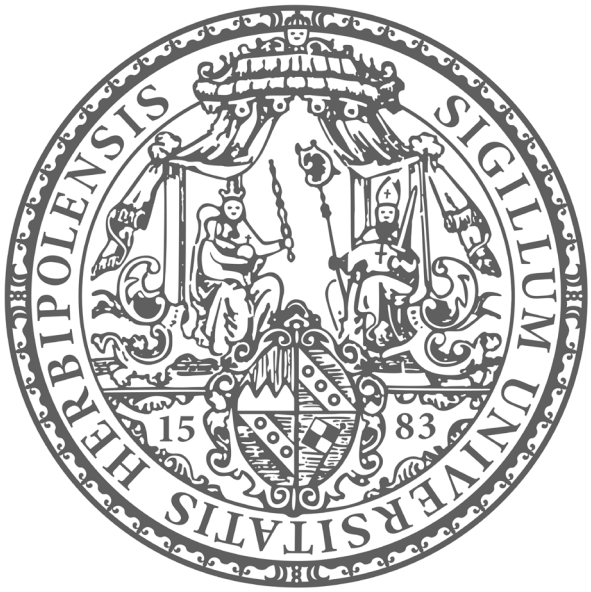
Molecular basis for product-specificity of DOT1 methyltransferases in *Trypanosoma brucei*



Dissertation zur Erlangung des
naturwissenschaftlichen Doktorgrades
der Julius-Maximilians-Universität Würzburg

vorgelegt von

Gülcin Dindar

aus München

Würzburg 2014

Eingereicht am: 30.06.2014

Mitglieder der Promotionskommission:

Vorsitzender: Prof. Dr. Ricardo Benavente
Erstgutachter: Prof. Dr. Christian Janzen
Zweitgutachter: PD Dr. Sandra Hake

Tag des Promotionskolloquiums: 29.8.2014

Doktorurkunde ausgehändigt am:

Eidesstattliche Erklärung

Hiermit erkläre ich ehrenwörtlich,

dass die vorliegende Dissertation von mir eigenständig angefertigt, d.h. insbesondere selbstständig und ohne unerlaubte Hilfe einer kommerziellen Promotionsberatung angefertigt und keine anderen als die von mir angegebenen Quellen oder Hilfsmittel benutzt wurden.

dass ich die Gelegenheit zum Promotionsvorhaben nicht kommerziell vermittelt bekommen habe und insbesondere keine Person oder Organisation eingeschaltet habe, die gegen Entgelt Betreuer bzw. Betreuerinnen für die Anfertigung der Dissertation sucht.

dass ich die Regeln der Universität Würzburg über gute wissenschaftliche Praxis eingehalten habe.

dass ich diese Dissertation weder in gleicher noch ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegt habe.

dass ich zuvor keinen anderen akademischen Doktorgrad erworben habe oder versucht habe zu erwerben.

Würzburg, 30.06.2014

……..

Gülcin Dindar

Parts of this thesis have been submitted to scientific journals for publication:

Dindar G, Anger AM, Mehlhorn C, Hake SB and Janzen CJ (2014). Structure-guided mutational analysis reveals the basis for product-specificity of DOT1 enzymes. Submitted to ***Nat. Communications***.

Pascoalino B*, Dindar G*, Vieira-da-Rocha JP, Machado CR, Janzen CJ and Schenkman S (2014). Characterization of two different Asf1 histone chaperones with distinct cellular localizations and functions in *Trypanosoma brucei*. ***Nucleic Acids Res.*** 42 (5), 2906-2918.
* These authors contributed equally to this work.

Parts of this thesis have been presented at international conferences:

2013 **In vitro reconstitution and modification of trypanosomal chromatin.** (Talk)
Gülcin Dindar (presenting author), Christine Mehlhorn, Ludmila Schneider, Andreas M. Anger, Sandra B. Hake, Christian J. Janzen

5th Kinetoplastid Molecular Cell Biology Meeting,
April 21-25, 2013, Woods Hole, USA.

2012 **An in vitro nucleosome reconstitution system to study changes in chromatin structure during differentiation of *Trypanosoma brucei*.** (Poster)
Gülcin Dindar (presenting author), Ferdinand Bucerius, Christine Mehlhorn, Andreas M. Anger, Ludmila Schneider, Sandra B. Hake and Christian J. Janzen

25th Annual Meeting of the German Society for Parasitology,
March 24-27, 2012, Heidelberg, Germany

Table of Contents

Acknowledgements

First of all I would like to thank Christian Janzen for giving me the opportunity to work in his lab on this extremely interesting project. I am grateful for his continuing support and advice. I appreciate the huge freedom you gave me to pursue my ideas and I really enjoyed working in your lab. My special thank goes to Sandra Hake for supporting me during the initial stages of my project and help in establishing the nucleosome reconstitution system. Thank you for many helpful comments on our manuscript and my thesis. I would like to thank Sergio Schenkman for our successful cooperation on the Asf1 project and Elisabeth Kremmer for providing antibodies.

Thanks to all members of the Janzen and Engstler labs for a great atmosphere and support during my time in Würzburg. Elisabeth Meyer-Natus, thank you for keeping the lab running, for sharing you experience and secrets regarding immunoflourescence microscopy and your skillful support with histone preparations. Lidia Schegelski, thank you so much for always being there for me. Thank you for sharing the good times and for cheering me up in hard times. You are a great friend and I am glad that I can always rely on your support and encouragement. George Githure, my office buddy during my time in Munich, thank you for sharing your knowledge regarding protein purifications and all the good discussions we had. Nicola Jones, thanks for all the fun we had in and outside the lab. I would also like to thank my talented and motivated bachelor student Ludmila Schneider for preparing and initial testing of DOT1 mutants.

I would like to thank my parents Aysel and Cemal Dindar for their support, encouragement and love throughout the years. The biggest thank goes to my beloved twin sister Elcin Dindar for her enthusiasm and her support in every moment. Thank you so much for always being there for me.

Andreas Anger, thank you for all your support, patience, continuing encouragement and never-ending enthusiasm. Thank you for all your love and your positive attitude in every moment. I can't wait for the next steps of our adventure, wherever they will lead us. I love you!

Zusammenfassung

Posttranslationale Histonmodifizierungen (PTMs), wie beispielsweise die Methylierung von Lysinseitenketten, beeinflussen maßgeblich die Struktur und Funktion von Chromatin. PTMs spielen eine wichtige Rolle in verschiedensten zellulären Prozessen, darunter DNA Replikation, Transkription oder Zelldifferenzierung. Darüber hinaus liegt ein verändertes PTM-Muster einer Vielzahl humaner Erkrankungen zugrunde, wie z.B. der akuten myeloischen Leukämie. DOT1-Enzyme sind hochkonservierte Histonmethyltransferasen, die für die Methylierung von Lysin 79 in Histon H3 (H3K79) verantwortlich sind. Im Gegensatz zu den meisten Eukaryoten, die lediglich ein einziges DOT1-Enzym besitzen, finden sich zwei homologe Proteine in afrikanischen Trypanosomen (DOT1A und DOT1B), die Lysin 76 in Histon H3 (H3K76) methylieren (H3K76 ist homolog zu H3K79 in anderen Organismen). DOT1A ist essentiell und katalysiert Mono- und Di-Methylierungen, wohin gegen DOT1B darüber hinaus eine Trimethylierung an H3K76 setzen kann. Derzeit fehlt jegliches mechanistische Verständnis darüber, wie beide Enzyme diese unterschiedliche Produktspezifität erreichen. Die vorliegende Dissertation macht sich den Umstand zunutze, dass Trypanosomen zwei DOT1-Methyltransferasen mit unterschiedlichen katalytischen Eigenschaften besitzen, um Einblicke in die molekulare Grundlage der unterschiedlichen Produktspezifität zu erlangen. Zunächst wurde ein Rekonstitutionssystem für Nukleosomen aus Trypanosomen etabliert, das es ermöglichte die Methyltransferase-Aktivitäten unter definierten in vitro Bedingungen zu analysieren. Homologiemodelle erlaubten die Identifikation von wichtigen Aminosäurepositionen innerhalb und außerhalb des katalytischen Zentrums der Enzyme, die einen Einfluss auf die Produktspezifität haben. Ein Austausch der Aminosäuren an diesen Positionen führte zu einer Umwandlung der Produktspezifität und offenbarte gleichzeitig DOT1A- und DOT1B-spezifische regulatorische Domänen, die an das katalytische Zentrum angrenzen. Diese Arbeit liefert erste Hinweise, dass wenige maßgebliche Aminosäuren in DOT1-Enzymen für den H3K76-Methylierungsgrad während der Katalyse entscheidend sind. Darüber hinaus haben die hier dargestellten Ergebnisse ebenfalls Konsequenzen für das funktionale Verständnis der homologen Enzyme in anderen Eukaryoten.

Abstract

Post-translational histone modifications (PTMs) such as methylation of lysine residues influence chromatin structure and function. PTMs are involved in different cellular processes such as DNA replication, transcription and cell differentiation. Deregulations of PTM patterns are responsible for a variety of human diseases including acute leukemia. DOT1 enzymes are highly conserved histone methyltransferases that are responsible for methylation of lysine 79 on histone H3 (H3K79). Most eukaryotes contain one single DOT1 enzyme, whereas African trypanosomes have two homologues, DOT1A and DOT1B, which methylate H3K76 (H3K76 is homologous to H3K79 in other organisms). DOT1A is essential and mediates mono- and di-methylations, whereas DOT1B additionally catalyzes tri-methylation of H3K76. However, a mechanistic understanding how these different enzymatic activities are achieved is lacking. This thesis exploits the fact that trypanosomes possess two DOT1 enzymes with different catalytic properties to understand the molecular basis for the differential product-specificity of DOT1 enzymes. A trypanosomal nucleosome reconstitution system was established to analyze methyltransferase activity under defined in vitro conditions. Homology modeling allowed the identification of critical residues within and outside the catalytic center that modulate product-specificity. Exchange of these residues transferred the product-specificity from one enzyme to the other and revealed regulatory domains adjacent to the catalytic center. This work provides the first evidence that few specific residues in DOT1 enzymes are crucial to catalyze methyl-state-specific reactions. These results have also consequences for the functional understanding of homologous enzymes in other eukaryotes.

Abbreviations

°C	Degree Celsius
A_{260}	Absorption at 260 nm wavelength
A_{280}	Absorption at 280 nm wavelength
aa	Amino acid
AdoHcy	S-adenosyl-L-homocysteine
ADP	Adenosine diphosphate
Asf1	Anti silencing factor 1
AT-rich	Adenine and thymine rich
bp	Base pair
Bre1	Brefeldin A sensitive 1
BSA	Bovine serum albumin
BSF	Blood stream form
BtnTg	Biotin tag
C-terminal	Carboxy-terminal
C. elegans	*Caenorhabditis elegans*
$CaCl_2$	Calcium chloride
CAF-1	Chromatin assembly factor 1
CENP-A	Centromere protein A
D. melanogaster	*Drosophila melanogaster*
DIM-5	Defective in DNA methylation 5
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DNase I	Desoxyribonuclease I
dNTPs	Deoxyribonucleotides
DOT1	Disruptor of telomeric silencing 1
DOT1L	DOT1-like
DOTcom	DOT complex
DTT	Dithiothreitol
E	Enzyme
E-value	Expect value
E. coli	*Escherichia coli*
E(Z)	Enhancer of zeste
EDTA	Ethylenediaminetetraacetic acid
FBS	Fetal bovine serum
for	Forward
g	Gravitational (g)-force
G2-phase	Gap 2 phase
h	Hour
H#	Histone #
H#V	Histone # variant

HAT	Histone acetyltransferase
HCl	Hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMM	Hidden Markov model
HWB	Histone wash buffer
HWB/T	Histone wash buffer/Tween
ID	Identifier
IgG	Immunglobulin G
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KCl	Potassium chloride
kDa	Kilodalton
KH_2PO_4	Monopotassium phosphate
KMTase	Lysine methyltransferase
KOH	Potassium hydroxide
L	Loop
LB	Luria Bertani
M	Molar
mA	Milliampere
mAb	Monoclonal antibody
mAU	Milli-absorbance units
MBP	Maltose binding protein
me1, me2, me3	Mono-, di-, trimethylation
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
min	Minute
MLA	methyl-lysine analog
mM	Millimolar
MWCO	Molecular weight cut-off
N-terminal	Amino-terminal
N_2	Nitrogen
Na(OAc)	Sodium acetate
Na_2HPO_4	Disodium phosphate
NaCl	Sodium chloride
NP-40	Nonidet P-40
O/N	Over night
OD_{600}	Optical density at 600 nm
pAb	Polyclonal antibody
PBS	Phosphate buffered saline
PBST	PBS-tween
PCF	Procyclic form
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
pH	*Potentia hydrogenii*
PHD	Plant homeo domain

PMSF	Phenylmethanesulfonylfluoride
PTM	Posttranslational modification
PVDF	Polyvinylidene fluoride
Rad53	Radiation-sensitive 53
RCF	Relative centrifugal force
rev	Reverse
RMSD	Root mean square deviation
RNA	Ribonucleic acid
rpm	Revolutions per minute
RT	Room temperature
Rtt109	Regulator of Ty1 transposition 109
S-phase	Synthesis phase
S. cerevisiae	*Saccharomyces cerevisiae*
S. pombe	*Schizosaccharomyces pombe*
SAM	S-adenosyl-L-methionine
SDS	Sodium dodecyl sulfate
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sec	Second
SET	Suppressor of variegation, enhancer of zeste and trithorax
SILAC	Stable isotope labeling by amino acids in cell culture
S_N2	Bimolecular nucleophilic substitution
SNAP	SILAC nucleosome affinity purification
SOC	Super optimal broth with catabolite repression
spp.	*Species pluralis*
Su(var)	Suppressor of variegation
t	Time
T. brucei	*Trypanosoma brucei*
TAE	Tris base, acetate, EDTA
TBE	Tris base, boric acid, EDTA
TCA	Trichloroacetic acid
TEMED	Tetramethylethylenediamine
TLK	Tousled-like kinase
Tris	Tris(hydroxymethyl)aminomethane
Trx	Trithorax
U	Unit
UV	Ultraviolet
V	Volt
v/v	Volume/volume percentage
VSG	Variant surface glycoprotein
w/v	Mass/volume percentage
WT	Wildtype
µM	Micromolar

1 Introduction

1.1 *Trypanosoma brucei*

Trypanosomes are unicellular parasitic eukaryotes that cause devastating human diseases, such as sleeping sickness and Chagas disease (also known as African and American trypanosomiasis, respectively). Parasites of the *Trypanosoma brucei* species are transmitted between different mammalian hosts by the tsetse fly (*Glossina spp.*) and persists in the blood of the infected mammal (Vickerman, 1985). To escape the host's immune response, *T. brucei* has to undergo multiple morphological changes. Inside the mammalian host, the bloodstream form (BSF) of *T. brucei* evades elimination by periodically switching expression of a variant surface glycoprotein (VSG) gene in a process called antigenic variation (Horn et al., 2010). Whereas, inside the tsetse midgut, the parasite differentiates from the bloodstream trypanosome into the procyclic form (PCF), replacing all VSG on their surface by the protein procyclin (Roditi et al., 1989). About 15 % of the trypanosomal genes belong to the VSG family (Barry et al., 2005). However, only one VSG gene is expressed at any time from BSF VSG expression sites that comprise large polycistronic units, transcribed by RNA polymerase I and are located in sub-telomeric regions of the genome (Günzl et al., 2003; Berriman et al., 2005). Understanding regulation of VSG transcription is a major area of trypanosome research, since this unique immune evasion system promises to be a prime target in the battle against trypanosomiasis. In order to survive in two different hosts, many cellular functions of the parasite, including metabolism, surface architecture, cell cycle control and antigenic variation require continuous adaptation and coordination. The adaption processes involves changes in chromatin structure that ultimately lead to altered protein expression patterns. Little is known about chromatin structure or epigenetic regulation by histone modification in *T. brucei* and up to now few functional studies have addressed chromatin remodeling (Hughes et al., 2007) or histone-modifying enzymes (Janzen et al., 2006b; Siegel et al., 2008) (reviewed in (Figueiredo et al., 2009)). Likewise, only several approaches have been made to understand the organization (Ersfeld et al., 1999; Hecker et al., 1994) and the function of chromatin in trypanosomes (Povelones et al., 2012). Importantly, the trypanosome genome sequence is available (Berriman et al., 2005), which gives the opportunity to search for homologues of chromatin-remodeling or histone-modifying enzymes. The access to a fully sequenced genome and many available molecular biological tools make this parasite an excellent system for reverse genetic experiments using gene knockout or RNA interference, as well as forward genetic methods to unravel essential and non-essential gene products by analysing the parasites phenotype (Alsford et al., 2012; Mazet et al., 2013). Gene knockouts can be introduced via homologous recombination that allows targeting of a specific gene and its replacement with a selective marker (Alsford et al., 2009; Oberholzer et al., 2006; Shen et al., 2001). Moreover, methods to monitor global gene expression and proteomic analysis have been successfully applied to understand the function of trypanosome gene products (Butter et al., 2013).

1.2 Chromatin structure

1.2.1 Core histones and the nucleosome

Every single organism is defined by the unique genetic information content that is stored in DNA molecules. Within the eukaryotic cell nucleus, DNA organization is achieved through assembly into DNA-protein complexes, called chromatin. Nucleosomes, the repetitive building block units of chromatin, are composed of a basic histone octamer complex containing two copies of each core histone (H3, H4, H2A, H2B) and 147 base pairs (bp) of DNA that are wrapped around it in 1.65 turns (Figure 1) (Andrews and Luger 2011; Kornberg and Lorch, 1999; Luger and Hansen, 2005). Nucleosomes are regularly spaced by linker DNA, resulting in an array that is able to form higher order structures, stabilized by the associated linker histone H1. Compaction of chromatin hinders the accessibility of DNA for proteins such as transcription factors to bind to specific nucleotide sequences to fulfill their function. Therefore relaxed and condensed chromatin states must be dynamic to allow regulated access to DNA and the transition between these states is crucial for many cellular processes like DNA replication, transcription and repair.

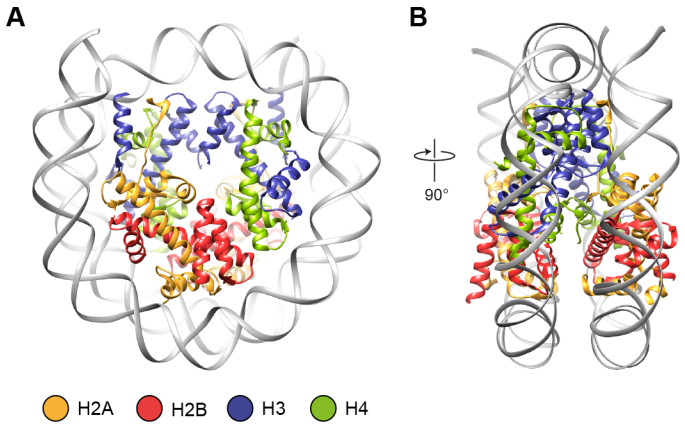


Figure 1. Overview of the nucleosome core structure
(**A**) Crystal structure of the nucleosome core particle (Luger et al., 1997). Individual histone proteins are colored as indicated and DNA is shown in grey. (**B**) Side-view of the particle related to (A) by a 90° rotation.

All four core histone proteins share a structurally conserved histone fold, which consist of about 70 amino acids (aa) (Figure 2A). The histone fold composed of three α-helices (α1-α3) that are connected by two short loops (L1 and L2) (Luger et al., 1997) (Figure 2A). Histone H3/H4 or H2A/H2B heterodimers are formed by an antiparallel arrangement of the two long α2 helices (Figure 2B). This arrangement creates three DNA interaction areas per dimer, two L1/L2 sites and one α1/α1 site (Luger et al., 1997). Two H3/H4 heterodimers can associate to form tetramers via a four-helix bundle between the H3 histones (Figure 2C). Similarly, a heterologous four-helix bundle structure is formed between H4 and H2B that tether one H2A/H2B dimer to one half of the histone H3/H4 tetramer (Luger et al., 1997) (Figure 2C). Outside the histone fold region, helical extensions are present either N- or C-terminally (αN or αC) that are not conserved and vary in length. These extensions are responsible for protein-protein interactions within the nucleosome and contribute to DNA binding and positioning on the surface of the particle (Luger et al., 1997). Moreover, flexible histone tails which protrude away from the ordered nucleosome structure have been shown to be important for higher order chromatin assembly by affecting inter-nucleosomal interactions (Hansen, 2002) or interaction with associated factors.

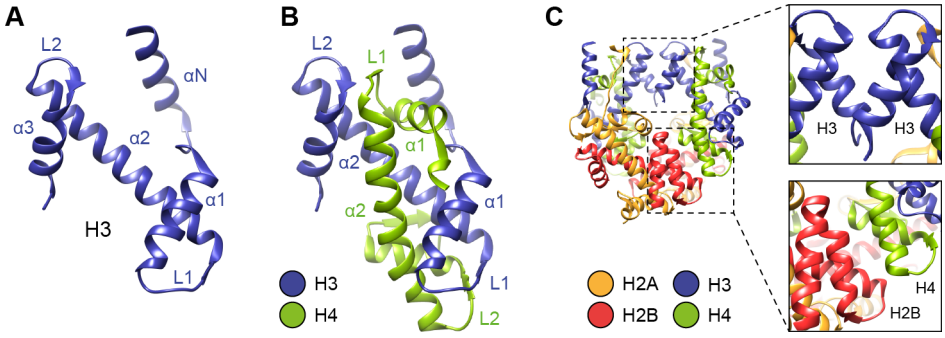


Figure 2. Protein interactions within the nucleosome
(**A**) Histone fold of H3 (Luger et al., 1997), α-helices (αN-α3) and loops (L1, L2) are indicated (**B**) The histone H3/H4 dimer (Luger et al., 1997). Secondary structure elements are colored as in (A). (**C**) Overview of histone octamer (Luger et al., 1997) showing four-helix bundle contacts between histones H3/H3 (α3-α3) (upper panel) and histones H2B/H4 (lower panel).

1.2.2 Histone variants

Gene expression, DNA replication, recombination and repair are regulated by incorporation of histone variants into chromatin as well as DNA and histone modifying enzymes. Many histone variants have been identified to be unique in vertebrates and others to be highly conserved among all eukaryotes (Malik and Henikoff, 2003; Weber and Henikoff, 2014). *T. brucei* contains four known histone variants (H2AZ, H2BV, H3V and H4V) (Lowell and Cross, 2004; Lowell et al., 2005) with individual functions (for a phylogenetic analysis of trypanosomal core histones and variants refer to (Alsford and Horn, 2004)). The variant H2AZ exists from *Tetrahymena* to humans, shares about 90 % sequence identity among different organisms (Dryhurst et al., 2004) and is linked to transcriptional activation (Adam et al., 2001; Santisteban et al., 2000), gene silencing (Dhillon and Kamakaka, 2000) and prevention of heterochromatin formation (Meneghini et al., 2003). Moreover, H2AZ is also involved in chromosome segregation (Rangasamy et al., 2004) by influencing the formation of pericentric heterochromatin (Rangasamy et al., 2003). Nucleosomes that contain the H2AZ variant have been shown to be less stable than nucleosomes with the canonical H2A in trypanosomes (Siegel et al., 2009), possibly by weakening the interface between H2AZ/H2BV dimers and the H3/H4 tetramer within the nucleosome (Suto et al., 2000). It has previously been shown that H2AZ promotes the folding of compact chromatin structures but prevents the ability of these arrays to further oligomerize in vitro (Fan et al., 2004). This indicates that H2AZ incorporation is used to avoid chromatin condensation beyond a certain point and to prime it for transcriptional activation (Siegel et al., 2009). The H2B variant exclusively dimerizes with H2AZ and appears to be a trypanosomatid-specific histone variant that is essential for viability (Lowell et al., 2005). Eukaryotes typically contain a histone H3 variant, called centromere protein A (CENP-A), which creates a specialized chromatin environment at centromeric regions and is essential for assembly of the kinetochore complex during mitosis (Mellone and Allshire, 2003). Notably, no CENP-A homologues can be found in the *T. brucei* genome (Berriman et al., 2005) and African trypanosomes have recently been shown to assemble a completely different set of kinetochore proteins on centromeres in comparison to other eukaryotes (Akiyoshi and Gull, 2014). Like CENP-A, the histone variants H2A.X and H3.3 that are involved in DNA repair and transcription, respectively appear to be absent in trypanosomes. The trypanosomal histone variant H4V dimerizes with H3V to form specific nucleosome complexes at

transcription termination sites in, whereas H2AZ and H2BV appear to be localized at transcription start sites (Siegel et al., 2009). However, the functional consequences of these histone variant localizations in trypansosomal chromatin are presently unknown.

1.3 Chromatin assembly and dynamics

The contacts between DNA and histones proteins within the nucleosome are mediated by electrostatic interactions of positively charged side chains (arginine, lysine and histidine) with the negatively charged phosphate backbone of the minor groove every 10.4 bp along the DNA double helix, leading to 14 relatively weak DNA-histone interactions per nucleosome. These individually weak contacts act cooperatively and collectively result in a stable position and limited mobility of the nucleosome on DNA. Notably, the histone octamer interacts with the sugar-phosphate backbone of the DNA in a sequence independent manner and allows the nucleosome to accommodate a broad spectrum of DNA sequences (Luger et al., 1997; Widom, 2001). Nevertheless, histones have been shown to exhibit a preference for certain DNA stretches (e.g. for adenine and thymine (AT)-rich sequences facing the histone octamer), which may provide positional cues for nucleosomes in certain contexts (e.g. in promoter regions) (Struhl and Segal, 2013). Eukaryotes contain a linker histone H1, which among other factors is involved in establishing higher order packaging of chromosomes (Harshman et al., 2013; Song et al., 2014). Assembly and disassembly of nucleosomes from newly synthesized and recycled histones as well as their repositioning along the DNA and exchange of individual histones by variants contribute to chromatin dynamics and are essential for fundamental processes in eukaryotic cells such as DNA replication, transcription and repair. The dynamic behavior of chromatin is a result of numerous associated factors like remodeling complexes and histone chaperones (Burgess and Zhang, 2010; 2013; Clapier and Cairns 2009; Liu and Churchill, 2012). Following DNA replication, newly synthesized histones have to be deposited onto the new DNA strands to mediate nucleosome formation and assembly of chromatin in a process termed replication-coupled nucleosome assembly. Here, nucleosomes are assembled by deposition of parental H3/H4 as well as newly synthesized H3/H4 on the nascent DNA to form $(H3\text{-}H4)_2$ tetramers, followed by rapid incorporation of H2A/H2B histone dimers. Both of these steps during DNA replication require specific histone chaperones (Burgess and Zhang, 2010; 2013). Chromatin assembly factor 1 (CAF-1) is a key player in the RC nucleosome assembly process and was first identified in human cells as a factor that promotes nucleosome assembly during DNA replication (Stillman, 1986). CAF-1 binds to H3/H4, assisted by the anti-silencing function 1 (Asf1) chaperone to deposit the heterodimer onto replicated DNA. Asf1 is a highly conserved histone chaperone that forms a complex with histone H3 and H4 and promotes dimer deposition and removal from chromatin (Donham et al., 2011; English et al., 2005). Moreover, Asf1 is also important for cell cycle regulation by interacting with the checkpoint kinase Rad53 in yeast and with Tousled-like kinase (TLK) during the DNA damage response (Silljé and Nigg, 2001). Asf1 was initially found to be a suppressor of gene silencing when overexpressed in budding yeast (Le et al., 1997; Singer et al., 1998). Later it was shown that Asf1 functions in replication-coupled as well as replication-independent nucleosome assembly (Tyler et al., 1999). Crystal structures of yeast and human Asf1 in a complex with the H3/H4 heterodimer revealed that Asf1 binds H3/H4 through the H3 α3 helix, that is involved in

$(H3\text{-}H4)_2$ tetramer formation. In addition, Asf1 also forms a short intermolecular β-sheet with H4 (English et al., 2006; Natsume et al., 2007). These interactions occur via the highly conserved N-terminal region of Asf1, whereas the C-terminus is more variable and has regulatory functions (Daganzo et al., 2003). In contrast to yeast and *Drosophila*, where only a single Asf1 protein is found, some eukaryotes including human and kinetoplastids possess two different Asf1 chaperones, Asf1A and Asf1B. Very little is known about the functions of the two Asf1 chaperones in trypanosomes. Asf1A is localized in the cytosol and translocates to the nucleus during early S phase, whereas Asf1B remains located in the nucleus throughout the cell cycle and might facilitate histone deposition during replication (Pascoalino et al., 2014). This spatial distribution suggests that the two chaperones might have distinct functions and share the labor during escort of newly synthesized histones from ribosomes to the nucleosome assembly sites in the nucleus, but details remain to be elucidated. Likewise it is unclear to what extend Asf1A and Asf1B influence post-translational modifications of newly synthesized histones prior deposition in chromatin. For instance, Asf1 has been shown to stimulate activity of the H3 histone acetyltransferase Rtt109 in a Rtt109-Asf1-H3-H4 complex in yeast (Tsubota et al., 2007).

1.4 Posttranslational histone modifications

1.4.1 Histone tail modifications

The flexible N-terminal histone tails protrude from the histone octamer and do not form an integral part of the nucleosome. They vary in length (ranging from 15 to 40 aa) among the four core histones and are the main targets for posttranslational modifications. In addition to the N-terminal tails found in all core histones, H2A contains a flexible C-terminal extension. Although, histones in general are highly conserved among different eukaryotes, sequence alignments reveal that histone tails are much less conserved than the remaining histone core domains (Luger, 2001). Histone tails neither interact with DNA, nor stabilize the nucleosome core particle (Luger, 2001), but in contrary possesses different roles depending on their state of modification and the presence of associated factors (Lennartsson and Ekwall, 2009; Strahl and Allis, 2000). It is well established that the degree of DNA packaging with proteins, depends on histone posttranslational modifications (PTMs) that contribute to the control of gene expression via their influence on chromatin dynamics (Zentner and Henikoff, 2013). The histone PTM repertoire includes acetylation, methylation, phosphorylation, ADP- ribosylation, sumoylation and ubiquitylation that might function individually or in a multitude of potential combinations, termed the histone-code (Strahl and Allis, 2000). Additional complexity results from alternative methylation states at histone lysine or arginine residues that can occur in various forms: mono-, di- or trimethylation for lysine, as well as mono- and dimethylation (symmetric or asymmetric) for arginines (Zhang and Reinberg, 2001). Histone modifications correlating with transcriptional activation include acetylation, methylation, phosphorylation and ubiquitylation, whereas methylation and ubiquitylation are also linked to repression (Kouzarides, 2007; Peterson and Laniel, 2004). This indicates that many, if not all histone modifications can lead to different biological outcomes, dependent on the specific context (Kouzarides, 2007). Different enzymes that are directly responsible for histone modifications have been identified for acetylation (Sterner and Berger, 2000), methylation (Zhang and Reinberg, 2001), phosphorylation (Nowak and Corces, 2004), ubiquitylation (Shilatifard, 2006), ADP-ribosylation (Hassa et al., 2006) and sumoylation (Nathan

et al., 2006). The enzymes that covalently add or remove histone modifications are commonly referred to as “writer” or “eraser”, respectively. Modified residues and their sequence context within the histone tails often serve as docking platforms for chromatin regulators that contain one or several so-called “reader” domains. For instance, recognition of methylation marks involves members of the Royal family of domains (e.g. chromo and tudor domains) as well as plant homeo domain (PHD) fingers (reviewed in (Patel and Wang, 2013; Ruthenburg et al., 2007; Taverna et al., 2007)).

1.4.2 Histone modifications in *T. brucei*

Knowledge about chromatin structure, histone modification and epigenetic regulation in kinetoplastids is still very limited. The study of histone PTM patterns and their associated molecular functions in trypanosomes will help to understand important biological processes like antigenic variation that are known to be influenced by changes in chromatin structure (Horn and McCulloch, 2010). Although core histones belong to the evolutionarily most conserved proteins known, substantial differences are evident when comparing histones of trypanosomes to those of other eukaryotic species (Thatcher and Gorovsky, 1994). These differences are most dramatic within the flexible histone tails that are the prime targets for PTMs. Previous studies have identified several histone modifications in trypanosomes that are common and different compared to higher eukaryotes (Horn, 2007; Janzen et al., 2006a; Mandava et al., 2007). Methylation of H3K76 in trypanosomes is homologous to the equivalent modification of H3K79 in other eukaryotes such as human and yeast (Nguyen and Zhang, 2011; van Leeuwen et al., 2002). Other evolutionary conserved modifications are acetylation of H4K14 (in *T. brucei*) / H4K16 (in yeast) and methylation of H4K18 (in *T. brucei*) / H4K20 (in yeast) (Mandava et al., 2007). Furthermore, *T. brucei* shows hyper-acetylation of the H2A C-terminus (Janzen et al., 2006a), which is rather found at the N-terminus of H2A in humans (Morales and Richard-Foy, 2000). Histone hyper-acetylation in general is thought to destabilize chromatin and likewise it might be used to create long euchromatin regions upon environmental changes in trypanosomes (Janzen et al., 2006a). A large number of studies showed that histone tails are piled with PTMs (reviewed in (Lennartsson and Ekwall, 2009; Strahl and Allis, 2000)). Interestingly trypanosomes appear to contain only a limited number of histone tail modifications. The lack of PTMs might be related to the limited repertoire of histone modifying enzymes that could be detected in the trypanosome genome project (Berriman et al., 2005). This includes 5 putative histone acetyltransferases (HATs), 20-27 putative SET (suppressor of variegation (Su(var)3-9), enhancer of zeste [E(Z)] and Trithorax (Trx)) domain containing histone methyltransferases, 2 non-SET enzymes of the DOT1 family and 5 possible histone arginine methyltransferases (Figueiredo et al., 2009). In contrast, about 90 and 37 putative histone modifying enzymes are encoded in the human and *Drosophila* genome, respectively (Cheng et al., 2005). Among those, the conserved H3K79 (H3K76 in *T. brucei*) methylating DOT1 enzymes are the main focus of this thesis and will be described in more detail below.

1.5 Histone lysine methylation

Within histones, lysine residues are subject to a variety of modifications including methylation acetylation, ubiquitylation and sumoylation. Lysine methylation in general is more complex than other modifications since it can occur in three states, namely mono- (me1), di- (me2) or trimethylation (me3). Depending on the degree of methylation, these modifications have been correlated with cellular processes including heterochromatin formation, X-chromosome inactivation, transcriptional regulation (Martin and Zhang, 2005) and a number of human cancers (Greer and Shi, 2012; Schneider et al., 2002). Histone lysine methylation is catalyzed by a group of lysine methyltransferases (KMTases) that can be divided into two classes based on differences in their catalytic domain (Zhang et al., 2011). Members of the first KMTase class contain the evolutionary conserved SET domain (Dillon et al., 2005) and include the KMT1-3 and KMT5-8 protein families, whereas the second class is formed by a single conserved non-SET protein named DOT1 (KMT4) (Feng et al., 2002; Lacoste et al., 2002; Ng, 2002; Singer et al., 1998; van Leeuwen et al., 2002). All histone KMTases utilize the cofactor S-adenosyl-L-methionine (SAM) as methyl group donor in the transfer reaction. The enzymes catalyze me1, me2 and me3 by mediating deprotonation of the target lysine ε-amine to produce a lone electron pair that can initiate the subsequent nucleophilic attack on the methylsulfonyl group of SAM. Products of this bimolecular nucleophilic substitution (S_N2) reaction are monomethyl-lysine and S-adenosyl-L-homocysteine (AdoHcy) (Figure 3) (Smith and Denu, 2009). The vast majority of SET-domain containing histone KMTases identified so far catalyze modification within the flexible N-terminal histone tails, whereas DOT1 enzymes target a single lysine residue (K76 in *T. brucei* and K79 in other eukaryotes) (Janzen et al., 2006b; van Leeuwen et al., 2002) of histone H3 that is located in the solvent exposed loop within the nucleosomal core (Luger et al., 1997) (Figure 2A).

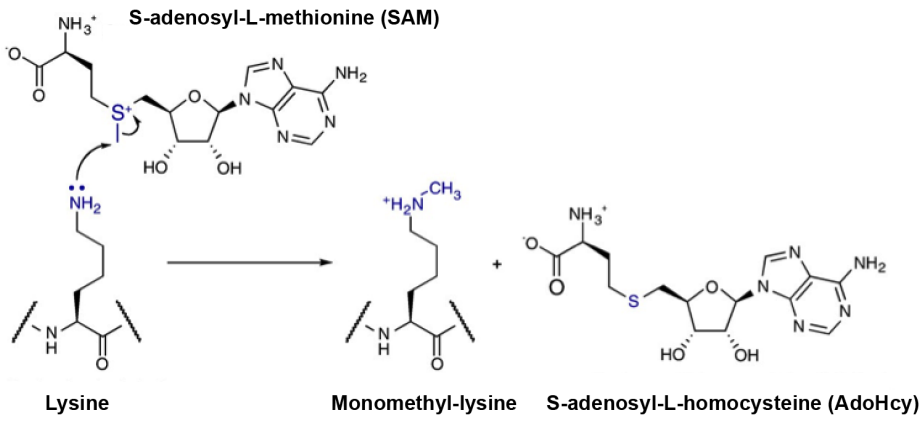


Figure 3. Chemical mechanism of SAM-dependent histone lysine methyltransferases
Methyl transfer by SAM-dependent KMTases proceeds via a nucleophilic attack on the methylsulfonyl group of SAM by a deprotonated target lysine to yield the products Monomethyl-lysine and AdoHcy. Figure modified from (Smith and Denu, 2009).

1.6 SET-domain histone lysine methyltransferases

Histone lysine methylation, with the exception of H3K79 (H3K76 in *T. brucei*), is catalyzed exclusively by the conserved SET domain family that was originally identified in *Drosophila melanogaster* (Jones and Gelbart, 1993; Stassen et al., 1995). SET domain containing KMTases have been predicted in a few bacterial and in a large number of eukaryotic proteomes with more than 60 members in humans (Binda, 2013). KMTases are subdivided into eight families (KMT1-8) and all members contain a SET domain with the KMT4 (DOT1) family being the sole exception (reviewed in (Zhang et al., 2011)). To date, several structures of SET-domain containing methyltransferases have been solved in various states that also include enzyme-substrate complexes (Campagna-Slater et al., 2011; Kwon et al., 2003; Manzur et al., 2003; Min et al., 2002; Nguyen et al., 2013; Schapira, 2011; Wilson et al., 2002; Wu et al., 2010; 2013; Xu et al., 2011; Zhang et al., 2003; 2002). The fact that many SET-domain enzymes methylate lysine residues within the flexible histone tails rather than on the nucleosome core surface, greatly facilitated the generation of enzyme-substrate complexes for structural studies by the use of short peptides that are sufficient to be recognized by the enzymes. A hallmark feature of SET-domain KMTases is that both the SAM-binding and the lysine-binding pockets are located in distinct clefts positioned on opposite surfaces of the enzyme. Both clefts meet at the center of the enzyme, where the methylation reaction takes place (Dillon et al., 2005). Methyl transfer from SAM to the substrate lysine ε-amine proceeds via a nucleophilic attack and only deprotonated lysine residues possess a free electron pair for this reaction (Smith and Denu, 2009). It has been proposed that lysine deprotonation is mediated by a conserved tyrosine residue at the active site of the enzymes (Jacobs et al., 2002; Kwon et al., 2003; Trievel et al., 2002; Xiao et al., 2003; Zhang et al., 2002). This speculation that was later supported by molecular dynamic simulations with the SET7/9 enzyme, indicating that deprotonation of the tyrosine residue by bulk water molecules converts this residue into a general base that becomes available to deprotonate the substrate lysine ε-amine (Guo and Guo, 2007). However, the exact deprotonation mechanism is not clear and requires further investigation. First insights into the structural determinants of product-specificity came from a comparison of the SET domain containing KMTases DIM-5 and SET7/9, which catalyze formation of distinct products (me1/2/3 for DIM-5 and me1 for SET7/9). A single position occupied by either a phenylalanine (in DIM-5) or a tyrosine (in SET7/9) determines product-specificity in these cases and swapping the residues changes product-specificity of the mutant enzymes (Zhang et al., 2003).

1.7 DOT1 methyltransferases

1.7.1 Overview

The DOT1 (disruptor of telomeric silencing; also called KMT4) enzyme is an evolutionarily conserved methyltransferase and was first identified in *Saccharomyces cerevisiae* (Dot1p) in a genetic screen for genes whose overexpression result in telomeric silencing defects (Singer et al., 1998). The DOT1 enzyme is responsible for methylation of histone H3K79 (Feng et al., 2002; Lacoste et al., 2002; van Leeuwen et al., 2002) that is located within the L1 loop of the globular domain of histone H3 and thus positioned on the surface of the assembled

nucleosome (Luger et al., 1997) (Figure 2A). In mammals the conserved homologous protein is called DOT1-like (DOT1L) (Feng et al., 2002). Great progress has been made in characterizing the enzymatic activities and the role of Dot1p/DOT1L mediated H3K79 methylation (Feng et al., 2002; Lacoste et al., 2002; Ng, 2002; van Leeuwen et al., 2002) that occurs in a nonprocessive manner in *S. cerevisiae* (Frederiks et al., 2008). African trypanosomes contain two DOT1 homologues, DOT1A and DOT1B, with different enzymatic activities. DOT1A mediates mono- and di-methylation of H3K76 (the homolog of H3K79 in other organisms), whereas DOT1B additionally catalyzes H3K76 tri-methylation (Frederiks et al., 2010; Janzen et al., 2006b). Moreover, five splice variants of DOT1 (DOT1a-Dot1e) have been predicted in mice (Zhang et al., 2004). However, if the enzymatic activities of these variants on H3K79 differ from one another is unknown at present. Notably, DOT1 enzymes cannot methylate free histone H3 and are strictly dependent on an intact nucleosomal core surface, indicating that DOT1 only methylates H3K79 when H3 is incorporated into chromatin (Fingerman et al., 2007; Janzen et al., 2006b; Lacoste et al., 2002; Sawada et al., 2004; van Leeuwen et al., 2002). Previous works have shown that DOT1 and its homologs appear to be the sole enzymes responsible for H3K79 methylation as deletion of DOT1 in yeast, flies and mice results in complete loss of H3K79 methylation (Jones et al., 2008; Shanower et al., 2005; van Leeuwen et al., 2002). H3K79 di-methylation (H3K79me2) appears to prevent heterochromatin formation since overexpression or deletion of DOT1 as well as mutations of H3K79 reduced the interaction with SIR (silent information regulator) proteins in budding yeast in vivo and therefore limits the ability to silence genes (Ng, 2002; van Leeuwen et al., 2002). Moreover, methylation has been shown to prevent Sir3 binding to the nucleosome in vitro (Onishi et al., 2007) and structural details of the Sir3 interaction with the nucleosome indicate how H3K79 methylation can result in a decreased Sir3 affinity for the nucleosome (Armache et al., 2011; Wang et al., 2013). As SIR proteins and DOT1 enzymes very likely compete for the same binding site on the nucleosome, the amount of SIR binding and H3K79 methylation levels have to be balanced to regulate heterochromatin formation e.g. at telomeric regions (Singer et al., 1998; van Leeuwen et al., 2002). In addition to regulation of telomere silencing, H3K79 methylation appears to influence transcription, since DOT1 has been found in transcription elongation complexes (reviewed in (Nguyen and Zhang, 2011). In yeast efficient H3K79 methylation requires ubiquitylation of histone H2B K123 (Briggs et al., 2002; Frederiks et al., 2008; 2010; Nakanishi et al., 2009; Ng, 2002) that is located on the same nucleosome surface approximately 30 Å away from H3K79 (Luger et al., 1997). It has been speculated that H2B ubiquitylation could serve as a spacer between adjacent nucleosomes to prevents compaction of chromatin (Fierz et al., 2011; Sun and Allis, 2002). This in turn could increase accessibility of the DOT1 binding site, although it has been shown that the absence of H2B K123 ubiquitylation does not influence DOT1 monomethylation of H3K79, but rather the progression to higher order methylation states (Frederiks et al., 2008; 2010; Shahbazian et al., 2005). Interestingly, while yeast Dot1p loses methylation activity in the absence of H2B K123 mono-ubiquitylation, trypanosomal DOT1 proteins are barely affected (Frederiks et al., 2010).

1.7.2 The conserved DOT1 core

Non-SET domain methyltransferases like DOT1 (KMT4) are characterized by a series of conserved motifs (I, I' and II) (Dlakić, 2001; Malone et al., 1995; Schubert et al., 2003). Residues from all motifs are involved in cofactor binding and/or contribute to formation of the target lysine-binding channel of DOT1 proteins (Min et al., 2003; Sawada et al., 2004). Interestingly, the human DOT1L

crystal structure revealed separate entrances for the SAM and lysine-binding pockets, located on opposite sides of the enzyme (Min et al., 2003). Since cofactor binding appears to be independent from substrate binding, it is possible that DOT1 enzymes are able to methylate H3K79 in a processive or in a non-processive (distributive) manner. DOT1 enzymes differ dramatically in size between different eukaryotes, ranging from 582 aa in yeast, over 1537 aa in humans to 2137 aa in the fruit fly (Figure 4) (Feng et al., 2002). In yeast the conserved KMTase domain is located at the C-terminus, whereas in human, *Caenorhabditis elegans*, *D. melanogaster* and *Anopheles gambiae* DOT1 homologs it is positioned at the N-terminal part of the protein (Figure 4). Length variations result from species-specific N- and C-terminal extensions outside the conserved histone KMTase domain. These additional overhangs have been shown to be responsible for nucleosomal targeting of the DOT1 enzyme in yeast and human via positively charged residues (Min et al., 2003; Sawada et al., 2004). *T. brucei* DOT1A (295 aa) and DOT1B (275 aa) appear to lack these extensions and are only comprised of the KMTase core (Figure 4).

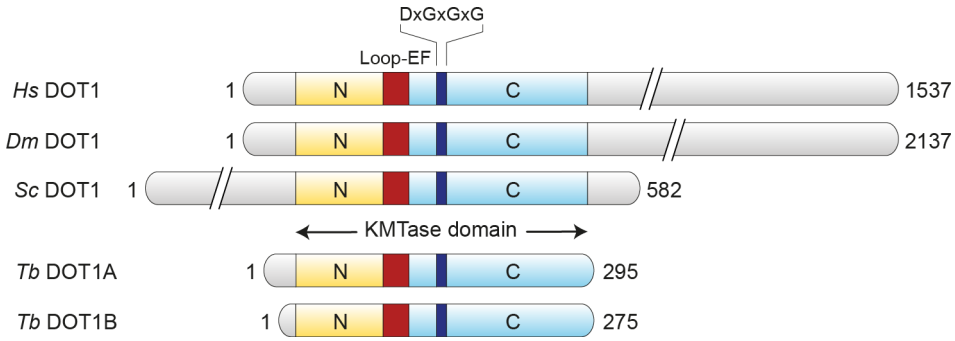


Figure 4. Schematic alignment of DOT1 enzymes from different species
The conserved KMTase core can be subdivided into an N- and C-terminal part (orange and light blue, respectively) connected by loop-EF (red). The DxGxGxG signature motif (dark blue) is essential for enzymatic activity.

The conserved KMTase domain of DOT1 proteins can be subdivided into a N- and C-terminal region. The N-terminal part consists of five α-helices (αA to αE) and two pairs of short β hairpins, while the C-terminal domain has an open α/β structure and is linked to the N-terminal domain via the loop-EF (Figure 5A) connecting helices αE and αF (Figure 5A). The domain is organized around a central β-sheet consisting of seven strands (β5 to β11), which is flanked by five α-helices (αF to αJ) (Figure 5B). Helices αF-αH and αI-αJ are packed against the central β-sheet from opposite sides (Figure 5B). The β-strands β5-β9 and β10 are running parallel to each other, whereas β11 has an antiparallel orientation with respect to the remaining strands. The N-terminal region makes extensive contacts with helices αF-αH of the C-terminal part of the KMTase domain with a major part of the interface being formed by helix αH and helices αD and αE (Figure 5A). The SAM cofactor is bound in a pocket of the enzyme that is build up by loop-EF and the C-terminal parts of β5 and β8. A part of the SAM-binding pocket is formed by the highly conserved DxGxGxG signature motif (motif I) (Min et al., 2003; Sawada et al., 2004) that has been shown to be essential for the enzyme activity (van Leeuwen et al., 2002). The methylsulfonyl group of SAM is aligned with an opening in the structure that has been proposed to form the substrate lysine-binding channel (Min et al., 2003) that creates the chemical environment required for methyl group transfer from SAM onto H3K79.

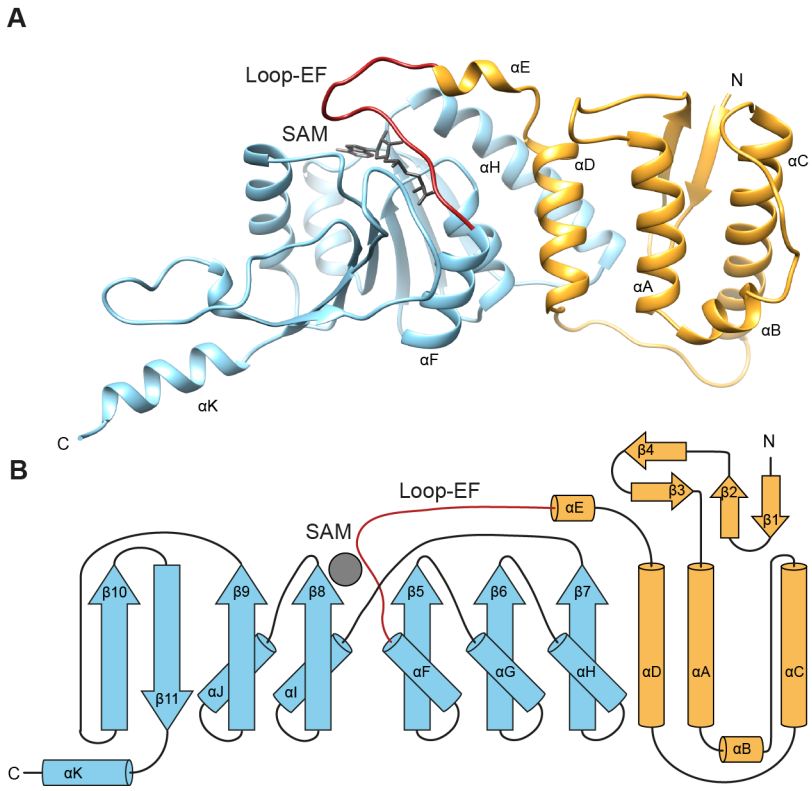


Figure 5. Overall structure of the human DOT1L methyltransferase
(**A**) Ribbon diagram of the human DOT1L crystal structure (protein data bank (PDB) accession number 1nw3) (Min et al., 2003). The N-terminal region is colored in orange, while the open α/β structure is shown in light blue, and the connecting loop-EF is depicted in red. Cofactor SAM is shown in grey. (**B**) Topological diagram of the catalytic core of the DOT1L enzymes. Secondary structure elements are colored as in (A). Helices are labeled alphabetically from αA to αK and β-strands are numbered β1 to β11 from the N- to the C-terminus. Figure modified from (Min et al., 2003).

1.7.3 SAM-binding pocket

The SAM-binding pocket is build by aa residues within the motifs I, I', II D1 and D2. This pocket forms a wide entrance that becomes narrower towards the center of the enzyme (Min et al., 2003; Sawada et al., 2004). The cofactor is bound and placed between the loop-EF and the carboxyl end of the open α/β structure (Min et al., 2003; Sawada et al., 2004), whereas the adenine group of SAM is pointing towards the outside and is surrounded by hydrophobic aa, the methylsulfonyl group enters deep into the enzyme, reaching the bottom of the lysine-binding pocket where it is available for the nucleophilic attack by the target lysine. The inside of the SAM-binding pocket contains a negatively charged environment that could contribute to substrate lysine deprotonation (Min et al., 2003; Sawada et al., 2004). Notably, this negatively charged environment is absent in the SAM-binding cleft of SET-domain methyltransferases. The cofactor SAM interacts with a great number of DOT1 residues (for details refer to (Min et al., 2003; Sawada et al., 2004)) and mutations at these positions abolishes enzymatic activity in yeast and human DOT1, verifying the critical importance of maintaining an intact SAM-binding pocket (Feng et al., 2002; Lacoste et al., 2002; Ng, 2002; van Leeuwen et al., 2002).

1.7.4 The substrate lysine-binding pocket

The entrances of the putative lysine-binding channel and the SAM-binding pocket are separated in DOT1 enzymes (Min et al., 2003; Sawada et al., 2004). As a consequence, cofactor and substrate can bind independently. Structural works have shown, that the lysine-binding pocket is lined with conserved hydrophobic residues, that allows the passage of the target lysine into the channel (Min et al., 2003; Sawada et al., 2004). This hydrophobic environment also potentially promotes a lowered lysine isoelectric point (pKa) (Westheimer, 1985), which favors deprotonation of the lysine ε-amine group. Deprotonation is essential to prime the substrate lysine for the subsequent methylation reaction, but the molecular details of this step are currently not well understood in DOT1 enzymes. Besides the influence on deprotonation, the lysine-binding pocket very likely contributes to the product-specificity of the DOT1 enzymes, since it creates the chemical environment for catalysis. Mutations focusing on residues within the lysine-binding pocket have been successful in revealing determinants for product-specificity of the SET domain containing methyltransferases PRMT5 (Sun et al., 2011) or DIM-5 (Zhang et al., 2003). However, no study has explored these aspects in DOT1 enzymes so far.

1.8 Motivation

T. brucei, the unicellular parasitic eukaryote that causes African sleeping sickness, will be exploited as a model system to shed light on the methylation state-specific activity of DOT1 enzymes. Interestingly, unlike most other eukaryotes, which possess only a single DOT1 enzyme, trypanosomes have two homologs, DOT1A and DOT1B that methylate histone H3 on lysine 76, the homologous residue to H3K79 in other organisms. Notably, DOT1A and DOT1B have been shown to differ in their product-specificity. While DOT1A mediates mono- and di-methylation of H3K76, only DOT1B is able to catalyze H3K76 tri-methylation. Different methylation levels exhibit completely different functions in trypanosomes. DOT1A-mediated H3K76me1/me2 seems to regulate replication initiation (Gassen et al., 2012), which has also been shown for higher eukaryotes recently (Fu et al., 2013; Kim et al., 2014). In contrast, H3K76me3 appears to be involved in transcriptional VSG switching during antigenic variation (Figueiredo et al., 2008) and developmental differentiation (Janzen et al., 2006b). Similar functions were also suggested for DOT1 enzymes during development of higher eukaryotes (Jones et al., 2008). In most cases, however, it is unclear if the mono-, di- or tri-methylated state is needed for the proper execution of these biological processes because only one enzyme is responsible for all different methylation states. Hence, *T. brucei* is an extremely useful model organism to unravel the function and regulation of individual histone lysine methylation states. Great progress has been made in understanding the structural and chemical properties of the substrate and cofactor binding sites in SET domain KMTases in the last years (Campagna-Slater et al., 2011; Kwon et al., 2003; Manzur et al., 2003; Min et al., 2002; Nguyen et al., 2013; Schapira, 2011; Wilson et al., 2002; Wu et al., 2010; 2013; Xu et al., 2011; Zhang et al., 2002; 2003). In comparison to this, the structure and function of DOT1 methyltransferases remains only poorly understood.

The aim of this thesis is to understand how different methylation levels are generated by trypanosome DOT1A and DOT1B methyltransferases. To this end, a combination of in vitro and in vivo approaches should be used to elucidate the molecular basis of specificity and activity of both

enzymes. The main tool will be a reconstitution system of trypanosomal nucleosomes to investigate DOT1A and DOT1B methyltransferase activities under defined in vitro conditions. Moreover, the unique feature of two different product specificities of the trypanosomal DOT1 proteins should help to gain insights into the molecular mechanism of these important enzymes. To unravel the function of trypanosome DOT1 enzymes would have long-reaching consequences for the functional understanding of homologous enzymes in higher eukaryotes and might offer new drug targets to fight the severe human diseases like African sleeping sickness and Chagas disease.

2 Material and Methods

All standard chemicals used in this study were obtained from AppliChem, Sigma, Roth, Merck, Serva, Sigma-Aldrich and Invitrogen, unless stated otherwise.

2.1 Molecular cloning

2.1.1 Expression vectors and bacterial strains

Cloning and expression vectors used in this study are summarized in Table 1. Bacterial strains for molecular cloning (*E. coli* TOP10) and protein expression (Rosetta Blue (DE3), BL21) are listed in Table 2. The plasmid pUC18 (Thermo Scientific) was used for amplification of the 601-sequence (Lowary and Widom, 1998).

Table 1. Cloning and expression vectors

Name	Selection marker	Origin of replication	Company
pUC18	Ampicillin	pMB1	Thermo Scientific
pET-21a(+)	Ampicillin	pBR322	Novagen
pMAL-c2X	Ampicillin	pBR322, M13	New England Biolabs

Table 2. Bacterial strains

E. coli Strain	Genotype	Resistance	Company
TOP10	F- mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 λ-	none	Invitrogen
BL21 (DE3)	F^- ompT gal dcm lon hsdS_B(r_B^- m_B^-) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])	none	New England Biolabs
Rosetta Blue (DE3)	endA1 hsdR17(rK12-mK12+) supE44 thi-1 recA1 gyrA96 relA1 (DE3) [F' proA+B+ lac1q ZΔ M::Tn10] pRARE (CamR, TetR)	Chloramphenicol, Tetracycline	Novagen

2.1.2 Bacterial growth

E. coli cells were grown in liquid Luria Bertani (LB) medium (5 g/l yeast extract (BD Bionutrients), 10 g/l bacto tryptone (AppliChem), 10 g/l sodium chloride (NaCl) or on LB agarose plates (LB medium containing additional 15 g/l agarose (BD Bionutrients)). The media was autoclaved at 121°C for 20 min and supplemented with appropriate antibiotics (34 µg/ml chloramphenicol and/or 100 µg/ml ampicillin) for plasmid selection prior use.

2.1.3 Polymerase chain reaction (PCR)

The polymerase chain reaction (PCR) was used to amplify double stranded DNA fragments. PCR reactions were conducted in Phusion buffer (Thermo Scientific) and contained specific forward and reverse primer (0.5 µM each) (synthesized by Sigma-Aldrich, Table 3), 200 µM dNTPs and template DNA (0.1 ng/µl - 2 ng/µl). PCR reactions were mixed with 2 U/µl Phusion High-Fidelity DNA polymerase (Thermo Scientific). The cycling conditions were adjusted to the length of the amplified DNA and melting temperature of the primers (Table 4). The obtained PCR fragments

were purified using the NucleoSpin Gel and PCR clean up Kit (Macherey-Nagel) according to the manufacturer's instructions. 50 µl EB buffer (Macherey-Nagel) were used in the final elution step.

Table 3. PCR primers

No.	Name	Sequence	Type	Project
GD17	DOT1A	ATG CCT GGA TTG CTA ATA TCC CGG	for	DOT1A WT cloning
GD18	DOT1A	CCCC AAG CTT TCA TCT CCG TCG GTG AAT GAA AAA AGG	rev	DOT1A WT cloning
GD19	DOT1B	ATG GAC GCA CGT GTT CAT CGT AGT AAG C	for	DOT1B WT cloning
GD20	DOT1B	CCCC AAG CTT TCA CGA TCG CTT GAT GTA AAG ATA AAA TGG	rev	DOT1B WT cloning
GD21	core601	TTA TGT GAT GGA CCC TAT ACG CGG CC	for	601 amplification
GD22	core601	CCC GAG TCG CTG TTC AAT ACA TGC A	rev	601 amplification
GD28	S218A	GCG AAC TTG TTG TTT CCA AAG TCC CTG AC	for	DOT1A mutagenesis
GD29	S218A	TAG AAG AAT TAC GGT TTT ACC CCT TTC TGA CTC A	rev	DOT1A mutagenesis
GD30	F246M	ATG GAC GAT CTT TAT CCG CAT AGC CGC	for	DOT1A mutagenesis
GD31	F246M	ACA TAG TAT CCT CGT TCC GCT TGG AAC C	rev	DOT1A mutagenesis
GD32	5'Biotin_601	(BtnTg) TTA TGT GAT GGA CCC TAT ACG CGG CC	for	601 amplification
GD37	Asf1A	CG GGA TCC ATG AGC ATA CAA CCA ATT GTA CAA C	for	Asf1A WT cloning
GD38	Asf1A	ACGC GTC GAC TCA TCT GGG TTC AAG TGC TTC	rev	Asf1A WT cloning
AK37	Asf1B	ATG ACC ACA GCC GGT CAG TC	for	Asf1B WT cloning
AK38	Asf1B	CCCC AAG CTT TTA ACG GTG GTG CTTT TCT TTC	rev	Asf1B WT cloning
CJ45	H2A	AGAT CAT ATG GCA ACA CCC AAA CAG GC	for	H2A cloning
CJ46	H2A	AT CTC GAG CTA GAC GCT TGG CGT CGC C	rev	H2A cloning
CJ47	H2B	AGAT CAT ATG GCC ACT CCT AAG AGC AC	for	H2B cloning
CJ48	H2B	AT CTC GAG CTA GCT GGA AGC GTG TGA CAC	rev	H2B cloning
CJ49	H3	AGAT CAT ATG TCG AGG ACC AAG GAA AC	for	H3 cloning
CJ50	H3	AT CTC GAG CTA TGC ACG TTC ACC GCG TAG	rev	H3 cloning
GD51	H4	CGC CAT ATG GCG AAG GGT AAG AAG AGT GGT	for	H4 cloning
GD52	H4	CG GAA TTC CTA TGC ATA ACC GTA CAG AAT CTT	rev	H4 cloning
GD58	H3V_NdeI	CGC CAT ATG ATG GCG CAA ATG AAG AAA ATA ACA C	for	H3V cloning
GD59	H3V_EcoRI	CG GAA TTC TTA GTT ACG CTC GCC TCG GAG	rev	H3V cloning
GD60	H4V_NdeI	CGC CAT ATG ATG GCA AAG GGC AAG AGA GTC	for	H4V cloning
GD61	H4V_EcoRI	CG GAA TTC TCA GTC GTA CCC GTA AAG AAC C	rev	H4V cloning
GD64	Nt-DOT1B+1A	GGC ACA TAG ATG ACG TGA CCC TTT GCA CGT GAC CTG TCG	for	DOT1A swapping
GD65	19bp1B +Ct-1A	CGA CAG GTC ACG TGC AAA GGG TCA CGT CAT CTA TGT GCC	rev	DOT1A swapping
GD66	L112V	GTG CTA CCT ACC TTT GTC TCG CGA ATG GTT CGC	for	DOT1A mutagenesis
GD67	L112V	GGA CTT GGC ACA TAG ATG ACG TGA CCC TGA CAC	rev	DOT1A mutagenesis
GD68	L112V_swap	GGA CTT GGC ACA TAG ATG ACG TGA CCC TTT GC	rev	DOT1A swapping
GD75	E100T	AAG CGC ATA ACG ACG GTG TCA GGG TCA CGT CAT CTA TGT GCC	for	DOT1A mutagenesis
GD76	F90Y	GAT ATC TAA GCG TCT ATA TGT GGC GTC CAA GAG ACG TTC G	rev	DOT1A mutagenesis
GD77	T99V_E100T	AAG CGC ATA GTC ACG GTG TCA GGG TCA CGT CAT CTA TGT GCC	for	DOT1A mutagenesis
GD78	T89C_F90Y_R92S	GAT ATC TAA CGA TCT ATA GCA GGC GTC CAA GAG ACG TTC GAA GAA GAC G	rev	DOT1A mutagenesis

Abbreviations: AK, Anika König; BtnTg, Biotin Tag; CJ, Christian Janzen; GD, Gülcin Dindar; for, forward; rev, reverse. Underlined sequences indicate restriction sites .

Table 4. Thermocycling conditions

Project	Procedure	Thermocycling protocol
DOT1 cloning DOT1 swapping Asf1 cloning Histone cloning	two-step	98°C/1 min - [98°C/10 sec - 72°C/1 min] x 35 - 72°C/10 min - 4°C/∞
DOT1 mutagenesis	two-step	98°C/1 min - [98°C/30 sec - 72°C/4 min] x 25 - 72°C/10 min - 4°C/∞
601 amplification	three-step	95°C/4 min - [95°C/45 sec - 58°C/30 sec - 72°C/1 min] x 35 - 72°C/10 min - 4°C/∞

2.1.4 Enzymatic digestion of DNA

Purified PCR fragments or plasmids were digested using commercially available restriction enzymes (New England Biolabs or Thermo Scientific), to obtain linearized plasmids and PCR fragments with appropriate ends for DNA ligation (see section 2.1.7). Digestions were performed according to the manufacturer's recommendations in a total volume of 20 µl with 1-2 µg DNA for 2-4 h at 37°C. The digested DNA was purified by agarose gel electrophoresis and gel extraction (see section 2.1.5).

2.1.5 Agarose gel electrophoresis and gel extraction

Agarose gel electrophoresis was used to separate DNA fragments according to size, for analytical (restriction digest control) and preparative (gel extraction) purposes. Gels contained 1 % or 0.8 % (w/v) agarose (AppliChem) in TAE buffer (40 mM Tris/hydrochloric acid (HCl) pH 8.0; 20 mM acetic acid; 1 mM ethylenediaminetetraacetic acid (EDTA)), respectively. DNA samples were supplemented with an appropriate amount of 6x loading dye (Thermo Scientific). Electrophoresis was conducted in TAE buffer at a constant voltage of 120 V. GeneRuler DNA ladder mix (Thermo Scientific) was used as molecular weight standard. DNA molecules were stained with ethidium bromide (0.1 µg/ml) and visualized under UV light on a Gel iX Imager (Intas). For extraction, gel pieces containing specific DNA fragments were cut out under UV light illumination (Vilber Lourmat) and processed with the Gel Extraction Kit (Macherey-Nagel) according to the manufacturer's protocol. 20 µl EB buffer was used in the final elution step.

2.1.6 DNA concentration determination

DNA concentrations were determined by light absorbtion at 260 nm (1 A_{260} corresponds to a DNA concentration of 50 µg/ml) using an Infinite M200 reader (Tecan) or estimated by agarose gel electrophoresis (see section 2.1.5) using GeneRuler DNA ladder mix as molecular weight standard with known concentration.

2.1.7 DNA ligation

Ligation was performed with purified plasmids and PCR fragments (see sections 2.1.5 and 2.1.3, respectively) at a molar ratio of 1:3. Reactions were set up in a total volume of 15 µl and supplemented with T4-Ligase and ligase buffer (Thermo Scientific) according to the manufacturer's protocol and incubated over night at 16°C.

2.1.8 Site directed mutagenesis

Site directed mutagenesis was used to introduce point and deletion mutations into the DNA sequence of a plasmid. Forward and/or reverse primers (Table 3) included the desired mutations and allowed primer extension away from the DNA sequence to be altered. Using a two-step thermocycling protocol (Table 4), the entire plasmid was amplified as a linear product with blunt ends containing the desired mutations. Prior recircularization, the template DNA was selectively degraded with the methylation dependent nuclease *Dpn*I (Thermo Scientific) over night at 37°C. Subsequently, the linearized DNA amplification product was precipitated with ethanol (see section 2.1.9), followed by ligation (see section 2.1.7) and transformation into magnesium chloride ($MgCl_2$) competent *E. coli* cells (see section 2.1.11).

2.1.9 Ethanol precipitation

Linearized plasmids (see section 2.1.4) were precipitated in 70 % (v/v) ethanol and 83 mM sodium acetate (Na(OAc)) over night at -20°C. The samples were centrifuged at 21.000 × g for 30 min at 4°C to pellet DNA. The pellet was washed once with 1 ml 70 % (v/v) ethanol, followed by a second centrifugation under the same condititions. The supernatant was discarded and the pellet was dried for 15 min at 40°C and further resuspended in 20 µl nuclease free water.

2.1.10 Preparation of magnesium chloride competent *E. coli* cells

A pre-culture of 5 ml LB medium was inoculated with a single colony of *E. coli* cells (Table 2) and incubated over night at 37°C under agitation (180 rpm). The pre-culture was diluted 1:100 and cultured until an optical density at 600 nm (OD_{600}) of 0.5-0.6 was reached. Cells were harvested in a 12169 rotor (Sigma) for 10 min at 900 × g at 4°C. The supernatant was discarded and the cell pellet stored on ice for 15 min. Afterwards, the pellet was carefully resuspended in 40 cell pellet volumes of pre-chilled LB medium, supplemented with 10 % (v/v) polyethylene glycol (PEG) 3350, 5 % (v/v) dimethyl sulfoxide (DMSO), 50 mM $MgCl_2$ under sterile conditions (the pH of the final modified LB medium was adjusted to pH 6.5). Competent *E. coli* cells were aliquoted at 4°C in pre-chilled (-80°C) 1.5 ml tubes, subsequently flash-frozen in liquid nitrogen and stored at -80°C.

2.1.11 DNA transformation into competent *E. coli* cells

Ligated plasmid DNA was transformed into $MgCl_2$ competent *E.coli* cells by heat shock. To this end, 100 µl competent *E.coli* cell suspension (see section 2.1.10 and Table 2) was mixed with 10-40 ng plasmid DNA and incubated on ice for 30 min. The cells were heat shocked at 42°C for 60 sec, followed by an incubation on ice for 2 min and recovery in 1 ml pre-warmed SOC-medium (5 g/l yeast extract (BD Bionutrients), 2 g/l tryptone (AppliChem), 10 mM NaCl, 2.5 mM potassium acetate (KCl), 10 mM $MgCl_2$, 10 mM magnesium sulfate ($MgSO_4$), 20 mM Glucose) for 1 h at 37°C. Cells were harvested gently by centrifugation using a table top centrifuge (Hermle) for 5 min at 1000 × g. 900 µl of the supernatant was discarded and cells were carefully resuspended in the remaining SOC-medium. The cells were plated on selective LB plates and incubated over night at 37°C.

2.1.12 Plasmid isolation

Plasmid DNA was purified from 5 ml *E. coli* (TOP10) over night cultures with appropriate antibiotics to select plasmid-transformed bacteria. DNA was extracted from the cells using NucleoSpin Plasmid Miniprep kit (Macherey-Nagel) according to the manufacturer's instructions and eluted in 50 µl EB buffer (Macherey-Nagel).

2.1.13 DNA sequencing

Plasmids (Table 6) were sent to the company GATC (Konstanz, Germany) for sequencing analysis. Samples contained at least 100 ng/µl and were prepared according to the company's instructions. Sequencing results were verified using CLC Software (CLC bio, Qiagen) and the ApE plasmid editor (http://biologylabs.utah.edu/jorgensen/wayned/ape/).

2.2 Protein sample analysis

2.2.1 Trichloroacetic acid (TCA) protein precipitation

Protein samples were filled up to a final volume of 1 ml with water and mixed with 0.015 % (w/v) Sodium-desoxycholate and 7.2 % (v/v) TCA solutions. Proteins were precipitated for 30 min on ice. Subsequently, the samples were centrifuged for 30 min at 21.000 × g at 4°C. The supernatant was removed and the pellet was washed once with 1 ml cold (-20°C) acetone. The centrifugation step was repeated for 10 min. The acetone was discarded and the sample was dried for 10 min at room temperature (RT). The precipitate was resuspended in 50 µl 2 x sodium dodecyl sulfate (SDS) loading buffer (100 mM Tris/HCl pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 0.2 % (w/v) bromphenol blue, and 200 mM β-mercaptoethanol, heated to 98°C for 10 min and further analysed by SDS-polyacrylamide gel electrophoresis (PAGE).

2.2.2 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used to separate proteins according to their molecular mass. Samples were mixed with 2 x SDS-loading buffer (100 mM Tris/HCl pH 6.8, 4 % (w/v) SDS, 20 % (v/v) glycerol, 0.2 % (w/v) bromphenol blue, and 200 mM β-mercaptoethanol), denaturated at 95°C for 10 min and subsequently loaded on polyacrylamide gels in a loading chamber (Biorad). Stacking gels consisted of 5 % (v/v) acrylamid/bisacrylamid (5:1) in stacking gel buffer (500 mM Tris/HCl pH 6.8, 0.4 % (w/v) SDS), while separation gels contained 15 % (v/v) acrylamid/bisacrylamide (5:1) in separation buffer (1.5 M Tris/HCl pH 8.8, 0.4 % (w/v) SDS). For size estimation of separated proteins, a prestained marker (Thermo Scientific) was used. The electrophoresis was performed in SDS-running buffer (25 mM Tris; 192 mM glycine; 0,1 % (w/v) SDS) at a constant voltage of 180 V using an EV261 electrophoresis power supply (PEQLAB).

2.2.3 Protein gel staining

Protein gels were transferred into 50 ml water and boiled for 60 sec using a mircowave oven. The water was discarded and the gels were stained with coomassie solution (0.12 % (w/v) Coomassie Brilliant Blue R250 (AppliChem), 20 % (v/v) ethanol, 10 % (v/v) ammonium sulfate,

10 % phosphoric acid) for 30 sec using a mircowave oven. Destaining followed in water until a clear background was obtained.

2.2.4 Western blot analysis

For western blot analysis, proteins were transferred from polyacrylamide gels to polyvinylidene fluoride (PVDF) membranes (Merck Millipore). The PVDF membrane was incubated in 100 % methanol for 10 sec and further stored in water until use. For the protein transfer, a semi-dry blotting sandwich was prepared from bottom to top: 2 sheets of filter paper (Albet) incubated in anode buffer (25 mM Tris/HCl pH 7.6, 20 % (v/v) methanol), PVDF membrane, polyacrylamide gel, 1 sheet of filter paper incubated in cathode buffer (300 mM Tris/HCl pH 7.6, 20 % (v/v) methanol, 40 mM ε-aminocaproic acid) and proteins were transferred in a semi-dry blotter (Biorad) for 1 h at 55 mA (approximately 1 mA/cm^2 gel). The membrane was blocked with 3 % (w/v) bovine serum albumine (BSA) (AppliChem) in phosphate buffered saline (PBS) (10 mM disodium phosphate (Na_2HPO_4); 1,8 mM monopotassium phosphate (KH_2PO_4) pH 7.4, 140 mM NaCl; 2,7 mM KCl) for 1 h at RT or over night at 4°C, to avoid unspecific binding of the antibodies. The BSA solution was discarded and the membrane washed with PBS-tween (PBST)/0.1 % (PBS supplemented with 0.1 % (v/v) Tween-20) for 2 min. The primary antibody was diluted in PBST/0.1 % and incubated with the membrane as described (Table 5.1). Next, the membrane was washed 3 x 10 min with PBST/0.2 % (PBS supplemented with 0.2 % (v/v) Tween-20) and incubated with secondary antibodies in PBST/0.1 % as described (Table 5.2). The membrane was washed 3 x 10 min with PBST/0.2 % and dried between filter papers. Secondary antibody signals were detected with an Odyssey infrared imaging system (LI-COR Biosciences).

2.2.5 Antibodies

The following primary (Table 5.1) and secondary (Table 5.2) antibodies were used for western blot analysis.

Table 5.1 Primary antibodies

Name	Animal	Type	Target	Dilution	Incubation	Reference
B1-Dia	rabbit	pAb	Tb H3K76me1	1 : 1000	over night, 4°C	Gassen et al., 2012
717.2	rabbit	pAb	Tb H3K76me2	1 : 2000	over night, 4°C	Janzen et al., 2006
446.2	rabbit	pAb	Tb H3K76me3	1 : 2000	over night, 4°C	Janzen et al., 2006
pH3	rabbit	pAb	Tb H3	1 : 10.000	1 h, RT	Gassen et al., 2012
pH4	rabbit	pAb	Tb H4	1 : 830	1 h, RT	Gassen et al., 2012
pH2A	guinea pig	pAb	Tb H2A	1 : 1000	1 h, RT	this study
MBP Nr. 8	mouse	mAb	MBP	1 : 5000	1 h, RT	this study
IGg2a MBP	rat	mAb	MBP	1 : 10.000	1 h, RT	Pascoalino et al., 2013

Abbreviations: mAb, monoclonal antibody; MBP, maltose binding protein; pAb, polyclonal antibody; RT, room temperature; Tb, *T. brucei*.

For detection of *T. brucei* histone H2A, a polyclonal guinea pig antibody was generated by the Company Pineda with the purified recombinant protein (see section 2.3.3.3). The monoclonal maltose binding protein (MBP) antibody was a gift from Elisabeth Kremmer (Helmholtz Center, Munich).

Table 5.2 Secondary antibodies

Name	Animal	Type	Target	Company	Dilution	Incubation
IRDye 680LT	goat	pAb	rabbit IgG	LI-COR	1 : 50.000	1 h, RT
IRDye 680LT	goat	pAb	mouse IgG	LI-COR	1 : 50.000	1 h, RT
IRDye 680LT	goat	pAb	rat IgG	LI-COR	1 : 50.000	1 h, RT
IRDye 680LT	donkey	pAb	rabbit IgG	LI-COR	1 : 50.000	1 h, RT
IRDye 800CW	goat	pAb	rabbit IgG	LI-COR	1 : 50.000	1 h, RT
IRDye 800CW	goat	pAb	rat IgG	LI-COR	1 : 50.000	1 h, RT
IRDye 800CW	donkey	pAb	guinea pig IgG	LI-COR	1 : 50.000	1 h, RT

Abbreviations: IgG, immunoglobulin G; mAb, monoclonal antibody; pAb, polyclonal antibody; RT, room temperature.

2.3 Protein expression and purification from *E. coli*

2.3.1 Heterologous protein test expression in *E. coli* Rosetta Blue (DE3) and BL21

Plasmids (Table 6) coding for trypanosomal proteins, were freshly transformed into *E. coli* Rosetta Blue (DE3) or BL21 cells (see section 2.1.11). Single colonies from the selective plate were used to inoculate 20 ml of selective LB medium, followed by incubation over night at 37°C, shaking at 180 rpm. Pre-cultures were diluted to an OD_{600} of 0.1 in 100 ml LB medium supplemented with appropriate antibiotics and grown until an OD_{600} of 0.5-0.6 was reached. Optimal conditions for protein expression were determined by induction with different amounts of isopropyl β-D-1-thiogalactopyranoside (IPTG) ranging from 0.1 - 1 mM for 1 h at RT, 28°C and 37°C.

2.3.2 Bacterial cell lysis

Cells were harvested by centrifugation using a 12169 rotor (SIGMA) for 1 min at 11.000 × g at 4°C or a Fiberlite F14-6 x 250y rotor (Thermo Scientific) for 20 min at 4000 × g at 4°C. Pellets were resuspended in lysis buffer (lysis buffer A or B for DOT1 or Asf1 proteins, respectively) (see sections 2.3.3.1 and 2.3.3.2) in a total volume of 2 ml. Shortly before the actual lysis step, cOmplete protease inhibitor cocktail mix without EDTA (Roche) was added according to the manufacturer's instruction. Cells were lysed with high power setting and a 30 sec on/off sonication cycle for 10 min in an ice water bath using a Biorutor (Diagenode). The soluble fraction (supernatant) was cleared from cell debris (pellet) by centrifugation (21.000 × g, 30 min, 4 °C).

2.3.3 Description of individual protein purification strategies

2.3.3.1 Purification of *T. brucei* DOT1A and DOT1B methyltransferases

Full length *T. brucei* DOT1A (TriTryp database identifier (ID): Tb927.8.1920) and DOT1B (Tb927.1.570) coding sequences were amplified by PCR from genomic DNA using specific primers (Table 3) and cloned into the pMAL-c2X expression vector (New England Biolabs) using *Hind*III and *Xmn*I restriction sites, resulting in N-terminal fusions to the MBP (Table 6). Mutations were introduced using site directed mutagenesis following standard procedures (see section 2.1.8). Vectors were transformed into *E. coli* Rosetta Blue (Table 2) and cells were grown in selective media (LB medium supplemented with 100 µg/ml ampicillin and 34 µg/ml chloramphenicol) to an OD_{600} of 0.5-0.6 prior to induction of protein expression with 1 mM IPTG for 2.5 h at 28°C. Cells were harvested (8000 × g, 15 min, 4°C) and lysed in buffer A (20 mM Tris/HCl pH 7.4, 200 mM

NaCl, 1 mM EDTA) using a Bioruptor (Diagenode) (see section 2.3.2). Cell debris was separated from the soluble protein fraction by centrifugation (21.000 × g, 30 min, 4 °C) and crude extract (2 ml) was incubated with 50 µg of washed amylose resin (New England Biolabs) for 1 h at 4°C, on a rotating wheel. Beads were harvested (1000 × g, 1 min, 4°C), followed by three washes with buffer A and elution with maltose buffer (Buffer A supplemented with 10 mM maltose). Aliquots of the eluate were mixed with 2x SDS-loading buffer, denaturated at 95°C for 10 min and analyzed by SDS-PAGE (see section 2.2.2). Purified DOT1 proteins were used fresh for subsequent experiments as no storage is possible without loss of activity.

Table 6. Protein expression plasmids

Project (s)	pMAL-c2X constructs	Tag (terminus)	Origin	Strain
DOT cloning	pMAL-c2X-DOT1A WT	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT cloning	pMAL-c2X-DOT1B WT	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A G138R	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1B G121R	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A S218A	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X- DOT1A F246M	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A L112V	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A S218A/F246M	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A S218A F246M/L112V	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A F90Y/E100T	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A F90Y/E100T/ S218A/F246M	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A T89C/F90Y/	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT mutagenesis	pMAL-c2X-DOT1A T89C/F90Y/R92S/	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT swapping	pMAL-c2X-Swap DOT1A WT	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT swapping	pMAL-c2X-Swap DOT1A S218A	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT swapping	pMAL-c2X-Swap DOT1A F246M	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT swapping	pMAL-c2X-Swap DOT1A L112V	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT swapping	pMAL-c2X-Swap DOT1A S218A/F246M	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
DOT swapping	pMAL-c2X-Swap DOT1A S218A/ F246M/L112V	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
Asf1 cloning	pMAL-c2X-Asf1A WT	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
Asf1 cloning	pMAL-c2X-Asf1B WT	MBP (N)	*T. brucei*	Rosetta Blue (DE3)
Project (s)	**pET-21a constructs**	**Tag (terminus)**	**Origin**	**Strain**
Histone cloning	pET-21a-H3	-	*T. brucei*	BL21 (DE3)
Histone cloning	pET-21a-H4	-	*T. brucei*	BL21 (DE3)
Histone cloning	pET-21a-H2A	-	*T. brucei*	BL21 (DE3)
Histone cloning	pET-21a-H2B	-	*T. brucei*	BL21 (DE3)
Histone cloning	pET-21a-H3V	-	*T. brucei*	BL21 (DE3)
Histone cloning	pET-21a-H4V	-	*T. brucei*	BL21 (DE3)

2.3.3.2 Purification of *T. brucei* Asf1A and Asf1B histone chaperones

Full length *T. brucei* Asf1A and Asf1B coding sequences were amplified by PCR from genomic DNA using specific primers (Table 3) and cloned into the expression vector pMAL-c2X using *BamH*I/*Sal*I (for Asf1A) or *Xmn*I/*Hind*III (for Asf1B) restriction sites to generate N-terminal MBP fusion proteins. Vectors were transformed into *E. coli* Rosetta Blue (DE3) and cells were grown in selective (100 µg/ml ampicillin and 34 µg/ml chloramphenicol) LB media (100 ml) to an OD_{600} of 0.3, followed by the addition of 10 mM benzylalcohol and continued incubation at 37°C. Once an OD_{600} of 0.5-0.6 was reached, protein expression was induced by adding 1 mM IPTG, followed by

incubation for 1 h at 37°C. Cells were lysed in buffer B (25 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)/potassium hydroxide (KOH) pH 7.5, 150 mM KCl, 10 % (v/v) glycerol, 12.5 mM $MgCl_2$, 1 mM dithiotreitol (DTT) and 1 mM PMSF) using a Bioruptor (Diagenode) (see section 2.3.2). Cell debris was separated from soluble protein fraction by centrifugation (21.000 × g, 30 min, 4°C). Crude extract was filtered through Amicon Ultra centrifuge tubes (Millipore) with a molecular weight cut-off (MWCO) of 100 kDa. MBP-fusion proteins were purified from the cleared lysate (2 ml) by affinity chromatography using 250 µg washed amylose resin (New England Biolabs). Beads and proteins were incubated for 1 h at 4°C on a rotating wheel, followed by three washes with 1 ml buffer B and elution with maltose buffer (Buffer B supplemented with 10 mM maltose). Aliquots of the eluate were mixed with 2x SDS-loading buffer, denaturated at 95°C for 10 min and analyzed by SDS-PAGE (see section 2.2.2). Purified Asf1 proteins were used fresh for subsequent experiments.

2.3.3.3 Purification of *T. brucei* histone proteins from inclusion bodies

T. brucei full-length open reading frames of histones H2A (TriTryp database ID: Tb927.7.2850), H2B (Tb927.10.10510), H3 (Tb927.1.2510), H4 (Tb927.5.4170) and histone variants H3V (Tb927.10.15350), H4V (Tb927.2.2670) were amplified by PCR from genomic DNA using specific primers (Table 3). PCR products were cloned into the pET21a (+) vector (Novagen) using *Nde*I/*Xho*I restriction sites for H2A, H2B and H3, as well as *Nde*I/*EcoR*I restriction sites for H4, H3V and H4V. All vectors were individually freshly transformed into *E. coli* BL21 (DE3) (see section 2.1.11) (Table 2). Cells were grown in selective media (LB medium supplemented with 100 µg/ml ampicillin, 1 % (v/v) glucose) to an OD_{600} of 0.6 prior to induction with 1 mM IPTG for 3-4 h at 37°C. The addition of 1% (w/v) glucose to the medium appeared to be beneficial for efficient expression of several histone proteins (H2B, H3V, H4 and H4V) (Table 7). Moreover, the use of buffled flasks as well as the ratio of medium to flask volume turned out to be critical for efficient expression. Individual expression conditions for each histone are summarized in Table 7.

Table 7. Optimization of expression protocols for trypanosome histone proteins

Histone	*E. coli* strain	Glucose (1 %)	buffled flask	LB-medium : buffled flask	IPTG	Expression
H2A	BL21	not required	only	1 : 2.5 to 1 : 10	1 mM	3 - 4h
H2B	BL21	essential	only	1 : 10	1 mM	4h
H3	BL21	not required	only	1 : 2.5 to 1 : 10	1 mM	3 - 4h
H3V	BL21	essential	only	1 : 5 to 1 : 10	1 mM	3 - 4h
H4	BL21	essential	only	1 : 10	1 mM	4h
H4V	BL21	essential	only	1 : 10	1 mM	3 - 4h

After protein expression, cells were harvested (8000 × g, 15 min, 4°C) using a Fiberlite F14-6x250y rotor (Thermo Scientific), washed once in histone wash buffer (HWB) (50 mM Tris/HCl pH 7.5, 100 mM NaCl, 1 mM EDTA, 1 mM benzamidine, 5 mM β-mercaptoethanol), and pelleted again (23.000 × g, 20 min, 4°C) using a Sorvall SS-34 rotor (Thermo Scientific). Cells were resuspended in HWB and lysed by sonification on ice (10 x 30 sec at high power with 30 sec pauses in between) using a Bioruptor (Diagenode) (see section 2.3.2). The inclusion bodies containing cell fraction was harvested (23.000 × g, 20 min, 4°C), resuspended in HWB supplemented with 1 % (v/v) Triton X-100 (HWB/T) and sonicated as described above. The purification steps (resuspension, sonification and harvest) were repeated once in HWB/T and twice in HWB. After the final purification step,

inclusion bodies were resuspended in unfolding buffer (7 M guanidinium chloride, 20 mM Tris/HCl pH 7.5, 10 mM DTT) for 1 h at RT and separated from remaining cell debris by centrifugation (23.000 × g, 15 min, 4°C). Unfolded proteins were dialyzed in Pur-A-Lyzer Maxi Dialysis kit (MWCO 6-8 kDa) (Sigma-Aldrich) twice against 1 l SAU-200 buffer (7 M Urea, 20 mM Na(OAc) pH 5.2, 200 mM NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol), for 1 h each, followed by a third dialysis step against 1 l SAU-200 buffer over night at 4°C. Proteins were separated from precipitates (formed during the dialysis step) by centrifugation (23.000 × g, 15 min, 4°C) and applied on a HiTrap SP HP (5 ml bed volume) cation exchange column (GE Healthcare) in SAU-200 buffer. The column run was performed with a constant flow rate of 0.5 ml/min. Histone proteins were eluted from the column during a salt gradient run from 200 to 600 mM NaCl, reaching 100 % SAU-600 buffer (7 M Urea, 20 mM Na(OAc) pH 5.2, 600 mM NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol) after 40 min, while collecting 500 µl sample fractions and recording the absorbtion profile at 280 nm wavelength (A_{280}). Protein fractions were analyzed by SDS-PAGE (see section 2.2.2), suitable fractions were pooled and dialyzed twice against 1 l of cold distilled water containing 2 mM β-mercaptoethanol, followed by a third dialysis step over night at 4°C. Pure histone proteins were lyophilized in 1 mg fractions using an Alpha 1-2 LD freeze dryer (Christ) and stored at -80°C.

2.4 Nucleosome assembly

2.4.1 Histone octamer assembly

Refolding of trypanosomal histone octamers was performed using lyophilized recombinant full-length histone proteins (see section 2.3.3.3). The four core histone proteins (H2A, H2B, H3, H4) were separately dissolved in unfolding buffer (7 M guanidinium chloride, 20 mM Tris/HCl pH 7.5, 10 mM DTT) and adjusted to a final concentration of 2 mg/ml each. The unfolding reaction was incubated for 30 min at RT. Histone proteins were subsequently mixed with a slight molar excess of H2A and H2B to ensure complete histone octamer formation (H2A : H2B : H3 : H4 in a molar ratio of 1.2 : 1.2 : 1 : 1). The total protein concentration of the mixed histone sample was adjusted to 1 mg/ml with unfolding buffer. For the refolding process, samples were dialyzed in 6-8 kDa MWCO dialysis tubing twice against 1 l of refolding buffer (10 mM Tris/HCl pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol) for 2 h at 4°C, followed by one dialysis against 1 l refolding buffer over night at 4°C. After the final dialysis step, samples were spun down (13.000 × g, 5 min, 4°C) and the supernatant containing assembled histone octamers was further concentrated via a Microcon-10 centrifugal filter unit (MWCO 10 KDa) (Millipore). Histone octamers were subsequently separated from dimers, tetramer and hexamers by size exclusion chromatography (see section 2.4.2).

2.4.2 Size exclusion chromatography

To purify octamers, refolded histone samples were applied on a Superdex 75 10/300 column or alternatively on a Superdex 200 HiLoad 16/600 (both GE Healthcare). Prior loading, the column was equilibrated with refolding buffer (10 mM Tris/HCl pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol). Gel filtration was performed at a flow rate of 0.5 ml/min while collecting 500 µl fractions and recording the A_{280} profile. Suitable fractions were analyzed by SDS-PAGE, combined and concentrated to approximately 1 mg/ml using a Microcon-30 centrifugal filter unit (MWCO 30 kDa) (Millipore). The final concentration was estimated by SDS-PAGE using BSA

samples with known concentration as a reference. Octamer aliquots were flash frozen in liquid nitrogen and stored at -80°C until use. In parallel, histone H2A/H2B dimers and histone tetramers H3/H4 were separately collected according to the A_{280} profile, analyzed by SDS-PAGE, pooled and concentrated using a Microcon-10 centrifugal filter unit (MWCO 10 kDa) (Millipore). Aliquots of histone dimers and tetramers were flash frozen in liquid nitrogen and stored at -80°C.

2.4.3 Histone tetramer refolding

Assembly of trypanosomal histone tetramers (H3/H4 or H3V/H4V) was performed using lyophilized pure recombinant full-length histone proteins (see section 2.3.3.3). Histone proteins (H3/H4 or H3V/H4V) were separately dissolved in unfolding buffer (7 M guanidinium chloride, 20 mM Tris/HCl pH 7.5, 10 mM DTT) and adjusted to a final concentration of 2 mg/ml each. The unfolding reaction was incubated for 30 min at RT. Histone proteins were subsequently mixed with a 1 : 1 molar ratio (H3 : H4 or H3V : H4V). The total protein concentration of the mixed histone sample was adjusted to 1 mg/ml with unfolding buffer. To refold histones, samples were dialyzed in 6-8 kDa MWCO dialysis tubing twice against 1 l of refolding buffer (10 mM Tris/HCl pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol) for 2 h at 4°C, followed by dialysis against 1 l refolding buffer over night at 4°C. After the final dialysis step, samples were spun down (13.000 × g, 5 min, 4°C) and the supernatant containing assembled histone tetramers were analyzed by SDS-PAGE and coomassie stain (see sections 2.2.2 and 2.2.3) and stored at 4°C.

2.4.4 Nucleosome assembly

For nucleosome assembly a pUC18 vector, containing 12 repeats of the 601 nucleosome positioning sequence (gift from Daniela Rhodes) was used (Lowary and Widom, 1998). The 204 bp 601 sequences were cut out from the vector using the *Ava*I restriction sites located between the repeats. DNA fragments were separated by electrophoresis using a 0.8 % (w/v) agarose gel and purified by gel extraction (see section 2.1.5). All components required for nucleosome assembly (histone octamers, excess H2A/H2B dimers and 601 DNA sequence) were mixed in a 1 : 1.2 : 1 molar ratio in reconstitution buffer (20 mM Tris/HCl pH 7.5, 10 mM DTT, 1 mM EDTA, 2 M KCl, 0.5 mM benzamidine) and incubated on ice for 30 min. The samples were subsequently placed into dialysis tubing with a 6-8 kDa MWCO and transferred into high salt buffer (10 mM Tris/HCl pH 7.5, 2 M KCl, 1 mM EDTA, 1 mM DTT, 0.5 mM benzamidine). A peristaltic pump with a flow rate of 210 µl/min was used to gradually replace the high salt buffer with low salt buffer (10 mM Tris/HCl pH 7.5, 250 mM KCl, 1 mM EDTA, 1 mM DTT, 0.5 mM benzamidine) during a time window of 50 h at 4°C. Afterwards, the dialysis tube was transferred to pre-cooled TCS buffer (20 mM Tris/HCl pH 7.5, 1 mM EDTA, 1 mM DTT) and incubated over night at 4°C. Nucleosome samples were supplemented with 15 % (v/v) glycerol and analyzed on a 5 % (w/v) native polyacrylamide gel in 0.5 x TBE pH 8.0 (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA) without any dye. To visualize the nucleosomal DNA, native acrylamide gels were stained with either diluted ethidium bromide (0.1 µg/ml) or SYBR Gold (Life technologies) solution in 0.5 x TBE buffer. Dye-nucleic complex were visualized under UV light on a Gel iX Imager (Intas).

2.4.5 Large-scale nucleosome assembly

All four lyophilized histones were separately solubilized in unfolding buffer (7 M guanidinium chloride, 20 mM Tris/HCl pH 7.5, 10 mM DTT), adjusted to a total protein concentration of 2 mg/ml each and incubated for 30 min at RT. Histone proteins were individually refolded in 6-8 kDa MWCO dialysis tubing twice against 1 l refolding buffer (10 mM Tris/HCl pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM β-mercaptoethanol) for 2 h at 4°C, followed by one additional dialysis step against 1 l refolding buffer over night at 4°C. Histone concentrations were determined by SDS-PAGE and coomassie staining (see sections 2.2.2 and 2.2.3), using BSA samples with known concentrations as a reference. The nucleosome reconstitution mixture contained 0.7 mg/ml of amplified 601 DNA fragment (see section 2.1.3) and 0.16 mg/ml of each histone (0.64 mg/ml protein in total) in assembly buffer (20 mM Tris/HCl pH 7.7, 2 M NaCl, 1 mM EDTA, 10 mM DTT, 0.5 mM benzamidine). The sample was incubated for 30 min at 4°C and subsequently transferred into 6-8 kDa MWCO dialysis tubing for dialysis against high salt buffer B (20 mM Tris/HCl pH 7.7, 2 M NaCl, 1 mM EDTA, 1 mM DTT, 0.5 mM benzamidine), followed by a gradual replacement for low salt (20 mM Tris/HCl pH 7.7, 250 mM NaCl, 1 mM EDTA, 1 mM DTT, 0.5 mM benzamidine) using a peristaltic pump with a flow rate of 210 µl/min over a time window of 50 h at 4°C. Afterwards, the dialysis tubing was transferred to pre-cooled zero salt buffer (20 mM Tris/HCl pH 7.7, 1 mM EDTA, 1 mM DTT, 0.5 mM benzamidine) and incubated over night at 4°C. Nucleosomes were supplemented with 15 % (v/v) glycerol and analyzed on a 5 % (w/v) native polyacrylamide in 0.5 x TBE pH 8.0 (44.5 mM Tris, 44.5 mM boric acid, 1 mM EDTA) gel without any dye. The acrylamide gel was stained with ethidium bromide (0.1 µg/ml) or SYBR GOLD in 0.5 x TBE buffer to visualize the nucleosomal DNA.

2.5 Functional analysis of purified proteins

2.5.1 Histone methyltransferase assay

To analyze methyltransferase activity, 22.5 ng/µl nucleosomes were incubated with 30 ng/µl of purified DOT1A or DOT1B enzymes in reaction buffer (50 mM Tris/HCl pH 8.0, 50 mM KCl, 5 mM calcium chloride ($CaCl_2$), 100 mM NaCl, 2.5 % (v/v) glycerol, 0.025 % (v/v) NP-40, 0.5 mM DTT) in the presence of 2.5 mM SAM in a total volume of 80 µl. Samples were incubated for 1 h, 3 h or until completion of the reaction at 30°C. Samples were analyzed by SDS-PAGE and western blot (see 2.2.2 and 2.2.4).

2.5.2 In vitro Asf1 pull-down assay

Equal amounts of purified MBP-tagged Asf1A or Asf1B (see section 2.3.3.2) (2.1 µM) were mixed with 2.1 µM histone tetramers (H3/H4) or histone tetramer variants (H3V/H4V) and reactions were adjusted to a final volume of 2 ml using zero salt buffer (25 mM HEPES/KOH pH 7.5, 10 % (v/v) glycerol, 12.5 mM $MgCl_2$, 1 mM DTT and 1 mM PMSF). Samples were incubated for 30 min on a rotating wheel at 4°C and supplemented with 50 µg amylose resin (New England Biolabs) to pull down MBP-tagged proteins. After an incubation for 1 h under rotation at 4°C, the bead suspension was transferred to a 5 ml column (Pierce). After three extensive washes with 1 ml buffer B (25 mM HEPES/KOH pH 7.5, 150 mM KCl, 10 % glycerol, 12.5 mM $MgCl_2$, 1 mM DTT, 1 mM PMSF and 1 x cOmplete EDTA free protease inhibitor cocktail (Roche)), proteins were eluted with maltose buffer (25 mM HEPES/KOH pH 7.5, 150 mM KCl, 10 % (v/v) glycerol, 12.5 mM

$MgCl_2$, 1 mM DTT and 1 mM PMSF, 10 mM maltose), TCA precipitated (see section 2.2.1) and analyzed by SDS-PAGE and western blot (see sections 2.2.2 and 2.2.4). MBP fusion proteins were detected by western blots analysis using a monoclonal rat MBP antibody (gift from E. Kremmer). Histones H3 and H4 (and their variants) were detected using a polyclonal rabbit antibody (Gassen et al., 2012) or a polyclonal peptide antibody (Siegel et al., 2008), respectively. Secondary antibodies (IRDye 800CW goat anti-rabbit and IRDye 680LT goat anti-rat (LI-COR Biosciences)) were detected with an Odyssey infrared imaging system (LI-COR).

2.5.3 Ex vivo Asf1 pull-down assay

Procyclic forms (PCF) of *T. brucei* strain 29-13 (Wirtz and Clayton, 1995) were cultivated in SDM-79 medium (Brun and Schönenberger, 1979) in the presence of 10 % fetal bovine serum (FBS) at 28°C. For pull-down assays with trypanosome whole cell extracts, $2x10^8$ PCF cells were harvested (1300 × g, 10 min, 4°C) and washed in cold PBS (10 mM Na_2HPO_4, 1,8 mM KH_2PO_4 pH 7.4, 140 mM NaCl, 2,7 mM KCl). Cell pellets were resuspended in buffer B (25 mM HEPES (pH 7.5), 150 mM KCl, 10 % glycerol, 12.5 mM $MgCl_2$, 1 mM DTT, 1 mM PMSF and 1 x cOmplete EDTA free protease inhibitor cocktail (Roche)) and lysed by three freeze and thaw cycles using liquid nitrogen, followed by sonification on ice (10 x 30 sec at high power with 30 sec pauses in between) using a Bioruptor (Diagenode). Cell extracts were cleared by centrifugation (10.000 × g, 10 min 4°C) and treated with 5 U DNase I (Thermo Scientific) for 10 min at 37°C. Purified MBP-tagged Asf1A or Asf1B were incubated with 50 µg amylose resin (New England Biolabs) and agitated for 1 h at 4°C. After two extensive washes with buffer B, immobilized Asf1A or Asf1B proteins were incubated with trypanosome cell extract for 1 h at 4°C. Next, bound protein complexes were washed five times with buffer B and eluted with maltose buffer (Buffer B + 10 mM maltose) and precipitated with TCA (see 2.2.1). Samples were analyzed as in section 2.5.2.

2.6 Homology modeling of DOT1 methyltransferases

Template searches were performed with HHpred (Söding et al., 2005), which is based on the pairwise comparison of profile hidden Markov models (HMMs) (Söding, 2005), using full-length *T. brucei* DOT1A and DOT1B sequences as input. Among the obtained results, the *H. sapiens* DOT1L crystal structure (PDB accession number 1NW3) (Min et al., 2003) was chosen as structural template for subsequent homology model building. The underlying alignments generated by HHpred had expect (E)-values of 1.1e-25 and 6.9e-22 for DOT1A and DOT1B, respectively. Note that the *S. cerevisiae* Dot1p crystal structure (PDB accession number 1U2Z) (Sawada et al., 2004) gave slightly better results in this search (E-values of 5.7e-26 and 4.7e-23 for DOT1A and DOT1B, respectively). However, the yeast template was not chosen because of the loop β10-β11 interaction with the substrate binding site observed in the Dot1p crystal structure (Sawada et al., 2004), which could pose unwanted restraints on the homology models. Similarity of *T. brucei* DOT1 enzymes with the DOT1L dropped considerably for the N-terminal part of the enzyme prior residues 79 and 62 for DOT1A and DOT1B, respectively, which prevented extension of the homology models beyond this point. The conserved CAKS motif (residues 108-111 and 91-94 for *T. brucei* DOT1A and DOT1B, respectively) was manually re-aligned under the D1 motif of the human DOT1L enzyme according to the structural argumentation described in the results section.

The DOT1 alignments obtained by HHpred were manually combined with the repositioned CAKS motif in Jalview (Waterhouse et al., 2009), exported in the PIR alignment file format and manually fed in Modeller (Eswar et al., 2008) with the human DOT1L crystal structure (PDB accession number 1NW3, residues 104-319) as template. The final *T. brucei* DOT1 homology models generated by Modeller covered residues 79-295 and 62-275 for DOT1A and DOT1B, respectively and have root mean square deviations (RMSD) of 1.36 and 1.04 Å with respect to the human DOT1L crystal structure (upon superposition of 212 α carbons in both cases) with good geometries as indicated by a validation using PROCHECK (Laskowski et al., 1993) (91.3/92.2% of DOT1A/DOT1B residues in most favored, 8.2/6.2% in additionally allowed, 0.5/1.0% in generously allowed and 0.0/0.5% in disallowed regions of the Ramachandran plot). All figures containing molecular structures were generated with UCSF Chimera (Pettersen et al., 2004).

3 Results

3.1 Reconstitution of trypanosomal nucleosomes

Recombinantly expressed histones do not contain any posttranslational modifications and thus can be obtained in a defined and pure form. In this work, the development of a trypanosomal nucleosome reconstitution system is prerequisite to characterize two methyltransferases DOT1A and DOT1B from *T. brucei* in vitro. Both enzymes methylate histone H3 on lysine 76 (H3K76) (Janzen et al., 2006b), which is located on the nucleosomal core surface and not in the histone tails where the majority of posttranslational modifications are placed. Methylation of H3K76 does neither work on purified H3, nor on histone H3 peptides. So far, in vitro studies of trypansosomal DOT1 enzymes were limited as no defined substrate exists and reconstitution of nucleosomes from trypanosomes remained unsuccessful. Although histones are in general highly conserved among different eukaryote species, the L1-loop of H3 containing the DOT1 target lysine K76 differs significantly in *T. brucei* (Frederiks et al., 2010; Janzen et al., 2006b) (Figure 6A). A closer examination reveals that aa differences extend beyond the H3 L1-loop and affect all four histones (Figure 6), resulting in a markedly different nucleosome surface in *T. brucei* (Figure 7). Moreover, it has been suggested that the human DOT1L enzyme binds to the nucleosome via interactions with all four core histones (Min et al., 2003), underlining the importance of a proper substrate for thorough enzymatic analysis of *T. brucei* DOT1 proteins.

3.1.1 Expression of trypanosomal histone proteins

Individual histones constructs (Table 6) were transformed into competent *E. coli* cells and full-length histone proteins were heterologously expressed. The intensity of histone expression has been found to depend on the *E. coli* strain type, cell density, temperature, culture volume, as well as IPTG and glucose concentration in the medium. Therefore, the optimal conditions for strongest protein expression were tested using different *E. coli* strains (BL21 cd (+), BL21 (DE3), Rosetta Blue (DE3), Rosetta (DE3) pLys). Strongest protein expression was achieved using BL21 (DE3), whereas only low protein yields were obtained with Rosetta Blue (DE3) cells (Figure 8) and all other strains tested (data not shown). Based on this result, all histone constructs were separately transformed into BL21 (DE3) cells and all parameters (cell density at time of induction, temperature, expression time) were kept constant (see section 2.3.3.3). Under those conditions, only histone H3 and H2A were expressed in a satisfactory manner, whereas initial attempts to express histones H2B, H4, H3V and H4V failed. Addition of glucose to the medium appeared to be beneficial for efficient histone expression and the use of buffled flasks, as well as the ratio of medium to flask volume turned out to be important for efficient expression (Table 7). Finally, all six full-length core histones and variants (H3, H4, H2A, H2B, H3V and H4V) were expressed and analyzed by SDS-PAGE (Figure 9A and B). After protein expression a two-step purification procedure that included the preparation of histone containing inclusion bodies (Figure 9C and data not shown) and ion exchange chromatography under denaturing conditions was performed. The collected fractions were analyzed by SDS-PAGE (Figure 9D and data not shown), pooled, dialyzed and lyophilized. This strategy yielded all histone proteins in a highly pure form.

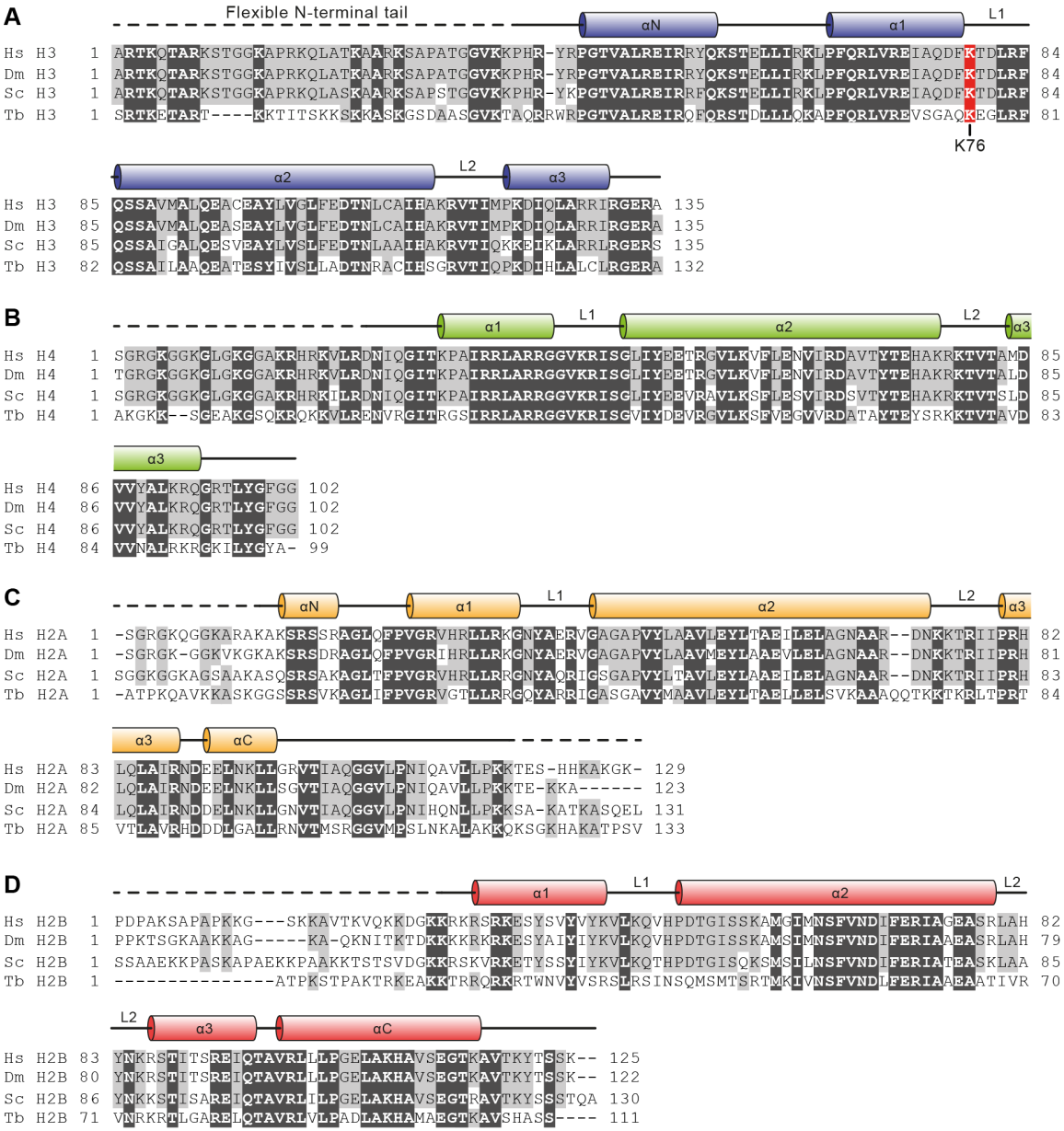


Figure 6. Sequence alignment of core histones

Sequence alignments of histone H3 (**A**), H4 (**B**), H2A (**C**) and H2B (**D**) from different species: Hs (*Homo sapiens*, gene identifiers (GI): 1568559 (H3), 119575932 (H4), 4504245 (H2A), 66912162 (H2B)), Dm (*Drosophila melanogaster*, GI: 24585677 (H3), 78707094 (H4), 78707122 (H2A), 24585671 (H2B)), Sc (*Saccharomyces cerevisiae*, GI: 6319482 (H3), 6319481 (H4), 398366187 (H2A), 6319471 (H2B)), and Tb (*Trypanosoma brucei*, GI: 115504299 (H3), 261328456 (H4), 72391222 (H2A), 71748422 (H2B)). Invariant amino acid positions are highlighted (white letters against black background), residues conserved in three organisms are shown with grey background. Secondary structure elements according to the yeast nucleosome crystal structure (Sawada et al., 2004) are indicated above the sequences. Flexible histone tails are displayed as dashed lines. Position of H3 lysine 76 (K76, *T. brucei* numbering) within loop 1 (L1) is shown in red. Sequence alignments were generated with Clustal Omega (Sievers et al., 2011).

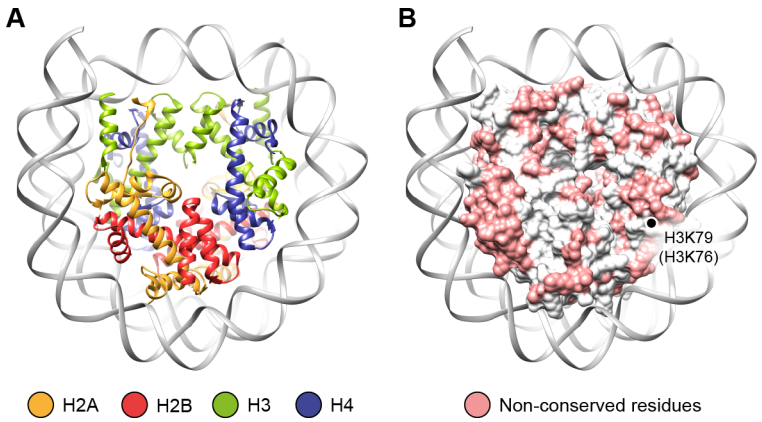


Figure 7. Nucleosome surface
(**A**) *S. cerevisiae* nucleosome structure (PDB accession code: 1ID3) (White et al., 2001), with individually colored histones as indicated below, DNA depicted in grey. (**B**) Surface representation of the yeast nucleosome with amino acid differences compared to *T. brucei* highlighted in pale red. Conserved residues and DNA are colored in grey. The *T. brucei* H3K76 equivalent position (H3K79 in yeast) is marked with a black dot.

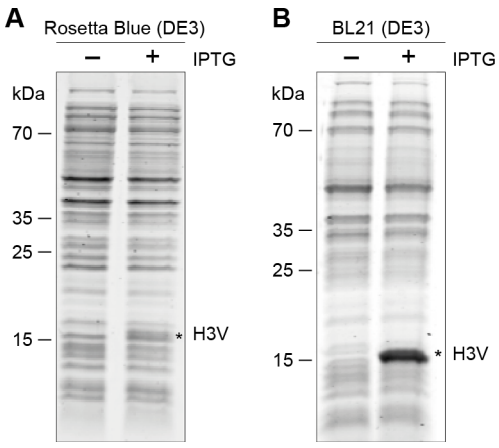


Figure 8. Histone expression using different *E. coli* strains
Representative expression of histone H3V (16 kDa), using Rosetta Blue (**A**) and BL21 (DE3) (**B**) competent cells. Bacterial whole cell extracts with and without IPTG induction for 3 h were analyzed by SDS-PAGE and coomassie staining. Full-length histone H3V proteins are marked with a black asterisk.

3.1.2 Nucleosome assembly

Histone octamers were refolded starting with lyophilized recombinant *T. brucei* full-length histone proteins. To this end, histone proteins (H3, H4, H2A, H2B) were separately dissolved under denaturing conditions and subsequently mixed in a molar ratio of 1 : 1 : 1.2 : 1.2 (H3 : H4 : H2A : H2B). The slight excess of H2A and H2B was included to ensure complete octamer formation. Refolding of histone octamer was achieved by slow removal of the denaturing compound (guanidinium chloride) during dialysis for several hours. After dialysis, concentrated octamer samples were applied on a size exclusion chromatography column to separate histone octamers from $(H3\text{-}H4)_2$ tetramers and (H2A-H2B) dimers (Figure 10A). The first peak in the size exclusion chromatogram is characteristic for higher molecular aggregates that formed during the refolding reaction. The relative large amount of aggregates suggests that the refolding reaction could

potentially be optimized e.g. by varying the ratios of H3/H4 to H2A/H2B in the refolding reaction. Octamer fractions were pooled and analyzed by SDS-PAGE (Figure 10A). Notably, histone H3 and H2A could not be separated on the gel due to almost identical molecular masses of 14.8 kDa (H3) and 14.2 kDa (H2A) and therefore run at the same position in the gel. However, the corresponding signal was twice as strong than those of histone H4 and H2B, which could be separated from each other (Figure 10A), suggesting that all four core histones are present in the octamer sample with equal stoichiometry. Histone octamers were further assembled into mono-nucleosomes by the addition of a specific 204 bp DNA fragment, the so called 601-sequence, which was found in a systematic evolution of ligands by exponential enrichment (SELEX) experiment in 1998 (Lowary and Widom, 1998) and possesses a high capacity to bind histone octamers at a single position with a defined conformation (Luger et al., 1999). A plasmid containing multiple repeats of the 601-sequence separated by a defined restriction site was digested until monomeric 601-sequence fragments were obtained (Figure 10B). Alternatively monomeric 601-sequences were amplified by PCR using specific primers (Table 3) (data not shown). In vitro reconstitution of nucleosomes relies on the assembly of histone H3/H4 tetramers combined with histone H2A/H2B dimers onto purified DNA. Assembly of this DNA-protein complex can be achieved through salt gradient deposition, which was successfully adapted for trypanosomal nucleosomes. Assembled nucleosomes were analyzed on native polyacrylamide gels and separated from unbound DNA (204 bp) and aggregates (Figure 10C).

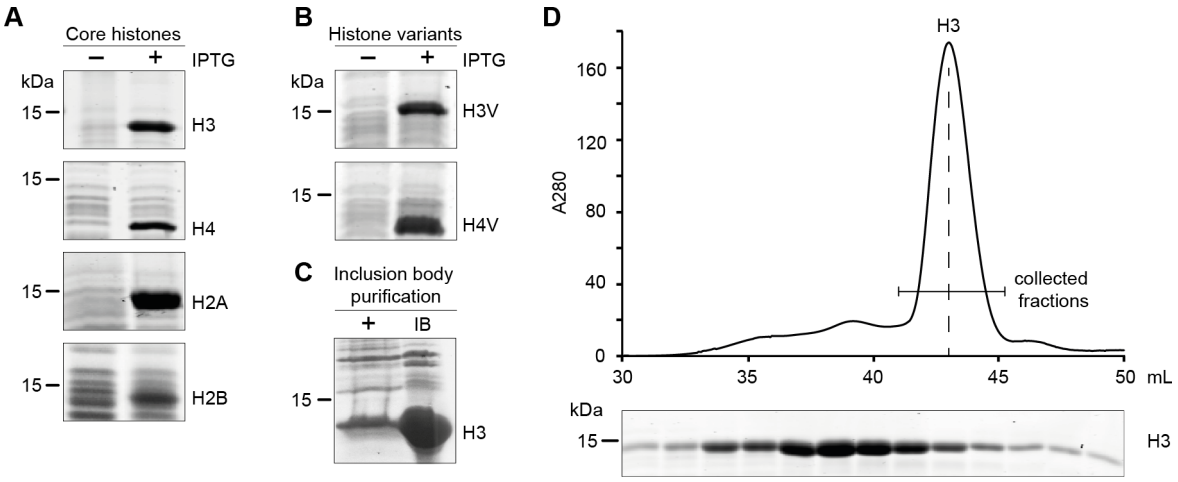


Figure 9. Histone expression and purification
T. brucei core histones (**A**) and histone variants (**B**) were individually expressed and analyzed by SDS-PAGE and coomassie staining (H3 (14.8 kDa), H2A (14.2 kDa), H2B (12.6 kDa), H4 (11.1 kDa), H3V (16 kDa), H4V (11.3 kDa). (**C**) Inclusion body purification of histone H3, analyzed as in (A). Results for the remaining histones are essentially identical. (**D**) Each histone inclusion body preparation was subjected to cation exchange chromatography under denaturing conditions to increase purity. Collected histone fractions (indicated as red line) were analyzed by SDS-PAGE and coomassie staining. A representative purification of histone H3 is shown. All other histone purifications gave essentially similar results.

In an alternative strategy to obtain large amounts of nucleosome core particles, all four histone proteins, together with DNA were mixed in one reaction, omitting the octamer assembly step described above (Lee and Narlikar, 2001). This strategy requires input of large amounts of histones and DNA, which are prone to aggregation during the salt dialysis step of the assembly protocol. However, the amount of precipitated material can be reduced by prolonged dialysis

over 50 h. After dialysis, the precipitates were separated by centrifugation and the supernatant was analyzed for presence of intact nucleosomes on a native polyacrylamide gel (Figure 11). Taken together, a trypanosomal nucleosome reconstitution system was successfully established, which yields mononucleosomes in a highly purified form.

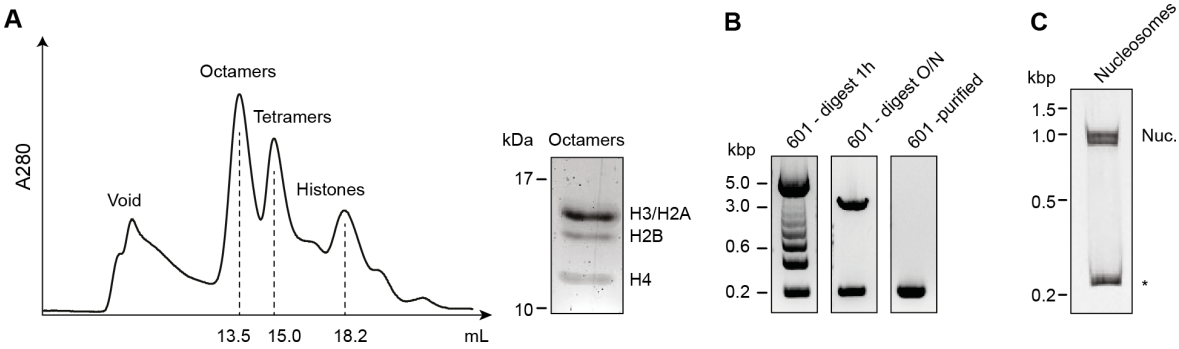


Figure 10. Nucleosome assembly
(**A**) Histone octamers were assembled from purified recombinant histones and subjected to size exclusion chromatography on a Superdex 200 HR 10/30 column to separate tetramers and free histones (left panel). The octamer fraction was separated on a 15 % polyacrylamide gel (right panel). Histones H3 and H2A can not be separated due to almost identical molecular masses (14.8 kDa and 14.2 kDa, respectively). (**B**) The pUC18 vector containing 12 x 601-sequence was digested with *Ava*I for 1 h at 37°C (left panel) and over night (O/N) until a single product band was obtained (middle panel). Monomeric 601-sequence fragments (204 bp) were separated by electrophoresis using a 0.8 % agarose gel and purified by gel extraction (right panel). (**C**) Nucleosome assembly was analyzed by gel mobility shift assay on a native 10 % (w/v) polyacrylamide gel. Nucleosomes (Nuc.) are separated from free DNA (marked with asterisk).

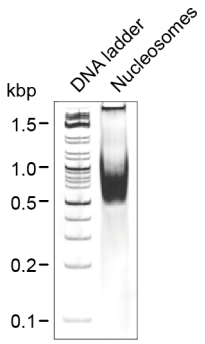


Figure 11. Large scale nucleosome assembly
Large-scale nucleosome assembly was analyzed by gel mobility shift assay on a 5 % (w/v) native polyacrylamide gel.

3.2 Asf1 histone chaperones

As an initial approach to determine the functionality of the assembled trypanosomal histones, H3/H4 complexes were used to test binding capacity of histone chaperones Asf1A and Asf1B under defined in vitro conditions. The anti-silencing function protein 1 (Asf1) is a molecular chaperone that binds to histones H3 and H4 heterodimers. Asf1 participates during several biological processes and facilitates H3/H4 dimer deposition and removal from chromatin during replication, transcription and DNA repair. Most eukaryotes possess two different Asf1 chaperones with potentially non-redundant functions. Interestingly, the two Asf1 chaperones that are found in *T. brucei* are more divergent than in other eukaryotes (Pascoalino et al., 2014). Asf1A is mainly localized in the cytosol but translocates to the nucleus during S phase of the cell cycle (Pascoalino et al., 2014). In contrast, Asf1B is predominantly localized in the nucleus (Pascoalino et al., 2014), and is proposed to be involved in the H3/H4 nuclear transport in human cells (Campos et al., 2010). To unravel possible different functions of the two Asf1 proteins in *T. brucei*, the trypanosomal nucleosome reconstitution system was used.

3.2.1 Purification of histone chaperones Asf1A and Asf1B

Full-length Asf1A and Asf1B proteins were cloned into pMAL-c2X expression vector (Table 6), followed by protein expression and purification from bacteria. Purified proteins were analyzed by SDS-PAGE and coomassie stain (Figure 12A and B). While robust amounts of monomeric Asf1B could be readily purified, Asf1A showed the tendency to form high molecular weight aggregates (Figure 12A and B). To solve this problem, different conditions during protein expression and purification were tested. These included different temperatures (10°C, 22°C, 30°C, 37°C), expression times (1 h, 2 h and 4 h), IPTG concentrations (0.001 mM, 0.01 mM, 0.1 mM and 1 mM) and varied salt concentrations. However, none of the tested conditions lead to a significant reduction of Asf1A aggregation (data not shown). Addition of benzylalcohol to the medium can induce an artificial stress response in *E. coli* by affecting membrane fluidity (Shigapova et al., 2005) and this effect has been utilized to induce cellular chaperones via the heat-shock response to improve heterologous protein expression and native folding (de Marco et al., 2005). The method turned out to be beneficial for Asf1A expression. Benzylacohol was added to the medium 30 min prior to Asf1A induction with IPTG and resulted in a significant decrease but not to a complete loss of Asf1A aggregates (Figure 12C). To further improve the Asf1A purification quality, the eluates were centrifuged through filters with a MWCO of 100 kDa, which effectively separated aggregates from full-length proteins (Figure 12C). Taken together, trypanosomal Asf1 proteins could be successfully purified and were further used to establish a histone pull-down assay to analyze the activity of Asf1A and Asf1B chaperones in vitro.

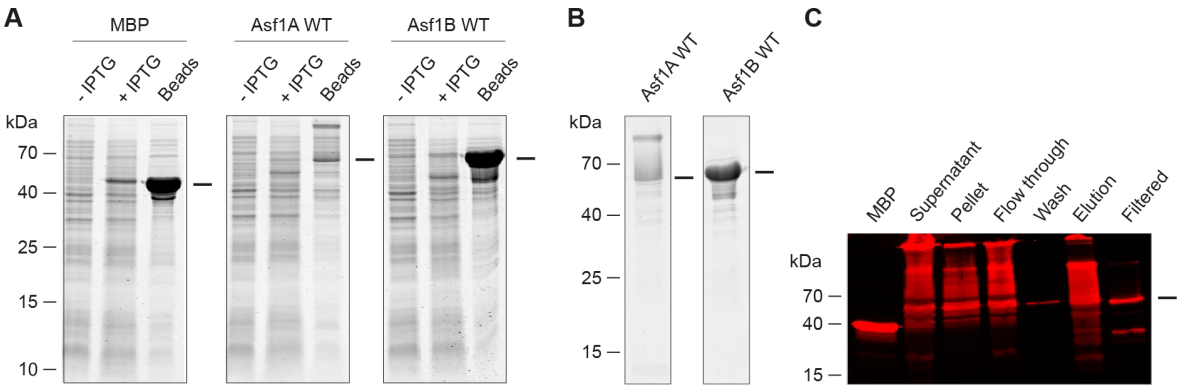


Figure 12. Bacterial expression and purification of *T. brucei* Asf1 chaperones
(**A**) MBP, Asf1A and Asf1B chaperones were expressed in *E. coli* (Rosetta Blue (DE3)). Protein expression was induced with 1 mM IPTG at a cell density of OD_{600} = 0.5-0.6. DOT1 proteins were expressed for 1 h at 37°C. Expression of proteins was analyzed on a 15 % SDS-PAGE and visualized with coomassie staining. MBP alone and MBP-Asf1 protein sizes are indicated with a line (MBP 42.5 kDa, MBP-Asf1A 61.62 kDa; MBP-Asf1B 67.25 kDa). MBP expression serves as a negative control. **(B)** Amylose resin bound MBP-tagged proteins were eluted with 10 mM maltose and analyzed by SDS-PAGE and visualized as in (A). Position of full-length MBP-Asf1 fusion proteins are indicated with a line. (**C**) Asf1A purification analyzed by western blot using a MBP specific antibody. Full-length MBP-Asf1A is indicated with a line.

3.2.2 Asf1A but not Asf1B binds to histone H3/H4 dimers in vitro

To elucidate direct interactions between histone heterodimers and Asf1 chaperones, a cell-free system based on recombinant proteins was chosen. Asf1A- or Asf1B-MBP fusion proteins were expressed in *E. coli*, purified and subsequently incubated with an equimolar amount (with respect to Asf1) of recombinant H3/H4 complexes. Possible interactions of Asf1 proteins with histones were analyzed by western blot using specific antibodies directed against MBP, H3 and H4 (Janzen et al., 2006b; Pascoalino et al., 2014; Siegel et al., 2009) after isolating the complexes via the MBP tag with amylose-coupled beads. Interestingly, only the cytosolic Asf1A but not the nuclear Asf1B could bind H3/H4 complexes in vitro (Figure 13A). To unravel potential binding preferences of both chaperones for histone variants, the experiment was repeated with heterodimer complexes of histone H3 and H4 variants (H3V/H4V). Similar to the results obtained for H3/H4 complexes, only Asf1A was able to bind recombinant histone variants (Figure 13B). Surprisingly, it was not possible to pull-down histone H4V in this assay. Histone H4V is probably lost in this assay because the variant dimer is less stable than the canonical H3/H4 complex as shown before for the H2BV/H2AZ variant nucleosome complex (Siegel et al., 2009). This result suggests a direct interaction of the cytosolic Asf1A with histone H3 or H3V (Natsume et al., 2007). The observation that Asf1B does not bind to recombinant histones in vitro at all might be due to the lack of posttranslational protein modifications on histones expressed in *E. coli*. To test the hypothesis that histone PTMs are a prerequisite for interaction with Asf1B, the pull-down assay was repeated with whole trypanosome cell extracts instead of purified recombinant histone proteins (Figure 13C). Interestingly, both chaperones could bind to histones from whole cell extracts with the same efficiency supporting the hypothesis that Asf1B binds specifically modified histone complexes in the nucleus.

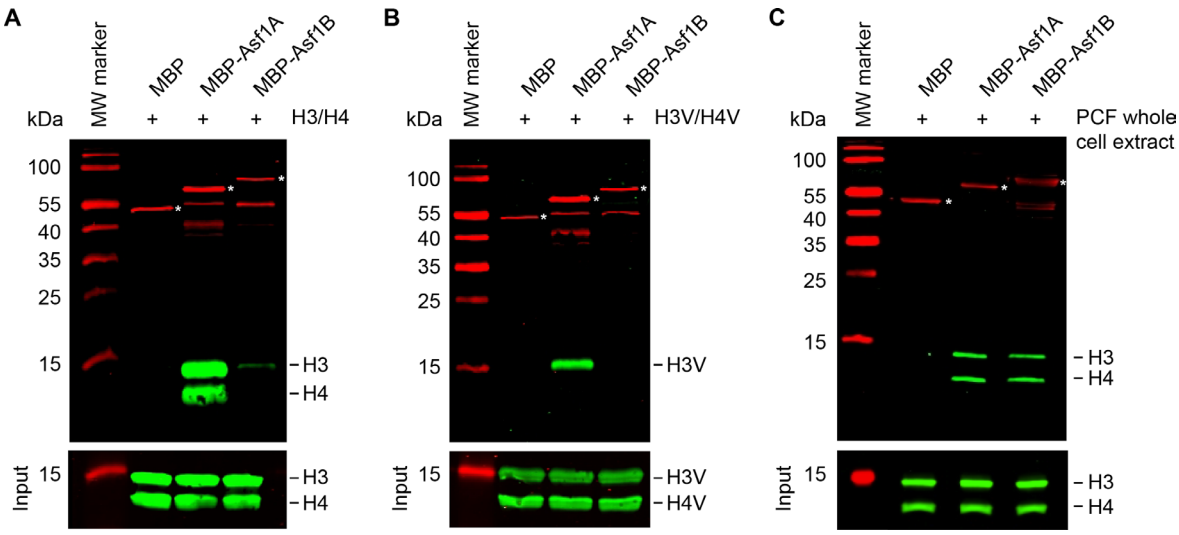


Figure 13. Asf1 pull-down assays

Cytosolic but not the nuclear Asf1 binds to histones in vitro. (**A**) Bacterial extracts containing MBP-tagged Asf1A or Asf1B were incubated with purified histone H3/H4 complexes. Protein interactions were tested by pull-down assays and monitored by western blot analysis (20 % of the total sample were loaded on each lane). H3/H4 input samples (2 %) are shown in the lower panel. MBP was used as control. Full-length fusion proteins are marked with asterisks (**B**) Interaction of MBP- tagged Asf1A and Asf1B with histone variant complexes (H3V/H4V) were analyzed as described in (A). (**C**) Purified MBP-tagged Asf1A or Asf1B were incubated with PCF trypanosome whole cell extracts from $3x10^7$cells for each pull-down. Protein interactions (20 % of the total sample) and input samples (5 %) were analyzed as described in (A).

3.3 Recombinant expression and purification of DOT1 enzymes

The pMAL-c2X expression vector was used to express and purify DOT1 methyltransferases from bacteria. This vector allows the insertion of the gene of interest downstream from the *mal*E gene from *E. coli*, which encodes the maltose-binding protein (MBP) and results in the expression of an N-terminal MBP fusion protein suitable for affinity purification. Full-length sequences of wild-type (WT) DOT1A and DOT1B and catalytic dead mutants (G138R and G121R, respectively) were cloned into pMAL-c2X expression vector (Table 6). High levels of MBP-fused DOT1 proteins could be expressed from *E. coli* (Figure 14A). Subsequent affinity purification relied on the tight binding of MBP to immobilized amylose on a column. MBP-fusion proteins were eluted from the affinity matrix by addition of free maltose, which is bound by MBP with high affinity. Purified samples were analyzed by SDS-PAGE and visualized with coomassie staining (Figure 14B). Attempts to store the active enzymes at different temperatures (4°C, -20°C or -80°C) in combination with different amounts of cryo-protecting glycerol (5 %, 10 %, 20 %) failed, resulting in either protein degradation and/or loss of activity (data not shown). Thus, fresh preparations of DOT1 proteins were used for all subsequent experiments. Nevertheless, *T. brucei* DOT1 proteins could be successfully purified and were further used to establish an in vitro histone methyltransferase assay to study the intrinsic activity of DOT1A and DOT1B enzymes.

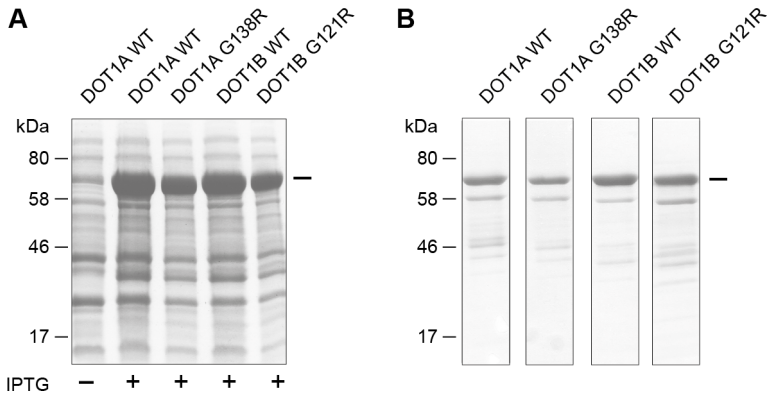


Figure 14. Recombinant expression and purification of *T. brucei* DOT1 proteins
(**A**) DOT1A and DOT1B methyltransferases and its catalytic inactive mutants (G138R and G121R, respectively) were expressed in *E. coli* (Rosetta Blue (DE3)). Protein expression was induced with 1 mM IPTG at a cell density of OD_{600} = 0.5-0.6. DOT1 proteins were expressed for 2.5 h at 28°C. Expression of proteins were analyzed on a 15 % SDS-PAGE and visualized by coomassie staining. Position of MBP-DOT1 proteins is indicated with a line (DOT1A and G138R, 75.9 kDa; DOT1B and G121R 73.4 kDa). The un-induced sample serves as a negative control. **(B)** Amylose resin bound MBP-tagged proteins were eluted with 10 mM maltose and analyzed by SDS-PAGE and visualized as in (A). Full-length MBP-DOT1 proteins are indicated with a line.

3.4 DOT1 histone methyltransferase assay

T. brucei histone H3K76 can be mono- (me1), di- (me2) and tri-methylated (me3) by DOT1 enzymes. To confirm that DOT1A and DOT1B methylate lysine 76 on histone H3 in vitro, the activities of both enzymes were tested on reconstituted mononucleosomes. As controls, the central glycine residue of the DxGxGxG signature motif was mutated to arginine (G138R and G121R for DOT1A and DOT1B, respectively), generating catalytically inactive enzymes (Frederiks et al., 2010; Janzen et al., 2006b). A histone methyltransferase assay was performed with freshly purified full-length DOT1 enzymes and nucleosomes. The H3K76 methylation levels (me1, me2, me3) were detected by western blot analysis using specific antibodies directed against H3K76me1/-me2/me3 and H3 (Gassen et al., 2012; Janzen et al., 2006b). Consistent with previous observations (Janzen et al., 2006b), DOT1A WT, but not its catalytically inactive G138R mutant, mono- and di-methylates H3K76 (Figure 15). In contrast, the DOT1B WT enzyme is able to catalyze all three methylation states on H3K76. These results provided new insights, since previous work showed that the DOT1B enzyme is only able to convert di-methylated nucleosomes to tri-methylated state ex vivo (Janzen et al., 2006b). The catalytic inactive mutants appear to possess weak residual mono- and di-methylation activity (Figure 15). However, this could be an artifact, since the negative control of non-methylated nucleosomes also shows very weak me1 and me2 signal. Moreover, methylation activity of both DOT1 enzymes was tested on purified histone H3, histone tetramers $(H3\text{-}H4)_2$ and octamers (data not shown). The results confirmed that DOT1 proteins exclusively methylate H3K76 only in a nucleosomal context. Further, they indicate that both enzymes are not dependent on additional histone PTMs or other regulatory factors (e.g. adaptor proteins). However, it cannot be excluded that such factors render trypanosomal DOT1 enzymes more efficient in vivo. To summarize, an in vitro trypanosomal nucleosome-based

methylation assay was successfully established to investigate the distinct product specificities of DOT1 enzymes.

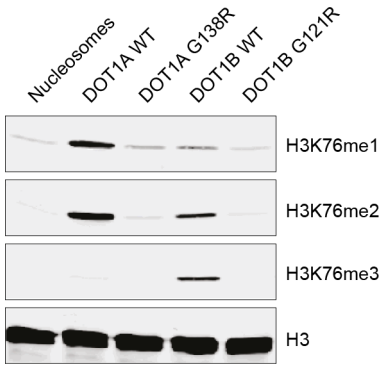


Figure 15. Histone methyltransferase assay
Enzymatic activity of purified *T. brucei* wild-type DOT1A and DOT1B and catalytically inactive mutants (DOT1A G138R and DOT1B G121R) were tested on reconstituted nucleosomes in vitro. As a negative control reconstituted nucleosomes without enzyme was loaded. After western blotting, different methylation states were detected with specific antibodies. H3 signals serve as loading control.

3.5 DOT1A and DOT1B are distributive enzymes in vitro

DOT1 enzymes can work either in a processive or in a distributive manner (Figure 16). Processive enzymes set multiple modifications on the substrate lysine without dissociation and therefore intermediate methylation states might not be observed (Figure 16A). In contrast, distributive enzymes are only able to transfer one methyl group to the substrate per binding event, resulting in detectable intermediates (Figure 16B). To determine the catalytic mode, a methyltransferase assay was performed with both trypanosomal DOT1 proteins using increasing enzyme concentrations for a fixed time window as previously described for yeast Dot1p (Frederiks et al., 2008). With a constant assay time, the two possible methylation patterns are characteristic for the enzyme's mode of action. Increasing amounts of a processive enzyme produce increasing amounts of methylated H3K76 but the relative abundance of the different methylation states should be invariant in the case of a slow enzyme (Figure 16A). For a fast processive enzyme the intermediate methylation states might not be observed. In contrast, a distributive enzyme is expected to first mono-methylate several lysines before it re-associates with mono-methylated lysines to introduce a second methyl group. Similarly, an accumulation of H3K76 di-methylation should be observed before tri-methylated lysine is detectable. Therefore, the relative abundance of the three different methylation states should change with increasing amount of distributive enzyme (Figure 16B). In the assay, low concentrations of DOT1A lead to a weak H3K76 mono-methylation signal, whereas increasing amounts result in the accumulation of me1, which is a prerequisite for the transition to me2 (Figure 17A). The independent H3K76me1 and -me2 product waves, hence, the different ratios of methylations states are characteristic for a distributive enzyme. Similar results were obtained for DOT1B, where H3K76me1, -me2 and -me3 products follow each other in consecutive waves with increasing enzyme concentrations (Figure 17B). Based on these results it can be concluded that both trypanosomal DOT1 KMTases

work in a distributive mechanism to methylate H3K76 in vitro. Taken together, the histone methyltransferase assay with purified components successfully served as a tool to reveal the mode of catalytic action of DOT1A and DOT1B under in vitro conditions. The data clearly demonstrates that both enzymes are distributive enzymes.

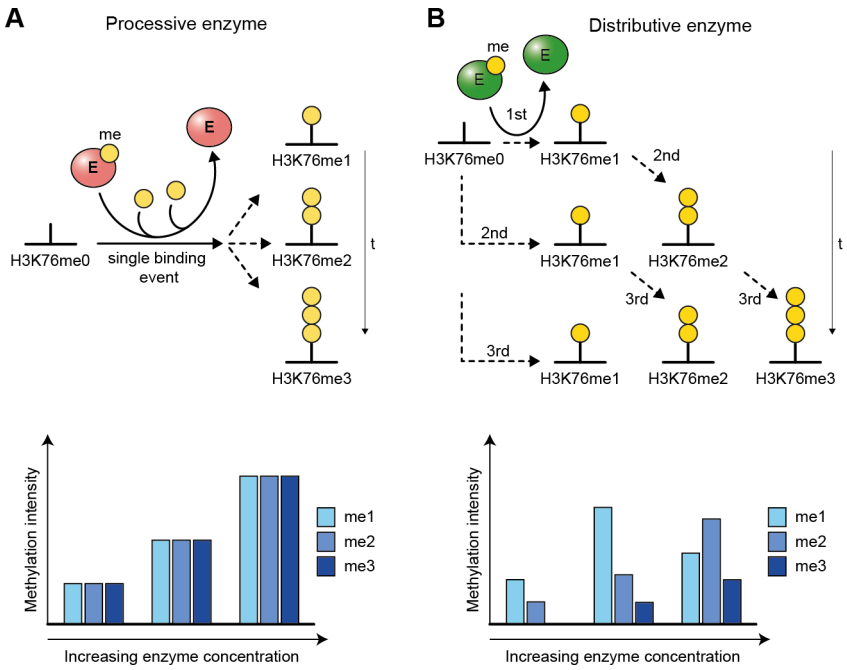


Figure 16. Schematic mechanism of processive and distributive enzymes
(**A**) Processive enzymes (E) catalyze multiple rounds of methylation without dissociating from the substrate (upper panel). With a constant incubation time increasing amounts of a processive enzyme are predicted to produce different amounts of only one methylation state (for a fast enzyme) or to produce methylation patterns in which invariant relative abundances of the different methylation states are observed (in case of a slow enzyme) (lower panel). (**B**) Distributive enzymes (E) dissociate and re-associate between individual methylation reactions to allow cofactor exchange (upper panel). With a constant incubation time increasing amounts of a distributive enzyme result in methylation patterns, in which the relative abundance of each methylation state depends on the amount of enzyme (lower panel).

3.6 DOT1 homology models

Compared to the human DOT1L enzyme, the catalytic cores of DOT1A and DOT1B share 20 % and 22 % sequence identity and 33 % and 32 % similarity, respectively (Figure 18). Based on this degree of conservation, homology models of the *T. brucei* proteins using the human DOT1L crystal structure (Min et al., 2003) as template were generated (for details on the modeling procedure see section 2.6). The models include residues 79-295 and 62-275 for DOT1A and DOT1B, respectively (Figure 19A) and cover the C-terminal open α/β domain, formed by seven β-strands (β5 to β11) that are flanked by five α-helices (αF to αJ) (Figure 19B and C). This C-terminal part of the KMTase domain is linked to N-terminal helices (αD and αE) via the loop-EF (Figure 18 and 19). The homology models end with helix αD, because no reliable template structures for the N-terminal part beyond helix αD are available at present.

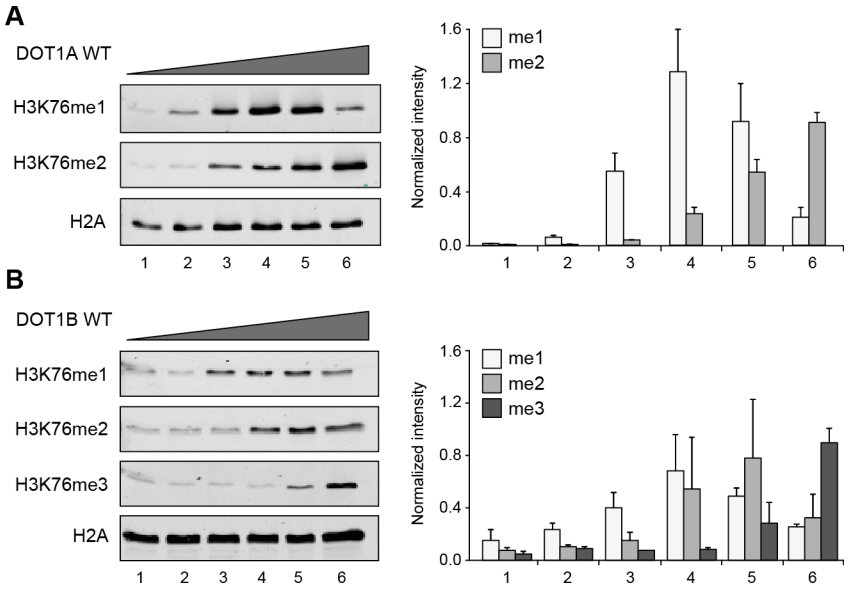


Figure 17. DOT1A and DOT1B are distributive enzymes

DOT1A and DOT1B are distributive enzymes *in vitro*. Histone methyltransferase assays were carried out for DOT1A (**A**) and DOT1B (**B**) with increasing enzyme concentrations (0.1-100 ng/µl, samples 1-6) and a fixed time of 2 h. Methylation states were analyzed by western blot (left) and quantified to allow statistical evaluation (right). Error bars indicate standard deviation of three independent experiments.

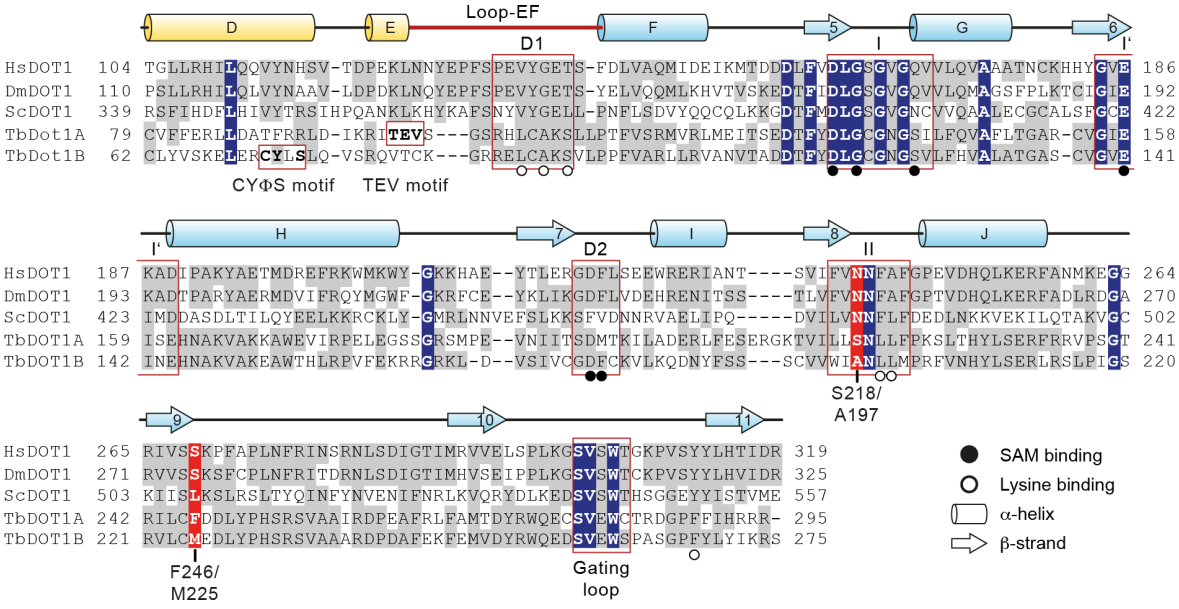


Figure 18. DOT1A and DOT1B sequence alignment of the catalytic domain

Structure-guided sequence alignment of the conserved DOT1 methyltransferase domain from Hs (*Homo sapiens*, gene identifier (GI): 22094135) (Min et al., 2003), Dm (*Drosophila melanogaster*, GI: 320542472), Sc (*Saccharomyces cerevisiae*, GI: 6320648) (Sawada et al., 2004) and Tb (*T. brucei*, GI: 72392449 (DOT1A), GI: 115503935 (DOT1B)). Invariant amino acid positions are highlighted (white letter against dark blue background), conserved residues are indicated with grey background. Secondary structure elements (α-helices D to J and β-strands 5 to 11) of the human enzyme are shown (Min et al., 2003). Conserved DOT1 motifs (I, I', II, D1, D2 and gating loop) are indicated. Open and closed circles mark residues implicated in lysine- and SAM-binding, respectively. Positions of amino acid exchanges (S218, F246 (DOT1A) and A197, M225 (DOT1B)) are highlighted in red.

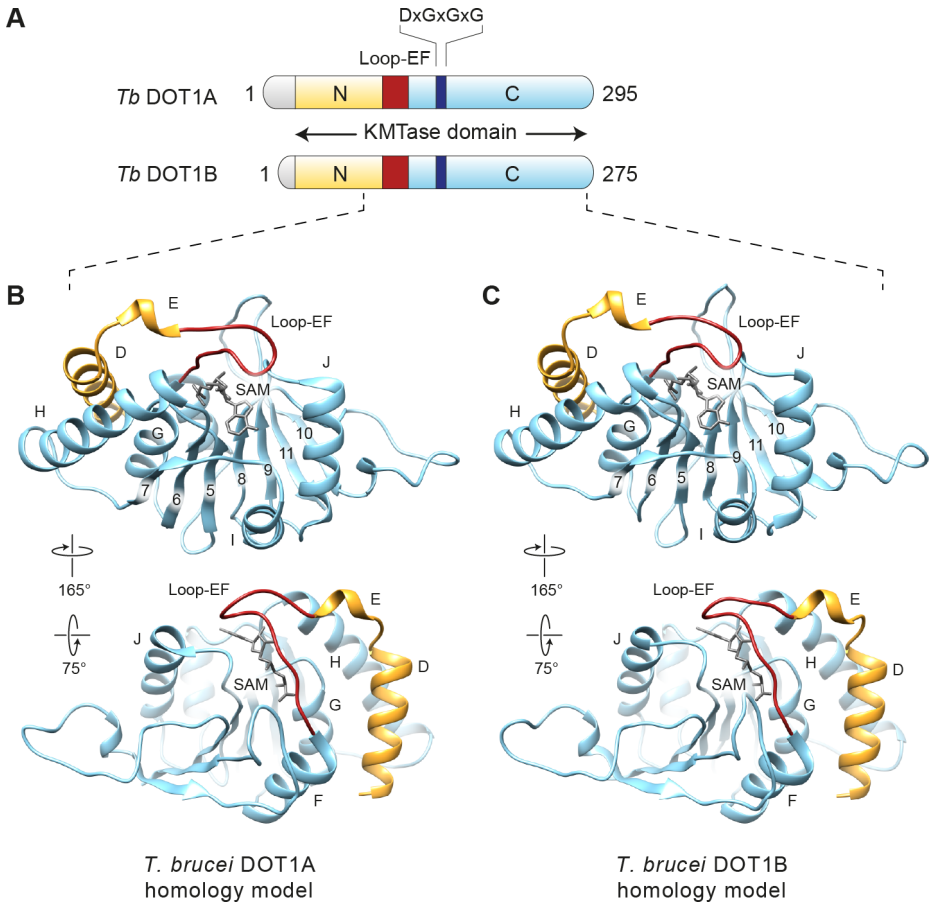


Figure 19. Homology modeling of DOT1A and DOT1B
(**A**) Schematic alignment of DOT1A and DOT1B. The conserved lysine methyltransferase (KMTase) core can be subdivided into an N- and C-terminal part (orange and light blue, respectively) connected by loop-EF (red). The DxGxGxG signature motif (dark blue) is essential for enzymatic activity. Homology model of the DOT1A (**B**) DOT1B (**C**) catalytic core domain viewed from the entrances of the SAM- (upper panels) and lysine-binding pockets (lower panels). Model is colored as in (A). Secondary structure elements are labeled (α-helices D to J and β-strands 5 to 11) according to (Min et al., 2003).

KMTases can be divided into two major classes, SET and non-SET enzymes. Non-SET domain methyltransferases like DOT1 are characterized by a series of conserved motifs (I, I' and II) (Dlakić, 2001; Schubert et al., 2003) and the corresponding regions are also present in *T. brucei* DOT1A and DOT1B (Figure 18 and Table 8). Two motifs, D1 and D2 appear to be specific for the DOT1 family of methyltransferases (Min et al., 2003). While the D2 motif is present but less conserved in *T. brucei* DOT1A and DOT1B, it has been proposed that the D1 motif is absent from trypanosomal DOT1 enzymes (Janzen et al., 2006b). Residues from all motifs (I, I', II, D1 and D2) are involved in co-factor binding and/or in formation of the target lysine-binding channel of DOT1 proteins (Min et al., 2003; Sawada et al., 2004) (Figure 18 and Table 8). The D1 motif forms a significant part of the substrate-binding pocket (Min et al., 2003; Sawada et al., 2004), which raises the question which part of the trypanosomal DOT1 enzymes could take over this function. Alignments allowed the identification of a sequence (CAKS, single letter aa code) that is nearly invariant across trypanosomes, suggesting conserved functional importance (Figure 20). In the *T. brucei* DOT1 homology models, the CAKS sequence was aligned under the D1 motif of other eukaryotes (YGET sequence) (Figure 18) based on the following structural arguments: (I) The small

alanine side chain (A109 (DOT1A), A92 (DOT1B)) is positioned where other eukaryotic DOT1 enzymes have a glycine residue that has been proposed to be critical to maintain an open lysine-binding channel (Min et al., 2003; Sawada et al., 2004) (Figure 18 and Table 8); (II) Placement of the lysine residue (K110 (DOT1A), K93 (DOT1B)) positions the positively charged side chain away from the substrate-binding channel that otherwise might cause electrostatic repulsion of the target-lysine (Figure 18); (III) The serine side chain (S111 (DOT1A) and S94 (DOT1B)) can stabilize the SAM carboxyl group via a hydrogen bond according to the function of threonine at this position in human DOT1L (Min et al., 2003) (Figure 18 and Table 8). Taken together, we propose that the CAKS sequence forms the trypanosomal D1 motif that generates part of the substrate-binding channel (Figure 18, 20 and Table 8).

Table 8. Functionally important residues of DOT1 methyltransferases

	H. sapiens DOT1L	*S. cerevisiae* Dot1p	*T. brucei* DOT1A	*T. brucei* DOT1B	Function
Motif D1 (loop-EF)	V135	V371	L107	L90	Lysine-binding pocket
	G137	G373	A109	A92	Lysine-binding pocket
	T139	L375	S111	S94	SAM-binding, lysine-binding pocket
Motif I	D161	D397	D134	D117	SAM-binding
	G163	G399	G136	G119	SAM-binding
	Q168	N404	S141	S124	SAM-binding
Motif I'	E186	E422	E158	E141	SAM-binding
Motif D2	D222	F460	D195	D177	SAM-binding
	F223	V461	M196	F178	SAM-binding
Motif II (loop-8/J)	N241	N479	S218*	A197*	Lysine-binding pocket
	F243	F481	L220	L199	Lysine-binding pocket
	A244	L482	L221	L200	Lysine-binding pocket
Loop-9/10	S269	L507	F246*	M225*	Lysine-binding pocket
Loop-10/11	Y312	Y550	F289	F268	Lysine-binding pocket

* Mutation target in this study.

3.7 Structure-guided mutations change the product-specificity of DOT1 enzymes

Several conserved amino acid residues of DOT1 are involved in cofactor SAM binding and some of them also contribute to the proposed lysine-binding pocket of the enzyme (Min et al., 2003). A systematic comparison including the available yeast and human DOT1 crystal structures (Min et al., 2003; Sawada et al., 2004) allowed the identification of eight residues that form a putative substrate-binding channel in *T. brucei* DOT1A (S111, S218, F246, F289, L107, A109, L220, L221) and DOT1B (S94, A197, M225, F268, L90, A92, L199, L200) that directly leads to the SAM methyl group as it has been proposed for yeast and human DOT1 proteins (Min et al., 2003; Sawada et al., 2004) (Figure 18, 20C and Table 8). Among these eight aa, two differ between both enzymes and are located between loop-8/9 (S218 (DOT1A), A197 (DOT1B)) and loop-9/10 (F246 (DOT1A), M225 (DOT1B)) (Figure 21 and Table 8). These two positions were exchanged in DOT1A with the corresponding residues of DOT1B to test the hypothesis that these aa positions are responsible for the different product specificities. The resulting mutant enzymes (DOT1A S218A

and F246M) were purified and tested in the in vitro methyltransferase assay. Samples were analyzed by SDS-PAGE and western blot and H3K76 methylation signals were detected with specified antibodies (Figure 21D). The results clearly demonstrate that an exchange of DOT1A serine 218 to alanine (DOT1A S218A) and DOT1A phenylalanine 246 to methionine (DOT1A F246M) are sufficient to confer DOT1B-specific tri-methylation activity to DOT1A (Figure 21D). Combination of both aa exchanges in a double mutant (DOT1A S218A/F246M) further increased the me3 signal to almost DOT1B WT level under these assay conditions (Figure 21D). Taken together, these results provide experimental support for the proposed structure of the substrate-lysine channel (Min et al., 2003; Sawada et al., 2004). Furthermore, they establish a novel function for two aa positions in DOT1 enzymes in determining the distinct methylation product-specificity.

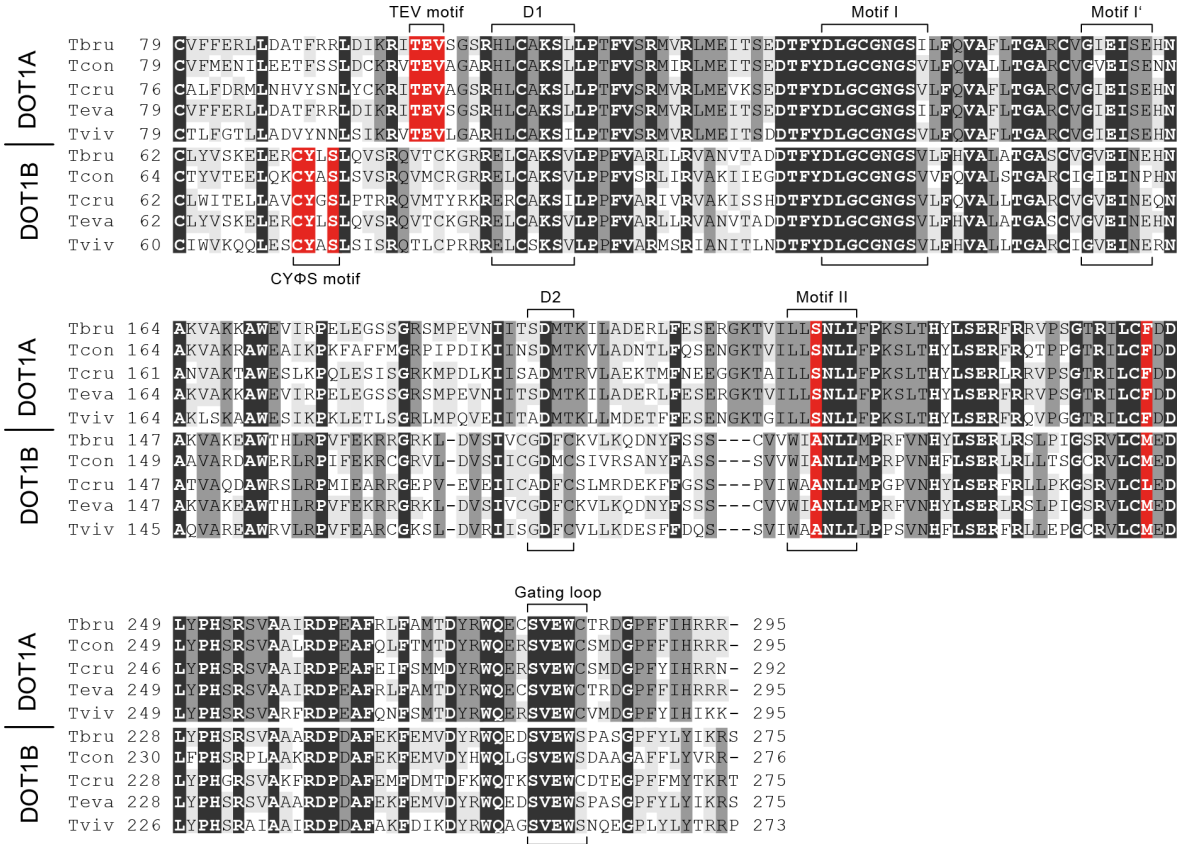


Figure 20. Sequence alignment of trypanosomal DOT1 enzymes

Sequence alignment of DOT1A and DOT1B methyltransferases from different trypanosomes: Tbru (*T. brucei*, TriTryp database IDs Tb927.8.1920 (DOT1A), Tb927.1.570 (DOT1B)) Tcon (*Trypanosoma congolense*, TriTryp database IDs TcIL3000_8_1870 (DOT1A), TcIL3000_1_140 (DOT1B)), Tcru (*Trypanosoma cruzi*, TriTryp database IDs TCSYLVIO_009826 (DOT1A), TCSYLVIO_007232 (DOT1B)), Teva (*Trypanosoma evansi*, TriTryp database IDs TevSTIB805.8.1850 (DOT1A), TevSTIB805.1.510 (DOT1B)) and Tviv (*Trypanosoma vivax*, TriTryp database IDs TvY486_0801310 (DOT1A), TvY486_0100040 (DOT1B)). Invariant amino acids in all DOT1 enzymes are highlighted with white letters against black background. Invariant residues between all DOT1A or DOT1B enzymes are indicated in dark grey and conserved residues shown with light grey backround. Locations of conserved DOT1 motifs are labeled. Positions of *T. brucei* DOT1 point mutations used in this study are highlighted in red.

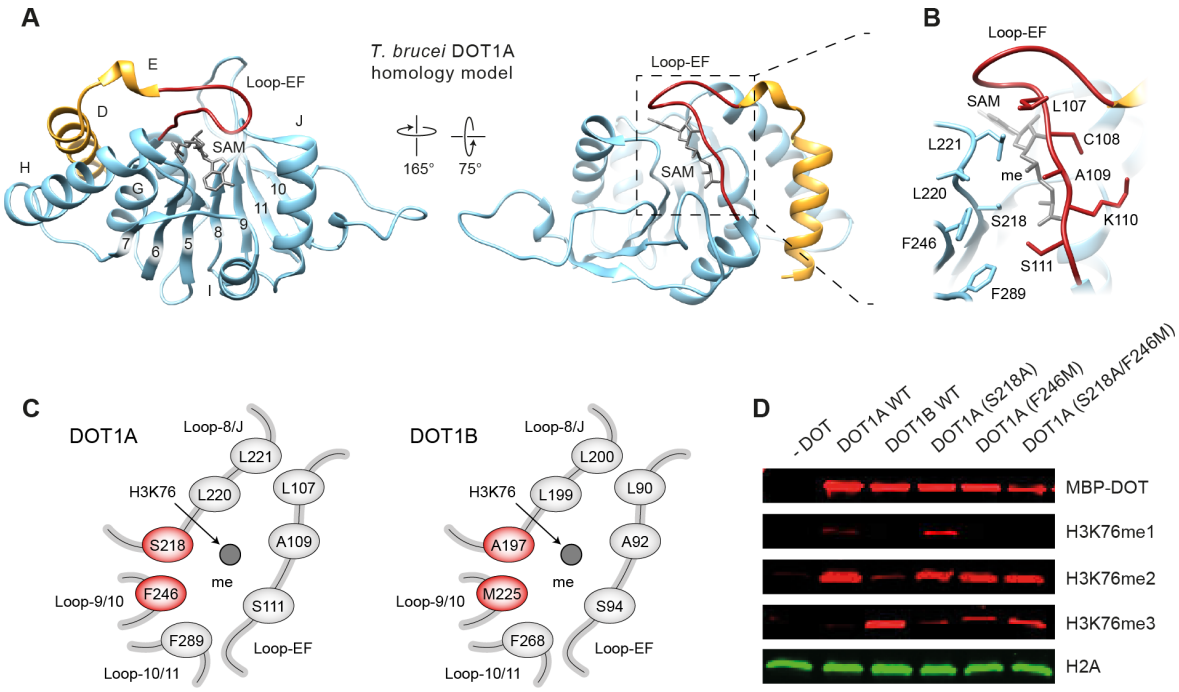


Figure 21. Structure-guided mutations identify residues within the catalytic center that influence product-specificity of DOT1A and DOT1B

(**A**) Homology model of the DOT1A catalytic core domain viewed from the entrances of the SAM- (left panel) and lysine- (right panel) binding pockets. N- and C-terminal domains of the catalytic core are colored in orange and blue, respectively. The connecting loop-EF is highlighted in red. (**B**) Close-up view of the lysine-binding pocket of DOT1A. (**C**) Schematic representation of the amino acids forming the H3K76-binding channel in *T. brucei* DOT1A (left) and DOT1B (right) leading to the methyl group (me) of SAM. Residues that differ between the two enzymes are highlighted in red. (**D**) Enzymatic activity of purified *T. brucei* DOT1A mutants (S218A, F246M and S218A/F246M) were tested on reconstituted nucleosomes *in vitro*. H2A signals serve as loading control.

Interestingly, DOT1A mutants containing the F246M exchange showed no H3K76me1 signal in comparison to DOT1A WT and the S218A mutant enzyme, indicating that the F246M mutation might selectively affect the reaction rate of H3K76 di-methylation (Figure 21D). Alternatively, the DOT1A F246M mutation could change the enzyme's catalytic mode from distributive to processive and such a change in principle could account for the observed methylation pattern (Figure 17A). To test the latter possibility, a methyltransferase assay was performed with increasing enzyme concentration for a fixed period of time as described above (see section 3.5). In this assay, low concentrations of DOT1A F246M lead to a weak H3K76 mono-methylation signal, whereas increasing amounts of the enzyme result in the accumulation of me1, which is a prerequisite for the transition to me2 (Figure 22). The independent me1 and me2 product waves are typical for a distributive enzyme and are comparable to the results obtained for DOT1A and DOT1B WT enzymes in the same experimental setup (Figure 17 and 22). Thus it can be concluded that the DOT1A F246M mutation does not change the enzyme's catalytic mode of action and the lack of me1 signal is rather caused by a selective modulation of the di-methylation reaction rate. To further explore this possiblility, H3K76me1 and me2 accumulation was monitored for DOT1A WT and the F246M mutant over time (Figure 23A and B). Interestingly, while the conversion of me0 to me1 showed no differences, the subsequent reaction to produce me2 appears to proceed significantly faster in the presence of the DOT1A F246M mutant (Figure 23D)

indicating that the phenylalanine at position 246 is responsible for selectively slowing down the reaction rate for H3K76 di-methylation by DOT1A. This in agreement with the in vitro results obtained for DOT1B WT, which contains a methionine at the equivalent position and appears to convert me1 faster to higher methylation states in comparison to DOT1A (Figure 15 and 22D). Importantly, end point samples of the reactions after 80 min incubation showed no H3K76me3 signals (Figure 23C), indicating that the observed course of me2 product formation is not influenced by the subsequent tri-methylation reaction catalyzed by the DOT1A F246M mutant (Figure 22D). In conclusion, these results establish a novel function for F246 in selectively modulating the dimethylation reaction rate and suggest that aa identity at this position is critical for determining the enzyme's speed in conversion of H3K76 me1 to higher methylation states.

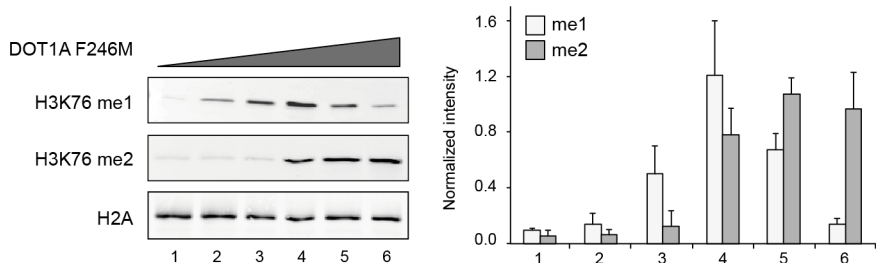


Figure 22. The DOT1A F246M mutant is a distributive enzyme in vitro
Histone methyltransferase assays was carried out for DOT1A F246M with increasing enzyme concentrations (0.1-100 ng/µl, samples 1-6) and a fixed time of 2 h. Methylation states were analyzed by western blot (left) and quantified to allow statistical evaluation (right). Error bars indicate standard deviation of three independent experiments.

3.8 The N-terminal region contributes to formation of the lysine-binding channel

Since the H3K76me3 signal generated by the DOT1A S218A/F246M double mutant did not reach DOT1B WT level (Figure 22D), the question arose, what additional components outside of the lysine-binding channel could influence the catalytic activity of the enzyme. Yeast and human DOT1 crystal structures revealed that the N- and C-terminal parts of the KMTase domain are in close vicinity to each other and that the interface appears to stabilize the lysine-binding pocket (Min et al., 2003; Sawada et al., 2004). Notably, the corresponding N-terminal parts of *T. brucei* DOT1A and DOT1B differ significantly between each other (Figures 18, 20 and 24A), possibly affecting the substrate-binding site geometries that contribute to the distinct product specificities. To test this hypothesis, the N-terminal part of DOT1A (residues 1-102) was exchanged with the corresponding sequence from DOT1B (residues 1-85) in WT and S218A/F246M mutant background. The chimeric WT enzyme produced only mono-methylated H3K76 and all chimeric mutants were catalytically inactive (Figure 24B). Importantly, the enzyme chimeras are very likely properly folded since me1 activity is maintained in the chimeric DOT1A WT protein (Figure 24B). These results indicate a contribution of the DOT1 N-terminus to the formation of the catalytic active site of the enzyme. However, exchanging the entire N-terminal region probably affects a wide range of contacts with the conserved C-terminal part (e.g.

interactions of helices αD and αH (Figure 18) and thus could explain the catalytically inactive mutant DOT1A chimeras (Figure 24B). This suggests that the enzyme's N-terminal part contributes to lysine binding channel formation via a small set of residues at the interface between the N- and C-terminal parts of the KMTase domain.

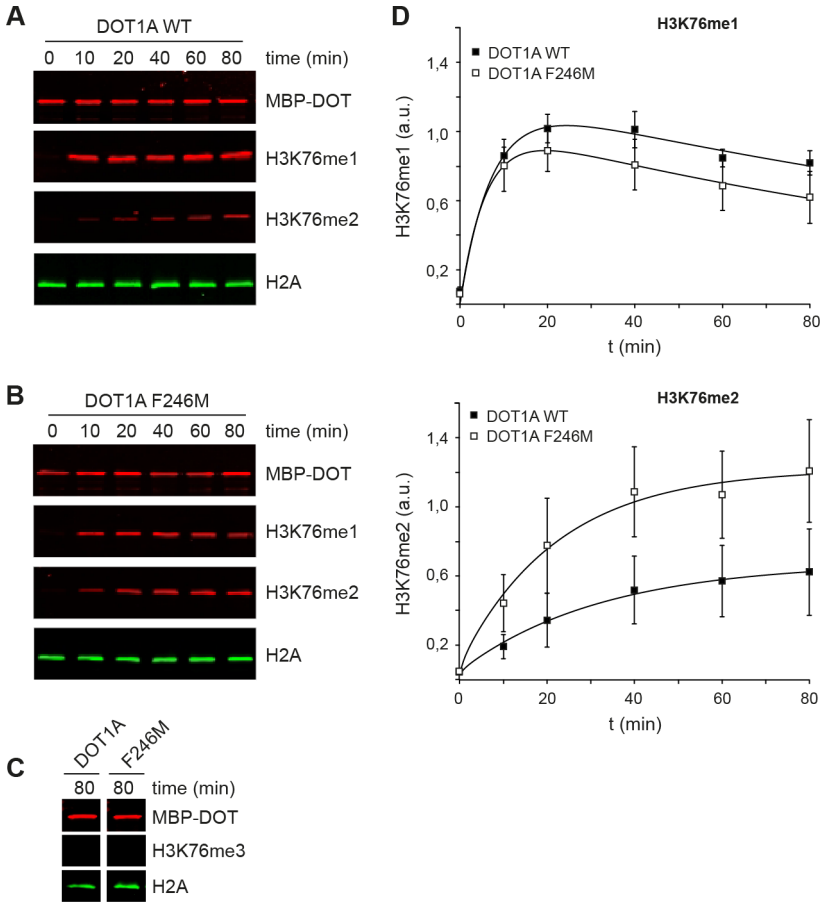


Figure 23. The DOT1A F246M mutation selectively alters the reaction rate for H3K76 di-methylation.
Histone methyltransferase assay was carried out for DOT1A WT (**A**) and DOT1A F246M (**B**) using constant amounts of enzyme and substrate. Aliquots were taken at different time points and H3K76me1 and me2 levels were analyzed by western blot. (**C**) End point samples of the reactions after 80 min incubation showed no H3K76me3 signals. Blot signals were quantified in Image Studio (LI-COR), normalized to the H2A signal and least square curve fits were determined in KaleidaGraph (Synergy Software). Error bars indicate standard deviation of three independent experiments.

Sequence alignment of several putative DOT1A and DOT1B sequences from various trypanosome species revealed two conserved residue stretches located N-terminally of the D1 motif. These motifs appear to be specific for either DOT1A (TEV sequence) or DOT1B (CYΦS sequence, where Φ represents hydrophobic aa) (Figure 18, 20 and 24A). Interestingly, the trypanosomal DOT1 homology models indicate that both of these stretches are pointing towards the common D1 motif (CAKS sequence) within loop-EF, which forms one side of the target lysine-binding channel (Figures 18 and 24A). The equivalent tyrosine of the CYΦS motif is conserved in yeast, fly and human DOT1 enzymes (Figure 18 and Table 9), where it participates in a hydrogen bond network to stabilize the D1 motif (Min et al., 2003; Sawada et al., 2004).

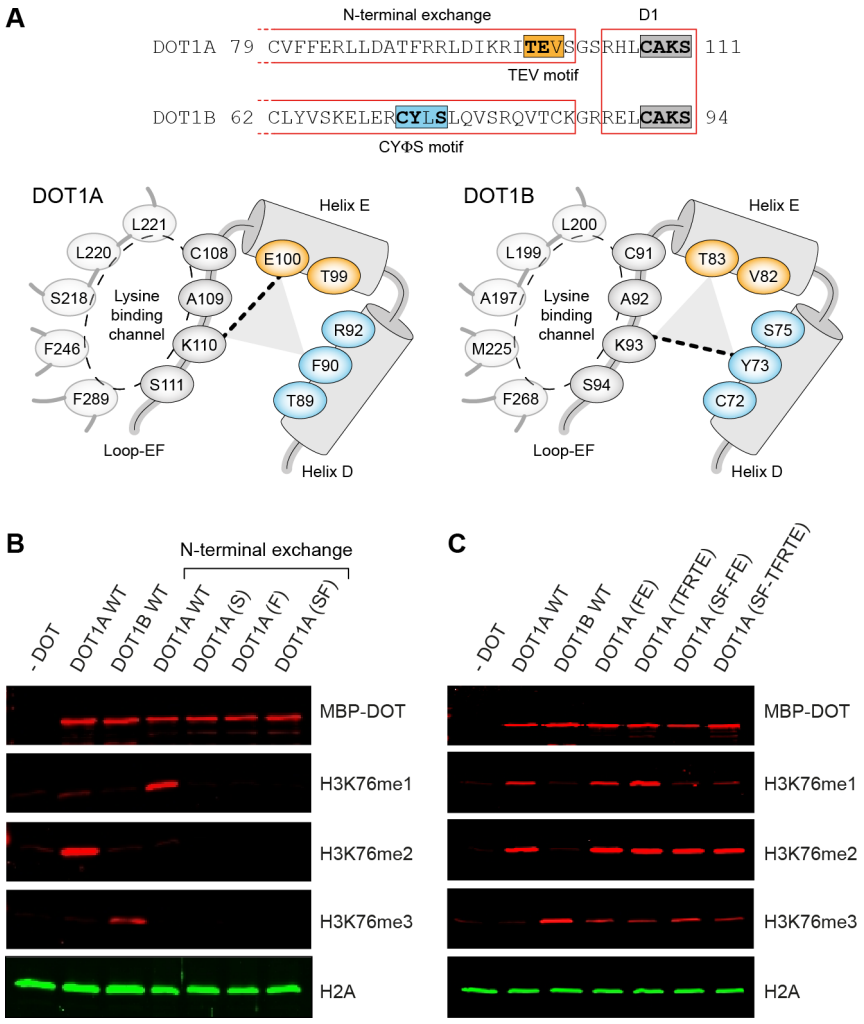


Figure 24. The N-terminal part of the DOT1 KMTase domain contributes to lysine-binding pocket formation.

(**A**) Amino acid sequence (upper panel) and schematic representation of the N-terminal regions with signature motifs (lower panel) of DOT1A and DOT1B enzymes. Highlighted are: CAKS (grey), TEV (orange) and CYΦS (blue). Range of the D1 motif and N-terminal exchange are marked with red boxes. The proposed differences in stabilization of the lysine-binding pocket by hydrogen bonds in DOT1A and DOT1B are indicated with dashed lines. (**B**) Enzymatic activity of purified DOT1A mutants (S218A, F246M and S218A/F246M) with residues 1-102 exchanged with the corresponding sequence from DOT1B (residues 1-85) was tested on reconstituted nucleosomes. H2A signals serve as loading control. (**C**) Activity of DOT1A double and quintuple mutants (DOT1A E100T/F90Y (DOT1A FE) and DOT1A T89C/F90Y/R92S/T99V/E100T (DOT1A TFRTE), respectively), in WT and S218A/F246M mutant background, were tested on reconstituted nucleosomes. H2A signals serve as loading control.

It can be speculated that the trypanosomal N-terminal sequence motifs are used to alter the lysine-binding channel geometry via different hydrogen bond interactions with the common central lysine residue (K110 in DOT1A and K93 in DOT1B) of the trypanosome D1 motif (Figure 21B). To test this hypothesis, the possible hydrogen bonding positions in DOT1A (F90/E100) were exchanged with the corresponding residues of DOT1B (Y73/T83) in WT and S218A/F246M mutant background (DOT1A (FE) and DOT1A (SF-FE) mutants, respectively) (Figure 24A). Swapping the putative hydrogen bonding residues conferred tri-methylation activity to DOT1A, revealing a significant influence of distant parts on the catalytic center of the enzyme (Figure 24C). Notably,

the F90Y mutation introduces a tyrosine at the exact same position where yeast and human DOT1 proteins contain their hydrogen-bonding tyrosine, indicating that this conserved residue is a universal hallmark of tri-methylating DOT1 enzymes. Furthermore, the majority of the DOT1B CYΦS motif was introduced in DOT1A, while simultaneously removing the TEV motif (T89C, F90Y, R92S, T99V, E100T) in the WT and S218A/F246M mutant background (DOT1A (TFRTE) and DOT1A (SF-TFRTE) mutants) to clarify whether residues beyond the possible hydrogen bonding positions also have an influence on the efficiency of tri-methylation. However, no apparent increase in tri-methylation signals can be observed for these multiple mutants (Figure 24C). Taken together, these results allowed definition of two sequence motifs (TEV and CYΦS) within the N-terminal region of the KMTase domain that appear to be specific for either DOT1A or DOT1B enzymes and are crucial in determining their distinct product specificities.

Table 9. N-terminal residues contacting the lysine-binding channel

	H. sapiens **DOT1L**	***S. cerevisiae*** **Dot1p**	***T. brucei*** **DOT1A**	***T. brucei*** **DOT1B**
Helix D (CYΦS motif in *T.b*. DOT1B)	V114	V349	T89	C72
	Y115	Y350	F90	Y73
	H117	R352	R92	S75
Helix E (TEV motif in *T.b*. DOT1A)	K124	K360	T99	V82
	L125	L361	E100	T83

4 Discussion

4.1 ASF1 histone chaperones

4.1.1 The trypanosomal Asf1A and Asf1B system

The two Asf1 histone chaperones, Asf1A and Asf1B are found in distinct cellular locations in the parasite cell. Asf1A is mainly found in the cytosol but accumulates in the nucleus during S phase, while Asf1B is located in the nucleus throughout the cell cycle (Pascoalino et al., 2014). The predominant cytosolic localization of Asf1A in *T. brucei* differs from several other eukaryotes, such as yeast and humans, where ScAsf1/HsAsf1a mainly reside in the nucleus (Campos et al., 2010; Sutton et al., 2001). Human Asf1b and *Drosophila* Asf1 have also been found in a complex with histone H3/H4 heterodimers in the cytosol (Campos et al., 2010; Moshkin et al., 2002). The predominant nuclear localization of Asf1 in non-trypanosomal species is assumed to be a consequence of the concerted import of the H3/H4-Asf1 complex after histone recognition (Keck and Pemberton, 2012). Nuclear import of histones in trypanosomes is not well understood. Asf1A possibly binds to free H3/H4 heterodimer complexes and is imported into the nucleus when histones are synthesized during S-phase. This process could be mediated by additional H3/H4-Asf1 associated proteins as it has been shown in yeast (Blackwell et al., 2007; Campos et al., 2010). Interestingly, results of this thesis revealed that only the cytosolic Asf1A but not the nuclear Asf1B is able to bind reconstituted histone H3/H4 dimers in *T. brucei* (Figure 13A and B). This result in combination with the observed Asf1 localizations in *T. brucei* make it tempting to speculate that only Asf1A mediates the delivery of newly synthesized H3/H4 complexes to the nucleus where they are handed over to Asf1B prior assembly into chromatin. The fact that Asf1B does not bind to recombinant histones at all might be due to the lack of posttranslational protein modifications on histones expressed in *E. coli*. Pull-down assays with whole trypanosome cell extracts confirmed that Asf1B is able bind to presumably modified histones (Figure 13C). However, the possibility that Asf1B itself has to be modified to allow H3/H4 binding cannot be excluded. For instance, Asf1 proteins are well known substrates of TLKs in humans, *Drosophila* and *C. elegans* (Carrera et al., 2003; Han et al., 2005; Silljé and Nigg, 2001) as well as in trypanosomes (Li et al., 2007; 2008). Moreover, additional yet unidentified components might also be necessary to enable Asf1B-histone complex formation in *T. brucei*, analogous to the situation in human (Tang et al., 2006). Interestingly, yeast and human Asf1 stimulate cytosolic acetylation of histones H3 and H4 on several lysines (reviewed in (Burgess et al., 2010)) and some of these modifications are required for nuclear import of histones (Mühlhäusser et al., 2001). A corresponsing acetylation site is also present in trypanosomal H4 (lysine 4) (Siegel et al., 2008). This allows speculations about a selective role of Asf1A in mediating H3/H4 modifications and nuclear import that could be a prerequisite for the H3/H4 transfer to Asf1B, which in turn might be selectively responsible for the final steps of histone deposition into chromatin in the trypanosomal nucleus (Figure 25).

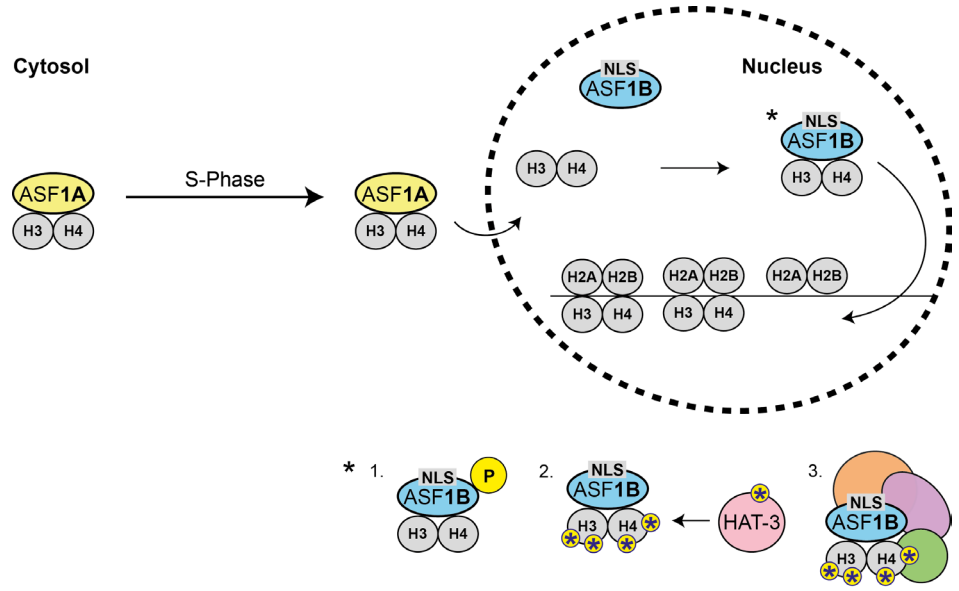


Figure 25. Schematic diagram of possible histone transport in *T. brucei*
The histone chaperone Asf1A binds to free H3/H4 complexes and is imported to the nucleus when histones are synthesized during S-phase, while Asf1B is located in the nucleus. Asf1B is not able to bind free histones but might acquire the ability through phosphorylation of Asf1B by TLKs (1), acetylation of histone H3 and H4 by the trypanosome histone acetyltransferase HAT-3 (2) or by Asf1B complex formation (3).

4.1.2 Conclusion and future perspective

T. brucei Asf1 proteins have distinguishable cellular localizations and different functions in the cell. However, the currently available results raise questions and demand important future experiments to elucidate the precise interplay and requirements for histone deposition by Asf1 chaperones in trypanosomes. Several questions can potentially be answered in vitro with the reconstituted nucleosome system presented in this thesis. Mutations of histone H3 and H4 at known acetylation sites from lysine to glutamine (to mimic acetylation) might be a strategy to analyze affinity changes of Asf1A and Asf1B towards the H3/H4 complex. Additionally, phosphorylation of Asf1 by TLKs could influence chaperone activation or change their binding behavior. In this context, phosphomimetic mutations (e.g. serine to aspartic acid) in histones provide an easily accessible option. However, due to the current lack of knowledge about the precise TLK phosphorylation sites in trypanosome H3 and H4 in vitro phosphorylation using purified TLK might be required. All of the described mutational strategies could be combined with pull down experiments using Asf1-H3/H4 complexes to reveal yet unidentified associated factors that are potentially involved in trypanosomal chromatin assembly in vivo.

4.2 DOT1 methyltransferases

4.2.1 The molecular basis for product-specificity of trypanosomal DOT1 enzymes

An important factor for the catalytic mechanism of DOT1 is deprotonation of the target lysine ε-amine to produce a lone electron pair that can initiate the nucleophilic attack on the methylsulfonyl group of SAM. The resulting methyl-lysine can further undergo consecutive rounds

of methylation that require repeated deprotonation in combination with a rotation of the target lysine side chain. Differences in the local environment of DOT1A and DOT1B that influence deprotonation could in principle account for the varying product specificities. One factor promoting deprotonation is an overall hydrophobic environment (Westheimer, 1995) in combination with negative charges in the proximity of the target lysine that could take up the subtracted proton. Although the molecular details remain to be determined, both *T. brucei* DOT1 enzymes are very likely to behave analogous to human and yeast DOT1s with respect to the deprotonation step. The lysine-binding channel has a hydrophobic entrance and an overall negative charge at the base in all DOT1 enzymes analyzed to date. No specific negatively charged residue is present in the channel, but the overall negative charge has been proposed to achieve deprotonation (Min et al., 2003; Sawada et al., 2004) analogous to the situation in SET domain-containing methyltransferases (Zhang et al., 2002). In addition, the SAM carboxylate group may contribute to deprotonation of the target lysine, since it is not compensated by a positive charge in its immediate proximity. Taken together, it can be proposed that differences in the lysine deprotonation step are very unlikely to exist for DOT1A and DOT1B and product-specificity is determined by other means.

First insights into the structural determinants of product-specificity came from comparisons of the SET domain containing KMTases DIM-5 and SET7/9, which catalyze formation of distinct products (me1/2/3 for DIM-5 and me1 for SET7/9) (Zhang et al., 2003). A single position occupied by either a phenylalanine (in DIM-5) or a tyrosine (in SET7/9) determines product-specificity in these cases and swapping the residues changes product-specificity of the mutant enzymes (Zhang et al., 2003). S218 in DOT1A could function in a related fashion by hydrogen bonding to the target lysine ε-amine and thus restricting the enzyme activity to mono- and di-methylation of H3K76. The S218A mutation removes the hydroxyl group, which would allow free rotation of the target lysine and consequently the mutant enzyme is able to set the tri-methylation mark. This scenario resembles the situation in the SET7/9 methyltransferase, where hydrogen bonding of two tyrosines restricts rotation of the substrate-lysine and thus limits the product to me1 (Xiao et al., 2003).

F246 of *T. brucei* DOT1A might restrict the target lysine movement or cause steric exclusion of a tri-methylated lysine via its bulky aromatic side chain. The F246M mutation introduces a smaller side chain and confers more space for the substrate-lysine, which might be a prerequisite for the tri-methylation activity of DOT1B. Surprisingly, the F246M mutation not only changes the product-specificity but also increases the reaction rate of di- but not mono-methylation, indicating that a mono-methylated lysine is a better substrate for the DOT1A F246M mutant compared to the WT enzyme. DOT1A F246 is located in close proximity to F116 and F289 on one side of the lysine-binding channel. Notably, recognition and binding of methylated lysines often involves so-called aromatic cages found in members of the Royal superfamily of folds (e.g. chromo- and tudor domains) and PHD fingers (Taverna et al., 2007). DOT1A F246 together with F116 and/or F289 might form an aromatic cage-like structure that could function as a temporal trap for H3K76me1. Such a trap close to the active site could effectively slow down the di-methylation reaction, while not affecting mono-methylation. In contrast, the F246 equivalent M225 in DOT1B cannot form an aromatic cage, which is consistent with the results showing that DOT1B is more effective in converting lower to higher methylation states compared to DOT1A.

This observation is perfectly compatible with the different functions of DOT1A and DOT1B in the parasite. DOT1A-mediated H3K76me1/2 appears slowly after incorporation of new H3 into the chromatin fiber restricting these marks to G2-phase and mitosis, whereas DOT1B seems to convert all H3K76 quickly to a tri-methylated state at the end of mitosis (Gassen et al., 2012).

4.2.2 Contribution of residues outside of the catalytic core

The importance of lysine-binding pocket stabilization by N-terminal parts of the KMTase domain has been previously established by structural work on yeast and human DOT1 enzymes (Min et al., 2003; Sawada et al., 2004). Here, the central glutamate (E374 in *S. cerevisiae* and E138 in *H. sapiens*) of the D1 motif is engaged in a hydrogen bond with an N-terminal tyrosine (Y350 and Y115 in *S. cerevisiae* and *H. sapiens*, respectively). Disrupting or changing the hydrogen bonding properties at this position by E374Q or E374A mutations abolishes enzymatic activity of the yeast Dot1p enzyme (Sawada et al., 2004). Moreover, a Y350F mutation in yeast dramatically reduces Dot1p activity, underlining the importance of the tyrosine hydroxyl group for positioning E374 via a hydrogen bond, thus maintaining a proper architecture of the lysine-binding pocket (Sawada et al., 2004). Results of this thesis clearly establish a novel function for the N-terminal part of the KMTase core in determining product-specificity of DOT1 enzymes by variation of the interaction pattern with the conserved D1 motif. This offers the interesting possibility to alter the D1 stabilization via the N-terminal region in organisms containing a single DOT1 enzyme to specifically change product-specificity. For example a DOT1 mutant that only mediates mono- or di-methylation of H3K79 in yeast could be an extremely useful tool to address the still discussed possibility of functional redundancy of different methylation levels.

4.2.3 Substrate-targeting of DOT1 enzymes

T. brucei DOT1A and DOT1B lack long N- or C-terminal extensions outside of the conserved KMTase domain that can be found in other eukaryotes such as yeast, fly and human (Figure 4). Parts of these extensions have been shown to be required for effective nucleosome interaction by providing a DNA binding interface (Min et al., 2003; Sawada et al., 2004). DNA binding is achieved via a (lysine- and arginine-rich) positively charged sequence in both the yeast and human DOT1 enzymes. Notably, this patch is positioned C-terminally (residues 361-416) with respect to the catalytic core domain in human DOT1L, while it is found N-terminally (residues 87-144) in yeast (Min et al., 2003; Sawada et al., 2004). This suggests, that chromatin recognition via DNA binding allows a certain degree of freedom and functions by bringing the enzyme in the proximity to the substrate but is likely not part of the DOT1-nucleosome interactions that properly align the enzyme for catalysis on the target lysine of H3. A close inspection of the N-terminal part of *T. brucei* DOT1A and DOT1B revealed lysine- and arginine-rich sequences in both enzymes (residues 8-44 and 4-20 in DOT1A and DOT1B, respectively) (Figure 26) that might function analogous to the DNA binding modules of yeast and human DOT1.

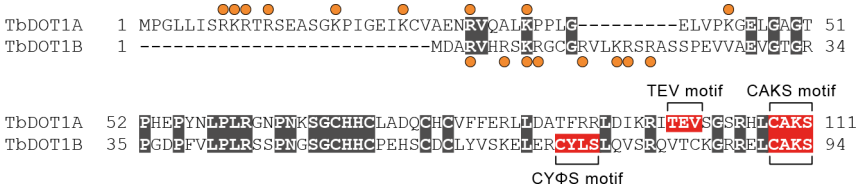


Figure 26. N-terminal sequence alignment of *T. brucei* DOT1A and DOT1B
Sequence alignment of the first 111 and 94 residues from DOT1A and DOT1B, respectively. Conserved positions are indicated in white with grey background. Characteristic sequence motifs are highlighted in red. Positively charged lysine and arginine residues within the N-terminal region are marked with orange dots above (for DOT1A) and below (for DOT1B) the sequences.

In addition, trypanosomal DOT1A and DOT1B have an acidic patch close to the active site (Figure 28), which is absent from DOT1 enzymes of other organisms and might be involved in nucleosome targeting and/or proper substrate recognition in trypanosomes. The patch is formed by D247/D248 and E226/D227 residues in *T. brucei* DOT1A and DOT1B, respectively, and is conserved among trypanosomal DOT1 enzymes (Figure 20).

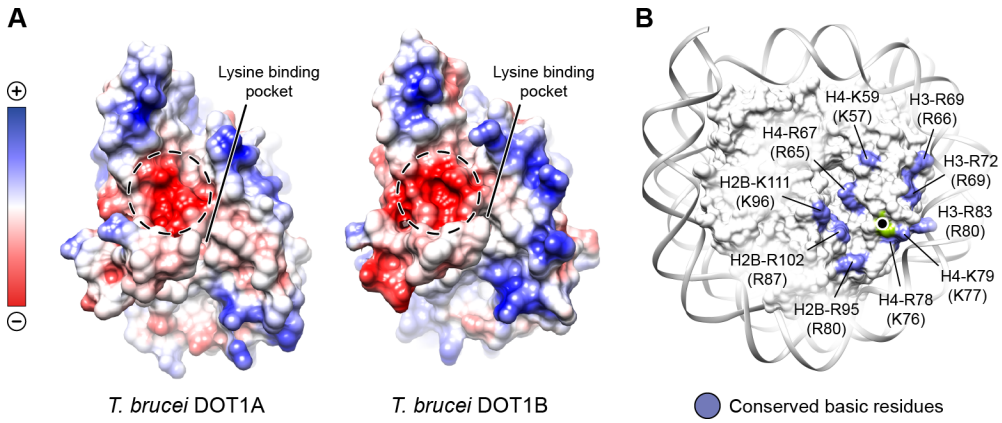


Figure 28. Surface of DOT1A and DOT1B enzymes
(**A**) Electrostatic surface potential representation of the *T. brucei* DOT1A and DOT1B homology models. Positive charges are depicted in blue and negative ones in red. Position of the lysine-binding pocket is indicated. A dashed circle marks the negatively charged patch next to the active site. (**B**) Surface representation of the yeast nucleosome with conserved positively charged residues around H3K79 (H3K76 in *T. brucei*, indicated with black dot) highlighted in blue. Amino acid numbering refers to *S. cerevisiae* histones with the equivalent *T. brucei* residues in parentheses.

4.2.4 Histone ubiquitylation

Although trypanosomes contain an ubiquitin system (Lowrie et al., 1993; Reverol et al., 1997), they lack the H2B K123 ubiquitylation, that is well known to stimulate H3 methylation by DOT1 enzymes in other organisms such as yeast and human (Briggs et al., 2002; Frederiks et al., 2008; 2010; McGinty et al., 2008; Nakanishi et al., 2009; Ng, 2002). In vivo methylation of H3K79 in a dot1Δ/bre1Δ yeast strain (Bre1p is the H2B K123 ubiquitin-ligase in yeast) by heterologously expressed enzymes showed no significant loss of activity for trypanosomal DOT1A and DOT1B, whereas methylation efficiency by yeast Dot1p is significantly decreased (Frederiks et al., 2010). It has previously been suggested that ubiquitin can serve as a spacer between adjacent nucleosomes to prevent chromatin condensation (Fierz et al., 2011; Sun and Allis, 2002). Thus,

stimulation of DOT1 activity might be due to facilitated access of the enzyme to the substrate within nucleosome arrays and not due to a direct interaction between ubiquitin and DOT1. It remains to be shown by future experiments, if trypanosomal DOT1 enzymes are dependent on chromatin decompaction in the cell, which in principle could be achieved by many other means than ubiquitylation of histone H2B. These for instance include binding by factors that disrupt close nucleosome contacts or chromatin remodeling by specialized molecular machines.

4.2.5 The trypanosomal DOT1A N-terminus as a potential drug target

The N-terminal part of the KMTase domain differs significantly between DOT1 enzymes from various eukaryotic species. Compared to human DOT1L, the N-terminal parts of *T. brucei* DOT1A and DOT1B share no similarity. Accordingly, our DOT1A and DOT1B homology models could not be extended towards the N-terminus beyond helix D. The finding that the N-terminal part influences the enzymatic activity of *T. brucei* DOT1A and DOT1B in combination with the essential character of DOT1A offers the possibility to use DOT1A N-terminus as a novel drug target. All attempts to inhibit human DOT1L enzymes in a strategy to cure acute myeloic leukemia have targeted the SAM-binding pocket via SAM analogs (Basavapathruni et al., 2012; Yu et al., 2012). Given the high degree of conservation of the SAM-binding pocket between human and trypanosomal DOT1 enzymes it is very unlikely that this would be a selective route to target the trypanosomal enzymes. However, it remains to be shown if the N-terminal part of DOT1A can selectively bind small molecules that have the potential to inhibit the enzyme's activity.

4.2.6 Conclusion and future perspective

Both trypanosomal DOT1 proteins work in a distributive manner to methylate H3K76. However, the in vitro results presented in this thesis cannot exclude the possibility that the enzymes are also regulated by posttranslational modifications or interactions with other proteins in vivo to modulate the enzymes' behavior towards a processive mode of action. Such modifications are not necessarily dependent on associated factors. It has been shown previously that an archaeal DOT1 homolog (aKMT4) is able to auto-methylate itself in the absence of a substrate (Niu et al., 2013). The in vitro approach appears to be promising tool to elucidate a potential self-methylating activity of DOT1 in trypanosomes. Recently, a DOT1-containing complex (DOTcom) could be isolated from human cells, which contains several transcription factors that potentially mediate specific chromatin targeting of the DOT1L methyltransferase (Mohan et al., 2010). This multisubunit complex consist of DOT1L together with proteins AF10, AF17, AF9, ENL, Skp1, TRRAP and ß-catenin and is associated with H3K79 di- and trimethylation (Mohan et al., 2010). DOT1-interacting factors in trypanosomes could be identified by pull-down experiments using DOT1 as bait. Unlike in other eukaryotes, trypanosomes might contain two DOTcom variants, specific for DOT1A or DOT1B. Analogous to the situation in humans, these complexes could include factors that mediate targeting to specific sites. Therefore, identification and characterization of such complexes in trypanosomes could reveal targeting differences for DOT1A and DOT1B that might explain how these enzymes mediate non-overlapping cellular functions. Based on sequence alignments it can be speculated that DOT1A- and DOT1B-specific interaction partners are likely to bind via the divergent N-terminal part of the enzymes rather then via the more similar C-terminal domain of the KMTase core. Methyl-lysine residues in a nucleosomal context have been shown to interact with factors (so-called readers) that execute biological functions (reviewed in (Martin

and Zhang, 2005; Maurer-Stroh et al., 2003; Patel and Wang, 2013; Taverna et al., 2007)). In trypanosomes, first attempts to isolate proteins that specifically interact with different H3K76 methylation states (H3K76me1, -me2 and -me3) using histone H3 peptides, remained unsuccessful (Christian Janzen, personal communication). This is very likely due to the location of H3K76 within the nucleosomal core and interacting factors might be dependent on both the nucleosomal context as well as on the H3K76 methylation mark. On the other hand, isolated nucleosomes from trypanosomes contain a mixture of different H3K76 methylations (Gassen et al., 2012) and are therefore not suitable as bait for methylation state specific H3K76 interactors in pull-down experiments. A defined biochemical system would overcome these problems. Results of this thesis demonstrate the use of purified DOT1 enzymes to create specific intermediate methylation states on nucleosomes in vitro. In principle, the system makes the H3K76me2 and me3 states accessible for subsequent pull-down experiments by using DOT1A and DOT1B in the methylation reaction, respectively (Figure 15). In addition, the DOT1A N-terminal swapping mutant identified in this work (Figure 24) appears to stop the reaction at the H3K76me1 level. However, initial tests to produce uniformly mono-, di- or tri-methylated nucleosomes in vitro via this strategy failed, since the in vitro reactions are difficult to drive to completion and vary in the degree of methylation. In consequence of this, the obtained nucleosomes were not uniform with respect to the H3K76 methylation state and further optimization will be required in future experiments. In an alternative chemical approach, extensions of an aminoethylation reaction are used to introduce methyl-lysine analogs (MLAs) into recombinant histone proteins prior nucleosome assembly (Simon et al., 2007). The method allows preparation of large amounts of uniformly modified histones and is dependent on the installation of a unique cysteine at the position that will later contain the MLA (Simon et al., 2007). In the case of *T. brucei* H3 this would require the K76C, C108A and C126A mutations, the latter two to prevent unwanted side reactions at these naturally occurring H3 cysteine residues. Uniformly modified nucleosomes could be introduced in a SILAC nucleosome affinity purification (SNAP) assay (Bartke et al., 2010) to identify factors that bind trypanosomal chromatin in the context of specific H3K76 methylation states.

In conclusion, the different enzymatic activities of trypanosome DOT1 homologues were exploited to identify several structural components that are responsible for product-specificity of these conserved histone methyltransferases. This does not only shed light on the mode of action of the parasite enzymes but might also help to understand the function and regulation of members of the DOT1 family in other eukaryotes. Information provided by this thesis could be useful to generate DOT1 mutants with specific enzymatic activities to solve long-standing questions about possible redundancy of different methylation states. To obtain these insights, development of the cell free methyltransferase assay based on reconstituted trypanosomal nucleosomes was essential. This new tool might also be very useful in future experiments to reveal additional aspects of the molecular mechanism and regulation of DOT1 enzymes in *T. brucei.*

5 References

Adam, M., Robert, F., Larochelle, M., and Gaudreau, L. (2001). H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. Mol. Cell. Biol. *21*, 6270–6279.

Akiyoshi, B., and Gull, K. (2014). Discovery of Unconventional Kinetochores in Kinetoplastids. Cell 156, 1247–1258.

Alsford, S., and Horn, D. (2004). Trypanosomatid histones. Mol. Microbiol. *53*, 365–372.

Alsford, S., duBois, K., Horn, D., and Field, M.C. (2012). Epigenetic mechanisms, nuclear architecture and the control of gene expression in trypanosomes. Expert Rev. Mol. Med. *14*, e13.

Alsford, S., Horn, D., and Glover, L. (2009). DNA breaks as triggers for antigenic variation in African trypanosomes. Genome Biol. *10*, 223.

Altaf, M., Utley, R.T., Lacoste, N., Tan, S., Briggs, S.D., and Côte, J. (2007). Interplay of chromatin modifiers on a short basic patch of histone H4 tail defines the boundary of telomeric heterochromatin. Mol. Cell *28*, 1002–1014.

Andrews, A.J., and Luger K. (2011). Nucleosome structure(s) and stability: variations on a theme. Annu Rev. Biophys. *40*, 99–117.

Armache, K.-J., Garlick, J.D., Canzio, D., Narlikar, G.J., and Kingston, R.E. (2011). Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. Science *334*, 977–982.

Barry, J.D., Marcello, L., Morrison, L.J., Read, A.F., Lythgoe, K., Jones, N., Carrington, M., Blandin, G., Böhme, U., Caler, E., et al. (2005). What the genome sequence is revealing about trypanosome antigenic variation. Biochem. Soc. Trans. *33*, 986–989.

Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S.C., Mann, M., and Kouzarides, T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell *143*, 470–484.

Basavapathruni, A., Jin, L., Daigle, S.R., Majer, C.R.A., Therkelsen, C.A., Wigle, T.J., Kuntz, K.W., Chesworth, R., Pollock, R.M., Scott, M.P., et al. (2012). Conformational adaptation drives potent, selective and durable inhibition of the human protein methyltransferase DOT1L. Chem. Biol. Drug Des. *80*, 971–980.

Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D.C., Lennard, N.J., Caler, E., Hamlin, N.E., Haas, B., et al. (2005). The genome of the African trypanosome *Trypanosoma brucei*. Science *309*, 416–422.

Binda, O. (2013). On your histone mark, SET, methylate! Epigenetics *8*, 457–463.

Blackwell, J.S., Wilkinson, S.T., Mosammaparast, N., and Pemberton, L.F. (2007). Mutational analysis of H3 and H4 N termini reveals distinct roles in nuclear import. J. Biol. Chem. *282*, 20142–20150.

Briggs, S.D., Xiao, T., Sun, Z.-W., Caldwell, J.A., Shabanowitz, J., Hunt, D.F., Allis, C.D., and Strahl, B.D. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. Nature *418*, 498.

Brun, R., and Schönenberger (1979). Cultivation and in vitro cloning or procyclic culture forms of *Trypanosoma brucei* in a semi-defined medium. Short communication. Acta Trop. *36*, 289–292.

Burgess, R.J., and Zhang, Z. (2010). Histones, histone chaperones and nucleosome assembly. Protein Cell *1*, 607–612.

Burgess, R.J., and Zhang, Z. (2013). Histone chaperones in nucleosome assembly and human disease. Nat. Struct. Mol. Biol. *20*, 14–22.

Burgess, R.J., Zhou, H., Han, J., and Zhang, Z. (2010). A role for Gcn5 in replication-coupled nucleosome assembly. Mol. Cell *37*, 469–480.

Butter, F., Bucerius, F., Michel, M., Cicova, Z., Mann, M., and Janzen, C.J. (2013). Comparative Proteomics of Two Life Cycle Stages of Stable Isotope-labeled *Trypanosoma brucei* Reveals Novel Components of the Parasite's Host Adaptation Machinery. Mol. Cell. Proteomics *12*, 172–179.

Campagna-Slater, V., Mok, M.W., Nguyen, K.T., Feher, M., Najmanovich, R., and Schapira, M. (2011). Structural chemistry of the histone methyltransferases cofactor binding site. J. Chem. Inf. Model. *51*, 612–623.

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

Carrera, P., Moshkin, Y.M., Gronke, S., Silljé, H.H.W., Nigg, E.A., Jackle, H., and Karch, F. (2003). Tousled-like kinase functions with the chromatin assembly pathway regulating nuclear divisions. Genes Dev. *17*, 2578–2590.

Cheng, X., Collins, R.E., and Zhang, X. (2005). Structural and sequence motifs of protein (histone) methylation enzymes. Annu. Rev. Biophys. Biomol. Struct. *34*, 267–294.

Clapier, C.R., Cairns, B.R. (2009). The biology of chromatin remodeling complexes. Annu. Rev. Biochem. *78*, 273–304.

Daganzo, S.M., Erzberger, J.P., Lam, W.M., Skordalakes, E., Zhang, R., Franco, A.A., Brill, S.J., Adams, P.D., Berger, J.M., and Kaufman, P.D. (2003). Structure and function of the conserved core of histone deposition protein Asf1. Curr. Biol. *13*, 2148–2158.

de Marco, A., Vigh, L., Diamant, S., and Goloubinoff, P. (2005). Native folding of aggregation-prone recombinant proteins in *Escherichia coli* by osmolytes, plasmid- or benzyl alcohol-overexpressed molecular chaperones. Cell Stress Chaperones *10*, 329–339.

Dhillon, N., and Kamakaka, R.T. (2000). A histone variant, Htz1p, and a Sir1p-like protein, Esc2p, mediate silencing at HMR. Mol. Cell *6*, 769–780.

Dillon, S.C., Zhang, X., Trievel, R.C., Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. Genome Biol. *6*,227.

Dlakić, M. (2001). Chromatin silencing protein and pachytene checkpoint regulator Dot1p has a methyltransferase fold. Trends Biochem. Sci. *26*, 405–407.

Donham, D.C., Scorgie, J.K., and Churchill, M.E.A. (2011). The activity of the histone chaperone yeast Asf1 in the assembly and disassembly of histone H3/H4-DNA complexes. Nucleic Acids Res. *39*, 5449–5458.

Dryhurst, D., Thambirajah, A.A., and Ausió, J. (2004). New twists on H2A.Z: a histone variant with a controversial structural and functional past. Biochem. Cell Biol. *82*, 490–497.

English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E.A., and Tyler, J.K. (2006). Structural Basis for the Histone Chaperone Activity of Asf1. Cell *127*, 495–508.

English, C.M., Maluf, N.K., Tripet, B., Churchill, M.E., and Tyler, J.K. (2005). ASF1 binds to a heterodimer of histones H3 and H4: a two-step mechanism for the assembly of the H3-H4 heterotetramer on DNA. Biochem. *44*, 13673–13682.

Ersfeld, K., Melville, S.E., and Gull, K. (1999). Nuclear and genome organization of Trypanosoma brucei. Parasitol. Today (Regul. Ed.) *15*, 58–63.

Eswar, N., Eramian, D., Webb, B., Shen, M.-Y., and Sali, A. (2008). Protein structure modeling with MODELLER. Methods Mol. Biol. *426*, 145–159.

Fan, J.Y., Rangasamy, D., Luger, K., and Tremethick, D.J. (2004). H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. Mol. Cell *16*, 655–661.

Feng, Q., Wang, H., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Struhl, K., and Zhang, Y. (2002). Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr. Biol. *12*, 1052–1058.

Fierz, B., Chatterjee, C., McGinty, R.K., Bar-Dagan, M., Raleigh, D.P., and Muir, T.W. (2011). Histone H2B ubiquitylation disrupts local and higher-order chromatin compaction. Nat. Chem. Biol. *7*, 113–119.

Figueiredo, L.M., Cross, G.A.M., and Janzen, C.J. (2009). Epigenetic regulation in African trypanosomes: a new kid on the block. Nat. Rev. Microbiol. *7*, 504–513.

Figueiredo, L.M., Janzen, C.J., and Cross, G.A.M. (2008). A histone methyltransferase modulates antigenic variation in African trypanosomes. PloS Biol. *6*, e161.

Fingerman, I.M., Li, H.-C., and Briggs, S.D. (2007). A charge-based interaction between histone H4 and Dot1 is required for H3K79 methylation and telomere silencing: identification of a new trans-histone pathway. Genes Dev. *21*, 2018–2029.

Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., et al. (2005). Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. Nat. Genet. *37*, 391–400.

Frederiks, F., Stulemeijer, I.J.E., Ovaa, H., and van Leeuwen, F. (2011). A modified epigenetics toolbox to study histone modifications on the nucleosome core. Chem. Biochem. *12*, 308–313.

Frederiks, F., Tzouros, M., Oudgenoeg, G., van Welsem, T., Fornerod, M., Krijgsveld, J., and van Leeuwen, F. (2008). Nonprocessive methylation by Dot1 leads to functional redundancy of histone H3K79 methylation states. Nat. Struct. Mol. Biol. *15*, 550–557.

Frederiks, F., van Welsem, T., Oudgenoeg, G., Heck, A.J.R., Janzen, C.J., and van Leeuwen, F. (2010). Heterologous expression reveals distinct enzymatic activities of two DOT1 histone methyltransferases of *Trypanosoma brucei*. J. Cell Sci. *123*, 4019–4023.

Fu, H., Maunakea, A.K., Martin, M.M., Huang, L., Zhang, Y., Ryan, M., Kim, R., Lin, C.M., Zhao, K., and Aladjem, M.I. (2013). Methylation of histone H3 on lysine 79 associates with a group of replication origins and helps limit DNA replication once per cell cycle. PLoS Genet. *9*, e1003542.

Gassen, A., Brechtefeld, D., Schandry, N., Arteaga-Salas, J.M., Israel, L., Imhof, A., and Janzen, C.J. (2012). DOT1A-dependent H3K76 methylation is required for replication regulation in *Trypanosoma brucei*. Nucleic Acids Res. 40, 10302-10311.

Guo, H.-B., and Guo, H. (2007). Mechanism of histone methylation catalyzed by protein lysine methyltransferase SET7/9 and origin of product specificity. Proc. Natl. Acad. Sci. U.S.A. *104*, 8797–8802.

Greer, E.L., and Shi, Y. (2012). Histone methylation: a dynamic mark in health, disease and inheritance. Nat. Rev. Genet. *13*, 343–357.

Günzl, A., Bruderer, T., Laufer, G., Schimanski, B., Tu, L.-C., Chung, H.-M., Lee, P.-T., and Lee, M.G.-S. (2003). RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in *Trypanosoma brucei*. Eukaryotic Cell *2*, 542–551.

Han, Z., Riefler, G.M., Saam, J.R., Mango, S.E., and Schumacher, J.M. (2005). The C. elegans Tousled-like kinase contributes to chromosome segregation as a substrate and regulator of the Aurora B kinase. Curr. Biol. *15*, 894–904.

Hansen, J.C. (2002). Conformational dynamics of the chromatin fiber in solution: determinants, mechanisms, and functions. Annu. Rev. Biophys. Biomol. Struct. *31*, 361–392.

Harshman, S.W., Young, N.L., Parthun, M.R., Freitas, M.A. (2013). H1 histones: current perspectives and challenges. Nucleic Acids Res. *41*, 9593-9609.

Hassa, P.O., Haenni, S.S., Elser, M., and Hottiger, M.O. (2006). Nuclear ADP-ribosylation reactions in mammalian cells: where are we today and where are we going? Microbiol. Mol. Biol. Rev. *70*, 789–829.

Hecker, H., Betschart, B., Bender, K., Burri, M., and Schlimme, W. (1994). The chromatin of trypanosomes. Int. J. Parasitol. *24*, 809–819.

Horn, D. (2007). Introducing histone modification in trypanosomes. Trends Parasitol. *23*, 239–242.

Horn, D., and McCulloch, R. (2010). Molecular mechanisms underlying the control of antigenic variation in African trypanosomes. Curr. Opin. Microbiol. *13*, 700–705.

Hughes, K., Wand, M., Foulston, L., Young, R., Harley, K., Terry, S., Ersfeld, K., and Rudenko, G. (2007). A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. EMBO J. *26*, 2400–2410.

Jacobs, S.A., Harp, J.M., Devarakonda, S., Kim, Y., Rastinejad, F., and Khorasanizadeh, S. (2002). The active site of the SET domain is constructed on a knot. Nat. Struct. Biol. *9*, 833–838.

Janzen, C.J., Fernandez, J.P., Deng, H., Diaz, R., Hake, S.B., and Cross, G.A.M. (2006a). Unusual histone modifications in Trypanosoma brucei. FEBS Lett. *580*, 2306–2310.

Janzen, C.J., Hake, S.B., Lowell, J.E., and Cross, G.A.M. (2006b). Selective di- or trimethylation of histone H3 lysine 76 by two DOT1 homologs is important for cell cycle regulation in *Trypanosoma brucei*. Mol. Cell *23*, 497–507.

Jones, B., Su, H., Bhat, A., Lei, H., Bajko, J., Hevi, S., Baltus, G.A., Kadam, S., Zhai, H., Valdez, R., et al. (2008). The Histone H3K79 Methyltransferase Dot1L Is Essential for Mammalian Development and Heterochromatin Structure. PLoS Genet. *4*, e1000190.

Jones, R.S., and Gelbart, W.M. (1993). The *Drosophila* Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. Mol. Cell. Biol. *13*, 6357–6366.

Keck, K.M., and Pemberton, L.F. (2012). Histone chaperones link histone nuclear import and chromatin assembly. Biochim. Biophys. Acta *1819*, 277–289.

Kim, W., Choi, M., and Kim, J.-E. (2014). The histone methyltransferase Dot1/DOT1L as a critical regulator of the cell cycle. Cell Cycle *13*.

Kornberg, R.D., and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell *98*, 285–294.

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693–705.

Kwon, T., Chang, J.H., Kwak, E., Lee, C.W., Joachimiak, A., Kim, Y.C., Lee, J., and Cho, Y. (2003). Mechanism of histone lysine methyl transfer revealed by the structure of SET7/9-AdoMet. EMBO J. *22*, 292–303.

Lacoste, N., Utley, R.T., Hunter, J.M., Poirier, G.G., and Côte, J. (2002). Disruptor of telomeric silencing-1 is a chromatin-specific histone H3 methyltransferase. J. Biol. Chem. *277*, 30421–30424.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton J.M. (1993). PROCHECK: a programm to check the stereochemical quality of protein structures. J. Appl. Cryst. *26*, 283–291.

Le, S., Davis, C., Konopka, J.B., and Sternglanz, R. (1997). Two new S-phase-specific genes from *Saccharomyces cerevisiae*. Yeast *13*, 1029–1042.

Lee, K.-M., and Narlikar, G. (2001). Assembly of Nucleosomal Templates by Salt Dialysis. Curr. Prot. Mol. Biol. 21, mb2106s54.

Lennartsson, A., and Ekwall, K. (2009). Histone modification patterns and epigenetic codes. Biochim. Biophys. Acta 1790, 863–868.

Li, Z., Gourguechon, S., and Wang, C.C. (2007). Tousled-like kinase in a microbial eukaryote regulates spindle assembly and S-phase progression by interacting with Aurora kinase and chromatin assembly factors. J. Cell Sci. *120*, 3883–3894.

Li, Z., Umeyama, T., and Wang, C.C. (2008). The chromosomal passenger complex and a mitotic kinesin interact with the Tousled-like kinase in trypanosomes to regulate mitosis and cytokinesis. PLoS ONE *3*, e3814.

Liu, W.H., and Churchill, M.E.A. (2012). Histone transfer among chaperones. Biochem. Soc. Trans. *40*, 357–363.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J. Mol. Biol. *276*, 19–42.

Lowell, J.E., and Cross, G.A.M. (2004). A variant histone H3 is enriched at telomeres in *Trypanosoma brucei*. J. Cell Sci. *117*, 5937–5947.

Lowell, J.E., Kaiser, F., Janzen, C.J., and Cross, G.A.M. (2005). Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in *Trypanosoma brucei*. J. Cell Sci. *118*, 5721–5730.

Lowrie, D.J., Giffin, B.F., and Ventullo, R.M. (1993). The ubiquitin-ligase system in *Trypanosoma brucei brucei*. Am. J. Trop. Med. Hyg. *49*, 545–551.

Luger, K. (2001). Nucleosomes: Structure and Function. In: eLS. John Wiley & Sons Ltd, Chichester. http://www.els.net

Luger, K., Dechassa, M.L., and Tremethick, D.J. (2012). New insights into nucleosome and chromatin structure: an ordered state or a disordered affair? Nat. Rev. Mol. Cell Biol. *13*, 436–447.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251–260.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome core particle from recombinant histones. Meth. Enzymol. *304*, 3–19.

Luger, K., and Hansen, J.C. (2005). Nucleosome and chromatin fiber dynamics. Curr. Opin. Struct. Biol. *15*, 188–196.

Malik, H.S., and Henikoff, S. (2003). Phylogenomics of the nucleosome. Nat. Struct. Biol. *10*, 882–891.

Malone, T., Blumenthal, R.M., and Cheng, X. (1995). Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. J. Mol. Biol. *253*, 618–632.

Mandava, V., Fernandez, J.P., Deng, H., Janzen, C.J., Hake, S.B., and Cross, G.A.M. (2007). Histone modifications in *Trypanosoma brucei*. Mol. Biochem. Parasitol. *156*, 41–50.

Manzur, K.L., Farooq, A., Zeng, L., Plotnikova, O., Koch, A.W., Sachchidanand, and Zhou, M.-M. (2003). A dimeric viral SET domain methyltransferase specific to Lys27 of histone H3. Nat. Struct. Biol. *10*, 187–196.

Martin, C., and Zhang, Y. (2005). The diverse functions of histone lysine methylation. Nat. Rev. Mol. Cell Biol. *6*, 838–849.

Maurer-Stroh, S., Dickens, N.J., Hughes-Davies, L., Kouzarides, T., Eisenhaber, F., and Ponting, C.P. (2003). The Tudor domain "Royal Family": Tudor, plant Agenet, Chromo, PWWP and MBT domains. Trends Biochem. Sci. *28*, 69–74.

Mazet, M., Morand, P., Biran, M., Bouyssou, G., Courtois, P., Daulouède, S., Millerioux, Y., Franconi, J.-M., Vincendeau, P., Moreau, P., et al. (2013). Revisiting the central metabolism of the bloodstream forms of *Trypanosoma brucei*: production of acetate in the mitochondrion is essential for parasite viability. PLoS Negl. Trop. Dis. *7*, e2587.

McGinty, R.K., Kim, J., Chatterjee, C., Roeder, R.G., and Muir, T.W. (2008). Chemically ubiquitylated histone H2B stimulates hDot1L-mediated intranucleosomal methylation. Nature *453*, 812–816.

Mellone, B.G., and Allshire, R.C. (2003). Stretching it: putting the CEN(P-A) in centromere. Curr. Opin. Genet. Dev. *13*, 191–198.

Meneghini, M.D., Wu, M., and Madhani, H.D. (2003). Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell *112*, 725–736.

Min, J., Feng, Q., Li, Z., Zhang, Y., and Xu, R.-M. (2003). Structure of the catalytic domain of human DOT1L, a non-SET domain nucleosomal histone methyltransferase. Cell *112*, 711–723.

Min, J., Zhang, X., Cheng, X., Grewal, S.I.S., and Xu, R.-M. (2002). Structure of the SET domain histone lysine methyltransferase Clr4. Nat. Struct. Biol. *9*, 828–832.

Mohan, M., Herz, H.-M., Takahashi, Y.-H., Lin, C., Lai, K.C., Zhang, Y., Washburn, M.P., Florens, L., and Shilatifard, A. (2010). Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom). Genes Dev. *24*, 574–589.

Morales, V., and Richard-Foy, H. (2000). Role of histone N-terminal tails and their acetylation in nucleosome dynamics. Mol. Cell. Biol. *20*, 7230–7237.

Moshkin, Y.M., Armstrong, J.A., Maeda, R.K., Tamkun, J.W., Verrijzer, P., Kennison, J.A., and Karch, F. (2002). Histone chaperone ASF1 cooperates with the Brahma chromatin-remodelling machinery. Genes Dev. *16*, 2621–2626.

Mühlhäusser, P., Müller, E.C., Otto, A., and Kutay, U. (2001). Multiple pathways contribute to nuclear import of core histones. EMBO Rep. *2*, 690–696.

Nakanishi, S., Lee, J.S., Gardner, K.E., Gardner, J.M., Takahashi, Y.-H., Chandrasekharan, M.B., Sun, Z.-W., Osley, M.A., Strahl, B.D., Jaspersen, S.L., et al. (2009). Histone H2BK123 monoubiquitination is the critical determinant for H3K4 and H3K79 trimethylation by COMPASS and Dot1. J. Cell Biol. *186*, 371–377.

Nathan, D., Ingvarsdottir, K., Sterner, D.E., Bylebyl, G.R., Dokmanovic, M., Dorsey, J.A., Whelan, K.A., Krsmanovic, M., Lane, W.S., Meluh, P.B., et al. (2006). Histone sumoylation is a negative regulator in *Saccharomyces cerevisiae* and shows dynamic interplay with positive-acting histone modifications. Genes Dev. *20*, 966–976.

Natsume, R., Eitoku, M., Akai, Y., Sano, N., Horikoshi, M., and Senda, T. (2007). Structure and function of the histone chaperone CIA/ASF1 complexed with histones H3 and H4. Nature *446*, 338–341.

Ng, H.H. (2002). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. Genes Dev. *16*, 1518–1527.

Ng, H.H., Xu, R.-M., Zhang, Y., and Struhl, K. (2002). Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. J. Biol. Chem. *277*, 34655–34657.

Nguyen, A.T., and Zhang, Y. (2011). The diverse functions of Dot1 and H3K79 methylation. Genes Dev. *25*, 1345–1358.

Nguyen, K.T., Li, F., Poda, G., Smil, D., Vedadi, M., and Schapira, M. (2013). Strategy to Target the Substrate Binding site of SET Domain Protein Methyltransferases. J. Chem. Inf. Model. *53*, 681–691.

Niu, Y., Xia, Y., Wang, S., Li, J., Niu, C., Li, X., Zhao, Y., Xiong, H., Li, Z., Lou, H., et al. (2013). A prototypic lysine methyltransferase 4 from Archaea with degenerate sequence specificity methylates chromatin proteins Sul7d and Cren7 in different patterns. J. Biol. Chem. 288, 13728–13740.

Nowak, S.J., and Corces, V.G. (2004). Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet. *20*, 214–220.

Oberholzer, M., Morand, S., Kunz, S., and Seebeck, T. (2006). A vector series for rapid PCR-mediated C-terminal in situ tagging of *Trypanosoma brucei* genes. Mol. Biochem. Parasitol *145*, 117–120.

Onishi, M., Liou, G.-G., Buchberger, J.R., Walz, T., and Moazed, D. (2007). Role of the conserved Sir3-BAH domain in nucleosome binding and silent chromatin assembly. Mol. Cell *28*, 1015–1028.

Pascoalino, B., Dindar, G., Vieira-da-Rocha, J.P., Machado, C.R., Janzen, C.J., and Schenkman, S. (2014). Characterization of two different Asf1 histone chaperones with distinct cellular localizations and functions in *Trypanosoma brucei*. Nucleic Acids Res. *42*, 2906–2918.

Patel, D.J., and Wang, Z. (2013). Readout of epigenetic modifications. Annu. Rev. Biochem. *82*, 81–118.

Peterson, C.L., and Laniel, M.-A. (2004). Histones and histone modifications. Curr. Biol. *14*, R546–R551.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera - a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612.

Povelones, M.L., Gluenz, E., Dembek, M., Gull, K., and Rudenko, G. (2012). Histone H1 plays a role in heterochromatin formation and VSG expression site silencing in *Trypanosoma brucei*. PLoS Pathog. *8*, e1003010.

Rangasamy, D., Berven, L., Ridgway, P., and Tremethick, D.J. (2003). Pericentric heterochromatin becomes enriched with H2A.Z during early mammalian development. EMBO J. *22*, 1599–1607.

Rangasamy, D., Greaves, I., and Tremethick, D.J. (2004). RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. Nat. Struct. Mol. Biol. *11*, 650–655.

Reverol, L., Chirinos, M., and Henriquez, D.A. (1997). Presence of an unusually high concentration of an ubiquitinated histone-like protein in *Trypanosoma cruzi*. J. Cell. Biochem. *66*, 433–440.

Roditi, I., Schwarz, H., Pearson, T.W., Beecroft, R.P., Liu, M.K., Richardson, J.P., Bühring, H.J., Pleiss, J., Bülow, R., and Williams, R.O. (1989). Procyclin gene expression and loss of the variant surface glycoprotein during differentiation of *Trypanosoma brucei*. J. Cell Biol. *108*, 737–746.

Ruthenburg, A.J., Li, H., Patel, D.J., and Allis, C.D. (2007). Multivalent engagement of chromatin modifications by linked binding modules. *8*, 983–994.

Santisteban, M.S., Kalashnikova, T., and Smith, M.M. (2000). Histone H2A.Z regulats transcription and is partially redundant with nucleosome remodeling complexes. Cell *103*, 411–422.

Sawada, K., Yang, Z., Horton, J.R., Collins, R.E., Zhang, X., and Cheng, X. (2004). Structure of the conserved core of the yeast Dot1p, a nucleosomal histone H3 lysine 79 methyltransferase. J. Biol. Chem. *279*, 43296–43306.

Schapira, M. (2011). Structural Chemistry of Human SET Domain Protein Methyltransferases. Curr. Chem. Genomics *5*, 85–94.

Schneider, R., Bannister, A.J., and Kouzarides, T. (2002). Unsafe SETs: histone lysine methyltransferases and cancer. Trends Biochem. Sci. *27*, 396–402.

Schubert, H.L., Blumenthal, R.M., and Cheng, X. (2003). Many paths to methyltransfer: a chronicle of convergence. Trends Biochem. Sci. *28*, 329–335.

Shahbazian, M.D., Zhang, K., Grunstein, M. (2005). Histone H2B ubiquitylation controls processive methylation but not monomethylation by Dot1 and Set1. Mol. Cell *19*, 271–277.

Shanower, G.A., Muller, M., Blanton, J.L., Honti, V., Gyurkovics, H., and Schedl, P. (2005). Characterization of the grappa gene, the *Drosophila* histone H3 lysine 79 methyltransferase. Genetics *169*, 173–184.

Shen, S., Arhin, G.K., Ullu, E., and Tschudi, C. (2001). In vivo epitope tagging of *Trypanosoma brucei* genes using a one step PCR-based strategy. Mol. Biochem. Parasitol. *113*, 171–173.

Shigapova, N., Török, Z., Balogh, G., Goloubinoff, P., Vigh, L., and Horváth, I. (2005). Membrane fluidization triggers membrane remodeling which affects the thermotolerance in *Escherichia coli*. Biochem. Biophys. Res. Comm. *328*, 1216–1223.

Shilatifard, A. (2006). Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annu. Rev. Biochem. *75*, 243–269.

Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S., and Cross, G.A.M. (2009). Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. Genes Dev. *23*, 1063–1076.

Siegel, T.N., Kawahara, T., DeGrasse, J.A., Janzen, C.J., Horn, D., and Cross, G.A.M. (2008). Acetylation of histone H4K4 is cell cycle regulated and mediated by HAT3 in *Trypanosoma brucei*. Mol. Microbiol. *67*, 762–771.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539.

Silljé, H.H., and Nigg, E.A. (2001). Identification of human Asf1 chromatin assembly factors as substrates of Tousled-like kinases. Curr. Biol. *11*, 1068–1073.

Simon, M.D., Chu, F., Racki, L.R., la Cruz, de, C.C., Burlingame, A.L., Panning, B., Narlikar, G.J., and Shokat, K.M. (2007). The site-specific installation of methyl-lysine analogs into recombinant histones. Cell *128*, 1003–1012.

Singer, M.S., Kahana, A., Wolf, A.J., Meisinger, L.L., Peterson, S.E., Goggin, C., Mahowald, M., and Gottschling, D.E. (1998). Identification of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae. Genetics *150*, 613–632.

Smith, B.C., and Denu, J.M. (2009). Chemical mechanisms of histone lysine and arginine modifications. Biochim. Biophys. Acta *1789*, 45–57.

Söding, J. (2005). Protein homology detection by HMM-HMM comparison. Bioinformatics *21*, 951–960.

Söding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. *33*, W244–W248.

Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.M., Zhu, P, and Li, G. (2014). Cryo-EM study of the chromatin fiber revaels a double helix twisted by tetranucleosomal units. Science *344*, 376–380.

Sterner, D.E., and Berger, S.L. (2000). Acetylation of histones and transcription-related factors. Microbiol. Mol. Biol. Rev. *64*, 435–459.

Stillman, B. (1986). Chromatin assembly during SV40 DNA replication in vitro. Cell *45*, 555–565.

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

Struhl, K., Segal, E. (2013). Determinants of nucleosome positioning. Nat. Struct. Mol. Biol. *20*, 267–273.

Sun, Z.-W., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. Nature *418*, 104–108.

Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. Nat. Struct. Biol. *7*, 1121–1124.

Sutton, A., Bucaria, J., Osley, M.A., and Sternglanz, R. (2001). Yeast ASF1 protein is required for cell cycle regulation of histone gene transcription. Genetics *158*, 587–596.

Tang, Y., Poustovoitov, M.V., Zhao, K., Garfinkel, M., Canutescu, A., Dunbrack, R., Adams, P.D., and Marmorstein, R. (2006). Structure of a human ASF1a-HIRA complex and insights into specificity of histone chaperone complex assembly. Nat. Struct. Mol. Biol. *13*, 921–929.

Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. Nat. Struct. Mol. Biol. *14*, 1025–1040.

Thatcher, T.H., and Gorovsky, M.A. (1994). Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. Nucleic Acids Res. *22*, 174–179.

Trievel, R.C., Beach, B.M., Dirk, L.M.A., Houtz, R.L., and Hurley, J.H. (2002). Structure and catalytic mechanism of a SET domain protein methyltransferase. Cell *111*, 91–103.

Tsubota, T., Berndsen, C.E., Erkmann, J.A., Smith, C.L., Yang, L., Freitas, M.A., Denu, J.M., and Kaufman, P.D. (2007). Histone H3-K56 acetylation is catalyzed by histone chaperone-dependent complexes. Mol. Cell *25*, 703–712.

Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T., and Kadonaga, J.T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. Nature *402*, 555–560.

van Leeuwen, F., Gafken, P.R., and Gottschling, D.E. (2002). Dot1p modulates silencing in yeast by methylation of the nucleosome core. Cell *109*, 745–756.

Vickerman, K. (1985). Developmental cycles and biology of pathogenic trypanosomes. Br. Med. Bull. *41*, 105–114.

Wang, F., Li, G., Altaf, M., Lu, C., Currie, M.A., Johnson, A., and Moazed, D. (2013). Heterochromatin protein Sir3 induces contacts between the amino terminus of histone H4 and nucleosomal DNA. Proc. Natl. Acad. Sci. U.S.A. *110*, 8495–8500.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191.

Weber, C.M., and Henikoff, S. (2014). Histone variants: dynamic punctuation in transcription. Genes Dev. *28*, 672–682.

Westheimer, F.H. (1995). Coincidences, decarboxylation, and electrostatic effects. Tetrahedron *51*, 3–20.

White, C.L., Suto, R.K., and Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. EMBO J. *20*, 5207–5218.

Widom, J. (2001). Role of DNA sequence in nucleosome stability and dynamics. Q. Rev. Biophys. *34*, 269–324.

Wilson, J.R., Jing, C., Walker, P.A., Martin, S.R., Howell, S.A., Blackburn, G.M., Gamblin, S.J., and Xiao, B. (2002). Crystal structure and functional analysis of the histone methyltransferase SET7/9. Cell *111*, 105–115.

Wirtz, E., and Clayton, C. (1995). Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. Science *268*, 1179–1183.

Wu, H., Min, J., Lunin, V.V., Antoshenko, T., Dombrovski, L., Zeng, H., Allali-Hassani, A., Campagna-Slater, V., Vedadi, M., Arrowsmith, C.H., et al. (2010). Structural biology of human H3K9 methyltransferases. PLoS ONE *5*, e8570.

Wu, H., Zeng, H., Dong, A., Li, F., He, H., Senisterra, G., Seitova, A., Duan, S., Brown, P.J., Vedadi, M., et al. (2013). Structure of the catalytic domain of EZH2 reveals conformational plasticity in cofactor and substrate binding sites and explains oncogenic mutations. PLoS ONE *8*, e83737.

Xiao, B., Jing, C., Wilson, J.R., Walker, P.A., Vasisht, N., Kelly, G., Howell, S., Taylor, I.A., Blackburn, G.M., and Gamblin, S.J. (2003). Structure and catalytic mechanism of the human histone methyltransferase SET7/9. Nature *421*, 652–656.

Yu, W., Chory, E.J., Wernimont, A.K., Tempel, W., Scopton, A., Federation, A., Marineau, J.J., Qi, J., Barsyte-Lovejoy, D., Yi, J., et al. (2012). Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors. Nat. Commun. *3*, 1288.

Zentner, G.E., and Henikoff, S. (2013). Regulation of nucleosome dynamics by histone modifications. Nat. Struct. Mol. Biol. *20*, 259–266.

Zhang, W., Hayashizaki, Y., and Kone, B.C. (2004). Structure and regulation of the mDot1 gene, a mouse histone H3 methyltransferase. Biochem. J. *377*, 641–651.

Zhang, X., Wen, H., and Shi, X. (2011). Lysine methylation: beyond histones. Acta. Biochim. Biophys. Sinica *44*, 14–27.

Zhang, X., Zhou, L., and Cheng, X. (2000). Crystal structure of the conserved core of protein arginine methyltransferase PRMT3. EMBO J. *19*, 3509–3519.

Zhang, X., Tamaru, H., Khan, S.I., Horton, J.R., Keefe, L.J., Selker, E.U., and Cheng, X. (2002). Structure of the Neurospora SET domain protein DIM-5, a histone H3 lysine methyltransferase. Cell *111*, 117–127.

Zhang, X., Yang, Z., Khan, S.I., Horton, J.R., Tamaru, H., Selker, E.U., and Cheng, X. (2003). Structural basis for the product specificity of histone lysine methyltransferases. Mol. Cell *12*, 177–185.

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

6 Publications

1. **Dindar G**, Anger AM, Mehlhorn C, Hake SB and Janzen CJ. (2014). Structure-guided mutational analysis reveals the functional requirements for product-specificity of DOT1 enzymes. ***submitted***

2. Pascoalino B*, **Dindar G***, Vieira-da-Rocha JP, Machado CR, J. Janzen CJ and Schenkman S (2014). Characterization of two different Asf1 histone chaperones with distinct cellular localizations and functions in *Trypanosoma brucei*. ***Nucleic Acids Res.*** 42 (5), 2906-2918.
   * These authors contributed equally to this work.

3. Armache JP, Jarasch A, Anger AM, Villa E, Becker T, Bhushan S, Jossinet F, Habeck M, **Dindar G**, Franckenberg S, Marquez V, Mielke T, Thomm M, Berninghausen O, Beatrix B, Söding J, Westhof E, Wilson DN, and Beckmann R (2010). Cryo-EM structure and rRNA model of a translating eukaryotic 80S ribosome at 5.5-Å resolution.
   Proc. Acad. Sci. USA 107, 19748-19753.

4. Armache JP, Jarasch A, Anger AM, Villa E, Becker T, Bhushan S, Jossinet F, Habeck M, **Dindar G**, Franckenberg S, Marquez V, Mielke T, Thomm M, Berninghausen O, Beatrix B, Söding J, Westhof E, Wilson DN, and Beckmann R (2010). Localization of eukaryote-specific ribosomal proteins in a 5.5-Å cryo-EM map of the 80S eukaryotic ribosome.
   Proc. Acad. Sci. USA 107, 19754-19759.