H2A.Z – a molecular guardian of RNA polymerase II transcription in African trypanosomes

H2A.Z – eine molekulare Wächterin der RNA Polymerase II Transkription in Afrikanischen Trypanosomen

Doctoral thesis for a doctoral degree
at the Graduate School of Life Sciences,
Julius-Maximilians-Universität Würzburg,
Section Infection and Immunity
submitted by

Amelie Johanna Kraus

from

Munich, Germany

Würzburg, 2021

Submitted on:

Members of the Doctoral Thesis Committee:

Chairperson:	Prof. Dr. Thomas Dandekar
Primary Supervisor:	Prof. T. Nicolai Siegel, PhD
Supervisor (Second):	Prof. Dr. Christian Janzen
Supervisor (Third):	Dr. Andreas Schlosser
Supervisor (Fourth):	Dr. Sonja Lorenz

Date of Public Defense:

Date of Receipt of Certificate:

"The important thing is not to stop questioning. Curiosity has its own reason for existing. One cannot help but be in awe when he contemplates the mysteries of eternity, of life, of the marvellous structure of reality. It is enough if one tries merely to comprehend a little of this mystery every day. Never lose a holy curiosity."

– Albert Einstein

Table of Content

Affidavit

I hereby confirm that my thesis entitled 'H2A.Z – a molecular guardian of RNA polymerase II transcription in African trypanosomes' is the result of my own work. I did not receive any help or support from commercial consultants. All sources and/or materials applied are listed and specified in the thesis.

Furthermore, I confirm that this thesis has not yet been submitted as part of another examination process neither in identical nor in similar form.

_____ _____

Place, Date Signature

Eidesstattliche Erklärung

Hiermit erkläre ich anEides statt, die Dissertation „H2A.Z – eine molekularer Wächterin der RNA Polymerase II Transkription in Afrikanischen Trypanosomen" eigenständig, d.h. insbesondere selbstständig und ohne Hilfsmittel eines kommerziellen Promotionsberaters, angefertigt und keine anderen als die von mir angegebenen Quellen und Hilfsmittel verwendet zu haben.

Ich erkläre außerdem, dass die Dissertation weder in gleicher noch in ähnlicher Form bereits in einem anderen Prüfungsverfahren vorgelegen hat.

_____ _____

Ort, Datum Unterschrift

Acknowledgements

First and foremost, I would like to express my deepest gratitude to my primary supervisor Prof. Nicolai Siegel. Throughout my time in his laboratory and beyond, he has accompanied my personal development with constant support, thoughtful guidance, invaluable feedback, and motivating badinages ✍ to become a successful scientist as well as the confident and ambitious woman, I am now. He realised my talents beyond science before I knew about them and gave me chances to strengthen them. Thank you!

Thesis Advisory Committee:

I would like to give a big thank you to my second supervisor Prof. Christian Janzen. In our joined meetings, he provided great and invaluable input for experiments, literature, and hypotheses around my PhD project. Moreover, his presence during my TAC meetings created a relaxed and joyful atmosphere that enabled fruitful and creative scientific discussions in this formal setting.

I would also like to thank Dr. Andreas Schlosser not only for being a member of my TAC but more importantly for the great collaboration we had to establish a method for quantifying histone acetylation levels. Without his work, thoughts and input, and the work and support of his lab my PhD project would not have been realisable.

I also would like to express my gratitude to Dr. Sonja Lorenz, who joined my TAC shortly after she started her lab in Würzburg. I could always pass by her office for advice, and she helped me out with *E. coli* strains, when I had problems with overexpressing a protein. Thank you!

Chairperson

I would like to express my gratitude to Prof. Dr. Thomas Dandekar for kindly agreeing to function as chairperson for the examination process of my doctoral thesis despite his numerous other responsibilities.

Siegel laboratory:

First of all, I would like to express gratitude from the deepest of my heart to the former and current members of the Siegel laboratory for the warm atmosphere, solidarity, the fun in- and outside of the lab and your friendship. Especially, I would like to thank:

- Raúl Oscar Cosentino (Ruli): Ruli has been the calm anchor in the lab. It doesn't matter how depressing a day in the lab was, he was there with cookies, jokes, and his calm nature. This made it impossible to be in a bad mood and helped to release the negativity. Moreover, his profound patience in helping me and others with our bioinformatic issues and teaching us the basics of scripting, handling our sequencing data and bioinformatic analyses is unique and is looking for its match.
- Ines Subota: Ines joined the Siegel lab in Munich. From the beginning, we had understood each other without words. We have the same character and share the same values, so a great

collaboration and a deep friendship was predestined. I am very grateful for Ines's presence in the lab. Her knowledge and training in working with trypanosomes was precious and she was always there with advice and an open ear.

- Vanessa Luzak: Vanessa brought a fresh breeze into the Siegel lab. She is fascinated by science, and she has used her fascination for helping everyone in the lab and for promoting all the projects. The discussions with her were always inspiring and often played a decisive role in finding solutions, when I thought there are none. Vanessa, you are a great scientist! I am very happy that we have developed a trustworthy friendship by working together. I am looking forward to following your professional and private path!
- Claudia Rabuffo: Thank you Claudia for the fun and joy we had together. We shared much laughter. Even though, we constantly teased each other, we had known the boundaries of the other one. Through this, a friendship has developed that I would not want to miss in my life.
- Benedikt Brink: After moving to Munich, Benedikt stepped in the footsteps of Konrad as bioinformatician, and he was a worthy successor. I am very grateful for your support, input and collaboration that were essential for both of my first-author papers. You were a great co-author!
- Claudia Thiel: Thank you for helping me to set up the lab in Munich. You were an eager learner and a loyal co-worker. Your help in doing experiments and administering the general lab stuff.
- Ramona, thank you for your professional support and for your accurate organisation and maintenance of the general lab stuff in Würzburg. Thank you also for helping me to set up the lab in Munich. Laura, thank you for the regular and much needed coffee breaks throughout the years.

Janzen Laboratory:

I also thank the former and present members of the Janzen lab, Zdenka, Helena, and Nicole for many valuable discussions in our joint Tuesday meetings.

Schlosser Laboratory:

I would like to express my deepest gratitude to all the former and current members of the Schlosser laboratory for the great collaboration and the development of FIPQuant, the method, which has laid the foundation for all the results presented in the PhD work. Moreover, I would like to say thank you for spontaneously accommodating me, when I had to perform the last FIPquant analyses.

- Dr. Jens T. Vanselow: I am very grateful for working with Jens, who was a driving force for developing FIPQuant. Thank you for your support, input, and equanimity, when explaining the analysis to me over and over again.
- Dr. Rasha ElBashir: I would like to thank Rasha ElBashir, who was brave enough to establish a novel and quantitative method during her PhD that sometimes seem to be unachievable.

- Stephanie Lamer: I am deeply thankful for the support of Stephanie, who took care of preparing my FIPQuant samples and analysing the mass spectrometry results following Rasha and Jens. Thank you, Stephanie, for spending so many hours on my samples, even if you had other things to do. I really appreciate your effort and support.
- Beate Vogt: I am really grateful for the time and effort of Beate, who supported Stephanie in preparing my FIPQuant samples and who taught me how to do it on my own. Thank you for helping me to prepare the gazillion samples on such a short notice and in such a short time.
- Christiane Winkler: I would like to thank Christiane for supporting me in preparing the FIPQuant samples and for the heartful welcome when I joined the laboratory for one week.

Departments:

- IMIB: I would like to thank all the fantastic people at the Institute of Molecular Infection Biology for the scientific environment and their helpfulness and solidarity. A very special thank you goes to my friends at the IMIB in Würzburg: Caroline, Charlotte, Clivia, Malvika, Usha, and Yanjie. Your company made time so great, and you enriched my life so much.
- I am deeply grateful for my former and current members of the CoreUnit SysMed at the IMIB, especially Margarete Göbel, Elena Katzowitsch, Dr. Thorsten Bischler, and Dr. Kristina Döring for their support in setting up and performing Next Generation Sequencing runs. I also would like to express my deepest gratitude to Prof. Konrad Förstner, who always made time for helping me in analysing my sequencing data and who thought me how to work with a HPC.
- TrypsClub: I would also like to thank the whole Tryps-community at Hubland. Thank you for your ideas you provided during our TrypClub meetings.
- I would like to thank all former and current members of the Department of Physiological Chemistry at the Biomedical Center Munich for the warm welcome that helped me to adapt to my new life in Munich, the scientific environment, and the constant support during the last two years of my PhD. I especially would like to point out Prof. Andreas Ladurner and Dr. Anton Eberharter for integrating us into the department and the great work environment. Morover, I would like to thank you for sharing with me your expertise in chromatin biology and for providing me highly appreciated input and feedback during the departmental seminars.
- I would like to thank my colleagues from the Chair for Experimental Parasitology for their support in administration and their enthusiasm in organising joint social activities I enjoyed so much.
- I would like to thank the lab of Prof. Michael Boshart for letting me use their cell culture and medium when I needed it.
- I would like to thank all the members of the Parasitology Meeting (Boshart lab and Meißner Lab) for providing thoughtful ideas, invaluable input, and encouraging feedback during my seminars.

- Many thanks go also to the Graduate School of Life Science for giving me the opportunity to join their PhD programme, for admitting me into their Mentoring programme and for their constant support during my PhD.

Family and Friends:

Last but not least, I would like to express my deepest gratitude to my family and friends that accompanied my PhD journey, especially my study fellows Nane, Lena, and Elisabeth, my botany friends Christof, Jenny and Sönke, and again my friends from the Siegel lab, IMIB, and BMC. We spent many unforgettable days and nights together and I am looking forward to what's coming.

Dr. Carolin Kaufhold-Wedel: Caro, I don't know what I would have done and what I would do without you. Only because of you, I am where I am now, and my life has developed like it has. I am so happy that I have you in my life. Your positivity lighted up every dark moment. Thank you for letting me into your life and for entering mine!

Dr. Mario Jaborsky: My heartfelt appreciation goes to you. Even though, we had just met, you were my solid rock in the turbulent waters after leaving my beloved Würzburg. You constantly have believed in me and more importantly, you have made it possible that I believe in myself. I enjoy every single second of our life and love you from the deepest of my heart.

Zum Schluss möchte ich noch die Gelegenheit nutzen, mich bei meinen Eltern, Angela und Werner, und meiner Schwester Lisabet für ihre bedingungslose Liebe und Unterstützung zu bedanken. Wenn ich euch brauchte, wart ihr sofort mit Rat und Tat zur Stelle. Eure Liebe und Wärme geben mir die Stärke und sind der Antrieb, meine eigenen Wege zu gehen. Euer Rückhalt ist der Grundstein für meinen Erfolg. Ich euch diese Arbeit, denn es ist zum Großteil euer Verdienst, dass ich so weit gekommen bin und auch weiß, was ich noch erreichen möchte. Ich habe euch lieb!

Abbreviation Index

Abbreviation	Full name
%	Percentage
°C	Degree Celsius
®	Registered trade mark
Å	Angström
ac	Acetylation
ADP	Adenosindiphosphat
Alanine	Amino terminal
Arginine	R
Asparagine	N
Aspartic acid	D
ATP	Adenosintriphosphat
BARP	bloodstream alanine-rich proteins
BDF	Bromodomain factor
BDF1	Bromodomain factor 1
BDF2	Bromodomain factor 2
BDF3	Bromodomain factor 3
BDF4	Bromodomain factor 4
BES	Bloodstream form expression site
BF	Bloodstream form
Ble	Phleomycin
Bp	Basepair
BSD	Blasticidin
C-terminal	Carboxyl terminal
CC	Closed complex
CenpAs	Centromeric protein A
CHD	Chromodomain helicase DNA-binding
ChIP	Chromatin immunoprecipitation
Chr	Chromosomes
CRC	Chromatin remodelling complexes
CRISPR/Cas9	Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated 9
CTD	Carboxyl-terminal domain
Cysteine	C
D. melanogaster	Drosophila melanogaster
Da	Dalton
DNA	Desoxyribonucleic acid
dNTP	Desoxyribonucleosid triphosphate
DOT1	Disruptor of telomeric silencing
Dox	Doxycycline
dsDNA	double stranded DNA
dTSS	divergent TSS
DTT	Dithiothreitol
E-value	Expect value
e.g.	exempli gratia
EC	Elongation complex
EDTA	Ethylenediaminetetraacetic acid
ERCC	External RNA Control Consortium
et al.	et alteres
FACT	Facilitates chromatin transcription
g	Units of times gravity
G418	Neomycin
gDNA	genomic DNA
Glutamic acid	E
Glutamine	Q
Glycine	G
GPI	Glycosylphosphatidylinositol
GTFs	General transcription factors
h	hour
H. sapiens	Homo sapiens
HAT	Human African Trypanosomiasis
HAT	Histone acetyltransferase
HAT1	Histone acetyltransferase 1
HAT2	Histone acetyltransferase 2
HCD	Higher-energy collisional dissociation
HDACs	Histone deacetylases
HF	High-fidelity
Histidine	H
HS	High-sensitivity
Hyg	Hygromycin
IgG	Immunoglobulin G
IgM	Immunoglobulin M
INO80	Inositol requiring 80
Abbreviation	**Full name**
Isoleucine	I

ISWI	Imitation switch
JBP2	base J binding protein
Kb	kilobase
KD	Knock down
kDA	Kilodalton
LC	Liquid chromatography
Leucine	L
log	logarithmic
Lysine	K
M	Molar
Mb	Megabase
me1	Monomethylation
me2	Dimethylation
me3	Trimethylation
Methionine	M
Mg	Magnesium
min	Minute
ml	Milliliter
ml	Meter
mM	Millimolar
MNase	Micrococcal nuclease
mRNA	messenger-RNA
MS	Mass-spectrometry
MYST	Moz, Ybf2/Sas3, Sas2, Tip60
mz	mass/charge
N-terminal	Amino terminal
NDR	Nucleosome-depleted region
ndTSS	non-divergent TSS
NEB	New England Biolabs
Neo	Neomycin
ng	Nanogramm
NuA4	nucleosome acetyltransferase of H4
OC	Open complex
ORF	Open reading frame
PAD1	Proteins Associated with Differentiation 1
PCR	Polymerase chain reaction
Abbreviation	**Full name**
PF	Procyclin form
PhD	Philosophical doctor
PIC	Preinitiation complex
PMSF	Phenylmethylsulfonylfluorid
poll	Polymerase
poly(A)-tail	Poly-Adenin-Tail
poly(Y)-tract	Polypyrimidine
ppm	parts per million
Proline	P
PTE	Polycistronische Transkriptionseinheiten
PTM	Post-translational modification
PTU	Polycistronic transcription unit
PUR	Puromycin
RNA	Ribonucleic acid
RNAi	RNA interference
RNAP	RNA polymerase
RNAPI	RNA polymerase 1
RNAPII	RNA polymerase 2
RNAPIII	RNA polymerase 3
Rpb	DNA-directed RNA polymerase II subunit
rRNA	ribosomal RNA
RT	Room temperature
S-phase	Synthesis phase
S. cerevisiae	Saccheromyces cerevisiae
SAGA	Spt-Ada-Gcn5 acetyltransferase
SANT/SLIDE	SWI3, ADA2, N-CoR, and TFIIIB/SANT-like ISWI domain
SAS	Splice acceptor site
SDS	Spice donor site
sec	Second
Serine	S
SL	Spliced Leader
SM	Single marker
SNF2	Sucrose non-fermentable
SWI	Switch
Swr1	SWI2/SNF2-Related 1
T. brucei	Trypanosoma brucei
Abbreviation	**Full name**
T. termophilia	Tetrahymena termophilia
TAE	Tris-Acetic acid-EDTA
TAFs	TBP-associated factors

Tb	Trypanosoma brucei
TBP	TATA-binding protein
TDB	Trypanosome dilution buffer
Tet	Tetracycline
TF	Transcription factor
Threonine	T
TLCK	Tosyl-L-lysyl-chloromethane hydrochloride
tRNA	transfer-RNA
Tryptophan	W
TSS	Transcription start site
Abbreviation	**Full name**
TTS	Transcription termination site
TY	TY-tag
Tyrosine	Y
U	Unit
UTR	Untranslated region
UV	Ultraviolet
Valine	V
VSG	Variant Surface Glycoprotein
WT	wild type
XML	Extensible Markup
µl	Microliter
µm	Micrometer
µM	Micromolar

List of Figures

List of Appendix Figures

List of Tables

Summary

In eukaryotes, the enormously long DNA molecules need to be packaged together with histone proteins into nucleosomes and further into compact chromatin structures to fit it into the nucleus. This nuclear organisation interferes with all phases of transcription that require the polymerase to bind to DNA. During transcription – the process in which the hereditary information stored in DNA is transferred to many transportable RNA molecules - nucleosomes form a physical obstacle for polymerase progression. Thus, transcription is usually accompanied by processes mediating nucleosome destabilisation, including post-translational histone modifications (PTMs) or exchange of canonical histones by their variant forms. To the best of our knowledge, acetylation of histones has the highest capability to induce chromatin opening. The lysine modification can destabilise histone-DNA interactions within a nucleosome and can serve as a binding site for various chromatin remodelers that can modify the nucleosome composition. For example, H4 acetylation can impede chromatin folding and can stimulate the exchange of canonical H2A histone by its variant form H2A.Z at transcription start sites (TSSs) in many eukaryotes, including humans. As histone H4, H2A.Z can be post-translationally acetylated and as acetylated H4, acetylated H2A.Z is enriched at TSSs suggested to be critical for transcription. However, thus far, it has been difficult to study the cause and consequence of H2A.Z acetylation.

Even though, genome-wide chromatin profiling studies such as ChIP-seq have already revealed the genomic localisation of many histone PTMs and variant proteins, they can only be used to study individual chromatin marks and not to identify all factors important for establishing a distinct chromatin structure. This would require a comprehensive understanding of all marks associated to a specific genomic locus. However, thus far, such analyses of locus-specific chromatin have only been successful for repetitive regions, such as telomeres.

In my doctoral thesis, I used the unicellular parasite *Trypanosoma brucei* as a model system for chromatin biology and took advantage of its chromatin landscape with TSSs comprising already 7% of the total *T. brucei* genome (humans: 0.00000156%). Atypical for a eukaryote, the protein-coding genes are arranged in long polycistronic transcription units (PTUs). Each PTU is controlled by its own ~10 kb-wide TSS, that lies upstream of the PTU. As observed in other eukaryotes, TSSs are enriched with nucleosomes containing acetylated histones and the histone variant H2A.Z. This is why I used *T. brucei* to particularly investigate the TSS-specific chromatin structures and to identify factors involved in H2A.Z deposition and transcription regulation in eukaryotes. To this end, I established an approach for locus-specific chromatin isolation that would allow me to identify the TSSs- and non-TSS-specific chromatin marks. Later, combining the approach with a method for quantifying lysine-specific histone acetylation levels, I found H2A.Z and H4 acetylation enriched in TSSs-nucleosomes and mediated by the histone acetyltransferases HAT1 and HAT2. Depletion of HAT2 reduced the levels of TSS-specific H4 acetylation, affected targeted H2A.Z deposition and shifted the sites of transcription initiation. Whereas HAT1

depletion had only a minor effect on H2A.Z deposition, it had a strong effect on H2A.Z acetylation and transcription levels. My findings demonstrate a clear link between histone acetylation, H2A.Z deposition and transcription initiation in the early diverged unicellular parasite *T. brucei*, which was thus far not possible to determine in other eukaryotes. Overall, my study highlights the usefulness of *T. brucei* as a model system for studying chromatin biology. My findings allow the conclusion that H2A.Z regardless of its modification state defines sites of transcription initiation, whereas H2A.Z acetylation is essential co-factor for transcription initiation. Altogether, my data suggest that TSS-specific chromatin establishment is one of the earliest developed mechanism to control transcription initiation in eukaryotes.

Zusammenfassung

In Eukaryoten muss die genomische DNA zusammen mit Histonproteinen zu Nukleosomen und weiter zu kompakten Chromatinstrukturen verpackt werden, damit sie in den Zellkern passt. Diese Organisation behindert die Transkription bei jedem Schritt, bei dem die Polymerase an der DNA bindet. Während der Transkription – dem Prozess bei dem die in der DNA gespeicherte Erbinformation in viele transportable RNA Molekülen umgewandelt wird – stellen Nukleosomen ein physikalisches Hindernis für das Vorankommen der Polymerase dar. Aus diesem Grund wird die Transkription üblicherweise von Prozessen begleitet, die für die Destabilisierung der Nukleosomen sorgen, wie zum Beispiel post-translationale Modifizierung (PTM) der Histone oder der Austausch von kanonischen Histonproteinen durch eine ihrer Varianten. Soweit bisher bekannt ist Histonacetylierung am besten dafür geeignet, eine offene Chromatinstruktur bereit zu stellen. Die Lysinmodifizierung kann Interaktionen zwischen der DNA und den Histonen innerhalb eines Nukleosomes destabilisieren und als Andockstelle für einige Proteinkomplexe sogenannte Chromatin-Modellierer fungieren, die die Zusammensetzung eines Nukleosomes verändern können. Zum Beispiel, kann Acetylierung am Histon H4 das „Zusammenfalten" des Chromatins erschweren und den Austausch von kanonischem H2A mit ihrer Variante H2A.Z an den Transkriptiosinitiationsstellen (TSSen) in vielen eukaryotischen Organismen, Menschen eingeschlossen, stimulieren. Wie Histon H4, kann auch H2A.Z post-translationell acetyliert werden und wie acetyliertes H4, findet man auch acetyliertes H2A.Z vor allem an TSSen. Deswegen geht man davon aus, dass es sehr wichtig für die Transkriptioninitiierung ist. Allerdings war es bisher nicht möglich, die Ursache und Wirkung von H2A.Z Acetylierung genauer zu untersuchen.

Genom-weite Chromatinprofilstudien wie z.B. ChIP-Seq ermöglichen es die genomische Lokalisierung von vielen Histon-Modifizierungen und -Varianten zu bestimmen. Dennoch reichen sie nicht dafür aus alle Faktoren, die für die Bildung einer bestimmten Chromatinstruktur notwendig sind, gleichzeitig herauszufinden. Das würde voraussetzen, dass man alle Merkmale der genomischen Stelle kennt. Bisher waren Analysen von spezifischen Chromatinstellen nur erfolgreich, wenn das Chromatin von einer repetitiven Region, wie z.B. Telomeren, stammt.

In meiner Doktorarbeit verwendete ich den einzelligen Parasiten *Trypanosoma brucei* als Modelsystem für Chromatinbiologie. Dabei machte ich mir dessen Chromatinorganisation zunutze, die eher untypisch für einen eukaryotische Organismus ist. TSSen machen hier ungefähr 7% des gesamten Genoms aus (Mensch: 0.00000156%). Protein-kodierende Gene sind in langen polycistronischen Transkriptionseinheiten (PTE) angeordnet. Jede dieser Einheiten besitzt eine eigene TSS, die vor der PTE liegt, und bis zu 10 kb lang sein kann. Jedoch, wie in anderen Eukaryoten, sind an den TSSen Nukleosomen angereichert, die sich durch acetylierte Histone und den Einbau der Histonvariante H2A.Z auszeichnen. Aus diesen Gründen verwendete ich *T. brucei*, um während meiner Doktorarbeit die Chromatinstrukturen, die TSSen auszeichnen, genauer zu untersuchen und die Faktoren, die bei der

H2A.Z Positionierung und dadurch bei der Transkriptionsregulation in Eukaryoten eine Rolle spielen, herauszufinden. Dafür etablierte ich zuerst eine Methode, mit der man Chromatin von einer bestimmten genomischen Stelle isolieren kann und die es mir ermöglichen würde, die Merkmale von TSS-spezifischen und -unspezifischen Chromatin zu identifizieren. Später konnte ich das entwickelte Protokoll mit einer Methode zur Quantifizierung von Lysin-spezifischen Histonacetylierung kombinieren. Dadurch konnte ich herausfinden, dass Nukleosomen an trypanosomischen TSSen stark acetyliertes H2A.Z und H4 enthalten und dass für diese Modifizierungen die Histonacetyltransferasen HAT1 und HAT2 verantwortlich sind. Eine Reduzierung der HAT2-Levels führte zu einer Reduzierung von H4 Acetylierung, verschlechterte die gezielte H2A.Z Positionierung und führte dazu, dass die Transkriptioninitiierung sich verlagerte. Wohingegen eine Reduzierung von HAT1, die zwar nur einen kleinen Einfluss auf die H2A.Z Positionierung hatte, eine sehr starke Verringerung von acetyliertem H2A.Z und der Transkriptionsrate zur Folge hatte. Durch meine Ergebnisse konnte ich zeigen, dass in *T. brucei*, einem evolutionär divergenten eukaryotischem Organismus, die Prozesse der Histonacetylierung, H2A.Z Positionierung und Transkriptionsinitiierung sehr stark miteinander verbunden sind. Meine Arbeit ist des weiteren ein Beweis dafür, dass *T. brucei* ein sehr wichtiger Modellorganismus für die Forschung an Chromatin ist. Insgesamt erlauben meine Ergebnisse die Schlussfolgerung, dass H2A.Z, egal ob modifiziert oder nicht, ein Herausstellungsmerkmal für TSSen ist, während acetyliertes H2A.Z essentiell für die Transkriptionsinitiierung darstellt. Zusammengefasst, weisen die Daten meiner Doktorarbeit darauf hin, dass die Etablierung von bestimmten Chromatinstrukturen an TSSen eines der frühesten entwickelten Mechanismen zur Kontrolle der Transkriptionsinitiierung in Eukaryoten ist.

Chapter 1: Introduction

Trypanosoma brucei – an attractive model organism

More than 300 million years ago, the salivarian trypanosomes –single-celled, protozoan parasites – emerged for the first time including the subspecies *T. brucei brucei* (Haag et al., 1998). Decades of research have proved *T. brucei* as a valuable model organism for many aspects of eukaryotic cell biology leading to discoveries such as Glycosylphosphatidylinositol(GPI)-mediated protein anchoring (Ferguson et al., 1988), RNA-editing (Benne et al., 1986) and *trans-splicing* (De Lange et al., 1984; Parsons et al., 1984).

One advantage of using *T. brucei* as a model organism is that the handling is rather simple. It can be easily cultured in liquid medium and grown to high densities (up to 4 x 10^7 cells/ml). In addition, many tools for genetic manipulation are available allowing gene knockouts, gene knockdowns using the intrinsic RNA-interference machinery, epitope-tagging of endogenous genes and overexpression of recombinant proteins. One recent breakthrough has been the adaptation of the CRISPR/Cas9-system allowing resistance marker-free genome editing at nucleotide resolution in *T. brucei* (Beneke et al., 2017; Rico et al., 2018; Vasquez et al., 2018).

For my PhD thesis, I decided to use *T. brucei* to study the most fascinating aspects of eukaryotic cell biology: DNA accessibility and transcription. In eukaryotes, DNA is not a free molecule, but packaged into a higher order structure, which makes it possible to fit the long range of DNA (2 m in *Homo sapiens*) into the small nucleus (~10 µM diameter) (Kornberg, 1974; Guo, 2014). The DNA compaction interferes with all DNA-templated processes including transcription (Venkatesh and Workman, 2015). While in most eukaryotes, transcription is controlled by many complex and inter-dependent processes, it has been assumed that in *T. brucei* transcription relies simply on the degree of DNA accessibility making *T. brucei* an excellent model organism to unravel the basic requirements for transcription.

General overview

T. brucei belongs together with its close relatives *Leishmania major* and *Trypanosoma cruzi* to the class of kinetoplastea (order: Trypanosomatida; family: Trypanosomatidae; genus: *Trypanosoma*). All three cause medically and economically relevant diseases, which are in case of *T. brucei* nagana in cattle or Human African Trypanosomiasis, better known as sleeping sickness. Thus, *T. brucei* parasites are also known as African trypanosomes. It should be noted that I will refer to *T. brucei* as trypanosomes in this document.

With exception of a few cases of asymptomatic patients or patients, who seem to have cleared the *T. brucei* infection (Jamonneau et al., 2012; Sudarshi et al., 2014; Capewell et al., 2016), untreated sleeping sickness is lethal. Since 2000, the monitored cases have reduced about 90% and in 2017, less than

1500 cases have been diagnosed. However, 57 million people in 36 countries are still at continuous risk of getting infected (Organization, 2018). Human African Trypanosomiasis occurs almost exclusively in Sub-Saharan Africa and is spatially limited to the presence of its arthropod vector, the tsetse fly of the *Glossina* species (Malvy and Chappuis, 2011). The species *T. brucei* includes three subspecies, *T. brucei gambiense* (causes chronic HAT; central and western Africa), *T. brucei rhodensiense* (causes acute HAT; southern and eastern Africa), and *T. brucei brucei* (animal trypanosomiasis; not infective for human) (Barrett et al., 2003). Drugs for treatment of human sleeping slickness are still unsatisfactory due to their high toxicity, low efficacy and high costs. The standard drugs, available on the market, are Suramin, Pentamidine, Melarsoprol, Eflornithine, and Nifurtimox, which is applied as a combination therapy with Eflornithine (Field et al., 2017).

Life cycle

Trypanosomes undergo a complex life cycle, which requires adaptation to different environments within the bloodstream of the mammal host and within the insect host, both accompanied by changes in morphology, energy pathway and surface proteins (**Figure 1.1**) (Vickerman, 1985).

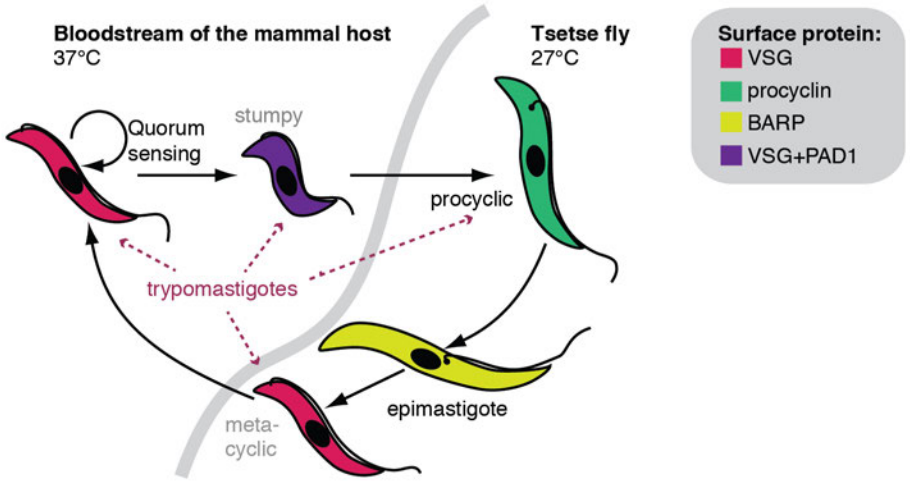


Figure 1.1. The life cycle of *Trypanosoma brucei*.
The scheme shows the different developmental stages in the mammal host (left) and in the insect vector (right). The colour of the parasite refers to the surface antigen, which is expressed in the life cycle stage. The kinetoplast and flagellar base are posterior to the nucleus (black circle) in the trypomastigotes (dotted arrows), and anterior in the epimastigote parasite. The illustration was modified from (Clayton, 2019).

Within the mammalian host, *T. brucei* lives extracellularly in the bloodstream as proliferative long-slender bloodstream forms (BF) that multiply by binary fission (Matthews et al., 2004). Increasing trypanosome-density elevates the levels of the stumpy inducing factor, which promotes trypanosome quorum sensing by oligopeptide signalling and stimulates the differentiation from proliferating BFs to non-dividing short-stumpy forms (Rojas et al., 2019). The tsetse fly takes up a mixture of long-slender and short-stumpy parasites during its bloodmeal. While it was thought that only the short-stumpy form can proceed successfully to the procyclic stage, the proliferating stage in the midgut of the tsetse fly, and the long-

slender parasites die after uptake because of the proteolytic environment in the insects' midgut (Matthews et al., 2004), recent findings suggest that BFs are also able to infect the tsetse fly (Schuster et al., 2019). The procyclics (PF) migrate from the midgut to the tsetse fly's salivary gland, where they differentiate into dividing epimastigotes and further into non-proliferating metacyclics. Metacyclics are injected in the mammal host during the next blood meal of the vector, where they differentiate to BFs (Vickerman, 1985).

In the bloodstream, *T. brucei* cells are under constant attack by the host's immune system. Nevertheless, trypanosomal infections can persist for years without successful eradication by the immune response (MacGregor et al., 2012). The immune response targets the surface coat proteins of the BFs generating specific IgM and IgG antibodies to neutralise the fluctuating parasites (Matthews et al., 2004; Malvy and Chappuis, 2011). The surface of *T. brucei* is coated with a dense, homogenous layer of ~10 million identical copies of one variant surface glycoprotein (VSG) (Fenn and Matthews, 2007) serving as a protection barrier against the immune response (Donelson, 2003; Matthews et al., 2004). By changing the surface coat through switching to another major VSG, trypanosomes can hide themselves from the circulating antibodies until the immune system produces another set of antibodies (Donelson, 2003). This process is known as antigenic variation and occurs stochastically and periodically leading to oscillating parasitaemia (**Figure 1.2**) (reviewed in (MacGregor et al., 2012)).

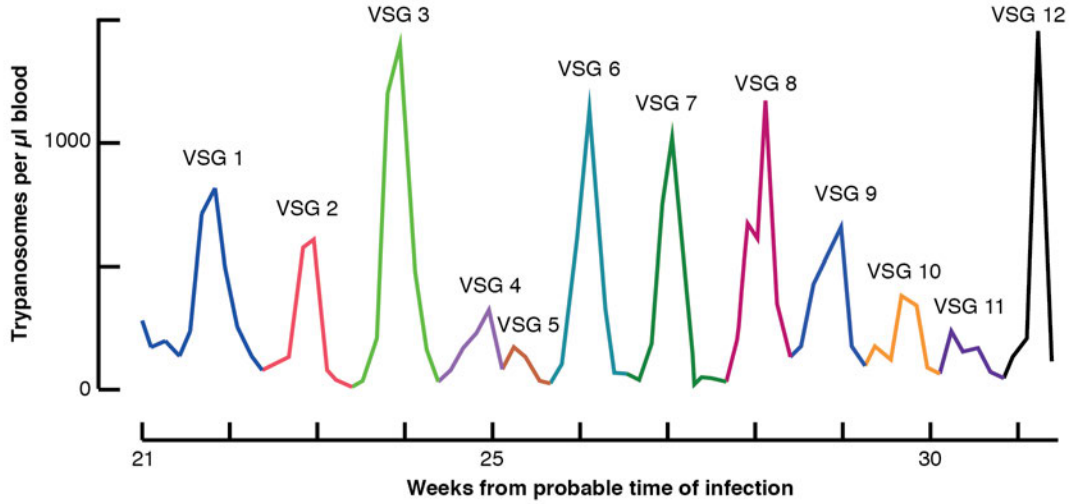


Figure 1.2. Antigenic variations causes oscillating parasitaemia.
Levels of parasitaemia are shown as the number of parasites per µl blood over a time period of around 80 days. The blood was taken from a 26-year-old patient from Northumberland, UK, who was infected with trypanosomes in N. E. Rhodesia (today Zambia) in September 1909. The concentration of parasites in the patient's blood increases and decreases periodically. Each wave of parasitaemia is a consequence of a switch in VSG expression, the major surface antigen of the parasite in the mammalian host, progressively exhausting the host's immune system (MacGregor et al., 2012). Illustration was redrawn from (Ross and Thomson, 1910) by T.N. Siegel.

VSGs are only transcribed if they are located in a telomere-adjacent polycistronic transcription unit (PTU), called expression site (Horn, 2014). In the bloodstream form parasite, around 15 bloodstream form expression sites (BESs) are present, but only one of them is transcriptionally active ensuring

mutually exclusive VSG expression (Borst, 2002). The trypanosomal genome harbours more than 2,500 VSG and pseudo-VSG genes (Cross et al., 2014) mostly found in the subtelomeres of the megabase chromosomes (Berriman et al., 2005) providing an extensive repertoire of surface proteins for trypanosomes.

Genome organisation in Trypanosoma brucei

The genome of the most commonly used laboratory strain *T. brucei* Lister 427 isolate has a total size of 42.25 Mb (Müller et al., 2018) consisting of 11 megabase chromosomes, stored in a nucleus with a diamenter of ~2.5 µM (Ersfeld et al., 1999), and harbours ~9,000 protein-coding genes (Berriman et al., 2005).

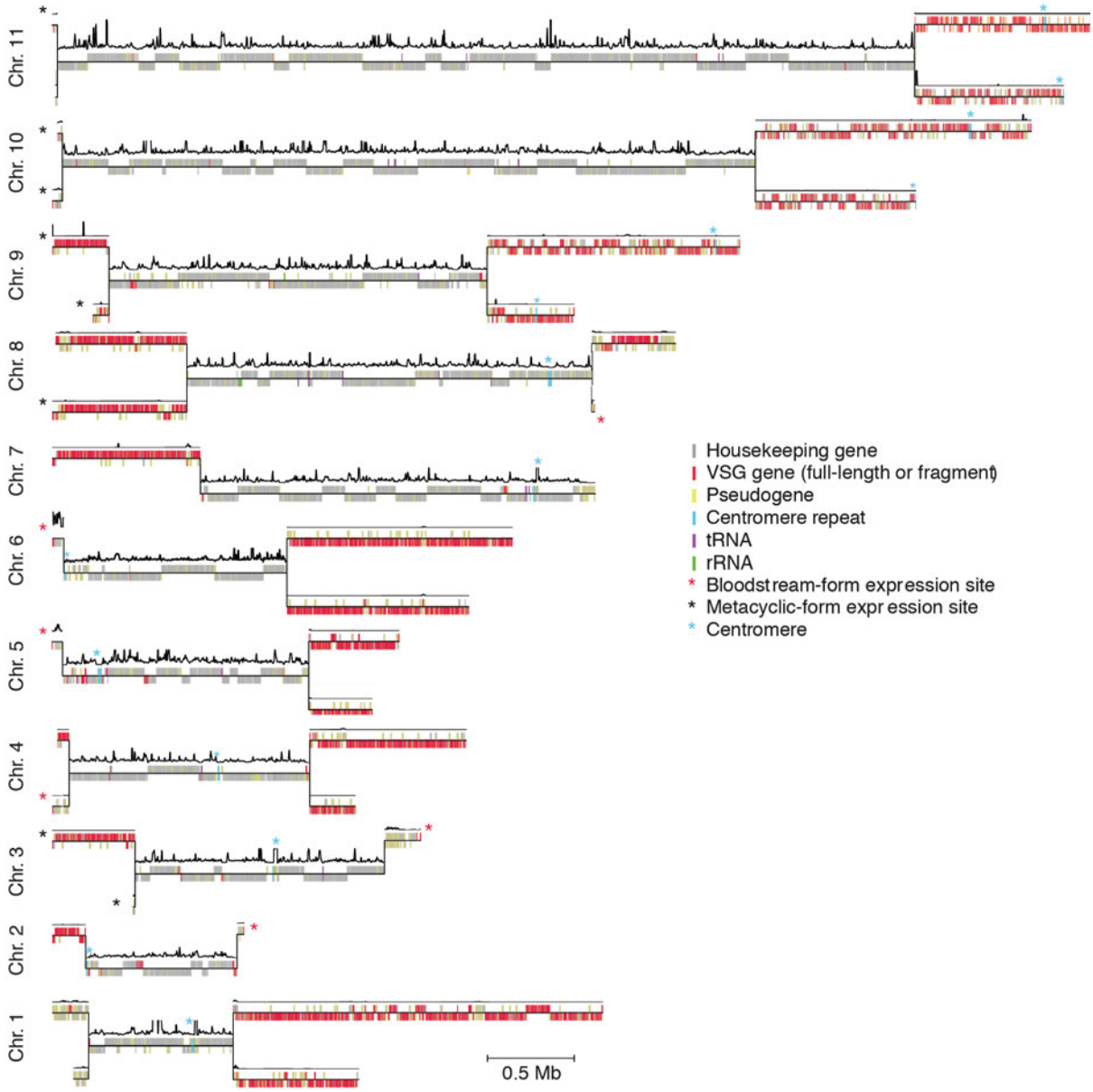


Figure 1.3. The genome architecture of *T. brucei*.
In 2018, the genome of the major laboratory *T. brucei* strain was *de novo* assembled. The scheme shows only one of the two homologues core region of each chromosome (chr.), whilst for the subtelomeric regions both heterozygous arms are shown. Above each chromosome line the relative transcript levels are shown. Specific genomic regions are highlighted in a

corresponding colour. Annotated centromeres, bloodstream- and metacyclic-form expression sites are assigned to the corresponding chromosomes (Asterix). The illustration was taken from (Müller et al., 2018).

The megabase chromosomes are compartmentalised into a transcriptionally active, homozygous core (22.71 Mb; diploid) and a transcriptionally silent, heterozygous, subtelomeric region (19.64 Mb; haploid) (**Figure 1.3**) (Müller et al., 2018). In addition, the nuclear genome contains around five intermediate (200 to 900 kb in length) (Berriman et al., 2005; Hertz-Fowler et al., 2007) and ~100 mini-chromosomes (50-150 kb in length) (Hertz-Fowler et al., 2007; Cross et al., 2014).

Unlike in other eukaryotes (excluding its kinetoplastid relatives) studied so far, the protein-coding genes of trypanosomes are arranged, functionally unrelated, in long arrays called polycistronic transcription units (PTUs) (Myler et al., 1999) (**Figure 1.4**). The genome of *T. brucei* contains ~200 PTUs, each transcribed as polycistronic pre-mRNA molecule (Martinez-Calvillo et al., 2010). Within an individual PTU, transcription occurs from the same strand, whereas two neighbouring PTUs can be transcribed also in the opposing direction. Strand-switch regions, arrays of rRNA or tRNA gene clusters separate two PTUs from each other. Within a strand-switch region, the transcription converges or diverges, depending on the transcription direction of the two neighbouring PTUs. In contrast to higher eukaryotes and yeast, transcription in trypanosomes is not initiated at a specific site, but instead within a broad region upstream of each PTU (Wedel et al., 2017), termed a transcription start site (TSS). Downstream of the last gene of the PTU, transcription terminates within a so-called transcription termination site (TTS) (Siegel et al., 2011).

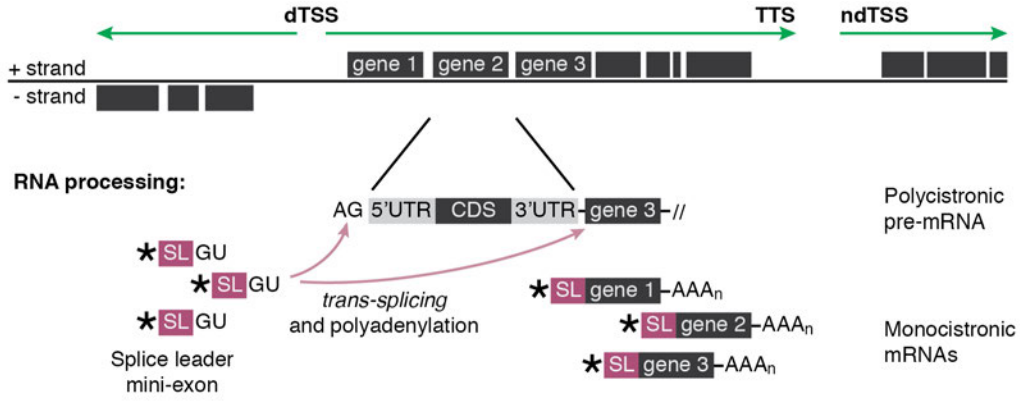


Figure 1.4. mRNA synthesis and maturation in *T. brucei*.
Schematic illustration of an exemplary RNAPII-transcribed PTU. Protein-coding genes are shown in black. Each gene consists of the 5´UTR (grey), CDS and 3´UTR (grey). Green arrows show the direction of transcription. Divergent TSSs are strand-switch regions, whereas the direction of transcription initiating from a non-divergent TSSs follows the direction of the preceding PTU. The genes within one PTU are transcribed polycistronically. RNA processing via *trans*-splicing separates the polycistronic pre-mRNA into the individual mRNA molecules. During *trans*-splicing, the spliced leader (SL) mini-exon is added to each 5´UTR, which also contains the cap-structure, and each 3´UTR is polyadenylated. After RNA-processing, the monocistronic mRNAs can be exported from the nucleus.

Thus far, there is very little understanding of how and why, in an organism that lacks transcriptional control, transcription initiates only within defined regions. Previous findings let us assume that the

degree of DNA packaging is important and transcription initiation is favoured by increased DNA accessibility at specific sites (McAndrew et al., 1998).

Chromatin

Within a eukaryotic nucleus, the genomic material is organised into a complex structure composed of DNA and proteins, called chromatin. Using light microscopy and basophil dyes, chromatin can be visualised and appears as scaffold of threads (Bickis and von Bertalanffy, 1956). Based on findings made in interphase nuclei, chromatin has been subdivided into two types according to its compaction level. Heterochromatin does not change its condensation state throughout the cell cycle and stays compacted, whereas euchromatin is completely decondensed during interphase. Heterochromatic regions are visible as darker stains at the periphery of the nucleus surrounding the lighter stained euchromatin in the nucleoplasm. Moreover, euchromatin is more susceptible to nuclease digestion (Weintraub and Groudine, 1976; McGhee et al., 1981; Sabo et al., 2006), suggesting that less proteins are bound to DNA within euchromatic regions compared to heterochromatin and thus easier accessible for proteins involved in DNA-templated processes, such as transcription (Campos and Reinberg, 2009). Indeed, TSSs and actively transcribed regions are enriched in euchromatin (Boyle et al., 2008; Shi et al., 2009), whereas heterochromatin is rather gene-poor and forms at repetitive elements such as telomeres and centromeric regions (Stralfors and Ekwall, 2011).

In contrast to other eukaryotes, trypanosomal chromatin cannot be simply classified into transcriptionally active euchromatin and transcriptionally silent heterochromatin. Electron microscopy of interphase nuclei showed areas in the nucleoplasm of different electron density and electron dense material clusters at the nuclear periphery, however less pronounced than in other eukaryotes (Ogbadoyi et al., 2000). Due to its distinct genome organisation and unregulated transcription, it has been suggested that trypanosomes do not possess much functional heterochromatin compared to complex eukaryotes. The only potential heterochromatic regions in *T. brucei* are the telomeres, subtelomeres (excluding the active BES), and the transcriptionally silent mini-chromosomes (Ogbadoyi et al., 2000). In general, trypanosomal chromatin seems to be more loosely packed than mammalian chromatin (Belli, 2000). In all eukaryotes, chromatin consists of repeating arrays of its fundamental unit: the nucleosome.

The nucleosome

The canonical nucleosome (Kornberg, 1974; Olins and Olins, 1974; Germond et al., 1976) consists of 147 base pairs of DNA wrapped 1.65 left-handed super-helical turns around a histone octamer containing two copies of each canonical histone H2A, H2B, H3 and H4 (Felsenfeld, 1978; Luger et al., 1997) (**Figure 1.5 a**). Canonical histones are exclusively expressed during S-phase from multiple gene arrays, which lack intron sequences, and are incorporated into chromatin in a replication-dependent manner (Buschbeck and Hake, 2017). Near the entry and exit site of the nucleosome, protecting the

inter-nucleosomal DNA, sits the linker histone H1, which stabilises the nucleosome (Hergeth and Schneider, 2015) and is required for higher-order chromatin architecture (Bednar et al., 1998).

Histones are small, basic proteins, which are evolutionary conserved. The four core histones share a common C-terminal histone fold domain, consisting of three α-helices that enable the formation of heterodimers between H2A with H2B and H3 with H4 (Venkatesh and Workman, 2015). Interaction between the H3 α-helices connects two H3-H4 dimers forming the central tetramer of the nucleosome (Luger et al., 1997). Further interactions of H2B with H4 bring two separate H2A-H2B-dimers to each side of the tetramer and, in the presence of DNA, completes the assembly of a single nucleosome (Kamakaka and Biggins, 2005) (**Figure 1.5 b**). As a result, removal of the H3-H4-tetramer requires preceding displacement of H2A-H2B-dimers (Crick, 1968).

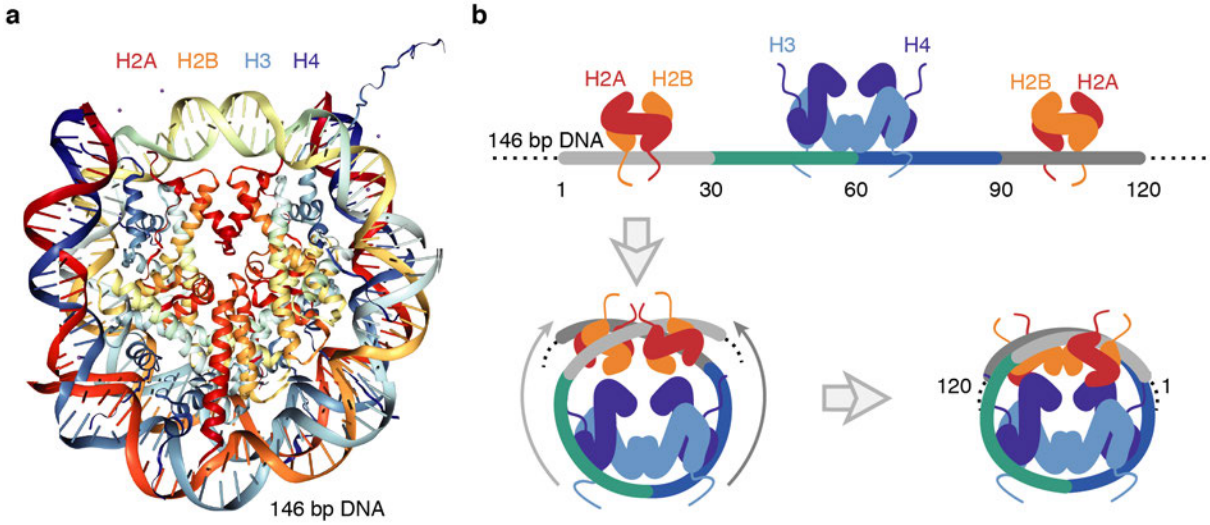


Figure 1.5. The canonical nucleosome.
a Crystal structure of the canonical nucleosome (Luger et al., 1997) downloaded from The Protein Data Bank rcsb.org (Berman et al., 2000). **b** The schematic illustration shows the step-wise assembly of the canonical nucleosome. 146 bp DNA wraps around a histone octamer consisting of two copies of two heterodimers H2A-H2B and H3-H4. The two H3-H4 heterodimers build the central tetramer. The interaction between H4 and H2B brings two separate H2A-H2B heterodimer to each side of the nucleosome. The interaction of the histones with DNA completes the nucleosome assembly. The illustration was adapted from (Luger and Richmond, 1998).

Within a nucleosome, the basic histone proteins have 14 contact points to the negatively charged DNA, making the nucleosome one of the most stable protein-DNA complexes (Luger et al., 1997). However, nucleosomes are not fixed components and regulation of nucleosomal composition accompanies all cellular processes which require accessible DNA as template (Li et al., 2007).

Histones of *T. brucei* are highly divergent

While mammal genomes contain only two major histone gene clusters (Marzluff and Duronio, 2002), the core histone gene arrays of trypanosomes are found in distinct clusters on different chromosomes. The histone H2A cluster is found on chromosome 7, H2B on chromosome 10, H3 on chromosome 1 and the cluster of H4 is located on chromosome 5 (Berriman et al., 2005). While histone proteins are

usually among the most highly conserved proteins in eukaryotic evolution, the core histones of trypanosomes are substantially divergent due to sequence differences of their N-terminal tails (**Figure 1.6**) (Thatcher and Gorovsky, 1994). The linker histone H1 of trypanosomatids is smaller compared to other eukaryotic H1 proteins and lacks the globular domain. However, the C-terminal domain is almost identical to other H1 proteins (Toro et al., 1993).

Nuclear run-on assays demonstrated that transcription rates of trypanosomal histone genes are not coupled to DNA replication (Soto et al., 2000; Recinos et al., 2001). Nevertheless, similar to other eukaryotes, the steady-state histone mRNA levels peak at S-phase through increased mRNA stability (Ersfeld et al., 1999; García-Salcedo et al., 1999). Interestingly, trypanosomal histone mRNAs are polyadenylated at their 3´-end, whereas in other eukaryotes histone mRNAs contain a stem-loop structure replacing the poly(A)-tail (Aslund et al., 1994; Marañón et al., 2000).

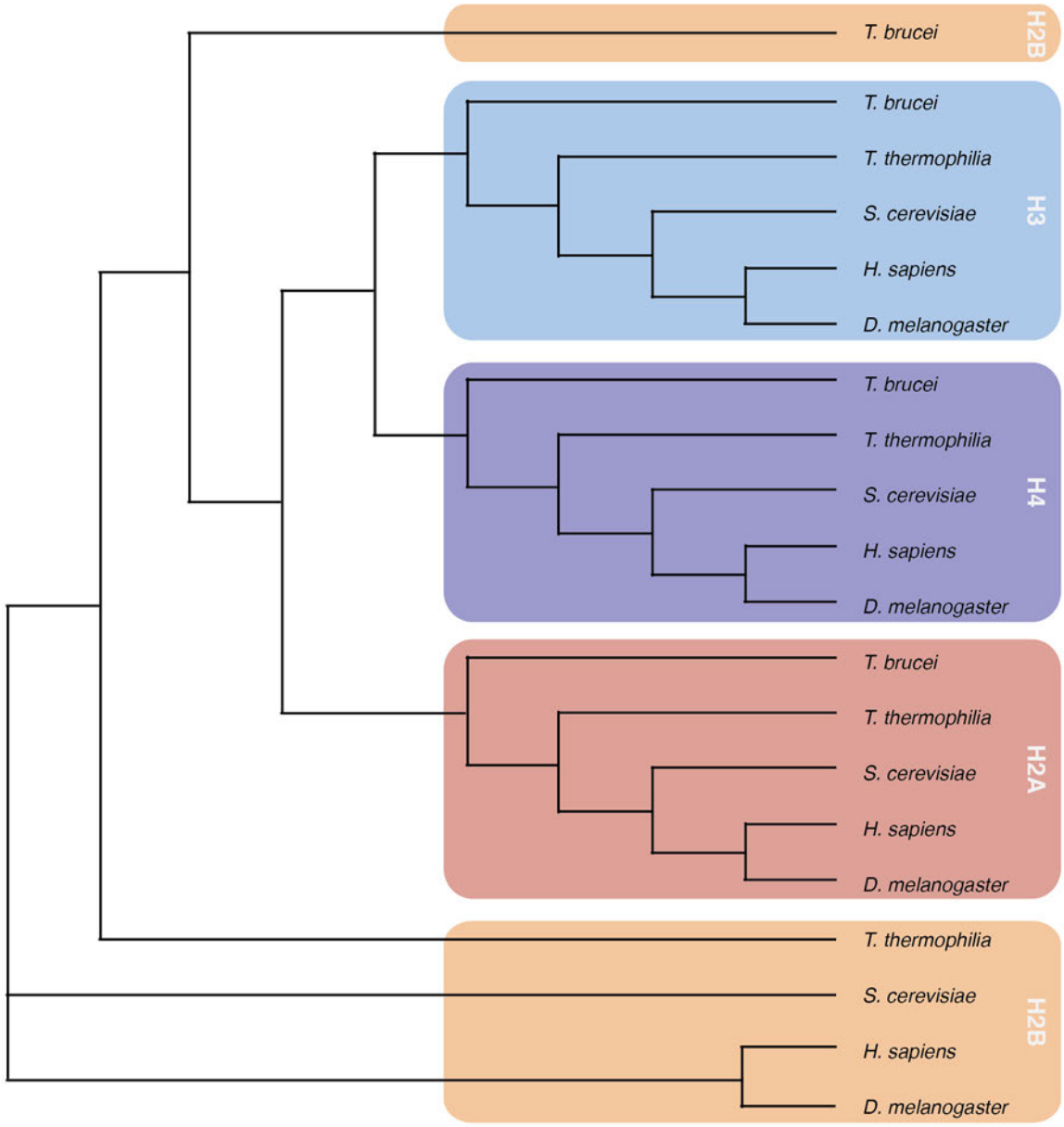


Figure 1.6. Phylogenetic analysis of the trypanosomal core histones.
The neighbouring-joining tree was generated using Clustal Omega with standard parameters (Sievers et al., 2011). The histone H2A (Tb927.7.2820), H2B (Tb927.10.10460), H3 (Tb927.1.2470) and H4 (Tb927.5.4170) protein sequences of *T.*

brucei were downloaded from TriTrypDB (Aslett et al., 2010). The histone H2A (sp|P35065|H2A1_TETTS), H2B (sp|P08993|H2B1_TETTS), H3 (sp|P69150|H31_TETTH) and H4 (sp|P69152|H4_TETTS) protein sequences of *T. thermophilia* were downloaded from UniProt (UniProt, 2019). The histone H2A (HTA1|YDR225W|SGDID:S000002633), H2B (HTB1|YDR224C|SGDID:S000002632), H3 (HHT1|YBR010W|SGDID:S000000214) and H4 (HHF1|YBR009C|SGDID:S000000213) protein sequences of *S. cerevisiae* were downloaded from yeastgenome.org (Cherry et al., 2012). The histone H2A (sp|P84051|H2A_DROME), H2B (sp|P02283|H2B_DROME), H3 (sp|P02299|H3_DROME) and H4 (sp|P84040|H4_DROME) protein sequences of *D. melanogaster* were downloaded from UniProt (UniProt, 2019). The histone H2A (tr|B2R5B3|B2R5B3_HUMAN), H2B (tr|B4DR52|B4DR52_HUMAN), H3 (tr|B4E380|B4E380_HUMAN) and H4 (sp|P62805|H4_HUMAN) protein sequences of *H. sapiens* were downloaded from UniProt (UniProt, 2019).

Nucleosomal rearrangements influence chromatin dynamics

The 'primary' chromatin structure is a concatenation of many nucleosomes looking like beads-on-a-string and is the fundament of higher-order chromatin formation (Luger et al., 2012), which is required for packaging entire eukaryotic genomes into single nuclei (Venkatesh and Workman, 2015). However, to fully package the genome, to organise specific chromatin regions and, to regulate DNA accessibility for proteins, nucleosomes need to be moved, ejected, incorporated or restructured. Thus, dynamic modulation of the chromatin state is an interplay of tightly regulated processes changing nucleosomal structure. From a basic view, all of these processes need to break the 14 DNA-histone contact points within a nucleosome to be able to modify the extent of DNA-wrapping, the composition of the histone octamer, or of both together (Zlatanova and Thakar, 2008).

The mechanisms which induce chromatin reorganisations are categorised into intrinsic or extrinsic dynamics (Campos and Reinberg, 2009). Intrinsic dynamics refer to changes in intra-nucleosomal components affecting nucleosome properties such as binding strength, mobility, size or conformation. Extrinsic effects are associated with interactional changes of two different nucleosomes opening or closing chromatin compaction. Many processes mediating chromatin reorganisation involve post-translational modifications of histones and incorporation of so-called histone variants and require specific chromatin modifying proteins, such as remodelling complexes or histone chaperones (Bönisch and Hake, 2012).

Post-translational histone modifications

Almost every histone possesses more than 60 different residues, which can be subjected to covalent post-transcriptional modifications (PTMs) (Kouzarides, 2007). These residues are found on the carboxyl-domain within the nucleosomal core and on the flexible amino-terminal tail protruding outward of the nucleosome (Fischle et al., 2003). The most prominent histone PTMs are arginine (R) methylation, lysine (K) acetylation, mono- /di- /trimethylation, ubiquitination, ADP-ribosylation and sumoylation, serine (S) and threonine (T) phosphorylation (Li et al., 2007) (**Figure 1.7 a**).

Deposition of distinct modifications to genomic regions is tightly regulated, especially during transcription. Histone modifications, which are associated to active transcription, are referred to as euchromatin modifications/marks and include for example histone lysine acetylation or H3K4 trimethylation (H3K4me3). Other modifications, such as H3K27me, are localised at transcriptionally silent sites and are therefore referred to as heterochromatin modifications/marks (Li et al., 2007).

Some histone PTMs, e.g. acetylation, can change the net charge of a nucleosome thereby changing intra-nucleosomal (such as histone-histone and/ or histone-DNA) or inter-nucleosomal interactions. The resulting alteration of the nucleosomal stability can in turn affect the chromatin composition (Kouzarides, 2007). For example, histone acetylation reduces the positive lysine charge and as a result the DNA-histone interaction gets weaker (Bannister and Kouzarides, 2011) (**Figure 1.7 b**). The enzymes mediating histone PTMs either act in a fairly non-selective manner or their catalytic activity is restricted to particular sites resulting in a unique function of the deposited chromatin mark (Kouzarides, 2007), as described for example for methylation and phosphorylation marks.

Besides directly inducing nucleosomal rearrangements, histone PTMs can serve as a binding platform for various effector proteins (**Figure 1.7 c**), such as histone chaperones or chromatin remodelling complexes. These proteins can contain characteristic domains, for example bromo- or chromo-domains, for recognising lysine acetyl and methyl marks, respectively (Zhang et al., 2015).

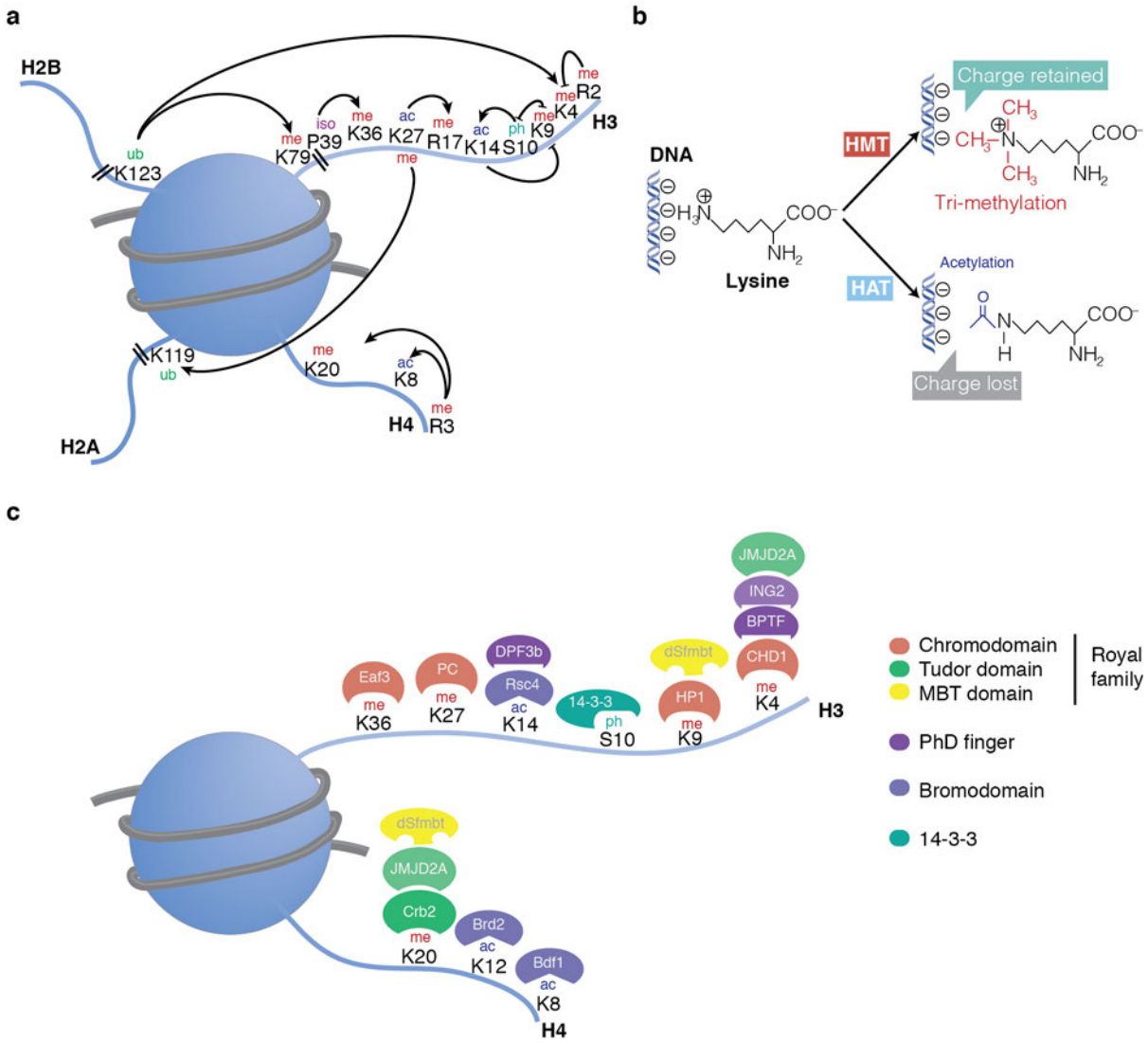


Figure 1.7. Histone PTMs can modify the nucleosome structure.
a Cross-talk of post-translational modifications found on yeast and human canonical histones. PTMs can have positive (arrowhead) and/or negative (flat head) effects on other histone modifications. Illustration was adapted from (Kouzarides,

2007; Bannister and Kouzarides, 2011). **b** Histone modifications can induce changes to the architecture and stability of a nucleosome. For example, lysine residues have a positive charge that can form salt bridges with the DNA, which is negatively charge. This interaction can stabilise the histone-DNA bonds within a nucleosome. Acetylation makes lysine less positive and results in reduced chemical forces between the DNA and the histone, thereby altering the nucleosomal stability. Whereas, lysine methylation does not alter the charge of the amino acid but depending on how many methyl groups are added to the lysine residue, can have steric implications on the histone-DNA interaction, that changes the nucleosomal stability. Illustration was redrawn from (Tessarz and Kouzarides, 2014). **c** Post-translational histone modifications serve as binding site for various proteins. Particular protein domains can recognise to a corresponding histone modification. Illustration was redrawn from (Bannister and Kouzarides, 2011).

Unconventional histone modifications in trypanosomes

As in other eukaryotes, histones are subjected to various post-translational modifications in trypanosomes, for example acetylation, methylation and phosphorylation. Studies have revealed that many of the evolutionary conserved PTMs are absent in *T. brucei*, but have found trypanosome specific PTMs (Saha, 2020).

One of the most remarkable difference of histone PTMs between *T. brucei* and other eukaryotes is the C-terminal acetylation of H2A, which has only been identified in trypanosomes (Janzen et al., 2006a; Mandava et al., 2007). The functionality of this PTM pattern is still not clear, but it might regulate the level of DNA accessibility for various DNA-binding proteins. In complex eukaryotes, the most prominent H2A modification occurs upon DNA damage and is known as γH2A. Serine phosphorylated H2A induces DNA repair by marking the break sites (Rogakou et al., 1998). In trypanosomes, γH2A is phosphorylated at threonine[130] and enriches *in vivo* at DNA damage sites (Glover and Horn, 2012).

Histone H2B of *T. brucei* is one of the most divergent histones. For example, it lacks the highly conserved lysine residue K120, which is monoubiquitinylated in various organisms and important for methylation of H3K4 and H3K79 (Sun and Allis, 2002). Interestingly, homologues to H3K4me (Wright et al., 2010) and H3K79me (Janzen et al., 2006a; Mandava et al., 2007; Wright et al., 2010) are found in trypanosomes, suggesting that the deposition of these marks is uncoupled from H2B ubiquitination in *T. brucei.* In bloodstream form parasites, H2B is acetylated to very low levels at K4, K12 and K16 (Mandava et al., 2007), which are absent in PF parasites (Janzen et al., 2006b) suggesting stage-specific variation in PTM distribution. To the best of our knowledge, histone H3 possesses the highest number and diversity of PTMs. For example, H3 carries at least 17 different kinds of PTMs on more than 30 different residues in humans (Xu et al., 2014) and many of them are important for transcriptional activation. For example, high levels of H3K4 methylation are found at TSSs and active promoters in humans (Kim et al., 2005) and also in trypanosomes (Wright et al., 2010). H3K76 methylation (the homologue to H3K79 methylation in other eukaryotes) is one of the best characterised PTMs in trypanosomes. In humans, H3K79 methylation is implicated in cell cycle regulation, DNA damage response and transcription (Nguyen and Zhang, 2011). H3K79 is mainly mono- and dimethylated and both marks are mediated by one DOT1 methyltransferase (Disruptor of Telomeric Silencing). In trypanosomes, H3K76 can be mono-, di- and trimethylated and the marks are mediated by two different

enzymes DOT1A and DOT1B (Janzen et al., 2006b). While DOT1A only mediates mono- and dimethylation and is involved in replication initiation (Gassen et al., 2012) and cell cycle progression (Frederiks et al., 2010), DOT1B can also mediate H3K76me3 (Janzen et al., 2006b), which is involved in differentiation and antigenic variation (Figueiredo et al., 2008; Frederiks et al., 2010). Histone H4 is the least divergent core histone in trypanosomes and has been shown to be acetylated at the N-terminal residues H4K2, H4K4, H4K5 and H4K10 (Janzen et al., 2006a), similar to the H4-tail in other eukaryotes.

Incorporation of histone variants

While core histones belong to the slowest evolving proteins known (Kamakaka and Biggins, 2005), so-called histone variants can exhibit sequence differences in histone tails, globular domains or in only a few amino acid residues compared to their canonical counterparts (Li et al., 2007). Except of histone H4 lineages, all canonical histones have numerous variant proteins in higher eukaryotes, e.g. humans (Malik and Henikoff, 2003) (**Figure 1.8**).

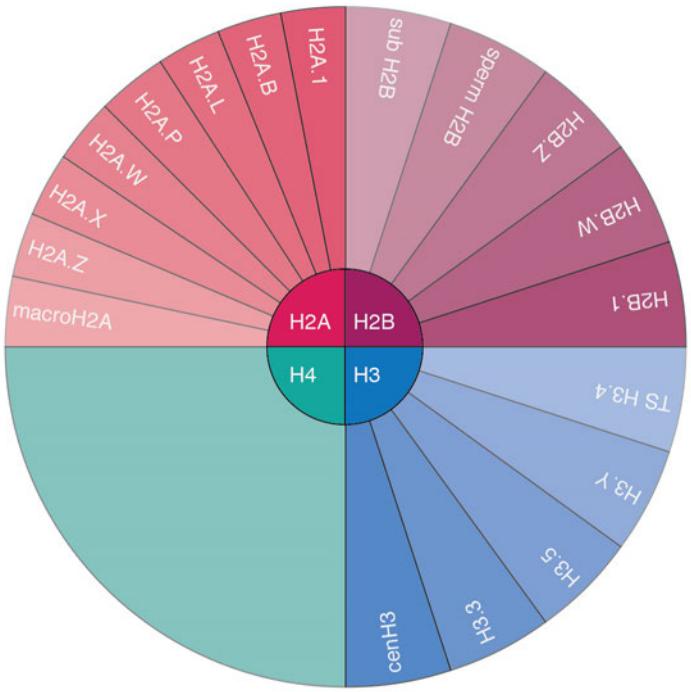


Figure 1.8. The human core histones and their variants.
Schematic illustrations of the described histone variants of human canonical histones. In contrast to H2A, H2B and H3 that have many different types of variant forms with different regulator function, H4 is rather invariant in most eukaryotes (Khare et al., 2012).

In contrast to canonical histones, the variant proteins are single-copy genes, and their transcripts are subjected to alternative splicing and polyadenylation (Biterge and Schneider, 2014). Histone variants are expressed and deposited into chromatin independently of DNA synthesis throughout the whole cell cycle (Kamakaka and Biggins, 2005). Specialised histone chaperones and other chromatin modifiers incorporate histone variants into nucleosomes, where they substitute pre-existing histones. Thus, they are often termed as “replacement” histones (Biterge and Schneider, 2014). Deposition of histone variants into a nucleosome introduces structural changes through altered histone/DNA and/or

histone/histone interaction. This can affect the nucleosomal stability and as a result the chromatin compaction (Henikoff, 2008). Consequently, exchange of canonical histones by variant proteins is an important way to modulate chromatin and adds another level of complexity to the regulation of DNA accessibility.

The changes on chromatin mediated by histone variants mediate can be profound. For example, specific histone H3 variants, so-called CenpAs or CenH3s, are the foundation of centromeric chromatin (Malik and Henikoff, 2003) and are important to distinguish centromeric DNA from DNA on the chromosome arms during cell division (Henikoff, 2008).

Histone variants in trypanosomes

Trypanosomes encode four histone variants H2A.Z, H2B.V, H3.V and H4.V, which are expressed from single copy genes (Lowell and Cross, 2004; Lowell et al., 2005). Analyses of their genome-wide distribution have shown that the histone variants are enriched at specific regulatory regions of RNAPII transcription in *T. brucei* (Siegel et al., 2009b). At RNAPII TTSs, the histone variants H3.V and H4.V are enriched (Siegel et al., 2009b). Additionally, H3.V is found at the transcriptionally silent telomeres (Lowell and Cross, 2004) and subtelomeric regions (Siegel et al., 2009b). Neither H3.V nor H4.V are essential for trypanosomal viability (Lowell and Cross, 2004; Siegel et al., 2009b). However, previous findings suggest that H3.V functions as a regulator of RNAPII transcription termination. For example, removal of H3.V affects transcription termination and increases expression of genes downstream of TSSs (Reynolds et al., 2016). Additionally, H3.V, together with H4.V, is important for genome architecture, maintaining the silent state of transcriptionally repressed BESs. For example, cells, where both histone variants were deleted, showed higher DNA accessibility on silent BESs and expressed a new major surface VSG compared to wild type cells (Müller et al., 2018). The variant forms of H2A and H2B are H2A.Z and H2B.V, respectively, and both are essential proteins for *T. brucei* (Lowell et al., 2005). H2A.Z and H2B.V dimerise exclusively together (Lowell et al., 2005) and are enriched at TSSs (Siegel et al., 2009b).

Though, ChIP-seq studies can reveal the genomic sites, where specific chromatin marks are incorporated, to understand its biological function, it is important to know how a specific mark is targeted to a specific genomic locus and which factors play a role for its targeted deposition. However, therefore, all marks present on the chromatin isolated from a specific genomic locus would need to be determined. Thus far, analyses of locus-specific chromatin to identify all chromatin marks present at specific genomic site, including PTMs, histone variants, histone modifying enzymes and chromatin remodelling complexes, have only been successfully performed on chromatin isolated from long repetitive regions such as telomeric repeats (Déjardin and Kingston, 2009).

Chromatin remodelling complexes

Generally, all chromatin dynamics require the activity of specialised ATP-dependent chromatin remodelling complexes (Clapier et al., 2017). They are important to ensure proper nucleosome spacing and density, to move or evict nucleosomes for DNA-accessibility, or to exchange canonical nucleosomes by histone variants at specific genomic regions (Clapier and Cairns, 2009). Many known chromatin remodelers have a SNF2-like domain their catalytic subunit and belong to the same class of RNA/DNA helicase superfamily 2, which is subdivided into four groups differentiating in their phylogenetic relationship (Flaus et al., 2006) and/or their functionalities: ISWI, CHD, SWI/SNF, INO80 (Becker and Hörz, 2002; Hargreaves and Crabtree, 2011; Becker and Workman, 2013; Narlikar et al., 2013; Bartholomew, 2014). Most of the studies analysing chromatin remodeler activities were performed in budding yeast and thus, the naming of the four classes is based on their function in yeast (Clapier et al., 2017).

The ISWI and CHD subfamily of chromatin remodelers mainly function in nucleosome assembly and organisation following replication (Gurard-Levin et al., 2014) or transcription. During both processes, the chromatin structure has to be dissolved and afterwards re-assembled. For example, during replication, nucleosomes are randomly deposited from the DNA upstream of the replication fork and following replication, the chromatin composition is re-established on the old and new double stranded DNA. At first, pre-nucleosomes are formed, which contain a complete histone octamer (Torigoe et al., 2011; Fei et al., 2015), but only 80 bp DNA wrapped around it. After that, chromatin remodelers assist to complete the formation of a full nucleosomes and establish the correct nucleosome spacing by their ATP-dependent sliding activities (Varga-Weisz et al., 1997; Corona et al., 1999).

The SWI/SNF superfamily contains predominantly chromatin remodelers, which assist in achieving DNA accessibility. For example, the RSC-complex exposes nucleosome DNA to generate a nucleosome depleted region at promoters without removing the associated histones (Parnell et al., 2008; Hartley and Madhani, 2009; Wippo et al., 2011). The remodelers have a striking binding pocket for one mononucleosomes (Asturias et al., 2002; Smith et al., 2003a; Leschziner et al., 2005). Interestingly, it has been shown that the remodelling activity of ISWI complexes is stimulated by nucleosomes containing the histone variant H2A.Z, which might enable other proteins to bind to chromatin (Goldman et al., 2010).

The INO80 family, defined by the two closely related ATPases Ino80 and Swr1, is known to edit the nucleosome composition in a replication-independent manner. Such remodelers form specialised chromatin regions by incorporating or evicting histone variant proteins (Morrison and Shen, 2009). An editing-remodeler needs to stabilise the hexasome following the removal of a histone dimers, while it incorporates the variant dimer. Thus, they need to be able to differentiate substrate and product nucleosomes (Clapier et al., 2017). A very prominent member of the INO80-family is the SWR1 complex,

which removes H2A-H2B dimers and incorporates H2A.Z-H2B dimers (Kobor et al., 2004; Mizuguchi et al., 2004). SWR1 contains proteins that specifically recognise H2A-H2B dimers (Wu et al., 2005; Hong et al., 2014). The replacement of H2A-H2B by H2A.Z-H2B occurs in a stepwise manner. That means that one H2A/H2B dimer is exchanged at a time. This creates heterotypic, intermediate nucleosomes containing H2A and H2A.Z (Luk et al., 2010). The stepwise replacement of H2A with H2A.Z by SWR1 is unidirectional. The reverse process requires the INO80 complex (Shen et al., 2003; Jin et al., 2009; Papamichos-Chronakis et al., 2011b). In contrast to the SWR1 complex (Ranjan et al., 2015), INO80 can additionally slide the nucleosome along a DNA strand (Udugama et al., 2011).

For a long time, it has been thought that each of the subfamilies has specific enzymatic mechanisms to modify the chromatin structure in its specific way (Clapier and Cairns, 2009). However, in a recent review, another model has been proposed suggesting that all remodelers use ATP-dependent DNA-translocation to change the chromatin structure. By associating with specialised proteins, such as TFs or histone chaperones, the remodelers are able to establish a specific chromatin structure at a specific genomic region. This model is termed “hourglass”-model and is based on the finding that all remodelers act similar to DNA translocases and possess an ATPase domain as the driving “motor” (Clapier et al., 2017). While different remodelers affect the nucleosome composition differently, they (i) are all more affine to nucleosomes than to naked DNA, (ii) have a single catalytic subunit containing the ATPase domain, (iii) associate with proteins regulating the ATPase subunit, and (iv) interact with other chromatin modifying proteins (Clapier et al., 2017).

Chromatin remodelling complexes are often recruited by histone PTMs (Suganuma and Workman, 2011; DesJarlais and Tummino, 2016) through PTM-reader proteins, for example bromodomain- (Hassan et al., 2002) or chromodomain-containing proteins for acetylation or methylation, respectively. Moreover, nucleosome-modifying proteins, such as HATs or histone chaperones, cooperate with remodelling complexes (Pollard and Peterson, 1998). For example, SAGA and NuA4 function along with SWR1 to promote gene activity (Mitra et al., 2006; Altaf et al., 2010; Cheng et al., 2015).

Chromatin remodelling complexes in T. brucei

The release of the genomes of the three trypanosomatids in 2005 (El-Sayed et al., 2005a; Ivens et al., 2005; Hertz-Fowler et al., 2007) opened the avenues to identify evolutionary conserved chromatin factors in trypanosomes based on sequence analyses. *T. brucei* encodes six different histone acetyltransferases (Ivens et al., 2005; Kawahara et al., 2008), seven histone deacetylases, around 25 histone methyltransferases, four putative demethylases (Ivens et al., 2005), five bromodomain proteins (Ivens et al., 2005; Siegel et al., 2009b; Schulz et al., 2015) and one potential chromodomain protein (Ivens et al., 2005). In contrast to bromo- or chromodomain proteins found in other eukaryotes, trypanosomal reader proteins have only a single PTM-binding site (Figueiredo et al., 2009). Moreover, *T. brucei* has 13 genes encoding for a potential SNF2-domain, the catalytic domain present in many

chromatin remodelling complexes and two of them have been characterised: TbISWI and JBP2 (base J binding protein) (DiPaolo et al., 2005).

Similar to other members of the ISWI family, TbISWI exhibits the typical SANT/SLIDE domain, which is important for nucleosome sliding activity of the remodeler (Hota et al., 2013). TbISWI is involved in silencing of *VSG* expression sites (Hughes et al., 2007). However, the remodelling activity of TbISWI that accomplish the ES downregulation is still unclear. Using tandem-affinity purification, five interacting partners of TbISWI have been identified, that do not show significant sequence similarity to the typical ISWI-associated proteins. ChIP-seq analyses suggest that TbISWI and three of the interacting proteins are enriched at TSSs (Stanne et al., 2011; Stanne et al., 2015). However, the function of the ISWI complex for RNAPII transcription or if it is stimulated by H2A.Z has not been studied.

Interestingly, except of the SNF2-domain, JBP2 lacks any remarkable chromatin-binding motif, but interacts with nuclear chromatin (DiPaolo et al., 2005). JBP2 is a thymidine hydroxylase that is responsible for *de novo* synthesis of the trypanosome-specific DNA modification "base J", which is a hypermodified thymidine residue (Cliffe et al., 2009). Base J co-localises with H3.V at telomeres and TTSs suggesting a role in transcription termination.

Next to putative chromatin remodelling proteins, *T. brucei* has a FACT complex composed of an orthologue to the large subunit Spt16 (Denninger et al., 2010) and the histone chaperone Pob3 (Denninger and Rudenko, 2014). Depletion of Spt16 and Pob3 results in growth arrest and derepression of silent VSG ESs, similar to depletion of TbISWI (Hughes et al., 2007), suggesting that some nucleosome remodelling activity is required for successful ES silencing by maintaining a silent chromatin state (Denninger and Rudenko, 2014). While studies in other eukaryotes have revealed that FACT interacts with the histone variant H2A.Z, this interaction has not been investigated in *T. brucei*.

The role of the chromatin structure in RNAPII transcription

In all living organisms, the hereditary material is encoded in multiple genes (Crick, 1958). However, the gene number itself does not define the complexity of an organism, more which genes are expressed in which cell and to what extent. For example, the human genome contains roughly 20,000 annotated protein-coded genes, but not all of them are expressed in each cell type (Velculescu et al., 1999).

Transcription, the initial part of gene expression, controls the 'on and off' status of a gene at the correct time and place. An approximately 550 kDa enzyme complex (Cramer et al., 2000), termed RNA polymerase, reads the DNA sequence of a gene and produces RNA, which is either protein-coding or non-coding (Vannini and Cramer, 2012). DNA-dependent RNA polymerases are classified by the RNA they produce, and RNA polymerase II (RNAPII) generates messenger RNAs (mRNAs) from protein-coding genes (RNAPI: ribosomal RNAs; RNAPIII: transfer RNAs) (Young, 1991), which are the basic molecules for protein biosynthesis.

Chromatin compaction interferes with all phases of RNAPII transcription that involve DNA-binding: initiation, pausing and elongation (Chen et al., 2018). Nucleosomes constitute a physical barrier for RNAPII progression (Bondarenko et al., 2006). Thus, transcription is accompanied by processes mediating nucleosome destabilisation that enable the polymerase to overcome the nucleosomal obstacle (Venkatesh and Workman, 2015) (**Figure 1.9**).

RNAPII transcription in eukaryotes

Generally, gene transcription by RNAPII starts with DNA binding of activator proteins adjacent to the actual transcription start site (TSS) that leads to RNAPII recruitment and its positioning at the TSS. The enzyme starts to synthesise a new mRNA molecule, thereby clearing the transcription start region and passing into productive elongation. At the end of a gene, DNA-sequence mediated transcription terminates, and the polymerase complex releases the complete transcript (Conaway et al., 2000).

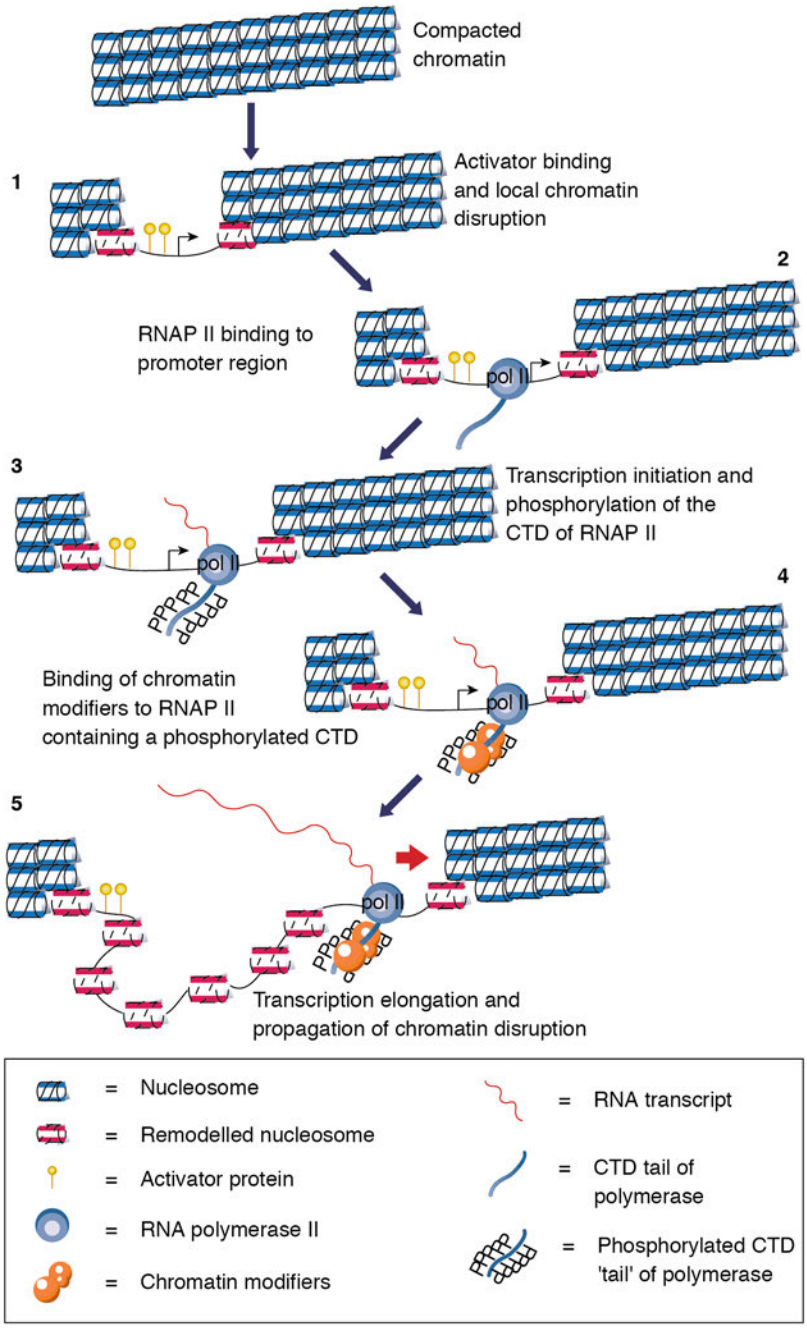


Figure 1.9. Chromatin remodelling accompanies transcription.
Chromatin modifiers associate to RNAPII during all steps of transcription that require DNA-binding of RNAPII. (1) Activator proteins at the promoter recruit chromatin modifiers to make the promoter-DNA accessible for RNAPII. (2) RNAPII binds to the promoter and (3) shortly after initiating transcription, the carboxyl-terminal domain (CTD) of RNAPII gets hyperphosphorylated. (4) The phosphor marks of RNAPII recruit additional chromatin modifiers (5) for opening up the chromatin during active elongation. These processes allow the transcription machinery to have constant access to the DNA during progression. The illustration was modified from (Orphanides and Reinberg, 2000).

Transcriptional regulation is one of the most fundamental requirements for controlling cell identity, growth, differentiation and development. For a long time, research has focused on transcription initiation, where major regulatory steps take place for correct RNAPII recruitment and TSS recognition. With disclosure of the RNAPII crystal structure (Cramer et al., 2000; Cramer et al., 2001), new, unexpected insights into the molecular mechanism of RNAPII pausing/arrest and transcription elongation

have been provided, which has permitted the study of the regulatory events leading to the transition from initiation, RNAPII pausing to productive elongation (Jonkers and Lis, 2015).

Overview of the RNAPII complex

RNAPII is the core of the transcription machinery and consists of a ~600 kDA enzyme complex, which includes 12 different subunits (Woychik et al., 1998) and is highly conserved from *Saccharomyces cerevisiae* to *H. sapiens*. The catalytic site of RNAPII lies at the end of a deep cleft, which is formed between the two largest subunits Rpb1 and Rpb2 and contains an essential Mg^{2+} ion (Cramer et al., 2000) (**Figure 1.10**).

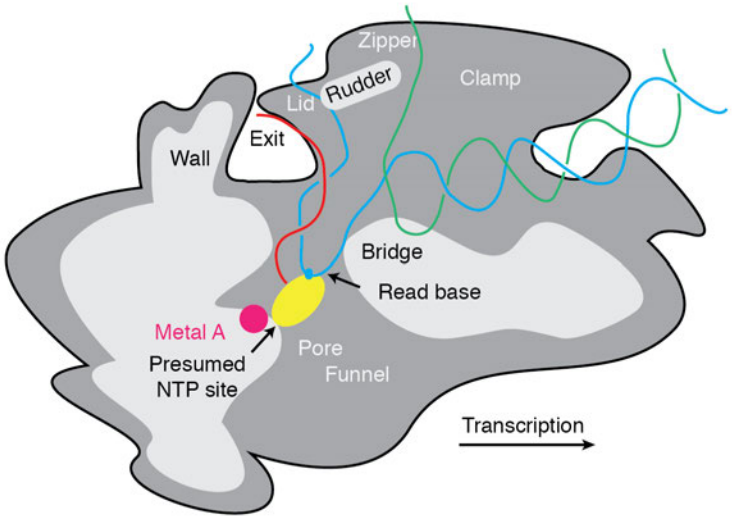


Figure 1.10. Conserved topology of transcribing RNA polymerase II.
Simplified cut-away drawing of the transcribing eukaryotic RNAPII (silver) based on a crystallographic model. The direction of transcription is indicated as a black arrow. The red strand is the nascent RNA molecule protruding out of the pore (yellow) with the active site (pink). The transcribed DNA is blue (non-coding strand) and green (coding strand). Illustration was adapted from (Cramer, 2004).

The cleft can harbour 8-9 base pairs of an RNA-DNA hybrid, maintaining the correct alignment of the 3´OH of the nascent RNA molecule to the metal ion, and ~20 nucleotides template DNA, which locate downstream of the catalytic site (Cramer et al., 2001). The entry site of the DNA-binding cleft is encircled by a pair of 'jaws', formed by parts of the subunits Rpb1, Rpb5 and Rpb9, that grip the entering DNA template (Cramer et al., 2000). Near the active site, the amino-terminal regions of Rpb1 and Rpb6 and the carboxyl-terminal part of Rpb2 form a clamp, referred to as 'hinged' domain (Fu et al., 1999), that holds the DNA at the active site in place. Downstream of the catalytic site, 10-20 bp of the nascent transcript are bound by potential grooves in the RNAPII complex (Rice et al., 1991), that stabilise the elongation. Within a cell, RNAPII processes 25 nucleotides per second, a rate which could be achieved only on naked DNA in *in vitro* studies (Orphanides and Reinberg, 2000).

The RNAPII complex of trypanosomes

Like other eukaryotes, trypanosomes employ three different RNA polymerases to transcribe their nuclear genome. Interestingly, trypanosomal RNAPI can also synthesise messenger RNAs from VSGs

and procylin genes that code for the major surface protein of PF parasites (Bütikofer et al., 1997), not only rRNAs (Günzl et al., 2003). Subtelomeric VSG expression sites are controlled by a RNAPI-specific promoter (Hertz-Fowler et al., 2008), excluding RNAPII from these sites. Additionally, transcription of the active expression site is compartmentalised into the 'expression site body' located adjacent to the nucleolus, where most of the RNAPI-guided transcription takes place (Navarro and Gull, 2001). All other protein-coding genes and small uracil-rich RNAs, which belong to the spliceosome complex and are important for processing of the pre-mRNAs, are transcribed by RNAPII.

The RNAPII-complex consists of 12 subunits (RPB1-RPB12) (Srivastava et al., 2018). RPB1, RPB2, RPB3, and RPB11 form the basic components of the enzyme complex and are homologues to subunits of the eukaryotic core RNAPII complex (Das et al., 2006). In the trypanosomal genome, RPB1, the largest subunit, is encoded by two non-allelic gene variants, termed polIIA and polIIB, which differ in four amino acids in the domain predicted for DNA-binding (Evers et al., 1989; Smith et al., 1989). Biochemical studies of RPB9-pull outs have revealed that *T. brucei* contains the full set of general transcription factors (GTFs), which are involved in transcription initiation in eukaryotes (Srivastava et al., 2018). However, the protein sequences of the GTFs are extremely divergent from those of other eukaryotes.

Transcription initiation

In general, gene transcription by RNAPII starts with the assembly of the pre-initiation complex (PIC) that consists of the GTFs – TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH – and the coactivator complex Mediator as stabilisator (Baptista et al., 2017; Hantsche and Cramer, 2017; Nogales et al., 2017). None of the known eukaryotic RNA polymerase complexes can initiate transcription on its own. They need the proteins of the PIC, which function in promoter recognition, polymerase recruitment, interaction with other regulatory proteins, DNA unwinding and polymerase positioning to the correct transcription initiation site (Grünberg and Hahn, 2013). PIC assembly starts with the formation of a closed complex (CC) by TFIIA, TFIIB and TFIIH at the promoter site (Hantsche and Cramer, 2017). With some exceptions, promoters are found upstream of each gene in eukaryotes and can extend ~35 bp upstream and/or downstream of the transcription initiation site. Core promoters interact directly with transcription factors and are characterised by specific DNA sequence elements, essential for recruitment of the transcription machinery (Grünberg and Hahn, 2013). The best studied core promoter element is the TATA-box and for a long time it has been suggested that TATA-promoter elements are largely invariant. However, it has become increasingly clear that the majority of promoters do not contain the canonical TATA-element (TATA-less promoters) (Smale and Kadonaga, 2003). TBP, previously classified as TATA-binding protein, associates with TATA-containing or TATA-less core promoter elements. The transcription factor TFIID, which consists of TBP and different TBP-associated factors (TAFs), functions in promoter recognition and tethers TFIIA to the core promoters (**Figure 1.11 a**) (Sainsbury et al., 2015).

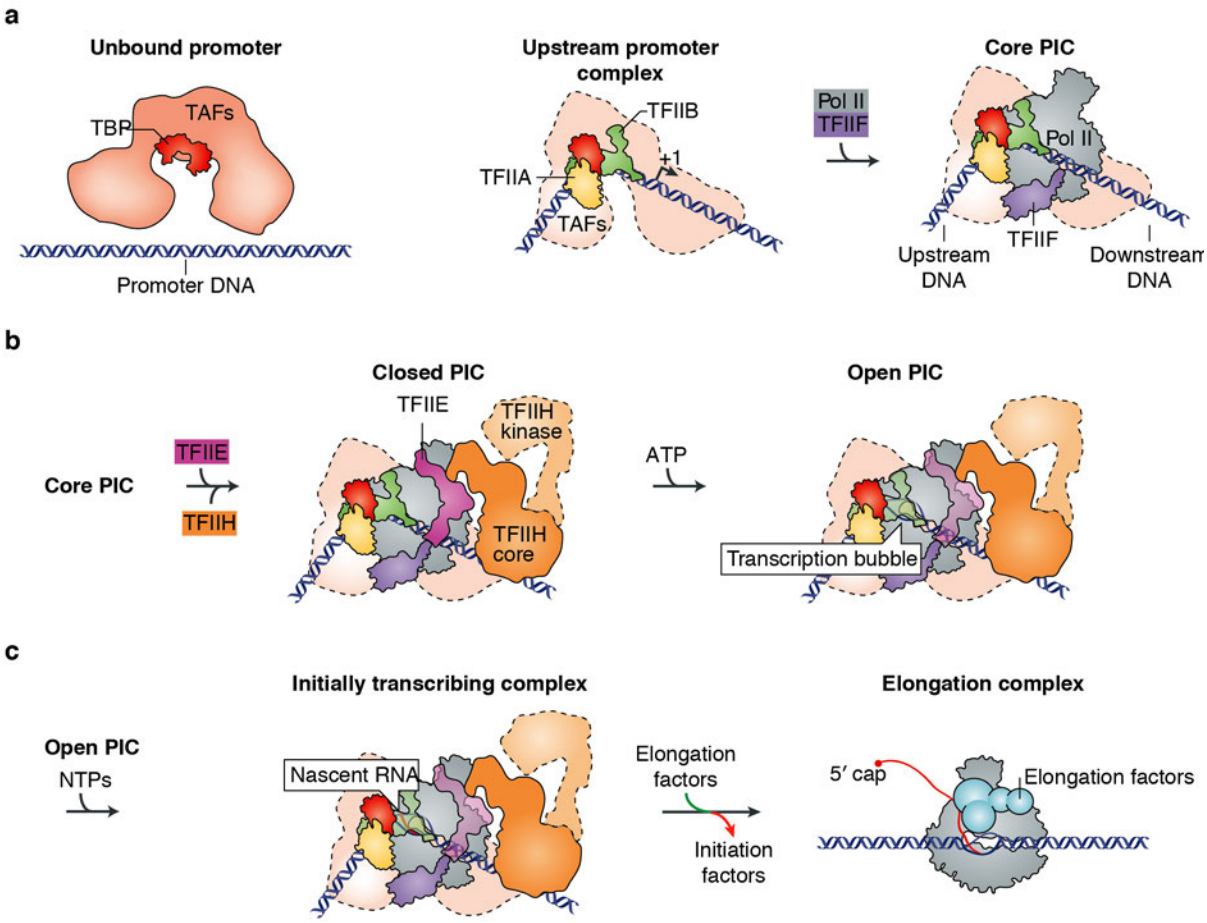


Figure 1.11. Model of RNAPII complex during transcription initiation in eukaryotes.
a The model of the stepwise PIC-assembly is shown from the GTFs (various colours) and RNAPII (grey) on canonical promoters. The TATA-binding proteins subunit (TBP) of TFIID binds to promoter DNA and induces DNA-bending. TFIIB and TFIIA bind to both sites of TBP stabilising the TBP-DNA complex resulting in the upstream promoter complex. TFIIF-bound RNAPII joins the complex and their binding initiate the formation of the core PIC. **b** Binding of TFIIE and TFIIH to the core PIC completes PIC assembly (closed PIC) and in the presence of ATP, the DNA is opened up forming the transcription bubble and the open PIC. **c** In the presence of NTPs, the open PIC starts transcription and forms the initially transcribing complex. The dissociation of the initiation factors and the binding of the elongation factors (blue) leads to the formation of the elongation complex. The illustration was modified from (Sainsbury et al., 2015).

TFIIA, TFIIB and TBP assemble upstream of the promoter sequence and recruit TFIIF-bound RNAPII, thereby accomplishing the closed PIC (**Figure 1.11 b**) (Conaway et al., 2000). The DNA is tightly wrapped around PIC through interaction with TFIIF, which inhibits RNAPII from subsequent promoter escape (Conaway et al., 2000). Additionally, TFIIF assists RNAPII in TSS selection and prevents binding of RNAPII from contacting not-promoter DNA (Sainsbury et al., 2015). Bound TFIIE recruits the helicase complex TFIIH, which mediates the ATP-dependent formation of an open promoter complex (OC) (**Figure 1.11 b**) (Grünberg and Hahn, 2013), most likely by unwinding the DNA downstream of the promotor (Conaway et al., 2000). TFIIF stimulates the synthesis of the first phosphodiester-bond (Sainsbury et al., 2015) and the transition of the OC to the initially transcribing complex (Hantsche and Cramer, 2017). After synthesising the first 10-15 nucleotides, RNAPII leaves the promoter and dissociates from its initiation factors (Conaway et al., 2000), forming an early elongation complex (EC) (**Figure 1.11 c**). Until the nascent transcript reaches a length of ~9 nucleotides, there is a constant competition between

dissolution of EC and synthesis of the next phosphodiester bond. It is thought that TFIIF counteracts transcriptional abortion during early elongation by stimulating nucleotide addition (Conaway et al., 2000). Before escaping the promoter region, very early EC can also arrest, which is suppressed by TFIIE and TFIIH, a process similar to open complex formation (Dvir, 2002).

PIC formation is absent from most TSSs in trypanosomes

The general features required for RNAPII transcription initiation are the general transcription factors for PIC formation, a core promoter as a defined landing platform for the initiation complex and a chromatin structure that guarantees an accessible promoter region. In trypanosomes, however, the RNAPII machinery dispenses with most of these regulatory features to manage transcription initiation. Through strand-specific nuclear run-on assays performed in *L. major* (Martínez-Calvillo et al., 2004), it could be shown that RNAPII transcription starts at divergent strand-switch regions between two neighbouring PTUs. Since 75% of *T. brucei* genes lie in the same genomic context as in *L. major* (El-Sayed et al., 2005b), divergent strand-switch regions have been suggested to be TSSs in *T. brucei* as well. Later, TSSs have been identified as sites of H2A.Z-enrichment (Siegel et al., 2009a) revealing additional TSSs in non-strand-switch regions between two neighbouring PTUs that have the same direction of transcription. TSSs can be up to 20 times larger as described for other eukaryotes (Kristiansson et al., 2009). RNAPII concentrates at the 5´end of the TSS, spanning a ~2-kb region, in which transcription initiates and transcription initiates from numerous sites within this ~2-kb region (Wedel et al., 2017). Trypanosomal TSSs lack typical eukaryotic core promoter motifs (Kolev et al., 2010). They contain GT-rich sequence elements at the coding strand which are capable of mediating transcription initiation and directionality (Wedel et al., 2017). Only promoters of the spliced leader (SL) genes exhibit a canonical promoter motif for RNAPII transcription (Gilinger and Bellofatto, 2001). The genome of trypanosomes encodes for around 100 SL genes, arranged in tandem repeats on chromosome 9 (Roberts et al., 1996). SL genes are the only known RNAPII-transcribed genes, which are transcribed monocistronically in trypanosomes, meaning each SL gene possesses its own promoter for transcription initiation. Analyses of proteins associated with SL synthesis have revealed that a complete set of RNAPII transcription factors is present in trypanosomes (Srivastava et al., 2018), including TBP-related factor 4 (Ruan et al., 2004), TFIIH helicases (Ivens et al., 2005; Lecordier et al., 2007; Lee et al., 2007; Lee et al., 2009), SNAPc (Das and Bellofatto, 2003; Schimanski et al., 2004), TBP (Das et al., 2005), TFIIB (Palenchar et al., 2006; Schimanski et al., 2006), TFIIE (Lee et al., 2009) and TFIIF (Srivastava et al., 2018). Most of the genes encoding for trypanosomal GTFs are highly divergent (Srivastava et al., 2018). Consequently, the RNAPII complex could only be identified using biochemical strategies not by homology search, with the exception of the TBP-related factor 4 and the two TFIIH helicases. PIC formation has only been experimentally associated to SL promoters and is still elusive for other TSSs (Martinez-Calvillo et al., 2010). Thus, it has been assumed that for the majority of protein-coding genes, transcription initiation

is probably not the major regulatory step to control gene expression in trypanosomes. Previous studies have shown that RNAPII can initiate from sites lacking any promoter characteristics, as long as transcription occurs nearby (McAndrew et al., 1998), suggesting that an 'open' chromatin is sufficient for transcription initiation in trypanosomes. Using micrococcal nuclease (MNase) digestion, which digests away DNA not protected by proteins, it has been observed that the DNA at TSSs is more prone to digestion, revealing that TSSs are indeed more accessible for proteins (Wedel et al., 2017).

Transcription elongation and RNA processing

During early elongation, RNAPII can be actively retained at the 3´end of the promoter region and concentrates ~30 to 60 bp downstream of the TSS (Jonkers and Lis, 2015), a regulatory step known as promoter-proximal pausing. Paused RNAPII stays stably associated to the nascent transcript and, in contrast to RNAPII that failed to leave the PIC or was blocked during early elongation, is competent to resume to productive elongation without rescue and does not need to reinitiate transcription (Adelman and Lis, 2012). Paused polymerase does not stay strictly in one position, but undergoes rounds of backtracking, progression and pausing, while the transcription factor TFIIS modifies the nascent transcript to keep the active site in place (Jonkers and Lis, 2015). To maintain promotor-proximal pausing, factors, such as the negative elongation factor (N-ELF), DSIF and the PAF1-complex, are necessary not only to prevent transcription termination directly after initiation, but also to antagonize RNAPII from premature release before recruitment of other regulatory proteins. Phosphorylated Ser5 of the carboxyl-terminal domain marks paused RNAPII and assists in recruiting and activating of capping enzymes (Chen et al., 2018).

RNAPII pausing, rate-limiting for transcription (Conaway et al., 2000), keeps promoters and genes in an active state and serves as a checkpoint for aberrant transcripts and for essential RNAPII modifications, before elongation can take place (Jonkers and Lis, 2015).

The canonical splicing factor SRSF2, which binds to unprotected RNA protruding from paused RNAPII, induces progression to productive elongation by activating the kinase P-TEFb. P-TEFb phosphorylates Ser2 within the conserved heptapeptide motif YSPTSPS of the RNAPII carboxyl-terminal domain (CTD). This phosphorylation step is accompanied by dephosphorylation of Ser5, and the negative regulators NELF and DSIF (Jonkers and Lis, 2015; Chen et al., 2018). Two classes of elongation factors are known, which are important for functional RNA synthesis: The SII family, which can reactivate arrested RNAPII by mediating the generation of a new 3´OH-end on the new RNA that is now correctly repositioned to the catalytic site, and the elongation factors, which counteract RNAPII stalling (TFIIF, Elongin, ELL) (Conaway et al., 2000). In the first few kilobases of a gene, elongation is not very efficient due to reorganisation and modification events of the transcription machinery. This slow elongation rate is important for coordination of elongation and co-transcriptional RNA processing (Jonkers and Lis, 2015) that includes modifications of the pre-cursor mRNA (pre-mRNA) molecule, such as 5´-capping,

removing of introns (non-coding regions) by *cis*-splicing and polyadenylation. RNA processing is required to generate a mature mRNA molecule that is stable and can be translated into a protein. Transcription of protein-coding genes terminates ≥500bp downstream of the poly(A)-signal and the transcription machinery releases the DNA strand and the synthesised RNA molecule (Rosonina et al., 2006).

For a long time, it has been believed that RNA synthesis and maturation are two independent processes. However, the processes of elongation and splicing are spatially and temporally aligned. Accurate coordination of both mechanisms is essential and results in efficient and accurate mRNA production (reviewed in (Bentley, 2014)). The amount of time the nascent transcript is available for splicing depends on the rate of transcription elongation (reviewed in (Carrillo Oesterreich et al., 2011)). Thus, perturbed elongation kinetics can lead to defects co-transcriptional splicing (de la Mata et al., 2003; Howe et al., 2003; Carrillo Oesterreich et al., 2010).

mRNA processing in trypanosomes

As mentioned earlier, transcription of protein-coding genes occurs polycistronically in trypanosomes. In most eukaryotes transcription is monicistronic meaning that each gene has its own TSS, where transcription initiates. Whereas, in polycistronic transcription one TSS regulates the transcription of multiple genes. Mature monocistronic mRNAs are generated during co-transcriptional RNA processing in trypanosomes, which includes a process called *trans*-splicing and 3´-polyadenylation. During *trans*-splicing, a 39 nt-long 'spliced leader' miniexon that contains the 5´-cap is joined to the 5´UTR of each ORF within the primary transcript (Clayton, 2019). *Trans*-splicing combines two exons derived from different RNA molecules, whereas *cis*-splicing, which is the process of RNA processing usually present in eukaryotes, removes the intronic regions and joins the exon within a monocistronic pre-mRNA molecule (reviewed in (Liang et al., 2003)). As *cis*-splicing, *trans*-splicing requires precise recognition of the joining to be successful and includes a 5´-splice donor site (SDS) on the splice leader miniexon, a 3´-splice acceptor site (SAS) next to the 5´end of the ORF and a polypyrimidine tract (poly(Y)-tract), that precedes the SAS. The SAS is usually the first AG dinucleotide downstream of the poly(Y)-tract (Siegel et al., 2005; Kolev et al., 2010).

While in most studied eukaryotes, the splicing complex has been shown to interact with the transcription elongation machinery (Antosz et al., 2017) and the transcribed chromatin (Neves et al., 2017), such a cross-talk between RNA synthesis and maturation machineries has not been described in trypanosomes. Very little is known about the proteins and kinetics of transcription elongation in *T. brucei*. Nuclear run-on analyses (Martínez-Calvillo et al., 2004) and mathematical modelling (Fadda et al., 2013) indicate that RNAPII transcription elongation is constitutive in the kinetoplasts. Interestingly, RNAPI elongation at the active BES seems to be regulated through selective recruitment of the elongation and RNA processing machinery (Vanhamme et al., 2000). Usually, RNAPII transcription elongation is accompanied by

phosphorylation and dephosphorylation of distinct serines in the CTD of RPB1. This modification switch is required for regulating the transcription cycle (Hsin and Manley, 2012). The modified serines are part of the structured heptapeptide motif YSPTSPS, that appears in 26-52 repeats within the CTD of most eukaryotes. Interestingly, the CTD in trypanosomes does not contain this conserved motif but seems to be phosphorylated during elongation (Chapman and Agabian, 1994). The trypanosomal CTD appears to essential for RNA synthesis of protein-coding genes (Das and Bellofatto, 2009) and is suggested to mediate the recognition of the genomic chromatin landscape (Das et al., 2017).

Nucleosome remodelling accompanies RNAPII transcription

To be able to handle environmental signals and meet cellular needs, the RNAPII transcription machinery demands rapid access to genes. However, the recognition sites of the GTFs are usually buried into chromatin. Nucleosome arrays are able to halt RNAPII progression and thus, RNAPII associates with factors, which decrease nucleosome occupancy to increase DNA accessibility in its path (Malik et al., 2017).

Transcription through chromatin does not directly imply a complete decomposition of the beads-on-a-string formation. As consequence of the nucleosomal structure, the displacement of the H3/H4 tetramer is rare during transcription, but loss or exchange of H2A/H2B dimers can occur rapidly and continuously (Kireeva et al., 2002; Venkatesh and Workman, 2015). Generally, the removal of only one H2A/H2B dimer is sufficient for transcription through a single nucleosome (Venkatesh and Workman, 2015). However, highly transcribed genes require more histone exchange activities and, in this case, the H3/H4 tetramers are removed by the histone chaperone FACT (facilitates chromatin transcription) (Venkatesh and Workman, 2015).

Overall, the nucleosome occupancy at promoters is lower than at coding regions. Nevertheless, before PIC assembly can start, activator proteins need to bind to, and switch on, a cascade of coordinated processes for chromatin decompaction at promoter regions (Li et al., 2007). Pre-deposited histone acetylation, historically linked to active transcription, is suggested to destabilise nucleosomes within promoter regions and to mediate their displacement by chromatin remodelers (**Figure 1.12 a**) (Workman, 2006). Active promoters are almost completely devoid of nucleosomes. This chromatin structure is characteristic for eukaryotic promoters and is known as a nucleosome-depleted region (NDR) (Haberle and Stark, 2018). Nucleosomes flanking NDRs, known as +1 or –1 nucleosome, are subjected to increased histone exchange, crucial for efficient promoter escape of the transcription machinery and for transition to productive elongation (**Figure 1.12 b**) (Venkatesh and Workman, 2015). In many organisms, like *H. sapiens*, *Drosophila melanogaster* or *S. cerevisiae*, the +1 and/or -1 nucleosomes are marked by the incorporation of the H2A histone variant H2A.Z (Albert et al., 2007; Barski et al., 2007; Mavrich et al., 2008; Tirosh and Barkai, 2008). Organisms which show H2A.Z deposition in both flanking nucleosomes, show bidirectional transcription initiation. Whereas in

organism, such as *D. melanogaster*, where H2A.Z is incorporated only in +1 nucleosomes, transcription initiates unidirectionally (Venkatesh and Workman, 2015).

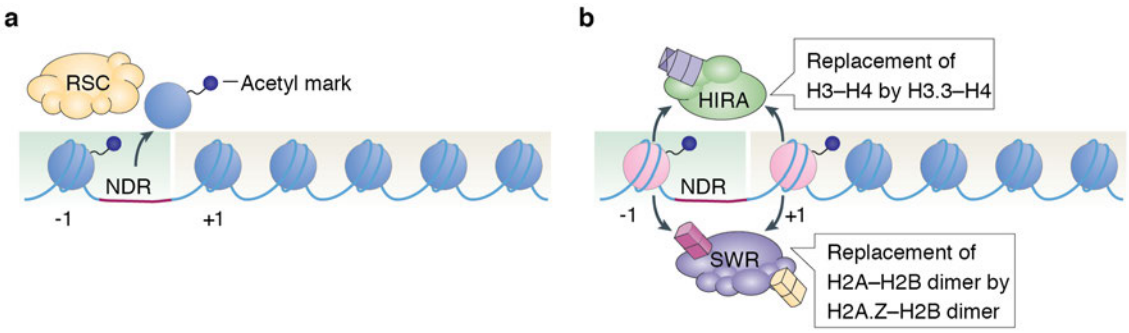


Figure 1.12. Nucleosome remodelling at yeast promoters.
a The RSC chromatin remodelling complex removes nucleosomes from the promotor to establish the nucleosome-depleted region (NDR). Histone acetylation ensures that RSC maintains the NDR at the promoter. **b** While the histone chaperone HIRA exchanges H3-H4 dimers by H3.3 dimers, the SWR1 chromatin remodelling complex replaces H2A-H2B with H2A.Z-H2B. The activity of HIRA and SWR1 is NDR-dependent. The alterations of the nucleosomal composition at promoters reduces the nucleosomal barrier and facilitates transcription elongation. The illustration was adapted from (Venkatesh and Workman, 2015).

Histone acetylation

Under all known PTMs, histone lysine acetylation has the highest capability to induce open chromatin (**Figure 1.13 a**) (Bannister and Kouzarides, 2011). Histone acetyltransferases (HATs) transfer an acetyl-group to side chains of lysines neutralizing their positive charge (**Figure 1.13 b**) (Bloch and Borek, 1946; Lipmann and Kaplan, 1946; Lipmann and Kaplan, 1947; Lipmann, 1954). In eukaryotes, HATs can be classified into type-A and type-B acetyltransferases. Type-B acetyltransferases are mainly cytoplasmic and acetylate free histones in the soluble pool. Type-A HATs are categorised into three subfamilies – GNAT, p300/CBP and MYST – and are able to modify lysines on chromatin incorporated histones (Roth et al., 2001). In yeast, the histone acetyltransferase NuA4 belonging to the MYST-proteins is responsible for H4 acetylation at promoter sites in yeast (Allard et al., 1999).

Histone acetylation is reversible, and the acetyl-group can be removed by histone deacetylases (HDACs), which are often involved in heterochromatin formation (Shahbazian and Grunstein, 2007). Particularly, unacetylated H4 tails are required for chromatin folding. The basic residues of the lysines form hydrogen bonds and salt bridges with the acidic patch of the H2A/H2B dimer that stabilise nucleosome interactions and chromatin folding. Thus, hyperacetylation of the H4 N-terminus is thought to counteract chromatin compaction (Annunziato et al., 1988; Tse et al., 1998; Dorigo et al., 2003; Siino et al., 2003; Shogren-Knaak et al., 2006) and is involved in transcription activation by promoting the establishment of a more open chromatin structure. For example, H4K16ac inhibits the formation of higher order chromatin structures by reducing inter-nucleosomal interactions (Shogren-Knaak et al., 2006) and changes the interplay of the H4 tail with chromatin modifying proteins (Zhang et al., 2016). Additionally, H4 acetylation is involved in the targeted deposition of H2A variants (Auger et al., 2008).

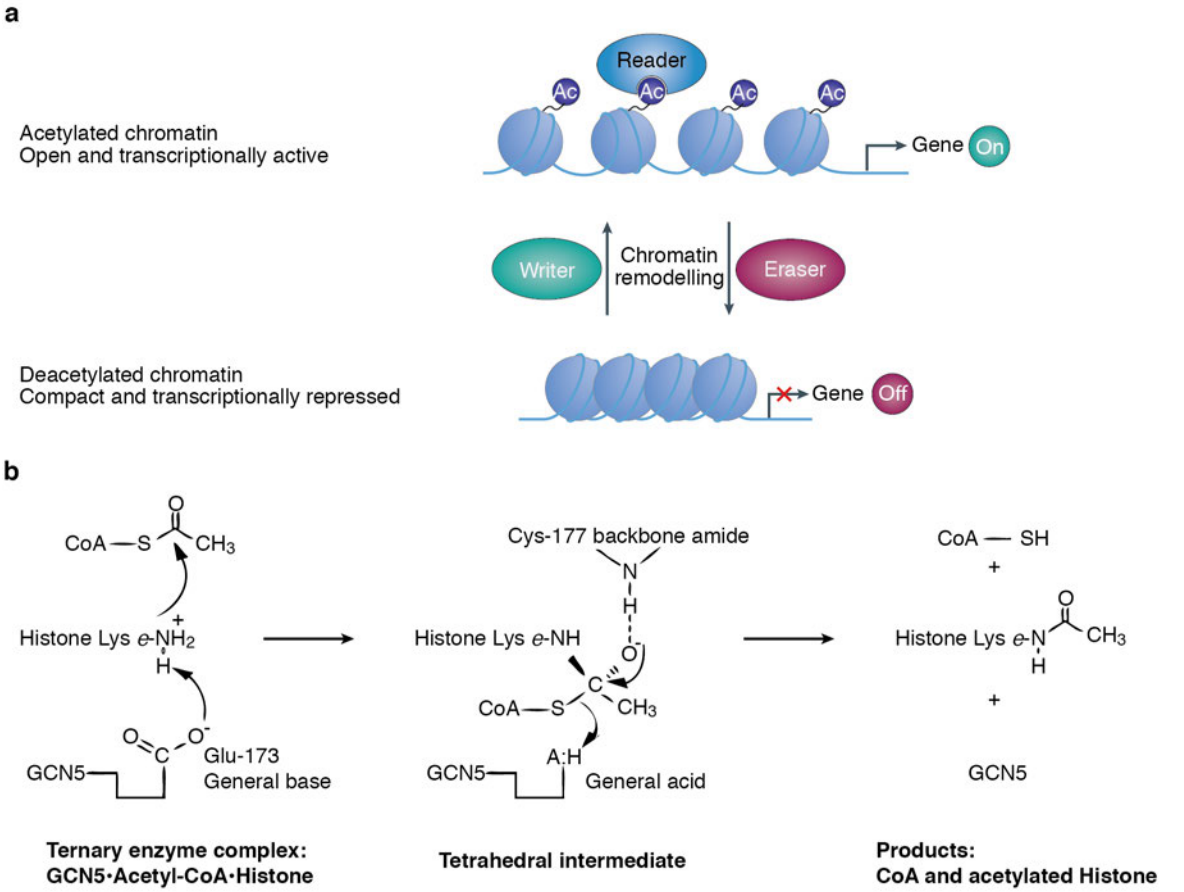


Figure 1.13. Histone acetylation is associated to active transcription.
a A string of nucleosomes is shown with different acetylation patterns. Acetylated chromatin is mainly open and transcriptionally active, because acetylated lysines leads to nucleosome destabilisation. Consequently, deacetylated chromatin is more compacted and transcriptionally repressed. The stages can be changed dynamically. So-called "writer" proteins can mediate lysine acetylation and "eraser" proteins can remove the modification. The illustration was adapted from (Verdin and Ott, 2015). **b** Scheme of catalytic and kinetic mechanism for lysine acetylation. The histone acetyltransferase GCN5 adds an acetyl group to a lysine on H3. GCN5 require the formation of a ternary complex (middle panel). The illustration was modified from (Roth et al., 2001).

Role of histone acetylation for transcription

More than 50 years ago, D.M. Phillips identified the presence of histone acetylation and Vincent Allfrey opened up the research field of its role in chromatin reorganisation and in transcription (Phillips, 1963; Allfrey et al., 1964). Already in 1964, he postulated "*dynamic and reversible mechanisms for activation as well as repression of RNA synthesis*". It took another 30 years until his hypothesis was validated.

Following Allfrey's observation that lysine acetylation reduces the barrier capability of histones for RNA synthesis, it could be shown that hyperacetylated histones are preferentially found in euchromatin and make DNA more susceptible to DNaseI digestion (Allfrey et al., 1964; Sealy and Chalkley, 1978; Vidali et al., 1978; Hebbes et al., 1988). However, the first direct link of histone acetylation and active transcription was described later revealing that mutations of H4 lysines to unacetylatable versions alter transcription in *S. cerevisiae* and global inhibition of mammalian HDACs causes increased gene expression (Durrin et al., 1991; Yoshida et al., 1995) (**Figure 1.14**). The non-selective nature of histone

acetyltransferases (Roth et al., 2001) suggested that the biological function of histone acetylation is more likely a result of the accumulation of multiple acetyl marks at a specific site rather than a single acetyl-mark (Dion et al., 2005).

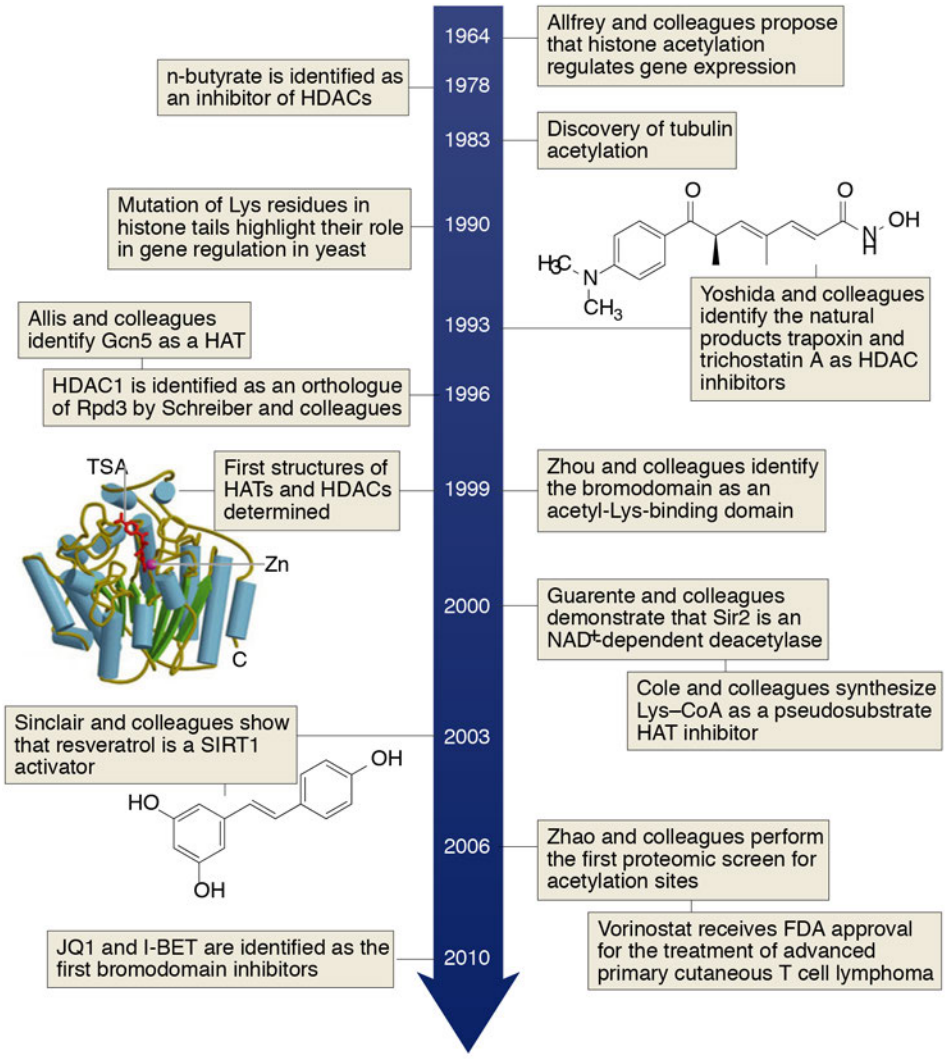


Figure 1.14. Almost 60 years of protein acetylation.
Time line showing the milestones of protein acetylation research including its role in transcription, the enzymes and inhibitors. FDA: Food and Drug Administration USA, HATs: histone acetyltransferases, HDACs: histone deacetylases, Sir2: silent information regulator 2, SIRT1: sirtuin 1. The illustration derives from (Verdin and Ott, 2015).

Acetylated lysines disrupt the electrostatic interactions between different nucleosomes, and between a histone and the phosphate group of the DNA strand within a nucleosome, by reducing the positive lysine charge (Bannister and Kouzarides, 2011; Verdin and Ott, 2015). As a result, an acetylated histone can be easier displaced from a nucleosome (Zhao et al., 2005). Thus, it is not surprising that acetylated histones have been mainly associated with sites of active transcription (Bannister and Kouzarides, 2011). In *S. cerevisiae*, acetylated H3 and H4 are enriched at promoters and the levels of acetylation correlate to the level of transcription (Pokholok et al., 2005). The histone acetyltransferases (HATs) Gcn5 and Esa1 are associated to promoters (Robert et al., 2004) and the SAGA complex interacts directly with an

activator protein at the GAL1 promoter prior PIC assembly (Brown et al., 2001). Upon activation of the PHO5 promoter, histones are hyperacetylated by SAGA and then displaced (Reinke and Hörz, 2003). There, the induced histone loss is mediated by the remodelling activity of the SWI/SNF complex (Workman, 2006), which is recruited to promoters by a transcriptional activator or by acetylated histones (Steger and Workman, 1996).

During the passage of RNAPII over gene bodies, pre-acetylated histones need to be rapidly deposited from the soluble pool into the transcribed chromatin for destabilizing nucleosomes in the path (Venkatesh and Workman, 2015). Moreover, the elongator complex itself possesses an intrinsic HAT activity and can acetylate histones in the path to guarantee rapid deposition and an RNAPII processing rate of 25 nt per seconds (**Figure 1.15**) (Conaway et al., 2000; Orphanides and Reinberg, 2000).

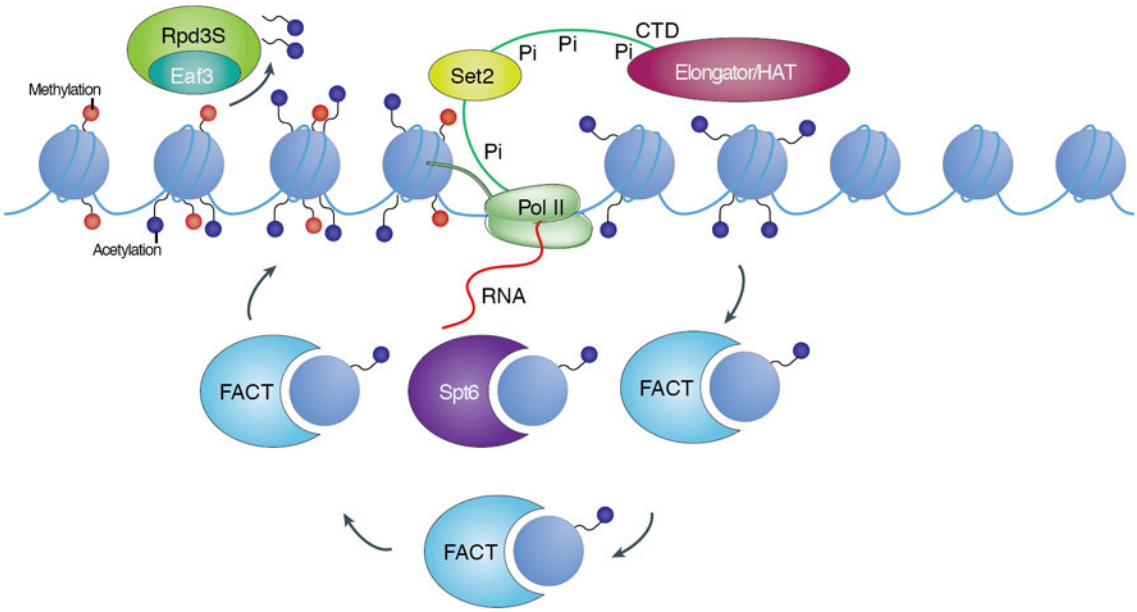


Figure 1.15. Histone acetylation is required for RNAPII progression.
The passage of RNAPII through chromatin requires displacement of nucleosome in the way of the transcription machinery. The elongation complex is associated to HATs that acetylate nucleosomes in front of RNAPII. This leads to destabilisation of the nucleosome and histones can be more easily displaced from the nucleosome. Through the activity of different histone chaperones (e.g. FACT) including the displaced are redeposited into the chromatin behind the transcription machinery. The displaced and reincorporated histones are hyperacetylated, which needs to be removed. Otherwise, the chromatin stays open and is unprotected. Set2-mediated H3K36 methylation recruits Rpd3S deacetylase through binding of the chromodomain Eaf3. Rpd3S removes the acetyl marks, which results in a stable nucleosome. The illustration was adapted from (Li et al., 2007).

Targeting of HAT complexes is not only promoted by transcriptional activators, but also by distinct lysine methyl marks, such as H3K4me or H3K36me. Interestingly, H3K4me is found mainly at the 5´end of genes, whereas H3K36me is distributed throughout coding regions (Krogan et al., 2003b; Ng et al., 2003). HAT complexes associate with chromodomain proteins – readers of lysine methylation. For example, SAGA recognises H3K4me through the chromodomain protein and chromatin remodeler Chd1 (Pray-Grant et al., 2005).

Reading histone acetylation

In addition to inducing chromatin unfolding, lysine acetylation also serves as a binding platform for proteins that are involved in transcription activation. Chromatin remodelling complexes, such as Swi-Snf or RSC, often contain bromodomains or interact with bromodomain proteins (BDFs) (Dhalluin et al., 1999; Hassan et al., 2002; Bennett and Peterson, 2015) that recognise and bind acetylated lysines. The Snf2 subunit of Swi-Snf has an intrinsic bromodomain that was shown to preferentially bind acetylated lysines at H3. Swi-Snf has been shown to mobilise nucleosome arrays and thereby, histone acetylation is required to stabilise the Swi-Snf/nucleosome binding. This process is known to be important for transcription activation (Hassan et al., 2001). At the *Pho* promoter, chromatin remodelling was reduced after removal of SAGA (Gregory et al., 1999). Moreover, established histone acetylation can maintain and enhance its presence by re-inducing the binding of the HAT through an intrinsic bromodomain, as described for Gcn5 in *S. cerevisiae* or TAFII250 in *H. sapiens* (Jacobson et al., 2000; Owen et al., 2000). Histone acetylation is also required for stimulating selective histone variant exchange. In *D. melanogaster*, nucleosomal phosphor-H2Av is acetylated by dTip60 and replaced with an unmodified H2Av (H2A.Z in other eukaryotes) at DNA lesions (Kusch et al., 2004). NuA4-mediated H4 acetylation induces H2A.Z incorporation at promoter regions in *S. cerevisiae* forming a specific chromatin structure, which facilitates transcription initiation. A similar process is present in humans (Auger et al., 2008; Altaf et al., 2010).

Histone acetylation in trypanosomes

Trypanosomes encode six different histone acetyltransferases (Ivens et al., 2005; Kawahara et al., 2008), seven histone deacetylases (Ivens et al., 2005) and five bromodomain proteins (Ivens et al., 2005; Siegel et al., 2009b; Schulz et al., 2015).

Out of the six putative HATs encoded in the *T. brucei* genome (Ivens et al., 2005; Kawahara et al., 2008), three also belong to the MYST-family, namely HAT1-3. All three HATs localise in the nucleus and while HAT1 and HAT2 are indispensable for growth (Kawahara et al., 2008), cells, in which both alleles of HAT3 were removed, are viable (Siegel et al., 2008). A systematic target screen has revealed that HAT3 mediates only a single, but highly abundant, histone acetyl mark at H4K4 (ElBashir et al., 2015) in a cell cycle-dependent manner (Siegel et al., 2008). Reduction of HAT1 levels compromises telomeric silencing and HAT2 depletion reduces the levels of H4K10ac (Kawahara et al., 2008), which is a TSSs-specific acetyl mark (Siegel et al., 2009b). No other target sites of HAT1 and HAT2 have thus far been identified. Both enzymes contain a chromodomain in their N-terminal domains suggesting that methyl marks might play a role in recruiting HAT1 and HAT2 to specific sites (Kawahara et al., 2008). Interestingly, H4K10ac co-localises with H3K4me3 at TSSs (Wright et al., 2010). Thus, it might be possible that H3K4 methylation at TSSs recruits HAT2, which modifies H4K10 at these sites. However, thus far it has not been analysed if H3K4me3 is a prerequisite for H4K10ac in trypanosomes. Next to

K4 and K10, H4 is acetylated at K2 and K5. Even though the primary sequence of trypanosomal histone H4 diverges from those of other eukaryotes, it is also extensively acetylated at the N-terminal tail (Janzen et al., 2006a; Mandava et al., 2007).

H2A.Z – a multiplayer in transcription regulation

H2A possesses the greatest variety of variant forms among the four core histones, which include H2A.Z, H2A.X, H2A.Bbd, and macroH2A (Wu et al., 1982; Pehrson and Fried, 1992; Chadwick and Willard, 2001). H2A variants differ from their canonical counterparts in the sequence or length of the globular domain or in the genomic distribution (Kamakaka and Biggins, 2005).

The histone variant H2A.Z is present in many eukaryotes and evolutionary highly conserved, being found in organisms diverse as *S. cerevisiae*, *D. melanogaster*, *H. sapiens* and also in smaller eukaryotes, like protozoan parasites (Zlatanova and Thakar, 2008). Despite being a histone variant, H2A.Z proteins are even less divergent in the eukaryotic lineage than canonical H2A proteins (West and Bonner, 1980; Thatcher and Gorovsky, 1994). H2A.Z has diverse biological functions and has been linked to, amongst others, regulation of heterochromatin silencing, cell cycle progression, chromosome segregation and gene expression (Allis et al., 1980; Stargell et al., 1993; Santisteban et al., 2000; Faast et al., 2001; Meneghini et al., 2003; Krogan et al., 2004; Mizuguchi et al., 2004; Rangasamy et al., 2004; Ridgway et al., 2004; Zhang et al., 2004; Swaminathan et al., 2005; Babiarz et al., 2006; Greaves et al., 2006; Deal et al., 2007; Greaves et al., 2007).

Histone H2A variant H2A.Z

The histone variant H2A.Z evolved early and only once during eukaryotic evolution (Thatcher and Gorovsky, 1994). It has only ~60% sequence identity to its canonical counterpart within one species, but between different species the sequence identity of H2A.Z is usually strikingly high (~80%). However, the H2A.Z proteins found in kinetoplasts do not share such a high sequence similarity to the H2A.Z proteins of other eukaryotes, e.g. the eukaryotic parasite *T. brucei* shares around 55% sequence similarity to the H2A.Z proteins of other eukaryotes (Lowell et al., 2005).

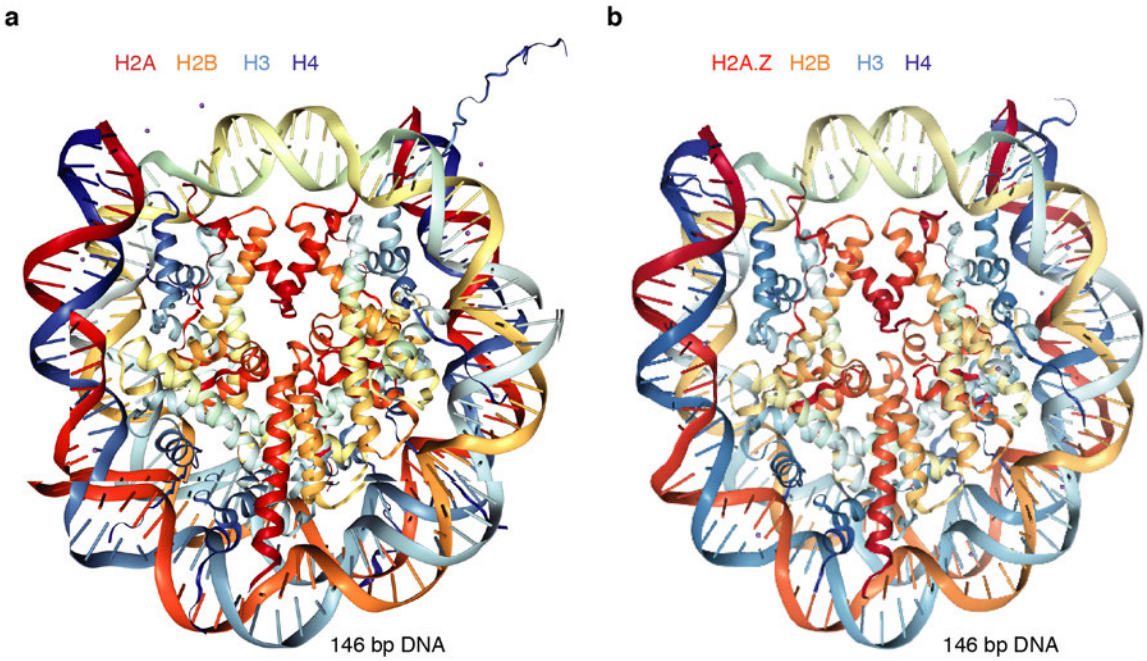


Figure 1.16. Nucleosomes containing H2A.Z resemble canonical nucleosomes.
a Crystal structure of the canonical nucleosome (Luger et al., 1997) downloaded from The Protein Data Bank rcsb.org (Berman et al., 2000). **b** Crystal structure of the H2A.Z-containing nucleosome (Suto et al., 2000) downloaded from The Protein Data Bank rcsb.org (Berman et al., 2000). The structure of the H2A.Z-containing nucleosome is similar to the nucleosome containing the major histone proteins. However, H2A.Z-containing nucleosomes have an altered surface with a metal ion. Moreover, the interaction between the (H2A.Z-H2B) dimer and the (H3-H4) tetramer are slightly destabilised through position changes (Suto et al., 2000).

The H2A.Z-containing nucleosome does not reflect the sequence difference between the variant and core histone. Only the region of H2A and H2A.Z which is important for dimer interaction within a nucleosome exhibits a subtle structural difference (**Figure 1.16 a and b**) (Luger et al., 1997; Suto et al., 2000).

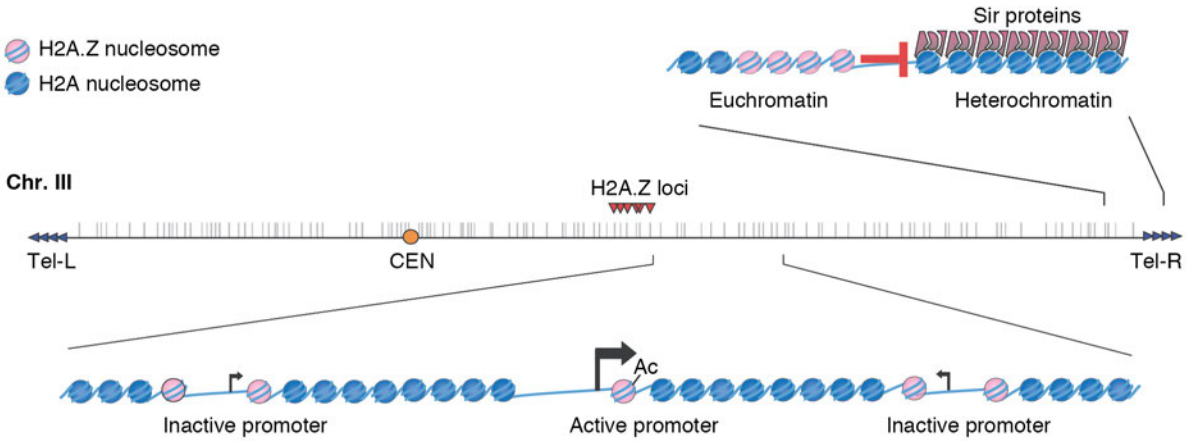


Figure 1.17. Scheme of H2A.Z enrichment sites in the yeast genome.
H2A.Z enrichment sites (grey lines) on chromosome 3 are shown with centromeric (CEN) and telomeric (Tel-L, Tel-R) regions. While active genes are marked by H2A.Z acetylated on K14, unmodified H2A.Z is flanking the NDRs of inactive promoters. Near telomeres, the H2A.Z enrichment sites are wider and K14 acetylation is depleted. There, H2A.Z is thought to prevent heterochromatin spreading. The illustration was modified from (Guillemette and Gaudreau, 2006).

In many organisms H2A.Z is essential for viability, for example in *Mus musculus* (Faast et al., 2001), *Tetrahymena thermophila* (Liu et al., 1996), *D. melanogaster* (van Daal and Elgin, 1992; Clarkson et al., 1999), *Xenopus leavis* (Iouzalen et al., 1996; Ridgway et al., 2004), and in *T. brucei* (Lowell et al., 2005). In *S. cerevisiae* and *Schizosaccharomyces pombe* loss of H2A.Z is not lethal but causes a severe growth defect (Carr et al., 1994; Jackson and Gorovsky, 2000).

In most eukaryotes, H2A.Z is a conserved feature of the TSS-chromatin structure and is implicated in many transcriptional regulatory processes. In *H. sapiens*, H2A.Z is found at promoters upstream and downstream of TSSs and its binding levels correlate with gene activity (Barski et al., 2007; Hardy et al., 2009). In *S. cerevisiae*, H2A.Z (*Htz1*) has been found at the 5´ends of active and inactive genes in euchromatic regions and marks inactive promoters (Guillemette et al., 2005; Raisner et al., 2005). A genome-wide map of H2A.Z (H2Av) in *D.* melanogaster has revealed that at least one H2A.Z-nucleosome is localised within 1 kb of the TSSs of 85% of coding genes. Additionally, the levels of gene expression correlate with the levels of the histone variant (Mavrich et al., 2008). H2A.Z at TSSs is subjected to a high histone turnover and efficient eviction of H2A.Z is dependent on PIC formation (Tramantano et al., 2016).

Despite its predominant localisation at TSSs, H2A.Z has also been shown at other genomic regions. In *S. cerevisiae,* H2A.Z is also found at gene clusters adjacent to telomeres and prevents telomeric silencing (**Figure 1.17**) (Meneghini et al., 2003; Babiarz et al., 2006).

In *D. melanogaster*, the localisation of H2A.Z is not restricted to transcriptionally active sites and is also found at heterochromatic regions and at non-coding, euchromatic genes (Leach et al., 2000). In mammals, H2A.Z is indeed absent from constitutive heterochromatin, but it is found on the inactive X-chromosome and in facultative heterochromatin. However, the heterochromatic H2A.Z is monoubiquitinylated (Sarcinella et al., 2007; Hardy et al., 2009). Furthermore, H2A.Z is recruited to gene bodies as a consequence of transcription repression (Hardy et al., 2009; Lashgari et al., 2017) (**Figure 1.18**).

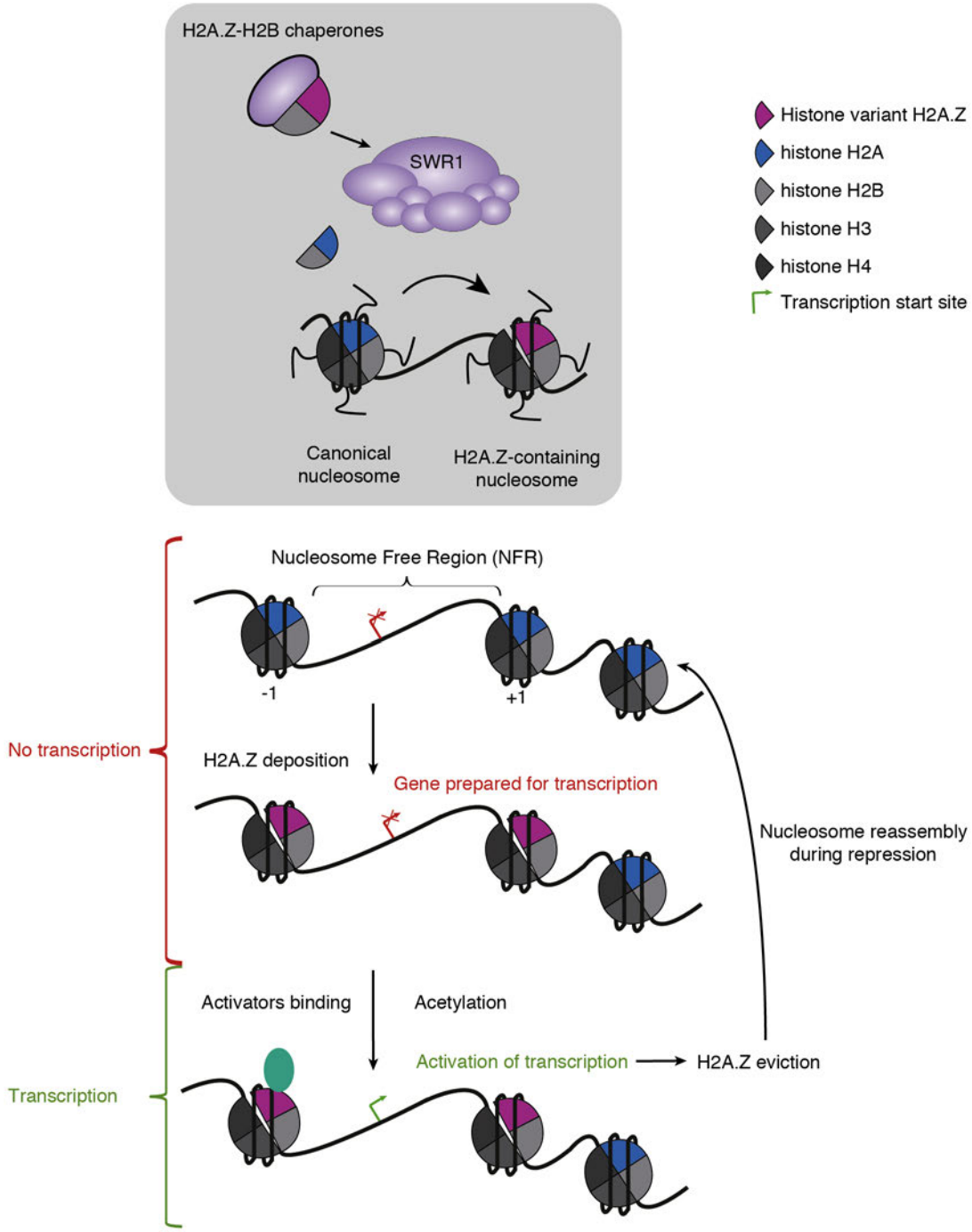


Figure 1.18. H2A.Z and its role in gene activation.
Incorporation of H2A.Z is mediated by the SWR1-complex in yeast and leads to changes in the nucleosomal structure, which are implicated in transcription. Unmodified H2A.Z is incorporated into the +1 and -1 nucleosomes flanking the NDR at promoters and poises the genes for rapid transcription activation. Acetylation of H2A.Z facilitates nucleosome eviction for the transcription machinery and is involved in transcription initiation. The illustration was modified from (Billon and Cote, 2013).

Role of acetylated H2A.Z and not acetylated H2A.Z

The complexity of the regulatory tasks that H2A.Z undertakes is, in part, a consequence of its diverse acetylation status. However, in contrast to unmodified H2A.Z, H2A.Zac does not serve as an epigenetic mark (Bruce et al., 2005).

In *S. cerevisiae*, H2A.Z itself is not sufficient to prevent heterochromatin spreading. H2A.Z needs to be acetylated by NuA4 following its incorporation and promotes further SAGA-mediated acetylation (Babiarz et al., 2006). In *Drosophila*, H2Av, the homologue of H2A.Z and H2A.X, is phosphorylated at sites of double strand breaks. As consequence, it is acetylated by dTip60 and can then be replaced by an unmodified version to re-establish the chromatin structure at those sites(Kusch et al., 2004). Deletion or mutation of H2A.Z to an unacetylable version results in chromosome entanglement and breakage during anaphase, demonstrating that H2A.Zac is essential for anaphase chromosome architecture in *S. pombe* (Kim et al., 2009).

The distinguishable functionalities of acetylated H2A.Z are caused by two modes of action, either generating a charge patch or a site-specific histone code. If the overall charge of a region mediates the regulatory function, mimicking lysine acetylation would not influence the cellular phenotype, but modulating the charge would have the same effect as deletion of the complete histone, as demonstrated for the highly basic histone H1 (Dou and Gorovsky, 2000). If a specific acetylation site defines the regulatory function, modifying this specific lysine would already affect the cellular phenotype. A mutagenesis screen in *Tetrahymena*, where H2A.Zac was mimicked or deleted, revealed that reducing the positive charge of the N-terminal lysine mediates the essential role of H2A.Z (Ren and Gorovsky, 2001).

In general, the H2A.Z-containing nucleosome resembles a canonical nucleosome (Suto et al., 2000), however H2A.Zac changes the structure and causes the tail to dissociate, reducing DNA-histone interactions and making the nucleosome more accessible for chromatin modifying proteins (Ren and Gorovsky, 2001). H2A.Zac with its anti-silencing functions is probably more important for transcription regulation than unmodified H2A.Z, as emphasised by many studies that associate acetylation of H2A.Z to active transcription. In chickens, the hyperacetylated form of H2A.Z is present at the 5´end of active genes but is absent at inactive genes (Bruce et al., 2005). In human prostate epithelial cells, the acetylated form of H2A.Z is a key feature of promoters of highly expressed genes (Valdes-Mora et al., 2012). The N-terminus of H2A.Z (Htz1) in yeast is acetylated at K3, K8, K10 and K14 and H2A.ZK14ac is exclusively enriched at active promoters (Millar et al., 2006). There, H2A.Zac is probably important for gene induction, as observed at the *GAL1* promoter (Halley et al., 2010). Additionally, acetylation prevents H2A.Z to be removed from promoter regions by the chromatin remodelling complex INO80 (Papamichos-Chronakis et al., 2011a).

H2A.Z deposition pathway

The histone variant H2A.Z is predominantly found at promoter regions and TSSs in many organisms (Guillemette et al., 2005; Raisner et al., 2005; Barski et al., 2007; Mavrich et al., 2008; Hardy et al., 2009).

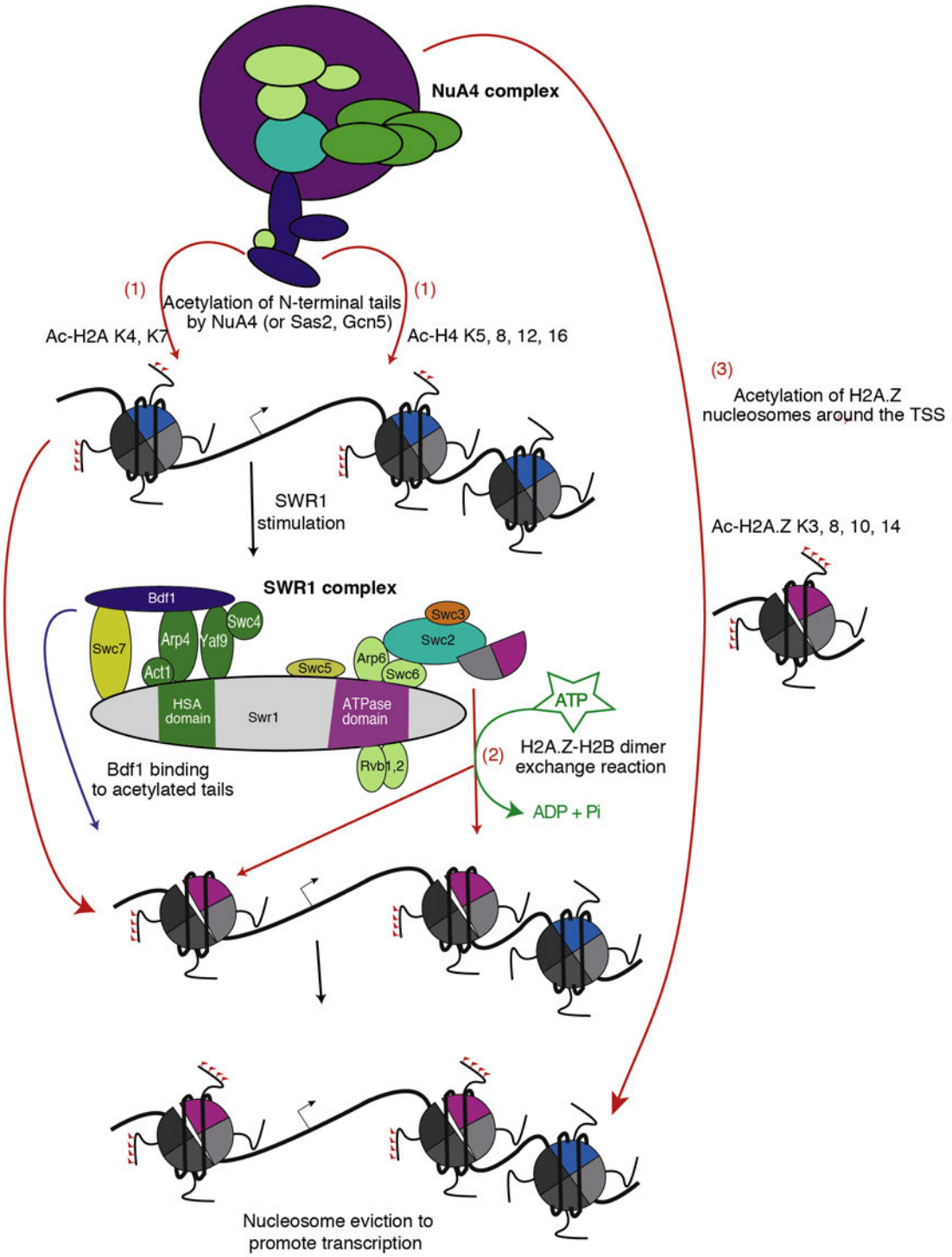


Figure 1.19. The SWR1-complexes exchanges H2A by H2A.Z at promoters in yeast.
Scheme of the processes underlying H2A.Z acetylation and deposition in yeast. The histone acetyltransferase complex NuA4 acetylation canonical H2A, H4 and H2A.Z. The acetyl marks on H2A and H4 recruit the chromatin remodeler SWR1 through interaction with the bromodomain subunit BDF1. SWR1 exchanges in an ATP-dependent manner each H2A-H2B dimer with an H2A.Z-H2B dimer step-by-step. Local NuA4-mediated H2A.Z acetylation at promoter sites is important for activation of gene expression. The illustration was modified from (Billon and Cote, 2013).

The evolutionary highly conserved SWR1-complex (SWR1-c), which mediates the targeted H2A.Z depositioning, was firstly described in *S. cerevisiae* (Krogan et al., 2003a; Mizuguchi et al., 2004). SWR1-c consists of 14 subunits including the Snf2-ATPase SWR1-C, which is responsible for the recruitment and incorporation of Htz1 (Kobor et al., 2004). The other 13 proteins are the non-essential proteins

BDF1, Arp6, Yaf9, Swc2, Swc3, Swc5, Swc6, and Swc7 and the five essential proteins Rvb1, Rvb2, Arp4, Swc4, and Act1. Some of these subunits are shared with other CRCs such as NuA4 and/or INO80 (reviewed in (Billon and Cote, 2013)). All three of these CRCs have been found to interact with H2A.Z (Kobor et al., 2004; Babiarz et al., 2006; Papamichos-Chronakis et al., 2011a) suggesting a functional link between these complexes. Homologues to the SWR1-c have been identified in *H. sapiens* with the enzyme complex SRCAP (Johnston et al., 1999; Ruhl et al., 2006) and in *Drosophila melanogaster* with the enzyme complex dTip60 (Kusch et al., 2004).

The SWR1-c evicts H2A and incorporates H2A.Z at specific genomic sites. It is believed that SWR1-c binding is stimulated by a specific DNA sequence in the nucleosome-free region (Ranjan et al., 2013) and that the binding is enhanced by NuA4-dependent acetylation of H4 and H2A (Altaf et al., 2010). The bromodomain factor BDF1, which is a subunit of SWR1-c, recognises and binds to acetylated lysines at H4 (Ladurner et al., 2003), is important for the site-specific incorporation of H2A.Z (Zhang et al., 2005; Durant and Pugh, 2007). The ATPase domain of Swr1-C mediates the stepwise replacement of H2A by H2A.Z in the nucleosomes flanking the NDR (Luk et al., 2010). The ATP-hydrolysis is only stimulated by incorporated H2A/H2B dimer and not by DNA (Mizuguchi et al., 2004) (**Figure 1.19**). Conversely, spreading of H2A.Z into gene bodies and/or mislocalisation of H2A.Z can be prevented by the FACT complex, supported by the histone chaperone Spt6, (Jeronimo et al., 2015) or the chromatin remodelling complex INO80, that can recognise unacetylated H2A.Z (Venkatesh and Workman, 2015) (**Figure 1.20**).

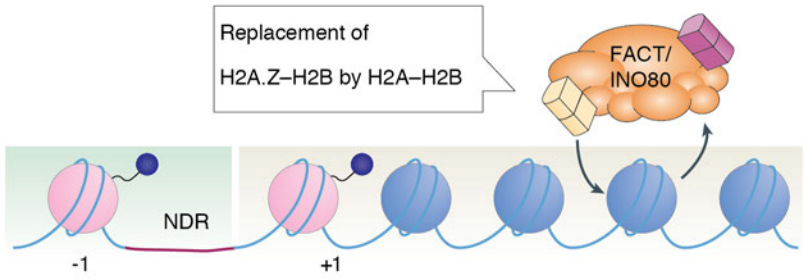


Figure 1.20. Mislocalised H2A.Z can be removed by FACT and/or INO80.
Mislocalised H2A.Z across the coding regions can be removed by the histone chaperone FACT and/or by the chromatin remodeler INO80. INO80 can only displace unmodified H2A.Z. Thus, to prevent removal of H2A.Z from the +1 and -1 nucleosome, H2A.Z acetylation ensures its retention at the promoter. The illustration was modified from (Venkatesh and Workman, 2015).

H2A.Z assists in RNA synthesis and maturation

The first evidence demonstrating that H2A.Z plays a role in active transcription regulation stems from research done in the protozoa *Tetrahymena thermophilia*, where H2A.Z is only found in the transcriptionally active macronucleus (Allis et al., 1980). Additionally, it has been shown that its presence correlates with gene activity in mammals (Bargaje et al., 2012) and in *S. cerevisiae*, H2A.Z is essential for optimal transcription initiation of repressed genes (Zhang et al., 2005). Since H2A.Z is associated with promoter regions and TSSs, initial studies mainly focused on its role in regulating transcription

initiation. However, in the past years many studies revealed that H2A.Z is implicated in various steps of the transcription process acting as a molecular caretaker for RNAPII (Subramanian et al., 2015).

During transcription initiation, H2A.Z marks the promoter sites for recruitment of transcription factors, such as the TATA-binding protein, and maintains persistent gene expression (Wan et al., 2009; Li et al., 2012). Additionally, genes with higher H2A.Z levels exhibit less RNAPII stalling and nucleosome occupancy. The incorporation of H2A.Z modulates the nucleosomal barrier of the +1 nucleosome to guarantee efficient RNAPII release from the pausing site, required for the transition to productive elongation (Weber et al., 2014). Incorporation of H2A.Z at TSSs does not only affect transcription initiation, elongation also requires H2A.Z-containing nucleosomes (Santisteban et al., 2011; Rudnizky et al., 2016). H2A.Z assists in RNAPII passaging probably by promoting nucleosome remodelling, e.g., by stimulating the activity of ISWI remodelers (Goldman et al., 2010), and elongation factor recruitment (Malagon et al., 2004). Deletion of H2A.Z reduces the elongation rate of RNAPII to 24% over the 9.5 kb *GAL10p-VPS13* gene in yeast. In the absence of H2A.Z, the nucleosome occupancy increases over the gene body. Additionally, the correct serine phosphorylation status of RNAPII and the assembly of the elongator complex depend on H2A.Z (Santisteban et al., 2011). With the knowledge that pre-mRNA splicing occurs co-transcriptionally, researchers started to question if H2A.Z (important for elongation which is itself a pre-requisite for splicing) is also required for RNA maturation. Indeed, two independent studies revealed that in *S. pombe*, H2A.Z occupancy promotes co-transcriptional splicing of suboptimal introns that may otherwise be discarded via proofreading ATPase (Nissen et al., 2017). Whereas in *S. cerevisiae*, H2A.Z aligns RNAPII progression and RNA processing by coordinating transcription elongation with spliceosome assembly (Neves et al., 2017).

H2A.Z in trypanosomes

Even though, H2A.Z is one of the least divergent histones in *T. brucei*, it only has a sequence identity of around 55% with other H2A.Z proteins and around 43% sequence identity with its canonical counterpart in *T. brucei*. Trypanosomal H2A.Z has a long N-terminal stretch, which is not present in other H2A.Z and dimerises exclusively with the H2B variant H2B.V (Lowell et al., 2005). This suggests that heterotypic H2A.Z/H2A nucleosome, which have been detected in other organisms (Luk et al., 2010), are absent in trypanosomes. Similar to trypanosomal H2A.Z, the N-terminus of H2B.V is longer compared to canonical H2B.

H2A.Z/H2B.V-containing nucleosomes are less stable than canonical nucleosomes and are enriched at TSSs (Siegel et al., 2009b). Thus, it is likely that the incorporation of H2A.Z and H2B.V results in a more open chromatin structure at TSSs, which facilitates the binding of the RNAPII transcription machinery for transcription initiation. While it has been shown that the GT-rich sequence elements at TSSs mediate H2A.Z deposition (Wedel et al., 2017), the mechanisms behind H2A.Z deposition in *T. brucei* are still elusive. Just recently, a novel SNF2 ATPase complex has been identified in *T. brucei* that has a potential

role in depositioning H2A.Z to TSSs in *T. brucei* (Vellmer et al., 2021). However, it is not clear whether the nucleosome-binding proteins that co-localise with H2A.Z at TSSs, including BDF1, BDF3, and BDF4 (Siegel et al., 2009b; Schulz et al., 2015) and the chromatin remodeler ISWI (Stanne et al., 2015), are important for H2A.Z deposition or shown to interact with H2A.Z. Moreover, a comprehensive analysis of PTMs on H2A.Z and H2B.V is missing.

Goal of the study

In almost every situation of life, the understanding of the cause and consequence of something, requires the knowledge of every single detail that is involved in it, generates it, influences it, is depend on it and originates from it. These can be simple things, such driving a car, or more complicated things, such as biological processes including transcription.

It is clear that the packaging of the DNA into chromatin affects transcription, because transcription requires DNA-binding and chromatin influences the accessibility of the DNA. However, this organisation is not only an obstacle for transcription, it also represents a very elegant way to regulate transcription. By opening up the chromatin at a specific gene particularly at its promoter or TSS, transcription of that gene can be switched on and by closing the chromatin at this site, it can be switched off. Transcription is fundamental for every organism and thus, it is controlled very tightly by various processes to ensure that it works fine. In line with this, transcription is usually accompanied by processes mediating chromatin opening, to provide DNA-accessibility for the transcription machinery. Since it is thought that transcription initiation is the most important regulatory step (Chen et al., 2018), it is not surprising that the chromatin structure at promoters and TSSs is very well studied in many eukaryotes, including yeast and humans. Chromatin at those sites is marked by acetylated histone H4 and the H2A variant H2A.Z. Moreover, it seems to be the case, that H2A.Z at those sites is also acetylated. While many findings showed that the TSS-specific chromatin marks are by themselves required for or involved in transcription initiation, whether they are acting together has thus far been difficult to disentangle. In higher eukaryotes and in yeast, the individual processes establishing TSS-specific chromatin structures are complex and also interdependent. With this in my mind, I asked at the beginning of my PhD how is TSS-chromatin established in *T. brucei*? Trypanosomes are unicellular, eukaryotic parasites that branched early from the eukaryotic lineage during evolution and has a rather simple genome organisation. Yet, the chromatin at TSSs is marked by acetylated histones and incorporated H2A.Z, as in other eukaryotes. Thus, I used *T. brucei* as a model to study TSS-specific chromatin establishment in eukaryotes and addressed following questions:

- What are the differences between TSS- and non-TSS-chromatin structures?
- Which factors are involved in establishing TSS-specific chromatin structures? In particular, does histone acetylation play a role?
- Are these factors required for establishing TSS-specific chromatin important for transcription regulation?
- And finally, how would these factors regulate transcription?

Answering these questions will not only shed on the role of TSS-specific chromatin in *T. brucei* but will also help to understand the link between histone modifications, histone variants and transcription regulation in higher eukaryotes.

Chapter 2: Material and Methods

Molecular biology strategies

Polymerase chain reaction

To amplify DNA fragments from genomic DNA (gDNA) extracted from *T. brucei* as template, in order to integrate it into a plasmid backbone afterwards, a 50 µl polymerase chain reaction (PCR) was set up containing 100 ng of gDNA, 0.2 mM of each dNTP (Thermo Scientific), 0.5 µM of the specific forward and reverse primer (synthesised by Sigma Aldrich), 1 U of Phusion High-Fidelity (HF) DNA Polymerase (Thermo Scientific), 1 x Phusion HF Buffer from a 5 x stock solution (Thermo Scientific) and filled up to 50 µl with milli-Q water. The cycling conditions were adjusted according to the melting temperature of the primers and the amplicon length following the manufacturer's instruction for the Phusion HF DNA Polymerase: 98 °C/30 sec, 25 cycles (98 °C/10 sec – X °C/30 sec – 72 °C/30 sec/kb), 72 °C/5 min, 12 °C/hold. To remove the salts of the PCR buffers, the 50 µl PCR reaction was cleaned up using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey&Nagel). To analyse the amplification efficiency, 500 ng of the purified PCR reaction were run on a 1.5% agarose gel using the Tris-Acetic acid-EDTA (TAE) buffer system. As size standards for the DNA fragments, 500 ng of the GeneRuler 1 kb or 100 bp DNA ladder were used (Thermo Scientific). The PCR fragments were visualised using a Ultraviolet (UV) illuminator after staining for 15 min in ethidium bromide (Roth) water bath. In case the PCR reaction contained undesired DNA fragments, the whole PCR reaction was run on a 1.5% agarose gel using the Tris-Acetic acid-EDTA (TAE) buffer system and the desired DNA fragment was cut out and purified using the manufacturer's instructions for DNA extraction from agarose gels of the NucleoSpin® Gel and PCR Clean-up Kit (Macherey&Nagel). Nucleotide concentrations were determined using a NanoDrop spectrometer (Thermo Scientific).

To amplify DNA fragments from plasmid DNA as template, in order to integrate it into a plasmid backbone afterwards, a 50 µl polymerase chain reaction (PCR) was set up containing 100 ng of gDNA, 0.2 mM of each dNTP (Thermo Scientific), 0.5 µM of the specific forward and reverse primer (synthesised by Sigma Aldrich), 1 U of Phusion High-Fidelity DNA Polymerase (Thermo Scientific), 1 x Phusion HF Buffer from a 5 x stock solution (Thermo Scientific) and filled up to 50 µl with milli-Q water. The cycling conditions were adjusted according to the melting temperature of the primers and the amplicon length following the manufacturer's instruction for the Phusion High-Fidelity DNA Polymerase: 98 °C/30 sec, 25 cycles (98 °C/10 sec – X °C/30 sec – 72 °C/30 sec/kb), 72 °C/5 min, 12 °C/hold. To remove the salts from the PCR buffers, the 50 µl PCR reaction was cleaned up using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey&Nagel). To analyse the amplification efficiency, the complete PCR reaction was run on a 1.5% agarose gel using the Tris-Acetic acid-EDTA (TAE) buffer system. As size standards for the DNA fragments, 500 ng of the GeneRuler 1 kb or 100 bp DNA ladder were used

(Thermo Scientific). The PCR fragments were visualised using a UV illuminator after staining for 15 min in ethidium bromide (Roth) water bath. The DNA fragment of interest was cut out from the gel and purified using the manufacturer's instructions for DNA extraction from agarose gels of the NucleoSpin® Gel and PCR Clean-up Kit (Macherey&Nagel) to remove the template plasmid DNA. Residual plasmid DNA would complicate further cloning efforts. Nucleotide concentrations were determined using a NanoDrop spectrometer (Thermo Scientific).

To test plasmids directly from bacterial colonies, for example after transformation (see below), a 20 µl reaction was set up containing 0.2 mM of each dNTP (Thermo Scientific), 0.5 µM of the specific forward and reverse primer (synthesised by Sigma Aldrich), 10% DMSO, 1 U of DreamTaq DNA Polymerase (Thermo Scientific), 1 x DreamTaq Buffer from a 10 x stock solution (Thermo Scientific) and filled up to 20 µl with milli-Q water. Per reaction one colony was picked and transferred to the PCR tube containing the PCR reaction. Remaining bacteria on the tip were dotted on a LB-agar plate supplemented with 50 µg/ml Ampicillin and incubated over night at 37 °C. The cycling conditions were adjusted according to the melting temperature of the primers and the amplicon length following the manufacturer's instruction for the DreamTaq DNA Polymerase Polymerase: 95 °C/5 min, 40 cycles (95 °C/3 min – X °C/30 sec – 72 °C/1 min/kb), 72 °C/5 min, 12 °C/hold. The complete PCR reaction was run on a 1.5% agarose gel using the Tris-Acetic acid-EDTA (TAE) buffer system. As size standards for the DNA fragments, 500 ng of the GeneRuler 1 kb or 100 bp DNA ladder were used (Thermo Scientific). The PCR fragments were visualised using a UV illuminator after staining for 15 min in ethidium bromide (Roth) water bath.

To verify the integration of a transfected DNA construct into the target site within the *T. brucei* genome so-called integration PCRs were performed using 2 µl of genomic DNA (see below) and DreamTaq DNA Polymerase as described above. The cycling conditions were adjusted according to the melting temperature of the primers and the amplicon length following the manufacturer's instruction for the DreamTaq DNA Polymerase Polymerase: 95 °C/3 min, 40 cycles (95 °C/30 sec – X °C/30 sec – 72 °C/1 min/kb), 72 °C/5 min, 12 °C/hold. The complete PCR reaction was run on a 1.5% agarose gel using the Tris-Acetic acid-EDTA (TAE) buffer system. As size standards for the DNA fragments, 500 ng of the GeneRuler 1 kb or 100 bp DNA ladder were used (Thermo Scientific). The PCR fragments were visualised using a UV illuminator after staining for 15 min in ethidium bromide (Roth) water bath.

Restriction digest

To prepare plasmid backbone and insert for ligation or InFusion reaction (see below), 2 µg of backbone or insert DNA were digested in a 50 µl reaction containing the respective restriction enzyme/s (New England Biolabs) according to the manufacturer's instructions for the enzyme, 1 x NEB buffer as recommended by the manufacturer and milli-Q water. The reaction was incubated for 1 h at the temperature according to the manufacturer's instructions for the enzyme. The backbone digest reaction was run on a 1.5% agarose gel using the Tris-Acetic acid-EDTA (TAE) buffer system. As size standards

for the DNA fragments, 500 ng of the GeneRuler 1 kb or 100 bp DNA ladder were used (Thermo Scientific). The DNA fragments were visualised using a UV illuminator after staining for 15 min in ethidium bromide (Roth) water bath. The DNA fragment of the backbone was cut out from the gel and purified using the manufacturer's instructions for DNA extraction from agarose gels of the NucleoSpin® Gel and PCR Clean-up Kit (Macherey&Nagel). To remove the salt of the digestion buffer, the insert digest reaction was purified using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey&Nagel) according to the manufacturer's instruction. Nucleotide concentrations were determined using a NanoDrop spectrometer (Thermo Scientific).

To linearise plasmids for transfection 50 µg of DNA were digested in a 100 µl reaction containing the respective restriction enzyme/s (New England Biolabs) according to the manufacturer's instructions for the enzyme, 1 x NEB buffer as recommended by the manufacturer and milli-Q water. The reaction was incubated for 3 h at the temperature according to the manufacturer's instructions for the enzyme. To remove the salt of the digestion buffer, the linearisation reaction was purified using the NucleoSpin® Gel and PCR Clean-up Kit (Macherey&Nagel) according to the manufacturer's instruction. Nucleotide concentrations were determined using a NanoDrop spectrometer (Thermo Scientific). To analyse the linearisation efficiency, 500 ng of the reaction were run on a 1.5% agarose gel using the Tris-Acetic acid-EDTA (TAE) buffer system. As size standards for the DNA fragments, 500 ng of the GeneRuler 1 kb or 100 bp DNA ladder were used (Thermo Scientific). The DNA fragments were visualised using a UV illuminator after staining for 15 min in ethidium bromide (Roth) water bath.

Plasmid generation

Ligation and transformation

To ligate the insert DNA fragment and the backbone DNA fragment, a 20 µl ligation reaction was set up containing 100 ng of the digested and purified backbone DNA, the corresponding amount of the digested and purified insert DNA (500 ng*insert size [bp] * backbone size [bp^{-1}]), 1 U of T4 DNA Ligase (Thermo Fisher), 1 x T4 DNA Ligase Buffer from a 10 x stock solution (Thermo Fisher) and filled up to 20 µl with milli-Q water. The ligation reaction was performed for 1.5 h at RT. To transform the ligation reaction into competent Top10 *Escherichia coli* K12 cells, 30 µl of gently thawed bacteria were transferred into a 1.5 ml reaction tube and 6 µl of the ligation reaction were added. The transformation reaction was reposed on ice for 30 min and afterwards a heat shock was performed for 1 min at 42 °C. After 1 min incubation on ice, 300 µl of SOC medium prewarmed to 37 °C was added to the transformation reaction and the bacteria were allowed to recover for 1 hour at 37 °C under constant shaking at and 200 rpm. 20 µl of the bacteria were spread on a LB-agar plate supplemented with 50 µg/ml ampicillin (final concentration). The remaining bacteria were centrifuged at 700 x g or 1 min, the supernatant was removed by decanting, the pellet was resuspended in the reflow and spread on a LB-agar plate supplemented with 50 µg/ml ampicillin (final concentration). Both agar plates were incubated over night at 37 °C. Bacterial colonies

were tested via colony PCR (see chapter 2.1.1) or used to set up a liquid culture in LB medium supplemented with 50 µg/ml Ampicillin (final concentration) to isolate plasmids using the NucleoSpin Plasmid Kits according to the instructions of the manufacturer (Macherey&Nagel).

InFusion and transformation

For InFusion reactions using In-Fusion® HD Cloning Plus CE (Takara Clontech), 5 µl InFusion reaction was set up containing 50 ng of the digested and purified backbone DNA, 25 ng of the purified insert DNA, generated by PCR using oligonucleotides with overhangs complementary to the first 15 bp of each end of the linearised backbone vector as primers designed according to the manufacturer's instructions, and 1 µl of InFusion HD Enzyme Premix and filled up with milli-Q water. The reaction was incubated for 15 min at 50 °C and placed on ice. Competent Stellar *E. coli* K12 cells were gently thawed on ice and 2.5 µl of InFusion reaction were added 50 µl of bacteria, which were transferred into a 15 ml reaction before the transformation reaction was reposed on ice for 30 min and afterwards a heat shock was performed for 45 sec at 42 °C. After 2 min incubation on ice, 500 µl of SOC medium were added and the bacteria were allowed to recover for 1 hour at 37 °C and under constant shaking at 200 rpm. 20 µl of the bacteria were spread on a LB-agar plate supplemented with 50 µg/ml Ampicillin (final concentration). The remaining bacteria were centrifuged at 3500 rpm for 1 min, the supernatant was removed by decanting, the pellet was resuspended in the reflow and spread on a LB-agar plate supplemented with 50 µg/ml Ampicillin (final concentration). Both agar plates were incubated over night at 37 °C. Bacterial colonies were tested via colony PCR (see Polymerase chain reaction) or used to set up a liquid culture in LB medium supplemented with 50 µg/ml Ampicillin (final concentration) to isolate plasmids using the NucleoSpin Plasmid Kits according to the instructions of the manufacturer (Macherey&Nagel).

Plasmid validation

The correctness of the generated plasmid constructs was analysed using Sanger sequencing. Therefore, 500 ng of plasmid DNA were mixed with 2.5 µM of specific primer. The volume was adjusted to 10 µl using milli-Q water and sent to the provided sequencing service.

Cell culture-based methods

Trypanosoma brucei culture

T. brucei wild type and genetically modified strains derived from one of the following strains: Lister 427 bloodstream-form Mitat 1.2 (clone 221a) strain, the derivative "single marker" (SM) bloodstream form strain expressing a T7 polymerase and a Tet repressor (Wirtz and Clayton, 1995) or from the derivative '2T1' strain expressing a Tet repressor and containing a puromycin-tagged ribosomal spacer as a landing pad for the transfection construct (Alsford and Horn, 2008). The bloodstream form *T. brucei* strains were cultivated in HMI-11 medium (HMI-9 medium (Hirumi and Hirumi, 1994) without serum-plus) at

37 °C and 5% CO_2. If required, the following drug concentrations (final) were used: 2 µg/ml G418, 5 µg/ml hygromycin, 0.1 µg/ml puromycin, 5 µg/ml blasticidin, 2.5 µg/ml phleomycin and 1 µg/ml doxycycline. Growth rates were monitored over 96h and cell densities were determined every 24 h. Transfections were performed using a Nucleofector (Amaxa), as described previously (Scahill et al., 2008).

Isolation of genomic DNA

For integration PCRs (see Polymerase chain reaction), gDNA from *T. brucei* BF cells were obtained using the Phusion Human Specimen Direct PCR Kit (Thermo Scientific). To do so, 1 x 10^6 cells grown to the mid log-phase ~ 1.0 – 1-5 x 10^6 were harvested by centrifugation for 10 min at 1,500 x g and RT. The supernatant was completely removed without disturbing the pellet and the cell pellet was resuspended in 20 µl of dilution buffer and 0.5 µl DNA release additive. The reaction was mixed, incubated for 5 min at RT and boiled for 2 min at 98 °C. Following centrifugation for 5 min at 2,000 x g and RT, the supernatant was transferred into a new 1.5 ml reaction tube and 1 µl of the supernatant was used as template for integration PCR.

To obtain pure gDNA from *T. brucei* BF cells, ~ 3 x 10^6 cells grown to the mid log-phase ~ 1.0 – 1.5 x 10^6 were harvested for 10 min at 1,500 x g and RT. The supernatant was removed by careful decanting and the cell pellet was resuspended in the residual liquid. The cell suspension was transferred into a 1.5 ml reaction tube and centrifuged for 4 min at 2,500 x g and RT. The supernatant was completely removed without disturbing the pellet and the cell pellet was resuspended in 500 µl DNAzol® (Invitrogen). Cell lysis was performed for 5 min at RT. The gDNA was precipitated by adding 250 µl 100% ethanol and following centrifugation for 5 min at 16,000 x g and 4 °C, the supernatant was completely removed without disturbing the pellet. The gDNA pellet was washed using 500 µl 70% EtOH and centrifuged 5 min at 16,000 x g and 4 °C. The supernatant was completely removed without disturbing the pellet. The gDNA pellet was immediately dissolved in 1 x TE-buffer (10 mM Tris-HCl, 1 mM EDTA pH 8). Nucleotide concentrations were determined using a NanoDrop spectrometer (Thermo Scientific). Isolated gDNA was stored at 4 °C.

Preparation of cell lysates for protein detection

To prepare cell lysates for protein detection using western blot analyses, 2 x 10^6 *T. brucei* cells grown to the mid log-phase ~ 1.0 – 1.5 x 10^6 were harvested for 10 min at 1,500 x g and RT. The supernatant was removed by careful decanting, the cell pellet was resuspended in the reflow and the cell suspension was transferred into a new 1.5 ml reaction tube. Following centrifugation for 5 min at 2,500 x g and RT, the supernatant was completely removed without disturbing the cell pellet. The pellet was resuspended in 1 ml 1x TDB buffer (5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$*$7H_2O$, 20 mM Na_2HPO_4, 2 mM NaH_2PO4*$2H_2O$, 20 mM glucose) and centrifuged for 5 min at 2,500 x g and RT. The supernatant was completely removed without disturbing the cell pellet. The cell pellet was resuspended in 10 µl cell lysis

buffer (1 ml 4x Laemmli sample buffer, 3 ml RIPA buffer, 8 µl 1 M DTT; Laemmli buffer: 40% glycerol, 0.02% bromphenol blue, 8% SDS, 250 mM Tris pH 6.5; RIPA buffer: 50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1% NP-40, 0.25% Na-Deoxycholate, 0.1% SDS) and incubated for 10 min at 100 °C to denature the proteins. The cell lysates were stored at -80 °C.

Generation of *T. brucei* BF mutants

Knock down of HAT1 and HAT2

Generation of constructs for inducible knock down of HAT1 and HAT2

The constructs for inducible knock down of HAT1 and HAT2 originated from pRPa$_{iSL}$ (Alsford and Horn, 2008) designed for efficient and stable RNA interference (RNAi) by stem-loop RNAs. Oligos for amplifying the targeting regions of HAT1 and HAT2 were ordered, as published previously (Kawahara et al., 2008). The forward oligos were extended with the sequences of the restriction enzymes ApaI and KpnI at their 5´ends and the reverse oligos were extended with the sequences of XbaI and BamHI at their 3´ends. The HAT1-targeting region was amplified with H1R5SL-GATCGGGCCCGGTACCCCCTTTTCGGAACATGAAGA- and H1R3SL-GATCTCTAGAGGATCCAGTGTGACGCAGTGACTT- from gDNA. The generated amplicon was digested using KpnI/BamHI to generate the sense targeting region and using XbaI/ApaI to generate the antisense targeting region. To generate, the sense targeting region was inserted into KpnI/BamHI-digested pRPa$_{iSL}$ and the antisense targeting region was inserted into the generated pRPa$_{iSL}$_HAT1_sense digested with XbaI/ApaI. The construct pRPa$_{iSL}$_HAT2_KD was generated as described for pRPa$_{iSL}$_HAT1_KD but using the HAT2-targeting region (amplified with H2R5SL- GATCGGGCCCGGTACCACTCACAGACTTGGGAGCA - and H2R3SL- GATCTCTAGAGGATCCTCATGTCATGTGCCCACTTT - from gDNA). For transfection, 20 µg of pRPa$_{iSL}$_HAT1_KD anord pRPa$_{iSL}$_HAT2_KD were linearised with AscI.

Generation of HAT1 and HAT2 knock down cell lines

The cell lines allowing inducible knock down of HAT1 and HAT2 derived from the 2T1 cell line, in which the knock down construct was integrated into a *PUR*-tagged *rRNA* locus for stable transfection efficiency (Alsford et al., 2005; Alsford and Horn, 2008). To generate 2T1_HAT1_KD, the AscI-linearised pRPa$_{iSL}$_HAT1_KD was transfected into 2T1 cells using a Nucleofector (Amaxa), as described previously (Scahill et al., 2008). The 2T1_HAT2_KD cell line was generated using AscI-linearised pRPa$_{iSL}$_HAT2_KD. Hygromycin and phleomycin were added to the culture medium to select for successful transfectants. Successful integration of the knock down constructs was analysed by PCR (CGCGTCTGCTGCTCCATACAAG and TCAAGCTCTAAATCGGGGGC) on gDNA purified from 2T1_HAT1_KD and 2T1_HAT2_KD cells, respectively. To phenotypically analyse the effect of HAT1 and HAT2 knock down, respectively, growth rates of induced 2T1_HAT1_KD and 2T1_HAT2_KD cells were monitored for 96h. Uninduced 2T1_HAT1_KD and 2T1_HAT2_KD cells served as control. To evaluate

the knock down efficiencies, the transcript levels of HAT1 (CAGAACGCGGAAGGAGAGTT and GCGAGAAGTACCACGTCTCC) and HAT2 (CCCACCACAGTGCGATACTT and TGCGTCGCTTAAAGGTGCTA) were analysed at different time points of RNAi-induction by quantitative real-time PCR (see Quantitative real-time PCR for testing RNAi efficiencies).

Knock down of different BDFs and chromatin modifiers

Generation of constructs for inducible knock down of different BDFs and chromatin modifiers

The constructs for inducible knock down of different chromatin modifiers (**Table 2.3**) originated from pRPa$_{iSL}$ (Alsford and Horn, 2008) designed for efficient and stable RNA interference by stem-loop RNAs. The RNAi-targeting regions of the chromatin modifiers were amplified from gDNA using a gene-specific forward oligo containing an ApaI/KpnI-5´-extension and reverse oligo containing a XbaI/BamHI-5´-extension (**Table 2.1**). The amplified targeting regions were digested using KpnI/BamHI to generate the sense targeting region and using XbaI/ApaI to generate the antisense targeting region. To generate pRPa$_{iSL}$"target_gene"_KD (**Table 2.3**), the sense targeting region was inserted into pRPa$_{iSL}$, digested with KpnI/BamHI and the antisense targeting region was inserted into the generated pRPa$_{iSL}$__"target_gene"_sense, which was digested with XbaI/ApaI. For transfection, 20 µg of pRPa$_{iSL}$_"target_gene"_KD were linearised with AscI.

Generation of knock down cell lines for inducible depletion of different BDFs and chromatin modifiers

The cell lines allowing inducible knock down of different chromatin modifiers derived from the 2T1 cell line, in which the knock down construct was integrated into a *PUR*-tagged *rRNA* locus for stable transfection efficiency (Alsford et al., 2005; Alsford and Horn, 2008). To generate 2T1_"target_gene"_KD (**Table 2.5**)., the AscI-linearised pRPa$_{iSL}$_"target_gene"_KD was transfected into 2T1 cells using a Nucleofector (Amaxa), as described previously (Scahill et al., 2008). Hygromycin and phleomycin were added to the culture medium to select for successful transfectants. To phenotypically analyse the effect of the "target_gene" knock down, growth rates of induced 2T1_"target_gene" _KD cells were monitored for 96h. Uninduced 2T1_"target_gene"_KD cells served as control. To evaluate the knock down efficiencies, the transcript levels of "target_gene" were analysed at different time points of RNAi-induction by quantitative real-time PCR, as described in Quantitative real-time PCR for testing RNAi efficiencies (for gene specific oligos see **Table 2.1**).

Overexpression of different chromatin modifiers

Generation of constructs for overexpression of different chromatin modifiers

The constructs for inducible overexpression of different chromatin modifiers (**Table 2.3**) originated from pRPaTag (Alsford and Horn, 2008). The coding region of the chromatin modifiers was amplified from gDNA using a gene-specific forward oligo containing a HindIII-5´-extension and reverse oligo containing a BamHI-5´-extension (**Table 2.1**). The amplified targeting regions were digested using HindIII/BamHI

and integrated into HindIII/BamHI-digested pRPaTag to generate pRPa_"target_gene"_OE (**Table 2.3**). For transfection, 20 µg of pRPa_"target_gene"_OE were linearised with AscI.

Generation of cell lines for inducible overexpression of different chromatin modifiers

The cell lines allowing inducible overexpression of different chromatin modifiers derived from the 2T1 cell line, in which the overexpression construct was integrated into a *PUR*-tagged *rRNA* locus for stable transfection efficiency (Alsford et al., 2005; Alsford and Horn, 2008). To generate 2T1_"target_gene"_OE (**Table 2.3**), the AscI-linearised pRPa_"target_gene"_OE was transfected into 2T1 cells using a Nucleofector (Amaxa), as described previously (Scahill et al., 2008). Hygromycin and phleomycin were added to the culture medium to select for successful transfectants. To phenotypically analyse the effect of the "target_gene" overexpression, growth rates of induced 2T1_"target_gene" _OE cells were monitored for 96h. Uninduced 2T1_"target_gene"_OE cells served as control. To evaluate the overexpression efficiencies, the transcript levels of "target_gene" were analysed at different time points of overexpression by quantitative real-time PCR, as described in Quantitative real-time PCR for testing RNAi efficiencies (for gene specific oligos see **Table 2.1**).

TY-tagging of RNAPII subunit *RPB9*

Generation of constructs for endogenous TY-tagging of both *RPB9* alleles

The tagging construct pPOTv3-BSD-2xTY originated from pPOTv3 (a precursor from the pPOT-constructs (Dean et al., 2015)). The YFP gene was exchanged by two succeeding, codon-optimised TY-tags using the restriction sites HindIII and SacI. The first one contained the start codon. The codon-optimised TY-tag-fragment was generated by annealing of two oligonucleotides (AGCTTATGGAGGTGCACACGAACCAGGACCCGCTGGACGCCGAGGTGCACACGAACCAGGACCCGCTGGAC GGATCCGGATCAGGATCTGGATCAGGATCGGGTAGTGAGCT and CACTACCCGATCCTGATCCAGATCCTGATCCGGATCCGTCAAGTGGATCTTGGTTAGTATGGACCTCGGCGTC AAGTGGGTCCTGGTTAGTATGGACCTCCATA).

To generate the constructs allowing endogenous TY-tagging of both *RPB9* alleles, the InFusion HD Cloning Plus reagents were used (Takara) according to the manufacturer's instructions. To generate the construct pPOTv3-BSD-2xTY_RPB9 for N-terminal 2xTY-tagging of *RPB9*, the upstream region of *RPB9* (amplified with CTATAGGGCGAATTGGGCCCTGTGTCAACCAACGGCGATA and GCATTATACGCGGCCGCGGTGGTTTTATTCCACCTTAGCTTCG from gDNA) was integrated into the ApaI/NotI-digested pPOTv3-BSD-2xTY. The downstream region, homologous to the 5´end of RPB9 CDS excluding the ATG (amplified with CAGGATCGGGTAGTGAGCTCGAGTCCACTTTGACGCAG and GCTTTTTCATGCTAGCGTTCACGAAGCATGTGACTTC) was inserted using SacI/NheI-digested pPOTv3-BSD-2xTY_RPB9_UR. To generate the construct pPOTv3-NEO-2xTY_RPB9 for N-terminal 2xTY-tagging of *RPB9*, the resistance cassette *BSD* was exchanged by a *NEO* cassette (amplified with

GCAACGAAGCTTGTTAACATGATTGAACAAGATGGATTGCAC and TGGGCAGGATCCGGATCAGAAGAACTCGTCAAGAAGGC from pyrFE-NEO (Kim et al., 2013)) using the restriction sites MluI/NotI. For transfection, the tagging constructs pPOTv3-BSD-2xTY_RPB9 and pPOTv3-NEO-2xTY_RPB9 were linearised using ApaI/NheI.

Generation of $^{2xTY/2xTY}$*RBP9* cell lines

The cell lines containing $^{2xTY/2xTY}$*RBP9* derived from the 2T1 cell line (Alsford et al., 2005; Alsford and Horn, 2008). For tagging of the first allele, ApaI/NheI-linearised pPOTv3-BSD-2xTY_RPB9 was transfected into 2T1 cells using a Nucleofector (Amaxa), as described previously (Scahill et al., 2008). Integration of the tagging construct was verified by PCR (TGCTAGCGCCTTATCACCAC and CACAAGGTCCCCCAGTAAAA; GCATCTTCACTGGTGTCAATGTA and CAAAGGTAGGCGGTGGGG) on gDNA purified from 2T1_$^{2xTY/x}$RBP9 cells. For tagging of the second allele, ApaI/NheI-linearised pPOTv3-NEO-2xTY_RPB9 was transfected into 2T1_$^{2xTY/x}$RBP9 cells using a Nucleofector (Amaxa). Integration of the tagging construct was verified by PCR (TGCTAGCGCCTTATCACCAC and AATATCACGGGTAGCCAACG; TGAATGAACTGCAGGACGAG and CAAAGGTAGGCGGTGGGG) on gDNA purified from 2T1_$^{2xTY/2xTY}$RBP9 cells. Successful Ty-tagging of both *RPB9* alleles was further evaluated by western blot analysis. For cultivation, the medium of the 2T1_$^{2xTY/2xTY}$RBP9 was supplemented with appropriate concentrations of puromycin, phleomycin, blasticidin and neomycin.

Generation of $^{2xTY/2xTY}$RBP9 cell lines allowing inducible depletion of HAT1 and HAT2

2T1_$^{2xTY/2xTY}$RBP9_HAT1_KD and 2T1_$^{2xTY/2xTY}$RBP9_HAT2_KD cells lines derived from the 2T1_$^{2xTY/2xTY}$RBP9. For transfection, the AscI-linearised pRPa$_{iSL}$_HAT1_KD and pRPa$_{iSL}$_HAT2_KD were used, respectively. 2T1_$^{2xTY/x}$RBP9_HAT1_KD and 2T1_$^{2xTY/x}$RBP9_HAT2_KD cells were cultivated adding hygromycin, phleomycin, blasticidin and neomycin to the medium. Successful integration of the knock down constructs was analysed by PCR (CGCGTCTGCTGCTCCATACAAG and TCAAGCTCTAAATCGGGGGC) on gDNA purified from 2T1_$^{2xTY/2xTY}$RBP9_HAT1_KD and 2T1_$^{2xTY/2xTY}$RBP9_HAT2_KD cells, respectively. To phenotypically analyse the effect of HAT1 and HAT2 knock down, respectively, growth rates of induced 2T1_$^{2xTY/2xTY}$RBP9_HAT1_KD and 2T1_$^{2xTY/2xTY}$RBP9_HAT2_KD cells were monitored for 96h. Uninduced 2T1_$^{2xTY/2xTY}$RBP9_HAT1_KD and 2T1_$^{2xTY/2xTY}$RBP9_HAT2_KD cells served as control. To evaluate the knock down efficiencies, the transcript levels of HAT1 (CAGAACGCGGAAGGAGAGTT and GCGAGAAGTACCACGTCTCC) and HAT2 (CCCACCACAGTGCGATACTT and TGCGTCGCTTAAAGGTGCTA) were analysed at different time points of RNAi-induction by quantitative real-time PCR (see Quantitative real-time PCR for testing RNAi efficiencies).

Quantitative real-time PCR for testing RNAi efficiencies

RNAi-mediated knock down of the target transcripts was induced over 0 h and 48 h using 1 µg/ml doxycycline in transgenic *T. brucei* cells. RNA was extracted from 5 x 10^7 cells from each time point

using the Macherey-Nagel, NucleoSpin® RNA kit according to the manufacturer's instruction and copyDNA (cDNA) was synthesised using the M-MLV reverse transcriptase system (Thermo Fisher). Quantitative real-time PCR on the cDNAs was performed using the iTaq system (Bio-Rad) according to the manufacturer's instructions. Therefore, gene-specific primers were chosen, that bind in the coding sequence of the target gene, but not in the RNAi-target sequence (**Table 2.1**). The $\Delta\Delta C_t$-method Methods (or-Livak-method) was applied that compares the expression of the target gene in the 48 h time point to the 0 h time point.

Lists of materials used in this study

Table 2.1. List of oligos used in this study.
Underlined sequences indicate InFusion overhangs; abbreviations: KD – Knock down, OE – Overexpression, Lig – Ligation, Inf – Infusion, RTR - RNAi targeting region, CDS – Coding sequence, qPCR – Quantitative real-time polymerase chain reaction, Neo – Neomycin/G418, ALD – Aldolase, UTR – Untranslated region.

Oligo name	Lab internal name	Purpose	DNA sequence 5´-3´
oAK_1	1270_BDF1_pRPaiSL_F	BDF1 RTR	GATCGGGCCCGGTACCAAGCTACCGGACTACCGGAT
oAK_2	1271_BDF1_pRPaiSL_R	BDF1 RTR	GATCTTCTAGAAGGATCCATTCCATTATGAGGCGGC
oAK_3	1272_BDF1_qPCR_F	BDF1 qPCR	CGCGCTGGCTTTTCTTCTTT
oAK_4	1273_BDF1_qPCR_R	BDF1 qPCR	TTTTCCACCCTCGCCTCTTC
oAK_5	842_BDF2_KD_pRPa_F	BDF2 RTR	GATCGGGCCCGGTACCTATCCCGTCTGTGGGATCTT
oAK_6	843_BDF2_KD_pRPa_R	BDF2 RTR	GATCTCTAGAGGATCCAGGTCCTCCTTGTTCCGTTT
oAK_7	814_BDF2_OE_infF	BDF2 CDS	TAAAATTCACAAGCTTGCAATGAGCAAGAACGAGCG
oAK_8	815_BDF2_OE_infR	BDF2 CDS	TAAATGGGCAGGATCCGTCGATACCAGTAGGCCGTG
oAK_9	939_qBDF2_F	BDF2 qPCR	GCGGAGAAAGATGTGCCTCT
oAK_10	940_qBDF2_R	BDF2 qPCR	TCGCTGCTGTCTACATCGTC
oAK_11	918_BDF3_KDlig_F	BDF3 RTR	GATCGGGCCCGGTACCAACCTGTCGATGCACTGAGG
oAK_12	919_BDF3_KDlig_F	BDF3 RTR	GATCTCTAGAGGATCCCCGCTGCTAAGGATGAGTGT
oAK_13	816_BDF3_OE_infF	BDF3 CDS	TAAAATTCACAAGCTTATGGTCCTCACAGGGGTCAA
oAK_14	817_BDF3_OE_infR	BDF3 CDS	TAAATGGGCAGGATCCTCACTCCAACTGCCCGTTTG
oAK_15	937_qBDF3_F	BDF3 qPCR	CACTCAGCCATCCACTACCG
oAK_16	938_qBDF3_R	BDF3 qPCR	GGCGCTCTCCTCTGCATAAA
oAK_17	1274_BDF4_pRPaiSL_F	BDF4 RTR	GATCGGGCCCGGTACCCGACATTGTGTGGACAGGAC
oAK_18	1275_BDF4_pRPaiSL_R	BDF4 RTR	GATCTCTAGAGGATCCGGACGGACGTGTTTCAGTTT
oAK_19	1276_BDF4_qPCR_F	BDF4 qPCR	TACCTTTTGCACGGCGTTTG
oAK_20	1277_BDF4_qPCR_R	BDF4 qPCR	CGCCACTTACACATCAGGGT
oAK_21	806_FACT_KD_ligF	FACT RTR	GATCGGGCCCGGTACCCGGTTGAAGTTGGACCCCTT
oAK_22	807_FACT_KD_ligR	FACT RTR	GATCTCTAGAGGATCCGCTTGGGGCATTTCGTTCC
oAK_23	804_FACT_OE_infF	FACT CDS	TAAAATTCACAAGCTTTTACCCTTTCCGCATCTCGC
oAK_24	805_FACT_OE_infR	FACT CDS	TAAATGGGCAGGATCCTTTCGTCCCCCTCGTTCTTG
oAK_25	808_qFACT_F	FACT qPCR	TGCAGCGCGACTTTACTACA
oAK_26	809_qFACT_R	FACT qPCR	ATTGAGTACTGCCCAACCGG
oAK_27	1751_H1R5SL	HAT1 RTR	GATCGGGCCCGGTACCCCCTTTTCGGAACATGAAGA
oAK_28	1752_H1R3SL	HAT1 RTR	GATCTCTAGAGGATCCAGTGTGACGCAGTGACTT
oAK_29	812_HAT1_OE_infF	HAT1 CDS	TAAAATTCACAAGCTTATGGTAATGTTTGAGGTGCGG
oAK_30	813_HAT1_OE_infR	HAT1 CDS	TAAATGGGCAGGATCCTTATACTTTCTCAGCTCTACGCA
oAK_31	941_qHAT1_F	HAT1 qPCR	CAGAACGCGGAAGGAGAGTT
oAK_32	942_qHAT1_R	HAT1 qPCR	GCGAGAAGTACCACGTCTCC

Oligo name	Lab internal name	Purpose	DNA sequence 5´-3´
oAK_33	1753_H2R5SL	HAT2 RTR	GATCGGGCCCGGTACCACTCACAGACTTGGGAGCA
oAK_34	1754_H2R3SL	HAT2 RTR	GATCTCTAGAGGATCCTCATGTCATGTGCCCACTTT
oAK_35	810_HAT2_OE_infF	HAT2 CDS	TAAAATTCACAAGCTTCCACCGGCAACATAGCTGTA
oAK_36	811_HAT2_OE_infR	HAT2 CDS	TAAATGGGCAGGATCCGAGGGACGACACATCAGCTC
oAK_37	475_qHAT2_F	HAT2 qPCR	CCCACCACAGTGCGATACTT
oAK_38	476_qHAT2_R	HAT2 qPCR	TGCGTCGCTTAAAGGTGCTA
oAK_39	800_ISWI_KD_ligF	ISWI RTR	GATCGGGCCCGGTACCTGACAGCTGGTTTGACACGG
oAK_40	801_ISWI_KD_ligR	ISWI RTR	GATCTCTAGAGGATCCTTTGTGCCCCTCCTTTTCGT
oAK_41	798_ISWI_OE_infF	ISWI CDS	TAAAATTCACAAGCTTGGGTTGCCTTACTCCCACTT
oAK_42	799_ISWI_OE_infR	ISWI CDS	TAAATGGGCAGGATCCCTCACACACGTCCCCTTTGA
oAK_43	802_qISWI_F	ISWI qPCR	GCTCGATGCGTGGCTTTAAG
oAK_44	803_qISWI_R	ISWI qPCR	TTGGGTGCGTTAAAAGCAGC
oAK_45	1860_pPOT_neoF	Neo CDS	GCAACGAAGCTTGTTAACATGATTGAACAAGATGGATTGCAC
oAK_46	1861_pPOT_neoR	Neo CDS	TGGGCAGGATCCGGATCAGAAGAACTCGTCAAGAAGGC
oAK_47	1862_pPOT_3ALDF	3´ ALD UTR	TCCGGATCCTGCCCATTTAG
oAK_48	1863_pPOT_3ALDR	3´ ALD UTR	AGCTTGAGCAACGCGTGG
oAK_49	1864_pPOT_5ALDF	5´ ALD UTR	AAAACCACCGCGGCCGCCGCGTTTCCTTACATATTTCT
oAK_50	1865_pPOT_5ALDR	5´ ALD UTR	AACAAGCTTCGTTGCAGTTGA
oAK_51	1303_Ty1_tagCO_S	2 x TY1-tag CDS	AGCTTATGGAGGTGCACACGAACCAGGACCCGCTGGACGCCGAGGT GCACACGAACCAGGACCCGCTGGACGGATCCGGATCAGGATCTGGA TCAGGATCGGGTAGTGAGCT
oAK_52	1304_Ty1_tagCO_AS	2 x TY1-tag CDS	CACTACCCGATCCTGATCCAGATCCTGATCCGGATCCGTCCAGCGGG TCCTGGTTCGTGTGCACCTCGGCGTCCAGCGGGTCCTGGTTCGTGTG CACCTCCATA

Table 2.2. List of parental plasmids used in this study.

Abbreviations: Amp – ampicillin, KD – Knock down, OE – Overexpression, Hygro – Hygromycin, Blas – Blasticidin, Neo – Neomycin/G418, UTR – Untranslated region.

Name	Purpose	Resistance in *E. coli*	Resistance in *T. brucei*	Reference
pPOTv3	Retrieve backbone for pPOTv3-BSD-2xTY	Amp	Blas	(Dean et al., 2015)
pPOTv3_TY-RPB9_BSD	Retrieve backbone for pPOTv3-TY-RPB9_Neo	Amp	Blas	(Wedel and Siegel, 2017)
pyrFEKO-NEO	Amplification of *NEO* resistance cassette	Amp	Neo	(Scahill et al., 2008)
pPOTv3_TY-RPB9_BSD	Amplification of 3´UTR of Aldolase	Amp	Blas	(Wedel and Siegel, 2017)
pPOTv3_TY-RPB9_BSD	Amplification of 5´UTR of Aldolase	Amp	Blas	(Wedel and Siegel, 2017)
$pRPa_{iSL}$	Retrieve backbone for pRPaiSL_"target_gene"_KD	Amp	Hygro	(Alsford and Horn, 2008)
$pRPa_{tag}$	Retrieve backbone for pRPa_"target_gene"_OE	Amp	Hygro	(Alsford and Horn, 2008)

Table 2.3. List of plasmids generated in this study.

Abbreviations: Amp – ampicillin, KD – Knock down, OE – Overexpression, Hygro – Hygromycin, Blas – Blasticidin, Neo – Neomycin/G418, Puro – Puromycin, Δ – deletion, UTR – Untranslated region.

Construct name	Purpose	Resistance in *E. coli*	Resistance in *T. brucei*	Reference
pRPaiSL_BDF1_KD	Inserts RNAi-targeting construct for inducible knock down of BDF1 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPaiSL_BDF2_KD	Inserts RNAi-targeting construct for inducible knock down of BDF2 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPa_BDF2_OE	Inserts construct for inducible overexpression of BDF2 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPaiSL_BDF3_KD	Inserts RNAi-targeting construct for inducible knock down of BDF3 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPa_BDF3_OE	Inserts construct for inducible overexpression of BDF3 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPaiSL_BDF4_KD	Inserts RNAi-targeting construct for inducible knock down of BDF4 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPaiSL_FACT_KD	Inserts RNAi-targeting construct for inducible knock down of FACT into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPa_FACT_OE	Inserts construct for inducible overexpression of FACT into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPaiSL_HAT1_KD	Inserts RNAi-targeting construct for inducible knock down of HAT1 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	(Kawahara et al., 2008; Kraus et al., 2020)
pRPa_HAT1_OE	Inserts construct for inducible overexpression of HAT1 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPaiSL_HAT2_KD	Inserts RNAi-targeting construct for inducible knock down of HAT2 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	(Kawahara et al., 2008; Kraus et al., 2020)
pRPa_HAT2_OE	Inserts construct for inducible overexpression of HAT2 into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPaiSL_ISWI_KD	Inserts RNAi-targeting construct for inducible knock down of ISWI into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pRPa_ISWI_OE	Inserts construct for inducible overexpression of ISWI into a *hygroΔstart::Puro* tagged *rRNA* locus on chromosome 2	Amp	Hygro	AJ Kraus
pPOTv3-BSD-2xTY	Endogenous N-terminal tagging of target gene with 2 x Ty1 epitope tag	Amp	Blas	AJ Kraus
pPOTv3_TY-RPB9_Neo	Endogenous N-terminal tagging of RPB9 with 2 x Ty1 epitope tag	Amp	Neo	(Kraus et al., 2020)

Table 2.4. List of cell lines used in this study.

Abbreviations: BF – bloodstream form, MITat1.2 – Molteno Institute Trypanozoon Antigen Type 1.2, SM – single marker; TETR – tetracycline repressor, T7RNAP – T7 RNA polymerase, Δ – deletion, NEO – aminoglycoside phosphotransferase, BSD – blasticidin S deaminase, PAC – puromycin N-acetyltransferase, HYG – hygromycin phosphotransferase, BLE – phleomycin resistance gene, Neo/G418 – Neomycin, Blas – Blasticidin, Puro – Puromycin, Hygro – Hygromycin, Phleo – phleomycin, Dox - Doxycyclin.

Name	Genotype	Constructs	Selection	Reference
Wt	BF Lister 427, MITat1.2, clone 221a	-	-	(Cross, 1975)
SM	BF Lister 427, MITat1.2, clone 221a, *TETR T7RNAP NEO*	pHD328, pLew114hyg5´	Neo	(Wirtz and Clayton, 1995)

2T1	BF Lister 427, MITat1.2, clone 221a, TetR::BLE::TUB, *hygΔstart::PAC::RRNA* locus on chromosome 2	pHD1313, ph3E	Puro, Phleo	(Alsford et al., 2005; Alsford and Horn, 2008)
BFJEL43	SM, RRNA::*Ty1-H2B.V BSD*, *ΔH2B.V::PAC* | *ΔH2B.V::HYG*	pJEL92, pJEL74, pJEL75	Neo, Blas, Puro, Hygro, Dox	(Lowell et al., 2005)
BFpFK8	SM, RRNA::*Ty1-H2A BLE*	pFK8	Neo, Phleo, Dox	(Lowell et al., 2005)
BFpFK3.4.5	SM; RRNA::*Ty1-H2A.Z BLE*, *ΔH2A.Z::PAC* | *ΔH2A.Z::HYG*	pFK3, pFK4, pFK5	Neo, Phleo, Puro, Hygro, Dox	(Lowell et al., 2005)

Table 2.5. List of transgenic cell lines generated in this study.

Abbreviations: KD – Knock down, OE - Overexpression NEO – aminoglycoside phosphotransferase, BSD – blasticidin S deaminase, PAC – puromycin N-acetyltransferase, HYG – hygromycin phosphotransferase, BLE – phleomycin resistance gene, Neo/G418 – Neomycin, Blas – Blasticidin, Puro – Puromycin, Hygro – Hygromycin, Phleo – phleomycin.

Name	Genotype	Constructs	Selection	Reference
2T1_BDF1_KD	2T1, *HYG::RRNA::BDF1_KD* on chromosome 2	pRPaiSL_BDF1_KD	Hygro, Phleo	AJ Kraus
2T1_BDF2_KD	2T1, *HYG::RRNA::BDF2_KD* on chromosome 2	pRPaiSL_BDF2_KD	Hygro, Phleo	AJ Kraus
2T1_BDF2_OE	2T1, *HYG::RRNA::BDF2_OE* on chromosome 2	pRPa_BDF2_OE	Hygro, Phleo	AJ Kraus
2T1_BDF3_KD	2T1, *HYG::RRNA::BDF3_KD* on chromosome 2	pRPaiSL_BDF3_KD	Hygro, Phleo	AJ Kraus
2T1_BDF3_OE	2T1, *HYG::RRNA::BDF3_OE* on chromosome 2	pRPa_BDF3_OE	Hygro, Phleo	AJ Kraus
2T1_BDF4_KD	2T1, *HYG::RRNA::BDF4_KD* on chromosome 2	pRPaiSL_BDF4_KD	Hygro, Phleo	AJ Kraus
2T1_FACT_KD	2T1, *HYG::RRNA::FACT_KD* on chromosome 2	pRPaiSL_FACT_KD	Hygro, Phleo	AJ Kraus
2T1_FACT_OE	2T1, *HYG::RRNA::FACT_OE* on chromosome 2	pRPa_FACT_OE	Hygro, Phleo	AJ Kraus
2T1_HAT1_KD	2T1, *HYG::RRNA::HAT1_KD* on chromosome 2	pRPaiSL_HAT1_KD	Hygro, Phleo	(Kawahara et al., 2008; Kraus et al., 2020)
2T1_HAT1_OE	2T1, *HYG::RRNA::HAT1_OE* on chromosome 2	pRPa_HAT1_OE	Hygro, Phleo	AJ Kraus
2T1_HAT2_KD	2T1, *HYG::RRNA::HAT2_KD* on chromosome 2	pRPaiSL_HAT2_KD	Hygro, Phleo	(Kawahara et al., 2008; Kraus et al., 2020)
2T1_HAT2_OE	2T1, *HYG::RRNA::HAT2_OE* on chromosome 2	pRPa_HAT2_OE	Hygro, Phleo	AJ Kraus
2T1_ISWI_KD	2T1, *HYG::RRNA::ISWI_KD* on chromosome 2	pRPaiSL_ISWI_KD	Hygro, Phleo	AJ Kraus
2T1_ISWI_OE	2T1, *HYG::RRNA:: ISWI_OE* on chromosome 2	pRPa_ISWI_OE	Hygro, Phleo	AJ Kraus
2T1_$^{2xTY/2xTY}$RBP9	2T1, *NEO::2xTy1::RPB9* *BSD::2xTy1::RPB9*	pPOTv3_TY-RPB9_Neo, pPOTv3_TY-RPB9_BSD	Puro, Phleo, Neo, Blas	(Kraus et al., 2020)
2T1_$^{2xTY/2xTY}$RBP9_HAT1KD	2T1_$^{2xTY/2xTY}$RBP9, *HYG::RRNA::HAT1_KD* on chromosome 2	pRPaiSL_HAT1_KD	Hygro, Phleo, Neo, Blas	(Kraus et al., 2020)
2T1_$^{2xTY/2xTY}$RBP9_HAT2KD	2T1_$^{2xTY/2xTY}$RBP9, *HYG::RRNA::HAT2_KD* on chromosome 2	pRPaiSL_HAT2_KD	Hygro, Phleo, Neo, Blas	(Kraus et al., 2020)

Quantification of site-specific acetylation levels

Mononucleosome preparation for mass-spectrometry analyses

Mononucleosomes were purified using established protocols (Lowell et al., 2005). To do so, cells were harvested at 4 °C and 1,800 x g for 10 min and washed using 40 ml of 1x TDB buffer (5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$*$7H_2O$, 20 mM Na_2HPO_4, 2 mM NaH_2PO4*$2H_2O$, 20 mM glucose). The cell pellet was resuspended in 1 ml of permeabilisation buffer (100 mM KCl, 10 mM Tris 8.0, 25 mM EDTA 8.0, 1 mM DTT, 1.46 µM pepstatin A, 4.7 µM leupeptin, 1 mM PMSF and 1 mM TLCK, 50 mM sodium butyrate, 0.5 µM anarcadic acid) and centrifuged at 4 °C and 1,800 x g for 10 min. The cell pellet was resuspended using 1 ml permeablisation buffer (see above). Digitonin was added to the cell suspension to a final concentration of 40 µM (stock solution: Digitonin was dissolved to 4 mM in permeabilisation buffer) and cell lysis was performed at RT for 15 min under gentle rotation. Following centrifugation, the pellet was washed three times in 1 ml ice-cold isotonic buffer (100 mM KCl, 10 mM Tris 8.0, 10 mM CaCl2, 5% Glycerol, 1 mM DTT, 1.46 µM pepstatin A, 4.7 µM leupeptin, 1 mM PMSF and 1 mM TLCK, 50 mM sodium butyrate, 0.5 µM anarcadic acid) for removing residual traces of EDTA. The cell pellet was resuspended in 1 ml isotonic buffer (see above) and mononucleosomes were generated by micrococcal nuclease digestion (MNase; Sigma) using 2.5 U of MNase and incubation for 15 min at 25 °C in a heating block. The reaction was stopped by adding 10 mM EDTA (pH 8.0) (final concentration). To the mononucleosomes-containing solution, 200 mM NaCl (final concentration), 0.05% NP-40 (final concentration) were added. The reaction was incubated for 5 min on ice. Afterwards, the suspension was mixed by vortexing (five times, 10 sec on, 30 sec break on ice). Following incubation on ice for 5 min without mixing, the mononucleosomes-containing suspension was cleared from other proteins by centrifugation at 4 °C and 11,000 x g for 10 min and the supernatant was transferred to a new 1.5 ml reaction tube.

Purification of histones from mononucleosomes using acid extraction

To identify all histone PTMs, irrespectively of their genomic location, histones were isolated by acid extraction from 1 x 10^9 cells per replicate. In total, three replicates were prepared per approach.

For the acid extraction of histones, 220 µl of 1 M H_2SO_4 were added to 1 ml mononucleosome-containing supernatant (see above) extracted from 2 x 10^6 bloodstream form trypanosomes according to a previously published protocol for acid extraction (Shechter et al., 2007). Following an overnight incubation at 4 °C on a rotation wheel, the acid-insoluble proteins were removed by centrifugation at 4 °C and 16,000 x g for 10 min. The supernatant containing the free, acid-soluble histones was transferred into a new 1.5 ml reaction tube and the histones were concentrated using the StrataClean® resin (Agilent Technologies) as described previously (Günzl and Schimanski, 2009). Therefore, 20 µl StrataClean® resin was added, and the suspension was incubated for 20 min at RT on a rotation wheel to allow binding

of the histones to the resin. The resin was collected by centrifugation at RT and 16,000 x g for 1 min. To release the histones, the resin was incubated in 60 µl 1 x NuPAGE® LDS Sample Buffer (Thermo Fisher) supplemented with 50 mM DTT at 70 °C for 10 min and afterwards collected by centrifugation at RT and 16,000 x g for 1 min. The supernatant containing the histones from 2 x 10^6 bloodstream form trypanosomes was transferred to a new 1.5 ml reaction tube. To reduce histones for mass spectrometry analysis, elution buffer was supplemented with 50 mM DTT (final concentration).

Locus-specific histone isolation

Non-TSS nucleosomes were purified from the cell line BFpFK8, which allows inducible overexpression of TY-tagged H2A from an *rDNA* locus (Lowell et al., 2005). For each replicate, mononucleosomes were isolated from 1 x 10^9 cells as described above. Immunoprecipitation of non-TSS nucleosomes was performed by adding 50 µl Dynabeads Protein G (Invitrogen) coupled to 10 µg monoclonal, purified BB2 mouse antibody (Bastin et al., 1996) to 1 ml mononucleosome-containing supernatant (2 x 10^8 cells). Following an overnight incubation at 4 °C on a rotation wheel, bound material was washed three times with 1 ml 1 x PBS and eluted using 40 µl 1 x NuPAGE LDS Sample Buffer (Thermo Fisher) at 70 °C for 10 min. To reduce histones for mass spectrometry analysis, the elution buffer was supplemented to contain 50 mM DTT (final concentration). The samples were frozen down in liquid nitrogen and stored at -80 °C.

To identify PTMs enriched at TSSs, TSS-nucleosomes were purified from cell line BFJEL41 (Lowell et al., 2005). In this cell line both endogenous H2A.Z alleles were deleted and an ectopic copy of TY1-tagged H2A.Z was constantly overexpressed. For each replicate, mononucleosomes were isolated from 1 x 10^9 cells as described above. Immunoprecipitation of TSS-nucleosomes was performed by adding 50 µl Dynabeads Protein G (Invitrogen) irreversibly crosslinked using dimethyl pimelidate (see Crosslinking of antibodies to Protein G beads) (Sigma-Aldrich) to 10 µg monoclonal, purified BB2 mouse antibody (Bastin et al., 1996) to 1 ml mononucleosome-containing supernatant (2 x 10^8 cells). Following an overnight incubation at 4 °C on a rotation wheel, bound material was washed three times with 1 ml 1 x PBS and eluted using 40 µl 1 x NuPAGE LDS Sample Buffer (Thermo Fisher) at 70°C for 10 min. To reduce histones, the elution buffer was supplemented to contain 50 mM DTT (final concentration). If necessary, the samples were frozen down in liquid nitrogen and stored at -80 °C.

Histone isolation for HAT1 and HAT2 target identification

To identify the target sites of HAT1 and HAT2, the 2T1_HAT1_KD and 2T1_HAT2_KD cell lines were used allowing the inducible downregulation of HAT1 and HAT2 by RNAi. Following depletion of HAT1 or HAT2 over a time period of 48 h, histones were isolated using the acid extraction of histones from mononucleosomes protocol (see above) or an adapted locus-specific histone isolation protocol was conducted. For immunoprecipitation of TSS-nucleosomes following depletion of HAT1 or HAT2, mononucleosomes were isolated from 1 x 10^9 2T1_HAT1_KD and 2T1_HAT2_KD cells, respectively, as

described above. 100 µl Dynabeads® M-280 sheep anti-rabbit IgG (Invitrogen), which were irreversibly crosslinked using dimethyl pimelidate (see Crosslinking of antibodies to Protein G beads) to 10 µg of a custom-made H2A.Z rabbit antibody (Wedel et al., 2017) were added to 1 ml mononucleosome-containing supernatant (2 x 10^8 cells). Following an overnight incubation at 4 °C on a rotation wheel, bound material was washed three times with 1 x PBS and eluted using 1 x NuPAGE® LDS Sample Buffer (Thermo Fisher) at 70 °C for 10 min. To reduce histones, elution buffer was supplemented with 50 mM DTT. If necessary, the samples can be frozen down in liquid nitrogen and stored at -80 °C.

Crosslinking of antibodies to Protein G beads

The required amount of Protein G beads was transferred to a 1.5 ml reaction tube and washed three times with 200 µl PBS-T (1 x PBS + 0.02% Tween 20). The beads were resuspended in 200 µl PBS-T (1 x PBS + 0.02% Tween 20) containing 10 µg of the required antibody. The antibodies were coupled to the beads for 2-4 h at RT on a rotation wheel. To remove the supernatant, the tube was placed on a magnetic rack for 2 min to bind the beads to the tube wall. The beads were washed twice using 500 µl PBS-T (1 x PBS + 0.02% Tween 20). Therefore, the reaction tube was removed from magnetic rack and the beads were resuspended in the solution. To separate the beads from the supernatant, the reaction tube was placed back onto the magnetic rack. The following washing steps were performed accordingly. After washing with PBS-T (1 x PBS + 0.02% Tween 20), the beads were washed once in 500 µl 1 x PBS and then twice using 500 µl 0.2 M sodium borate pH 9.0 (Stock solution: 1 M sodium borate pH 9; 1 M boric acid and the pH was adjusted with NaOH). The beads were resuspended in 500 µl 0.2 M sodium borate pH 9.0 containing freshly added dimethyl pimelimidate (Sigma) to a final concentration of 25 mM (6.5 mg/ml). The antibodies were irreversibly crosslinked to the beads for 45 min at RT on a rotation wheel. Afterwards, the beads were washed once in 500 µl 0.2 M sodium borate pH 9.0. The beads were resuspended in 500 µl 0.2 M ethanolamine pH 8 (Sigma; Stock solution 1 M; pH adjusted with NaOH) and incubated for 2 h at RT on a rotation wheel to block all the N-reactive groups. Afterwards, the beads were washed twice in 1 x PBS. washed once with 500 µl 0.3 M NaOAc, washed once with 500 µl 0.2 M sodium borate pH 9.0 and twice using 1 x PBS. Afterwards, the beads are resuspended in 200 µl PBS and can be used for immunoprecipitation reactions or stored for 24 h at 4 °C.

Quantification of site-specific acetylation levels

In the section “Quantification of site-specific acetylation levels”, I aim to comprehensively describe the experimental and computational methodologies required to determine the site-specific acetylation levels on trypanosomal histones, including sample preparation for mass-spectrometry analysis, mass-spectrometry analysis of histone PTMs, Data processing and Fragment Ion Patchwork Quantification, estimation of histone methylation levels, and identification of novel histone PTMs. The experimental and computational work was performed by Dr. Rasha ElBashir, Dr. Jens Vanselow, Stephanie Lamer, Beate

Vogt, and Dr. Andreas Schlosser in the laboratory of Dr. Andreas Schlosser at the Rudolf Virchow Center for Integrative and Translational Bioimaging of the University Würzburg. The technological details, I describe, were published in (ElBashir et al., 2015):

ElBashir R, Vanselow JT, Kraus A, Janzen CJ, Siegel NT, Schlosser A. (2015). Fragment Ion Patchwork Quantification for Measuring Site-Specific Acetylation Degrees. Anal Chem 87:9939–9945.

I include this section for reasons of completeness and coherence of my study. To avoid any mistakes in the description, it closely resembles the original.

Sample preparation for mass-spectrometry analysis

Sample processing was performed, as described previously (ElBashir et al., 2015). Enriched and with DTT reduced histones were alkylated using 4.8 µl of 1 M freshly prepared Iodacetamide (final concentration: 120 mM) for 20 min at room temperature without exposure to light. For size-separation on NuPAGE Novex 4—12% Bis-Tris gels (Thermo Fisher) using the MOPS buffer system according to manufacturer's instructions, each replicate was separated over three lanes. As molecular weight standard the Novex Sharp Prestained Protein Standard (Thermo Fisher) was used to monitor protein migration. The gels were run at 200 V for around 15 min at room temperature until the 30 kDa ladder had separated. Gels were washed three times for 5 min with water and stained for 45 min with Simply Blue Safe Stain (Thermo Fisher). Following destaining of the gel in desalted water for 2 h, the histone-containing gel regions were excised. Gel slices were chopped into small pieces and prepared for mass-spectrometry analysis. For complete removal of the loading dye, the gel slices were destained using 400 µl destaining buffer (70% acetonitrile in 100 mM NH_4HCO_3 (pH 8)), equilibrated in 400 µl fresh equilibration buffer (100 mM ammonium bicarbonate), shrunk using 200 µl 100% acetonitrile for 2 min on a vortexer plate until the gel pieces were white, and dried in a vacuum concentrator (Eppendorf). For in-gel-acetylation, dried gel pieces were resuspended in a mixture of 10 µl of acetic anhydride-1,1′-$^{13}C^2$ and 50 µl of 1 M sodium acetate-1-^{13}C (pH 7) and incubated for 3.5 h at 25 °C on a shaker. To refresh the reaction solutions, the supernatants were carefully removed by pipetting and fresh reaction solutions were added. Incubations were continued for 3.5 h at 25 °C on a shaker. Supernatants were removed, gel pieces were quenched in 100 µl 500 mM ammonium bicarbonate and reactions were washed applying three washing steps. For the first washing step, 250 µl 500 mM ammonium bicarbonate were added to the quenched cell pieces and the reaction was incubated for 15 min at room temperature on a vortexer plate. Afterwards, the washed gel slices were shrunk in 200 µl 100% acetonitrile for 2 min on a vortexer plate until the gel pieces were white and dried in a vacuum concentrator (Eppendorf). For the second and third washing step 250 µl 250 mM and 100 mM ammonium bicarbonate were used, respectively. The second and the third washing step were performed as the first washing step. The washed and dried gel pieces were stored at 4 °C until mass-spectrometry analysis was performed.

Mass-spectrometry analysis of histone PTMs

To analyse the extracted histones using mass-spectrometry, the washed and dried gel pieces were resuspended in 200 µl 100 mM NH_4HCO_3 (pH 8) containing 0.1 µg of a protease (elastase, thermolysin, or papain). Per replicate, all the proteases were used. Proteolytic digests were performed overnight at 37 °C. L-Cysteine (5 mM) was added to papain digests in order to activate the protease. Peptides were extracted from the gel slices with 100 µl 5% formic acid and transferred to LC vials. NanoLC-MS/MS analyses were performed on an Orbitrap Fusion (Thermo Fisher) equipped with an EASY-Spray Ion Source and coupled to an EASY-nLC 1000 (Thermo Fisher). Peptides were loaded on a trapping column (2 cm x 75 µm ID, PepMap C-18 3 µm particles, 100 Å pore size) and separated on an EASY-spray column (25 cm x 75 µm ID, PepMap C-18 2 µm particles, 100 Å pore size) with a 120 min linear gradient from 3 to 30% acetonitrile and 0.1% formic acid. Both MS and MS/MS scans were acquired in the Orbitrap analyser with a resolution of 60.000. HCD fragmentation with 35% normalised collision energy was applied. A top speed data-dependent MS/MS method with a fixed cycle time of 3 sec was used. Dynamic exclusion was applied with a repeat count of 3 and an exclusion duration of 45 sec. Minimum signal threshold for precursor selection was set to 50.000, and the width of the precursor selection window was set to 6 Da. Predictive AGC was used with a target value of 2 $x10^5$ for MS scans and 5 x 10^4 MS/MS scans. EASY-IC was used for internal calibration.

Data processing

Data processing was performed as described by ElBashir et al. (ElBashir et al., 2015). R scripts used for the analysis and the software can be downloaded from github [https://github.com/JTVanselow/Patchwork_AcK_Quant] and were deposited at Zenodo DOI 10.5281/zenodo.3662776 [https://doi.org/10.5281/zenodo.3662776]. Each raw data file was processed with Mascot Distiller 2.4 (Matrix Science) using two different types of parameter settings. One parameter set (ID) generated peak lists for database searching with Mascot, the other parameter set (Quant) generated peak lists for quantitative analysis. ID peak lists were searched with Mascot Server 2.4 against a costume database, compiled from a *T. brucei* database downloaded from https://tritrypdb.org/tritrypdb/. In addition to all histone sequences, the custom database contained the top 300 of proteins identified in our histone samples. Acetylation with $_{13}C_1$-acetyl is a proxy for in vivo unmodified lysines. Thus, all Ks were replaced by an O in all protein sequences of the database fasta file. O was defined as a new amino acid with the chemical composition $_{13}CC_7H_{14}N_2O_2$ (monoisotopic mass 171.108883 Da). For N-terminal acetylation, J was defined as a new amino acid with the chemical composition of $_{13}C_1$-acetyl ($_{13}CCH_2O$, monoisotopic mass 43.01392 Da). J was added to all protein N-termini of the database fasta file. The fasta file contains each protein sequence in two versions, one starting with J and the other starting with JM. All variable protein N-terminal and lysine modifications were defined relative to $_{13}C_1$-acetyl (see below). Database searching was performed for the following

modifications: oxidation at M (O), light acetyl at O ($_{13}C-_{1}C$), light acetyl at protein N-terminus ($_{13}C-_{1}C$), monomethyl at O (CH_2), monomethyl at protein N-terminus (CH_2), dimethyl at O ($_{13}C-_{1}CH_2O-1$), dimethyl at protein N-terminus ($_{13}C-_{1}CH_2O-1$), trimethyl at O ($_{13}C-_{1}C_2H_4O-1$), and trimethyl at protein N-terminus ($_{13}C-_{1}C_2H_4O-1$), taking into account that also monomethylated amino groups react with acetic anhydride. In addition, carbamidomethyl at C (C_2H_3NO) was set as a fixed modification. Database searches were performed excluding protease specificity (enzyme: none), 7 ppm mass tolerance for the precursor, and 0.02 Da mass tolerance for fragment ions. The Mascot search results were exported as XML file and the start and end positions of the peptides were included in the export file additional to the standard parameters. Peptides, which have a score <15 (corresponding to average E-values of 0.05), were removed.

Fragment ion patchwork quantification

To perform MS2-isotope pattern-based quantification, scripts programmed in R version 3.1.2 using parallel processing when possible (doParallel) were executed. XML files containing Mascot identifications and peak list files with complete isotope patterns of fragment ions (exported as mzData-files from Mascot Distiller after processing with the Quant options, see above), were parsed and combined. For each peptide, fragment ion masses were calculated with OrgMassSpecR, and fragment ions with one, two, or three acetylation sites (incl. N-terminal) were kept for further analysis. Data from the same experiment from different digests or technical replicates were combined. The site information from the Mascot search was considered as reliable, if the delta score was higher than 10, and from the peptides, which were clearly identified, a list of reliable modification sites was generated. These were compared to the set of modifications of each isoform of peptides with delta scores <10. If the number of ambiguous peptide-isoforms could be reduced to one by considering only reliable modifications, all other ambiguous isoforms were eliminated, and the peptide was considered as “non-ambiguous”. In case that none or more than one isoform could be explained by reliable modifications alone, all isoforms were used for further analysis, but marked as “ambiguous” in the results table. In the first quantification round, theoretical isotope patterns for fragment ions with a single acetylation site were calculated using Rdisop, one heavy ($_{13}C_1$) acetylated and one light (natural) acetylated isotope pattern for each fragment. The isotope purity of the heavy acetic anhydride (99.5% according to certificate of analysis) was considered for the calculation of heavy isotope patterns. For each fragment ion, calculated theoretical isotope patterns were combined to “heavy-to-light” proportions from 0% to 100% in 0.01% steps. Resulting mixed isotope patterns were compared with the corresponding experimental isotope pattern. For each mixing ratio, the deviation from the experimental pattern was calculated. A mixing ratio with a minimum deviation from the experimental pattern was assumed as equal to the acetylation degree of this site. To remove calculated isotope patterns, which are inferred by other fragment ions or are only partial due to a low intensity, the Pearson correlation r between experimental and theoretical isotope

pattern >0.99 and the maximum difference in the relative abundance of a single isotope peak <20 were applied as quality criteria using additionally the -1 isotope peak of the isotope patterns. Since the acetylation degree of particular sites can be influenced by other modifications in the same peptide, the acetylation degree for each fragment ion covering a particular site was calculated using the peptide context information about other modifications to group acetylation degrees during outlier removal and averaging. Outliers (outside 1.5 x interquartile range, IQR) from the distribution of acetylation degrees were removed. The median acetylation degrees for each site (and modification context) were used for calculating the acetylation degree of fragment ions containing two acetylation sites. Here, heavy and light isotope patterns were calculated with one of the acetylation sites treated as acetylated with the median acetylation degree from the first quantification round. For comparing the mixed theoretical isotope patterns with the experimental patterns, all peaks from the monoisotopic to the +4-isotope peak were used to identify the best matching mixing ratio and the acetylation degree of this particular site. The same quality criteria as described above were applied but using the -1 to the +4 peaks. Outliers were removed as above. For each fragment ion containing two acetylation sites, an acetylation degree for both sites is calculated, if a median value from the first round was available for both sites. The analysis of fragment ions with three acetylation sites was performed accordingly, but the theoretical isotope patterns were calculated using the median acetylation degrees of two of the three acetylation sites determined in the second round.

Estimation of histone methylation levels

Methylation levels were estimated by spectra counting. To this end the number of spectra for each modification type containing a specific modification site (e.g. dimethylated on K-4) were counted. The following modification forms were considered: unmodified, monomethyl, dimethyl, trimethyl as described before (ElBashir et al., 2015). Spectra counting does not consider the differences in ionisation efficiencies between differentially modified peptides (e.g. mono- and dimethylated form of the same peptide) and is thus less accurate than FIPQuant.

Identification of novel PTMs

Data from FIPQuant were analyzed for additional PTMs with PEAKS studio X (Bioinformatics Solutions Inc., Canada). Raw data refinement was performed with the following settings: Merge Options: no merge, Precursor Options: corrected, Charge Options: no correction, Filter Options: no filter, Process: true, Default: true, Associate Chimera: yes. De novo sequencing and PEAKS database searching were performed with Parent Mass Error Tolerance set to 8 ppm. Fragment Mass Error Tolerance was set to 0.02 Da, and Enzyme was set to none. The following variable modifications have been used: Oxidation (M), Acetylation (K, protein N-terminal), Acetylation $_{13}C_1$ (K, protein N-terminal), combined Methylation with Acetylation $_{13}C_1$ (K, protein N-terminal), Trimethylation (K, protein N-terminal), Dimethylation (K, protein N-terminal), Phosphorylation (STY). Carbamidomethylation (C) was set as fixed modification. A

maximum of 5 variable PTMs were allowed per peptide. A custom database with all *T. brucei* histone sequences as well as common contaminants was used for database searching. Results were filtered to 0.5% PSM-FDR.

Next-generation-sequencing approaches

MNase-ChIP sequencing

MNase-ChIPs were performed using WT cells, the cell line BFpFK8, which is capable of inducible overexpression TY-tagged H2A from an *rDNA* locus (Lowell et al., 2005), and using 2T1 cell lines, in which HAT1 or HAT2 can be inducibly depleted, as previously described (Wedel and Siegel, 2017). In brief, 2 x 10^8 cells were harvested at 1,400 x g and the cells were crosslinked in 1% formaldehyde solution consisting of 40 ml fresh HMI-11 medium and 4 ml Formaldehyde (Sigma) for 20 min at RT and regular mixing (every 5 min). The crosslink was stopped by adding 2.5 ml of 2 M glycine and the crosslinked cells were centrifuged at 4 °C and 4,000 x g for 20 min. Afterwards, the pellet was washed in 30 ml cold 1 x TDB-buffer and spun down at 4 °C and 4,000 x g for 20 min. The supernatant was decanted, the cell pellet was dissolved in 1 ml of permeabilisation buffer (100 mM KCl, 10 mM Tris 8.0, 25 mM EDTA 8.0, 1.46 µM pepstatin A, 4.7 µM leupeptin, 1 mM PMSF and 1 mM TLCK), transferred to a 1.5 ml DNA lowbind reaction tube and Digitonin was added to the cell suspension to a final concentration of 200 µM (stock solution: Digitonin was dissolved to 4 mM in permeabilisation buffer and heated to 98 °C for dissolution). Cell lysis was performed at RT for 15 min under gentle rotation and afterwards, the cells were collected by centrifugation at 4 °C and 1,400 x g for 5 min. To remove any traces of EDTA, the cell pellet was washed three times in ice-cold NP-S buffer (0.5 mM Spermidine, 0.075% (v/v) IGEPAL, 50 mM NaCl, 10 mM Tris-HCl pH 7.5, 5 mM $MgCl_2$, 1 mM $CaCl_2$, 1.46 µM pepstatin A, 4.7 µM leupeptin, 1 mM PMSF and 1 mM TLCK). The cell pellet was always dissolved by vortexing. Following the last washing step, the cell pellet was dissolved in 400 µl ice-cold NP-S buffer and the chromatin was fragmented by adding 1 U/µl MNase (Sigma-Aldrich) and incubating the mixture for 5 min at 25 °C in a heating block. The digestion was stopped by adding 8 µl of 0.5 M EDTA pH 8 (final concentration: 10 mM).

To purify H2A.Z-associated DNA, immunoprecipitation was performed using Dynabeads® M-280 sheep anti-rabbit IgG (Invitrogen) coupled to 10 µg polyclonal affinity-purified H2A.Z rabbit antibody (Wedel et al., 2017). To purify TY1-tagged H2A and TY1-tagged H2A.Z, immunoprecipitation was performed using Dynabeads® Protein G (Invitrogen) coupled to 10 µg mono-clonal, purified BB2 mouse antibody (Bastin et al., 1996). To purify H3-associated DNA, immunoprecipitation was performed using 100 µl Dynabeads® M-280 sheep anti-rabbit IgG (Invitrogen) coupled to 33 µl of polyclonal H3 rabbit antiserum (Gassen et al., 2012). Following overnight incubation at 4 °C in the presence of 0.05% SDS on a rotation wheel, bound material was washed and eluted. DNA-protein cross-links were reversed at 65 °C for ~11

h in the presence of 300 mM NaCl. ChIP-seq libraries were constructed using 35 ng of immunoprecipitated or input DNA, as described previously (Wedel and Siegel, 2017), and library concentrations were determined in duplicates using Qubit dsDNA HS Assay Kit (Invitrogen) for Qubit 2.0 Fluorometer (Invitrogen). ChIP-sequencing libraries were quantified using the KAPA Library Quantification Kit (KAPA Biosystems) according to the manufacturer's instruction and sequenced in paired-end mode with 2 x 76 cycles using an Illumina® Miseq or an Illumina® NextSeq 500 platform.

RNAPII ChIP sequencing

RNAPII ChIPs were performed in 2T1 cell lines, in which both alleles of *RPB9*, a DNA-bound subunit of the RNAPII complex in *T. brucei* were N-terminally TY1-tagged and, if required, allowed inducible RNAi-mediated depletion of HAT1 or HAT2. Except for minor changes, the ChIPs were performed as described previously (Wedel et al., 2017) and as explained above. In brief, 3 x 10^8 cells were harvested, crosslinked in 1% formaldehyde and lysed using digitonin (200 µM final). After centrifugation, the pellet was resuspended in 1000 µl NP-S buffer supplemented with 10 µl of 10% SDS solution (final concentration: 0.1% SDS) and sonicated using Covaris® S220 focused-ultrasonication (Peak Intensity: 150, intensity: 5, Duty Factor: 10, Counts Per Burst: 200, Time: 10 min). After centrifugation, the supernatant was used for immunoprecipitation and 60 µl of the supernatant were separated as input. Immunoprecipitation of fragmented TY1-RPB9-bound DNA was performed using 50 µl Dynabeads® Protein G (Invitrogen), coupled to 10 µg mono-clonal, purified BB2 mouse antibody (Bastin et al., 1996). Following overnight incubation at 4 °C in presence of 300 mM NaCl on a rotation wheel, bound material was washed, and eluted, and DNA-protein cross-links were reversed at 65 °C for ~11 h in the presence of 300 mM NaCl. ChIP-seq libraries were constructed using ~5 ng of immunoprecipitated or 35 ng input DNA as described previously (Wedel and Siegel, 2017), size-selected (150 bp – 350 bp) using a native PAGE gel system and library concentrations were determined in duplicates using the Qubit dsDNA HS Assay Kit (Invitrogen) for Qubit 2.0 Fluorometer (Invitrogen). ChIP-sequencing libraries were quantified using the KAPA Library Quantification Kit (KAPA Biosystems/Roche) according to the manufacturer's instruction and sequenced in paired-end mode with 2 x 76 cycles using an Illumina® NextSeq 500 platform.

RNA sequencing

RNA sequencing was performed in triplicates for WT cells and 2T1 cells, in which HAT1 or HAT2 were inducibly depleted over 48 h, as previously described (Kraus et al., 2019). In brief, total RNA was extracted from 45 million cells, which were grown to a density up to 0.9 x 10^6 cells/ml, using the NucleoSpin RNA kit (Macherey-Nagel). Therefore, the cell lysis buffer was supplemented with 3.8 µl 1 M RNAse-free dithiothreitol (Sigma-Aldrich) and 1 µl of 1:10 Ambion ERCC RNA Spike-In Mix (Thermo Fisher). Depletion of ribosomal RNAs from 2 µg of total RNA was performed using hybridisation of biotinylated oligos designed for efficient and specific depletion of trypanosomal rRNA and removal of rRNA-oligo hybrids using streptavidin-coated paramagnetic beads (Kraus et al., 2019). Afterwards, cDNA

was synthesised using the NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England Biolabs) and strand specific RNA-seq libraries were generated as described (Kraus et al., 2019). Library concentrations were determined in duplicates using Qubit dsDNA HS Assay Kit (Invitrogen, cat. no. Q32854) for Qubit 2.0 Fluorometer (Invitrogen, cat. no. Q32866). Strand-specific RNA-sequencing libraries were quantified using the KAPA Library Quantification Kit (KAPA Biosystems/Roche) according to the manufacturer's instruction and sequenced in paired-end mode on an Illumina NextSeq 500 sequencer 2 x 76 cycles.

Bioinformatic approaches to analyse sequencing data sets

Mapping, normalisation, and visualisation of ChIP sequencing data

Processing of ChIP sequencing data was performed as described previously (Wedel and Siegel, 2017). In brief, adapter sequences were removed using Cutadapt (Martin, 2011) and the ChIP-sequencing datasets were mapped to *T. brucei* Lister 427 genome version 9 (HGAP3_Tb427v9), which was published by Müller-Hübner and Cosentino *et al.* (Müller et al., 2018) using Bowtie 2 (Li et al., 2009). The HGAP3_Tb427v9 fasta file (HGAP3_Tb427v9.fasta) and annotation file (HGAP3_Tb427v9.gff) can be downloaded from Zenodo (https://doi.org/10.5281/zenodo.823671). Following SAM to BAM conversion, the aligned reads were sorted and indexed using SAMtools version 1.8 (Li et al., 2009). The number of reads was normalised to counts per billion mapped reads and coverage files were generated in the wiggle format using COVERnant version 0.3.0 with the subcommand *ratio*, as previously described (Wedel et al., 2017). For visualisation, regions of interests were extracted from the wiggle files and the coverage was illustrated using GraphPad Prism version 7.0c.

The coverage for multiple regions was extracted and averaged using COVERnant *extract*. The resulting average plots were visualised plotting the median without zeros of the output-matrices to the indicated locations and illustrated using GraphPad Prism version 7.0c.

Mapping, normalisation, and visualisation of RNA sequencing data

RNA-seq data sets were mapped, processed, and analysed similar to as described previously (Kraus et al., 2019). In short, adapter sequences were removed using Cutadapt (Martin, 2011) and the RNA-seq datasets were mapped to a hybrid genome, which contains the *T. brucei* Lister 427 genome version 9 (HGAP3_Tb427v9; from Zenodo https://doi.org/10.5281/zenodo.823671) and the sequences of the 92 ERCC spike-in transcripts, using BWA-mem (Li et al., 2009). Following SAM to BAM conversion, the aligned reads were sorted and indexed using SAMtools version 1.8 (Li et al., 2009). Thereby, unmapped, PCR or optical duplicate, not primary aligned and supplementary aligned reads were removed from the alignment files using the SAM flag *3332* (Li et al., 2009).

Using BAM files, reads per *T. brucei* CDS were counted with the HGAP3_Tb427v9.gff as annotation file and the reads per ERCC transcripts were counted with ERCC92.gff using the GenomicAlignments

package (Lawrence et al., 2013) in R (Team). Differential gene expression analysis was conducted using the DESeq2 package (Love et al., 2014) from R/Bioconductor normalising the counts per *T. brucei* gene to the ERCC spike-in counts and comparing the normalised gene counts of HAT1 KD 48h or HAT2 KD 48 h to wild type. The fold changes (log_2) were determined and considered as differentially expressed if the adjusted P value (calculated based on Wald test and adjusted for multiple testing using the procedure of Benjamini and Hochberg) was below 0.1 (Benjamini et al., 2001). To analyse the depletion efficiency for HAT1 and HAT2, the CDS information of HAT1 and HAT2 in the annotation file was split into the RNAi target, the upstream target and the downstream target region.

Additionally, the number of reads was normalised to reciprocal ERCC factors, calculated using the ERCC factors from the DESeq2 analysis, to generate coverage files in the wiggle format using COVERnant version 0.3.0 with the subcommand *ratio* including the additional parameters *–factor_numerator* and *–factor_denominator*. For visualisation, regions of interests were extracted from the wiggle files and the coverage was illustrated using GraphPad Prism version 7.0c. Meta-plots were generated by plotting the median without zeros of the output-matrices to the indicated locations and illustrated using GraphPad Prism version 7.0c.

Meta-plot generation for binning PTUs

To be able to average the sequencing coverage data for multiple polycistronic transcription units with different length, COVERnant's subcommand *extract* was extended by the optional argument *--meta_plots*, which calculates the coverage as a fraction relative to the length (https://github.com/bgbrink/COVERnant).

List of antibodies used in this study

Table 2.6. List of antibodies used in this study

Antibody	Information	Reference
H2A.Z	primary; polyclonal rabbit antibody raised against trypanosomal H2A.Z; purified from serum	Wedel et al., 2017
H3	primary; polyclonal rabbit antibody raised against trypanosomal H3; serum	Gassen et al., 2012
BB2	primary; monoclonal mouse antibody for detecting the TY1-tag; purified from hybridoma cells	Bastin et al., 1996
Anti-mouse IgG Sheep Polyclonal Antibody (HRP (Horseradish Peroxidase))	secondary; polyclonal sheep antibody raised against mouse IgG; purified from serum and coupled to HRP	VWR
Anti-rabbit IgG Donkey Polyclonal Antibody (HRP (Horseradish Peroxidase))	secondary; polyclonal donkey antibody raised against rabbit IgG; purified from serum and coupled to HRP	VWR

List of bioinformatic software used in this study

Table 2.7. List of bioinformatic software used in this study

Software	Reference
Bowtie 2 (version 2.2.9)	(Langmead and Salzberg, 2012); Fast gapped-read alignment with Bowtie 2. Nat Methods 9(4):357-9. (2012)
BWA-mem (version 0.7.16)	(Li and Durbin, 2010); Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25(14):1754-60. (2009)
COVERnant v0.3.0	(Wedel et al., 2017); https://github.com/konrad/COVERnant
cutadapt (version 1.15)	(Martin, 2011); Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. EMBnet.journal 17, 10-12. (2011)
GraphPad Prism version 7.0c	http://www.graphpad.com
IGV browser (version 2.4.10)	(Robinson et al., 2011); http://software.broadinstitute.org/software/igv
ImageJ	(Schneider et al., 2012); NIH Image to ImageJ: 25 years of image analysis. Nature methods 9(7): 671-675. (2012)
Python (version 3.6)	(van Rossum and de Boer, 1991); Interactively Testing Remote Servers Using the Python Programming Language. CWI Quarterly 4(4):283-303. (1991)
R (version 3.5.0)	(Team); A language and environment for statistical computing. R Foundation for Statistical Computing. https://www.R-project.org/ (2018)
R package GenomicFeatures	(Lawrence et al., 2013); Software for computing and annotating genomic ranges. PLoS Comput Biol 9, e1003118 . (2013).
R package DESeq2 (version 1.20.0)	(Love et al., 2014); Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550 (2014)
R package RColorBrewer	(Neuwirth, 2014); Package 'RColorBrewer'. https://cran.rstudio.com/web/packages/RColorBrewer/RColorBrewer.pdf. (2014)
R package gplots	(Warnes et al., 2016); Package 'gplots'. https://cran.r-project.org/web/packages/gplots/gplots.pdf. (2016)
SAMtools (version 1.8)	(Li et al., 2009); The Sequence Alignment/Map format and SAMtools. Bioinformatics 25(16):2078-9. (2009)

Information about sequencing data discussed in this study

Table 2.8: Information about discussed sequencing data.
The sequencing data discussed in this study and generated for the publication by Kraus et al. have been deposited in the Gene Expression Omnibus (GEO) and are through the accession number GSE145812. The source of external sequencing data is accessible over the corresponding reference. Unpublished sequencing data discussed in this study are administered by Prof. T. N. Siegel and are available upon request. Information about their processing is listed in **Table 2.9**.

	Sequencing					
Reference	**Fastq name**	**Type of seq**	**Ilumina sequencer**	**Seq mode**	**Read length**	**Seq reads**
[1]	NS263_WT_ChIP_H2AZ_R1.fq.bz2; NS263_WT_ChIP_H2AZ_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3015378
[1]	NS276_2T1_HAT2_KD_ChIP_0h_1_R1.fq.gz; NS276_2T1_HAT2_KD_ChIP_0h_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3380265
[1]	NS277_2T1_HAT2_KD_ChIP_24h_1_R1.fq.gz; NS277_2T1_HAT2_KD_ChIP_24h_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3357305
[1]	NS278_2T1_HAT2_KD_ChIP_48h_1_R1.fq.gz; NS278_2T1_HAT2_KD_ChIP_48h_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3453220
[1]	NS283_HAT2_KD_2T1_ChIP_0h_2_R1.fq.gz; NS283_HAT2_KD_2T1_ChIP_0h_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2619228
[1]	NS284_HAT2_KD_2T1_ChIP_24h_2_R1.fq.gz; NS284_HAT2_KD_2T1_ChIP_24h_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2695096
[1]	NS285_HAT2_KD_2T1_ChIP_48h_2_R1.fq.gz; NS285_HAT2_KD_2T1_ChIP_48h_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2276000
[1]	NS286_HAT2_KD_2T1_ChIP_72h_1_R1.fq.gz; NS286_HAT2_KD_2T1_ChIP_72h_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1080736
[1]	NS287_HAT2_KD_2T1_ChIP_72h_2_R1.fq.gz; NS287_HAT2_KD_2T1_ChIP_72h_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	7344525
[1]	NS291_HAT2_KD_2T1_Input_0h_2_R1.fq.gz; NS291_HAT2_KD_2T1_Input_0h_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2788620
[1]	NS292_HAT2_KD_2T1_Input_24h_2_R1.fq.gz; NS292_HAT2_KD_2T1_Input_24h_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2709181
[1]	NS293_HAT2_KD_2T1_Input_48h_2_R1.fq.gz; NS293_HAT2_KD_2T1_Input_48h_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3004490
[1]	NS294_HAT2_KD_2T1_Input_72h_1_R1.fq.gz; NS294_HAT2_KD_2T1_Input_72h_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2816922
[1]	NS295_HAT2_KD_2T1_Input_72h_2_R1.fq.gz; NS295_HAT2_KD_2T1_Input_72h_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2084397
[1]	NS476_2T1_HAT1_KD_24h_Input_1_R1.fq.gz; NS476_2T1_HAT1_KD_24h_Input_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	7211097
[1]	NS477_2T1_HAT1_KD_48h_Input_1_R1.fq.gz; NS477_2T1_HAT1_KD_48h_Input_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	6611792
[1]	NS484_2T1_HAT1_KD_0h_ChIP_1_R1.fq.gz; NS484_2T1_HAT1_KD_0h_ChIP_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	8091270
[1]	NS485_2T1_HAT1_KD_24h_ChIP_1_R1.fq.gz; NS485_2T1_HAT1_KD_24h_ChIP_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	8246854
[1]	NS486_2T1_HAT1_KD_48h_ChIP_1_R1.fq.gz; NS486_2T1_HAT1_KD_48h_ChIP_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	11620011
[1]	NS487_2T1_HAT1_KD_0h_Input_1_R1.fq.gz; NS487_2T1_HAT1_KD_0h_Input_1_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	10018772
[1]	NS491_2T1_HAT1_KD_0h_ChIP_2_R1.fq.gz; NS491_2T1_HAT1_KD_0h_ChIP_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3980805
[1]	NS492_2T1_HAT1_KD_24h_ChIP_2_R1.fq.gz; NS492_2T1_HAT1_KD_24h_ChIP_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	4484259
[1]	NS493_2T1_HAT1_KD_48h_ChIP_2_R1.fq.gz; NS493_2T1_HAT1_KD_48h_ChIP_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	12761757
[1]	NS494_2T1_HAT1_KD_0h_Input_2_R1.fq.gz; NS494_2T1_HAT1_KD_0h_Input_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	5350241
[1]	NS495_2T1_HAT1_KD_24h_Input_2_R1.fq.gz; NS495_2T1_HAT1_KD_24h_Input_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	5922089
[1]	NS496_2T1_HAT1_KD_48h_Input_2_R1.fq.gz; NS496_2T1_HAT1_KD_48h_Input_2_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	5596528
[1]	NS531_HAT1KD_0h_H2AZ_ChIP_3_R1.fq.bz2; NS531_HAT1KD_0h_H2AZ_ChIP_3_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2967118
[1]	NS532_HAT1KD_0h_H2AZ_ChIP_4_R1.fq.bz2; NS532_HAT1KD_0h_H2AZ_ChIP_4_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3597753

	Sequencing					
Reference	Fastq name	Type of seq	Ilumina sequencer	Seq mode	Read length	Seq reads
[1]	NS533_HAT1KD_48h_H2AZ_ChIP_3_R1.fq.bz2; NS533_HAT1KD_48h_H2AZ_ChIP_3_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3738898
[1]	NS534_HAT1KD_48h_H2AZ_ChIP_4_R1.fq.bz2; NS534_HAT1KD_48h_H2AZ_ChIP_4_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3015430
[1]	NS535_HAT1KD_0h_H2AZ_Input_3_R1.fq.bz2; NS535_HAT1KD_0h_H2AZ_Input_3_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3587464
[1]	NS536_HAT1KD_0h_H2AZ_Input_4_R1.fq.bz2; NS536_HAT1KD_0h_H2AZ_Input_4_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3176148
[1]	NS537_HAT1KD_48h_H2AZ_Input_3_R1.fq.bz2; NS537_HAT1KD_48h_H2AZ_Input_3_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3429705
[1]	NS538_HAT1KD_48h_H2AZ_Input_4_R1.fq.bz2; NS538_HAT1KD_48h_H2AZ_Input_4_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2808329
[1]	NS539_HAT2KD_0h_H2AZ_ChIP_3_R1.fq.bz2; NS539_HAT2KD_0h_H2AZ_ChIP_3_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2060120
[1]	NS540_HAT2KD_0h_H2AZ_ChIP_4_R1.fq.bz2; NS540_HAT2KD_0h_H2AZ_ChIP_4_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3065918
[1]	NS541_HAT2KD_48h_H2AZ_ChIP_3_R1.fq.bz2; NS541_HAT2KD_48h_H2AZ_ChIP_3_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3308527
[1]	NS542_HAT2KD_48h_H2AZ_ChIP_4_R1.fq.bz2; NS542_HAT2KD_48h_H2AZ_ChIP_4_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3561204
[1]	NS543_HAT2KD_0h_H2AZ_Input_3_R1.fq.bz2; NS543_HAT2KD_0h_H2AZ_Input_3_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2884766
[1]	NS544_HAT2KD_0h_H2AZ_Input_4_R1.fq.bz2; NS544_HAT2KD_0h_H2AZ_Input_4_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2783881
[1]	NS545_HAT2KD_48h_H2AZ_Input_3_R1.fq.bz2; NS545_HAT2KD_48h_H2AZ_Input_3_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2961602
[1]	NS546_HAT2KD_48h_H2AZ_Input_4_R1.fq.bz2; NS546_HAT2KD_48h_H2AZ_Input_4_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2643620
[1]	NS470_BF_2T1_TyRPB9_ChIP_1_R1.fq.gz; NS470_BF_2T1_TyRPB9_ChIP_1_R2.fq.gz	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	13626619
[1]	NS471_BF_2T1_TyRPB9_Input_1_R1.fq.gz; NS471_BF_2T1_TyRPB9_Input_1_R2.fq.gz	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	12737249
[1]	NS474_BF_2T1_TyRPB9_ChIP_2_R1.fq.gz; NS474_BF_2T1_TyRPB9_ChIP_2_R2.fq.gz	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	12737249
[1]	NS475_BF_2T1_TyRPB9_Input_2_R1.fq.gz; NS475_BF_2T1_TyRPB9_Input_2_R2.fq.gz	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	11622367
[1]	NS512_HAT1_KD_TyRPB9_0h_Input_1_R1.fq.bz2; NS512_HAT1_KD_TyRPB9_0h_Input_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	6396887
[1]	NS513_HAT1_KD_TyRPB9_24h_Input_1_R1.fq.bz2; NS513_HAT1_KD_TyRPB9_24h_Input_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	8439618
[1]	NS514_HAT1_KD_TyRPB9_48h_Input_1_R1.fq.bz2; NS514_HAT1_KD_TyRPB9_48h_Input_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	8144610
[1]	NS515_HAT2_KD_TyRPB9_0h_Input_1_R1.fq.bz2; NS515_HAT2_KD_TyRPB9_0h_Input_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	7919934
[1]	NS516_HAT2_KD_TyRPB9_24h_Input_1_R1.fq.bz2; NS516_HAT2_KD_TyRPB9_24h_Input_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	7887257
[1]	NS517_HAT2_KD_TyRPB9_48h_Input_1_R1.fq.bz2; NS517_HAT2_KD_TyRPB9_48h_Input_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	4348814
[1]	NS518_HAT1_KD_TyRPB9_0h_ChIP_1_R1.fq.bz2; NS518_HAT1_KD_TyRPB9_0h_ChIP_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	9739561
[1]	NS519_HAT1_KD_TyRPB9_24h_ChIP_1_R1.fq.bz2; NS519_HAT1_KD_TyRPB9_24h_ChIP_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	9440267
[1]	NS520_HAT1_KD_TyRPB9_48h_ChIP_1_R1.fq.bz2; NS520_HAT1_KD_TyRPB9_48h_ChIP_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	8647732
[1]	NS521_HAT2_KD_TyRPB9_0h_ChIP_1_R1.fq.bz2; NS521_HAT2_KD_TyRPB9_0h_ChIP_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	11248028
[1]	NS522_HAT2_KD_TyRPB9_24h_ChIP_1_R1.fq.bz2; NS522_HAT2_KD_TyRPB9_24h_ChIP_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	9259488
[1]	NS523_HAT2_KD_TyRPB9_48h_ChIP_1_R1.fq.bz2; NS523_HAT2_KD_TyRPB9_48h_ChIP_1_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	7396188
[1]	NS567_2T1_TyRPB9_HAT1KD_0h_ChIP_2_R1.fq.bz2; NS567_2T1_TyRPB9_HAT1KD_0h_ChIP_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	4848659
[1]	NS568_2T1_TyRPB9_HAT1KD_24h_ChIP_2_R1.fq.bz2; NS568_2T1_TyRPB9_HAT1KD_24h_ChIP_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	1494552
[1]	NS569_2T1_TyRPB9_HAT1KD_48h_ChIP_2_R1.fq.bz2; NS569_2T1_TyRPB9_HAT1KD_48h_ChIP_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	3314592
[1]	NS570_2T1_TyRPB9_HAT1KD_0h_Input_2_R1.fq.bz2; NS570_2T1_TyRPB9_HAT1KD_0h_Input_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	2378211

	Sequencing					
Reference	Fastq name	Type of seq	Ilumina sequencer	Seq mode	Read length	Seq reads
[1]	NS571_2T1_TyRPB9_HAT1KD_24h_Input_2_R1.fq.bz2; NS571_2T1_TyRPB9_HAT1KD_24h_Input_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	2024278
[1]	NS572_2T1_TyRPB9_HAT1KD_48h_Input_2_R1.fq.bz2; NS572_2T1_TyRPB9_HAT1KD_48h_Input_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	2469782
[1]	NS573_2T1_TyRPB9_HAT2KD_0h_ChIP_2_R1.fq.bz2; NS573_2T1_TyRPB9_HAT2KD_0h_ChIP_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	3295209
[1]	NS574_2T1_TyRPB9_HAT2KD_24h_ChIP_2_R1.fq.bz2; NS574_2T1_TyRPB9_HAT2KD_24h_ChIP_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	3306559
[1]	NS575_2T1_TyRPB9_HAT2KD_48h_ChIP_2_R1.fq.bz2; NS575_2T1_TyRPB9_HAT2KD_48h_ChIP_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	8180258
[1]	NS578_2T1_TyRPB9_HAT2KD_0h_Input_2_R1.fq.bz2; NS578_2T1_TyRPB9_HAT2KD_0h_Input_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	2331259
[1]	NS579_2T1_TyRPB9_HAT2KD_24h_Input_2_R1.fq.bz2; NS579_2T1_TyRPB9_HAT2KD_24h_Input_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	3043255
[1]	NS580_2T1_TyRPB9_HAT2KD_48h_Input_2_R1.fq.bz2; NS580_2T1_TyRPB9_HAT2KD_48h_Input_2_R2.fq.bz2	RPB9 ChIPseq	NextSeq 500	paired-end	2x76 bp	3042335
[1]	NS320_ssRNAd_WT1_R1.fq.bz2; NS320_ssRNAd_WT1_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	16976623
[1]	NS350_ssRNAd_WT2_R1.fq.bz2; NS350_ssRNAd_WT2_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	18535853
[1]	NS351_ssRNAd_WT3_R1.fq.bz2; NS351_ssRNAd_WT3_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	25632003
[1]	NS352_ssRNAd_HAT2KD2T1_48h_1_R1.fq.bz2; NS352_ssRNAd_HAT2KD2T1_48h_1_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	21049801
[1]	NS353_ssRNAd_HAT2KD2T1_48h_2_R1.fq.bz2; NS353_ssRNAd_HAT2KD2T1_48h_2_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	20642909
[1]	NS354_ssRNAd_HAT2KD2T1_48h_3_R1.fq.bz2; NS354_ssRNAd_HAT2KD2T1_48h_3_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	23392972
[1]	NS524_ssRNAseq_HAT1KD2T1_48h_1_R1.fq.bz2; NS524_ssRNAseq_HAT1KD2T1_48h_1_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	21549125
[1]	NS525_ssRNAseq_HAT1KD2T1_48h_2_R1.fq.bz2; NS525_ssRNAseq_HAT1KD2T1_48h_2_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	20683939
[1]	NS526_ssRNAseq_HAT1KD2T1_48h_3_R1.fq.bz2; NS526_ssRNAseq_HAT1KD2T1_48h_3_R2.fq.bz2	RNAseq	NextSeq 500	paired-end	2x76 bp	18908839
[2]	Lowell2005_TyH2AZ_32bp_fastq.txt	H2A.Z ChIPseq	Solexa		32 bp	5809554
[1]	NS563_SM_TyH2A_ChIP_1_R1.fq.bz2; NS563_SM_TyH2A_ChIP_1_R2.fq.bz2	H2A ChIPseq	NextSeq 500	paired-end	2x76 bp	3867340
[1]	NS564_SM_TyH2A_ChIP_2_R1.fq.bz2; NS564_SM_TyH2A_ChIP_2_R2.fq.bz2	H2A ChIPseq	NextSeq 500	paired-end	2x76 bp	2877590
[1]	NS565_SM_TyH2A_Input_1_R1.fq.bz2; NS565_SM_TyH2A_Input_1_R2.fq.bz2	H2A ChIPseq	NextSeq 500	paired-end	2x76 bp	3189383
[1]	NS566_SM_TyH2A_Input_2_R1.fq.bz2; NS566_SM_TyH2A_Input_2_R2.fq.bz2	H2A ChIPseq	NextSeq 500	paired-end	2x76 bp	1159079
	L1500085-BDF1_KD_0h_R1.fq.gz; L1500085-BDF1_KD_0h_R1.fq.gz	H2A ChIPseq	NextSeq 500	paired-end	2x76 bp	4520515
	L1500087-BDF1_KD_48h_R1.fq.gz; L1500087-BDF1_KD_48h_R1.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	4970377
	L1700062_S7_BDF2_KD_2T1_IP_0h_#2_R1_001.fastq.gz; L1700062_S7_BDF2_KD_2T1_IP_0h_#2_R2_001.fastq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2416451
	L1700064_S9_BDF2_KD_2T1_IP_48h_#2_R1_001.fastq.gz; L1700064_S9_BDF2_KD_2T1_IP_48h_#2_R2_001.fastq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2650846
	L1700066_S11_BDF3KD_2T1_IP_0h_#2_R1_001.fastq.gz; L1700066_S11_BDF3KD_2T1_IP_0h_#2_R2_001.fastq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2368563
	L1700068_S13_BDF3KD_2T1_IP_48h_#2_R1_001.fastq.gz; L1700068_S13_BDF3KD_2T1_IP_48h_#2_R2_001.fastq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2954172
	L1500088-BDF4_KD_0h_R1.fq.gz; L1500088-BDF4_KD_0h_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	4467325
	L1500090-BDF4_KD_48h_R1.fq.gz; L1500090-BDF4_KD_48h_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	5264138
	HAT1_OE_0_R1.fq.gz; HAT1_OE_0_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	3195731
	ID-003919-NS137_HAT1_OE_IP_48h_R1.fq.gz; ID-003919-NS137_HAT1_OE_IP_48h_R2.fq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1360428
	ID-003532-NS107_HAT2_OE_IP_0_h_R1.fq.bz2; ID-003532-NS107_HAT2_OE_IP_0_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1079814
	ID-003910-NS128_HAT2_OE_IP_48h_pool_R1.fq.bz2; ID-003910-NS128_HAT2_OE_IP_48h_pool_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	

	Sequencing					
Reference	Fastq name	Type of seq	Ilumina sequencer	Seq mode	Read length	Seq reads
	ID-003534-NS109_BDF2_OE_IP_0_h_R1.fq.bz2; ID-003534-NS109_BDF2_OE_IP_0_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1224225
	ID-003535-NS110_BDF2_OE_IP_48_h_R1.fq.bz2; ID-003535-NS110_BDF2_OE_IP_48_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1138566
	ID-003537-NS112_BDF3_OE_IP_0_h_R1.fq.bz2; ID-003537-NS112_BDF3_OE_IP_0_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1035139
	ID-003539-NS114_BDF3_OE_IP_48_h_R1.fq.bz2; ID-003539-NS114_BDF3_OE_IP_48_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1164467
	ID-003540-NS115_FACT_KD_IP_0_h_R1.fq.bz2; ID-003540-NS115_FACT_KD_IP_0_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1040682
	ID-003541-NS116_FACT_KD_IP_48_h_R1.fq.bz2; ID-003541-NS116_FACT_KD_IP_48_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1337861
	ID-003542-NS117_ISWI_KD_IP_0_h_R1.fq.bz2; ID-003542-NS117_ISWI_KD_IP_0_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	1191899
	ID-003543-NS118_ISWI_KD_IP_48_h_R1.fq.bz2; ID-003543-NS118_ISWI_KD_IP_48_h_R2.fq.bz2	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	954958
	L1700058_S3_WT_I-BET151_IP_0h_#2_R1_001.fastq.gz; L1700058_S3_WT_I-BET151_IP_0h_#2_R2_001.fastq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2615755
	L1700060_S5_WT_I-BET151_IP_48h_#2_R1_001.fastq.gz; L1700060_S5_WT_I-BET151_IP_48h_#2_R2_001.fastq.gz	H2A.Z ChIPseq	NextSeq 500	paired-end	2x76 bp	2637565
[3]	L1800150_H3_V_Ty_MNase_ChIP_1_R1.fq.gz; L1800150_H3_V_Ty_MNase_ChIP_1_R2.fq.gz	H3.V ChIPseq	NextSeq 500	paired-end	2x76 bp	2840144
[3]	L1800151_H3_V_Ty_MNase_Input_1_R1.fq.gz; L1800151_H3_V_Ty_MNase_Input_1_R2.fq.gz	H3.V ChIPseq	NextSeq 500	paired-end	2x76 bp	4176616
[3]	Mueller2018_ATACseq_Wt_20_R1.fastq.gz; Mueller2018_ATACseq_Wt_20_R2.fastq.gz	ATACseq	NextSeq 500	paired-end	2x76 bp	13590213

[1] (Kraus et al., 2020); [2] (Lowell et al., 2005); [3] (Müller et al., 2018)

Information about the processing of the sequencing data discussed in this study

Table 2.9: Information about processing of the discussed sequencing data.
The computational data analysis was implemented as a Unix shell script in a Slurm Workload Manager environment (Yoo et al., 2003). Workflows and custom-made scripts to reproduce the sequencing data are deposited at Zenodo (DOI 10.5281/zenodo.3662776 [https://doi.org/10.5281/zenodo.3662776]).

	Sequencing						
Figure	Description	Mapping software	Read alignments	Reference genome	Wig file ws ss	Average ws ss	Statistics
Figure 4.1 a, c; Figure 5.1 d; Figure 5.2 d; Figure 6.1 a, b; Figure 6.2 c; Figure 6.3 a, b, c, d	WT_ChIP_H2AZ	Bowtie2 local	99.08%	HGAP3_Tb427v9	denominator - ws 5001 ss 1001; ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS276_2T1_HAT2_KD_ChIP_0h_1	Bowtie2 local	97.95%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS277_2T1_HAT2_KD_ChIP_24h_1	Bowtie2 local	98.39%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS278_2T1_HAT2_KD_ChIP_48h_1	Bowtie2 local	98.62%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median

	Sequencing						
Figure	**Description**	**Mapping software**	**Read alignments**	**Reference genome**	**Wig file ws ss**	**Average ws ss**	**Statistics**
Appendix figure 5 b	NS283_HAT2_KD_2T1_ChIP_0h_2	Bowtie2 local	97.67%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS284_HAT2_KD_2T1_ChIP_24h_2	Bowtie2 local	96.62%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.2 d	NS285_HAT2_KD_2T1_ChIP_48h_2	Bowtie2 local	95.92%	HGAP3_Tb427v9	numerator - ws 5001 ss 1001; ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS286_HAT2_KD_2T1_ChIP_72h_1	Bowtie2 local	96.91%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS287_HAT2_KD_2T1_ChIP_72h_2	Bowtie2 local	97.61%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
-	NS291_HAT2_KD_2T1_Input_0h_2	Bowtie2 local	93.45%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS292_HAT2_KD_2T1_Input_24h_2	Bowtie2 local	96.05%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS293_HAT2_KD_2T1_Input_48h_2	Bowtie2 local	94.67%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS294_HAT2_KD_2T1_Input_72h_1	Bowtie2 local	95.57%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS295_HAT2_KD_2T1_Input_72h_2	Bowtie2 local	91.39%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS476_2T1_HAT1_KD_24h_Input_1	Bowtie2 local	97.38%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS477_2T1_HAT1_KD_48h_Input_1	Bowtie2 local	96.91%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 a	NS484_2T1_HAT1_KD_0h_ChIP_1	Bowtie2 local	97.65%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 a	NS485_2T1_HAT1_KD_24h_ChIP_1	Bowtie2 local	97.97%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.1 d	NS486_2T1_HAT1_KD_48h_ChIP_1	Bowtie2 local	97.97%	HGAP3_Tb427v9	numerator - ws 5001 ss 1001; ws 1 ss 1	ws 501 ss 101	median
-	NS487_2T1_HAT1_KD_0h_Input_1	Bowtie2 local	97.73%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 a	NS491_2T1_HAT1_KD_0h_ChIP_2	Bowtie2 local	97.99%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 a	NS492_2T1_HAT1_KD_24h_ChIP_2	Bowtie2 local	99.10%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 a	NS493_2T1_HAT1_KD_48h_ChIP_2	Bowtie2 local	97.75%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
-	NS494_2T1_HAT1_KD_0h_Input_2	Bowtie2 local	96.52%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS495_2T1_HAT1_KD_24h_Input_2	Bowtie2 local	96.69%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS496_2T1_HAT1_KD_48h_Input_2	Bowtie2 local	97.81%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 a	NS531_HAT1KD_0h_H2AZ_ChIP_3	Bowtie2 local	98.04%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 a	NS532_HAT1KD_0h_H2AZ_ChIP_4	Bowtie2 local	98.24%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 a	NS533_HAT1KD_48h_H2AZ_ChIP_3	Bowtie2 local	98.14%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median

	Sequencing						
Figure	**Description**	**Mapping software**	**Read alignments**	**Reference genome**	**Wig file ws ss**	**Average ws ss**	**Statistics**
Appendix figure 5 a	NS534_HAT1KD_48h_H2AZ_ChIP_4	Bowtie2 local	97.86%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
-	NS535_HAT1KD_0h_H2AZ_Input_3	Bowtie2 local	97.15%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS536_HAT1KD_0h_H2AZ_Input_4	Bowtie2 local	97.31%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS537_HAT1KD_48h_H2AZ_Input_3	Bowtie2 local	97.86%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS538_HAT1KD_48h_H2AZ_Input_4	Bowtie2 local	96.60%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS539_HAT2KD_0h_H2AZ_ChIP_3	Bowtie2 local	95.11%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS540_HAT2KD_0h_H2AZ_ChIP_4	Bowtie2 local	95.45%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS541_HAT2KD_48h_H2AZ_ChIP_3	Bowtie2 local	98.65%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 5 b	NS542_HAT2KD_48h_H2AZ_ChIP_4	Bowtie2 local	98.50%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
-	NS543_HAT2KD_0h_H2AZ_Input_3	Bowtie2 local	94.38%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS544_HAT2KD_0h_H2AZ_Input_4	Bowtie2 local	95.24%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS545_HAT2KD_48h_H2AZ_Input_3	Bowtie2 local	97.52%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS546_HAT2KD_48h_H2AZ_Input_4	Bowtie2 local	95.74%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.1 a; Figure 6.2 a; Figure 6.3 a; Figure 6.4 a	NS470_BF_2T1_TyRPB9_ChIP_1	Bowtie2 local	95.48%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Figure 6.1 a; Figure 6.2 a; Figure 6.3 a; Figure 6.4 a	NS471_BF_2T1_TyRPB9_Input_1	Bowtie2 local	95.55%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
-	NS474_BF_2T1_TyRPB9_ChIP_2	Bowtie2 local	94.08%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
-	NS475_BF_2T1_TyRPB9_Input_2	Bowtie2 local	94.25%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Appendix figure 9 b	NS512_HAT1_KD_TyRPB9_0h_Input_1	Bowtie2 local	97.72%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Appendix figure 9 b	NS513_HAT1_KD_TyRPB9_24h_Input_1	Bowtie2 local	97.68%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Figure 6.1 a; Figure 6.2 a	NS514_HAT1_KD_TyRPB9_48h_Input_1	Bowtie2 local	96.54%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median

	Sequencing						
Figure	**Description**	**Mapping software**	**Read alignments**	**Reference genome**	**Wig files ss**	**Average ws ss**	**Statistics**
Appendix figure 9 a	NS515_HAT2_KD_TyRPB9_0h_Input_1	Bowtie2 local	97.70%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Appendix figure 9 a	NS516_HAT2_KD_TyRPB9_24h_Input_1	Bowtie2 local	98.03%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Figure 6.1 a; Figure 6.2 a	NS517_HAT2_KD_TyRPB9_48h_Input_1	Bowtie2 local	97.40%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Appendix figure 9 b	NS518_HAT1_KD_TyRPB9_0h_ChIP_1	Bowtie2 local	95.54%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 b	NS519_HAT1_KD_TyRPB9_24h_ChIP_1	Bowtie2 local	97.55%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Figure 6.3 a; Figure 6.4 a	NS520_HAT1_KD_TyRPB9_48h_ChIP_1	Bowtie2 local	93.72%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 a	NS521_HAT2_KD_TyRPB9_0h_ChIP_1	Bowtie2 local	90.78%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 a	NS522_HAT2_KD_TyRPB9_24h_ChIP_1	Bowtie2 local	97.02%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Figure 6.3 a; Figure 6.4 a	NS523_HAT2_KD_TyRPB9_48h_ChIP_1	Bowtie2 local	91.75%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 b	NS567_2T1_TyRPB9_HAT1KD_0h_ChIP_2	Bowtie2 local	12.33%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 b	NS568_2T1_TyRPB9_HAT1KD_24h_ChIP_2	Bowtie2 local	12.32%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 b	NS569_2T1_TyRPB9_HAT1KD_48h_ChIP_2	Bowtie2 local	6.83%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 b	NS570_2T1_TyRPB9_HAT1KD_0h_Input_2	Bowtie2 local	85.80%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Appendix figure 9 b	NS571_2T1_TyRPB9_HAT1KD_24h_Input_2	Bowtie2 local	83.29%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Appendix figure 9 b	NS572_2T1_TyRPB9_HAT1KD_48h_Input_2	Bowtie2 local	87.63%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Appendix figure 9 a	NS573_2T1_TyRPB9_HAT2KD_0h_ChIP_2	Bowtie2 local	15.08%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 a	NS574_2T1_TyRPB9_HAT2KD_24h_ChIP_2	Bowtie2 local	8.75%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 a	NS575_2T1_TyRPB9_HAT2KD_48h_ChIP_2	Bowtie2 local	7.88%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Appendix figure 9 a	NS578_2T1_TyRPB9_HAT2KD_0h_Input_2	Bowtie2 local	84.82%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median

	Sequencing						
Figure	**Description**	**Mapping software**	**Read alignments**	**Reference genome**	**Wig file ws ss**	**Average ws ss**	**Statistics**
Appendix figure 9 a	NS579_2T1_TyRPB9_HAT2KD_24h_Input_2	Bowtie2 local	86.00%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Appendix figure 9 a	NS580_2T1_TyRPB9_HAT2KD_48h_Input_2	Bowtie2 local	91.39%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Figure 6.1 b; Figure 6.2 b; Figure 6.3 b, c, d; Figure 6.4 a, c	NS320_ssRNAd_WT1	BWA mem	98.56%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.1 b; Figure 6.2 b; Figure 6.3 b, c, d; Figure 6.4 a, c	NS350_ssRNAd_WT2	BWA mem	99.55%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.1 b; Figure 6.2 b; Figure 6.3 b, c, d; Figure 6.4 a, c	NS351_ssRNAd_WT3	BWA mem	99.35%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.3 b, c, d; Figure 6.4 a, c	NS352_ssRNAd_HAT2KD2T1_48h_1	BWA mem	99.15%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.3 b, c, d; Figure 6.4 a, c	NS353_ssRNAd_HAT2KD2T1_48h_2	BWA mem	99.27%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.3 b, c, d; Figure 6.4 a, c	NS354_ssRNAd_HAT2KD2T1_48h_3	BWA mem	99.16%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.1 b; Figure 6.2 b	NS524_ssRNAseq_HAT1KD2T1_48h_1	BWA mem	97.81%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.1 b; Figure 6.2 b	NS525_ssRNAseq_HAT1KD2T1_48h_2	BWA mem	98.22%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.1 b; Figure 6.2 b	NS526_ssRNAseq_HAT1KD2T1_48h_3	BWA mem	98.77%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 4.1 c	Lowell2005_TyH2AZ	Bowtie2 local solexa-quals,	87.15%	HGAP3_Tb427v9	numerator - ws 5001 ss 1001		

	Sequencing						
Figure	**Description**	**Mapping software**	**Read alignments**	**Reference genome**	**Wig file ws ss**	**Average ws ss**	**Statistics**
		int-quals					
Figure 4.1 a	NS563_SM_TyH2A_ChIP_1	Bowtie2 local	98.15%	HGAP3_Tb427v9	numerator - ws 5001 ss 1001; ws 1 ss 1	ws 501 ss 101	median
-	NS564_SM_TyH2A_ChIP_2	Bowtie2 local	98.08%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
-	NS565_SM_TyH2A_Input_1	Bowtie2 local	96.89%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
-	NS566_SM_TyH2A_Input_2	Bowtie2 local	94.65%	HGAP3_Tb427v9	denominator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.4 a	L1500085-BDF1_KD_0h	Bowtie2 local	98.33%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.4 a	L1500087-BDF1_KD_48h	Bowtie2 local	97.64%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.4 b	L1700062_S7_BDF2_KD_2T1_IP_0h	Bowtie2 local	93.98%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.4 b	L1700064_S9_BDF2_KD_2T1_IP_48h_#2	Bowtie2 local	97.60%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.4 c	L1700066_S11_BDF3KD_2T1_IP_0h_#2	Bowtie2 local	98.20%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.4 c	L1700068_S13_BDF3KD_2T1_IP_48h_#2	Bowtie2 local	92.20%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.4 d	L1500088-BDF4_KD_0h	Bowtie2 local	98.11%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 5.4 d	L1500090-BDF4_KD_48h	Bowtie2 local	96.96%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 e	HAT1_OE_0h	Bowtie2 local	97.99%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 e	ID-003919-NS137_HAT1_OE_IP_48h	Bowtie2 local	97.27%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 f	ID-003532-NS107_HAT2_OE_IP_0_h	Bowtie2 local	97.19%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 f	ID-003910-NS128_HAT2_OE_IP_48h	Bowtie2 local		HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 g	ID-003534-NS109_BDF2_OE_IP_0_h	Bowtie2 local	97.98%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 g	ID-003535-NS110_BDF2_OE_IP_48_h	Bowtie2 local	95.06%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 h	ID-003537-NS112_BDF3_OE_IP_0_h	Bowtie2 local	93.77%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 h	ID-003539-NS114_BDF3_OE_IP_48_h	Bowtie2 local	95.49%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 c	ID-003540-NS115_FACT_KD_IP_0_h	Bowtie2 local	97.55%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 c	ID-003541-NS116_FACT_KD_IP_48_h	Bowtie2 local	75.06%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 d	ID-003542-NS117_ISWI_KD_IP_0_h	Bowtie2 local	97.73%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Appendix figure 8 d	ID-003543-NS118_ISWI_KD_IP_48_h	Bowtie2 local	96.92%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
-	L1700058_S3_WT_I-BET151_IP_0h_#2	Bowtie2 local	97.67%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median

	Sequencing						
Figure	**Description**	**Mapping software**	**Read alignments**	**Reference genome**	**Wig file ws ss**	**Average ws ss**	**Statistics**
-	L1700060_S5_WT_I-BET151_IP_48h_#2	Bowtie2 local	98.54%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median
Figure 6.1 a, b	L1800150_H3_V_Ty_MNase_ChIP_1	Bowtie2 local	96.77%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (numerator)	ws 501 ss 101	median
Figure 6.1 a, b	L1800151_H3_V_Ty_MNase_Input_1	Bowtie2 local	88.18%	HGAP3_Tb427v9	ratio - ws 1 ss 1 (denominator)	ws 501 ss 101	median
Figure 6.3 c	Mueller2018_ATACseq_Wt_20	BWA mem	100%	HGAP3_Tb427v9	numerator - ws 1 ss 1	ws 501 ss 101	median

Extraction of chromatin-associated proteins

The protocol for extraction of chromatin-associated proteins was kindly provided by Joana Faria (David Horn Lab; School of Life Science; University of Dundee). Analysis of chromatin-associated proteins was performed for 2T1 or 2T1_$^{2xTY/2xTY}$*RPB9* cells and 2T1 or 2T1_$^{2xTY/2xTY}$*RPB9* cells, in which HAT1 or HAT2 were depleted over a time period of 48 h. For each sample 4 million cells were harvested at 4 °C and 1.500 x g for 10 min and washed in 1 ml of 1 x TDB (5 mM KCl, 80 mM NaCl, 1 mM $MgSO_4$, 20 mM Na_2HPO_4, 2 mM NaH_2PO_4, 20 mM glucose pH 7.4) at 4°C and 1.500x g for 10 min. The resulting cell pellet was resuspended in 20 µl freshly prepared CSK-buffer (100 mM NaCl, 0.1% Triton X-100, 300 mM Sucrose, 1 mM $MgCl_2$, 1 mM EGTA, 10 mM PIPES (pH 6.8; with NaOH) supplemented with 1 mM Pepstatin A, 5 mM Leupeptin, 100 mM PMSF and 100 mM TLCK) and incubated on ice for 10 min. To separate the soluble (cytosol + nucleosoluble) from the insoluble (chromatin + nuclear matrix) fraction, cell suspension was centrifuged at 2.550 x g and 4 °C for 5 min. The supernatant containing the soluble fraction was separated and the pellet was washed using 20 µl CSK-buffer at 2.550 x g and 4 °C for 5 min. The supernatant from the washing step was separated from the pellet, which contains the insoluble fraction. To 23 µl soluble fraction 7 µl 4 x sample buffer (4 x Laemmli buffer supplemented with 2.5% ß-mercaptoethanol, 1 mM pepstatin A, 5 mM leupeptin, 100 mM PMSF and 100 mM TLCK) were added and the pellet containing the insoluble fraction was resuspended in 30 µl 1 x sample buffer. Proteins were reduced and denatured at 90 °C for 10 min. The samples were shock-frozen in liquid nitrogen and stored at -80 °C. The insoluble and soluble fraction of 2 million cells were analysed by western blot similarly as described before (Siegel et al., 2008). As control, a total protein extract from 2 million cells from 2T1 or 2T1_$^{2xTY/2xTY}$*RPB9* cells and 2T1 or 2T1_$^{2xTY/2xTY}$*RPB9*cells, in which HAT1 or HAT2 were depleted over a time period of 48 h, was used.

Western blot analysis for target protein detection

To evaluate successful protein tagging or to compare protein amounts from different conditions, western blot analysis was conducted using an established protocol (Siegel et al., 2008). Therefore, protein samples were loaded on a 15% SDS-PAGE gel and separated using the Mini-PROTEAN® Tetra Vertical Electrophoresis Cell (Bio-Rad) according to the manufacturer's instructions. For protein transfer the Mini Trans-Blot® Electrophoretic Transfer Cell system (Bio-Rad) was used according to the manufacturer's instructions. The membranes were stained for 5 min at room temperature under constant shaking in Amidoblack staining solution (0.5% amido black in 10% acetic acid) to evaluate the protein transfer. Following destaining in destaining solution (24% isopronaol and 10% acetic acid) for 5 min at room temperature under constant shaking, the membranes were blocked for 1 h in PBS-T (1 x PBS + 0.1% Tween 20) containing 3% BSA. Afterwards the membranes were hybridised with the primary antibody. For detection of H3, the membranes were incubated overnight at 4 °C under constant rotation in 1%

milk solved in 1 x PBS-Tween20 (0.1%) supplemented with the polyclonal H3 rabbit antiserum (1:20,000) (Gassen et al., 2012). For detection of H2A.Z, the membranes were incubated overnight at 4 °C under constant rotation in 1% milk solved in 1 x PBS-Tween20 (0.1%) supplemented with the polyclonal, affinity-purified H2A.Z rabbit antibody (1:750; stock concentration 1.6 µg/µl) (Wedel et al., 2017). For detection of TY-RPB9, the membranes were incubated overnight at 4°C under constant rotation in 1% milk solved in 1 x PBS-Tween20 (0.1%) supplemented with the monoclonal, affinity-purified BB2 mouse antibody (1:750; stock concentration 1 µg/µl) (Bastin et al., 1996). The secondary antibodies Sheep Anti-Mouse IgG Polyclonal Antibody Horseradish Peroxidase (VWR) or Donkey Anti-Rabbit IgG Polyclonal Antibody Horseradish Peroxidase (VWR) were diluted to 1:10,000 in 1% milk solved in 1 x PBS-Tween20 (0.1%). The membranes were incubated in the secondary antibody solution for 1.5 h at room temperature under constant rotation. For signal detection, the Immobilon Western HRP Substrate (Merck) was applied according to the manufacturer's instruction.

Quantification of protein amounts

Quantification of protein amounts from western blots was performed as described by Hossein Davarinejad (Quantifications of Western Blots with ImageJ; http://www.yorku.ca/yisheng/Internal/Protocols/ImageJ.pdf). ImageJ (Schneider et al., 2012) was downloaded from https://imagej.nih.gov/ij/download.html. Western blot quantification was performed under the measurement criteria "Grey Mean Value". Thereby, the signal of H3 was used as loading control. In Image, the individual regions of interests (ROI) of H2A.Z, TY1-RPB9 and H3 were selected using the *rectangle* tool. The H3 signal and background were measured of each lane using the H3 ROI. The H2A.Z signal and background of each lane using the H2A.Z ROI and the TY1-RPB9 signal and background for each lane using the TY1-RPB9 ROI. The inverted value was calculated for each protein and background value by subtracting the measured value from 255. For calculating the net protein value, the inverted background value was subtracted from the inverted protein value. For relative quantification, the ratio of the net H2A.Z value or net TY1-RPB9 value over the net H3 value (loading control) was calculated. The H2A.Z/H3 ratios from the 48 h HAT1- or HAT2-depleted chromatin fractions were compared to the H2A.Z/H3 ratio from 0 h HAT1- or HAT2-depleted chromatin fractions. The TY1-RPB9/H3 ratios from the from the 48 h HAT1- or HAT2-depleted chromatin fractions were compared to the TY1-RPB9/H3 ratio 0 h HAT1- or HAT2-depleted chromatin fractions.

Chapter 3: Genome-wide acetylation pattern in *Trypanosoma brucei*

Summary

Previous studies have shown that trypanosomal core histones are acetylated and methylated at various sites (Janzen et al., 2006a; Mandava et al., 2007). However, due to technical limitations, these analyses could not cover the full-length protein sequence of the four core histones. Additionally, the PTM-patterns of the four histone variants have not been studied. Thus, I performed a global PTM-analysis of histones, isolated from whole cell extracts of *T. brucei* wild type cells, applying a standard acid extraction protocol (Shechter et al., 2007). While other nuclear proteins precipitate in acidic environments, histones are acid-soluble due to their highly basic character (Luger et al., 1997). Given the apparent link of histone acetylation and H2A.Z deposition, I was mainly interested in identifying histone acetylation marks and thus, I applied the previously published mass-spectrometry based method for quantification of lysine-specific acetylation degrees (ElBashir et al., 2015).

This strategy allowed me to identify 157 different histone modifications in *T. brucei* BFs combining the results of different histone isolation approaches and only counting PTMs present in at least three independent experiments. While previous studies mostly detected N-terminal modifications, I could generate a comprehensive dataset of histone modifications covering the protein sequence of all core histones and their variant forms.

Introduction

Post-translational modifications of histones are an important means to regulate chromatin compaction and influence many biological processes that require DNA as template (Kouzarides, 2007). Each histone contains more than 60 residues that can be modified. Histone PTMs, such as acetylation or phosphorylation, can directly alter the nucleosomal architecture and/or function as binding platforms for other chromatin modifying proteins, such as remodeler, chaperones and transcriptional regulators. Interestingly, many proteins that mediate specific histone modifications also contain domains for recognising them and this can further enhance the chromatin modification state (Zhang et al., 2015).

Even though trypanosomal histone proteins are highly divergent (Thatcher and Gorovsky, 1994; Alsford and Horn, 2004), various evolutionary conserved PTMs have been identified. For example, the N-terminal tail of H4 is extensively acetylated in *T. brucei* as observed in other organisms and H3K76, the trypanosomal homologue to H3K79 of other eukaryotes, can be mono-, di- and trimethylated as well. These marks have been found to be implicated in replication initiation (Gassen et al., 2012), cell cycle progression and differentiation (Frederiks et al., 2010) similar to H3K79 methyl marks.

The genome of trypanosomes encodes many ‘writers’ and ‘readers’ for histone PTMs including six different histone acetyltransferases (Ivens et al., 2005; Kawahara et al., 2008), seven histone

deacetylases, around 25 histone methyltransferases and four putative demethylases (Ivens et al., 2005; Figueiredo et al., 2009), five bromodomain proteins (Ivens et al., 2005; Siegel et al., 2009b; Schulz et al., 2015) and one potential chromodomain (Ivens et al., 2005). It is likely that most of them play a regulatory role in trypanosomes, given the importance of histone PTMs for many fundamental biological processes and epigenetic inheritance in other organisms. However, to be able to study their physiological function in *T. brucei*, it is essential to know which histone PTMs are present and how they are distributed along the genome.

In this chapter, I describe how I generated a comprehensive dataset of histone PTMs present in *T. brucei* combining different histone isolation strategies with **F**ragment **I**on **P**atchwork **Quant**ification (FIPQuant), a mass-spectrometry approach for quantification of site-specific lysine acetylation degrees and identification of lysine-specific methylation patterns.

Results

Fragment ion patchwork quantification

To utilise a eukaryotic genome that is packaged into chromatin as template, regulating chromatin opening by histone acetylation is fundamental for the gene expression machinery. Programming events of histone acetylation accompany the development of specific gene expression patterns, that characterise healthy or diseased cells for example (Fischle et al., 2003). Traditionally, modification-specific antibodies were used in chromatin immunoprecipitation analyses to study the appearance or disappearance of acetyl marks under different signalling conditions (Kouzarides, 2007). However, such studies strongly depend on the target-specificity of the antibody. Even though, modification-specific antibodies are typically raised to bind exclusively to one acetyl mark, in many studies, it could be observed that they can recognise multiply acetylated peptides or are unable to differentiate combinatorial motifs (Fuchs et al., 2011; Rothbart et al., 2012). Moreover, antibody-based detection strategies cannot obtain a comprehensive picture of all acetyl marks characterizing a specific signalling condition.

Thus, mass-spectrometry (MS) based detection of lysine acetylation is usually the approach of choice to circumvent such issues. Many MS-based techniques, such as SILAC (Bonenfant et al., 2007) or label-free quantification (Lee et al., 2008; Lundby et al., 2012), have been developed, which allow the relative quantification of acetylation sites. While these methods only measure the acetylation fold-change between different sample, site-specific degrees of lysine acetylation cannot be accurately determined. The acetylation degree of a specific lysine is simply determined by comparing the fraction of acetylated peptides with the fraction of unacetylated peptides.

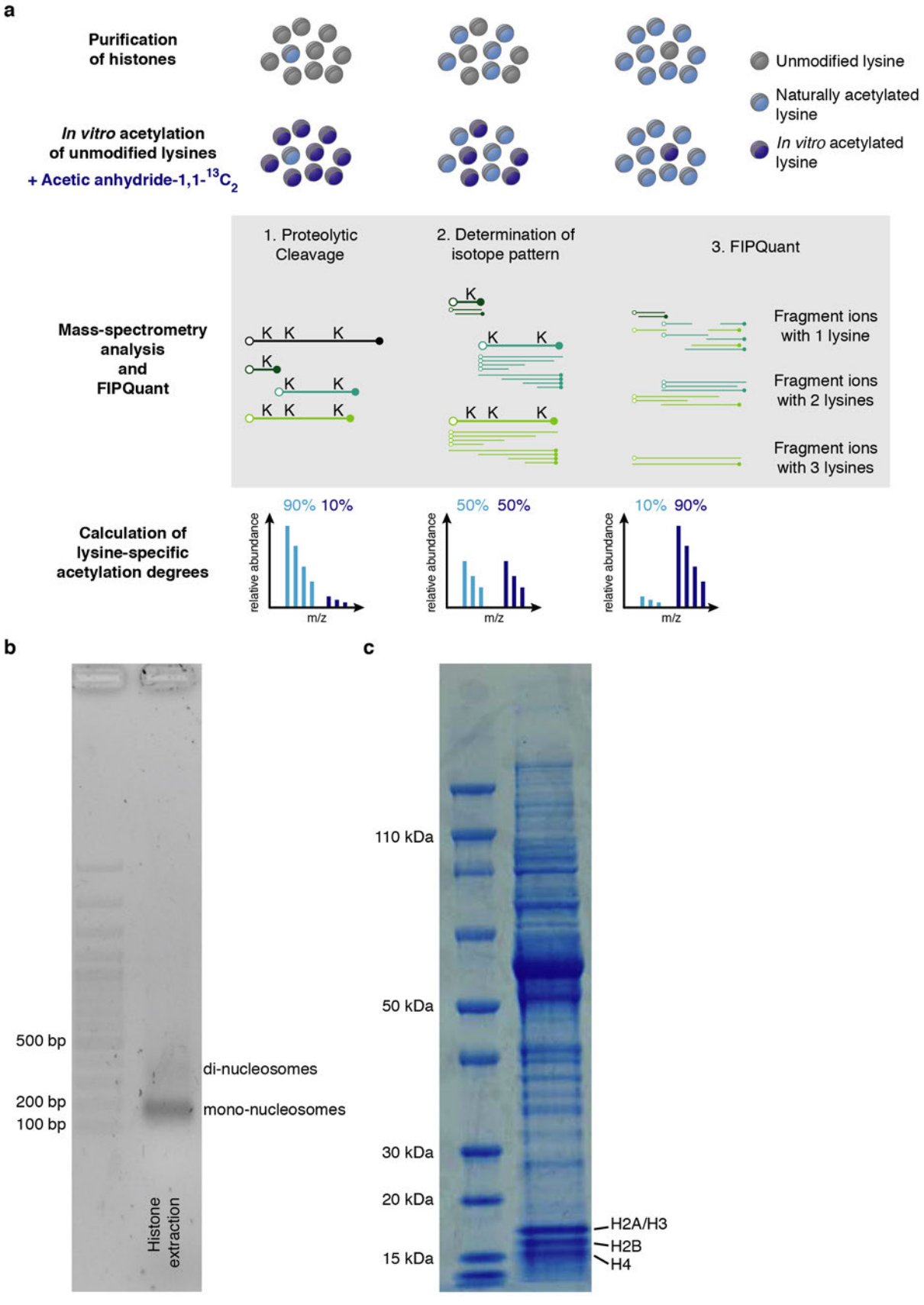

a
Purification of histones
In vitro acetylation of unmodified lysines
+ Acetic anhydride-1,1-13C2
Unmodified lysine
Naturally acetylated lysine
In vitro acetylated lysine
1. Proteolytic Cleavage
2. Determination of isotope pattern
3. FIPQuant
Mass-spectrometry analysis and FIPQuant
Fragment ions with 1 lysine
Fragment ions with 2 lysines
Fragment ions with 3 lysines
Calculation of lysine-specific acetylation degrees
90% 10%
50% 50%
10% 90%
relative abundance
m/z
b
500 bp
200 bp
100 bp
di-nucleosomes
mono-nucleosomes
Histone extraction
c
110 kDa
50 kDa
30 kDa
20 kDa
15 kDa
H2A/H3
H2B
H4

Figure 3.1. FIPQuant allows quantification of lysine-specific acetylation degrees.
a Outline of Fragment Ion Patchwork Quantification (FIPQuant) methodology. Following histone purification, unmodified lysines were in-gel-^{13}C-acetylated and digested in multiple fragments using low-specificity proteases. The isotope pattern of each fragment was determined by mass-spectrometry analysis and the site-specific degree of lysine acetylation was quantified using FIPQuant. **b** Evaluation of mononucleosome preparation. Agarose-gel of DNA extracted from the nucleosome-containing supernatant shows that mononucleosomes were enriched. Since over-digestion had to be avoided, a minor fraction of di-nucleosomes needs to be detectable. **c** Mononucleosome preparation and acid-extraction allow extraction of all histones. Coomassie-gel of whole histone extract shows the purification of all canonical histones. The minor fraction of histone variants is not visible due to the gel-resolution.

However, this is challenging because proteolytic cleavage and ionisation efficiencies of acetylated and non-acetylated peptides are different with the consequence that they cannot be directly compared anymore.

However, peptides can be modified *in vitro* by chemical acetylation of unmodified and monomethylated lysines yielding in physicochemical properties similar to naturally acetylated peptides (Smith et al., 2003b; Smith, 2005; Hersman et al., 2012; Fiedler et al., 2013; Baeza et al., 2015). Recently, D_3-acetyl derivatisation using deuterated acetic anhydride has been successfully applied in combination with a targeted MS-approach to determine site-specific acetylation of histones following depletion of different HATs and HDACs (Feller et al., 2015). However, adding a deuterated mass tag causes also a retention time shift during liquid chromatography (LC) thereby biasing the MS-analysis. To overcome this problem, the authors introduced a LC-MS-correction factor calculated from measurements of synthetic peptides containing different combinatorial motifs of histone acetyl- and methyl marks, which were previously identified (Feller et al., 2015). While this strategy is very useful for analysing known PTM motifs under different conditions, it still harbours the risk to miss unknown PTMs or PTM patterns for which no synthetic peptides were integrated.

Thus, I applied FIPQuant for determination of site-specific levels of lysine acetylation to analyse histone PTMs of trypanosomes in an unbiased manner. This strategy is based on converting all unmodified amino groups to acetyl-amino groups with a mass tag ($^{13}C_1$) using stable, isotope-labelled acetic anhydride for chemical acetyl derivatisation on the protein level (ElBashir et al., 2015). Thereby, the physicochemical properties stay similar between the naturally acetylated and the *in vitro* acetylated peptide and consequently, the proteolytic cleavage and ionisation efficiencies during the LC-MS-workflow stay comparable.

For FIPQuant (**Figure 3.1 a**), intact histones are modified by in-gel-$^{13}C_1$-acetyl-derivatisation using acetic anhydride-1,1-$^{13}C_2$ changing all initially unmodified/monomethylated to “heavy-acetylated” lysines. To obtain complete coverage of histone protein sequences, low-specificity proteases, such as elastase, papain and thermolysin, are used for proteolytic cleavage yielding in higher numbers of different peptides compared to sequence-specific proteases, such as trypsin. For calculation of lysine-specific acetylation levels, peptides with up to three lysines were considered. During the first MS-step, the full isotope pattern of an acetylated peptide is determined. The full isotope pattern is a mixture of isotope patterns derived from different isotopologues containing naturally or heavy acetylated lysines. FIPQuant

is performed during the second MS-step consisting of three succeeding rounds of quantification. During the first quantification round, experimentally determined isotope patterns from fragmented ions with a single acetylation site are compared to a mixture of corresponding, theoretically determined isotope patterns. The theoretical isotope patterns are calculated from one 'heavy-acetyl' and one natural-acetyl fragment ion covering ~10,000 different heavy-to-natural proportions. The site-specific acetylation degree is calculated by determining the fractions of 'heavy' and natural acetyl marks in total number of acetylated fragment ions. The median acetylation degree determined in the first quantification round is then used to obtain the theoretical isotope patterns of fragment ions with two acetylation sites and to calculate the acetylation degree of fragment ions using the experimentally determined isotope patterns with two acetylation sites. In case both acetylation sites could be quantified in the first round, the second quantification round is also performed for both sites. To analyse fragment ions with three acetylation sites correspondingly to the second quantification round, the theoretical isotope patterns were calculated using for two of the three acetylation sites the median degrees obtained in the second round. During FIPQuant, a large number of fragment ions derived from numerous, overlapping peptides are analysed, resulting in a robust and highly accurate quantification of a site-specific lysine acetylation degree. In addition, site-specific methylation marks can be semi-quantitatively determined using peptide counting. To identify PTMs, such as phosphorylation, the data from the FIPQuant analysis was further examined using the peptide identification software PEAKS studio X (Zhang et al., 2012).

To generate a comprehensive list of acetylation and methylation marks present on histones of *T. brucei*, I isolated mononucleosomes from BF *T. brucei* wild type (WT) cells (**Figure 3.1 b**), extracted the histone proteins from the mononucleosome-containing supernatant using a standard acid extraction protocol (**Figure 3.1 c**) (Shechter et al., 2007) and performed FIPQuant analysis. To identify the PTMs of the histone variants H2A.Z and H2B.V, I applied an immunoprecipitation-based histone isolation method, which will be described in the following chapter (see Characterisation of TSS-chromatin modifications), because the coverage of H2A.Z and H2B.V was too low in the whole histone isolates (see **Appendix figure 1**).

Post-translational modifications of H2A and its variant H2A.Z

The core histone H2A

While the histone-fold domain of the *T. brucei* H2A protein shares sequence similarity to other H2A sequences, the sequences of the N-terminal tail and the C-terminal domain are extremely different suggesting that homologues histone H2A PTMs were absent (**Figure 3.2 a**). Indeed, the C-terminus of H2A was highly acetylated in *T. brucei*, as previously identified (Janzen et al., 2006a; Mandava et al., 2007). Such an acetylation pattern has not been described in other eukaryotes, except for *T. cruzi*, a close relative of *T. brucei* (de Jesus et al., 2016; Picchi et al., 2017). In most eukaryotes, the C-terminus

of H2A is marked by ubiquitination of K119, which is important for the DNA damage response (Khare et al., 2012) but not characterised by acetyl- and methyl marks (Zhao and Garcia, 2015).

In addition to already described acetyl and methyl marks of H2A in *T. brucei* wild type cells, I detected various undescribed methyl marks (**Figure 3.2 b**), such as the monomethylation marks of K8 and K9 and the trimethyllation marks of K4, K68, and K115 using FIPQuant. Moreover, combining the results from different histone isolation strategies and only counting PTMs present in at least three experiments, I identified additional methyl marks including A1me2 (N-terminal α-amino group), K12me3, K75me1 and various methyl marks at the C-terminal lysines of H2A in addition to the acetyl marks (**Figure 3.2c**). Open PTM search using PEAKS confirmed that H2A was phosphorylated at T2 (Urbaniak et al., 2013), and revealed an additional phosphorylation at S132.

Taken together, combining FIPQuant, peptide counting and PEAKS analysis, I could confirm 11 of the 12 known H2A PTMs missing the low-abundant acetyl mark on K4 (1%; (Mandava et al., 2007)) and identify 13 previously undescribed PTMs.

The histone variant H2A.Z

In most eukaryotes, histone H2A possess the highest number of histone variants. For example, thus far, 19 different H2A variants are described in *H. sapiens* (Khare et al., 2012). In *T. brucei*, however, only H2A.Z is present, which is the most abundant H2A variant present in a majority of eukaryotes. In *T. brucei*, H2A.Z shared only ~40% sequence identity with canonical H2A (**Figure 3.2 d**). Both proteins differed in the length of the N-terminal tail, which was 40 aa longer in H2A.Z, and the sequences of the histone-fold and C-terminal domains. Sequence alignments revealed, that H2A.Z in *T. brucei* shares a higher sequence similarity to H2A.Z proteins (~55%) than to its canonical counterpart ((Lowell et al., 2005) and **Figure 3.2 e**). Similar to canonical H2A, trypanosomal H2A.Z differed to other H2A.Z proteins mainly in the N-terminal tail that was around 50 aa longer than for example the human H2A.Z. As all histones, H2A.Z can be post-translationally modified and the function of different H2A.Z PTMs has already been studied in *S. cerevisiae*, *H. sapiens* and *T. thermophila* (Giaimo et al., 2019).

Using FIPQuant, I could identify that also *T. brucei* H2A.Z was marked by various post-translational modifications. The two lysines (K31 and K35) of the N-terminal tail and the seven adjacent lysines (K43, K45, K47, K49, K51, K55, K59) were extensively acetylated from (40-90%). Additionally, the N-terminal serine was acetylated to ~5% (**Figure 3.2 f**). A similar acetylation pattern of H2A.Z has also been described in *T. cruzi*. Four lysines (K43, K49, K51, and K55) might have homologues in other eukaryotes and these lysines have been shown to be acetylated. For example, K49ac of *T. brucei* might be a homologue to K8ac in *S. cerevisiae* (Keogh et al., 2006) and K10ac in *T. thermophila* (Ren and Gorovsky, 2001).

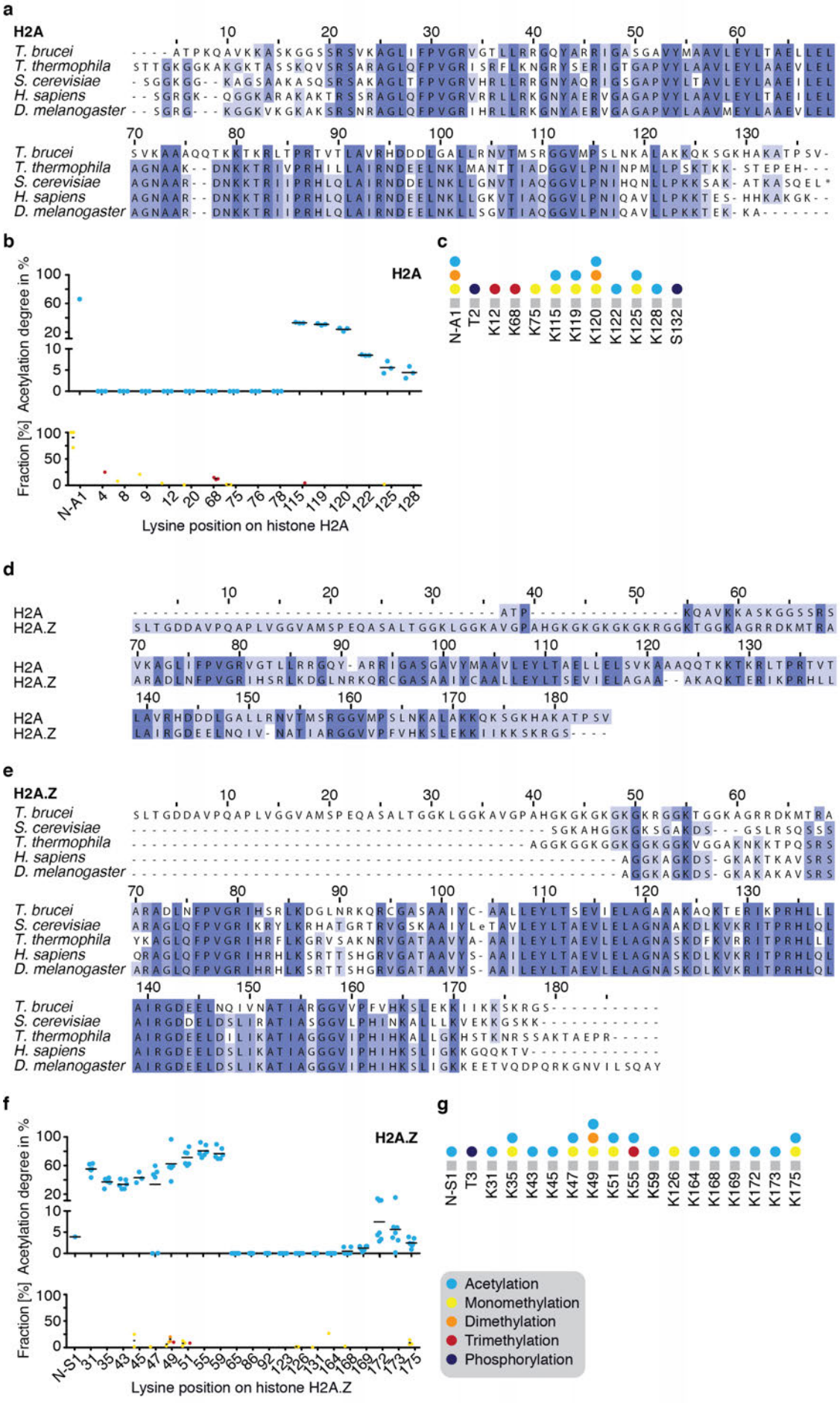

a
H2A
T. brucei
T. thermophila
S. cerevisiae
H. sapiens
D. melanogaster
b
H2A
Acetylation degree in %
Fraction [%]
Lysine position on histone H2A
c
N-A1
T2
K12
K68
K75
K115
K119
K120
K122
K125
K128
S132
d
H2A
H2A.Z
e
H2A.Z
T. brucei
S. cerevisiae
T. thermophila
H. sapiens
D. melanogaster
f
H2A.Z
Acetylation degree in %
Fraction [%]
Lysine position on histone H2A.Z
g
N-S1
T3
K31
K35
K43
K45
K47
K49
K51
K55
K59
K126
K164
K168
K169
K172
K173
K175
Acetylation
Monomethylation
Dimethylation
Trimethylation
Phosphorylation

Figure 3.2. The variant H2A.Z contains many N-terminal acetyl marks in contrast to canonical H2A.
a H2A of *T. brucei* contains a conserved histone-fold domain. Alignment of H2A protein sequences from T. brucei, *T. thermophila*, *S. cerevisiae*, H. sapiens and D. melanogaster shows that the histone-fold domains share many conserved residues (dark blue) and residues with similar properties (light blue). In contrast, the N-terminal and C-terminal domain of trypanosomal H2A only shares little sequence identity with the other H2A proteins. The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **b** Trypanosomes' H2A carries many different N-terminal methyl marks. Lysine-specific acetylation degrees (blue) are shown for core histone H2A and were determined by FIPQuant using whole histone extracts of WT cells (upper panel). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (lower panel). **c** Summary of all acetylation, methylation and phosphorylation marks present on trypanosomal H2A detected in this study. **d** H2A.Z contains a long N-terminal domain in contrast to H2A in T. brucei. The alignment of the histone variant H2A.Z and its canonical counterpart H2A shows that the N-terminus of H2A.Z is 40 aa longer than the N-terminus of H2A. The histone-fold and the C-terminal domains share many residues with similar properties (light blue) and some fully conserved residues (dark blue). The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **e** H2A.Z of T. brucei contains a conserved histone-fold domain. Alignment of H2A.Z protein sequences from T. brucei, S. cerevisiae, T. thermophila, H. sapiens and D. melanogaster shows that the histone-fold domains share many conserved residues (dark blue) and residues with similar properties (light blue). In contrast to the other H2A.Z proteins, trypanosomes' H2A.Z contains a long N-terminal stretch. The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **f** The N-terminus of trypanosomes' H2A.Z is hyperacetylated. Lysine-specific acetylation degrees are shown for histone variant H2A.Z (blue) and were determined by FIPQuant using immunoprecipitation of TSS-nucleosomes (upper panel). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (lower panel). **g** Summary of all acetylation, methylation and phosphorylation marks present on trypanosomal H2A.Z detected in this study.

K55ac might be a homologue to K7ac in *H. sapiens* and K16ac in *T. termophila* (Ren and Gorovsky, 2001). Interestingly, K55ac might also be a homologue to K14ac in *S. cerevisiae*, which is exclusively enriched at active promoters (Keogh et al., 2006). Moreover, the C-terminal lysines of H2A.Z were also acetylated. Such an acetylation pattern is also present on *T. cruzi* H2A.Z. Next to the acetylation marks, I also could detect various mono-, di- and trimethyl marks. Most lysines were modified either by acetylation or monomethylation. However, K49 harboured all three methylation states next the acetylation mark and K51 was acetylated, mono- or trimethylated. Combining the results from different histone isolation strategies and only counting PTMs present in at least three experiments, I identified 16 acetyl marks, 6 mono-, 1 di- and 1 trimethyl marks present on H2A.Z. Using PEAKs analysis, I could detect that T3 is phosphorylated (**Figure 3.2 g**).

My results showed that H2A.Z was extensively acetylated in *T. brucei* and some of these marks might have direct homologues in other eukaryotes. Since H2A.Zac might be more important for transcription than unmodified H2A.Z, I decided to study the role of H2A.Zac in *T. brucei*. The corresponding results are described in chapter five (Role of histone acetylation for TSS-specific chromatin) and six (Chromatin composition at TSSs plays a critical role for transcription).

Post-translational modifications of H2B and its variant H2B.V

The core histone H2B

Similar to H2A, trypanosomal H2B had the highest sequence identity to H2B proteins of other organisms in its histone-fold domain (**Figure 3.3 a**). The N-terminal tail and C-terminal domain were different.

However, in contrast to H2A, the H2B protein of *T. brucei* was rather short compared to other H2B proteins and consisted of only 112 aa. In contrast, H2B of *S. cerevisiae* is 131 aa long and in *H. sapiens* H2B consist of 167 aa. Generally, H2B is not frequently post-translationally modified in eukaryotes. The most prominent PTMs of H2B is ubiquitination at a C-terminal lysine, which is important for transcriptional activation (Weake and Workman, 2008). H2B ubiquitination has not been identified in *T. brucei*. While others showed that H2B can be acetylated at K4, K12 and K16 to minor levels (~1%) (Janzen et al., 2006a; Mandava et al., 2007), I could only detect acetylation at the N-terminal A1 to 25%. Moreover, A1 was mono- or dimethylated, as described previously (Janzen et al., 2006a; Mandava et al., 2007) (**Figure 3.3 b**). Combining the results from different histone isolation strategies and only counting PTMs present in at least three experiments, various previously undescribed H2B methyl marks could be identified (**Figure 3.3 c**).

Taken together, I could observe that H2B was modified rather by methylation than by acetylation marks. Thus far, no homologous lysine methyl marks were found in *T. cruzi* or higher eukaryotes suggesting that the lysine methyl marks of H2B are *T. brucei*-specific. But, in both *T. brucei* and *T. cruzi*, H2B carried a phosphorylation mark at T2 (**Figure 3.3 c** and (Urbaniak et al., 2013; Picchi et al., 2017)).

The histone variant H2B.V

Similar to H2A.Z, H2B.V had a long N-terminal stretch compared to canonical H2B in *T. brucei* and both histones shared only ~36% sequence identity (**Figure 3.3 d**). H2B.V dimerises exclusively with H2A.Z and heterodimers with H2A have not been detected in *T. brucei* (Lowell et al., 2005). Consequently, H2B.V has been found enriched at TSSs (Siegel et al., 2009b). In general, H2B proteins are rather invariant compared to H2A and H3 (Talbert and Henikoff, 2010). For example, in *H. sapiens*, only 11 H2B variants with unknown function have been identified (Albig et al., 1999; Bonenfant et al., 2006). Interestingly, the Apicomplexan parasite *Plasmodium falciparum* contains the H2B variant H2B.Z, which dimerises as well with H2A.Z and is enriched at promoters (Hoeijmakers et al., 2013) similar to H2A.Z/H2B.V-nucleosomes in trypanosomes. Alignment of *T. brucei* H2B.V to the human H2B variant H2BA and H2B.Z of *P. falciparum* showed that H2B.V shared only ~36% sequence identity with them (**Figure 3.3 e**) but ~96% sequence identity with H2B.V from *T. cruzi* suggesting that the type of H2B.V protein is trypanosome-specific.

In contrast to variants of H2A, very little is known about histone H2B variants and even less is known about their post-translational modifications. Using, FIPQuant, I could detect that H2B.V is extensively acetylated at its N-terminal (**Figure 3.3 f**) and some of these marks are found in *T. cruzi*, too (Picchi et al., 2017). However, in *T. cruzi* the N-terminal P1 is not modified, which is the case for H2B.V of *T. brucei*. There, P1 was found to be mono- or dimethylated. Overall, combining the results of different histone isolation strategies and only counting PTMs found in at least three experiments, six acetyl-, four monomethyl and four dimethyl marks were detected on H2B.V in *T. brucei* (**Figure 3.3 g**).

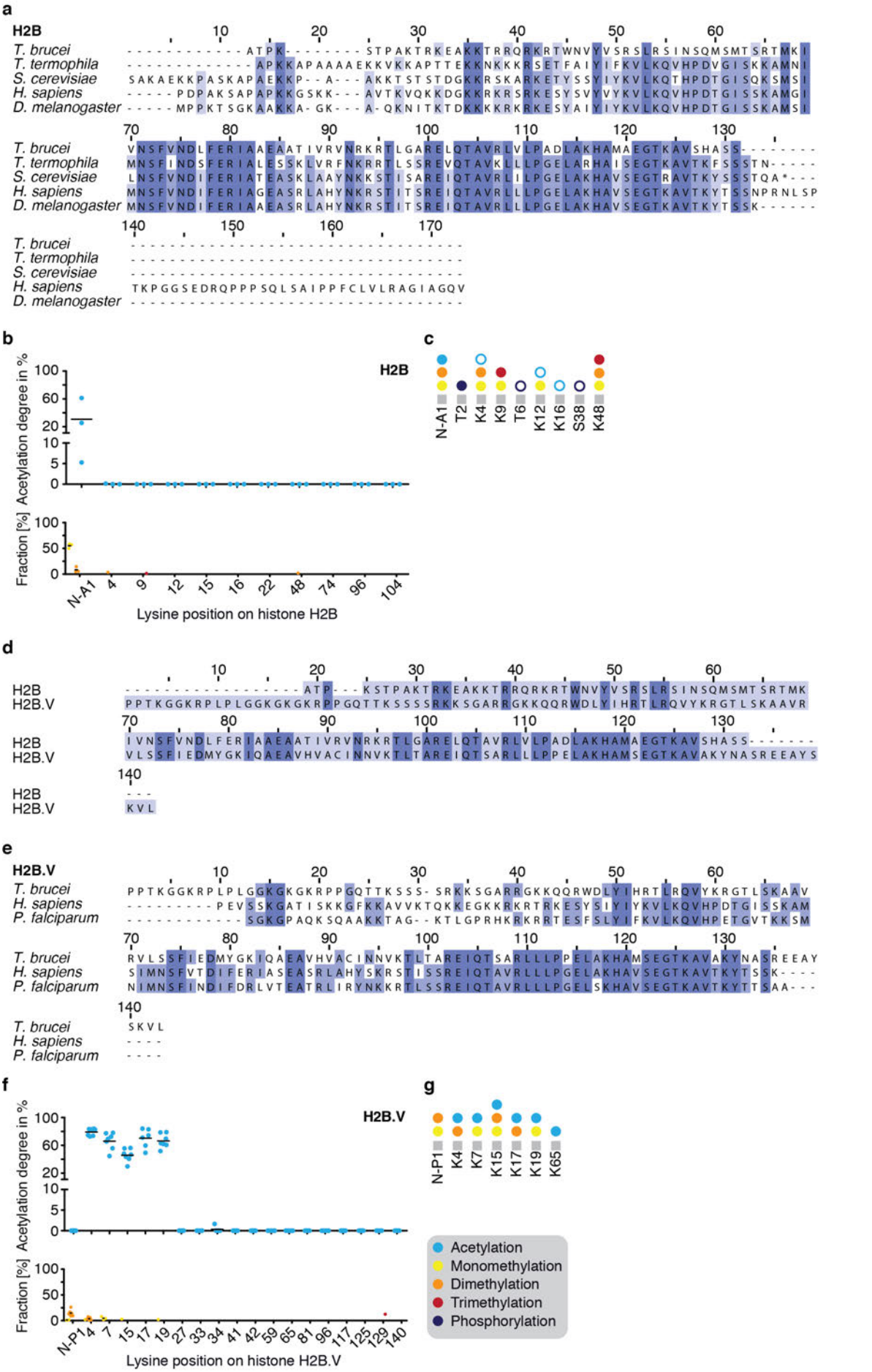

a
H2B
T. brucei
T. termophila
S. cerevisiae
H. sapiens
D. melanogaster
b
H2B
Acetylation degree in %
Fraction [%]
N-A1
Lysine position on histone H2B
c
H2B
N-A1
T2
K4
K9
T6
K12
K16
S38
K48
d
H2B
H2B.V
e
H2B.V
T. brucei
H. sapiens
P. falciparum
f
H2B.V
Acetylation degree in %
Fraction [%]
N-P1
Lysine position on histone H2B.V
g
H2B.V
N-P1
K4
K7
K15
K17
K19
K65
Acetylation
Monomethylation
Dimethylation
Trimethylation
Phosphorylation

Figure 3.3. Trypanosomes' H2B.V and H2B are not extensively modified.
a H2B of T. brucei contains many conserved protein residues. Alignment of H2B protein sequences from T. brucei, T. thermophila, S. cerevisiae, H. sapiens and D. melanogaster shows that the N-terminal, the histone-fold and the C-terminal domains share many conserved residues (dark blue) and residues with similar properties (light blue). The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **b** Trypanosomes' H2B is acetylated and methylated at its N-terminal α-amino group. Lysine-specific acetylation degrees (blue) are shown for core histone H2B and were determined by FIPQuant using whole histone extracts of WT cells (upper panel). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (lower panel). **c** Summary of all acetylation, methylation and phosphorylation marks present on trypanosomal H2B detected in this study. **d** H2B.V shares many residues with canonical H2B in T. brucei. The alignment of the histone variant H2B.V and its canonical counterpart H2B shows that the N-terminus of H2B.V is 19 aa longer than the N-terminus of H2B. The histone-fold and the C-terminal domains share many residues with similar properties (light blue) and some fully conserved residues (dark blue). The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **e** H2B.V of T. brucei contains a conserved C-terminal domain. Alignment of H2B.V protein sequences from T. brucei, H. sapiens and P. falciparim shows that the histone-fold domains share many conserved residues (dark blue) and residues with similar properties (light blue). The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **f** The N-terminus of trypanosomes' H2B.V carries many acetyl marks. Lysine-specific acetylation degrees are shown for histone variant H2B.V (blue) and were determined by FIPQuant using immunoprecipitation of TSS-nucleosomes (upper panel). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (lower panel). **g** Summary of all acetylation, methylation and phosphorylation marks present on trypanosomal H2B.V detected in this study.

Taken together, my data revealed that H2B.V, similar to H2A.Z, was hyperacetylated at its N-terminus. Overall, H2B.V was not marked by many different PTMs, as observed already for canonical H2B.

Post-translational modifications of H3 and its variant H3.V

The histone H3

Histone H3 belongs to the core component of a canonical nucleosome and its protein sequence has been highly conserved during evolution. Among different model organisms, including yeast, mouse, fruit fly and humans, H3 proteins are 80% - 90% identical (**Figure 3.4 a**). In contrast, the H3 protein of *T. brucei* shares only ~55% sequence identity with them. Of all core histones, H3 harbours the maximum number of post-translational modifications in eukaryotes (Zhao and Garcia, 2015) and they were reported to be involved in many different biological processes, such as transcription activation, chromosome condensation or heterochromatin formation.

Acetylation and methylation of H3 are frequent and well characterised in many eukaryotes (Kouzarides, 2007). It has not been possible to determine the modification status of H3 in *T. brucei* since its N-terminus had been blocked for Edman-degradation and MS-based sequencing of trypsin-digested H3 did not cover the complete peptide sequence (Janzen et al., 2006a; Mandava et al., 2007). Using FIPQuant, I could identify that *T. brucei* H3 was barely acetylated, which is untypical for a eukaryote. Analysing whole histone extracts from wild type cells, I found only the N-terminal S1 of H3 to be acetylated (~5%) (**Figure 3.4 b**). Combining the results of different histone isolation strategies and only counting PTMs found in at least three experiments, I observed that each lysine of H3 was marked by at least one modification (**Figure 3.4 c**). Nevertheless, only one lysine acetyl mark on K23 could be

detected on H3 of *T. brucei*. Otherwise, lysines of H3 were marked by methylation. K17, K19, K32 and K119 were only monomethylated, whereas K16 and K61 harboured also dimethyl marks and K23 was di- or trimethylated. All three methylation states were found on K4, K10, K11, K20. H3K4, H3K23, H3K76 and H3K119 were identified as direct homologues to lysines in other eukaryotes, where their function has already been extensively studied. For example, in *H. sapiens*, H3K4me1-3 and H3K23ac are involved in transcriptional activation (Ng et al., 2003; Zhao and Garcia, 2015). Since in *T. brucei* nucleosomes at TSSs are enriched with H3K4me3 (Wright et al., 2010), this mark has been associated to active transcription, too. However, nothing is known about the genomic distribution and the role of H3K23ac in *T. brucei*.

Taken together, I was able to analyse for the first time the modification status of all lysines of H3 and could determine that in *T. brucei* the major H3 was only acetylated on K23, which is also present in other eukaryotes. However, H3 was marked by various mono-, di- and trimethyl marks in *T. brucei*.

The histone variant H3.V

Like H2A, H3 has usually a high number of variant proteins. They can differ from their canonical counterpart only in a few amino acids or in the complete protein sequence. For example, the H3 variant CenpA that marks centromeric chromatin (Malik and Henikoff, 2003), is 35 aa longer than canonical H3 but the variant H3.3 found at promoters, enhancers and gene bodies and associated to active transcription (Mito et al., 2005; Jin et al., 2009) differs only in five amino acids that do not alter the nucleosomal structure (Tachiwana et al., 2011). In contrast to other eukaryotes, *T. brucei* has only one H3.V, which shares only 45% sequence identity with H3 and has a remarkably different N-terminal tail compared to the major H3 (**Figure 3.4 d** and (Lowell and Cross, 2004)). While H3.V is to ~60% identical to CenpA, ChIP-seq data could not relate H3.V to centromeric regions (Müller et al., 2018). H3.V is not homologue to other eukaryotic H3 variants and shares only ~61% sequence identity to *T. cruzi* H3.V. Since in *T. brucei* H3.V is mainly found at subtelomeric regions (Lowell and Cross, 2004; Müller et al., 2018) and its knock out causes derepression of silent *VSG* expression sites (Müller et al., 2018), it is likely that it diverged away from other H3.V proteins during evolution to take over this specific role.

FIPQuant analysis revealed that in a whole histone extracts from wild type cells H3.V, similar to major H3, did not harbour any acetyl marks (**Figure 3.4 e**). Peptide counting showed that H3.V was methylated at K14 (mono-, di- and trimethylation) and at K82 (mono- and trimethylation). By taking into account all experiments I performed for this study and only considering PTMs present in at least three experiments, three additional lysine methyl marks could be identified at K4 and K82 (**Figure 3.4 f**). The N-terminal A1 was not modified on H3.V from wild type histone extracts but acetylated, mono- and dimethylated on H3.V from other histone isolation procedures.

My data showed that, in *T. brucei*, H3.V harboured only a few PTMs.

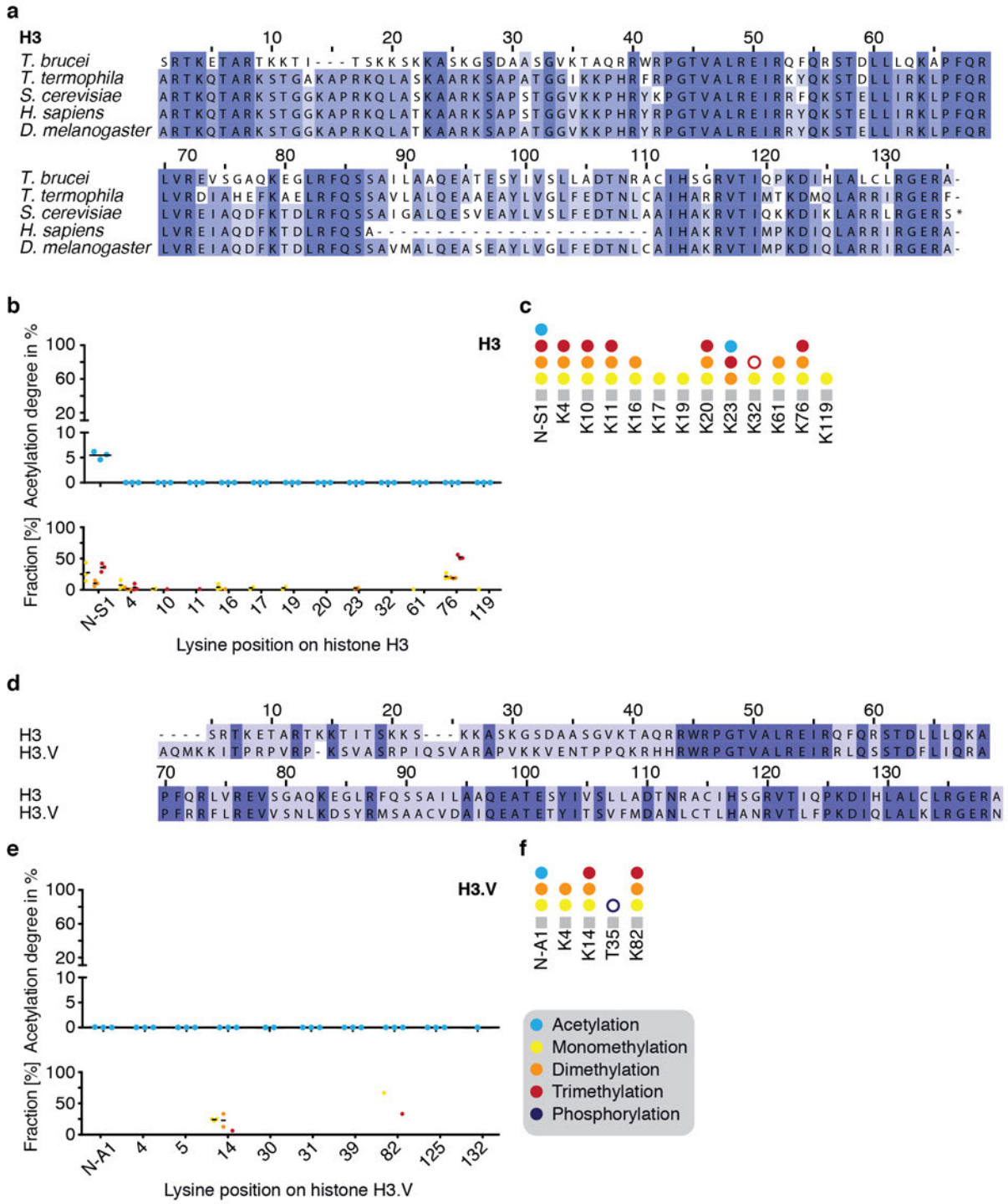


Figure 3.4. Trypanosomes' H3.V is a highly divergent histone protein.
a H3 of T. brucei contains many conserved protein residues. Alignment of H3 protein sequences from T. brucei, T. thermophila, S. cerevisiae, H. sapiens and D. melanogaster shows that the N-terminal, the histone-fold and the C-terminal domains share many conserved residues (dark blue and blue) and residues with similar properties (light blue). The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **b** Trypanosomes' H3 does not carry acetyl marks. Lysine-specific acetylation degrees (blue) are shown for core histone H3 and were determined by FIPQuant using whole histone extracts of WT cells (upper panel). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (lower panel). **c** Summary of all acetylation, methylation and phosphorylation marks present on trypanosomal H3 detected in this study. **d** H3.V shares many residues with canonical H3 in T. brucei. The alignment of the histone variant H3.V and its canonical counterpart H3 shows that the N-terminus of H3.V exhibit little sequence identities. The histone-fold and the C-terminal domains share many

identical residues (dark blue) and residues with similar properties (light blue). The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **e** H3.V only contains few PTMs. Lysine-specific acetylation degrees are shown for histone variant H3.V (blue) and were determined by FIPQuant using whole histone extracts of WT cells (upper panel). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (lower panel). **f** Summary of all acetylation, methylation and phosphorylation marks present on trypanosomal H3.V detected in this study.

Post-translational modifications of H4 and its variant H4V

The histone H4

Acetylation of H4 is involved in many different processes and is associated to active transcription, for example it restricts chromatin folding to a higher-order structure and leaves loosely packed and easily accessible chromatin (Dorigo et al., 2003). Moreover, H4 acetylation is important for transcription initiation by recruiting different transcription factors or mediating chromatin remodeling (Shia et al., 2006; Altaf et al., 2010; Draker et al., 2012). Among all core histones in *T. brucei*, H4 has the highest homology to other H4 proteins and shares around ~58% sequence identity (**Figure 3.5 a**). Additionally, the N-terminal tail of H4 is also extensively modified in *T. brucei*. H4 is acetylated at K2, K4, K5 and K10, with K2ac being the most abundant marks (~80%) (Janzen et al., 2006a; Mandava et al., 2007; ElBashir et al., 2015). The N-terminal A1 and K2 are monomethylated and all three methylation states are present on K17 and K18. Using FIPQuant, I could confirm the previously identified H4 PTMs and, moreover, I identified that A1 was acetylated and K14 can be mono-, di- and tri methylated (**Figure 3.5 b** and ElBashir et al., 2015).

Interestingly, adding the information from other experiments, H4 was also mono-, di- and trimethylated at A1, K2, K4, K5 and K10 and K57 and K77 were monomethylated (**Figure 3.5 c**). These marks were not present on H4 in whole histone extracts from wild type cells suggesting they might be localised at specific genomic regions, similar to H4K10ac, which is enriched at TSSs (Siegel et al., 2009b). H4 of *T. brucei* is ~97% identical to H4 of *T. cruzi* and most of the PTMs identified on H4 are also present in *T. cruzi* (da Cunha et al., 2006; Picchi et al., 2017) suggesting that they might be involved in general regulatory processes.

The histone variant H4.V

Thus far, only a few H4 variants have been identified. The H4G protein in humans (Long et al., 2019) and the H4 variant H4.V in *T. brucei* (Berriman et al., 2005) and *T. cruzi* (Ivens et al., 2005). In *T. brucei*, H4.V is almost identical to major H4 (~85%) (Siegel et al., 2009b). H4 and H4.V differed only in 15 aa, which were distributed throughout the whole protein sequence (**Figure 3.5 d**). In addition to the high sequence similarity, H4 and H4.V shared also a similar modification pattern (**Figure 3.5 e-f**). The major acetyl mark was found at H4.VK4, similar to H4 (**Figure 3.5 e**). Furthermore, K2 was monomethylated and all three methylation states were found on K14, K17 and K18. However, K93 of H4.V carried a

monomethyl mark, which was not present on canonical H4. Interestingly, while the N-terminal A1 of H4 was highly modified, A1 of H4.V was only monomethylated (**Figure 3.5 f**). Overall, H4.V carried less PTMs than its canonical counterpart.

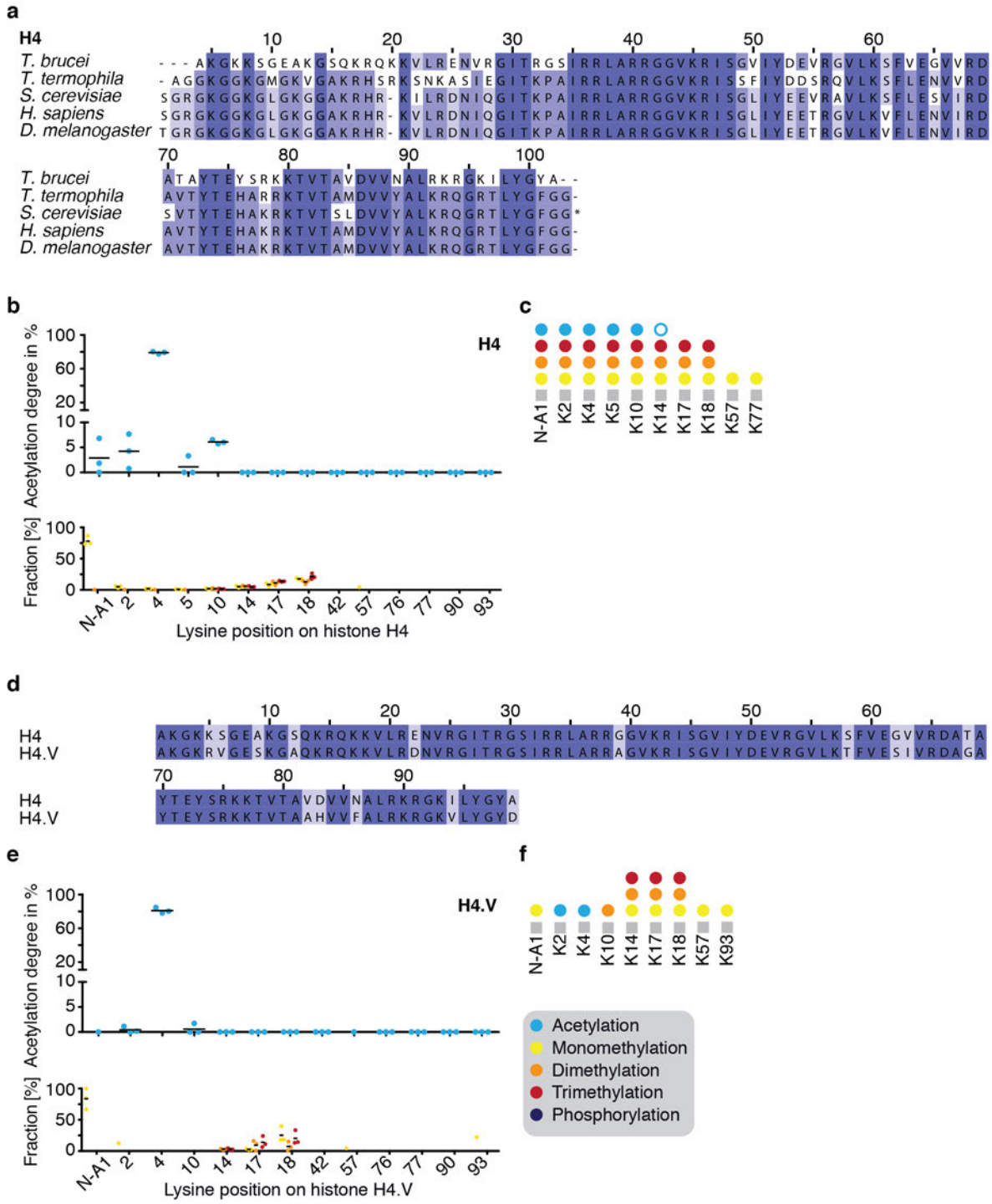


Figure 3.5. Trypanosomes' H4 contains many conserved residues.
a H4 of T. brucei contains many conserved protein residues. Alignment of H4 protein sequences from T. brucei, T. thermophila, S. cerevisiae, H. sapiens and D. melanogaster shows that the N-terminal, the histone-fold and the C-terminal domains share many conserved residues (dark blue and blue) and residues with similar properties (light blue). The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **b** The N-terminus of H4 contains many

conserved acetyl marks. Lysine-specific acetylation degrees (blue) are shown for core histone H4 and were determined by FIPQuant using whole histone extracts of WT cells (upper panel). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (lower panel). **c** Summary of all acetylation, methylation and phosphorylation marks present on trypanosomal H4 detected in this study. **d** H4.V differs only in 15 aa from canonical H4 in T. brucei. The alignment of the histone variant H4.V and its canonical counterpart H4 shows that their protein sequences share many identical residues (dark blue and blue) and residues with similar properties (light blue). The alignment was conducted with Clustal Omega using standard parameters (Sievers et al., 2011). **e** H4.V is highly acetylated at K4 similar to canonical H4. Lysine-specific acetylation degrees are shown for histone variant H4.V (blue) and were determined by FIPQuant using whole histone extracts of WT cells (upper panel). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (lower panel). **f** Summary of all acetylation, methylation and phosphorylation marks present on trypanosomal H4.V detected in this study.

Conclusion

Previous studies have already revealed that histones in trypanosomes are highly modified (Janzen et al., 2006a; Mandava et al., 2007; ElBashir et al., 2015). In this chapter, I presented a comprehensive list of PTMs (acetylation, methylation, phosphorylation) present on trypanosomal histones covering for the first time the protein sequence of all four canonical histones and histone variants.

Combining FIPQuant, peptide counting and PEAKS analysis, I could confirm most of the previously detected PTMs, except for some low-abundant marks at H2A, H2B and H3. Moreover, I identified in total 157 PTMs, of which 126 PTMs have not been reported before in *T. brucei*, including prominent methyl marks on the N-terminal tail of H3.

Chapter 4: Characterisation of TSS-chromatin modifications

Summary

While genome-wide chromatin profiling approaches, such as ChIP-seq, have thus far been the method of choice to analyse the genomic enrichment sites of a chromatin marks, including histone acetyl marks, they require the knowledge of the existence of the mark as a pre-requisite. Thus, they do not allow to comprehensively characterise the chromatin structure of a specific genomic site in a completely unbiased manner. Understanding of how a specific chromatin structure is established at a specific genomic site would be possible if using the reverse approach were possible, e. g. the identification of all chromatin marks associated with a specific genomic locus. Thus far, such locus-specific chromatin analyses have been successfully performed only for chromatin from long repetitive regions, such as telomeric repeats (Wierer and Mann, 2016).

To be able to characterise the TSS-specific chromatin in *T. brucei* and to identify the differences between TSS- and non-TSS-chromatin, I firstly established an approach that allowed enrichment of locus-specific chromatin. Therefore, I could take advantage of the mutually exclusive distribution of H2A.Z (TSS-nucleosomes) and H2A (non-TSS-nucleosomes), that allowed me to isolate TSS-chromatin and non-TSS chromatin in an almost completely unbiased manner. Combining the approach with FIPQuant analysis, I could observe that TSS-nucleosomes were highly acetylated in contrast to non-TSS-nucleosomes. At TSSs, the N-terminal lysines of H4 were hyperacetylated, in contrast to non-TSSs H4. Only the prominent acetyl mark at H4K4 was present at both sites. Moreover, the histone variants H2A.Z and H2B.V were extensively acetylated at their N-terminus. The modification pattern of H2A.Z and H2B.V are described and discussed in the previous chapter.

Introduction

Using genome-wide chromatin profiling approaches, such as ChIP-seq, the genomic distribution of various chromatin marks, including histone PTMs, histone variants and chromatin modifiers, has already been identified (Li et al., 2007). For example, in all studied eukaryotes, TSSs are marked by acetylated histones and by deposition of the histone variant H2A.Z (Haberle and Stark, 2018; Giaimo et al., 2019), which probably mediate a more open chromatin structure to facilitate the binding of the transcription machinery. However, ChIP-seq studies rely on knowledge whether a chromatin mark exists and the availability of antibody to detect it. Furthermore, the understanding of how distinct marks are deposited to various sites, and of the proteins which are involved, requires to know all chromatin features of a specific genomic site. Even though, locus-specific chromatin analyses could provide a comprehensive picture of the enriched chromatin marks, they were only successfully applied to analyse repetitive sequences such as telomeric regions (Wierer and Mann, 2016).

Despite their divergence from other eukaryotes, trypanosomes possess an extensively acetylated H4 tail and a histone H2A variant at TSSs (Siegel et al., 2009b). Since commercially available antibodies cannot be used for genome-wide chromatin profiling in trypanosomes, H4K10ac and H3K4me3 are until now the only histone PTMs which have been confirmed to be enriched at TSSs.

To be able to analyse the functional link between histone acetylation, H2A.Z deposition and transcription initiation in a eukaryotic organism, I needed to identify the TSS-specific chromatin marks first. In this chapter, I describe the method I applied to enrich TSS-specific chromatin and how the locus-specific chromatin isolation was a pre-requisite to identify 58 PTMs that are specific for TSS-chromatin in *T. brucei*.

Results

Establishment of TSS-specific chromatin immunoprecipitation

While heterotypic nucleosomes containing H2A.Z together with canonical H2A have been found, for example in yeast (Luk et al., 2010), previous studies have showed that in *T. brucei* H2A.Z does not co-immunoprecipitate with H2A (Lowell et al., 2005) and that H2A.Z is 300-fold enriched at TSSs over other genomic sites (Siegel et al., 2009b). Both observations point to very low levels or even to an absence of heterotypic nucleosomes in trypanosomes.

To validate the previous findings, I analysed the genome-wide distribution of TY1-tagged H2A and repeated the ChIP-seq analysis of H2A.Z. In doing so, I could observe that H2A was strongly depleted at TSSs in contrast to H2A.Z suggesting that the genomic distribution of the histone and its variant is to a large extent mutually exclusive (**Figure 4.1 a**). To test if the genomic distribution pattern of H2A.Z and H2A would allow enrichment of either TSS-nucleosomes, those containing H2A.Z, and non-TSS nucleosomes, those containing H2A, I isolated either H2A.Z-containing or TY1-tagged H2A-containing nucleosomes from the cell line that allows overexpression of TY1-tagged H2A and analysed the co-immunoprecipitated histone by western blotting. No H2A.Z co-immunoprecipitated with TY1-tagged H2A, as reported before (Lowell et al., 2005), and no H2A signal could be detected following H2A.Z co-immunoprecipitation (**Figure 4.1 b**). Altogether, the findings described above suggest that H2A and its variant do not co-localise or co-exist in the same nucleosome in *T. brucei* that should allow to isolate specifically TSS-nucleosomes and non-TSS-nucleosomes.

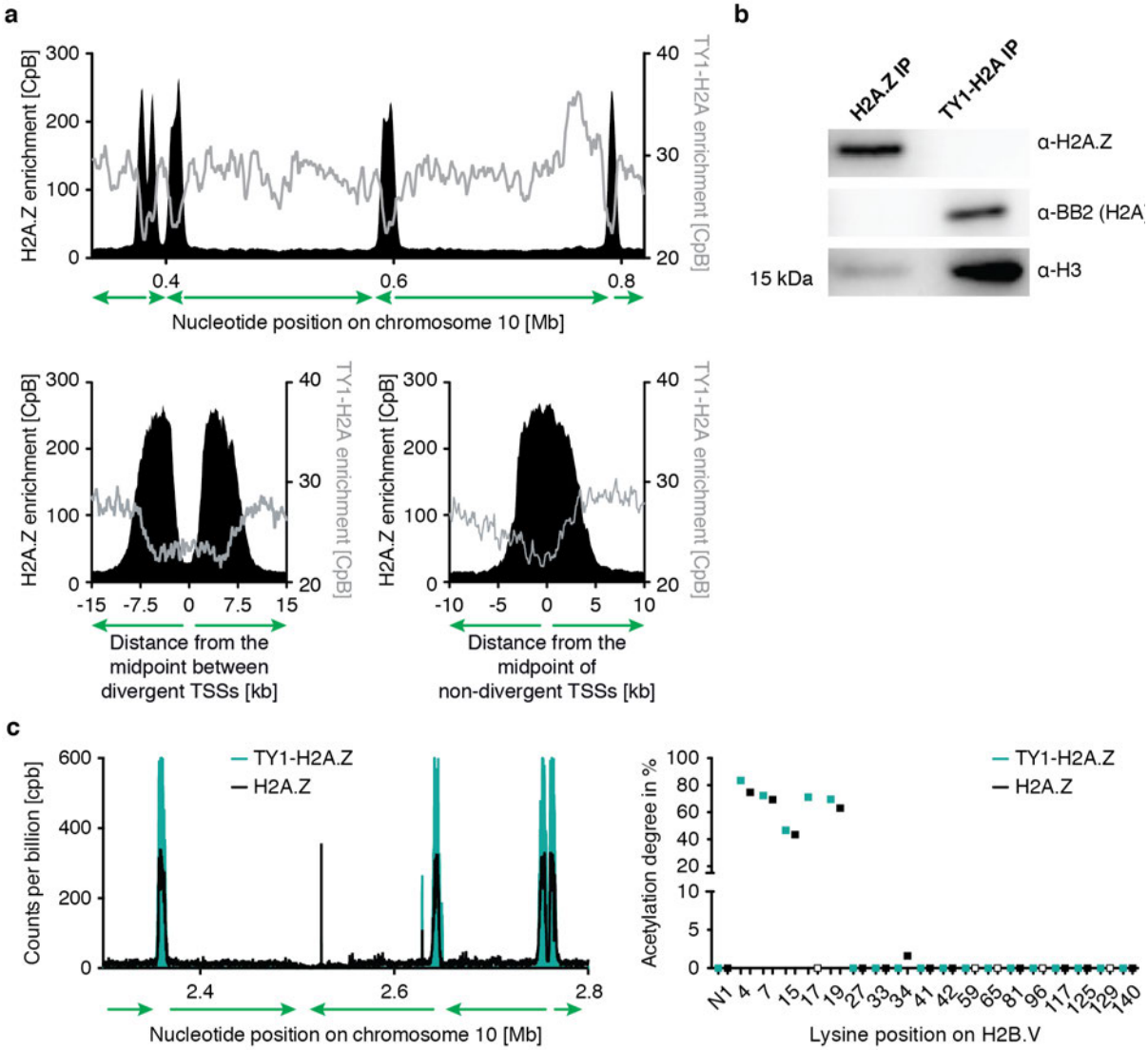


Figure 4.1. H2A.Z replaces H2A in nucleosomes at transcription start sites.
a TSSs are depleted of H2A-containing nucleosomes. MNase-ChIP-seq data of the histone variant H2A.Z (black) and of the overexpressed and TY-tagged canonical histone H2A (grey) are shown across a representative region on chromosome 10 (upper panel) and averaged across divergent (n=81) and non-divergent (n=49) TSSs (lower panel). **b** Western blot of proteins co-immunoprecipitated with H2A.Z (left lane) or TY1-H2A (right lane). Mononucleosomes were extracted from a cell line, which allows the inducible overexpression of ectopic TY1-H2A and were immunoprecipitated using either a custom-made antibody against untagged H2A.Z (n=1) or the BB2 antibody recognising TY1-tagged H2A (n=1). Immunoprecipitated material was then examined by western blotting for the presence of TY1-H2A or untagged H2A.Z along with histone H3 as loading control. **c** The N-terminal TY-tag does not affect the nucleosomal composition. MNase-ChIP-seq data of untagged histone variant H2A.Z (black) from wild type cells (WT) and of the overexpressed and TY-tagged histone variant H2A.Z (teal) are shown across a representative region on chromosome 10 (left panel). Lysine-specific acetylation degrees are shown of H2B.V from TSS-nucleosomes enriched by immunoprecipitation of untagged H2A.Z (black; n=1) and overexpressed and TY-tagged H2A.Z (teal; n=6). The acetylation degrees of K17, K59, K65, K96, K129 could not be determined for TSS-nucleosomes enriched by immunoprecipitation of untagged H2A.Z.

For the experiments described above, I used a previously generated cell line expressing only exogenous TY1-tagged H2A.Z from an *rDNA* locus. To confirm that overexpression of a TY1-tagged histone did not influence its genomic localisation, I compared ChIP-seq analyses of TSS-nucleosomes isolated from *T. brucei* BF wild type cells using a polyclonal antibody (Wedel et al., 2017) detecting untagged H2A.Z or from cells overexpressing TY1-tagged H2A.Z using an antibody that recognises the TY1-tag (Bastin et al., 1996). I could observe that neither the TY1-tag nor overexpression did affect the genomic localisation of H2A.Z. Moreover, I obtained identical PTM patterns of H2B.V analysing Co-IPs of TY1-tagged or untagged H2A.Z by FIPQuant (**Figure 4.1 c**).

Since H2B.V dimerises exclusively with H2A.Z in trypanosomes, the findings of the ChIP-seq and FIPQuant analysis confirmed that overexpression of TY1-taggged H2A.Z did not affect the genomic distribution nor the PTM pattern of TSS-nucleosomes. Since H2A exists in an array of multiple genes on chromosome 7 (Berriman et al., 2005), I used a previously published cell line overexpressing a TY-tagged H2A from a random rRNA locus in addition to endogenous, untagged H2A (Lowell et al., 2005) to enrich for non-TSS-nucleosomes. To verify that isolation of TY1-tagged H2A-nucleosomes allows to recover the histones, I analysed the proteins co-immunoprecipitated with TY1-tagged H2A on a denaturing SDS page and observed robust co-immunoprecipitation of H2B, H3 and H4 (**Figure 4.2 a**).

Therefore, I extracted uncrosslinked mononucleosomes from both cell lines and immunoprecipitated either non-TSS- or TSS-nucleosomes using an antibody against the TY-tag (**Figure 4.2 a**). Despite the presence of untagged H2A, IP of non-TSS-nucleosome pulling on TY-H2A co-immunoprecipitated successfully H2B, H3 and H4 (**Figure 4.2 b**), whereas IP of TSS-nucleosomes pulling on TY-H2A.Z reconstituted H2B.V, H3 and H4 (**Figure 4.2 b**).

The observed findings strongly indicated that taking advantage of the mutually exclusive distribution of H2A and its histone variant H2A.Z, would allow me to specifically enrich TSS-nucleosomes and non-TSS-nucleosomes and thus, to perform locus-specific chromatin isolation from a non-repetitive region.

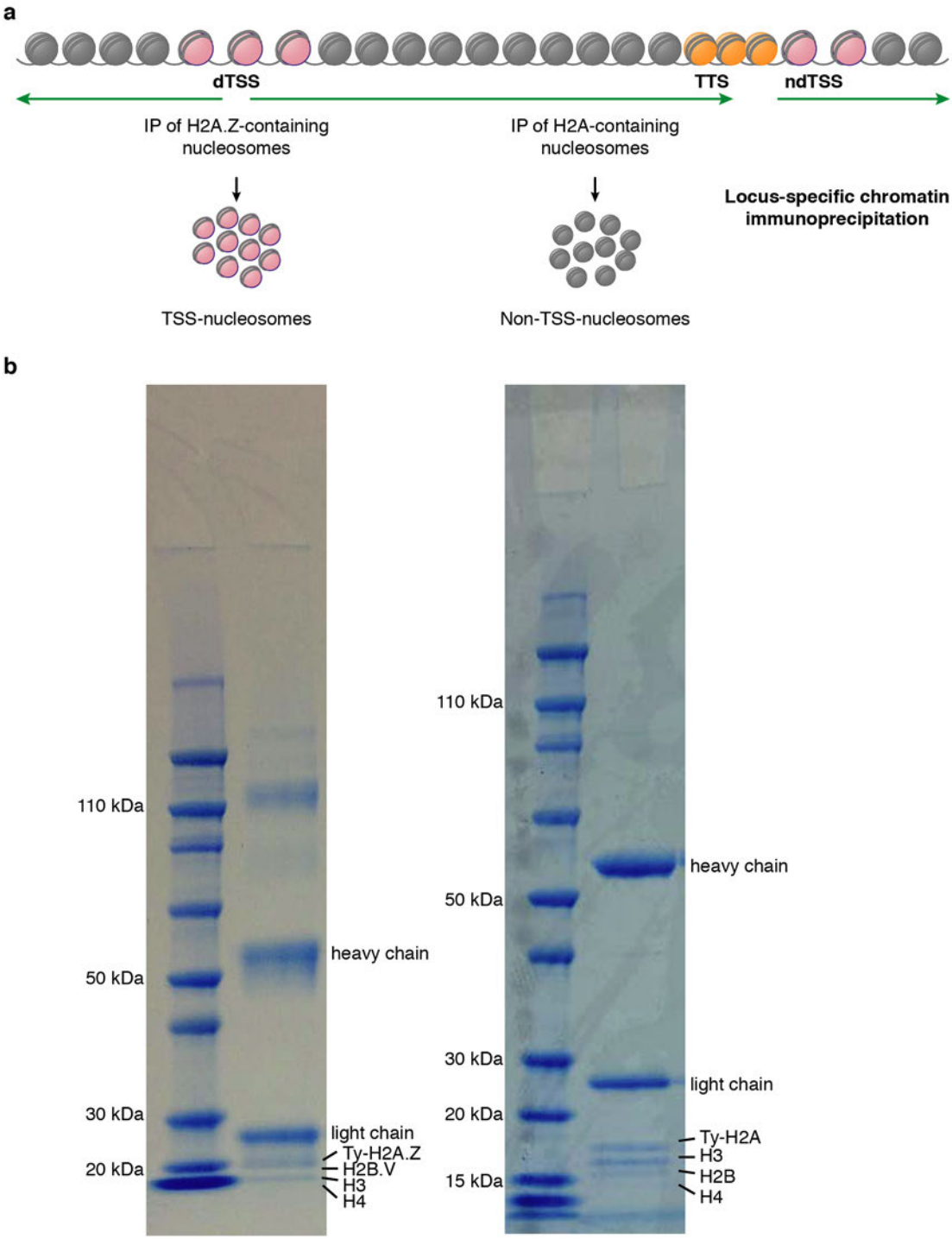


Figure 4.2. Locus-specific chromatin isolation allows for specific enrichment of TSS- or non-TSS-nucleosomes.
a Chromatin organisation of T. brucei. By taking advantage of the mutually exclusive distribution of H2A.Z and H2A, nucleosomes from TSSs and non-TSSs, respectively, can be specifically isolated. **b** Immunoprecipitation of TY-H2A.Z and TY-H2A allow co-purification of all other histones of TSS- and non-TSS-nucleosomes, respectively. Coomassie gel of enriched TSS-nucleosomes shows the co-immunoprecipitation of H2B.V, H3 and H4 with TY-H2A.Z (left panel). Coomassie gel of enriched non-TSS-nucleosomes shows the co-immunoprecipitation of H2B, H3 and H4 with TY-H2A.

Lysine modifications at transcription start regions

Using ChIP-seq analyses, it has already been shown that H3K4me3 (Wright et al., 2010) and H4K10ac are enriched at TSSs in *T. brucei* (Siegel et al., 2009b). While I could already observe that the TSS-specific histone variants are extensively acetylated (see Genome-wide acetylation pattern in *Trypanosoma brucei*), combining my approach to enrich specifically for TSS-nucleosomes or non-TSS-nucleosomes with FIPQuant analysis should allow me to detect site-specific differences in the PTM patterns of H3 and H4 as well.

In doing so, I confirmed the previously described PTMs H3K4me3 (Wright et al., 2010) and H4K10ac (Siegel et al., 2009b) to be enriched at TSSs. These findings validated my approach, additionally. Overall, I could identify 58 different histone PTMs that were specific for TSSs and observed that both H3 and H4 exhibited a different modification pattern in TSS-nucleosomes than in non-TSS-nucleosomes. Since the PTM analysis of wild type histone extracts has already suggested that H3 acetyl marks are relatively rare in *T. brucei*, in contrast to other eukaryotes, it was not surprising that H3 was not extensively acetylated in TSS- nor non-TSS-nucleosomes as well. H3K23 was acetylated to ~5% at TSSs and notably was the only lysine acetyl mark found on H3 in this study (**Figure 4.3 a**). Comparison of the methylation pattern revealed that at TSSs H3K4 could also be dimethylated. In addition, dimethyl and trimethyl marks on H3K10 and H3K11 were enriched at TSSs. In contrast, monomethylation of H3K16 and H3K19 and acetylation of the N-terminal S1 (~7%) were only found at non-TSS-nucleosomes (**Figure 4.3 a**).

Analysing the site-specific PTM pattern of H4, I could assign the previously identified H4 acetyl marks at K2 and K5 (Janzen et al., 2006a; Mandava et al., 2007; ElBashir et al., 2015) to be TSS-specific as H4K10ac (Siegel et al., 2009b). At TSSs, H4K2 was acetylated to ~70%, H4K5 to ~40% and H4K10ac to 90% (**Figure 4.3 b**). Only the high abundant H4K4 acetyl-mark (~80%) was found on both TSS- and non-TSS-nucleosomes. While TSS- and non-TSS-nucleosomes showed particularly distinctive differences in acetylation, the differences in methyl marks were relatively modest. I detected the methyl marks at A1, K2, K5, K10, K14, K17 and K18 on H4 of TSS- and non-TSS-nucleosomes. Only H4K4me1 and H4K10me3 were non-TSS-specific, whereas H4K2me1 was exclusive found in TSS-nucleosomes (**Figure 4.3 b**).

In summary, combining locus-specific chromatin isolation and FIPQuant, I could dissect the PTM patterns of TSS-nucleosomes and non-TSS-nucleosomes. My findings clearly demonstrated that the N-terminus of H4 is extensively acetylated and that this acetylation pattern is TSS-specific in *T. brucei*.

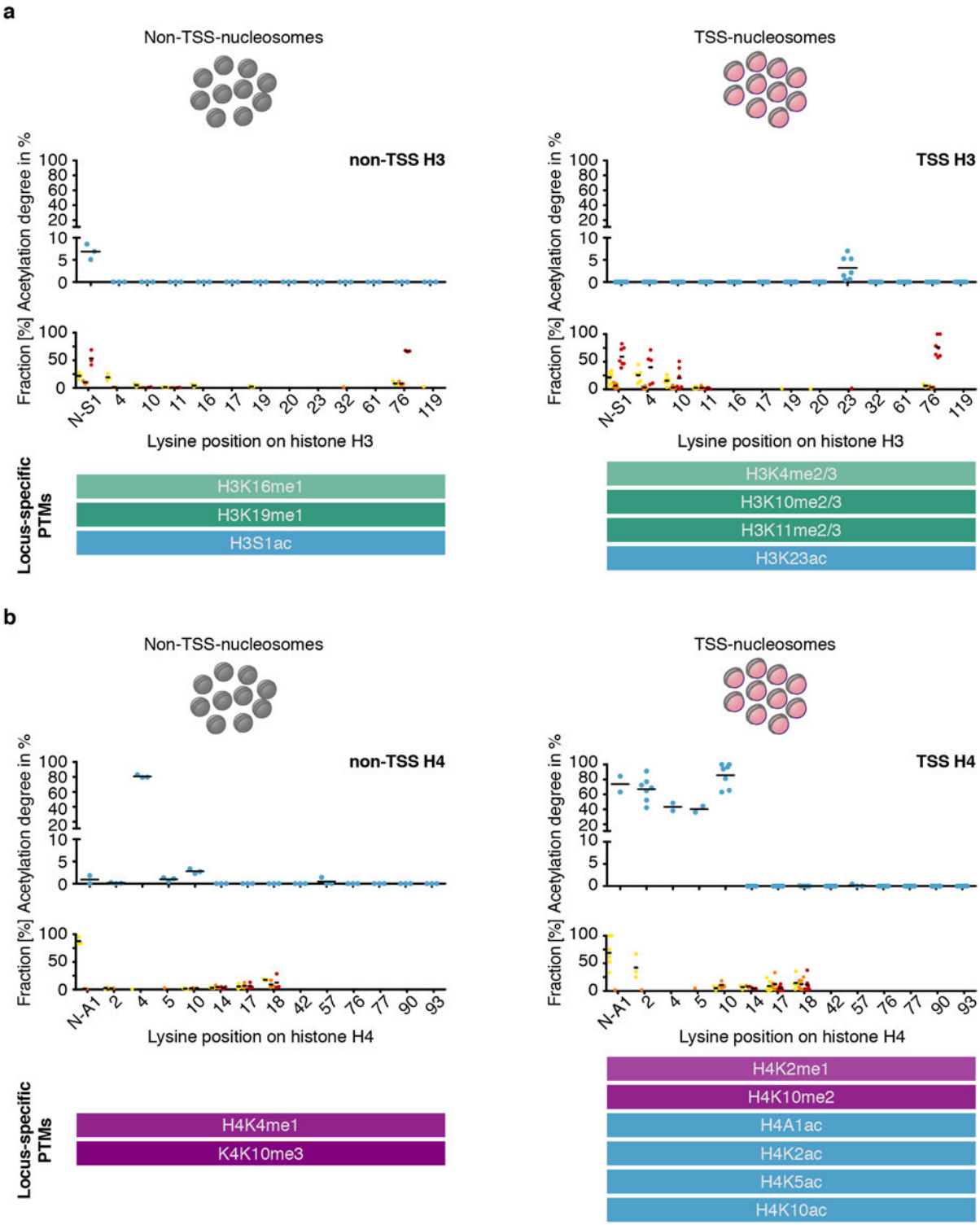


Figure 4.3. TSS-nucleosomes are enriched with specific acetyl- and methyl marks.
a H3K23 is specifically acetylated at TSSs. The PTM-patterns of H3 from non-TSS- (left panel) and TSS-nucleosomes (right panel) are shown. The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (middle panels). The lower panels show the summaries of locus-specific histone modifications H3. **b** The N-terminal acetyl marks are TSS-specific. The PTM-patterns of H4 from non-TSS- (left panel) and TSS-nucleosomes (right panel) are shown. The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. The fractions of mono- (yellow), di- (orange) and trimethylated (red) peptides identified in each replicate are plotted. The black bar indicates the average fraction of mono-, di- and trimethylated peptides (middle panels). The lower panels show the summaries of locus-specific histone modifications of H4.

Conclusions

Thus far, ChIP-seq analysis is the method of choice to identify the genomic localisation of a specific chromatin mark. However, to fully understand the biological role of specific mark and/or the protein it associates with, it is important to determine all PTMs in a specific nucleosomal context or isolated from a specific genomic locus.

In this chapter, I described a strategy combining locus-specific chromatin isolation with FIPQuant, which allowed me to identify TSS-specific and non-TSS-specific PTMs in *T. brucei*. Overall, TSS-nucleosomes were marked extensively by histone acetylation. However, while the histone variants H2A.Z and H2B.V and the canonical histone H4 were hyperacetylated at TSSs, mostly at their N-terminal tail, H3 was only acetylated at a single. Notably, H3K23ac at TSSs was the only H3 acetyl mark found in this study. But, H3 was marked by site-specific methylation patterns. Taken together, I identified 58 TSS-specific histone PTMS and demonstrated that also in an early diverged eukaryote, that lacks any regulation of RNAPII transcription initiation, histone acetylation dominates sites of transcription initiation.

Chapter 5: Role of histone acetylation for TSS-specific chromatin

Summary

Acetylation of H4 and H2A.Z represents a conserved feature of TSSs in many eukaryotes (Haberle and Stark, 2018; Giaimo et al., 2019). It has been shown that at TSSs, H4 acetylation stimulates targeted H2A.Z deposition and that acetylated H2A.Z destabilises the nucleosome to facilitate transcription (Venkatesh and Workman, 2015). Moreover, acetylation of H2A.Z can stimulate its own deposition (Keogh et al., 2006). Most complexes that mediate the exchange the H2A/ H2B-dimer with an H2A.Z/ H2B-dimer (Krogan et al., 2003a; Kobor et al., 2004; Mizuguchi et al., 2004) are functionally associated to histone acetyltransferase complexes and/or contain a bromodomain factor (BDF) (Johnston et al., 1999; Kusch et al., 2004; Ruhl et al., 2006). Most histone acetyltransferases that are linked to H2A.Z deposition belong to the MYST-family of acetyltransferases (Roth et al., 2001).
Three MYST-homologues are encoded by the *T. brucei* genome, HAT1, HAT2 and HAT3 (Ivens et al., 2005). A systematic target screen revealed the non-essential HAT3 to mediate exclusively H4K4ac (ElBashir et al., 2015), the only H4 acetyl mark not found enriched at TSSs.
Using FIPQuant and different next-generation sequencing analysis, I identified that HAT1 and HAT2 were required for establishing the TSS-specific chromatin composition. While HAT1 was responsible for acetylating the histone variants H2A.Z and H2B.V, and was not required for H2A.Z deposition, HAT2 acetylated H4 at TSSs and depletion of HAT2 resulted in reduced H2A.Z levels at TSSs. In addition, I observed that three BDFs were involved in the targeted H2A.Z deposition in *T. brucei*.

Introduction

Histone acetylation is a key regulator in inducing chromatin rearrangements (Struhl, 1998). It mediates nucleosome destabilisation by decreasing the DNA-histone interaction within a nucleosome and by inducing histone variant incorporation. Many chromatin remodelling complexes mediating H2A.Z deposition contain a bromodomain protein, that can bind to acetylated lysines (Bannister and Kouzarides, 2011). Moreover, histone acetyltransferase complexes are often associated to such chromatin remodelling complexes. In all eukaryotic organisms, acetylation of H4 and H2A.Z is found at TSSs and fundamental for transcription initiation (Haberle and Stark, 2018; Giaimo et al., 2019). H4 acetylation plays an additional role during the targeted deposition of H2A.Z. It serves as a binding platform for the chromatin remodelling complexes that exchange H2A with H2A.Z. The interaction of the bromodomain-containing protein stabilises the binding of the remodelling complex to these genomic sites (Brownell and Allis, 1996) demonstrating an intrinsic link between H2A.Z incorporation and histone acetylation.

In *S. cerevisiae*, Esa1, the catalytic component of the NuA4 histone acetyltransferase complex, mediates acetylation of H4 (Allard et al., 1999) as well as H2A.Z acetylation (Keogh et al., 2006). Similarly, the TIP60 acetyltransferase complex of p400 in *H. sapiens* acetylates both H4 and H2A.Z (Kimura and Horikoshi, 1998; Giaimo et al., 2019). Thus, the cause and consequence of H2A.Z acetylation could not be uncovered because of these interdependent processes.

NuA4 and TIP60, belong to the MYST-family of histone acetyltransferases. The genome of *T. brucei* encodes for three MYST-homologues, HAT1, HAT2 and HAT3 (Ivens et al., 2005). HAT3 is a non-essential protein and using a systematic target screen it has been shown that HAT3 acetylates exclusively H4K4 (ElBashir et al., 2015). Whereas HAT1 and HAT2 are essential for parasite growth. No such screen has been performed for HAT1 and HAT2 and, except HAT2-mediated TSS-specific H4K10ac, their targets sites are not known (Kawahara et al., 2008).

In this chapter I describe how I identified that HAT1 and HAT2 are important for establishing the TSS-specific chromatin structure in *T. brucei*. Therefore, I analysed the target sites of HAT1 and HAT2 using FIPQuant analysis and studied the effect of depleting HAT1 and HAT2 on targeted H2A.Z deposition in *T. brucei* by ChIP-seq. Since the intrinsic link of histone acetylation and H2A.Z deposition is established by BDF recruitment, I investigated whether different BDFs were involved in H2A.Z deposition as well. For these analyses, I generated cell lines for efficient and inducible depletion of HAT1, HAT2 or BDF1-4 transcripts based on RNA interference. The used RNAi system allows conditional overexpression of intramolecular stem-loop double-stranded RNA that mediate the degradation of the target transcripts (Durand-Dubief et al., 2003). Analyses to determine depletion efficiency and growth rate following RNAi-mediated depletion of HAT1, HAT2 or BDF1-4 transcripts over a time period of 48 h were conducted in replicates (see **Appendix figure 3** and **Appendix figure 4**). There, the results gained from knock down or overexpression analyses of different chromatin modifiers can be found (see **Appendix figure 8**).

Results

HAT1 acetylates specifically the TSS-specific histone variants H2A.Z and H2B.V

To identify the acetylation sites of HAT1, I depleted HAT1 in *T. brucei* using RNAi over a time period of 48 h, isolated histones from mononucleosomes by acid extraction or by immunoprecipitation of TSS-nucleosomes and analysed the acetylome changes by FIPQuant.

Following HAT1-depletion, I observed that in histone extracts the levels of the TSS-specific H4 acetyl marks on K5 and K10 were not affected compared to wild type levels. Whereas the levels of H4K2ac were reduced, also a TSS-specific acetyl mark (**Figure 5.1 a**). A similar pattern could be detected for H4 immunoprecipitated with TSS-nucleosomes (**Figure 5.1 b**). While HAT1-depletion did not affect TSS-specific H4 acetylation, I could observe a clearly distinct pattern for the TSS-specific histone variants H2A.Z and H2B.V.

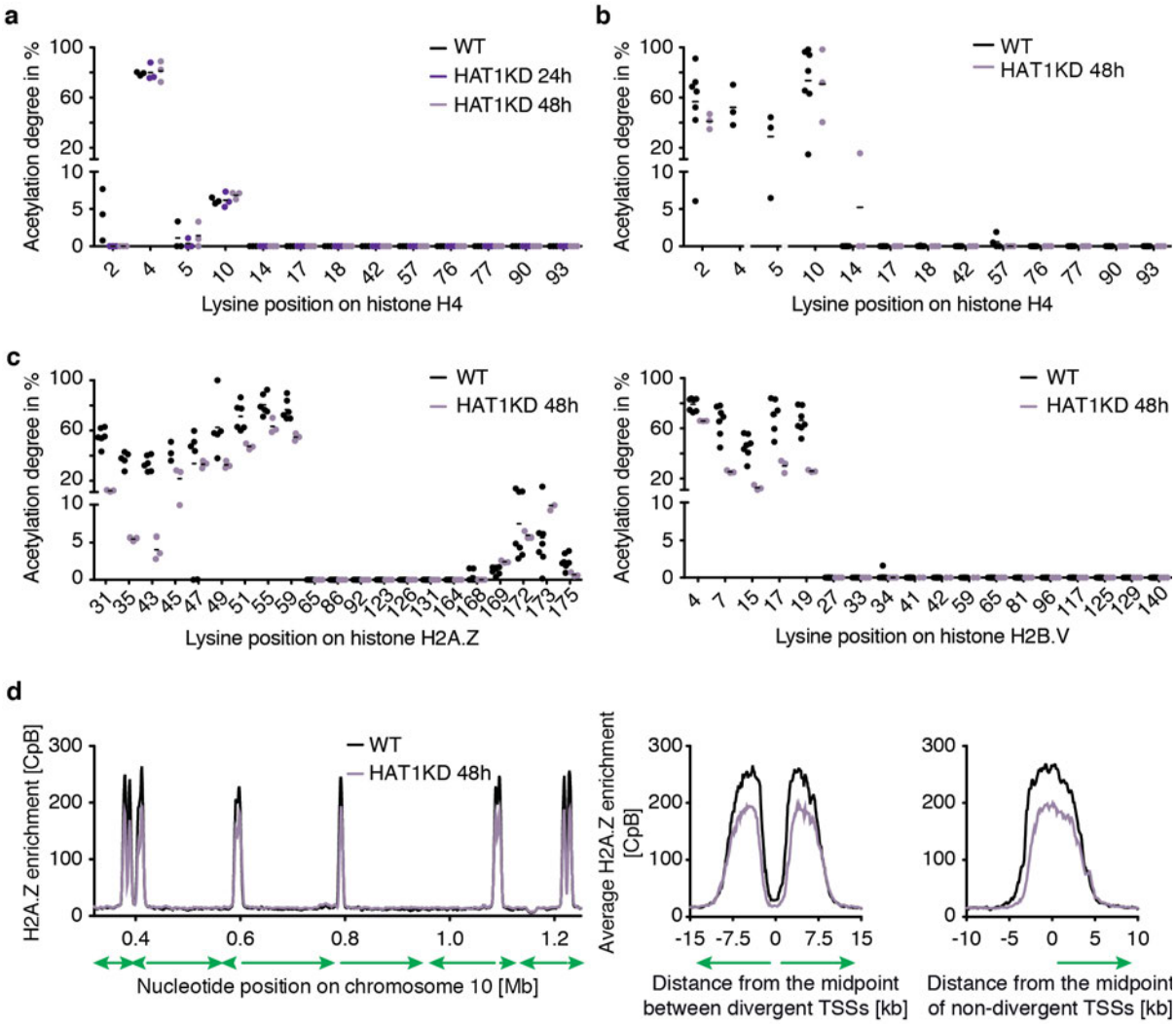


Figure 5.1. HAT1 acetylates the TSS-specific histone variants H2A.Z and H2B.V.
a HAT1 depletion reduces H4K2 acetylation in whole histone extracts. Lysine-specific acetylation degrees for histone H4 for wild type cells (WT) and after depletion of HAT1. H4 acetyl marks were quantified by FIPQuant using histones extracted from WT cells (black) and from 2T1 cells, which were depleted of HAT1 for 24 h (dark purple) or 48 h (purple; n=3). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. **b** HAT1 depletion does not affect acetyl marks of H4 co-immunoprecipitated with TSS-nucleosomes. Lysine-specific acetylation degrees for histone H4 at TSS-nucleosomes are shown for wild type cells (WT) and after depletion of HAT1. H4 acetyl marks were quantified using histones from immunoprecipitated TSS-nucleosomes of WT cells (black; n≥7) and 2T1 cells, depleted for HAT1 for 48 h (purple; n≥3). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. **c** Acetylation of H2A.Z and H2B.V is mediated by HAT1. H2A.Z and H2B.V acetyl marks were quantified using histones from immunoprecipitated TSS-nucleosomes of WT cells (black) and 2T1 cells, depleted for HAT1 for 48 h (purple). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. **d** HAT1 depletion does not affect H2A.Z deposition. MNase-ChIP-seq data of histone variant H2A.Z before (black) and after HAT1 depletion over a time period of 48 h (purple) are shown across a representative region on chromosome 10 (left panel) and averaged across divergent (n=81) and non-divergent (n=49) TSSs (right panel).

The N-terminal acetyl marks of both histone variants were remarkably reduced following HAT1 depletion (**Figure 5.1 c**) demonstrating that H2A.Z and H2B.V acetylation is mediated by HAT1. Since it has been shown that H2A.Z deposition can be stimulated by intrinsic acetyl marks, I performed MNase-ChIP-seq analyses to determine the genome-wide H2A.Z distribution following 0 h (n=4), 24 h (n=2) and 48 h (n=4) of RNAi-mediated HAT1-depletion. Compared to wild type, H2A.Z deposition was not extremely affected in HAT1-depleted cells (**Figure 5.1 d** and **Appendix figure 5 a**). While some replicates showed a decrease of the H2A.Z peak height to 80% compared to wild type or after 0 h HAT1-depletion, the width of the H2A.Z peak was identical compared to wild type.

Altogether, I could demonstrate that the histone acetyltransferase HAT1 mediated H2A.Z and H2B.V acetylation in *T. brucei*. Moreover, HAT1 depletion had only a mild effect on H2A.Z deposition revealing that H2A.Z acetylation might not be involved in its deposition in *T. brucei*.

HAT2 is a key player for H2A.Z depositioning

Analysing TSS-specific acetylation, targeted H2A.Z deposition and the total amount of chromatin-bound H2A.Z following HAT2-depletion revealed a clearly distinct pattern for HAT2 compared to HAT1.

FIPQuant analyses of whole histone extracts generated following HAT2-depletion over a time period of 48 h revealed that the levels of all three TSS-specific H4 acetyl marks (K2, K5 and K10) decreased drastically (**Figure 5.2 a**). Whereas FIPQuant analysis of H4 co-immunoprecipitated with H2A.Z revealed that following 48 h of HAT2-depletion H4 incorporated at TSSs was marked by similar levels of H4K2 and H4K10 acetylation than wild type (**Figure 5.2 b**). Moreover, H2A.Z and H2B.V acetyl marks were not affected by HAT2-depletion (**Figure 5.2 c**). The observation that H2A.Z-containing nucleosomes were marked by high levels of TSS-specific H4 acetylation even if HAT2 was depleted but in whole histone extracts TSS-specific H4 acetyl marks were reduced might already point to an intrinsic link of H4 acetylation and H2A.Z deposition in *T. brucei*.

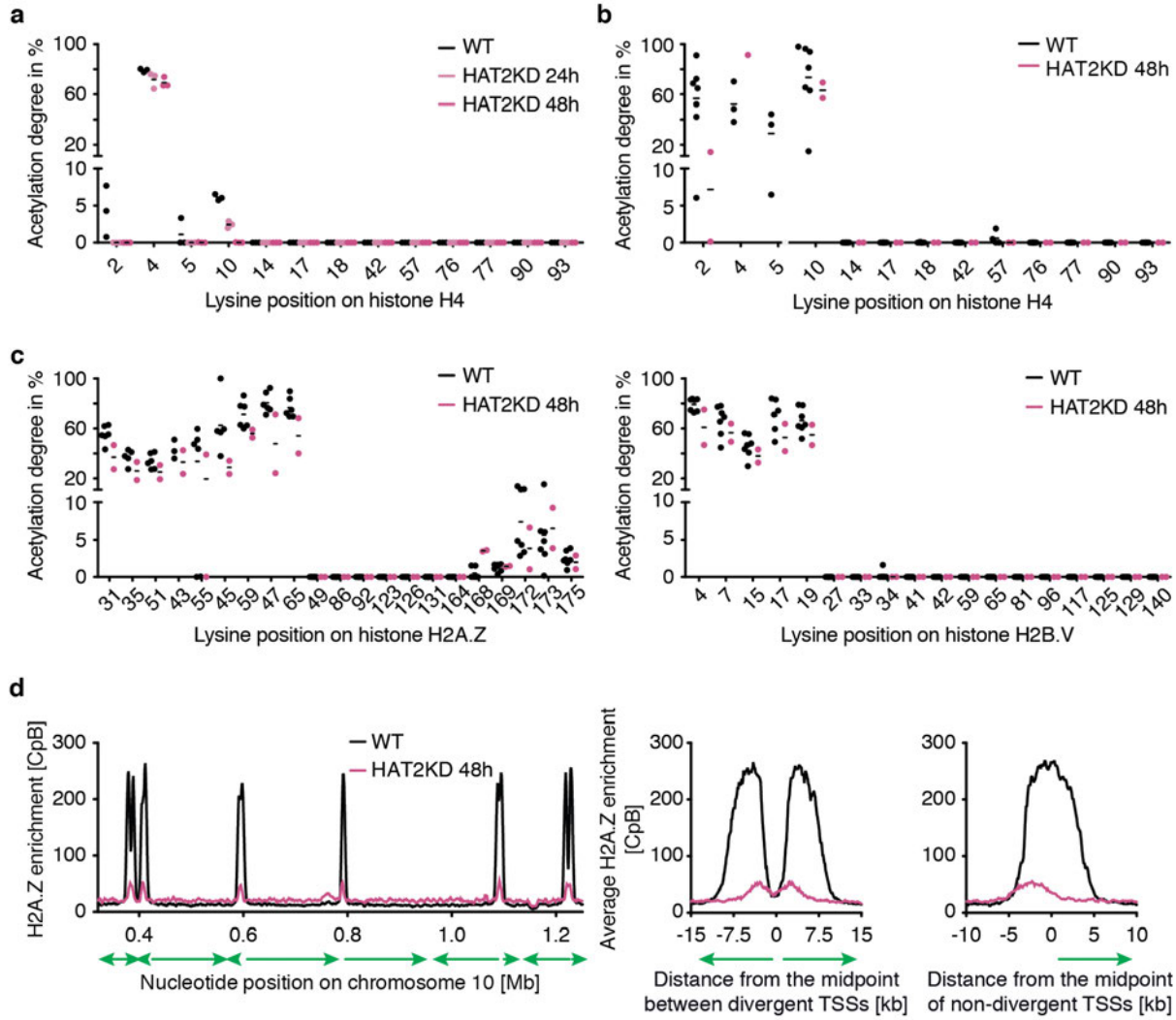


Figure 5.2. HAT2-mediated H4 acetylation is important for H2A.Z deposition.
a TSS-specific H4 acetylation decrease following HAT2 depletion. Lysine-specific acetylation degrees for histone H4 for wild type cells (WT) and after depletion of HAT2. H4 acetyl marks were quantified by FIPQuant using histones extracted from WT cells (black) and from 2T1 cells, which were depleted of HAT2 for 24 h (rose; n=3) or 48 h (magenta; n=3). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. **b** H4 acetyl marks co-purified with H2A.Z are still present after HAT2 depletion. Lysine-specific acetylation degrees for histone H4 at TSS-nucleosomes are shown for wild type cells (WT) and after depletion of HAT2. H4 acetyl marks were quantified using histones from immunoprecipitated TSS-nucleosomes of WT cells (black; n≥7) and 2T1 cells, depleted for HAT2 for 48 h (magenta; n≥2). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. **c** HAT2 depletion does not affect H2A.Z and H2B.V acetylation. H2A.Z and H2B.V acetyl marks were quantified using histones from immunoprecipitated TSS-nucleosomes of WT cells (black) and 2T1 cells, depleted for HAT2 for 48 h (magenta). The plotted acetylation degrees represent the median values of the quantified fragment numbers, calculated for each replicate. The black bar indicates the mean calculated from the median values. **d** H2A.Z deposition changes after HAT2 depletion. MNase-ChIP-seq data of histone variant H2A.Z before (black) and after HAT2 depletion over a time period of 48 h (magenta) are shown across a representative region on chromosome 10 (left panel) and averaged across divergent (n=81) and non-divergent (n=49) TSSs (right panel).

Thus, I sought to investigate this link in more detail and performed MNase-ChIP-seq analyses of the genomic H2A.Z distribution after 0 h (n=4), 24 h (n=2), 48 h (n=4) and 72 h (n=2) of RNAi-mediated HAT2-depletion. In doing so, I observed that following reduction of HAT2 levels H2A.Z incorporation at TSSs was impaired (**Figure 5.2 d** and **Appendix figure 5 b**). Compared to wild type, the peak width

decreased to ~50% and the peak height was reduced fivefold. While the replicates differed in the time scale of the effect, they all showed reduced levels of deposited H2A.Z latest after 48 h of HAT2. These findings suggested a clear link between TSS-specific H4 acetylation and H2A.Z deposition in *T. brucei*. Even though, standard ChIP-seq analysis can reveal whether a distinct chromatin mark is there or not, it cannot be used to yield information about the absolute amount of deposited protein. Thus, from my ChIP-seq analyses, I could not conclude whether the change of the H2A.Z levels at TSSs following HAT2-depletion was caused by less H2A.Z incorporated at TSSs or more H2A.Z incorporated at non-TSSs (**Figure 5.3 a**).

To differentiate these two possibilities, I quantified the fraction of chromatin-bound H2A.Z following depletion of HAT2 over a time period of 48 h by western blot using H3 levels as reference. In case H2A.Z levels at TSSs were reduced because of a loss of H2A.Z deposition, the fraction of chromatin-bound H2A.Z should also decrease. Whereas increased H2A.Z deposition to non-TSSs should not alter the fraction of chromatin-bound H2A.Z. Moreover, I analysed the total H2A.Z levels after 48 h of HAT2 depletion to exclude that the described effect was caused by reduced H2A.Z synthesis. In doing so, I observed that following 48 h of HAT2-depletion the levels of chromatin-bound H2A.Z decreased to 30.5% (**Figure 5.3 b** and **Appendix figure 6**). Whereas the total H2A.Z levels were not affected (**Figure 5.3 c** and **Appendix figure 7**) demonstrating that the change of the H2A.Z levels at TSSs following HAT2-depletion was caused by less H2A.Z incorporated at TSSs.

In summary, I discovered that TSS-specific H4 acetylation was mediated by HAT2 in *T. brucei*. Moreover, I showed that HAT2-mediated acetylation of H4 at TSSs is required for the targeted H2A.Z incorporation revealing a clear, functional link of H4 acetylation and H2A.Z deposition similar to observations in other eukaryotes.

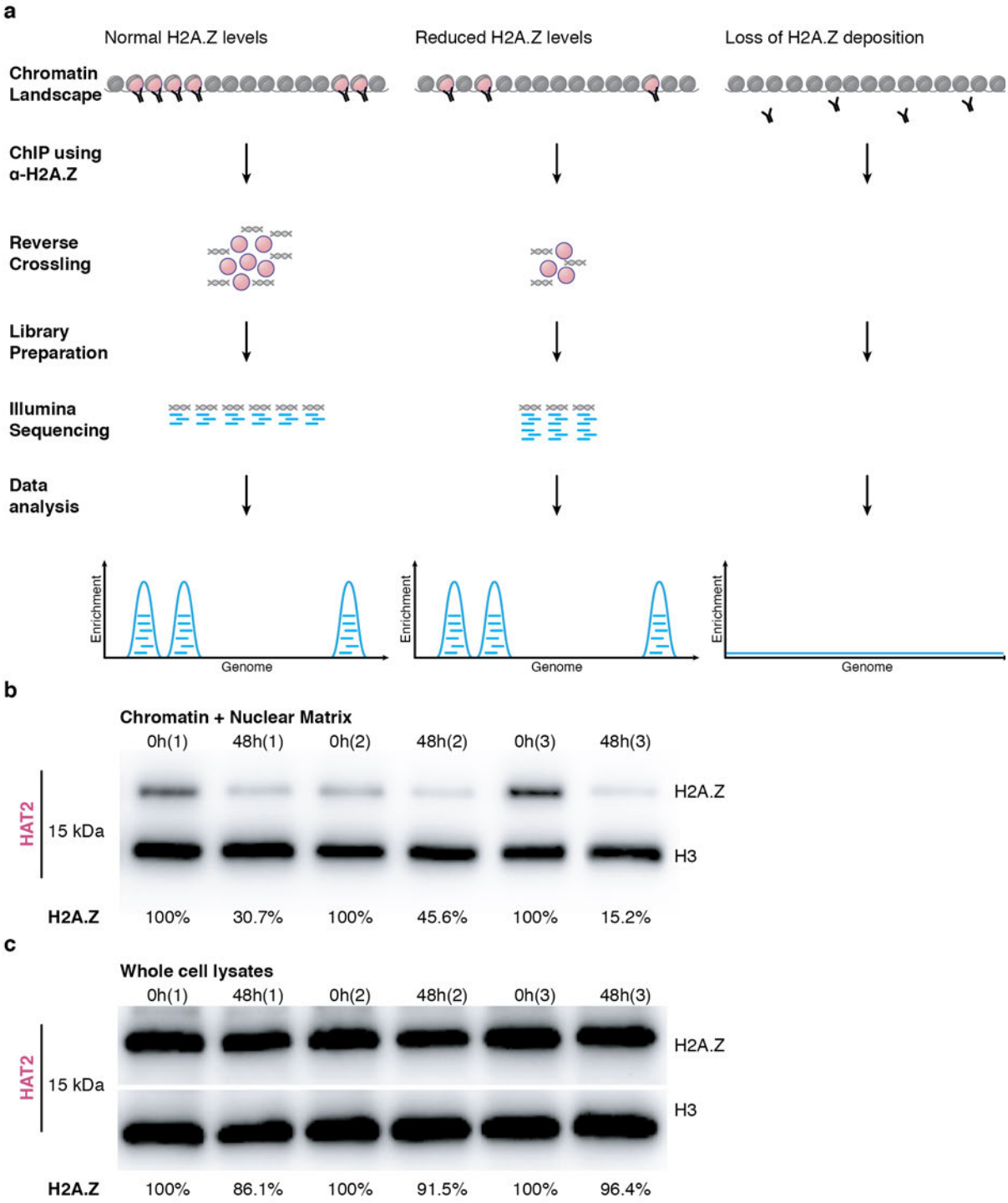


Figure 5.3. HAT2 depletion reduces H2A.Z incorporation at TSSs.
a Using H2A.Z-ChIP-Seq analysis, it is not possible to determine the absolute amount of chromatin-bound H2A.Z. Low H2A.Z signals results in similar sequencing output as high H2A.Z signals. **b** Western blot of chromatin-associated proteins extracted from 2 x 10^6 2T1 cells, which were depleted of HAT2 (n=3) for a time period of 48 h. Loaded are the insoluble fractions, containing chromatin-bound and nuclear matrix material. The H2A.Z percentages for each replicate were calculated by comparing the H2A.Z/H3 ratio after 48 h HAT2-depletion to the H2A.Z/H3 ratio of 0 h HAT2-depletion, which was set to 100%. H2A.Z/H3 were calculated by quantifying the H2A.Z and H3 signal over the background for each lane signal using ImageJ. **c** Western blot of whole cell extracts extracted from 2 x 10^6 from 2T1 cells, which were depleted of HAT2 (n=3) for a time period of 48 h. The H2A.Z percentages for each replicate were calculated by comparing the H2A.Z/H3 ratio after 48 h HAT2-depletion to the H2A.Z/H3 ratio of 0 h HAT2-depletion, which was set to 100%. H2A.Z/H3 were calculated by quantifying the H2A.Z and H3 signal over the background for each lane signal using ImageJ.

Bromodomain-containing proteins mediate proper H2A.Z incorporation

Histone acetylation plays key role for chromatin remodelling during gene activation. Nearly, all known transcriptional co-activators have either an intrinsic HAT activity (Struhl, 1998) or contain a bromodomain, which are conserved, ~110 amino-acid modules within proteins (Haynes et al., 1992) that bind specifically to acetylated lysines (Jeanmougin et al., 1997; Ornaghi et al., 1999). For example, the human TFIID contains a subunit, which has a double bromodomain recognising acetylated H4 (Jacobson et al., 2000). In yeast, the histone acetyltransferase Gcn5 possesses a bromodomain binding to H3 and H4 acetyl marks (Ornaghi et al., 1999), which is required for stabilizing the binding of the SWI/SNF chromatin remodelers and SAGA-complexes at promoters (Hassan et al., 2002). SWR-1, a SWI/SNF chromatin remodelling complex, essential for H2A.Z deposition (Mizuguchi et al., 2004) is composed of many different functional subunits including BDF1 (Kobor et al., 2004), which has been shown to bind acetylated lysines of H3 and H4 (Ladurner et al., 2003). While NuA4-mediated acetylation of H4 mediates SWR-1 chromatin remodelling, BDF1 is essential for transferring the stimulatory signal of NuA4 acetylation (Altaf et al., 2010).

In *T. brucei*, six bromodomain-containing proteins are encoded ranging from 220-660 amino acids (Siegel et al., 2009b). ChIP-seq assays analysing the genomic distribution of BDF1-4 revealed that three of them BDF1, BDF3 and BDF4 are enriched at TSSs (Siegel et al., 2009b; Schulz et al., 2015) and are essential for the parasite (**Appendix figure 4**). Moreover, it has been shown that the HAT2-enzyme complex, which is required for H2A.Z deposition, contains BDF3 (Vellmer et al., 2021). In contrast, BDF2 is not found at TSSs, and its genomic localisation is still elusive. BDF2 is not essential but is important for maintaining the BF stage integrity and immune evasion (Schulz et al., 2015).

To identify if BDFs are involved in the targeted H2A.Z deposition in *T. brucei*, I analysed the genomic localisation of H2A.Z following depletion of BDF1, BDF2, BDF3 or BDF4 using ChIP-seq and observed that all BDFs enriched at TSSs were implicated in H2A.Z deposition (**Figure 5.4** and **Appendix figure 8**). Whereas reduction of BDF2 that is not found TSSs over a time period of 48 h did not affect H2A.Z distribution (**Figure 5.4 b**). After 48 h of BDF1 depletion, the H2A.Z peak width increased from ~10 kb to ~15 kb (**Figure 5.4 a**). Following BDF4 depletion over 48 h, the H2A.Z peak width increased to ~20 kb (**Figure 5.4 d**) and to even ~30 kb after 48 h depletion of BDF3 (**Figure 5.4 c**). These findings revealed that following BDF1, BDF2 and BDF3 depletion H2A.Z was not anymore exclusively incorporated at TSSs. While depletion of all three BDFs led to a widening of the H2A.Z peak, the extend differed among them. Even though, the strongest effect was observed after BDF3 depletion, depletion of either BDF1 or BDF4 caused already a widening of the H2A.Z peak.

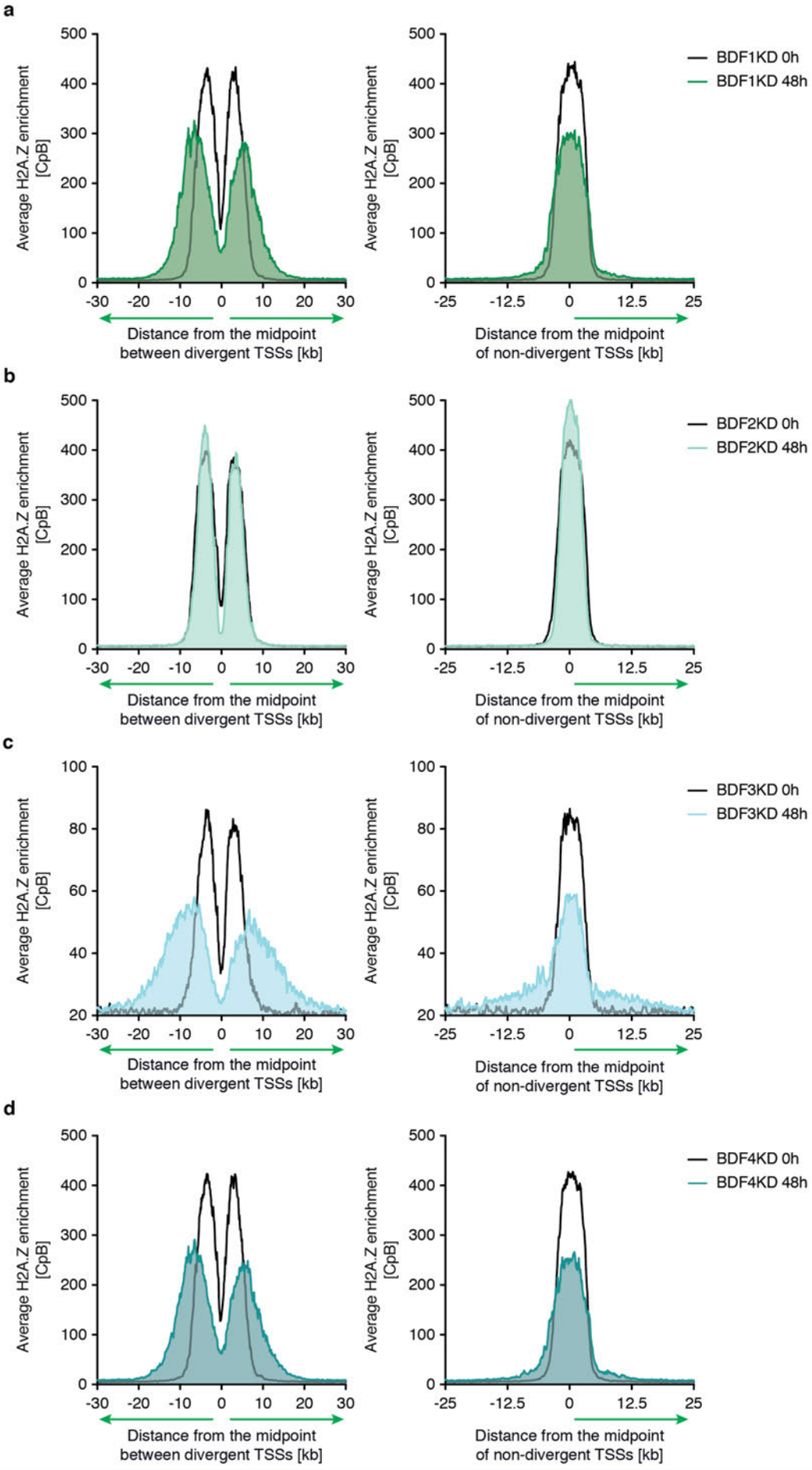

a
b
c
d
Average H2A.Z enrichment [CpB]
Distance from the midpoint between divergent TSSs [kb]
Distance from the midpoint of non-divergent TSSs [kb]
BDF1KD 0h
BDF1KD 48h
BDF2KD 0h
BDF2KD 48h
BDF3KD 0h
BDF3KD 48h
BDF4KD 0h
BDF4KD 48h

Figure 5.4. TSS-specific BDFs mediate targeted H2A.Z deposition.
a H2A.Z-containing regions change after BDF1 depletion. MNase-ChIP-seq data following 0 h (black) and 48 h (green) BDF1 depletion are averaged across divergent (n=81; left panel) and non-divergent (n=49; right panel) TSSs (right panel). **b** BDF2 depletion does not affect H2A.Z deposition. MNase-ChIP-seq data following 0 h (black) and 48 h (light green) BDF2 depletion are averaged across divergent (n=81; left panel) and non-divergent (n=49; right panel) TSSs (right panel). **c** Depletion of BDF3 leads to spreading of H2A.Z into PTUs. MNase-ChIP-seq data following 0 h (black) and 48 h (turquoise) BDF3 depletion are averaged across divergent (n=81; left panel) and non-divergent (n=49; right panel) TSSs (right panel). **d** Depletion of BDF4 widens the H2A.Z-containing regions. MNase-ChIP-seq data following 0 h (black) and 48 h (petrol) BDF4 depletion are averaged across divergent (n=81; left panel) and non-divergent (n=49; right panel) TSSs (right panel).

Taken together, my data revealed a clear link between BDF1, BDF3 and BDF4 and targeted H2A.Z deposition in *T. brucei*. Since all three BDFs were implicated in H2A.Z deposition, it might be that they act synergistically to ensure precise H2A.Z incorporation at TSSs.

Conclusions

Histone acetylation and H2A.Z incorporation are conserved characteristics of TSS-chromatin and critical for transcription initiation in almost all organisms (Haberle and Stark, 2018). The enzyme complexes acetylating histones and mediating H2A.Z deposition have already been identified in *S. cerevisiae*, *D. melanogaster* and *H. sapiens* revealing a highly conserved mechanism for establishing TSS-chromatin (Johnston et al., 1999; Krogan et al., 2003a; Kobor et al., 2004; Kusch et al., 2004; Mizuguchi et al., 2004; Ruhl et al., 2006).

Histone acetylation is a pre-requisite for targeted H2A.Z deposition because it stimulates recruitment of the responsible chromatin remodeler by interaction with a bromodomain-containing protein (Shia et al., 2006; Altaf et al., 2010).

It has already been shown that in *T. brucei* TSSs are enriched with acetylated histones, BDFs and H2A.Z (Siegel et al., 2009b). However, a direct link between this chromatin marks, as described for other eukaryotes, was still elusive. In this chapter, I described three major findings. Firstly, I identified that both MYST-family homologues HAT1 and HAT2 mediated TSS-specific histone acetylation, but at different target histones. HAT1 was responsible for hyperacetylation of H2A.Z and H2B.V, whereas HAT2 mediated TSS-specific H4 acetylation. Acetylation marks on the other histones were not affected (see **Appendix figure 2**). My second finding was that HAT2 depletion reduced H2A.Z levels at TSSs suggesting that the TSS-specific acetyl marks at H4 might be involved in targeted H2A.Z deposition. Whereas, HAT1 depletion had only a minor effect on H2A.Z deposition. Since H4K2 acetylation was affected after HAT1 depletion and HAT2 depletion, it might be that the slight decrease of H2A.Z deposition observed after HAT1 depletion was a consequence of the loss of H4K2ac. Thirdly, I could identify that the incorporation of H2A.Z in *T. brucei* involved bromodomain-containing proteins.

Taken together, my data revealed a clear link between histone acetylation and targeted H2A.Z deposition in *T. brucei*, as described for other eukaryotes, and confirmed that the establishment of TSS-chromatin is highly conserved in eukaryotes.

Chapter 6: Chromatin composition at TSSs plays a critical role for transcription

Summary

Unlike in other eukaryotes studied so far, the process of RNAPII transcription is not regulated at the level of transcription initiation in *T. brucei* and differential gene expression is caused by post-transcriptional processing of transcripts (Clayton, 2002; Clayton, 2019). Given the lack of canonical promoter sequences at TSSs, it has been suggested that acetylated histones and the histone variant H2A.Z mediate an open chromatin structure at TSSs to facilitate RNAPII binding and transcription initiation in trypanosomes (Siegel et al., 2009b; Wedel et al., 2017).

In many eukaryotes, H2A.Z is a conserved feature of the TSS-chromatin structure and is implicated in many transcriptional processes with sometimes contradictive roles. While this is true for unmodified H2A.Z, acetylated H2A.Z is clearly linked to active transcription (Giaimo et al., 2019). In many eukaryotes, including *S. cerevisiae* and *H. sapiens*, the enzyme complexes that acetylate H2A.Z, also mediate H4 acetylation (Kimura and Horikoshi, 1998; Allard et al., 1999; Keogh et al., 2006; Giaimo et al., 2019), a fundamental prerequisite for H2A.Z deposition (Altaf et al., 2010; Giaimo et al., 2019). Thus, studies of the cause and consequence of H2A.Z acetylation at TSSs are complicated in these organisms and the understanding whether the deposition or the acetylation of H2A.Z is important for transcriptional regulation is very little.

I identified that in *T. brucei* H4 and H2A.Z were acetylated by two different enzymes, HAT2 and HAT1, at TSSs. Since only HAT2-mediated H4 acetylation was required for targeted H2A.Z deposition and HAT1-mediated H2A.Z acetylation not, I anticipated that the processes of H2A.Z deposition and acetylation are functionally independent in *T. brucei*.

Combining different ChIP-seq approaches, RNA-seq and western blot analysis of chromatin-bound RNAPII-levels following HAT1 or HAT2 depletion, I was able to disentangle the role of H2A.Z acetylation and deposition. Even though a loss of H2A.Z deposition at TSSs deposition mediated by HAT2-depletion changed the sites of transcription initiation and reduced RNAPII occupancy throughout the PTUs, the genome-wide transcript levels were not affected. In contrast, decreasing H2A.Z acetylation through HAT1 depletion did not affect the sites of transcription initiation, but led to a drastic reduction of genome-wide RNAPII occupancy and transcription.

Introduction

Thus far, H2A.Z is the only histone variant found in almost all studied eukaryotes and has been shown to be involved in many different biological processes, such as DNA repair, chromosome segregation and transcription (Giaimo et al., 2019). While H2A.Z is predominantly localised at TSSs and plays an

important regulatory role for transcription initiation, it has also been found at genomic regions which are associated to silent transcription (Guillemette and Gaudreau, 2006; Marques et al., 2010). These contradicting findings have made it difficult to clearly associate the histone variant to active transcription (Zlatanova and Thakar, 2008).

As all histones, H2A.Z can be post-translationally modified and acetylated H2A.Z has been exclusively associated to sites of active transcription in contrast to unmodified H2A.Z. For example, in chicken, hyperacetylated H2A.Z marks active genes (Bruce et al., 2005). In *S. cerevisiae*, H2A.Z is found at the 5´-end of active and inactive genes (Raisner et al., 2005). But H2A.ZK14ac is specifically enriched at promoters of active genes (Keogh et al., 2006) and only unmodified H2A.Z is found at inactive promoters (Guillemette et al., 2005). A similar pattern has been described at active promoters in human cells, where H2A.Zac is enriched and unmodified H2A.Z depleted (Valdes-Mora et al., 2012).

Overall, these findings suggest that acetylation of H2A.Z is more important for transcription initiation than unmodified H2A.Z in eukaryotes. However, studying the role of H2A.Zac is difficult in these organisms because the histone acetyltransferase acetylated H2A.Z is also important for stimulating its deposition. For example, in yeast, H2A.Z acetylation is mediated by NuA4 (Allard et al., 1999; Keogh et al., 2006). However, NuA4 also acetylates H4 at TSSs, which is required to target the chromatin remodeler SWR1 to these sites that then exchanges H2A with H2A.Z (Krogan et al., 2003a; Kobor et al., 2004; Mizuguchi et al., 2004). Similarly, in humans, the TIP60 complex acetylates both H4 and H2A.Z (Kimura and Horikoshi, 1998; Giaimo et al., 2019).

In the previous chapter, I demonstrated that in *T. brucei*, the histone acetyltransferase HAT1 and HAT2, both important for TSS-specific acetylation of H2A.Z and H4, respectively, are functionally independent. This allowed me to separately analyse the role of H2A.Z deposition and acetylation for transcription initiation in a eukaryotic organism. In this chapter, I describe how I could unravel the functional link between histone acetylation, H2A.Z deposition and transcription initiation in an organism that lacks defined sites of transcription initiation.

Results

H2A.Z acetylation is required for transcription initiation

To understand the role of H2A.Z acetylation for transcription initiation in eukaryotes, I analysed the effect of HAT1 depletion, which mediates H2A.Z acetylation in *T. brucei*, on the genome-wide RNAPII distribution by ChIP-seq. Therefore, I generated a 2T1 cell line, which expressed only TY1-tagged RPB9, a DNA-binding subunit of the RNAPII complex (Devaux et al., 2006) and integrated afterwards the construct allowing inducible, RNAi-mediated degradation of HAT1 transcripts.

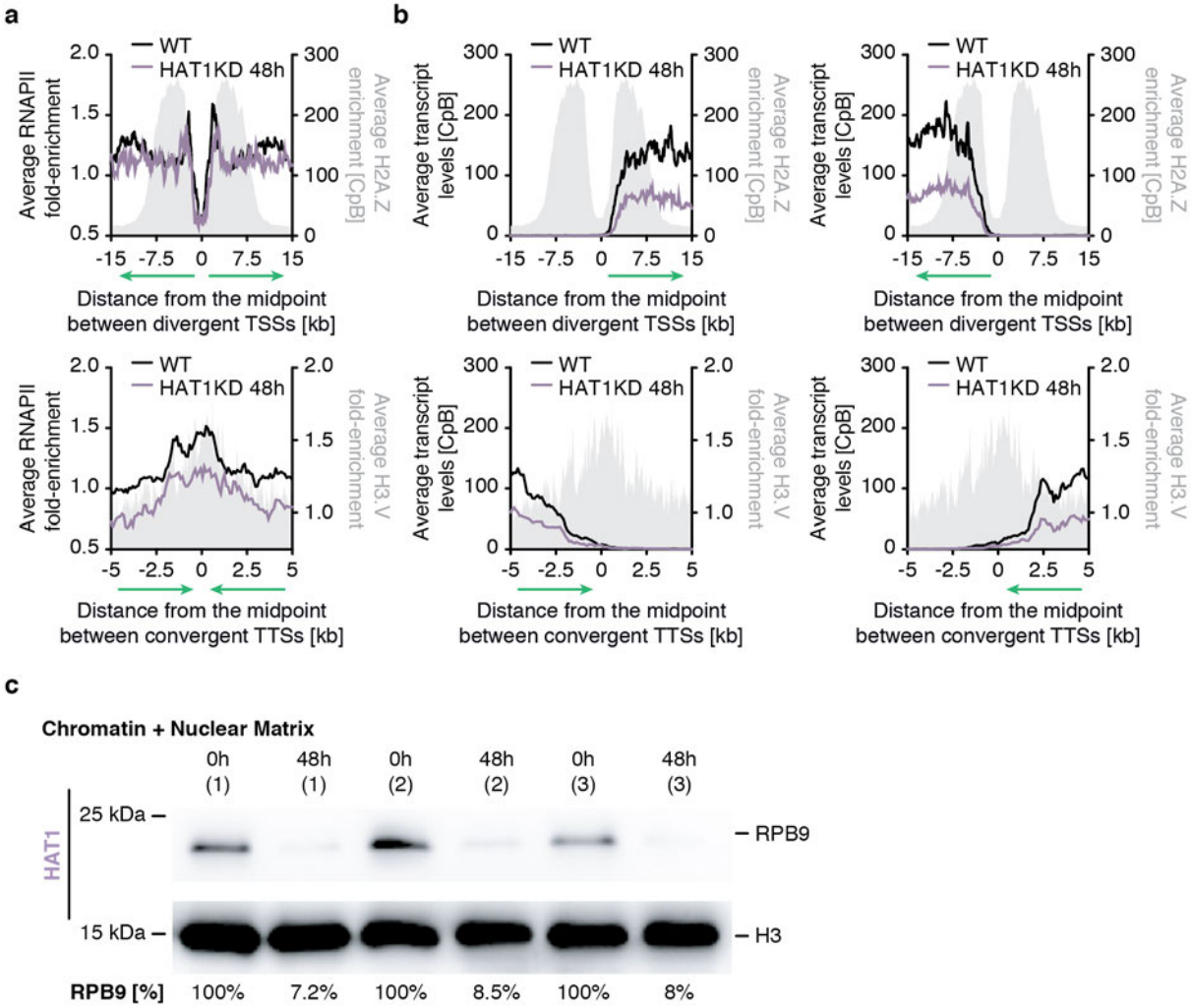


Figure 6.1. Reduction of H2A.Z acetylation strongly reduces RNAPII transcription.
a HAT1 depletion strongly reduces levels of transcription initiation. ChIP-seq data of RNAPII before (black) and after HAT1 depletion over a time period of 48 h (purple) are averaged across divergent (n=81) TSSs (left panel). Strand-specific RNA-seq data before (black) and after HAT1 depletion over a time period of 48 h (purple) are averaged across divergent (n=81) TSSs (+strand: middle panel; -strand: right panel). To illustrate the width of the TSSs, MNase-ChIP-seq data of H2A.Z is shown in grey. **b** RNAPII-enrichment is still present at TTSs after HAT1 depletion. ChIP-seq data of RNAPII before (black) and after HAT1 depletion over a time period of 48 h (purple) are averaged across convergent (n=41) TTSs (left panel). Strand-specific RNA-seq data before (black) and after HAT1 depletion over a time period of 48 h (purple) are averaged across convergent (n=41) TTSs (+strand: middle panel; -strand: right panel). To illustrate the width of the TTSs, MNase-ChIP-seq data of H3.V is shown in grey. **c** Western blot of chromatin-associated proteins extracted from 2 x 10^6 2T1 cells, which were depleted of HAT1 (n=3) for a time period of 48 h. Loaded are the insoluble fractions, containing chromatin-bound and nuclear matrix material. The TY1-RPB9 percentages for each replicate were calculated by comparing the TY1-RPB9 /H3 ratio after 48 h HAT1-depletion to the TY1-RPB9 /H3 ratio of 0 h HAT1-depletion, which was set to 100%. TY1-RPB9 /H3 were calculated by quantifying the H2A.Z and H3 signal over the background for each lane signal using ImageJ.

Similar to previous findings (Wedel et al., 2017), I found RNAPII to be strongly enriched at the 5´-end of TSSs in wild type cells. Moreover, transcription termination was accompanied by RNAPII accumulation in *T. brucei* demonstrated by the enrichment of RNAPII at TTSs in wild type cells (**Figure 6.1 a** and **Appendix figure 9 b**). Even though, this pattern of RNAPII enrichment did not change much upon depletion of HAT1 over a time period of 48 h, I observed a strong, 2-fold decrease in the total transcript levels following HAT1 depletion at TSSs and TTSs (**Figure 6.1 b**). Thus, I analysed the levels of

chromatin-bound RNAPII after 48 h HAT1 depletion and observed that they only comprised 8% of the levels after 0h HAT1 depletion (**Figure 6.1 c** and **Appendix figure 10**). These findings prompted me to analyse whether the RNAPII occupancy across the PTUs was affected by HAT1 depletion and indeed, the RNAPII levels at the 5´-end of PTUs were higher than at the 3´-end after 48 h of HAT1 depletion (**Figure 6.2 a**). To complete my findings, I analysed the RNA transcript levels across the PTUs, which were also remarkably reduced (**Figure 6.2 b**). To be able to compare absolute transcript levels among different experiments, I added a set of 92 synthetic transcripts (ERCC) of known concentrations as spike-ins to my samples before extracting total RNA for the RNA-seq analyses.

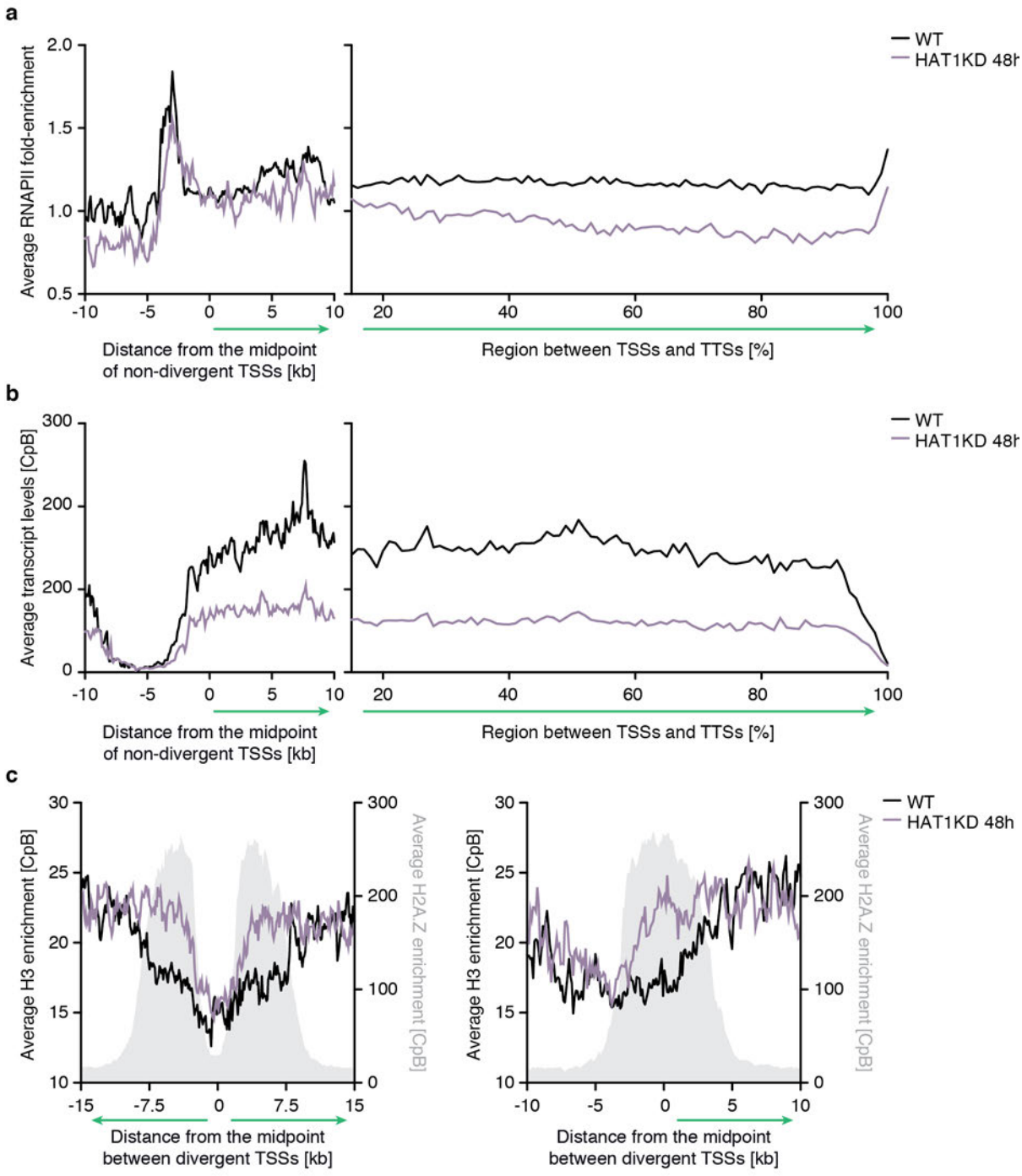


Figure 6.2. Reduction of H2A.Z acetylation increases nucleosomal barrier at TSSs.
a Genome-wide RNAPII-mediated transcript levels are strongly affected following HAT1 depletion. Outline of chromatin organisation in *T. brucei* (Upper panel). Nucleosomes containing H2A.Z and H2B.V characterise TSSs (rose) and at TTSs H3 and H4 are replaced by H3.V and H4.V in nucleosomes (orange). ChIP-seq data of RNAPII before (black) and after HAT1 depletion over a time period of 48 h (purple) are averaged across non-divergent (n=49) TSSs and across regions between TSSs and TTSs on the +strand (n=109; middle panel). **b** Strand-specific RNA-seq data before (black) and after HAT1 depletion over a time period of 48 h (purple) are averaged are averaged across non-divergent (n=49) TSSs and across regions between TSSs and TTSs on the +strand (n=109; lower panel). **c** Nucleosome occupancy at TSSs increases after HAT1 depletion. MNase-ChIP-seq data of canonical H3 are averaged across divergent (n=81; left panel) and non-divergent (n=49; right panel) TSSs. To illustrate the width of the TSSs, MNase-ChIP-seq data of H2A.Z is shown in grey.

Together, my results revealed that transcription was impaired following the reduction of H2A.Z acetylation levels. Since H2A.Z acetylation is known to destabilise the nucleosomal core particle in synergy with acetylated core histones (Ishibashi et al., 2009) and is important to modulate the charge patch of H2A.Z's N-terminal tail (Ren and Gorovsky, 2001), it is possible that H2A.Z acetylation affects the DNA-H2A.Z interaction within the nucleosome in *T. brucei*. Previously, it has been shown, that the DNA at TSSs is more accessible in *T. brucei* (Wedel et al., 2017). Since modulation of the nucleosomal stability is known to directly affect DNA accessibility (Fenley et al., 2018), I analysed whether reduced H2A.Z acetylation would affect the DNA accessibility at TSSs in *T. brucei*. Therefore, I evaluated the nucleosome occupancy following HAT1 depletion over a time period of 48 h using an already published H3-ChIP-seq protocol (Wedel et al., 2017) and observed that the nucleosome occupancy was higher than in wild type (**Figure 6.2 c**). Since the degree of nucleosome occupancy can be used as a marker for DNA accessibility, my results suggested that reduction of H2A.Z acetylation made the DNA less accessible at TSSs.

Taken together, my data revealed a link between H2A.Z acetylation and transcription initiation. I found that reducing acetylated H2A.Z at TSSs had tremendous effect on the RNA transcript levels and on the levels of chromatin-bound RNAPII. Even though, my findings are somewhat preliminary, they suggest DNA accessibility at TSSs is decreased following a reduction of H2A.Z acetylation levels at TSSs.

H2A.Z serves as a landing pat for RNA polymerase II

Given the apparent absence of the core features of transcription initiation, such as RNAPII promoter motifs and PIC assembly, at TSSs in *T. brucei*, it was hypothesised that a reduced chromatin compaction at TSSs is sufficient for RNAPII transcription initiation. The reduced chromatin compaction could be established by specifically deposited histone acetyl marks and incorporated histone variants (McAndrew et al., 1998; Siegel et al., 2009b; Wedel et al., 2017). While my finding that H2A.Z acetylation facilitates transcription initiation by reducing the nucleosome barrier at TSSs shows that this is indeed the case, the role of H2A.Z deposition remains unclear. It might be that H2A.Z pre-defines the sites for RNAPII recruitment as suggested before (Adam et al., 2001; Neves et al., 2017; Nissen et al., 2017).

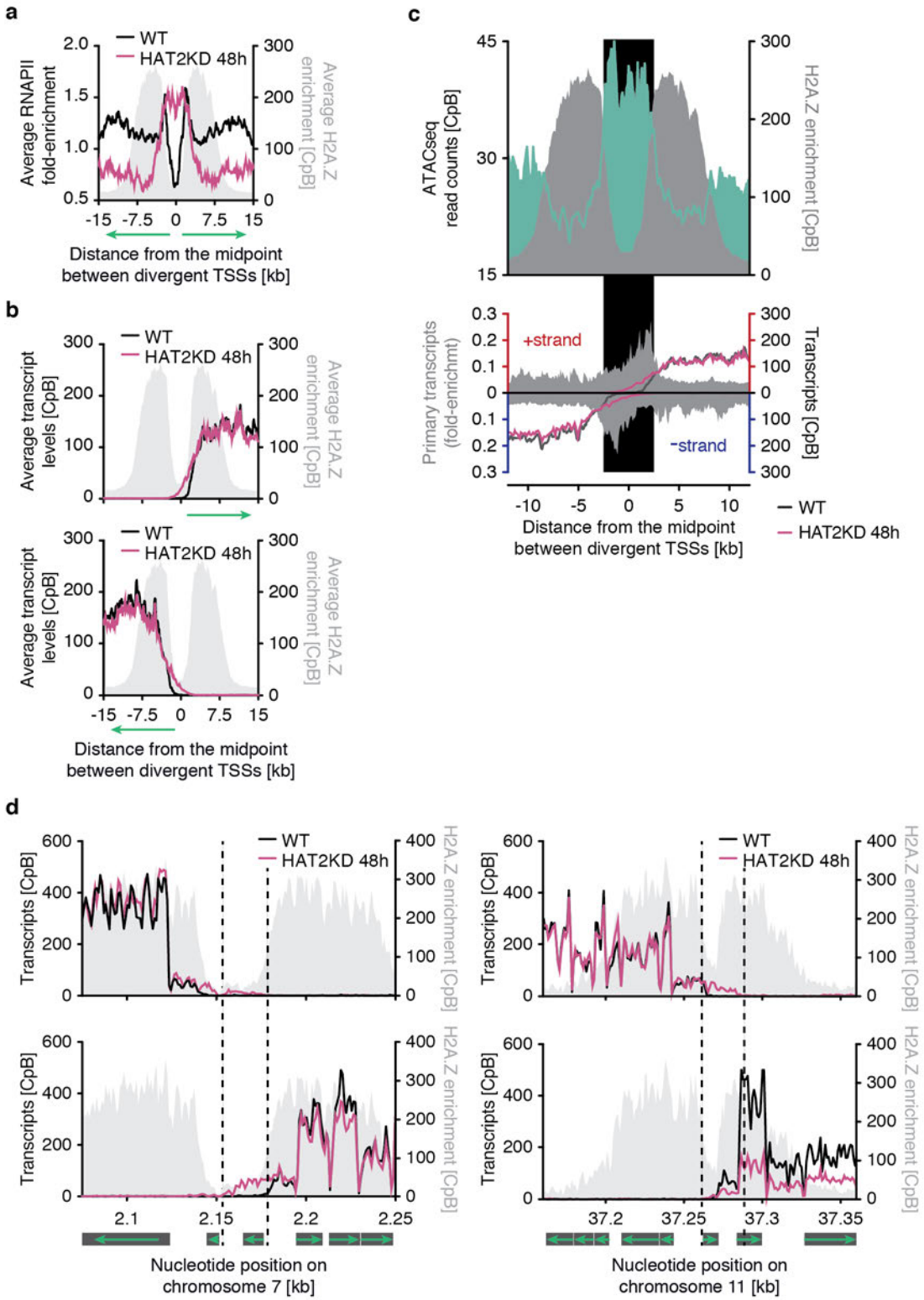


Figure 6.3. Loss of H2A.Z deposition shifts transcription initiation.
a ChIP-seq data of RNAPII before (black) and after HAT2 depletion over a time period of 48 h (magenta) are averaged across divergent (n=81) TSSs (left panel). **b** Strand-specific RNA-seq data before (black) and after HAT2 depletion over a time period of 48 h (magenta) are averaged across divergent (n=81) TSSs (+strand: middle panel; -strand: right panel). To illustrate the width of the TSSs, MNase-ChIP-seq data of H2A.Z is shown in grey. **c** The ATAC-seq data (green) and MNase-ChIP-seq data of H2A.Z (grey) are derive from wild type (WT) cells and are averaged across divergent TSSs (n=37; upper panel). The ATAC-seq (Müller et al., 2018) and MNase-ChIP-seq data are normalised to the total number of reads plotted as counts per billion reads [CpB]. RNA-seq data showing transcripts derived from the +strand or the -strand before (black) and after 48 h HAT2 depletion (magenta) are averaged across divergent TSSs (n=37; lower panel). The data are normalised to a spike-in control to account for differences in total RNA levels per cell after HAT2 depletion and plotted as counts per

billion reads [CpB]. **d** HAT2 depletion causes antisense transcription at TSSs. Strand-specific RNA-seq data (-strand: upper panel; +strand: lower panel) before (black) and after HAT2 depletion over a time period of 48 h (magenta) are shown across representative TSSs on chromosome 7 (left panel) and chromosome 11 (right panel). To illustrate the width of the TSSs, MNase-ChIP-seq data of H2A.Z is shown in grey.

Thus, a change in the H2A.Z distribution might affect the genome-wide RNAPII distribution. Given the ability to influence H2A.Z deposition by reducing TSS-specific H4 acetylation, I investigated the link between H2A.Z deposition and RNAPII recruitment in *T. brucei* following HAT2 depletion.

Therefore, I generated a 2T1 cell line that expressed only TY1-tagged RPB9 and allowed inducible, RNA-mediated degradation of HAT2 transcripts and determined the genome-wide enrichment of RNAPII by ChIP-seq analysis following 48 h of HAT2 depletion. Compared to wild-type conditions, I observed a marked upstream shift of the RNAPII peaks at TSSs following HAT2 depletion. Moreover, the RNAPII peaks increased in width (**Figure 6.3 a** and **Appendix figure 9 a**). In line with this observation, RNA-seq analysis revealed an upstream shift of transcription initiation to the same extend as the RNAPII peaks (**Figure 6.3 b**). To understand the preference of RNAPII to bind to TSS-adjacent sites, I analysed a previously published ATAC-seq dataset (Müller et al., 2018) from *T. brucei* BF wild-type cells and investigated the TSS-chromatin for its transposase accessibility. Interestingly, I observed that the DNA regions upstream of the canonical TSSs enriched with H2A.Z were more accessible then the DNA at the TSSs itself. Thus, reduced H2A.Z deposition at TSSs shifted transcription initiation to sites of increased DNA accessibility (**Figure 6.3 c**).

While *T. brucei* is suggested to live without any regulation of transcription initiation, defined sites of transcription initiation and termination may be essential to avoid downregulation of essential genes by the highly efficient RNAi-pathway of *T. brucei*. Given the upstream shift of transcription initiation and the read-through of transcription at termination sites after HAT2 depletion over 48 h, I analysed if antisense transcripts of essential genes were produced. Indeed, I detected antisense transcripts for essential genes with ORFs in the regions upstream of TSSs (**Figure 6.3 d**). However, while the change of H2A.Z deposition had a strong effect on transcription initiation, RNAPII distribution and increased antisense transcription at TSSs, the genome-wide RNA transcript levels across PTUs were only slightly reduced following 48 h of HAT2 depletion (**Figure 6.4 a**). In other eukaryotes, the processes of RNAPII transcription initiation are tightly regulated not only to ensure functional RNAPII progression but also to align the processes of transcription to the ones of RNA processing (Carrillo Oesterreich et al., 2011). Thus, I wondered whether the changes of transcription after HAT2 depletion had an effect on RNA maturation in trypanosomes. Therefore, I compared the quantity of *trans*-splicing events in wild type cells and following 48 h of HAT2 depletion. Since my total RNA-seq datasets were generated from rRNA-depleted and not from poly(A)-enriched RNA that only includes transcript of mRNAs, I could use them as input to perform a preliminary analysis of *trans*-splicing through *in silico* enrichment of transcripts associated to the splice leader mini-exon (**Figure 6.4 b**). In doing so, I observed that *trans*-splicing was shifted to sites far more upstream from the first gene of the PTU following HAT2 depletion compared

to wild type (**Figure 6.4 c**), in line with the upstream shift of the RNAPII peaks and transcription initiation sites. Moreover, across PTUs, I could detect less *trans*-splicing events than found in wild type (**Figure 6.4 d**).

Altogether, my observations supported my hypothesis that H2A.Z deposition is important for RNAPII recruitment, since loss of H2A.Z shifted the sites of transcription initiation. The shift of transcription initiation might be deleterious for *T. brucei* because it caused synthesis of antisense RNA, which can be used by the RNAi-machinery to downregulate expression of essential genes. Moreover, it is possible that correct transcription initiation is required for the processes of *trans*-splicing in *T. brucei*.

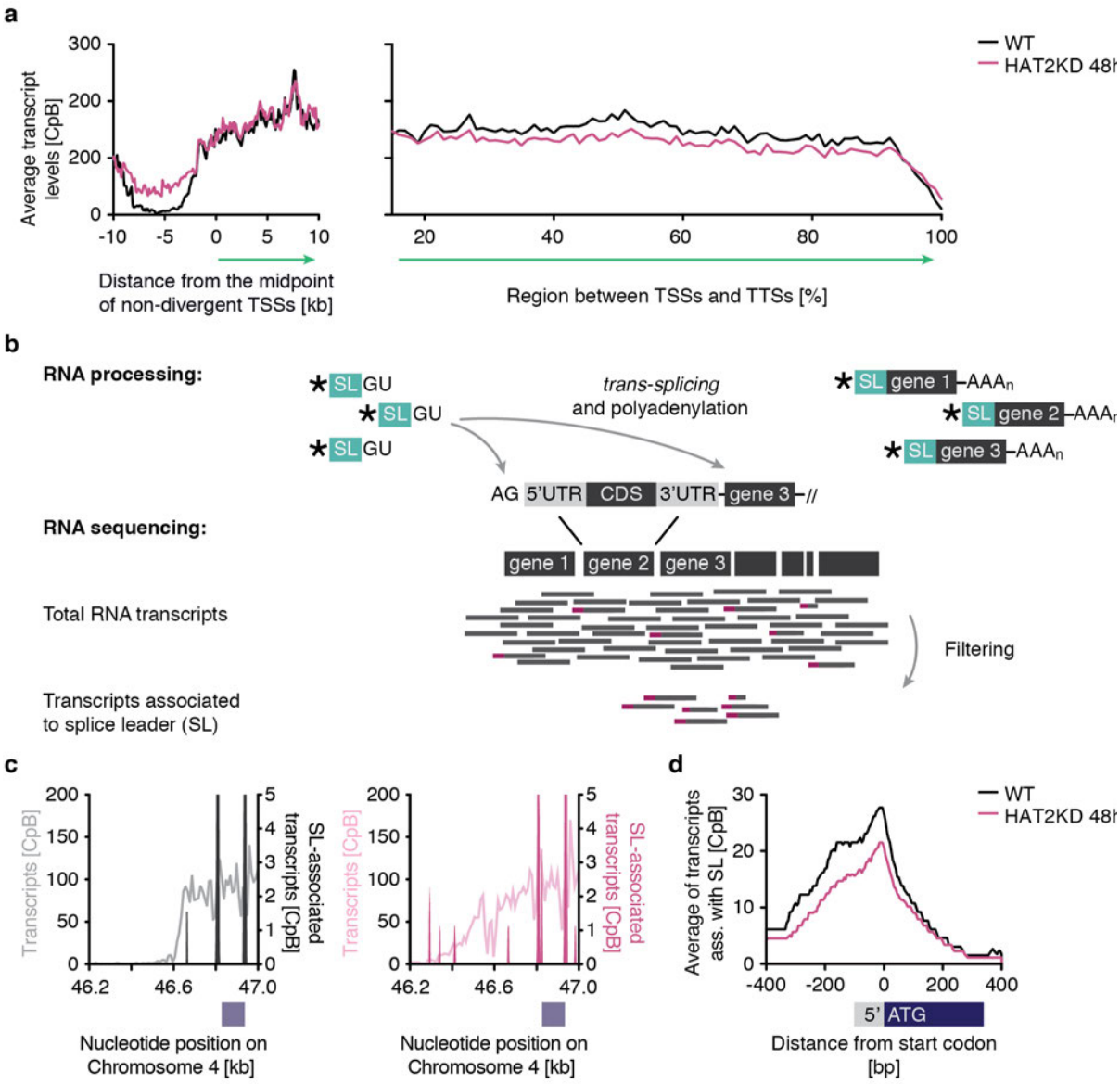


Figure 6.4. TSS-chromatin is critical for coordinating RNA synthesis and processing.
a Outline of genome organisation in *T. brucei* (Upper panel). RNAPII-transcribed genes are organised in long polycistronic transcription units (PTUs). The first gene of a PTU is represented in a light blue box and the residual genes in a dark blue box. Nucleosomes containing H2A.Z and H2B.V characterise TSSs (rose) and at TTSs H3 and H4 are replaced by H3.V and H4.V in nucleosomes (orange). ChIP-seq data of RNAPII before (black) and after HAT2 depletion over a time period of 48 h (magenta) are averaged across 5´UTRs of all residual genes (n=8798; left panel). MNase-ChIP-seq data of canonical H3 before (black) and after HAT2 depletion over a time period of 48 h (magenta) are averaged across 5´UTRs of all residual genes (n=8798; middle panel). RNA-seq data before (black) and after HAT2 depletion over a time period of 48 h (magenta) are averaged across 5´UTRs of all residual genes (n=8798; right panel). **b** Workflow of in silico filtering of transcripts

associated to the spliced-leader mini-exon. Each mRNA contains the 39 nt-mini-exon of the spliced-leader (SL). By filtering the raw reads of RNA-seq datasets using the spliced-leader sequences as query, mRNAs can be enriched in silico over unprocessed or degraded RNAs. **c** *Trans*-splicing shifts together with transcription initiation following HAT2 depletion. RNA-seq data (grey) and SL-associated transcripts (black) before HAT2 depletion are shown across a representative region on chromosome 4 (left panel). RNA-seq data (rose) and SL-associated transcripts (magenta) after 48 h HAT2 depletion are shown across a representative region on chromosome 4 (right panel). The light blue box represents the first gene of the downstream PTU. **d** *Trans*-splicing events are reduced across PTUs following HAT2 depletion. SL-associated transcripts before (black) and after 48 h (magenta) HAT2 depletion are averaged across 5´UTRs of all residual genes (n=8798). The grey box represents the average length of a 5´UTR and the dark blue box represent the ORF.

Conclusions

For a long time, it has been suggested that acetylated H2A.Z is more important for active transcription than unmodified H2A.Z by mediating nucleosome destabilisation and DNA accessibility for the transcription machinery. Thus far, it has been difficult to study *in vivo* the role of H2A.Z acetylation and deposition for transcription, since in many cases, the enzymes acetylating H2A.Z are also responsible for H4 acetylation, which is a pre-requisite for H2A.Z deposition.

Since I found that in *T. brucei*, the H4 and H2A.Z are acetylated by different enzymes at TSSs and H2A.Z deposition relies H4 but not on H2A.Z acetylation, I could disentangle the roles of H2A.Z acetylation and deposition for transcription, to my knowledge for the first time in a eukaryotic organism.

My data clearly show that H2A.Z acetylation and deposition was critical for transcription in *T. brucei*. Regardless of its modification state, H2A.Z was required for defining the correct initiation sites for the transcription machinery. However, unacetylated H2A.Z can perturb transcription initiation probably by increasing the nucleosomal barrier for RNAPII, whereas shifted transcription initiation sites do not affect initiation per se but affect other regulatory pathways. Overall, I could reveal that integrity of TSS-chromatin is critical for efficient gene expression even in an organism that lives without transcriptional control.

Chapter 7: Discussion

In almost all organisms that contain histones, acetylation of histone H4 and the H2A variant H2A.Z are present at TSSs of RNAPII and critical for efficient transcription initiation (Haberle and Stark, 2018). However, the understanding of the cause and consequence of H2A.Z acetylation for transcription is limited, because the enzyme complexes acetylating H2A.Z can be also responsible for TSS-specific H4-acetylation, an essential premise for H2A.Z deposition.

Taking advantage of the unusual genome organisation, the lack of transcriptional regulation and the manageable number of histone-modifying proteins in *T. brucei*, I was able to disentangle the processes of H2A.Z acetylation and deposition demonstrating, for the first time, a comprehensive link between H2A.Z acetylation, deposition and transcription regulation in a eukaryotic organism.

Histones are extensively modified in *T. brucei*

By combining different histone enrichment strategies with mass-spectrometry approaches for quantitative and qualitative protein PTM analyses, I was able to generate an updated catalogue of trypanosomal histone PTMs containing 157 different histone PTMs present in *T. brucei* BF parasites including acetylation, methylation and phosphorylation. While previous datasets have been incomplete mostly due to technical limitations, the approaches used in this study covered the protein sequences of all trypanosomal histone proteins including canonical H3 and the four histone variants. Moreover, combining FIPQuant analysis and peptide counting and a newly established approach for locus-specific chromatin isolation, I could identify a large number of TSS-specific acetylation and methylation marks. Interestingly, I identified many marks associated to transcriptional activation, whereas I could only detect a few possible silencing PTMs e.g., monomethylated H4K57. This mark could be a potential homologue to methylated H4K59 in *S. cerevisiae* that functions in silent chromatin formation (Zhang et al., 2003). The low number of silencing histone PTMs could be a consequence of the trypanosome-specific histone variants H3.V and H4.V that are enriched at transcriptionally silent chromatin sites (Lowell and Cross, 2004; Siegel et al., 2009b; Müller et al., 2018). In line with that, I observed no remarkable acetylation marks on H3.V and H4.V. It might be a possibility that while higher eukaryotes possess many histone PTMs associated to repressed transcription on the canonical histones H3 and H4, *T. brucei* just expresses to specific histone variants.

While it was not surprising, that H3.V was lacking histone acetylation – an important regulator of active transcription –, I had expected a higher number of methyl marks present on H3.V. This suggests that H3.V itself is sufficient for mediating regulatory functions without the need of further modification. Nevertheless, it would be interesting, if some marks, for example H3.VK14me and H3.VK82me, are enriched at specific genomic sites. Since H3.V marks very important regulatory regions, such as TTSs,

subtelomeres and silent BESs (Lowell and Cross, 2004; Siegel et al., 2009b; Müller et al., 2018), it could be that the modification pattern of H3.V differs among those sites.

Here, I described for the first time the acetylation and methylation pattern of a eukaryotic H4 variant. Similar to H3.V, H4.V is associated to transcriptionally silent chromatin and colocalises with H3.V at TTSs and to a lesser extend also at subtelomeres (Siegel et al., 2009b; Müller et al., 2018). Thus, it was not surprising that H4.V did not contain many acetyl marks as well.

Despite the large number of PTMs, some previously reported marks, including the N-terminal acetyl marks on H2B (K4, K12 and K16), were not detected. However, out of the 157 PTMs, 126 have not been described before, including many methyl marks of H3 and the N-terminal hyperacetylation of H2A.Z and H2B.V.

In comparison to mammals, where more than 400 different histone modifications have identified up to date just on the canonical histones (Huang et al., 2014; Sidoli et al., 2019), the total number of PTMs present on all four core histones and histone variants in *T. brucei* was rather small. However, the total number was surprisingly high given the small number of histone-modifying enzymes (Ivens et al., 2005; Figueiredo et al., 2009) and the lack of RNAPII transcription regulation (Clayton, 2002).

In *T. brucei*, TSSs do not possess a defined site of RNAPII transcription initiation comparable to canonical promoters and transcription can initiate at several sites within a rather wide region (~2 kb) at the 5´end of TSSs (Wedel et al., 2017). While RNAPII transcription is usually precisely regulated in eukaryotes, in *T. brucei* RNAPII transcription lacks any regulation and initiates at similar rates at each TSS (Clayton, 2002). It has been assumed that a wider and more open chromatin structure at TSSs compared to the rest of genome would facilitate RNAPII binding (Siegel et al., 2009b; Wedel et al., 2017). One third of the identified PTMs, mainly acetyl marks, were enriched at TSSs and, assuming that many of them serve as binding sites, they probably recruit other proteins to TSSs. These properties indicate that it is important for the parasite that transcription initiation occurs at defined regions highlighted by different PTMs and the histone variants H2A.Z and H2B.V.

Across all eukaryotes, histone H3 contains the largest number of modification sites including many conserved acetyl marks, which are important for transcriptional activation (Li et al., 2007). Thus, it was quite unexpected that in *T. brucei* almost not H3 acetylation was detected except for the low abundant (~5%) H3K23 acetyl-mark at TSSs. Previous attempts to analyse the PTM pattern of H3 in *T. brucei* failed, because the N-terminus was blocked for Edman-sequencing, indicating that H3 is highly modified (Janzen et al., 2006a; Mandava et al., 2007). Indeed, I could identify all lysines of H3 as methylation sites. Interestingly, the genome of *T. brucei* encodes for 25 histone methyltransferases and four demethylases (Ivens et al., 2005), which comprise together the largest group of all putative chromatin modifiers in the parasite. This suggests that histone methylation is involved in defining different chromatin structures. Since it has been shown that the genome of *T. brucei* is compartmentalised in

regions differing in the degree of chromatin compaction (Müller et al., 2018), it would be worth to analyse if some of the described methyl marks are responsible for establishing the nuclear architecture. While the histones of *T. brucei* are quite divergent compared to those from more complex eukaryotes, they share a high sequence similarity (~90%) to histones of *T. cruzi*, another member of the trypanosomatidae family (Aslett et al., 2010). However, most of the identified PTMs were not detected in *T. cruzi* suggesting that they either might have a *T. brucei*-specific role or have not yet been identified. For example, the phosphorylated S130 has not been detected in *T. cruzi*. Serine-phosphorylated H2A (γH2A) is a hallmark of the DNA damage response in many eukaryotes (Rogakou et al., 1998). However, it has been shown that in *T. brucei*, sites of DNA damage are enriched with H2A phosphorylated at Thr^{130} (Glover and Horn, 2012). Nevertheless, it could be that Ser^{130}-phosphorylation is also involved in the DNA damage signalling pathway.

Other marks, such as the C-terminal acetyl marks of H2A, could be identified in both trypanosome species, but have no obvious homologues in other eukaryotes (Zhao and Garcia, 2015).

The trypanosomal genome organisation and gene expression pattern results in genomic regions marked by identical chromatin structures. For example, ChIP-seq data have suggested that the 200 TSSs of *T. brucei* have similar levels of H2A.Z, H2B.V and histone acetylation (Siegel et al., 2009b). Using the parasite as a model system for chromatin biology, it should be possible to shed light on many fundamental processes that involve chromatin reorganisation and in their evolutionary development. Thus, the comprehensive histone PTM dataset of *T. brucei* generated in this study provides valuable information to the eukaryotic chromatin biology.

Moreover, studies of the parasite-specific modifications regarding their genomic distribution and their associated proteins should increase the understanding of chromatin-related processes in *T. brucei*, such as derepression of a VSG expression site. It should be noted here, that for analysing the stage-specific role of histone PTMs, it is important to complete my dataset with a comprehensive PTM list found on *T. brucei* PF histones.

The process of H2A.Z deposition is highly conserved in eukaryotes

While the process of targeted H2A.Z deposition has been extensively studied in many model organisms, such as *S. cerevisiae*, *D. melanogaster* and *H. sapiens*, very little is known about the mechanism in *T. brucei*. The lessons learned from my data let conclude that H2A.Z depositioning relies on the similar mechanisms to other eukaryotes with some exceptions. While in many organisms, H4 as well as H2A.Z acetylation can stimulate H2A.Z deposition, I only could observe a defect in H2A.Z depositioning after loss of H4 acetylation. H2A.Z acetylation seems to be not involved. Moreover, in other organisms, HAT-complexes can mediate both H2A.Z and H4 acetylation. Whereas in *T. brucei*, I could identify that two HATs are required. HAT1 was responsible for H2A.Z acetylation and HAT2 for H4 acetylation. Both HATs

belong to the MYST-family similar to the HATs involved in H2A.Z depositioning in other eukaryotes suggesting that the mechanisms mediating H4 and H2A.Z acetylation have also been highly conserved during eukaryotic evolution, not only the link between H4 acetylation and targeted H2A.Z deposition. In To my opinion, this could mean that incorporation of H2A.Z into a specific chromatin region is one of the most relevant processes in eukaryotic cell biology, otherwise I would expect more mechanistic variations.

That H4 acetylation stimulates H2A.Z deposition would not work efficiently without bromodomain factors – the reader proteins of acetyl marks.

Consequently, to obtain a comprehensive understanding of the H2A.Z deposition process, I analysed the role of BDFs for H2A.Z deposition in *T. brucei*. It was not surprising that only the TSSs-enriched BDFs – BDF1, BDF3 and BDF4 – had a large effect on the H2A.Z incorporation pattern resulting in wider H2A.Z-containing peaks at TSSs that spread into the 5´end of the PTUs. However, the extent of reduction in H2A.Z levels was less compared to the levels after HAT2 depletion. H2A.Z was still enriched at TSSs following BDF1, BDF3 or BDF4 depletion suggesting that they are not key players in the targeted H2A.Z deposition process in *T. brucei*. Does that mean that BDFs are not as relevant as anticipated? I actually don't think so. Because of my findings, I hypothesise that BDF1, BDF3 and BDF4 are synergistically implicated in H2A.Z depositioning. Thus, depletion of only one out of the three would not lead to too detrimental effects because the other two can compensate the loss.

Even though depleting BDF1, BDF3 as well as BDF4 resulted in similar patterns, the extent of effect varied among them with the strongest effect following BDF3 depletion. Since BDF3 has been shown to be part of the HAT2-enzyme complex (Vellmer et al., 2021) and HAT2-mediated H4 acetylation is a pre-requisite of H2A.Z deposition, it is possible that BDF3 is mainly involved in targeting HAT2 to the TSSs or to prevent HAT2 from binding outside of the TSSs. This could mean that BDF1 and BDF4 are more important for targeting H2A.Z to TSSs by recognising H4 acetylation.

To confirm this, the H2A.Z deposition following co-depletion of BDF1 and BDF4 could be analysed by ChIP. In general, it would be interesting to uncover the processes underlying the interaction of the BDFs with TSS-nucleosomes. Therefore, the acetyl-sites critical for BDF-binding could be identified using peptide competition assays with acetylated H4 or acetylated H2A.Z peptides as bait, as described previously (Draker et al., 2012). It might be possible that BDF1 and BDF4 recognise acetylated H4 to induce H2A.Z deposition, whereas H2A.Z acetylation may be important for BDF3-recruitment. For example, in *H. sapiens*, H2A.Z acetylation engages BDF-binding at the nucleosome (Draker et al., 2012).

Despite the growing knowledge of proteins involved in H2A.Z deposition in *T. brucei*, the mechanistic processes behind H2A.Z deposition remain unclear. So far, no homologue to the catalytically active ATPase domain of SWR-1 has been identified in *T. brucei*. Since the histone chaperone FACT has been shown to be important for removing wrongly incorporated H2A.Z, for example from gene bodies

(Jeronimo et al., 2015), and that H2A.Z-containing nucleosomes stimulate the activity of ISWI remodelers (Goldman et al., 2010), I had analysed the trypanosomal homologues of FACT and ISWI for their role in H2A.Z depositioning (**Appendix figure 8 c** and **d**). However, due to the high lethality caused by FACT or ISWI depletion, the results gained from this analysis were not reproducible. Thus, I did not continue with analysing their role in H2A.Z dispositioning.

To screen for factors mediating H2A.Z deposition in *T. brucei*, I would suggest co-immunoprecipitation experiments isolating proteins associated to TSSs. Since H2A.Z-containing nucleosomes are unstable (Siegel et al., 2009b), I would not use H2A.Z as bait for such analyses. It is likely that TSS-nucleosomes fall apart during the experiment leading to dissociation of the bound proteins. Furthermore, interaction between the chromatin remodelling complex and H2A.Z might be transient. I rather recommend analysing proteins bound to HAT1, HAT2, BDF1, BDF3 and BDF4, respectively, because they are all required for H2A.Z deposition and acetylation in *T. brucei*. Moreover, to reveal whether FACT and ISWI homologue play a role in H2A.Z incorporation, I would suggest to perform H2A.Z-ChIP-experiments at earlier time points than applied in this study, for example after 6 h and 12 h of depletion.

H2A.Z - a guardian of RNAPII transcription

Given the absence of precise sites of RNAPII transcription initiation at TSSs and identical transcription initiation rates among all PTUs (Wedel et al., 2017), it has been anticipated that RNAPII-mediated transcription is not regulated in *T. brucei* (Clayton, 2002) and that RNAPII transcription initiation occurs due to a higher DNA-accessibility at TSSs compared to other genomic sites (Wedel et al., 2017).

While my work does not provide sufficient data to prove the opposite, it revealed that the role of TSS-specific chromatin composition goes beyond facilitating DNA-binding of RNAPII. H2A.Z acetylated or not acetylated defines the sites of transcription initiation for correct RNAPII binding probably functioning as an epigenetic marker of TSSs, as described other eukaryotes. Whereas H2A.Z acetylation was important for the transcription process itself allowing efficient RNAPII initiation. It seems plausible that reducing H2A.Z acetylation levels increase the nucleosomal barrier at TSSs. Consequently, bound RNAPII becomes stuck at those sites and transcription cannot initiate efficiently.

In general, acetylated H2A.Z makes the nucleosome better accessible for chromatin modifying proteins (Ren and Gorovsky, 2001). While a nucleosome containing unmodified H2A.Z resembles a canonical nucleosome (Suto et al., 2000), H2A.Z acetylation changes the nucleosomal structure leading to dissociation of the tail, thereby reducing DNA-histone interactions. In many eukaryotes including *S. cerevisiae* (Millar et al., 2006), *G. gallus* (Bruce et al., 2005), *T. thermophila* (Ren and Gorovsky, 2001) and also *T. brucei* (Lowell et al., 2005), the first half of the H2A.Z protein contains a stretch of multiple lysines, which can form an “acidic patch” because of the natural, positive charge of lysine. Regarding the composition of a nucleosome, the acidic patch of unmodified H2A.Z could increase the molecular

forces of the histone-DNA-interaction stabilizing the nucleosomal core particle. For example, in *T. thermophilia* reducing of the acidic patch of H2A.Z is essential for the organism's viability (Ren and Gorovsky, 2001), similar to what I describe in this study.

It is also possible that specific acetyl marks of H2A.Z are important for the regulatory function in *T. brucei*. To analyse, if in *T. brucei* acetylated H2A.Z functions by generating a charge patch or if specific acetyl marks are important for transcription regulation, mutagenesis screens can be performed mimicking or removing lysine acetylation (Dou and Gorovsky, 2000). If an acetyl marks mediates a site-specific histone code, mimicking lysine acetylation would still influence the cellular phenotype. In case the overall charge is important, modulation of all acetylation sites to unacetylable amino acids would have similar effects than HAT1 depletion. Based on different studies measuring higher levels of acetylated than unmodified H2A.Z at active promoters (Bruce et al., 2005; Millar et al., 2006; Valdes-Mora et al., 2012), it has been hypothesised that H2A.Z acetylation has anti-silencing functions required for active transcription. Consequently, H2A.Z acetylation can be more critical for an organism than unmodified H2A.Z. Overall, this is in line with what I also observed. Thus, my findings confirm a clear link between H2A.Z acetylation and active transcription in an early diverged eukaryote indicating that the role of H2A.Z acetylation has been highly conserved during evolution.

H2A.Z acetylation and deposition are essential for transcription in *T. brucei* since interfering with both processes is lethal for the parasite. However, my data clearly showed that the role of H2A.Z deposition goes beyond bringing H2A.Z to TSSs, so that it can be acetylated. Incorporating H2A.Z at TSSs seems to be a relevant prerequisite for correct RNAPII recruitment. Even though, mislocalised RNAPII might not change the transcription process itself, since the overall transcript levels were not strongly affected, it results in harmful side effects. This suggests that even in a parasite that lacks any regulatory control of transcription defined sites of transcription – sites marked by H2A.Z incorporation and histone acetylation – are required. A shift of transcription initiation led to antisense transcription of some essential genes, which are located at the 5´end of TSSs. Given the presence of a highly efficient RNAi-pathway in *T. brucei*, it is very likely that the antisense transcripts trigger silencing of the corresponding genes.

RNA interference is an important pathway for targeted mRNA degradation. RNAi is induced by gene-specific, double-stranded RNA (dsRNA), which are processed by the two ribonucleases Dicer and Argonaute protein (Ago). Dicer generates small RNAs (sRNAs), which loaded into Ago guide the RNA-induced silencing complex (RISC) to the corresponding mRNA. The small RNAs anneal to the transcripts and RISC cleaves across the formed duplex. Two classes of small, non-coding RNAs are involved in the RNAi-pathway, the small interfering RNAs (siRNAs) processed from long dsRNAs and micro RNAs (miRNAs), which derive from regions of the mRNA transcript that form short hairpin structures on

themselves. Antisense transcription generates complementary RNA, which bind to the corresponding sense transcript forming RNA duplexes that can be fed into the RNAi-pathway.

The genome of *T. brucei* encodes for a single *Ago1* homologue and two Dicer homologues, the nuclear *DCL2* and the cytoplasmic *DCL1*. While the nuclear Dicer is mainly involved in silencing transcripts synthesised from retrotransposon hotspots or repeat region, the role of the cytoplasmic DCL1 is less clear, since deletion of *DCL1* does not shut down the endogenous RNAi-response. In contrast, ablation of *Ago1* completely disables RNAi (Shi et al., 2004).

Thus, if the antisense transcripts observed after HAT2 depletion induce RNAi-mediated degradation of essential genes, silencing of the RNAi-pathway of *T. brucei*, for example by deletion of *Ago1*, should recover the growth phenotype of HAT2-depleted cells. However, the HAT2 depletion approach applied in this study includes overexpression of intramolecular, stem-loop dsRNA and requires an intact RNAi-machinery to degrade HAT2-specific transcripts. To overcome this problem, a cell line can be generated, in which one allele of *HAT2* and one allele of *Ago1* were deleted and the second allele of both genes can be removed in a conditional manner. Moreover, the previously established toolkit for CRISPR/Cas9-mediated genome editing in *T. brucei* (Rico et al., 2018) can be used for conditional silencing of *HAT2* and *Ago1.* Using the CRISPR/Cas9 approach would allow to remove both alleles of *HAT2* and *Ago1* in a single step. Overall, both approaches can reveal whether the shift of transcription initiation following HAT2 depletion triggers degradation of essential genes and confirm that the growth phenotype following HAT2 depletion is a consequence of the increased antisense transcription at TSSs.

Moreover, it might be that the changes in transcription start site selection perturb many downstream processes of transcription. Even though, my findings are preliminary, they suggest that *trans*-splicing is not fully functional after loss of H2A.Z deposition. Interestingly, previous studies propose a role of H2A.Z in co-transcriptional splicing (Neves et al., 2017). For example, H2A.Z is important for spliceosome assembly and promotes splicing of weak introns in *S. cerevisiae* (Neves et al., 2017; Nissen et al., 2017).

During *trans*-splicing, the mRNA cap-structure supplied by the spliced leader mini exon is added to the 5´splice site at the 5´UTR of each gene (Clayton, 2019). Since the spliced leader synthesis is RNAPII-dependent (Gilinger and Bellofatto, 2001), it is likely that the *trans*-spliceosome assembly may be coupled to RNAPII. So far, based on the finding that RNAPI- and RNAPII-synthesised transcripts are processed with the similar efficiencies, it has been suggested that *trans-splicing* and RNAPII transcription are uncoupled processes in *T. brucei* (Stewart et al., 2010). However, my findings, even though they are preliminary, suggest actually the opposite. Thus, more experiments are required to unravel whether the processes of RNAPII transcription and *trans*-splicing are coupled or not and whether H2A.Z plays a role in aligning them to each other, as shown in yeast (Neves et al., 2017; Nissen et al., 2017). To determine the impact of reduced H2A.Z levels on *trans*-splicing, the efficiency of the

spliceosome assembly can be assessed by analysing the enrichment sites of different components of the spliceosome. Moreover, splice leader trapping (SLT (Nilsson et al., 2010)) can be used to analyse the *trans*-splicing efficiency after HAT2 depletion. Thereby, the RNA should be spiked with external control RNAs. However, the ERCC spike-ins should be avoided, because the second strand synthesis of SLT is based on a spliced-leader specific oligo.

In this regard, it is important to investigate the role of H2A.Z for RNAPII pausing and transcription elongation in *T. brucei*. It seems plausible that changing transcription initiation has consequences on all processes relying on efficient RNAPII progression. Moreover, H2A.Z associates to elongation factors (Santisteban et al., 2011).

To determine the rate of RNAPII progression in dependency of H2A.Z deposition at TSSs, I would suggest to measure the levels of nascent transcripts in a cell line, which expresses a reporter gene, such as the firefly luciferase, under control of the spliced leader promoter and allows inducible RNAi against HAT2. By taking advantage of the spliced leader promoter, it should be possible to obtain high transcript levels of the reporter gene.

Overall, my findings clearly reveal that the TSS-chromatin structure is essential for efficient RNAPII progression. While it has been suggested that an open chromatin structure is sufficient for transcription initiation in *T. brucei*, the drastic effects on RNA synthesis and processing observed after changing the composition of TSS-nucleosomes indicate that transcription by RNAPII is more complex than previously assumed.

Conclusion

The chromatin structure is one of the most important means for transcriptional regulation in all eukaryotic organisms. Local and global changes of the chromatin structure influence the DNA accessibility resulting in more compacted and less accessible or more open and well accessible genomic sites. The experiments I performed during my PhD revealed that in eukaryotes transcription initiation is strongly dependent on a specific chromatin structure at TSSs not only to mediate DNA accessibility for the transcription machinery but also to ensure precise coordination of RNA synthesis and maturation. While the chromatin structure at TSSs is highly conserved throughout the eukaryotic lineage, complex and interdependent regulatory processes have restricted further studies of its regulatory role. Thus, my findings also approve the early diverged unicellular parasite *T. brucei* as a valuable model system for studying eukaryotic chromatin biology and let hypothesise that TSS-specific chromatin represents one of the earliest developed regulator of transcription initiation.

Whilst I disentangled the role of H2A.Z deposition and acetylation for gene expression, it would be interesting to know if they are important for the genome architecture in *T. brucei*. Since transcription is not regulated and is always on, it is likely that the TSS-specific chromatin structure is involved in

positioning the TSSs within the nucleus to generate genomic compartments easily accessible for the transcription machinery.

Chapter 8: Bibliography

1. Adam M, Robert F, Larochelle M, 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.
2. Adelman K, Lis JT. 2012. Promoter-proximal pausing of RNA polymerase II: emerging roles in metazoans. *Nat Rev Genet* **13**: 720–731.
3. Albert I, Mavrich TN, Tomsho LP, Qi J, Zanton SJ, Schuster SC, Pugh BF. 2007. Translational and rotational settings of H2A.Z nucleosomes across the Saccharomyces cerevisiae genome. *Nature* **446**: 572–576.
4. Albig W, Trappe R, Kardalinou E, Eick S, Doenecke D. 1999. The human H2A and H2B histone gene complement. *Biol Chem* **380**: 7–18.
5. Allard S, Utley RT, Savard J, Clarke A, Grant P, Brandl CJ, Pillus L, Workman JL, Côté J. 1999. NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *EMBO J* **18**: 5108–5119.
6. Allfrey VG, Faulkner R, Mirsky AE. 1964. Acetylation and methylation of histones and their possible role in the regulation of RNA synthesis. *Proc Natl Acad Sci U S A* **51**: 786–794.
7. Allis CD, Glover CV, Bowen JK, Gorovsky MA. 1980. Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, Tetrahymena thermophila. *Cell* **20**: 609–617.
8. Alsford S, Horn D. 2004. Trypanosomatid histones. *Mol Microbiol* **53**: 365–372.
9. Alsford S, Horn D. 2008. Single-locus targeting constructs for reliable regulated RNAi and transgene expression in Trypanosoma brucei. *Mol Biochem Parasitol* **161**: 76–79.
10. Alsford S, Kawahara T, Glover L, Horn D. 2005. Tagging a T. brucei RRNA locus improves stable transfection efficiency and circumvents inducible expression position effects. *Mol Biochem Parasitol* **144**: 142–148.
11. Altaf M, Auger A, Monnet-Saksouk J, Brodeur J, Piquet S, Cramet M, Bouchard N, Lacoste N, Utley RT, Gaudreau L, Cote J. 2010. NuA4-dependent acetylation of nucleosomal histones H4 and H2A directly stimulates incorporation of H2A.Z by the SWR1 complex. *J Biol Chem* **285**: 15966–15977.
12. Annunziato AT, Frado LL, Seale RL, Woodcock CL. 1988. Treatment with sodium butyrate inhibits the complete condensation of interphase chromatin. *Chromosoma* **96**: 132–138.
13. Antosz W, Pfab A, Ehrnsberger HF, Holzinger P, Köllen K, Mortensen SA, Bruckmann A, Schubert T, Längst G, Griesenbeck J, Schubert V, Grasser M, Grasser KD. 2017. The Composition of the Arabidopsis RNA Polymerase II Transcript Elongation Complex Reveals the Interplay between Elongation and mRNA Processing Factors. *Plant Cell* **29**: 854–870.
14. Aslett M, Aurrecoechea C, Berriman M, Brestelli J, Brunk BP, Carrington M, Depledge DP, Fischer S, Gajria B, Gao X, Gardner MJ, Gingle A, Grant G, Harb OS, Heiges M, Hertz-Fowler C, Houston R, Innamorato F, Iodice J, Kissinger JC, Kraemer E, Li W, Logan FJ, Miller JA, Mitra S, Myler PJ, Nayak V, Pennington C, Phan I, Pinney DF, Ramasamy G, Rogers MB, Roos DS, Ross C, Sivam D, Smith DF, Srinivasamoorthy G, Stoeckert CJ, Subramanian S, Thibodeau R, Tivey A, Treatman C, Velarde G, Wang H. 2010. TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res* **38**: D457–62.
15. Aslund L, Carlsson L, Henriksson J, Rydåker M, Toro GC, Galanti N, Pettersson U. 1994. A gene family encoding heterogeneous histone H1 proteins in Trypanosoma cruzi. *Mol Biochem Parasitol* **65**: 317–330.
16. Asturias FJ, Chung WH, Kornberg RD, Lorch Y. 2002. Structural analysis of the RSC chromatin-remodeling complex. *Proc Natl Acad Sci U S A* **99**: 13477–13480.
17. Auger A, Galarneau L, Altaf M, Nourani A, Doyon Y, Utley RT, Cronier D, Allard S, Côté J. 2008. Eaf1 is the platform for NuA4 molecular assembly that evolutionarily links chromatin acetylation to ATP-dependent exchange of histone H2A variants. *Mol Cell Biol* **28**: 2257–2270.
18. Babiarz JE, Halley JE, Rine J. 2006. Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in Saccharomyces cerevisiae. *Genes Dev* **20**: 700–710.
19. Baeza J, Smallegan MJ, Denu JM. 2015. Site-specific reactivity of nonenzymatic lysine acetylation. *ACS Chem Biol* **10**: 122–128.
20. Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell Res* **21**: 381–395.
21. Baptista T, Grünberg S, Minoungou N, Koster MJE, Timmers HTM, Hahn S, Devys D, Tora L. 2017. SAGA Is a General Cofactor for RNA Polymerase II Transcription. *Mol Cell* **68**: 130–143.e5.
22. Bargaje R, Alam MP, Patowary A, Sarkar M, Ali T, Gupta S, Garg M, Singh M, Purkanti R, Scaria V, Sivasubbu S, Brahmachari V, Pillai B. 2012. Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. *Nucleic Acids Res* **40**: 8965–8978.
23. Barrett MP, Burchmore RJ, Stich A, Lazzari JO, Frasch AC, Cazzulo JJ, Krishna. 2003. The trypanosomiases. *Lancet* **362**: 1469–1480.
24. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. *Cell* **129**: 823–837.
25. Bartholomew B. 2014. Regulating the chromatin landscape: structural and mechanistic perspectives. *Annu Rev Biochem* **83**: 671–696.
26. Bastin P, Bagherzadeh Z, Matthews KR, Gull K. 1996. A novel epitope tag system to study protein targeting and organelle biogenesis in Trypanosoma brucei. *Mol Biochem Parasitol* **77**: 235–239.
27. Becker PB, Hörz W. 2002. ATP-dependent nucleosome remodeling. *Annu Rev Biochem* **71**: 247–273.
28. Becker PB, Workman JL. 2013. Nucleosome remodeling and epigenetics. *Cold Spring Harb Perspect Biol* **5**:
29. Bednar J, Horowitz RA, Grigoryev SA, Carruthers LM, Hansen JC, Koster AJ, Woodcock CL. 1998. Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc Natl Acad Sci U S A* **95**: 14173–14178.
30. Belli SI. 2000. Chromatin remodelling during the life cycle of trypanosomatids. *Int J Parasitol* **30**: 679–687.
31. Beneke T, Madden R, Makin L, Valli J, Sunter J, Gluenz E. 2017. A CRISPR Cas9 high-throughput genome editing toolkit for kinetoplastids. *R Soc Open Sci* **4**: 170095.

32. Benjamini Y, Drai D, Elmer G, Kafkafi N, Golani I. 2001. Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* **125**: 279–284.
33. Benne R, Van den Burg J, Brakenhoff JP, Sloof P, Van Boom JH, Tromp MC. 1986. Major transcript of the frameshifted coxII gene from trypanosome mitochondria contains four nucleotides that are not encoded in the DNA. *Cell* **46**: 819–826.
34. Bennett G, Peterson CL. 2015. SWI/SNF recruitment to a DNA double-strand break by the NuA4 and Gcn5 histone acetyltransferases. *DNA Repair (Amst)* **30**: 38–45.
35. Bentley DL. 2014. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet* **15(3)**: 163–175.
36. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. *Nucleic Acids Res* **28**: 235–242.
37. Berriman M, Ghedin E, Hertz-Fowler C, Blandin G, Renauld H, Bartholomeu DC, Lennard NJ, Caler E, Hamlin NE, Haas B, Bohme U, Hannick L, Aslett MA, Shallom J, Marcello L, Hou L, Wickstead B, Alsmark UC, Arrowsmith C, Atkin RJ, Barron AJ, Bringaud F, Brooks K, Carrington M, Cherevach I, Chillingworth TJ, Churcher C, Clark LN, Corton CH, Cronin A, Davies RM, Doggett J, Djikeng A, Feldblyum T, Field MC, Fraser A, Goodhead I, Hance Z, Harper D, Harris BR, Hauser H, Hostetler J, Ivens A, Jagels K, Johnson D, Johnson J, Jones K, Kerhornou AX, Koo H, Larke N, Landfear S, Larkin C, Leech V, Line A, Lord A, Macleod A, Mooney PJ, Moule S, Martin DM, Morgan GW, Mungall K, Norbertczak H, Ormond D, Pai G, Peacock CS, Peterson J, Quail MA, Rabbinowitsch E, Rajandream MA, Reitter C, Salzberg SL, Sanders M, Schobel S, Sharp S, Simmonds M, Simpson AJ, Tallon L, Turner CM, Tait A, Tivey AR, Van Aken S, Walker D, Wanless D, Wang S, White B, White O, Whitehead S, Woodward J, Wortman J, Adams MD, Embley TM, Gull K, Ullu E, Barry JD, Fairlamb AH, Opperdoes F, Barrell BG, Donelson JE, Hall N, Fraser CM, Melville SE, El-Sayed NM. 2005. The genome of the African trypanosome Trypanosoma brucei. *Science* **309**: 416–422.
38. Bickis I, von Bertalanffy L. 1956. Identification of cytoplasmic basophilia (ribonucleic acid) by fluorescence microscopy. *J Histochem Cytochem* **4**: 481–493.
39. Billon P, Cote J. 2013. Precise deposition of histone H2A.Z in chromatin for genome expression and maintenance. *Biochim Biophys Acta* **1819**: 290–302.
40. Biterge B, Schneider R. 2014. Histone variants: key players of chromatin. *Cell Tissue Res* **356**: 457–466.
41. Bloch K, Borek E. 1946. Biological acetylation of natural amino acids. *J Biol Chem* **164**: 483.
42. Bondarenko VA, Steele LM, Ujvari A, Gaykalova DA, Kulaeva OI, Polikanov YS, Luse DS, Studitsky VM. 2006. Nucleosomes can form a polar barrier to transcript elongation by RNA polymerase II. *Mol Cell* **24**: 469–479.
43. Bonenfant D, Coulot M, Towbin H, Schindler P, van Oostrum J. 2006. Characterization of histone H2A and H2B variants and their post-translational modifications by mass spectrometry. *Mol Cell Proteomics* **5**: 541–552.
44. Bonenfant D, Towbin H, Coulot M, Schindler P, Mueller DR, van Oostrum J. 2007. Analysis of dynamic changes in post-translational modifications of human histones during cell cycle by mass spectrometry. *Mol Cell Proteomics* **6**: 1917–1932.
45. Bönisch C, Hake SB. 2012. Histone H2A variants in nucleosomes and chromatin: more or less stable. *Nucleic Acids Res* **40**: 10719–10741.
46. Borst P. 2002. Antigenic variation and allelic exclusion. *Cell* **109**: 5–8.
47. Boyle AP, Davis S, Shulha HP, Meltzer P, Margulies EH, Weng Z, Furey TS, Crawford GE. 2008. High-resolution mapping and characterization of open chromatin across the genome. *Cell* **132**: 311–322.
48. Brown CE, Howe L, Sousa K, Alley SC, Carrozza MJ, Tan S, Workman JL. 2001. Recruitment of HAT complexes by direct activator interactions with the ATM-related Tra1 subunit. *Science* **292**: 2333–2337.
49. Brownell JE, Allis CD. 1996. Special HATs for special occasions: linking histone acetylation to chromatin assembly and gene activation. *Curr Opin Genet Dev* **6**: 176–184.
50. Bruce K, Myers FA, Mantouvalou E, Lefevre P, Greaves I, Bonifer C, Tremethick DJ, Thorne AW, Crane-Robinson C. 2005. The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. *Nucleic Acids Res* **33**: 5633–5639.
51. Buschbeck M, Hake SB. 2017. Variants of core histones and their roles in cell fate decisions, development and cancer. *Nat Rev Mol Cell Biol* **18**: 299–314.
52. Bütikofer P, Ruepp S, Boschung M, Roditi I. 1997. GPEET' procyclin is the major surface protein of procyclic culture forms of Trypanosoma brucei brucei strain 427. *Biochem J* **326**: 415–423.
53. Campos EI, Reinberg D. 2009. Histones: annotating chromatin. *Annu Rev Genet* **43**: 559–599.
54. Capewell P, Cren-Travaillé C, Marchesi F, Johnston P, Clucas C, Benson RA, Gorman TA, Calvo-Alvarez E, Crouzols A, Jouvion G, Jamonneau V, Weir W, Stevenson ML, O'Neill K, Cooper A, Swar NK, Bucheton B, Ngoyi DM, Garside P, Rotureau B, MacLeod A. 2016. The skin is a significant but overlooked anatomical reservoir for vector-borne African trypanosomes. *Elife* **5**:
55. Carr AM, Dorrington SM, Hindley J, Phear GA, Aves SJ, Nurse P. 1994. Analysis of a histone H2A variant from fission yeast: evidence for a role in chromosome stability. *Mol Gen Genet* **245**: 628–635.
56. Carrillo Oesterreich F, Bieberstein N, Neugebauer KM. 2011. Pause locally, splice globally. *Trends Cell Biol* **21**: 328–335.
57. Carrillo Oesterreich F, Preibisch S, Neugebauer KM. 2010. Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol Cell* **40**: 571–581.
58. Chadwick BP, Willard HF. 2001. Histone H2A variants and the inactive X chromosome: identification of a second macroH2A variant. *Hum Mol Genet* **10**: 1101–1113.
59. Chapman AB, Agabian N. 1994. Trypanosoma brucei RNA polymerase II is phosphorylated in the absence of carboxyl-terminal domain heptapeptide repeats. *J Biol Chem* **269**: 4754–4760.
60. Chen FX, Smith ER, Shilatifard A. 2018. Born to run: control of transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **19**: 464–478.
61. Cheng X, Auger A, Altaf M, Drouin S, Paquet E, Utley RT, Robert F, Côté J. 2015. Eaf1 Links the NuA4 Histone Acetyltransferase Complex to Htz1 Incorporation and Regulation of Purine Biosynthesis. *Eukaryot Cell* **14**: 535–544.
62. Cherry JM, Hong EL, Amundsen C, Balakrishnan R, Binkley G, Chan ET, Christie KR, Costanzo MC, Dwight SS, Engel SR, Fisk DG, Hirschman JE, Hitz BC, Karra K, Krieger CJ, Miyasato SR, Nash RS, Park J, Skrzypek MS, Simison

M, Weng S, Wong ED. 2012. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res* **40**: D700–5.
63. Clapier CR, Cairns BR. 2009. The biology of chromatin remodeling complexes. *Annu Rev Biochem* **78**: 273–304.
64. Clapier CR, Iwasa J, Cairns BR, Peterson CL. 2017. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**: 407–422.
65. Clarkson MJ, Wells JR, Gibson F, Saint R, Tremethick DJ. 1999. Regions of variant histone His2AvD required for Drosophila development. *Nature* **399**: 694–697.
66. Clayton C. 2019. Regulation of gene expression in trypanosomatids: living with polycistronic transcription. *Open Biol* **9**: 190072.
67. Clayton CE. 2002. Life without transcriptional control? From fly to man and back again. *EMBO J* **21**: 1881–1888.
68. Cliffe LJ, Kieft R, Southern T, Birkeland SR, Marshall M, Sweeney K, Sabatini R. 2009. JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes. *Nucleic Acids Res* **37**: 1452–1462.
69. Conaway JW, Shilatifard A, Dvir A, Conaway RC. 2000. Control of elongation by RNA polymerase II. *Trends Biochem Sci* **25**: 375–380.
70. Corona DF, Längst G, Clapier CR, Bonte EJ, Ferrari S, Tamkun JW, Becker PB. 1999. ISWI is an ATP-dependent nucleosome remodeling factor. *Mol Cell* **3**: 239–245.
71. Cramer P. 2004. RNA polymerase II structure: from core to functional complexes. *Curr Opin Genet Dev* **14**: 218–226.
72. Cramer P, Bushnell DA, Fu J, Gnatt AL, Maier-Davis B, Thompson NE, Burgess RR, Edwards AM, David PR, Kornberg RD. 2000. Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* **288**: 640–649.
73. Cramer P, Bushnell DA, Kornberg RD. 2001. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**: 1863–1876.
74. Crick FH. 1958. On protein synthesis. *Symp Soc Exp Biol* **12**: 138–163.
75. Crick FH. 1968. The origin of the genetic code. *J Mol Biol* **38**: 367–379.
76. Cross GA. 1975. Identification, purification and properties of clone-specific glycoprotein antigens constituting the surface coat of Trypanosoma brucei. *Parasitology* **71**: 393–417.
77. Cross GA, Kim HS, Wickstead B. 2014. Capturing the variant surface glycoprotein repertoire (the VSGnome) of Trypanosoma brucei Lister 427. *Mol Biochem Parasitol* **195**: 59–73.
78. da Cunha JP, Nakayasu ES, de Almeida IC, Schenkman S. 2006. Post-translational modifications of Trypanosoma cruzi histone H4. *Mol Biochem Parasitol* **150**: 268–277.
79. Das A, Banday M, Fisher MA, Chang YJ, Rosenfeld J, Bellofatto V. 2017. An essential domain of an early-diverged RNA polymerase II functions to accurately decode a primitive chromatin landscape. *Nucleic Acids Res* **45**: 7886–7896.
80. Das A, Bellofatto V. 2003. RNA polymerase II-dependent transcription in trypanosomes is associated with a SNAP complex-like transcription factor. *Proc Natl Acad Sci U S A* **100**: 80–85.
81. Das A, Bellofatto V. 2009. The non-canonical CTD of RNAP-II is essential for productive RNA synthesis in Trypanosoma brucei. *PLoS One* **4**: e6959.
82. Das A, Li H, Liu T, Bellofatto V. 2006. Biochemical characterization of Trypanosoma brucei RNA polymerase II. *Mol Biochem Parasitol* **150**: 201–210.
83. Das A, Zhang Q, Palenchar JB, Chatterjee B, Cross GA, Bellofatto V. 2005. Trypanosomal TBP functions with the multisubunit transcription factor tSNAP to direct spliced-leader RNA gene expression. *Mol Cell Biol* **25**: 7314–7322.
84. de Jesus TC, Nunes VS, Lopes MC, Martil DE, Iwai LK, Moretti NS, Machado FC, de Lima-Stein ML, Thiemann OH, Elias MC, Janzen C, Schenkman S, da Cunha JP. 2016. Chromatin Proteomics Reveals Variable Histone Modifications during the Life Cycle of Trypanosoma cruzi. *J Proteome Res* **15**: 2039–2051.
85. de la Mata M, Alonso CR, Kadener S, Fededa JP, Blaustein M, Pelisch F, Cramer P, Bentley D, Kornblihtt AR. 2003. A slow RNA polymerase II affects alternative splicing in vivo. *Mol Cell* **12**: 525–532.
86. De Lange T, Michels PA, Veerman HJ, Cornelissen AW, Borst P. 1984. Many trypanosome messenger RNAs share a common 5' terminal sequence. *Nucleic Acids Res* **12**: 3777–3790.
87. Deal RB, Topp CN, McKinney EC, Meagher RB. 2007. Repression of flowering in Arabidopsis requires activation of flowering locus c expression by the histone variant H2A.Z. *Plant Cell* **19**: 74–83.
88. Dean S, Sunter J, Wheeler RJ, Hodkinson I, Gluenz E, Gull K. 2015. A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol* **5**: 140197.
89. Déjardin J, Kingston RE. 2009. Purification of proteins associated with specific genomic Loci. *Cell* **136**: 175–186.
90. Denninger V, Fullbrook A, Bessat M, Ersfeld K, Rudenko G. 2010. The FACT subunit TbSpt16 is involved in cell cycle specific control of VSG expression sites in Trypanosoma brucei. *Mol Microbiol* **78**: 459–474.
91. Denninger V, Rudenko G. 2014. FACT plays a major role in histone dynamics affecting VSG expression site control in Trypanosoma brucei. *Mol Microbiol* **94**: 945–962.
92. DesJarlais R, Tummino PJ. 2016. Role of Histone-Modifying Enzymes and Their Complexes in Regulation of Chromatin Biology. *Biochemistry* **55**: 1584–1599.
93. Devaux S, Lecordier L, Uzureau P, Walgraffe D, Dierick JF, Poelvoorde P, Pays E, Vanhamme L. 2006. Characterization of RNA polymerase II subunits of Trypanosoma brucei. *Mol Biochem Parasitol* **148**: 60–68.
94. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, Zhou MM. 1999. Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**: 491–496.
95. Dion MF, Altschuler SJ, Wu LF, Rando OJ. 2005. Genomic characterization reveals a simple histone H4 acetylation code. *Proc Natl Acad Sci U S A* **102**: 5501–5506.
96. DiPaolo C, Kieft R, Cross M, Sabatini R. 2005. Regulation of trypanosome DNA glycosylation by a SWI2/SNF2-like protein. *Mol Cell* **17**: 441–451.
97. Donelson JE. 2003. Antigenic variation and the African trypanosome genome. *Acta Trop* **85**: 391–404.
98. Dorigo B, Schalch T, Bystricky K, Richmond TJ. 2003. Chromatin fiber folding: requirement for the histone H4 N-terminal tail. *J Mol Biol* **327**: 85–96.

99. Dou Y, Gorovsky MA. 2000. Phosphorylation of linker histone H1 regulates gene expression in vivo by creating a charge patch. *Mol Cell* **6**: 225–231.
100. Draker R, Ng MK, Sarcinella E, Ignatchenko V, Kislinger T, Cheung P. 2012. A combination of H2A.Z and H4 acetylation recruits Brd2 to chromatin during transcriptional activation. *PLoS Genet* **8**: e1003047.
101. Durand-Dubief M, Kohl L, Bastin P. 2003. Efficiency and specificity of RNA interference generated by intra- and intermolecular double stranded RNA in Trypanosoma brucei. *Mol Biochem Parasitol* **129**: 11–21.
102. Durant M, Pugh BF. 2007. NuA4-directed chromatin transactions throughout the Saccharomyces cerevisiae genome. *Mol Cell Biol* **27**: 5327–5335.
103. Durrin LK, Mann RK, Kayne PS, Grunstein M. 1991. Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* **65**: 1023–1031.
104. Dvir A. 2002. Promoter escape by RNA polymerase II. *Biochim Biophys Acta* **1577**: 208–223.
105. El-Sayed NM, Myler PJ, Bartholomeu DC, Nilsson D, Aggarwal G, Tran AN, Ghedin E, Worthey EA, Delcher AL, Blandin G, Westenberger SJ, Caler E, Cerqueira GC, Branche C, Haas B, Anupama A, Arner E, Aslund L, Attipoe P, Bontempi E, Bringaud F, Burton P, Cadag E, Campbell DA, Carrington M, Crabtree J, Darban H, da Silveira JF, de Jong P, Edwards K, Englund PT, Fazelina G, Feldblyum T, Ferella M, Frasch AC, Gull K, Horn D, Hou L, Huang Y, Kindlund E, Klingbeil M, Kluge S, Koo H, Lacerda D, Levin MJ, Lorenzi H, Louie T, Machado CR, McCulloch R, McKenna A, Mizuno Y, Mottram JC, Nelson S, Ochaya S, Osoegawa K, Pai G, Parsons M, Pentony M, Pettersson U, Pop M, Ramirez JL, Rinta J, Robertson L, Salzberg SL, Sanchez DO, Seyler A, Sharma R, Shetty J, Simpson AJ, Sisk E, Tammi MT, Tarleton R, Teixeira S, Van Aken S, Vogt C, Ward PN, Wickstead B, Wortman J, White O, Fraser CM, Stuart KD, Andersson B. 2005a. The genome sequence of Trypanosoma cruzi, etiologic agent of Chagas disease. *Science* **309**: 409–415.
106. El-Sayed NM, Myler PJ, Blandin G, Berriman M, Crabtree J, Aggarwal G, Caler E, Renauld H, Worthey EA, Hertz-Fowler C, Ghedin E, Peacock C, Bartholomeu DC, Haas BJ, Tran AN, Wortman JR, Alsmark UC, Angiuoli S, Anupama A, Badger J, Bringaud F, Cadag E, Carlton JM, Cerqueira GC, Creasy T, Delcher AL, Djikeng A, Embley TM, Hauser C, Ivens AC, Kummerfeld SK, Pereira-Leal JB, Nilsson D, Peterson J, Salzberg SL, Shallom J, Silva JC, Sundaram J, Westenberger S, White O, Melville SE, Donelson JE, Andersson B, Stuart KD, Hall N. 2005b. Comparative genomics of trypanosomatid parasitic protozoa. *Science* **309**: 404–409.
107. ElBashir R, Vanselow JT, Kraus A, Janzen CJ, Siegel TN, Schlosser A. 2015. Fragment ion patchwork quantification for measuring site-specific acetylation degrees. *Anal Chem* **87**: 9939–9945.
108. Ersfeld K, Melville SE, Gull K. 1999. Nuclear and genome organization of Trypanosoma brucei. *Parasitol Today* **15**: 58–63.
109. Evers R, Hammer A, Köck J, Jess W, Borst P, Mémet S, Cornelissen AW. 1989. Trypanosoma brucei contains two RNA polymerase II largest subunit genes with an altered C-terminal domain. *Cell* **56**: 585–597.
110. Faast R, Thonglairoam V, Schulz TC, Beall J, Wells JR, Taylor H, Matthaei K, Rathjen PD, Tremethick DJ, Lyons I. 2001. Histone variant H2A.Z is required for early mammalian development. *Curr Biol* **11**: 1183–1187.
111. Fadda A, Färber V, Droll D, Clayton C. 2013. The roles of 3'-exoribonucleases and the exosome in trypanosome mRNA degradation. *RNA* **19**: 937–947.
112. Fei J, Torigoe SE, Brown CR, Khuong MT, Kassavetis GA, Boeger H, Kadonaga JT. 2015. The prenucleosome, a stable conformational isomer of the nucleosome. *Genes Dev* **29**: 2563–2575.
113. Feller C, Forné I, Imhof A, Becker PB. 2015. Global and specific responses of the histone acetylome to systematic perturbation. *Mol Cell* **57**: 559–571.
114. Felsenfeld G. 1978. Chromatin. *Nature* **271**: 115–122.
115. Fenley AT, Anandakrishnan R, Kidane YH, Onufriev AV. 2018. Modulation of nucleosomal DNA accessibility via charge-altering post-translational modifications in histone core. *Epigenetics Chromatin* **11**: 11.
116. Fenn K, Matthews KR. 2007. The cell biology of Trypanosoma brucei differentiation. *Curr Opin Microbiol* **10**: 539–546.
117. Ferguson MA, Homans SW, Dwek RA, Rademacher TW. 1988. Glycosyl-phosphatidylinositol moiety that anchors Trypanosoma brucei variant surface glycoprotein to the membrane. *Science* **239**: 753–759.
118. Fiedler KL, Bheda P, Dai J, Boeke JD, Wolberger C, Cotter RJ. 2013. A quantitative analysis of histone methylation and acetylation isoforms from their deuteroacetylated derivatives: application to a series of knockout mutants. *J Mass Spectrom* **48**: 608–615.
119. Field MC, Horn D, Fairlamb AH, Ferguson MAJ, Gray DW, Read KD, De Rycker M, Torrie LS, Wyatt PG, Wyllie S, Gilbert IH. 2017. Anti-trypanosomatid drug discovery: an ongoing challenge and a continuing need. *Nat Rev Microbiol* **15**: 447.
120. Figueiredo LM, Cross GA, Janzen CJ. 2009. Epigenetic regulation in African trypanosomes: a new kid on the block. *Nat Rev Microbiol* **7**: 504–513.
121. Figueiredo LM, Janzen CJ, Cross GA. 2008. A histone methyltransferase modulates antigenic variation in African trypanosomes. *PLoS Biol* **6**: e161.
122. Fischle W, Wang Y, Allis. 2003. Histone and chromatin cross-talk. *Curr Opin Cell Biol* **15**: 172–183.
123. Flaus A, Martin DM, Barton GJ, Owen-Hughes T. 2006. Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* **34**: 2887–2905.
124. Frederiks F, van Welsem T, Oudgenoeg G, Heck AJ, Janzen CJ, van Leeuwen F. 2010. Heterologous expression reveals distinct enzymatic activities of two DOT1 histone methyltransferases of Trypanosoma brucei. *J Cell Sci* **123**: 4019–4023.
125. Fu J, Gnatt AL, Bushnell DA, Jensen GJ, Thompson NE, Burgess RR, David PR, Kornberg RD. 1999. Yeast RNA polymerase II at 5 A resolution. *Cell* **98**: 799–810.
126. Fuchs H, Gailus-Durner V, Adler T, Aguilar-Pimentel JA, Becker L, Calzada-Wack J, Da Silva-Buttkus P, Neff F, Götz A, Hans W, Hölter SM, Horsch M, Kastenmüller G, Kemter E, Lengger C, Maier H, Matloka M, Möller G, Naton B, Prehn C, Puk O, Rácz I, Rathkolb B, Römisch-Margl W, Rozman J, Wang-Sattler R, Schrewe A, Stöger C, Tost M, Adamski J, Aigner B, Beckers J, Behrendt H, Busch DH, Esposito I, Graw J, Illig T, Ivandic B, Klingenspor M, Klopstock T, Kremmer E, Mempel M, Neschen S, Ollert M, Schulz H, Suhre K, Wolf E, Wurst W, Zimmer A, Hrabě de Angelis M. 2011. Mouse phenotyping. *Methods* **53**: 120–135.

127. García-Salcedo JA, Gijón P, Pays E. 1999. Regulated transcription of the histone H2B genes of Trypanosoma brucei. *Eur J Biochem* **264**: 717–723.
128. Gassen A, Brechtefeld D, Schandry N, Arteaga-Salas JM, Israel L, Imhof A, Janzen CJ. 2012. DOT1A-dependent H3K76 methylation is required for replication regulation in Trypanosoma brucei. *Nucleic Acids Res* **40**: 10302–10311.
129. Germond JE, Bellard M, Oudet P, Chambon P. 1976. Stability of nucleosomes in native and reconstituted chromatins. *Nucleic Acids Res* **3**: 3173–3192.
130. Giaimo BD, Ferrante F, Herchenröther A, Hake SB, Borggrefe T. 2019. The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin* **12**: 37.
131. Gilinger G, Bellofatto V. 2001. Trypanosome spliced leader RNA genes contain the first identified RNA polymerase II gene promoter in these organisms. *Nucleic Acids Res* **29**: 1556–1564.
132. Glover L, Horn D. 2012. Trypanosomal histone γH2A and the DNA damage response. *Mol Biochem Parasitol* **183**: 78–83.
133. Goldman JA, Garlick JD, Kingston RE. 2010. Chromatin remodeling by imitation switch (ISWI) class ATP-dependent remodelers is stimulated by histone variant H2A.Z. *J Biol Chem* **285**: 4645–4651.
134. Greaves IK, Rangasamy D, Devoy M, Marshall Graves JA, Tremethick DJ. 2006. The X and Y chromosomes assemble into H2A.Z-containing [corrected] facultative heterochromatin [corrected] following meiosis. *Mol Cell Biol* **26**: 5394–5405.
135. Greaves IK, Rangasamy D, Ridgway P, Tremethick DJ. 2007. H2A.Z contributes to the unique 3D structure of the centromere. *Proc Natl Acad Sci U S A* **104**: 525–530.
136. Gregory PD, Schmid A, Zavari M, Münsterkötter M, Hörz W. 1999. Chromatin remodelling at the PHO8 promoter requires SWI-SNF and SAGA at a step subsequent to activator binding. *EMBO J* **18**: 6407–6414.
137. Grünberg S, Hahn S. 2013. Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem Sci* **38**: 603–611.
138. Guillemette B, Bataille AR, Gevry N, Adam M, Blanchette M, Robert F, Gaudreau L. 2005. Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. *PLoS Biol* **3**: e384.
139. Guillemette B, Gaudreau L. 2006. Reuniting the contrasting functions of H2A.Z. *Biochem Cell Biol* **84**: 528–535.
140. Günzl A, Bruderer T, Laufer G, Schimanski B, Tu LC, Chung HM, Lee PT, Lee MG. 2003. RNA polymerase I transcribes procyclin genes and variant surface glycoprotein gene expression sites in Trypanosoma brucei. *Eukaryot Cell* **2**: 542–551.
141. Günzl A, Schimanski B. 2009. Tandem affinity purification of proteins. *Curr Protoc Protein Sci* **Chapter 19**: Unit 19.19.
142. Guo J. 2014. Transcription: the epicenter of gene expression. *J Zhejiang Univ Sci B* **15**: 409–411.
143. Gurard-Levin ZA, Quivy JP, Almouzni G. 2014. Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annu Rev Biochem* **83**: 487–517.
144. Haag J, O'hUigin C, Overath P. 1998. The molecular phylogeny of trypanosomes: evidence for an early divergence of the Salivaria. *Mol Biochem Parasitol* **91**: 37–49.
145. Haberle V, Stark A. 2018. Eukaryotic core promoters and the functional basis of transcription initiation. *Nat Rev Mol Cell Biol* **19**: 621–637.
146. Halley JE, Kaplan T, Wang AY, Kobor MS, Rine J. 2010. Roles for H2A.Z and its acetylation in GAL1 transcription and gene induction, but not GAL1-transcriptional memory. *PLoS Biol* **8**: e1000401.
147. Hantsche M, Cramer P. 2017. Conserved RNA polymerase II initiation complex structure. *Curr Opin Struct Biol* **47**: 17–22.
148. Hardy S, Jacques PE, Gevry N, Forest A, Fortin ME, Laflamme L, Gaudreau L, Robert F. 2009. The euchromatic and heterochromatic landscapes are shaped by antagonizing effects of transcription on H2A.Z deposition. *PLoS Genet* **5**: e1000687.
149. Hargreaves DC, Crabtree GR. 2011. ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. *Cell Res* **21**: 396–420.
150. Hartley PD, Madhani HD. 2009. Mechanisms that specify promoter nucleosome location and identity. *Cell* **137**: 445–458.
151. Hassan AH, Neely KE, Workman JL. 2001. Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* **104**: 817–827.
152. Hassan AH, Prochasson P, Neely KE, Galasinski SC, Chandy M, Carrozza MJ, Workman JL. 2002. Function and selectivity of bromodomains in anchoring chromatin-modifying complexes to promoter nucleosomes. *Cell* **111**: 369–379.
153. Haynes SR, Dollard C, Winston F, Beck S, Trowsdale J, Dawid IB. 1992. The bromodomain: a conserved sequence found in human, Drosophila and yeast proteins. *Nucleic Acids Res* **20**: 2603.
154. Hebbes TR, Thorne AW, Crane-Robinson C. 1988. A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J* **7**: 1395–1402.
155. Henikoff S. 2008. Nucleosome destabilization in the epigenetic regulation of gene expression. *Nat Rev Genet* **9**: 15–26.
156. Hergeth SP, Schneider R. 2015. The H1 linker histones: multifunctional proteins beyond the nucleosomal core particle. *EMBO Rep* **16**: 1439–1453.
157. Hersman E, Nelson DM, Griffith WP, Jelinek C, Cotter RJ. 2012. Analysis of Histone Modifications from Tryptic Peptides of Deuteroacetylated Isoforms. *Int J Mass Spectrom* **312**: 5–16.
158. Hertz-Fowler C, Figueiredo LM, Quail MA, Becker M, Jackson A, Bason N, Brooks K, Churcher C, Fahkro S, Goodhead I, Heath P, Kartvelishvili M, Mungall K, Harris D, Hauser H, Sanders M, Saunders D, Seeger K, Sharp S, Taylor JE, Walker D, White B, Young R, Cross GA, Rudenko G, Barry JD, Louis EJ, Berriman M. 2008. Telomeric expression sites are highly conserved in Trypanosoma brucei. *PLoS One* **3**: e3527.
159. Hertz-Fowler C, Renauld H, Berriman M. 2007. The Genome of *Trypanosoma brucei.* In *Trypanosomes: After The Genome*. (eds. Barry D, R McCulloch, J Mottram, A Acosta-Serrano), pp. 5–48. Horizon Bioscience,
160. Hirumi H, Hirumi K. 1994. Axenic culture of African trypanosome bloodstream forms. *Parasitol Today* **10**: 80–84.
161. Hoeijmakers WA, Salcedo-Amaya AM, Smits AH, Françoijs KJ, Treeck M, Gilberger TW, Stunnenberg HG, Bártfai R. 2013. H2A.Z/H2B.Z double-variant nucleosomes

inhabit the AT-rich promoter regions of the Plasmodium falciparum genome. *Mol Microbiol* **87**: 1061–1073.
162. Hong J, Feng H, Wang F, Ranjan A, Chen J, Jiang J, Ghirlando R, Xiao TS, Wu C, Bai Y. 2014. The catalytic subunit of the SWR1 remodeler is a histone chaperone for the H2A.Z-H2B dimer. *Mol Cell* **53**: 498–505.
163. Horn D. 2014. Antigenic variation in African trypanosomes. *Mol Biochem Parasitol* **195**: 123–129.
164. Hota SK, Bhardwaj SK, Deindl S, Lin YC, Zhuang X, Bartholomew B. 2013. Nucleosome mobilization by ISW2 requires the concerted action of the ATPase and SLIDE domains. *Nat Struct Mol Biol* **20**: 222–229.
165. Howe KJ, Kane CM, Ares M. 2003. Perturbation of transcription elongation influences the fidelity of internal exon inclusion in Saccharomyces cerevisiae. *RNA* **9**: 993–1006.
166. Hsin JP, Manley JL. 2012. The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev* **26**: 2119–2137.
167. Huang H, Sabari BR, Garcia BA, Allis CD, Zhao Y. 2014. SnapShot: histone modifications. *Cell* **159**: 458–458.e1.
168. Hughes K, Wand M, Foulston L, Young R, Harley K, Terry S, Ersfeld K, Rudenko G. 2007. A novel ISWI is involved in VSG expression site downregulation in African trypanosomes. *EMBO J* **26**: 2400–2410.
169. Iouzalen N, Moreau J, Méchali M. 1996. H2A.ZI, a new variant histone expressed during Xenopus early development exhibits several distinct features from the core histone H2A. *Nucleic Acids Res* **24**: 3947–3952.
170. Ishibashi T, Dryhurst D, Rose KL, Shabanowitz J, Hunt DF, Ausió J. 2009. Acetylation of vertebrate H2A.Z and its effect on the structure of the nucleosome. *Biochemistry* **48**: 5007–5017.
171. Ivens AC, Peacock CS, Worthey EA, Murphy L, Aggarwal G, Berriman M, Sisk E, Rajandream MA, Adlem E, Aert R, Anupama A, Apostolou Z, Attipoe P, Bason N, Bauser C, Beck A, Beverley SM, Bianchettin G, Borzym K, Bothe G, Bruschi CV, Collins M, Cadag E, Ciarloni L, Clayton C, Coulson RM, Cronin A, Cruz AK, Davies RM, De Gaudenzi J, Dobson DE, Duesterhoeft A, Fazelina G, Fosker N, Frasch AC, Fraser A, Fuchs M, Gabel C, Goble A, Goffeau A, Harris D, Hertz-Fowler C, Hilbert H, Horn D, Huang Y, Klages S, Knights A, Kube M, Larke N, Litvin L, Lord A, Louie T, Marra M, Masuy D, Matthews K, Michaeli S, Mottram JC, Muller-Auer S, Munden H, Nelson S, Norbertczak H, Oliver K, O'neil S, Pentony M, Pohl TM, Price C, Purnelle B, Quail MA, Rabbinowitsch E, Reinhardt R, Rieger M, Rinta J, Robben J, Robertson L, Ruiz JC, Rutter S, Saunders D, Schafer M, Schein J, Schwartz DC, Seeger K, Seyler A, Sharp S, Shin H, Sivam D, Squares R, Squares S, Tosato V, Vogt C, Volckaert G, Wambutt R, Warren T, Wedler H, Woodward J, Zhou S, Zimmermann W, Smith DF, Blackwell JM, Stuart KD, Barrell B, Myler PJ. 2005. The genome of the kinetoplastid parasite, Leishmania major. *Science* **309**: 436–442.
172. Jackson JD, Gorovsky MA. 2000. Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. *Nucleic Acids Res* **28**: 3811–3816.
173. Jacobson RH, Ladurner AG, King DS, Tjian R. 2000. Structure and function of a human TAFII250 double bromodomain module. *Science* **288**: 1422–1425.
174. Jamonneau V, Ilboudo H, Kaboré J, Kaba D, Koffi M, Solano P, Garcia A, Courtin D, Laveissière C, Lingue K, Büscher P, Bucheton B. 2012. Untreated human infections by Trypanosoma brucei gambiense are not 100% fatal. *PLoS Negl Trop Dis* **6**: e1691.
175. Janzen CJ, Fernandez JP, Deng H, Diaz R, Hake SB, Cross GA. 2006a. Unusual histone modifications in Trypanosoma brucei. *FEBS Lett* **580**: 2306–2310.
176. Janzen CJ, Hake SB, Lowell JE, Cross GA. 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.
177. Jeanmougin F, Wurtz JM, Le Douarin B, Chambon P, Losson R. 1997. The bromodomain revisited. *Trends Biochem Sci* **22**: 151–153.
178. Jeronimo C, Watanabe S, Kaplan CD, Peterson CL, Robert F. 2015. The Histone Chaperones FACT and Spt6 Restrict H2A.Z from Intragenic Locations. *Mol Cell* **58**: 1113–1123.
179. Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. 2009. H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nat Genet* **41**: 941–945.
180. Johnston H, Kneer J, Chackalaparampil I, Yaciuk P, Chrivia J. 1999. Identification of a novel SNF2/SWI2 protein family member, SRCAP, which interacts with CREB-binding protein. *J Biol Chem* **274**: 16370–16376.
181. Jonkers I, Lis JT. 2015. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**: 167–177.
182. Kamakaka RT, Biggins S. 2005. Histone variants: deviants. *Genes Dev* **19**: 295–310.
183. Kawahara T, Siegel TN, Ingram AK, Alsford S, Cross GA, Horn D. 2008. Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. *Mol Microbiol* **69**: 1054–1068.
184. Keogh MC, Mennella TA, Sawa C, Berthelet S, Krogan NJ, Wolek A, Podolny V, Carpenter LR, Greenblatt JF, Baetz K, Buratowski S. 2006. The Saccharomyces cerevisiae histone H2A variant Htz1 is acetylated by NuA4. *Genes Dev* **20**: 660–665.
185. Khare SP, Habib F, Sharma R, Gadewal N, Gupta S, Galande S. 2012. HIstome--a relational knowledgebase of human histone proteins and histone modifying enzymes. *Nucleic Acids Res* **40**: D337–42.
186. Kim HS, Li Z, Boothroyd C, Cross GA. 2013. Strategies to construct null and conditional null Trypanosoma brucei mutants using Cre-recombinase and loxP. *Mol Biochem Parasitol* **191**: 16–19.
187. Kim HS, Vanoosthuyse V, Fillingham J, Roguev A, Watt S, Kislinger T, Treyer A, Carpenter LR, Bennett CS, Emili A, Greenblatt JF, Hardwick KG, Krogan NJ, Bähler J, Keogh MC. 2009. An acetylated form of histone H2A.Z regulates chromosome architecture in Schizosaccharomyces pombe. *Nat Struct Mol Biol* **16**: 1286–1293.
188. Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B. 2005. A high-resolution map of active promoters in the human genome. *Nature* **436**: 876–880.
189. Kimura A, Horikoshi M. 1998. Tip60 acetylates six lysines of a specific class in core histones in vitro. *Genes Cells* **3**: 789–800.
190. Kireeva ML, Walter W, Tchernajenko V, Bondarenko V, Kashlev M, Studitsky VM. 2002. Nucleosome

remodeling induced by RNA polymerase II: loss of the H2A/H2B dimer during transcription. *Mol Cell* **9**: 541–552.
191. Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, Link AJ, Madhani HD, Rine J. 2004. A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. *PLoS Biol* **2**: E131.
192. Kolev NG, Franklin JB, Carmi S, Shi H, Michaeli S, Tschudi C. 2010. The transcriptome of the human pathogen Trypanosoma brucei at single-nucleotide resolution. *PLoS Pathog* **6**: e1001090.
193. Kornberg RD. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science* **184**: 868–871.
194. Kouzarides T. 2007. Chromatin modifications and their function. *Cell* **128**: 693–705.
195. Kraus AJ, Brink BG, Siegel TN. 2019. Efficient and specific oligo-based depletion of rRNA. *Sci Rep* **9**: 12281.
196. Kraus AJ, Vanselow JT, Lamer S, Brink BG, Schlosser A, Siegel TN. 2020. Distinct roles for H4 and H2A.Z acetylation in RNA transcription in African trypanosomes. *Nat Commun* **11**: 1498.
197. Kristiansson E, Thorsen M, Tamás MJ, Nerman O. 2009. Evolutionary forces act on promoter length: identification of enriched cis-regulatory elements. *Mol Biol Evol* **26**: 1299–1307.
198. Krogan NJ, Baetz K, Keogh MC, Datta N, Sawa C, Kwok TC, Thompson NJ, Davey MG, Pootoolal J, Hughes TR, Emili A, Buratowski S, Hieter P, Greenblatt JF. 2004. Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. *Proc Natl Acad Sci U S A* **101**: 13513–13518.
199. Krogan NJ, Keogh MC, Datta N, Sawa C, Ryan OW, Ding H, Haw RA, Pootoolal J, Tong A, Canadien V, Richards DP, Wu X, Emili A, Hughes TR, Buratowski S, Greenblatt JF. 2003a. A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell* **12**: 1565–1576.
200. Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S, Greenblatt J. 2003b. Methylation of histone H3 by Set2 in Saccharomyces cerevisiae is linked to transcriptional elongation by RNA polymerase II. *Mol Cell Biol* **23**: 4207–4218.
201. Kusch T, Florens L, Macdonald WH, Swanson SK, Glaser RL, Yates JR, Abmayr SM, Washburn MP, Workman JL. 2004. Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science* **306**: 2084–2087.
202. Ladurner AG, Inouye C, Jain R, Tjian R. 2003. Bromodomains mediate an acetyl-histone encoded antisilencing function at heterochromatin boundaries. *Mol Cell* **11**: 365–376.
203. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**: 357–359.
204. Lashgari A, Millau JF, Jacques PÉ, Gaudreau L. 2017. Global inhibition of transcription causes an increase in histone H2A.Z incorporation within gene bodies. *Nucleic Acids Res* **45**: 12715–12722.
205. Lawrence M, Huber W, Pagès H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. 2013. Software for computing and annotating genomic ranges. *PLoS Comput Biol* **9**: e1003118.
206. Leach TJ, Mazzeo M, Chotkowski HL, Madigan JP, Wotring MG, Glaser RL. 2000. Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster. *J Biol Chem* **275**: 23267–23272.
207. Lecordier L, Devaux S, Uzureau P, Dierick JF, Walgraffe D, Poelvoorde P, Pays E, Vanhamme L. 2007. Characterization of a TFIIH homologue from Trypanosoma brucei. *Mol Microbiol* **64**: 1164–1181.
208. Lee JH, Jung HS, Günzl A. 2009. Transcriptionally active TFIIH of the early-diverged eukaryote Trypanosoma brucei harbors two novel core subunits but not a cyclin-activating kinase complex. *Nucleic Acids Res* **37**: 3811–3820.
209. Lee JH, Nguyen TN, Schimanski B, Günzl A. 2007. Spliced leader RNA gene transcription in Trypanosoma brucei requires transcription factor TFIIH. *Eukaryot Cell* **6**: 641–649.
210. Lee JR, Lee J, Kim SK, Kim KP, Park HS, Yeo WS. 2008. Mass spectrometry signal amplification method for attomolar detection of antigens using small-molecule-tagged gold microparticles. *Angew Chem Int Ed Engl* **47**: 9518–9521.
211. Leschziner AE, Lemon B, Tjian R, Nogales E. 2005. Structural studies of the human PBAF chromatin-remodeling complex. *Structure* **13**: 267–275.
212. Li B, Carey M, Workman JL. 2007. The role of chromatin during transcription. *Cell* **128**: 707–719.
213. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**: 589–595.
214. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 GPDPS. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
215. Li Z, Gadue P, Chen K, Jiao Y, Tuteja G, Schug J, Li W, Kaestner KH. 2012. Foxa2 and H2A.Z mediate nucleosome depletion during embryonic stem cell differentiation. *Cell* **151**: 1608–1616.
216. Liang XH, Haritan A, Uliel S, Michaeli S. 2003. trans and cis splicing in trypanosomatids: mechanism, factors, and regulation. *Eukaryot Cell* **2**: 830–840.
217. Lipmann F. 1954. Development of the acetylation problem, a personal account. *Science* **120**: 855–865.
218. Lipmann F, Kaplan NO. 1946. Report on a coenzyme for acetylation. *Fed Proc* **5**: 145.
219. Lipmann F, Kaplan NO. 1947. Coenzyme for acetylation, a pantothenic acid derivative. *J Biol Chem* **167**: 869.
220. Liu X, Li B, GorovskyMA. 1996. Essential and nonessential histone H2A variants in Tetrahymena thermophila. *Mol Cell Biol* **16**: 4305–4311.
221. Long M, Sun X, Shi W, Yanru A, Leung STC, Ding D, Cheema MS, MacPherson N, Nelson CJ, Ausio J, Yan Y, Ishibashi T. 2019. A novel histone H4 variant H4G regulates rDNA transcription in breast cancer. *Nucleic Acids Res* **47**: 8399–8409.
222. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**: 550.
223. Lowell JE, Cross GA. 2004. A variant histone H3 is enriched at telomeres in Trypanosoma brucei. *J Cell Sci* **117**: 5937–5947.
224. Lowell JE, Kaiser F, Janzen CJ, Cross GA. 2005. Histone H2AZ dimerizes with a novel variant H2B and is enriched at repetitive DNA in Trypanosoma brucei. *J Cell Sci* **118**: 5721–5730.

225. Luger K, Dechassa ML, Tremethick DJ. 2012. New insights into nucleosome and chromatin structure: an ordered state or a disordered affair. *Nat Rev Mol Cell Biol* **13**: 436–447.
226. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**: 251–260.
227. Luger K, Richmond TJ. 1998. DNA binding within the nucleosome core. *Curr Opin Struct Biol* **8**: 33–40.
228. Luk E, Ranjan A, Fitzgerald PC, Mizuguchi G, Huang Y, Wei D, Wu C. 2010. Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. *Cell* **143**: 725–736.
229. Lundby A, Lage K, Weinert BT, Bekker-Jensen DB, Secher A, Skovgaard T, Kelstrup CD, Dmytriyev A, Choudhary C, Lundby C, Olsen JV. 2012. Proteomic analysis of lysine acetylation sites in rat tissues reveals organ specificity and subcellular patterns. *Cell Rep* **2**: 419–431.
230. MacGregor P, Szöőr B, Savill NJ, Matthews KR. 2012. Trypanosomal immune evasion, chronicity and transmission: an elegant balancing act. *Nat Rev Microbiol* **10**: 431–438.
231. Malagon F, Tong AH, Shafer BK, Strathern JN. 2004. Genetic interactions of DST1 in Saccharomyces cerevisiae suggest a role of TFIIS in the initiation-elongation transition. *Genetics* **166**: 1215–1227.
232. Malik HS, Henikoff S. 2003. Phylogenomics of the nucleosome. *Nat Struct Biol* **10**: 882–891.
233. Malik I, Qiu C, Snavely T, Kaplan CD. 2017. Wide-ranging and unexpected consequences of altered Pol II catalytic activity in vivo. *Nucleic Acids Res* **45**: 4431–4451.
234. Malvy D, Chappuis F. 2011. Sleeping sickness. *Clin Microbiol Infect* **17**: 986–995.
235. Mandava V, Fernandez JP, Deng H, Janzen CJ, Hake SB, Cross GA. 2007. Histone modifications in Trypanosoma brucei. *Mol Biochem Parasitol* **156**: 41–50.
236. Marañón C, Thomas MC, Puerta C, Alonso C, López MCL. 2000. The stability and maturation of the H2A histone mRNAs from Trypanosoma cruzi are implicated in their post-transcriptional regulation. *Biochim Biophys Acta* **1490**: 1–10.
237. Marques M, Laflamme L, Gervais AL, Gaudreau L. 2010. Reconciling the positive and negative roles of histone H2A.Z in gene transcription. *Epigenetics* **5**: 267–272.
238. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal; Vol 17, No 1: Next Generation Sequencing Data Analysis*
239. Martínez-Calvillo S, Nguyen D, Stuart K, Myler PJ. 2004. Transcription initiation and termination on Leishmania major chromosome 3. *Eukaryot Cell* **3**: 506–517.
240. Martinez-Calvillo S, Vizuet-de-Rueda JC, Florencio-Martinez LE, Manning-Cela RG, Figueroa-Angulo EE. 2010. Gene expression in trypanosomatid parasites. *J Biomed Biotechnol* **2010**: 525241.
241. Marzluff WF, Duronio RJ. 2002. Histone mRNA expression: multiple levels of cell cycle regulation and important developmental consequences. *Curr Opin Cell Biol* **14**: 692–699.
242. Matthews KR, Ellis JR, Paterou A. 2004. Molecular regulation of the life cycle of African trypanosomes. *Trends Parasitol* **20**: 40–47.
243. Mavrich TN, Jiang C, Ioshikhes IP, Li X, Venters BJ, Zanton SJ, Tomsho LP, Qi J, Glaser RL, Schuster SC, Gilmour DS, Albert I, Pugh BF. 2008. Nucleosome organization in the Drosophila genome. *Nature* **453**: 358–362.
244. McAndrew M, Graham S, Hartmann C, Clayton C. 1998. Testing promoter activity in the trypanosome genome: isolation of a metacyclic-type VSG promoter, and unexpected insights into RNA polymerase II transcription. *Exp Parasitol* **90**: 65–76.
245. McGhee JD, Wood WI, Dolan M, Engel JD, Felsenfeld G. 1981. A 200 base pair region at the 5' end of the chicken adult beta-globin gene is accessible to nuclease digestion. *Cell* **27**: 45–55.
246. Meneghini MD, Wu M, Madhani HD. 2003. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. *Cell* **112**: 725–736.
247. Millar CB, Xu F, Zhang K, Grunstein M. 2006. Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes Dev* **20**: 711–722.
248. Mito Y, Henikoff JG, Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nat Genet* **37**: 1090–1097.
249. Mitra D, Parnell EJ, Landon JW, Yu Y, Stillman DJ. 2006. SWI/SNF binding to the HO promoter requires histone acetylation and stimulates TATA-binding protein recruitment. *Mol Cell Biol* **26**: 4095–4110.
250. Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* **303**: 343–348.
251. Morrison AJ, Shen X. 2009. Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat Rev Mol Cell Biol* **10**: 373–384.
252. Müller LSM, Cosentino RO, Förstner KU, Guizetti J, Wedel C, Kaplan N, Janzen CJ, Arampatzi P, Vogel J, Steinbiss S, Otto TD, Saliba AE, Sebra RP, Siegel TN. 2018. Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature*
253. Myler PJ, Audleman L, deVos T, Hixson G, Kiser P, Lemley C, Magness C, Rickel E, Sisk E, Sunkin S, Swartzell S, Westlake T, Bastien P, Fu G, Ivens A, Stuart K. 1999. Leishmania major Friedlin chromosome 1 has an unusual distribution of protein-coding genes. *Proc Natl Acad Sci U S A* **96**: 2902–2906.
254. Narlikar GJ, Sundaramoorthy R, Owen-Hughes T. 2013. Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell* **154**: 490–503.
255. Navarro M, Gull K. 2001. A pol I transcriptional body associated with VSG mono-allelic expression in Trypanosoma brucei. *Nature* **414**: 759–763.
256. Neuwirth E. 2014. R Package 'Colour Brewer'.
257. Neves LT, Douglass S, Spreafico R, Venkataramanan S, Kress TL, Johnson TL. 2017. The histone variant H2A.Z promotes efficient cotranscriptional splicing in S. cerevisiae. *Genes Dev* **31**: 702–717.
258. Ng HH, Robert F, Young RA, Struhl K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell* **11**: 709–719.
259. Nguyen AT, Zhang Y. 2011. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev* **25**: 1345–1358.
260. Nilsson D, Gunasekera K, Mani J, Osteras M, Farinelli L, Baerlocher L, Roditi I, Ochsenreiter T. 2010. Spliced

leader trapping reveals widespread alternative splicing patterns in the highly dynamic transcriptome of Trypanosoma brucei. *PLoS Pathog* **6**: e1001037.
261. Nissen KE, Homer CM, Ryan CJ, Shales M, Krogan NJ, Patrick KL, Guthrie C. 2017. The histone variant H2A.Z promotes splicing of weak introns. *Genes Dev* **31**: 688–701.
262. Nogales E, Louder RK, He Y. 2017. Structural Insights into the Eukaryotic Transcription Initiation Machinery. *Annu Rev Biophys* **46**: 59–83.
263. Ogbadoyi E, Ersfeld K, Robinson D, Sherwin T, Gull K. 2000. Architecture of the Trypanosoma brucei nucleus during interphase and mitosis. *Chromosoma* **108**: 501–513.
264. Olins AL, Olins DE. 1974. Spheroid chromatin units (v bodies). *Science* **183**: 330–332.
265. Ornaghi P, Ballario P, Lena AM, González A, Filetici P. 1999. The bromodomain of Gcn5p interacts in vitro with specific residues in the N terminus of histone H4. *J Mol Biol* **287**: 1–7.
266. Orphanides G, Reinberg D. 2000. RNA polymerase II elongation through chromatin. *Nature* **407**: 471–475.
267. Owen DJ, Ornaghi P, Yang JC, Lowe N, Evans PR, Ballario P, Neuhaus D, Filetici P, Travers AA. 2000. The structural basis for the recognition of acetylated histone H4 by the bromodomain of histone acetyltransferase gcn5p. *EMBO J* **19**: 6141–6149.
268. Palenchar JB, Liu W, Palenchar PM, Bellofatto V. 2006. A divergent transcription factor TFIIB in trypanosomes is required for RNA polymerase II-dependent spliced leader RNA transcription and cell viability. *Eukaryot Cell* **5**: 293–300.
269. Papamichos-Chronakis M, Watanabe S, Rando OJ, Peterson CL. 2011a. Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* **144**: 200–213.
270. Papamichos-Chronakis M, Watanabe S, Rando OJ, Peterson CL. 2011b. Global regulation of H2A.Z localization by the INO80 chromatin-remodeling enzyme is essential for genome integrity. *Cell* **144**: 200–213.
271. Parnell TJ, Huff JT, Cairns BR. 2008. RSC regulates nucleosome positioning at Pol II genes and density at Pol III genes. *EMBO J* **27**: 100–110.
272. Parsons M, Nelson RG, Watkins KP, Agabian N. 1984. Trypanosome mRNAs share a common 5' spliced leader sequence. *Cell* **38**: 309–316.
273. Pehrson JR, Fried VA. 1992. MacroH2A, a core histone containing a large nonhistone region. *Science* **257**: 1398–1400.
274. Phillips DM. 1963. The presence of acetyl groups of histones. *Biochem J* **87**: 258–263.
275. Picchi GF, Zulkievicz V, Krieger MA, Zanchin NT, Goldenberg S, de Godoy LM. 2017. Post-translational Modifications of Trypanosoma cruzi Canonical and Variant Histones. *J Proteome Res* **16**: 1167–1179.
276. Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA. 2005. Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* **122**: 517–527.
277. Pollard KJ, Peterson CL. 1998. Chromatin remodeling: a marriage between two families. *Bioessays* **20**: 771–780.
278. Pray-Grant MG, Daniel JA, Schieltz D, Yates JR, Grant PA. 2005. Chd1 chromodomain links histone H3 methylation with SAGA- and SLIK-dependent acetylation. *Nature* **433**: 434–438.
279. Team RC. A language and environment for statistical computing. *R Foundation for Statistical Computing*
280. Raisner RM, Hartley PD, Meneghini MD, Bao MZ, Liu CL, Schreiber SL, Rando OJ, Madhani HD. 2005. Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell* **123**: 233–248.
281. Rangasamy D, Greaves I, Tremethick DJ. 2004. RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. *Nat Struct Mol Biol* **11**: 650–655.
282. Ranjan A, Mizuguchi G, FitzGerald PC, Wei D, Wang F, Huang Y, Luk E, Woodcock CL, Wu C. 2013. Nucleosome-free region dominates histone acetylation in targeting SWR1 to promoters for H2A.Z replacement. *Cell* **154**: 1232–1245.
283. Ranjan A, Wang F, Mizuguchi G, Wei D, Huang Y, Wu C. 2015. H2A histone-fold and DNA elements in nucleosome activate SWR1-mediated H2A.Z replacement in budding yeast. *Elife* **4**: e06845.
284. Recinos RF, Kirchhoff LV, Donelson JE. 2001. Cell cycle expression of histone genes in Trypanosoma cruzi. *Mol Biochem Parasitol* **113**: 215–222.
285. Reinke H, Hörz W. 2003. Histones are first hyperacetylated and then lose contact with the activated PHO5 promoter. *Mol Cell* **11**: 1599–1607.
286. Ren Q, Gorovsky MA. 2001. Histone H2A.Z acetylation modulates an essential charge patch. *Mol Cell* **7**: 1329–1335.
287. Reynolds D, Hofmeister BT, Cliffe L, Alabady M, Siegel TN, Schmitz RJ, Sabatini R. 2016. Histone H3 Variant Regulates RNA Polymerase II Transcription Termination and Dual Strand Transcription of siRNA Loci in Trypanosoma brucei. *PLoS Genet* **12**: e1005758.
288. Rice GA, Kane CM, Chamberlin MJ. 1991. Footprinting analysis of mammalian RNA polymerase II along its transcript: an alternative view of transcription elongation. *Proc Natl Acad Sci U S A* **88**: 4245–4249.
289. Rico E, Jeacock L, Kovářová J, Horn D. 2018. Inducible high-efficiency CRISPR-Cas9-targeted gene editing and precision base editing in African trypanosomes. *Sci Rep* **8**: 7960.
290. Ridgway P, Brown KD, Rangasamy D, Svensson U, Tremethick DJ. 2004. Unique residues on the H2A.Z containing nucleosome surface are important for Xenopus laevis development. *J Biol Chem* **279**: 43815–43820.
291. Robert F, Pokholok DK, Hannett NM, Rinaldi NJ, Chandy M, Rolfe A, Workman JL, Gifford DK, Young RA. 2004. Global position and recruitment of HATs and HDACs in the yeast genome. *Mol Cell* **16**: 199–209.
292. Roberts TG, Dungan JM, Watkins KP, Agabian N. 1996. The SLA RNA gene of Trypanosoma brucei is organized in a tandem array which encodes several small RNAs. *Mol Biochem Parasitol* **83**: 163–174.
293. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nat Biotechnol* **29**: 24–26.
294. Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem* **273**: 5858–5868.
295. Rojas F, Silvester E, Young J, Milne R, Tettey M, Houston DR, Walkinshaw MD, Pérez-Pi I, Auer M, Denton H, Smith TK, Thompson J, Matthews KR. 2019.

Oligopeptide Signaling through TbGPR89 Drives Trypanosome Quorum Sensing. *Cell* **176**: 306–317.e16.
296. Rosonina E, Kaneko S, Manley JL. 2006. Terminating the transcript: breaking up is hard to do. *Genes Dev* **20**: 1050–1056.
297. Ross R, Thomson D. 1910. A case of sleeping sickness studied by precise enumerative methods: Regular periodical increase of the parasites disclosed. *Proc. R. Soc. Lond. B.* **82**: 411–415.
298. Roth SY, Denu JM, Allis CD. 2001. Histone acetyltransferases. *Annu Rev Biochem* **70**: 81–120.
299. Rothbart SB, Krajewski K, Strahl BD, Fuchs SM. 2012. Peptide microarrays to interrogate the "histone code". *Methods Enzymol* **512**: 107–135.
300. Ruan JP, Arhin GK, Ullu E, Tschudi C. 2004. Functional characterization of a Trypanosoma brucei TATA-binding protein-related factor points to a universal regulator of transcription in trypanosomes. *Mol Cell Biol* **24**: 9610–9618.
301. Rudnizky S, Bavly A, Malik O, Pnueli L, Melamed P, Kaplan A. 2016. H2A.Z controls the stability and mobility of nucleosomes to regulate expression of the LH genes. *Nat Commun* **7**: 12958.
302. Ruhl DD, Jin J, Cai Y, Swanson S, Florens L, Washburn MP, Conaway RC, Conaway JW, Chrivia JC. 2006. Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. *Biochemistry* **45**: 5671–5677.
303. Sabo PJ, Kuehn MS, Thurman R, Johnson BE, Johnson EM, Cao H, Yu M, Rosenzweig E, Goldy J, Haydock A, Weaver M, Shafer A, Lee K, Neri F, Humbert R, Singer MA, Richmond TA, Dorschner MO, McArthur M, Hawrylycz M, Green RD, Navas PA, Noble WS, Stamatoyannopoulos JA. 2006. Genome-scale mapping of DNase I sensitivity in vivo using tiling DNA microarrays. *Nat Methods* **3**: 511–518.
304. Saha S. 2020. Histone Modifications and Other Facets of Epigenetic Regulation in Trypanosomatids: Leaving Their Mark. *mBio* **11**:
305. Sainsbury S, Bernecky C, Cramer P. 2015. Structural basis of transcription initiation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**: 129–143.
306. Santisteban MS, Hang M, Smith MM. 2011. Histone variant H2A.Z and RNA polymerase II transcription elongation. *Mol Cell Biol* **31**: 1848–1860.
307. Santisteban MS, Kalashnikova T, Smith MM. 2000. Histone H2A.Z regulats transcription and is partially redundant with nucleosome remodeling complexes. *Cell* **103**: 411–422.
308. Sarcinella E, Zuzarte PC, Lau PN, Draker R, Cheung P. 2007. Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol Cell Biol* **27**: 6457–6468.
309. Scahill MD, Pastar I, Cross GA. 2008. CRE recombinase-based positive-negative selection systems for genetic manipulation in Trypanosoma brucei. *Mol Biochem Parasitol* **157**: 73–82.
310. Schimanski B, Brandenburg J, Nguyen TN, Caimano MJ, Günzl A. 2006. A TFIIB-like protein is indispensable for spliced leader RNA gene transcription in Trypanosoma brucei. *Nucleic Acids Res* **34**: 1676–1684.
311. Schimanski B, Laufer G, Gontcharova L, Günzl A. 2004. The Trypanosoma brucei spliced leader RNA and rRNA gene promoters have interchangeable TbSNAP50-binding elements. *Nucleic Acids Res* **32**: 700–709.
312. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **9**: 671–675.
313. Schulz D, Mugnier MR, Paulsen EM, Kim HS, Chung CW, Tough DF, Rioja I, Prinjha RK, Papavasiliou FN, Debler EW. 2015. Bromodomain Proteins Contribute to Maintenance of Bloodstream Form Stage Identity in the African Trypanosome. *PLoS Biol* **13**: e1002316.
314. Schuster S, Subota I, Lisack J, Zimmermann H, Reuter C, Morriswood B, Engstler M. 2019. A modification to the life cycle of the parasite *Trypanosoma brucei*. *bioRxiv* 717975.
315. Sealy L, Chalkley R. 1978. DNA associated with hyperacetylated histone is preferentially digested by DNase I. *Nucleic Acids Res* **5**: 1863–1876.
316. Shahbazian MD, Grunstein M. 2007. Functions of site-specific histone acetylation and deacetylation. *Annu Rev Biochem* **76**: 75–100.
317. Shechter D, Dormann HL, Allis CD, Hake SB. 2007. Extraction, purification and analysis of histones. *Nat Protoc* **2**: 1445–1457.
318. Shen X, Ranallo R, Choi E, Wu C. 2003. Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Mol Cell* **12**: 147–155.
319. Shi B, Guo X, Wu T, Sheng S, Wang J, Skogerbø G, Zhu X, Chen R. 2009. Genome-scale identification of Caenorhabditis elegans regulatory elements by tiling-array mapping of DNase I hypersensitive sites. *BMC Genomics* **10**: 92.
320. Shi H, Ullu E, Tschudi C. 2004. Function of the Trypanosome Argonaute 1 protein in RNA interference requires the N-terminal RGG domain and arginine 735 in the Piwi domain. *J Biol Chem* **279**: 49889–49893.
321. Shia WJ, Li B, Workman JL. 2006. SAS-mediated acetylation of histone H4 Lys 16 is required for H2A.Z incorporation at subtelomeric regions in Saccharomyces cerevisiae. *Genes Dev* **20**: 2507–2512.
322. Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL. 2006. Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**: 844–847.
323. Sidoli S, Lopes M, Lund PJ, Goldman N, Fasolino M, Coradin M, Kulej K, Bhanu NV, Vahedi G, Garcia BA. 2019. A mass spectrometry-based assay using metabolic labeling to rapidly monitor chromatin accessibility of modified histone proteins. *Sci Rep* **9**: 13613.
324. Siegel TN, Gunasekera K, Cross GA, Ochsenreiter T. 2011. Gene expression in Trypanosoma brucei: lessons from high-throughput RNA sequencing. *Trends Parasitol* **27**: 434–441.
326. Siegel TN, Hekstra DR, Kemp LE, Figueiredo LM, Lowell JE, Fenyo D, Wang X, Dewell S, Cross GAM. 2009b. Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. *Genes & Development* **23**: 1063–1076.
327. Siegel TN, Kawahara T, Degrasse JA, Janzen CJ, Horn D, Cross GA. 2008. Acetylation of histone H4K4 is cell cycle regulated and mediated by HAT3 in Trypanosoma brucei. *Mol Microbiol* **67**: 762–771.
328. Siegel TN, Tan KS, Cross GA. 2005. Systematic study of sequence motifs for RNA trans splicing in Trypanosoma brucei. *Mol Cell Biol* **25**: 9586–9594.
329. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation

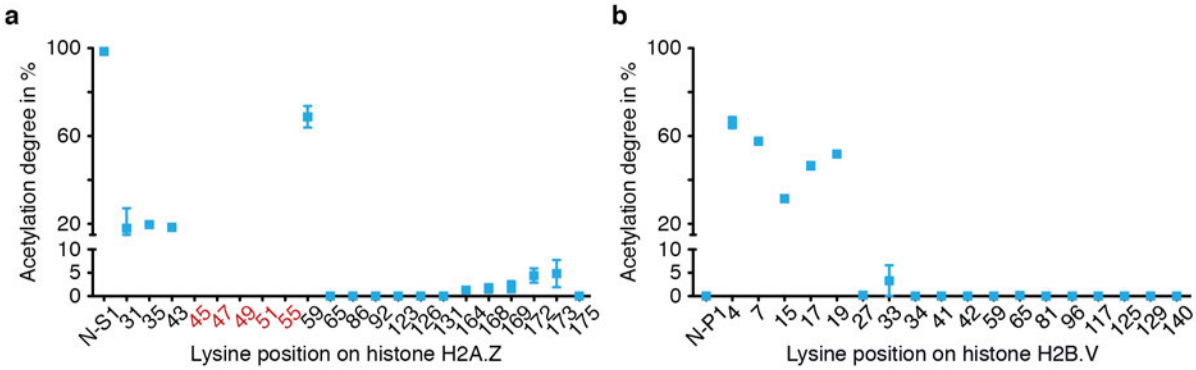
of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**: 539.
330. Siino JS, Yau PM, Imai BS, Gatewood JM, Bradbury EM. 2003. Effect of DNA length and H4 acetylation on the thermal stability of reconstituted nucleosome particles. *Biochem Biophys Res Commun* **302**: 885–891.
331. Smale ST, Kadonaga JT. 2003. The RNA polymerase II core promoter. *Annu Rev Biochem* **72**: 449–479.
332. Smith CL, Horowitz-Scherer R, Flanagan JF, Woodcock CL, Peterson CL. 2003a. Structural analysis of the yeast SWI/SNF chromatin remodeling complex. *Nat Struct Biol* **10**: 141–145.
333. Smith CM. 2005. Quantification of acetylation at proximal lysine residues using isotopic labeling and tandem mass spectrometry. *Methods* **36**: 395–403.
334. Smith CM, Gafken PR, Zhang Z, Gottschling DE, Smith JB, Smith DL. 2003b. Mass spectrometric quantification of acetylation at specific lysines within the amino-terminal tail of histone H4. *Anal Biochem* **316**: 23–33.
335. Smith JL, Levin JR, Ingles CJ, Agabian N. 1989. In trypanosomes the homolog of the largest subunit of RNA polymerase II is encoded by two genes and has a highly unusual C-terminal domain structure. *Cell* **56**: 815–827.
336. Soto M, Quijada L, Alonso C, Requena JM. 2000. Histone synthesis in Leishmania infantum is tightly linked to DNA replication by a translational control. *Biochem J* **346 Pt 1**: 99–105.
337. Srivastava A, Badjatia N, Lee JH, Hao B, Günzl A. 2018. An RNA polymerase II-associated TFIIF-like complex is indispensable for SL RNA gene transcription in Trypanosoma brucei. *Nucleic Acids Res.* **46**: 1695–1709.
338. Stanne TM, Kushwaha M, Wand M, Taylor JE, Rudenko G. 2011. TbISWI regulates multiple polymerase I (Pol I)-transcribed loci and is present at Pol II transcription boundaries in Trypanosoma brucei. *Eukaryot Cell* **10**: 964–976.
339. Stanne TM, Narayanan MS, Ridewood S, Ling A, Witmer K, Kushwaha M, Wiesler S, Wickstead B, Wood J, Rudenko G. 2015. Identification of the ISWI Chromatin Remodeling Complex of the Early Branching Eukaryote Trypanosoma brucei. *J Biol Chem* **290**: 26954–26967.
340. Stargell LA, Bowen J, Dadd CA, Dedon PC, Davis M, Cook RG, Allis CD, Gorovsky MA. 1993. Temporal and spatial association of histone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila. *Genes Dev* **7**: 2641–2651.
341. Steger DJ, Workman JL. 1996. Remodeling chromatin structures for transcription: what happens to the histones. *Bioessays* **18**: 875–884.
342. Stewart FJ, Ottesen EA, DeLong EF. 2010. Development and quantitative analyses of a universal rRNA-subtraction protocol for microbial metatranscriptomics. *ISME J* **4**: 896–907.
343. Stralfors A, Ekwall K. 2011. Heterochromatin and Euchromatin - Organization, Boundaries, and Gene Regulation. In *Reviews in Cell Biology and Molecular Medicine*. (eds. pp. American Cancer Society,
344. Struhl K. 1998. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev* **12**: 599–606.
345. Subramanian V, Fields PA, Boyer LA. 2015. H2A.Z: a molecular rheostat for transcriptional control. *F1000Prime Rep* **7**: 01.
346. Sudarshi D, Lawrence S, Pickrell WO, Eligar V, Walters R, Quaderi S, Walker A, Capewell P, Clucas C, Vincent A, Checchi F, MacLeod A, Brown M. 2014. Human African trypanosomiasis presenting at least 29 years after infection--what can this teach us about the pathogenesis and control of this neglected tropical disease. *PLoS Negl Trop Dis* **8**: e3349.
347. Suganuma T, Workman JL. 2011. Signals and combinatorial functions of histone modifications. *Annu Rev Biochem* **80**: 473–499.
348. Sun ZW, Allis CD. 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**: 104–108.
349. Suto RK, Clarkson MJ, Tremethick DJ, Luger K. 2000. Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat Struct Biol* **7**: 1121–1124.
350. Swaminathan J, Baxter EM, Corces VG. 2005. The role of histone H2Av variant replacement and histone H4 acetylation in the establishment of Drosophila heterochromatin. *Genes Dev* **19**: 65–76.
351. Tachiwana H, Osakabe A, Shiga T, Miya Y, Kimura H, Kagawa W, Kurumizaka H. 2011. Structures of human nucleosomes containing major histone H3 variants. *Acta Crystallogr D Biol Crystallogr* **67**: 578–583.
352. Talbert PB, Henikoff S. 2010. Histone variants--ancient wrap artists of the epigenome. *Nat Rev Mol Cell Biol* **11**: 264–275.
353. Tessarz P, Kouzarides T. 2014. Histone core modifications regulating nucleosome structure and dynamics. *Nat Rev Mol Cell Biol* **15**: 703–708.
354. Thatcher TH, Gorovsky MA. 1994. Phylogenetic analysis of the core histones H2A, H2B, H3, and H4. *Nucleic Acids Res* **22**: 174–179.
355. Tirosh I, Barkai N. 2008. Two strategies for gene regulation by promoter nucleosomes. *Genome Res* **18**: 1084–1091.
356. Torigoe SE, Urwin DL, Ishii H, Smith DE, Kadonaga JT. 2011. Identification of a rapidly formed nonnucleosomal histone-DNA intermediate that is converted into chromatin by ACF. *Mol Cell* **43**: 638–648.
357. Toro GC, Galanti N, Hellman U, Wernstedt C. 1993. Unambiguous identification of histone H1 in Trypanosoma cruzi. *J Cell Biochem* **52**: 431–439.
358. Tramantano M, Sun L, Au C, Labuz D, Liu Z, Chou M, Shen C, Luk E. 2016. Constitutive turnover of histone H2A.Z at yeast promoters requires the preinitiation complex. *Elife* **5**:
359. Tse C, Sera T, Wolffe AP, Hansen JC. 1998. Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. *Mol Cell Biol* **18**: 4629–4638.
360. Udugama M, Sabri A, Bartholomew B. 2011. The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Mol Cell Biol* **31**: 662–673.
361. UniProt C. 2019. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* **47**: D506–D515.
362. Urbaniak MD, Martin DM, Ferguson MA. 2013. Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of Trypanosoma brucei. *J Proteome Res* **12**: 2233–2244.
363. Valdes-Mora F, Song JZ, Statham AL, Strbenac D, Robinson MD, Nair SS, Patterson KI, Tremethick DJ, Stirzaker C, Clark SJ. 2012. Acetylation of H2A.Z is a key epigenetic modification associated with gene deregulation

and epigenetic remodeling in cancer. *Genome Res* **22**: 307–321.
364. van Daal A, Elgin SC. 1992. A histone variant, H2AvD, is essential in Drosophila melanogaster. *Mol Biol Cell* **3**: 593–602.
365. van Rossum G, de Boer J. 1991. Interactively testing remote servers using the Python programming language. *CWI Quarterly* **4**: 283–304.
366. Vanhamme L, Poelvoorde P, Pays A, Tebabi P, Van Xong H, Pays E. 2000. Differential RNA elongation controls the variant surface glycoprotein gene expression sites of Trypanosoma brucei. *Mol Microbiol* **36**: 328–340.
367. Vannini A, Cramer P. 2012. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol Cell* **45**: 439–446.
368. Varga-Weisz PD, Wilm M, Bonte E, Dumas K, Mann M, Becker PB. 1997. Chromatin-remodelling factor CHRAC contains the ATPases ISWI and topoisomerase II. *Nature* **388**: 598–602.
369. Vasquez JJ, Wedel C, Cosentino RO, Siegel TN. 2018. Exploiting CRISPR-Cas9 technology to investigate individual histone modifications. *Nucleic Acids Res* **46**: e106.
370. Velculescu VE, Madden SL, Zhang L, Lash AE, Yu J, Rago C, Lal A, Wang CJ, Beaudry GA, Ciriello KM, Cook BP, Dufault MR, Ferguson AT, Gao Y, He TC, Hermeking H, Hiraldo SK, Hwang PM, Lopez MA, Luderer HF, Mathews B, Petroziello JM, Polyak K, Zawel L, Kinzler KW. 1999. Analysis of human transcriptomes. *Nat Genet* **23**: 387–388.
371. Vellmer T, Hartleb L, Sola AF, Kramer S, Meyer-Natus E, Butter F, Janzen CJ. 2021. A novel SNF2 ATPase complex in *Trypanosoma brucei* with a role in H2A.Z-mediated chromatin remodelling. *bioRxiv* 2021.04.06.438560.
372. Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol*
373. Verdin E, Ott M. 2015. 50 years of protein acetylation: from gene regulation to epigenetics, metabolism and beyond. *Nat Rev Mol Cell Biol* **16**: 258–264.
374. Vickerman K. 1985. Developmental cycles and biology of pathogenic trypanosomes. *Br Med Bull* **41**: 105–114.
375. Vidali G, Boffa LC, Bradbury EM, Allfrey VG. 1978. Butyrate suppression of histone deacetylation leads to accumulation of multiacetylated forms of histones H3 and H4 and increased DNase I sensitivity of the associated DNA sequences. *Proc Natl Acad Sci U S A* **75**: 2239–2243.
376. Wan Y, Saleem RA, Ratushny AV, Roda O, Smith JJ, Lin CH, Chiang JH, Aitchison JD. 2009. Role of the histone variant H2A.Z/Htz1p in TBP recruitment, chromatin dynamics, and regulated expression of oleate-responsive genes. *Mol Cell Biol* **29**: 2346–2358.
377. Warnes GR, Bolker B, Bonebakker L, Gentleman R, Huber W, Liaw A, Lumley T, Maechler M, Magnusson A, Moeller S, Schwartz M, B V. 2016. R package 'gplots' - Various R Programming Tools for Plotting Data.
378. Weake VM, Workman JL. 2008. Histone ubiquitination: triggering gene activity. *Mol Cell* **29**: 653–663.
379. Weber CM, Ramachandran S, Henikoff S. 2014. Nucleosomes Are Context-Specific, H2A.Z-Modulated Barriers to RNA Polymerase. *Mol Cell* **53**: 819–830.
380. Wedel C, Förstner KU, Derr R, Siegel TN. 2017. GT-rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes. *EMBO J*
381. Wedel C, Siegel TN. 2017. Genome-wide analysis of chromatin structures in Trypanosoma brucei using high-resolution MNase-ChIP-seq. *Exp Parasitol* **180**: 2–12.
382. Weintraub H, Groudine M. 1976. Chromosomal subunits in active genes have an altered conformation. *Science* **193**: 848–856.
383. West MH, Bonner WM. 1980. Histone 2A, a heteromorphous family of eight protein species. *Biochemistry* **19**: 3238–3245.
384. Wierer M, Mann M. 2016. Proteomics to study DNA-bound and chromatin-associated gene regulatory complexes. *Hum Mol Genet* **25**: R106–R114.
385. Wippo CJ, Israel L, Watanabe S, Hochheimer A, Peterson CL, Korber P. 2011. The RSC chromatin remodelling enzyme has a unique role in directing the accurate positioning of nucleosomes. *EMBO J* **30**: 1277–1288.
386. Wirtz E, Clayton C. 1995. Inducible gene expression in trypanosomes mediated by a prokaryotic repressor. *Science* **268**: 1179–1183.
387. Workman JL. 2006. Nucleosome displacement in transcription. *Genes Dev* **20**: 2009–2017.
388. Organization WH. 2018. Human African Trypanosomiasis.
389. Woychik RP, Klebig ML, Justice MJ, Magnuson TR, Avner ED, Avrer ED. 1998. Functional genomics in the post-genome era. *Mutat Res* **400**: 3–14.
390. Wright JR, Siegel TN, Cross GA. 2010. Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in Trypanosoma brucei. *Mol Biochem Parasitol* **172**: 141–144.
391. Wu RS, Tsai S, Bonner WM. 1982. Patterns of histone variant synthesis can distinguish G0 from G1 cells. *Cell* **31**: 367–374.
392. Wu WH, Alami S, Luk E, Wu CH, Sen S, Mizuguchi G, Wei D, Wu C. 2005. Swc2 is a widely conserved H2AZ-binding module essential for ATP-dependent histone exchange. *Nat Struct Mol Biol* **12**: 1064–1071.
393. Xu S, Grullon S, Ge K, Peng W. 2014. Spatial clustering for identification of ChIP-enriched regions (SICER) to map regions of histone methylation patterns in embryonic stem cells. *Methods Mol Biol* **1150**: 97–111.
394. Yoo AB, Jette MA, Grondona M. 2003. SLURM: Simple Linux Utility for Resource Management. 44–60.
395. Yoshida M, Horinouchi S, Beppu T. 1995. Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. *Bioessays* **17**: 423–430.
396. Young RA. 1991. RNA polymerase II. *Annu Rev Biochem* **60**: 689–715.
397. Zhang H, Richardson DO, Roberts DN, Utley R, Erdjument-Bromage H, Tempst P, Côté J, Cairns BR. 2004. The Yaf9 component of the SWR1 and NuA4 complexes is required for proper gene expression, histone H4 acetylation, and Htz1 replacement near telomeres. *Mol Cell Biol* **24**: 9424–9436.
398. Zhang H, Roberts DN, Cairns BR. 2005. Genome-wide dynamics of Htz1, a histone H2A variant that poises repressed/basal promoters for activation through histone loss. *Cell* **123**: 219–231.
399. Zhang J, Xin L, Shan B, Chen W, Xie M, Yuen D, Zhang W, Zhang Z, Lajoie GA, Ma B. 2012. PEAKS DB: de

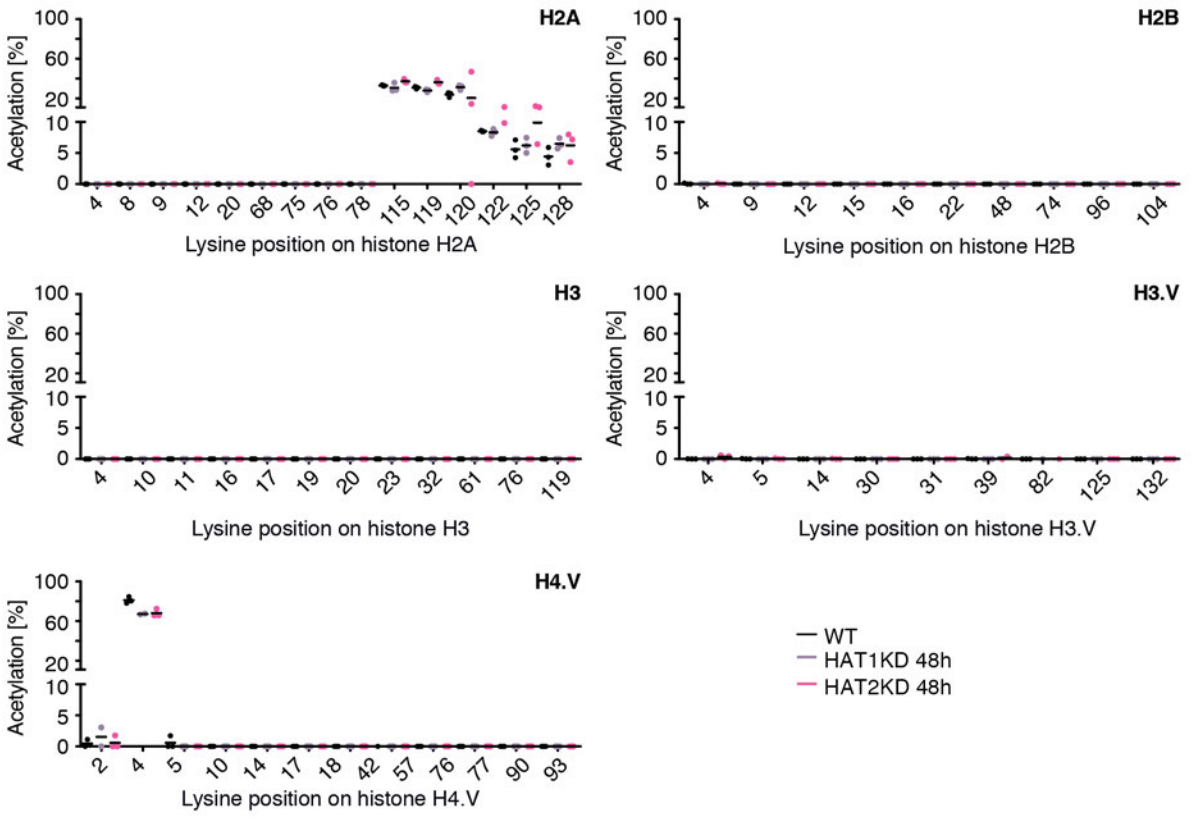
novo sequencing assisted database search for sensitive and accurate peptide identification. *Mol Cell Proteomics* **11**: M111.010587.
400. Zhang K, Xu W, Wang C, Yi X, Zhang W, Su Z. 2016. Differential deposition of H2A.Z in combination with histone modifications within related genes in rice callus and seedling. *Plant J*
401. Zhang L, Eugeni EE, Parthun MR, Freitas MA. 2003. Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma* **112**: 77–86.
402. Zhang T, Cooper S, Brockdorff N. 2015. The interplay of histone modifications - writers that read. *EMBO Rep* **16**: 1467–1481.
403. Zhao J, Herrera-Diaz J, Gross DS. 2005. Domain-wide displacement of histones by activated heat shock factor occurs independently of Swi/Snf and is not correlated with RNA polymerase II density. *Mol Cell Biol* **25**: 8985–8999.
404. Zhao Y, Garcia BA. 2015. Comprehensive Catalog of Currently Documented Histone Modifications. *Cold Spring Harb Perspect Biol* **7**: a025064.
405. Zlatanova J, Thakar A. 2008. H2A.Z: view from the top. *Structure* **16**: 166–179.

Chapter 9: Appendix

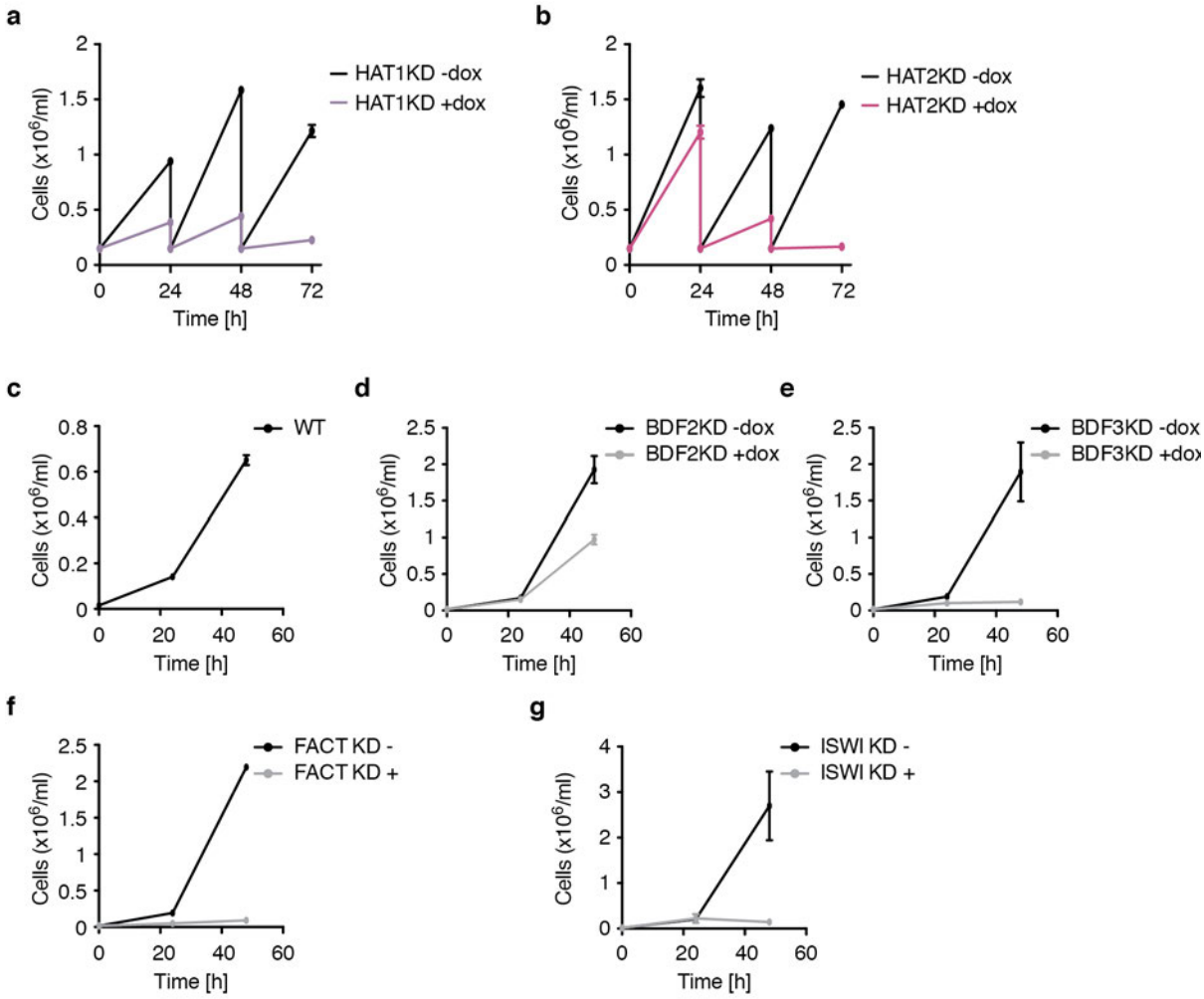
Appendix figures



Appendix figure 1. FIPQuant analysis of H2A.Z and H2B.V purified from wild type histone extracts.
a Protein sequence of H2A.Z is not fully covered following histone extraction from wild type cells. Lysine-specific acetylation degrees are shown for histone variant H2A.Z (blue) and were determined by FIPQuant using histones isolated by acid extraction from wild type (WT) cells. The acetylation percentages [%] represent the averages of the median values from each of the independent experiments (n = 3) determined by FIPQuant. Error bars indicate standard deviations. Lysine positions that could not be detected are labelled in red. **b** Quantification of lysine acetyl marks on H2B.V following histone extraction from wild type cells. Lysine-specific acetylation degrees are shown for histone variant H2B.V (blue) and were determined by FIPQuant using histones isolated by acid extraction from wild type (WT) cells. The acetylation percentages [%] represent the averages of the median values from each of the independent experiments (n = 3) determined by FIPQuant. Error bars indicate standard deviations.

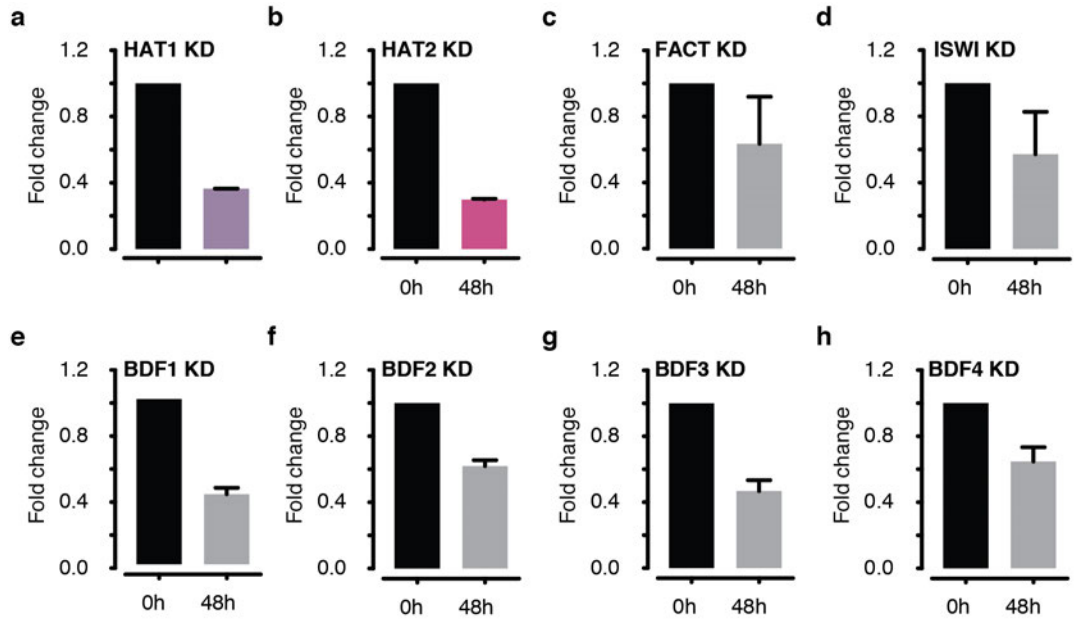


Appendix figure 2. Only TSS-specific histone acetylation is affected by HAT1 and HAT2 depletion.
Lysine-specific acetylation levels for histone H2A, H2B, H3, H3.V and H4.V are shown for wild type (WT; n=3) cells and after depletion of HAT1 (purple; n=3) and HAT2 (rose; n=3). Acetyl marks were quantified by FIPQuant using histones extracted from WT cells (black) and from 2T1 cells, which were depleted of HAT1 (purple) or HAT2 (rose) for 48 h. The plotted acetylation percentages represent the average percentages of the quantified fragment numbers, calculated for each replicate. The black bars indicate the average percentages calculated from the median values.

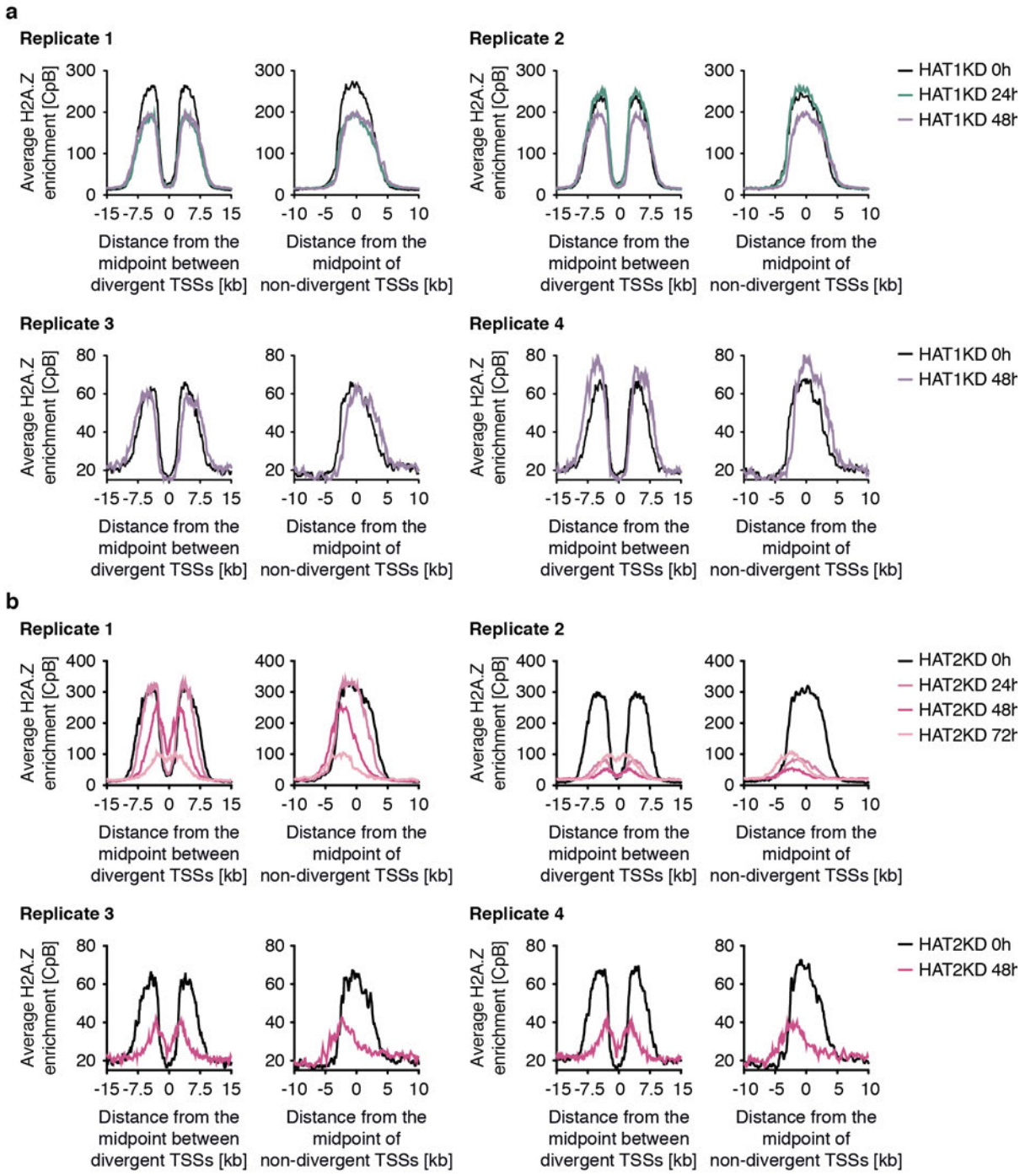


Appendix figure 3. Growth phenotypes observed following depletion of different chromatin modifiers.
a Growth analyses of 2T1 cells, induced for HAT1 depletion (+dox; purple; n=3) or not induced (-dox; black; n=3), are shown over a time period of 72 h. Cells were diluted to 0.15 x 10^6 cells per ml every 24 h. Plotted are the averages of the cell concentrations measured for each replicate (n=3). Error bars indicate standard deviations. **b** Growth analyses of 2T1 cells, induced for HAT2 depletion (+dox; rose; n=3) or not induced (-dox; black; n=3), are shown over a time period of 72 h. Cells were diluted to 0.15 x 10^6 cells per ml every 24 h. Plotted are the averages of the cell concentrations measured for each replicate (n=3). Error bars indicate standard deviations. **c** Growth analysis of wild type cells (black; n=3) is shown over a time period of 48 h. The cell density at 0 h was 0.15 x 10^6 cells per ml and was measured every 24 h. Plotted are the averages of the cell concentrations measured for each replicate (n=3). Error bars indicate standard deviations. **d** Growth analysis 2T1 cells, induced for BDF2 depletion (+dox; grey; n=3) or not induced (-dox; black; n=3), are shown over a time period of 48 h. The cell density at 0 h was 0.15 x 10^6 cells per ml and was measured every 24 h. Plotted are the averages of the cell concentrations measured for each replicate (n=3). Error bars indicate standard deviations. **e** Growth analysis 2T1 cells, induced for BDF3 depletion (+dox; grey; n=3) or not induced (-dox; black; n=3), are shown over a time period of 48 h. The cell density at 0 h was 0.15 x 10^6 cells per ml and was measured every 24 h. Plotted are the averages of the cell concentrations measured for each replicate (n=3). Error bars indicate standard deviations. **f** Growth analysis 2T1 cells, induced for FACT depletion (+dox; grey; n=3) or not induced (-dox; black; n=3), are shown over a time period of 48 h. The cell density at 0 h was 0.15 x 10^6 cells per ml and was measured every 24 h. Plotted are the averages of the cell concentrations measured for each replicate (n=3). Error bars indicate standard deviations. **g** Growth analysis 2T1 cells, induced for ISWI depletion (+dox; grey; n=3) or not induced (-dox; black; n=3), are shown over a time period of 48 h. The

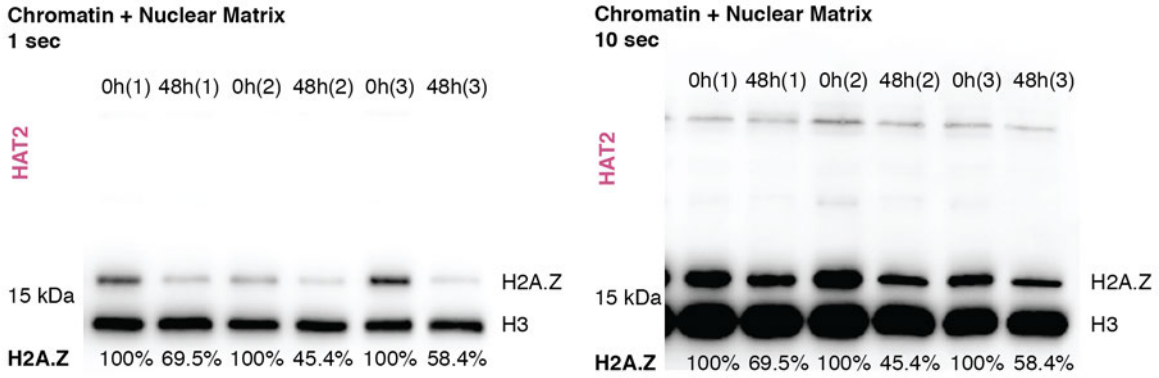
cell density at 0 h was 0.15 x 10^6 cells per ml and was measured every 24 h. Plotted are the averages of the cell concentrations measured for each replicate (n=3). Error bars indicate standard deviations.



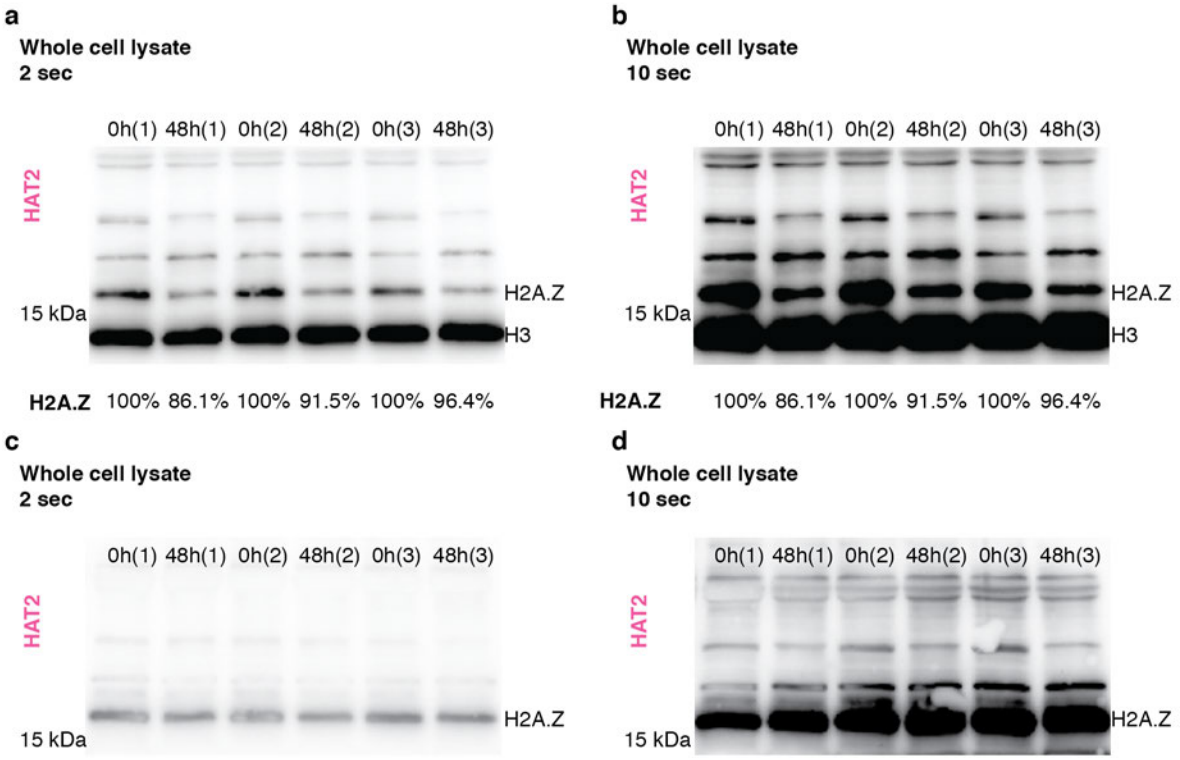
Appendix figure 4. Fold change in expression of different chromatin modifiers before and after RNAi-mediated depletion. Shown are the fold change in transcription of different chromatin modifiers before (0 h) and after RNAi-mediated depletion (48 h). The fold change in expression were determined using quantitative real-time PCR (see Quantitative real-time PCR for testing RNAi efficiencies) and calculated using the ΔΔCt-method Methods (or-Livak-method) that compares the expression of the target gene in the 48 h time point to the 0 h time point. Fold change in expression of **a** HAT1 KD 0h (black) and HAT1 KD 48 h (purple). **b** HAT2 KD 0h (black) and HAT2 KD 48 h (rose). **c** FACT KD 0h (black) and FACT KD 48 h (grey). **d** ISWI KD 0h (black) and ISWI KD 48 h (grey). **e** BDF1 KD 0h (black) and BDF1 KD 48 h (grey). **f** BDF2 KD 0h (black) and BDF2 KD 48 h (grey). **g** BDF3 KD 0h (black) and BDF3 KD 48 h (grey). **h** BDF4 KD 0h (black) and BDF4 KD 48 h (grey).



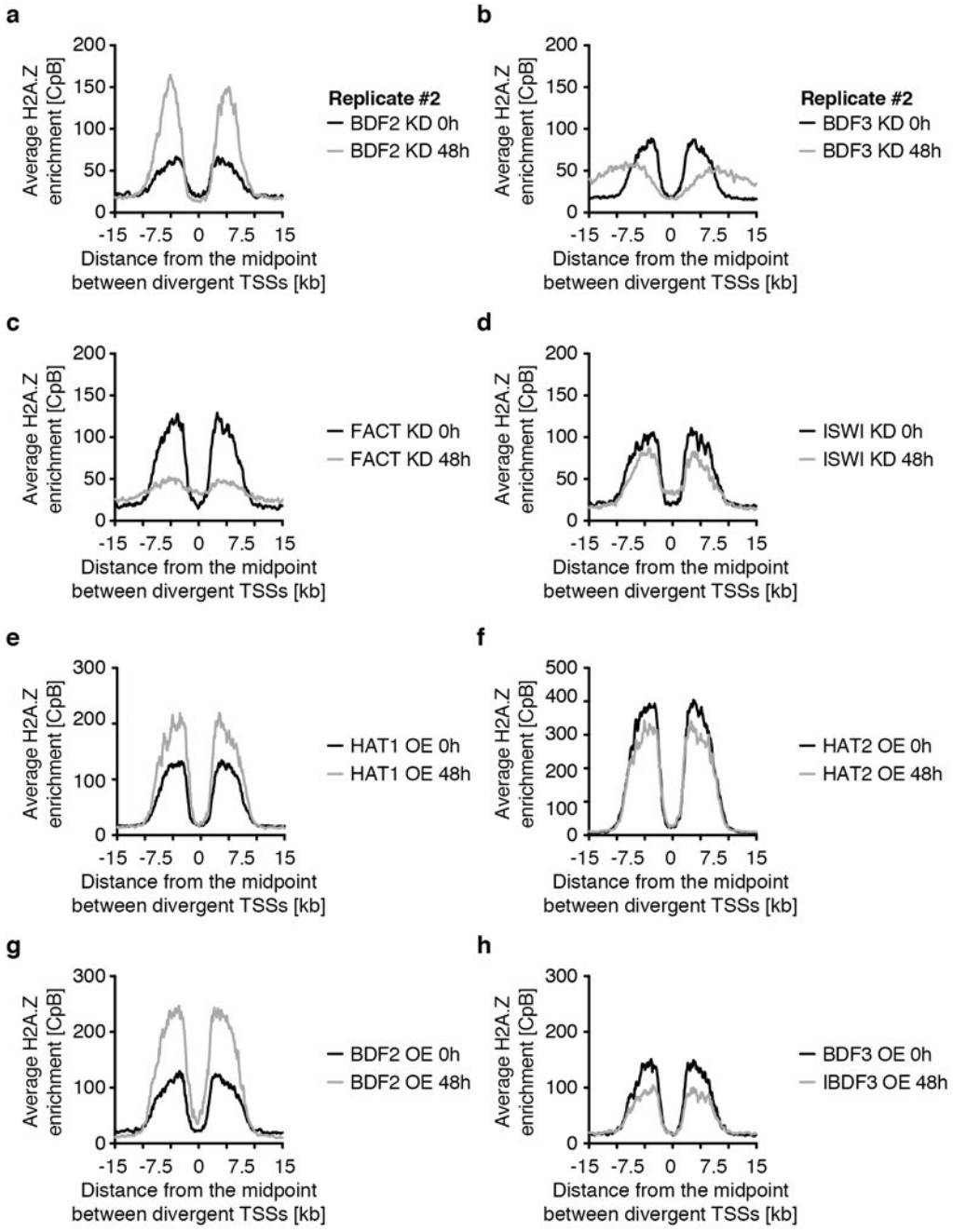
Appendix figure 5. H2A.Z-enrichment at TSSs following depletion of HAT1 and HAT2.
a MNase-ChIP-seq datasets of histone variant H2A.Z after 0 h (black; n=4), 24 h (green; n=2), and 48 h (purple; n=4) HAT1 depletion are averaged across divergent (n=37) and non-divergent (n=49) TSSs. The ChIP-seq data are normalized to the total number of reads and plotted as counts per billion reads [CpB]. **b** MNase-ChIP-seq datasets of histone variant H2A.Z after 0 h (black; n=4), 24 h (light rose; n=2), 48 h (magenta; n=4), and 72 h (rose; n=2) HAT2 depletion are averaged across divergent (n=37) and non-divergent (n=49) TSSs. The ChIP-seq data are normalized to the total number of reads and plotted as counts per billion reads [CpB].



Appendix figure 6. Levels of chromatin-bound H2A.Z decrease following HAT2 depletion.
Western blot of chromatin-associated proteins extracted from 2 x 10^6 2T1 cells, which were depleted of HAT2 (n=3) for a time period of 48 h. Loaded are the insoluble fractions, containing chromatin-bound and nuclear matrix material. The H2A.Z percentages for each replicate were calculated by comparing the H2A.Z/H3 ratio after 48 h HAT2-depletion to the H2A.Z/H3 ratio of 0 h HAT2-depletion, which was set to 100%. H2A.Z/H3 were calculated by quantifying the H2A.Z and H3 signal over the background for each lane signal using ImageJ. Exposure time: 1 sec (left) and 10 sec (right).

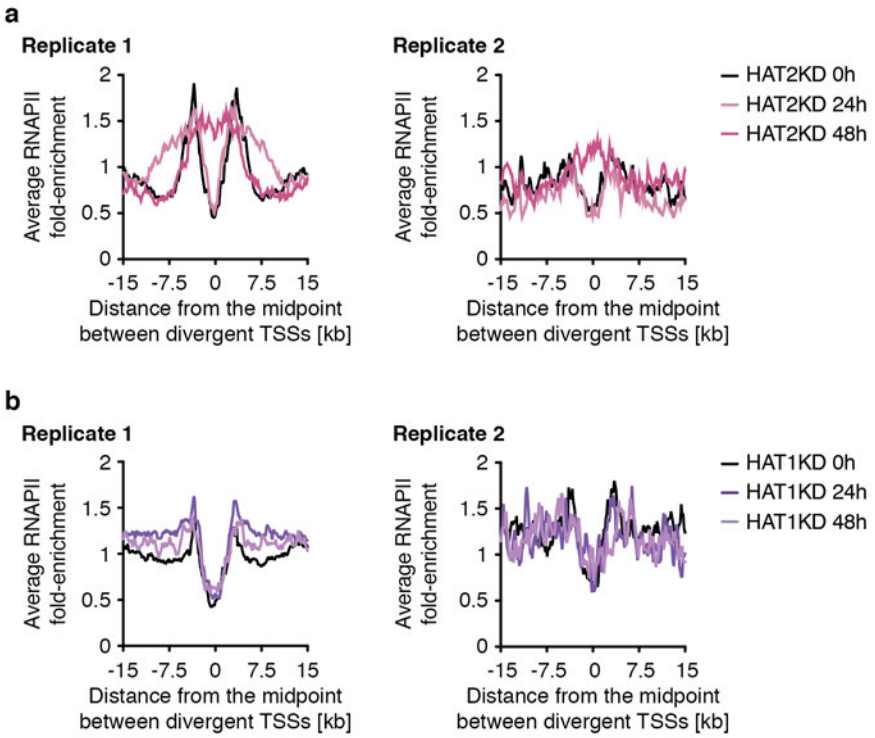


Appendix figure 7. Total protein levels of H2A.Z are not affected by HAT2 depletion.
a and **b** Western blot of whole cell extracts extracted from 2 x 10^6 from 2T1 cells, which were depleted of HAT2 (n=3) for a time period of 48 h. The H2A.Z percentages for each replicate were calculated by comparing the H2A.Z/H3 ratio after 48 h HAT2-depletion to the H2A.Z/H3 ratio of 0 h HAT2-depletion, which was set to 100%. H2A.Z/H3 were calculated by quantifying the H2A.Z and H3 signal over the background for each lane signal using ImageJ. Exposure time: 2 sec (**a)** and 10 sec (**b**). **c** and **d** Western blot of whole cell extracts extracted from 2 x 10^6 from 2T1 cells, which were depleted of HAT2 (n=3) for a time period of 48 h using an antibody that recognises H2A.Z. Exposure time: 2 sec (**c)** and 10 sec (**d**).

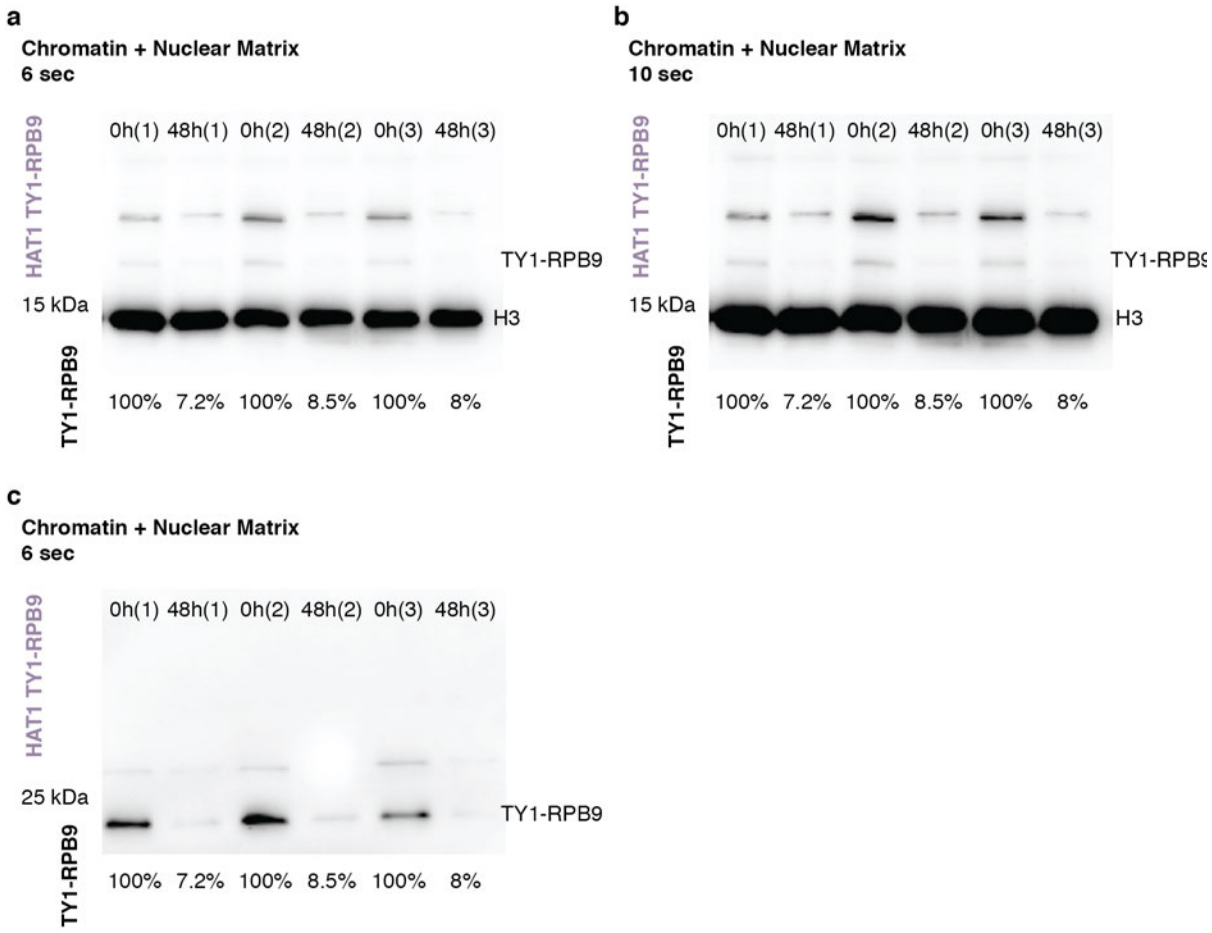


Appendix figure 8. H2A.Z-enrichment at TSSs following depletion or overexpression of different chromatin-modifying proteins.

MNase-ChIP-seq datasets of histone variant H2A.Z after 0 h (black; n=1) and 48 h (grey; n=1) depletion or overexpression (OE) of different chromatin modifying proteins are averaged across divergent (n=37) TSSs. The ChIP-seq data are normalized to the total number of reads and plotted as counts per billion reads [CpB]. **a** BDF2 KD (replicate) **b** BDF3 KD (replicate) **c** FACT KD **d** ISWI KD **e** HAT1 OE **f** HAT2 OE **g** BDF2 OE **h** BDF3 OE.



Appendix figure 9. RNAPII-enrichment at TSSs following depletion of HAT1 and HAT2.
a ChIP-seq data of RNAPII after 0 h (black), 24 h (rose), and 48 h (magenta) of HAT2 depletion are averaged across divergent (n=81) TSSs. right: replicate 1; left: replicate 2. **b** ChIP-seq data of RNAPII after 0 h (black), 24 h (light purple), and 48 h (purple) of HAT1 depletion are averaged across divergent (n=81) TSSs. right: replicate 1; left: replicate 2. **b**



Appendix figure 10. Depletion of HAT1 drastically reduces the levels of chromatin-bound RNAPII.
a and **b** Western blot of chromatin-associated proteins extracted from 2 x 10^6 2T1 cells, which were depleted of HAT1 (n=3) for a time period of 48 h. Loaded are the insoluble fractions, containing chromatin-bound and nuclear matrix material. The TY1-RPB9 percentages for each replicate were calculated by comparing the TY1-RPB9 /H3 ratio after 48 h HAT1-depletion to the TY1-RPB9 /H3 ratio of 0 h HAT1-depletion, which was set to 100%. TY1-RPB9 /H3 were calculated by quantifying the H2A.Z and H3 signal over the background for each lane signal using ImageJ. Exposure time: 6 sec (**a**) and 10 sec (**b**). **c** Western blot of chromatin-associated RNAPII. Chromatin-bound proteins were extracted from 2 x 10^6 2T1 cells, which were depleted of HAT1 (n=3) for a time period of 48 h. Loaded are the insoluble fractions, containing chromatin-bound and nuclear matrix material.

Publication List

Publications containing parts of this thesis:

1. **Kraus AJ**, Vanselow JT, Lamer S, Brink BG, Schlosser A, Siegel TN. (2020). Distinct roles for H4 and H2A.Z acetylation in RNA transcription in African trypanosomes. *Nat Commun*. DOI: 10.1038/s41467-020-15274-0. PMID: 32198348; PMCID: PMC7083915.
2. **Kraus AJ**, Brink BG, Siegel TN. (2019). Efficient and specific oligo-based depletion of rRNA. *Sci Rep*. DOI: 10.1038/s41598-019-48692-2. PMID: 31439880; PMCID: PMC6706579.
3. ElBashir R, Vanselow JT, **Kraus A**, Janzen CJ, Siegel TN, Schlosser A. (2015). Fragment ion patchwork quantification for measuring site-specific acetylation degrees. *Anal Chem*. DOI: 10.1021/acs.analchem.5b02517. PMID: 26335048.

Other publications:

1. Rosón JN, de Oliveira Vitarelli M, Costa-Silva HM, Pereira KS, da Silva Pires D, de Sousa Lopes L, Cordeiro B, **Kraus AJ**, Cruz KNT, Calderano SG, Fragoso SP, Siegel TN, Elias MC, da Cunha, JPC. (2021). Histone H2B.V demarcates strategic regions in the Trypanosoma cruzi genome, associates with a bromodomain factor and affects parasite differentiation and host cell invasion. *bioRxiv*. DOI: 10.1101/2021.06.08.447515.
2. **Kraus AJ**, Cosentino RO. (2019). Ribosome Profiling in Trypanosomatids. *Methods Mol Biol*. DOI: 10.1007/978-1-4939-9210-2_5. PMID: 30980300.

Curriculum Vitae

Attended conferences and courses

Conferences

2019 28th Annual Meeting of the German Society for Parasitology 2018, Berlin, Germany; Talk: 'Histone acetylation and histone variants during transcription initiation in *Trypanosoma brucei*'

2017 4th Symposium of the Centre of Toxins, Immune-Response and Cell Signalling „DNA metabolism and histone modifications in trypanosomatids' São Paolo, Brasil; Talk: 'Elucidating the epigenome of *Trypanosoma brucei*'

2017 7th Kinetoplastid Molecular Cell Biology Meeting 2017, Woods Hole, Massachusetts, USA, Poster: 'Histone acetylation and its associated proteins are major players in defining transcription initiation sites in *T. brucei*'

2016 Eureka 11th International GSLS Symposium 2015, Würzburg, Germany, Poster: 'Histone acetylation is a major player in defining transcription initiation sites in *T. brucei*'

2015 Eureka 10th International GSLS Symposium 2015, Würzburg, Germany, Poster: 'Unravelling the mechanism of targeted histone variant deposition in *T. brucei*'

2015 6th Kinetoplastid Molecular Cell Biology Meeting 2015, Woods Hole, Massachusetts, USA, Talk: 'Quantitative Analysis of the Histone Acetylome and its Role in Histone Variant Deposition'

Transferable skills training

2015 – 2016 Mentoring Life Sciences Program, GSLS

- Skill Profiling and Self-Promotion
- Tools and Strategies for a Successful Time- and Self-Management
- Small Talk at International Conferences
- Conflict Management
- Dos and Taboos of Business and Table Etiquette

GSLS workshops

- Introduction into Biotech Industries, Christian Grote-Westrick
- Project management in Biotech industries, Christian Grote-Westrick
- Software Carpentry for beginners, Dr. Konrad Förstner, Malvika Sharan, Markus Ankenbrand
- Cover letter & CV, Robert Zaal
- Good Scientific Practice, Dr. Stephan Schröder-Köhne
- Self-awareness coaching, Cornelia C. Fink
- Scientific writing for PhD students, Dr. Andrew Davis

Other workshops

- Vektorgrafiken erstellen mit Adobe Illustrator, Rechenzentrum Uni Würzburg
- Scientific Computing Skills, Graduiertenzentrum Weihenstephan, Dr. Konrad Förstner, Florian Thöle