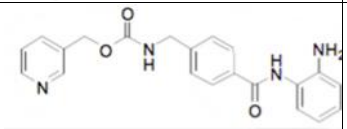
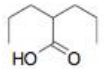
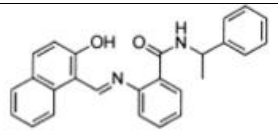
In humans there are four classes of HDACs grouped on the basis of sequence similarity and co-factor usage [85]. *P. falciparum* parasites have been found to have at least 3 of these classes, designated class I, II and III. Class I and II HDACs utilise zinc as a co-factor and exhibit sequence similarity to the yeast HDACs (Rpd3 and Hda1) [30]. Class III HDACs utilise NAD^+ and are homologous to yeast ***s***ilencing ***i***nformation ***r***egulator 2 (Sir2)[30]; these HDACs are also referred to as ***sir***tuins [86]. The sirtuins in *Plasmodium*, PfSir2A and PfSir2B, have been associated with heterochromatin and *var* gene silencing [86, 87]. From comparative genomic studies, no class IV HDACs were found in *Plasmodium* species [88].

In totality, five HDACs have been found in *P. falciparum* parasites, namely PfHDAC1, PfHDAC2, PfHDAC3, PfSir2A and PfSir2B [85]. PfHDAC2 and PfHDAC3 belong to class II, while the latter two belong to class III. In addition to its deacetylase ability PfSir2A also possesses ADP-ribosylation activity [86]. PfHDAC1, which has been studied extensively, is a class I HDAC [85] expressed in merozoites, trophozoites and schizonts [89], where it localises in the nucleus [90]. PfHDAC1 homology models have a conserved active site, though the substrate entrance to this site differs from that of human HDACs [91]. The active site contains a cationic zinc molecule co-ordinated to histidine and aspartic acid moieties. Deacetylation occurs with the use of a water molecule, which is deprotonated by the zinc cation, and the proton is transferred to the co-ordinating residues [92]. Gametocytes express both PfHDAC2 and PfHDAC3, while trophozoites do not express the latter [89].

1.4 Effect of HDAC inhibitors used against *P. falciparum* parasites

As HDACs are important for gene expression and chromatin organisation, they also play an important role in cell cycle regulation. In cancer cells, hypoacetylation (or increased HDAC activity) leads to under expression of genes essential to cell cycle regulation, which increases cancer cell growth [93]. Conversely, when HDAC inhibitors are used, hyperacetylation (or decreased HDAC activity) arrests cancer cell growth [93]. The similarities between cancer cells and parasitised erythrocytes may extend to the effects of HDAC inhibition as well, which is why many HDAC inhibitors used against cancer cells have been tested against *P. falciparum* parasites with promising results. Table 1.1 summarizes the current inhibitors tested against *P. falciparum* HDAC classes and will be discussed below.

Table 1.1: Class I, II and III HDAC inhibitors and their effects in *P. falciparum* parasites

Class of HDAC inhibited	Type of HDAC inhibitor	Name of inhibitor	Structure	IC_{50} in *P. falciparum*	Advantages/Disadvantages	Ref	Other inhibitors in this group
Class I and II	Cyclic Tetrapeptides	Apicidin		125 ng/ml	– Lack of selectivity for parasite HDACs – Quinolone derivatives 200-fold more selective for *P. falciparum* parasites than mammalian cells	[94, 95]	FR235222
	Hydroxamates	TSA		11 nM	– IC_{50} (nM) at which parasite histones are hyperacetylated – Less selective than SBHA	[94, 96]	ABHA SBHA Triazoylphenyl based (compound **10c,** 2-aminosuberic acid, L-cysteine, WR301801)
		SAHA		0.1-0.3 µM	– FDA-approved for lymphoma treatment – More selective than TSA, more active than SBHA	[85]	
	Benzamides	MS 275		8.3-10 µM	– Modest activity against *P. falciparum* parasites – Selective for human HDAC1 over human HDAC2	[91, 97]	MGCD0103
	Short-chain fatty acid	Valproic acid		>100 µM	– Non-HDAC targeted effects	[85]	Butyrate
Class III		Sirtinol		9.1-12.9 µM	– Low homology between mammalian and *Plasmodium* class III HDACs – Modest activity	[91]	Splitomycin Surfactin Hyperforin

1.4.1 Class I and II HDAC inhibitors

HDAC inhibitors against class I and II HDACs in *P. falciparum* parasites have been researched extensively, particularly because of successful recombinant expression of *P. falciparum* HDAC1 (PfHDAC1) [97] and the high homology shared with human class I HDACs [85]. The archetypal HDAC inhibitor (Figure 1.4) has been used to design class I and II HDAC inhibitors. The cap group is a structural group necessary for specific action against the targeted HDAC, and can also be altered to increase the potency of a compound [89, 97, 98]. The metal chelating group binds the zinc molecules that are necessary for HDAC activity [89, 97, 98]. The hydrophobic linker group links the two functional groups that make these compounds HDAC specific, however, the length of the linker group is also a contributing factor to the compounds' inhibitory activity [89, 97, 98].

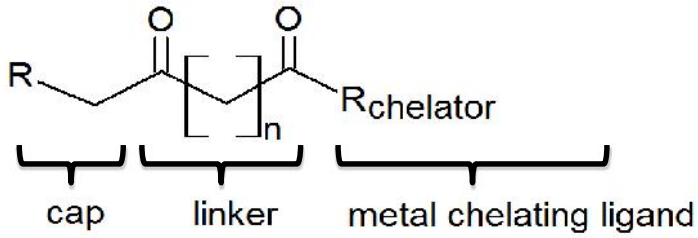


Figure 1.4: An archetypal class I and II HDAC inhibitor. This typical inhibitor contains a cap group, a hydrophobic linker and a metal chelating ligand for inhibition of the zinc-requiring HDACs. Adapted from [89, 97, 98].

The *Fusarium* fungal metabolite, apicidin (Table 1.1), is a cyclic tetrapeptide with nanomolar activity against *P. falciparum* and other Apicomplexan parasites [94]. Apicidin was confirmed as an HDAC inhibitor in *P. falciparum* parasites, which perturbed gene expression within the intra-erythrocytic developmental cycle [30]. *In vivo* efficacy of apicidin in *P. berghei*-infected mice was seen but the compound lacks parasite selectivity over mammalian cells, thus reducing its potential to be considered an anti-malarial [94].

Lack of selectivity of apicidin is thought to be due to the presence of a 2-amino-8-oxodecanoic acid side chain on apicidin, which resembles that of the Ɛ-amino group on lysine residues thereby affecting both mammalian and parasite HDAC enzymes [94]. Apicidin derivatives with a more basic moiety resembling the basicity of histones had increased activity against *P. falciparum* HDACs [89]. Quinolone derivatives of apicidin, such as methoxyquinolone and 4-hydroxyquinolone *N*-substituted derivatives, were the most effective and were found to have up to 200-fold more selectivity for *P. falciparum* parasites than mammalian cells (Table 1.1) [99]. Other cyclic tetrapeptides tested against *P. falciparum* parasites were HC-toxin and FR235222, both of which had nanomolar activities [89]. Like apicidin, HC-toxin had cytotoxic effects against the parasite [89]. Another fungal-derived peptide, FR235222, from the *Acremonium* species has also been found to target HDAC3 in *Toxoplasma gondii* affecting stage-specific gene expression [95].

Known mammalian hydroxamate-based HDAC inhibitors and their analogues tested against *P. falciparum* parasites include trichostatin A (TSA), azelaic bishydroxamic acid (ABHA), suberohydroxamic acid (SBHA), aryltriazolylhydroxamates and suberoylanilide hydroxamic acid (SAHA) amongst others (Table 1.1) [96]. TSA showed nanomolar IC_{50} values against *P. falciparum* parasites (Table 1.1) due to its ability to hyperacetylate *Plasmodium* parasite histones [94]. ABHA and SBHA also inhibited parasite proliferation with SBHA showing greater activity [96]. SBHA treatment showed cytostatic effects against *P. berghei*-infected mice [96], but its selectivity for transformed tumour cells over normal cells make it an important parent compound for structure-based drug design of more potent analogues [96]. SAHA has been FDA-approved for treatment of T-cell cutaneous lymphoma and has also shown greater inhibitory activity against *P. falciparum* parasite infected erythrocytes than SBHA, as well as better selectivity for the parasite than TSA (Table 1.1) [85].

Aryltriazolylhydroxamates pose IC_{50} values against *P. falciparum* parasites in the nanomolar range [100] with the most active compounds being those that had a linker group of five to six methylene groups; a large cap group with either a biphenyl, naphthalene or quinolone groups; and a triazole ring at the *meta*- position [100]. Furthermore, these compounds exhibited activity similar to TSA and at least 5-15 fold more than SAHA [100]. Compounds based on 2-aminosuberic acid and L-cysteine are also hydroxamate-based. These compounds are selective for transformed cancer cells over normal cells and when tested against *P. falciparum* parasites inhibited both multi-drug resistant and sensitive strains of *P. falciparum* parasites in the nanomolar range [89]. The 2-aminosuberic acid compounds were more effective than L-cysteine based compounds, and were able to halt life cycle progression when used on various asexual stage parasites, where treated parasites were unable to progress to the next developmental stage [91].

Triazoylphenyl hydroxamate inhibitors (Table 1.1), e.g. WR301801, exhibited IC_{50} values against *P. falciparum* parasites of less than 3 nM with selectivity indices of greater than 600 [101] and resultant histone hyperacetylation [101]. Treatment of *P. berghei*-infected mice with high doses of WR301801 (640mg/kg/day) stabilised parasitaemia, however, combination therapy of WR301801 (52 mg/kg/day) and chloroquine (Table 1.1) was required for a curative effect [101]. Structure-activity relationships indicated that amino or hydroxyl groups at the *meta*- position of the phenyl group gave the most potent compounds, whereas *ortho*- or *para*- substitutions, as well as tertiary amino or methyl groups decreased the potency of the compound [101]. Loss of activity could also be correlated to the loss of the hydroxamate group due to metabolism in the body [101]. Another triazoylphenyl-based hydroxamate inhibitor (compound **10c**), showed equivalent potency to TSA against multi-drug resistant strains of *P. falciparum* parasites (C235 and C2A) and compound **10c**'s potency was greater than that of chloroquine and mefloquine [92]. This compound was also

more selective against the parasite than against a variety of human pancreatic cancer cell lines (Hup T3, Bx-PC-3, and SU 86.86) [92, 94].

Short chain fatty-acid type HDAC inhibitors include valproic acid, butyrate and their derivatives (Table 1.1). Even though they do have inhibitory effects against HDACs, they were also found to have off-target effects [85]. This property, together with their minimal inhibition of *P. falciparum* parasites *in vitro*, makes them poor anti-malarials [85].

Benzamide-containing HDAC inhibitors that do not fall within the major groups and exhibit human HDAC1 selectivity include MS 275 and MGCD0103 (Table 1.1) [102]. These 2 compounds possess nanomolar IC_{50} values against HeLa cells [102]. MS 275 has since been advanced to Phase 2 clinical trials [102]. However, when used against *P. falciparum* parasites the effect of MS 275 was not as potent, and IC_{50} values of 8.3 µM and 7.8 µM were obtained in chloroquine sensitive and resistant strains, respectively [91]. MS 275 is unique in its directed HDAC1 inhibition as compared to TSA and SAHA, which are pan (unselective) HDAC inhibitors [102].

1.4.2 Class III HDAC inhibitors

The above-mentioned inhibitors have been found to inhibit class I and II HDACs. Other inhibitors of class III HDACs or sirtuins have been found [97]. When mammalian class III HDAC inhibitors (sirtinol, surfactin, splitomycin, nicotinamide and hyperforin) were used against *P. falciparum* parasites, micro- to millimolar IC_{50} values were determined (Table 1.1) [85, 89, 97]. This modest activity against the parasite is probably due to low homology between human sirtuins and PfSir2 (Table 1.1) [85].

1.5 Problem statement and objective

The gene encoding PfHDAC1 was found to be essential to the *P. falciparum* parasite [103], making this a possible drug target for anti-malarials. The present work therefore investigates the inhibition of PfHDAC1. Inhibition of HDAC enzymes results in hyperacetylation of H3K9 and of histone H4 N-terminal residues; putative HDAC inhibitors must satisfy this criterion. However, the detection of hyperacetylation is routinely done with Western Blot analysis, though quantitative mass spectrometry (MS) analysis can possibly distinguish hyperacetylated profiles of treated samples at a much more sensitive level. Previous studies have shown that MS analysis of *P. falciparum* parasite histones can detect a number of modifications [49]. However, an unbiased method, which not only detects all 8 histone proteins, but the abundant number of modifications, which could exist on *P. falciparum* parasite histones remains to be developed. Further elucidation of these PTMs would provide valuable epigenetic information in *P. falciparum* parasites. The objective of this study was therefore the development of a histone isolation method, which allows quantitative,

comparative and comprehensive MS analysis of histone PTMs to aid the characterisation of PfHDAC1 inhibition in *P. falciparum* parasites by novel anti-malarial compounds.

1.6 Aims

1. Optimise a histone isolation method to determine the effect on histone acetylation by HDAC inhibitors using MS analysis.
2. Identify potential HDAC inhibitors against a homology model of PfHDAC1 using *in silico* screening.
3. Develop an HDAC activity assay to confirm the inhibitory activity on *P. falciparum* parasite HDACs.

1.7 Outputs

This work was presented at the following conferences and symposia:

1. Gordon Research Conference on Malaria (Molecular and Cell Biology of Malaria). August 4-9, 2013, Renaissance Tuscany Il Ciocco Resort, Lucca (Barga), Italy
2. The 6th Multilateral Initiative on Malaria (MIM) PAN African Malaria Conference, October 6-11, 2013, Durban International Convention Centre (ICC), Durban, South Africa
3. UPCSMC (University of Pretoria Centre for Sustainable Malaria Control) project showcase day for the National Department of Health (NDoH), October 2, 2013, University of Pretoria, FABI Auditorium; Pretoria, South Africa

Manuscript in preparation:

Nabila Ismail, Abraham I. Louw and Lyn-Marie Birkholtz

Rapid, inexpensive technique for the isolation and detection of histones and novel histone post-translational modifications in the malaria parasite, *Plasmodium falciparum*

Journal: BMC Epigenetics and Chromatin

Chapter 2

Materials and Methods

2.1 Cultivation and maintenance of parasite cultures

2.1.1 Collection of blood

Blood (O^+ or A^+) was collected in a blood bag (Fenwal Primary container with citrate phosphate adenine anticoagulant, Adcock Ingram), and aliquoted into Falcon tubes aseptically. The blood was left overnight at 4ºC to allow the serum to separate from the erythrocytes. Removal of the buffy coat is essential as leukocytes are capable of destroying parasites in culture. Erythrocytes were washed at least twice in phosphate buffered saline (PBS; 10 mM Na_2HPO_4; 1.4 mM KH_2PO_4, 137 mM NaCl, 2 mM KCl (Merck); pH 7.4) by centrifugation at 3000*g* (Hermle Z320 centrifuge) for 5 minutes at room temperature until all visible leukocytes were removed. An equal volume of culture medium [RPMI-1640 (Sigma) supplemented with 23.81 mM sodium bicarbonate (Merck); 0.024 mg/ml gentamycin (Thermo Scientific); 25 mM HEPES (Sigma); 0.2% ($^w/_v$) glucose,(Sigma); 0.2 mM hypoxanthine (Sigma) and 2.5 g/l Albumax II] was added to generate a 50% haematocrit solution, which was maintained at 4ºC and used for parasite cultivation. Albumax II is used in place of human serum to prevent contaminating immunoglobulins from entering the culture. Gentamycin prevents bacterial contamination and glucose maintains osmolarity of the solution and is a key carbon source for malaria parasites [Methods in Malaria Research, www.mr4.org].

2.1.2 Thawing and maintenance of parasite cultures

The use of human erythrocytes and malaria parasites has been ethically cleared by the University of Pretoria Faculty of Natural and Agricultural Sciences Ethics Committee (ECI20821-077). Parasite stocks of *P. falciparum* strain 3D7 (a chloroquine-sensitive strain) were thawed from cryopreservation (-180ºC) at 37ºC for 5 minutes. To this, 12% ($^w/_v$) NaCl (Merck) was added, re-suspended and followed by addition of 0.6% ($^w/_v$) NaCl. This parasite suspension was centrifuged at 3000*g* for 5 minutes and the supernatant aspirated. Washed erythrocytes (50% haematocrit) were added to the pelleted parasite suspension and pre-warmed (37°C) culture medium was added to obtain a 5% haematocrit culture. This culture was transferred to a 75 cm^3 Cellstar culture flask (Greiner bio-one) and gassed for 30 seconds with a unique gas mixture containing 90% nitrogen, 5% oxygen and 5% carbon dioxide (Afrox) to mimic *in vivo* growth conditions. The culture was incubated at 37ºC and rotated at 58 rpm to minimize multiple infection of erythrocytes, prevent localised lactic acid build up and ensure even distribution of nutrients [104]. Parasite cultures were

maintained daily to monitor parasitaemia levels (between 5 to 10%), to replenish spent culture medium and add fresh erythrocytes when necessary [105]. Parasitaemia was calculated after counting 1000 cells from Giemsa-stained blood smears using light microscopy techniques (Nikon Labophot, total magnification of a 1000×). Giemsa stains the nucleus of the parasite purple and the cytoplasm blue, allowing for effective stage identification and determination of parasitaemia [106].

2.2 Sorbitol synchronisation

Synchronisation of *P. falciparum* parasite proliferation to similar life cycle stages was performed using sorbitol, which lyses erythrocytes infected with mature stages of the parasite but does not affect erythrocytes infected with ring forms and uninfected erythrocytes [107]. The parasite culture (> 2% rings, 5% haematocrit) was centrifuged, spent culture medium aspirated and the cells were re-suspended in 5% ($^{w}/_{v}$) D-sorbitol (Sigma) and incubated at 37ºC for 15 minutes. The cells were washed with complete culture medium followed by centrifugation at 3000*g*, after which the cells were re-established in culture as described in Section 2.1.2. Cultures were synchronised at the ring-stage and had to have undergone at least 3 consecutive synchronisation cycles prior to use in an assay to ensure ~90% synchronicity.

The following methods will be separated in two parts according to the project Aims for clarity.

PART I

Optimise histone isolation method to determine the effect of HDAC inhibitors on histone acetylation using MS analysis.

2.3 Isolation of *P. falciparum* parasite histone proteins

PTMs of histones can be detected with MS, giving detailed information regarding the PTMs that influence epigenetic regulation within the parasite [108]. This requires isolation of histones from blood stage parasites, such as trophozoites, by removing contaminating erythrocyte and parasite membrane proteins, dissociating core histones from DNA and enriching acid-soluble protein fractions containing histones. Three published histone isolation methods were compared as shown in Figure 2.1 below. An explanation of all reagents and techniques used are provided in Appendix-Table 1A. Modifications were then made to each of the methods in attempts to obtain high protein yields for efficient MS analysis.

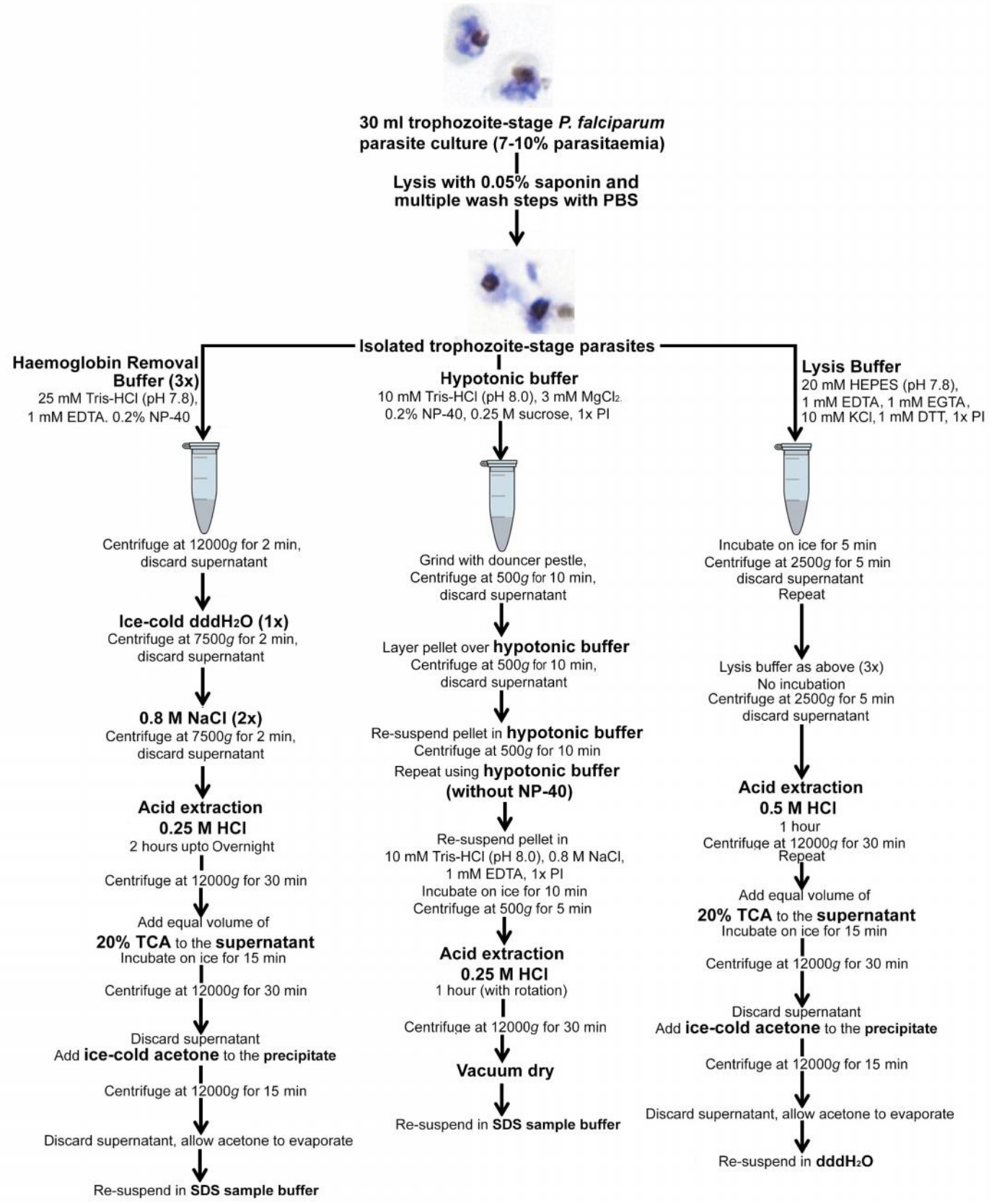


Figure 2.1: Isolation of *P. falciparum* parasite histones from trophozoite stage parasites using three different published methods. The method of Issar and Scherf [www.mr4.org] is shown on the left, the Trelle method in the middle [49] and Merrick method on the right [86, 109]. Text in bold is indicative of the sample preparation step as well as the fraction it was performed on.

2.3.1 Histone isolation (Merrick, [86, 109])

A 30 ml trophozoite stage culture (7-10% parasitaemia, 5% haematocrit) was lysed with 0.05% saponin (Sigma) in PBS and incubated on ice for 10 minutes. Lysed parasites were then centrifuged at 315*g* for 5 minutes (BOECO centrifuge, C-28 A), the supernatant aspirated and the pellet washed thrice with ice cold PBS. Ice cold lysis buffer (20 mM HEPES (Sigma); pH 7.8, 10 mM KCl (Merck), 1 mM EDTA (Merck), 1 mM EGTA (Sigma), 1 mM DTT (Merck), 1× Protease Inhibitor cocktail (Roche), 0.65% ($^{v}/_{v}$) Nonidet P-40 (Roche)) was added to the parasite pellet and incubated on ice for 5 minutes followed by centrifugation at 2500*g* for 5 minutes at 4°C (Eppendorf centrifuge 5415 R). The supernatant containing cytoplasmic proteins was discarded. Re-suspension in lysis buffer, incubation and centrifugation was then repeated on the pellet to enrich for nuclear proteins. Subsequently, the nuclear protein rich pellet was washed thrice with lysis buffer until the supernatant was clear. To isolate histones from the obtained nuclear pellet, acid extraction with 0.5 M HCl (Merck) was performed for 1 hour at 4°C. A douncer pestle was used to fragmentise the pellet allowing HCl to act on the particles. The solution was then centrifuged at 12000*g* for 30 minutes, and the pellet was extracted as before. Both acid extraction fractions were combined thereafter and these resulting fractions were incubated with an equal volume of 20% trichloroacetic acid (TCA, Sigma) for 15 minutes on ice. TCA-precipitated proteins were harvested by centrifugation at 12000*g* for 30 minutes followed by a wash with ice-cold acetone (Merck), and a final centrifugation at 12000*g* for 15 minutes. Excess acetone was removed and allowed to evaporate and the protein precipitate was re-suspended in SDS sample buffer (0.063 M Tris-HCl (Sigma); pH 6.8, 8.7% ($^{v}/_{v}$) Glycerol (Sigma), 2% ($^{w}/_{v}$) SDS (Merck), excluding bromophenol blue and 2-β mercaptoethanol) and stored at -70°C until use or used immediately. Re-suspension of protein samples in $ddd H_2O$ (Figure 2.1 as published) was not performed due to solubility issues faced (Data not shown). Modifications that were made to the published isolation procedure include: performing a continuous 2 hour acid extraction procedure instead of two separate acid extractions and exclusion of the acetone wash step.

2.3.2 Histone isolation (Issar and Scherf, www.mr4.org, Methods in Malaria Research)

A 30 ml trophozoite stage culture (7-10% parasitaemia, 5% haematocrit) was isolated, and washed thrice with PBS as described above (Section 2.3.1). The resulting pellet was then washed thrice with haemoglobin extraction buffer (25 mM Tris-HCl (Sigma); pH 7.8, 1 mM EDTA (Merck), 0.2% ($^{v}/_{v}$) Nonidet P-40 (Roche)) to remove contaminating membranes and haemoglobin from the parasite lysate. Each wash step was followed by a centrifugation step at 12000*g* for 1 minute. The pellet was then washed once with ice-cold $ddd H_2O$, and centrifuged at 7500*g* for 1 minute. A 0.8 M NaCl (Merck) solution was then used to wash the pellet twice (7500*g* for 2 minutes), followed by acid extraction with 0.25 M HCl (Merck). The pellet was homogenised using a douncer pestle and left to extract for 2 hours or overnight at 4°C. The solution was then centrifuged at 12000*g* for 30

minutes and the supernatant incubated on ice for 15 minutes after addition of an equal volume of 20% TCA solution. The mixture was then centrifuged again for 30 minutes at 12000*g* and the supernatant removed. The precipitate was then washed with ice-cold acetone and centrifuged at 12000*g* for 15 minutes. The acetone was removed and allowed to evaporate and the protein precipitate was re-suspended in SDS sample buffer (without bromophenol blue and 2-β mercaptoethanol) and stored at -70°C until use. Since the published method (Figure 2.1) recommends an acid extraction period from 2 hours to overnight, these two outlying periods were used to determine if they could affect the protein yield. A modification made to the published isolation procedure (Figure 2.1) was the exclusion of the acetone extraction step.

2.3.3 Histone isolation (modified from Trelle, [49])

Isolated trophozoite stage parasites were washed thrice with PBS as described above (Section 2.3.1). A hypotonic buffer (10 mM Tris-HCl (Sigma); pH 8.0, 3 mM $MgCl_2$ (Merck), 0.2% ($^{v}/_{v}$) Nonidet P-40 (Roche), 0.25 M sucrose (Melford Chemicals), 1× Protease Inhibitor cocktail (Roche)) was added to the isolated parasites, followed by homogenisation with a douncer pestle. The pellet was obtained by centrifugation at 500*g* for 10 minutes at 4°C (Thermo Scientific centrifuge SL 16R), and was layered over the hypotonic buffer. This was followed by centrifugation at 500*g* for 5 minutes. The pellet was then washed once more in the hypotonic buffer followed by a wash in a buffer with the same composition, though without Nonidet P-40. Pelleted nuclei were then obtained by centrifugation at 500*g* for 5 minutes. These nuclei were slowly re-suspended in a high salt concentration buffer (0.8 M NaCl (Merck), 10 mM Tris-HCl (Sigma); pH 8.0, 1 mM EDTA (Merck), 1× Protease Inhibitor cocktail (Roche)) and incubated on ice for 15 minutes. This solution was then pelleted at 500*g* for 5 minutes and the pellet treated with 0.25 M HCl. The acid solution was then homogenised with a douncer pestle, followed by rotation for 1 hour at 4°C (Rotamix). To obtain the acid soluble histones, the homogenate was vacuum dried for approximately 2 hours (BACHOFER, vacuum concentrator). Proteins were re-suspended in SDS sample buffer (excluding 2-β mercaptoethanol and bromophenol blue). Modifications that were made to the published isolation procedure (Figure 2.1) were: including the TCA precipitation and acetone wash steps (as in Section 2.3.1, as these steps were not in the original method [49]) and exclusion of the acetone wash step from the modified method.

All histones isolated using the modified methods as well as the published methods had their protein concentrations determined using the Pierce® BCA Protein Assay kit (Section 2.4), after which 10 µg of protein was loaded onto a BioRad AnyKD ™ Mini-PROTEAN® TGX™ (Section 2.5), and analysed.

2.4. Protein concentration determination

Histone proteins showed poor solubility in water, thus a buffer that solubilises proteins, such as SDS sample buffer (0.063 M Tris-HCl (Sigma); pH 6.8, 8.7% ($^{v}/_{v}$) Glycerol (Sigma), 2% ($^{w}/_{v}$) SDS (Merck), 0.7% ($^{v}/_{v}$) 2-β mercaptoethanol and 0.2% ($^{w}/_{v}$) bromophenol blue (Sigma)) had to be used. The high concentrations of SDS in the buffer interfered with the Bradford reagent and thus other kits like the 2D Quant Kit (GE Healthcare, Life Sciences) and the Pierce® BCA Protein Assay kit (Thermo Scientific) were tried. The Pierce® BCA Protein Assay kit was chosen since the SDS concentration within the buffer is well tolerated in this assay. 2-β Mercaptoethanol and bromophenol blue were excluded during the protein concentration determination (as they interfered with the BCA reagent), and introduced only when the proteins had to be analysed on SDS-PAGE. The BCA assay is based on bicinchoninic acid for determination of protein content within a sample. A series of standards of known protein concentration was made using bovine serum albumin (BSA, Roche, 10 mg/ml) to give a range of concentrations from 25 to 2000 µg/ml. These standards were diluted with SDS sample buffer (without 2-β mercaptoethanol and bromophenol blue) to minimise the interference from the sample solution buffer and to give an accurate protein concentration determination. Each standard solution was pipetted into a 96-well microplate (10 µl, in triplicate) as well as the sample buffer, which served as a blank. Each sample of unknown protein concentration was also pipetted into the microplate (10 µl). Working reagent was made from two solutions, as provided in the kit, in a ratio of 50:1 (Part A: Part B). The working reagent (200 µl) was then added to each well containing either sample or standard solutions. The plate was incubated for 30 minutes to 2 hours at 37 °C, after which it was cooled to room temperature. The absorbance of each of the wells was measured using the Multiskan Ascent V1.24 Plate reader with a filter of 595 nm.

2.5 SDS-PAGE analysis

BioRad AnyKD™ Mini-PROTEAN® TGX™ precast gels were used for electrophoretic separation of proteins. Histone proteins in SDS sample buffer (excluding bromophenol blue and 2-β mercaptoethanol) were mixed in a 1:2 ratio with 0.7% ($^{v}/_{v}$) 2-β mercaptoethanol and 0.2% ($^{w}/_{v}$) bromophenol blue (Sigma), boiled at 95ºC for 5 minutes and loaded onto the gels. Gels were placed into the Mini-PROTEAN® Tetra and run for 25-30 minutes under a constant voltage of 200 V generated by PowerPac™ Basic. The running buffer used for electrophoresis had a composition of 25 mM Tris, 192 mM glycine, 0.1% SDS. Once separated, gels were stained overnight at 37ºC (0.1% Coomassie Brilliant Blue R-250 (Sigma) ($^{w}/_{v}$), 40% methanol (Merck) ($^{v}/_{v}$) and 10% acetic acid ($^{v}/_{v}$) (Merck)), followed by an overnight de-stain with 25% methanol. Gels were then stored in 1% acetic acid solution until analysis.

2.6 Comparison of the three histone isolation methods

On evaluation of the three published histone isolation methods and the modifications made to them, the modifications that comprised of effective sample preparation steps and returned high protein yields were chosen to isolate histones from untreated *P. falciparum* parasites. These isolated histones were evaluated using MS analysis. MS analysis required the isolation of high concentrations of histones from each of the three methods and 15 µg of histones from three biological repeats for each of the three isolation methods were loaded onto BioRad AnyKD™ Mini-PROTEAN® TGX™ gels. All of these samples were analysed at the Central Analytic Facilities (CAF, University of Stellenbosch) and compared in terms of the histones identified as well as the number of PTMs found. All data generated by Sequest in excel files (Section 2.6.1) were analysed manually for site localisation of PTMs. Significant PTMs were those that were identified in two or three of the biological repeats and therefore considered to be biologically relevant. Comparison of the optimal modified method to previously identified PTMs was then carried out.

2.6.1 LC/MS/MS analysis

This service was provided by Dr Salomie Smit at the CAF Facilities (University of Stellenbosch).

Gel pieces containing the proteins of interest were washed with water then 50% acetonitrile followed by ammonium bicarbonate. Gel pieces were then incubated in 100% acetonitrile after which they were reduced with DTT and alkylated with 55 mM iodoacetamide for 1 hour in the dark. The gel pieces were in-gel trypsin digested and the resulting peptides were extracted with 70% acetonitrile in 0.1% trifluoroacetic acid and then dried. The peptides were analysed with a Thermo Scientific EASY-nLC II connected to a LTQ Orbitrap Velos mass spectrometer (Thermo Scientific, Bremen, Germany) equipped with a nano-electrospray source. MS spectra were acquired from *m/z* 400- 2000. Thermo Proteome Discoverer 1.3 (Thermo Scientific, Bremen, Germany) was used to identify proteins via automated database searching (Sequest) of all tandem mass spectra against the PlasmoDB 7.2 database. Carbamidomethylation of cysteine was set as fixed modification, and deamidation, oxidation, acetylation, methylation and phosphorylation were set as dynamic modifications. Manual analysis of the histones and PTMs results were compared to identify the method that detected the highest number of histones and PTMs.

PART II

Identification of potential HDAC inhibitors against a homology model of PfHDAC1 using *in silico* screening and development of an HDAC activity assay to confirm the inhibitory activity on *Plasmodium* HDACs

2.7 IC_{50} determination of suberoylanilide hydroxamic acid (SAHA)

Suberoylanilide hydroxamic acid (SAHA) is a known HDAC inhibitor of human HDAC enzymes as well as *P. falciparum* parasite HDAC enzymes [85]. The IC_{50} against intra-erythrocytic *P. falciparum* 3D7 parasites was determined using the Malaria SYBR Green Fluorescence assay [110, 111].

Synchronised intra-erythrocytic ring stage parasites (1% parasitaemia, 2% haematocrit) were treated with various concentrations of SAHA (2 fold dilutions ranging from 1.6 µM to 12.5 nM). Parasites were treated with 0.5 µM chloroquine disulphate or left untreated for use as positive and negative controls for drug treatment, respectively. Treated cells were grown statically for 96 hours in 96-well plates in a gas chamber at 37ºC. The cell suspensions were subsequently re-suspended with equal volumes (100 µl) with SYBR Green I lysis buffer (0.2 µl/ml 10 000× SYBR Green I, (Invitrogen Inc); 20 mM Tris-HCl (Sigma), pH 7.5; 5 mM EDTA (Merck); 0.008% ($^{w}/_{v}$) saponin (Sigma); 0.08% ($^{v}/_{v}$) Triton X-100 (Sigma)) and incubated for 1 hour after which the fluorescence was measured using a Fluoroskan Ascent FL microplate fluorometer (Thermo Scientific, excitation at 485 nm and emission at 538 nm). After subtraction of the averaged fluorescence emitted from the negative control, the data was expressed as a percentage of the averaged untreated control, to give percentage parasite proliferation. Dose-response curves were derived for each compound from which the IC_{50} values could be determined. Non-linear regression and statistical analyses were performed using GraphPad Prism.

2.8 Comparison of histones isolated from treated and untreated *P. falciparum* parasites

A trophozoite stage culture (180 ml, ~7% parasitaemia), was split into two 90 ml cultures. One of the cultures was treated for 2 hours with approximately 2× IC_{50} (665 nM) of the known HDAC inhibitor, SAHA, and the other with 0.2% DMSO. After treatment, the parasites were split into three 30 ml cultures and isolated using all three histone isolation methods, with paired SAHA and DMSO treated samples. The protein concentrations of isolated histone samples were determined using the Pierce® BCA Protein Assay, and 25 µg of protein was loaded of each sample. All six samples were analysed at the CAF Proteomics Facility (University of Stellenbosch).

2.9 *In silico* screening to identify potential HDAC inhibitors

This work was done in collaboration with Dr P. Burger, postdoctoral fellow, University of Pretoria.

2.9.1 Homology Modelling

The sequences for the crystal structure of human HDAC2 complexed with an N-(2-aminophenyl) benzamide (PDB ID: 3MAX) as well as two other human HDAC crystal structures (PDB ID: 4A69 and 1T64) were obtained from the Protein Data Bank (PDB) [112]. Sequences for *P. falciparum* HDAC1 were obtained from PlasmoDB (PDB ID: PFI1260C). The PfHDAC1 sequence (PFI1260c) together with the three resolved HDAC crystal structure sequences (PDB IDs: 3MAX, 4A69 and 1T64) were aligned to the PFAM seed alignment, PF00850 using alignment program *Muscle* [113]. PfHDAC1 homology models were generated with Modeller 9v1 [114] using the human HDAC2 (PDB ID: 3MAX) as template. Protein quality assessment was performed with the PDBsum generator [115].

2.9.2 Protein Preparation

The PfHDAC1 homology model was prepared for docking using the Protein Preparation Wizard in the Maestro software suite (Schrodinger, Inc). Side-chain optimization and the assigning of protonation states was performed to relieve energetically unfavourable constraints, followed by an *Impref* minimization run with root mean square deviation (RMSD) set to 0.3 Å.

2.9.3 Library Preparation

Compounds from the MMV Malaria Box, which is a chemical library containing 200 drug- and 200 probe-like compounds [7], were prepared for docking using the Prepare Ligand module of the Discovery Studio 3.0 software suite (Accelrys Inc.). Ionisation states were calculated in the pH range 6.5 to 8.5 before tautomers and isomers were generated for all compounds resulting in 2971 entries that were subsequently used in docking studies.

2.9.4 Molecular Docking

N-(2-aminophenyl) benzamide was used to define the centre for the active site sphere and its radius set to 12 Å for all structures. CDOCKER was used as a docking method generating 10 random conformations from equilibration dynamics and minimisation of the starting ligand structure and refined using simulated annealing (ridged protein and flexible ligand) [116]. The final minimisation was performed using a full force field potential (CHARMm) and consisted of 50 steps of steepest descent followed by up to 200 steps of conjugate-gradient using an energy tolerance of 0.001 kcal/mol [116]. The top 2 poses were saved based on the total docking energy (including the

intra-molecular energy for ligands and the ligand-protein interactions) [116]. The binding free energies for the docked compounds were determined using the Calculate Binding Energy method in Discovery Studio 3.0 (Accelrys Inc.).

2.9.5 Compound Selection

The virtual screen resulted in 2827 poses and were sorted based on their best docking scores. A visual inspection of the top scoring compounds (CDOCKER energy <-30 kcal/mol) was performed and the compounds with the best docking scores were selected for *in vitro* screening.

2.10 Molecular properties of the screened compounds

The Malaria Box is a library of 400 compounds and their IC_{50} values against *P. falciparum* parasites, logP values and the molecular properties that contribute to Lipinski's rule of five, were provided by the MMV [7]. Lipinski's Rule of 5 is an indication of the drug-likeness of a molecule, in terms of hydrogen bond donors and acceptors, lipophilicity and molecular weight [117]. It states that a compound should have no more than 5 hydrogen bond donors (nitrogen or oxygen atoms with one or more hydrogen atoms); no more than 10 hydrogen bond acceptors (nitrogen or oxygen atoms); a molecular mass of less than 500 daltons; and an octanol-water partition coefficient log *P* not greater than 5 [117].

2.11 Screening of MMV compounds against intra-erythrocytic *P. falciparum* parasites

The ten compounds obtained from *in silico* screening were used in the Malaria SYBR Green Fluorescent assay described above (Section 2.7). Synchronised intra-erythrocytic ring stage parasites (1% parasitaemia, 2% haematocrit) were treated with 5 µM of each of the compounds for 96 hours. The remainder of the assay format including the controls included, re-suspension with lysis buffer, measuring of fluorescence and data analysis were performed as described in Section 2.7.

2.12 HDAC activity assays (Adapted from [118])

The HDAC activity assay [118] utilises a fluorogenic HDAC substrate, (*S*)-[5-Acetylamino-1-(4-methyl-2-oxo-2H-chromen-7-ylcarbamoyl)-pentyl]-carbamic acid benzyl ester [119], also termed Z-MAL (kindly provided by Prof M. Jung, University of Freiburg, Germany). This substrate has an acetyl group, which is removed by either the action of zinc-requiring HDACs or by sirtuins in the presence of NAD^+. Only the deacetylated substrate releases the fluorescent molecule, 7-amino-4-methylcoumarin (355/460 nm), upon trypsin cleavage. When HDACs are inhibited by HDAC inhibitors, the acetylated substrate is not recognised by trypsin, resulting in no fluorescence. The entire reaction scheme for the HDAC assay is depicted in Figure 2.2 below.

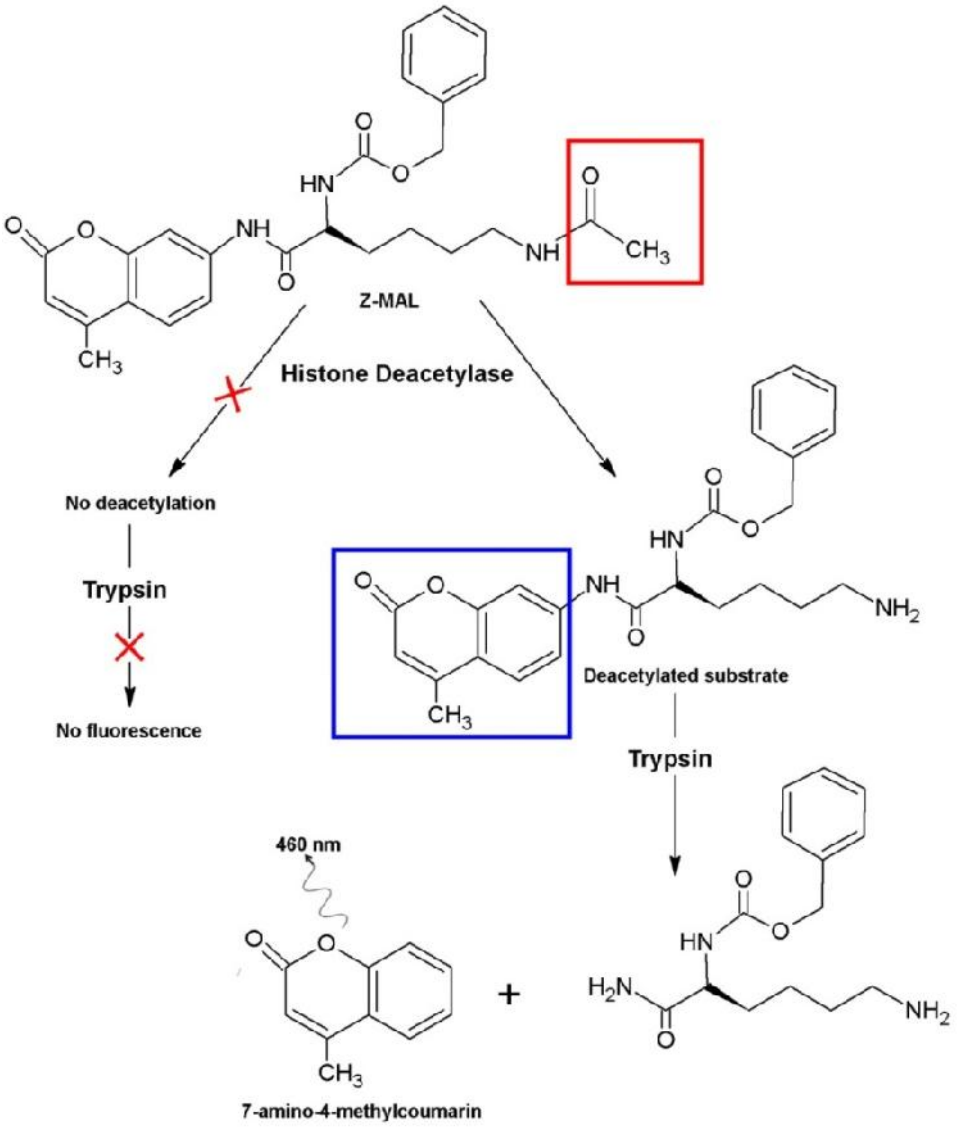


Figure 2.2: Reaction scheme of the HDAC assay using the non-isotopic substrate Z-MAL. Z-MAL is deacetylated by HDAC enzymes, and the deacetylated substrate can be cleaved by the protease trypsin to release the fluorescent molecule, 7-amino-4methylcoumarin (AMC). This molecule can then be detected with a fluorometer at 460 nm after it is excited at 355 nm.

2.12.1 HDAC assay on recombinant enzyme

The efficacy of this assay was determined on recombinantly expressed PfHDAC1 (Sigma). A 1:120 dilution of enzyme (0.48 µg/µl; Sigma) in assay buffer (45 mM Tris-HCl (Sigma); pH 8.0, 137 mM NaCl (Merck), 2.7 mM KCl (Merck); 1 mM $MgCl_2$ (Merck) and 0.1 mg/ml BSA (Roche)) was either treated with SAHA or assay buffer for 10 minutes at 37°C. These served as positive and negative controls of HDAC inhibition, respectively. The HDAC substrate, Z-MAL (0.126 mM), was added to each well containing enzyme (reaction sample) and to wells containing 110 µl of assay buffer (blank for subtraction of background). The final contents of the wells in a black 96-well microplate were therefore; 60 µl of enzyme solution (2 ng/µl of PfHDAC1), 45 µl of assay buffer; 5 µl of inhibitor or vehicle (positive and negative controls for HDAC inhibition respectively) and 10 µl of *Z*-MAL substrate (10.5 µM); giving a total well volume of 120 µl. A 100% control for substrate

cleavage was included by adding 10 µl of 0.126 mM 7-amino-4-methylcoumarin (AMC) to 110 µl of buffer. The reaction was incubated overnight at 37°C. A comparison of incubation length was also conducted i.e. 30 minute *vs.* overnight incubation at 37ºC. After this incubation, the reactions were terminated by the addition of a stop solution (2.75 µM SAHA and 1 mM trypsin in trypsin buffer; 50 mM Tris-HCl, 100 mM NaCl; pH 8.0). Reactions that were uninhibited are terminated due to the presence of SAHA in the stop solution while trypsin acts on the deacetylated substrate. The fluorescence was measured using a Fluoroskan Ascent FL microplate fluorometer (Thermo Scientific, excitation at 355 nm and emission at 460 nm) after a 20 minute incubation period at 37˚C. The data after subtraction of background were expressed as average relative fluorescence units (RFU) of at least two replicates, and when expressed as a percentage of 100% control, indicate the amount of substrate conversion. *P*-values were calculated with GraphPad InStat using an unpaired t-test.

2.12.1.1 Calculation of assay parameters for assay optimisation

The assay conditions to carry out further experiments were decided upon after calculating the signal to noise and signal to background ratios for each assay conducted. It was then found that an overnight incubation as compared to a 30 minute incubation period provided the highest signal to noise ratios.

Signal-to-noise ratios are calculated using the formula below [120]:

Signal to noise ratio= (Mean signal- Mean background)/(Std. Dev. of background)

This ratio takes into account the average fluorescent signal that is obtained from the uninhibited sample and subtracts the average fluorescent signal from the background control (Z-MAL and the assay buffer). In addition to this, the standard deviation of the background is also accounted for to overcome any large variations in the background noise. Signal to background ratio, on the other hand is less sensitive and only takes into account the average fluorescent units of the uninhibited samples and the blank, as can be seen in the formula below [120]:

Signal to background ratio= Mean signal/ Mean background

Both these parameters are important for the validation of the assay and to ensure that results obtained are easily detectable.

2.12.1.2 Testing MMV compounds on recombinant enzyme

Use of an overnight incubation step together with 2 µg of PfHDAC1 were the conditions used to test the effect of the ten *in silico* screened MMV compounds on PfHDAC1 activity. This involved

using two concentrations of the compounds (10 and 100 µM) and incubating them with the enzyme (as for SAHA) for 10 minutes prior to addition of substrate followed by an overnight incubation at 37°C. The plates were then removed, the reactions terminated and the fluorescence measured as above (Section 2.12.1)

2.12.2 HDAC assay on isolated *P. falciparum* parasites

2.12.2.1 Comparing various parasite concentrations for assay

Synchronisation of parasite cultures at the ring stage enables trophozoite stage parasites to be isolated on alternate days. Thus, a 20 ml trophozoite stage culture (7-10% parasitaemia, 5% haematocrit) was lysed with 0.05% saponin as above (Section 2.3.1) using a glucose-containing saline solution (125 mM NaCl (Merck), 5 mM KCl (Merck), 25 mM HEPES (Sigma), 20 mM glucose (Sigma), 1mM $MgCl_2$ (Merck)). The saponin isolated parasites were then washed until the supernatant remained clear. The parasite pellet was then re-suspended in 1 ml of saline solution and determination of parasites/ml was conducted using a haemocytometer (counting 10 fields of parasites in a 1:10 dilution of parasite solution). A number of dilutions were made of 1×10^6, 1×10^7, 1×10^8 and 2.5×10^8 parasites/ml in the phosphate assay buffer (Section 2.12.1). The microplate set up was as in Section 2.12.1 (replacing enzyme solution with parasite solution) using a single Z-MAL concentration (10.5 µM). Corresponding blanks containing phosphate buffer and the respective Z-MAL concentrations were included, as well as a 100% AMC control (10.5 µM). The plates were incubated, stop solution added and the fluorescence measured as in Section 2.12.1. The data were expressed as average relative fluorescence units (RFU) of at least two replicates ± SD as above, or as a percentage of 100% control to indicate the amount of substrate conversion.

2.12.2.2 Testing for HDAC inhibitory effects with SAHA and CQ

Parasites were prepared as described in section 2.13.2 and 1×10^8 parasites/ml were added in duplicate to a number of wells containing a range of varying SAHA or CQ concentrations (10 µM, 1 µM, 100 nM and 10 nM). Uninhibited parasites, a 100% control (10.5 µM AMC) as well as the blank containing phosphate buffer and Z-MAL (10.5 µM) were included. The plates were incubated overnight, the stop solution added and the fluorescence measured as described in section 2.12.1. The data, after subtraction of background, were expressed as a percentage of the uninhibited control of at least 2 replicates ± SEM after three independent experiments.

2.12.2.3 Testing the MMV compounds on isolated parasites in the HDAC assay

Using the favourable conditions determined, i.e. 1×10^8 parasites/ml, 10.5 µM Z-MAL, and an overnight incubation, the microplate was set up as above, using two concentrations of the ten MMV compounds (10 µM and 100 µM). The compounds were incubated for 10 minutes at 37^{o}C with the

parasites prior to substrate addition. Positive and negative controls for HDAC inhibition were included (SAHA or phosphate buffer respectively), and the reaction was maintained overnight at 37ºC. The data, after subtraction of background, were expressed as a percentage of the uninhibited control of at least 3 replicates ± SEM after three independent experiments.

2.12.2.4 Calculation of limit of blank and limit of detection

The limit of blank (LoB) and limit of detection (LoD) were calculated using the formulae described in literature [121]. The limit of blank is defined to be "highest *apparent* analyte concentration expected to be found when replicates of a blank sample containing no analyte are tested" [121]. In terms of these experiments, the LoB will refer to the highest RFU values for the blank samples. For calculation of LoB, two to six replicate values of the Z-MAL blank (10.5 µM) from a series of eight experiments were used, and their mean and standard deviation substituted into the formulae below:

$$LoB = mean_{blank} + 1.645(SD_{blank})$$

The value of 1.645 is used as it is the *Z*-multiplier for a 90% confidence interval, to obtain a 95% confidence interval for the LoB value, a *Z*-multiplier of 1.96 should be used, and to obtain a 99% confidence interval a *Z*-multiplier of 2.576 [122]. The LoB for all three confidence intervals were calculated.

LoD is defined as "the lowest analyte concentration likely to be reliably distinguished from the LoB and at which detection is feasible" [121]. Here, the LoD will refer to the lowest RFU values that distinguish analyte from the blank. To calculate the LoD values, the LoB values were used in the formula below:

$$LoD = LoB + 1.645(SD_{low\ concentration\ sample})$$

In this formula, the LoB value was substituted as well as the standard deviation of the samples with the lowest concentration sample (10.5 µM) using 1×10^8 parasites/ml from a series of 4 identical experiments were used.

Chapter 3

Results

PART I

Optimisation of a histone isolation method suitable to determine the effect of HDAC inhibitors on histone acetylation using MS analysis

3.1 Protein isolation and analysis

The aim of comparing and modifying the histone isolation procedures was to select an optimum method for the isolation of all the canonical histones (H2A, H2B, H3, H4) and their variants (H2A.Z, H3 centromeric, H3.3 and H2B.Z) for subsequent identification of their PTMs by mass spectrometry. This requires the separation of known concentrations of proteins using SDS-PAGE analysis. Following histone isolation, the protein concentration of each sample was determined using a BSA standard curve (Figure 3.1) generated using the Pierce® BCA Protein Assay kit. An R^2 value above 0.95 is representative of an accurate fit of the regression line to the data. Subsequently, a known protein concentration was quantitatively loaded onto SDS-PAGE gels (Figures 3.2-3.5).

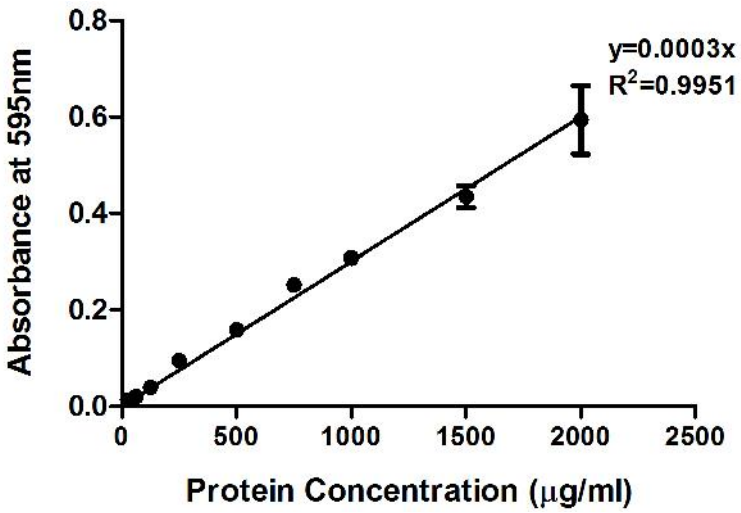


Figure 3.1: Representative BSA standard curve for protein concentration determination. This curve was obtained using a series of BSA standards with the Pierce® BCA Protein concentration kit. After incubating with a BCA colour-developing reagent for 30 minutes up to 2 hours, the absorbance is read at 595 nm.

3.2 Optimisation of histone isolation

Histone isolation methods were compared based on the respective protein concentrations obtained from each of the sample preparation methods. Protein yields needed to be 20 µg per sample for

subsequent MS analysis. The composition of sample buffers and the sample preparation steps were also taken into account. The solutions used in the various isolation techniques are described in Appendix Table 1A. Since histone concentration increases from the late trophozoite towards the schizont stages, probably due to DNA synthesis and concurrent nucleosome assembly [123], only trophozoite stage parasites were investigated. All histone isolations using the three different methods and their modifications were performed on homogeneous late trophozoite stage cultures with the same number of parasites/ml per sample, in order to derive comparable data. Table 3.1 shows the variations made to each method, their respective protein concentrations and the amount of protein yielded per sample.

For the Issar and Scherf method, the original method calls for a 2 hour up to an overnight acid extraction and therefore these two outlying acid extraction periods were investigated. The lack of TCA precipitation, as called for in the Trelle method [49], did not explain how acidic proteins were removed from the solution. Following personal communication with the author, Dr M. Trelle, it was found that histones in the acid solution were vacuum dried. The modification made to this method was the use of TCA precipitation, followed by inclusion of an acetone wash step. The original Merrick method called for two separate 1 hour acid extraction steps. The modification made to this was to perform a continuous 2 hour acid extraction step. The acetone wash step was excluded as a modification in all three histone methods to determine whether use of this step would affect protein yield.

Table 3.1: Protein yields and concentrations of histone samples isolated using three published methods and a number of modifications. Values are indicated as proteins isolated from a total of ~2× 10^8 parasites/ml and total protein yield in a final volume of 40 µl. Unmodified, original published methods are ndicated by the asterices.

		[Prot.] (µg/ml)	Yield (prot. (µg) in 40 µl)
Issar and Scherf	*Overnight acid extraction	902	36
	*2 hour acid extraction	827	33
	Exclusion of acetone wash step	747	30
Trelle	*Vacuum Dry	1767	71
	Inclusion of TCA precipitation and acetone wash step	557	22
	Exclusion of acetone wash step	857	34
Merrick	*2x 1 hour acid extractions	942	38
	2 hour acid extraction	1322	53
	Exclusion of acetone wash step	862	34

According to Table 3.1, modification of the Issar and Scherf method returned only a marginally different total protein yield (30 µg) as compared to the original methods (i.e. overnight and 2 hour acid extractions), which yielded 36 µg and 33 µg, respectively. The total protein yield for the Trelle protocol with vacuum dry variation was the highest compared to the other 8 samples at 71 µg

(Table 3.1). Comparatively, modifications made to the Trelle method (such as TCA precipitation and inclusion of acetone washing) resulted in the lowest protein concentrations overall. For the Merrick protocol, the use of two separate acid extraction procedures led to a decreased protein yield (38 µg, Table 3.1), as compared to a single acid extraction step for the same duration (53 µg). Exclusion of the acetone wash step yielded similar protein concentrations for all the samples (~30 µg Table 3.1). From Table 3.1, it can be seen that any of these isolation methods could be used as they yield sufficient protein for downstream MS analysis. However, to determine whether the protein contained within each sample correlates to possible histone proteins, 10 µg of each sample was analysed using SDS-PAGE. In the figures below, the modifications made to each method were compared and three histone isolation methods were chosen for evaluation with MS.

The ability of the modified and unmodified (published) Issar and Scherf methods to isolate histones was subsequently evaluated with SDS-PAGE (Figure 3.2).

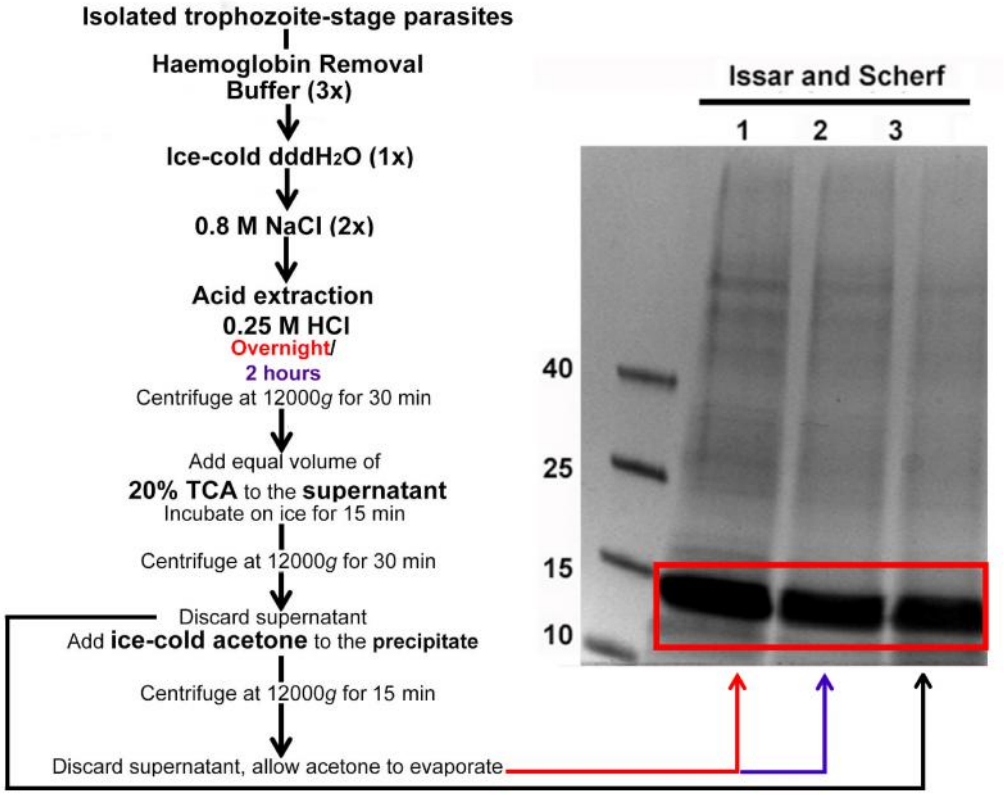


Figure 3.2: SDS-PAGE analysis showing comparison of the published Issar and Scherf method (Methods in Malaria Research, www.mr4.org) and modifications made to it. Text in bold in the flow diagram is indicative of the sample preparation step as well as the fraction it was performed on. Possible histones were loaded onto a BioRad AnyKD™ Mini-PROTEAN TGX™ gel. Isolated histones are indicated by the red box. The published method included; performing an overnight acid extraction (10 µg, Lane 1) or a 2 hour acid extraction step (10 µg, Lane 2), while exclusion of the acetone wash step (10 µg, Lane 3) was the modification made to the published method.

In Figure 3.2 it can be seen that there are bands corresponding to 15-10 kDa; the known sizes of histone proteins. Thus, in terms of isolation of possible histone proteins, all three method variations seemed to perform equally well. The 2 hour acid extraction modification (Lane 2, Figure 3.2) was

preferred over the overnight acid extraction (Lane 1, Figure 3.2), as acid labile modifications were more likely to stay intact with the shortened extraction procedure [124]. Furthermore, lane 2 can also be compared to lane 3 to determine whether any observable differences are present between the inclusion and exclusion of the acetone wash step. There seemed to be no observable difference on the SDS-PAGE gel and thus the use of acetone was preferred since it assists in the removal of any TCA salts and sample drying, which may impact downstream MS analysis. Thus for further analysis, the Issar and Scherf method with the 2 hour acid extraction, together with the acetone wash step was used.

When analysing the isolated proteins obtained with the modified and unmodified (published) Trelle method, it was found that all three samples contained possible histone proteins (Figure 3.3).

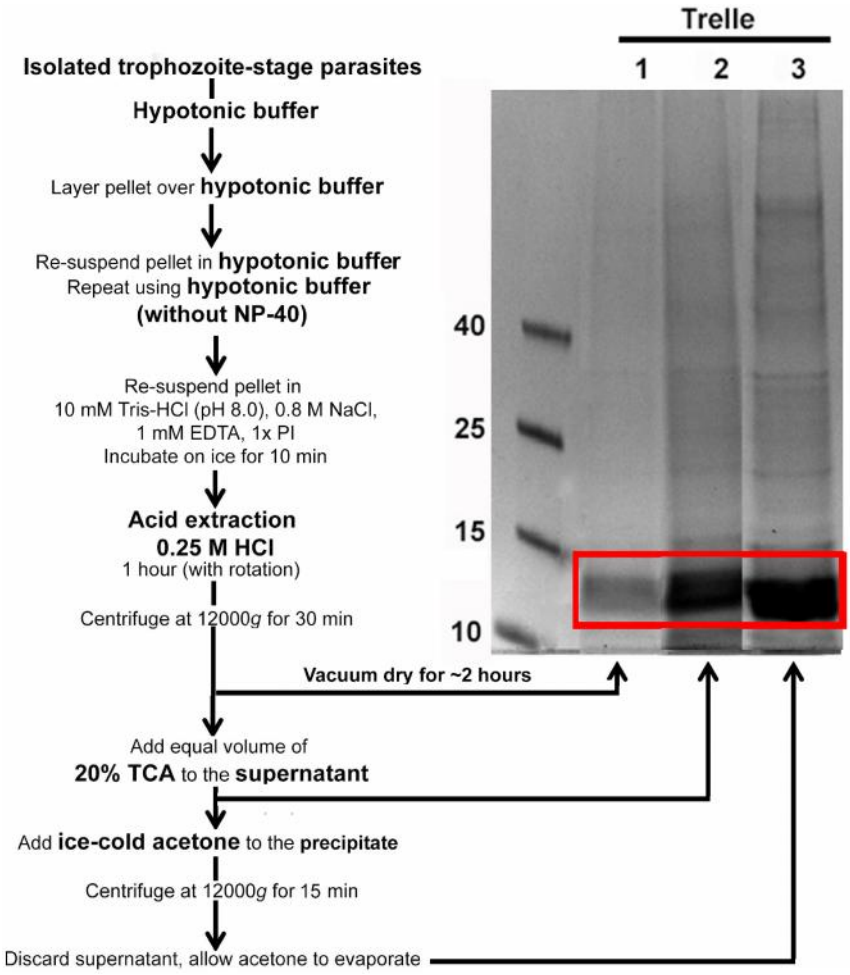


Figure 3.3: SDS-PAGE analysis showing comparison of the published Trelle method [49] and modifications made to it. Text in bold in the flow diagram is indicative of the sample preparation step as well as the fraction it was performed on. Histones were loaded onto a BioRad AnyKD™ Mini-PROTEAN TGX™ gel. Possible isolated histones are indicated by the red box. The published Trelle method [49] utilises vacuum dried acid extracted histones (10 µg, Lane 1), Modifications made to this method included TCA precipitation and acetone washing as additional steps to the protocol (10 µg, Lane 2) and performing the TCA precipitation step while excluding the acetone wash step (10 µg, Lane 3).

A discrepancy was seen only with the vacuum dried sample where 10 µg was loaded but this did not correlate to what was seen on the SDS-PAGE gel (Lane 1, Figure 3.3). Since prolonged exposure to acid means that acid-labile modifications could be removed, TCA precipitation was preferred immediately after acid extraction of the nuclear pellet, thus the modification made to this method was to include the TCA precipitation and acetone wash step (Lane 2, Figure 3.3). Exclusion of the acetone wash step showed no observable difference in the possible histone proteins isolated. However, for the sake of downstream analysis, acetone washing is preferred.

Similar to what was seen with the Issar and Scherf and Trelle methods, the Merrick methods were also capable of isolating proteins, which correlate to the size of histones (Figure 3.4). Here too, the inclusion of the acetone wash step showed no observable, deleterious effects on the proteins yielded (Lane 2, Figure 3.4). A continuous 2 hour extraction step was favoured compared to two separate 1 hour acid extractions as the sample handling was lowered and had greater ease of use.

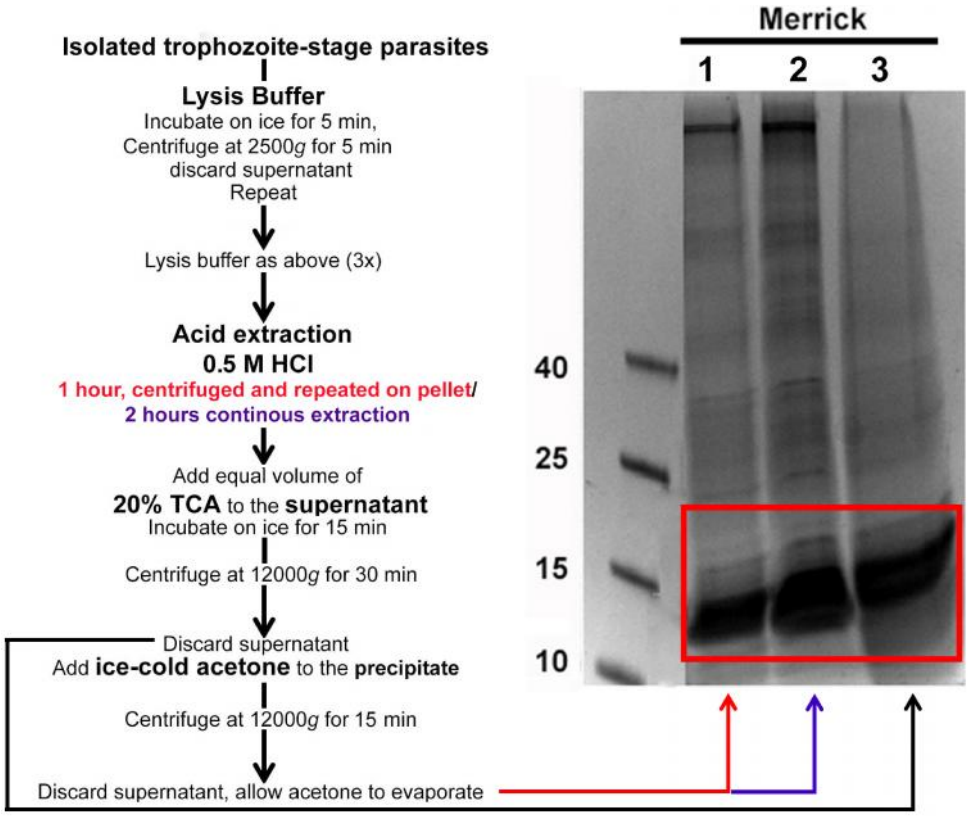


Figure 3.4: SDS-PAGE analysis showing comparison of the published Merrick method [86, 109] and modifications made to it. Text in bold in the flow diagram is indicative of the sample preparation step as well as the fraction it was performed on. Possible histone proteins were loaded onto a BioRad AnyKD™ Mini-PROTEAN TGX™ gel. Isolated histones are indicated by the red box. Histones isolated using the published Merrick method [86, 109] where two 1 hour acid extractions were performed are shown in Lane 1 (10 µg). Performing a continuous 2 hour acid extraction step (10 µg, Lane 2), and exclusion of an acetone wash step on isolated histones (10 µg, Lane 3) were the two modifications made to the published Merrick method.

Since all the methods and modifications made to them resulted in the isolation of possible histone proteins as evaluated by SDS-PAGE analysis, the sample preparation steps were also considered. For the Trelle protocol, use of vacuum dried sample was not feasible, as there appeared to be

inaccurate protein quantification and vacuum drying calls for increased exposure of the proteins to acid. For the Issar and Scherf protocol an overnight acid extraction step could remove acid labile modifications and the protein concentration did not differ greatly from the other 2 samples, i.e. 36 *vs*. ~30 µg. For the Merrick method, using two separate acid extraction steps led to increased sample handling. Acetone washing needs to be included to facilitate effective MS analysis.

Therefore, based on the results obtained, methods were chosen to evaluate further their efficacy in MS identification of histone proteins. For the Issar and Scherf method, this included the 2 hour acid extraction step; for the Trelle method, all modifications made were included (TCA precipitation and acetone washing) and for the Merrick protocol, a continuous 2 hour acid extraction method was used.

3.3 Determining optimal histone isolation method with MS analysis of biological repeats

Three histone isolation methods that were determined to yield sufficiently high histone concentrations and comprised of effective sample preparation steps (Section 3.2) were used for histone isolation from trophozoite stage parasites. Each of the three isolation methods were used for histone isolation on three biologically independent trophozoite cultures and analysed with SDS-PAGE (Figure 3.5). In a representative gel depicting two of the biological repeats (Figure 3.5), the potential histone proteins are highlighted. These possible histone proteins identified on the SDS-PAGE gel were then excised from the gel and compared with MS analysis.

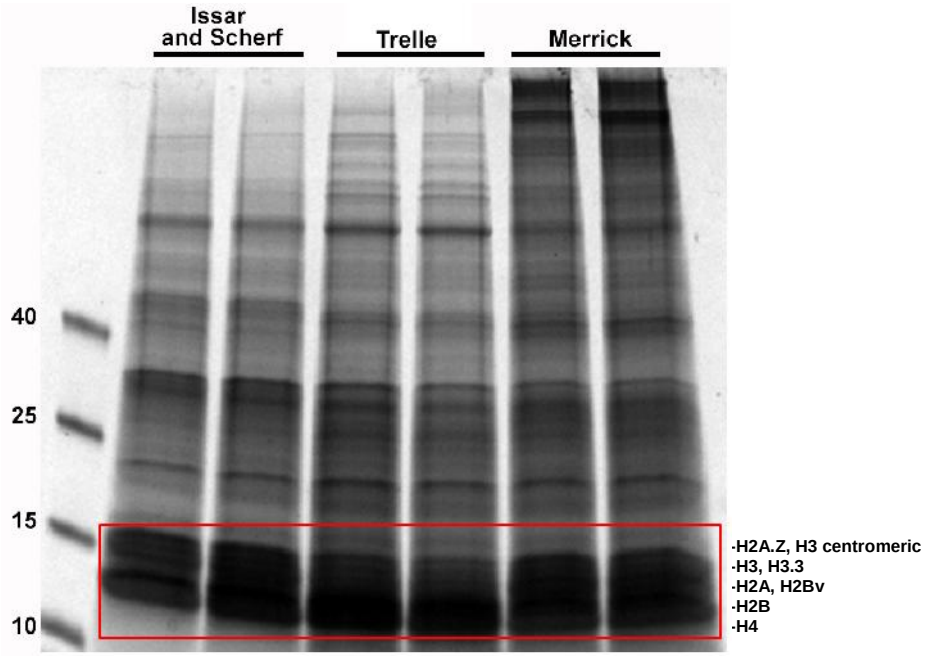


Figure 3.5: SDS-PAGE analysis of histones isolated using three different isolation methods. Possible histone proteins (15 µg/lane) were loaded onto a BioRad AnyKD™ Mini-PROTEAN TGX™ gel. Lanes are demarcated according to the methods used for histone isolation, i.e. Issar and Scherf, modified Merrick and modified Trelle methods. Potential histone proteins are highlighted (red box). Approximate positions of core and variant histone proteins are indicated to the right of the gel.

Within the putative histone protein bands, the lower bands should correspond to H2B and H4, while the middle bands to H2A and H2Bv and the top bands to H3, H3.3, H2A.Z and H3 centromeric [49, 50]. The faint line in each of the protocols is indicative of H2A.Z and the low abundant H3 centromeric. The Issar and Scherf method has a more intense band around the position of H2B, while the Merrick method appears to have uniform enrichment for all the protein bands that are present in the region. For the modified Trelle method, the bands corresponding to H2B and H4 have a more intense staining than the remainder of the bands.

To confirm the identity of the proteins highlighted (Figure 3.5), MS analysis was performed. Table 3.2 below shows, which histone proteins were identified as well as their sequence coverage in each of the biological repeats. Sequence coverage refers to the peptides identified that span the length of the parent sequence. A higher coverage provides greater confidence that peptide identification was accurate.

Table 3.2: MS analysis of each histone isolated using three histone isolation methods. Absence of histone proteins is indicated by a dash (-). Average coverage is calculated by dividing the sum of the obtained coverage percentages for each repeat divided by the number of repeats that the histone was identified in. Data is from three independent biological replicates.

	Issar and Scherf				**Trelle**				**Merrick**			
	Coverage (%)			**Ave.**	**Coverage (%)**			**Ave.**	**Coverage (%)**			**Ave.**
H2A	71	46	46	**54**	61	73	58	**64**	90	66	17	**58**
H2A.Z	71	78	39	**63**	73	72	49	**65**	96	78	45	**73**
H2B	97	95	38	**77**	100	92	47	**80**	94	98	51	**81**
H2B.Z	89	95	36	**73**	89	94	40	**74**	84	92	26	**67**
H3.1	60	61	37	**53**	61	62	35	**53**	65	64	29	**53**
H3.3	63	69	29	**54**	67	61	35	**54**	73	66	29	**56**
H3 cen	**-**	**-**	14	**14**	19	47	-	**33**	-	-	-	**-**
H4	90	90	59	**80**	96	87	51	**78**	84	93	42	**73**

The modified Trelle method detected all eight *P. falciparum* histones in all the biological replicates except for H3 centromeric, which was detected in two of the three biological replicates. However, H3 centromeric was not identified at all with the modified Merrick protocol and was only detected once with the modified Issar and Scherf protocol (Table 3.2). The average sequence coverage for the remaining histones ranged between 50-80% and showed low variation between the three protocols.

Once the isolated proteins were correctly identified with MS as histones, the PTMs further identified with MS on the various peptides were mapped onto their corresponding parent sequence. To reiterate, in this study, three biological replicates were used for each modified isolation method and the histones and their PTMs identified in each of the three biological replicates per method. This uniquely allowed determination of the biological relevance of each

PTM identified. To our knowledge, this is the first study that used biological replicates for assignment of a PTM's significance. As such, if a PTM was detected and identified in only one of the biological samples, this was considered not representative of the biological situation and therefore described as non-significant (Appendix Table 2A). However, PTMs that were consistently identified in at least two of the biological samples were seen to be a true reflection of PTMs occurring in the samples and therefore having biological relevance and were subsequently described as significant (Table 3.3). This was done for each of the three modified histone isolation protocols, independently. Due to the complexity of the large number of modifications, only the significant modifications were taken into account in Table 3.3. For a more detailed overview, all modifications are provided in Appendix-Table 3A.

Table 3.3: MS analysis detection of histone PTMs detected on histones isolated using three different isolation methods (n=3). I+S=Issar and Scherf. Trelle and Merrick represent methods, which have been modified from these original publications. Highly significant modifications are defined as present in all three biological repeats for each method (red) and significant if detected in two biological repeats (orange). The number of significant modifications are presented as a fraction of the total number of modifications (in bold) from all three biological repeats. The total number of modifications includes the non-significant modifications detected as well, which are presented in Appendix-Table 3A.

	Acetylation			Monomethylation			Dimethylation			Trimethylation			Phosphorylation		
H2A	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
	N-term	N-term	N-term		K3	K3		K3	K3	K3	K3	K3	S1	S1	S1
	K3	K3	K3	K5	K5	K5	K5	K5	K5		K5		T6	T6	T6
	K5	K5	K5	R 8			R 8	R 8	R 8			K9	S12	S12	S12
	K9	K9	K9	K9	K9	K9	K9	K9	K9	K10		K10	T15	T15	T15
	K10	K10	K10	K10	K10	K10	K10	K10	K10	K13	K13	K13	S16	S16	S16
		K13		K13	K13	K13	K13	K13	K13	K20	K20	K20	S18	S18	S18
	K20	K20		K20	K20		K20	K20	K20				S76	S76	S76
			K35		R 29		K75						T79	T79	T79
			K41			K38		R 77							
	K75	K75	K75		R 77			R 81	R 81						
				R 88	R 88	R 88									
	7/9	**8/8**	**8/11**	**7/11**	**9/12**	**7/14**	**7/14**	**9/15**	**8/14**	**4/9**	**4/8**	**5/7**	**8/8**	**8/9**	**8/8**
	Acetylation			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
H2A.Z	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
	K10	K10	K10			K14		K18		K10		K10	T30	T30	T30
	K14	K14	K14		K24		K27			K14			S32	S32	S32
	K18	K18	K18	K27	K27		K29	K29	K29	K18			T35	T35	T35
	K24	K24	K24	K29	K29	K29	K34	K34	K34			K24	S41	S41	S41
	K27	K27	K27	K34	K34	K34	K36	K36	K36	K27	K27	K27	S44	S44	S44
	K29	K29	K29	K36	K36	K36	K37	K37	K37	K29	K29	K29			S61
	K34	K34	K34	K37	K37	K37	R 42	R 42	R 42	K34	K34	K34	S64		
	K36	K36	K36	R 42	R 42	R 42		R 45	R 45	K36	K36	K36			S65
	K37	K37	K37	R 45	R 45	R 45	R 54			K37	K37	K37	T105		T105
	K153				R 54									T149	T149
					K60										
					K155										
	10/13	**9/15**	**9/17**	**7/13**	**11/17**	**7/14**	**7/16**	**7/14**	**6/13**	**8/9**	**5/10**	**7/8**	**7/10**	**6/9**	**9/11**
	Acetylation			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
H2B	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
			K3		K3				K7		K7		S2	S2	
		K4			K4		K9	K9				K9	T11	T11	T11
	K7		K7	K9				K10	K10	K10	K10	K10	T13	T13	T13

			K9		K10	K10	K18	K18			K18	K18	S24	S24	
	K10	K10		K18	K18	K18	K19	K19	K19		K19	K19	T44	T44	
	K18	K18	K18			K19			R 21	K20			S47		
	K19	K19	K19		K20		K23			K112		K112			S82
	K20		K20	R 21			R 48				K116		S83		
		K22		K23			K100							T88	
	K64	K64	K64			R 48	K116		K116				T111		
		K74			K108								T114	S115	
	K108	K108			K112										
		K112			K116										
	7/15	**9/15**	**7/17**	**4/14**	**8/15**	**4/15**	**7/14**	**4/16**	**5/15**	**3/6**	**5/11**	**5/11**	**9/16**	**7/15**	**3/12**
H2B.Z	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
		N-term	N-term	K8	K8		K3		K3	K8	K8		S1	S1	S1
	K3	K3	K3	K13	K13		K8	K8	K8	K13	K13	K13	S9	S9	S9
	K8	K8	K8			K14	K14	K14	K14	K14	K14	K14	T15	T15	T15
	K13	K13	K13			K18	K18	K18	K18	K18		K18	T19	T19	T19
	K14	K14	K14		R 23	R 23		R 23	R 23	K42			T30	T30	
	K18	K18	K18	K25	K25			K25	K25		K52			S34	
		K25			R 26		R 26		R 26		K104			T48	
	K27			R 88				K27						T51	
	K42			R 95	R 95	R 95			R 68					T84	
	K112						R 88							S86	
		K104	K104											S87	
															T84
													S87		S87
													T92		T92
														S108	
	8/12	**8/11**	**7/11**	**5/13**	**6/17**	**4/13**	**6/15**	**6/17**	**8/12**	**5/11**	**5/9**	**3/7**	**7/12**	**12/17**	**7/10**
H3 cen	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
											K23			S21	
											K26			T24	
														T27	
														S38	
														S39	
	0/1	**0/5**		**0/3**	**0/5**		**0/2**	**0/5**			**2/4**		**0/4**	**5/8**	
H3.1	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
		N-term			R 2	R 2	R 2	R 2				K4	T3	T3	T3

		R 2		K4	K4	K4	K4		K4	K9		K9	T6	T6	T6
	K4	K4	K4	R 8	R 8	R 8	R 8	R 8	R 8	K14	K14	K14	S10	S10	S10
	K9	K9	K9	K9	K9	K9	K9	K9	K9		K18		T11	T11	T11
	K14	K14	K14	K14	K14	K14		K14	K14	K23	K23		S22	S22	S22
	K18	K18	K18	R 17	R 17	R 17	R 17	R 17	R 17	K27	K27		S28	S28	
	K23	K23	K23	K18	K18	K18	K18	K18	K18				S32	S32	S32
	K27	K27	K27	K23	K23		K23	K23	K23					T33	
		K36	K36		R 26	R 26			R 26						S57
					K27			K27							T58
				K36		K36			K36						
				K37											
	6/9	**8/9**	**7/11**	**9/13**	**10/12**	**9/15**	**7/10**	**8/14**	**9/15**	**4/8**	**4/6**	**3/6**	**7/8**	**8/10**	**8/11**

H3.3	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
		N-term			R 2	R 2	R 2	R 2		K4		K4	T3	T3	T3
	K4	K4	K4		K4	K4	K4		K4	K9	K9	K9	T6	T6	T6
	K9	K9	K9	R 8	R 8	R 8	R 8	R 8			K14	K14	S10	S10	S10
	K14	K14	K14	K9	K9	K9	K9	K9	K9				T11	T11	T11
	K18	K18	K18	K14	K14	K14	K14	K14	K14	K23		K23	S22	S22	S22
	K23	K23	K23	R 17	R 17	R 17	R 17	R 17	R 17	K27	K27	K27	S28	S28	S28
	K27	K27	K27	K18	K18	K18	K18	K18	K18	K36			S32		S32
		K36	K36	K23	K23	K23		K23			K37	K37		T33	T33
		K53			R 26			K27							T45
	K115				K27	K27		K36							
					K36				R 49						
				K37											
	7/11	**9/12**	**7/10**	**7/13**	**11/18**	**9/16**	**7/12**	**9/16**	**6/16**	**5/7**	**4/6**	**6/9**	**7/11**	**7/11**	**9/11**

H4	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
			N-term	R 3		R 3	R 3		R 3		K5		S1	S1	S1
	K5	K5	K5			K5	K5	K5	K5		K12	K12			
	K8	K8	K8		K8	K8	K8		K8	K16			S47		
	K12	K12	K12	K12	K12	K12	K12	K12	K12	K20	K20	K20			T80
	K16	K16	K16		R 35	R 35	K16	K16			K31				T82
		K31	K31	R 39					K31				S89	S89	S89
			K44	R 45					R 35						
	K79		K79	K67			R 40								
				R 78					K77						
					K91		K79								
	6/14	**7/14**	**9/16**	**6/14**	**4/10**	**5/13**	**7/14**	**3/6**	**7/11**	**2/7**	**4/4**	**2/5**	**3/7**	**2/4**	**4/6**

From Table 3.3 it can be seen that a vast majority of the modifications were indeed identified as significant. For clarity, these data were further simplified in Table 3.4 where the following was summarised: 1) the total number of PTMs identified using each isolation method; 2) the significant number of PTMs identified using each isolation method; 3) the different PTM types, i.e. acetylation, methylation (mono-, di- and tri-) and phosphorylation, expressed as percentages of either the total number of PTMs or the number of significant modifications.

Table 3.4: Summary of PTMs identified with MS using the three modified histone isolation methods. The total number of PTMs and the number of significant PTMs (those present in >2 of the biological replicates) identified by MS in each method is indicated. Additionally, the percentage distribution of the various PTM types in relation to the total PTMs identified are indicated.

	Issar and Scherf				Modified Trelle				Modified Merrick			
			Sig.				Sig.				Sig.	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Total No. of PTMs	409		222		435		250		406		226	
Acetylation	80	20	50	23	85	20	56	22	88	22	53	23
Monomethylation	94	23	45	20	106	24	59	24	100	25	45	20
Dimethylation	97	24	48	22	103	24	46	18	96	24	49	22
Trimethylation	57	14	31	14	58	13	34	14	53	13	31	14
Phosphorylation	81	20	48	22	83	19	55	22	69	17	48	21

The Issar and Scherf and modified Merrick methods identified approximately the same number of modifications in total (409 *vs.* 406 PTMs) as well as those identified as significant (222 *vs.* 226, respectively). By contrast, the modified Trelle method detected a larger number of PTMs in totality (435 *vs.* ~400 from the other two methods) and was also able to more successfully identify PTMs with biological significance (250 significant PTMs *vs.* ~220 for the other two methods). Thus, it is evident that the modified Trelle method is capable of isolating histones with the highest number of modifications.

The various types of PTMs that were detected with MS, i.e. acetylation, methylation (mono-, di- and tri-) and phosphorylation are presented in Table 3.4 and for visual clarity a comparison of the significant percentages are graphically presented in Figure 3.6. This graphical distribution of PTM types in each of the three protocols show that with the exception of trimethylation, there was no bias or skewing towards the detection of a certain type of PTM. When looking at only the significant PTMs, no drastic difference was seen in the acetylation profile of histones isolated with the three protocols (22% for Issar and Scherf and modified Trelle methods and 21% for the Merrick method). The Issar and Scherf and modified Merrick methods detected identical percentages for mono- and dimethylation (each at 20% and 22%, respectively). The modified Trelle method differed marginally from the other two methods; detecting a higher amount of monomethylation (24% *vs.* 20%) with a lower detection of dimethylation (18% *vs.* 22%). Trimethylation percentages were identical for all

three methods (14%) and highly similar for phosphorylation (22% for Issar and Scherf and modified Trelle methods and 21% for the modified Merrick method).

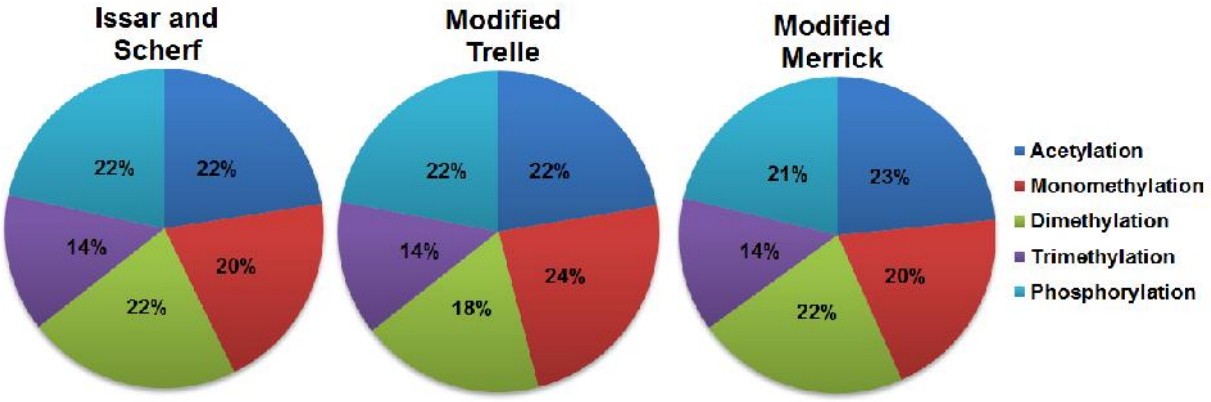


Figure 3.6: Percentage distribution of the various PTM types identified on histones isolated using three different histone isolation methods. The percentage of acetylation, methylation (mono-, di- and tri-) as well as phosphorylation modifications detected significantly (≥2 times) is depicted.

Based on the data depicted in Table 3.4, the modified Trelle method detected the highest number of PTMs, both in totality and significantly (435 and 250 PTMs, respectively). PTMs identified on histones isolated with the modified Trelle method were subsequently compared to the total of 100 modifications previously identified in various literature reports [37, 49, 50, 65, 66] (Figure 1.3). This was performed both at the level of total PTMs but also comparing the PTMs assigned as having biological significance in the modified Trelle method. Of the 100 PTMs previously detected in various literature reports, only 20 PTMs were not detected using the modified Trelle method (Figure 3.7). In addition, a further 16 were also not detected significantly. The remaining 64 PTMs identified previously were all also detected with the modified Trelle method, and these were assigned biological significance with the latter, indicating indeed that these are core modifications that will have biological relevance in *P. falciparum* histones. However, the modified Trelle method was able to identify an additional 186 novel, significant PTMs that have not been described previously (Figure 3.7).

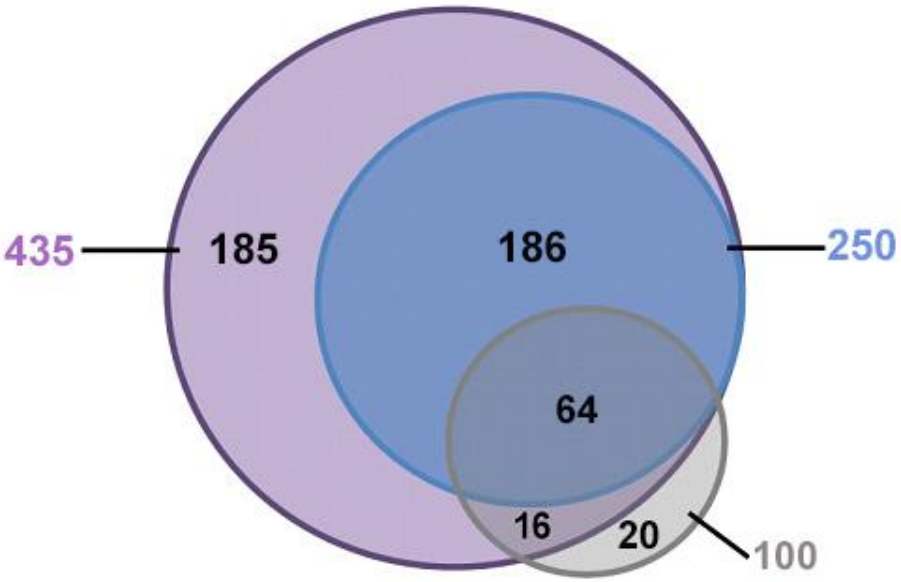


Figure 3.7: Venn diagram showing a comparison of the total and significant number of PTMs detected with the modified Trelle method to those previously identified in literature. The largest purple bubble is representative of the total 435 PTMs detected using the modified Trelle method. From these 435 PTMs, the 250 PTMs that were detected significantly are shown with the blue circle. The 100 known PTMs that were previously identified are shown with the grey circle. Numbers in black are representative of the difference found after taking into account the overlapped regions.

A comprehensive comparison of the known PTMs, their literature references and whether they were detected using the modified Trelle method are presented in Appendix-Table 4A. Together with Figure 3.7, this data is also graphically presented in Figure 3.8 where the PTMs previously identified in literature are indicated on each of the histones and those that were significantly detected with the modified Trelle method are highlighted in yellow and those detected non-significantly are highlighted in green. Thus, of the 100 known PTMs in literature, 80 are highlighted; 16 (non-significant detection) are indicated in green, 64 (significant detection) are in yellow and 20 are not highlighted (Figure 3.8). Furthermore, the modifications made to the published Trelle method improved on the histone PTMs detected. Whereas the published Trelle method only detected 44 modifications when used for MS analysis previously [49], 250 PTMs were detected using the modified Trelle method.

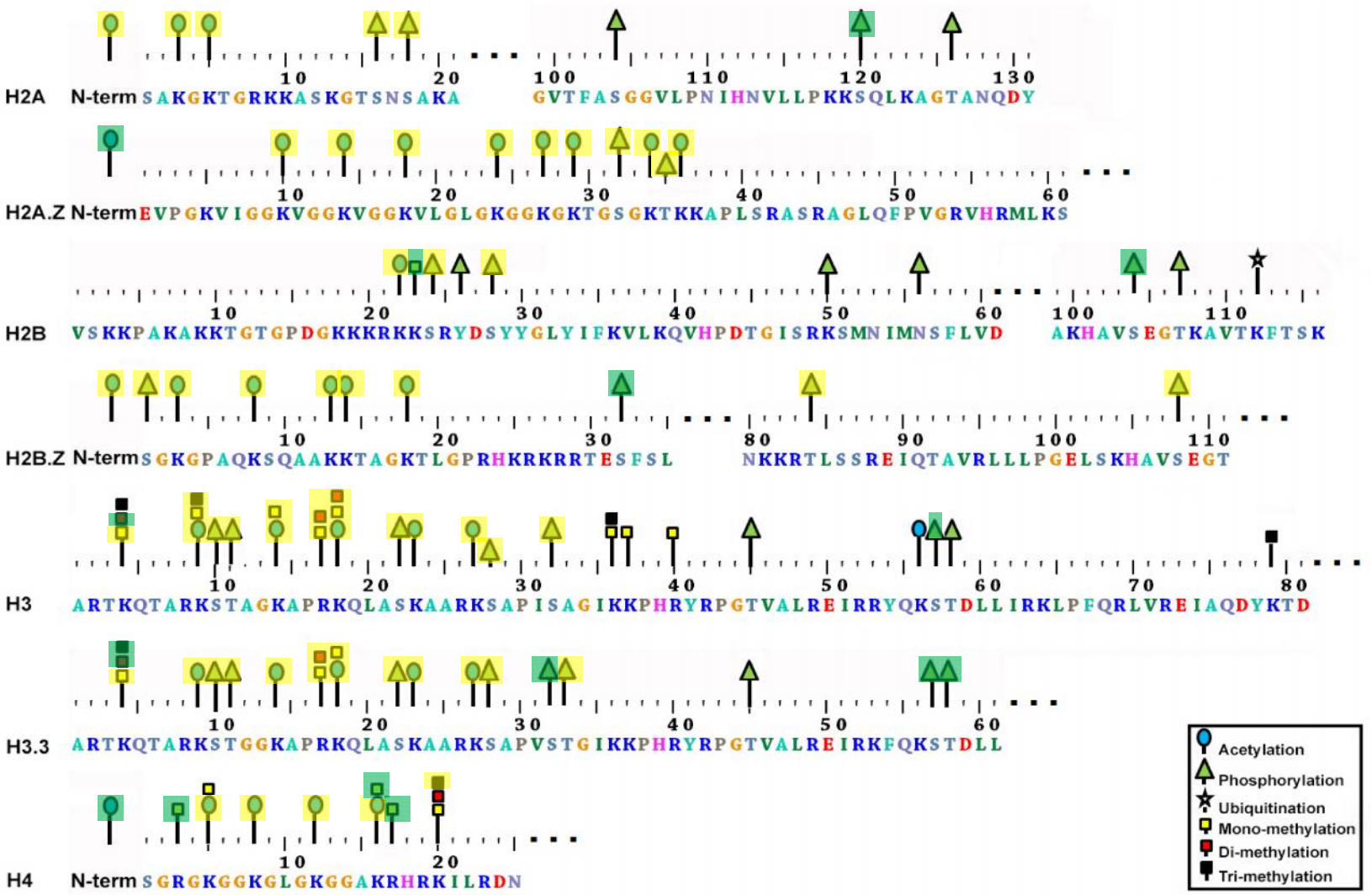


Figure 3.8: Comparison of the PTMs identified in the modified Trelle method as compared to literature. The PTMs detected by the modified Trelle method that correlate to those found in literature are highlighted. The PTMs that were detected significantly are highlighted in yellow and those detected non-significantly are highlighted in green. Information compiled from [37, 49, 50, 65, 66].

When evaluating the various methods and the results they returned, the modified Trelle method was chosen for further experiments due to the following: 1) it effectively identified all 8 histone proteins; 2) returned protein concentrations, which were sufficient for subsequent MS analysis; 3) showed no drastic skewing or bias for any of the various PTM types (compared to the other two isolation methods); 4) identified the highest number of PTMs and 5) has significantly detected 64 previously identified PTMs. Furthermore, in terms of sample preparation, the use of protease inhibitors, the shortened acid extraction procedure and the ease of use of this protocol make for an effective and rapid histone isolation method.

PART II

Identification of potential HDAC inhibitors against a homology model of PfHDAC1 using *in silico* screening and development of an HDAC activity assay to confirm the inhibitory activity on *Plasmodium* HDACs

3.4 IC_{50} determination of SAHA

To confirm previously published IC_{50} values of SAHA [81, 85, 96] against the *P. falciparum* 3D7 strain of parasites cultured in our lab, the Malaria SYBR Green Fluorescent assay [110, 125] was used to establish a dose-response curve thereof. The IC_{50} is the concentration at which a compound is able to inhibit the *in vitro* proliferation of *P. falciparum* parasites by 50%. A dose-response curve was obtained after SAHA treatment (1.6 µM to 12.5 nM, Figure 3.9) resulting in an IC_{50} value of 80.8 ± 6.6 nM for SAHA (n=3 ± SEM), corresponding to the IC_{50} of SAHA from literature (100-300 nM).

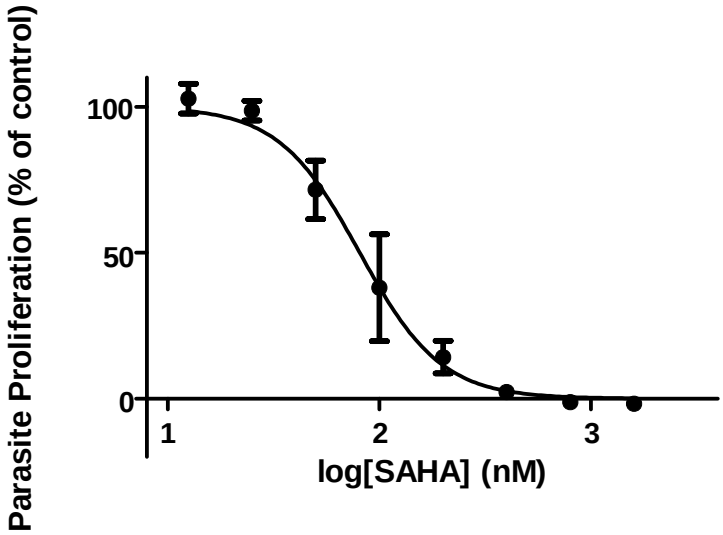


Figure 3.9: Dose-response curve of a known HDAC inhibitor, SAHA, against intra-erythrocytic *P. falciparum* 3D7 parasites. The effect of SAHA on parasite proliferation was determined using the MSF assay after 96 hour incubation. The compound was serially diluted from a 1.6 µM to 12.5 nM concentration. Data are representative of three independent experiments performed in six technical repeats with standard error of the mean (n=3, ±SEM).

3.5 Comparison of histone PTMs from treated and untreated *P. falciparum* parasites

Differences in histone PTMs have been known to occur following treatment with HDAC inhibitors. Generally, these differences are quantified using Western blot techniques with antibodies directed against tetra-acetylated H4 and H3K9Ac. Immunodetection shows increased antibody binding to these PTMs on histones isolated from treated parasites. Moreover, the presence of a greater number of these acetylated residues, known as hyperacetylation, has been found to occur as early as 2 hour after treatment of *P. falciparum* parasites with SAHA and other HDAC inhibitors [30, 126]. Thus, histone proteins were isolated after a 2 hour *in vitro* treatment of *P. falciparum* parasites with either a known HDAC inhibitor, SAHA (665 nM), or 0.2% DMSO (vehicle-treated control). After SDS-PAGE analysis and excision of bands corresponding to potential histone proteins (Figure 3.10, red box), MS analysis was performed to determine whether differences in histone PTM landscapes could be detected. The isolation method that was chosen was the modified Trelle method, as discussed in Section 3.2 above.

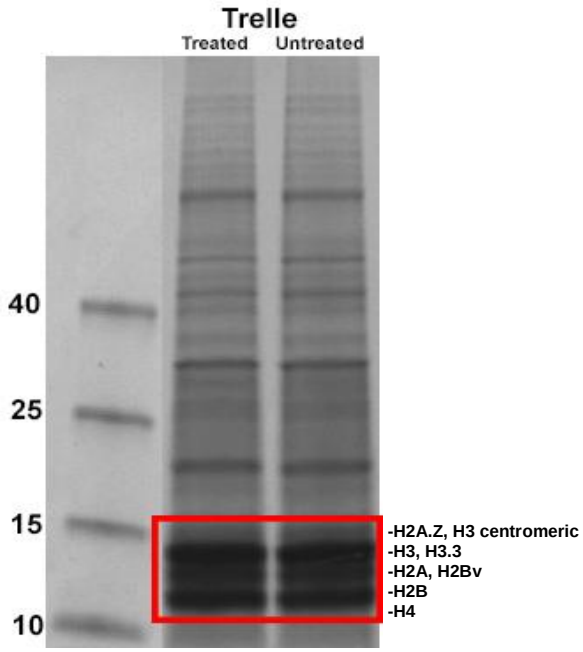


Figure 3.10: SDS-PAGE analysis of histones isolated from treated and untreated *P. falciparum* parasites. Histones were isolated using the modified Trelle method (Section 3.2) from *P. falciparum* trophozoite stage parasites after a 2 hour incubation with the known HDAC inhibitor, SAHA (625 nM, treated) or with 0.2% DMSO (untreated/ vehicle control). Twenty five micrograms of protein were quantitatively loaded onto a BioRad AnyKD ™ Mini-PROTEAN® TGX™ gel. Approximate distributions of the eight histone proteins are indicated to the right of the gel.

In Figure 3.10 no observable difference in the band distribution could be seen between the histones isolated from both treated and untreated *P. falciparum* parasites. Both samples appeared to have enrichment for histones H3.1, H3.3 and H4 as compared to the other histones. The histone PTMs identified with MS analysis on the various peptides detected were then mapped to the corresponding histone parent sequence. A tabulated version of each modification detected is made available in Appendix-Table 4A, where differences between treated and untreated parasite histones are highlighted. This data is graphically represented (Figure 3.12) in order to visualise the

different PTM landscapes from treated and untreated parasite histones. To explain the means in which the graphical representations were generated, Figure 3.11 depicts the steps taken for data analysis. Acetylation is known to impede methylation, since the Ɛ-amino group of lysine is bound to the carbonyl carbon of the acetyl group, leaving no available bonds for further modification. Methylation occurs on the lysine Ɛ-amino group as well as on the nitrogen within the guanidinium group of arginine. Methylation occurs in succession, with monomethylation occurring first, followed by di- and trimethylation. Thus as in Figure 3.11, if a trimethyl group was detected on a residue, this residue must have been previously mono- and dimethylated. The graphical depiction whereby a residue has an acetyl group as well as a trimethyl group is therefore an indication that two peptides were found with either one of the mutually exclusive PTMs.

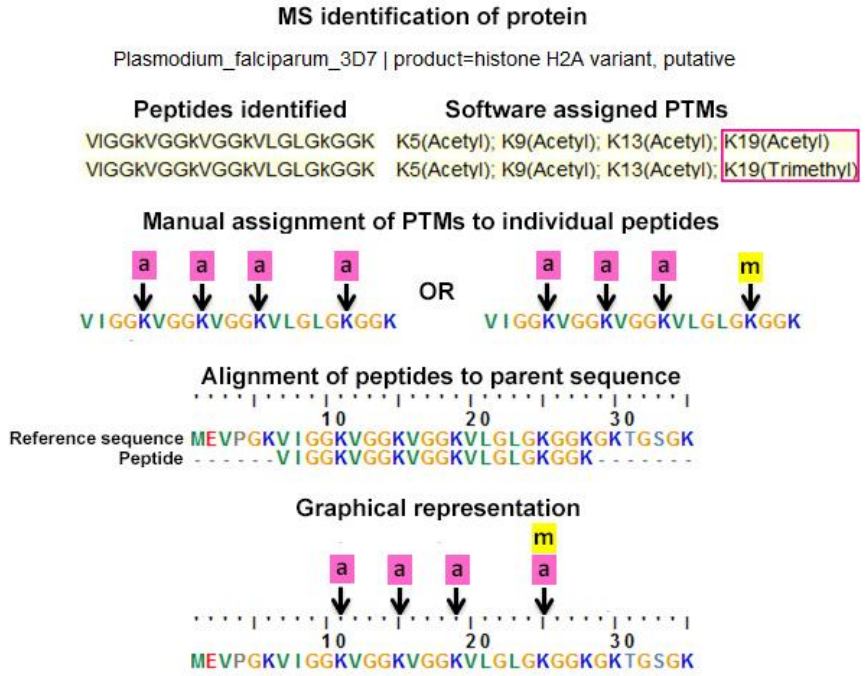
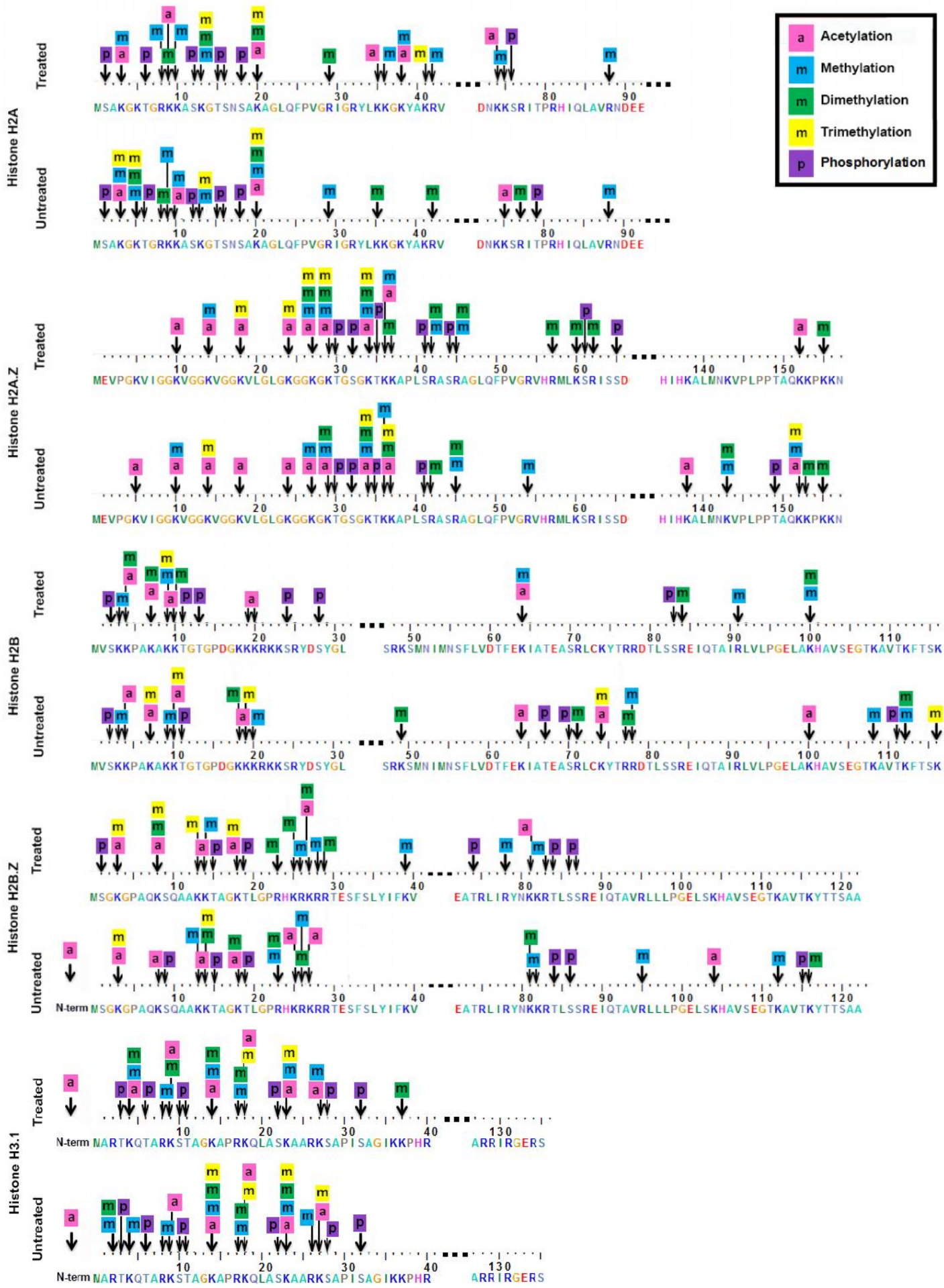


Figure 3.11: Graphical depiction of the process undertaken for assignment of PTMs to the parent sequence. Histones were identified based on the peptides identified and the sequence coverage obtained. PTMs were assigned by MS software onto identified peptides followed by manual assignment of the PTMs to their respective residues. An example of two mutually exclusive PTMs, are the acetyl and trimethyl groups on K19 (highlighted in pink box). Thus, when these peptides were aligned to the reference sequence the two PTMs were depicted above one another.

Using the process depicted in Figure 3.11, the PTM landscape of histones isolated from treated or untreated parasites is graphically represented and evaluated for possible differences, which may have occurred due to treatment (Figure 3.12). It should be noted that only one biological replicate was used for this evaluation.



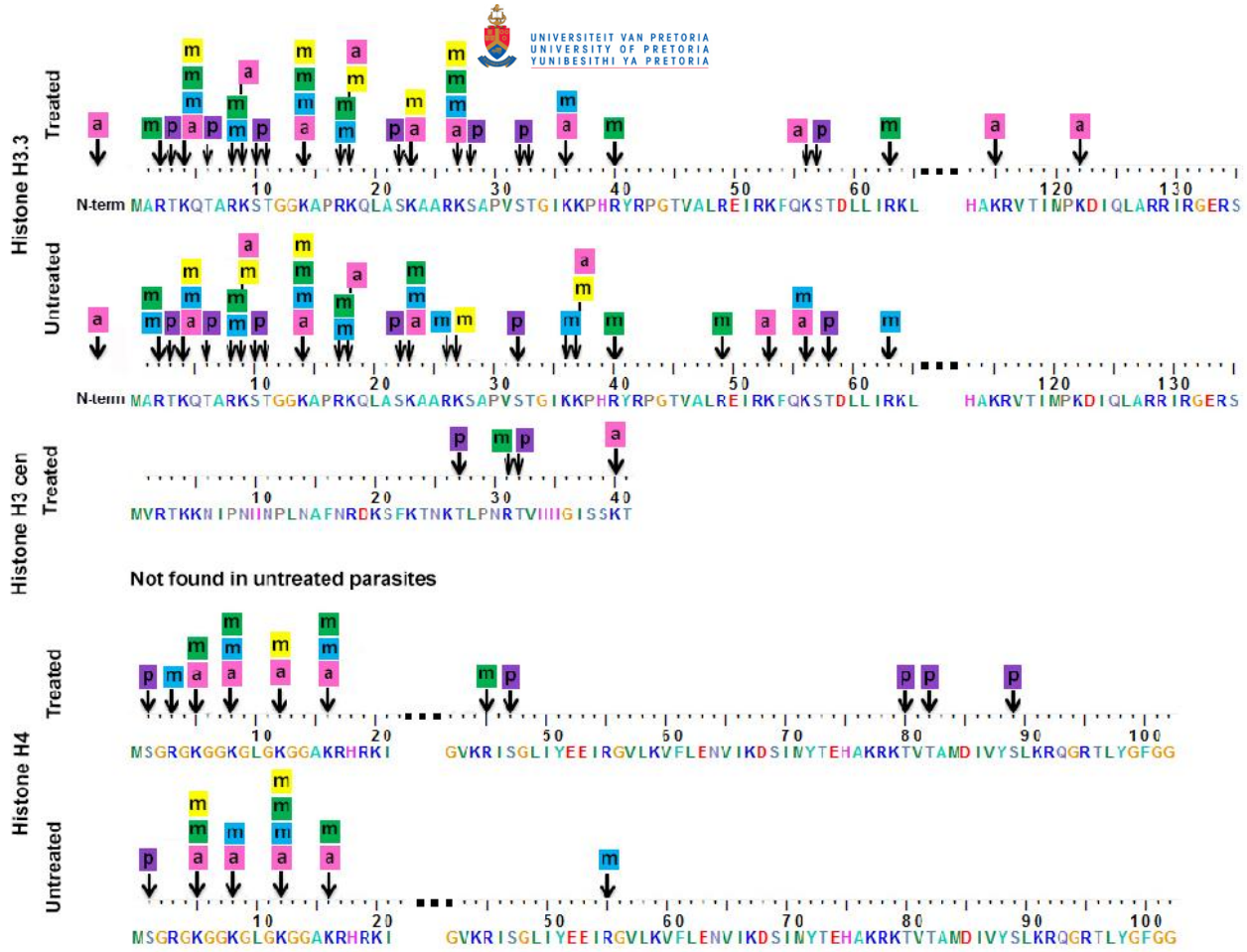


Figure 3.12: PTMs detected on histones isolated from treated or untreated *P. falciparum* parasites. Histones were treated with a known HDAC inhibitor, SAHA (625 nM) or with vehicle DMSO (0.2%) for 2 hours (treated and untreated, respectively). Histones were isolated using the modified Trelle method, analysed with SDS-PAGE, followed by MS analysis of bands containing possible histone proteins. Histone PTMs were mapped onto their corresponding parent sequence as depicted in Figure 3.11. Acetylation occurs as a mutually exclusive PTM compared to other PTMs and where trimethylation was detected, it is preceded by mono- and dimethylation. Ellipses are indicative of portions of sequence where no PTMs were found and were thus excluded for simplicity.

On comparison of the histones present in treated and untreated samples, Histone H3 centromeric was not detected in the untreated sample. A key observation when looking at these proteins was that the majority of the PTMs occurred at the N-termini and this was where most of the differences were detected when comparing the 2 samples. For ease of analysis, only marks known to have epigenetic consequences will be highlighted. An example of this is the tetra-acetylation of H4 (K5, K8, K12 and K16), which serves as activation marks for transcription. It was seen that the acetyl groups remained unchanged in both treated and untreated samples, though there were more methyl groups on the untreated samples in these positions as compared to the treated. Similarly, H3K9ac and H3K14ac, two other activation marks, exhibited unchanged acetylation but additional methyl groups were seen in these positions for the treated samples at K9 and for untreated samples at K14. Looking at marks associated with *var* genes on H3.1, histones from untreated parasites exhibited differences at both K4 and K9, where an additional methyl group was found, altering the methyl profile from mono- to dimethylation. The repression mark, H3K9me3 was not present in both samples. Other known marks that were not present in either sample were H3K4me3, H3K56ac, H3K79me3, H4K20me1, H4K20me3 and H4R3me2. In addition to looking at

the landscapes of the histones, an overview table of the two histone samples is presented in Table 3.5. This table contains the following data; 1) The total number of histone PTMs detected for each sample; 2) the number of histones identified for each sample and 3) the number of each different type of modification detected in the two samples.

Table 3.5: Number of histone proteins and PTMs (detected by MS analysis), isolated from SAHA-treated or untreated *P. falciparum* parasites.

	Modified Trelle Method	
	Treated	**Untreated**
Acetylation	52	52
Monomethylation	51	60
Dimethylation	47	45
Trimethylation	21	27
Phosphorylation	50	37
Total	236	225
No of histones	8	7

The number of acetylated residues remains unchanged between treated and untreated samples (Table 3.5). The greatest difference observed was that of phosphorylated residues, where treated samples had 13 more phosphorylated residues than untreated (50 *vs.* 37 respectively). This is followed by increased mono- and trimethylation in untreated samples (a difference of 9 and 6, respectively) and decreased dimethylation in untreated samples (a difference of 2). In totality, more PTMs were detected on the treated samples as compared to the untreated. No other quantifiable data were obtained on these histone proteins. Therefore, it can be seen that for the treated and untreated parasite histones, no major differences occur between the number of PTMs but there are some vast differences on the PTM landscapes of these two samples.

3.6 *In silico* screening to identify potential HDAC inhibitors

As depicted in Figure 1.4, potential HDAC inhibitors should have two crucial characteristics; firstly, binding of the zinc molecule required for the enzyme's activity and secondly, a capping group (chemical structural group that can be modified for enzyme recognition) that binds the enzyme selectively in the rim near the surface of the active site. Thus, to identify putative HDAC inhibitors, *in silico* analyses were used. This technique is beneficial as compounds are detected faster by virtual screening of a library of potential inhibitors using the set criteria conforming to basic properties of such inhibitors.

3.6.1 Generation of a PfHDAC1 homology model for screening of MMV Malaria Box

No crystal structure exists for PfHDAC1 and therefore, a homology model based on a human class I HDAC enzyme had to be built. The HDAC protein family alignment showed a 62% sequence identity and 80% sequence similarity between PfHDAC1 and the human HDAC2 (PDB ID: 3MAX;

2.05 Å resolution), which were subsequently used in model generation (Figure 3.13). Protein quality assessment of the homology models showed an overall G value (statistical value of geometric parameters) of -0.09 for the PfHDAC1 model and 0.14 for the human HDC2 (PDB ID: 3MAX; scores above -0.5 are considered as acceptable). The human HDAC2 enzyme was crystallised with the HDAC inhibitor *N*-(2-amino-5-substituted phenyl) benzamide, as well as calcium and zinc ions in the enzyme active site, and these molecules are shown in the homology model below.

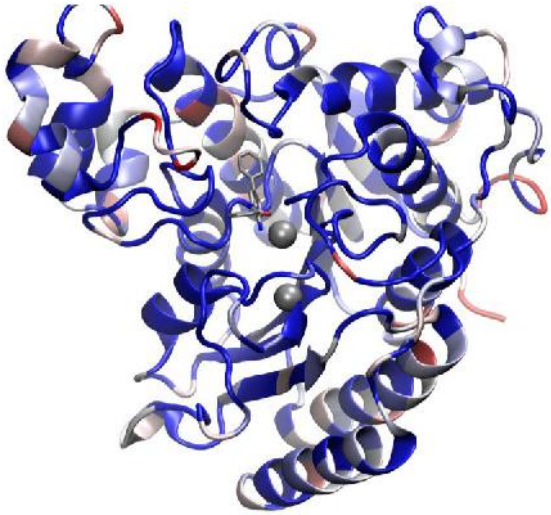
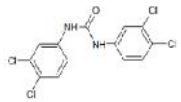
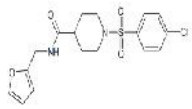
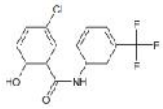
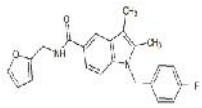
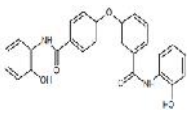
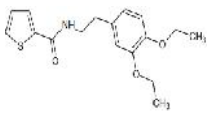
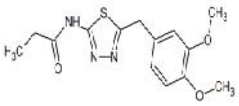
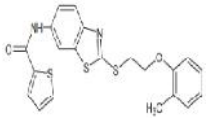
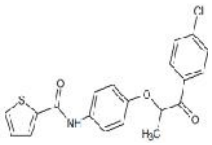
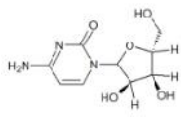
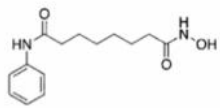


Figure 3.13: Homology model of PfHDAC1 using Human HDAC2 (PDB ID: 3MAX) as a template. Dark blue regions indicate identical residues, light blue and white highlights conserved residues and regions in red and light red indicate non-conserved residues (Blosum65). The active site contains the stick form of *N*-(2-amino-5-substituted phenyl) benzamide together with bound zinc and calcium.

The homology model was then used to screen the MMV Malaria Box, a collection of 400 compounds, of which half are drug-like and the other half, probe-like [7]. Ten compounds were subsequently selected based on their CDOCKER scores (less than -30 kcal/mol) and visual inspection of the compounds. Many of the 10 compounds that were considered as “hits” (Table 3.6) have amide groups (i.e. carboxamides, proponamide and benzamides). Compound 10 was an exception as this compound lacked the amide group and instead had a pyrimidinone. Since CDOCKER scores take into account hydrogen bonding, hydrophobic effects or interactions as well as ligand strain (energy of the ligand conformation), the CDOCKER scores are an indication of how well a molecule may bind to the enzyme active site. A cut off of -30 kcal/mol was chosen and compounds that returned scores of higher than -30 kcal/mol were considered to be poor binders. Compounds 1 and 5 returned the best CDOCKER scores (-13.1 and -14.2 kcal/mol, respectively), while the remainder of the compounds exhibited scores in the same range as SAHA (~-10 kcal/mol). Since *in silico* docking is not always an indication of compounds that are actual inhibitors of the active site, visual inspection of the top-scoring compounds was required. A few criteria had to be taken into consideration; 1) general poses, this is an indication of how a compound binds into the enzyme active site; 2) if the compound interacts with residues known to be involved in interactions within the active site; 3) exclusion of large molecules (500 kDa cut off) and long,

flexible molecules that gave better docking scores since they bind non-specifically to the protein due to their size and ability to conform to the active site.

Table 3.6: Top scoring compounds selected from screening the MMV Malaria Box against a PfHDAC1 homology model. Designated compound numbers are mentioned in brackets.

Structure	Compound Name	Chemical Name	CDOCKER Score (kcal/mol)
	**SJ000159330 (1)**	5-(2,4-dichlorophenyl)-*N*-(2-hydroxyphenyl)furan-2-carboxamide	-13.1
	**GNF-Pf-4692 (2)**	1-[(4-chlorophenyl)sulfonyl]-*N*-(furan-2-ylmethyl)piperidine-4-carboxamide	-9.8
	**SJ000010289(3)**	5-chloro-2-hydroxy-*N*-[3-(trifluoromethyl)phenyl]benzamide	-10.4
	**GNF-Pf-4602 (4)**	1-[(4-fluorophenyl)methyl]-*N*-(furan-2-ylmethyl)-2,3-dimethylindole-5-carboxamide	-10.0
	**TCMDC-125220 (5)**	*N*-(2-hydroxyphenyl)-3-[4-[(2-hydroxyphenyl)carbamoyl]phenoxy]benzamide	-14.2
	**GNF-Pf-4376(6)**	*N*-[2-(3,4-diethoxyphenyl)ethyl]thiophene-2-carboxamide	-9.8
	**TCMDC-125741(7)**	*N*-[5-[(3,4-dimethoxyphenyl)methyl]-1,3,4-thiadiazol-2-yl]proponamide	-8.8
	**GNF-Pf-3312 (8)**	*N*-[2-[2-(2-methylphenoxy)ethylsulfanyl]-1,3-benzothiazol-6-yl]thiophene-2-carboxamide	-10.4
	**TCMDC-125438 (9)**	*N*-(4-{[(2*R*)-1-(4-chlorophenyl)-1-oxopropan-2-yl]oxy}phenyl)thiophene-2-carboxamide	-10.0
	**MMV638723 (10)**	4-amino-1-[(3*R*,4*S*,5*R*)-3,4-dihydroxy-5-(hydroxymethyl)oxolan-2-yl]pyrimidin-2-one	-9.3
	**SAHA**	Suberoylanilide hydroxamic acid	-10.2

3.6.2 Biological properties of the ten potential HDAC inhibitors

For the compounds in the MMV Malaria Box, the logP values, IC_{50} values against *P. falciparum* parasites and their properties as drug-like or probe-like molecules had already been determined by the MMV and their partners. These properties are displayed in Table 3.7 below. These compounds have IC_{50} values against *P. falciparum* parasites in the low micromolar range, showing their potency against the parasite at minimal concentrations. In general, a compound should have a balanced hydrophobicity-hydrophilicity relationship, as this affects its ADME properties. Hydrogen bond donors and acceptors play an important role in determining the hydrophobicity and hydrophilicity of a compound. LogP is a partition coefficient of a compound between an organic and aqueous phase at a pH where all compounds are in the neutral form [117]. The molecular weight is an important factor to determine the permeability of the compound through the lipid bilayers [117]. Lipinski's rule of five assists in this regard by taking into account all these factors contributing to a compound's drug-likeness. Thus, drug-like compounds have a logP value of less than 5, a molecular weight of not more than 500, less than 5 hydrogen bond donors and less than 10 hydrogen bond acceptors [117], enabling them to be suitable drug candidates. Probe-like compounds on the other hand are said to have lower molecular weight, lower hydrophobicity and higher solubility [127], or they violate Lipinski's rule of five.

Table 3.7: Biological properties of the ten potential HDAC inhibitors identified by *in silico* screening
All biological properties were determined by the MMV.

Comp	MW (g/mol)	IC_{50} [µM]			HBD	HBA	LogP[a]	Ro5 viol.[b]	Prop.[c]
		3D7	W2	K1					
1	348.18	2.5	-	-	2	4	4.906	0	Drug-like
2	382.9	0.5-0.9	0.4	-	1	6	2.056	0	Drug-like
3	315.7	1.0	-	1.20	2	3	3.979	0	Drug-like
4	376.4	0.08	0.4	-	1	4	4.723	0	Drug-like
5	440.4	0.8	-	-	4	7	4.475	0	Drug-like
6	319.4	0.9	0.6	-	1	4	3.561	0	Drug-like
7	307.4	0.9	-	-	1	6	2.018	0	Drug-like
8	426.6	0.5-0.7	1.6	1.04	1	4	5.853	1	Probe-like
9	385.9	0.8-1.03	-	-	1	4	4.956	0	Probe-like
10	243.22	0.4	-	-	4	8	-2.396	0	Probe-like

[a] **logP: Partition coefficient of the compound between an organic phase (octanol) and water, it is a measure of the lipophilicity**
[b] **Lipinski's rule of five (Ro5): Number of violations of this rule**
[c] **Properties of the compounds:** This is an indication of whether the molecule is drug-like or probe-like based on its molecular characteristics.

From the 10 compounds, only GNF-Pf-3312 violates Lipinski's rule of five (Table 3.7), due to the fact that its logP value is greater than 5. Similarly, MMV638723 and TCMDC-125438 are also probe-like molecules (Table 3.7), which have the advantage of being structurally diverse from the remaining molecules [7]. Seven of the compounds are drug-like (Table 3.7), meaning that these

"compounds have sufficiently acceptable ADME properties and sufficiently acceptable toxicity properties to survive through the completion of human Phase I clinical trials" [128].

3.7 Determining inhibition of *P. falciparum* parasite proliferation with the ten potential HDAC inhibitors

In order for compounds in the MMV Malaria Box to be included into the library, they had to exhibit activity equal to or lower than 4 µM against intra-erythrocytic *P. falciparum* parasites [7]. This was determined (by the MMV's partners) using a high throughput fluorescent high content imaging assay [129]. With this in mind, the potential inhibitors obtained from *in silico* screening against PfHDAC1 were screened using a fluorescence-based assay for their ability to inhibit the proliferation of intra-erythrocytic *P. falciparum* parasites at 5 µM. Figure 3.14 shows that of the ten compounds screened, six inhibited parasite proliferation by more than 30%, with four of the compounds (compounds 2, 3, 5 and 6) showing a significant inhibition of parasite proliferation (P<0.001, n=3, paired student t-test). Compounds 4 and 10 inhibited parasite proliferation by ≥30%. Compounds 1 and 7 showed virtually no inhibition of parasite proliferation while compounds 8 and 9 showed a 10% and 20% inhibition, respectively.

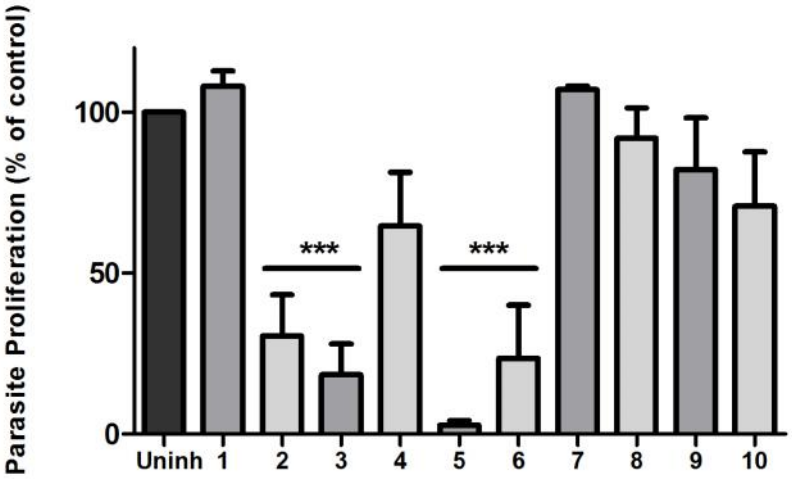


Figure 3.14: The inhibitory effect of a series of 10 putative HDAC inhibitors on intra-erythrocytic *P. falciparum* parasites. Parasite proliferation was monitored in the presence of 5 µM of each of the ten compounds from the MMV Malaria Box at 37°C for 96 hours. DNA levels were measured as relative fluorescence units based on SYBR Green I binding. Data are expressed as a percentage of the untreated control to determine cell proliferation, n=3 ±SEM. Significance was indicated at P<0.001 (***), Student's t-test (paired).

3.8 HDAC activity assays against recombinant PfHDAC1

Ensuing the observed difference in IC_{50} values when tested on intra-erythrocytic *P. falciparum* parasites, the compounds were tested against a recombinant enzyme (PfHDAC1) using an assay adapted from Heltweg *et al.* [118]. This *in vitro* screening of the compounds allows for the confirmation of their mechanism of action against a particular protein target. The HDAC assay

reaction scheme shown in Figure 2.2 involves the use of a synthetic acetylated HDAC substrate, trypsin (for cleavage of deacetylated substrate molecules), and the fluorescent molecule AMC. The substrate Z-MAL is able to be cleaved by both class 1 and 2 human HDAC families [118] and since the homology between *Plasmodium* and human class 1 HDACs are high [90], this specification was applied to *P. falciparum* HDACs as well. The assay was optimised for use against the recombinant PfHDAC1 to confirm *in silico* screening results.

3.8.1 Comparison of negative and positive controls of HDAC inhibition

One of the initial steps taken in the development of this assay was to determine the efficacy of this assay to distinguish between true positive signals (reflecting HDAC activity) compared to background values and assay noise (Section 2.12.1.1). Recombinantly expressed PfHDAC1 enzyme was therefore either inhibited with SAHA (positive control for inhibition) or uninhibited when incubated in buffer alone (negative control for inhibition). The blank consisted of buffer and the Z-MAL substrate and the fluorescence readings from the blank were subtracted from all other readings. In addition, the assay was performed for two periods at 37°C, a 30 minutes incubation or overnight. A 30 minute incubation period was recommended by the enzyme manufacturer, however, since the substrate that was used (Z-MAL) differed from that recommended by the manufacturer, an overnight incubation was included to ensure sufficient deacetylation. The use of the overnight incubation resulted in a statistically significant difference between inhibited and uninhibited PfHDAC1 samples (Figure 3.15, $P<0.01$, unpaired t-test) compared to the samples incubated for only 30 minutes.

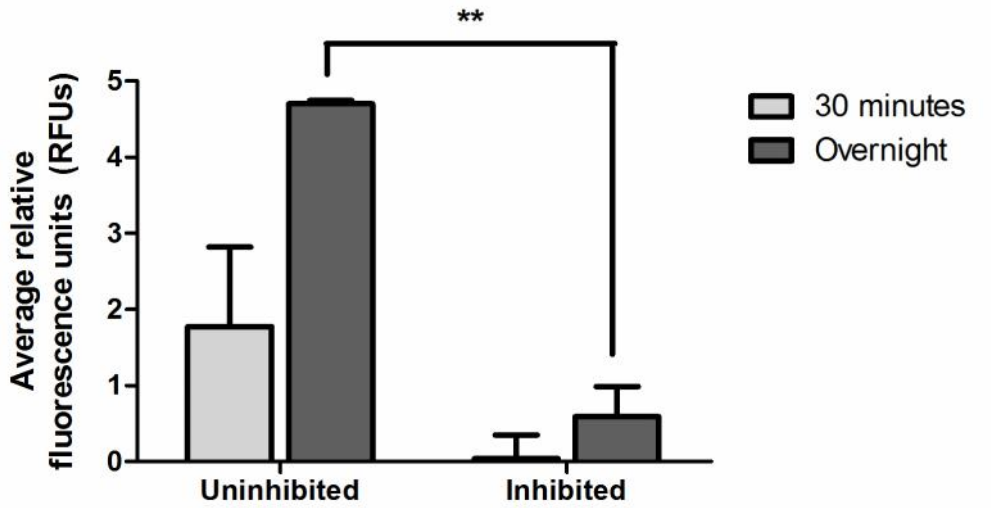


Figure 3.15: Investigation of HDAC assay conditions to determine difference in fluorescence of inhibited and uninhibited PfHDAC1 enzyme. Recombinantly expressed PfHDAC1 was incubated overnight or for 30 minutes at 37°C with the Z-MAL substrate (10.5 µM) and treated with SAHA (10 µM) or the assay buffer (positive and negative controls for HDAC inhibition, respectively). The reaction was terminated by addition of a trypsin-containing stop solution. Fluorescence was measured (355/460 nm) and averaged relative fluorescence units are indicative of the fluorescence emitted from each of the samples once the background has been subtracted. $P<0.01$ (**), unpaired t-test (n=1 ±SD).

Signal to noise and signal to background ratios provide a rapid means for determining whether any signals detected are from the sample or are only due to the background fluorescence. Any signal to background ratio higher than 1 is generally preferred. Signal to noise ratios take into account the standard deviation of the background and a lower standard deviation is preferred as it affords a higher ratio.

Table 3.8: Signal to noise and signal to background ratios before and after optimisation of HDAC assay conditions

	Signal to background ratio= Mean signal/Mean background			**Signal to noise ratio=** (Mean signal-Mean background) /(Standard deviation of background)		
	Incubation times	**Inhibited**	**Uninhibited**	**Incubation times**	**Inhibited**	**Uninhibited**
Assay conditions	30 minutes	0.8	1.05	30 minutes	-1.06	0.33
	Overnight	1.08	1.7	Overnight	3.9	31.4

Again, the assay with an overnight incubation step provided the highest overall signal to background ratio (Table 3.8; 0.8 *vs.* 1.08 for the inhibited sample) as well as signal to noise ratio (31.4 *vs.* 0.33 for the uninhibited sample). The assay was also able to discriminate between uninhibited protein activities compared to proteins inhibited with SAHA (Figure 3.15). In particular, recombinant PfHDAC1 inhibited with SAHA during an overnight incubation period showed a highly significant ($P<0.01$, $n=1$) decrease in enzyme activity of ~3-fold (4.9 RFU untreated *vs.* 1.8 RFU treated). The adapted assay could therefore be used to test the ten MMV compounds previously identified as potential HDAC inhibitors.

3.8.2 Inhibition of recombinant PfHDAC1 by MMV compounds

Once it was possible to distinguish between an enzyme that is inhibited or uninhibited, the ten putative HDAC inhibitors were assayed against PfHDAC1. From these 10 compounds, compounds 5 and 10 show more than 95% inhibition at a 100 µM concentration (Figure 3.16). This inhibition occurred to a greater extent compared to the inhibition caused by any of the other compounds. These effects were also concentration dependent; for example with compound 10's inhibition increasing from ~60% at 10 µM to a ~96% inhibition at 100 µM. Compounds 6, 7, 8 and 9 all exhibited ~50% inhibition, with only compound 6 displaying concentration dependent HDAC inhibition. Compounds 7 and 8 exhibit HDAC inhibition properties where concentration differences do not seem to play a role in the extent of HDAC inhibition. Compounds 1 to 4 appear to have no effect on HDAC activity as compared to the known HDAC inhibitor, SAHA. Thus, at this stage of investigation, compounds 5 and 10 were identified as putative HDAC inhibitors.

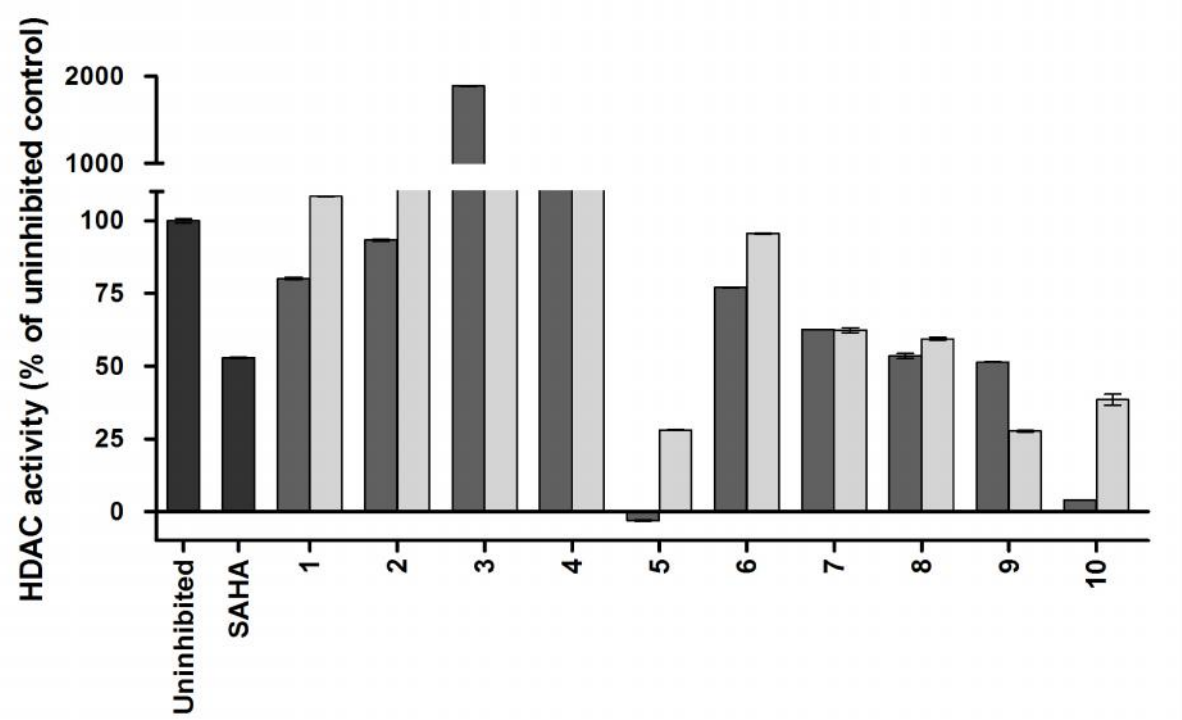


Figure 3.16: Determination of the effect of ten MMV compounds on PfHDAC1 activity. Ten hits from screening the MMV Malaria Box were exposed at a 10 and 100 µM concentration (light and dark grey bars respectively) to 2ng of PfHDAC1 overnight at 37 ºC. Negative and positive controls for HDAC inhibition were also included. The reaction was terminated by addition of a stop solution containing trypsin and the fluorescence was measured (355/460 nm). Data was expressed as a percentage of the uninhibited control after subtraction of background. (n=1, ± SD).

3.9 HDAC activity assay against isolated trophozoite stage parasites

PfHDAC1 is a nuclear enzyme that is expressed primarily in mature asexual stages and gametocytes [90]. Endogenous HDAC activity can thus be detected in intact parasites; however, this necessitated the optimisation of the HDAC assay as described below. Intact, isolated parasites were used for these assays, which were obtained by removal of the erythrocyte membrane trophozoite stage parasites using saponin.

3.9.1 Optimisation of number of parasites/ml for HDAC assay

The linearity of the HDAC assay was determined over a range of parasite cell numbers with a single Z-MAL concentration (substrate concentration as published in [118]). Figure 3.17 shows a drastic increase in the average fluorescence units as an indicator of HDAC activity and this increase correlates to increased numbers of parasites, reaching a maximal 10-fold increase when the parasite numbers are increased from 1×10^7 to 1×10^8. The single concentration of 10.5 µM is sufficient in allowing description of HDAC activity in this assay, when used together with 1×10^8 parasites/ml.

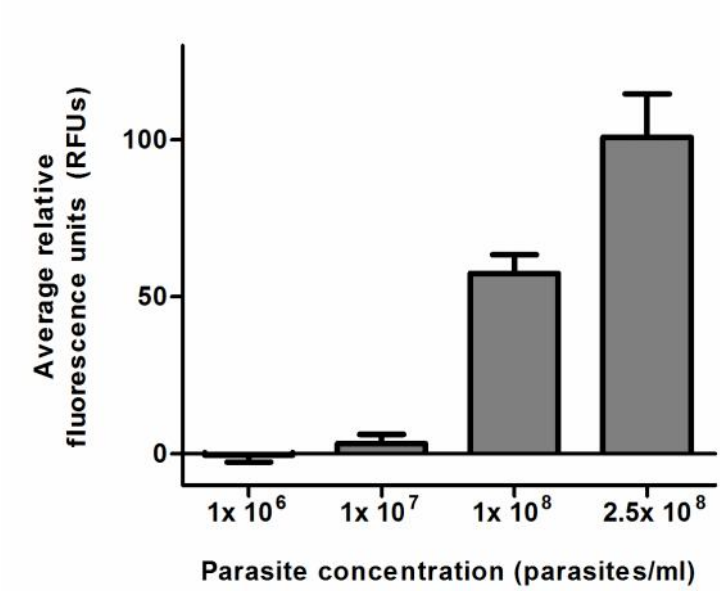


Figure 3.17: HDAC activity assay showing the effect of increasing parasite numbers at a single substrate concentration. Parasites were incubated overnight at 37°C with 10.5 µM Z-MAL substrate. The reaction was terminated by addition of a stop solution containing trypsin and the fluorescence was measured (355/460 nm). Averaged RFU were indicative of the fluorescence emitted from each of the samples once the background has been subtracted. Data are obtained from one biological repeat, performed in triplicate (n=1 ± SD).

As a measure of assay efficiency, the percentage conversion of substrate to product (deacetylation of Z-MAL) was determined (Table 3.9). At a concentration of 10.5 µM substrate, the percentage conversion increased to a maximal increase of ~16-fold when parasite numbers were increased from $1x10^7$ to 1×10^8 parasites.

Table 3.9: Conversion of Z-MAL substrate (10.5 µM) by various parasite concentrations Substrate conversion was calculated by expressing the substrate fluorescence from uninhibited samples as a percentage of AMC control.

Parasites/ml	Percentage of substrate conversion
1×10^6	≤0%
1×10^7	0.6%
1×10^8	10%
2.5×10^8	17.7%

3.9.2 Assessment of statistical factors for the HDAC assay

To determine whether an assay is generating results that are within a 90-99% confidence interval, and that the values produced are not false positives, the LoB values must be determined. Seeing that the control (comprising of the assay buffer and Z-MAL substrate) included in this assay did produce fluorescence measurements that were independent of the release of AMC, it was important to exclude any background fluorescence, which may have led to false negatives from the assay. The LoD is also an important factor in assay development, since the lowest possible

concentration of substrate which can be detected should be calculated. In Table 3.10 below, all the calculated LoB and LoD values at various confidence intervals are shown.

Table 3.10: Limit of Blank (LoB) values obtained from 8 independent experiments and Limit of Detection (LoD) values obtained from 4 independent experiments, and the average values obtained for both of these limits (± SD). CI=confidence interval. Different *k*-values were used to calculate the LoB values for the various confidence intervals.

	90% CI	**95% CI**	**99% CI**
Limit of Blank: $LoB = mean_{blank} + k(SD_{blank})$ k = 1.645 (90% CI) k = 1.96 (95% CI) k = 2.576 (99% CI)	5.8	6.0	6.4
	6.5	6.6	6.8
	5.5	5.5	5.6
	11.2	11.5	12.3
	15.2	15.8	16.9
	16.6	16.9	17.5
	5.8	6.0	6.4
Average LoB ± SD	10.1 ± 5.0 RFU	10.4 ± 5.1 RFU	10.9 ± 5.4 RFU
Limit of Detection: $LoD = LoB + k(SD_{low\ concentration\ sample})$ k = 1.645 (90% CI) k = 1.96 (95% CI) k = 2.576 (99% CI)	19.7	21.8	26.0
	21.9	24.4	29.3
	19.3	21.3	25.2
	19.3	21.3	25.2
	14.3	15.4	17.5
	20.7	23.0	27.4
Average LoD ± SD	19.2 ± 2.6 RFU	21.2 ± 2.8 RFU	25.1 ± 3.7 RFU

The LoB and LoD values were calculated to ensure that false positives were not included in the results. Generally, a 95% confidence interval value is considered statistically sound, and the LoB values for this interval were 10.4 ± 5.1 RFU. This means that when performing this experiment an average fluorescence of approximately 10 RFU will be emitted that is solely due to the blank, and not due to the presence of AMC. Using this LoB to calculate the LoD, it was found that at a 95% confidence interval, the LoD value was 21.2 ± 2.8 RFU. This means that an average RFU of at least 21 had to be emitted from test samples releasing AMC to overcome any interference from the blank.

3.9.3 Comparison of parasites treated with SAHA and CQ

SAHA is a known HDAC inhibitor, while CQ (even though it is a known anti-malarial) has non-HDAC inhibiting effects. The use of these two compounds is a means of determining whether the assay is able to distinguish between HDAC inhibitors as compared to non-inhibitors. Isolated parasites, at a concentration of 1×10^8 parasites/ml were treated for 10 minutes, with various

concentrations of either SAHA or CQ. The assay set-up was carried out as described in Section 3.7.1 (replacing recombinantly expressed PfHDAC1 with 1× 10^8 parasites/ml), and the treated samples were compared to untreated samples. After three independent experiments it was found that only two of the four SAHA concentrations significantly decreased HDAC activity in isolated parasites (1 µM and 10 µM with P<0.01 and P<0.001 respectively, paired t-test, Figure 3.18 A), while none of the CQ concentrations significantly cause HDAC inhibition (P>0.05, Figure 3.18 B). Once it was determined that a difference could be detected using this HDAC assay between HDAC inhibitors and non-HDAC inhibitors, the ten MMV compounds were tested against isolated trophozoite stage parasites using the favourable assay conditions determined.

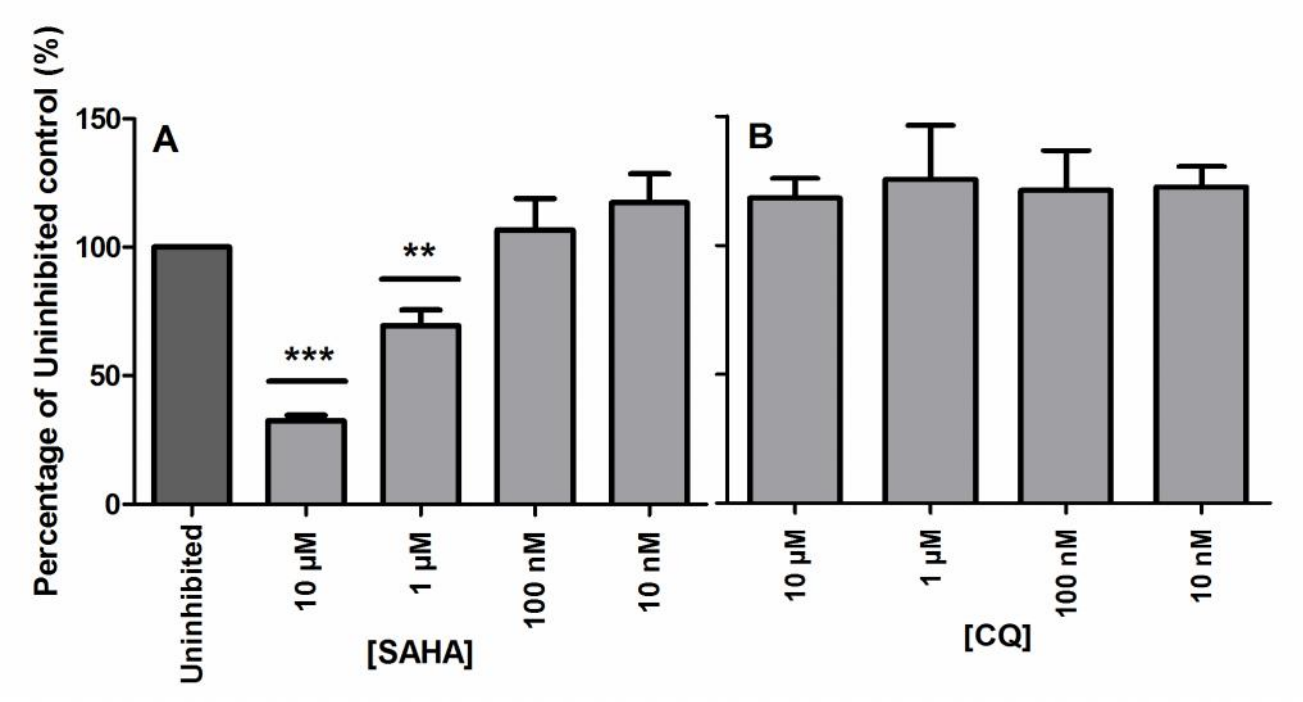


Figure 3.18: HDAC activity assay using isolated trophozoite stage parasites with varying (A) SAHA and (B) CQ concentrations. A fixed parasite concentration of 1× 10^8 parasites/ml was incubated for 10 minutes at 37°C with varying concentrations of SAHA or CQ (10 µM- 10 nM) prior to addition of substrate, followed by an overnight incubation at 37°C. The reaction was terminated by addition of a stop solution containing trypsin and the fluorescence was measured (355/460 nm). The data were expressed as a percentage of the uninhibited control (n=3 ± SEM). Significance was indicated for each sample at P<0.001 (***) and P<0.01(**), paired t-test.

3.9.4 Testing MMV compounds on isolated trophozoite stage parasites in the HDAC assay

To correlate the effects seen on the recombinant PfHDAC1 with the HDAC enzymes in intact parasites, the 10 MMV compounds were tested against isolated trophozoite stage parasites. No significant difference (P>0.05, Figure 3.19) was seen between the uninhibited control and parasites treated at two concentrations of the ten MMV compounds. The only difference seen between inhibited and uninhibited samples is between the control samples, where treatment with SAHA resulted in the significant inhibition (P<0.001, Figure 3.19) of HDAC activity.

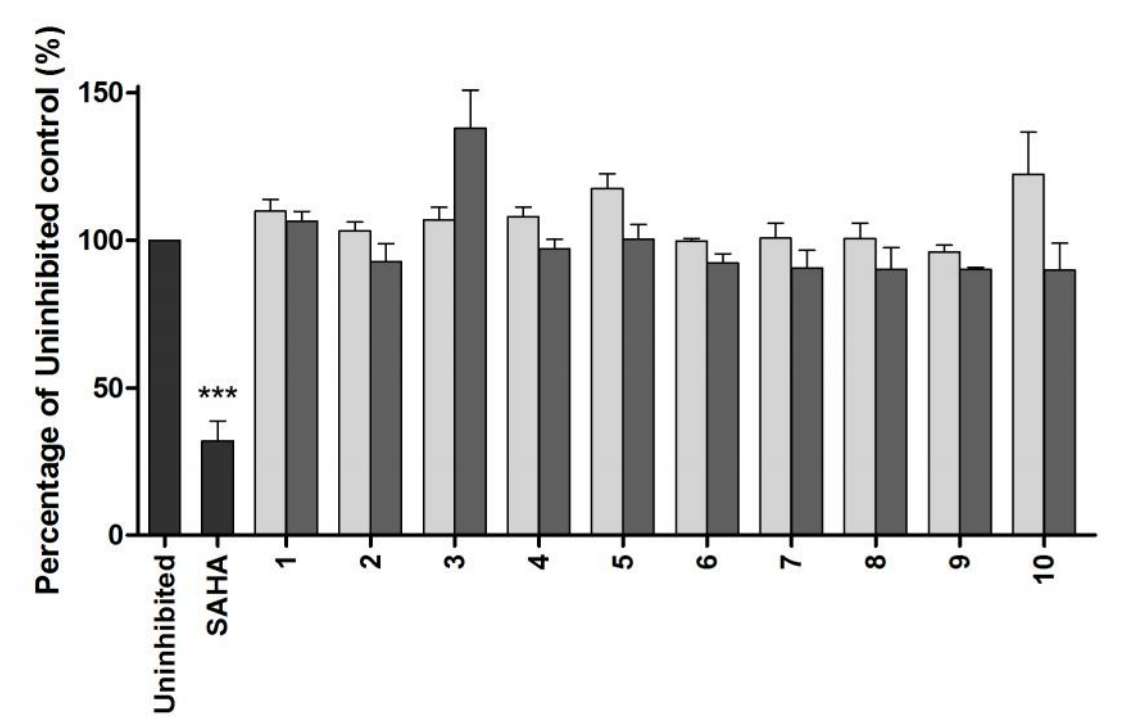


Figure 3.19: Determination of the effect of ten MMV compounds on HDAC activity. Ten putative inhibitors from screening the MMV Malaria Box were exposed at a 10 µM and 100 µM concentration (light and dark grey bars respectively) to 1× 10^8 parasites/ml overnight at 37ºC. Negative and positive controls for HDAC inhibition were also included. The reaction is terminated by addition of a stop solution containing trypsin and the fluorescence is measured (355/460 nm). Data is expressed as a percentage of the uninhibited control after subtraction of background. (n=3, ± SEM).

Although various assays were performed using the MMV Malaria Box compounds, the summation of this data provides valuable results. Table 3.11 shows the inhibition values obtained for each assay. The various end points measured for each assay were a) effect on parasite proliferation in the MSF assay, b) inhibition of HDAC activity of recombinant enzyme and 3) endogenous HDAC activity.

Table 3.11: Various assays performed using the putative HDAC inhibitors (from the MMV Malaria Box), together with approximate percentage inhibition values. Lack of inhibition is indicated by×. The concentrations indicated are the concentrations of each compound used for screening. Figure numbers refer to the assay results from which the data in this table is compiled.

Compound	**MSF assay** Figure 3.14	**HDAC assay on rec. enzyme** Figure 3.16		**HDAC assay on isolated trophozoite stage parasites** Figure 3.19	
	5 µM	**100 µM**	**10 µM**	**100 µM**	**10 µM**
1	×	20%	×	×	×
2	70%	10%	×	×	×
3	80%	×	×	×	×
4	40%	×	×	×	×
5	100%	100%	80%	×	×
6	20%	20%	×	×	×
7	×	40%	40%	×	×
8	10%	50%	50%	×	×
9	20%	50%	50%	×	×
10	30%	90%	70%	×	×

Compound 1 showed little inhibition of parasite proliferation or HDAC activity and thus is an unsuitable lead. Compound 2 was highly active in the MSF assay (70% inhibition of parasite proliferation), but showed poor HDAC inhibition (10% and no inhibition respectively). This was the same for compound 3. Compound 4 showed slight inhibition values of parasite proliferation, though no inhibition of HDAC activity in both subsequent assays. Compound 5 was the only compound that was highly active in its inhibition of parasite proliferation as well as in the HDAC assay against the recombinant enzyme at two concentrations. Compound 7 surprisingly showed no inhibition of parasite proliferation, but exhibited a concentration independent inhibition of the PfHDAC1 recombinant enzyme. Compounds 8, 9 and 10 all showed higher inhibition values against the recombinant enzyme at both concentrations than they did in the MSF assay. Neither the 10 nor the 100 µM concentration of any of the compounds affected endogenous HDAC activity measured in isolated trophozoite stage parasites.

Chapter 4

Discussion

P. falciparum parasites belong to the most virulent *Plasmodium* species, with biological characteristics such as a high rate of replication and antigenic variation that are key contributing factors to its pathogenicity [2]. Both of these mechanisms are also dependent on gene expression and regulation. *Plasmodium* gene regulation has been found to have a strong epigenetic influence, which a number of histone modifying and chromatin remodelling enzymes play a role in. A key enzyme of interest is a member of the histone deacetylase family, PfHDAC1. This enzyme mediates changes in lysine acetylation marks on *P. falciparum* parasite histones and inhibition of this enzyme by a number of human HDAC inhibitors has been found to lead to gene expression perturbation and ultimately cell death [85]. Hyperacetylation of *P. falciparum* parasite histones is characteristic of treatment with an HDAC inhibitor. In this study, the effect of inhibition of an enzyme within the histone deacetylase family and selected histone PTMs of *P. falciparum* trophozoite stages were investigated.

Detection of hyperacetylation is commonly carried out using Western blot techniques, where histones from treated parasites are isolated. Western blotting techniques require that the protein of interest is transferred to a membrane and detected using antibodies directed against modified histones [130]. In the case of histone PTMs, the modification on the protein of interest can also be detected. Commonly detected modifications when investigating hyperacetylated histones are H4K8Ac, tetra-acetylated H4 and H3K9ac. With immunodetection of these PTMs, the effect of HDAC inhibitors on histone acetylation can then be determined. However, commonly found drawbacks of this method include: false positives derived from non-specific binding, lack of stability of certain antibodies, requirement of optimisation of transfer and blotting procedures and lack of accurate quantification of detected proteins [131].

Mass spectrometry analysis on the other hand generates gas-phase ions from samples, separates them according to their mass-to-charge ratio (*m/z*) and generates a record of their abundance [132]. Like any method, drawbacks do exist for MS, such as lengthy sample and data analysis and missed proteolytic cleavages in proteins [133]. However, MS also offers a much more sensitive, accurate and quantitative means to detect molecular weights, amino acid sequences and PTMs (using the change in mass of a peptide) of peptides and proteins. The key advantage of MS analysis for PTM elucidation is that the site of modification does not need to be known, rather the modification is mapped based on the mass shifts within the peptide [134]. Thus, in an effort to

quantitatively detect histone PTMs in *P. falciparum* parasites, and extend this to detect hyperacetylation, a major area of this research involved optimisation of histone isolation and subsequent MS analysis to identify histones and their PTMs and compare this to what has been found previously.

The identification of the eight *P. falciparum* parasite histones (4 canonical and 4 variant) and their PTMs relied heavily on the ability to isolate pure histone samples and to reproducibly analyse the proteomic data obtained from these samples. To achieve this, multiple optimisation steps had to be performed and compared between three known isolation procedures. It should be noted that this is the first study, to our knowledge, where the Issar and Scherf and Merrick methods were adapted to isolate histones for subsequent analyses with MS techniques. The modified Trelle method identified unprecedented PTM numbers, and detected modifications on the low abundant H3 centromeric, which were not previously reported in literature. One possible explanation for the greater numbers obtained with the modified method as compared to the other isolation methods are the sample preparation steps taken. Firstly, the use of protease inhibitors (PIs), which are responsible for the inactivation of enzymes, which may destroy protein samples and affect histones bearing PTMs, were present in the Merrick and modified methods' buffers. Secondly, the combined use of salt and acid extraction procedures in the Issar and Scherf method, which also features in the Trelle method, could be another contributing factor. Even though acid extraction alone (at or below pH 0.7) is believed to solubilise all histones, disassociating them from DNA [135], salt extraction does the converse; excluding insoluble histones from DNA through electrostatic competition [136]. A combination of both these acid and salt extraction enables histones from euchromatic and heterochromatic environments to be isolated [124]. The Trelle method thus combines three effective sample preparation steps, i.e. using buffers containing PIs, and both acid and salt extraction steps. The Merrick method comprises of the former two while the Issar and Scherf method comprises only the latter two. A shorter acid extraction step could be an additional reason why the modified Trelle method could detect more PTMs, as acid labile PTMs were unaffected.

The inconsistencies that exist between the modified and unmodified (published) Trelle methods such as the discrepancy in PTM numbers and obtained protein concentrations could also be explained by differences in sample preparation. As vacuum drying of samples is lengthy, it leads to extended exposure of histones to acid, possibly reducing PTM numbers detected. Furthermore, the vacuum dried sample yielded high protein concentrations, which did not correlate to SDS-PAGE analysis. This inaccurate protein quantification could have been due to chloride ions interfering with the BCA reagent or protein hydrolysis, which may have occurred during vacuum drying. Due to the incompatibility of the vacuum dried sample with the BCA reagent and to shorten protein exposure to acid (by performing TCA precipitation immediately), the modified Trelle method was preferred. In

addition to the ease of use obtained with the modified Trelle method, this method obtained similar sequence coverage for all 8 histones and detected higher PTM numbers compared to the original Trelle method (44 *vs.* 250) [49].

There was a lack of bias seen when investigating the significant PTM types detected with all three histone isolation methods. Whether the marginal difference seen with the modified Trelle method for mono and dimethylation compared to the other two protocols is an indication of the actual PTM landscape that exists on *P. falciparum* trophozoite stage histones is yet to be further investigated. A postulation for the increase in monomethylation could be due to the loss of a single methyl group from a small percentage of dimethylated residues. For example, H4R3me2 (a PTM occurring in high levels in trophozoite stages [137]), was significantly detected with the Issar and Scherf and modified Merrick methods. However, only H4R3me was detected with the modified Trelle method. Trimethylation is the only modification that returns a lower percentage distribution compared to the other modifications. The current hypothesis is that arginine trimethylation does not occur, thus only trimethylated lysine residues were taken into account. Paradoxically, acetylation also only occurs on lysine residues, yet a higher percentage of acetylated residues were detected. This correlates with the fact that malaria parasites have a highly euchromatic genome [38], favouring acetylation rather than trimethylation [49]. Phosphorylation is believed to be a low abundance, acid-labile modification [65], however, the detection of phosphorylation at approximately similar percentage distributions to the remaining modifications is in disagreement with this. An inherent problem with the varied numbers and types of PTMs detected in multiple studies lies both in parasite complexity and sample preparation techniques. Here, validation of PTMs is a point of concern- with MS being a sensitive and accurate technique; few studies have employed the use of biological repeats and have opted rather for manual annotation of spectra, computational validation using site-localisation scores or the development of algorithms to confirm site-localisation [49, 50, 65].

Nevertheless, the PTM landscape and the distribution of the various types of PTMs hold valuable information regarding parasite epigenetic regulation mechanisms. *P. falciparum* parasites have been found to exhibit a “bet-hedging” strategy, where they are prepared to adapt to various perturbations in the environment [31]. This evolutionary advantage could explain the numerous PTMs identified. The existence of highly studded N-termini of the remaining histones indicates a strong epigenetic control of the underlying genomic regions, either to silence or activate their transcription. Since different combination of marks result in different means of genetic regulation, it is possible that parasites possess various combinations leading to functionally different outcomes. An example of this is the presence of both H3K9me3 and H3K9ac, where the former is responsible for silenced *var* genes (genes associated with parasite antigenic variation) and the latter with *var* gene activation [77, 138]. Furthermore, a single mark is not responsible for *var* gene activation and H3K9ac together with H3K4me2 or H3K4me3 must replace H3K9me3 for the activation to occur

[138]. A similar example of this is seen in human embryonic stem cells, where enhancers of genes containing a combination of H3K4me1 and H3K27me3 are poised, while active enhancers replace the trimethylation with H3K27ac [139]. This is characteristic of genes involved in developmental processes [139]. Should the parasite possess both possibilities for transcriptional activation and repression, the sheer numbers of PTMs then required for epigenetic control would be extremely vast. It is then likely that parasites with a broad range of poised and active probabilities will have improved survival and an evolutionary advantage.

A further level of complexity is added when taking into account the nucleosomal level of epigenetic regulation. For example, H2A.Z has been found to co-localise almost perfectly with H3K9ac and H3K4me3 when using ChIP-sequencing technologies to study these interactions [54]. Thus, PTMs are also necessary for the recruitment of various histone variants. One would assume that the modification patterns vary between H3.1 and H3.3, since the replacement of the core with the variant should result in an alteration of the gene expression profile. However, due to the high sequence similarity of these two proteins, the PTM landscape appears highly similar. Since a key epigenetic mechanism is the ability of H3.3 to form an unstable nucleosome with H2A.Z forming euchromatic regions, further divergence from the core H3.1 was probably unnecessary. By contrast, the sequence and PTM landscape of H3 centromeric diverges greatly from that of the core histone H3.1 and unlike the remaining histone proteins H3 centromeric does not exhibit highly decorated N-terminal histone tails. An intriguing observation was the detection of mainly phosphorylation marks on H3 centromeric, since this histone plays a key role in delimiting centromeres and in chromosome segregation and phosphorylation is known to play a role in mitosis, DNA replication and chromosome condensation, a combined detection of these two factors could pose a possible epigenetic relationship.

Studies such as those carried out by Trelle *et al.* [49] prompted the comparison of the plethora of histone PTMs identified on *P. falciparum* parasite histones to those from other eukaryotes. As the techniques involved in investigation of parasite epigenetics evolve, the data obtained becomes increasingly complex. Multi-level approaches would be useful when investigating intricacies such as the histone PTM landscape. Such is the case with a recent study by Gupta *et al.* [140], where ChIP-on-chip analysis was used to associate twelve PTMs to chromosomal regions, crucial information regarding the positioning of PTMs to either coding or non-coding DNA regions was obtained. Without an integrative approach such as this, an incomplete understanding of histone PTMs is conferred. However, what should be noted here, is that only known PTMs can be probed prior to designing and implementing in-depth strategies such as this. Thus, the data obtained can fulfil a niche in the epigenetic field. This exposes a vulnerable regulatory mechanism where drugs targeting the enzymes that “read”, “erase” or “write” certain histone PTMs can be inhibited to induce deleterious effects. An epigenetic “erasing” family of interest are the histone deacetylases.

Previous studies have shown that treatment of *P. falciparum* parasites with SAHA can result in more than 2 fold increase in expression of up to 21% of the parasite's genome within a 2 hour period, along with hyperacetylation of parasite histones [126]. Exposure of *P. falciparum* parasites to SAHA for 2 hours surprisingly revealed no change in the number of acetylated residues of treated parasite histones. Additionally, investigation of H4K8Ac, tetra-acetylated H4 and H3K9ac marks, which are associated with HDAC inhibition, were detected on histones from untreated and treated parasites. Concurrently, a change in phosphorylation was detected, which could be due to a stress response or DNA damage, where phosphorylated residues could be signalling for DNA repair [65]. However, changes in PTM numbers are hardly an indication for what could be occurring on the histone PTM landscape and here drastic changes were seen. H3.1 marks showed an increase in the methylation state of residues associated with *var* genes was seen. This is common with heterochromatin formation, i.e. an increased methylation state leading to gene silencing [79].

Compared to the untreated sample, H3 centromeric was detected in treated samples. This histone is known to show increased expression towards the later schizont stages, plays an important role in chromosome segregation and has a key epigenetic role in the maintenance of centromeres [57]. It is known that HDAC inhibitors like apicidin possibly affect gene expression where genes known to be expressed in the schizont stage are expressed earlier [30]. However, since the modified method is capable of detecting H3 centromeric in untreated samples, more quantitative data is required for further hypotheses regarding SAHA treatment to be drawn. Save for these postulations, definitive conclusions cannot be drawn from MS data generated from histones isolated from treated parasites due to the lack of biological validation, transcriptional analysis and quantifiable data. To obtain a more encompassing, accurate view of the effects of HDAC treatment would probably require the use of quantitative MS, which requires the use of unmodified histone peptides as standards to which modified peptides can be compared [141] and changes in intensity or abundance of histone PTMs can be detected. Optimisation of techniques such as this would facilitate the confirmation of compounds as HDAC inhibitors.

Putative inhibitors of the PfHDAC1 active site, resulting in epigenetic deregulation could provide leads for anti-malarials. *In silico* analysis enabled putative HDAC inhibitors to be identified and investigated through the screening of a library of compounds into the enzyme active site. The 10 compounds obtained possess functional groups such as hydroxamates and benzamides, which function possibly as the zinc-binding portion of the molecule in the enzyme catalytic centre. This zinc ion in conjunction with a side chain histidine play crucial roles as they enable nucleophillic attack of the water molecule in the active site, followed by attack of the carbonyl carbon of acetyl lysine, resulting in deacetylation [142].

A requirement for the inclusion of the ten compounds into the MMV Malaria Box was the possession of *in vitro* inhibition values of less than 4 µM against whole cell *P. falciparum* parasites. The re-evaluation of their activity was bound to result in discrepancies arising from varying sample quality of the compounds [7] or the use of different assays, which often return different inhibition concentrations. The high throughput screening assay used by the MMV employs the DAPI (4'-6-diamidino-2-phenylindole) stain in a fluorescent high content imaging assay [129]. On comparison of this assay to the MSF assay, three apparent differences exist: the incubation times (72 hour *vs.* 96 hour, respectively), detection of the fluorescent signal (confocal imaging system *vs.* a fluorometer) and the end point measured (DNA content *vs.* phenotypic effects) [125, 129]. Certain compounds, such as compound 2 and 3 showed a significant decrease of parasite proliferation. However, when tested against the recombinant enzyme, these compounds showed low inhibitory activity. This could be an indication of a different mechanism of action, unrelated to inhibition of HDAC activity. The converse was seen with compounds like compound 10, which showed low inhibition of parasite proliferation, but high HDAC inhibitory activities. This could be due to the lack of drug-like properties of compound 10, where its hydrophilic nature could be hindering entry into the parasite, however, when exposed to the enzyme exogenously, it exhibits activity.

Compound 5 holds promise as a possible anti-plasmodial as it inhibits parasite proliferation almost completely in the MSF assay and effectively inhibits PfHDAC1 activity in a concentration-dependent manner. What lends confidence to the identification of compound 5 as a lead is that other benzamide containing compounds, such as MS-275, are known for their histone deacetylase activity. This HDAC inhibitor has previously shown modest activity against *P. falciparum* parasite proliferation *in vitro* (8.3 µM) [91] and has ~1 µM activity against PfHDAC1 [97]. Although *in silico* docking was used here as a guide and a filter for narrowing down the MMV Malaria Box, compound 5 demonstrated the lowest CDOCKER score, an indication that its binding to the enzyme active site probably shows a strong interaction. It is therefore probable that compound 5 selectively inhibits PfHDAC1 activity, which leads to an inhibition of parasite proliferation and ultimately cell death, as PfHDAC1 is an essential enzyme for the parasite [103]. The inactivity seen with the remaining compounds in the HDAC assay could be due to a time-dependent dosing effect, which was seen previously with HDAC inhibitors requiring a longer period to bind to the active site zinc molecule positioned deep within the HDAC binding site [143]. Since these inhibitors were only exposed to the enzyme for 10 minutes, a longer incubation time could be used to determine whether their potency could be intensified.

The endogenous HDAC activity when investigated in isolated trophozoite-stage parasites was not inhibited by any of the compounds. However, correlating enzyme inhibition to potency against whole cells is always disconcerting. This is due to the fact that there are multiple parameters to take into consideration, such as enzyme selectivity, membrane permeability, compound

metabolism, cellular localisation of target enzyme and presence of efflux pumps for xenobiotics. Foregoing the effect of the MMV compounds on isolated trophozoites in the HDAC assay, assay parameters such as limits of detection as well as optimal parasite numbers were validated. Since late trophozoite stages are known to show increased expression of PfHDAC1 [90], increased parasite numbers in the HDAC assay probably contributed to an increased concentration of the HDAC enzyme. This would explain the concurrent increase in RFU with parasite number. Parasite controls (isolated parasites and phosphate buffer) were also found to exhibit fluorescence readings that were lower than the Z-MAL control, thus showing that the readings measured were not due to the light scattering properties of the brown coloured parasites. The substrate concentration was also varied to determine whether increased substrate concentrations were required for a constant parasite concentration (which correlates to enzyme concentration). The ability of the assay to differentiate between HDAC inhibitors and non-HDAC inhibitors was also seen on comparison of treatment with SAHA and CQ. A noteworthy fact is that SAHA, like many other HDAC inhibitors, is a pan-inhibitor (indiscriminate inhibition of HDAC enzymes) of zinc-requiring HDACs functions [144]. Therefore, lack of inhibition seen by treatment with MMV compounds as compared to SAHA could be the result of PfHDAC1 selective inhibition compared to inhibition of multiple HDACs.

The ability of the compounds to reach the parasites is also of concern. With trophozoite stages the presence of new permeability pathways (NPP) induced in the erythrocyte membrane after infection, allow for entry of multiple substances to reach the parasite [145]. Thus, removal of the erythrocyte surrounding the parasite could affect the entry of these substances. However, as cellular permeability is an important criterion for these MMV compounds, computational analyses were performed to determine their permeability prior to selection [7]. The other factor that influences activity is the location of the enzyme, since PfHDAC1 is localised in the nucleus of the parasite, it is important to determine whether the compound actually reaches the nucleus to have an effect. It could be possible that the inhibition of PfHDAC1 is a non-specific effect of the compound and that the primary mode of action is otherwise.

Thus, what can be derived from this work is the heavy dependence the parasite places on the utilisation of chromatin mediated epigenetic strategies to control its gene expression. An enigmatic fact though is the corroborating evidence as to which PTM combinations occur to mediate epigenetic effects on transcription. Together with the methodological developments presented, the finding of putative HDAC inhibitors could lead to novel drugs, providing a cornerstone for inhibiting and understanding vital mechanisms in the malaria parasite.

Chapter 5

Conclusion

The histone deacetylase family in *P. falciparum* parasites exposes an epigenetic target for novel anti-malarials. The objectives of this work were therefore to inhibit a member of this family (PfHDAC1) and develop means of assessing the effect of drug treatment on the parasite, either with an HDAC assay or the investigation of histone PTM landscapes. Since the approach used was unable to fully elucidate the effects of HDAC inhibition, a more quantitative approach is required to determine, which PTMs are abundant following HDAC inhibition. The development of a rapid, reproducible histone isolation technique allows for high yields of histones to be isolated, which can be used in applications other than MS (such as Western blotting and as substrates for enzymatic assays). Furthermore, the extension of this technique to the investigation of other *P. falciparum* life-stages, as well as other *Plasmodium* species, facilitates an in-depth understanding of the *Plasmodium* epigenome. The plethora of histone PTMs identified exclusively with this technique is evidence of the complexity of the epigenetic code, which once elucidated may allow further understanding of the biological mechanisms in *P. falciparum* parasites, thereby exposing these to exploitation for therapeutic benefit.

An alternative means for assessing HDAC inhibition was through the development of an HDAC assay, which showed promise when used with the recombinant PfHDAC1 enzyme. Here, biological validation is still outstanding but the identification of putative HDAC inhibitors is a positive avenue for exploring novel anti-plasmodials. HDAC inhibitors could reveal a promising class of novel epidrugs (drugs targeting epigenetics) due to their attenuation of HDAC activity, which affects transcription of various genes. Further studies on these compounds would be required to progress these compounds along the drug development pipeline. These include microarray analysis to determine the effects on transcriptional profiles of treated parasites, confirmatory assays with Western blotting techniques, and the development of resistant parasites (to identify affected gene regions of mutant parasites).

In conclusion, epigenetic mechanisms in the malaria parasite are key targets for drug discovery. Herewith, we have moved a step closer to understanding epigenetics in *P. falciparum* parasites, contributing assays and techniques, which may aid in epigenetic studies and discovery of lead compounds that could be further developed to populate the anti-malarial pipeline.

References

1 Muller, I. B. and Hyde, J. E. (2010) Antimalarial drugs: modes of action and mechanisms of parasite resistance. Future Microbiol. **5**, 1857-1873

2 Heddini, A. (2002) Malaria pathogenesis: a jigsaw with an increasing number of pieces. International journal for parasitology. **32**, 1587-1598

3 WHO. (2011) World Malaria Report 2011. ed.)^eds.)

4 Birkholtz, L. M., Bornman, R., Focke, W., Mutero, C. and de Jager, C. (2012) Sustainable malaria control: transdisciplinary approaches for translational applications. Malar J. **11**, 431

5 Guinovart, C., Navia, M. M., Tanner, M. and Alonso, P. L. (2006) Malaria: burden of disease. Curr Mol Med. **6**, 137-140

6 Miller, L. H., Ackerman, H. C., Su, X. Z. and Wellems, T. E. (2013) Malaria biology and disease pathogenesis: insights for new treatments. Nat Med. **19**, 156-167

7 Spangenberg, T., Burrows, J. N., Kowalczyk, P., McDonald, S., Wells, T. N. and Willis, P. (2013) The open access malaria box: a drug discovery catalyst for neglected diseases. PloS one. **8**, e62906

8 Salcedo-Amaya, A. M., Hoeijmakers, W. A., Bartfai, R. and Stunnenberg, H. G. (2010) Malaria: could its unusual epigenome be the weak spot? Int J Biochem Cell Biol. **42**, 781-784

9 Lehninger, A. L., Nelson, D.L. and Cox, M.M. (2004) Principles of Biochemistry. W.H. Freeman

10 Lodish, H. F. (2008) Molecular cell biology. W.H. Freeman, New York

11 Hammarton, T. C., Mottram, J. C. and Doerig, C. (2003) The cell cycle of parasitic protozoa: potential for chemotherapeutic exploitation. Prog Cell Cycle Res. **5**, 91-101

12 Gerald, N., Mahajan, B. and Kumar, S. (2011) Mitosis in the human malaria parasite Plasmodium falciparum. Eukaryot Cell. **10**, 474-482

13 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. Int J Parasitol. **41**, 71-80

14 Prudencio, M., Rodriguez, A. and Mota, M. M. (2006) The silent path to thousands of merozoites: the *Plasmodium* liver stage. Nat Rev Microbiol. **4**, 849-856

15 Arnot, D. E. and Gull, K. (1998) The *Plasmodium* cell-cycle: facts and questions. Ann Trop Med Parasitol. **92**, 361-365

16 Leete, T. H. and Rubin, H. (1996) Malaria and the cell cycle. Parasitol Today. **12**, 442-444

17 Lobo, C. A. and Kumar, N. (1999) Differential transcription of histone genes in asexual and sexual stages of *Plasmodium falciparum*. International journal for parasitology. **29**, 1447-1449

18 Jacobberger, J. W., Horan, P. K. and Hare, J. D. (1992) Cell cycle analysis of asexual stages of erythrocytic malaria parasites. Cell Prolif. **25**, 431-445

19 Gritzmacher, C. A. and Reese, R. T. (1984) Protein and nucleic acid synthesis during synchronized growth of *Plasmodium falciparum*. J Bacteriol. **160**, 1165-1167

20 Janse, C. J., Haghparast, A., Speranca, M. A., Ramesar, J., Kroeze, H., del Portillo, H. A. and Waters, A. P. (2003) Malaria parasites lacking eef1a have a normal S/M phase yet grow more slowly due to a longer G1 phase. Mol Microbiol. **50**, 1539-1551

21 Naughton, J. A. and Bell, A. (2007) Studies on cell-cycle synchronization in the asexual erythrocytic stages of Plasmodium falciparum. Parasitology. **134**, 331-337

22 Baker, D. A. (2010) Malaria gametocytogenesis. Molecular and biochemical parasitology. **172**, 57-65

23 Voet, D. and Voet, J. G. (2011) DNA Replication, Repair, and Recombination. In Biochemistry ed.)^eds.). p. 1202, John Wiley and Sons, New York.

24 Carvalho, T. G., Doerig, C. and Reininger, L. (2013) Nima- and Aurora-related kinases of malaria parasites. Biochim Biophys Acta. **1834**, 1336-1345

25 Mitra, P., Deshmukh, A. S. and Dhar, S. K. (2012) DNA replication during intra-erythrocytic stages of human malarial parasite *Plasmodium falciparum*. CURRENT SCIENCE. **102**, 725

26 Hakimi, M. A. and Deitsch, K. W. (2007) Epigenetics in Apicomplexa: control of gene expression during cell cycle progression, differentiation and antigenic variation. Current opinion in microbiology. **10**, 357-362

27 Horrocks, P., Wong, E., Russell, K. and Emes, R. D. (2009) Control of gene expression in *Plasmodium falciparum* - ten years on. Molecular and biochemical parasitology. **164**, 9-25

28 Bozdech, Z., Llinas, 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**, E5

29 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

30 Chaal, B. K., Gupta, A. P., Wastuwidyaningtyas, B. D., Luah, Y. H. and Bozdech, Z. (2010) Histone deacetylases play a major role in the transcriptional regulation of the *Plasmodium falciparum* life cycle. PLoS Pathogens. **6**, e1000737

31 Rovira-Graells, N., Gupta, A. P., Planet, E., Crowley, V. M., Mok, S., Ribas de Pouplana, L., Preiser, P. R., Bozdech, Z. and Cortes, A. (2012) Transcriptional variation in the malaria parasite *Plasmodium falciparum*. Genome Res. **22**, 925-938

32 Templeton, T. J., Iyer, L. M., Anantharaman, V., Enomoto, S., Abrahante, J. E., Subramanian, G. M., Hoffman, S. L., Abrahamsen, M. S. and Aravind, L. (2004) Comparative analysis of apicomplexa and genomic diversity in eukaryotes. Genome Res. **14**, 1686-1695

33 Bischoff, E. and Vaquero, C. (2010) In silico and biological survey of transcription-associated proteins implicated in the transcriptional machinery during the erythrocytic development of *Plasmodium falciparum*. BMC Genomics. **11**, 34

34 Duffy, M. F., Selvarajah, S., Josling, G. A. and Petter, M. (2012) The role of chromatin in *Plasmodium* gene expression. Cellular microbiology

35 Merrick, C. J. and Duraisingh, M. T. (2010) Epigenetics in Plasmodium: what do we really know? Eukaryot Cell. **9**, 1150-1158

36 Holliday, R. (1987) The inheritance of epigenetic defects. Science. **238**, 163-170

37 Cui, L. and Miao, J. (2010) Chromatin-mediated epigenetic regulation in the malaria parasite *Plasmodium falciparum*. Eukaryot Cell. **9**, 1138-1149

38 Hoeijmakers, W. A., Stunnenberg, H. G. and Bartfai, R. (2012) Placing the *Plasmodium falciparum* epigenome on the map. Trends Parasitol. **28**, 486-495

39 Weiner, A., Dahan-Pasternak, N., Shimoni, E., Shinder, V., von Huth, P., Elbaum, M. and Dzikowski, R. (2011) 3D nuclear architecture reveals coupled cell cycle dynamics of chromatin and nuclear pores in the malaria parasite *Plasmodium falciparum*. Cell Microbiol. **13**, 967-977

40 Cortes, A., Crowley, V. M., Vaquero, A. and Voss, T. S. (2012) A view on the role of epigenetics in the biology of malaria parasites. PLoS Pathog. **8**, e1002943

41 Sims, J. S., Militello, K. T., Sims, P. A., Patel, V. P., Kasper, J. M. and Wirth, D. F. (2009) Patterns of gene-specific and total transcriptional activity during the *Plasmodium falciparum* intraerythrocytic developmental cycle. Eukaryot Cell. **8**, 327-338

42 Ponts, N., Harris, E. Y., Lonardi, S. and Le Roch, K. G. (2011) Nucleosome occupancy at transcription start sites in the human malaria parasite: a hard-wired evolution of virulence? Infect Genet Evol. **11**, 716-724

43 Brivanlou, A. H. and Darnell, J. E., Jr. (2002) Signal transduction and the control of gene expression. Science. **295**, 813-818

44 Boschet, C., Gissot, M., Briquet, S., Hamid, Z., Claudel-Renard, C. and Vaquero, C. (2004) Characterization of PfMyb1 transcription factor during erythrocytic development of 3D7 and F12 *Plasmodium falciparum* clones. Mol Biochem Parasitol. **138**, 159-163

45 Painter, H. J., Campbell, T. L. and Llinas, M. (2011) The Apicomplexan AP2 family: integral factors regulating *Plasmodium* development. Mol Biochem Parasitol. **176**, 1-7

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

47 Chene, A., Vembar, S. S., Riviere, L., Lopez-Rubio, J. J., Claes, A., Siegel, T. N., Sakamoto, H., Scheidig-Benatar, C., Hernandez-Rivas, R. and Scherf, A. (2012) PfAlbas constitute a new eukaryotic DNA/RNA-binding protein family in malaria parasites. Nucleic Acids Res. **40**, 3066-3077

48 Goyal, M., Alam, A., Iqbal, M. S., Dey, S., Bindu, S., Pal, C., Banerjee, A., Chakrabarti, S. and Bandyopadhyay, U. (2012) Identification and molecular characterization of an Alba-family protein from human malaria parasite *Plasmodium falciparum*. Nucleic Acids Res. **40**, 1174-1190

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

50 Miao, J., Fan, Q., Cui, L. and Li, J. (2006) The malaria parasite *Plasmodium falciparum* histones: organization, expression, and acetylation. Gene. **369**, 53-65

51 Inselburg, J. and Banyal, H. S. (1984) Synthesis of DNA during the asexual cycle of *Plasmodium falciparum* in culture. Mol Biochem Parasitol. **10**, 79-87

52 Felsenfeld, G. and Groudine, M. (2003) Controlling the double helix. Nature. **421**, 448-453

53 Petter, M., Lee, C. C., Byrne, T. J., Boysen, K. E., Volz, J., Ralph, S. A., Cowman, A. F., Brown, G. V. and Duffy, M. F. (2011) Expression of *P. falciparum* var genes involves exchange of the histone variant H2A.Z at the promoter. PLoS Pathog. **7**, e1001292

54 Bartfai, R., Hoeijmakers, W. A., Salcedo-Amaya, A. M., Smits, A. H., Janssen-Megens, E., Kaan, A., Treeck, M., Gilberger, T. W., Francoijs, K. J. and Stunnenberg, H. G. (2010) H2A.Z demarcates intergenic regions of the plasmodium falciparum epigenome that are dynamically marked by H3K9ac and H3K4me3. PLoS Pathog. **6**, e1001223

55 Petter, M., Selvarajah, S. A., Lee, C. C., Chin, W. H., Gupta, A. P., Bozdech, Z., Brown, G. V. and Duffy, M. F. (2013) H2A.Z and H2B.Z double-variant nucleosomes define intergenic regions and dynamically occupy var gene promoters in the malaria parasite *Plasmodium falciparum*. Mol Microbiol. **87**, 1167-1182

56 Hoeijmakers, W. A., Salcedo-Amaya, A. M., Smits, A. H., Francoijs, K. J., Treeck, M., Gilberger, T. W., Stunnenberg, H. G. and Bartfai, R. (2013) H2A.Z/H2B.Z double-variant nucleosomes inhabit the AT-rich promoter regions of the *Plasmodium falciparum* genome. Mol Microbiol. **87**, 1061-1073

57 Hoeijmakers, W. A., Flueck, C., Francoijs, K. J., Smits, A. H., Wetzel, J., Volz, J. C., Cowman, A. F., Voss, T., Stunnenberg, H. G. and Bartfai, R. (2012) *Plasmodium falciparum* centromeres display a unique epigenetic makeup and cluster prior to and during schizogony. Cell Microbiol. **14**, 1391-1401

58 Clark, R. S., Bayir, H. and Jenkins, L. W. (2005) Posttranslational protein modifications. Critical care medicine. **33**, S407-409

59 Yang, X. J. and Seto, E. (2007) HATs and HDACs: from structure, function and regulation to novel strategies for therapy and prevention. Oncogene. **26**, 5310-5318

60 Hoffman, M. D., Sniatynski, M. J. and Kast, J. (2008) Current approaches for global post-translational modification discovery and mass spectrometric analysis. Analytica chimica acta. **627**, 50-61

61 Workman, J. L. and Kingston, R. E. (1998) Alteration of nucleosome structure as a mechanism of transcriptional regulation. Annual review of biochemistry. **67**, 545-579

62 Strahl, B. D. and Allis, C. D. (2000) The language of covalent histone modifications. Nature. **403**, 41-45

63 Miller-Jensen, K., Dey, S. S., Schaffer, D. V. and Arkin, A. P. (2011) Varying virulence: epigenetic control of expression noise and disease processes. Trends in biotechnology. **29**, 517-525

64 Munshi, A., Shafi, G., Aliya, N. and Jyothy, A. (2009) Histone modifications dictate specific biological readouts. Journal of genetics and genomics = Yi chuan xue bao. **36**, 75-88

65 Dastidar, E. G., Dzeyk, K., Krijgsveld, J., Malmquist, N. A., Doerig, C., Scherf, A. and Lopez-Rubio, J. J. (2013) Comprehensive histone phosphorylation analysis and identification of Pf14-3-3 protein as a histone H3 phosphorylation reader in malaria parasites. PloS one. **8**, e53179

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

67 Chung, D. W., Ponts, N., Cervantes, S. and Le Roch, K. G. (2009) Post-translational modifications in *Plasmodium*: more than you think! Molecular and biochemical parasitology. **168**, 123-134

68 Philip, N. and Haystead, T. A. (2007) Characterization of a UBC13 kinase in *Plasmodium falciparum*. Proceedings of the National Academy of Sciences of the United States of America. **104**, 7845-7850

69 Weake, V. M. and Workman, J. L. (2008) Histone ubiquitination: triggering gene activity. Molecular cell. **29**, 653-663

70 Issar, N., Roux, E., Mattei, D. and Scherf, A. (2008) Identification of a novel post-translational modification in *Plasmodium falciparum*: protein sumoylation in different cellular compartments. Cell Microbiol. **10**, 1999-2011

71 David, G., Neptune, M. A. and DePinho, R. A. (2002) SUMO-1 modification of histone deacetylase 1 (HDAC1) modulates its biological activities. The Journal of biological chemistry. **277**, 23658-23663

72 Gill, G. (2003) Post-translational modification by the small ubiquitin-related modifier SUMO has big effects on transcription factor activity. Current opinion in genetics & development. **13**, 108-113

73 Berger, S. L. (2010) Cell signaling and transcriptional regulation via histone phosphorylation. Cold Spring Harb Symp Quant Biol. **75**, 23-26

74 Solyakov, L., Halbert, J., Alam, M. M., Semblat, J. P., Dorin-Semblat, D., Reininger, L., Bottrill, A. R., Mistry, S., Abdi, A., Fennell, C., Holland, Z., Demarta, C., Bouza, Y., Sicard, A., Nivez, M. P., Eschenlauer, S., Lama, T., Thomas, D. C., Sharma, P., Agarwal, S., Kern, S., Pradel, G., Graciotti, M., Tobin, A. B. and Doerig, C. (2011) Global kinomic and phospho-proteomic analyses of the human malaria parasite *Plasmodium falciparum*. Nat Commun. **2**, 565

75 Dastidar, E. G., Dayer, G., Holland, Z. M., Dorin-Semblat, D., Claes, A., Chene, A., Sharma, A., Hamelin, R., Moniatte, M., Lopez-Rubio, J. J., Scherf, A. and Doerig, C. (2012) Involvement of *Plasmodium falciparum* protein kinase CK2 in the chromatin assembly pathway. BMC Biol. **10**, 5

76 Cui, L., Fan, Q. and Miao, J. (2008) Histone lysine methyltransferases and demethylases in *Plasmodium falciparum*. International journal for parasitology. **38**, 1083-1097

77 Lopez-Rubio, J. J., Gontijo, A. M., Nunes, M. C., Issar, N., Hernandez Rivas, R. and Scherf, A. (2007) 5' flanking region of var genes nucleate histone modification patterns linked to phenotypic inheritance of virulence traits in malaria parasites. Mol Microbiol. **66**, 1296-1305

78 Martin, C. and Zhang, Y. (2005) The diverse functions of histone lysine methylation. Nature reviews. Molecular cell biology. **6**, 838-849

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

80 Perez-Toledo, K., Rojas-Meza, A. P., Mancio-Silva, L., Hernandez-Cuevas, N. A., Delgadillo, D. M., Vargas, M., Martinez-Calvillo, S., Scherf, A. and Hernandez-Rivas, R. (2009) *Plasmodium falciparum* heterochromatin protein 1 binds to tri-methylated histone 3 lysine 9 and is linked to mutually exclusive expression of var genes. Nucleic Acids Res. **37**, 2596-2606

81 Agbor-Enoh, S., Seudieu, C., Davidson, E., Dritschilo, A. and Jung, M. (2009) Novel inhibitor of *Plasmodium* histone deacetylase that cures *P. berghei*-infected mice. Antimicrobial agents and chemotherapy. **53**, 1727

82 Jenuwein, T. and Allis, C. D. (2001) Translating the histone code. Science. **293**, 1074

83 Ji, D. D. and Arnot, D. E. (1997) A *Plasmodium falciparum* homologue of the ATPase subunit of a multi-protein complex involved in chromatin remodelling for transcription. Molecular and biochemical parasitology. **88**, 151-162

84 Cui, L. and Miao, J. (2007) Cytotoxic effect of curcumin on malaria parasite *Plasmodium falciparum*: inhibition of histone acetylation and generation of reactive oxygen species. Antimicrobial agents and chemotherapy. **51**, 488-494

85 Andrews, K. T., Haque, A. and Jones, M. K. (2012) HDAC inhibitors in parasitic diseases. Immunology and cell biology. **90**, 66-77

86 Merrick, C. J. and Duraisingh, M. T. (2007) *Plasmodium falciparum* Sir2: an unusual sirtuin with dual histone deacetylase and ADP-ribosyltransferase activity. Eukaryot Cell. **6**, 2081-2091

87 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 Biol. **7**, e84

88 Ledent, V. and Vervoort, M. (2006) Comparative genomics of the class 4 histone deacetylase family indicates a complex evolutionary history. BMC biology. **4**, 24

89 Andrews, K. T., Tran, T. N., Wheatley, N. C. and Fairlie, D. P. (2009) Targeting histone deacetylase inhibitors for anti-malarial therapy. Current Topics in Medicinal Chemistry. **9**, 292-308

90 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

91 Andrews, K. T., Tran, T. N., Lucke, A. J., Kahnberg, P., Le, G. T., Boyle, G. M., Gardiner, D. L., Skinner-Adams, T. S. and Fairlie, D. P. (2008) Potent antimalarial activity of histone deacetylase inhibitor analogues. Antimicrobial agents and chemotherapy. **52**, 1454-1461

92 Chen, Y., Lopez-Sanchez, M., Savoy, D. N., Billadeau, D. D., Dow, G. S. and Kozikowski, A. P. (2008) A series of potent and selective, triazolylphenyl-based histone deacetylases inhibitors with activity against pancreatic cancer cells and *Plasmodium falciparum*. Journal of medicinal chemistry. **51**, 3437-3448

93 Casero Jr, R. A. and Woster, P. M. (2009) Recent Advances in the Development of Polyamine Analogues as Antitumor Agents Journal of medicinal chemistry. **52**, 4551-4573

94 Darkin-Rattray, S. J., Gurnett, A. M., Myers, R. W., Dulski, P. M., Crumley, T. M., Allocco, J. J., Cannova, C., Meinke, P. T., Colletti, S. L., Bednarek, M. A., Singh, S. B., Goetz, M. A., Dombrowski, A. W., Polishook, J. D. and Schmatz, D. M. (1996) Apicidin: a novel antiprotozoal agent that inhibits parasite histone deacetylase. Proceedings of the National Academy of Sciences of the United States of America. **93**, 13143-13147

95 Bougdour, A., Maubon, D., Baldacci, P., Ortet, P., Bastien, O., Bouillon, A., Barale, J. C., Pelloux, H., Menard, R. and Hakimi, M. A. (2009) Drug inhibition of HDAC3 and epigenetic control of differentiation in Apicomplexa parasites. The Journal of experimental medicine. **206**, 953-966

96 Andrews, K. T., Walduck, A., Kelso, M. J., Fairlie, D. P., Saul, A. and Parsons, P. G. (2000) Anti-malarial effect of histone deacetylation inhibitors and mammalian tumour cytodifferentiating agents. International journal for parasitology. **30**, 761-768

97 Patel, V., Mazitschek, R., Coleman, B., Nguyen, C., Urgaonkar, S., Cortese, J., Barker Jr, R. H., Greenberg, E., Tang, W. and Bradner, J. E. (2009) Identification and characterization of small molecule inhibitors of a class I histone deacetylase from *Plasmodium falciparum*. Journal of medicinal chemistry. **52**, 2185-2187

98 Mwakwari, S. C., Guerrant, W., Patil, V., Khan, S. I., Tekwani, B. L., Gurard-Levin, Z. A., Mrksich, M. and Oyelere, A. K. (2010) Non-peptide macrocyclic histone deacetylase inhibitors derived from tricyclic ketolide skeleton. Journal of medicinal chemistry. **53**, 6100

99 Meinke, P. T., Colletti, S. L., Doss, G., Myers, R. W., Gurnett, A. M., Dulski, P. M., Darkin-Rattray, S. J., Allocco, J. J., Galuska, S., Schmatz, D. M., Wyvratt, M. J. and Fisher, M. H. (2000) Synthesis of apicidin-derived quinolone derivatives: parasite-selective histone deacetylase inhibitors and antiproliferative agents. Journal of medicinal chemistry. **43**, 4919-4922

100 Patil, V., Guerrant, W., Chen, P. C., Gryder, B., Benicewicz, D. B., Khan, S. I., Tekwani, B. L. and Oyelere, A. K. (2010) Antimalarial and antileishmanial activities of histone deacetylase inhibitors with triazole-linked cap group. Bioorganic & medicinal chemistry. **18**, 415-425

101 Dow, G. S., Chen, Y., Andrews, K. T., Caridha, D., Gerena, L., Gettayacamin, M., Johnson, J., Li, Q., Melendez, V., Obaldia, N., 3rd, Tran, T. N. and Kozikowski, A. P. (2008) Antimalarial activity of phenylthiazolyl-bearing hydroxamate-based histone deacetylase inhibitors. Antimicrobial agents and chemotherapy. **52**, 3467-3477

102 Khan, N., Jeffers, M., Kumar, S., Hackett, C., Boldog, F., Khramtsov, N., Qian, X., Mills, E., Berghs, S. C., Carey, N., Finn, P. W., Collins, L. S., Tumber, A., Ritchie, J. W., Jensen, P. B., Lichenstein, H. S. and Sehested, M. (2008) Determination of the class and isoform selectivity of small-molecule histone deacetylase inhibitors. The Biochemical journal. **409**, 581-589

103 Tran, T. N. (2009) *Plasmodium falciparum* histone deacetylases as novel antimalarial drug targets. ed.)^eds.). p. 158, Griffith University

104 Allen, R. J. W. and Kirk, K. (2004) The membrane potential of the intraerythrocytic malaria parasite *Plasmodium falciparum*. Journal of Biological Chemistry. **279**, 11264

105 Trager, W. and Jensen, J. B. (1976) Human malaria parasites in continuous culture. J Parasitol. **91**, 484-486

106 Cox, F. E. (2010) History of the discovery of the malaria parasites and their vectors. Parasit Vectors. **3**, 5

107 Lambros, C. and Vanderberg, J. P. (1979) Synchronization of *Plasmodium falciparum* erythrocytic stages in culture. J Parasitol. **65**, 418-420

108 Salcedo Amaya, A. M. (2012) Malaria epigenome : how histone post-translational modifications contribute to gene expression in *Plasmodium falciparum*. ed.)^eds.). p. 141, Radboud Universiteit Nijmegen

109 Voss, T. S., Mini, T., Jenoe, P. and Beck, H. P. (2002) *Plasmodium falciparum* possesses a cell cycle-regulated short type replication protein A large subunit encoded by an unusual transcript. J Biol Chem. **277**, 17493-17501

110 Bennett, T. N., Paguio, M., Gligorijevic, B., Seudieu, C., Kosar, A. D., Davidson, E. and Roepe, P. D. (2004) Novel, rapid, and inexpensive cell-based quantification of antimalarial drug efficacy. Antimicrob Agents Chemother. **48**, 1807-1810

111 Smilkstein, M., Sriwilaijaroen, N., Kelly, J. X., Wilairat, P. and Riscoe, M. (2004) Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening. Antimicrobial agents and chemotherapy. **48**, 1803

112 Berman, H. M., Bhat, T., Bourne, P. E., Feng, Z., Gilliland, G., Weissig, H. and Westbrook, J. (2000) The Protein Data Bank and the challenge of structural genomics. Nature Structural Biology. **7**, 957-959

113 Edgar, R. C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**, 1792-1797

114 Sali, A. and Blundell, T. (1994) Comparative protein modelling by satisfaction of spatial restraints. IOS Press, Amsterdam, The Netherlands

115 Laskowski, R. A. (2009) PDBsum new things. Nucleic Acids Res. **37**, D355-359

116 Wu, G., Robertson, D. H., Brooks, C. L., 3rd and Vieth, M. (2003) Detailed analysis of grid-based molecular docking: A case study of CDOCKER-A CHARMm-based MD docking algorithm. J Comput Chem. **24**, 1549-1562

117 Lipinski, C. A., Lombardo, F., Dominy, B. W. and Feeney, P. J. (1996) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev. **46**, 3-26

118 Heltweg, B., Trapp, J. and Jung, M. (2005) *In vitro* assays for the determination of histone deacetylase activity. Methods. **36**, 332-337

119 Heltweg, B., Dequiedt, F., Verdin, E. and Jung, M. (2003) Nonisotopic substrate for assaying both human zinc and NAD-dependent histone deacetylases. Analytical biochemistry. **319**, 42-48

120 Zhang, J. H., Chung, T. D. and Oldenburg, K. R. (1999) A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. J Biomol Screen. **4**, 67-73

121 Armbruster, D. A. and Pry, T. (2008) Limit of blank, limit of detection and limit of quantitation. The Clinical Biochemist Reviews. **29**, S49

122 Samuels, M. L., Witmer, J. A. and Schaffner, A. (2003) Statistics for the life sciences. Pearson Education

123 Ponts, N., Harris, E. Y., Prudhomme, J., Wick, I., Eckhardt-Ludka, C., Hicks, G. R., Hardiman, G., Lonardi, S. and Le Roch, K. G. (2010) Nucleosome landscape and control of transcription in the human malaria parasite. Genome Res. **20**, 228-238

124 Shechter, D., Dormann, H. L., Allis, C. D. and Hake, S. B. (2007) Extraction, purification and analysis of histones. Nat Protoc. **2**, 1445-1457

125 Smilkstein, M., Sriwilaijaroen, N., Kelly, J. X., Wilairat, P. and Riscoe, M. (2004) Simple and inexpensive fluorescence-based technique for high-throughput antimalarial drug screening. Antimicrob Agents Chemother. **48**, 1803-1806

126 Andrews, K. T., Gupta, A. P., Tran, T. N., Fairlie, D. P., Gobert, G. N. and Bozdech, Z. (2012) Comparative gene expression profiling of *P. falciparum* malaria parasites exposed to three different histone deacetylase inhibitors. PloS one. **7**, e31847

127 Oprea, T. I., Allu, T. K., Fara, D. C., Rad, R. F., Ostopovici, L. and Bologa, C. G. (2007) Lead-like, drug-like or "Pub-like": how different are they? J Comput Aided Mol Des. **21**, 113-119

128 Lipinski, C. A. (2000) Drug-like properties and the causes of poor solubility and poor permeability. J Pharmacol Toxicol Methods. **44**, 235-249

129 Duffy, S. and Avery, V. M. (2012) Development and optimization of a novel 384-well anti-malarial imaging assay validated for high-throughput screening. Am J Trop Med Hyg. **86**, 84-92

130 Towbin, H., Staehelin, T. and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc.Natl.Acad.Sci.USA. **6**, 4350-4354

131 Kurien, B. T. and Scofield, R. H. (2006) Western blotting. Methods. **38**, 283-293

132 Glish, G. L. and Vachet, R. W. (2003) The basics of mass spectrometry in the twenty-first century. Nat Rev Drug Discov. **2**, 140-150

133 Aebersold, R., Burlingame, A. L. and Bradshaw, R. A. (2013) Western Blots versus Selected Reaction Monitoring Assays: Time to Turn the Tables? Mol Cell Proteomics. **12**, 2381-2382

134 Arnaudo, A. M. and Garcia, B. A. (2013) Proteomic characterization of novel histone post-translational modifications. Epigenetics Chromatin. **6**, 24

135 Murray, K. (1966) The acid extraction of histones from calf thymus deoxyribonucleoprotein. J Mol Biol. **15**, 409-419

136 Murray, K. (1969) Stepwise removal of histones from native deoxyribonucleoprotein by titration with acid at low temperature and some properties of the resulting partial nucleoproteins. J Mol Biol. **39**, 125-144

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

138 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

139 Zentner, G. E., Tesar, P. J. and Scacheri, P. C. (2011) Epigenetic signatures distinguish multiple classes of enhancers with distinct cellular functions. Genome Res. **21**, 1273-1283

140 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 Pathog. **9**, e1003170

141 Mackeen, M. M., Kramer, H. B., Chang, K. H., Coleman, M. L., Hopkinson, R. J., Schofield, C. J. and Kessler, B. M. (2010) Small-molecule-based inhibition of histone demethylation in cells assessed by quantitative mass spectrometry. J Proteome Res. **9**, 4082-4092

142 Lombardi, P. M., Cole, K. E., Dowling, D. P. and Christianson, D. W. (2011) Structure, mechanism, and inhibition of histone deacetylases and related metalloenzymes. Curr Opin Struct Biol. **21**, 735-743

143 Bressi, J. C., Jennings, A. J., Skene, R., Wu, Y., Melkus, R., De Jong, R., O'Connell, S., Grimshaw, C. E., Navre, M. and Gangloff, A. R. (2010) Exploration of the HDAC2 foot pocket: Synthesis and SAR of substituted N-(2-aminophenyl)benzamides. Bioorg Med Chem Lett. **20**, 3142-3145

144 Marks, P. A. (2007) Discovery and development of SAHA as an anticancer agent. Oncogene. **26**, 1351-1356

145 Ginsburg, H., Krugliak, M., Eidelman, O. and Cabantchik, Z. I. (1983) New permeability pathways induced in membranes of *Plasmodium falciparum* infected erythrocytes. Mol Biochem Parasitol. **8**, 177-190

146 Makkar, H. S., Siddhuraju, P. and Becker, K. (2007) Saponins. In Plant Secondary Metabolites. pp. 93-100, Humana Press

147 Deutscher, M. P. (1990) [8] Maintaining protein stability. In Methods in Enzymology (Murray, P. D., ed.). pp. 83-89, Academic Press

148 Rodriguez-Collazo, P., Leuba, S. H. and Zlatanova, J. (2009) Robust methods for purification of histones from cultured mammalian cells with the preservation of their native modifications. Nucleic acids research. **37**, e81-e81

149 Keith Wilson and John Walker, ed. (2010) Principles and Techniques of Biochemistry and Molecular Biology. Cambridge University Press

150 Blobel, G. and Potter, V. R. (1966) Nuclei from rat liver: isolation method that combines purity with high yield. Science. **154**, 1662-1665

151 Bashkin, J., Hayes, J. J., Tullius, T. D. and Wolffe, A. P. (1993) Structure of DNA in a nucleosome core at high salt concentration and at high temperature. Biochemistry. **32**, 1895-1898

152 Daly, M. M., Mirsky, A. E. and Ris, H. (1951) The amino acid composition and some properties of histones. J Gen Physiol. **34**, 439-450

153 Sivaraman, T., Kumar, T., Jayaraman, G. and Yu, C. (1997) The mechanism of 2, 2, 2-trichloroacetic acid-induced protein precipitation. Journal of protein chemistry. **16**, 291-297

154 Rajalingam, D., Loftis, C., Xu, J. J. and Kumar, T. K. (2009) Trichloroacetic acid-induced protein precipitation involves the reversible association of a stable partially structured intermediate. Protein Sci. **18**, 980-993

Appendix

Appendix Table 1A: Detailed explanation of the buffers and reagents used for histone isolation methods shown in Figure 2.1 (main text)

Appendix Table 2A: MS-detected histone PTMs from three individual biological replicates using three histone isolation methods (n=3).

Appendix Table 3A: MS-detected histone PTMs from a combination of three individual biological replicates using three histone isolation methods (n=3).

Appendix Table 4A: PTMs detected with the modified Trelle method compared to those found in literature.

Appendix Table 5A: Comparison of treated and untreated histones using the modified Trelle method (n=1).

Appendix Table 1A: Detailed explanation of the buffers and reagents used for histone isolation methods shown in Figure 2.1

Reagent/Technique Used	Method used in	Mechanism of action/ Used for	Ref.
Saponin (0.5%)	All 3	Saponins consist of a steroid and sugar group, which give them surface active qualities. They can therefore interact non-specifically with membrane proteins, phospholipids and cholesterol in erythrocytes, resulting in haemolysis.	[146]
Neutral pH	All 3	Maintenance of pH in a neutral range ensures least amount of protein denaturation or inactivation, since this mimics the cellular environment. Ensures protection of histone PTMs that are acid-labile.	[124, 147]
dddH2O	Issar and Scherf	Allows proteins to form hydrophobic globules, which protects them from the action of any denaturing molecules, and allows them to maintain intermolecular interactions.	[148]
NP-40	Issar and Scherf, Merrick and Trelle	A non-ionic detergent for solubilisation of membranes to aid in protein purification.	[149]
0.25M sucrose	Trelle	Higher density nuclei move towards the bottom of the eppendorf tube due to the application of centrifugal force in a high density sucrose gradient.	[150]
EDTA	Merrick and Trelle	Inhibitor of metallo-proteases, due to its ability to chelate divalent cations	[149]
EGTA	Merrick	Inhibitor of calcium-requiring metallo-proteases.	[149]
DTT	Merrick	Causes reduction of disulphide bridges within proteins.	[149]
Protease Inhibitor Cocktail	Merrick and Trelle	Contains a mixture of serine, cysteine and metallo-proteases, to prevent degradation of proteins that are being isolated by proteolytic enzymes that were contained within the cell's lysosomes.	[149]
0.8M NaCl	Issar and Scherf, and Trelle	Histone octamer is released from the chromatin complex due to the high ionic strength of this solution. Maintains neutral pH, protecting PTMs that are acid-labile.	[124, 151]
0.25M HCl /0.5 M HCl	All 3	Histones are highly basic molecules that have pI values between 10 and 11, this means that they will precipitate out of basic solutions, but will be solubilised in acidic solutions.	[152]
Homogenisation with douncer pestle	Issar and Scherf, Merrick and Trelle	Molecular dissociation of acid-soluble proteins from the chromatin complex due to shear force.	[149]
Combination of acid and salt extraction	Issar and Scherf, and Trelle	Makes it possible to obtain histones that are free within solution and those that are bound in nucleosomes.	[124]
20% TCA	Issar and Scherf, Merrick and Trelle	The trichloro group of TCA assists in protein precipitation by inducing unfolding of proteins due to disruption of electrostatic interactions. This leads to exposure of hydrophobic patches causes proteins to precipitate out of solution.	[153] [154]
Acetone	Issar and Scherf, Merrick and Trelle	Removes excess TCA from the pellet without dissolving any protein.	[124]

Appendix Table 2A: MS-detected histone PTMs from three individual biological replicates using three histone isolation methods (n=3). Each subcolumn contains PTMs detected in a single biological replicate, with the shaded column indicating a combination of the three biological replicates, depicted in Appendix Table 3A. Significance is determined as being present ≥2 times in the three biological repeats. Highly significant modifications are shown in red (3 times) and significant modifications in orange (2 times). Non-significant modifications are depicted in black.

	Issar and Scherf																			
	Acetylation				Monomethylation				Dimethylation				Trimethylation				Phosphorylation			
H2A	N-term	N-term	N-term	N-term	K5	K3	K3	K3	K3	K5	K5	K3	K3	K9	K3	K3	S1	S1	S1	S1
	K3	K3	K3	K3	R 8	K5	K5	K5	R 8	R 8	R 8	K5	K10	K10	K5	K5	T6	T6	T6	T6
	K10	K5	K5	K5	K13	R 8	K9	R 8	K9	K9	K10	R 8	K13	K20	K10	K10	S12	S12	S12	S12
		K9	K9	K9	K20	K9	K10	K9	R 77	K10	K13	K9		K123	K13	K13	T15	T15	T15	T15
		K10	K10	K10	R 29	K10	K13	K10	R 88	K13	K20	K10			K20	K20	S18	S16	S16	S16
		K20	K13	K13	R 88	K13	K20	K13		K20	K74	K13				K123	S76	S18	S18	S18
		K74	K20	K20		K20	R 77	K20		K38	K75	K20					T79		S76	S76
		K75	K35	K35		R 88	R 81	R 29		K41	R 81	K38							T79	T79
		K119	K75	K75			R 88	R 77		K75		K41								
								R 81				K74								
								R 88				K75								
												R 77								
												R 81								
												R 88								
H2A.Z	K10	K10	K10	K10	K10	K14	K27	K10	K27	K24	K5	K5	K14	K10	K10	K10	T30	T30	T30	T30
	K14	K14	K14	K14	K24	K27	K29	K14	K29	K27	K24	K24	K24	K14	K18	K14	S32	S32	S32	S32
	K18	K18	K18	K18	K27	K29	K34	K24	K34	K29	K27	K27	K29	K27	K27	K18	T35	T35	T35	T35
	K24	K24	K24	K24	K29	K34	K36	K27	K36	K34	K29	K29	K34	K29	K29	K27	S41	S41	S41	S41
	K27	K27	K27	K27	K34	K36	R 42	K29	K37	K36	K34	K34	K36	K34	K34	K29	S61	S44	S44	S44
	K29	K29	K29	K29	K36	K156	R 45	K34	R 42	K37	K36	K36		K37	K37	K34	S64	S64	S61	S61
	K34	K34	K34	K34	K37			K36	R 54	R 42	K37	K37			K143	K36		S65	S64	S64
	K36	K36	K36	K36	R 42			K37	K153	R 45	R 42	R 42				K37		T105	T105	S65
	K155	K37	K102	K37	R 45			R 42		R 54	R 54	R 45				K143		T126		T105
		K153	K138	K102	R 62			R 45		K60	R 57	R 54								T126
			K153	K138	R 68			R 62		K138	R 107	R 57								
				K153				R 68		K156	K156	K60								
				K155				K156				R 107								
												K138								
												K153								
												K156								
H2B	K7	K7	N-term	N-term	K3	K4	K7	K3	K4	K3	K7	K3	K10	K9	K4	K4	S2	S2	S2	S2
	K10	K10	K4	K4	K9	K9	K18	K4	K18	K9	K9	K4	K20	K10	K10	K9	T11	T11	T11	T11
	K19	K18	K9	K7	K18	K18	R 21	K7	K19	K10	K10	K7	K112	K20	K112	K10	T13	T13	T13	T13
	K20	K19	K10	K9	K19	K20	K23	K9	R 48	K19	K18	K9		K22		K20	S24	S24	S24	S24
	K64	K20	K18	K10	K23	R 21	R 48	K18	R 77	K23	K19	K18				K22	T44	T44	S28	S28
		K49	K19	K18	R 91	K23	K100	K19	K100	K74	K20	K19				K112	S47	T67	T44	T44

		K64	K20	K19		R 77		K20	K116	K100	K23	K20					S83	T76	S47	S47
		K74	K23	K20		R 78		R 21			R 48	K23					T111	T80	S115	T67
		K108	K38	K23				K23			K49	R 48					T114	S82		T76
		K116	K49	K38				R 48			K116	K49						S83		T80
			K108	K49				R 77				K74						T88		S82
				K64				R 78				R 77						T111		S83
				K74				R 91				K100						T114		T88
				K108				K100				K116								T111
				K116																T114
																				S115
H2B.Z	K3	K3	K3	K3	K8	K8	K13	K8	K4	K3	K3	K3	K8	K8	K8	K8	S1	S1	S1	S1
	K8	K8	K8	K8	K13	K13	K18	K13	K8	K8	K8	K4	K13	K14	K13	K13	S9	S9	S9	S9
	K13	K13	K13	K13	R 23	K14	K82	K14	K14	K14	K13	K8	K14	K18	K18	K14	T15	T15	T15	T15
	K14	K14	K14	K14	K25	K18	R 95	K18	K18	K18	K14	K13	K18	K27	K25	K18	T19	T19	T19	T19
	K18	K18	K18	K18	R 28	K25	K104	R 23	K25	R 23	K18	K14	K104	K42	K42	K25	T30	T48	T30	T30
	K25	K25	K25	K25	R 29	K27		K25	R 26	R 26	R 26	K18			K52	K27	S32	T51	T74	S32
	K27	K27	K27	K27	K52	R 88		K27	R 28	K52	K81	R 23			K112	K42		S87	T84	T51
	K104	K42	K42	K42	R 88	R 95		R 29	R 29	R 88	K82	K25			K116	K52		T92	S86	T74
		K112	K52	K52	R 95			K52			R 83	R 26				K104			S87	T84
			K82	K82				K82			R 88	R 28				K112			T92	S86
			K112	K104				R 88				R 29				K116			S103	S87
				K112				R 95				K52								T92
								K104				K81								S103
												K82								
												R 83								
												R 88								
H3 cen	K20			K20	K20			K20	R 31			R31					S21			S21
					K26			K26	R 166			R166					T24			T24
					R 166			R 166									T27			T27
																	T32			T32
H3.1	K4	K4	K4		R 2	R 8	K4	R 2	K4	R 2	R 2	R 2	K9	K4	K9	K4	T3	T3	T3	T3
	K9	K9	K9	K4	K4	K9	K14	K4	K14	R 8	K4	K4	K14	K14	K14	K9	S10	T6	T6	T6
	K14	K14	K14	K9	R 8	K14	R 17	R 8	R 17	K9	R 8	R 8	K18	K23	K23	K14	T11	S10	S10	S10
	K18	K18	K18	K14	K9	R 17	K18	K9	K18	R 17	K9	K9	K27	K36	K27	K18	S22	T11	T11	T11
	K23	K23	K23	K18	K14	K18	K23	K14		K23	R 17	K14			K64	K23	S32	S22	S22	S22
	K27	K27	K27	K23	R 17	K23	K36	R 17			K18	R 17				K27		S28	S28	S28
	K122	K36	K64	K27	K18	K27	K37	K18			K23	K18				K36		S32	S32	S32
				K36	K23	K79		K23			K36	K23				K64			S57	S57
				K64	R 26			R 26			R 63	K36								
				K122	K36			K27				R 63								
					K37			K36												
					R 40			K37												

								R 40											
H3.3	K4	K4	K4	K4	R 2	R 8	K4	R 2	K4	R 2	R 2	R 2	K9	K4	K4	K4	T3	T3	T3	T3
	K9	K9	K9	K9	K14	K9	R 8	K4	R 8	K4	K4	K4	K14	K9	K9	K9	S10	T6	T6	T6
	K14	K14	K14	K14	R 17	K14	K9	R 8	K9	K9	R 8	R 8	K18	K23	K23	K18	T11	S10	S10	S10
	K18	K18	K18	K18	K18	R 17	K14	K9	R 17	K14	K9	K9	K27	K27	K27	K23	S22	T11	T11	T11
	K23	K23	K23	K23	K23	K18	R 17	K14	K18	R 17	K14	K14		K36	K36	K27	S32	S22	S22	S22
	K115	K27	K27	K27	R 26	K23	K18	R 17		K18	R 17	R 17			K64	K36		S28	S28	S28
	K122		K36	K36	K37	K27	K23	K18		K23	K18	K18				K64		S32	S32	S32
			K37	K37		K36	K37	K23			K36	K23							T33	T33
			K64	K64		K37	R 40	R 26			K37	K36							S57	S57
			K115	K115			K56	K36			R 63	K37							T58	T58
				K122				K37			K122	R 63							T113	T113
								R 40				K122								
								K56												
H4	K5	N-term	K5	N-term		R 3	R 3	R 3	R 3	R 3	R 3	R 3	K16	K5	K20	K5	S1	S1	S1	S1
	K8	K5	K8	K5		K5	K8	K5	K5	K5	K5	K5	K91	K8		K8	T30	S47	S47	T30
	K12	K8	K12	K8		K12	K12	K8	K8	K12	K8	K8		K16		K16	S89	T80	S69	S47
	K16	K12	K16	K12		R 17	K16	K12	K16	K16	K12	K12		K20		K20		T82		S69
	K77	K16		K16		R 35	R 23	K16	R 17	K31	K16	K16		K31		K31		S89		T80
	K79	K31		K31		R 39	R 39	R 17	K79	R 35	K20	R 17		K79		K79				T82
		K44		K44		R 40	K44	R 23	R 92	R 40	R 40	K20				K91				S89
		K79		K77		R 45	R 45	R 35		K44	K77	K31								
		K91		K79		K67	R 78	R 39		K79	K79	R 35								
				K91				R 40				R 40								
								K44				K44								
								R 45				K77								
								K67				K79								
								R 78				R 92								

Trelle

H2A	**Acetylation**				**Monomethylation**				**Dimethylation**				**Trimethylation**				**Phosphorylation**			
	K3	N-term	N-term	N-term	K3	K3	K3	K3	K5	K3	K3	K3	K3	K3	K5	K3	S1	S1	S1	S1
	K10	K3	K3	K3	K5	K5	K9	K5	R 8	K5	K5	K5	K5	K5	K9	K5	T6	T6	T6	T6
	K20	K5	K5	K5	K9	R 8	K10	R 8	K9	R 8	R 8	R 8	K13	K13	K10	K9	S12	S12	S12	S12
	K75	K9	K9	K9	K10	K9	K13	K9	K20	K9	K9	K9	K20	K20	K13	K10	T15	T15	T15	T15
		K10	K10	K10	K13	K13	K20	K10	K35	K10	K10	K10			K20	K13	S16	S16	S16	S16
		K13	K13	K13	K20	K20	R 29	K13	R 42	K13	K13	K13			K119	K20	S18	S18	S18	S18
		K20	K20	K20	R 29	R 29	K74	K20	R 77	K20	K20	K20			K123	K119	T79	S76	S76	S76
			K75	K75	R 88	K36	R 77	R 29		R 29	K74	R 29				K123			T79	T79
						K75	R 88	K36		R 81	K75	K35							S120	S120
						R 77		K75			R 77	R 42								
						R 88		R 77			R 81	K74								
								R 88			R 88	K75								

												R 77							
												R 81							
												R 88							
H2A.Z	K5	K10	K10	K5	K10	K24	K24	K10	K29	K5	K18	K5	K14	K27	K18	K14	T30	T30	T30	T30
	K10	K14	K14	K10	K27	K27	K27	K24	K34	K18	K24	K18	K34	K29	K24	K18	S32	S32	S32	S32
	K14	K18	K18	K14	K29	K29	K29	K27	K36	K27	K29	K24	K36	K34	K27	K24	T35	T35	T35	T35
	K18	K24	K24	K18	K34	K34	K34	K29	K37	K29	K34	K27	K37	K36	K29	K27	S41	S41	S41	S41
	K24	K27	K27	K24	K36	K36	K36	K34	R 42	K34	K36	K29	K152	K37	K34	K29	T149	S44	S44	S44
	K27	K29	K29	K27	R 45	K37	K37	K36	R 45	K36	K37	K34			K36	K34		S61	S64	S61
	K29	K34	K34	K29	R 54	R 42	R 42	K37	K143	K37	R 45	K36			K37	K36		T149	S65	S64
	K34	K36	K36	K34	K143	R 45	R 45	R 42		R 42	R 54	K37			K60	K37				S65
	K36	K37	K37	K36	K152	K60	R 54	R 45		R 152	K155	R 42				K60				T149
	K37	K155	K60	K37	K153	R 68	R 57	R 54				R 45				K152				
	K138		K153	K60	K155		K60	R 57				R 54								
	K152			K138			K155	K60				K143								
				K152			K156	R 68				R 152								
				K153				K152				K155								
				K155				K153												
								K155												
								K156												
H2B	K4	K9	K3	K3	K3	K3	K3	K3	K18	K3	K7	K3	K7	K4	K9	K4	S2	S2	S2	S2
	K75	K10	K4	K4	K4	K4	K4	K4	K49	K9	K9	K7	K10	K7	K10	K7	T11	T11	T11	T11
	K10	K18	K75	K9	K9	K10	K10	K9	R 71	K10	K10	K9	K19	K18	K18	K9	T67	T13	T13	T13
	K18	K19	K9	K10	K10	K18	K18	K10	R 77	K18	K19	K10	K74	K19	K19	K10	S70	S24	S24	S24
	K19	K22	K10	K18	K20	K20	K19	K18	R 78	K19		K18	K116		K20	K18	T111	T44	S28	S28
	K64	K35	K18	K19	R 78	K108	K23	K19	K112	K20		K19			K22	K19		T80	T44	T44
	K74	K49	K19	K20	K108	K116	K35	K20		R 21		K20			K23	K20		T88	S47	S47
	K100	K64	K20	K22	K112		R 48	K23		K35		R 21			K116	K22		S115	T76	S70
		K74	K22	K35			K100	K35		R 84		K35				K23			S83	T76
		K108	K64	K49			K112	R 48		K100		K49				K74			T88	T80
		K112	K108	K64			K116	R 78				R 71				K116			S104	S83
			K112	K74				K100				R 77							S115	T88
				K100				K108				R 78								S104
				K108				K112				R 84								T111
				K112				K116				K100								S115
												K112								
H2B.Z	N-term	N-term	K3	N-term	K13	K8	K8	K8	K14	K8	K3	K3	K3	K8	K8	K3	S9	S1	S1	S1
	K3	K3	K8	K3	R 23	K13	K14	K13	K18	K13	K8	K8	K14	K13	K13	K8	T15	S9	S9	S9
	K8	K8	K13	K8	R 26	K18	R 23	K14	R 23	K14	K18	K13		K14	K14	K13	T84	T15	T15	T15
	K13	K13	K14	K13	K81	R 23	K25	K18	K25	K18	K25	K14		K25	K18	K14	S86	T19	T19	T19
	K14	K14	K18	K14	K82	K25	R 26	R 23	R 26	R 23	R 29	K18		K27	K52	K18	T115	T30	T30	T30
	K18	K18	K25	K18	R 95	R 26	K39	K25	K27	K25	K39	R 23		K52	K104	K25		S34	S32	S32

	K25	K25	K42	K25	K112	R 28	K52	R 26	K81	R 26	K53	K25		K104		K27		T48	S34	S34
	K27	K104	R 75	K27		R 29	R 95	R 28	K116	K27		R 26				K52		T51	T48	T48
	K104		K104	K42		K52	K116	R 29		R 28		K27				K104		T71	T51	T51
			K112	K104		R 75		K39		R 68		R 28						T84	S86	T71
				K112		R 95		K52		R 83		R 29						S87	S87	T84
						K104		R 75		K112		K39						S108	T92	S86
								K81				K53							S108	S87
								K82				R 68							T111	T92
								R 95				K81							T119	S108
								K104				R 83								T111
								K112				K112								T115
								K116				K116								T119
H3.1	N-term	N-term	K4	N-term	R 2	R 2	K4	R 2	R 2	R 8	R 2	R 2	K14	K14	K9	K9	T3	T3	T3	T3
	K9	K4	K9	K4	K4	K4	R 8	K4	K14	K9	K4	K4	K18	K18	K14	K14	T6	T6	T6	T6
	K14	K9	K14	K9	R 8	R 8	K9	R 8	R 17	K14	R 8	R 8	K23	K23	K27	K18	S10	S10	S10	S10
	K18	K14	K18	K14	K9	K14	K14	K9	K18	R 17	K9	K9	K27	K27		K23	T11	T11	T11	T11
	K23	K18	K23	K18	K14	R 17	R 17	K14	K23	K18	K14	K14		K37		K27	S22	S22	S22	S22
	K27	K23	K27	K23	R 17	K18	K18	R 17		K23	R 17	R 17				K37	S28	S28	S28	S28
		K27	K36	K27	K18	K23	K23	K18		K27	K18	K18					S32	T33	S32	S32
		K36	K122	K36	K23	R 26	K27	K23		R 40	R 26	K23						S135	T33	T33
				K122	R 26	K27		R 26		K53	K27	R 26							S57	S57
						R 128		K27			K36	K27								S135
						R 129		R 128			R 42	K36								
								R 129				R 40								
												R 42								
												K53								
H3 cen		K4		K4	R 31			R 31	R 2			R 2	K20	K23		K20		S21	T3	T3
		K20		K20	R 46			R 46	K26			K26	K23	K26		K23		T24	S21	S21
		K23		K23	R 68			R 68	R 31			R 31	K26	K69		K26		T27	T24	T24
		K26		K26	R 72			R 72	K40			K40				K69		T32	T27	T27
		K69		K69	R 168			R 168	R 51			R 51						S38	S38	T32
																		S39	S39	S38
																			T41	S39
																				T41
H3.3	N-term	N-term	N-term	N-term	R 2	R 2	K4	R 2	R 2	R 8	R 2	R 2	K4	K8	K9	K4	T3	T3	T3	T3
	K4	K4	K4	K4	K4	K4	R 8	K4	R 8	K9	K4	K4	K9	K14	K14	K9	T6	T6	T6	T6
	K9	K9	K9	K9	R 8	R 8	K9	R 8	K9	K14	R 8	R 8	K14	K18	K27	K14	S10	S10	S10	S10
	K14	K14	K14	K14	K9	K14	K14	K9	K14	R 17	K9	K9	K27	K27		K18	T11	T11	T11	T11
	K18	K18	K18	K18	K14	R 17	R 17	K14	R 17	K18	K14	K14	K37	K37		K27	S22	S22	S22	S22
	K23	K23	K23	K23	R 17	K18	K18	R 17	K18	K23	R 17	R 17				K37	S32	S28	S28	S28
	K37	K27	K27	K27	K18	K23	K23	K18	K23	K27	K18	K18					T58	T33	S32	S32
	K53	K36	K36	K36	K23	R 26	K27	K23	R 41	R 40	R 26	K23						S135	T33	T33

	K56		K53	K37	R 26	K27	K36	R 26	R 49	K53	K27	R 26						S57	S57
			K122	K53	K35	R 128	R 52	K27			K36	K27							T58
				K56	K36	R 129		K35			R 42	K36							S135
				K122	K37			K36				R 40							
					K56			K37				R 41							
					R 63			R 52				R 42							
								K56				R 49							
								R 63				K53							
								R 128											
								R 129											
H4	K5	N-term	K5	N-term	K8	R 3	K12	R 3	K5	K5	K31	K5	K5	K20	K5	K5	S1	T73	S1	S1
	K8	K5	K8	K5	K12	K8	R 35	K8	K12	K12	R 78	K12	K12	K31	K12	K12		T80	S89	T73
	K12	K8	K12	K8	R 50	K12	R 40	K12	K16	K16		K16			K20	K20		S89		T80
	K16	K12	K16	K12		K16	K91	K16		R 36		K31			K31	K31				S89
		K16	K31	K16		R 17		R 17				R 36								
		K31	K67	K31		R 35		R 35				R 78								
		K44		K44		R 55		R 40												
		K59		K59		K91		R 50												
		K79		K67				R 55												
				K79				K91												

Merrick

	Acetylation				Monomethylation				Dimethylation				Trimethylation				Phosphorylation			
H2A	N-term	N-term	N-term	N-term	K5	K3	K3	K3	K3	K5	K3	K3	K10	K3	K3	K3	S1	S1	S1	S1
	K3	K3	K3	K3	K9	K5	K5	K5	K5	R 8	K5	K5	K13	K9	K5	K5	T6	T6	T6	T6
	K5	K5	K5	K5	K10	K9	R 8	R 8	R 8	K9	R 8	R 8	K20	K10	K9	K9	S12	S12	S12	S12
	K9	K10	K9	K9	K13	K10	K9	K9	K9	K10	K9	K9	K38	K13	K13	K10	T15	T15	T15	T15
	K35	K13	K10	K10	K20	K13	K10	K10	K10	K13	K10	K10		K20		K13	S18	S16	S16	S16
	K41	K20	K35	K13	K38	R 32	K13	K13	K13	K20	K20	K13				K20		S18	S18	S18
		K35	K36	K20	R 88	K74	R 29	K20	K20	R 29	K36	K20				K38		S76	S76	S76
		K75	K41	K35		K75	K36	R 29		K75	K74	R 29						T79	T79	T79
			K75	K36		R 88	K38	R 32		R 77	R 81	K36								
				K41			R 88	K36		R 81		K74								
				K75				K38		R 88		K75								
								K74				R 77								
								K75				R 81								
								R 88				R 88								
H2A.Z	K10	K10	K10	K10	K14	K18	K14	K14	K27	K29	K5	K5	K10	K24	K10	K10	T30	T30	T30	T30
	K14	K14	K14	K14	K29	K29	K27	K18	K29	K34	K24	K24	K24	K27	K27	K24	S32	S32	S32	S32
	K18	K18	K18	K18	K34	K34	K29	K27	K34	K36	K29	K27	K27	K34	K29	K27	T35	T35	T35	T35
	K24	K24	K24	K24	K36	K36	K34	K29	K36	K37	K34	K29	K29	K36	K34	K29	S41	S41	S41	S41
	K27	K27	K27	K27	K37	K37	K36	K34		R 42	K36	K34	K34	K37	K37	K34	S44	S44	S44	S44
	K29	K29	K29	K29		R 42	K37	K36	R 42	R 45	K37	K36	K36	K143		K36	T105	S61	S61	S61

	K34	K34	K34	K34		R 45	R 42	K37	R 45		R 45	K37	K37			K37		S65	S65	S65
	K36	K36	K37	K36		R 54	R 45	R 42			K60	R 42				K143		S71	T105	S71
	K37	K37	K60	K37		K97	R 57	R 45			R 62	R 45						T72	T149	T72
		K102	K152	K60			K60	R 54			R 68	K60						T105		T105
		K138	K153	K102			R 156	R 57			K156	R 62						T149		T149
		K143		K138				K60				R 68								
		K155		K143				K97				K156								
		K156		K152				R 156												
				K153																
				K155																
				K156																
H2B	K3	K9	K3	K3	K18	K10	K4	K4	K7	K7	K4	K4	K4	K10	K9	K4	S2	T11	T11	S2
	K7	K10	K4	K4	K23	K18	K10	K10	K10	K10	K7	K7	K7	K18	K10	K7	T11	T13	T13	T11
	K10	K19	K7	K7	R 84	K19	K18	K18	K22	K19	K9	K9	K9	K38	K19	K9	T13	S24	T67	T13
	K18	K20	K9	K9		R 48	K19	K19	K23	K20	K10	K10	K18	K108	K74	K10	S82	S28	T80	S24
	K19	K22	K10	K10			K35	K23	K49	R 21	K19	K19	K19	K112		K18	S104	S50	S82	S28
	K20	K64	K18	K18			R 48	K35		K64	R 21	K20	K112	K116		K19	S115	S104	S115	S50
	K64	K108	K19	K19			K59	R 48		K116	K38	R 21				K38		T114		T67
	K100	K112	K35	K20			R 71	K59			K59	K22				K74		S115		T80
		K116	K59	K22			R 77	R 71			K74	K23				K108				S82
			K64	K35			K100	R 77			K116	K38				K112				S104
			K74	K59				R 84				K49				K116				T114
			K109	K64				K100				K59								S115
				K74								K64								
				K100								K74								
				K108								K116								
				K109																
				K116																
H2B.Z	K8	N-term	N-term	N-term	K13	K3	K8	K3	K3	K3	K3	K3	K13	K13	K8	K8	T15	S1	S1	S1
	K13	K3	K3	K3	K14	K14	K14	K8	K14	K8	K8	K8	K14	K14	K13	K13	T19	S9	S9	S9
	K14	K8	K8	K8	K18	R 23	K18	K13	K18	K13	K14	K13	K82	K18	K18	K14	T48	T15	T15	T15
	K18	K13	K13	K13	R 95	K25	R 23	K14	R 23	K14	K18	K14	K104	K25		K18	T71	T19	T19	T19
	K104	K14	K14	K14		K82	R 68	K18	R 26	K18	R 23	K18				K25	T84	T84	T74	T48
		K18	K18	K18		R 83	R 95	R 23	K42	R 23	K25	R 23				K82		S87	T84	T71
		K25	K53	K25		R 95		K25	K52	K25	R 26	K25				K104		T92	S86	T74
		K27	K112	K27				R 68	R 75	R 68	K53	R 26							S87	T84
		K104		K53				K82			R 68	K52							T92	S86
				K104				R 83			R 88	K53								S87
				K112				R 95				R 68								T92
												R 75								

												R 88								
H3.1	K4	K4	N-term	N-term	R 2	R 8	R 2	R 2	R 8	K4	R 2	R 2	K9	K4	K4	K4	T6	T3	T3	T3
	K9	K9	K4	K4	K4	K14	K4	K4	K9	R 8	K4	K4	K14	K9	K9	K9	S10	T6	T6	T6
	K14	K14	K9	K9	R 8	R 17	R 8	R 8	K14	K9	K9	R 8		K14	K14	K14	T11	S10	S10	S10
	K18	K18	K14	K14	K9	K18	K9	K9	R 17	K14	K14	K9		K36	K18	K18	S22	T11	T11	T11
	K23	K23	K18	K18	K14	K23	R 17	K14	K18	R 17	R 17	K14		K40	K23	K23	S32	S22	S22	S22
	K27	K27	K23	K23	R 17	K27	K18	R 17	K23	K18	R 26	R 17			K27	K36	S57	S28	S32	S28
	K36	K36	K27	K27	K18	K36	R 26	K18	R 83	K23	K36	K18				K40	T58	S32	S57	S32
	K79	K37		K36	R 26	K79	K36	K23		R 26	K37	K23						T45	T58	T45
		K40		K37	K36		R 40	R 26		K36		R 26								S57
				K40			R 49	K27		R 49		K36								T58
				K79			R 52	K36		R 52		K37								
								R 40		R 134		R 49								
								R 49				R 52								
								R 52				R 83								
								K79				R 134								
H3.3	K4	K4	N-term	N-term	R 2	R 2	R 2	R 2	K4	K4	R 2	R 2	K14	K4	K4	K4	S10	T3	T3	T3
	K9	K9	K4	K4	K4	K4	K4	K4	K9	R 8	K4	K4	K23	K9	K9	K9	T11	T6	T6	T6
	K14	K14	K9	K9	R 8	K14	R 8	R 8	K14	K9	K9	R 8	K27	K14	K14	K14	S22	S10	S10	S10
	K18	K18	K14	K14	K9	R 17	K9	K9	R 17	K14	K14	K9	K37	K23	K18	K18	S28	T11	T11	T11
	K23	K23	K18	K18	K14	K18	K14	K14		R 17	R 17	K14		K37	K23	K23	S32	S22	S22	S22
	K27	K27	K23	K23	R 17	K27	R 17	R 17		K18	K18	R 17			K27	K27	T33	S28	S28	S28
		K36	K27	K27	K18	R 40	K18	K18		K23	K27	K18			K36	K36		S32	S32	S32
		K37	K36	K36	K23	K122	K23	K23		R 26	R 49	K23			K53	K37		T33	T33	T33
		K53		K37	K36		R 26	R 26		K36	K56	R 26				K53		T45	T45	T45
				K53			K27	K27		R 40		K27						T118	T58	T58
							K37	K36		R 49		K36								T118
							R 42	K37		R 116		R 40								
							R 49	R 40		R 134		R 49								
								R 42				K56								
								R 49				R 116								
								K122				R 134								
H4	N-term	K5	N-term	N-term	R 3	R 3	R 3	R 3	R 3	R 3	R 3	R 3	K12	K12	K8	K8	S1	S1	S1	S1
	K5	K8	K5	K5	K5	K8	K5	K5	K5	K8	K5	K5		K20	K16	K12	T80	T30	T80	T30
	K8	K12	K8	K8	K12	K12	K8	K8	K8	K12	K8	K8			K20	K16	T82	S69	T82	S69
	K12	K16	K12	K12		K31	K12	K12	K12	K31	K31	K12			K44	K20	S89		S89	T80
	K16	K31	K16	K16		R 35	R 17	R 17	K16	R 35	R 35	K16				K44				T82
	K79	K44	K20	K20			R 19	R 19	K79	K77	R 40	K31								S89
		K67	K31	K31			R 35	K31			K59	R 35								
			K44	K44			R 39	R 35			K77	R 40								
			K79	K67			R 45	R 39				K59								

			K91	K79			R 78	R 45				K77							
				K91			K91	R 78				K79							
							R 92	K91											
								R 92											

Appendix Table 2A: MS-detected histone PTMs from a combination of three individual biological replicates using three histone isolation methods (n=3). Significance is determined as being present ≥2 times in the three biological repeats (as in Table 2A). Highly significant modifications are shown in red (3 times) and significant modifications in orange (2 times). Non-significant modifications are depicted in black. The number of significant modifications are presented (in bold red) as a fraction over the total number of modifications detected.

	Acetylation			Monomethylation			Dimethylation			Trimethylation			Phosphorylation		
H2A	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
	N-term	N-term	N-term	K3	K3	K3	K3	K3	K3	K3	K3	K3	S1	S1	S1
	K3	K3	K3	K5	K5	K5	K5	K5	K5	K5	K5	K5	T6	T6	T6
	K5	K5	K5	R 8	R 8	R 8	R 8	R 8	R 8		K9	K9	S12	S12	S12
	K9	K9	K9	K9	K9	K9	K9	K9	K9	K10	K10	K10	T15	T15	T15
	K10	K10	K10	K10	K10	K10	K10	K10	K10	K13	K13	K13	S16	S16	S16
	K13	K13	K13	K13	K13	K13	K13	K13	K13	K20	K20	K20	S18	S18	S18
	K20	K20	K20	K20	K20	K20	K20	K20	K20			K38	S76	S76	S76
	K35		K35	R 29	R 29	R 29		R 29	R 29		K119		T79	T79	T79
			K36			R 32		K35		K123	K123			S120	
			K41		K36	K36			K36						
	K75	K75	K75			K38	K38								
						K74	K41								
					K75	K75		R 42							
				R 77	R 77		K74	K74	K74						
				R 81			K75	K75	K75						
				R 88	R 88	R 88	R 77	R 77	R 77						
							R 81	R 81	R 81						
							R 88	R 88	R 88						
	7/9	**8/8**	**8/11**	**7/11**	**9/12**	**7/14**	**7/14**	**9/15**	**8/14**	**4/9**	**4/8**	**5/7**	**8/8**	**8/9**	**8/8**
	Acetylation			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
H2A.Z	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
		K5		K10	K10		K5	K5	K5	K10		K10	T30	T30	T30
	K10	K10	K10	K14		K14		K18		K14	K14		S32	S32	S32
	K14	K14	K14			K18	K24	K24	K24	K18	K18		T35	T35	T35
	K18	K18	K18	K24	K24		K27	K27	K27		K24	K24	S41	S41	S41
	K24	K24	K24	K27	K27	K27	K29	K29	K29	K27	K27	K27	S44	S44	S44
	K27	K27	K27	K29	K29	K29	K34	K34	K34	K29	K29	K29	S61	S61	S61
	K29	K29	K29	K34	K34	K34	K36	K36	K36	K34	K34	K34	S64	S64	
	K34	K34	K34	K36	K36	K36	K37	K37	K37	K36	K36	K36	S65	S65	S65
	K36	K36	K36	K37	K37	K37	R 42	R 42	R 42	K37	K37	K37			S71
	K37	K37	K37	R 42	R 42	R 42	R 45	R 45	R 45		K60				T72
		K60	K60	R 45	R 45	R 45	R 54	R 54		K143		K143	T105		T105
	K102		K102		R 54	R 54	R 57				K152		T126		
	K138	K138	K138		R 57	R 57	K60		K60					T149	T149
			K143		K60	K60			R 62						
		K152	K152	R 62					R 68						

	K153	K153	K153	R 68	R 68		R 107									
	K155	K155	K155	K156		K97	K138									
			K156		K152			K143								
					K153			R 152								
					K155		K153	K155								
					K156	R 156										
							K156		K156							
	10/13	**9/15**	**9/17**	**7/13**	**11/17**	**7/14**	**7/16**	**7/14**	**6/13**	**8/9**	**5/10**	**7/8**	**7/10**	**6/9**	**9/11**	
H2B	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**			
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	
	N-term			K3	K3		K3	K3		K4	K4	K4	S2	S2	S2	
		K3	K3	K4	K4	K4	K4		K4		K7	K7	T11	T11	T11	
	K4	K4	K4	K7			K7	K7	K7	K9	K9	K9	T13	T13	T13	
	K7		K7	K9	K9		K9	K9	K9	K10	K10	K10	S24	S24	S24	
	K9	K9	K9		K10	K10		K10	K10		K18	K18	S28	S28	S28	
	K10	K10	K10	K18	K18	K18	K18	K18			K19	K19	T44	T44		
	K18	K18	K18	K19	K19	K19	K19	K19	K19	K20	K20		S47	S47		
	K19	K19	K19	K20	K20		K20	K20	K20	K22	K22				S50	
	K20	K20	K20	R 21				R 21	R 21		K23		T67		T67	
		K22	K22	K23	K23	K23			K22			K38		S70		
	K23				K35	K35	K23		K23		K74	K74	T76	T76	T80	
		K35	K35	R 48	R 48	R 48		K35				K108	T80	T80	S82	
	K38					K59			K38	K112		K112	S82			
	K49	K49				R 71	R 48				K116	K116	S83	S83		
			K59	R 77		R 77	K49	K49	K49				T88	T88		
	K64	K64	K64	R 78	R 78				K59					S104	S104	
	K74	K74	K74			R 84			K64				T111	T111		
		K100	K100	R 91				R 71					T114	S115	T114	
	K108	K108	K108	K100	K100	K100	K74		K74				S115		S115	
			K109		K108		R 77	R 77								
		K112			K112			R 78								
	K116		K116		K116			R 84								
							K100	K100								
								K112								
							K116		K116							
	7/15	**9/15**	**7/17**	**4/14**	**8/15**	**4/15**	**7/14**	**4/16**	**5/15**	**3/6**	**5/11**	**5/11**	**9/16**	**7/15**	**3/12**	
H2B.Z	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**			
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	
		N-term	N-term			K3	K3	K3	K3		K3		S1	S1	S1	
	K3	K3	K3	K8	K8	K8	K4			K8	K8	K8	S9	S9	S9	
	K8	K8	K8	K13	K13	K13	K8	K8	K8	K13	K13	K13	T15	T15	T15	

	K13	K13	K13	K14	K14	K14	K13	K13	K13	K14	K14	K14	T19	T19	T19
	K14	K14	K14	K18	K18	K18	K14	K14	K14	K18	K18	K18	T30	T30	
	K18	K18	K18	R 23	R 23	R 23	K18	K18	K18	K25	K25	K25	S32	S32	
	K25	K25	K25	K25	K25	K25	R 23	R 23	R 23	K27	K27			S34	
	K27	K27	K27		R 26		K25	K25	K25	K42				T48	T48
	K42	K42		K27			R 26	R 26	R 26	K52	K52		T51	T51	
	K52				R 28			K27				K82		T71	T71
	K82		K53	R 29	R 29		R 28	R 28		K104	K104	K104	T74		T74
	K104	K104	K104		K39		R 29	R 29		K112			T84	T84	T84
	K112	K112	K112	K52	K52			K39		K116			S86	S86	S86
						R 68	K52		K52				S87	S87	S87
					R 75			K53	K53				T92	T92	T92
					K81			R 68	R 68				S103		
				K82	K82	K82			R 75					S108	
						R 83	K81	K81						T111	
				R 88			K82							T115	
				R 95	R 95	R 95	R 83	R 83						T119	
				K104	K104		R 88		R 88						
					K112			K112							
					K116			K116							
	8/12	**8/11**	**7/11**	**5/13**	**6/17**	**4/13**	**6/15**	**6/17**	**8/12**	**5/11**	**5/9**	**3/7**	**7/12**	**12/17**	**7/10**
H3 cen	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
		K4		K20				R 2			K20			T3	
	K20	K20		K26				K26			K23		S21	S21	
		K23			R 31		R 31	R 31			K26		T24	T24	
		K26			R 46			K40			K69		T27	T27	
		K69			R 68			R 51					T32	T32	
					R 72		R 166							S38	
				R 166										S39	
					R 168									T41	
	0/1	**0/5**		**0/3**	**0/5**		**0/2**	**0/5**			**2/4**		**0/4**	**5/8**	
H3.1	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
		N-term	N-term	R 2	R 2	R 2	R 2	R 2	R 2	K4		K4	T3	T3	T3
	K4	K4	K4	K4	K4	K4	K4	K4	K4	K9	K9	K9	T6	T6	T6
	K9	K9	K9	R 8	R 8	R 8	R 8	R 8	R 8	K14	K14	K14	S10	S10	S10
	K14	K14	K14	K9	K9	K9	K9	K9	K9	K18	K18	K18	T11	T11	T11
	K18	K18	K18	K14	K14	K14	K14	K14	K14	K23	K23	K23	S22	S22	S22
	K23	K23	K23	R 17	R 17	R 17	R 17	R 17	R 17	K27	K27		S28	S28	S28
	K27	K27	K27	K18	K18	K18	K18	K18	K18	K36		K36	S32	S32	S32
	K36	K36	K36	K23	K23	K23	K23	K23	K23		K37			T33	

			K37	R 26	R 26	R 26		R 26	R 26	K64					T45
			K40	K27	K27	K27		K27					S57	S57	S57
	K64			K36		K36	K36	K36	K36						T58
			K79	K37					K37					S135	
	K122	K122		R 40		R 40		R 40							
						R 49		R 42							
						R 52			R 49						
						K79			R 52						
					R 128			K53							
					R 129		R 63								
									R 83						
									R 134						
	6/9	**8/9**	**7/11**	**9/13**	**10/12**	**9/15**	**7/10**	**8/14**	**9/15**	**4/8**	**5/6**	**3/6**	**7/8**	**8/10**	**8/11**
H3.3	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
		N-term	N-term	R 2	R 2	R 2	R 2	R 2	R 2	K4	K4	K4	T3	T3	T3
	K4	K4	K4	K4	K4	K4	K4	K4	K4	K9	K9	K9	T6	T6	T6
	K9	K9	K9	R 8	R 8	R 8	R 8	R 8	R 8		K14	K14	S10	S10	S10
	K14	K14	K14	K9	K9	K9	K9	K9	K9	K18	K18	K18	T11	T11	T11
	K18	K18	K18	K14	K14	K14	K14	K14	K14	K23		K23	S22	S22	S22
	K23	K23	K23	R 17	R 17	R 17	R 17	R 17	R 17	K27	K27	K27	S28	S28	S28
	K27	K27	K27	K18	K18	K18	K18	K18	K18	K36		K36	S32	S32	S32
	K36	K36	K36	K23	K23	K23	K23	K23	K23		K37	K37	T33	T33	T33
	K37	K37	K37	R 26	R 26	R 26		R 26	R 26			K53			T45
		K53	K53		K27	K27		K27	K27	K64			S57	S57	
		K56			K35		K36	K36	K36				T58	T58	T58
	K64			K36	K36	K36	K37						T113		
	K115			K37	K37	K37		R 40	R 40						T118
	K122	K122		R 40		R 40		R 41						S135	
						R 42		R 42							
						R 49		R 49	R 49						
					R 52			K53							
				K56	K56				K56						
					R 63		R 63								
						K122			R 116						
					R 128		K122								
					R 129				R 134						
	7/11	**9/12**	**7/10**	**7/13**	**11/18**	**9/16**	**7/12**	**9/16**	**6/16**	**5/7**	**4/6**	**6/9**	**7/11**	**7/11**	**9/11**
H4	**Acetylation**			**Monomethylation**			**Dimethylation**			**Trimethylation**			**Phosphorylation**		
	I+S	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**	**I+S**	**Trelle**	**Merrick**
	N-term	N-term	N-term	R 3	R 3	R 3	R 3		R 3	K5	K5		S1	S1	S1

	K5	K5	K5	K5		K5	K5	K5	K5	K8		K8	T30		T30
	K8	K8	K8	K8	K8	K8	K8		K8		K12	K12	S47		
	K12	K12	K12	K12	K12	K12	K12	K12	K12	K16		K16	S69		S69
	K16	K16	K16	K16	K16		K16	K16	K16	K20	K20	K20		T73	
			K20	R 17	R 17	R 17	R 17			K31	K31		T80	T80	T80
	K31	K31	K31			R 19	K20					K44	T82		T82
	K44	K44	K44	R 23			K31	K31	K31	K79			S89	S89	S89
		K59				K31	R 35		R 35	K91					
		K67	K67	R 35	R 35	R 35		R 36							
	K77			R 39		R 39	R 40		R 40						
	K79	K79	K79	R 40	R 40		K44								
	K91		K91	K44					K59						
				R 45		R 45	K77		K77						
					R 50		K79	R 78							
					R 55				K79						
				K67			R 92								
				R 78		R 78									
					K91	K91									
						R 92									
	5/10	**5/10**	**8/11**	**6/14**	**4/10**	**5/13**	**7/14**	**3/6**	**7/11**	**2/7**	**4/4**	**2/5**	**3/7**	**2/4**	**4/6**

Appendix Table 4A: PTMs detected with the modified Trelle method compared to those found in literature. Only the PTMs that correlate to those found in literature are listed. Sub-column one indicates all the modifications detected significantly in red or orange (3 times or 2 times respectively) or non-significantly in pink (1 time). Sub-column two shows the modifications found in literature and sub-column three shows the literature reference. Information compiled from [37, 49, 50, 65, 66].

	Acetylation			Monomethylation			Dimethylation			Trimethylation			Phosphorylation			Ubiquitination		
H2A	N-term	N-term	[49]										S16	S16	[65]			
	K3	K3	[49]										S18	S18	[65]			
	K5	K5	[49]											S104	[66]			
													S120	S120	[65, 66]			
														T126	[65, 66]			
H2A.Z	K5	N-term	[49]										S32	S32	[65]			
	K10	K10	[49]										T35	T35	[65]			
	K14	K14	[49]															
	K18	K18	[49]															
	K24	K24	[49, 50, 65]															
	K27	K27	[49, 50, 65]															
	K29	K29	[49, 50, 65]															
	K34	K34	[49, 50]															
	K36	K36	[65]															
H2B	K22	K22	[65]	K23	K23	[65]							S24	S24	[65]		K112	[49]
														Y26	[65]			
													S28	S28	[66]			
														S50	[66]			
														S56	[66]			
													S104	S104	[65]			
														T107	[65]			
H2B.Z	N-term	N-term	[49, 50]										S1	S1	[65]			
	K3	K3	[49, 65]										S32	S32	[66]			
	K8	K8	[49, 65]										T84	T84	[50]			
	K13	K13	[49, 50, 65]										S108	S108	[65]			
	K14	K14	[49]															
	K18	K18	[49, 50]															
H3.1	K9	K9	[49, 50, 65]	K4	K4	[49]	K4	K4	[49]		K4	[49]	S10	S10	[65]			
	K14	K14	[49, 50, 65]	K9	K9	[37]	R 17	R 17	[49]	K9	K9	[49]	T11	T11	[65]			
	K18	K18	[49, 50, 65]	K14	K14	[50]	K18	K18	[65]		K36	[37]	S22	S22	[65]			
	K23	K23	[49, 65]	R 17	R 17	[49, 50, 65]					K79	[37]	S28	S28	[65, 66]			
	K27	K27	[49, 50, 65]	K18	K18	[65]							S32	S32	[65, 66]			

		K56	[49]		K36	[65]								T45	[65]			
					K37	[65]							S57	S57	[65]			
					R 40	[65]								T58	[65]			
H3.3	K9	K9	[49, 50, 65]	K4	K4	[49]	K4	K4	[49]	K4	K4	[49]	S10	S10	[65]			
	K14	K14	[49, 50, 65]	R 17	R 17	[49]	R 17	R 17	[49]				T11	T11	[65]			
	K18	K18	[49, 65]	K18	K18	[65]							S22	S22	[65]			
	K23	K23	[49, 65]										S28	S28	[65, 66]			
	K27	K27	[49, 65]										S32	S32	[65, 66]			
													T33	T33	[65]			
														T45	[65]			
													S57	S57	[65, 66]			
													T58	T58	[65, 66]			
H4	N-term	N-term	[49]	R 3	R 3	[49]		K20	[37]	K20	K20	[49]						
	K5	K5	[49]		K5	[50]												
	K8	K8	[49, 50]	K16	K16	[50]												
	K12	K12	[49, 50]	R 17	R 17	[50]												
	K16	K16	[49]		K20	[37]												

Appendix Table 5A: Comparison of treated and untreated histones using the modified Trelle method (n=1). The number of modifications detected is shown in red text. The data below are summarised in Table 3.6 (main text). UT=Untreated and T=Treated. Differences detected between the PTMs detected are highlighted in yellow.

	Modified Trelle method									
H2A	Acetylation		Monomethylation		Dimethylation		Trimethylation		Phosphorylation	
	UT	T	UT	T	UT	T	UT	T	UT	T
	K3	K3	K3	K3	K5		K3		S1	S1
		K9		R 8	R 8	R 8	K5		T6	T6
	K10		K5		K9	K9	K13	K13	S12	S12
	K20	K20	K9			K10	K20	K20	T15	T15
		K35	K10	K10		K13		K41	S16	S16
		K38	K13	K13	K20	K20			S18	S18
		K74	K20			R 29				S76
	K75		R 29		K35				T79	
				K36	R 42					
				K38	R 77					
				R 42						
				K74						
				K75						
			R 88	R 88						
	4	6	8	10	7	6	4	3	7	7
H2A.Z	Acetylation		Monomethylation		Dimethylation		Trimethylation		Phosphorylation	
	UT	T	UT	T	UT	T	UT	T	UT	T
	K5		K10			K26	K14		T30	T30
		K9		K14	K29	K29		K18	S32	S32
	K10		K26	K26	K34	K34		K24	T35	T35
	K14	K14	K29	K29	K36	K36		K26	S41	S41
	K18	K18	K34	K34	K37	K37		K29		S44
	K24	K24	K36	K36	R 42	R 42	K34	K34		S61
	K26	K26		R 42	R 45	R 45	K36			S65
	K29	K29	R 45	R 45		R 57	K37		T149	
	K34	K34	R 54			K60	K152			
	K36	K36	K143			R 62				
	K37		K152		K143					
	K138		K153			K155				
	K152	K152	K155							
	12	9	11	7	7	11	5	5	5	7
H2B	Acetylation		Monomethylation		Dimethylation		Trimethylation		Phosphorylation	
	UT	T	UT	T	UT	T	UT	T	UT	T
	K4	K4	K3	K3		K4	K7		S2	S2
	K7	K7	K4	K4		K7		K9	T11	T11
		K9	K9	K9		K10	K10			T13
	K10	K10			K18		K19			S24
	K18		K10		K49		K74			S28
	K19	K19	K20		R 71		K116		T67	
		K20			R 77				S70	
	K64	K64		K64	R 78					S83
	R 71		R 78			R 84			T111	
	K74			R 91		K100				
		R 78		K100	K112					
		R 91	K108							
	K100		K112							
	9	9	8	6	6	5	5	1	5	6
H2B.Z	Acetylation		Monomethylation		Dimethylation		Trimethylation		Phosphorylation	
	UT	T	UT	T	UT	T	UT	T	UT	T

	N-term		K13			K8	K3	K3		S2
	K3	K3		K14	K14			K8	S9	
	K8	K8	R 23		K18			K13	T15	T15
	K13	K13		K25	R 23	R 23	K14			T19
	K14	K14	R 26	R 26	K25	K25		K18		T74
	K18	K18		K27	R 26				T84	T84
	K25			R 28	K27	K27			S86	S86
	K27	K27		K39	K81	R 29				S87
		R 75		R 78	K116				T115	
	R 78		K81	K81						
		K81	K82							
	K104			R 83						
			R 95							
			K112							
	10	**8**	**7**	**9**	**8**	**5**	**2**	**4**	**5**	**7**
H3.1	**Acetylation**		**Monomethylation**		**Dimethylation**		**Trimethylation**		**Phosphorylation**	
	UT	**T**	**UT**	**T**	**UT**	**T**	**UT**	**T**	**UT**	**T**
	N-term	N-term	R 2		R 2		K14		T3	T3
		R 2	K4	K4		K4	K18	K18	T6	T6
		K4	R 8	R 8		K9	K23	K23	S10	S10
	R 8	R 8	K9	K9	K14	K14	K27		T11	T11
	K9	K9	K14	K14	R 17	R 17			S22	S22
	K14	K14	R 17	R 17	K18	K18			S28	S28
	R 17	R 17	K18	K18	K23				S32	S32
	K18	K18	K23	K23		K37				
	K23	K23	R 26							
		R 26		K27						
	K27	K27								
	R 128									
	9	**11**	**9**	**8**	**5**	**6**	**4**	**2**	**7**	**7**
H3.3	**Acetylation**		**Monomethylation**		**Dimethylation**		**Trimethylation**		**Phosphorylation**	
	UT	**T**	**UT**	**T**	**UT**	**T**	**UT**	**T**	**UT**	**T**
	N-term	N-term	R 2		R 2	R 2	K4	K4	T3	T3
	R 2	R 2	K4	K4		K4	K9		T6	T6
	K4	K4	R 8	R 8	R 8	R 8	K14	K14	S10	S10
	R 8		K9	K9	K9	K9		K18	T11	T11
	K9	K9	K14	K14	K14	K14		K23	S22	S22
	K14	K14	R 17	R 17	R 17	R 17	K27	K27		S28
	R 17	R 17	K18	K18	K18	K18	K37		S32	S32
	K18	K18	K23		K23					T33
	K23	K23	R 26			K27				S57
		K27		K27		R 40			T58	
		K36	K35		R 41					
	K37		K36	K36	R 49					
	R 52		K37			R 63				
	K53		K56							
	K56	K56	R 63							
		K115								
		K122								
	R 128									
	14	**13**	**14**	**8**	**9**	**10**	**5**	**5**	**7**	**9**
H3 cen	**Acetylation**		**Monomethylation**		**Dimethylation**		**Trimethylation**		**Phosphorylation**	
	UT	**T**	**UT**	**T**	**UT**	**T**	**UT**	**T**	**UT**	**T**
		K40								T27
										T32
		1								**2**

	Acetylation		Monomethylation		Dimethylation		Trimethylation		Phosphorylation	
	UT	T	UT	T	UT	T	UT	T	UT	T
H4		R 3		R 3	K5	K5	K5		S1	S1
	K5	K5	K8	K8		K8	K12	K12		S47
	K8	K8		K16	K12					T80
	K12	K12	K12		K16	K16				T82
	K16	K16	R 50			K45				S89
	4	5	3	3	3	4	2	1	1	5