



**New insights into the histone variant H2A.Z incorporation pathway in  
*Trypanosoma brucei***

---

**Neue Erkenntnisse zum Einbau der Histonvariante H2A.Z in *Trypanosoma brucei***

**Doctoral thesis**

for a doctoral degree

at the Graduate School of Life Sciences,  
Julius-Maximilians-Universität Würzburg,

Section Infection and Immunity

submitted by

**Tim Vellmer**

from

**Holzminden**

**Würzburg 2021**

**Submitted on:**

**Members of the Thesis Committee:**

**Chairperson: Prof. Dr. Dandekar**

**Primary Supervisor: Prof. Dr Christian J. Janzen**

**Second Supervisor: Prof. Dr. Tim Nicolai Siegel**

**Third Supervisor: Prof. Dr. Klaus Brehm**

**Date of Public Defence:**

**Date of Receipt of Certificates:**

*„Wenn wir wüssten, was wir tun, würde man es nicht  
Forschung nennen.“*

*Albert Einstein (1879 - 1955)*

## Affidavit

I hereby confirm that my thesis entitled "New insights into the histone variant H2A.Z incorporation pathway in *Trypanosoma brucei*" 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.

Würzburg

Place, Date

Tim Vellmer

## Eidesstattliche Erklärung

Hiermit erkläre ich an Eides statt, die Dissertation „Neue Erkenntnisse zum Einbau der Histonvariante H2A.Z in *Trypanosoma brucei*“ eigenständig, d.h. insbesondere selbständig und ohne Hilfe 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.

Würzburg

Ort, Datum

Tim Vellmer

## Table of contents

Members of the Thesis Committee .....	2
Affidavit .....	4
Table of contents .....	5
Acknowledgement .....	12
Summary .....	13
Zusammenfassung .....	14
Abbreviations .....	16
1. Introduction .....	23
1.1 Chromatin and its structure.....	23
1.2 The nucleosome structure.....	24
1.3 Nucleosome modifications and chromatin structure .....	24
1.4 Chromatin remodelling complexes .....	27
1.4.1 Chromatin remodelling complexes of the Inositol non-fermenting 80 subfamily ....	29
1.5 The histone variant H2A.Z.....	32
1.5.1 The role of H2A.Z in transcription .....	34
1.5.2 Factors of the H2A.Z transcription regulation machinery that interact with the RNAP II transcription pre-initiation complex.....	35
1.6 <i>Trypanosoma brucei</i> , a small overview .....	37
1.6.1 <i>T. brucei</i> as a model organism for <i>Kinetoplastids</i> .....	37
1.6.2 Transcription and H2A.Z in <i>T. brucei</i> .....	38
1.6.3 Transcription-associated post-translational modifications in <i>T. brucei</i> .....	39
1.6.4 Chromatin remodelling and its complexes in <i>T. brucei</i> .....	40
1.7 Aim of the study .....	41
2. Materials and Methods.....	43
2.1 Bacterial culture .....	43
2.1.1 Bacteria growth .....	43
2.1.2 Transformation of chemically competent <i>E. coli</i> .....	43
2.1.3 Isolation of plasmid DNA from <i>E. coli</i> .....	43
2.2 Trypanosomatid methods.....	43
2.2.1 <i>T. brucei</i> strains .....	43

## Table of contents

2.2.2 <i>In vitro</i> cultivation of <i>T. brucei</i> .....	47
2.2.3 Growth curves .....	47
2.2.4 Genetic modification of Trypanosomes .....	48
2.2.5 Freezing and thawing.....	48
2.2.6 Isolation of genomic DNA from <i>T. brucei</i> cells .....	48
2.2.7 Purification of DNA for diagnostic polymerase chain reaction .....	49
2.2.8 Luciferase assay .....	49
2.3 Flow cytometry analyses.....	49
2.3.1 Cell cycle analysis.....	49
2.3.2 Live-dead analysis.....	50
2.4 DNA methods and DNA plasmids .....	50
2.4.1 Polymerase chain reaction .....	50
2.4.2 Agarose gel electrophoresis.....	55
2.4.3 Restriction digest.....	56
2.4.4 Measurement of DNA concentration.....	56
2.4.5 Recombination cloning .....	56
2.4.6 Ligation.....	59
2.4.7 <i>In vivo</i> Ligation.....	59
2.4.8 DNA sequencing .....	59
2.4.9 Plasmids used and generated in course of the project.....	59
2.5 RNA methods.....	63
2.5.1 Northern Blot analysis.....	63
2.6 Protein methods .....	63
2.6.1 Preparation of whole cell lysates for sodium dodecyl sulfate polyacrylamide gel electrophoresis .....	63
2.6.2 SDS-PAGE .....	63
2.6.3 Western Blot .....	64
2.6.4 Extraction of chromatin-associated proteins .....	65
2.6.5 Co-Immunoprecipitation .....	65
2.6.6 Biotin identification.....	66
2.6.7 Mass spectrometry and data analyses.....	67

## Table of contents

2.6.8 MNase-ChIP sequencing.....	67
2.6.9 Bioinformatics analysis.....	68
2.6.10 Reads processing and mapping .....	68
2.6.11 Differential binding analysis.....	68
2.7 Microscopy.....	69
2.7.1 Fluorescence microscopy analysis.....	69
2.7.2 EM sample preparation and imaging.....	69
3. Results.....	71
3.1 Identification of a potential SWR1-like complex.....	71
3.1.1 Characterization of SNF2 proteins in <i>T. brucei</i> .....	71
3.1.2 Identification of a new SNF2 CRC of the INO80 subfamily.....	73
3.1.3 The identified SNF2 complex has characteristics of a SWR1-like complex with an unusual complex composition .....	79
3.2 The <i>Tb</i> SWR1 complex regulates H2A.Z incorporation, transcriptional activity and chromatin structure .....	83
3.2.1 Characterisation of <i>Tb</i> SWR1 .....	83
3.2.2 <i>Tb</i> SWR1 is required for H2A.Z incorporation in <i>T. brucei</i> .....	85
3.2.3 Depletion of <i>Tb</i> SWR1 leads to reduced transcriptional activity .....	88
3.2.4 Loss of chromatin-associated H2A.Z affects chromatin structure .....	91
3.3 The HAT1 and the HAT2 complex.....	93
3.3.1 The HAT2 complex has a so far undescribed complex composition .....	94
3.3.2 The HAT1 complex has characteristics of a NuA4 complex .....	99
3.3.3 H2A.Z appears to be involved in the DNA DSB response in <i>T. brucei</i> .....	102
3.4. Identification of potential COMPASS and SAGA complex subunits .....	103
4. Discussion, Conclusion and Outlook.....	107
4.1 Analysis of the composition of the <i>Tb</i> SWR1, HAT1 and HAT2 complex.....	107
4.1.1 The <i>Tb</i> SWR1 complex .....	108
4.1.2 The HAT2 complex .....	110
4.1.3 The HAT1 complex .....	111
4.1.4 Evaluation of the experimental approach used to identify the <i>Tb</i> SWR1, HAT2 and HAT1 complex .....	112
4.2 <i>Tb</i> SWR1 and the loss of RNA.....	114

## Table of contents

4.3 The role of <i>Tb</i> SWR1 in genome integrity.....	115
4.4 Bringing together the differences: the interplay of HAT2, <i>Tb</i> SWR1 and HAT1 in course of different cellular processes.....	116
4.4.1 The potential interplay of HAT2, <i>Tb</i> SWR1 and HAT1 in course of H2A.Z-mediated transcription regulation .....	116
4.4.2 The potential interplay of the HAT2 and <i>Tb</i> SWR1 complex in course of H2A.Z-mediated DNA DSB repair .....	118
4.4.3 The presence of transcription factors and enzymes responsible for post-translational modifications in close proximity to <i>Tb</i> SWR1 .....	118
4.5 Conclusion and Outlook:.....	121
5. References .....	122
6. Appendix.....	137
6.1 Supplementary figures .....	137
Curriculum Vitae.....	157



**Figure Index:**

Figure 1: Schematic depiction of the chromatin structure. ....	24
Figure 2: Example of cross interactions of histone PTMS in eukaryotes. ....	26
Figure 3: The four subclasses of SNF2 ATPases .....	27
Figure 4: The role of each of the four SNF2 ATPase subfamilies in chromatin remodelling ....	29
Figure 5: A comparison between the SWR1 and INO80 complex. ....	31
Figure 6: Overview of the interplay between H2A.Z and its remodelling complexes in DNA DSB.....	33
Figure 7: Depiction of a eukaryotic transcription start site.....	34
Figure 8: Interaction network of PTMs, proteins and protein complexes that regulate transcription initiation.....	36
Figure 9: Genome organization of <i>T. brucei</i> .....	40
Figure 10: Overview of SNF2 ATPases in <i>T. brucei</i> .....	72
Figure 11: Representative western blots of the RuvB2 co-IPs.....	73
Figure 12: Control of the co-IPs that identified the <i>TbSWR1</i> complex .....	76
Figure 13: Core region of Tb927.9.8510 ( <i>TbSWRC3</i> ) is homologous to the human H3K4 specific SET7/9 histone methyltransferases.....	80
Figure 14: The <i>T. brucei</i> SNF2 ATPase complex exhibits characteristics of a SWR1 complex....	82
Figure 15: Loss of <i>TbSWR1</i> leads to cell death.....	83
Figure 16: Depletion of <i>TbSWR1</i> leads to the generation of anucleated cells.....	84
Figure 17: Loss of <i>TbSWR1</i> leads to a reduction of chromatin-associated H2A.Z.....	86
Figure 18: Loss of <i>TbSWR1</i> (Tb927.11.10730) leads to a reduced H2A.Z deposition at TSS ....	87
Figure 19: Depletion of <i>TbSWR1</i> and H2A.Z caused a decrease of reporter luciferase activity within a PTU.....	89
Figure 20: Depletion of <i>TbSWR1</i> leads to a reduction of mRNA and rRNA.....	90
Figure 21: Loss of <i>TbSWR1</i> and RNAP II leads to chromatin condensation.....	92
Figure 22: Test Western Blot of the co-IP experiments with HA-HAT2 and Ty1-Bdf3 .....	94
Figure 23: Predicted secondary structure of the C-terminal end of Tb927.4.2340 is homologous to known bromo-domains in <i>T. cruzi</i> and <i>L. donovani</i> .....	96
Figure 24: Loss of Bdf3 leads to cell death and a reduction of chromatin-associated H2A.Z... 98	98
Figure 25: Test Western Blot of the co-IP experiments with HA- <i>TbEaf6</i> and Ty1- <i>TbEaf3</i> .....	99
Figure 26: Chromatin-associated H2A.Z level increase after phleo-mycin treatment.....	102
Figure 27: Classification of the proteins identified in course of the <i>TbSWRC1</i> BioID.....	104
Figure 28: A disadvantage of MS-coupled co-IP experiments.....	113
Figure 29: Putative interplay of the HAT2, <i>TbSWR1</i> and HAT1 complex in course of H2A.Z-mediated transcription regulation .....	117
Figure 30: Extended model of a putative H2A.Z-mediated transcription regulation mechanism in <i>T. brucei</i> .....	120

**Figure Index:**

Table 1: Overview of SNF2 complexes of the INO80 subfamily in different organisms.....	28
Table 2: Parental and transgenic <i>T. brucei</i> cell lines used and generated in this study .....	44
Table 3: Transgenic <i>T. brucei</i> cell lines generated and used in this study.....	44
Table 4: Drug selection used for <i>T. brucei</i> cell culture.....	47
Table 5: Primers used for <i>T. brucei</i> methods.....	51
Table 6: Primers used for diagnostic PCRs.....	55
Table 7: Primers used for RNAi constructs.....	56
Table 8: Primers used for sequencing.....	59
Table 9: Used and generated plasmids.....	59
Table 10: DNA probes used for Northern Blot analysis .....	63
Table 11: Used antibodies .....	65
Table 12: Summary of the RuvB2 co-IP.....	74
Table 13: summary of co-IP experiments that identified the <i>TbSWR1</i> complex.....	78
Table 14: Summary of the HA-HAT2 and Ty1-Bdf3 co-IP .....	97
Table 15: Summary of the co-IPs that identified the HAT1 complex.....	100
Table 16: List of proteins with potential involvement in transcription regulation identified in the <i>TbSWRC2</i> BioID .....	106

**Appendix Figure Index:**

Figure S1: Alignment of the histones H3 and H4 sequences..... 137

Figure S2: Control of the cell lines used for the co-IP experiments to identify the *Tb*SWR1 complex..... 138

Figure S3: Volcano plots of the co-IPs that identified the *Tb*SWR1 complex part 1..... 139

Figure S4: Volcano plots of the co-IPs that identified the *Tb*SWR1 complex part 2..... 140

Figure S5: Growth curves of *Tb*SWRC3 RNAi cell lines..... 141

Figure S6: Depletion of *Tb*SWRC1 and *Tb*SWRC2 reduce the amount of chromatin-associated H2A.Z..... 142

Figure S7: ChiP-seq results of *Tb*SWR1-depleted cells, chromosomes 1-4..... 143

Figure S8: ChiP-seq results of *Tb*SWR1-depleted cells, chromosomes 5-8..... 144

Figure S9: ChiP-seq results of *Tb*SWR1-depleted cells, chromosomes 9-11 ..... 145

Figure S10: Depletion of the histone acetyltransferase HAT1 and HAT2 caused a decrease of reporter luciferase activity within a PTU ..... 146

Figure S11: Growth curves of RNAi cell lines used for the luciferase assay ..... 147

Figure S12: Evaluation of the cell lines used for the co-IP experiments that identified the HAT1 and HAT2 complex ..... 148

Figure S13:Volcano plots of the co-IPs that identified the HAT2 complex ..... 149

Figure S14: Volcano plots of the co-IPs that identified the HAT1 complex..... 150

Figure S15: Predicted secondary structure of the C-terminal end of Tb927.9.13320 is homologous to an FHA-Domain ..... 151

Figure S16: Representative western blots of the *Tb*SWRC2 BioID and its control..... 152

Figure S17: PANTHER analysis of the proteins identified in the *Tb*SWRC2 BioID, part 1..... 153

Figure S18: PANTHER analysis of the proteins identified in the *Tb*SWRC2 BioID, part 2..... 154

Figure S19: Core region of Tb927.9.8100 is homologous to the SET-domain of the H3K4 specific COMPASS SET-methyltransferase..... 155

## **Acknowledgement**

First and foremost, I would like to express my genuine gratitude to my primary supervisor, Christian Janzen, for his guidance and support in course of the five years of my PhD. Thank you for the dozens of “in the hallway and in between” talks that we had, thank you for helping me to stay focused and thank you for putting your trust in me when I developed this project.

I am also very grateful to Nicolai Siegel and Klaus Brehm, who kindly agreed to be my second and third supervisor. Thank you for the great input you provided in our thesis advisory committee meetings.

A special thanks goes to Falk Butter, Albert Fradera Sola and the whole research group Butter. Without your efforts this whole project would still be nothing more than an idea. My gratitude also goes to the present and former lab members of the Janzen lab Elisabeth, Elisa, Nadine, Nico, Laura, Helena, Nicole, Marie and Zdenka. Thank you for a great and professional atmosphere with lots of discussions and also solidarity during the hard times. Thank you, Helena and Nicole, for guiding me through the first months of my PhD and thank you Laura for being such a strong support in the last month of my PhD. A special thanks goes to Elisabeth the calm and patient “backbone” of our lab for all her work in the electron microscopy facility.

In this respect I also need to thank the whole research group of Nicolai Siegel for the joined lab meetings at the beginning of my PhD. Amelie without your work it would not have been impossible to assemble this whole thesis the way I did. Caro you always had an open ear and you were such a great help when introducing me to the world of ChIP-seq analysis. I would like to thank the whole Zoologie I department for many productive discussions in course of our TrypClub sessions, for Friday afternoon beers and social events. A special thanks goes to Susanne Kramer for her work for my publication and Markus Engstler for his interest in my work and his interest in the progress I’ve made in course of the years.

Marie and Carina, both of you became incredibly good friends. Especially during times when I was sitting between a rock and a hard place I could always count on you and for this I am more than grateful.

My deep-felt gratitude goes to my girlfriend, Birte, for your interest in my work and the “endless” discussions about my ideas and problems. No matter how tired or exhausted you have been, I always had your ear. Without your constant support and your never-ending patience nothing of this would have ever been possible. Thank you for travelling this long and very rocky road with me. Thank you for your love and strength especially during the last months.

„Last but not least“ gilt mein tiefer Dank meinen Eltern, die mich immer wieder aufgefangen haben, wenn es schlecht lief, sowie meiner „Schwester“, die sich immer wieder erkundigt hat, wie ich vorankomme. Danke für alles. Ohne eure langjährige Unterstützung wäre das alles nicht möglich gewesen

## Summary

The histone variant H2A.Z is a key player in transcription regulation in eukaryotes. Histone acetylations by the NuA4/TIP60 complex are required to enable proper incorporation of the histone variant and to promote the recruitment of other complexes and proteins required for transcription initiation. The second key player in H2A.Z-mediated transcription is the chromatin remodelling complex SWR1, which replaces the canonical histone H2A with its variant. By the time this project started little was known about H2A.Z in the unicellular parasite *Trypanosoma brucei*. Like in other eukaryotes H2A.Z was exclusively found in the transcription start sites of the polycistronic transcription units where it keeps the chromatin in an open conformation to enable RNA-polymerase II-mediated transcription. Previous studies showed the variant co-localizing with an acetylation of lysine on histone H4 and a methylation of lysine 4 on histone H3. Data indicated that HAT2 is linked to H2A.Z since it is required for acetylation of lysine 10 on histone H4. A SWR1-like complex and a complex homologous to the NuA4/TIP60 could not be identified yet.

This study aimed at identifying a SWR1-like remodelling complex in *T. brucei* and at identifying a protein complex orthologous to NuA4/TIP60 as well as at answering the question whether HAT2 is part of this complex or not. To this end, I performed multiple mass spectrometry-coupled co-immunoprecipitation assays with potential subunits of a SWR1 complex, HAT2 and a putative homolog of a NuA4/TIP60 subunit. In the course of these experiments, I was able to identify the *Tb*SWR1 complex. Subsequent cell fractionation and chromatin immunoprecipitation-coupled sequencing analysis experiments confirmed, that this complex is responsible for the incorporation of the histone variant H2A.Z in *T. brucei*. In addition to this chromatin remodelling complex, I was also able to identify two histone acetyltransferase complexes assembled around HAT1 and HAT2. In the course of my study data were published by the research group of Nicolai Siegel that identified the histone acetyltransferase HAT2 as being responsible for histone H4 acetylation, in preparation to promote H2A.Z incorporation. The data also indicated that HAT1 is responsible for acetylation of H2A.Z. According to the literature, this acetylation is required for proper transcription initiation. Experimental data generated in this study indicated, that H2A.Z and therefore *Tb*SWR1 is involved in the DNA double strand break response of *T. brucei*. The identification of the specific complex composition of all three complexes provided some hints about how they could interact with each other in the course of transcription regulation and the DNA double strand break response. A proximity labelling approach performed with one of the subunits of the *Tb*SWR1 complex identified multiple transcription factors, PTM writers and proteins potentially involved in chromatin maintenance. Overall, this work will provide some interesting insights about the composition of the complexes involved in H2A.Z incorporation in *T. brucei*. Furthermore, it is providing valuable information to set up experiments that could shed some light on RNA-polymerase II-mediated transcription and chromatin remodelling in *T. brucei* in particular and *Kinetoplastids* in general.

## Zusammenfassung

Die Histonvariante H2A.Z ist ein Schlüsselement bei der Transkriptionsregulation in Eukaryoten. Histonacetylierungen die vom NuA4/Tip60 Komplex prozessiert werden, sind für den korrekten Einbau der Variante unerlässlich. Darüber hinaus erlauben diese posttranslationalen Modifikationen die Rekrutierung weiterer Proteine und Komplexe die für die Transkription notwendig sind. Ein weiteres Schlüsselement der mittels H2A.Z regulierten Transkription ist der Komplex zur Umstrukturierung des Chromatins SWR1, welcher das kanonische Histon H2A gegen seine Variante austauscht. Zu Beginn dieses Projektes war der Wissenstand bezüglich der Histonvariante H2A.Z in dem einzelligen Parasiten *Trypanosoma brucei* limitiert. Wie in anderen eukaryotischen Organismen wurde die Variante ausschließlich an den Startpunkten der polyzistronischen Transkriptionseinheiten gefunden, an denen es für die Öffnung des Chromatins verantwortlich ist und so die Transkription mittels RNA-Polymerase II ermöglicht. Vorangegangene Studien konnten zeigen, dass die Variante mit einer Acetylierung des Lysins 10 im Histon H4 und einer Methylierung des Lysins 4 im Histon H3 co-lokalisiert. Einige Daten lieferten den Hinweis, dass die Histon-Acetyltransferase HAT2 mit H2A.Z in Zusammenhang steht, da diese die Acetylierung des Lysins 10 im Histon H4 prozessiert. Komplexe die in ihrer Funktion dem SWR1 oder dem NuA4/TIP60 Komplex entsprechen, konnten bisher noch nicht gefunden werden.

Die vorliegende Arbeit zielt darauf ab Komplexe zu identifizieren, die in ihrer Funktion dem SWR1 sowie dem NuA4/TIP60 Komplex entsprechen. Zudem soll die Frage geklärt werden ob HAT2 Teil eines möglichen NuA4/TIP60 Komplexes ist. In diesem Zusammenhang habe ich mehrere Massenspektrometrie gekoppelte Co-Immunopräzipitationen mit potenziellen Untereinheiten eines SWR1 Komplexes sowie HAT2 und einem Protein welches ortholog zu einer NuA4/TIP60 Untereinheit ist, durchgeführt. Im Verlauf dieser Experimente konnte der SWR1 Komplex in *T. brucei* (*TbSWR1*) identifiziert werden. Anschließende Zellfraktionierungen sowie Chromatin Immunopräzipitationen gekoppelte Sequenzanalysen konnten bestätigen, dass der identifizierte Komplex für den Einbau der Histonvariante H2A.Z zuständig ist. Darüber hinaus konnten neben diesem Komplex noch zwei weitere Komplexe identifiziert werden, die jeweils die Histonacetyltransferasen HAT1 und HAT2 als Kernkomponenten enthalten. Im Verlauf meiner Arbeit wurden von der Arbeitsgruppe von Nicolai Siegel Daten publiziert die zeigten, dass die Histonacetyltransferase HAT2, in Vorbereitung auf den Einbau von H2A.Z, für die Acetylierung des Histons H4 verantwortlich ist. Im Gegenzug ist HAT1 für die Acetylierung von H2A.Z notwendig, welche wiederum für die korrekte Initiation der Transkription benötigt wird. Damit entspricht die Funktion der Acetylierung von H2A.Z in *T. brucei* der in der Literatur beschriebenen Funktion. Experimentelle Daten die im Verlauf dieser Arbeit generiert wurden, lieferten einen Hinweis darauf, dass H2A.Z auch an der Reparatur von DNS Doppelstrangbrüchen beteiligt ist.

## Zusammenfassung

Die Aufschlüsselung der spezifischen Zusammensetzung aller drei Komplexe gab einige Hinweise darauf, wie sie sowohl während der Transkriptionsregulation als auch der Reparatur von DNS Doppelstrangbrüchen miteinander interagieren.

Im Zuge einer molekularen Umgebungskartierung, die mit einer der Untereinheiten des *Tb*SWR1 Komplexes durchgeführt wurde, konnten mehrere Transkriptionsfaktoren und Enzyme zur Histonmodifizierung identifiziert werden. Dabei wurden auch einige Proteine identifiziert, welche möglicherweise mit der Umformung des Chromatins in Zusammenhang stehen. Abschließend ist festzuhalten, dass diese Arbeit einige äußerst interessante Einsichten über die Zusammensetzung der Komplexe, die am H2A.Z Einbau in *T. brucei* beteiligt sind, liefern konnte. Darüber hinaus stellt sie einige wertvolle Informationen zur Verfügung. Diese könnten zur gezielten Planung von Experimenten genutzt werden, um mehr über RNA-Polymerase II vermittelte Transkription und Chromatin Umstrukturierung in *T. brucei* im speziellen und in Kinetoplastiden im Allgemeinen zu erfahren.

## Abbreviations

2T1	VSG221-expressing, tagged, clone1
ADP	Adenosine diphosphate
AF9	ALL1 (Leukemia, Acute Lymphocytic, Susceptibility To, 1) fused gene from chromosome 9
Amp	Ampicillin
Anp32e	Acidic Nuclear Phosphoprotein 32 Family Member E
APS	Ammonium persulfate
ARP/Arp(5)	Actin-related protein (5)
ASF1	Anti-Silencing Factor 1
asn	Asparagine
ATAC	Assay for Transposase-Accessible Chromatin
ATP	Adenosine triphosphate
BCNT	Bucentaur
Bdf(1)	Bromodomain factor (1)
BioID	Proximity-dependent biotin identification
Blas	Blasticidin
<i>BLE</i>	Bleomycin resistance gene
bp	Base pair
BSA	Bovine serum albumin
<i>BSD</i>	Blasticidin-S deaminase resistance gene
BSF	Bloodstream form
btb	BR-C, ttk and bab
C	Celsius
<i>C. elegans</i>	<i>Caenorhabditis elegans</i>
CAF1	chromatin assembly factor 1
Cbf5	Centromere-binding factor 5
CHD	Chromodomain helicase DNA-binding
ChIP	Chromatin-immunoprecipitation
ChIP-seq	Chromatin-immunoprecipitation coupled sequencing
ch. n	Charged negative
ch. p.	Charged positive
CITFA3	subunit 3 of the class I transcription factor A
CLMS	Cross-linking mass spectrometry
cm <sup>2</sup>	Square centimeter
co-IP	Co-immunoprecipitation
COMPASS	Complex proteins associated with SET1
CRC(s)	Chromatin-remodelling complexes
C-terminal	Carboxy-terminal
<i>D. melanogaster</i>	<i>Drosophila melanogaster</i>
DAMP1	DNA Methyltransferase 1-associated Protein 1



## Abbreviations

DBD	DNA-binding domain
DMAP1	DNA methyltransferase 1-associated protein 1
DNA	Deoxyribonucleic acid
dNTPs	Deoxyribonucleoside triphosphate(s)
DOT1	Disruptor of telomeric silencing 1
DSB	Double-strand break
dTIP60	Drosophila TIP60 (Tat interactive protein 60 kDa)
DTT	Dithiothreitol
DUF1935	Domain-of-unknown-function 1935
<i>E. coli</i>	<i>Escherichia coli</i>
Eaf(x)	ESA1-associated factor (x)
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylenglycol-bis(aminoethylether)-N,N,N',N'-tetraessigsäure
EM	Electron microscopy
Ep1	Enhancer of polycomb-like 1
ER $\alpha$	Estradiol-receptor Alpha
Esa1	Essential SAS2-related acetyltransferase 1
EtOH	Ethanol
eYFP	Enhanced yellow fluorescent protein
FACT	Facilitates chromatin transcription
FCS	Enhanced yellow fluorescent protein
FHA	Forkhead
fw	Forward
g	Gram
<i>g</i>	Relative centrifugal force
GAS41	Glioma amplified sequence 41
Gcn5	General control non-de-repressible 5
gDNA	Genomic DNA
GO	Gene ontology
GOI	Gene of interest
G-C	Guanine-cytosine
h	Hours
<i>H. sapiens</i>	<i>Homo sapiens</i>
HA	Hemagglutinin
HAT(1)	Histone-acetyltransferase (1)
Hda1	Histone deacetylase 1
HDAC(1)	Histone deacetylase (1)
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIRA	Histone regulator A
HIT	Histidine triad 1
HMI-9	Hirumi's modified Iscove's medium 9

## Abbreviations

HP1	Heterochromatin-protein-1
HSA	Helicase/SANT-associated
HSS	HAND (named by the crystal structure of the domain)-SANT (Swi3, Ada2, N-Cor, and TFIIB ) -SLIDE (SANT-like SWI-domain)
Htz1	Histone H2A.Z in <i>Saccharomyces cerevisiae</i>
hy	Hydrophobic
Hyg	Hygromycin
ICP4	Infected cell protein 4
les(x)	Ino eighty subunit (x)
IMDM	Iscoe's modified Dulbecco's medium
INO80	Inositol non-fermenting 80
ISWI	Imitation switch
i.d.	Identity
J-Base (Base-J)	$\beta$ -D-glucosyl-hydroxymethyluracil
JBP(x)	J-Base binding protein (x)
kbp	Kilobasepair
kDa	Kilodalton
KO	Knockout
Ks	Kelvin per second
Ku80	Ku autoantigen 80 kDa
kV	Kilovolt
l	Liter
LB	Lysogeny broth
<i>L. donovani</i>	<i>Leishmania donovani</i>
<i>M. musculus</i>	<i>Mus musculus</i>
mA	Milliampere
Mb	Megabase
MeOH	Methanol
mg	Milligram
min	Minutes
ml	Milliliter
MLL(x)	Mixed-lineage leukemia (x)
mM	Millimolar
MNase	Micrococcal nuclease
MOPS	3-(N-morpholino) propanesulfonic acid
MORF4	Mortality factor 4
MRG	MORF4 related gene
mRNA	Messenger RNA
MRX	Mre11 (Meiotic-recombination 11), Rad50 (Radiation 50), and Xrs2 (X-ray sensitive 2)
MS	Mass spectrometry

## Abbreviations

MSL(3)	Male specific lethal (complex subunit 3)
MYST	MOZ (monocytic leukemia zinc finger protein), Ybf2 (yeast binding factor 1), Sas2 (Something About Silencing 2), TIP60 (Tat-interacting protein, 60 kDa)
μl	Microliter
μm	Micrometer
μM	Micromolar
NCBI	National Center for Biotechnology Information
Neo	Neomycin
NES	Nuclear enrichment score
ng	Nanogram
Nhp10	Non-histone protein 10
nm	Nanometer
N-terminal	Carboxy-terminal
NuA4	Nucleosome acetyltransferase of H4
ORF	Open reading frame
PAC	Puromycin N-acetyltransferase
PAGE	Polyacrylamide gel electrophoresis
PAPA-1	PAP-1 (Pim1 (Proviral integration site for Moloney murine leukemia virus-1)-associated protein-1)-associated protein-1
PBS	Phosphate-buffered saline
PCF	Procyclic form
PCNA	Proliferating cell nuclear antigen
PCR	Polymerase chain reaction
PDB ID	Protein Data Bank Identification code
PFR	Paraflagellar rod
pH	Potential of hydrogen
PHD	Plant homeodomain
Phleo	Phleomycin
PI	Propidium iodide
PIC	Preinitiation Complex
Pih1	Protein interacting with Hsp (heat-shock protein) 90 1
po. un.	Polar uncharged
PTM	Post-translational modification
PVDF	Polyvinylidene fluoride
p.i	Post induction
R2TP	Rvb1-Rvb2-Tah1 (TPR (tetratricopeptide repeat)-containing protein associated with Hsp (heat-shock protein) 90 1)-Pih1 (Protein interacting with Hsp (heat-shock protein) 90 1)
RCSB PDB	Research Collaboratory for Structural Bioinformatics Protein Data Bank
rev	Reverse
RING	Really Interesting New Gene

## Abbreviations

RNA	Ribonucleic acid
RNAi	RNA interference
RNAP (II)	DNA-dependent RNA Polymerase (II)
RNase A	Ribonuclease A
RPB1	RNA Polymerase II Subunit 1
rRNA	Ribosomal RNA
RT	Room temperature
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>
SAGA	Spt-Ada-Gcn5 acetyltransferase
SAM	S-adenosyl methionine
SANT	Swi3 (switching-defective protein 3), Ada2 (Adaptor 2), N-Cor (nuclear receptor co-repressor), and TFIIIB (Transcription factor II B)
SDM	Semi-defined medium
SDS	Sodium dodecyl sulfate
sec	Second(s)
ser	Serine
SET	Su(var)3-9, Enhancer-of-zeste and trithorax
SETD7	SET Domain Containing 7
SF2	Superfamily 2
SHPRH	SNF2, histone linker, PHD, RING, helicase
SILAC	Stable isotope labeling by amino acids in cell culture
SL	Splice leader
SM	Single marker
SnAC	SNF2-ATP-coupling
SNF(2)	Sucrose Non-Fermentable (2)
snoRNP	Small nucleolar ribonucleoproteins
SOC	Super optimal broth with catabolite repression
<i>spec.</i>	Species
SRCAP	Snf2 Related CREBBP (CREB (cAMP-response element binding) binding Protein) Activator Protein
ssDNA	Single-strand DNA
SSR(s)	Strand-switch region(s)
SU(VAR)	Suppressor of variegation
Suv39H1	Suppressor of Variegation 3-9 Homolog 1
SWC2	<i>SWR1</i> -complex <i>protein</i> (2)
SWI	Switch
SWR1	SWI2/SNF2-Related 1
<i>T. brucei</i>	<i>Trypanosoma brucei</i>
<i>T. cruzi</i>	<i>Trypanosoma cruzi</i>
Taf(14)	TBP-associated factor (14)

## Abbreviations

Tah1	TPR (tetratricopeptide repeat)-containing protein associated with Hsp (heat-shock protein) 90
TATA	TATA-nucleotide sequence
TBL	Turbo-biotinligase
<i>Tb</i> ARP(x)	<i>Trypanosoma brucei</i> actin-related protein (x)
<i>Tb</i> Eaf(x)	<i>Trypanosoma brucei</i> ESA1-associated factor (x)
TBP	<i>TATA-binding protein</i>
<i>Tb</i> SWR1	<i>Trypanosoma brucei</i> SWI2/SNF2-Related 1
<i>Tb</i> SWRC(x)	<i>Trypanosoma brucei</i> SWR1-complex protein (x)
tD	Population doubling time
TDB	Trypanosome dilution buffer
TE	Tris/EDTA
TEMED	Tetramethylethylenediamine
TETR	Tetracycline responsive element
TF(s)	Transcription factor(s)
TFII(x)	Transcription factor II (x)
thr	Threonine
TIP60	Tat interactive protein 60 kDa
Tm	Annealing temperature
TSS	Transcription start site
TTS	Transcription termination site
TUB	Tubulin
T7RNAP	T7 RNA polymerase
UTR	Untranslated region
UV	Ultraviolet
V	Volt
Vps(71)	Vacuolar protein sorting-associated protein (71)
VSG	Variant Surface Glycoprotein
v/v	Volume/volume
WB	Western blot
WD40	WD stands for the amino acids tryptophan and aspartic acid that build the end of each WD repeat; the repeat normally consists of 40 amino acids
WHO	World Health Organisation
WT	Wild type
wt/vol	Weight/volume
Y	Tyrosine
Yaf9	Yeast AF9 (ALL1 (Leukemia, Acute Lymphocytic, Susceptibility To, 1) fused gene from chromosome 9)
YEATS	Yaf9-ENL (Eleven-nineteen-leukemia) -AF9 -Taf14 -Sas5 (Something About Silencing 5)
YTA7	Yeast Tat-binding homolog 7

## Abbreviations

ZNHIT1	Zinc Finger HIT-Type Containing 1
$\gamma$ H2A(X)	Phosphorylated H2A(X) that occurs in course of DNA damage

## 1. Introduction

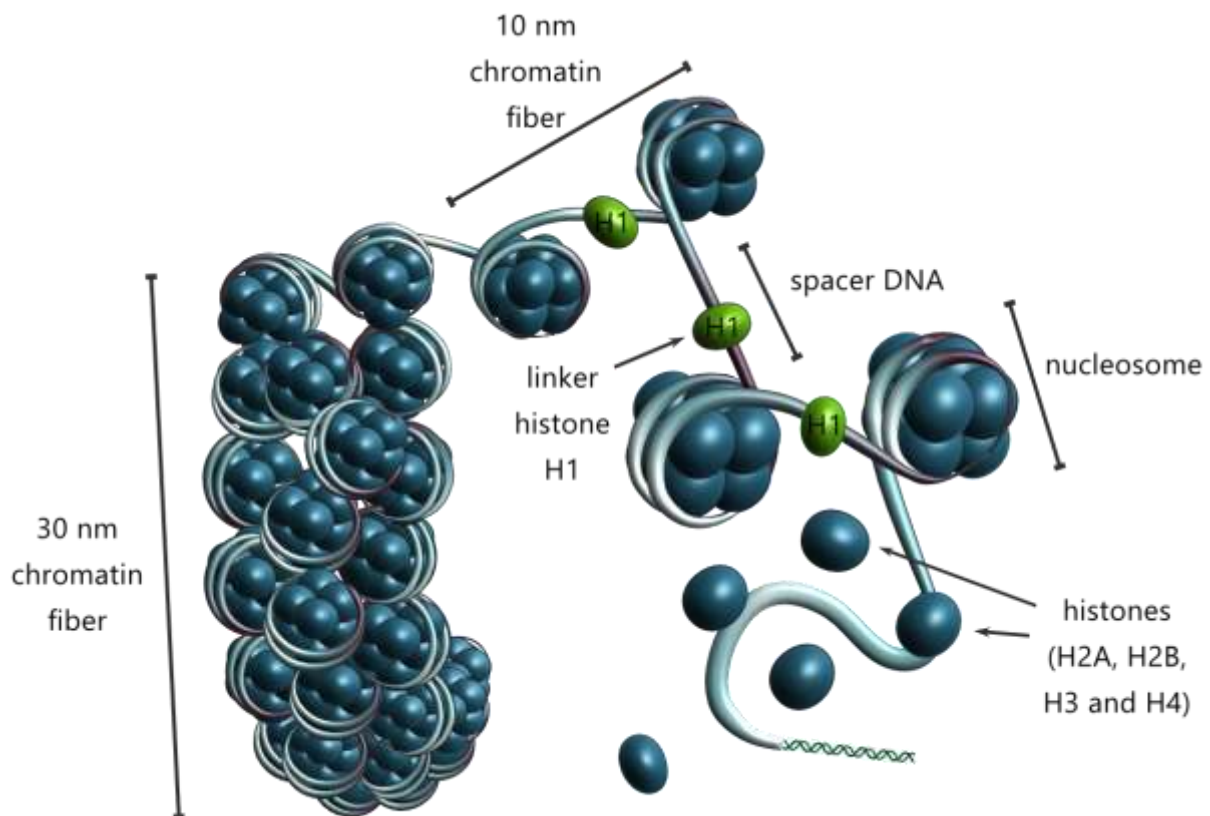
The following chapters will give an overview about nucleosomes, their composition and modifications as well as chromatin remodelling complexes to understand their interdependency in the course of the incorporation pathway of the histone variant H2A.Z. This will build the basis to describe how histone H2A.Z regulates important cellular processes with a special focus on transcription regulation. The last part of this introduction will summarise the present knowledge about histone H2A.Z related PTMs and chromatin remodelling in the unicellular parasite *Trypanosoma brucei*.

### 1.1 Chromatin and its structure

Genomic information is encoded in the sequence of 'Deoxyribonucleic acid' (DNA) strands. In eukaryotic organisms, nuclear DNA is wrapped around nucleosomes. The entirety of these repeating nucleoprotein complexes is called chromatin (Hewish & Burgoyne, 1973; Kornberg & Thomas, 1974). In general chromatin can occupy two major conformations. A dense conformation called heterochromatin and a loose one called euchromatin. Regulation of the chromatin structure is required for regulation of cellular processes like DNA repair, replication or transcription (Campos & Reinberg, 2009). While duplication of the DNA requires access to the whole genome to ensure the transfer of all genetic information into new cells, DNA repair merely requires access at sites of occurring DNA damage. DNA accessibility for transcription into RNA which is essential for protein production must be highly flexible (Crick, 1970; Owen-Hughes & Workman, 1994). While some genes have to be constantly transcribed to maintain general cellular processes, other genes are only activated in response to intra- or extracellular signals (Hill & Treisman, 1995; Brivanlou & Darnell, 2002). Actively-transcribed genes are mostly located in open 'euchromatic' areas of the genome. Transcriptional inactive regions (like centromeres for example) were found to have a more condensed, 'heterochromatic' structure (Weisbrod, 1982). Therefore, heterochromatin formation can be used to systematically silence genes or whole gene arrays (reviewed in Weisbrod, 1982; Wutz, 2011). Replication, DNA repair and transcription require a well-regulated assembly and dis-assembly of nucleosomes with plenty of factors identified until today, that affect nucleosome dynamics and therefore chromatin (Owen-Hughes & Workman, 1994; MacAlpine & Almouzni, 2013; Stadler & Richly, 2017). Post-translational modifications (PTMs) of nucleosomes were identified as being as relevant for the chromatin structure as for the composition of the nucleosome (Katan-Khaykovich & Struhl, 2005; Altaf *et al.*, 2010; Migliori *et al.*, 2012). Protein complexes, so called chromatin remodelling complexes (referred to as CRCs in the following), are required for nucleosome assembly, restructuring and dis-assembly (reviewed in Clapier & Cairns, 2009). There are several examples that describe the interdependency of chemical modifications, nucleosome components and CRCs (Altaf *et al.*, 2010), reviewed in (Hoffmann & Spengler, 2019; Clapier, 2021). In the following chapters I will take a closer look at PTMs, nucleosomes and their composition as well as CRCs to comprehend their interdependency.

## 1.2 The nucleosome structure

A nucleosome is composed of two copies of each core histone H2A, H2B, H3 and H4. 147 bp of DNA are wrapped 1.65 times around a single nucleosome (Luger *et al.*, 1997) (Fig. 1.1). The histone H1 acts as a linker and binds 20 bp of the DNA that leaves the entry and exit site to complete the second wrap. Together nucleosome and histone H1 build the chromatin. Multiple chromatin fibers build the so-called 10 nm fibre (Fig 1) which is also known as "beads-on-a-string-array" (Widom, 1998). Although *in vitro* analysis could show that the 10-nm fibre can build a higher order structure, the 30-nm fibre (Fig. 1) multiple investigations revealed that this structure could not be observed *in vivo* (Tremethick, 2007 Joti *et al.*, 2012; Chen *et al.*, 2016).



**Figure 1: Schematic depiction of the chromatin structure.**

Modified from Figueiredo *et al.*, 2009

## 1.3 Nucleosome modifications and chromatin structure

PTMs of proteins were shown to play important roles in cellular processes. Ubiquitination of proteins for instance was shown to be crucial for protein degradation while phosphorylations play a central role in almost every signal cascade (Graves & Krebs, 1999; Kornitzer & Ciechanover, 2000). Therefore, it is not surprising that PTMs are equally important in regulating nucleosome and chromatin structure (reviewed in Bannister & Kouzarides, 2011). The following paragraph focusses on two histone PTMs, their role in chromatin structure regulation and how they are linked with H2A.Z



## 1. Introduction

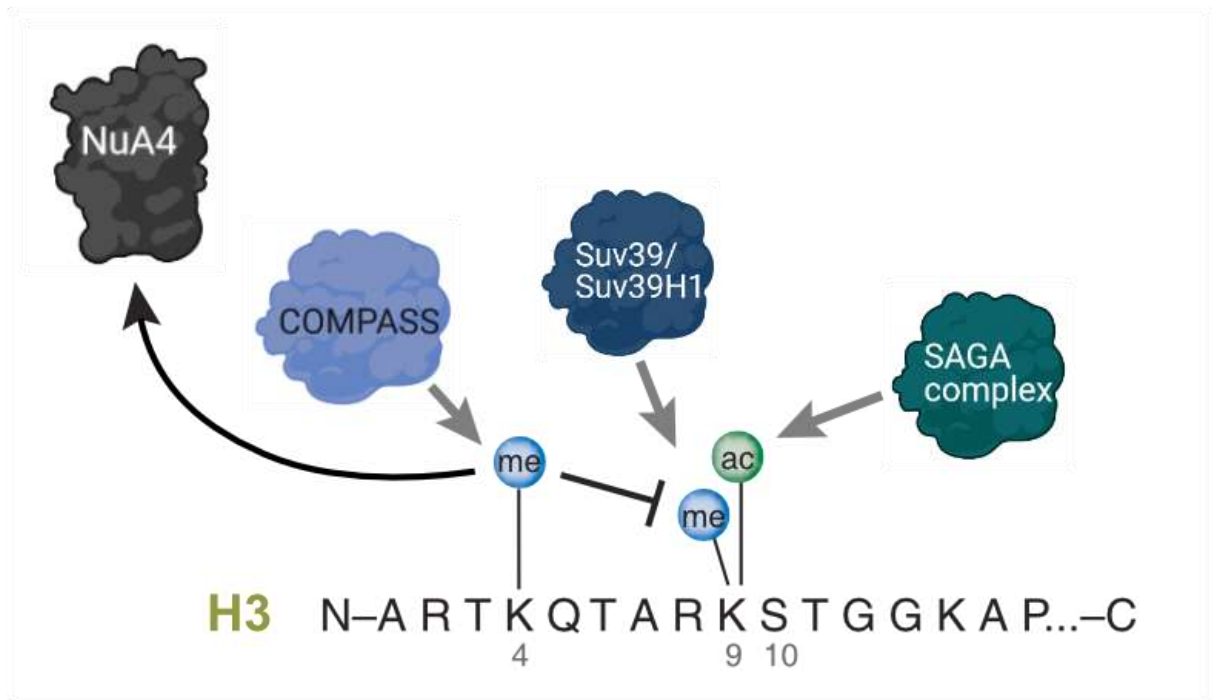
PTMs can be divided into those that alter the chromatin/nucleosome structure in an active manner and histone PTMs that passively regulate the chromatin/nucleosome structure by recruiting proteins that assist a remodelling step.

Active PTMs like the acetylation of lysine 56 on histone H3 (H3K56) alter the structure of a histone, nucleosome or chromatin due to their chemical properties. *In vitro* assays showed that H3K56 acetylation prevents oligomerization of nucleosomes (Watanabe *et al.*, 2010) and could be associated with transcription regulation and DNA damage repair (Masumoto *et al.*, 2005; Xu *et al.*, 2005). Passive PTMs require interactions with other proteins to affect the nucleosome/chromatin structure. Such proteins, often as part of CRCs, bind a PTM with a specific binding motif. With an increasing importance of high throughput methods, microarrays as well as mass spectrometry-coupled 'stable isotope labelling by/with amino acids in cell culture' (SILAC) assays were performed to identify different readers/effectors and their domains. Histone acetylations were shown to be bound by bromo and tandem 'plant homeodomain' (PHD) domains (Dhalluin *et al.*, 1999; Lange *et al.*, 2008). Nine different domains could be linked with binding of methylated histones (reviewed in Yun *et al.*, 2011). Some of the best described domains that bind histone methylations are chromo, PHD, Tudor and WD40 (WD stands for the amino acids tryptophan and aspartic acid that build the end of each WD repeat) domains (Bannister *et al.*, 2001; Wysocka *et al.*, 2005; Sanchez & Zhou, 2011; Lu & Wang, 2013). One of the earliest studied passive PTMs is the tri-methylation of histone H3 on lysine 9 (H3K9) which is involved in heterochromatin formation (Nakayama *et al.*, 2001). 'Heterochromatin protein 1' (HP1) binds to H3K9 tri-methylation with its chromodomain, thus promoting chromatin condensation (Bannister *et al.*, 2001). In contrast, acetylation of H3K9 was associated with active gene transcription (Pokholok *et al.*, 2005; Karmodiya *et al.*, 2012). 'General control non-de-repressible 5' (Gcn5) as part of the 'Spt-Ada-Gcn5 acetyltransferase' (SAGA) complex (Grant *et al.*, 1997) acetylates H3K9 (Jin *et al.*, 2011) and requires previous methylation of lysine 4 on histone H3 (H3K4; Bian *et al.*, 2011).

Cross-regulation of PTMs on H3K9 in which one chemical modification prevents addition of a different one is called *in situ* cross regulation (Fig 2; reviewed in Latham & Dent, 2007). H3K9 methylation does not only antagonize H3K9 acetylation, it might also be associated with removal of H3K9 acetylation. The H3K9 methyltransferase 'Suppressor of variegation 3-9' (SU(VAR)3-9; *Drosophila*) and 'Suppressor of Variegation 3-9 Homolog 1' (Suv39H1; human), which are both responsible for di-methylation of H3K9 10949293, interact with the 'histone deacetylases' (HDAC) HDAC1 in *Drosophila* and HDAC1/2/3 in human cells (Czermin *et al.*, 2001; Vaute *et al.*, 2002) therefore creating a 'feedback-loop'. Besides acetylation of H3K9, methylation of H3K4 which is processed by the 'Complex proteins associated with SET1' (COMPASS) complex in *Saccharomyces cerevisiae* (Miller *et al.*, 2001) is one of the most known histone marks linked with active transcription in yeast and metazoans (Santos-Rosa *et al.*, 2002; Schübeler *et al.*, 2004). Metazoans contain various non-redundant and complex-associated methyltransferases (SET1A, SET1B, MLL1-4, SET7/9) which share H3K4 as their substrate (reviewed in Eissenberg & Shilatifard, 2009; Shilatifard, 2012).

## 1. Introduction

Aside from its role in SAGA recruitment H3K4 methylation by 'Su(var)3-9, Enhancer-of-zeste and Trithorax' (SET) 7/9 also inhibits H3K9 methylation by Suv39H1 in human (Nishioka *et al.*, 2002). Methylation of H3K4 is associated with activity of 'nucleosome acetyltransferase of H4' (NuA4; Ginsburg *et al.*, 2009; Ginsburg *et al.*, 2014), which is acetylating histones H4 and H2A prior to H2A.Z incorporation by the SWR1 CRC (Altaf *et al.*, 2010; A detailed description of the interdependency of NuA4, SAGA and COMPASS will follow in chapter 1.5.1).

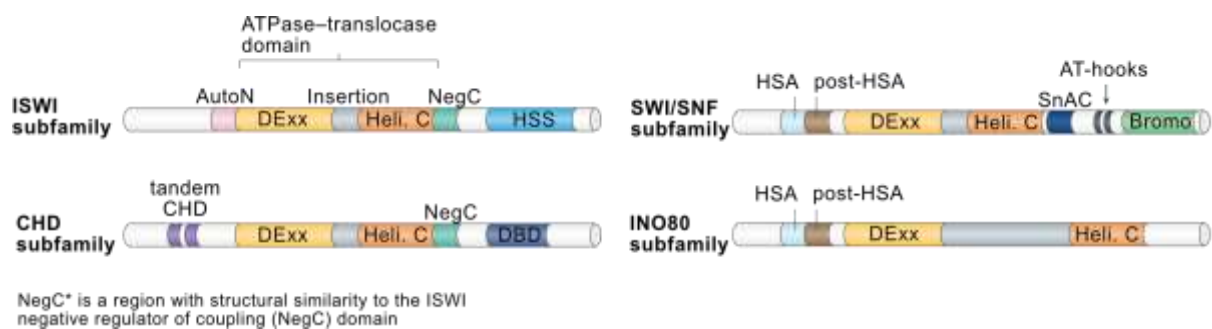


**Figure 2: Example of cross interactions of histone PTMs in eukaryotes.**

Black arrows indicate an enhancing function, black lines indicate an inhibition, grey arrows indicate the generation of a PTM by a complex; modified from Latham & Dent, 2007

## 1.4 Chromatin remodelling complexes

Histone chaperones like 'anti-silencing factor 1' (ASF1), 'histone regulator A' (HIRA) or 'chromatin assembly factor 1' (CAF1), could be associated with transport and incorporation of histones in a replication-dependent or independent manner (reviewed in Gurard-Levin *et al.*, 2014). In addition to these chaperones eukaryotic cells possess CRCs assembled around DNA-dependent ATPases of the 'sucrose non-fermentable 2' (SNF2) family. These complexes play a crucial role in ATP-dependent chromatin remodelling (reviewed in Narlikar *et al.*, 2013). The SNF2 family is part of the DNA/RNA superfamily 2 helicases (SF2; Flaus *et al.*, 2006). The SNF2 family can be divided into four subfamilies: 1.) 'imitation switch' (ISWI), 2.) 'chromodomain helicase DNA-binding' (CHD), 3.) 'switch/sucrose non-fermentable' (SWI/SNF) and 4.) 'inositol non-fermenting 80' (INO80, Fig 3; reviewed in Clapier & Cairns, 2009).



**Figure 3: The four subclasses of SNF2 ATPases**

Domain organization of the four SNF2 ATPase subfamilies. The ATPase-translocase domain of all remodelers contains the DExx- motif, an insertion of variable length and the helicase C domain (Heli. C.); 'helicase/SANT-associated' (HSA), 'SNF2-ATP-coupling' (SnAC), DNA-binding domain (DBD), auto-inhibitory N-terminal domain (AutoN); modified from Clapier *et al.*, 2017

Phylogenetic analysis enabled the classification of these four subfamilies into several smaller subgroups with conserved structural motifs that characterize members of each subgroup (Flaus *et al.*, 2006). Studies in several model organisms revealed an evolutionary conservation of each subfamily but individual complex compositions can vary between organisms (Table 1; reviewed in Clapier & Cairns, 2009). Nevertheless, SNF2 CRCs have several common characteristics: 1.) An ATPase domain that possesses a DExx-Box (DEAD-Box) and a helicase C domain 2.) proteins with ATPase-regulating characteristics, 3.) affinity to the nucleosome itself, often associated with additional DNA binding domains, 4.) a domain that facilitates recognition of histone PTMs and 5.) proteins with protein-protein interaction interfaces to bind other factors (reviewed in Clapier & Cairns, 2009).

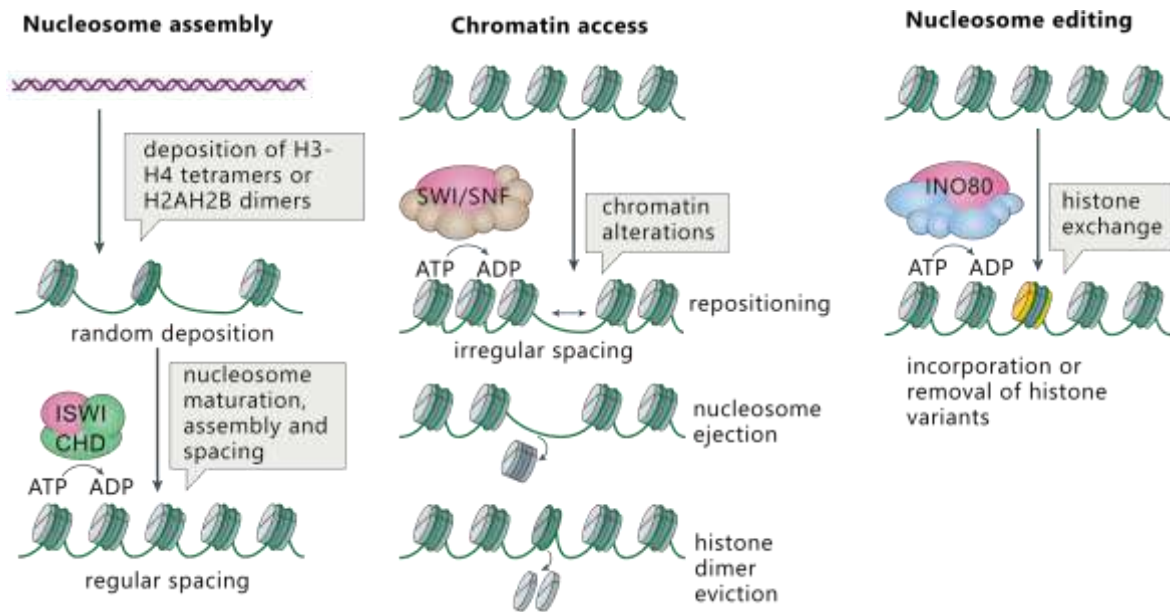
**Table 1: Overview of SNF2 complexes of the INO80 subfamily in different organisms**

Family and composition		organism				
		Yeast		Human		
INO80	<b>Complex</b>	<b>INO80</b>	<b>SWR1</b>	<b>INO80</b>	<b>SRCAP</b>	<b>TIP60</b>
	ATPase	Ino80	Swr1	hIno80	SRCAP	p400
	Non-catalytic homologous subunits	Rvb1, 2		RUVBL1, 2 / Tip49a, b		
		Arp5, 8	Arp6	BAF53a		
		Arp4, Actin 1		Arp5, 8	Arp6	Actin
		Taf14	Yaf9		GAS41	
		les2, 6		hles2, 6		
			SWC4/Eaf2		DMAP1	
			SWC2/Vps72		YL-1	
			Bdf1		Brd8/TRC/p120	
			H2A.Z, H2B		H2A.Z, H2B	
			SWC6/Vps71		Znf-HIT1	
						TRRAP
						Tip60
						MRG15 MRGX
				FLJ11730		
				MRGBP		
				EPC1, EPC-like		
				ING3		
Unique	les1, les3-5, Nhp10	SWC3, 5, 7	*			

\*Amida, NFRKB, MCRS1, UCH37, FLJ90652, FLJ20309; modified from Clapier & Cairns, 2009

The list displays all components of a certain complex and its orthologs in other organisms

The ATPase domain of SNF2 proteins dissolves histone-DNA contacts prior to remodelling. Both ATPase sub-domains are required to establish an ATP binding–hydrolysis-dependent conformational cycle of ‘recombinase A’ (RecA)-like lobes that drives DNA translocation (reviewed in Clapier *et al.*, 2017). In course of this translocation, nucleosomes can be 1.) slid along the DNA, 2.) can be removed from the DNA or can 3.) be modified (Fig 4; reviewed in Clapier *et al.*, 2017). Experiments performed in various eukaryotic model organisms showed that the four SNF2 subfamilies contribute to highly specified remodelling processes. In general, all SNF2 chromatin remodelers play a key role in shaping the nucleosomal landscape contributing to cellular processes such as DNA replication, DNA repair and especially in transcription (reviewed in Clapier *et al.*, 2017).



**Figure 4: The role of each of the four SNF2 ATPase subfamilies in chromatin remodelling**

A brief description of the function of each of the four SNF2 ATPase subfamilies; figure modified from Clapier *et al.*, 2017

#### 1.4.1 Chromatin remodelling complexes of the Inositol non-fermenting 80 subfamily

The two ATPases INO80 and 'SWI2/SNF2-Related 1' (SWR1) belong to the INO80 subfamily. Members of this subfamily can be characterized by two specific features. 1.) the DExx domain possesses a DEXQ motif (Flaus *et al.*, 2006), 2.) the DExx domain and the helicase c domain that are separated by an insertion, required for association with a hetero-multimeric RuvB (the protein has been named by the *ruvB* locus in *E. coli*, mutations in this locus were associated with a higher susceptibility against UV-irradiation; Otsuji *et al.*, 1974) ring that is characteristic for the INO80 subfamily (Fig 4; reviewed in Gerhold & Gasser, 2014). INO80 and SWR1 are involved in the maintenance of genome integrity, transcription regulation, replication and DNA damage repair (reviewed in Poli *et al.*, 2017, Morrison & Shen, 2009 Aslam *et al.*, 2019). First described in *S. cerevisiae*, SWR1 and its complex were associated with the exchange of canonical H2A with its variant H2A.Z (Krogan *et al.*, 2003 Mizuguchi *et al.*, 2004). A stimulation by free H2A.Z-H2B dimers and H2A-containing nucleosomes is required to facilitate dimer exchange (Luk *et al.*, 2010). This mechanism and a rather low ATPase activity in comparison to other ATPases prevents removal of nucleosome-associated H2A.Z-H2B dimers.

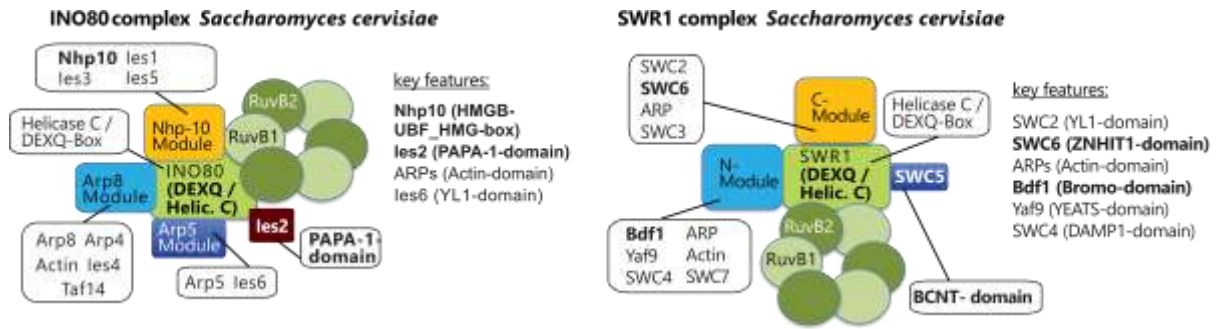
However, removal of H2A.Z-H2B dimers by SWR1 has been reported in response to acetylation of lysine 56 on histone H3 (Watanabe *et al.*, 2013). Like SWR1, INO80 was first described in *S. cerevisiae* where the ATPase and the eponymous complex are involved in removal of H2A.Z and  $\gamma$ H2A (Ebbert *et al.*, 1999; Papamichos-Chronakis *et al.*, 2006; Alatwi & Downs, 2015; Brahma *et al.*, 2017). INO80 is also participating in nucleosome sliding in an ISWI-related manner (Udugama *et al.*, 2011). The INO80 complex mainly consists of four different modules (Fig 1.5; Tosi *et al.*, 2013).

## 1. Introduction

The “foot” is built by an Arp8-module which mainly consists of actin, the two ‘actin-related proteins’ (ARPs) Arp8, Arp4 and the ‘Yaf9, ENL, AF9, Taf14, and Sas5’ (YEATS) domain protein ‘TBP-associated factor 14’ (TAF14) which may contribute to binding of acetylated histones (Schulze *et al.*, 2010; Tosi *et al.*, 2013). The major task of the Arp8 module is to increase nucleosome affinity (reviewed in Gerhold & Gasser, 2014). In yeast interaction of INO80 with  $\gamma$ H2A is mediated by Arp4 (Downs *et al.*, 2004), in mammalian INO80 Arp8 mediates interactions with  $\gamma$ H2AX (Kashiwaba *et al.*, 2010; Gerhold *et al.*, 2012). The ‘Non-histone protein 10’ (Nhp10) module, forming the body of the INO80 complex consists of Nhp10 and the three auxiliary proteins ‘Ino eighty subunit 1’ (les1), les3 and les5 (Tosi *et al.*, 2013). It also provides stable interaction between  $\gamma$ H2A and the INO80 complex (Morrison *et al.*, 2004) while building a platform for nucleosome binding (Tosi *et al.*, 2013). Nucleosome remodelling is mediated by an Arp5 module (Tosi *et al.*, 2013) that consists of the eponymous Arp and les6 which harbours a YL1 domain that is associated with H2A.Z binding (Wu *et al.*, 2005). A RuvB hetero-dodecamer consisting of RuvB1 and RuvB2 is bundling the head of the INO80 complex and provides 5’ to 3’ helicase activity (Gribun *et al.*, 2008).

Compared to the INO80 complex, the SWR1 complex consists of three different modules assembled around the SNF2 ATPase (Fig 5; Nguyen *et al.*, 2013). While INO80 possesses a dodecamer RuvB1/RuvB2 ring (Tosi *et al.*, 2013) the SWR1 complex possesses a hexameric ring (Nguyen *et al.*, 2013). The N- and C-module are associated with the corresponding termini of the ATPase (Wu *et al.*, 2009; Nguyen *et al.*, 2013). The N-module, which binds the ‘helicase/SANT (Swi3, Ada2, N-Cor, and TFIIIB)-associated’ (HSA) domain, is required for nucleosome binding and recognition of acetylated histones. Arp4 was associated with nucleosome binding (Galarneau *et al.*, 2000; Kashiwaba *et al.*, 2010) while ‘bromodomain factor 1’ (Bdf1) and the YEATS domain containing ‘yeast af9’ (Yaf9) were associated with binding of acetylated histone tails (Durant & Pugh, 2007; Schulze *et al.*, 2010; Li *et al.*, 2014). In addition to Arp4, Bdf1 and Yaf9 the N-module is composed of actin, the SANT domain protein ‘SWR1 complex protein 4’ (SWC4) and the auxiliary protein SWC7 (Nguyen *et al.*, 2013). The C-module is composed of Arp6 and the three SWC proteins 2, 3 and 6 (Nguyen *et al.*, 2013). SWC2 has a YL1 domain and acts as a molecular lock for H2A.Z to prevent eviction of the variant (Wu *et al.*, 2005). SWC6 (formerly known as ‘Vacuolar protein sorting-associated protein 71’ (Vps71)) harbours a ‘zinc-finger histidine triad 1’ (Znf-HIT1) domain and is essential for H2A.Z deposition (Krogan *et al.*, 2003), which might be dependent on its association with SWC2 (Wu *et al.*, 2005). The C-module mediates H2A.Z deposition but is also required for nucleosome binding. Deletion of Arp6 reduces nucleosome affinity of the SWR1 complex around 10-fold (Wu *et al.*, 2005).

## 1. Introduction



**Figure 5: A comparison between the SWR1 and INO80 complex.**

**INO80 complex *S. cerevisiae*:** The Nhp-10 module forms a platform for nucleosome interaction. The Arp8 module is the nucleosome binding module, while the Arp5 module, which contains the YL1-domain protein les6 is responsible for the nucleosome remodelling step (Tosi *et al.*, 2013). les2 with the PAPA-1-domain plays a structural role within the complex (Chen *et al.*, 2013; Tosi *et al.*, 2013). INO80 specific domains are highlighted in bolt. **SWR1 complex *S. cerevisiae*:** Proteins listed under key features are essential for H2AZ incorporation (Gerhold & Gasser, 2014). The C-module of the SWR1 complex, which contains the Bromo-, YEATS- and SANT/DAMP1-domain mediates nucleosome affinity, while the N-module with the YL1- and Zinc finger ZNHIT1-domain is involved in the histone variant exchange reaction (Wu *et al.*, 2005; Gerhold & Gasser, 2014). SWR1 specific domains are highlighted in bolt letters (Gerhold & Gasser, 2014; Willhoft & Wigley, 2020).

Both complexes share a high degree of similarity (Fig 5; RuvB helicases, a YEATS domain containing protein, actin as well as multiple actin related proteins) but a few unique and non-conserved features separate the two complexes from each other. A Znf-HIT1 domain-containing protein, a bromodomain-containing protein or a BCNT domain-containing factor appear to be exclusively present in SWR1-like complexes (reviewed in Willhoft & Wigley, 2020). A 'Pim-1-associated protein-1 (PAP-1)-associated protein-1' (papa-1) domain-containing factors (INO80B in humans, Les2 in *S. cerevisiae*) in turn has only been described in INO80-like complexes (reviewed in Willhoft & Wigley, 2020). While yeast cells have two distinct complexes for H2A/H4 acetylation and H2A.Z incorporation (NuA4 and SWR1), higher eukaryotes possess the fusion complex 'TAT-interactive protein, 60kDa' (TIP60)/p400 (dTIP60/Domino in *Drosophila*) that combines both functions (Kusch *et al.*, 2004; Gévry *et al.*, 2007). In addition to the fusion complex humans possess a SWR1-like SRCAP complex, which is also capable of incorporating H2A.Z (Ruhl *et al.*, 2006; Wong *et al.*, 2007). Why human cells possess two complexes and how they are regulated remains elusive. Homologs of the NuA4 and SWR1 complex were identified in all eukaryotic model organisms making the cascade of histone H2A/H4 acetylation by an acetyltransferase complex, followed by H2A.Z incorporation via SWR1 a highly conserved process in eukaryotes.

## 1.5 The histone variant H2A.Z

Variants of the four canonical histones play an essential role in the regulation or maintenance of the chromatin structure (reviewed in Ausió, 2006). Canonical histones are exclusively expressed and incorporated in the course of the S-Phase. In contrast, histone variants are expressed and incorporated throughout the cell cycle (reviewed in Martire & Banaszynski, 2020). Evolutionary conserved histone variants can be identified in every eukaryote, but histone sequences highly differ between organisms (Molaro & Drinnenberg, 2018). Thus, is the sequence of the canonical H2A and H3 in *S. cerevisiae* more similar to mammalian H2A.X and H3.3 than to their mammalian canonical counterparts (Malik & Henikoff, 2003).

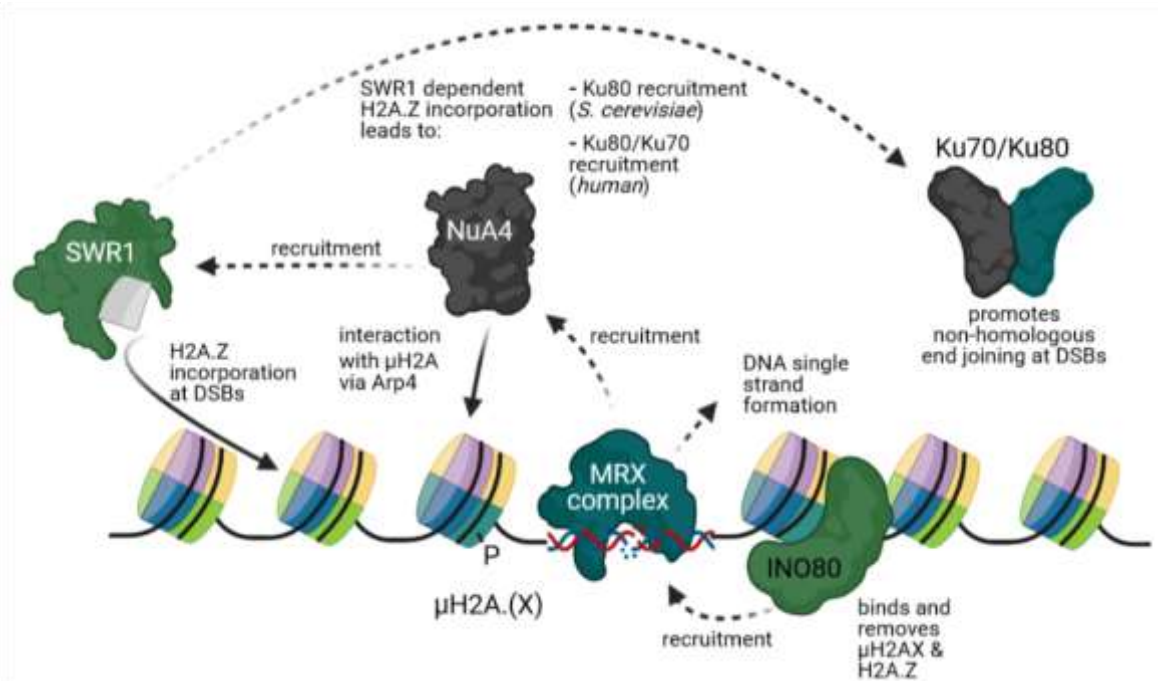
The histone H2A exhibits by far the biggest group of histone variants (reviewed in Bönisch & Hake, 2012; Jiang *et al.*, 2020) and since its identification in 1980 in mice (West & Bonner, 1980) the histone variant H2A.Z (Htz1 in yeast) became the most investigated histone variant. Only a few organisms lack this highly conserved variant (e.g. *Giardia*, *Trichomonas* and *Entamoeba* spp. Dalmasso *et al.*, 2011). H2A.Z is required for heterochromatin and centromere formation (Greaves *et al.*, 2007; Ryan & Tremethick, 2018) and anti-silencing at heterochromatin boundaries (Babiarz *et al.*, 2006; Zhou *et al.*, 2010), but it also forms the boundaries of 'nucleosome-depleted regions' (NDRs; Raisner *et al.*, 2005). Its role in nucleosome stability and transcription regulation can be contrary. It was linked with an increase (Park *et al.*, 2004; Thambirajah *et al.*, 2006) as well as a decreases of nucleosome stability (Flaus *et al.*, 2004; Jin & Felsenfeld, 2007) and affects transcription in a positive and negative manner (reviewed in Giaimo *et al.*, 2019; a detailed description of how H2A.Z affects transcription will follow in chapter 1.5.1). In *S. cerevisiae* the incorporation of H2A.Z requires acetylation of the histones H2A and H4 (Shia *et al.*, 2006; Altaf *et al.*, 2010). Acetylation of lysines 5, 8 and 12 of histone H4 and lysine 4, 7, and 13 of histone H2A appear to have accumulative but non-redundant function (Altaf *et al.*, 2010). Acetylation of lysine 16 on H4 by SAS2 appears to be specific for H2A.Z at sub-telomeric regions (Shia *et al.*, 2006). While complete loss of acetylations on H2A and H4 is lethal, loss of individual acetylations on one of the histones only led to a reduced incorporation of H2A.Z (Altaf *et al.*, 2010). Experiments performed with human cells and *Drosophila* showed some evidence that the NuA4 homologs TIP60 and dTIP60 (Doyon *et al.*, 2004) also acetylate H2A.Z in addition to H2A and H4 (Kimura & Horikoshi, 1998; Kusch *et al.*, 2004; Giaimo *et al.*, 2018; Lee *et al.*, 2020).

Aside from its role in transcription regulation H2A.Z also plays an essential role in DNA damage response. Damage of genetic information represents a threat to every organism and requires an efficient signalling and repair mechanism that involves histones H2A.X/H2A and multiple protein complexes (Fig 6; reviewed in Dinant *et al.*, 2008). Mammalian H2A.X that flanks the site of damaged DNA is reversibly phosphorylated at serine 139 (Rogakou *et al.*, 1998). Histone phosphorylation subsequent to occurring DNA damage is consistent among all eukaryotes, but the histone being phosphorylated differs.



## 1. Introduction

In *Drosophila melanogaster* the histone variant H2A.Z is phosphorylated on serine 137 (Madigan *et al.*, 2002) while in *S. cerevisiae* as well as *Schizosaccharomyces pombe* serine 128/129 on canonical H2A is phosphorylated in order to signal DNA damage (Downs *et al.*, 2000; Nakamura *et al.*, 2004). In *S. cerevisiae*, SWR1-dependent incorporation of H2A.Z is required for loading of Ku80 and for localisation of DSBs to the nuclear periphery in G1- and S-Phase cells (van Attikum *et al.*, 2007; Horigome *et al.*, 2014). The 'Mre11-Rad50-Xrs2' (MRX; MRN in mammals) complex, which facilitates single-strand DNA (ssDNA) formation at DSBs (Tsukuda *et al.*, 2005; Cannavo & Cejka, 2014) recruits NuA4 to DNA DSBs. NuA4 in turn recruits the SWR1 complex (Altaf *et al.*, 2010). Interaction of the NuA4 complex with  $\gamma$ H2A is interfered by Arp4 (Downs *et al.*, 2004). H2A.Z as well as  $\gamma$ H2A are removed from the nucleosome in course of the repair of DNA DSB by 'Acidic Nuclear Phosphoprotein 32 Family Member E' (Anp32e) and INO80 (Morrison *et al.*, 2004; Papamichos-Chronakis *et al.*, 2006; Alatwi & Downs, 2015; Gursoy-Yuzugullu *et al.*, 2015). Data from experiments in human cells showed that the incorporation and removal of H2A.Z at DSBs is tightly controlled and timed. The amount of H2A.Z only increases in the first eight hours after induction of a DSB, before it reaches a plateau (Xu *et al.*, 2012) indicating that H2A.Z is only required in the initial DNA DSB response.

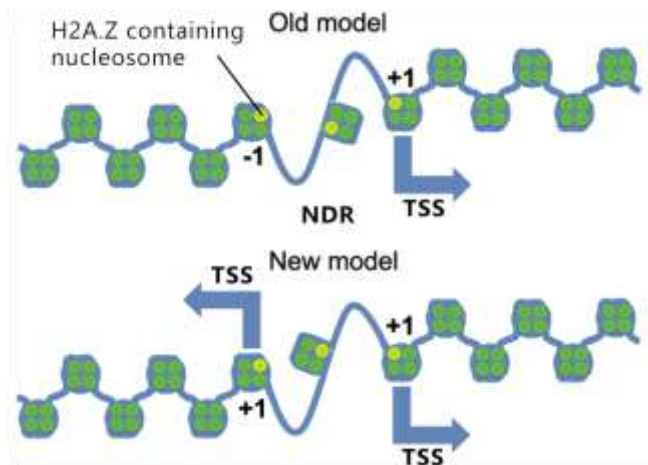


**Figure 6: Overview of the interplay between H2A.Z and its remodelling complexes in DNA DSB.**

Dashed arrows indicate recruitment or a process while sustained arrows represent a direct interaction; the figure has been designed with BioRender©.

### 1.5.1 The role of H2A.Z in transcription

The histone variant H2A.Z is a key player in chromatin organisation and transcription regulation. First hints for a link between H2A.Z and chromatin organisation/transcription were found in *S. cerevisiae*. A well-positioned H2A.Z containing nucleosome (+1 and -1 nucleosome) forms the borders of the NDR of transcription start sites (TSS; Guillemette *et al.*, 2005; Albert *et al.*, 2007; Bagchi *et al.*, 2020). The size of nucleosome-depleted TSSs can range from 100 base pairs (bp) up to 10 kilo base pairs (kbp; Carninci *et al.*, 2006; Siegel *et al.*, 2009). Recent data derived from experiments in *S. cerevisiae* could disprove the assumption of the existence of +1 and -1 nucleosomes. An RNA-seq approach revealed that the -1 nucleosome is in fact a +1 nucleosome that marks a TSS where bi-directional transcription occurs (Fig 7; Bagchi *et al.*, 2020).



**Figure 7: Depiction of a eukaryotic transcription start site**  
+1 defines the first nucleosome that follows upstream of the TSS. -1 defines the first nucleosome that follows downstream of the TSS. Modified from Bagchi *et al.*, 2020

This observation is in accordance with data from *S. cerevisiae*, *Drosophila* and humans where H2A.Z was shown to only form the +1 nucleosome of TSSs (Mavrich *et al.*, 2008; Schones *et al.*, 2008). NDRs of human TSSs that were originally assumed to be nucleosome free, harbour highly instable H2A.Z/H3.3 containing nucleosomes (Jin *et al.*, 2009). In contrast to TSSs and enhancers, H2A.Z is actively excluded from coding regions (Farris *et al.*, 2005) by the RNAP II-associated remodeler 'facilitates chromatin transcription' (FACT) and the histone chaperone Spt6 (Jeronimo *et al.*, 2015). A loss of the variant leads to a reduction of 'RNA-Polymerase II' (RNAP II) recruitment at TSSs in *S. cerevisiae* and human cells (Adam *et al.*, 2001; Hardy *et al.*, 2009) and reduced 'TATA-binding protein' (TBP) occupancy (Wan *et al.*, 2009). The interaction between the SWR1 complex and acetylated histone tails is mediated by Bdf1 (Matangkasombut & Buratowski, 2003; Durant & Pugh, 2007). Despite of a conserved localisation of H2A.Z, its role in transcription regulation is a controversial one.

## 1. Introduction

While H2A.Z blocks Notch and  $\Delta$ Np63a target genes (Gallant-Behm *et al.*, 2012; Giaimo *et al.*, 2018) it is essential for 'estrogen receptor alpha' (ER $\alpha$ ) signalling (Gévry *et al.*, 2009). In sharp contrast to the ER $\alpha$  signalling, H2A.Z negatively regulates transcription of the Cyclin D1 gene during transcriptional initiation in response to estradiol (Dalvai *et al.*, 2013). In *Drosophila* as well as in human the variant could be linked with occurring transcription events (Barski *et al.*, 2007; Kusch *et al.*, 2014). The link to transcription repression might be explained by nucleosome surface alterations caused by H2A.Z which promote binding of HP1-Alpha (an isoform of HP1) and therefore chromatin fibre folding (Fan *et al.*, 2004). PTMs were shown to be the key factors in determining the role of H2A.Z. Ubiquitination and de-ubiquitination for instance, were linked with X-Chromosome silencing and androgen receptor-mediated gene activation in mammals (Sarcinella *et al.*, 2007; Draker *et al.*, 2011). Active transcription was first linked with acetylation of H2A.Z in 2005 when acetylated H2A.Z was identified at actively-transcribed genes in chicken (Bruce *et al.*, 2005). Since then, histone acetylation became a hallmark for increased transcriptional activity in eukaryotes (Millar *et al.*, 2006; Halley *et al.*, 2010; reviewed in Giaimo *et al.*, 2019). Dynamic regulation of transcription in response to intra- or extracellular signals, requires active deacetylation of H2A.Z. In 2010 the research group of Michael-Christopher Keogh showed that 'histone deacetylase 1' (Hda1) actively deacetylates H2A.Z and that acetylations of lysine 8, 10 and 14 are redundant (Mehta *et al.*, 2010). Systematic inhibition of deacetylases via Trichostatin-A in mice led to an increased level of H2A.Z acetylations without increasing the overall level of chromatin-associated H2A.Z (Narkaj *et al.*, 2018) indicating that H2A.Z is also actively de-acetylated in higher eukaryotes. A specific histone deacetylase being responsible for deacetylation of H2A.Z in higher eukaryotes remains elusive.

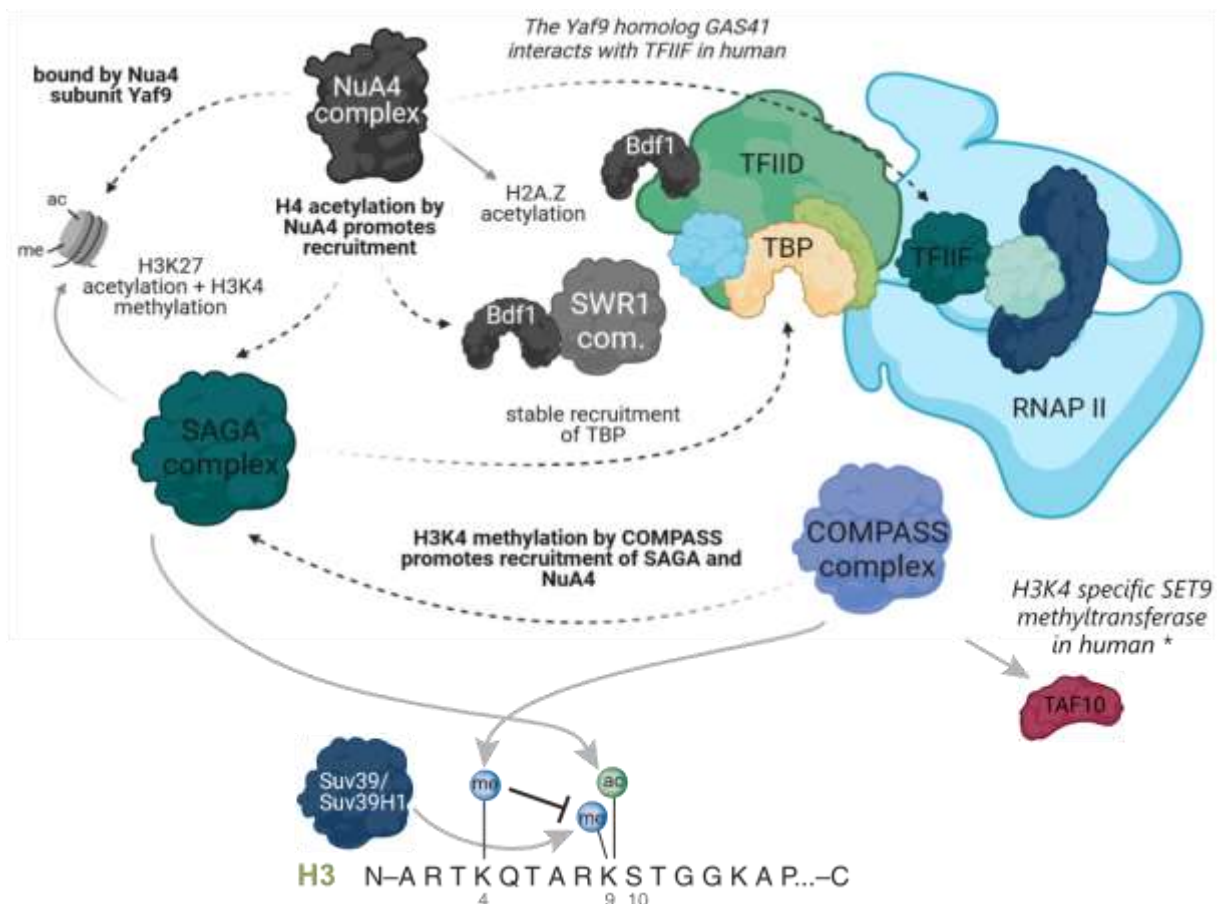
### **1.5.2 Factors of the H2A.Z transcription regulation machinery that interact with the RNAP II transcription pre-initiation complex**

Transcription regulation by H2A.Z is highly complex and involves multiple PTMs and protein complexes (Fig 8). The following paragraph shows how H2A.Z, the previously mentioned acetylation of Histone H4, methylation and acetylation of histone H3K9 and at least four different proteins complexes interact with each other in *S. cerevisiae* to regulate transcription initiation.

Bdf1 a subunit of the SWR1 complex (Altaf *et al.*, 2010; Nguyen *et al.*, 2013) could promote a direct way of transcription activation since it is also a loosely associated component of the general transcriptional initiation factor 'transcription factor II D' (TFIID) in *S. cerevisiae* (Matangkasombut *et al.*, 2000). Whether an exchange of Bdf1 between the SWR1 complex and TFIID occurs or if separate Bdf1 proteins bind to the complexes is not known yet. TFIID recognizes promoters and recruits the PIC which then RNAP II (Orphanides *et al.*, 1996; Papai *et al.*, 2011; Warfield *et al.*, 2017).

## 1. Introduction

NuA4 acetylates H4 (Altaf *et al.*, 2010) which stimulates, together with methylation of H3K4, the recruitment of the bromodomain-containing SAGA complex (Ginsburg *et al.*, 2014; Ringel *et al.*, 2015). The SAGA complex is required for stable binding of TBP to the promoter region (Baptista *et al.*, 2017) and acetylates H3K27 and H3K9 residues among others (Riss *et al.*, 2015). The role of Gcn5 in acetylation of H3K9 and H3K27 appears to be conserved since Gcn5 in *Arabidopsis thaliana* is also required for both acetylations (Benhamed *et al.*, 2006). Acetylation of the histones H4 by NuA4 in turn is dependent on methylation of H3K4 by COMPASS (Ginsburg *et al.*, 2014) but also on SAGA-mediated acetylation since Yaf9 actively binds to acetylated H3K27 (Klein *et al.*, 2018). The Yaf9 homolog GAS41 also binds to TFIIF (Heisel *et al.*, 2010). TFIIF is required for loading of RNAP II into the PIC (reviewed in Luse, 2012). Due to the positive effect of H3K4 methylation by COMPASS on H3K9 and H3K27 acetylation (Bian *et al.*, 2011) COMPASS is indirectly linked with the Bdf1/NuA4/SWR1/SAGA transcription initiation and H2A.Z incorporation cascade in *S. cerevisiae* (Fig 8). Therefore, COMPASS promotes a feedback loop in which SAGA and NuA4 support the recruitment of each other.



**Figure 8: Interaction network of PTMs, proteins and protein complexes that regulate transcription initiation.**

Grey sustained lines indicate the generation of a PTM by a complex, black sustained lines indicate enhancing (arrow) or repressing effects (straight line), dashed arrows indicate a recruitment process or an interaction; the figure has been designed with BioRender©. \* = Kouskouti *et al.*, 2004

## 1.6 *Trypanosoma brucei*, a small overview

The eugleozoan (formerly known as flagellated protozoan) parasite *Trypanosoma brucei* belongs to the class of *Kinetoplastids*, to the order of *Trypanosomatida* and the family of *Trypanosomatidae*. It is the causative agent of African trypanosomiasis. This vector-borne disease is known as sleeping sickness in human and as Nagana in cattle. The epithet “sleeping sickness” originates from the ability of the parasite to pass the blood-brain barrier and to disturb the sleep cycle, whereas Nagana is characterized by impairments in growth, fertility and work output. In 2009, around 10,000 human cases have been reported with approximately 3000 annual infections per year, which decreased to less than 1000 new cases in 2019 (WHO 2021). For transmission the parasite depends on its vector, the tsetse fly (*Glossina spec.*; Malvy & Chappuis, 2011), which restricts the disease to the vector habitat of sub-Saharan Africa. The sleeping sickness is caused by the two subspecies *Trypanosoma brucei gambiense* and *Trypanosoma brucei rhodensiense* (Malvy & Chappuis, 2011). The third subspecies *Trypanosoma brucei brucei*, which infects cattle (Courtin *et al.*, 2008) is the preferred model organism for trypanosome biology.

### 1.6.1 *T. brucei* as a model organism for *Kinetoplastids*

The global warming increases the importance of research that focuses on *Kinetoplastids* as more and more areas become endemic for several subspecies like *Leishmania spec.* or individual parasites like *Trypanosoma cruzi* (Short *et al.*, 2017; Guarner, 2019). A reliable model organism is required to understand unique features of the *Kinetoplastid* biology to develop new and potent drugs to counter this development. In course of the past decades, *T. brucei* became such a model organism as plenty of molecular tools became available to study molecular processes. A conventional gene ‘knockout’ (KO) of non-essential genes in which the ‘gene of interest’ (GOI) is replaced by an ‘open reading frame’ (ORF) of a selectable resistance marker uses the highly effective homologous recombination machinery of this organism. Only around 30 bp of homologous DNA sequence are needed for homologous recombination in *T. brucei*, which allows the utilization of fast polymerase chain reaction (PCR)- based techniques (Shen *et al.*, 2001; Arhin *et al.*, 2004). The establishment of an inducible system that includes the bacteriophage T7 polymerase and a tetracycline repressor (TETR) system allows protein overexpression (Wirtz *et al.*, 1999). In contrast to *T. cruzi* and most *Leishmania spec.* the RNA interference (RNAi) mechanism is available in *T. brucei*, which allows targeted downregulation of essential proteins in combination with the TETR system (DaRocha *et al.*, 2004; Bellofatto & Palenchar, 2008; Lye *et al.*, 2010).

### 1.6.2 Transcription and H2A.Z in *T. brucei*

The genome of *T. brucei* is around 35 Mb in size (haploid) and contains ~9,000 genes distributed among its eleven chromosomes (Berriman *et al.*, 2005; Muller *et al.*, 2018) and its organization is highly unique. Most protein-coding genes in *Kinetoplastids* are organized in large polycistronic transcription units (PTUs) which are transcribed by RNA pol II (Fig 1.11; Campbell *et al.*, 2003; Clayton, 2016). Polycistronic transcription has also been described in other organisms but in a far less extent. In *Caenorhabditis elegans* only around 15% of the protein-coding genes are arranged in PTUs (Blumenthal *et al.*, 2002). PTUs in *T. brucei* were shown to be separated by rRNA and tRNA gene arrays (transcribed by RNAP I and RNAP III) or by so-called strand switch regions (SSRs) in which the transcription changes its direction (Fig 10; Muller *et al.*, 2018, reviewed in Siegel *et al.*, 2011). Factors that define the direction of transcription could not be identified so far. SSRs harbour either two transcription start sites (divergent TSS; TSSs with a previous transcription termination site (TTS) are called non-divergent TSS) or two TTSs. During a *trans*-splicing reaction that requires a 'splice leader' (SL) RNA and polyadenylation, the large polycistronic mRNAs are processed into individual mature mRNAs (reviewed in Clayton, 2019). The sequence coding for the SL is arranged in a multi-copy array on chromosome 9 and independently-transcribed by the RNAP II (Gilinger & Bellofatto, 2001; Lee *et al.*, 2007). Transcription of the SL is regulated by a bipartite sequence element located upstream of the SL-array and an initiator (Gilinger & Bellofatto, 2001). The general transcription factors TFIIA, TFIIB, and TFIIF appear to be involved in SL RNA transcription regulation but not in RNAP II-mediated transcription in general (Schimanski *et al.*, 2005a; Schimanski *et al.*, 2006; Lee *et al.*, 2007).

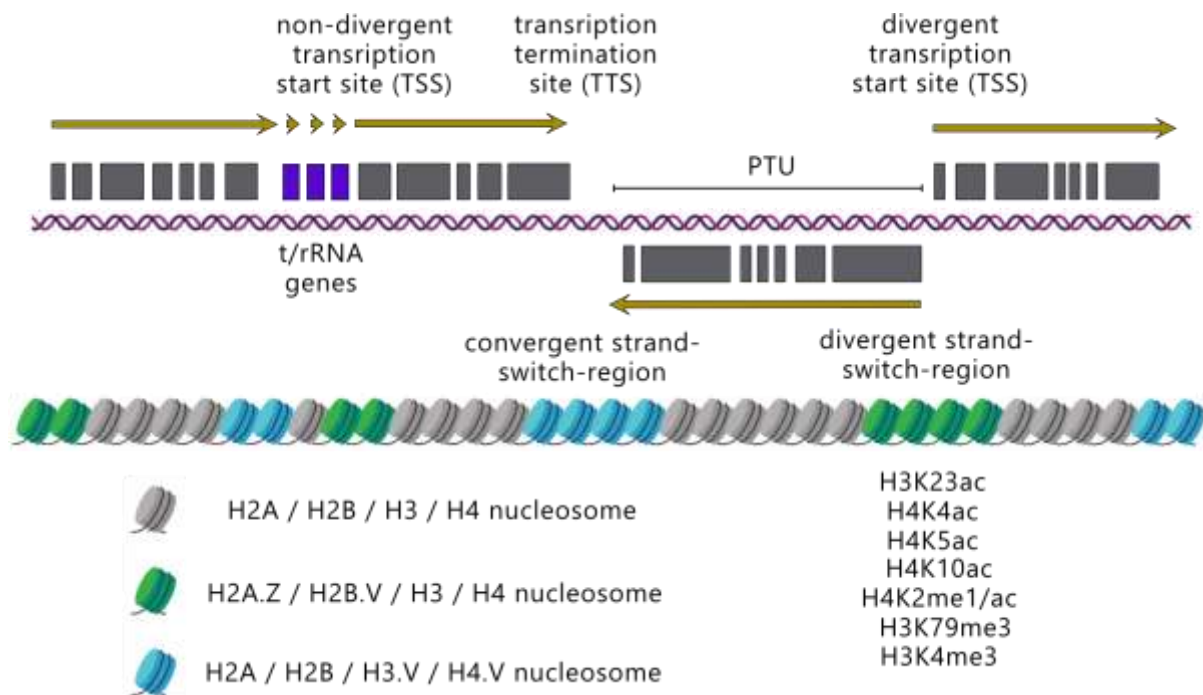
Just like in other organisms H2A.Z appears to be a key player in transcription regulation. H2A.Z, together with the histone variant H2B.V, is located in the TSSs of *T. brucei* (Fig 10; Siegel *et al.*, 2009) and loss of H2A.Z acetylation leads to reduced RNAP II transcription (Kraus *et al.*, 2020). The length of TSS in *T. brucei* can cover a range of up to 10 kb, therefore forming more of a transcription start region (TSR) than a TSS. Aside from its role in transcription H2A.Z could also be linked with alterations of the nucleosome and chromatin structure. H2A.Z/H2B.V-containing nucleosomes possess a reduced nucleosome stability (Siegel *et al.*, 2009) and areas of the genome that are enriched in H2A.Z showed a higher susceptibility towards 'micrococcal nuclease' (MNase) treatment indicating that incorporation of H2A.Z leads to a more open chromatin structure (Wedel *et al.*, 2017). The molecular processes underlying transcription initiation in *T. brucei* remain elusive and no hints for regulation of RNAP II-mediated transcription could be found so far (reviewed in Clayton, 2019).

While promoter-like elements remained unidentified for a long time, recent research work provided hints that G-C rich DNA sequences can act as promoter elements and recruit H2A.Z indicating that transcription initiation and chromatin structure are linked with each other (Wedel *et al.*, 2017). Recent data from Nicolai Siegel's research group supported this hypothesis (Wedel *et al.*, 2017).

### **1.6.3 Transcription-associated post-translational modifications in *T. brucei***

Experiments performed in the research group of George Cross in 2006 and 2007 identified the first histone PTMs in *T. brucei*. While some of them correspond to conserved PTMs in other eukaryotic organisms (Mandava *et al.*, 2007) a few PTMs (acetylations on the C-terminus of the histone H2A) were identified as *T. brucei* specific (Janzen *et al.*, 2006). Two of the conserved PTMs were methylation of H3K4 and acetylation of H4K10. Alignments of *T. brucei* histone H4 show that the *T. brucei* H4 lysine 10 corresponds to the H4 lysine 12 in other eukaryotes (Fig S1). Recent investigations of the Siegel research group identified more than 150 histone PTMs, many of them exclusively located on H2A.Z- and H2B.V-containing nucleosomes. Besides the already known and TSS-associated H3K4 tri-methylation and H4K10 acetylation (Siegel *et al.*, 2009; Wright *et al.*, 2010), acetylation of H3K23 (homologous to H3K27 in other eukaryotes, Fig S1), H4K4 and H4K5 were identified as well (Fig 9; Kraus *et al.*, 2020). A knockdown of the 'MOZ, Ybf2, Sas2, TIP60' (MYST) 'histone acetyltransferase 2' (HAT2) resulted in a reduction of H4K2, H4K4 and H4K10 acetylation and reduced chromatin association of H2A.Z (Kraus *et al.*, 2020). Acetylations on the histone variants H2A.Z and H2B.V are dependent on HAT1 activity (Kraus *et al.*, 2020). The identification of conserved transcription-associated PTMs like H3K23 acetylation or H3K4 methylation raise the question whether complexes like COMPASS or SAGA exist in *T. brucei*. Many PTM synthesizing enzymes need to be identified and how certain PTMs might contribute to transcription regulation or to the chromatin landscape must be further investigated. Even though it is known that the bromodomain-containing Bdf3 co-localizes with H2A.Z at TSSs (Siegel *et al.*, 2009), the role of this factor (and PTM-recognizing proteins in general) in H2A.Z-dependent transcription regulation remains also elusive. Drug development approaches in the recent past focused on Bdfs as potential new targets (Jeffers *et al.*, 2017). Given their role in transcription regulation (Matangkasombut *et al.*, 2000; Altaf *et al.*, 2010; Donczew & Hahn, 2021) especially when linked with the SWR1 and NuA4 complex (Altaf *et al.*, 2010; Nguyen *et al.*, 2013) and given the localisation of *T. brucei* Bdf3, the identification of a SWR1-like and a NuA4-like complex in *T. brucei* might provide new potential drug targets.





**Figure 9: Genome organization of *T. brucei***

Green nucleosomes represent nucleosomes at the transcription start site (TSS), blue nucleosomes represent nucleosomes at the transcription termination site; polycistronic transcription unit (PTU); PTMs of the TSSs origin from (Kraus *et al.*, 2020); the figure has been designed with BioRender©.

### 1.6.4 Chromatin remodelling and its complexes in *T. brucei*

Even though various SNF2 proteins can be found in *T. brucei*, only limited information is available about SNF2 CRCs. Only two SNF2 remodelling complexes could be identified so far. An ISWI-like complex was identified in 2015 (Stanne *et al.*, 2015) and deregulation of this complex caused an expression de-regulation of the 'Variant Surface Glycoproteins' (VSGs) of the parasite (Hughes *et al.*, 2007; Stanne *et al.*, 2011). In addition to its role in RNAP I-mediated expression of VSGs, the ISWI was identified at the boundaries of RNAP II-transcribed PTUs (Stanne *et al.*, 2015). This indicates a general role in chromatin remodelling in *T. brucei* as well. The other SNF2 complex is the 'J-Base binding protein 2' (JBP2) complex (Kieft *et al.*, 2020). Base J is a modified DNA nucleotide that was first identified in *T. brucei* in 1993 (Gommers-Ampt *et al.*, 1993). The JBP2 complex together with the histone variant H3.V are involved in transcription termination (Reynolds *et al.*, 2016; Kieft *et al.*, 2020). Important SNF2 CRCs especially INO80- and SWR1-like complexes remain elusive. Nevertheless, other important non-SNF2 complexes were identified in *T. brucei*. The FACT complex was identified by the research group of Gloria Rudenko and could be linked to chromatin dynamics in course of expression regulation of VSGs (Denninger & Rudenko, 2014). Homology searches within the *T. brucei* genome could identify several proteins that might be involved in establishing an open chromatin structure.



While histone acetyltransferases were identified (Kawahara *et al.*, 2008; Siegel *et al.*, 2008) and linked with H2A.Z-dependent chromatin remodelling (Kraus *et al.*, 2020), several histone methyltransferases (Figueiredo *et al.*, 2009) and Bdfs (Siegel *et al.*, 2009) could not be linked with chromatin remodelling yet (Staneva *et al.*, 2021). In regard of lacking SNF2 complexes of the INO80-subfamily and the lack of RNAP II transcription regulation in *T. brucei* many aspects of H2A.Z dependent transcription regulation need to be further investigated.

### 1.7 Aim of the study

Several studies performed in various eukaryotic model organisms revealed that transcription regulation by the histone variant H2A.Z is highly conserved. It requires a well-organized interplay of multiple protein complexes and PTMs. H2A.Z in *T. brucei* was found to play an essential role in transcription regulation but important transcription-associated complexes like SWR1, INO80, NuA4, COMPASS or SAGA could not be identified so far. This raises several questions that need to be answered: 1a.) Does a SWR1- and INO80-like complex exist in *T. brucei* that is responsible for incorporation and removal of H2A.Z or 1b.) Does a specialized histone chaperone that incorporates the variant? 2.) Does a complex exist which is homologous to the NuA4/TIP60 complex? 3.) Does the H2A.Z transcription regulation pathway involve other complexes like homologs of the SAGA and/or COMPASS complex? According to these questions the tasks for my project were:

- Identification of complexes of the INO80 subfamily and/or chaperones that are involved in H2A.Z incorporation
  - Characterization of complex components to classify their contribution to H2A.Z incorporation
  - Performing ChIP-seq experiments to determine nucleosome-associated H2A.Z levels
- Identification of a complex homologous to the NuA4/TIP60 complex
  - Characterization of complex components to assess their contribution to H2A.Z incorporation
- Identification of complexes homologous to SAGA or COMPASS that are involved in transcription regulation

Answering these questions will provide new insights into H2A.Z-dependent transcription regulation in the unicellular parasite *T. brucei* and might also provide new information about *Kinetoplastid*-specific RNAP II transcription regulation in general.

## 1. Introduction

## 2. Materials and Methods

### 2.1 Bacterial culture

#### 2.1.1 Bacteria growth

*Escherichia coli* (*E. coli*) (strain: TOP10; genotype: F-mcrA  $\Delta$ (mrr-hsdRMS-mcrBC)  $\phi$ 80lacZ $\Delta$ M15  $\Delta$ lacX74 nupG recA1 ara D139  $\Delta$ (ara-leu)7697 galE15 galK16 rpsL(StrR) endA1 l) were either grown in liquid 'lysogeny broth' (LB) medium (1 % tryptone, 0.5 % yeast extract, 0.5 % NaCl, pH 7.5) in baffled flasks overnight shaking at 160 'revolutions per minute' (rpm) or on LB agar plates (LB medium containing 1.5 % agar) at 37 ° Celsius (C). In both cases, required antibiotics were added for selection (100 microgram ( $\mu$ g)/milliliter (ml) ampicillin (amp), 100  $\mu$ g/ml kanamycin).

#### 2.1.2 Transformation of chemically competent *E. coli*

100 microliter ( $\mu$ l) chemically competent *E. coli* cells were thawed on ice. DNA (2  $\mu$ l BP cloning reaction, 10  $\mu$ l ligation reaction or 1 nanogram (ng) plasmid DNA) was added and mixed by flicking the reaction tube five times. Mixture was placed on ice for 30 minutes (min) prior to a heat shock at 42°C for 45 seconds (sec) and further incubation on ice for 2 min. 500  $\mu$ l of 'room temperature' (RT) 'super optimal broth with catabolite repression' (SOC) medium (2% tryptone, 0.5% yeast extract, 10 millimolar (mM) NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM glucose) was added and bacteria were grown for 1 hour (h) shaking at 160 rpm and at 37°C. This recovery is required for the cells to express the antibiotic resistance coded on the transformed plasmid. Cells were centrifuged (500 'relative centrifugal force' (g) for 1 min at RT) and 400  $\mu$ l of supernatant was removed. The remaining bacteria were resuspended by pipetting and plated on pre-warmed LB plates containing the corresponding antibiotics.

#### 2.1.3 Isolation of plasmid DNA from *E. coli*

High-copy plasmids were isolated from 3 ml to 15 ml overnight bacterial cultures using the NucleoSpin® Plasmid Mini kit or from 50 ml to 100 ml using the NucleoBond® PC100 Midi kit (Macherey-Nagel) according to the manufacturer's instructions.

## 2.2 Trypanosomatid methods

### 2.2.1 *T. brucei* strains

Table 2 shows the parental cell lines that were used for the genetic modifications listed in Table 3. All studies were conducted with the *Trypanosoma brucei brucei* substrain, referred to as *T. brucei* in the following.

**Table 2: Parental and transgenic *T. brucei* cell lines used and generated in this study**

Name	Genotype	Construct	Selection	Reference
<b>Bloodstream form cell lines (BSF)</b>				
SM	Strain Lister 427, MiTat1.2, clone 221a, TETR T7RNAP NEO	pLEW114hyg5' pHD 328	G418/neomycin	Wirtz <i>et al.</i> , 1999
2T1	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::hygΔstart PAC	pHD1313, ph3EP	phleo, puro	Alsford <i>et al.</i> , 2005

**Table 3: Transgenic *T. brucei* cell lines generated and used in this study**

Name	Genotype	Construct	Selection	Reference
<b>Bloodstream form cell lines (BSF)</b>				
2T1 Ty1-H2A.Z	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::hygΔstart PAC Ty1::H2A.Z BSD	pHD1313, ph3EP	Phleo, Puro, Blas	T. Vellmer
2T1 Ty1:H2A.Z TbSWR1 RNAi <sup>Ti</sup>	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::TbSWR1 (2476 - 2876) <sup>Ti</sup> HYG   Ty1::H2A.Z BSD	pHD1313, ph3EP, pPOTv7_Ty1	Phleo, Hyg, Blas	T. Vellmer
2T1 Ty1:H2A.Z TbSWRC1 RNAi <sup>Ti</sup>	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::TbSWRC1 (-99 - 528) <sup>Ti</sup> HYG   Ty1::H2A.Z BSD	pHD1313, ph3EP, pPOTv7_Ty1	Phleo, Hyg, Blas	T. Vellmer
2T1 TbSWRC2 RNAi <sup>Ti</sup>	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::TbSWRC2 (-28 - 528) <sup>Ti</sup> HYG	pHD1313, ph3EP,	Phleo, Hyg	T. Vellmer
2T1 TbSWRC3_1 RNAi <sup>Ti</sup>	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::TbSWRC2 (875 - 1467) <sup>Ti</sup> HYG	pHD1313, ph3EP,	Phleo, Hyg	T. Vellmer
2T1 TbSWRC3_2 RNAi <sup>Ti</sup>	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::TbSWRC2 (93 - 612) <sup>Ti</sup> HYG	pHD1313, ph3EP,	Phleo, Hyg	K. Köthe
2T1 TbSWRC3_3 RNAi <sup>Ti</sup>	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::TbSWRC2 (1488 - 1988) <sup>Ti</sup> HYG	pHD1313, ph3EP,	Phleo, Hyg	K. Köthe
2T1 HAT1 RNAi <sup>Ti</sup>	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::HAT1 (648 - 1184) <sup>Ti</sup> HYG	pHD1313, ph3EP,	Phleo, Hyg	T. Vellmer
2T1 HAT2 RNAi <sup>Ti</sup>	Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::HAT2 (499 - 1063) <sup>Ti</sup> HYG	pHD1313, ph3EP,	Phleo, Hyg	T. Vellmer

## 2. Materials and Methods

<b>Name</b>	<b>Genotype</b>	<b>Construct</b>	<b>Selection</b>	<b>Reference</b>
2T1 Ty1:H2A.Z H2A.Z RNAi <sup>Ti</sup>	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::H2A.Z (155 - 504)<sup>Ti</sup> HYG   Ty1::H2A.Z BSD</i>	pHD1313, ph3EP, pPOTv7_Ty1	Phleo, Hyg, Blas	T. Vellmer
2T1 Ty1:H2A.Z RPB1 RNAi <sup>Ti</sup>	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::RPB1 (3804 - 4259)<sup>Ti</sup> HYG   Ty1::H2A.Z BSD</i>	pHD1313, ph3EP, pPOTv7_Ty1	Phleo, Hyg, Blas	T. Vellmer
2T1 Ty1:H2A.Z TbSWR1 RNAi <sup>Ti</sup> Rluc	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::TbSWR1(2476 - 2876)<sup>Ti</sup> HYG   Ty1::H2A.Z BSD   TUB::Rluc NEO</i>	pHD1313, ph3EP, pPOTv7_Ty1 pNG14	Phleo, Hyg, Blas, G418/Neo	T. Vellmer
2T1 HAT1 RNAi <sup>Ti</sup> Rluc	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::HAT1 (648 - 1184)<sup>Ti</sup> HYG   TUB::Rluc NEO</i>	pHD1313, ph3EP, pPOTv7_Ty1 pNG14	Phleo, Hyg, Blas, G418/Neo	T. Vellmer
2T1 HAT2 RNAi <sup>Ti</sup> Rluc	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::HAT2 (499 - 1063)<sup>Ti</sup> HYG   TUB::Rluc NEO</i>	pHD1313, ph3EP, pNG14	Phleo, Hyg, G418/Neo	T. Vellmer
2T1 Ty1:H2A.Z RPB1 RNAi <sup>Ti</sup> Rluc	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::RPB1 (3804 - 4259)<sup>Ti</sup> HYG   Ty1::H2A.Z BSD   TUB::Rluc NEO</i>	pHD1313, ph3EP, pPOTv7_Ty1 pNG14	Phleo, Hyg, Blas, G418/Neo	T. Vellmer
2T1 HA:TbSWR1sT	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::hygΔstart PAC   HA::TbSWR1 HYG</i>	pHD1313, ph3EP, pPOTv7_HA	Phleo, Puro, Hygro,	T. Vellmer
2T1 TbSWR1:HA sT	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::hygΔstart PAC   HA::TbSWR1 HYG</i>	pHD1313, ph3EP, pMOTag4h	Phleo, Puro, Hygro,	T. Vellmer
2T1 HA:TbSWRC1	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::hygΔstart PAC   HA::TbSWRC1 HYG   ΔTbSWRC1 NEO</i>	pHD1313, ph3EP, pPOTv7_HA	Phleo, Puro, Hygro, G418/Neo	T. Vellmer
2T1 HA:TbSWRC2	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::hygΔstart PAC   HA::TbSWRC2 HYG   ΔTbSWRC2 NEO</i>	pHD1313, ph3EP, pPOTv7_HA	Phleo, Puro, Hygro, G418/Neo	T. Vellmer
2T1 TBL:TbSWRC2	<i>Strain Lister 427, MiTat1.2, clone 221a, TUB::TETR BLE   RRNA::hygΔstart PAC   TBL::TbSWRC2 HYG   ΔTbSWRC2 NEO</i>	pHD1313, ph3EP, pPOTv4_TBL	Phleo, Puro, Hygro, G418/Neo	T. Vellmer

## 2. Materials and Methods

Name	Genotype	Construct	Selection	Reference
2T1 HA:TbSWRC3	Strain <i>Lister 427</i> , <i>MiTat1.2</i> , clone <i>221a</i> , TUB::TETR BLE   RRNA::hygΔstart PAC   HA::TbSWRC3 HYG	pHD1313, ph3EP, pPOTv7_HA	Phleo, Puro, Hygro	T. Vellmer
2T1 TbSWRC3:HA	Strain <i>Lister 427</i> , <i>MiTat1.2</i> , clone <i>221a</i> , TUB::TETR BLE   RRNA::hygΔstart PAC   TbSWRC3::HA HYG	pHD1313, ph3EP, pMOTag4H	Phleo, Puro, Hygro	T. Vellmer
2T1 ΔTbSWRC3	Strain <i>Lister 427</i> , <i>MiTat1.2</i> , clone <i>221a</i> , TUB::TETR BLE   RRNA::hygΔstart PAC   ΔTbSWRC3 HYG	pHD1313, ph3EP	Phleo, Puro, Hygro	T. Vellmer
2T1 TbSWRC4:HA	Strain <i>Lister 427</i> , <i>MiTat1.2</i> , clone <i>221a</i> , TUB::TETR BLE   RRNA::hygΔstart PAC   TbSWRC4::HA HYG   TbSWRC4::HA NEO	pHD1313, ph3EP, pMOTag3H/4H	Phleo, Puro, Hygro, G418/Neo	T. Vellmer
2T1 HA:HAT2	Strain <i>Lister 427</i> , <i>MiTat1.2</i> , clone <i>221a</i> , TUB::TETR BLE   RRNA::hygΔstart PAC   HA:HAT2 HYG   ΔEaf6-like NEO	pHD1313, ph3EP, pPOTv7_HA	Phleo, Puro, Hygro, G418/Neo	T. Vellmer
2T1 HA:TbEaf6- like	Strain <i>Lister 427</i> , <i>MiTat1.2</i> , clone <i>221a</i> , TUB::TETR BLE   RRNA::hygΔstart PAC   HA::Eaf6- like HYG   ΔEaf6-like NEO	pHD1313, ph3EP, pPOTv7_HA	Phleo, Puro, Hygro, G418/Neo	T. Vellmer
SM Ty1:TbEaf3- like	Strain <i>Lister 427</i> , <i>MiTat1.2</i> , clone <i>221a</i> , TETR T7RNAP NEO   Ty1::Eaf3-like BSD	pHD1313, ph3EP, pPOTv7_Ty1	G418/Neo, Blas	T. Vellmer
SM Ty1:Bdf3	Strain <i>Lister 427</i> , <i>MiTat1.2</i> , clone <i>221a</i> , TETR T7RNAP NEO   Ty1::Bdf3-like BSD	pHD1313, ph3EP, pPOTv7_HA	G418/Neo, Blas	L. Hartleb

### Procylic form cell lines (PCF)

Name	Genotype	Construct	Selection	Reference
427	Strain <i>Lister 427</i>	-	-	S. Kramer
29-13	Strain <i>Lister 427</i> , TETR T7RNAP NEO HYG	pLEW29, pLEW13	G418, Hygro	Wirtz <i>et al.</i> , 1999
427 RuvB2:HA	Strain <i>Lister 427</i> , <i>RuvB2::HA</i> PAC   <i>RuvB2::HA</i> HYG	pMOTag3H/4H	Hygro, Puro	T. Vellmer

Abbreviations: MiTat1.2 (Molteno institute antigen type 1.2), SM (single marker), TUB (tubulin), TETR (tetracycline repressor protein) T7RNAP (T7 RNA polymerase), Blas (blasticidin), BLE (bleomycin resistance gene), BSD (blasticidin-S deaminase resistance gene), eYFP (enhanced yellow fluorescent protein), G418/Neo (neomycin), HA (hemagglutinin), HYG (hygromycin phosphotransferase resistance gene), Hygro (hygromycin), NEO (aminoglycoside phosphotransferase resistance gene), PAC (puromycin N-acetyltransferase resistance gene), Phleo (phleomycin), Puro (puromycin), 2T1 (VSG221-expressing, tagged, clone1).

### 2.2.2 *In vitro* cultivation of *T. brucei*

Monomorphic *T. brucei* bloodstream form (BSF) cells were cultivated in 'Hirumi's modified Iscove's medium 9' (HMI-9 medium; for 10 liter (L): 176.6 gram (g) Iscove's modified Dulbecco's medium (IMDM), 30.24 g NaHCO<sub>3</sub>, 143 µl β-mercaptoethanol, 100 ml penicillin/streptomycin solution (Invitrogen), 282 milligram (mg) bathocuproine sulfonate, 390 mg thymidine, 1.36 g hypoxanthine, 1.82 g L-cysteine, 1L heat-inactivated 'fetal calf serum' (FCS), pH 7.5) (Hirumi and Hirumi 1989) at 37°C and 5% CO<sub>2</sub> in humidified air. *T. brucei* procyclic form (PCF) were cultured in modified 'semi-defined medium' (SDM)-79 (Brun and Schonberger 1979) (for 10 L: 20 g NaHCO<sub>3</sub>, 18.642 g dextrose, 5.134 g L-glutamine, 6.15 g L-proline, 1 g sodium pyruvate, 4.068 g L-threonine, 105 mg sodium acetate, 224.2 mg L-glutamic acid, pH 7.3; 900 ml basic SDM-79 medium was complemented with 7.5 mg/ml hemin, 10 ml penicillin/streptomycin solution (Invitrogen) and 100 ml heat-inactivated FCS prior usage) at 27 °C and 5 % CO<sub>2</sub> in humidified air. For routine cultivation of BSF and PCF cells were maintained in 10 ml medium in vented T25 flasks. Cells were always kept in the logarithmic growth phase. *Tb*BSFs were kept below 1x10<sup>6</sup> cells/ml and *Tb*PCFs were kept between 5x10<sup>5</sup> cells/ml to 1x10<sup>7</sup> cells/ml. A Coulter Counter Z2 particle counter (Beckman Coulter) was used to measure cell densities. If dilutions below 5x10<sup>5</sup> cells/ml were required for *Tb*PCF cells, the cells were cultivated in 20 % conditioned medium. Conditioned SDM-79 medium was previously prepared by growth of *Tb*PCF cells to >1x10<sup>7</sup> cells/ml to mimic a cell density of 2x10<sup>6</sup> in 20 % conditioned medium. Cells were centrifuged (1,500 x g for 10 min at RT) and the supernatant was sterilized by filtration with a 0.2 micrometer (µm) filter. Drug selections were added to the culture media as indicated in table 2.2 if required.

**Table 4: Drug selection used for *T. brucei* cell culture**

Drug	Gene	Stock	Concentration <i>Tb</i> BSF (mg/ml)	Concentration <i>Tb</i> PCF (mg/ml)	Company
G418	NEO	10	2	15	AppliChem
Hygromycin	HYG	10	2.5	25	AppliChem
Phleomycin	BLE	10	2.5	2.5	InvivoGen
Puromycin	PAC	1	0.1	1	AppliChem
Blasticidin	BSD	10	5	5	InvivoGen
Tetracycline	-	10	1	0.005-1	AppliChem

### 2.2.3 Growth curves

For preparation of a growth curve logarithmically growing BSF cells were diluted down to 8x10<sup>4</sup> cells/ml in 10 ml medium. Cell density was measured after diluting the cells using the Coulter Counter Z2 particle counter (Beckman Coulter). Three times 2 ml were transferred into separate wells of a 24-well plate to perform the analysis in triplicates. For 96 – 120 h the cells were counted and diluted (if necessary) down to 8x10<sup>4</sup> in a total volume of 2 ml every day., The dilution factors were multiplied by the corresponding cell numbers to calculate cumulative growth values. The population doubling times (tD) were determined by calculating the growth rates (g) as stated below:

## 2. Materials and Methods

$$g = \frac{\ln(x1 - x2)}{t1 - t0} \quad tD = \frac{\ln 2}{g}$$

### 2.2.4 Genetic modification of Trypanosomes

10 µg linearized plasmid DNA or 7 µg PCR product were transferred into a BTX cuvette in preparation of the transfection.  $2 \times 10^7$  BSF cells were harvested (1,500 x *g* for 10 min at RT) and washed once with 10 ml pre-warmed 'trypanosome dilution buffer' (TDB; 5 mM KCl, 80 mM NaCl, 1 mM MgSO<sub>4</sub>, 20 mM Na<sub>2</sub>PO<sub>4</sub>, 2 mM NaH<sub>2</sub>PO<sub>4</sub>, 20 mM glucose, pH 7.4; 1,500 x *g* for 10 min at RT). The cell pellet was resuspended in 400 µl pre-warmed transfection buffer (90 mM Na<sub>2</sub>PO<sub>4</sub>, 5 mM KCl, 0.15 mM CaCl<sub>2</sub>, 50 mM HEPES, pH 7.3) and transferred to the BTX-cuvette that was prepared with the DNA. The BTX cuvette with DNA-cell mixture was electroporated using the AMAXA nucleofector (Lonza) choosing the program 'X001 free choice'. The transfected cells were then put in pre-warmed HMI-9 medium that contains the parental selection. A 1:5, 1:10 and 1:100 dilution was prepared and plated on two 24-well plates (1 ml/well). 8 h post transfection, drugs were added to select for positively transfected cells. The 8 hours recovery time make allow the cells divide at least one time to guarantee proper integration of the DNA construct and expression of the resistance marker. PCF cells were transfected as described above with a few minor adaptations. PCF were washed in 10 ml 'phosphate-buffered saline' (PBS; 10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub> pH 7.4, 140 mM NaCl, 2.7 mM KCl). After the electroporation, PCF cells were transferred into 20% conditioned SDM-79 medium and diluted 1:10, 1:100 and 1:1000. Each dilution was plated on one 24-well plate (1 ml/well). The selection was added 12-15 h after the transfection.

### 2.2.5 Freezing and thawing

Logarithmically growing cells were harvested by centrifugation (1,500 x *g* for 10 min at RT) and  $2 \times 10^6$  BSF, or  $2 \times 10^7$  PCF cells per vial were resuspended in 500 µl ice-cold freezing medium (HMI-9 for BSF cells or SDM-79 for PCF cells with 10 % glycerol). Cells were frozen in cryogenic tubes (Sarstedt) in Styrofoam at -80 °C. To retrieve cells, they were thawed in a 37°C water bath, washed with 10 ml pre-warmed HMI-9 or SDM-79 medium (1,500 x *g* for 10 min at RT) and transferred into a culture flask that contains the corresponding drug with selective drugs as required.

### 2.2.6 Isolation of genomic DNA from *T. brucei* cells

To purify genomic DNA of *T. brucei* for PCR amplifications the High Pure PCR Template Preparation kit (Roche) was used. Briefly,  $5 \times 10^6$  cells were harvested (1,500 x *g* for 10 min at RT) and the pellet resuspended in 200 µl PBS buffer. 200 µl of Binding Buffer and 40 µl Proteinase K were added, mixed with subsequent incubation at 70°C for 10 min. After incubating 100 µl of isopropanol were added, mixed and the sample loaded onto a column. Genomic DNA was bound to the column by centrifugation (8,000 x *g* for 1 min at RT) and was subsequently washed with 500 µl of Inhibitor Removal Buffer (8,000 x *g* for 1 min at RT). A second wash step with 500 µl of Wash Buffer (8,000 x *g* for 1 min at RT) was performed.



## 2. Materials and Methods

The DNA was dried by full-speed centrifugation (21,000 x *g* for 10 sec at RT) and eluted with 200 µl pre-warmed Elution Buffer (8,000 x *g* for 1 min at RT). DNA concentration was measured with an Infinite M200 plate reader (TECAN).

### 2.2.7 Purification of DNA for diagnostic polymerase chain reaction

The Phusion Human specimen direct PCR kit (Thermo) was used according to the manufacturer's instructions.  $1 \times 10^6$  cells (PCF or BSF) cells were centrifuged (1,500 x *g* for 10 min at RT), the supernatant was discarded and the cell pellet was resolved in 20 µl Dilution Buffer. 1 µl DNARelease was added. The solution was incubated at RT for five minutes and then boiled (98°C for 2 min). Afterwards cell debris was pelleted by centrifugation (2,500 x *g* for 5 min at RT). 1 µl of the supernatant that contained the genomic DNA was subsequently used for a diagnostic PCR. Primers used for diagnostic PCR were synthesized by Sigma-Aldrich (Table 2.3).

### 2.2.8 Luciferase assay

Luciferase assays were performed using the dual luciferase assay system (Promega).  $1 \times 10^6$  cells were centrifuged (1,500 x *g* for 10 min at 4°C) and washed with 1ml ice-cold PBS (10 mM Na<sub>2</sub>HPO<sub>4</sub>, 1.8 mM KH<sub>2</sub>PO<sub>4</sub> pH 7.4, 140 mM NaCl, 2.7 mM KCl). After centrifugation (1,500 x *g* for 10 min at 4°C) the supernatant was discarded and the cell pellet was dissolved in 155 µl Passive Lysis Buffer. The 155 µl were immediately transferred into a transparent flat bottom 96-well plate. 55 µl of Stop&Glow solution were added in the dark and the lid was covered to prevent light incidence. The plate was incubated for 2 minutes at RT. The samples were analysed with an Infinite 200M plate reader (TECAN). The integration time for the measurement was set to 1000 ms.

## 2.3 Flow cytometry analyses

### 2.3.1 Cell cycle analysis

$5 \times 10^6$  BSF cells were centrifuged (1,500 x *g* for 10 min at 4°C) and washed once with 5 ml ice-cold TDB. After subsequent centrifugation (1,500 x *g* for 10 min at 4°C) the cells were resuspended in 1 ml ice cold PBS/2 mM EDTA. The cells were fixed by dropwise adding 2.5 ml ice cold 100 % ethanol or methanol while vortexing the cells at lowest speed possible with subsequent storage at 4°C for 1 h. After a centrifugation step (1,500 x *g* for 10 min at RT) the cells were washed with 1 ml PBS/2 mM EDTA, centrifuged (at RT, 1,500 x *g* for 10 min) and resuspended in 1 ml PBS/EDTA. 1 µl RNase (10 µg/µl) and 10 µl propidiumiodide (1 µg/µl) were added to the cell suspension and incubated for 30 min at 37°C. Samples were stored at 4°C in the dark until analysis with a BD FACSCalibur using the FL-2 detector channel. 50,000 events were analysed with an average flow rate of 1,000 events per second. The resulting cell cycle profiles were then evaluated with the CellQuest Pro software (BD Bioscience).

### 2.3.2 Live-dead analysis

1x10<sup>6</sup> BSF cells were centrifuged (1,500 x *g* for 10 min at 4 °C) and washed twice with 1 ml pre-cooled TDB (stored on ice) with subsequent centrifugation (1,500 x *g* for 10 min at 4 °C). The cells were resuspended in 400 µl TDB and incubated with 1 µl propidium iodide (1 mg/ml with a final concentration of 2.5 µg/ml) for 10 min on ice, in the dark. After the staining, the cells were analysed with the FL-2 detector channel of a BD FACSCalibur. 50,000 events were analysed with an average flow rate of 1,000 events per second. The resulting cell cycle profiles were then evaluated with the CellQuest Pro software (BD Bioscience).

## 2.4 DNA methods and DNA plasmids

### 2.4.1 Polymerase chain reaction

Polymerase chain reaction (PCR) was used to amplify specific DNA sequences from plasmids or from 'genomic DNA' (gDNA) of *T. brucei*. PCR reactions contained 1x HF buffer, 200 µM 'deoxyribonucleoside triphosphates' (dNTPs) (Thermo), 0.5 µM forward (fw) primer, 0.5 µM reverse (rev) primer, 10 ng plasmid DNA or 50 ng gDNA per 50 µl reaction and 0.02 Units (U)/µl Phusion High Fidelity DNA polymerase (Thermo). In case of a primer GC- content above 60 % 1x GC-buffer was used instead of the HF-buffer. For optimizing the amplification 1,5 % were added to the reaction. DMSO binds at the cytosine residue and changes its conformation which makes the DNA more unstable which mainly effects GC-rich primers resulting in a reduction of the annealing temperature. Primer sequences used for construction of the RNAi constructs were downloaded from the online platform 'RNAi' (Redmond *et al.*, 2003). Primers were synthesized by Sigma-Aldrich and are listed in table 2.6 and 2.1. Their 'melting temperature' <sup>TM</sup> and the according annealing temperature was determined using Tm calculator online software from NEB. The software calculated the Tm by utilizing thermodynamic data from Santa Lucia (SantaLucia, 1998) and the salt correction of Owczarzy (Owczarzy *et al.*, 2004). For Phusion DNA Polymerases, the salt correction of Schildkraut (Schildkraut, 1965; Owczarzy *et al.*, 2004) was used. The cycling conditions were optimized for the length of the amplified DNA fragments. Tm was set according to the calculated Tm of the corresponding primer pair resulting in the following PCR protocol: 98 °C/30 sec – 30 cycles [98 °C/20 sec, Tm/20 sec, 72 °C 30 sec/kb] – 72 °C/5 min – 4 °C/hold. Reactions were conducted in the T100 Thermo Cycler (BIO-RAD).

**Table 5: Primers used for *T. brucei* methods**

<b>Name</b>	<b>Sequence 5' to 3'</b>	<b>Purpose</b>
TV067	aaaagactttctcttcggtgaggcagaccgagggcgagg atgtcttacggtgcgagtgatgagaat <b>tgaggaagagcag</b> GGTACCGGGCCCCCCTCGAG	forward primer for PCR-based C-terminal HA in situ tagging of RuvB2 (Tb927.4.2000) with pMOTag vector series
TV068	aagctctatatttatctttttctttttatgtgctgccatttactca cgccaccctctttccctggagaaaaggaaacaTGGCGGC CGCTCTAGAACTAGTGGAT	reverse primer for PCR-based C-terminal HA in situ tagging of RuvB2 (Tb927.4.2000) with pMOTag vector series
TV160	gtcagctagcttaagcttgggctagaactagtgATGATTG AACAAATGGATTGCA	forward primer for amplification of a <i>NEO</i> ORF from the pC-PTP-neo vector for subsequent cloning into the pMOTag vector; contains a HindIII restriction site; suitable for <i>in vivo</i> cloning
TV161	tactggatccgcttcggtcggcatctactTCAGAAGAA CTCGTCAAGAAGG	reverse primer for amplification of a <i>NEO</i> ORF from the pC-PTP-neo vector for subsequent cloning into the pMOTag vector; contains a BamHI restriction site; suitable for <i>in vivo</i> cloning
TV162	gtcagctagcttaagcttgggctagaactagtgATGCCTT TGTCTCAAGAAGAATC	forward primer for amplification of a <i>BSD</i> ORF from the pPOT vector for subsequent into the pMOTag vector; contains a HindIII restriction site, suitable for <i>in vivo</i> cloning
TV163	tactggatccgcttcggtcggcatctactTTAGCCCTCC CACACATAAC	reverse primer for amplification of a <i>BSD</i> ORF from the pC-PTP-neo vector for subsequent cloning into the pPOT vector; contains a BamHI restriction site, suitable for <i>in vivo</i> cloning
TV164	gtcagctagcttaagcttgggctagaactagtgATGAAA AAGCCTGAACTCACC	forward primer for amplification of a <i>NEO</i> resistance cassette from the pC-PTP-neo vector for subsequent into the pMOTag vector; contains a HindIII restriction site, suitable for <i>in vivo</i> cloning
TV165	tactggatccgcttcggtcggcatctactTTATTCCTTT GCCCTCGG	reverse primer for amplification of a <i>NEO</i> ORF from the pC-PTP-neo vector for subsequent cloning into the pMOTag vector; contains a BamHI restriction site, suitable for <i>in vivo</i> cloning
TV178	tgccaggtttgctgggaggtgtggtactctcaaggcctca cttaaaaacttgcgagtgacgcttctcaatgaaataGGT ACCGGGCCCCCCTCGAG	forward primer for PCR-based C-terminal HA in situ tagging of <i>TbSWRC4</i> (Tb927.7.4040) with the pMOTag vector series
TV179	cgacgatagtacaatatgtgccagactgagagatgctgg ccaacaaatgggtaaataggagtgggtgagcagcaatccT GGCGGCCGCTCTAGAACTAGTGGAT	reverse primer for PCR-based C-terminal HA in situ tagging of <i>TbSWRC4</i> (Tb927.7.4040) with the pMOTag vector series

## 2. Materials and Methods

<b>Name</b>	<b>Sequence 5' to 3'</b>	<b>Purpose</b>
TV184	cagatcaagcttATGTACCCTTACGATGTGCCT	forward primer for amplification of the 3xHA-Tag from the pMOTag plasmid for subsequent ligation into the pPOT vector; contains a HindIII restriction site
TV185	atgactggatcccgcgtaactctggcacgctg <b>gtacgggtaa</b> <b>gcatagtccggaacatcgtatggata</b> TGCGTAATCG GGCACATC	reverse primer for amplification of the 3xHA-Tag from the pMOTag plasmid for subsequent ligation into the pPOT vector; contains a BamHI restriction site
TV190	tgcggtgctgtgatagtggggaacctaggggcaaggtgc tggaatagagggaacgctgtgtggtgtagtgaagtca GTATAATGCAGACCTGCTGC	forward primer for PCR-based N-terminal HA in situ tagging of <i>TbSWRC1</i> (Tb927.10.11690) with the pPOT vector series
TV191	gtggcacctgtgtaaactacatattcgtgtccccgtgatgg aaacagaagatccagtagattagaagtgccctgc <b>atctc</b> <b>agggggggcccgtacc</b> CGCGTAATCTGGCACGT C	reverse primer for PCR-based N-terminal HA in situ tagging of <i>TbSWRC1</i> (Tb927.10.11690) with the pPOT vector series
TV192	caaaatttatctaccatttctcacgctagttacttcattaccgtg agccttcagactggtcgtgcacacttcaactGTATAAT GCAGACCTGCTGC	forward primer for PCR-based N-terminal HA in situ tagging of <i>TbSWRC2</i> (Tb927.11.5830) with the pPOT vector series
TV193	ttttggagcatattgccacggttgccacgcccggcggtc gtcagcgacattccagcgggagtcgtcaacaccgtccat <b>ctc</b> <b>gagggggggcccgtacc</b> CGCGTAATCTGGCACG TC	reverse primer for PCR-based N-terminal HA in situ tagging of <i>TbSWRC2</i> (Tb927.11.5830) with the pPOT vector series
TV193a	ttttggagcatattgccacggttgccacgcccggcggtc gtcagcgacattccagcgggagtcgtcaacaccgtccatG GATCCTTCAAGAGGTGGTC	reverse primer for PCR-based in situ tagging with the pPOT vector series for generation of a TBL- <i>TbSWRC2</i> (Tb927.11.5830) fusion protein
TV194	tgcggtgctgtgatagtggggaacctaggggcaaggtgc tggaatagagggaacgctgtgtggtgtagtgaagtca ATGATTGAACAAGATGGATTGCA	forward primer for the amplification of the <i>NEO</i> ORF from pC-PTP neo or pMOTag 3H with 80 bp of the 5'UTR of <i>TbSWRC1</i> (Tb927.10.11690) for recombination-based KO approach
TV195	acggtaaatgccgctcgttggttaacgcccttcaacatccta aacctgtaacctgaccacacgcaaggaacacactccTCA GAAGAACTCGTCAAGAAGG	reverse primer for the amplification of the <i>NEO</i> ORF from pC-PTP neo or pMOTag 3H with 80 bp of the 5'UTR of <i>TbSWRC1</i> (Tb927.10.11690) for recombination-based KO approach
TV196	caaaatttatctaccatttctcacgctagttacttcattaccgtg agccttcagactggtcgtgcacacttcaactATGATTG ACAAGATGGATTGCA	forward primer for the amplification of the <i>NEO</i> ORF from pC-PTP neo or pMOTag 3H with 80 bp of the 5'UTR of <i>TbSWRC2</i> (Tb927.11.5830) for recombination-based KO approach
TV197	cagcctccagatcgatgtgaatgcaccctctgtcaacgatct gtgttactacacgaacatcatcaacatcaccggaTCA GAAGAACTCGTCAAGAAGG	reverse primer for the amplification of the <i>NEO</i> ORF from pC-PTP neo or pMOTag 3H with 80 bp of the 5'UTR of <i>TbSWRC2</i> (Tb927.11.5830) for recombination-based KO approach

## 2. Materials and Methods

Name	Sequence 5' to 3'	Purpose
TV206	attatacacggacacgaacgaataaaaactcactggtttattc agttgtacacagtaaccacatttaaagcacttgcgtcGTAT AATGCAGACCTGCTGC	forward primer for PCR-based N-terminal Ty1 in situ tagging of H2A.Z with the pPOT vector series
TV207	agggcactcgctgctcgggactcatgccacgccgcccgcac gaggggggctgagggcaccgcatcatcacctgtaagagac <b>tcgaggggggcccgtacc</b> CCTTGGGTCAAGTGG GTC	reverse primer for PCR-based N-terminal Ty1 in situ tagging of H2A.Z with the pPOT vector series
TV212	tctggtgaacgctgagcgttgccacagtgggtaatagactg taagaactcacacgactttagaggcgcgagctggttgcGT ATAATGCAGACCTGCTGC	forward primer for PCR-based N-terminal HA in situ tagging of <i>TbEaf6</i> (Tb927.9.2910) with the pPOT vector series
TV213	agagtggcactggcttttacttccctttgctccctgggggcg ccgtaccctctaccccgacctgacccctaccggactc <b>ga</b> <b>gggggggcccgtacc</b> CGCGTAATCTGGCACGTC	reverse primer for PCR-based N-terminal HA in situ tagging of <i>TbEaf6</i> (Tb927.9.2910) with the pPOT vector series
TV218	cccaccggcaacatagctgtagcggttttgtgatcggaggtt gagagtctgttgcctagtggatgagaggttgcgcaGT ATAATGCAGACCTGCTGC	forward primer for PCR-based N-terminal HA in situ tagging of HAT2 (Tb927.11.11530) with the pPOT vector series
TV219	aagagtgttgagctgatccttctcagacgggccccca ttgatgtcacctcgtctttttgtgctaacgacg <b>cctcgag</b> <b>gggggggcccgtacc</b> CGCGTAATCTGGCACGTC	reverse primer for PCR-based N-terminal HA in situ tagging of HAT2 (Tb927.11.11530) with the pPOT vector series
TV228	ggttct <b>aa</b> gcttATGGAGGTCCATACTAACCAGGA C	forward primer for amplification of the 2xTy1-Tag from the pMOTag 2T plasmid for subsequent ligation into the pPOT vector; contains a HindIII restriction site
TV229	cagtccgatcc <b>ttcaagaggtggtcctgtaccgtcaagt</b> <b>ggtcctggttagtatggacctcccttg</b> GGTCAAGTGG GTCCTGGTTAG	reverse primer for amplification of the 2xTy1-Tag from the pMOTag 2T plasmid for subsequent ligation into the pPOT vector; contains a BamHI restriction site
TV230	cccaccggcaacatagctgtagcggttttgtgatcggaggtt gagagtctgttgcctagtggatgagaggttgcgcaAT GATTGAACAAGATGGATTGCA	forward primer for the amplification of the <i>NEO</i> ORF from pC-PTP neo or pMOTag 3H with 80 bp of the 5'UTR of HAT2 (Tb927.11.11530) for recombination-based KO approach
TV231	acagaaaccgcagaagcactgactctgggaaaggttaagtt agaaataaaagaccatttatattgtggagggacgacacaT CAGAAGAAGCTCGTCAAGAAGG	reverse primer for the amplification of the <i>NEO</i> ORF from pC-PTP neo or pMOTag 3H with 80 bp of the 5'UTR of HAT2 (Tb927.11.11530) for recombination-based KO approach
TV232	tctggtgaacgctgagcgttgccacagtgggtaatagactg taagaactcacacgactttagaggcgcgagctggttgcAT GATTGAACAAGATGGATTGCA	forward primer for the amplification of the <i>NEO</i> ORF from pC-PTP neo or pMOTag 3H with 80 bp of the 5'UTR of <i>TbEaf6</i> (Tb927.9.2910) for recombination-based KO approach

## 2. Materials and Methods

Name	Sequence 5' to 3'	Purpose
TV233	caacgtgcatcaactggtccaggaacgaaaacggt taacacgccgagacatgtcctttcatcgccacgaccataTCA GAAGAACTCGTCAAGAAGG	reverse primer for the amplification of the <i>NEO</i> ORF from pC-PTP neo or pMOTag 3H with 80 bp of the 5'UTR of <i>TbEaf6</i> (Tb927.9.2910) for recombination-based KO approach
TV247	gactgaattcATGAAAAGCCTGAACTCACC	forward primer for amplification of a <i>HYG</i> ORF from the pMOTag 4H vector for subsequent cloning into the pPOTv7 vector; contains an <i>EcoRI</i> restriction site
TV248	gactaccatggTTATTCCTTTGCCCTCGG	reverse primer for amplification of a <i>HYG</i> ORF from the pMOTag 4H vector for subsequent cloning into the pPOTv7 vector; contains a <i>NcoI</i> restriction site
TV259a	gactattcatcggtttatattagcaacagtaggtactagcacc actaacaacaacaacaagcacttctattttatcATGATT GAACAAGATGGATTGCA	forward primer for amplification of a luciferase reporter construct from pFG14 for subsequent integration into the tubulin locus; primer consists of the last 80bp of the 5'UTR of $\beta$ -Tubulin
TV260a	gagacagaaaacaattcacaaaagaagaagaagaac ataaatgaaaacctacacatggtgacgtgtcacactttTT ATTGTTCATTTTGAGAACTCGC	reverse primer for amplification of a luciferase reporter construct from pFG14 for subsequent integration into the tubulin locus; primer consists of the first 80bp of the 3'UTR of $\beta$ -Tubulin
TV261	gactgaattcATGACCGAGTACAAGCCCACGG	forward primer for amplification of a <i>PAC</i> ORF from the pMOTag 2H vector for subsequent cloning into the pPOTv7 vector; contains an <i>EcoRI</i> restriction site
TV262	gactaccatggTCAGGCACCGGGCTTGCG	reverse primer for amplification of a <i>PAC</i> ORF from the pMOTag 2H vector for subsequent cloning into the pPOTv7 vector; contains a <i>NcoI</i> restriction site
TV300	tggttccggttccggttctaagcttATGAAGGACAACA CGGTGCC	forward primer for amplification of a TBL biotin ligase for subsequent cloning into the pPOTv4 vector; contains an <i>EcoRI</i> restriction site; suitable for in vivo cloning
TV301	ctgatccagatcctgatccggatcctcaagaggtggtcct <b>gtacc</b> TGCGTAGTCGGGCACGTC	reverse primer for amplification of a TBL biotin ligase for subsequent cloning into the pPOTv4 vector; contains an <i>EcoRI</i> restriction site; suitable for in vivo cloning
TV328	ctccctttttcccccttccctccttttaagtgttattcactt agccattgttcttttaccgcggtgcaggtgaGTATAATG CAGACCTGCTGC	forward primer for PCR-based N-terminal Ty1 in situ tagging of <i>TbEaf3</i> (Tb927.1.650) with the pPOT vector series
TV329a	cgcggttcatgggatgatccactgctcaaatgattgcgc aatatttgcagcaaccgaggggcccactgtatctcgtcctcg <b>agggggggcccggtacc</b> CCTTGGGTCAAGTGGT C	reverse primer for PCR-based N-terminal Ty1 in situ tagging of <i>TbEaf3</i> (Tb927.1.650) with the pPOT vector series

## 2. Materials and Methods

Name	Sequence 5' to 3'	Purpose
TV346	gagagcaaaaaaagaattgtgtaccgaagtcacacattg agaaaacctagtggaaccagtaagtattgcgcgcatcgC GTATAATGCAGACCTGCTGC	forward primer for PCR-based N-terminal Ty1 in situ tagging of Bdf3 (Tb927.11.10070) with the pPOT vector series
TV347	agtggatggctgagtgccagtcgatgtacctgatccaata gtttatatcttcagcaagtttgaccctgtgaggaccatCCTT GGGTCAAGTGGGTC	reverse primer for PCR-based N-terminal Ty1 in situ tagging of Bdf3 (Tb927.11.10070) with the pPOT vector series

Sequences written in capital letters indicate primer annealing sites, small letters indicate sequences for homologous recombination, underlined and small letters indicate restriction enzyme cut sites and sequences in bold and italic represent linker sequences

**Table 6: Primers used for diagnostic PCRs**

Name	Sequence 5' to 3'	Purpose
TV103	CATATTTGTTTCCCCCTACACG	anneals in the 5'UTR of RuvB2 (Tb927.4.2000)
TV104	CTTTTTATGTGCTGCCATTCTAC	anneals in the 3'UTR of RuvB2 (Tb927.4.2000)
TV168	GTGCTGGAATAGAGGGAGAA	anneals in the 5'UTR of <i>TbSWRC1</i> (Tb927.10.11690)
TV169	CAACATCCTAACCTGTAACCA	anneals in the 3'UTR of <i>TbSWRC1</i> (Tb927.10.11690)
TV170	CATTACCGTGTAGCCTTTCAG	anneals in the 5'UTR of <i>TbSWRC2</i> (Tb927.11.5830)
TV171	CATGTTGATAATGCGCCAGC	anneals in the 3'UTR of <i>TbSWRC2</i> (Tb927.11.5830)
TV182	GGATGAATGTGTAAGTAGGTTGC	anneals in the 5'UTR of <i>TbSWRC4</i> (Tb927.7.4040)
TV183	GTGCCAGACTTGAGAGATG	anneals in the 3'UTR of <i>TbSWRC4</i> (Tb927.7.4040)
TV220	CGGACACGAACGAATAAAAACT	anneals in the 5'UTR of H2A.Z
TV221	CGGGGCCCTTCTCTATTA	anneals in the 3'UTR of H2A.Z
TV224	CCGCCTTCATTCACTTCTCC	anneals in the 5'UTR of HAT2 (Tb927.11.11530)
TV225	GAGTAAGACAGAAACCGCAGA	anneals in the 3'UTR of HAT2 (Tb927.11.11530)
TV226	CCGTCGTAAGGTAGTTCAGAT	anneals in the 5'UTR of <i>TbEaf6</i> (Tb927.9.2910)
TV227	GTTAAAAGGTGGAATGCCCA	anneals in the 3'UTR of <i>TbEaf6</i> (Tb927.9.2910)
TV330	GGCTCCCACTCTAGTTGT	anneals in the 5'UTR of <i>TbEaf3</i> (Tb927.1.650)
TV331	CGACAGCAAACAGTACCG	anneals in the 3'UTR of <i>TbEaf3</i> (Tb927.1.650)
TV348	AGAAAACCTAGTGGAACCCAGT	anneals in the 5'UTR of Bdf3 (Tb927.11.10070)
TV349	AACCATGCCGCTGTCTACA	anneals in the 3'UTR of Bdf3 (Tb927.11.10070)

Sequences written in capital letters indicate primer annealing sites.

### 2.4.2 Agarose gel electrophoresis

For visualization and purification DNA fragments were separated according to their length by agarose gel electrophoresis. Depending on the length of the DNA fragments that ought to be analysed 0.5 %, 1 % or 1.5 % agarose were solved in 'Tris-acetate- ethylenediaminetetraacetic acid' (EDTA)' (TAE) buffer (40 mM Tris-HCl pH 8.0, 40 mM NaOAc, 1 mM EDTA). The gel solution was either supplemented with 0.1 µg/ml ethidium bromide or SYBR Safe DNA Gel Stain (1:50,000) (Thermo). DNA samples were mixed with DNA loading dye (Thermo) and were loaded with the GeneRuler DNA Ladder Mix (Thermo) to assess the length of the DNA fragments. Electrophoresis was carried out 120 Volt (V) for 20-40 min depending on the migration of the tracking dye.

## 2. Materials and Methods

For gels with a higher agarose content more time was required to separate the fragments. DNA stained with ethidium bromide was visualized by ultraviolet (UV) light with a Gel iX Imager (Intas). DNA stained with SYBR Safe was visualized by UV light with the iBright Imaging System (Thermo).

### 2.4.3 Restriction digest

DNA from cleaned-up PCR samples or purified plasmid DNA was digested with restriction endonucleases according to the manufacturer's instructions (Thermo).

### 2.4.4 Measurement of DNA concentration

Concentration and purity of DNA was determined by measuring the absorption at 260 nm and 280 nm with the InfiniteM200 Reader (TECAN).

### 2.4.5 Recombination cloning

For construction of the pGL2084 RNAi plasmids the Gateway BP Clonase II Enzyme mix (Thermo) was used according to the manufacturer's instructions. In brief 100ng of the PCR product that was amplified with the corresponding primer pairs (Table 7) was added to 150ng of the pGL2084 plasmid. 1 µl of the BP clonase was added and the reaction was filled up to 5 µl with Tris/EDTA buffer (TE-Buffer; 1mM EDTA, 10mM Tris-HCL ph 8.0). The reaction was incubated (25°C, 1 hour) and 2,5 µl were transformed into competent *E. coli* cells.

**Table 7: Primers used for RNAi constructs**

Name	Sequence 5' to 3'	Purpose
TV152	ggggacaagtttgatacaaaaaagcaggctTGCTTCAGCTT CCCAGACAG	forward primer for amplification of <i>TbSWR1</i> RNAi target (bp 2476 – 2876 of ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV153	ggggaccactttgtacaagaagctgggtGTCACCGGCTT TGTTTGCC	reverse primer for amplification of <i>TbSWR1</i> RNAi target (bp 2476 – 2876 of ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV158	ggggacaagtttgatacaaaaaagcaggctTCAGTACTGGT CGTGACAC	forward primer for amplification of <i>TbSWRC2</i> RNAi target (bp -28 – 528 of the 5' UTR and ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV159	ggggaccactttgtacaagaagctgggtCCGTTTCTTTG CAGCTGCTT	reverse primer for amplification of <i>TbSWRC2</i> RNAi target (bp -28 – 528 of the 5' UTR and ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV249	ggggacaagtttgatacaaaaaagcaggctCCCACCACAG TGCGATACTT	forward primer for amplification of <i>HAT2</i> RNAi target (bp 499 – 1063 of ORF) from gDNA; primer with attB1 site for recombination into pGL2084



## 2. Materials and Methods

<b>Name</b>	<b>Sequence 5' to 3'</b>	<b>Purpose</b>
TV250	ggggaccactttgtacaagaaagctgggtAAGTGACGAT GCATGCCAGA	reverse primer for amplification of HAT2 RNAi target (bp 499 – 1063 of ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV251c	ggggacaagtttgtacaaaaaagcaggctAGCGAGGAGG TAAACTGGC	forward primer for amplification of H2A.Z RNAi target (bp 155 – 504 of ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV252c	ggggaccactttgtacaagaaagctgggtTTCCAGGCTCT TGTGCACAA	reverse primer for amplification of H2A.Z RNAi target (bp 155 – 504 of ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV257	ggggacaagtttgtacaaaaaagcaggctGAAGAAGCTG AAGATCAGCT	forward primer for amplification of <i>Tb</i> SWRC1 RNAi target (bp -99 – 502 of the 5' UTR and ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV258	ggggaccactttgtacaagaaagctgggtGAAAGCCGTTA GGAAGAGT	reverse primer for amplification of <i>Tb</i> SWRC1 RNAi target (bp -99 – 502 of the 5' UTR and ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV282	ggggacaagtttgtacaaaaaagcaggctAACCGAAGCTT TAGAGGCC	forward primer for amplification of the 1 <sup>st</sup> <i>Tb</i> SWRC3 RNAi target (bp 875 – 1467 of ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV283	ggggaccactttgtacaagaaagctgggtCGTCGTATTCC TCGCCGTAA	reverse primer for amplification of the 1 <sup>st</sup> <i>Tb</i> SWRC3 RNAi target (bp 875 – 1467 of ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV284	ggggacaagtttgtacaaaaaagcaggctCGAAGGAGCT GACTCGATCC	forward primer for amplification of RPB1 RNAi target (bp 3804 – 4259 of ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV285	ggggaccactttgtacaagaaagctgggtGACGTATCGG AGCGGTTGAT	reverse primer for amplification of RPB1 RNAi target (bp 3804 – 4259 of ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV286	ggggacaagtttgtacaaaaaagcaggctGACAGGGGAG ATGGGAGAGT	forward primer for amplification of the 2 <sup>nd</sup> <i>Tb</i> SWRC3 RNAi target (bp 93 - 612 of ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV287	ggggaccactttgtacaagaaagctgggtACGTTCTGTTCCG TACCATGCT	reverse primer for amplification of the 2 <sup>nd</sup> <i>Tb</i> SWRC3 RNAi target (bp 93 - 612 of ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV292	ggggacaagtttgtacaaaaaagcaggctCCACGTCCGTC TAGTTTGCT	forward primer for amplification of HAT1 RNAi target (bp 648 – 1184 of ORF) from gDNA; primer with attB1 site for recombination into pGL2084

## 2. Materials and Methods

<b>Name</b>	<b>Sequence 5' to 3'</b>	<b>Purpose</b>
TV293	ggggaccactttgtacaagaaagctgggtGTTCGCATGAC ATCCGCTTC	reverse primer for amplification of HAT1 RNAi target (bp 648 – 1184 of ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV316	ggggacaagtttgtacaaaaagcaggctGTTTGTGATGG GCACCTTG	forward primer for amplification of the 3 <sup>rd</sup> <i>Tb</i> SWRC3 RNAi target (bp 1488 – 1988 of ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV317	ggggaccactttgtacaagaaagctgggtGTGGCATGAA CCACAGCG	reverse primer for amplification of the 3 <sup>rd</sup> <i>Tb</i> SWRC3 RNAi target (bp 1488 – 1988 of ORF) from gDNA; primer with attB2 site for recombination into pGL2084
TV344	ggggacaagtttgtacaaaaagcaggctCACTCAGCCAT CCACTACCG	forward primer for amplification of Bdf3 RNAi target (bp 65 - 565 of the ORF) from gDNA; primer with attB1 site for recombination into pGL2084
TV345	ggggaccactttgtacaagaaagctgggtTGAGTGTGGG TCTTCACGG	reverse primer for amplification of Bdf3 RNAi target (bp 65 - 565 of the ORF) from gDNA; primer with attB2 site for recombination into pGL2084

Primer annealing sites are in upper case, sequences for homologous recombination are in lower case.

### 2.4.6 Ligation

Ligation reactions were prepared with a 3:1 molar ratio of insert DNA over linearized vector DNA. Ratio was calculated with the NEBio Calculator web tool. Reactions were carried out in a final volume of 20 µl. T4 DNA ligase Buffer and T4 DNA ligase (Thermo) were used according to the manufacturer instructions. 10 µl of the reactions were used for transformation of *E. coli*.

### 2.4.7 *In vivo* Ligation

Ligation reactions were prepared with a 3:1 molar ratio of insert DNA over linearized vector DNA. Ratio was calculated with the NEBio Calculator web tool. Reactions were carried out in a final volume of 20 µl. T4 DNA ligase Buffer and T4 DNA ligase (Thermo) were used according to the manufacturer instructions. 10 µl of the reactions were used for transformation of *E. coli*.

### 2.4.8 DNA sequencing

The company GATC sequenced plasmid DNA, which was prepared as instructed by the company. Briefly, 20 µl of 100 ng/µl purified plasmid DNA and 20 µl of 10 µM sequencing primer (Table 8) were sent to the company. Sequencing results were analysed with the CLC workbench software (CLC bio, Quiagen).

**Table 8: Primers used for sequencing**

Name	Sequence 5' to 3'	Purpose
TV202a	CAAGTATACCAACAAGCCCG	sequencing of the tag region of pPOTv7 and pPOTv4 plasmids
MBS34	TAAAAGTAGCGCTTACGG	RNAi construct sequencing forward
MBS35	TGCCTGCACTAACACTAC	RNAi construct sequencing reverse

### 2.4.9 Plasmids used and generated in course of the project

**Table 9: Used and generated plasmids**

Plasmid name	Description	Selection	Reference
pC-PTP-neo	plasmid contains a C-terminal tandem affinity purification tag within a construct designed for genome integration into trypanosomes.	amp ( <i>E. coli</i> ) G418/Neo ( <i>T. brucei</i> )	Schimanski <i>et al.</i> , 2005b
pFG14	The original pFG14 (received from L. Figueiredo) encodes a Renilla luciferase reporter gene and contains a hygromycin and neomycin resistance cassette. Hygromycin resistance was removed by Helena Reis. The plasmid was used to amplify the Luciferase ORF and the neomycin resistance for cloning the construct into the Tubulin locus as a reporter for transcriptional activity.	amp ( <i>E. coli</i> ) G418/Neo ( <i>T. brucei</i> )	H. Reis

## 2. Materials and Methods

Plasmid name	Description	Selection	Reference
pGL2084	Tetracycline-inducible, stem-loop RNAi vector. RNAi target fragments can be cloned into pGL2084 in a single Gateway BP reaction. After Sgsl linearization, constructs have terminal hygΔstop and RRNA spacer sequences for targeting the hygΔstart::PAC::RRNA locus of 2T1 cells (Alsford and Horn 2008).	amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	Jones <i>et al.</i> 2014
pGL2084_ <i>Tb</i> SWR1	401 bp of the <i>Tb</i> SWR1 ORF (bp 2476 – 2876) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV152 /TV153 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>Sgsl</i> // <i>Ascl</i> .	amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer
pGL2084_ <i>Tb</i> SWRC1	601 bp of the <i>Tb</i> SWRC1 ORF (bp -99 – 502) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV257 / TV258 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>Sgsl</i> // <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer
pGL2084_ <i>Tb</i> SWRC2	556 bp of the <i>Tb</i> SWRC2 ORF (bp -28 – 528) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV158 / TV159 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>Sgsl</i> // <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer
pGL2084_ <i>Tb</i> SWRC3_1	593 bp of the <i>Tb</i> SWRC3 ORF (bp 875 - 1467) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV282 / TV283 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>Sgsl</i> // <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer
pGL2084_ <i>Tb</i> SWRC3_2	520 bp of the <i>Tb</i> SWRC3 ORF (bp 93 - 612) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV286 / TV287 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>Sgsl</i> // <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	K. Köthe
pGL2084_ <i>Tb</i> SWRC3_3	501 bp of the <i>Tb</i> SWRC3 ORF (bp 1488 - 1988) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV316/TV317 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>Sgsl</i> // <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	K. Köthe
pGL2084_ HAT1	537 bp of the HAT1 ORF (bp 648 – 1184) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV292 / TV293 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>Sgsl</i> // <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer
pGL2084_ HAT2	565 bp of the HAT2 ORF (bp 499 – 1063) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV249 / TV250 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>Sgsl</i> // <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer

## 2. Materials and Methods

Plasmid name	Description	Selection	Reference
pGL2084_H2A.Z	352 bp of the H2A.Z ORF (bp 155 – 504) were used as RNAi target sequence and was amplified from genomic DNA with primer pair TV251c /TV252c containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>SgsI</i> / <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer
pGL2084_RPB1	456 bp of the RPB1 ORF (bp 3804 – 4259) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV284 / TV285 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>SgsI</i> / <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer
pGL2084_Bdf3	501 bp of the Bdf3 ORF (bp 65 - 565) were used as RNAi target sequence and were amplified from genomic DNA with primer pair TV344 / TV345 containing <i>attB</i> sites for BP-cloning into pGL2084. Linearization with <i>SgsI</i> / <i>Ascl</i> .	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	T. Vellmer
pLEW100v5	Dual promoter expression vector. T7 promoter drives BLE and tetracycline-inducible RRNA promoter drives ectopic expression of GOI. Integration into the RRNA locus.	Amp ( <i>E. coli</i> ) Phleo ( <i>T. brucei</i> )	G. Cross
pLew100_Myc_BirA*	Dual promoter expression vector. T7 promoter drives BLE and tetracycline-inducible GPEET promoter drives ectopic expression of Myc tag and modified bacterial biotin ligase BirA*. Integration into the RRNA locus	Amp ( <i>E. coli</i> ) Phleo ( <i>T. brucei</i> )	Morriswood <i>et al.</i> 2013
pLEW100_DOT1B:Myc:BirA*	DOT1A ORF was amplified from genomic DNA with NE112/NE113 primer pair containing HindIII and NdeI restriction enzyme sites for cloning into pLew100_Myc_BirA*. Linearization with NotI	Amp ( <i>E. coli</i> ) Phleo ( <i>T. brucei</i> )	T. Vellmer
pMOTag 2T	Plasmid for PCR-based Ty1 C-terminal in situ tagging of <i>T. brucei</i> genes	Amp ( <i>E. coli</i> ) Phleo ( <i>T. brucei</i> )	Oberholzer <i>et al.</i> 2006
pMOTag 2H	Plasmid for PCR-based 3'-end in situ tagging of GOI with 3xHA. The plasmid was used for amplification of the 3xHA tag to construct the pPOTv7_B/B_HA. 3xHA sequence was amplified with the primers TV xxx /TV xxx which contain a <i>BamHI</i> and <i>HindIII</i> cut site for subsequent cloning into pPOTv7.	Amp ( <i>E. coli</i> ) Puro ( <i>T. brucei</i> )	Oberholzer <i>et al.</i> 2006
pMOTag 3H	Plasmid for PCR-based 3'-end in situ tagging of GOI with 3xHA. Plasmid was used as template for amplification of the neomycin ORF with TV230 / TV231; TV232 / TV233 primer pair, containing 80 bp of the <i>TbSWRC1</i> ; <i>TbSWRC2</i> UTR for a recombination-based KO approach.	amp ( <i>E. coli</i> ) G418/Neo ( <i>T. brucei</i> )	Oberholzer <i>et al.</i> 2006

## 2. Materials and Methods

Plasmid name	Description	Selection	Reference
pMOTag 4H	Plasmid for PCR-based 3'-end in situ tagging of GOI with 3xHA. The plasmid was used as template for amplification of the HYG ORF with TV247 / TV248 primer pair, containing an <i>EcoRI</i> and <i>NcoI</i> cut site for ligation into pPOTv7_B/B_HA	Amp ( <i>E. coli</i> ) Hygro ( <i>T. brucei</i> )	Oberholzer <i>et al.</i> 2006
pPOTv4	Plasmid for PCR-based 5'- or 3'-end in situ tagging of GOI with eYFP.	Amp ( <i>E. coli</i> ) Blas (N-ter) hygro (C-ter.) ( <i>T. brucei</i> )	Dean <i>et al.</i> 2015
pPOTv4_B/H_TBL	Plasmid for PCR-based 5'- or 3'-end in situ tagging of GOI with a codon optimized turbo biotin ligase (TBL). Plasmid was constructed by removal of the eYFP ORF with <i>BamHI</i> and <i>HindIII</i> and subsequent <i>in vivo</i> ligation of the TBL ORF which was amplified via PCR. For <i>in vivo</i> ligation the 5' end of the TBL PCR product contained 25 bp homologous to the 5' region of the <i>BamHI</i> cut site and the 3' end contained 25 bp homologous to the 3' region of the <i>HindIII</i> cut site	Amp ( <i>E. coli</i> ) Blas (N-ter) Hygro (C-ter.) ( <i>T. brucei</i> )	K. Köthe
pPOTv7	Plasmid for PCR-based 5'- or 3'-end in situ tagging of GOI with eYFP.	Amp ( <i>E. coli</i> ) Hygro (N-ter) Blas (C-ter.) ( <i>T. brucei</i> )	Dean <i>et al.</i> 2015
pPOTv7_B/B_HA	Plasmid for PCR-based 5'- or 3'-end in situ tagging of GOI with an 3xHA Tag. Plasmid was constructed by removal of the eYFP ORF with <i>BamHI</i> and <i>HindIII</i> and subsequent ligation of a 3xHA ORF which was amplified via PCR from pMOTag2H. Primers contained <i>BamHI</i> and <i>HindIII</i> for subsequent cloning into pPOTv7.	amp ( <i>E. coli</i> ) Hygro (N-ter) Blas (C-ter.) ( <i>T. brucei</i> )	T. Vellmer
pPOTv7_H/B_HA	Plasmid for PCR-based 5'- or 3'-end in situ tagging of GOI with an 3xHA Tag. The blasticidin resistance cassette was replaced with a hygromycin resistance cassette via cut ligation with <i>EcoRI</i> and <i>NcoI</i>	Amp ( <i>E. coli</i> ) Hygro (N-ter) Blas (C-ter.) ( <i>T. brucei</i> )	T. Vellmer
pPOTv7_B/B_Ty1	Plasmid for PCR-based 5'- or 3'-end in situ tagging of GOI with an 2xTy1 Tag. Plasmid was constructed by removal of the eYFP ORF with <i>BamHI</i> and <i>HindIII</i> and subsequent ligation of a 2xTy1 ORF which was amplified via PCR from pMOTag2H. Primers contained <i>BamHI</i> and <i>HindIII</i> for subsequent cloning into pPOTv7.	Amp ( <i>E. coli</i> ) Hygro (N-ter) Blas (C-ter.) ( <i>T. brucei</i> )	T. Vellmer

Abbreviations: Blas (blasticidin), BLE (bleomycin resistance gene), BSD (blasticidin-S deaminase resistance gene), eYFP (enhanced yellow fluorescent protein), G418/Neo (neomycin), HA (hemagglutinin), HYG (hygromycin phosphotransferase resistance gene), Hygro (hygromycin), NEO (aminoglycoside phosphotransferase resistance gene), PAC (puromycin N-acetyltransferase resistance gene), Phleo (phleomycin), Puro (puromycin)

## 2.5 RNA methods

### 2.5.1 Northern Blot analysis

Northern blots were done as previously described in Kramer *et al.*, 2008. mRNA was prepared with the RNeasy kit (Qiagen). 18S rRNA and 5.8S rRNA were detected with antisense oligos coupled to IRDye 800. Total mRNA and SL RNA were detected by an oligo antisense to the mini-exon sequence coupled to IRDye 700 (Table 10). Blot images were obtained with the Odyssey Infrared Imaging System (LI-COR Biosciences) and quantified with the Image Studio Lite Software. Background subtraction was performed by defining areas for normalization above and below the measured area (for 5.8S rRNA) or manually by defining a square in between the lanes (total mRNA).

**Table 10: DNA probes used for Northern Blot analysis**

Name	Sequence 5' to 3'
18S rRNA antisense oligonucleotide	CCTTCGCTGTAGTTCGTCTTGGTGCGGTCTAAGAATTC
5.8S rRNA antisense oligonucleotide	ACTTTGCTGCGTCTTCAACGAAATAGGAAGCCAAGTC
mini-exon antisense oligonucleotide	CAATATAGTACAGAAACTGTTCTAATAATAGCGTT

## 2.6 Protein methods

### 2.6.1 Preparation of whole cell lysates for sodium dodecyl sulfate polyacrylamide gel electrophoresis

Cells were harvested by centrifugation (1,500 x *g* for 10 min at 4°C) and washed with 1 ml ice-cold PBS (PCF) or TDB (BSF) (1,500 x *g* for 10 min at 4°C). Cell pellets were resuspended in 'sodium dodecyl sulfate' (SDS) loading buffer (126 mM Tris-HCl pH 6.8, 20 % glycerol, 4 % SDS, 0.02 % bromophenol blue, 60 mM dithiothreitol (DTT)) to a final concentration of 2x10<sup>5</sup> cells/μl. After boiling the samples for 10 min at 98 °C, they were either used directly for SDS-'polyacrylamide gel electrophoresis' (PAGE) or stored at -20°C.

### 2.6.2 SDS-PAGE

SDS-PAGE was performed to separate proteins according to their molecular weight. SDS gels were hand-casted with the Mini-PROTEAN Tetra Cell Casting Stand. SDS gels consisted out of a separating gel (10 %, 13 % or 15 % acrylamide/bisacrylamide (37.5:1), 375 mM Tris-HCl pH 8.8, 0.1% SDS) and a stacking gel (4% acrylamide/bisacrylamide (37.5:1), 125 mM Tris-HCl pH 6.8, 0.1% SDS). Polymerization of gels was initiated by adding ammonium persulfate (APS) (3.34x10<sup>-4</sup> % in separating gel or 5x10<sup>-4</sup> % in stacking gel) and tetramethylethylenediamine (TEMED) (1.34x10<sup>-4</sup> % in separating gel or 1x10<sup>-3</sup> % in stacking gel). After gel polymerization samples (*T. brucei* lysates of 3x10<sup>6</sup> or 8x10<sup>6</sup> cells (3x10<sup>6</sup> by default; and 8x10<sup>6</sup> for eluted proteins derived from co-IPs)) were loaded. 3 μl of the PageRuler Prestained Protein Ladder (Thermo) were also loaded onto a separate lane to assess the molecular weight of the separated proteins. The electrophoresis was carried out in SDS running buffer (25 mM Tris-HCl pH 8.5, 192 mM glycine, 0.1% SDS) at 120 V for 90 min using the EV265 electrophoresis power supply (Consort). Separated proteins were then analysed by western blot analysis (WB).

### 2.6.3 Western Blot

An Immobilon 'polyvinylidene fluoride' (PVDF) membrane (Merck Millipore) was used to transfer the proteins separated by SDS-PAGE via semi-dry or wet blotting technique. Prior to the blotting step for both techniques, the PVDF membrane was activated in 100% methanol (MeOH) for 15 sec and stored in distilled H<sub>2</sub>O. For semi-dry blotting the blotting sandwich was assembled from bottom to top with two sheets of 1.5 mm thick Whatman paper (Albet) soaked in anode buffer (25 mM Tris-HCl pH 7.6, 20% MeOH), the PVDF membrane, the protein gel and one 1.5 mm thick Whatman paper soaked in cathode buffer (300 mM Tris-NaOH pH 10.4, 20% MeOH, 40 mM  $\epsilon$ -aminocaproic acid). The transfer was conducted with 0.8 milliampere (mA)/cm<sup>2</sup> gel for 90 min using the EV265 electrophoresis power supply (Consort).

For the wet blot technique, the blotting sandwich was assembled from bottom to top with two sheets of 0.35 mm thick Whatman paper (Albet) soaked in transfer-buffer (125 mM Tris-HCl pH 7.6, 20% MeOH), the protein gel, the PVDF membrane and again two sheets of 0.35 mm thick Whatman paper (Albet) soaked in transfer-buffer. The sandwich was assembled in a Mini Trans-Blot Cell (BioRad). The transfer was conducted with 150 milliampere per gel for 180 min using the EV265 electrophoresis power supply (Consort) at 4°C.

The membrane was incubated either with a 10 ml solution of 5% milk/PBS or with 3% 'bovine serum albumin' (BSA)/PBS in 50 ml reaction tubes for 1 h at RT or overnight at 4 °C rotating to prevent unspecific binding of the antibodies. The blocking solution was selected according to the required antibody. After blocking, the membrane was incubated with the primary antibodies (Table 2.8) in 0.1% Tween/PBS for 1 h at RT (depending on the used antibody, the antibody was either diluted in 5 ml or 10 ml solution in 50 ml reaction tubes). After three washing steps with 10 ml 0.2 % Tween/PBS for 5 min secondary antibodies were diluted in 5 ml of a 0.1% Tween/0.02% SDS/PBS solution. The membranes were incubated with the secondary antibody solution for 1 h at RT protected from light. After three additional washing steps the membrane was dried between two Whatman papers. The membranes were analysed and quantified using an Odyssey CLx (LICOR) with the ImageStudio software (LI-COR).



**Table 11: Used antibodies**

Name	Isotype	Clonality	Blocking	Dilution WB	Reference
<b>Primary antibody</b>					
Anti-PFR1,2 (L13D6)	mouse	monoclonal	5 % milk/PBS	1:200	K. Gull
Anti-HA (12CA5)	mouse	monoclonal	5 % milk/PBS	1:1,000	E. Kremmer
Anti-HA (3F10)	rat	monoclonal	5 % milk/PBS	1:1,000	Roche
Anti-H3	guinea pig	polyclonal	5 % milk/PBS	1:1,000	Gassen <i>et al.</i> , 2012
Anti-H3	rabbit	polyclonal	5 % milk/PBS	1:50,000	Gassen <i>et al.</i> , 2012
Anti- $\gamma$ H2A	rabbit	polyclonal	3 % BSA/PBS	1:2,000	R. McCulloch
Anti-H2A.Z	rabbit	polyclonal	5 % milk/PBS	1:1,000	Kraus <i>et al.</i> , 2020
Anti-Ty1 (BB2)	mouse	monoclonal	5 % milk/PBS	1:500	K. Gull
<b>Secondary antibody</b>					
IRDye 80 LT anti-guinea pig IgG	goat	polyclonal	0.1%Tween/ 0.02%SDS/PBS	1:20,000	Li-COR
IRDye 680LT anti-mouse IgG	goat	polyclonal	0.1%Tween/ 0.02%SDS/PBS	1:20,000	Li-COR
IRDye 680LT anti-rabbit IgG	goat	polyclonal	0.1%Tween/ 0.02%SDS/PBS	1:20,000	Li-COR
IRDye 680LT anti-rat IgG	goat	polyclonal	0.1%Tween/ 0.02%SDS/PBS	1:20,000	Li-COR
IRDye 800CW Streptavidin	-	-	0.1%Tween/ 0.02%SDS/PBS	1:20,000	Li-COR

Abbreviations: PFR (Paraflagellar rod), IgG (Immunoglobulin G).

#### 2.6.4 Extraction of chromatin-associated proteins

To analyse the amount of chromatin-associated histones we followed the extraction protocol described by Kraus *et al.* 2020. In brief, cells were harvested by centrifugation ( $1,500 \times g$  for 10 min at 4 °C) and washed in 1 ml of trypanosome dilution buffer (TDB; 5 mM KCl, 80 mM NaCl, 1 mM MgSO<sub>4</sub>, 20 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 mM NaH<sub>2</sub>PO<sub>4</sub>, 20 mM glucose, pH 7.4) followed by an additional centrifugation ( $1,500 \times g$  for 10 min at 4 °C). The cell pellet was solubilized in CSK-buffer (100 mM NaCl, 0.1% Triton X-100, 300 mM Sucrose, 1 mM MgCl<sub>2</sub>, 1 mM EGTA, 10 mM PIPES (pH 6.8; with NaOH) supplemented to contain 1x concentration of *Roche cOmplete Protease Inhibitor Cocktail EDTA-free*) and incubated for 10 min at 4°C. To separate the soluble from the insoluble fraction, the suspension was centrifuged ( $2,550 \times g$  for 5 min at 4 °C). The supernatant was removed and the pellet was resuspended with CSK-buffer and centrifuged ( $2,550 \times g$  for 5 min at 4 °C). The pellet with the chromatin fraction was resuspended in SDS-loading buffer supplemented with protease inhibitor cocktail. Samples were boiled at 90°C for 10 min, separated via SDS-PAGE and analysed via WB.

#### 2.6.5 Co-Immunoprecipitation

Prior to the co-immunoprecipitation (co-IP), 30  $\mu$ l of Protein G Sepharose 4 Fast Flow beads (GE Healthcare) per replicate were washed with 1 ml PBS ( $1,000 \times g$  for 1 min at 4°C) and twice in PBS/1 % bovine serum albumin (BSA). Unspecific binding sites were blocked by incubation with PBS/1 % BSA for 1 h at 4°C on an orbital mixer with subsequent centrifugation ( $500 \times g$  for 1 min at 4 °C).

## 2. Materials and Methods

The supernatant was removed, and the corresponding antibody (50  $\mu$ l anti-HA 12CA5 monoclonal mouse IgG antibody 750  $\mu$ l ml anti-TY BB2 monoclonal mouse IgG) diluted in 1.5 ml PBS, was added. Beads and antibody were incubated overnight at 4°C. Unbound antibody was removed by washing the beads three times with 1 ml PBS/0.1 % BSA (500 x *g* for 1 min at 4 °C). Coupled antibodies were stored on ice if the co-immunoprecipitation step was performed at the same day or supplemented with sodium azide (final concentration of 0.02%) for long time storage at 4°C. Before adding the lysate for immunoprecipitation, the beads were washed with 1 ml IP-Buffer buffer (150 mM NaCl, 0.5 % IGEPAL CA-630, 20 mM Tris-HCl, pH 8.0, 10 mM MgCl<sub>2</sub>, 1 mM dithiothreitol (DTT), protease inhibitor cocktail (*Roche cOmplete*<sup>™</sup>)) and subsequent centrifugation (at 4°C, 1000 x *g* for 1 min). The Lysis protocol was performed according to Lowell and colleagues (Lowell et al. 2005) with minor changes. Per replicate, 1x10<sup>8</sup> to 3x10<sup>8</sup> cells were harvested and washed with 10 ml ice-cold TDB (BSF), or with PBS (PCF), and incubated with 1 ml IP buffer for 20 min on ice. Cells were lysed by sonication (5 cycles, each 30 s on and 30 s off) using a Biorupter (Diagenode). A centrifugation step (20,000 x *g* for 15 min at 4 °C) was performed to separate the soluble and insoluble fraction. The soluble fraction was then added to the antibody-coupled beads and incubated at 4°C for 3 h. Beads were washed twice with 1 ml IP buffer for 10 min at 4°C. Proteins were eluted by incubating the beads in 65  $\mu$ l sample buffer (1xNuPAGE<sup>®</sup> LDS Sample buffer with 100 mM DTT) at 70°C for 10 min. Beads were centrifuged (1,000 x *g* for 1 min at RT) and the supernatant was transferred with a Hamilton syringe to new reaction tubes. Eluates were stored at -20°C until analysis by mass spectrometry (MS). 5  $\mu$ l of the precipitated material were used for WB analysis.

### 2.6.6 Biotin identification

The proximity labelling technique Biotin identification (BioID), was carried out as described in Morriswood et al. 2013, with minor changes. Quadruplicates of 1x10<sup>8</sup> 2T1 TBL:*TbSWRC2* cells and of 2T1 WT cells were harvested (1,500 x *g* for 5 min at 4°C). Cells were washed three times with 10 ml ice-cold PBS (1,500 x *g* for 5 min at 4°C) prior to lysis in 1 ml ice-cold IP buffer for 20 min and sonication (6 cycles: 30 sec on, 30 sec off; settings: high) with the Biorupter (Diagenode). Lysates were cleared by centrifugation (15,000 x *g* for 10 min at 4°C), the soluble fraction of each replicate was transferred into new reaction tubes and stored on ice. Pierce Streptavidin Agarose beads (Thermo) were washed twice with 1 ml ice-cold Binding Buffer (50 mM Na<sub>2</sub>HPO<sub>4</sub>, 50 mM NaH<sub>2</sub>PO<sub>4</sub>, 150 mM NaCl) and equilibrated with 1 ml ice-cold IP buffer (500 x *g* for 1 min at 4°C). 50  $\mu$ l beads were added to each of the soluble fractions. The reactions were incubated for 4 h under mild agitation at 4°C. The unbound material was separated from the beads by centrifugation (500 x *g* for 5 min at 4°C) and beads were washed twice with 1 ml ice-cold IP buffer for 5 min on ice (500 x *g* for 5 min at 4°C). Protein elution was carried out as described in chapter 2.5.5.

### 2.6.7 Mass spectrometry and data analyses

In cooperation with his research group F. Butter (IMB, Mainz) carried out mass spectrometry and data analyses. Briefly, samples were run on a Novex Bis-Tris 4-12% gradient gel (Thermo) with '3-(N-morpholino) propanesulfonic acid' (MOPS) buffer (Thermo) for 10 min at 180 V. The gel was stained with Coomassie blue G250 dye (Biozym) prior to cutting each gel lane into pieces, the gel lanes were minced and destained in 50% EtOH/water. The gel pieces were dehydrated with pure acetone, reduced with 10 mM DTT (Sigma Aldrich) and alkylated with 55 mM iodoacetamide in the dark. The dried gel pieces were rehydrated with 1  $\mu$ M trypsin for an in-gel digestion overnight at 37°C. On the following day the digested peptides were desalted and stored on StageTips Rappsilber *et al.*, 2007 for further analysis. Using a C18 reverse phase column that was previously packed in-house with Reprosil C18 (Dr. Maisch GmbH) the peptides were separated along a 240 min gradient on an EasyLC 1000 UHPLC system. The column was enclosed into a column oven (Sonation) and peptides were sprayed into a Q Exactive Plus mass spectrometer (Thermo), which was operating in a data-dependent acquisition mode using a top10 method. Spray voltage was set to approximately 2.4 kilovolt (kV). The acquired raw files were processed with MaxQuant (version 1.5.8.2) Cox & Mann, 2008 using the *T. brucei* protein database downloaded from TriTrypDB. Contaminants, reverse hits, protein groups that were only identified by site and protein groups with less than two peptides (one of them unique) were removed prior to bioinformatics analysis.

All mass spectrometry data can be found on the USB-Stick in the back of this thesis.

### 2.6.8 MNase-ChIP sequencing

The MNase-ChIP experiments were performed with 2T1 cells and the 2T1 *TbSWR1* RNAi cell line with a single Ty1-tagged H2A.Z allele. Functionality of the Ty1-tagged H2A.Z has been described in Kraus *et al.*, 2020. In brief,  $2 \times 10^8$  cells were harvested, crosslinked in 1 % formaldehyde, and subsequently lysed using 200  $\mu$ M digitonin (final concentration). Chromatin was fragmented by incubating the cells with 4 U  $\mu$ l<sup>-1</sup> MNase (Sigma-Aldrich) for 10 min at 25°C. For a detailed ChIP protocol see Wedel & Siegel, 2017. Immunoprecipitation was performed using Dynabeads M-280 sheep anti-mouse coupled to 10  $\mu$ g monoclonal, purified anti-Ty1 (BB2) mouse antibody Bastin *et al.*, 1996, overnight (~16 h) at 4°C in the presence of 0.05% (final concentration) sodium dodecyl sulfate (SDS). Immunoprecipitated material was washed with RIPA-Buffer (50 mM HEPES-KOH (pH 7.5), 500 mM LiCl, 1 mM EDTA, 1% (vol/vol) IGEPAL CA-630, 0.7% (wt/vol) Na-Deoxycholate. LiCl and Na-Deoxycholate dissolved separately in, mix and add remaining components. Store at 4 °C.) and eluted with 200 $\mu$ l elution buffer (50 mM Tris-HCl (pH 8.0), 10 mM EDTA, 1% (wt/vol) SDS) at 65°C for 30 min. Cross-links were reversed at 65°C for ~16 h in the presence of 300 mM NaCl (final concentration). 4  $\mu$ l of 10 mg/ml RNaseA were added to the ChIP sample and incubated at 37°C for 2 h. 4  $\mu$ l of 10 mg/ml proteinase K were added to the ChIP sample and incubated at 55 C for 2 h in a heat block. The tubes were centrifuged (10,000 x g for 10 min at RT).

## 2. Materials and Methods

DNA was purified with the Macherey & Nagel NucleoSpin Gel and PCR Clean-up kit (the NTB buffer was used instead of the NTI buffer due to the high SDS concentration within the samples). Sample purification was performed according to the manufacturer's instructions. The ChIP sample was eluted with 16  $\mu$ l and the input sample with 26  $\mu$ l of NE buffer. The DNA library preparation was performed using NuGEN's Ovation Ultralow System V2 (M01379 v5). Libraries were prepared with a starting amount 7ng of DNA and were amplified in 8 PCR cycles. Libraries were profiled with a 2100 Bioanalyzer (Agilent technologies) and quantified using the Qubit dsDNA HS Assay Kit, in a Qubit 2.0 Fluorometer (Life technologies) set to high sensitivity. All 14 samples were pooled in equimolar ratio and sequenced on a NextSeq500 High output Kit, PE for 2x 42 cycles plus 8 cycles for the index read.

The ChIP-seq data generated in this thesis can be downloaded with the SRA tool kit from <https://dataview.ncbi.nlm.nih.gov/object/PRJNA744383?reviewer=kah6sgb628dqn15orde25j2n7>.

### 2.6.9 Bioinformatics analysis

All bioinformatic analysis were performed by Albert Fradera-Sola (IMB, Mainz)

### 2.6.10 Reads processing and mapping

Library quality was assessed with FastQC version 0.11.8 before being aligned against the *T. Brucei* genome assembly TriTrypDB-48\_TbruceiLister427\_2018\_Genome.fasta and the TriTrypDB-48\_TbruceiLister427\_2018.gff annotation file (Aslett *et al.*, 2010; Muller *et al.*, 2018). Such alignment was performed with bowtie2 aligner (Langmead & Salzberg, 2012) version 2.3.4 (options: --very-sensitive --phred33 --fr --maxins 1000 --minins 0 --end-to-end ). Multimapping reads were filtered out and the resulting unique reads were sorted and indexed using SAMtools (Li *et al.*, 2009). Peaks mapping to annotated features in the .GFF file were quantified with MACS2 (Zhang *et al.*, 2008) version 2.1.2 (options: --g 35000000 --bw 150 --min-length 150 --format BAMPE --keep-dup --auto) with an FDR cutoff of 5%. Chromosome coverage tracks were generated with deepTools (Ramírez *et al.*, 2016) version 3.1 (bamCoverage, options: --binSize 10 --normalizeUsing CPM) and plotted using Gviz (Hahne & Ivanek, 2016) in an R framework (Team, 2014). Finally, signal at TSS was quantified (computeMatrix reference-point, options: --referencePoint TSS -b 25000 -a 25000 --skipZeros) and plotted (plotProfile, options: --perGroup) using deepTools and a custom .BED file containing transcript annotations at the peaks.

### 2.6.11 Differential binding analysis

Further filtering and an exploratory analysis was performed in an R framework including ggplot2 (Wickham, 2009). Overall experimental quality was assessed via IP strength (Diaz *et al.*, 2012). Differential binding comparisons were performed with DiffBind package (Stark & Brown, 2011) and differential peaks were selected with a 5% FDR. Finally, differential peaks were functionally annotated with ChIPseeker package (Yu *et al.*, 2015) were a -3000/ + 3000 region around the TSS was considered as promoter.

### 2.7 Microscopy

#### 2.7.1 Fluorescence microscopy analysis

$1 \times 10^7$  BSF cells were harvested ( $1,500 \times g$  for 10 min at RT) and resuspended in 1 ml HMI-9. The cell suspension was fixed with 2 % (v/v) formaldehyde in HMI-9 for 5 min at RT. The fixed cells were washed three times with 1 ml PBS and resuspended in 500  $\mu$ L PBS. A total of 100  $\mu$ L of cells were added to poly-L-lysine-coated slides. The cells settled down on the slide for 20 min at RT. Attached trypanosomes were permeabilized with 0.2 % IGEPAL CA-630 (v/v) in PBS for 5 min at RT. After washing twice with PBS cells were blocked with 1 % BSA in PBS for 1 h at 37 °C. After 3 washes with PBS, Hoechst was applied for 45 min at RT to stain the DNA. After three subsequent washing steps with PBS, cells were mounted on glass slides using Vectashield (Vecta Laboratories Inc.), and images were captured by using an IMIC microscope (TILL Photonics). Slides were stored at 4 °C in the dark until images were captured with a Leica DMI 6000B microscope. Images were processed with the software Fiji

#### 2.7.2 EM sample preparation and imaging

The EM sample preparation protocol can be found in (Goos *et al.*, 2019) and was adapted from (Weimer, 2006; Höög *et al.*, 2010; Markert *et al.*, 2017).  $3 \times 10^7$  BSF cells were centrifuged ( $750 \times g$  for 3 min at RT). All but 2 ml medium were removed and 2 ml heat-inactivated fetal calf serum was added as a cryoprotectant. Cells were centrifuged ( $750 \times g$  for 3 min at RT) and the pellet was transferred to a PCR tube and further compacted (10 s, minifuge). A drop of the final pellet (around 1.5  $\mu$ L) was transferred to the freezing container (specimen carriers type A, 100  $\mu$ m, covered with specimen carriers type B, 0  $\mu$ m, Leica Microsystems). High pressure freezing was done in an EM HPM100 (Leica Microsystems) at a freezing speed  $>20\,000 \text{ Ks}^{-1}$  and a pressure  $>2100$  bar. The samples were stored in liquid nitrogen until freeze substitution in an EM AFS2 freeze substitution system (Leica Microsystems). For embedding in Epon, samples were incubated in 0.1% (w/v) tannic acid and 0.5% (v/v) glutaraldehyde in anhydrous acetone at  $-90^\circ\text{C}$  for 96 h (with one change in solution after 24 h), washed four times for 1 h with anhydrous acetone at  $-90^\circ\text{C}$  and fixed in 2%  $\text{OsO}_4$  (w/v) in anhydrous acetone at  $-90^\circ\text{C}$  for 28 h. Then the temperature was gradually raised to  $-20^\circ\text{C}$  within 14 h, kept at  $-20^\circ\text{C}$  for 16 h and gradually raised to  $4^\circ\text{C}$  within 4 h. Afterwards samples were immediately washed with anhydrous acetone at  $4^\circ\text{C}$  four times at 0.5 h intervals, followed by gradually increasing the temperature to  $20^\circ\text{C}$  within 1 h. Subsequently, samples were transferred for embedding into increasing concentrations of Epon (50% Epon in acetone for 3 h at room temperature, 90% Epon in acetone overnight at  $4^\circ\text{C}$ , followed by two times 100% Epon at room temperature for 2 h, all solutions were freshly prepared). Epon infiltrated samples were polymerized for 72 h at  $60^\circ\text{C}$ . For staining and contrasting, Epon-embedded sections were incubated in 2% aqueous uranyl acetate for 10 min followed by incubation in Reynolds lead citrate for 5 min. LR White-embedded sections were incubated in 2% aqueous uranyl acetate for 5 min followed by incubation in Reynolds lead citrate for 1.5 min.

## 2. Materials and Methods

A 200 kV JEM-2100 (JEOL) transmission electron microscope or a 30 kV JEOL JSM-7500F Scanning Electron Microscope equipped with a TemCam F416 4k x 4k camera (Tietz Video and Imaging Processing Systems) was used for imaging.

### 3. Results

Like in other eukaryotic organisms, transcription regulation in *T. brucei* is dependent on the histone variant H2A.Z. Several PTMs like H3K4 methylation or H3K23 acetylation were exclusively found on H2A.Z containing nucleosomes (Kraus *et al.*, 2020). Two different histone acetyltransferases are required for H4 and H2A.Z acetylation which is in sharp contrast to any other eukaryotic model organism. In addition to that, complexes responsible for incorporation and removal of H2A.Z could not be identified so far. This leaves several questions about the molecular mechanisms underlying H2A.Z based transcription regulation in *T. brucei* unanswered. To answer these questions, I started with a systematic search for SNF2 proteins that possess the molecular characteristics of the INO80 subfamily followed by co-immunoprecipitation (co-IP) experiments to identify new CRCs of this subfamily. Furthermore, I used co-IP experiments to identify potential NuA4-like complexes in *T. brucei*. These experiments focused on HAT2 and a homolog of a NuA4 subunit.

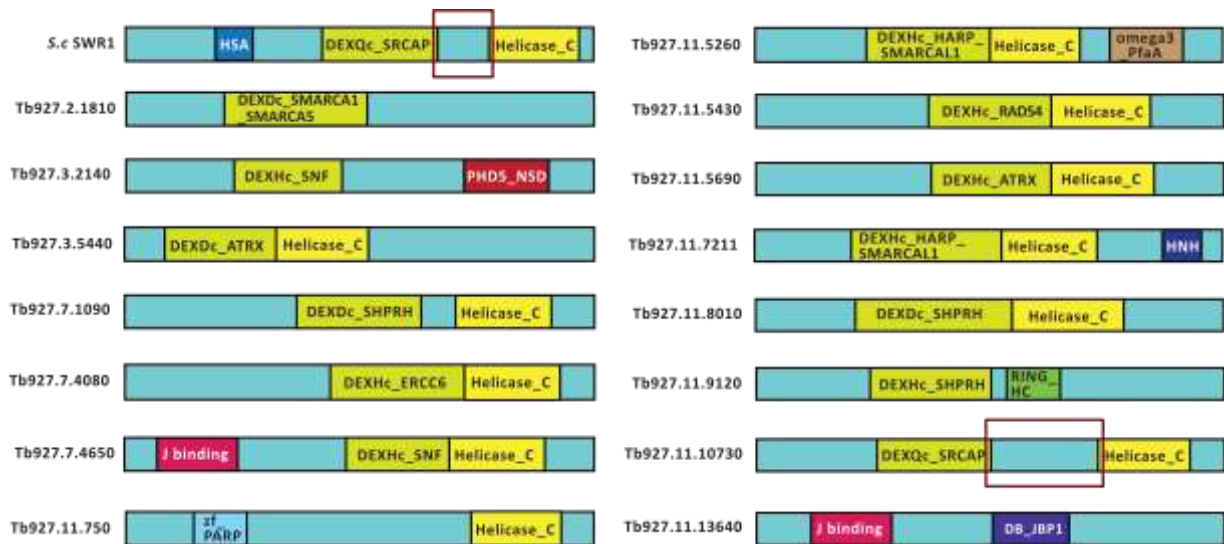
#### 3.1 Identification of a potential SWR1-like complex

##### 3.1.1 Characterization of SNF2 proteins in *T. brucei*

The project started with a systematic analysis of the 15 SNF2 proteins that are listed in the Tritryp database. It aimed towards the identification of potential candidates of the INO80 subfamily for subsequent co-IP analysis. The amino acid sequences of all 15 proteins were analysed with the NCBI and InterPro 'Basic Local Alignment Search Tool' (BLAST). In focus were especially those proteins that possess an insertion between the DExx-motif and the helicase C domain since this insertion is characteristic for members of the INO80 subfamily (Gerhold & Gasser, 2014; chapter 1.4.1). The proteins Tb927.7.1090 and Tb927.11.10730 were the only proteins that fulfilled this criterion (Fig 10). The DExx-motif of Tb927.7.1090 was identified as being a DEXDc motif related to SHPRH.

SHPRH is a nucleosome-E3 ubiquitin ligase in humans that is linked with DNA repair. It is homologous to yeast Rad5 that is linked with PCNA ubiquitylation in the template switch pathway (Myung & Smith, 2008). Tb927.7.1090 was identified as being homologous to the *S. cerevisiae* Rad5 (Dattani & Wilkinson, 2019) and therefore discarded as a potential member of the INO80 subfamily. BLAST analysis of Tb927.11.10730 in contrast identified the DExx-motif as a DEXQc motif that is homologous to the DEXQc motif of SRCAP. Since DEXQc motifs can only be found in members of the INO80 subfamily (Flaus *et al.*, 2006) and since SRCAP is the human homolog of the SWR1 complex (Table 1; Clapier & Cairns, 2009), Tb927.11.10730 became a very interesting candidate for subsequent mass spectrometry (MS)-coupled co-IP experiments. For this purpose, one allele of Tb927.11.10730 was tagged with an HA-epitope tag and the remaining allele was replaced by a resistance gene. Tryptag is a database that provides information on whether *T. brucei* proteins can be tagged at the N- or C-terminus (Dean *et al.*, 2017). Since Tryptag had no information about Tb927.11.10730, I decided to tag the protein at the C-terminus.

### 3. Results



**Figure 10: Overview of SNF2 ATPases in *T. brucei***

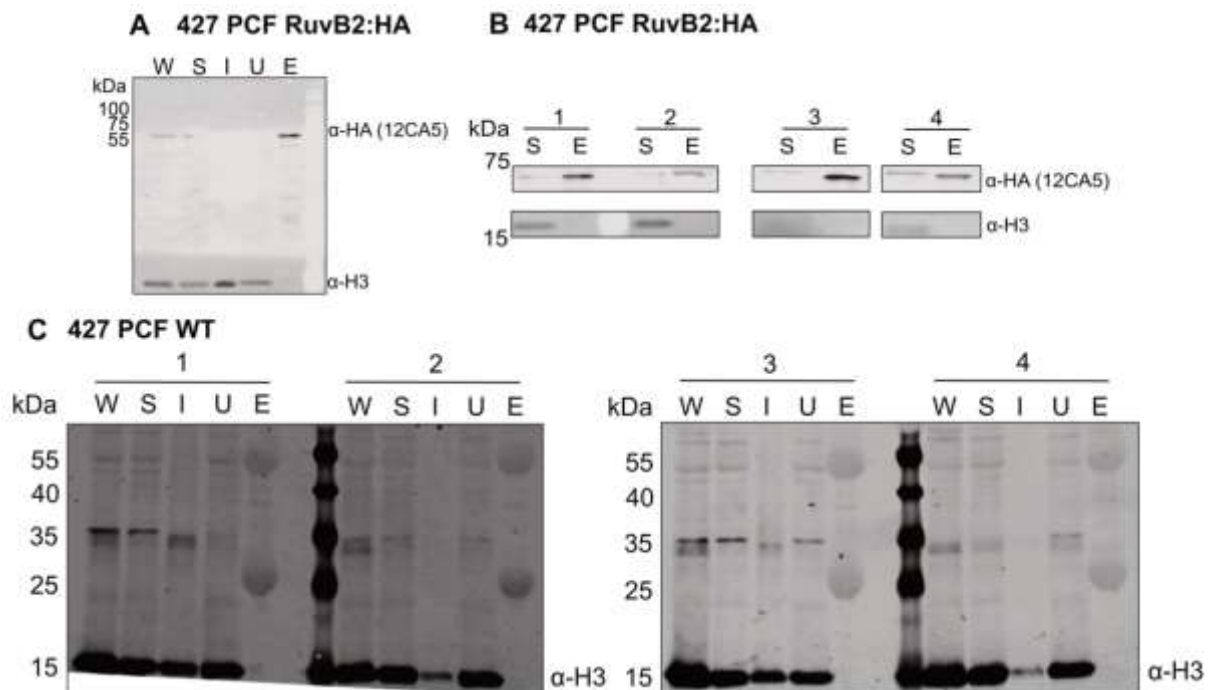
Schematic depiction of the domain composition of the *S. cerevisiae* SWR1 and all 15 SNF2 ATPases in *T. brucei*. The subdomains of each ATPase–translocase domain is labelled in green (DEX- motif) and yellow (Helicase C domain).

Only one of the Tb927.11.10730 alleles could be tagged. Tagging of the second allele or efforts to replace it with a resistance marker failed multiple times. The attempt to tag or to knock-out the second allele in a cell line with a N-terminally tagged Tb927.11.10730 failed as well. This indicated that tagging of Tb927.11.10730 on both termini might lead to the expression of a non-functional protein. Western Blot analysis revealed a signal between 100 and 130kDA only for N-terminally tagged Tb927.11.10730 but not for the C-terminally tagged protein (Fig S2). The estimated molecular weight of the protein including the HA-epitope tag would be around 140kDA. Given that the wild type control revealed no signal, it is likely that the signal is derived from the tagged protein. Since it was not possible to ensure the expression of a functional version of Tb927.11.10730 I had to use an alternative strategy to identify CRCs of the INO80 subfamily.



### 3.1.2 Identification of a new SNF2 CRC of the INO80 subfamily

The DNA helicases RuvB1 and RuvB2 are important and conserved parts of the INO80 as well as the SWR1 complex (chapter 1.4.1; Willhoft & Wigley, 2020). If Tb927.11.10730 or any other SNF2 ATPase is part of one of these complexes, a co-IP experiment with RuvB2 should identify the protein. Tb927.4.1270 and Tb927.4.2000 are the two RuvB helicases in *T. brucei* and I will refer to these two proteins as RuvB1 and RuvB2 in the following. RuvB2 was used for the tagging approach and data from the TrypTag database indicated that it is possible to tag the protein at both termini. Due to the availability of pMOTag tagging vectors with different selection markers, I decided to tag both RuvB2 alleles C-terminally. Proper integration of the tagging constructs was verified by diagnostic PCR (Fig S2C). Expression of the tagged protein and evaluation of the efficiency of the co-IP protocol as well as successful precipitation of the tagged protein were controlled via Western Blot analysis (Fig 11A). The expected molecular weight of the protein was around 56 kDa. According to this the protein was visible in the whole cell lysate (W) and in the soluble fraction (S). The absence of any signal in the insoluble fraction (I) indicated that the cell lysis worked well and that the amount of detergent used in the lysis buffer was sufficient to enable solubilization of the target protein. The absence of a signal in the sample of the unbound fraction (U), which was taken after the incubation with the  $\alpha$ -HA-antibody-coupled beads indicated that the immunoprecipitation step was successful. This was also confirmed by the strong signal that could be detected in the sample of the eluted material (E, Fig 11A).



**Figure 11: Representative western blots of the RuvB2 co-IPs**

**(A)** Protein extracts of RuvB2:HA (~ 56 kDa) were used for the test co-IP; **(B)** four replicates for MS analysis and **(C)** 427 PCF WT control cells as a control. W (whole cell extract), S (soluble supernatant), I (insoluble pellet), U (unbound material), E (eluate: 10-fold compared to the other samples). Blots were probed with 12CA5 anti-HA and anti-H3 antibodies. Vellmer *et al.*, 2021 in revision

### 3. Results

Since I could not confirm the efficiency of the co-IP protocol, I performed control WB analysis only with samples of the soluble fraction and the eluted material. The co-IP experiment was performed in quadruplets with  $1,2 \times 10^8$  cells per replicate. Four co-IPs using cell lysates of  $1,2 \times 10^8$  427 PCF wild type cells served as a negative control (Fig 11C). Successful precipitation of RuvB2-HA was controlled by WB analysis (Fig 11B). MS analysis of precipitated RuvB2-HA identified Tb927.10.10730, RuB2 and 13 additional proteins that were significantly enriched ( $p < 0,01$ ; fold-enrichment above 1.0; Table 12, Fig S3A, MS data set 1)

**Table 12: Summary of the RuvB2 co-IP**

Accession number	Annotation	Fold enrichment	p-value	NES
Tb927.4.2000 (RuvB2)	RuvB-like DNA helicase, putative	8.10	1.50E-08	0.52
Tb927.11.10730	SWI/SNF-related helicase, putative	7.62	7.08E-08	5.04
Tb927.4.980	actin	6.65	1.92E-05	5.71
Tb927.11.5830	YL1 nuclear protein	6.63	5.94E-07	-
Tb927.7.4040	hypothetical protein	6.62	1.56E-07	5.52
Tb927.4.1270 (RuvB1)	RuvB-like DNA helicase, putative	6.52	3.21E-05	0.69
Tb927.10.11690	YEATS family, putative	6.24	1.70E-05	4.89
Tb927.10.2000	actin like Protein	6.11	1.11E-05	5.31
Tb927.6.2570	putative SUMO-interacting motif containing protein	5.80	6.16E-06	7.26
Tb927.8.600	Bucentaur or craniofacial development, putative	5.44	5.06E-07	2.52
Tb927.3.3020	actin like protein	4.80	5.41E-04	3.43
Tb927.11.6290	HIT zinc finger, putative	4.70	1.02E-08	4.38
Tb927.11.16370	SHQ1 protein, putative	4.05	2.68E-07	-
Tb927.9.5320	nucleolar RNA binding protein	3.63	6.18E-07	2.33
Tb927.10.170	pseudouridine synthase, Cbf5p	1.90	9.47E-09	2.51

15 proteins with a positive or unknown nuclear enrichment score were identified by MS in the RuvB2-HA co-IP (Fig S3A, MS data set 1). The nuclear enrichment score (NES; 28727848 indicates the probability of a nuclear localisation based on cell fractionation combined with quantitative MS analysis. The "Annotation" column indicates the curated annotation that was found for the corresponding accession number on the TriTyp database. In the "p-value" column a probabilistic confidence measure (*P-value*) is assigned to each identified protein. The fold enrichment compared to the WT control is stated for every identified protein (Vellmer *et al.*, 2021 in revision).

As I intended to identify a CRC, I was only interested in proteins that are enriched in the nucleus, therefore having a positive nuclear enrichment score (NES). The nuclear enrichment score indicates whether a protein is located in the nucleus (positive score) or in another cellular compartment (indicated by a negative NES; Goos *et al.*, 2017). Only Tb927.11.10730 and no other SNF2 ATPase was co-precipitated in the RuvB2 co-IP. BLAST analysis of Tb927.7.4040 the only protein having no annotation revealed no conserved DEXQ motif or helicase C domain. The presence of actin, actin-like proteins, a YEATS and a YL1-domain containing protein hints towards the presence of an SNF2 complex of the INO80 subfamily (Table 1; Willhoft & Wigley, 2020). A more detailed analysis of the 14 proteins indicated the presence of a second protein complex.

### 3. Results

This complex might be assembled around the SHQ1 protein (Tb927.11.16370; formerly known as Yil104cm now SHQ1 originates from 'required for small nucleolar RNAs of the box H/ACA family quantitative accumulation'; Yang *et al.*, 2002) and the 'Centromere-binding factor 5' (Cbf5) pseudouridine synthetase (Tb927.10.170). RuvB1 and RuvB2 as part of the R2TP chaperone complex 21925213 were shown to interact with NAP57 (also known as dyskerin which is Cbf5 in *S. cerevisiae*) and SHQ1 in course of H/ACA 'small nucleolar ribonucleoproteins' (snoRNP) biogenesis (Machado-Pinilla *et al.*, 2012). The nucleolar RNA binding protein Tb927.9.5320 might therefore be part of such a complex.

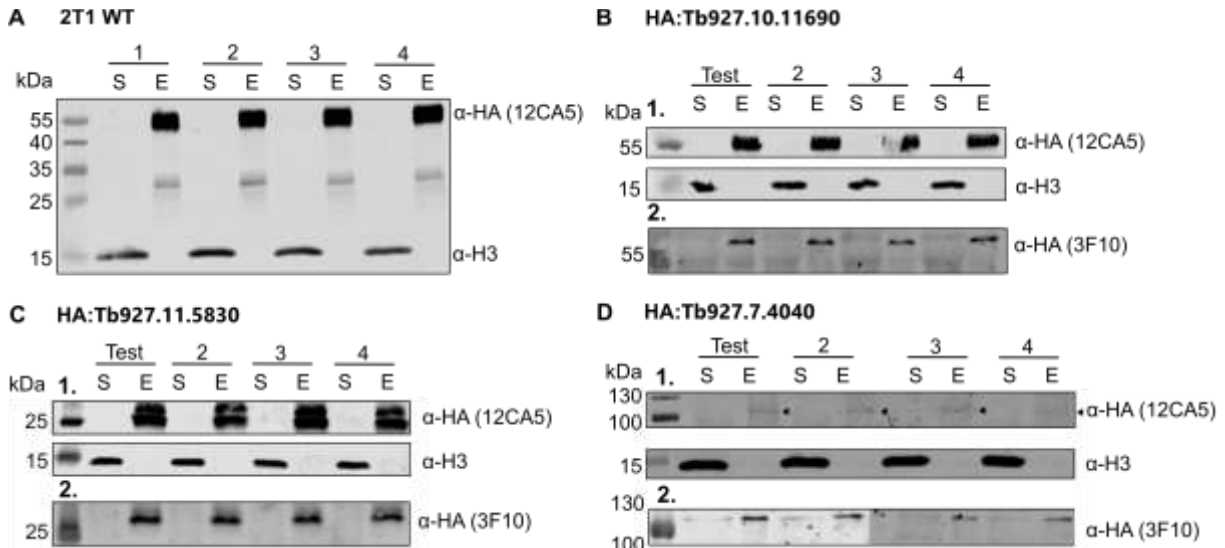
To separate the two putative complexes from each other and to learn more about a complex assembled around the SNF2 ATPase Tb927.11.10730 I decided to perform a set of reciprocal co-IP experiments. For this second set of co-IPs, I selected the YEATS-domain containing protein Tb927.10.11690, the YL-1 domain containing protein Tb927.11.5830 and the hypothetical protein Tb927.7.4040. Based on the orthology and synteny data on the TriTryp database Tb927.7.4040 could be a unique complex component since it can only be found in *T. brucei*, *T. cruzi* and *Blechnomonas ayalai*. The YEATS and YL1-domain containing proteins are conserved subunits of CRCs of the INO80 subfamily and consequently representing good candidates to confirm or to confute the initial RuvB2 co-IP results.

I used again the TrypTag database to issue the best tagging strategy for all three proteins. For Tb927.7.4040 I chose the same tagging strategy as for RuvB2. Tb927.10.11690 and Tb927.11.5830 were tagged N-terminally with an HA-epitope tag. For amplification of the tagging construct via PCR, a modified pPOTv7 vector (ref) was used. By using conventional restriction enzyme-based cloning, I replaced the eYFP cassette with an HA-tag that consists of three epitopes. The second allele of both genes was replaced with a resistance marker. Each protein was modified in a separate cell line. Correct integration of the tagging constructs was verified by diagnostic PCR and expression of the tagged protein was confirmed by WB analysis (Fig 12B-D & Fig S2). No signals for HA-tagged Tb927.7.4040, Tb927.10.11690 and Tb927.11.5830 could be detected in the samples of the soluble fraction that was used as the input material (Fig 12B-D). Due to the possibility that a low expression of these proteins could have caused the absence of signals in the WB analysis, I performed test co-IP experiments to assess, whether the tagged proteins can be precipitated (Fig 12B-D).

The expected molecular weights of the HA-tagged proteins were estimated to be approximately 57 kDa (Tb927.10.11690), around 30 kDa (Tb927.11.5830) and around 96 kDa (Tb927.7.4040). Samples taken from the eluted material of each co-IP revealed bands at the corresponding molecular weight of the tagged proteins, therefore confirming expression of the proteins (Fig 12B-D). Additional bands migrating at around 25 and 55 kDa could be observed in the HA:Tb927.11.5830 and HA:Tb927.10.11690 co-IP. IgG antibodies consist of two different peptides, a so-called "light chain" and "heavy chain" with molecular weights of 25 and 50 kDa (Schroeder & Cavacini, 2010). Therefore, the additional bands appear to originate from the 12CA5  $\alpha$ HA-antibody that was used for the co-IPs. A control co-IP with 2T1 wild type confirmed this assumption (Fig 12A).

### 3. Results

In the following, I repeated the WB analysis using the mouse 3F10  $\alpha$ HA-antibody to avoid detection of the rabbit 12CA5 antibody. The eluted material of the test co-IP was stored for subsequent MS analysis. In the following three additional co-IPs were performed to obtain a quadruplet for subsequent MS analysis.



**Figure 12: Control of the co-IPs that identified the *TbSWR1* complex**

Representative Western blots of samples taken during the co-IP of (A) 2T1 WT cells, (B) HA:Tb927.10.11690 (~57 kDa), (C) HA:Tb927.11.5830 (~30 kDa) and (D) HA:Tb927.7.4040 (~96 kDa). S (soluble supernatant, D2: 2-fold compared to other samples), E (eluate: 10-fold compared to the other samples, D2: 20-fold compared to the other samples). Blots were probed with 12CA5 anti-HA and anti-H3 antibodies (B1 – D1) or 3F10 anti-HA antibody (B2 – D2) to avoid signals from the eluted 12CA5 antibody (Vellmer et al., 2021 in revision).

The eluates of these three co-IP experiments were also analysed via WB before the eluted material was sent for MS analysis (Fig 12B-D, MS data set 1). MS of the data revealed that Cbf5p, SHQ1 and the nucleolar RNA binding protein which were identified in the initial RuvB2 co-IP could not be identified again (Table 13, Fig S3B & S4). The remaining 12 proteins that were identified in the initial RuvB2 co-IP, except for the YL1-domain containing protein Tb927.11.5830 which was not identified in the co-IP with HA-tagged Tb927.10.11690, were identified in all co-IPs of the second set (Table 13, Fig S3B & S4, MS data set 1). This supported the hypothesis that Cbf5p, SHQ1 and the nucleolar RNA binding protein could in fact build a separate RuvB-dependent protein complex. Nevertheless, subsequent co-IP experiments with one of these three proteins will be needed to confirm this. The three canonical histones H2B, H3 and H4 were identified in all co-IPs of the second set. H2A and H2A.Z were only identified in the co-IP experiments with tagged Tb927.7.4040 and Tb927.11.5830 (Table 3.2, Fig S4). Mass spectrometry analysis of the eluates derived from the co-IPs performed with HA-tagged Tb927.7.4040 also identified a protein that possesses the domain of unknown function DUF1935 (Tb927.10.2610) and the hypothetical protein Tb927.9.12800. Interestingly the hypothetical protein Tb927.9.8510 was identified in all reciprocal co-IPs (Table 13, Fig S3B & S4) indicating that it is also part of the complex. The identification of all four canonical histones and the variant H2A.Z provided a hint that the identified protein complex is a CRC of the INO80 subfamily.

### 3. Results

The complex assembled around the SNF2 ATPase Tb927.11.10730 appeared to consist of at least 13 proteins (blue labelled proteins Table 3.2), since these proteins could be independently identified in at least three of four co-IP experiments (Fig S3 & 4). To learn more about the 13 putative complex components, I analysed each protein with the Phyre2 homology modelling web tool (Kelley *et al.*, 2015).

### 3. Results

**Table 13: summary of co-IP experiments that identified the *Tb*SWR1 complex**

Gene ID	Annotation	Identified domains	Phyre2 modelling	NES	Co-IP 1	Co-IP 2	Co-IP 3	Co-IP 4
Tb927.3.3020 ( <i>Tb</i> ARP1)	actin like Protein, putative	Actin	Arpx (Cov. ~98 % Conf. 100 %)	3.43	X	X	X	X
Tb927.4.980	actin	Actin	-	5.71	X	X	X	X
Tb927.4.1270 ( <i>RuvB1</i> )	ruvB-like DNA helicase, putative	TIP-49 domain	ruvb-like protein 1/2 (Cov. ~99 % Conf. 100 %)	0.69	X	X	X	X
Tb927.4.2000 ( <i>RuvB2</i> ) co-IP 1	ruvB-like DNA helicase, putative	TIP-49 domain	ruvb-like protein 1/2 (Cov. ~94% Conf. 100%)	0.52	X	X	X	X
Tb927.6.2570 ( <i>Tb</i> ARP2)	SUMO-interacting motif-containing protein	-	Arpx8/9 (Cov. ~79 % Conf. ~98 %)	7.26	X	X	X	X
Tb927.7.4040 ( <i>Tb</i> SWRC4) co-IP 4	hypothetical protein	-	SANT (Cov. 12 % Conf. 97 %)	5.52	X	X	X	X
Tb927.8.600 ( <i>Tb</i> SWRC5)	Bucentaur or craniofacial development, put.	BCNT-domain	-	2.52	X	X	X	X
Tb927.9.8510 ( <i>Tb</i> SWRC3)	hypothetical protein, conserved	-	hist. Methyltransf. set7/9 (Cov. 33 % Conf. 96 %)	N/A	/	X	X	X
Tb927.10.2000 ( <i>Tb</i> ARP3)	actin like Protein, putative	Actin	Arpx (Cov. ~87 % Conf. 100 %)	5.31	X	X	X	X
Tb927.10.11690 ( <i>Tb</i> SWRC1) co-IP 2	YEATS family, putative	YEATS-domain	Yaf-9 (Cov. 25% Conf. 98 %)	4.89	X	X	X	X
Tb927.11.5830 ( <i>Tb</i> SWRC2) co-IP 3	YL1 nuclear protein, putative	YL1	SWC2 (Cov. 24% Conf.98 %)	-	X	/	X	X
Tb927.11.6290 ( <i>Tb</i> SWRC6)	HIT zinc finger, putative	ZNHIT1	SWC6 (Cov. 27 % Conf. 99 %)	4.38	X	X	X	X
Tb927.11.10730 ( <i>Tb</i> SWR1)	SWI/SNF-related helicase, putative	DEXQ-Box SRCAP; Helic. C	INO80/CHD1/SWR1/RAD54 (Cov. ~60 % Conf. ~100 %)	5.04	X	X	X	X
Tb927.9.12800	hypothetical protein	-	-	N/A	/	X	/	/
Tb927.10.2610	domain of unknown function (DUF1935)	-	-	N/A	/	X	/	/
Tb927.11.16370	SHQ1 protein, putative	-	-	N/A	X	/	/	/
Tb927.9.5320	nucleolar RNA binding protein	-	-	2.33	X	/	/	/
Tb927.10.170	pseudouridine synthase, Cbf5p	-	-	2.51	X	/	/	/
Tb927.7.2940	histone H2A	-	-	1.02	/	X	/	X
Tb927.10.10590	histone H2B	-	-	0.04	/	X	X	X
Tb927.1.2550	histone H3	-	-	2.04	/	X	X	X
Tb927.5.4260	histone H4	-	-	-0,37	/	X	X	X
Tb927.7.6360	histone H2A.Z	-	-	1,57	/	X	/	X

### 3. Results

23 proteins with a positive or non-applicable (N/A) NES were identified by MS analysis in at least three of the four co-IP experiments. The initial co-IP was performed with RuvB2 (Tb927.4.2000, the reciprocal co-IPs with the proteins Tb927.10.11690, Tb927.11.5830 and Tb927.7.4040 were performed to confirm the RuvB2 co-IP data). The proteins labelled in blue were found in at least three of four co-IP experiments. The “Annotation” column indicates the curated annotation that was found for the corresponding accession number in the TriTryp database. The “identified domains” column displays the domains that were found by BLAST search using the NCBI database. The Phyre2 modelling column indicates proteins that were identified by homology modelling. Coverage (Cov.) indicates the coverage in percent between query and template. The confidence (Conf.) represents the relative probability in percent (from 0 to 100) that the match between query and template is a true homology. The last column shows in which of the four co-IP experiments the protein could be identified.

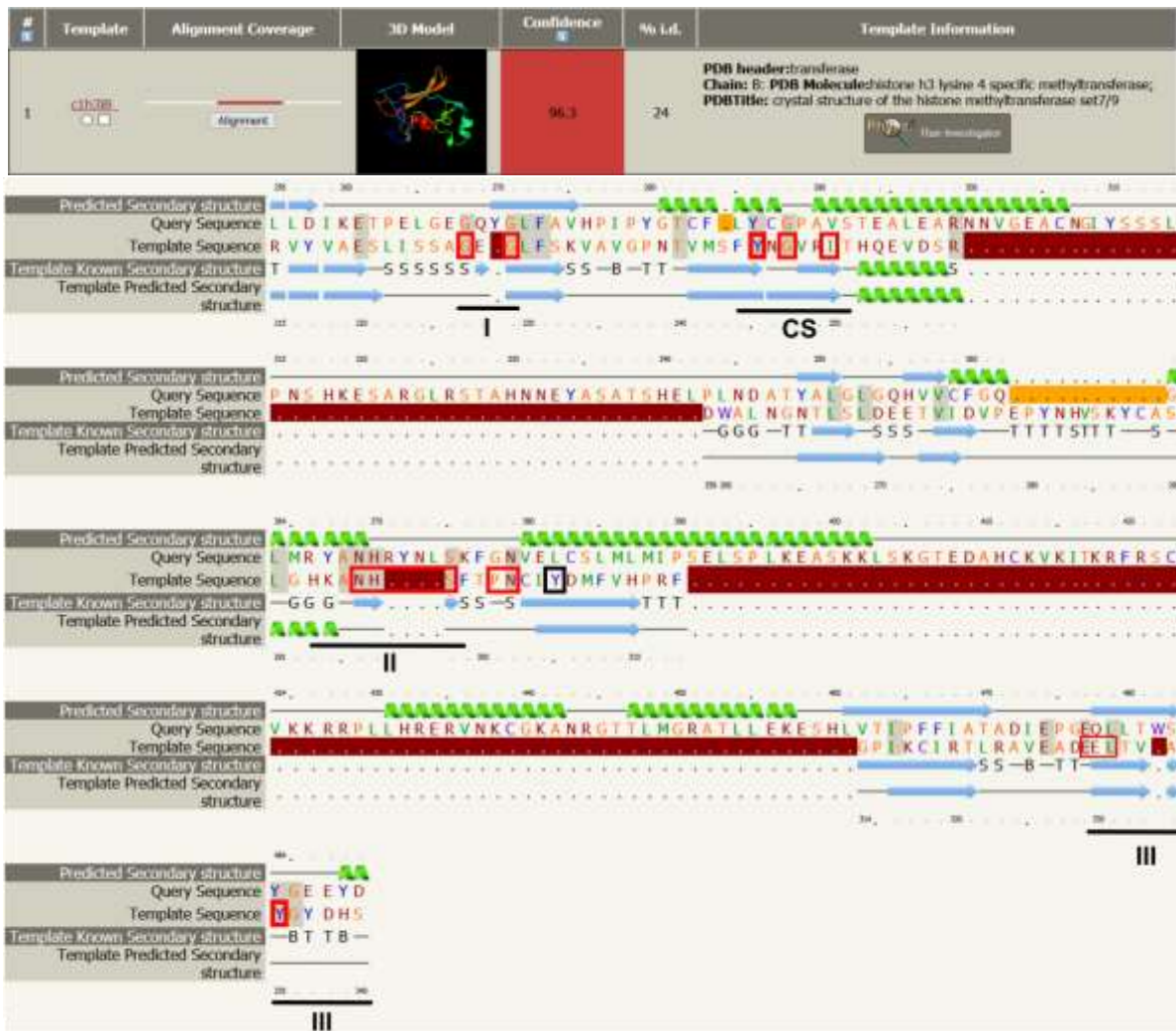
#### **3.1.3 The identified SNF2 complex has characteristics of a SWR1-like complex with an unusual complex composition**

Briefly, the Phyre2 homology modelling web tool uses the amino acid sequence of the protein of interest to predict its secondary structure. This secondary structure is then modelled onto known secondary structures of other proteins, therefore giving information about potential homologs in other organisms (Kelley *et al.*, 2015).

The first three proteins I analysed with Phyre2 were the three co-IP candidates of the second set, the hypothetical protein Tb927.7.4040, the YEATS-domain containing protein Tb927.10.11690 and the YL1-domain containing protein Tb927.11.5830. The YEATS-domain protein was identified as being partially homologous to Yaf-9 while homology modelling of the YL1-domain containing protein could relate it with SWC2. Both proteins are part of the SWR1 complex (Table 1; Willhoft & Wigley, 2020). A small part of Tb927.7.4040 was identified as being homologous to a SANT-domain of ‘DNA methyltransferase 1-associated protein 1’ (DMAP1). A SANT-domain can be found in the SWR1 complex subunit SWC4 (Liu *et al.*, 2020). The ZNHIT1 domain containing protein Tb927.11.6290 was modelled onto the protein sequence of the yeast SWR complex component SWC6. Homology modelling results of these four proteins indicated that the identified complex is a SWR1-like complex. The presence of two SWR1 specific factors, the BCNT-domain-containing Tb927.8.600 and the Zf-HIT-domain-containing Tb927.11.6290 supported this assumption. However, the homology modelling of the SNF2 ATPase did not support this assumption, since the ATPase was related to multiple members of the SNF2 ATPase family (Table 13). Characterisation of the ‘small ubiquitin-related modifier’ (SUMO)-interacting motif-containing protein Tb927.6.5270 and the hypothetical protein Tb927.9.8510 revealed some interesting results. With a coverage of almost 80% Tb927.6.5270 could be modelled onto the structures of the *S. cerevisiae* proteins ‘actin related protein 5’ (Arp5) and Arp8. The amino acid sequences of Tb927.3.3020 and Tb927.10.2000 could be modelled onto multiple Arps. Together with Tb927.6.5270 the complex appears to possess three different actin-related proteins which is in sharp contrast to other SWR1 complexes which only possess two Arps (Table 1; Willhoft & Wigley, 2020). According to these data I named the proteins *TbARP1* (Tb927.3.3020), *TbARP2* (Tb927.6.5270) and *TbARP3* (Tb927.10.2000). INO80 complexes on the other hand consist of three Arps (Willhoft & Wigley, 2020).

### 3. Results

The core of Tb927.9.8510 was modelled onto the SET-domain of the human H3K4-specific methyltransferase SET7/9 (Fig 13; PDB ID: 1H3I). SET-domains possess four major motifs. Besides the catalytic site (CS; Fig 13) the domain harbours three sites to bind the S-adenosyl-L-methionine (SAM; Fig 13 I-III) which supplies the methyl group for the methylation reaction (Dillon *et al.*, 2005). A so-called F/Y-switch (named by phenylalanine (F) and tyrosine (Y) that build the switch) can be found in close proximity to the second and third SAM binding motif. The F/Y switch determines whether the target lysine is mono-, di- or tri-methylated (Collins *et al.*, 2005).



**Figure 13: Core region of Tb927.9.8510 (TbSWRC3) is homologous to the human H3K4 specific SET7/9 histone methyltransferases**

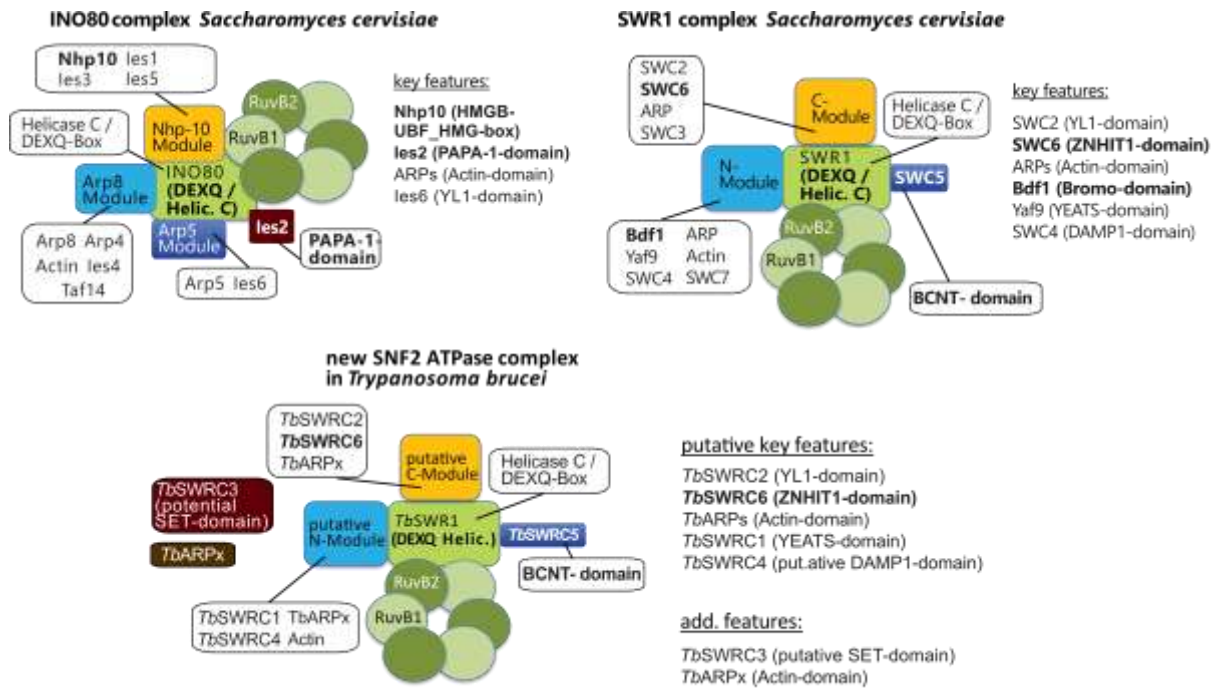
Phyre2 modelling of Tb927.9.8510 could link the core of the protein to the SET7/9 methyltransferase. In the alignment of Tb927.9.8510 (Query) with the human SETD7 (Template, PDB ID: 1H3I) the three S-adenosylmethionin binding sites (I-III) as well as the catalytic site (CS) are highlighted. Dark-red areas represent sequence gaps in the template, yellow areas represent sequence gaps in the Query. Predicted and known β-sheets are depicted in light-blue. Predicted and known α-helices are depicted in light-green. The alignment was created by the Phyre2 online modelling platform (56). The figure listed in the “Template” column is giving the RCSB PDB ID of the protein that has been used for homology modelling. The “Alignment Coverage” column shows to which part of the query the template could be aligned (depicted in red). % i.d. represents the sequence identity between query and template (Vellmer *et al.*, 2021 in revision)



### 3. Results

A detailed investigation of the Phyre2 alignment revealed that only two of three conserved amino acids of the CS are present. While the Tb927.9.8510 CS possesses tyrosine and glycine just like the CS of SETD7, the conserved isoleucine is replaced by valine in Tb927.9.8510 (Fig 13; conserved amino acids are rendered with red boxes). Since valine and isoleucine have the same chemical properties, it can be assumed that this difference is not affecting the processivity of the CS. The three SAM binding motifs can be identified in Tb927.9.8510 but each motif possesses an insertion of one or four amino acids. In addition, an insertion of 43 amino acids is localized between SAM binding motif I and II and a 71 amino acid large insertion between SAM binding motif II and III (Fig 13). Based on the Phyre2 homology modelling data and the presence of proteins with a BCNT and ZNHIT1 domain (Tb927.8.600 and Tb927.11.6290), domains that were exclusively identified in SWR1 complexes (Fig 14; Willhoft & Wigley, 2020), I assumed that the newly identified SNF2 CRC is a SWR1-like complex. According to my assumption I renamed the majority of the 13 complex subunits (Table 13). Nevertheless, a further molecular characterization was required to confirm this hypothesis. For this purpose, I started to analyse RNAi cell lines of Tb927.11.10730 (in the following referred to as *TbSWR1*) and Tb927.9.8510 (in the following referred to as *TbSWRC3*). Since a SET-methyltransferase has never been described as part of a SWR1-like complex, I focused on the contribution of *TbSWRC3* to H2A.Z incorporation. Although multiple RNAi cell lines with different target sites in the *TbSWRC3* mRNA were generated, depletion of *TbSWRC3* showed only a slight or no phenotype (Bachelor thesis Kostantin Köthe; Fig S5). The deletion of both wild type alleles was not successful (Bachelor thesis Kostantin Köthe), which indicated that the protein is essential for the organism. As the results for *TbSWRC3* were inconclusive, I continued with the generation of *TbSWR1* RNAi cell lines. Downregulation of *TbSWR1* should result in a decrease of chromatin-associated H2A.Z if *TbSWR1* is the SWR1-like ATPase in *T. brucei*. In case of *TbSWR1* being an INO80-like ATPase the opposite phenotype should be observable. Since H2A.Z was shown to be required for proper transcription in *T. brucei* (Kraus *et al.*, 2020), *TbSWR1* depletion should result in a loss of transcriptional.

### 3. Results



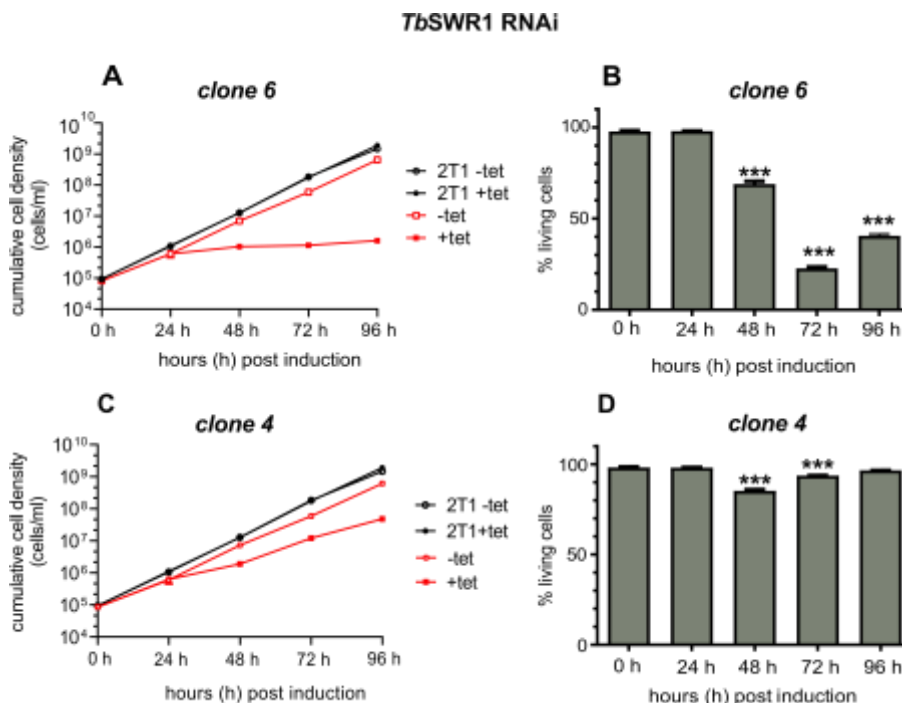
**Figure 14: The *T. brucei* SNF2 ATPase complex exhibits characteristics of a SWR1 complex**

A comparison of the modular composition of the identified SNF2 complex with the SWR1 and INO80 complexes of *S. cerevisiae*. **INO80 complex *S. cerevisiae***: The Nhp-10 module is required for nucleosome interaction. The Arp5 module, which contains the YL1-domain protein les6 is responsible for the nucleosome remodelling step, while the Arp8 module binds the nucleosome (Tosi *et al.*, 2013). The PAPA-1-domain of les2 plays a structural role within the complex (Chen *et al.*, 2013; Tosi *et al.*, 2013). INO80 specific domains are highlighted in bold. **SWR1 complex *S. cerevisiae***: Proteins listed under key features are essential for H2AZ incorporation (Gerhold & Gasser, 2014). The C-module of the SWR1 complex, which contains the Bromo-, YEATS- and SANT/DAMP1-domain mediates nucleosome affinity, while the N-module with the YL1- and Zinc finger ZNHIT1-domain is involved in the histone variant exchange reaction (Wu *et al.*, 2005; Gerhold & Gasser, 2014). SWR1 specific domains are highlighted in bold letters (Gerhold & Gasser, 2014; Willhoft & Wigley, 2020). **New SNF2 ATPase complex *T. brucei***: The identification of a BCNT- as well as a ZNHIT1 domain (highlighted in bold letters) hint toward the complex being a SWR1-like complex (Gerhold & Gasser, 2014; Willhoft & Wigley, 2020). Structure of the SWI2/SNF2 complex and its modules is only putative. The potential interaction interface of the species-specific proteins *TbSWRC3* and *TbARPx* is unknown (Vellmer *et al.*, 2021 in revision)

## 3.2 The *TbSWR1* complex regulates H2A.Z incorporation, transcriptional activity and chromatin structure

### 3.2.1 Characterisation of *TbSWR1*

For construction of an inducible RNAi cell line a section of the ORF (around 500 bp) of the 'gene of interest' GOI was amplified by PCR and cloned into two different sites of the pGL2084 plasmid. One fragment was integrated 5' to 3' the other one 3' to 5' to facilitate transcription of a hairpin mRNA structure to activate the RNAi system. The plasmid was linearized and transfected into 2T1 cells. The cell line possesses an incomplete *HYG* ORF. Correct integration of the plasmid completes the *HYG* ORF while removing a *PAC* ORF at the same time. This allows site specific integration of RNAi constructs to prevent clonal differences. Successful transfection will therefore, be controlled via counter-selection with puromycin since false integration will preserve the puromycin resistance. Expression of the hairpin mRNA is controlled by the so called "Tet off" system (Gossen & Bujard, 1992). A tetracycline repressor binds to the promoter region of the target sequence which prevents active transcription. The repressor disassembles after the addition of tetracycline or one of its derivatives to the parasite culture, resulting in transcriptional activation of the target sequence. Transcription of the RNAi construct was induced with 1µg/ml tetracycline. Analysis of *TbSWR1* RNAi cell lines showed that depletion of the SNF2 ATPase caused a severe phenotype with an almost complete growth arrest after 24 hours (Fig 15A).



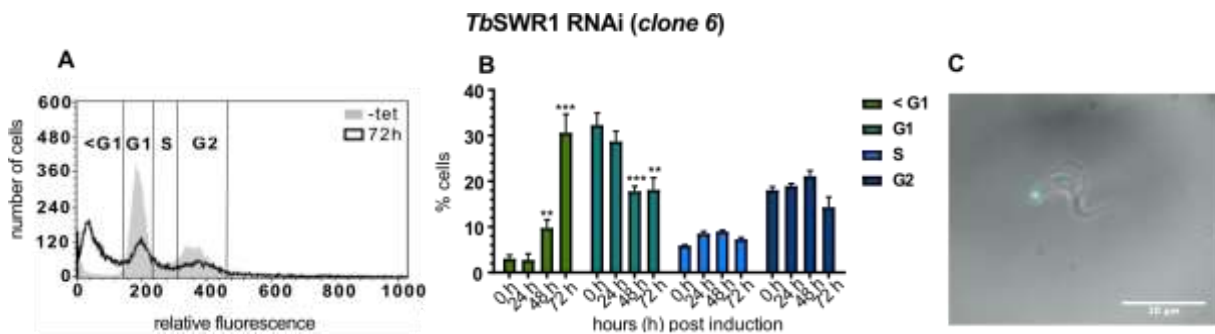
**Figure 15: Loss of *TbSWR1* leads to cell death**

**(A)** Growth of parasites was monitored for 96 hours after RNAi-mediated depletion of *TbSWR1* (Tb927.11.10730) using tetracycline (tet). The parental 2T1 cell line was used as a control (n=3). **(B)** Quantification of live/dead staining with propidium iodide of *TbSWR1*-depleted cells at the indicated timepoints post-induction. Analysis was done by flow cytometry (n=3). Vellmer *et al.*, 2021 in revision

### 3. Results

This is in contrast to previous data received from a large RNAi screen performed by Nicola Baker and colleagues who identified *TbSWR1* as a non-essential protein (Baker *et al.*, 2011). While ‘fluorescence activated cell sorting’ (FACS) analysis of propidium iodide-stained cells showed almost unaffected viability of the parasites 24 h subsequent to RNAi induction, the number of living cells decreased to around 70 % after 48 h and down to around 30 % after 72 h. At 96 h a slight increase of cell viability could be observed, which is consistent with an increase of cell growth (Fig 3.7A & B). Despite using the optimized 2T1 cell line that should reduce clonal differences, variations in the intensity of the phenotype could be observed between different clones. Analysis of a second *TbSWR1* RNAi clone confirmed my initial observation even though with less severe phenotypes (Fig 15B & C).

In the following I performed a cell cycle analysis to investigate how the loss of *TbSWR1* affects cell cycle progression in *T. brucei*. As part of the depletion of *TbSWR1a* significant increase of cells with a DNA content below that of G1 (sub-G1) cells could be observed (Fig 16A & B). FACS data also revealed a significant increase of anucleated cells, so-called “zoids”, within the sub-G1 population. The presence of zoids was confirmed by fluorescence microscopy (Fig 16C). 72 h post induction (p.i.) the number of sub-G1 cells increased up to 30 % while the number of cells in the G1- and G2-Phase decreased about 15 % and 5 %.



**Figure 16: Depletion of *TbSWR1* leads to the generation of anucleated cells**

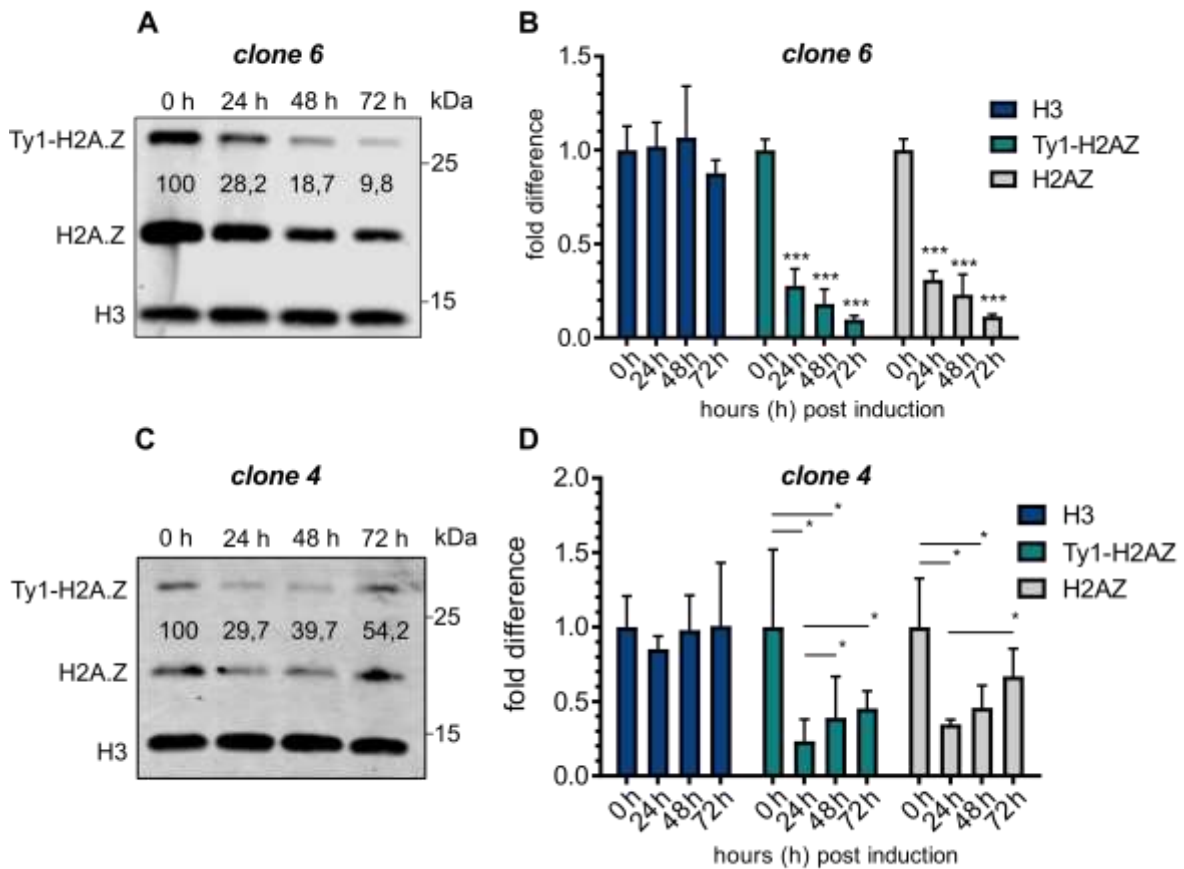
**(A)** Exemplary cell cycle profile of bloodstream form cells without (grey line) *TbSWR1* (Tb927.11.10730) depletion and after 72 h of protein depletion (black line). The Gates show the different populations of sub G1-, G1-, S- and G2-Phase cells. **(B)** Data of three triplicates, sub G1 Phase cells (green), G1-Phase cells (green-blue), S-Phase cells (light blue) and G2-Phase cells (dark blue). The data show a decrease of cells in G1 and G2 Phase in addition to the increase of sub G1-Phase cells (n=3 for all depicted experiments; \*\*\* = p-value <0.001; \*\* = p-value 0.001-0.01; \* = p-value 0.01-0.05). **(C)** Light microscopy images (N=1) of a BSF cell after 72h of *TbSWR1* depletion. Scale bar 10µm. Vellmer *et al.*, 2021 in revision

### 3.2.2 *TbSWR1* is required for H2A.Z incorporation in *T. brucei*

Since the *in-silico* analysis I performed could not completely confirm that the identified SNF2 complex is a SWR1-like complex, I decided to conduct WB-coupled cell fractionation experiments to determine nucleosome-associated H2A.Z levels. A decrease or increase of chromatin-associated H2A.Z in comparison to core histones following depletion of *TbSWR1* will show whether the SNF2 ATPase and the associated complex is INO80- or SWR1-related. By using an MNase based ChIP-seq analysis (Wedel & Siegel, 2017; Kraus *et al.*, 2020), I intended to investigate how H2A.Z levels develop at each TSS in the course of *TbSWR1* depletion. For both approaches cell lines were generated in which I tagged one allele of H2A.Z with a Ty1 tagging construct. The construct was amplified from a pPOTv7 vector (Dean *et al.*, 2015) in which the eYFP cassette was replaced by a Ty1 tag with two tandem epitopes. The amplified PCR product was then transfected into the clones C4 and C6 of the *TbSWR1* RNAi cell line. Functionality of the Ty1 tag was confirmed in course of the investigations that focused on HAT1 and HAT2 (Kraus *et al.*, 2020). The insoluble fraction of the cell fractionation that contains the chromatin was analysed by WB with antibodies against histone H3 and the variant H2A.Z. The cell fractionation was performed in triplicates for each clone. While levels of Ty1-tagged and non-tagged H2A.Z started to decrease significantly after 24 h of *TbSWR1* depletion, H3 levels appeared to be unaffected (Fig 17B & D). After 48 h RNAi induction the level of chromatin-associated H2A.Z was reduced around 60% (clone 4; Fig 17C) and 80 % (clone 6; Fig 17A) when compared to uninduced cells. This indicated that *TbSWR1* is required for active incorporation of H2A.Z into chromatin. While clones 4 and 6 showed an equally significant reduction of the H2A.Z signals after 24 hours, the signal progression in both clones was different. The H2A.Z signals continued to decrease at 48 and 72 h p.i. in clone 6 but the H2A.Z signals in clone 4 started to recover after 48 h p.i. (Fig 17B & D). In this regard the data correlate with the increased cell viability that was observed for clone 4 after 48 h of *TbSWR1* depletion. Cell fractionation experiments with *TbSWRC1*- and *TbSWRC2*-depleted cells confirmed the *TbSWR1* RNAi phenotype (Fig S5) strengthening my assumption that the identified complex is the SWR1 complex in *T. brucei*. I decided to perform the MNase ChIP-seq analysis with clone 4 of the 2T1 Ty1:H2A.Z *TbSWR1* RNAi<sup>TI</sup> cell line after 48 h of *TbSWR1* depletion. Since clone 6 aggregated a larger number of dead cells (Fig 15) over time, I intended to reduce any possible negative secondary effects that the decreasing cell fitness might have on chromatin integrity as much as possible.

### 3. Results

#### *Tb*SWR1 RNAi

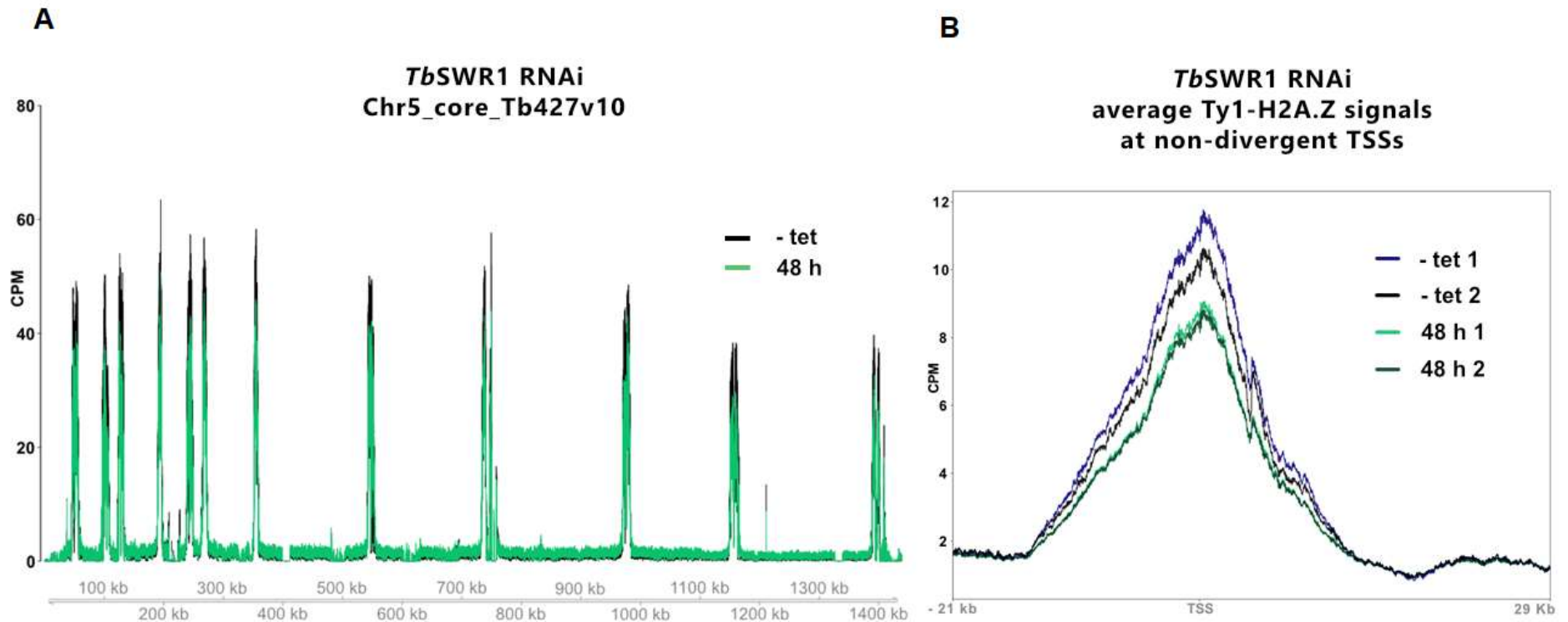


**Figure 17: Loss of *Tb*SWR1 leads to a reduction of chromatin-associated H2A.Z**

(A+C) Western blot analysis of the insoluble nuclear fraction with antibodies specific for histone H3 and the histone variant H2A.Z. Lysates from an equal number of cells ( $2 \times 10^6$  per lane) were analysed for each timepoint. Figures indicate the total amount of H2A.Z for each timepoint (B+D) Quantification of chromatin-associated H3 (dark blue), Ty1-H2A.Z (turquoise) and H2AZ (grey) (N=3 for all depicted experiments; \*\*\* = p-value <0.001; \*\* = p-value 0.001-0.01; \* = p-value 0.01-0.05 (Vellmer *et al.*, 2021 in revision)

The experiment was conducted in duplicates with the BB2 antibody to precipitate Ty1-tagged H2A.Z as only insufficient amounts of the  $\alpha$ -H2A.Z antibody were available. I used the MNase ChIP-seq protocol that has recently been used to investigate H2A.Z kinetics in course of HAT2 and HAT1 depletion with some minor adjustments (see chapter 2.6.8; Wedel & Siegel, 2017; Kraus *et al.*, 2020). MNase digestion efficiency was controlled by agarose gel electrophoresis. The evaluation showed that the input material mostly contained mono-nucleosomes (Fig S9B). In line with the data received from the cell fractionation experiments, both ChIP-seq replicates showed a reduction of Ty1-H2A.Z at TSSs after *Tb*SWR1 depletion (Fig 18; Fig S7-9; ChIP-seq data set 1). Based on the averaged total number of reads at TSSs, the amount of Ty1-H2A.Z decreased by around 30% (Fig 18B). These observations clearly confirmed that *Tb*SWR1 and the identified complex are SWR1-like and responsible for incorporation of the histone variant H2A.Z into chromatin in *T. brucei*.

### 3. Results



**Figure 18: Loss of *TbSWR1* (Tb927.11.10730) leads to a reduced H2A.Z deposition at TSS**

**(A)** ChIP-Seq analysis of distribution of Ty1-tagged H2A.Z before (black) and after (green) RNAi-mediated depletion of *TbSWR1* (48h post induction) revealed a reduction of chromatin associated Ty1-H2A.Z. Depicted is a representative region of chromosome 5. Data (n=2) were normalised to the total number of reads and plotted as counts per million reads (CPM).

**(B)** Average Ty1-H2A.Z signal across non-divergent TSSs. Peaks derived from two non-tetracycline induced reference samples depicted in blue and black. Peaks derived from two samples after 48 h of RNAi-mediated depletion of *TbSWR1* depicted in green and dark-green. Vellmer *et al.*, 2021 in revision

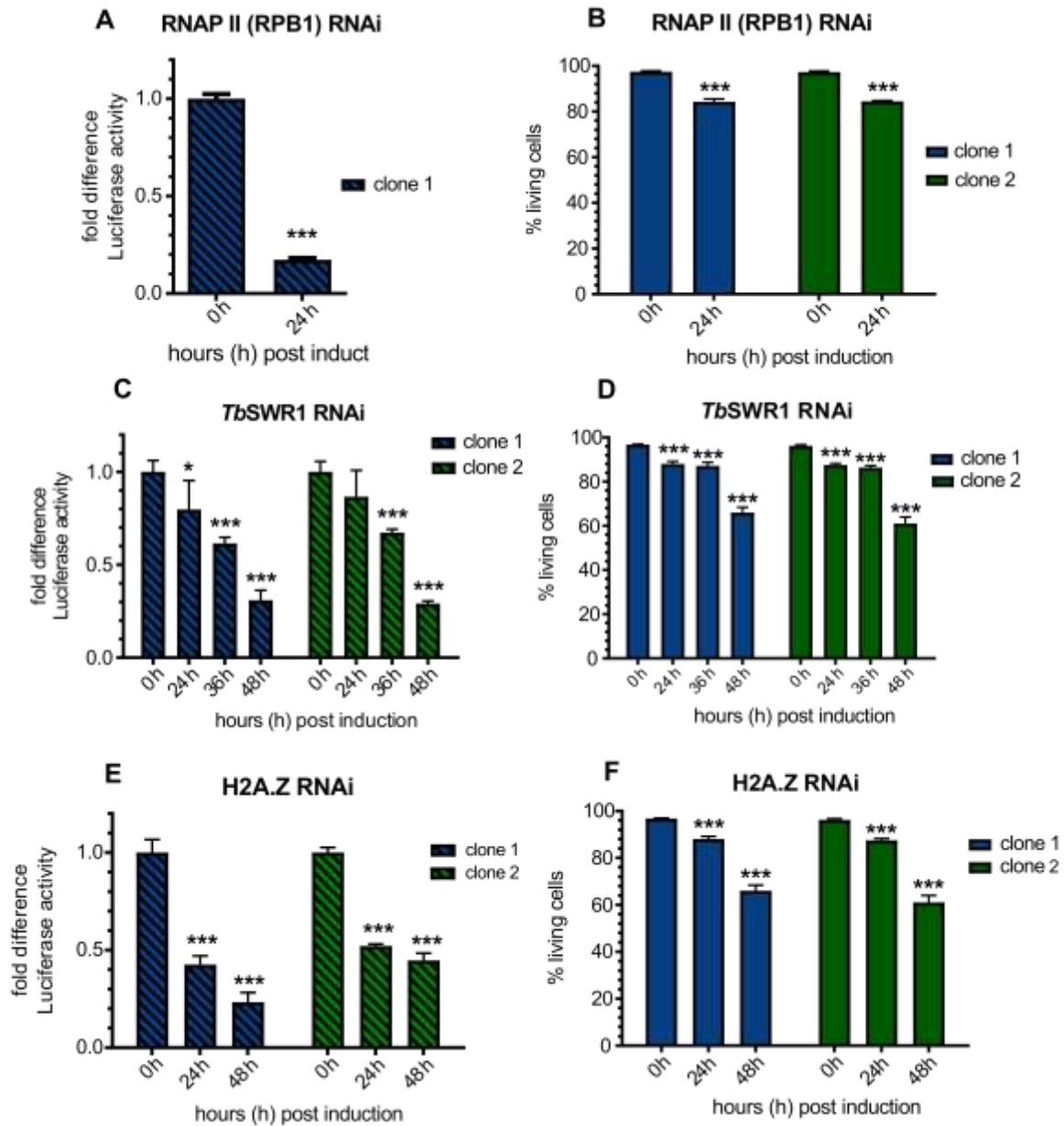
### 3.2.3 Depletion of *TbSWR1* leads to reduced transcriptional activity

Data from the research group of Nicolai Siegel could show HAT2 depletion caused a reduction of chromatin-associated H2A.Z and that HAT2 and HAT1 depletion led to a reduced amount of RNAP II derived transcripts (Kraus *et al.*, 2020). Based on the dependency of RNAP II-mediated transcription on H2A.Z acetylation and given that this PTM is directly dependent on the presence of nucleosome-associated H2A.Z, I intended to investigate if *TbSWR1* activity is required for RNAP II-mediated transcription. For this purpose, a single luciferase reporter construct was used, to investigate transcriptional activity subsequent to *TbSWR1* depletion. In course of investigations of VSG expression dynamics in *T. brucei* luciferase reporter assays became a highly reliable tool to investigate transcriptional activity (Davies *et al.*, 1997; Navarro & Cross, 1998; Batram *et al.*, 2014; López-Farfán *et al.*, 2014; Reis *et al.*, 2018). The polycistronic transcription in *T. brucei* allows almost genome-wide integration of the luciferase reporter construct to monitor RNAP II activity. I decided to integrate the reporter construct into the tubulin array of clone 6 of the 2T1 Ty1:H2A.Z *TbSWR1* RNAi<sup>Ti</sup> cell line. The tubulin array consists of multiple copies of  $\alpha$ - and  $\beta$ -tubulin ORFs, therefore increasing the chance of successful recombination of the reporter construct into this locus. I decided to generate single reporter cell lines that only contain the Renilla luciferase and to normalize the activity to the number of cells used for the experiment. Generation of a dual reporter cell line in which the Renilla activity is normalized to the activity of a Firefly luciferase located in an RNAP I-transcribed region for example, might lead to false results. A loss of RNAP II activity might affect RNAP I-mediated transcription since all genes that code for RNAP I subunits are transcribed by RNAP II. To monitor how a loss of RNAP II activity affects the luciferase activity, a cell line in which RPB1 can be depleted was generated. Depletion of RPB1 led to an expected decrease of luciferase activity. After 24 h the luciferase activity decreased more than 80 %. Live-dead analysis in contrast showed only around 20 % dead cells (Fig 19). Growth curves of two clones of the RPB1 RNAi cell line showed a growth arrest after 24 h but also a recovery after 72 h (Fig S11). In the following luciferase assays with *TbSWR1* RNAi cells revealed a reduction of luciferase activity of 70% after 48 h depletion of *TbSWR1* with only 40 % of the cells being dead (Fig 19). The data clearly showed that the loss of luciferase activity was always higher than the number of dead cells proving that the decrease in luciferase activity is a direct effect of the protein depletion and not caused by the reduced cell viability. Since mRNA levels are down regulated subsequent to HAT1 and HAT2 depletion (even though to a lesser extent in HAT2-depleted cells; Kraus *et al.*, 2020) I also integrated luciferase reporter constructs into HAT1 and HAT2 RNAi cell lines to compare the luciferase activity after *TbSWR1*, HAT1 and HAT2 depletion. After 48 h of HAT1 and HAT2 depletion the luciferase activity decreased by around 40 % and 30 %, while the number of dead cells did not drop below 85 % (Fig S10). Luciferase assays performed with *TbSWR1*-depleted cells indicated that RNAP II-mediated transcription is dependent on H2A.Z. To confirm this, I finally generated an H2A.Z RNAi luciferase reporter cell line.



### 3. Results

Subsequent to 48 h of H2A.Z depletion only around 30 % of the original luciferase activity could be detected while 60 % of the cells were still alive (Fig 19).



**Figure 19: Depletion of *TbSWR1* and H2A.Z caused a decrease of reporter luciferase activity within a PTU**

A single luciferase reporter construct was integrated into the tubulin array of all RNAi cell lines. The *TbSWR1* RNAi clone 6 was used for transfection (Fig 15). Samples for the luciferase assay were normalised to cell numbers. **(A)** As a positive control, luciferase activity of the same reporter construct was measured in an RPB I RNAi cell line. **(C+E)** Luciferase activity was monitored for 48 h after induction of RNAi in two independent clones. Values of non-induced cells were set to 1. **(B+D+F)** Live/dead staining of each RNAi cell line was performed in triplicates at the same time points. (n=3 for all depicted experiments; \*\*\* = p-value <0.001; \*\* = p-value 0.001-0.01; \* = p-value 0.01-0.05). Vellmer *et al.*, 2021 in revision

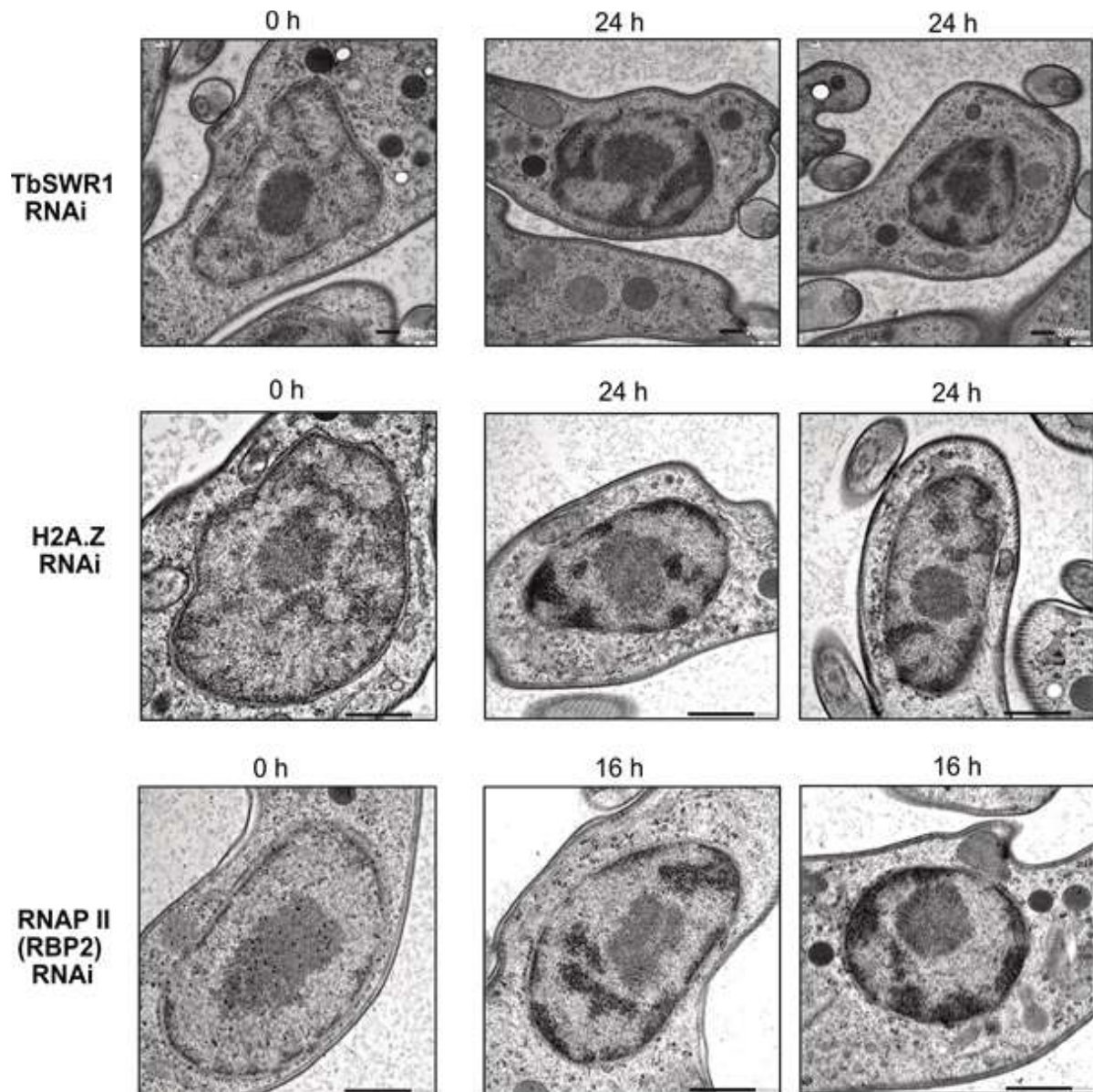
These data are in line with the data from the luciferase assays performed with *TbSWR1* RNAi cells. Combined with the data of the cell fractionation experiments and the ChIP-seq analysis, the luciferase assays I performed clearly show that the loss of transcriptional activity is caused by the loss of chromatin-associated H2A.Z.



#### **3.2.4 Loss of chromatin-associated H2A.Z affects chromatin structure**

H2A.Z plays an important role in limiting heterochromatin formation in eukaryotes and is linked with an open chromatin structure in *T. brucei* (Wedel *et al.*, 2017; Ryan & Tremethick, 2018), therefore I wanted to investigate potential changes in the chromatin structure after *Tb*SWR1 and H2A.Z depletion. Due to the availability of an electron microscopy (EM) facility, I decided to investigate potential changes in the chromatin structure via EM instead of performing HI-C (derived from 3C (chromosome conformation capture) sequencing or an 'Assay for Transposase-Accessible Chromatin' (ATAC) sequencing approach. EM sample preparation was performed with the assistance of Elisabeth Meyer-Natus. Evaluation of EM images of the nuclei of *Tb*SWR1 and H2A.Z-depleted cells showed a clear chromatin condensation after 24 h (Fig 21A & B). In the following the RPB1 RNAi cell line was used to exclude that the chromatin condensates in response to a lack of RNAP II. Chromatin condensation in response to a complete loss of RNAP II-mediated transcription has been reported in mice as well as humans (Marinozzi & Fiume, 1971; Roeder, 1976; Episkopou *et al.*, 2009). The evaluation of EM analysis of nuclei of cells in which RBP2 was depleted for 16 h also revealed a condensation of the chromatin indicating that the chromatin condensates in response to the cessation of RNAP II-mediated transcription and not in response to a loss of chromatin-associated H2A.Z. Nevertheless, it cannot be excluded that the chromatin condensation in *Tb*SWR1- and H2A.Z-depleted cells is partly caused by the loss of chromatin-associated H2A.Z.

### 3. Results



**Figure 21: Loss of *TbSWR1* and RNAP II leads to chromatin condensation**

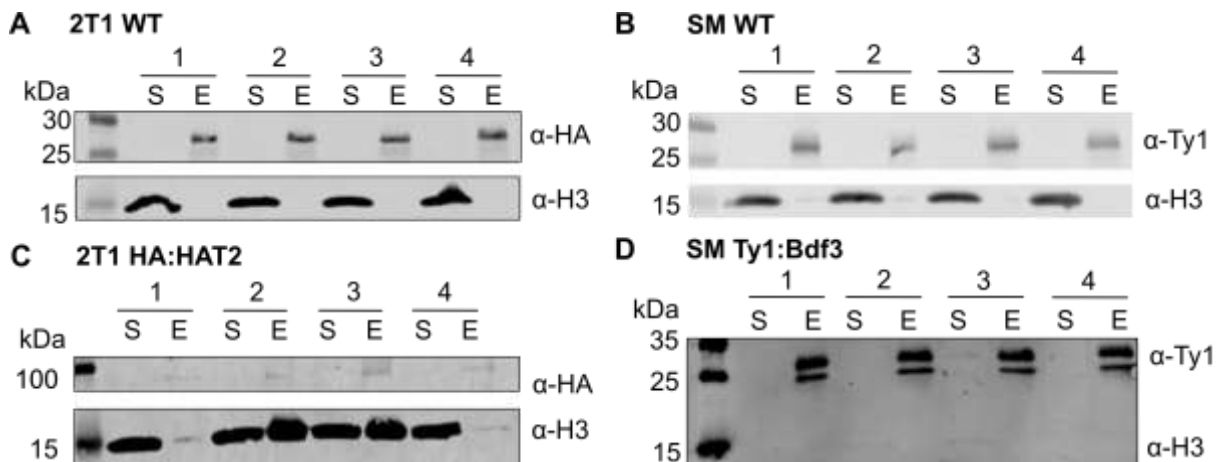
Illustrative electron microscopy images of nuclei of *TbSWR1*-, H2A.Z- and RNAP II-depleted and non-depleted parasites as indicated. Depletion of the proteins led to large black patches of condensed chromatin. Scale bar, 200 nm for *TbSWR1* RNAi and H2A.Z RNAi images, 500 nm for RNAP II (RBP2) RNAi images. The images of *TbSWR1* RNAi cells were obtained using a transmission electron microscope, images of RNAP II RNAi cells with a scanning transmission electron microscope. (Vellmer *et al.*, 2021 in revision)

#### 3.3 The HAT1 and the HAT2 complex

One of the main goals of this project was the identification of a NuA4-like complex in *T. brucei* that acetylates histones H2A and or H4 prior to H2A.Z incorporation by *Tb*SWR1. At the start of my search for a NuA4-like complex, only little information was available about H2A/H4 acetylation and the incorporation of H2A.Z. The histone acetyltransferase HAT2 was identified as being responsible for acetylation of H4K10 which was found to be enriched at TSSs (Kawahara *et al.*, 2008; Siegel *et al.*, 2009) H4K10 corresponds to H4K12 in other eukaryotic organisms (Fig S1; Kawahara *et al.*, 2008) and acetylation of H4K12 was shown to be essential for correct incorporation of H2A.Z by SWR1. A search in the TriTryp database for potential subunits of a NuA4 complex identified the protein Tb927.9.2910 as homologous to the 'ESA1-associated factor 6' (Eaf6) subunit of the *S. cerevisiae* NuA4 complex (In the following referred to as *Tb*Eaf6). Since HAT2 corresponds to 'essential SAS2-related acetyltransferase 1' (Esa1) in its function I assumed that HAT2 and *Tb*Eaf6 form a complex that possesses characteristics of a NuA4 complex. Based on these findings individual cell lines with N-terminally HA-tagged HAT2 and N-terminally HA-tagged *Tb*Eaf6 were generated. In both cell lines I replaced one wild type allele of HAT2 and *Tb*Eaf6-like with a resistance marker and tagged the remaining allele with an HA-tagging construct. The construct was amplified via PCR from the pPOTv7-HA vector that has previously been used for the *Tb*SWR1 co-IPs. Correct integration of the tagging construct was confirmed by diagnostic PCR (Fig S12A & C). With subsequent co-IP experiments I intended to identify a histone acetyltransferase complex with characteristics of a NuA4 complex assembled around HAT2 and *Tb*Eaf6. In the context of this approach data published by the research group of Nicolai Siegel revealed that HAT2 only acetylates lysines on histone H4 and is required for correct H2A.Z incorporation indicating that HAT2 corresponds to the NuA4 histone acetyltransferase Esa1. In addition, they could show that acetylations on H2A.Z are catalysed by HAT1 and not HAT2. These data suggest the presence of two distinct histone acetyltransferase complexes assembled around HAT1 and HAT2 or the presence of a hybrid complex assembled around both HATs.

### 3.3.1 The HAT2 complex has a so far undescribed complex composition

Prior to the HAT2 co-IP I performed a control co-IP with the HA-antibody 12CA5 and unmodified 2T1 cells as a negative control (Fig 22A). The WB analysis once again revealed the characteristic bands at around 25kDa that originate from the light chain of the antibody (Fig 22A). WB analysis of the soluble fraction of 2T1 HA:HAT2 cells revealed no signal (Fig 22C). Only a thin band migrating below 100 kDa could be detected in the elution of the HAT2 co-IP (Fig 22C). Given that the molecular weight of the protein including the 3xHA tag is around 71 kDa, the band appears to correspond to the tagged HAT2.



**Figure 22: Test Western Blot of the co-IP experiments with HA-HAT2 and Ty1-Bdf3**

Representative western blots of samples taken during the co-IP of (A) 2T1 WT cells, (B) of SM WT cells, (C) of 2T1 HA:HAT2 cells (~71 kDa) and (D) of SM Ty1:Bdf3 (~31 kDa). S (soluble supernatant), E (eluate: 10-fold compared to the other samples). Blots were probed with 12CA5 anti-HA and anti-H3 antibodies (A+C) or BB2 anti-Ty1 and anti-H3 antibodies (B+D). (Vellmer *et al.*, 2021 in revision).

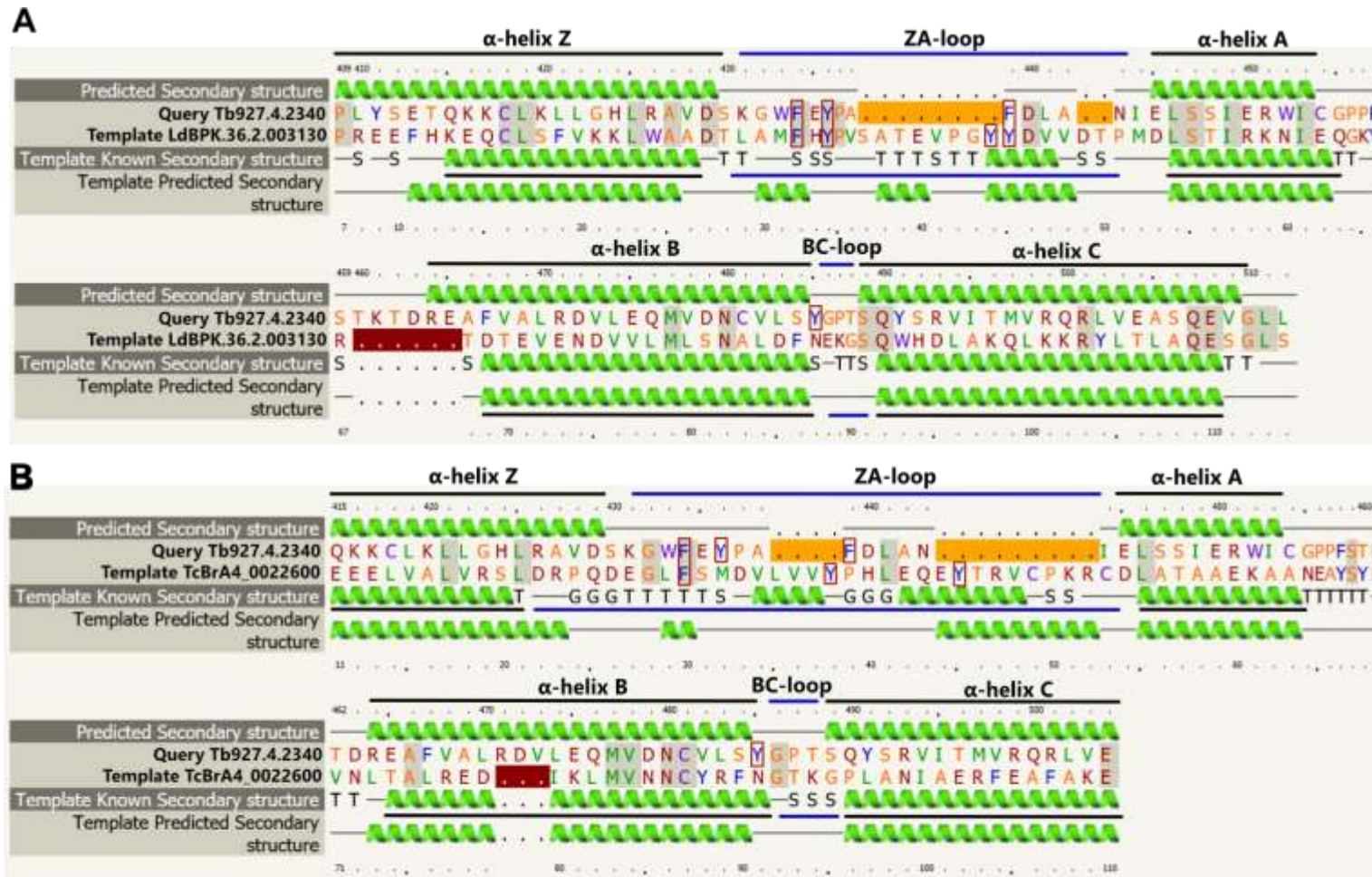
Mass spectrometry analysis of the eluates received from the HA-HAT2 co-IP experiments identified HAT2 and 11 additional proteins. Two of these proteins had a negative NES indicating that these proteins were contaminants (Table 14; Fig S13A; MS data set 2). Among the nine proteins that had a positive NES were the bromodomain factors (Bdfs) Bdf3 and Bdf5 (Table 14; Fig S13A; MS data set 2). Besides SUMO six hypothetical proteins were enriched in the HAT2 co-IP. Phyre2 analysis of these six proteins linked parts of three proteins with known secondary structures of certain domains (Table 3.3). A part of the core of Tb927.3.4140 was identified as being homologous to a btb-domain which is known to mediate protein-protein interactions (Stogios *et al.*, 2005). The coverage of the generated model includes only 7 % of the amino acid sequence and since the confidence in the modelled secondary structure is below 80 %, it is questionable whether this structure corresponds to a btb-domain or not. With a coverage of 26 % and a confidence of almost 94 % the modelled secondary structure of the C-terminus of Tb927.9.13320 was identified as being homologous to the 'forkhead' (FHA; Fig S15) domain which is associated with the binding of phosphopeptides (Durocher *et al.*, 2000 reviewed in Almawi *et al.*, 2017). The most interesting result was the identification of a potential third Bdf within the HAT2 complex.

### 3. Results

Phyre2 modelled 98 amino acids of the C-terminus of Tb927.4.2340 onto the structure of the bromodomain of the *Leishmania donovani* protein LdBPK.36.2.003130 (confidence ~97% and a coverage of 20 %) 87 amino acids (confidence ~97% and coverage of 17 %) were modelled onto the bromodomain structure of the *Trypanosoma cruzi* bromodomain factor TcBrA4\_0022600 respectively (Fig 23). Since the presence of a third bromodomain factor appeared doubtful I analysed the modelled sequence in more detail. A typical bromodomain consists of four  $\alpha$ -helices named Z, A, B and C (from N- to C-terminus; Zeng & Zhou, 2002; Mujtaba *et al.*, 2007). Hydrophobic amino acids like tyrosines and phenylalanines are the key amino acids required for binding of the acetylated lysine and can be found in the ZA loop that link the corresponding  $\alpha$ -helices (Zeng & Zhou, 2002; Mujtaba *et al.*, 2007). The BC- loop often contains a mix of amino acids with polar uncharged side chains like serine (ser), threonine (thr) or asparagine (asn) and non-polar amino acids like glycine (Zeng & Zhou, 2002; Mujtaba *et al.*, 2007). Both Phyre2 alignments indicate that the C-terminus of Tb927.4.2340 forms the typical secondary structure of a bromodomain including the ZA and the BC loop (Fig 23). The predicted ZA loop of Tb927.4.2340 contains two phenylalanines and one tyrosine. Even though the composition of the amino acids that potentially form the BC-loop in Tb927.4.2340 differs from the composition of the BC-loops from LdBPK.36.2.003130 and TcBrA4\_0022600 (Fig 23), the predicted BC-loop of Tb927.4.2340 contains two polar but uncharged amino acids. According to the Phyre2 data I named Tb927.4.2340 Bdf6 in the following. To confirm these initial co-IP results, I intended to perform a reciprocal co-IP experiment with Bdf3.



### 3. Results



**Figure 23: Predicted secondary structure of the C-terminal end of Tb927.4.2340 is homologous to known bromodomains in *T. cruzi* and *L. donovani***

Phyre2 modelling of Tb927.4.2340 could link the core of the protein to the bromodomains of LdBPK.36.2.003130 (A) and TcBrA4\_0022600 (B). In the alignments the four α-helices Z, A, B and C are highlighted with black lines, the linking loops are highlighted with blue lines. The conserved amino acids required for binding of the acetylation are highlighted with red boxes. The chemical properties of the amino acids located in the BC-loop are indicated. hy= hydrophobic, po. un. = polar but uncharged, ch. p = charged positive, ch. n. = charged negative, - = non-polar. Predicted and known α-helices are depicted in light-green. The alignment was created using the Phyre2 online modelling platform (Kelley *et al.*, 2015).

**C**

	<u>BC-loop</u>					<u>BC-loop</u>					
	hy	-	-	po.	po.	hy	-	-	po.	po.	
				un.	un.				un.	un.	
Tb927.4.2340	Tyr	Gly	Pro	Thr	Ser	Tb927.4.2340	Tyr	Gly	Pro	Thr	Ser
LdBPK.36.2.003130	Asn	Glu	Lys	Gly	Ser	TcBrA4_0022600	Asn	Gly	Thr	Lys	Gly
	po.	ch.n	ch.p	-	-		po.	-	po.	ch.p	-
	un.						un.		un.		



### 3. Results

For the reciprocal co-IP experiment, a cell line was generated in which one Bdf3 allele was N-terminally tagged with two tandem epitopes of a Ty1-tag (Fig S12B). For Ty1-tagging of the target proteins the 3xHA-tag in the pPOTv7-HA vector was exchanged with two epitopes of a Ty1-tag. The tagging construct was amplified by PCR. The amplified tagging construct was then transfected into the SM cell line. Since only one Bdf3 allele was tagged, I used  $2 \times 10^8$  cells instead of  $1 \times 10^8$  cells for the co-IP to compensate for the Ty1-Bdf3 expression level. Prior of the Bdf3 co-IP, I performed a control co-IP with the BB2-antibody and unmodified 2T1 (Fig 22D). Here the WB analysis also revealed the characteristic bands at around 25kDa that originate from the light chain of the BB2 antibody (Fig 22B). WB analysis of the Ty1-Bdf3 co-IP revealed two bands migrating above 25 kDa. Given that the 2xTy1-tagged Bdf3 has a molecular weight of around 31 kDa the upper band is likely to originate from this protein (Fig 22D). Except for the protein with the accession number Tb927.3.4140, MS analysis of the Bdf3 co-IP identified the same set of proteins already identified in the HAT2 co-IP experiment, including Bdf5 and Bdf6 (Table 14; Fig S12B; MS data set 2).

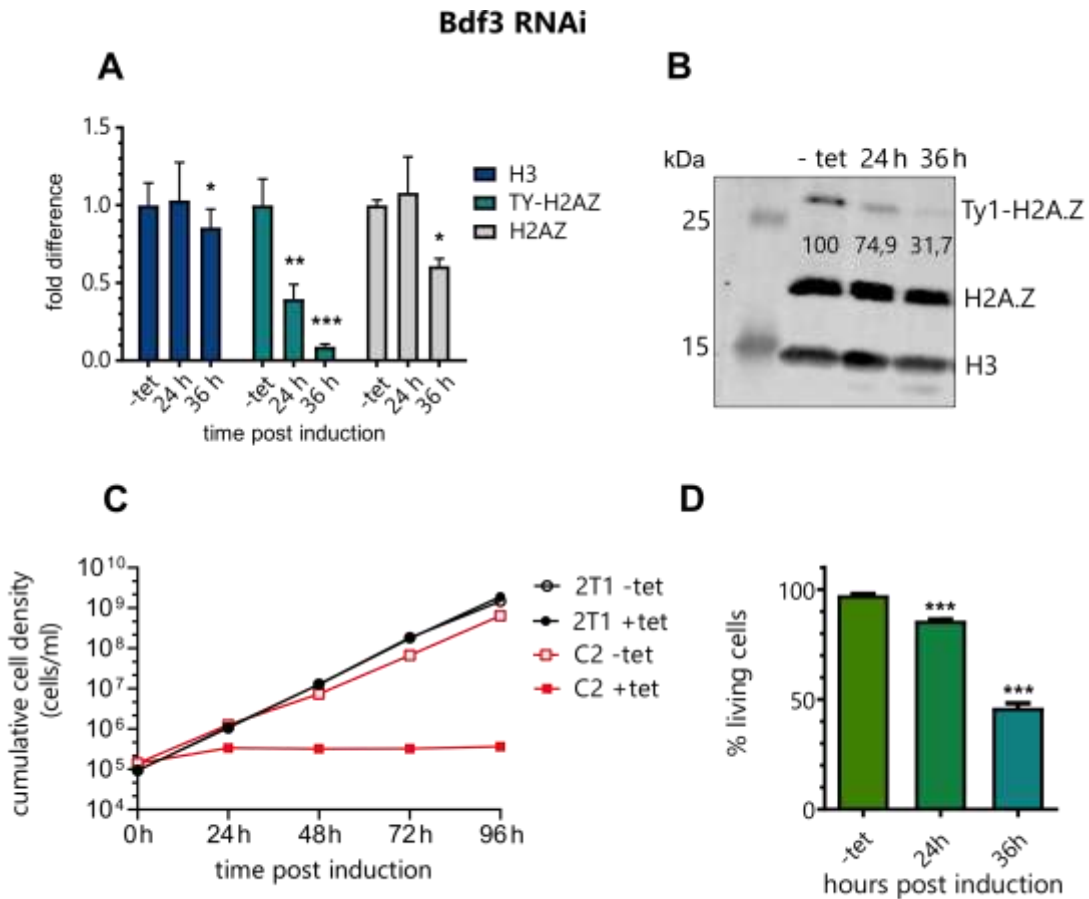
**Table 14: Summary of the HA-HAT2 and Ty1-Bdf3 co-IP**

Co-IP	Gene ID	Annotation	Identified domains	Phyre2 modelling	NES	Ident. in Co-IP No.
No. 1	Tb927.11.11530	Histone acetyl-transferase HAT2	SAS2 superfamily	C-terminus: TIP60 (Cov. 67% Conf. 100%)	5.35	1+2
No. 2	Tb927.11.10070	Bromodomain, putative, Bdf3	Bromo-domain	-	2.3	1+2
	Tb927.3.4140	hypothetical protein		btb domain (Cov. 7%; Conf. 73.3)	1.55	1
	Tb927.4.2340	hypothetical protein	-	- bromo-domain from <i>Leishmania donovani</i> complexed with bromosporine (Cov. 20%; Conf. 97%); - crystal structure of tcbdf5 (Cov. 17%; Conf. 97%)	3.25	1+2
	Tb927.5.3210	small ubiquitin-related modifier	UBQ/SUMO	-	3.05	1+2
	Tb927.6.1070	hypothetical protein	-	-	4.82	1+2
	Tb927.7.2770	hypothetical protein	-	-	5.81	1+2
	Tb927.9.13320	hypothetical protein		SMAD/FHA domain (Cov. 26%; Conf. 93%)	2.66	1+2
	Tb927.11.5230	hypothetical protein	-	ENT-domain of <i>T. brucei</i> (Cov. 35%; Conf. 100%)	2.32	1+2
	Tb927.11.13400	Bromodomain, putative, Bdf5	Bromo-domain	-	3.18	1+2

10 proteins were identified via mass spectrometry in two Co-IP experiments. The initial Co-IP was performed with Tb927.11.11530, the reciprocal Co-IP with the protein Tb927.11.10070 was performed to confirm the Tb927.4.2000 co-IP data. Only Tb927.3.4140 could be identified in the initial but not in the reciprocal Co-IP experiment. The "Annotation" column indicates the curated annotation that was found for the corresponding accession number in the TriTyp database. The "Identified domains" column displays the domains that were found by BLAST search using the NCBI database. The Phyre2 modelling column indicates proteins that were identified by homology modelling. Coverage (Cov.) indicates the coverage in percent between query and template. The confidence (Conf.) represents the relative probability in percent (from 0 to 100) that the match between query and template is a true homology. The nuclear enrichment score (NES) indicates a nuclear localization if positive. The last column shows in which of the two Co-IPs the protein could be identified (modified from Vellmer *et al.*, 2021).

### 3. Results

In the following I analysed if Bdf3 is involved in the incorporation of H2A.Z. Therefore I generated a Bdf3 RNAi construct and used it for the transfection of a 2T1 Ty1:H2A.Z cell line. A growth curve of RNAi induced cells revealed a growth arrest after 24 h which is comparable to *TbSWR1*-depleted cells (Fig 24C). Cell fractionation experiments revealed a significant reduction of chromatin-associated H2A.Z subsequent to the depletion of Bdf3 (Fig 3.15A & B).



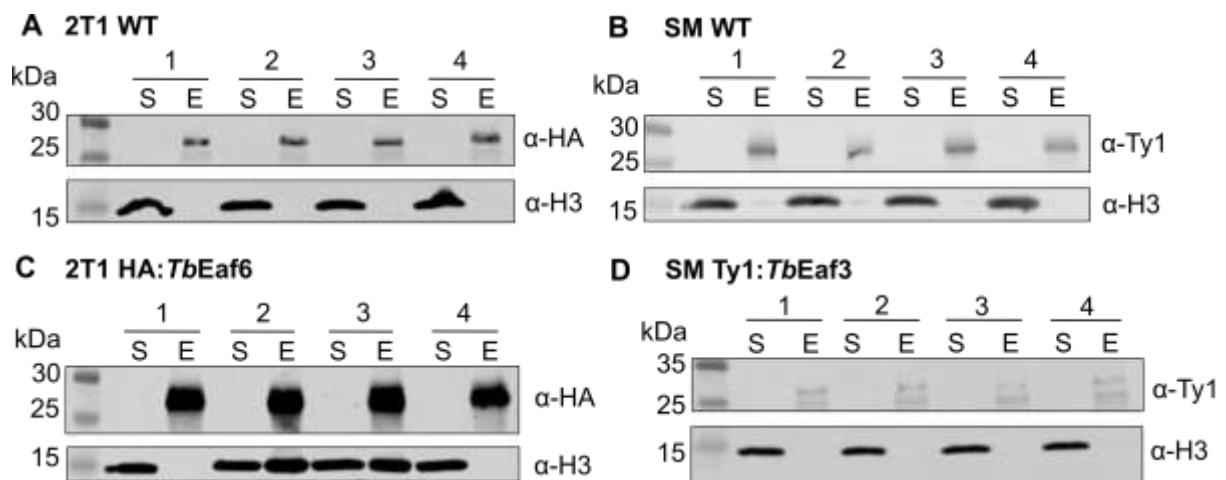
**Figure 24: Loss of Bdf3 leads to cell death and a reduction of chromatin-associated H2A.Z**

**(A)** Quantification of chromatin-associated H3 (dark blue), Ty1-H2A.Z (turquoise) and H2AZ (grey); (N=3 for all depicted experiments; \*\*\* = p-value <0.001; \*\* = p-value 0.001-0.01; \* = p-value 0.01-0.05). **(B)** Western blot analysis of the insoluble nuclear fraction with antibodies specific for histone H3 and the histone variant H2A.Z. Figures on the WB image indicate the alterations of the H2A.Z levels in comparison to -tet in percentage **(C)** Growth of parasites was monitored for 96 hours after RNAi-mediated depletion of Bdf3 using tetracycline (tet) induction. The parental 2T1 cell line was used as a control (n=3). **(D)** Quantification of live/dead staining with propidium iodide of Bdf3-depleted cells. Analysis was done by flow cytometry (n=3). Vellmer *et al.*, 2021 in revision

While a significant decrement of Ty1-tagged H2A.Z can be observed after 24 h a significant decrement of non-tagged H2A.Z is only detectable after 36 h of RNAi induction. After 24 h RNAi induction the amount of chromatin-associated H2A.Z was reduced around 25 % and around 65 % after 36 h (Fig 24A). The number of living cells dropped down to almost 45 % after 36 h of Bdf3 depletion (Fig 24). Analysis of cell fractionation experiments performed by Laura Hartleb with Bdf5 and Bdf6 RNAi showed that depletion of Bdf5 led to a reduction of chromatin-associated H2A.Z while depletion of Bdf6 did not (personal communication with Laura Hartleb).

### 3.3.2 The HAT1 complex has characteristics of a NuA4 complex

The HA-*TbEaf6*-like co-IP was performed together with the HAT2 co-IP, so the 2T1 negative control also acted as a negative control for the HA:*TbEaf6* co-IP (Fig 25A). WB analysis of the soluble fraction derived from 2T1 HA:*TbEaf6* cells did not show any signal of HA:*TbEaf6* (Fig 25C). WB analysis of the precipitated material then revealed a large “double” band migrating above 25 kDa (Fig 25C) but no signal below 25 kDa which was surprising as the expected molecular weight of the HA-tagged *TbEaf6* would be around 23 kDa. Due to the presence of PTMs on *TbEaf6* it might be possible that the protein is migrating above its calculated weight therefore migrating together with the light chain of the 12CA5 antibody and creating a large/double band in the WB analysis (Fig 25C; Carruthers *et al.*, 2015).



**Figure 25: Test Western Blot of the co-IP experiments with HA-*TbEaf6* and Ty1-*TbEaf3***

Representative western blots of samples taken during the co-IP of (A) 2T1 WT cells, (B) of SM WT cells, (C) of 2T1 HA:*TbEaf6* cells (~23 kDa) and (D) of SM Ty1:*TbEaf3* (~25 kDa). S (soluble fraction), E (eluate: 10-fold compared to the other samples). Blots were probed with 12CA5 anti-HA and anti-H3 antibodies (A+C) or BB2 anti-Ty1 and anti-H3 antibodies (B+D) (Vellmer *et al.*, 2021 in revision).

Mass spectrometry analysis of the *TbEaf6* co-IP identified the protein, therefore confirming that tagging of the protein was successful (Table 15; Fig S14; MS data set 3). Besides *TbEaf6*, 10 additional proteins with a positive or non-defined NES were identified (Table 15; Goos *et al.*, 2017). Among these proteins were the two histone acetyltransferases HAT1 and HAT3. A YEATS-domain containing protein, a PHD-Zinc-finger-domain containing, 5 hypothetical proteins and a conserved protein were identified as well. NCBI BLAST and Interpro analysis of these 10 proteins only revealed a bromodomain in Tb927.1.3400 and a chromo-barrel-domain in the N-terminus of HAT1 (Table 15; Fig S14; MS data set 3). In response to these results, I once again used the Phyre2 modelling tool to learn more about the hypothetical proteins and the conserved protein and the results were very interesting since five of the 11 proteins could be linked to known subunits of the NuA4 complex.

### 3. Results

**Table 15: Summary of the co-IPs that identified the HAT1 complex**

Co-IP	Gene ID	Annotation	NES	Ident. in co-IP No.	Identified domain	Phyre2 modelling	Yeast NuA4 subunit	Domain(s)
No. 2	<i>TbEaf3</i> (Tb927.1.650)	conserved protein, unknown function	4.38	1+2	-	MSL3 like MRG domain (Cov. 96% Conf. 95%)	Eaf3/MORF4	Chromo- & MRG-domain
	Tb927.7.4560	Histone acetyltransferase 1	2.84	1+2	tudor-knot Chromo-like domain MYST HAT	C-terminus: TIP60 (Cov. 66% Conf. 100%) N-terminus: knotted tudor domain of Esa1 (Cov. 15% Conf. 99%)	Esa1	Tudor-knot Chromo-like domain & MYST HAT
	Tb927.7.5310	YEATS family, putative	6.59	1+2	YEATS-Domain	Yaf9 /GAS41 (Cov. 14% Conf. 99%)	Yaf9	YEATS
No. 1	<i>TbEaf6</i> (Tb927.9.2910)	histone acetyltransferase subunit NuA4	4.59	1+2	Eaf6	n.p	Eaf6	Eaf6
	Tb927.10.14190	hypothetical protein, conserved	5.72	1+2	-	Epl1 (Cov. 63% Conf. 98%)	Epl1	EpcA
	Tb927.1.3400	hypothetical protein, conserved	5.40	1+2	Bromodomain	Bdf5 T.c (Cov. 37% Conf. 97%)		
	Tb927.6.1240	hypothetical protein, conserved	N/A	1		-		
	Tb927.8.5320	hypothetical protein, conserved	2.23	1+2		-		
	Tb927.10.8310	histone acetyltransferase 3	N/A	1	MYST HAT	n.p		
	Tb927.10.9930	PHD-zinc-finger like domain	N/A	1		n.p		
	Tb927.11.3430	hypothetical protein, conserved	N/A	1+2	-	-		

11 proteins were identified via mass spectrometry in two co-IP experiments. The initial Co-IP was performed with Tb927.9.2910, the reciprocal co-IP with the protein Tb927.1.650 was performed to confirm the Tb927.9.2910 co-IP data. Tb927.6.1240, Tb927.10.8310 and Tb927.10.9930 could only be identified in the initial but not in the reciprocal co-IP experiments. The "Annotation" column indicates the curated annotation that was found for the corresponding accession number in the TriTyp database. Proteins labelled in green exhibit a homolog in the *S. cerevisiae* NuA4 complex. The nuclear enrichment score (NES) indicates a nuclear localization if positive. The "Ident. in co-IP" column shows in which of the two Co-IPs the protein could be identified. The "Identified domains" column displays the domains that were found by BLAST search using the NCBI / Interpro database. The "Phyre2 modelling" column indicates proteins that were identified by homology modelling. Coverage (Cov.) indicates the coverage in percent between query and template. The confidence (Conf.) represents the relative probability in percent (from 0 to 100) that the match between query and template is a true homology. The "Yeast NuA4 subunit" column states the corresponding NuA4 complex subunit with its domain ("domain(s)" column) to which the identified trypanosome protein is homologous to (modified from Vellmer *et al.*, 2021)

### 3. Results

While the predicted secondary structure of the N-terminus of HAT1 appears to be homologous to the knotted-tudor-domain (a chromo-like-domain) of Esa1, the acetyltransferase subunit of the *S. cerevisiae* NuA4 complex, the predicted secondary structure of the C-terminus of HAT1 was identified as being homologous to various histone acetyltransferases including TIP60 the human equivalent to the yeast histone acetyltransferase Esa1 (Doyon *et al.*, 2004). Phyre2 modelling of Tb927.10.14190 could align 63% of its predicted secondary structure with 'enhancer of polycomb-like 1' (Epl1). Epl1 as a part of the NuA4 complex is required for acetylation of the histones H4 and H2A and in cooperation with Esa1 for chromatin structure regulation (Searle *et al.*, 2017). Loss of function of Epl1 is associated with suppression of the telomere position effect (Boudreault *et al.*, 2003). The predicted secondary structure of Tb927.1.650 was identified as being homologous to the "mortality factor 4' (MORF4) related gene' (MRG)-domain of "male specific lethal' (MSL) complex subunit 3' (MSL3)/Esa3 which is also part of the NuA4 complex (Eisen *et al.*, 2001). The MRG-domain is required for protein interactions within the complex and the secondary structure of the protein shares similarities with known DNA-binding domains (Bowman *et al.*, 2006). The putative secondary structure of the YEATS-domain of Tb927.7.5310 appears to be homologous to Yaf-9/GAS41 which are subunits of the SWR1 and NuA4 complex (Yaf-9) as well as the TIP60 complex (GAS41) and are responsible for binding of the acetylation of lysine 27 on histone H3 (H3K27; Klein *et al.*, 2018 reviewed in Lu *et al.*, 2009). Phyre2 results clearly indicate the presence of a complex with characteristics of a NuA4-like complex. Due to the identification of HAT1 and HAT3 in the initial co-IP, I hypothesized that the initial co-IP precipitated two distinct HAT complexes that share at least TbEaf6-like as a common subunit. To test this hypothesis, I decided to perform a second co-IP with the Eaf3-like protein Tb927.1.650 which I named *TbEaf3* in the following. I generated a cell line in which one *TbEaf3* allele was N-terminally tagged with two tandem epitopes of a Ty1-tag (Fig S11D). For amplification of the tagging construct I used the pPOTv7-Ty vector, that was constructed in course of the generation of the SM Ty1:Bdf3 cell line. Like for the Ty1:Bdf3 co-IP I used  $2 \times 10^8$  cells instead of  $1 \times 10^8$  cells for the co-IP since only one of the two alleles was tagged. WB analysis of the eluted *TbEaf3* co-IP material revealed a band migrating at 25 kDa and a second one slightly above (Fig 25D). The expected molecular weight of tagged *TbEaf3* would be around 25 kDa. Since the band that migrates slightly above the 25 kDa marker was also detected in the WT control, the lower band appeared to be the tagged *TbEaf3*. Like for HAT2, *TbEaf6* and Bdf3 no tagged protein could be observed in the WB analysis of the soluble fraction. Mass spectrometry analysis of the material precipitated in course of the *TbEaf3* co-IP confirmed my initial co-IP results. Except for HAT3, the PHD-Zinc-finger like domain containing protein Tb927.10.9930 and the hypothetical protein Tb927.6.1240 all other proteins identified in the first co-IP could be identified again, which supported my hypothesis of two distinct HAT complexes.

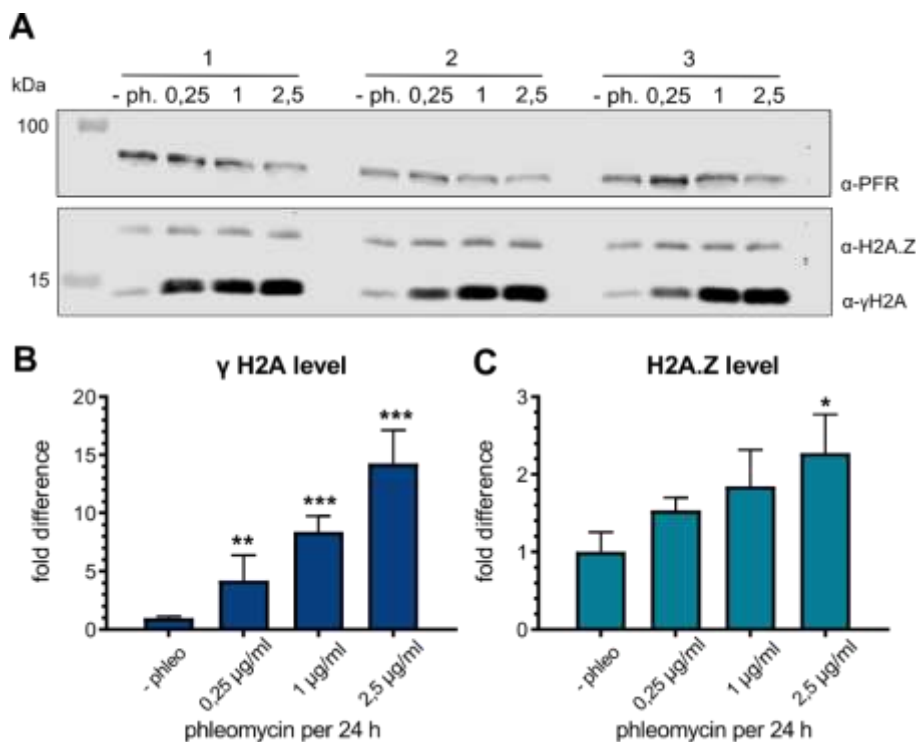
### 3. Results

At this point it appears that *TbEaf6* is part of the HAT1 and HAT3 complex while *TbEaf3* is a unique complex component of the HAT1 complex. Further investigations are required to assess whether the two complexes share additional subunits.

#### 3.3.3 H2A.Z appears to be involved in the DNA DSB response in *T. brucei*

The NuA4 complex and H2A.Z play a key role not only in transcription regulation but also in DNA DSB response. HAT1 appears to be homologous to the NuA4 complex but does not regulate H2A.Z incorporation into chromatin (Kraus *et al.*, 2020). Regarding the activation of the H2A.Z incorporation/removal pathway in course of occurring DNA DSBs, the identification of two distinct acetyltransferase complexes responsible for H2A.Z incorporation and H2A.Z acetylation raises a lot of questions. The most important question is whether H2A.Z in *T. brucei* is also involved in DNA DSB repair or not. If the histone variant is involved in the DNA DSB repair response it must be analysed whether both HAT complexes are involved or if the DNA repair pathway only requires activity of the HAT2 complex.

To answer the first question, I treated SM WT cells with different concentrations of phleomycin for 24 h. Phleomycin is a glycopeptide antibiotic of the bleomycin family (Maeda *et al.*, 1956) and generates DNA DSBs (Dorr, 1992). With subsequent cell fractionation experiments I wanted to analyse whether the amount of chromatin-associated H2A.Z increases in response to occurring DNA DSBs or not. I normalized  $\gamma$ H2A and H2A.Z signals with the signal of the paraflagellar rod protein (PFR) since it was reported that histone degradation can occur in course of increasing DNA damage (Hauer *et al.*, 2017). WB analysis of the insoluble chromatin fraction revealed an increase of H2A.Z levels correlating with the increase of the phleomycin concentration used (Fig 26). Nevertheless, it has to be mentioned that the increase of H2A.Z levels was lower than the increase of  $\gamma$ H2A levels.



**Figure 26: Chromatin-associated H2A.Z level increase after phleomycin treatment.**

**(A)** Western blot analysis of the insoluble nuclear fraction with antibodies specific for PFR, H2A.Z and  $\gamma$ H2A lysates from an equal number of cells ( $2 \times 10^6$  per lane) were analysed for each concentration ( $n=3$ ). **(B)** Quantification of chromatin-associated  $\gamma$ H2A **(C)** Quantification of chromatin-associated H2A.Z (N=3 for all depicted experiments; \*\*\* = p-value <0.001; \*\* = p-value 0.001-0.01; \* = p-value 0.01-0.05).

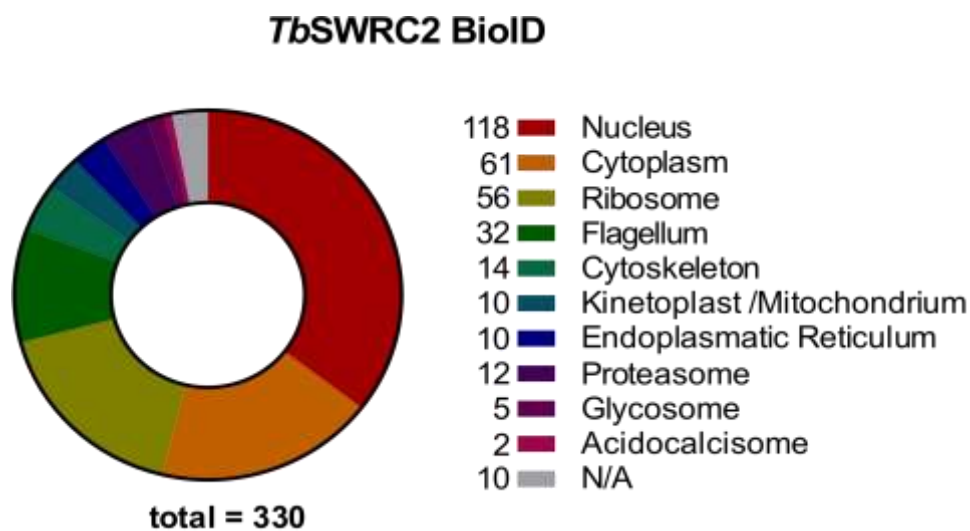
#### 3.4. Identification of potential COMPASS and SAGA complex subunits

The data published by Kraus et al, revealed various new PTMs in *T. brucei*. One of the most interesting PTMs was the acetylation on lysine 23 on histone H3 (H3K23). *T. brucei* H3K23 corresponds to H3K27 in other eukaryotes (Fig S1) which is acetylated by GCN5, the catalytical subunit of the SAGA complex. Since it is unknown which methyltransferase is responsible for methylation of H3K4 or if more than one methyltransferase is capable of methylating H3K4 like in human, it is assumable that an unidentified COMPASS like complex could be existing in *T. brucei*. Given the importance of SAGA and COMPASS in H2A.Z-mediated transcription regulation I started a search for subunits of potential SAGA and COMPASS-like complexes in *T. brucei*.

Biotinylation based proximity labelling approaches (BioID) with the *E. coli* biotinligase (BL) BirA have proven to be useful to identify new flagellar subdomains in *T. brucei* (Morriswood *et al.*, 2013; Vélez-Ramírez *et al.*, 2021). Recent studies from our lab could show that this proximity labelling approach is also suitable to identify new complex in the highly dynamic environment of the nucleus (Eisenhuth *et al.*, 2020). In order to identify potential subunits of a SAGA-like or a COMPASS-like complex and to identify other proteins that might contribute to transcription regulation in *T. brucei*, I intended to perform a BioID with endogenous tagged *TbSWRC2*. Constant expression of the fusion protein will lead to an increased chance to identify transient interactions in close proximity to the *TbSWR1* complex. Unlike to the 'Disruptor of telomeric silencing 1' (DOT1)B-BioID in which a codon optimized *E.coli* BirA was used (Morriswood *et al.*, 2013), I planned to use a Turbo-biotinligase (TBL), an optimized version of the BirA-BL (Branon *et al.*, 2018). While BirA requires an additional supplement of Biotin, the TBL does not. The 13 µg/l that are available in the HMI-9 medium in are enough to enable efficient proximity biotinylation (Susanne Kramer personal communication). Like in the 2T1 HA:*TbSWRC2* cell line I intended to tag one *TbSWRC2* alleles with a tagging construct amplified from one of the pPOT vectors and replace the remaining wild type allele with a resistance marker. A codon optimized TBL was kindly provided by Susanne Kramer and the Bachelor student Konstantin Köthe cloned the TBL sequence into the pPOTv4 vector (Dean *et al.*, 2015; Bachelor thesis Konstantin Köthe). A diagnostic PCR was performed after each transfection to confirm correct integration of the tagging and the knock-out construct (Fig S16C). The final purification of biotinylated proteins and their associated interaction partners from the soluble fraction of 2T1 TBL:*TbSWRC2* cell lysates was performed in quadruplicates. Cell lysates of non-transfected 2T1 wild type cells were used as a control (Fig S16A). WB samples of the soluble fraction of 2T1 TBL:*TbSWRC2* cell lysates starting material (S) and the eluted material (E) were prepared in course of the protein purification. WB analysis with IRDye 800CW Streptavidin from Li-Cor, revealed a large number of biotinylated proteins in the starting material and in the eluted material (Fig S16B). MS analysis of the eluted material identified 330 significantly enriched and 3 significantly depleted proteins in total (for a detailed list of purified proteins see MS data set 4).

### 3. Results

Based on the curated 'gene ontology' (GO) components annotation found in the TriTryp database, 118 of the 330 proteins of the were classified as nuclear (Fig 27). Aside these 118 proteins various contaminants from other cellular compartments were identified as well. 61 proteins with a cytoplasm localisation were identified, 57 ribosomal proteins and 33 flagellar proteins among others (Fig 27). PANTHER analysis could only link 124 proteins with a molecular function and only three proteins could be identified as chromatin binding proteins (Fig S17 & 18). Except for *TbSWRC2* all 13 subunits of the *TbSWR1* complex could be purified which indicates that the BioID was successful. In addition, three subunits of the HAT2 complex and all subunits of the HAT1 complex could be identified. Here it should be mentioned that neither HAT3 nor Tb927.6.1240 or Tb927.10.9930 could be identified in the BioID. These three proteins were only identified in the initial *TbEaf6* co-IP but not in the reciprocal *TbEaf3* co-IP.



**Figure 27: Classification of the proteins identified in course of the *TbSWRC1* BioID**

For most of the proteins more than one curated GO component annotation can be found in the TriTryp database. This chart displays only the first and therefore most probable hit.

Among the 118 proteins were seven proteins that might be of interest in terms of transcription regulation in *T. brucei* (Table 3.5). Beside Tb927.10.15950 a TATA-box binding protein, Bdf2, three transcription factors (TFs), two methyltransferases and AAA ATPase were identified. According to the InterPro analysis the AAA ATPase Tb927.11.6350 appears to be homologous to 'yeast Tat-binding homolog 7' (YTA7), a protein that is involved in regulation of nucleosome density and gene regulation via its interaction with the FACT complex subunit Spt16 (Gradolatto *et al.*, 2008; Lombardi *et al.*, 2011). YTA7 just like Tb927.11.6350 harbours a bromodomain but an association with a specific histone acetylation could not be confirmed (Gradolatto *et al.*, 2009). The identification of Tb927.11.1410 the 'subunit 3 of the class I transcription factor A' (CITFA3) was surprising since CITFA3 was so far only associated with regulation of the surface proteins of the parasite (Nguyen *et al.*, 2014). The second and third TFs identified were TFIIA (Tb927.10.4840) and Tb927.10.11720 a CW-type Zinc Finger protein.



### 3. Results

While the latter one was identified as being homologous to 'Infected cell protein 4' (ICP4), a TF that is known to bind the TATA-binding protein and TFIIB (Smith *et al.*, 1993), TFIIA is part of the PIC (reviewed in Allen & Taatjes, 2015). In *T. brucei* TFIIA could only be linked with the regulation of SL transcription but not RNAP II transcription in general (Schimanski *et al.*, 2005a; Schimanski *et al.*, 2006). The identification of Bdf2 in close proximity to SWR1 came to my surprise, especially due to the fact that Bdf2 could be linked with CITFA in course of a mass spectrometry-coupled co-IP experiment (Staneva *et al.*, 2021). NCBI and InterPro BLAST analysis identified Tb927.10.8100 and Tb927.9.13470 as SET-Methyltransferases. Data published in Staneva et al., 2021 could link Tb927.10.8100 with the CW-type Zinc Finger protein Tb927.10.11720.

### 3. Results

**Table 16: List of proteins with potential involvement in transcription regulation identified in the *TbSWRC2* BioID**

Accession Number	Product Description	Curated GO Components	Curated GO Functions (Phyre2, Interpro & NCBI BLASTp results)	Log <sub>2</sub> fold-dif.	Log <sub>10</sub> pvalue
Tb927.10.7420	bromodomain factor 2 protein, putative	nucleus	histone binding	7,1701	6,98E+14
Tb927.10.15950	TATA-box-binding protein	cytoplasm; nucleoplasm; nucleus	RNA polymerase II general transcription initiation factor activity	6,4947	7,44E+14
Tb927.10.4840	transcription factor II a, putative	cytoplasm; nuclear lumen	N/A	5,2202	7,70E+14
Tb927.11.6350	AAA ATPase, putative	nucleoplasm; nucleus	ATPase activity; chromatin binding; histone binding; (Interpro: bromodomain containing, YTA7 like; involved in heterochromatin formation and gene regulation <sup>1,2</sup> )	4,2407	4,65E+14
Tb927.10.11720	CW-type Zinc Finger, putative	nucleoplasm; nucleus	N/A (NCBI: ICP4 related: A TF that binds TBP and TFIIIB) potential interaction with Tb927.10.8100 <sup>3</sup> )	3,9218	3,57E+14
Tb927.10.8100	hypothetical protein, conserved	nucleoplasm	N/A (NCBI BLAST/InterPro: SET-Methyltransferase domain containing protein; Phyre2: COMPASS related SET Methyltransferase; also known as SET26 <sup>3</sup> )	3,7316	5,85E+14
Tb927.11.1410	class I transcription factor A, subunit 3	nucleoplasm; nucleus	VSG regulation <sup>4</sup> PMID: 17972917	1,7776	2,93E+14
Tb927.9.13470	hypothetical protein, conserved	nucleoplasm; nucleus	N/A (NCBI BLAST/InterPro: SET-Methyltransferase domain containing protein; also known as SET27 <sup>3</sup> )	0,0340	5,80E+13

<sup>1</sup> Lombardi *et al.*, 2011 <sup>2</sup> Gradolatto *et al.*, 2008 <sup>3</sup> Staneva *et al.*, 2021

This table depicts selected candidate proteins that are potentially involved in H2A.Z-mediated transcription regulation. Based on the information in the Trytrip database the "curated GO Components" column states the cellular compartment in which the protein can be found. The "curated GO Functions" column states the biological function of the protein.

## 4. Discussion, Conclusion and Outlook

H2A.Z and the SWR1-like complexes that incorporate the histone variant in eukaryotes are key players in transcription regulation (Raisner *et al.*, 2005; Altaf *et al.*, 2010; Brunelle *et al.*, 2015; Giaimo *et al.*, 2018). Even though H2A.Z is present in *T. brucei* the molecular machinery that regulates H2A.Z dynamics have not been identified yet. Neither a SWR1 nor a NuA4/TIP60-like complex could be identified so far. The goal of my PhD project was to identify homologs of these complexes in the unicellular parasite *T. brucei* by using mass spectrometry-coupled co-IPs experiments. Moreover, I wanted to investigate the role of HAT2 in H2A.Z incorporation and whether the histone acetyltransferase is part of a NuA4/TIP60-like complex. By using a proximity labeling approach, I intended to identify components of other protein complexes that might play a role in transcription regulation, just like the SAGA or the COMPASS complex. With a set of co-IP experiments it was possible to identify a protein complex with characteristics of a SWR1-like complex that appears to consist of 13 subunits. The complex is assembled around the SNF2-ATPase Tb927.11.10730 and subsequent ChIP-seq and cell fractionation experiments with Tb927.11.10730-depleted cells could confirm that the SNF2 ATPase is required for correct H2A.Z incorporation in *T. brucei*. Generated data also demonstrate that the loss of *TbSWR1* results in a loss of RNA Polymerase I- and II-dependent transcripts and a condensation of chromatin. Initial co-IP experiments with HAT2 and an Eaf6-like factor could identify two protein complexes assembled around HAT1 and HAT2 and reciprocal co-IP approaches confirmed their existence. While the HAT1 complex appears to be homologous to the *S. cerevisiae* NuA4 complex, the HAT2 complex bears a unique and so far, undescribed complex structure that includes the two bromodomain factors Bdf3 and Bdf5 as well as multiple other proteins. Proximity labelling approaches could identify various transcription factors and SET-methyltransferases that might play a role in RNAP II transcription regulation. Even though the data I generated in course of this project could answer a couple of questions about H2A.Z-mediated transcription regulation, new questions regarding RNAP II transcription regulation and DNA DSB response came up.

### 4.1 Analysis of the composition of the *TbSWR1*, HAT1 and HAT2 complex

In the following chapters I will discuss the composition of the three complex and will also evaluate the quality of the experimental approach and the received data. By focusing on some of the unique features of each of the complexes I will try to assess how these three complexes could interact with each other in the course of different cellular processes to create a *Kinetoplastid*-specific model of transcription regulation via HAT2, SWR1 and HAT1.

### 4.1.1 The *TbSWR1* complex

The initial approach to tag *TbSWR1* on either the N-terminus or on the C-terminus failed. Even though N-terminal HA-tagging of *TbSWR1* led to the translation of a tagged version of the protein, it was not possible to replace the wild type allele. This indicated that the tagged version of the protein is not functional. The N-terminus of SWR1 in *S. cerevisiae* is binding directly to H2A.Z and provides a platform for the stable assembly of other complex subunits like Arp4 and Yaf-9 (Wu *et al.*, 2009). Since the C-module of the complex is assembled around the C-terminus of SWR1 (Nguyen *et al.*, 2013 reviewed in Gerhold & Gasser, 2014), it is likely that tagging of any of the termini blocks H2A.Z interaction and/or proper complex assembly.

The co-IP experiment with the DNA-helicase RuvB2 that was performed in the following already gave a good overview about the *TbSWR1* complex since 12 of the 13 subunits could be identified (Table 12, Fig S3A, MS data set 1). The absence of SHQ1, Cbf5p and the nucleolar RNA binding protein in the reciprocal co-IPs with *TbSWRC1*, *TbSWRC2* and *TbSWRC4* confirmed my initial assumption that these three proteins are RuvB2-specific interaction partners and no components of the *TbSWR1* complex. Even though *TbSWRC2* and *TbSWRC3* were only identified in three of the four co-IPs (Table 3.2, MS data set 1), it can be assumed that they are stable subunits of the *TbSWR1* complex and that their absence in one of the co-IPs only reflects a technical issue. This assumption is based on the fact that *TbSWRC1* was co-purified in the *TbSWRC2* co-IP. If the co-IP with *TbSWRC1* indicated the presence of a *TbSWRC2*-free complex then the *TbSWRC2* co-IP should not have co-purified *TbSWRC1*. Even though it cannot be excluded that *TbSWRC3* forms a separate complex without RuvB2 it is very unlikely considering the conserved role of RuvB2 within the SWR1 complex (reviewed in Willhoft & Wigley, 2020). Another hint that *TbSWRC2* is in fact a part of the *TbSWR1* complex and that its absence in the *TbSWRC1* co-IP is based on a technical issue, is the observation that *TbSWRC2*, just like *TbSWRC1* contributes to the H2A.Z incorporation (Fig S6)

In course of *in-silico* analysis using NCBI and InterPro BLAST almost all 13 subunits of the *TbSWR1* complex could be related to domains/proteins that have been described as components/subunits of the *S. cerevisiae* SWR1 complex. The only exceptions were the proteins *TbSWRC4* (Tb927.7.4040), *TbSWRC3* (Tb927.10.8510) and *TbARP2* (Tb927.6.2570). *TbSWRC4* which was linked to a SANT/DAMP-1 domain in course of Phyre2 homology modelling and appears to be exclusively found in *Blechnomonas ayalai* and the subspecies of *T. brucei* and *T. cruzi*. Given the conserved role of SANT/DAMP-1 domains in SWR1-like complexes it is possible that a protein with a more conserved SANT/DAMP-1 domain fulfills the role of *TbSWRC4* in other *Kinetoplastids*.

Despite a model confidence of almost 100% the Phyre2 homology modelling of *TbSWRC3* raised some questions. While the conserved CS of SET-Methyltransferases can be found almost unaltered in *TbSWRC3* all three potential SAM-binding sites of *TbSWRC3* possess an insertion of one or more amino acids (Fig 13). These insertions and insertions between the SAM-binding sites might alter the structure of the SAM-binding pocket.

#### 4. Discussion, Conclusion and Outlook

Without further structural data and/or data derived from a methyltransferase assay it is impossible to assess whether *TbSWRC3* is a functional SET-methyltransferase or not. Nevertheless, data generated in the course of this project suggested that *TbSWRC3* is essential. The missing response to RNAi induction against *TbSWRC3* via tetracycline might be explained by the leakage of a lethal RNAi. If the system is slightly activated even without adding tetracycline, tetracycline responsive clones would have been lost during the cloning process, therefore selecting dor clones with a faulty RNAi system. Even though the 2T1 cell line was optimized to reduce leaky RNAi, it can still occur (Rico *et al.*, 2017; Briggs *et al.*, 2019). The loss of response to tetracycline that was observed for the first RNAi cell line generated might (Fig S5) support the assumption of a leaking RNAi system.

*TbARP2* (Tb927.6.2570) is annotated as a potential SUMO-interacting motif-containing protein. Therefore, it was very surprising that Phyre2 modeled almost 80% of the predicted secondary structure of the protein onto the secondary structure of Arp8 and Arp9. The SWR1 complex normally consists of the two Arp proteins Arp4 and Arp6 (reviewed 31838293). Arps in general play an important role in the nucleosome association of chromatin remodeling complexes (Shen *et al.*, 2003). Arp4 for example prefers unmodified (Harata *et al.*, 1999) and phosphorylated H2A (Downs *et al.*, 2004) while Arp8 prefers the H3/H4 tetramer (Gerhold *et al.*, 2012; Saravanan *et al.*, 2012). The third ARP might therefore be required for correct positioning of *TbSWRC3* on histone H3. Nevertheless, with the available data it is not possible to assess whether all three ARPs are part of the *TbSWR1* complex or not. *In vitro* reconstitution assays with recombinant proteins are required to answer this question. Regarding the relevance of the Phyre2 data all results must be seen as valid, since the confidence in the modeled secondary structures of all three proteins were above 95 %.

Another very important subunit of SWR1 complexes, a bromodomain-containing protein, could not be identified in any of the four co-IP experiments. The Bdf within the SWR1 complex is required for recruitment of the complex since it interacts with acetylated histone H4 (Matangkasombut & Buratowski, 2003). Besides the possibility that *TbSWR1* does not possess a Bdf due to the complex composition of the HAT2 complex, another reason for the absence of such a factor might be the already mentioned ionization efficiency (Hu *et al.*, 2016; Liigand *et al.*, 2019). Relating to the function of *TbSWR1* I could not confirm that the complex incorporates H2A.Z as a dimer together with H2B.V, it is known that the two histone variants are located in nucleosomes of the TSS (Siegel *et al.*, 2009) and we identified H2B.V in the *TbSWRC2* proximity labeling approach. Nevertheless, I could not show that incorporation of H2B.V is dependent on *TbSWR1*. Cell fractionation experiments with subsequent WB analysis and ChIP-seq analysis are required to confirm this.

### 4.1.2 The HAT2 complex

Our mass spectrometry data indicated that the HAT2 complex consists of 9 subunits. The hypothetical protein Tb927.3.4140 was only identified in the initial HAT2 co-IP (Fig S13; MS data set 2) but data published by Staneva *et al.*, 2021 indicate that it is in fact part of the HAT2 complex. The protein was identified in co-IPs with YFP-tagged Bdf3, Bdf5 and HAT2. Therefore, it is likely that the complex consists of 10 instead of 9 subunits. Nevertheless, further experiments are required to confirm these data (examples of subsequent experiments are listed in 4.1.4).

In course of this project the research group of Nicolai Siegel, identified HAT2 as the acetyltransferase being responsible for acetylations on histone H4 but not on H2A (Kraus *et al.*, 2020). Based on this function HAT2 is homologous to Esa1. Nevertheless, none of the HAT2 complex subunits (Fig S13; MS data set 2) is homologous to known subunits of the NuA4-like complex (Table 14). The identification of Bdf3, Bdf5 and Tb927.4.2340, a potential third Bdf, as subunits of the HAT2 complex were unexpected. Immunofluorescence microscopy and ChIP-seq analysis already indicated a potential link between HAT2, Bdf3, H4K10 acetylation and H2A.Z (Siegel *et al.*, 2009). Therefore, it was not surprising to find Bdf3-associated with HAT2 and to be involved in H2A.Z incorporation (Fig 24). Previous experiments could not link Bdf5 with the TSSs (Siegel *et al.*, 2009), but cell fractionation experiments revealed its contribution to H2A.Z incorporation as part of the HAT2 complex (personal communication Laura Hartleb). In contrast to this, cell fractionation experiments with Bdf6-depleted cells did not show any alterations in chromatin-associated H2A.Z levels (personal communication Laura Hartleb). Without having further information about the complex structure and the role of individual complex components it is difficult to speculate about a reason for this.

One explanation for Bdf6 being not involved in H2A.Z incorporation is the possibility that the protein and therefore the HAT2 complex is involved in a different/additional cellular process. The potential presence of an FHA-domain containing protein supports this assumption and might hint towards an involvement of the HAT2 complex in DNA damage repair (I will explain this in more detail in chapter 4.3). An analysis of the Phyre2 alignment of the predicted FHA-domain revealed that, despite some differences within the composition of  $\alpha$ -helices and  $\beta$ -sheets, all conserved amino acids that are required for the binding of a phosphorylated peptide, are present in Tb927.9.13320 (Fig S15). Nevertheless, the presence of multiple bromodomain factors within the HAT2 complex might be the most important aspect since it could hint towards the binding dynamics of the HAT2 complex. One of the major differences between the *S. cerevisiae* NuA4 and the human TIP60/p400 complex is the dual function of the TIP60/p400 complex. While NuA4 is only responsible for the acetylation step of H4/H2A and H2A.Z (Keogh *et al.*, 2006; Altaf *et al.*, 2010), the TIP60/p400 complex is also responsible for subsequent incorporation of H2A.Z (Kusch *et al.*, 2004; Gévry *et al.*, 2007; Giaimo *et al.*, 2018; Lee *et al.*, 2020).

The TIP60/p400 must be kept in position to perform the histone exchange step, and the bromodomain factor Brd8 is required to keep the complex in position. Loss of Brd8 leads to a reduction of chromatin-associated p400 and H2A.Z (Couture *et al.*, 2012).

In light of these data it can be presumed that HAT2 has to be kept in position for subsequent incorporation of H2A.Z (I will explain this in more detail in chapter 4.4). RNAi experiments with the hypothetical proteins that were identified as subunits of the HAT2 complex might help to learn more about regulation of HAT2 and H2A.Z incorporation and additional functions of this complex.

### 4.1.3 The HAT1 complex

Co-IP experiments with *TbEaf6* (Tb927.9.2910) the homolog of *S. cerevisiae* Eaf6 and the reciprocal co-IP I performed with Tb927.1.650 identified two different histone acetyltransferase complexes assembled around HAT1 and HAT3. Based on the Phyre2 homology modelling especially of Tb927.10.14190 and Tb927.1.650 it is likely to assume that the HAT1 complex is homologous to the *S. cerevisiae* NuA4 complex. Nevertheless, there are some important differences the HAT1 and the NuA4 complex that have to be mentioned here. While Yaf-9 is a shared subunit of the *S. cerevisiae* NuA4 and SWR1 complex (Willhoft & Wigley, 2020), the *T. brucei* SWR1 and HAT1 complex possess two different YEATS-domain containing proteins (Table 13 & 15, Fig S3 & 11, MS data sets 1 & 3) which might indicate that they both bind to different substrates. This circumstance might reflect an adaptation to prevent HAT1 binding in course of DNA DSBs (I will explain this in more detail in chapter 4.4). The second difference is the localisation of a histone methylation-binding domain within the complex. While a chromo-domain together with an MRG domain can be found in Eaf3, *TbEaf3* appears to lack such a domain. In contrast to this HAT1 does possess a tudor-knot chromo-like domain. This domain might play an important role in nucleosome binding of the HAT1 complex subsequent to H2A.Z incorporation by *TbSWR1* (I will explain this in more detail in chapter 4.4). The presence of a bromodomain factor within the HAT1 complex raises some questions.

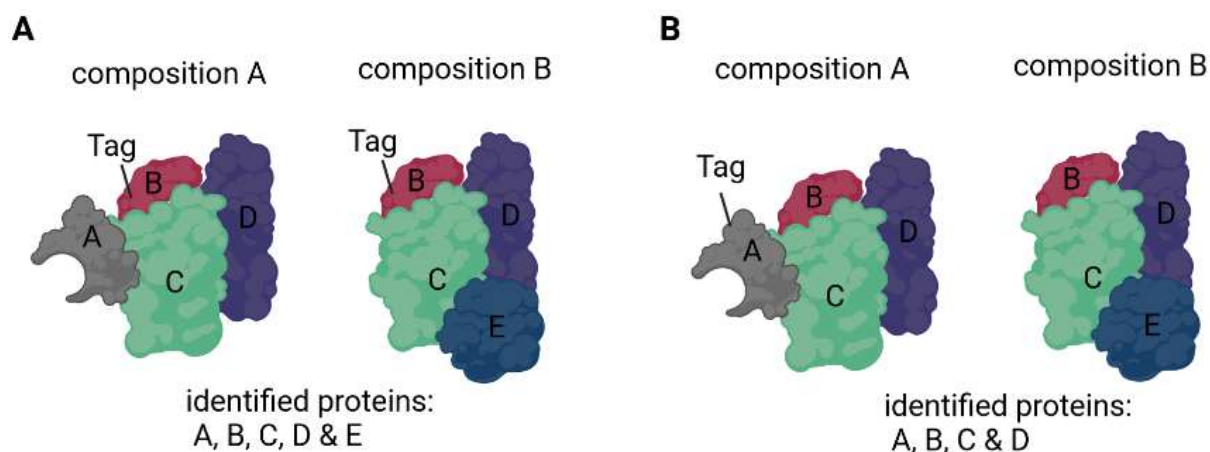
As for the HAT2 complex the bromodomain factor in the HAT1 complex might be required to keep the complex in position for recruitment of proteins that are required for subsequent transcription. To learn more about the role of the bromodomain factors, especially those in the HAT2 complex, it might be useful to perform an RNAi assay that is coupled with a gene replacement approach. Replacing one wild type allele with a mutated and recoded version of the *Bdf* coding gene might allow a systematic depletion of the wild type protein via RNAi, leaving the expression of the recoded and mutated protein unaltered. With systematic mutation of conserved amino acids that are required for binding of the acetylated amino acids, it is possible to assess whether the bromodomain is required for H2A.Z incorporation (in case of the HAT2 Bdfs). Standard RNAi approaches might be inconclusive since a negative phenotype can either result from an affected complex assembly (if the depleted protein has structural relevance) or from a loss of function due to a missing domain.

As for the hypothetical proteins of the HAT2 complex standard RNAi experiments might be useful to get initial information about the function and the role of the hypothetical proteins of the HAT1 complex.

### **4.1.4 Evaluation of the experimental approach used to identify the *Tb*SWR1, HAT2 and HAT1 complex**

The label free liquid chromatography (LC) mass spectrometry approach that we used to identify the *Tb*SWR1, the HAT2 and the HAT1 represents a pretty reliable method to identify protein complexes and protein-protein interactions. Even though it provides a high sensitivity (Easterling *et al.*, 2020) it is still possible that information about potential binding partners is lost during the sample analysis. Differences within the ionization efficiency of peptides can lead to a loss of information in data-dependent and data-independent acquisition (Hu *et al.*, 2016; Liigand *et al.*, 2019). Besides such technical issues in course of the MS analysis the performance of the co-IP itself can cause severe differences in the subsequent MS analysis. While the use of ionic detergents like SDS in the co-IP buffer would lead to the denaturation of proteins and therefore to a disruption of the complex (Otzen, 2002), the use of non-ionic and mild detergents might lead to the identification of false positives. Especially in course of the precipitation of chromatin-associated proteins it is possible to precipitate proteins that do not directly interact with the protein of interest but with the DNA that was co-precipitated as part of the co-IP (Nguyen & Goodrich, 2006). Evaluation of the data received in course of the MS analysis can also affect the number of potential interaction partners. In the current case four replicates were analysed and the identified proteins were not transferred between the different MS runs. As a consequence, only the most abundant proteins in each of the replicates were listed. Moreover, co-IP experiments with a few selected candidates might not be sufficient to identify varying complex conformations (Fig 28). A sucrose gradient centrifugation coupled with native-PAGE and subsequent silver staining can be used to answer the question, if there are several SWR1/HAT1/HAT2 complex of different compositions.





**Figure 28: A disadvantage of MS-coupled co-IP experiments**

**(A)** An exemplary Co-IP experiment with tagged protein B and the estimated result, **(B)** exemplary co-IP with tagged protein A the estimated result. The figure has been designed with BioRender©.

In the present case a gradient from 5 to 20 % should be sufficient to gain a more detailed insight into the composition of all three complexes since their molecular weight is above 200 kDa and below 800 kDa (Fernandez-Martinez *et al.*, 2016; based on the MS data the *TbSWR1* should have a weight of around 750 kDa, the HAT1 complex a weight of around 295 kDa and the HAT2 complex a molecular weight between around 460 kDa and 520 kDa). The fact that Staneva *et al.*, published the same complex compositions for the HAT1 and HAT2 complex in an independent approach supports the data presented in this work.

*In vitro* experiments could be the proper choice to investigate the role of putative PTM readers and writers identified in the HAT2 and *TbSWR1* complex. A methyltransferase assay with recombinant *TbSWRC3* will be required to confirm its methyltransferase activity. Such an assay could be performed with recombinant histones, since it was shown that SET-methyltransferases can methylate histones without a nucleosomal context (Tachibana *et al.*, 2001). Even though the ionization process might be negatively affected by PTMs, tandem MS analysis of the methyltransferase assay products could provide precise information about which amino acid is modified (Larsen *et al.*, 2006). Cross-linking mass spectrometry could be used to learn more about the interaction interface of each complex subunit.

### 4.2 *TbSWR1* and the loss of RNA

H2A.Z was described as a transcription regulation feature in many organisms (Brunelle *et al.*, 2015; Giaimo *et al.*, 2018) and a recent publication validated this for *T. brucei* (Kraus *et al.*, 2020). The data derived from luciferase experiments revealed a loss of transcriptional activity subsequent to the loss of *TbSWR1*. Data of the control experiments performed with H2A.Z and HAT2-depleted cells suggest that this phenotype is H2A.Z-dependent. A Northern Blot assay performed by PD Dr. Susanne Kramer confirmed a loss of RNA polymerase I and II transcripts after the downregulation of *TbSWR1* (Fig 20). In combination with the data from the luciferase experiments, loss of *TbSWR1* appears to lead to a downregulation of transcription, which is in line with the role of H2A.Z in other organisms. Since it was shown that acetylation of H2A.Z via HAT1 is required for proper transcription via RNAP II, it is likely that the loss of the HAT1 substrate leads to a loss of acetylation and therefore to a reduction of transcription. Despite a downregulation of total mRNA and rRNA the splice leader mRNA appears to be unaffected by the downregulation of *TbSWR1* (Fig 20). If the loss of luciferase activity subsequent to *TbSWR1* downregulation would be a consequence of translational defects, an overall reduction of RNA levels would be detected and not a selective loss.

Regulation of the splice leader transcription differs from the regulation of other PTUs and involves the TFIIA, TFIIH and upstream regulatory elements (Günzl *et al.*, 1997; Lee *et al.*, 2007) among others, which would explain why splice leader mRNA levels are not affected. Loss of ribosomal RNA came to our surprise, but since an equal loss of rRNA was observed subsequent to RNAP II depletion, it is likely that the loss of rRNA represents a side effect (Fig 20). RNAP I subunits, as part of a PTU, are transcribed by RNAP II. A lack of RNAP II activity might therefore result in the reduced amount of RNAP I transcripts that finally leads to reduced rRNA.

Based on the cell cycle analysis data it appears that the cells keep dividing subsequent to the depletion of *TbSWR1* (Fig 16). This requires a doubling of the rRNA, which aggravates the loss of rRNA within the cell. It can also be excluded that the reduction of RNA is caused by the death of the cells. While the number of dead cells remains almost unaltered after 24 and 36 h *TbSWR1* depletion, the luciferase activity is reduced from 80 % of the original activity down to 60 % after 36 h (Fig 19). General loss of mRNA that occurs in course of the loss of chromatin-associated H2A.Z could also explain the recovery in growth that was observed after *TbSWR1*, H2A.Z and RPB1 depletion (Fig S11) and the recovery of the H2A.Z signal observed in *TbSWRC1*- and *TbSWRC2*-depleted cells (Fig S6). A reduction of mRNAs coding for the proteins of the RNAi system would affect its efficiency, therefore suspending the RNAi effect after a while. This effect has already been described in course of investigations regarding the RNAi machinery (Shi *et al.*, 2007). Different dynamics between the cell lines can be explained by varying targets and efficiencies of the RNAi system. Due to the lack of data concerning protein and mRNA stability it is not possible to assess whether a fast or a slow RNAi response is favoring a quick recovery of the cells.

The chromatin condensation observed in *TbSWR1*-, H2A.Z- and RPB1-depleted cells appears to be based on the loss of RNAP II activity. Treatment of cells with Alpha-Amanitin, a characterized RNAP II and RNAP III inhibitor (reviewed in Bensaude, 2011), leads to a condensation of chromatin (Marinozzi & Fiume, 1971; Roeder, 1976; Episkopou *et al.*, 2009). There is evidence that Alpha-Amanitin treatment of *T. brucei* cells also leads to chromatin condensation. In course of a project that focused on transcription regulation effects of different UTRs Alpha-Amanitin was used as a control to systematically shut down RNAP II and III transcription (McAndrew *et al.*, 1998). Surprisingly, it also affected genes, located in RNAP II-transcribed regions, that were under the regulation of a T3 Promoter (McAndrew *et al.*, 1998). These results hint towards the existence of a RNAP II transcription-dependent chromatin condensation mechanism which is in line with the EM images of the *TbSWR1*, H2A.Z and RPB1-depleted cells (Fig 21).

### **4.3 The role of *TbSWR1* in genome integrity**

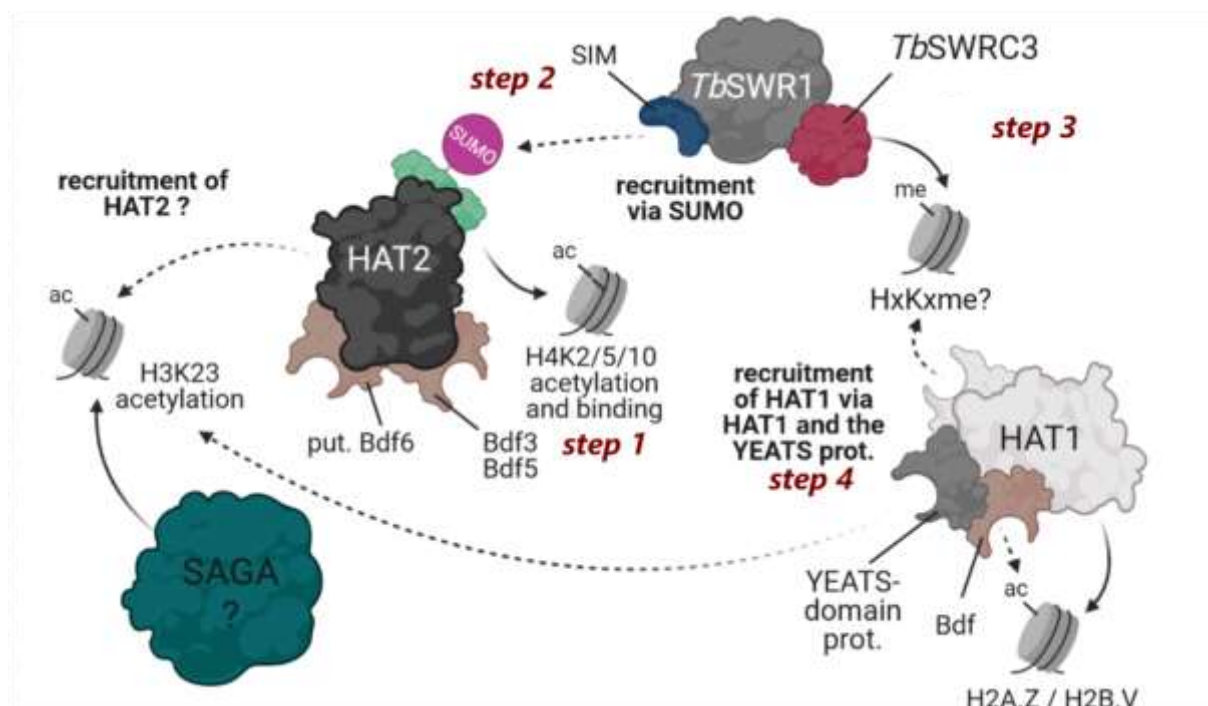
H2A.Z is a very important component of the centromere and is responsible to maintain the three-dimensional structure of the centromere (Greaves *et al.*, 2007). Experiments performed in *S. pombe* could show that H2A.Z is required for regulation of cohesion dynamics and loss of H2A.Z leads to chromosome entanglement and breakage in anaphase (Kim *et al.*, 2009; Tapia-Alveal *et al.*, 2014). In light of these data it can be assumed that H2A.Z plays a similar role in *T. brucei* since cell cycle analysis revealed a large number of cells with a reduced DNA content subsequent to *TbSWR1* depletion (Fig 16).

#### **4.4 Bringing together the differences: the interplay of HAT2, *TbSWR1* and HAT1 in course of different cellular processes.**

Incorporation of H2A.Z either in course of transcription regulation or in course of the DNA DSB response requires a well-orchestrated interplay between NuA4-like histone acetyltransferases and SWR1-like complexes (chapter 1.5 and 1.5.1; Altaf *et al.*, 2010). The presence of two distinct acetyltransferase complexes responsible for H4 and H2A.Z acetylation, requires a well-regulated interaction between the HAT2 and *TbSWR1* complex as well as between the *TbSWR1* and HAT1 complex. This interaction network has to ensure efficient incorporation of H2A.Z in course of transcription regulation and DNA DSB response. In addition to that, the system must have the possibility to differentiate between both cellular processes to prevent transcription initiation at DNA DSBs. In the following chapter I will introduce a putative H2A.Z incorporation model for each process. Each model is assembled around a *TbSWR1* complex with a unique protein composition. The idea of *TbSWR1* complexes with altering compositions is based on the deliberation that SWR1-like complexes normally consist of two instead of three ARPs.

##### **4.4.1 The potential interplay of HAT2, *TbSWR1* and HAT1 in course of H2A.Z-mediated transcription regulation**

One of the major differences between the HAT2 and NuA4 complex is the presence of multiple Bdfs in the HAT2 complex and a potential sumoylation on one of the HAT2 subunits. As I already stated in chapter 4.1.2 the presence of Bdfs indicate that the HAT2 complex is kept in position subsequent to acetylation of histone H4 (Fig 29 step 1). Given that I could co-purify SUMO in the HAT2 and Bdf3 co-IP (Table 14, Fig S12, MS data set 2) and given that the *TbSWR1* complex might possess a protein with a potential SUMO-interacting motif (Table 13, Fig S3 & S4, MS data set 1), the sumoylation of one of the HAT2 subunits might be required for the recruitment of the *TbSWR1* complex to the TSSs (Fig 29 step 2). This would explain the absence of a Bdf within the *TbSWR1* complex. It was already shown that sumoylation plays an important role in the assembly of CRCs of the INO80 subfamily (Cox *et al.*, 2017). CLMS could be a suitable approach to investigate this putative interaction between the two complexes. One of the acetylations identified on the core of H2A (Kraus *et al.*, 2020) might be bound by *TbSWRC1* to support binding of its complex. YEATS-domains were shown to bind to acetylated histones 25417107. In the following *TbSWRC3* methylates histone H3 subsequent to the incorporation of H2A.Z (Fig 29 step 3). Binding of this methylation by the tudor-knot chromo-like domain of HAT1 is the first step required for the recruitment of the HAT1 complex. The second step involves binding of the acetylation on H3K23 by the YEATS-domain containing protein Tb927.7.5310 to acetylate H2A.Z and H2B.V (Fig 29 step 4). Acetylated H3K23 marks the nucleosomes of the TSSs. Without this PTM cannot bind therefore preventing binding of the HAT1 complex at DNA DSBs (chapter 4.4.2).



**Figure 29: Putative interplay of the HAT2, *TbSWR1* and HAT1 complex in course of H2A.Z-mediated transcription regulation**

Dashed lines indicate recruitment or a binding process while sustained lines indicate the generation of a PTM by a complex; the figure has been designed with BioRender©

Subsequent to the acetylation of H2A.Z and H2B.V, the HAT1 remains at the nucleosome by binding an acetylation on H2A.Z or H2B.V for further recruitment of transcription initiation factors. Recent data could identify various histone methylations associated with nucleosomes in the TSSs, therefore it is likely that *TbSWR1* (Table 16; MS data set 4) complex are responsible for one of these methylations. One of these methylations might be bound by the Given that SET-methyltransferases are also required to methylate transcription factors (Kouskouti *et al.*, 2004) it might be worth to perform CLMS analysis with tagged versions of some of these candidates. Here it should be mentioned that one of the key switches in chromatin regulation in eukaryotes, the acetylation/methylation switch on H3K9, appears to be absent in *T. brucei* (Kraus *et al.*, 2020). This hints towards the presence of a much simpler transcription regulation pathway in *T. brucei* than in other eukaryotic organisms. Whether H4K2 which was reported to be acetylated and methylated, corresponds to H3K9 in *T. brucei* is questionable since both modifications of H4K2 were identified on nucleosomes of the TSS (Kraus *et al.*, 2020).

An important question that occurred as a part of this work concerns the regulation of the HAT2 complex. The NuA4 complex is dependent on methylation of H3K4 by COMPASS (Ginsburg *et al.*, 2009; Ginsburg *et al.*, 2014) which depends on the ubiquitination of H2B (Hsu *et al.*, 2019). Due to the binding of Yaf-9 to H3K27 acetylation (Li *et al.*, 2014) NuA4 is also dependent on SAGA, since SAGA acetylates H3K27. The HAT2 complex appears to possess multiple histone acetylation-binding proteins but no protein for binding methylated histones. Therefore, the regulation of HAT2 must be different to the regulation of NuA4. Whether acetylation of H3K23 (which corresponds to H3K27, Fig S1) plays a role in HAT2 recruitment must be assessed via ChIP-seq experiments with mutated versions of H3.

#### **4.4.2 The potential interplay of the HAT2 and *TbSWR1* complex in course of H2A.Z-mediated DNA DSB repair**

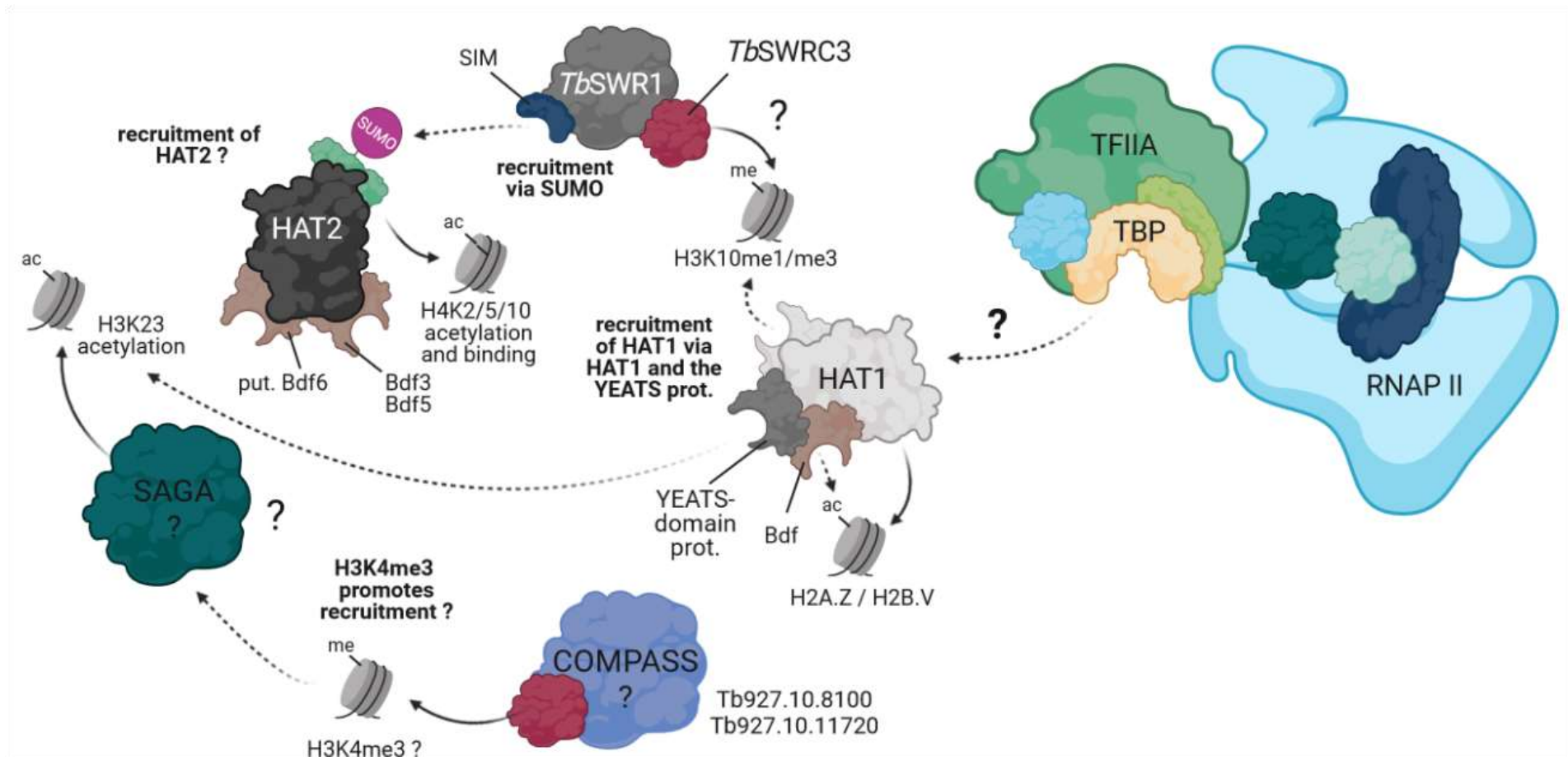
Data derived from investigations regarding the DNA damage response indicated an increase of chromatin-associated H2A.Z in response to occurring DNA DSBs. This implies that HAT2 is also involved in the DNA damage response but likely via a different pathway. The presence of a potential FHA-domain containing protein within the complex (Table 14, Fig S13, MS data set 2) might be an explanation for this. The FHA-domain would allow a direct binding of  $\gamma$ H2A. In the following *TbSWR1* is recruited to exchange  $\gamma$ H2A with H2A.Z. The absence of the TSS-associated acetylation on H3K23 now prevents binding of the HAT1 complex and therefore unintentional initiation of transcription. Removal of H2A.Z in course of the DNA DSB response could be performed by the Anp32e-like protein Tb927.5.1270. Anp32e has been reported to remove H2A.Z from chromatin in course of the DNA damage response (Obri *et al.*, 2014; Gursoy-Yuzugullu *et al.*, 2015).

#### **4.4.3 The presence of transcription factors and enzymes responsible for post-translational modifications in close proximity to *TbSWR1***

Even though no potential SAGA complex subunit could be identified in the *TbSWRC1* BioID the presence of an TSS-associated acetylation on H3K23 clearly hints towards the presence of a SAGA complex or a Gcn5-related acetyltransferase. The presence of such a complex would provide an answer about how HAT1 is recruited exclusively to TSSs. The presence of at least two SET-methyltransferases (MS data set 4) could be a hint towards the presence of a COMPASS-like complex. Especially the identification of Tb927.10.8100 as a COMPASS related SET-methyltransferase by Phyre2 homology modelling analysis (Fig S19) and the potential association with the CW-type Zinc-finger-motif containing protein Tb927.10.11720 (Staneva *et al.*, 2021) supports this assumption. The CW-type Zinc Finger protein Cfdp1 is a subunit of the COMPASS complex and important for recruitment of the complex (Beurton *et al.*, 2019). Whether Tb927.10.8100 is part of a COMPASS like complex has to be investigated. The presence of multiple TFs, especially TFIIA and subunit 3 of CITFA as well as the TBP TRF4 raised several questions about RNAP II transcription regulation in *T. brucei*.

#### 4. Discussion, Conclusion and Outlook

So far CITFA could only be linked with the transcriptional activity of RNAP I (Nguyen *et al.*, 2012). TFIIA and TRF4 were only shown to play a role in SL transcription (Ruan *et al.*, 2004; Schimanski *et al.*, 2005a; Lee *et al.*, 2007). Given that the SL sequence is only transcribed from a specific locus on chromosome 9 (Berriman *et al.*, 2005; Muller *et al.*, 2018) it is highly unlikely that the presence of TFIIA and TRF4 in close proximity to the *Tb*SWR1 complex is a coincidence. The bromodomain factor in the HAT1 complex might be required to keep the complex in position so that it can recruit subunits of a potential PIC, like TRF4 and TFIIA to the TSSs. Given that TFIIA and the TATA-binding protein TRF4 were highly enriched in the *Tb*SWRC2 BioID (MS data set 4) it is highly unlikely that they are not involved in regulation of RNAP II transcription in general only regulating SL transcription. YTA7 is linked with regulation of the nucleosome density (Gradolatto *et al.*, 2009; Lombardi *et al.*, 2011). Therefore, the identification of the YTA7 homolog Tb927.11.6350 (also known as Bdf7; Staneva *et al.*, 2021) and an additional SET-Methyltransferase in close proximity to the *Tb*SWR1 complex, raised the question if these proteins are also involved in transcription regulation and regulation of the chromatin structure. The absence of *Tb*SWRC2 in the mass spectrometry data can be explained by the nature of the experiment. The Biotin Streptavidin interaction is very resistant against pH, temperature and denaturants (Stayton *et al.*, 1999). Since the BioID samples were boiled to elute the biotinylated proteins from the beads, it is possible that some of the proteins were not properly eluted. A trypsin digestion performed with the matrix-associated proteins, would improve this. Even though some interesting candidates might have been missed, the *Tb*SWRC2 BioID revealed some pretty interesting candidates for further investigations underlying the molecular mechanisms of RNAP II transcription regulation. Due to the identification of potential COMPASS complex components and several transcription factors like TFIIA or a TATA-binding protein, I adapted my putative model of H2A.Z-dependent transcription regulation (Fig 29).



**Figure 30: Extended model of a putative H2A.Z-mediated transcription regulation mechanism in *T. brucei***

Dashed lines indicate recruitment or a binding process while sustained lines indicate the generation of a PTM by a complex; the figure has been designed with BioRender©



### 4.5 Conclusion and Outlook:

The identification of a SWR1-like complex in *T. brucei* that is responsible for incorporation of the histone variant H2A.Z closed the gap between the two histone acetyltransferases HAT2 and HAT1. The Luciferase experiments performed in course of this project showed a clear reduction of RNAP II transcription activity subsequent to the depletion of *Tb*SWR1, H2A.Z HAT2 and HAT1 confirming previous findings made in *T. brucei*. Northern Blot analysis confirmed these observations by showing a clear reduction of mRNA in *Tb*SWR1-depleted cells. SL RNA levels are not affected by *Tb*SWR1 depletion which underlines that SL transcription by RNAP II is regulated individually. Taken together these data create a coherent picture which is in line with the role of H2A.Z in transcription regulation: without H2A.Z, acetylation of H2A.Z cannot occur. Since this acetylation in turn is essential for transcription initiation its absence leads to a complete loss of RNAP II-mediated transcription. Nevertheless, it has to be mentioned that the data presented here are contradictory to the data from the research group of Nicolai Siegel that showed no reduction of mRNA levels subsequent to a loss of chromatin-associated H2A.Z due to HAT2 depletion.

In course of this project I found no evidence for the presence of an INO80 complex which might be the consequence of a permanent localization of H2A.Z in the TSSs. The Anp32e-like protein Tb927.5.1270 on the other hand might be an interesting candidate for further investigations regarding its ability to remove H2A.Z in course of the DNA DSB response.

The unique complex composition of the HAT2 complex, as well as the identification of some interesting features within the *Tb*SWR1 and HAT1 complex raised the question how these complexes interact with each other and how the parasite can distinguish between H2A.Z at a DSB and at a TSS.

Investigations of the transcription factors, PTM "writers" and "readers" as well as potential chromatin remodeling factors could provide some interesting information about the regulation of the chromatin landscape and transcription regulation in *T. brucei* in particular and maybe in *Kinetoplastids* in general. Taken together this project provided some new insights in H2A.Z-dependent transcription regulation and provides some very interesting approaches for new research projects.

## 5. Publication List

- 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(18), 6270-6279.
- Alatwi, H. E., & Downs, J. A. (2015). Removal of H2A.Z by INO80 promotes homologous recombination. *EMBO Rep*, 16(8), 986-994.
- Albert, I., Mavrich, T. N., Tomsho, L. P., Qi, J., Zanton, S. J., Schuster, S. C., & Pugh, B. F. (2007). Translational and rotational settings of H2A.Z nucleosomes across the *Saccharomyces cerevisiae* genome. *Nature*, 446(7135), 572-576.
- Allen, B. L., & Taatjes, D. J. (2015). The Mediator complex: a central integrator of transcription. *Nat Rev Mol Cell Biol*, 16(3), 155-166.
- Almawi, A. W., Matthews, L. A., & Guarné, A. (2017). FHA domains: Phosphopeptide binding and beyond. *Progress in biophysics and molecular biology*, 127, 105-110.
- 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(2), 142-148.
- Altaf, M., Auger, A., Monnet-Saksouk, J., Brodeur, J., Piquet, S., Cramet, M., Bouchard, N., Lacoste, N., Utley, R. T., Gaudreau, L., & Côté, 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(21), 15966-15977.
- Arhin, G. K., Shen, S., Ullu, E., & Tschudi, C. (2004). A PCR-based method for gene deletion and protein tagging in *Trypanosoma brucei*. *Methods Mol Biol*, 270, 277-286.
- Aslam, M., Fakher, B., Jakada, B. H., Cao, S., & Qin, Y. (2019). SWR1 Chromatin Remodeling Complex: A Key Transcriptional Regulator in Plants. *Cells*, 8(12).
- Aslett, M., Aurrecochea, C., Berriman, M., Brestelli, J., Brunk, B. P., Carrington, M., Depledge, D. P., Fischer, S., Gajria, B., Gao, X., Gardner, M. J. *et al.*, (2010). TriTrypDB: a functional genomic resource for the Trypanosomatidae. *Nucleic Acids Res*, 38(Database issue), D457-462.
- Ausió, J. (2006). Histone variants--the structure behind the function. *Brief Funct Genomic Proteomic*, 5(3), 228-243.
- Babiarz, J. E., Halley, J. E., & Rine, J. (2006). Telomeric heterochromatin boundaries require NuA4-dependent acetylation of histone variant H2A.Z in *Saccharomyces cerevisiae*. *Genes Dev*, 20(6), 700-710.
- Bagchi, D. N., Battenhouse, A. M., Park, D., & Iyer, V. R. (2020). The histone variant H2A.Z in yeast is almost exclusively incorporated into the +1 nucleosome in the direction of transcription. *Nucleic Acids Res*, 48(1), 157-170.
- Baker, N., Alsford, S., & Horn, D. (2011). Genome-wide RNAi screens in African trypanosomes identify the nifurtimox activator NTR and the eflornithine transporter AAT6. *Mol Biochem Parasitol*, 176(1), 55-57.
- Bannister, A. J., & Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res*, 21(3), 381-395.
- Bannister, A. J., Zegerman, P., Partridge, J. F., Miska, E. A., Thomas, J. O., Allshire, R. C., & Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature*, 410(6824), 120-124.
- Baptista, T., Grünberg, S., Minoungou, N., Koster, M. J. E., Timmers, H. T. M., Hahn, S., Devys, D., & Tora, L. (2017). SAGA Is a General Cofactor for RNA Polymerase II Transcription. *Mol Cell*, 68(1), 130-143.e135.
- Barski, A., Cuddapah, S., Cui, K., Roh, T. Y., Schones, D. E., Wang, Z., Wei, G., Chepelev, I., & Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell*, 129(4), 823-837.
- Bastin, P., Bagherzadeh, Z., Matthews, K. R., & Gull, K. (1996). A novel epitope tag system to study protein targeting and organelle biogenesis in *Trypanosoma brucei*. *Mol Biochem Parasitol*, 77(2), 235-239.

## 5. Publication List

- Batram, C., Jones, N. G., Janzen, C. J., Markert, S. M., & Engstler, M. (2014). Expression site attenuation mechanistically links antigenic variation and development in *Trypanosoma brucei*. *Elife*, *3*, e02324.
- Bellofatto, V., & Palenchar, J. B. (2008). RNA interference as a genetic tool in trypanosomes. *Methods Mol Biol*, *442*, 83-94.
- Benhamed, M., Bertrand, C., Servet, C., & Zhou, D. X. (2006). Arabidopsis GCN5, HD1, and TAF1/HAF2 interact to regulate histone acetylation required for light-responsive gene expression. *Plant Cell*, *18*(11), 2893-2903.
- Bensaude, O. (2011). Inhibiting eukaryotic transcription: Which compound to choose? How to evaluate its activity? *Transcription*, *2*(3), 103-108.
- Berriman, M., Ghedin, E., Hertz-Fowler, C., Blandin, G., Renauld, H., Bartholomeu, D. C., Lennard, N. J., Caler, E., Hamlin, N. E., Haas, B. *et al.*, (2005). The genome of the African trypanosome *Trypanosoma brucei*. *Science*, *309*(5733), 416-422.
- Beurton, F., Stempor, P., Caron, M., Appert, A., Dong, Y., Chen, R. A. J., Cluet, D., Couté, Y., Herbette, M., Huang, N. *et al.*, (2019). Physical and functional interaction between SET1/COMPASS complex component CFP-1 and a Sin3S HDAC complex in *C. elegans*. *Nucleic Acids Res*, *47*(21), 11164-11180.
- Bian, C., Xu, C., Ruan, J., Lee, K. K., Burke, T. L., Tempel, W., Barysyt, D., Li, J., Wu, M., Zhou, B. O. *et al.*, (2011). Sgf29 binds histone H3K4me2/3 and is required for SAGA complex recruitment and histone H3 acetylation. *EMBO J*, *30*(14), 2829-2842.
- Blumenthal, T., Evans, D., Link, C. D., Guffanti, A., Lawson, D., Thierry-Mieg, J., Thierry-Mieg, D., Chiu, W. L., Duke, K., Kiraly, M., & Kim, S. K. (2002). A global analysis of *Caenorhabditis elegans* operons. *Nature*, *417*(6891), 851-854.
- Bönisch, C., & Hake, S. B. (2012). Histone H2A variants in nucleosomes and chromatin: more or less stable? *Nucleic Acids Res*, *40*(21), 10719-10741.
- Boudreault, A. A., Cronier, D., Selleck, W., Lacoste, N., Utley, R. T., Allard, S., Savard, J., Lane, W. S., Tan, S., & Côté, J. (2003). Yeast enhancer of polycomb defines global Esa1-dependent acetylation of chromatin. *Genes & development*, *17*(11), 1415-1428.
- Bowman, B. R., Moure, C. M., Kirtane, B. M., Welschhans, R. L., Tominaga, K., Pereira-Smith, O. M., & Quijcho, F. A. (2006). Multipurpose MRG domain involved in cell senescence and proliferation exhibits structural homology to a DNA-interacting domain. *Structure (London, England : 1993)*, *14*(1), 151-158.
- Brahma, S., Udugama, M. I., Kim, J., Hada, A., Bhardwaj, S. K., Hailu, S. G., Lee, T. H., & Bartholomew, B. (2017). INO80 exchanges H2A.Z for H2A by translocating on DNA proximal to histone dimers. *Nat Commun*, *8*, 15616.
- Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N., & Ting, A. Y. (2018). Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol*, *36*(9), 880-887.
- Briggs, E., Crouch, K., Lemgruber, L., Hamilton, G., Lapsley, C., & McCulloch, R. (2019). *Trypanosoma brucei* ribonuclease H2A is an essential R-loop processing enzyme whose loss causes DNA damage during transcription initiation and antigenic variation. *Nucleic Acids Res*, *47*(17), 9180-9197.
- Brivanlou, A. H., & Darnell, J. E., Jr. (2002). Signal transduction and the control of gene expression. *Science*, *295*(5556), 813-818.
- Bruce, K., Myers, F. A., Mantouvalou, E., Lefevre, P., Greaves, I., Bonifer, C., Tremethick, D. J., Thorne, A. W., & 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*(17), 5633-5639.
- Brunelle, M., Nordell Markovits, A., Rodrigue, S., Lupien, M., Jacques, P., & Gérvy, N. (2015). The histone variant H2A.Z is an important regulator of enhancer activity. *Nucleic Acids Res*, *43*(20), 9742-9756.
- Campbell, D. A., Thomas, S., & Sturm, N. R. (2003). Transcription in kinetoplastid protozoa: why be normal? *Microbes Infect*, *5*(13), 1231-1240.
- Campos, E. I., & Reinberg, D. (2009). Histones: annotating chromatin. *Annu Rev Genet*, *43*, 559-599.

## 5. Publication List

- Cannavo, E., & Cejka, P. (2014). Sae2 promotes dsDNA endonuclease activity within Mre11-Rad50-Xrs2 to resect DNA breaks. *Nature*, *514*(7520), 122-125.
- Carninci, P., Sandelin, A., Lenhard, B., Katayama, S., Shimokawa, K., Ponjavic, J., Semple, C. A., Taylor, M. S., Engström, P. G., Frith, M. C. *et al.*, (2006). Genome-wide analysis of mammalian promoter architecture and evolution. *Nat Genet*, *38*(6), 626-635.
- Carruthers, N. J., Parker, G. C., Gratsch, T., Caruso, J. A., & Stemmer, P. M. (2015). Protein Mobility Shifts Contribute to Gel Electrophoresis Liquid Chromatography Analysis. *J Biomol Tech*, *26*(3), 103-112.
- Chen, C., Lim, H. H., Shi, J., Tamura, S., Maeshima, K., Surana, U., & Gan, L. (2016). Budding yeast chromatin is dispersed in a crowded nucleoplasm in vivo. *Mol Biol Cell*, *27*(21), 3357-3368.
- Chen, L., Conaway, R. C., & Conaway, J. W. (2013). Multiple modes of regulation of the human Ino80 SNF2 ATPase by subunits of the INO80 chromatin-remodeling complex. *Proc Natl Acad Sci U S A*, *110*(51), 20497-20502.
- Clapier, C. R. (2021). Sophisticated Conversations between Chromatin and Chromatin Remodelers, and Dissonances in Cancer. *International journal of molecular sciences*, *22*(11).
- Clapier, C. R., & Cairns, B. R. (2009). The biology of chromatin remodeling complexes. *Annu Rev Biochem*, *78*, 273-304.
- Clapier, C. R., Iwasa, J., Cairns, B. R., & Peterson, C. L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol*, *18*(7), 407-422.
- Clayton, C. (2019). Regulation of gene expression in trypanosomatids: living with polycistronic transcription. *Open Biol*, *9*(6), 190072.
- Clayton, C. E. (2016). Gene expression in Kinetoplastids. *Curr Opin Microbiol*, *32*, 46-51.
- Collins, R. E., Tachibana, M., Tamaru, H., Smith, K. M., Jia, D., Zhang, X., Selker, E. U., Shinkai, Y., & Cheng, X. (2005). In vitro and in vivo analyses of a Phe/Tyr switch controlling product specificity of histone lysine methyltransferases. *J Biol Chem*, *280*(7), 5563-5570.
- Courtin, D., Berthier, D., Thevenon, S., Dayo, G. K., Garcia, A., & Bucheton, B. (2008). Host genetics in African trypanosomiasis. *Infect Genet Evol*, *8*(3), 229-238.
- Couture, J.-P., Nolet, G., Beaulieu, E., Blouin, R., & Gévry, N. (2012). The p400/Brd8 chromatin remodeling complex promotes adipogenesis by incorporating histone variant H2A.Z at PPAR $\gamma$  target genes. *Endocrinology*, *153*(12), 5796-5808.
- Cox, E., Hwang, W., Uzoma, I., Hu, J., Guzzo, C. M., Jeong, J., Matunis, M. J., Qian, J., Zhu, H., & Blackshaw, S. (2017). Global Analysis of SUMO-Binding Proteins Identifies SUMOylation as a Key Regulator of the INO80 Chromatin Remodeling Complex. *Molecular & cellular proteomics : MCP*, *16*(5), 812-823.
- Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*, *26*(12), 1367-1372.
- Crick, F. (1970). Central dogma of molecular biology. *Nature*, *227*(5258), 561-563.
- Czermin, B., Schotta, G., Hülsmann, B. B., Brehm, A., Becker, P. B., Reuter, G., & Imhof, A. (2001). Physical and functional association of SU(VAR)3-9 and HDAC1 in Drosophila. *EMBO Rep*, *2*(10), 915-919.
- Dalmaso, M. C., Sullivan, W. J., Jr., & Angel, S. O. (2011). Canonical and variant histones of protozoan parasites. *Front Biosci (Landmark Ed)*, *16*, 2086-2105.
- Dalvai, M., Fleury, L., Bellucci, L., Kocanova, S., & Bystricky, K. (2013). TIP48/Reptin and H2A.Z requirement for initiating chromatin remodeling in estrogen-activated transcription. *PLoS Genet*, *9*(4), e1003387.
- DaRocha, W. D., Otsu, K., Teixeira, S. M., & Donelson, J. E. (2004). Tests of cytoplasmic RNA interference (RNAi) and construction of a tetracycline-inducible T7 promoter system in *Trypanosoma cruzi*. *Mol Biochem Parasitol*, *133*(2), 175-186.
- Dattani, A., & Wilkinson, S. R. (2019). Deciphering the interstrand crosslink DNA repair network expressed by *Trypanosoma brucei*. *DNA repair*, *78*, 154-166.
- Davies, K. P., Carruthers, V. B., & Cross, G. A. (1997). Manipulation of the vsg co-transposed region increases expression-site switching in *Trypanosoma brucei*. *Mol Biochem Parasitol*, *86*(2), 163-177.

## 5. Publication List

- Dean, S., Sunter, J., Wheeler, R. J., Hodkinson, I., Gluenz, E., & Gull, K. (2015). A toolkit enabling efficient, scalable and reproducible gene tagging in trypanosomatids. *Open Biol*, *5*(1), 140197.
- Dean, S., Sunter, J. D., & Wheeler, R. J. (2017). TrypTag.org: A Trypanosome Genome-wide Protein Localisation Resource. *Trends Parasitol*, *33*(2), 80-82.
- Denninger, V., & Rudenko, G. (2014). FACT plays a major role in histone dynamics affecting VSG expression site control in *Trypanosoma brucei*. *Mol Microbiol*, *94*(4), 945-962.
- Dhalluin, C., Carlson, J. E., Zeng, L., He, C., Aggarwal, A. K., & Zhou, M. M. (1999). Structure and ligand of a histone acetyltransferase bromodomain. *Nature*, *399*(6735), 491-496.
- Diaz, A., Nellore, A., & Song, J. S. (2012). CHANCE: comprehensive software for quality control and validation of ChIP-seq data. *Genome Biol*, *13*(10), R98.
- Dillon, S. C., Zhang, X., Trievel, R. C., & Cheng, X. (2005). The SET-domain protein superfamily: protein lysine methyltransferases. *Genome biology*, *6*(8), 227.
- Dinant, C., Houtsmuller, A. B., & Vermeulen, W. (2008). Chromatin structure and DNA damage repair. *Epigenetics Chromatin*, *1*(1), 9.
- Donczew, R., & Hahn, S. (2021). BET family members Bdf1/2 modulate global transcription initiation and elongation in *Saccharomyces cerevisiae*. *Elife*, *10*.
- Dorr, R. T. (1992). Bleomycin pharmacology: mechanism of action and resistance, and clinical pharmacokinetics. *Seminars in oncology*, *19*(2 Suppl 5), 3-8.
- Downs, J. A., Allard, S., Jobin-Robitaille, O., Javaheri, A., Auger, A., Bouchard, N., Kron, S. J., Jackson, S. P., & Côté, J. (2004). Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. *Mol Cell*, *16*(6), 979-990.
- Downs, J. A., Lowndes, N. F., & Jackson, S. P. (2000). A role for *Saccharomyces cerevisiae* histone H2A in DNA repair. *Nature*, *408*(6815), 1001-1004.
- Doyon, Y., Selleck, W., Lane, W. S., Tan, S., & Côté, J. (2004). Structural and functional conservation of the NuA4 histone acetyltransferase complex from yeast to humans. *Mol Cell Biol*, *24*(5), 1884-1896.
- Draker, R., Sarcinella, E., & Cheung, P. (2011). USP10 deubiquitylates the histone variant H2A.Z and both are required for androgen receptor-mediated gene activation. *Nucleic Acids Res*, *39*(9), 3529-3542.
- Durant, M., & Pugh, B. F. (2007). NuA4-directed chromatin transactions throughout the *Saccharomyces cerevisiae* genome. *Mol Cell Biol*, *27*(15), 5327-5335.
- Durocher, D., Taylor, I. A., Sarbassova, D., Haire, L. F., Westcott, S. L., Jackson, S. P., Smerdon, S. J., & Yaffe, M. B. (2000). The molecular basis of FHA domain:phosphopeptide binding specificity and implications for phospho-dependent signaling mechanisms. *Mol Cell*, *6*(5), 1169-1182.
- Easterling, L. F., Yerabolu, R., Kumar, R., Alzarieni, K. Z., & Kenttämaa, H. I. (2020). Factors Affecting the Limit of Detection for HPLC/Tandem Mass Spectrometry Experiments Based on Gas-Phase Ion-Molecule Reactions. *Anal Chem*, *92*(11), 7471-7477.
- Ebbert, R., Birkmann, A., & Schüller, H. J. (1999). The product of the SNF2/SWI2 paralogue INO80 of *Saccharomyces cerevisiae* required for efficient expression of various yeast structural genes is part of a high-molecular-weight protein complex. *Mol Microbiol*, *32*(4), 741-751.
- Eisen, A., Utley, R. T., Nourani, A., Allard, S., Schmidt, P., Lane, W. S., Lucchesi, J. C., & Cote, J. (2001). The yeast NuA4 and *Drosophila* MSL complexes contain homologous subunits important for transcription regulation. *J Biol Chem*, *276*(5), 3484-3491.
- Eisenhuth, N., Vellmer, T., Butter, F., & Janzen, C. J. (2020). A DOT1B/Ribonuclease H2 protein complex is involved in R-loop processing, genomic integrity and antigenic variation in *Trypanosoma brucei*. *bioRxiv*, 2020.2003.2002.969337.
- Eissenberg, J. C., & Shilatifard, A. (2009). Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev Biol*.
- Episkopou, H., Kyrtopoulos, S. A., Sfrikakis, P. P., Fousteri, M., Dimopoulos, M. A., Mullenders, L. H., & Souliotis, V. L. (2009). Association between transcriptional activity, local chromatin structure, and the efficiencies of both subpathways of nucleotide excision repair of melphalan adducts. *Cancer Res*, *69*(10), 4424-4433.
- Fan, J. Y., Rangasamy, D., Luger, K., & Tremethick, D. J. (2004). H2A.Z alters the nucleosome surface to promote HP1alpha-mediated chromatin fiber folding. *Mol Cell*, *16*(4), 655-661.

## 5. Publication List

- Farris, S. D., Rubio, E. D., Moon, J. J., Gombert, W. M., Nelson, B. H., & Krumm, A. (2005). Transcription-induced chromatin remodeling at the c-myc gene involves the local exchange of histone H2A.Z. *J Biol Chem*, *280*(26), 25298-25303.
- Fernandez-Martinez, J., LaCava, J., & Rout, M. P. (2016). Density Gradient Ultracentrifugation to Isolate Endogenous Protein Complexes after Affinity Capture. *Cold Spring Harbor protocols*, *2016*(7).
- Figueiredo, L. M., Cross, G. A., & Janzen, C. J. (2009). Epigenetic regulation in African trypanosomes: a new kid on the block. *Nat Rev Microbiol*, *7*(7), 504-513.
- Flaus, A., Martin, D. M., Barton, G. J., & Owen-Hughes, T. (2006). Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res*, *34*(10), 2887-2905.
- Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N., & Owen-Hughes, T. (2004). Sin mutations alter inherent nucleosome mobility. *EMBO J*, *23*(2), 343-353.
- Galarneau, L., Nourani, A., Boudreault, A. A., Zhang, Y., Héliot, L., Allard, S., Savard, J., Lane, W. S., Stillman, D. J., & Côté, J. (2000). Multiple links between the NuA4 histone acetyltransferase complex and epigenetic control of transcription. *Mol Cell*, *5*(6), 927-937.
- Gallant-Behm, C. L., Ramsey, M. R., Bensard, C. L., Nojek, I., Tran, J., Liu, M., Ellisen, L. W., & Espinosa, J. M. (2012).  $\Delta$ Np63 $\alpha$  represses anti-proliferative genes via H2A.Z deposition. *Genes Dev*, *26*(20), 2325-2336.
- Gassen, A., Brechtefeld, D., Schandry, N., Arteaga-Salas, J. M., Israel, L., Imhof, A., & Janzen, C. J. (2012). DOT1A-dependent H3K76 methylation is required for replication regulation in *Trypanosoma brucei*. *Nucleic Acids Res*, *40*(20), 10302-10311.
- Gerhold, C. B., & Gasser, S. M. (2014). INO80 and SWR complexes: relating structure to function in chromatin remodeling. *Trends Cell Biol*, *24*(11), 619-631.
- Gerhold, C. B., Winkler, D. D., Lakomek, K., Seifert, F. U., Fenn, S., Kessler, B., Witte, G., Luger, K., & Hopfner, K. P. (2012). Structure of Actin-related protein 8 and its contribution to nucleosome binding. *Nucleic Acids Res*, *40*(21), 11036-11046.
- Gévry, N., Chan, H. M., Laflamme, L., Livingston, D. M., & Gaudreau, L. (2007). p21 transcription is regulated by differential localization of histone H2A.Z. *Genes Dev*, *21*(15), 1869-1881.
- Gévry, N., Hardy, S., Jacques, P. E., Laflamme, L., Svtelisl, A., Robert, F., & Gaudreau, L. (2009). Histone H2A.Z is essential for estrogen receptor signaling. *Genes Dev*, *23*(13), 1522-1533.
- Giaimo, B. D., Ferrante, F., Herchenröther, A., Hake, S. B., & Borggreffe, T. (2019). The histone variant H2A.Z in gene regulation. *Epigenetics Chromatin*, *12*(1), 37.
- Giaimo, B. D., Ferrante, F., Vallejo, D. M., Hein, K., Gutierrez-Perez, I., Nist, A., Stiewe, T., Mittler, G., Herold, S., Zimmermann, T. *et al.*, (2018). Histone variant H2A.Z deposition and acetylation directs the canonical Notch signaling response. *Nucleic Acids Res*, *46*(16), 8197-8215.
- 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*(7), 1556-1564.
- Ginsburg, D. S., Anlembom, T. E., Wang, J., Patel, S. R., Li, B., & Hinnebusch, A. G. (2014). NuA4 links methylation of histone H3 lysines 4 and 36 to acetylation of histones H4 and H3. *J Biol Chem*, *289*(47), 32656-32670.
- Ginsburg, D. S., Govind, C. K., & Hinnebusch, A. G. (2009). NuA4 lysine acetyltransferase Esa1 is targeted to coding regions and stimulates transcription elongation with Gcn5. *Mol Cell Biol*, *29*(24), 6473-6487.
- Gommers-Ampt, J. H., Van Leeuwen, F., de Beer, A. L., Vliegenthart, J. F., Dizdaroglu, M., Kowalak, J. A., Crain, P. F., & Borst, P. (1993). beta-D-glucosyl-hydroxymethyluracil: a novel modified base present in the DNA of the parasitic protozoan *T. brucei*. *Cell*, *75*(6), 1129-1136.
- Goos, C., Dejung, M., Janzen, C. J., Butter, F., & Kramer, S. (2017). The nuclear proteome of *Trypanosoma brucei*. *PLoS One*, *12*(7), e0181884.
- Goos, C., Dejung, M., Wehman, A. M., E, M. N., Schmidt, J., Sunter, J., Engstler, M., Butter, F., & Kramer, S. (2019). Trypanosomes can initiate nuclear export co-transcriptionally. *Nucleic Acids Res*, *47*(1), 266-282.
- Gossen, M., & Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline-responsive promoters. *Proc Natl Acad Sci U S A*, *89*(12), 5547-5551.

## 5. Publication List

- Gradolatto, A., Rogers, R. S., Lavender, H., Taverna, S. D., Allis, C. D., Aitchison, J. D., & Tackett, A. J. (2008). *Saccharomyces cerevisiae* Yta7 regulates histone gene expression. *Genetics*, *179*(1), 291-304.
- Gradolatto, A., Smart, S. K., Byrum, S., Blair, L. P., Rogers, R. S., Kolar, E. A., Lavender, H., Larson, S. K., Aitchison, J. D., Taverna, S. D., & Tackett, A. J. (2009). A noncanonical bromodomain in the AAA ATPase protein Yta7 directs chromosomal positioning and barrier chromatin activity. *Mol Cell Biol*, *29*(17), 4604-4611.
- Grant, P. A., Duggan, L., Cote, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F. *et al.*, (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes Dev*, *11*(13), 1640-1650.
- Graves, J. D., & Krebs, E. G. (1999). Protein phosphorylation and signal transduction. *Pharmacol Ther*, *82*(2-3), 111-121.
- Greaves, I. K., Rangasamy, D., Ridgway, P., & Tremethick, D. J. (2007). H2A.Z contributes to the unique 3D structure of the centromere. *Proc Natl Acad Sci U S A*, *104*(2), 525-530.
- Gribun, A., Cheung, K. L., Huen, J., Ortega, J., & Houry, W. A. (2008). Yeast Rvb1 and Rvb2 are ATP-dependent DNA helicases that form a heterohexameric complex. *J Mol Biol*, *376*(5), 1320-1333.
- Guarner, J. (2019). Chagas disease as example of a reemerging parasite. *Seminars in diagnostic pathology*, *36*(3), 164-169.
- Guillemette, B., Bataille, A. R., Gévry, 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*(12), e384.
- Günzl, A., Ullu, E., Dörner, M., Fragoso, S. P., Hoffmann, K. F., Milner, J. D., Morita, Y., Nguu, E. K., Vanacova, S., Wünsch, S. *et al.*, (1997). Transcription of the *Trypanosoma brucei* spliced leader RNA gene is dependent only on the presence of upstream regulatory elements. *Mol Biochem Parasitol*, *85*(1), 67-76.
- Gurard-Levin, Z. A., Quivy, J.-P., & Almouzni, G. (2014). Histone chaperones: assisting histone traffic and nucleosome dynamics. *Annual review of biochemistry*, *83*, 487-517.
- Gursoy-Yuzugullu, O., Ayrapetov, M. K., & Price, B. D. (2015). Histone chaperone Anp32e removes H2A.Z from DNA double-strand breaks and promotes nucleosome reorganization and DNA repair. *Proc Natl Acad Sci U S A*, *112*(24), 7507-7512.
- Hahne, F., & Ivanek, R. (2016). Visualizing Genomic Data Using Gviz and Bioconductor. *Methods Mol Biol*, *1418*, 335-351.
- Halley, J. E., Kaplan, T., Wang, A. Y., Kobor, M. S., & Rine, J. (2010). Roles for H2A.Z and its acetylation in GAL1 transcription and gene induction, but not GAL1-transcriptional memory. *PLoS Biol*, *8*(6), e1000401.
- Harata, M., Oma, Y., Mizuno, S., Jiang, Y. W., Stillman, D. J., & Wintersberger, U. (1999). The nuclear actin-related protein of *Saccharomyces cerevisiae*, Act3p/Arp4, interacts with core histones. *Molecular biology of the cell*, *10*(8), 2595-2605.
- Hardy, S., Jacques, P. E., Gévry, N., Forest, A., Fortin, M. E., 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*(10), e1000687.
- Hauer, M. H., Seeber, A., Singh, V., Thierry, R., Sack, R., Amitai, A., Kryzhanovska, M., Eglinger, J., Holcman, D., Owen-Hughes, T., & Gasser, S. M. (2017). Histone degradation in response to DNA damage enhances chromatin dynamics and recombination rates. *Nature structural & molecular biology*, *24*(2), 99-107.
- Heisel, S., Habel, N. C., Schuetz, N., Ruggieri, A., & Meese, E. (2010). The YEATS family member GAS41 interacts with the general transcription factor TFIIF. *BMC Mol Biol*, *11*, 53.
- Hewish, D. R., & Burgoyne, L. A. (1973). Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem Biophys Res Commun*, *52*(2), 504-510.
- Hill, C. S., & Treisman, R. (1995). Transcriptional regulation by extracellular signals: mechanisms and specificity. *Cell*, *80*(2), 199-211.

## 5. Publication List

- Hoffmann, A., & Spengler, D. (2019). Chromatin Remodeling Complex NuRD in Neurodevelopment and Neurodevelopmental Disorders. *Front Genet*, *10*, 682.
- Höög, J. L., Gluenz, E., Vaughan, S., & Gull, K. (2010). Ultrastructural investigation methods for *Trypanosoma brucei*. *Methods Cell Biol*, *96*, 175-196.
- Horigome, C., Oma, Y., Konishi, T., Schmid, R., Marcomini, I., Hauer, M. H., Dion, V., Harata, M., & Gasser, S. M. (2014). SWR1 and INO80 chromatin remodelers contribute to DNA double-strand break perinuclear anchorage site choice. *Mol Cell*, *55*(4), 626-639.
- Hsu, P. L., Shi, H., Leonen, C., Kang, J., Chatterjee, C., & Zheng, N. (2019). Structural Basis of H2B Ubiquitination-Dependent H3K4 Methylation by COMPASS. *Mol Cell*, *76*(5), 712-723.e714.
- Hu, A., Noble, W. S., & Wolf-Yadlin, A. (2016). Technical advances in proteomics: new developments in data-independent acquisition. *F1000Research*, *5*.
- 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*(9), 2400-2410.
- Janzen, C. J., Fernandez, J. P., Deng, H., Diaz, R., Hake, S. B., & Cross, G. A. (2006). Unusual histone modifications in *Trypanosoma brucei*. *FEBS Lett*, *580*(9), 2306-2310.
- Jeffers, V., Yang, C., Huang, S., & Sullivan, W. J., Jr. (2017). Bromodomains in Protozoan Parasites: Evolution, Function, and Opportunities for Drug Development. *Microbiol Mol Biol Rev*, *81*(1).
- Jeronimo, C., Watanabe, S., Kaplan, C. D., Peterson, C. L., & Robert, F. (2015). The Histone Chaperones FACT and Spt6 Restrict H2A.Z from Intragenic Locations. *Mol Cell*, *58*(6), 1113-1123.
- Jiang, X., Soboleva, T. A., & Tremethick, D. J. (2020). Short Histone H2A Variants: Small in Stature but not in Function. *Cells*, *9*(4).
- Jin, C., & Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. *Genes & development*, *21*(12), 1519-1529.
- 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*(8), 941-945.
- Jin, Q., Yu, L.-R., Wang, L., Zhang, Z., Kasper, L. H., Lee, J.-E., Wang, C., Brindle, P. K., Dent, S. Y. R., & Ge, K. (2011). Distinct roles of GCN5/PCAF-mediated H3K9ac and CBP/p300-mediated H3K18/27ac in nuclear receptor transactivation. *EMBO J*, *30*(2), 249-262.
- Joti, Y., Hikima, T., Nishino, Y., Kamada, F., Hihara, S., Takata, H., Ishikawa, T., & Maeshima, K. (2012). Chromosomes without a 30-nm chromatin fiber. *Nucleus*, *3*(5), 404-410.
- Karmodiya, K., Krebs, A. R., Oulad-Abdelghani, M., Kimura, H., & Tora, L. (2012). H3K9 and H3K14 acetylation co-occur at many gene regulatory elements, while H3K14ac marks a subset of inactive inducible promoters in mouse embryonic stem cells. *BMC Genomics*, *13*, 424.
- Kashiwaba, S., Kitahashi, K., Watanabe, T., Onoda, F., Ohtsu, M., & Murakami, Y. (2010). The mammalian INO80 complex is recruited to DNA damage sites in an ARP8 dependent manner. *Biochem Biophys Res Commun*, *402*(4), 619-625.
- Katan-Khaykovich, Y., & Struhl, K. (2005). Heterochromatin formation involves changes in histone modifications over multiple cell generations. *EMBO J*, *24*(12), 2138-2149.
- Kawahara, T., Siegel, T. N., Ingram, A. K., Alsford, S., Cross, G. A., & Horn, D. (2008). Two essential MYST-family proteins display distinct roles in histone H4K10 acetylation and telomeric silencing in trypanosomes. *Mol Microbiol*, *69*(4), 1054-1068.
- Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N., & Sternberg, M. J. E. (2015). The Phyre2 web portal for protein modeling, prediction and analysis. *Nat Protoc*, *10*(6), 845-858.
- Keogh, M. C., Mennella, T. A., Sawa, C., Berthelet, S., Krogan, N. J., Wolek, A., Podolny, V., Carpenter, L. R., Greenblatt, J. F., Baetz, K., & Buratowski, S. (2006). The *Saccharomyces cerevisiae* histone H2A variant Htz1 is acetylated by NuA4. *Genes Dev*, *20*(6), 660-665.
- Kieft, R., Zhang, Y., Marand, A. P., Moran, J. D., Bridger, R., Wells, L., Schmitz, R. J., & Sabatini, R. (2020). Identification of a novel base J binding protein complex involved in RNA polymerase II transcription termination in trypanosomes. *PLoS Genet*, *16*(2), e1008390.



## 5. Publication List

- Kim, H.-S., Vanoosthuysse, V., Fillingham, J., Roguev, A., Watt, S., Kislinger, T., Treyer, A., Carpenter, L. R., Bennett, C. S., Emili, A. *et al.*, (2009). An acetylated form of histone H2A.Z regulates chromosome architecture in *Schizosaccharomyces pombe*. *Nature structural & molecular biology*, *16*(12), 1286-1293.
- Kimura, A., & Horikoshi, M. (1998). Tip60 acetylates six lysines of a specific class in core histones in vitro. *Genes Cells*, *3*(12), 789-800.
- Klein, B. J., Ahmad, S., Vann, K. R., Andrews, F. H., Mayo, Z. A., Bourriquen, G., Bridgers, J. B., Zhang, J., Strahl, B. D., Côté, J., & Kutateladze, T. G. (2018). Yaf9 subunit of the NuA4 and SWR1 complexes targets histone H3K27ac through its YEATS domain. *Nucleic Acids Res*, *46*(1), 421-430.
- Kornberg, R. D., & Thomas, J. O. (1974). Chromatin structure; oligomers of the histones. *Science*, *184*(4139), 865-868.
- Kornitzer, D., & Ciechanover, A. (2000). Modes of regulation of ubiquitin-mediated protein degradation. *J Cell Physiol*, *182*(1), 1-11.
- Kouskouti, A., Scheer, E., Staub, A., Tora, L., & Talianidis, I. (2004). Gene-specific modulation of TAF10 function by SET9-mediated methylation. *Mol Cell*, *14*(2), 175-182.
- Kramer, S., Queiroz, R., Ellis, L., Webb, H., Hoheisel, J. D., Clayton, C., & Carrington, M. (2008). Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2(alpha) phosphorylation at Thr169. *J Cell Sci*, *121*(Pt 18), 3002-3014.
- Kraus, A. J., Vanselow, J. T., Lamer, S., Brink, B. G., Schlosser, A., & Siegel, T. N. (2020). Distinct roles for H4 and H2A.Z acetylation in RNA transcription in African trypanosomes. *Nat Commun*, *11*(1), 1498.
- Krogan, N. J., Keogh, M. C., Datta, N., Sawa, C., Ryan, O. W., Ding, H., Haw, R. A., Pootoolal, J., Tong, A., Canadien, V., Richards, D. P., Wu, X., Emili, A., Hughes, T. R., Buratowski, S., & Greenblatt, J. F. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Mol Cell*, *12*(6), 1565-1576.
- Kusch, T., Florens, L., Macdonald, W. H., Swanson, S. K., Glaser, R. L., Yates, J. R., 3rd, Abmayr, S. M., Washburn, M. P., & Workman, J. L. (2004). Acetylation by Tip60 is required for selective histone variant exchange at DNA lesions. *Science*, *306*(5704), 2084-2087.
- Kusch, T., Mei, A., & Nguyen, C. (2014). Histone H3 lysine 4 trimethylation regulates cotranscriptional H2A variant exchange by Tip60 complexes to maximize gene expression. *Proc Natl Acad Sci U S A*, *111*(13), 4850-4855.
- Lange, M., Kaynak, B., Forster, U. B., Tönjes, M., Fischer, J. J., Grimm, C., Schlesinger, J., Just, S., Dunkel, I., Krueger, T., Mebus, S., Lehrach, H., Lurz, R., Gobom, J., Rottbauer, W., Abdelilah-Seyfried, S., & Sperling, S. (2008). Regulation of muscle development by DPF3, a novel histone acetylation and methylation reader of the BAF chromatin remodeling complex. *Genes Dev*, *22*(17), 2370-2384.
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods*, *9*(4), 357-359.
- Larsen, M. R., Trelle, M. B., Thingholm, T. E., & Jensen, O. N. (2006). Analysis of posttranslational modifications of proteins by tandem mass spectrometry. *BioTechniques*, *40*(6), 790-798.
- Latham, J. A., & Dent, S. Y. R. (2007). Cross-regulation of histone modifications. *Nature structural & molecular biology*, *14*(11), 1017-1024.
- Lee, J. H., Nguyen, T. N., Schimanski, B., & Günzl, A. (2007). Spliced leader RNA gene transcription in *Trypanosoma brucei* requires transcription factor TFIIH. *Eukaryot Cell*, *6*(4), 641-649.
- Lee, K. K., Zhang, Y., Tirado-Magallanes, R., Rajagopalan, D., Bhatia, S. S., Ng, L., Desi, N., Tham, C. Y., Teo, W. S., Hoppe, M. M. *et al.*, (2020). TIP60 acetylates H2AZ and regulates doxorubicin-induced DNA damage sensitivity through *RAD51* transcription. *bioRxiv*, 2020.2006.2010.145193.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., & Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078-2079.
- Li, Y., Wen, H., Xi, Y., Tanaka, K., Wang, H., Peng, D., Ren, Y., Jin, Q., Dent, S. Y., Li, W., Li, H., & Shi, X. (2014). AF9 YEATS domain links histone acetylation to DOT1L-mediated H3K79 methylation. *Cell*, *159*(3), 558-571.
- Liigand, P., Kaupmees, K., & Kruve, A. (2019). Influence of the amino acid composition on the ionization efficiencies of small peptides. *Journal of mass spectrometry : JMS*, *54*(6), 481-487.

## 5. Publication List

- Liu, J. C., Li, Q. J., He, M. H., Hu, C., Dai, P., Meng, F. L., Zhou, B. O., & Zhou, J. Q. (2020). Swc4 positively regulates telomere length independently of its roles in NuA4 and SWR1 complexes. *Nucleic Acids Res*, *48*(22), 12792-12803.
- Lombardi, L. M., Ellahi, A., & Rine, J. (2011). Direct regulation of nucleosome density by the conserved AAA-ATPase Yta7. *Proc Natl Acad Sci U S A*, *108*(49), E1302-1311.
- López-Farfán, D., Bart, J. M., Rojas-Barros, D. I., & Navarro, M. (2014). SUMOylation by the E3 ligase TbSIZ1/PIAS1 positively regulates VSG expression in *Trypanosoma brucei*. *PLoS Pathog*, *10*(12), e1004545.
- Lu, P. Y. T., Lévesque, N., & Kobor, M. S. (2009). NuA4 and SWR1-C: two chromatin-modifying complexes with overlapping functions and components. *Biochemistry and cell biology = Biochimie et biologie cellulaire*, *87*(5), 799-815.
- Lu, R., & Wang, G. G. (2013). Tudor: a versatile family of histone methylation 'readers'. *Trends Biochem Sci*, *38*(11), 546-555.
- Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F., & Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*, *389*(6648), 251-260.
- Luk, E., Ranjan, A., Fitzgerald, P. C., 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*(5), 725-736.
- Luse, D. S. (2012). Rethinking the role of TFIIF in transcript initiation by RNA polymerase II. *Transcription*, *3*(4), 156-159.
- Lye, L.-F., Owens, K., Shi, H., Murta, S. M. F., Vieira, A. C., Turco, S. J., Tschudi, C., Ullu, E., & Beverley, S. M. (2010). Retention and loss of RNA interference pathways in trypanosomatid protozoans. *PLoS Pathog*, *6*(10), e1001161.
- MacAlpine, D. M., & Almouzni, G. (2013). Chromatin and DNA replication. *Cold Spring Harb Perspect Biol*, *5*(8), a010207.
- Machado-Pinilla, R., Liger, D., Leulliot, N., & Meier, U. T. (2012). Mechanism of the AAA+ ATPases pontin and reptin in the biogenesis of H/ACA RNPs. *RNA*, *18*(10), 1833-1845.
- Madigan, J. P., Chotkowski, H. L., & Glaser, R. L. (2002). DNA double-strand break-induced phosphorylation of *Drosophila* histone variant H2Av helps prevent radiation-induced apoptosis. *Nucleic Acids Res*, *30*(17), 3698-3705.
- Maeda, K., Kosaka, H., Yagishita, K., & Umezawa, H. (1956). A new antibiotic, phleomycin. *The Journal of antibiotics*, *9*(2), 82-85.
- Malik, H. S., & Henikoff, S. (2003). Phylogenomics of the nucleosome. *Nat Struct Biol*, *10*(11), 882-891.
- Malvy, D., & Chappuis, F. (2011). Sleeping sickness. *Clin Microbiol Infect*, *17*(7), 986-995.
- Mandava, V., Fernandez, J. P., Deng, H., Janzen, C. J., Hake, S. B., & Cross, G. A. (2007). Histone modifications in *Trypanosoma brucei*. *Mol Biochem Parasitol*, *156*(1), 41-50.
- Marinozzi, V., & Fiume, L. (1971). Effects of -amanitin on mouse and rat liver cell nuclei. *Exp Cell Res*, *67*(2), 311-322.
- Markert, S. M., Bauer, V., Muenz, T. S., Jones, N. G., Helmprobst, F., Britz, S., Sauer, M., Rössler, W., Engstler, M., & Stigloher, C. (2017). 3D subcellular localization with superresolution array tomography on ultrathin sections of various species. *Methods Cell Biol*, *140*, 21-47.
- Martire, S., & Banaszynski, L. A. (2020). The roles of histone variants in fine-tuning chromatin organization and function. *Nat Rev Mol Cell Biol*, *21*(9), 522-541.
- Masumoto, H., Hawke, D., Kobayashi, R., & Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature*, *436*(7048), 294-298.
- Matangkasombut, O., Buratowski, R. M., Swilling, N. W., & Buratowski, S. (2000). Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. *Genes Dev*, *14*(8), 951-962.
- Matangkasombut, O., & Buratowski, S. (2003). Different sensitivities of bromodomain factors 1 and 2 to histone H4 acetylation. *Mol Cell*, *11*(2), 353-363.
- Mavrich, T. N., Jiang, C., Ioshikhes, I. P., Li, X., Venters, B. J., Zanton, S. J., Tomsho, L. P., Qi, J., Glaser, R. L., Schuster, S. C., Gilmour, D. S., Albert, I., & Pugh, B. F. (2008). Nucleosome organization in the *Drosophila* genome. *Nature*, *453*(7193), 358-362.

## 5. Publication List

- 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*(1), 65-76.
- Mehta, M., Braberg, H., Wang, S., Lozsa, A., Shales, M., Solache, A., Krogan, N. J., & Keogh, M. C. (2010). Individual lysine acetylations on the N terminus of *Saccharomyces cerevisiae* H2A.Z are highly but not differentially regulated. *J Biol Chem*, *285*(51), 39855-39865.
- Migliori, V., Müller, J., Phalke, S., Low, D., Bezzi, M., Mok, W. C., Sahu, S. K., Gunaratne, J., Capasso, P., Bassi, C. *et al.*, (2012). Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nat Struct Mol Biol*, *19*(2), 136-144.
- Millar, C. B., Xu, F., Zhang, K., & Grunstein, M. (2006). Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes Dev*, *20*(6), 711-722.
- Miller, T., Krogan, N. J., Dover, J., Erdjument-Bromage, H., Tempst, P., Johnston, M., Greenblatt, J. F., & Shilatifard, A. (2001). COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc Natl Acad Sci U S A*, *98*(23), 12902-12907.
- Mizuguchi, G., Shen, X., Landry, J., Wu, W. H., Sen, S., & Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science*, *303*(5656), 343-348.
- Molaro, A., & Drinnenberg, I. A. (2018). Studying the Evolution of Histone Variants Using Phylogeny. *Methods Mol Biol*, *1832*, 273-291.
- Morrison, A. J., Highland, J., Krogan, N. J., Arbel-Eden, A., Greenblatt, J. F., Haber, J. E., & Shen, X. (2004). INO80 and gamma-H2AX interaction links ATP-dependent chromatin remodeling to DNA damage repair. *Cell*, *119*(6), 767-775.
- Morrison, A. J., & Shen, X. (2009). Chromatin remodelling beyond transcription: the INO80 and SWR1 complexes. *Nat Rev Mol Cell Biol*, *10*(6), 373-384.
- Morriswood, B., Havlicek, K., Demmel, L., Yavuz, S., Sealey-Cardona, M., Vidilaseris, K., Anrather, D., Kostan, J., Djinojic-Carugo, K., Roux, K. J., & Warren, G. (2013). Novel bilobe components in *Trypanosoma brucei* identified using proximity-dependent biotinylation. *Eukaryot Cell*, *12*(2), 356-367.
- Mujtaba, S., Zeng, L., & Zhou, M. M. (2007). Structure and acetyl-lysine recognition of the bromodomain. *Oncogene*, *26*(37), 5521-5527.
- Muller, L. S. M., Cosentino, R. O., Forstner, K. U., Guizetti, J., Wedel, C., Kaplan, N., Janzen, C. J., Arampatzi, P., Vogel, J., Steinbiss, S., Otto, T. D., Saliba, A. E., Sebra, R. P., & Siegel, T. N. (2018). Genome organization and DNA accessibility control antigenic variation in trypanosomes. *Nature*, *563*(7729), 121-125.
- Myung, K., & Smith, S. (2008). The RAD5-dependent postreplication repair pathway is important to suppress gross chromosomal rearrangements. *Journal of the National Cancer Institute. Monographs*(39), 12-15.
- Nakamura, T. M., Du, L. L., Redon, C., & Russell, P. (2004). Histone H2A phosphorylation controls Crb2 recruitment at DNA breaks, maintains checkpoint arrest, and influences DNA repair in fission yeast. *Mol Cell Biol*, *24*(14), 6215-6230.
- Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D., & Grewal, S. I. (2001). Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science (New York, N.Y.)*, *292*(5514), 110-113.
- Narkaj, K., Stefanelli, G., Wahdan, M., Azam, A. B., Ramzan, F., Steininger, C. F. D., Jr., Walters, B. J., & Zovkic, I. B. (2018). Blocking H2A.Z Incorporation via Tip60 Inhibition Promotes Systems Consolidation of Fear Memory in Mice. *eNeuro*, *5*(5).
- Narlikar, G. J., Sundaramoorthy, R., & Owen-Hughes, T. (2013). Mechanisms and functions of ATP-dependent chromatin-remodeling enzymes. *Cell*, *154*(3), 490-503.
- Navarro, M., & Cross, G. A. (1998). In situ analysis of a variant surface glycoprotein expression-site promoter region in *Trypanosoma brucei*. *Mol Biochem Parasitol*, *94*(1), 53-66.
- Nguyen, T. N., & Goodrich, J. A. (2006). Protein-protein interaction assays: eliminating false positive interactions. *Nature methods*, *3*(2), 135-139.

## 5. Publication List

- Nguyen, T. N., Muller, L. S., Park, S. H., Siegel, T. N., & Gunzl, A. (2014). Promoter occupancy of the basal class I transcription factor A differs strongly between active and silent VSG expression sites in *Trypanosoma brucei*. *Nucleic Acids Res*, *42*(5), 3164-3176.
- Nguyen, T. N., Nguyen, B. N., Lee, J. H., Panigrahi, A. K., & Günzl, A. (2012). Characterization of a novel class I transcription factor A (CITFA) subunit that is indispensable for transcription by the multifunctional RNA polymerase I of *Trypanosoma brucei*. *Eukaryot Cell*, *11*(12), 1573-1581.
- Nguyen, V. Q., Ranjan, A., Stengel, F., Wei, D., Aebersold, R., Wu, C., & Leschziner, A. E. (2013). Molecular architecture of the ATP-dependent chromatin-remodeling complex SWR1. *Cell*, *154*(6), 1220-1231.
- Nishioka, K., Chuikov, S., Sarma, K., Erdjument-Bromage, H., Allis, C. D., Tempst, P., & Reinberg, D. (2002). Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes & development*, *16*(4), 479-489.
- Obri, A., Ouararhni, K., Papin, C., Diebold, M.-L., Padmanabhan, K., Marek, M., Stoll, I., Roy, L., Reilly, P. T., Mak, T. W., Dimitrov, S., Romier, C., & Hamiche, A. (2014). ANP32E is a histone chaperone that removes H2A.Z from chromatin. *Nature*, *505*(7485), 648-653.
- Orphanides, G., Lagrange, T., & Reinberg, D. (1996). The general transcription factors of RNA polymerase II. *Genes & development*, *10*(21), 2657-2683.
- Otsuji, N., Iyehara, H., & Hideshima, Y. (1974). Isolation and characterization of an *Escherichia coli* *ruv* mutant which forms nonseptate filaments after low doses of ultraviolet light irradiation. *J Bacteriol*, *117*(2), 337-344.
- Otzen, D. E. (2002). Protein unfolding in detergents: effect of micelle structure, ionic strength, pH, and temperature. *Biophysical journal*, *83*(4), 2219-2230.
- Owczarzy, R., You, Y., Moreira, B. G., Manthey, J. A., Huang, L., Behlke, M. A., & Walder, J. A. (2004). Effects of sodium ions on DNA duplex oligomers: improved predictions of melting temperatures. *Biochemistry*, *43*(12), 3537-3554.
- Owen-Hughes, T., & Workman, J. L. (1994). Experimental analysis of chromatin function in transcription control. *Crit Rev Eukaryot Gene Expr*, *4*(4), 403-441.
- Papai, G., Weil, P. A., & Schultz, P. (2011). New insights into the function of transcription factor TFIID from recent structural studies. *Curr Opin Genet Dev*, *21*(2), 219-224.
- Papamichos-Chronakis, M., Krebs, J. E., & Peterson, C. L. (2006). Interplay between Ino80 and Swr1 chromatin remodeling enzymes regulates cell cycle checkpoint adaptation in response to DNA damage. *Genes Dev*, *20*(17), 2437-2449.
- Park, Y.-J., Dyer, P. N., Tremethick, D. J., & Luger, K. (2004). A new fluorescence resonance energy transfer approach demonstrates that the histone variant H2AZ stabilizes the histone octamer within the nucleosome. *J Biol Chem*, *279*(23), 24274-24282.
- Pokholok, D. K., Harbison, C. T., Levine, S., Cole, M., Hannett, N. M., Lee, T. I., Bell, G. W., Walker, K., Rolfe, P. A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D. K., & Young, R. A. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, *122*(4), 517-527.
- Poli, J., Gasser, S. M., & Papamichos-Chronakis, M. (2017). The INO80 remodeler in transcription, replication and repair. *Philos Trans R Soc Lond B Biol Sci*, *372*(1731).
- Raisner, R. M., Hartley, P. D., Meneghini, M. D., Bao, M. Z., Liu, C. L., Schreiber, S. L., Rando, O. J., & Madhani, H. D. (2005). Histone variant H2A.Z marks the 5' ends of both active and inactive genes in euchromatin. *Cell*, *123*(2), 233-248.
- Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., Heyne, S., Dündar, F., & Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res*, *44*(W1), W160-165.
- Rappsilber, J., Mann, M., & Ishihama, Y. (2007). Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat Protoc*, *2*(8), 1896-1906.
- Redmond, S., Vadivelu, J., & Field, M. C. (2003). RNAi: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei*. *Mol Biochem Parasitol*, *128*(1), 115-118.

## 5. Publication List

- Reis, H., Schwebs, M., Dietz, S., Janzen, C. J., & Butter, F. (2018). TelAP1 links telomere complexes with developmental expression site silencing in African trypanosomes. *Nucleic Acids Res*, *46*(6), 2820-2833.
- Reynolds, D., Hofmeister, B. T., Cliffe, L., Alabady, M., Siegel, T. N., Schmitz, R. J., & 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*(1), e1005758.
- Rico, E., Ivens, A., Glover, L., Horn, D., & Matthews, K. R. (2017). Genome-wide RNAi selection identifies a regulator of transmission stage-enriched gene families and cell-type differentiation in *Trypanosoma brucei*. *PLoS Pathog*, *13*(3), e1006279.
- Ringel, A. E., Cieniewicz, A. M., Taverna, S. D., & Wolberger, C. (2015). Nucleosome competition reveals processive acetylation by the SAGA HAT module. *Proc Natl Acad Sci U S A*, *112*(40), E5461-5470.
- Riss, A., Scheer, E., Joint, M., Trowitzsch, S., Berger, I., & Tora, L. (2015). Subunits of ADA-two-A-containing (ATAC) or Spt-Ada-Gcn5-acetyltransferase (SAGA) Coactivator Complexes Enhance the Acetyltransferase Activity of GCN5. *J Biol Chem*, *290*(48), 28997-29009.
- Roeder, R. G. (1976). *Eukaryotic nuclear RNA polymerases* (Vol. 285). Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S., & Bonner, W. M. (1998). DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J Biol Chem*, *273*(10), 5858-5868.
- Ruan, J. P., Arhin, G. K., 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*(21), 9610-9618.
- Ruhl, D. D., Jin, J., Cai, Y., Swanson, S., Florens, L., Washburn, M. P., Conaway, R. C., Conaway, J. W., & Chrivia, J. C. (2006). Purification of a human SRCAP complex that remodels chromatin by incorporating the histone variant H2A.Z into nucleosomes. *Biochemistry*, *45*(17), 5671-5677.
- Ryan, D. P., & Tremethick, D. J. (2018). The interplay between H2A.Z and H3K9 methylation in regulating HP1 $\alpha$  binding to linker histone-containing chromatin. *Nucleic Acids Res*, *46*(18), 9353-9366.
- Sanchez, R., & Zhou, M.-M. (2011). The PHD finger: a versatile epigenome reader. *Trends in biochemical sciences*, *36*(7), 364-372.
- SantaLucia, J., Jr. (1998). A unified view of polymer, dumbbell, and oligonucleotide DNA nearest-neighbor thermodynamics. *Proc Natl Acad Sci U S A*, *95*(4), 1460-1465.
- Santos-Rosa, H., Schneider, R., Bannister, A. J., Sherriff, J., Bernstein, B. E., Emre, N. C., Schreiber, S. L., Mellor, J., & Kouzarides, T. (2002). Active genes are tri-methylated at K4 of histone H3. *Nature*, *419*(6905), 407-411.
- Saravanan, M., Wuerges, J., Bose, D., McCormack, E. A., Cook, N. J., Zhang, X., & Wigley, D. B. (2012). Interactions between the nucleosome histone core and Arp8 in the INO80 chromatin remodeling complex. *Proc Natl Acad Sci U S A*, *109*(51), 20883-20888.
- Sarcinella, E., Zuzarte, P. C., Lau, P. N., Draker, R., & Cheung, P. (2007). Monoubiquitylation of H2A.Z distinguishes its association with euchromatin or facultative heterochromatin. *Mol Cell Biol*, *27*(18), 6457-6468.
- Schildkraut, C. (1965). Dependence of the melting temperature of DNA on salt concentration. *Biopolymers*, *3*(2), 195-208.
- Schimanski, B., Brandenburg, J., Nguyen, T. N., Caimano, M. J., & Gunzl, A. (2006). A TFIIB-like protein is indispensable for spliced leader RNA gene transcription in *Trypanosoma brucei*. *Nucleic Acids Res*, *34*(6), 1676-1684.
- Schimanski, B., Nguyen, T. N., & Gunzl, A. (2005a). Characterization of a multisubunit transcription factor complex essential for spliced-leader RNA gene transcription in *Trypanosoma brucei*. *Mol Cell Biol*, *25*(16), 7303-7313.
- Schimanski, B., Nguyen, T. N., & Gunzl, A. (2005b). Highly efficient tandem affinity purification of trypanosome protein complexes based on a novel epitope combination. *Eukaryot Cell*, *4*(11), 1942-1950.
- Schones, D. E., Cui, K., Cuddapah, S., Roh, T. Y., Barski, A., Wang, Z., Wei, G., & Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell*, *132*(5), 887-898.

## 5. Publication List

- Schroeder, H. W., & Cavacini, L. (2010). Structure and function of immunoglobulins. *The Journal of allergy and clinical immunology*, 125(2 Suppl 2), S41-52.
- Schübeler, D., MacAlpine, D. M., Scalzo, D., Wirbelauer, C., Kooperberg, C., van Leeuwen, F., Gottschling, D. E., O'Neill, L. P., Turner, B. M., Delrow, J., Bell, S. P., & Groudine, M. (2004). The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote. *Genes & development*, 18(11), 1263-1271.
- Schulze, J. M., Wang, A. Y., & Kobor, M. S. (2010). Reading chromatin: insights from yeast into YEATS domain structure and function. *Epigenetics*, 5(7), 573-577.
- Searle, N. E., Torres-Machorro, A. L., & Pillus, L. (2017). Chromatin Regulation by the NuA4 Acetyltransferase Complex Is Mediated by Essential Interactions Between Enhancer of Polycomb (Epl1) and Esa1. *Genetics*, 205(3), 1125-1137.
- Shen, S., Arhin, G. K., Ullu, E., & Tschudi, C. (2001). In vivo epitope tagging of *Trypanosoma brucei* genes using a one step PCR-based strategy. *Mol Biochem Parasitol*, 113(1), 171-173.
- Shen, X., Ranallo, R., Choi, E., & Wu, C. (2003). Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Mol Cell*, 12(1), 147-155.
- Shi, H., Tschudi, C., & Ullu, E. (2007). Depletion of newly synthesized Argonaute1 impairs the RNAi response in *Trypanosoma brucei*. *RNA*, 13(7), 1132-1139.
- Shia, W. J., Li, B., & Workman, J. L. (2006). SAS-mediated acetylation of histone H4 Lys 16 is required for H2A.Z incorporation at subtelomeric regions in *Saccharomyces cerevisiae*. *Genes Dev*, 20(18), 2507-2512.
- Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annual review of biochemistry*, 81, 65-95.
- Short, E. E., Caminade, C., & Thomas, B. N. (2017). Climate Change Contribution to the Emergence or Re-Emergence of Parasitic Diseases. *Infectious diseases*, 10, 1178633617732296.
- Siegel, T. N., Gunasekera, K., Cross, G. A., & Ochsenreiter, T. (2011). Gene expression in *Trypanosoma brucei*: lessons from high-throughput RNA sequencing. *Trends Parasitol*, 27(10), 434-441.
- Siegel, T. N., Hekstra, D. R., Kemp, L. E., Figueiredo, L. M., Lowell, J. E., Fenyó, D., Wang, X., Dewell, S., & Cross, G. A. (2009). Four histone variants mark the boundaries of polycistronic transcription units in *Trypanosoma brucei*. *Genes Dev*, 23(9), 1063-1076.
- Siegel, T. N., Kawahara, T., Degrasse, J. A., Janzen, C. J., Horn, D., & Cross, G. A. (2008). Acetylation of histone H4K4 is cell cycle regulated and mediated by HAT3 in *Trypanosoma brucei*. *Mol Microbiol*, 67(4), 762-771.
- Smith, C. A., Bates, P., Rivera-Gonzalez, R., Gu, B., & DeLuca, N. A. (1993). ICP4, the major transcriptional regulatory protein of herpes simplex virus type 1, forms a tripartite complex with TATA-binding protein and TFIIIB. *Journal of virology*, 67(8), 4676-4687.
- Stadler, J., & Richly, H. (2017). Regulation of DNA Repair Mechanisms: How the Chromatin Environment Regulates the DNA Damage Response. *Int J Mol Sci*, 18(8).
- Staneva, D. P., Carloni, R., Auchynnika, T., Tong, P., Arulanandam, J. A., Rappsilber, J., Matthews, K. R., & Allshire, R. (2021). A systematic analysis of *Trypanosoma brucei* chromatin factors identifies novel protein interaction networks associated with sites of transcription initiation and termination. *bioRxiv*, 2021.2002.2009.430399.
- Stanne, T., Narayanan, M. S., 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(45), 26954-26967.
- Stanne, T. M., Kushwaha, M., Wand, M., Taylor, J. E., & 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(7), 964-976.
- Stark, R., & Brown, G. (2011). DiffBind differential binding analysis of ChIP-Seq peak data. *In R package version*, 100.
- Stayton, P. S., Freitag, S., Klumb, L. A., Chilkoti, A., Chu, V., Penzotti, J. E., To, R., Hyre, D., Le Trong, I., Lybrand, T. P., & Stenkamp, R. E. (1999). Streptavidin-biotin binding energetics. *Biomolecular engineering*, 16(1-4), 39-44.

## 5. Publication List

- Stogios, P. J., Downs, G. S., Jauhal, J. J. S., Nandra, S. K., & Privé, G. G. (2005). Sequence and structural analysis of BTB domain proteins. *Genome biology*, 6(10), R82.
- Tachibana, M., Sugimoto, K., Fukushima, T., & Shinkai, Y. (2001). Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J Biol Chem*, 276(27), 25309-25317.
- Tapia-Alveal, C., Lin, S.-J., Yeoh, A., Jabado, O. J., & O'Connell, M. J. (2014). H2A.Z-dependent regulation of cohesin dynamics on chromosome arms. *Mol Cell Biol*, 34(11), 2092-2104.
- Team, R. D. C. (2014). R: A language and environment for statistical computing. In.
- Thambirajah, A. A., Dryhurst, D., Ishibashi, T., Li, A., Maffey, A. H., & Ausió, J. (2006). H2A.Z stabilizes chromatin in a way that is dependent on core histone acetylation. *J Biol Chem*, 281(29), 20036-20044.
- Tosi, A., Haas, C., Herzog, F., Gilmozzi, A., Berninghausen, O., Ungewickell, C., Gerhold, C. B., Lakomek, K., Aebersold, R., Beckmann, R., & Hopfner, K. P. (2013). Structure and subunit topology of the INO80 chromatin remodeler and its nucleosome complex. *Cell*, 154(6), 1207-1219.
- Tremethick, D. J. (2007). Higher-order structures of chromatin: the elusive 30 nm fiber. *Cell*, 128(4), 651-654.
- Tsukuda, T., Fleming, A. B., Nickoloff, J. A., & Osley, M. A. (2005). Chromatin remodelling at a DNA double-strand break site in *Saccharomyces cerevisiae*. *Nature*, 438(7066), 379-383.
- Udugama, M., Sabri, A., & Bartholomew, B. (2011). The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Mol Cell Biol*, 31(4), 662-673.
- van Attikum, H., Fritsch, O., & Gasser, S. M. (2007). Distinct roles for SWR1 and INO80 chromatin remodeling complexes at chromosomal double-strand breaks. *EMBO J*, 26(18), 4113-4125.
- Vaute, O., Nicolas, E., Vandel, L., & Trouche, D. (2002). Functional and physical interaction between the histone methyl transferase Suv39H1 and histone deacetylases. *Nucleic Acids Res*, 30(2), 475-481.
- Vélez-Ramírez, D. E., Shimogawa, M. M., Ray, S. S., Lopez, A., Rayatpisheh, S., Langousis, G., Gallagher-Jones, M., Dean, S., Wohlschlegel, J. A., & Hill, K. L. (2021). APEX2 Proximity Proteomics Resolves Flagellum Subdomains and Identifies Flagellum Tip-Specific Proteins in *Trypanosoma brucei*. *mSphere*, 6(1).
- Vellmer, T., Hartleb, L., Sola, A. F., Kramer, S., Meyer-Natus, E., Butter, F., & Janzen, C. J. (2021). A novel SNF2 ATPase complex in *Trypanosoma brucei* with a role in H2A.Z-mediated chromatin remodelling. *bioRxiv*, 2021.2004.2006.438560.
- Wan, Y., Saleem, R. A., Ratushny, A. V., Roda, O., Smith, J. J., Lin, C. H., Chiang, J. H., & Aitchison, J. D. (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(9), 2346-2358.
- Warfield, L., Ramachandran, S., Baptista, T., Devys, D., Tora, L., & Hahn, S. (2017). Transcription of Nearly All Yeast RNA Polymerase II-Transcribed Genes Is Dependent on Transcription Factor TFIID. *Mol Cell*, 68(1), 118-129.e115.
- Watanabe, S., Radman-Livaja, M., Rando, O. J., & Peterson, C. L. (2013). A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme. *Science*, 340(6129), 195-199.
- Watanabe, S., Resch, M., Lilyestrom, W., Clark, N., Hansen, J. C., Peterson, C., & Luger, K. (2010). Structural characterization of H3K56Q nucleosomes and nucleosomal arrays. *Biochim Biophys Acta*, 1799(5-6), 480-486.
- Wedel, C., Förstner, K. U., Derr, R., & Siegel, T. N. (2017). GT-rich promoters can drive RNA pol II transcription and deposition of H2A.Z in African trypanosomes. *EMBO J*, 36(17), 2581-2594.
- Wedel, C., & Siegel, T. N. (2017). Genome-wide analysis of chromatin structures in *Trypanosoma brucei* using high-resolution MNase-ChIP-seq. *Exp Parasitol*.
- Weimer, R. M. (2006). Preservation of *C. elegans* tissue via high-pressure freezing and freeze-substitution for ultrastructural analysis and immunocytochemistry. *Methods Mol Biol*, 351, 203-221.
- Weisbrod, S. (1982). Active chromatin. *Nature*, 297(5864), 289-295.
- West, M. H., & Bonner, W. M. (1980). Histone 2A, a heteromorphous family of eight protein species. *Biochemistry*, 19(14), 3238-3245.
- WHO (2021) [https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-\(sleeping-sickness\)](https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness))

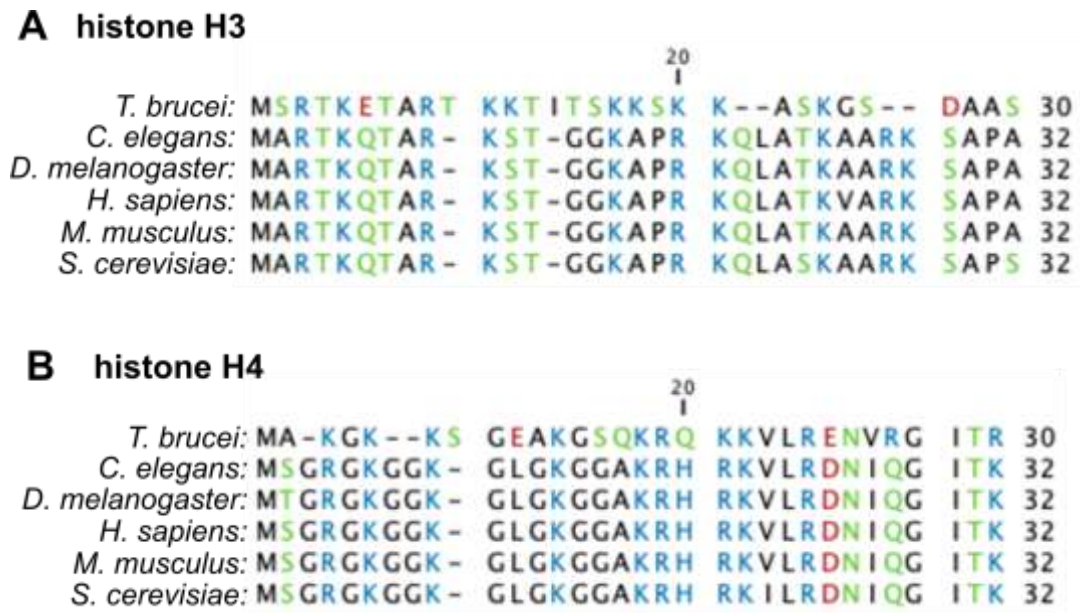
## 5. Publication List

- Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. Springer Publishing Company, Incorporated.
- Widom, J. (1998). Structure, dynamics, and function of chromatin in vitro. *Annu Rev Biophys Biomol Struct*, 27, 285-327.
- Willhoft, O., & Wigley, D. B. (2020). INO80 and SWR1 complexes: the non-identical twins of chromatin remodelling. *Curr Opin Struct Biol*, 61, 50-58.
- Wirtz, E., Leal, S., Ochatt, C., & Cross, G. A. (1999). A tightly regulated inducible expression system for conditional gene knock-outs and dominant-negative genetics in *Trypanosoma brucei*. *Mol Biochem Parasitol*, 99(1), 89-101.
- Wong, M. M., Cox, L. K., & Chrivia, J. C. (2007). The chromatin remodeling protein, SRCAP, is critical for deposition of the histone variant H2A.Z at promoters. *J Biol Chem*, 282(36), 26132-26139.
- Wright, J. R., Siegel, T. N., & Cross, G. A. (2010). Histone H3 trimethylated at lysine 4 is enriched at probable transcription start sites in *Trypanosoma brucei*. *Mol Biochem Parasitol*, 172(2), 141-144.
- Wu, W. H., Alami, S., Luk, E., Wu, C. H., 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(12), 1064-1071.
- Wu, W. H., Wu, C. H., Ladurner, A., Mizuguchi, G., Wei, D., Xiao, H., Luk, E., Ranjan, A., & Wu, C. (2009). N terminus of Swr1 binds to histone H2AZ and provides a platform for subunit assembly in the chromatin remodeling complex. *J Biol Chem*, 284(10), 6200-6207.
- Wutz, A. (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. *Nature Reviews Genetics*, 12(8), 542-553.
- Wysocka, J., Swigut, T., Milne, T. A., Dou, Y., Zhang, X., Burlingame, A. L., Roeder, R. G., Brivanlou, A. H., & Allis, C. D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. *Cell*, 121(6), 859-872.
- Xu, F., Zhang, K., & Grunstein, M. (2005). Acetylation in histone H3 globular domain regulates gene expression in yeast. *Cell*, 121(3), 375-385.
- Xu, Y., Ayrapetov, M. K., Xu, C., Gursoy-Yuzugullu, O., Hu, Y., & Price, B. D. (2012). Histone H2A.Z controls a critical chromatin remodeling step required for DNA double-strand break repair. *Mol Cell*, 48(5), 723-733.
- Yang, P. K., Rotondo, G., Porras, T., Legrain, P., & Chanfreau, G. (2002). The Shq1p.Naf1p complex is required for box H/ACA small nucleolar ribonucleoprotein particle biogenesis. *J Biol Chem*, 277(47), 45235-45242.
- Yu, G., Wang, L. G., & He, Q. Y. (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*, 31(14), 2382-2383.
- Yun, M., Wu, J., Workman, J. L., & Li, B. (2011). Readers of histone modifications. *Cell Res*, 21(4), 564-578.
- Zeng, L., & Zhou, M. M. (2002). Bromodomain: an acetyl-lysine binding domain. *FEBS Lett*, 513(1), 124-128.
- Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., Nusbaum, C., Myers, R. M., Brown, M., Li, W., & Liu, X. S. (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol*, 9(9), R137.
- Zhou, B. O., Wang, S. S., Xu, L. X., Meng, F. L., Xuan, Y. J., Duan, Y. M., Wang, J. Y., Hu, H., Dong, X., Ding, J., & Zhou, J. Q. (2010). SWR1 complex poises heterochromatin boundaries for antisilencing activity propagation. *Mol Cell Biol*, 30(10), 2391-2400.



## 6. Appendix

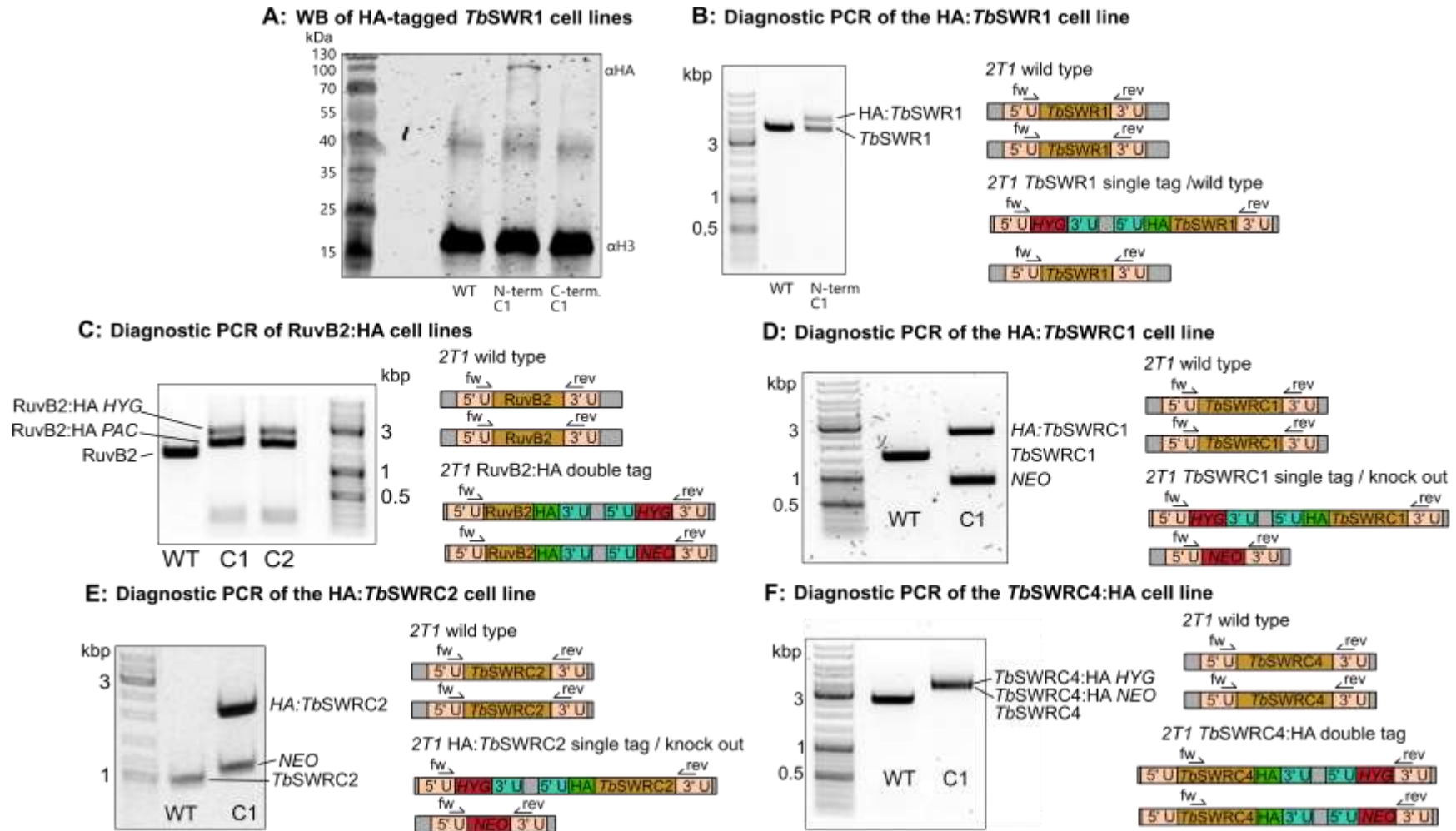
### 6.1 Supplementary figures



**Figure S1: Alignment of the histones H3 and H4 sequences**

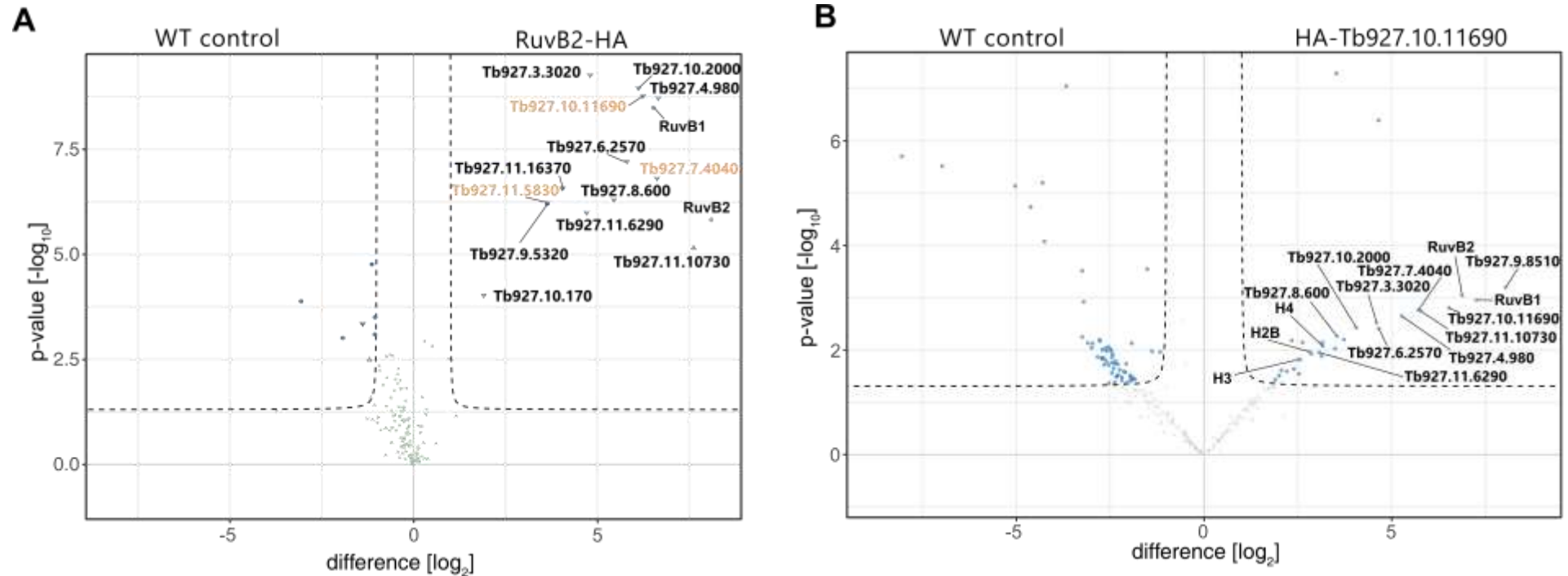
The first 30 amino acids of *T. brucei* histone H3 and histone H4 aligned with H3 and H4 of *Caenorhabditis elegans*, *Drosophila melanogaster*, *Homo sapiens*, *Mus musculus* and *Saccharomyces cerevisiae*. Amino acids labelled in blue possess a positive charge, amino acids labelled in red possess a negative charge, amino acids labelled in green possess a polar uncharged side chain, amino acids labelled in black possess hydrophobic side chains.

6. Appendix



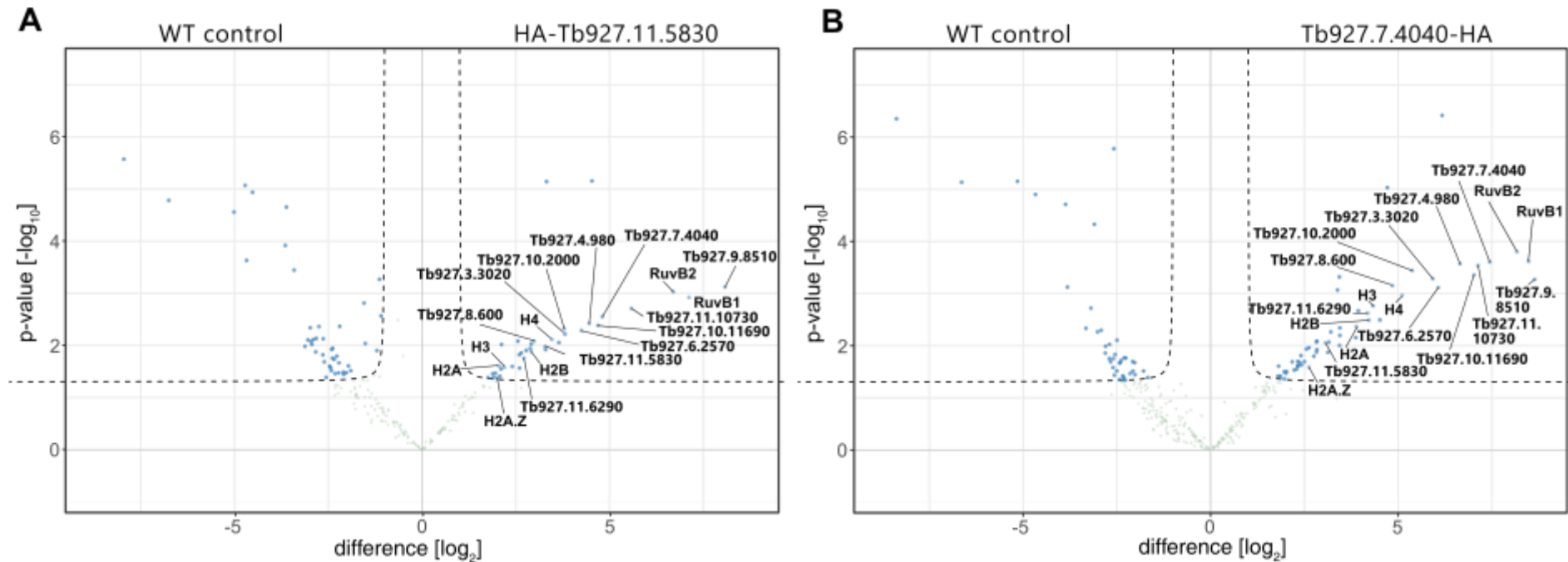
**Figure S2: Control of the cell lines used for the co-IP experiments to identify the *TbSWR1* complex**

(A) Western Blot analysis of cell lines with single HA-tagged *TbSWR1*, (B) Diagnostic PCR of the single HA-tagged *TbSWR1* cell lines, (C) Diagnostic PCR of the 427 PCF RuvB2:HA cell line, (D) Diagnostic PCR of the 2T1 HA:*TbSWRC1* cell line, (E) Diagnostic PCR of the 2T1 HA:*TbSWRC2* cell line, (F) Diagnostic PCR of the 2T1 *TbSWRC4*:HA cell line; fw indicates the relative binding site of the forward primer, rev indicates the relative binding site of the reverse primer, 5'U indicates the 5' UTR and 3'U indicates the 3' UTR, ORF represents the open reading frame of the protein, *NEO* indicates the location of the neomycin resistance cassette, *HYG* indicates the location of the neomycin resistance cassette, HA represents the location of the HA-tag. Vellmer *et al.*, 2021 in revision



**Figure S3: Volcano plots of the co-IPs that identified the *Tb*SWR1 complex part 1**

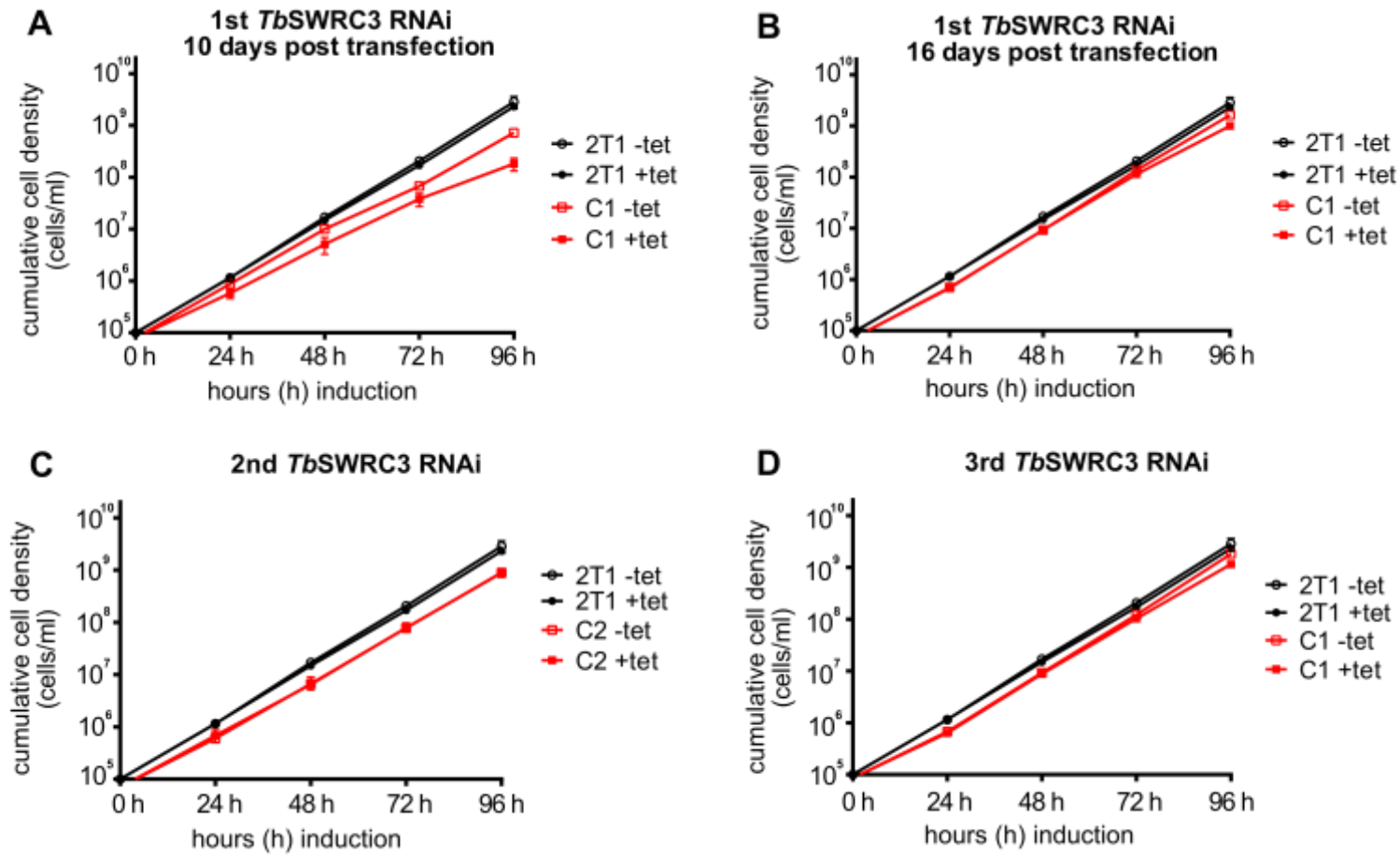
Volcano blot of co-purified proteins after Volcano blot of co-purified proteins after **(A)** WT control vs. HA-RuvB2 (Tb927.4.2000), **(B)** WT control vs. HA-*Tb*SWRC1 (Tb927.10.11690) co-IPs obtained by MS analysis of four biological replicates. Green dots represent purified proteins with a p-value of  $> 0.01$  or with a fold-enrichment of  $= / < 1$ . Blue dots represent the proteins with a p-value  $= / < 0.01$  or with a fold-enrichment of  $> 1$  that could be identified in at least three of the four co-IPs. Proteins labelled in orange indicate the three proteins that were selected for the reciprocal co-IP approach. Vellmer *et al.*, 2021 in revision



**Figure S4: Volcano plots of the co-IPs that identified the *TbSWR1* complex part 2**

Volcano blot of co-purified proteins after **(A)** WT control vs. HA-*TbSWRC2* (Tb927.11.5830) and **(B)** WT control vs. *TbSWRC4*-HA (Tb927.7.4040) co-IPs obtained by MS analysis of four biological replicates. Green dots represent purified proteins with a p-value of > 0.01 or with a fold-enrichment of = / < 1. Blue dots represent the proteins with a p-value = / < 0.01 or with a fold-enrichment of > 1 that could be identified in at least three of the four co-IPs. Vellmer *et al.*, 2021 in revision

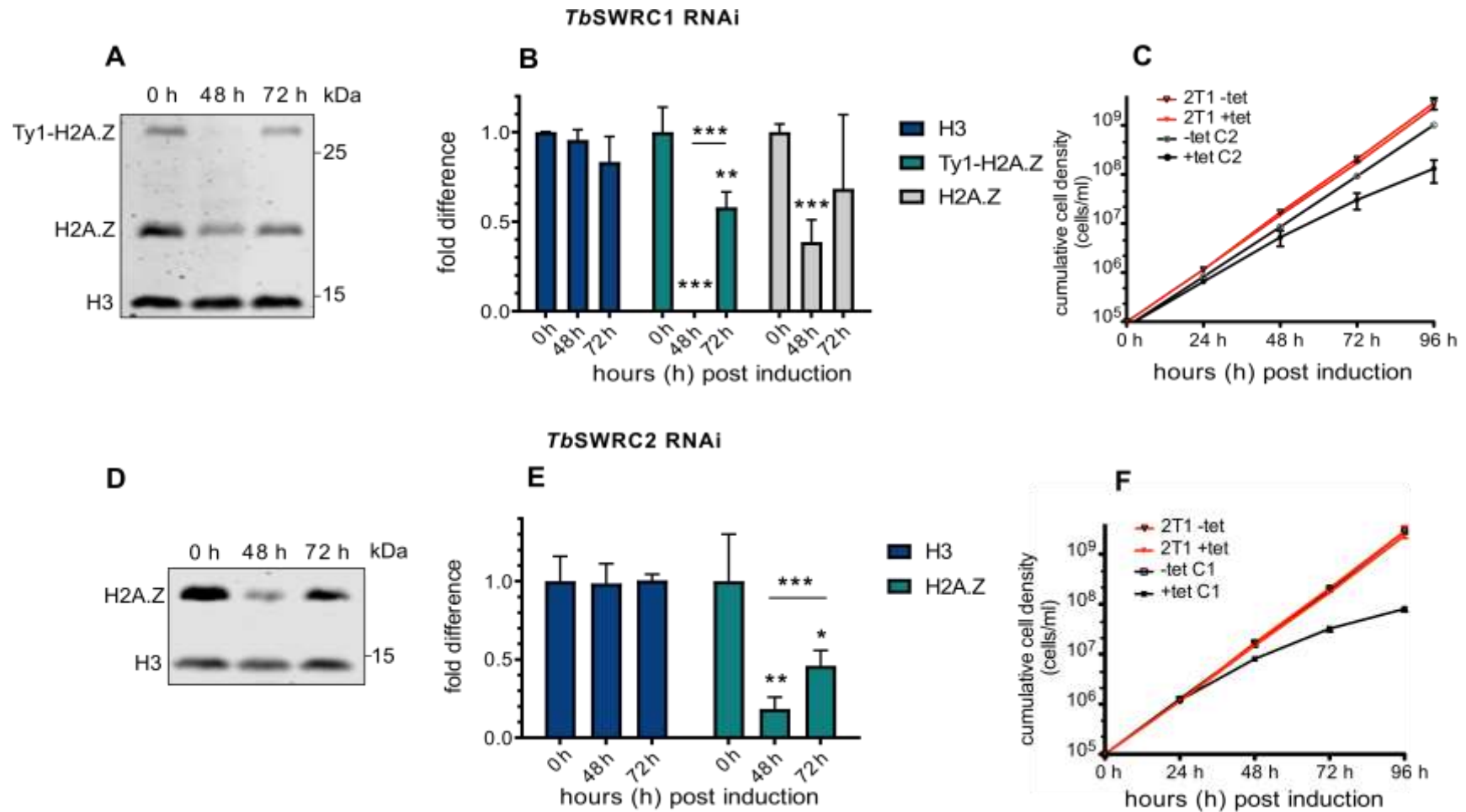
6. Appendix



**Figure S5: Growth curves of *TbSWRC3* RNAi cell lines**

Growth of parasites was monitored for 96 hours after RNAi-mediated depletion of *TbSWRC3*. (A) growth curve of the cell line transfected with the first RNAi target sequence 10 days post transfection, (B) growth curve of the cell line transfected with the first RNAi target sequence 16 days post transfection, (C) growth curve of the cell line transfected with the second RNAi target sequence (10 days post transfection), (D) growth curve of the cell line transfected with the third RNAi target sequence (10 days post transfection). Parental 2T1 cell line was measured for 96h as a control. (N=3 for all depicted experiments).

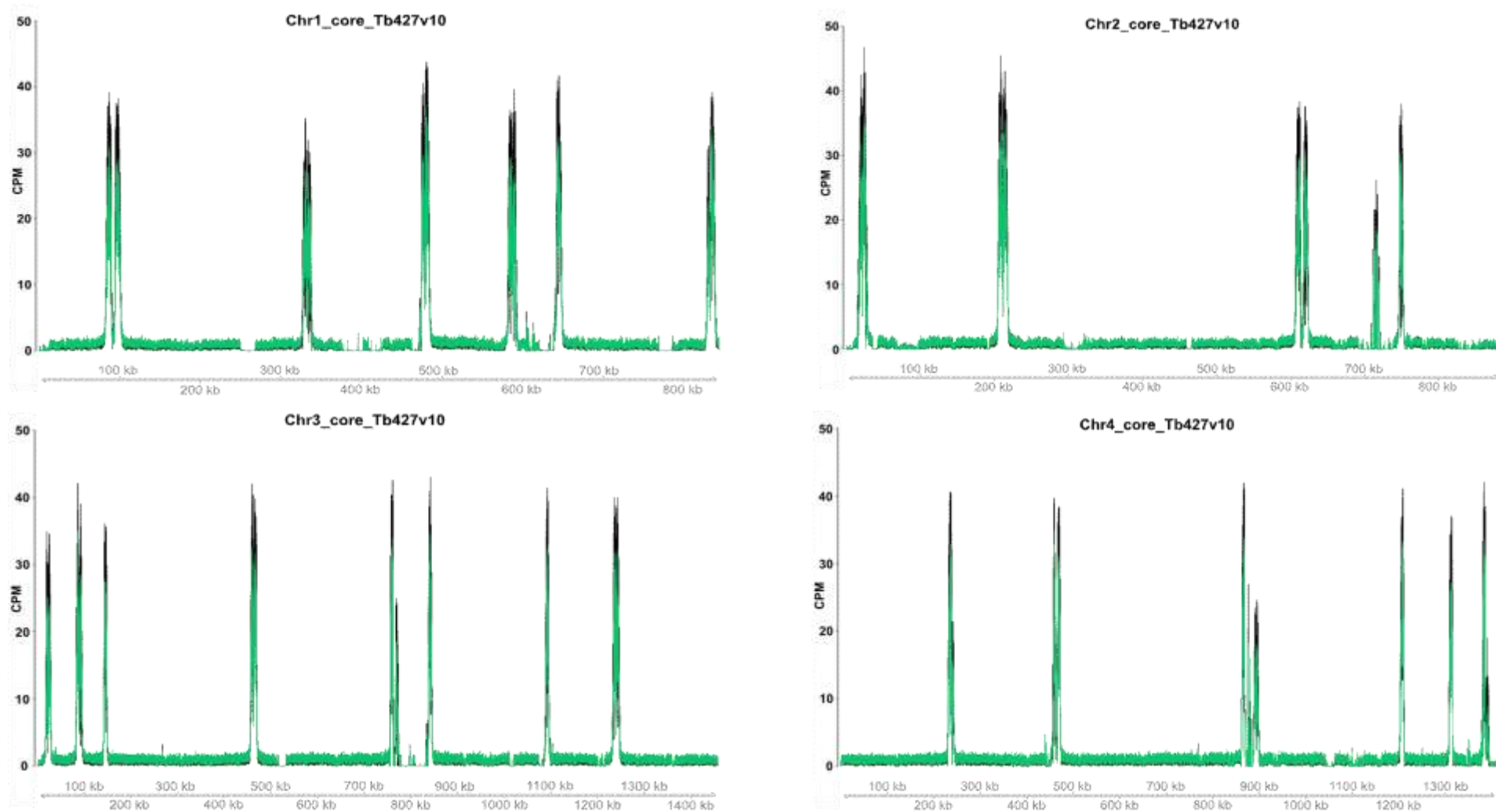
6. Appendix



**Figure S6: Depletion of *TbSWRC1* and *TbSWRC2* reduce the amount of chromatin-associated H2A.Z**

Exemplary Western Blot analysis of the nuclear fraction with antibodies against histone H3 and the histone variant H2A.Z of *TbSWRC1*- (**A**) and *TbSWRC2*- (**D**) depleted cells. An equal amount of cell equivalent was loaded for each timepoint. (**B+E**) The development of chromatin-associated H3 (dark blue), Ty1-H2A.Z (green-blue) and H2A.Z (grey) in course of *TbSWRC1* (**B**) and *TbSWRC2* (**E**) depletion is plotted (N=3). Growth of parasites was monitored for 96 hours after RNAi-mediated depletion of *TbSWRC1* (**C**) and *TbSWRC2* (**F**) using tetracycline (tet). Growth of tet induced and non-induced parental 2T1 cells was measured for 96h and acts as a reference (N=3 for all depicted experiments; \*\*\* = p-value <0.001; \*\* = p-value 0.001-0.01; \* = p-value 0.01-0.05). Vellmer *et al.*, 2021 in revision

## 6. Appendix

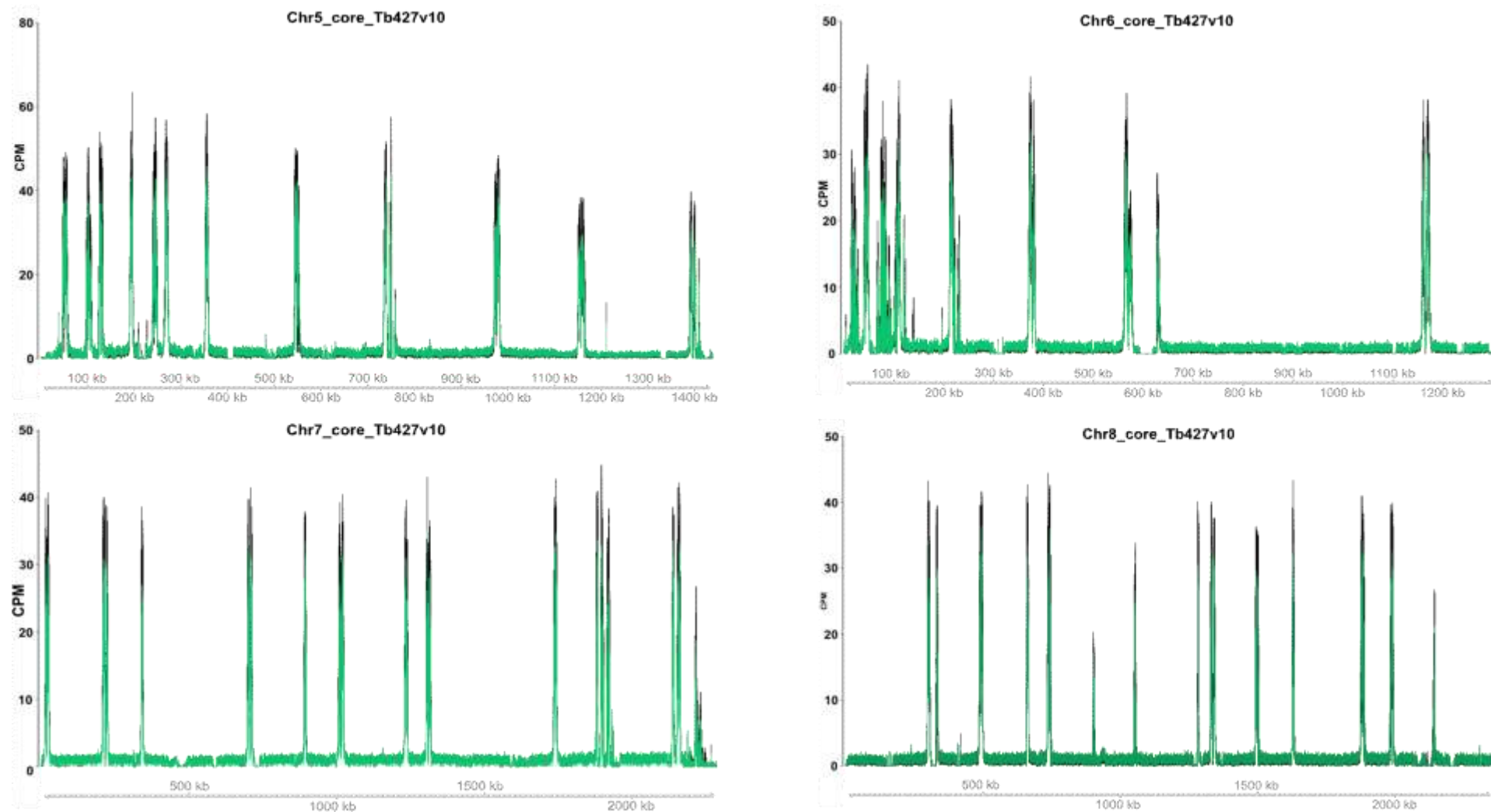


**Figure S7: ChIP-seq results of *TbSWR1*-depleted cells, chromosomes 1-4**

ChIP-Seq analysis of distribution of Ty1-tagged H2A.Z before (black) and after (green) RNAi-mediated depletion of *TbSWR1* (48h post induction) revealed a reduction of chromatin-associated Ty1-H2A.Z. Depicted is a representative region of chromosome 1 - 4. Data (n=2) were normalised to the total number of reads and plotted as counts per million reads (CPM). Vellmer *et al.*, 2021 in revision



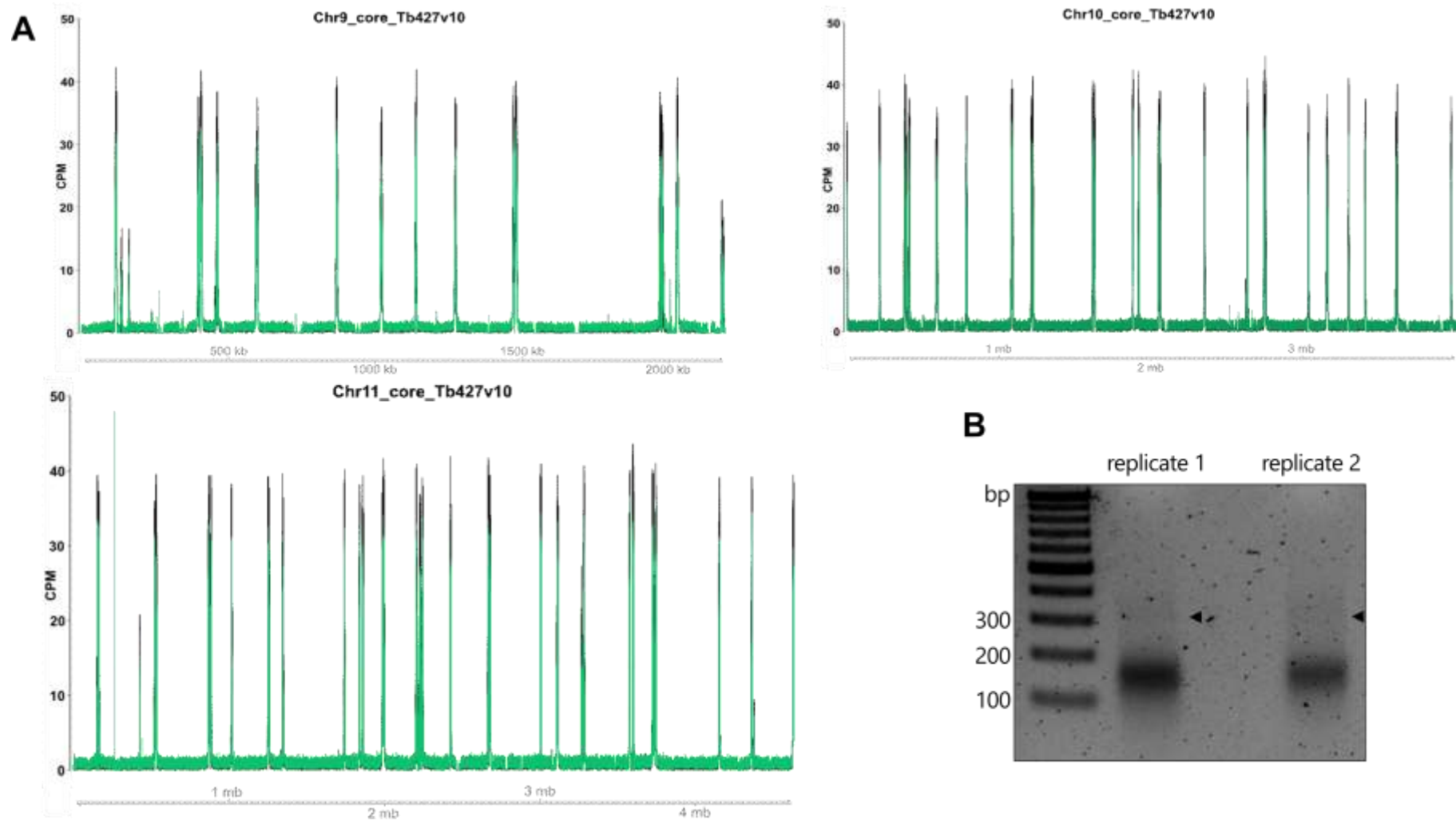
## 6. Appendix



**Figure S8: ChIP-seq results of *TbSWR1*-depleted cells, chromosomes 5-8**

ChIP-Seq analysis of distribution of Ty1-tagged H2A.Z before (black) and after (green) RNAi-mediated depletion of *TbSWR1* (48h post induction) revealed a reduction of chromatin-associated Ty1-H2A.Z. Depicted is a representative region of chromosome 5 - 8. Data (n=2) were normalised to the total number of reads and plotted as counts per million reads (CPM). Vellmer *et al.*, 2021 in revision

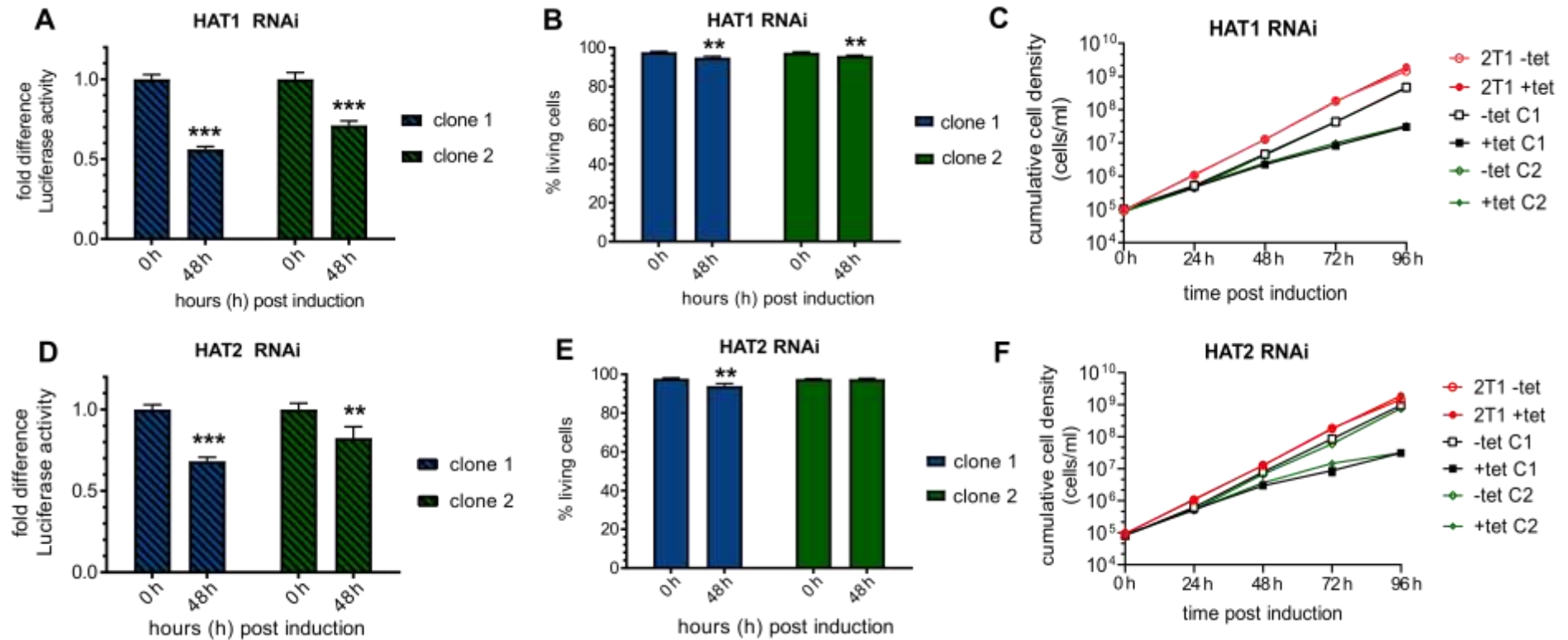




**Figure S9: ChIP-seq results of *TbSWR1*-depleted cells, chromosomes 9-11**

**(A)** ChIP-Seq analysis of distribution of Ty1-tagged H2A.Z before (black) and after (green) RNAi-mediated depletion of *TbSWR1* (48h post induction) revealed a reduction of chromatin-associated Ty1-H2A.Z. Depicted is a representative region of chromosome 9 - 11. Data (n=2) were normalised to the total number of reads and plotted as counts per million reads (CPM). **(B)** Agarose gel electrophoresis picture of the MNase-digested ChIP-seq input material of both replicates, 150ng DNA were separated on each lane. Black arrows indicate a faint band running above 300 bp. Vellmer *et al.*, 2021 in revision

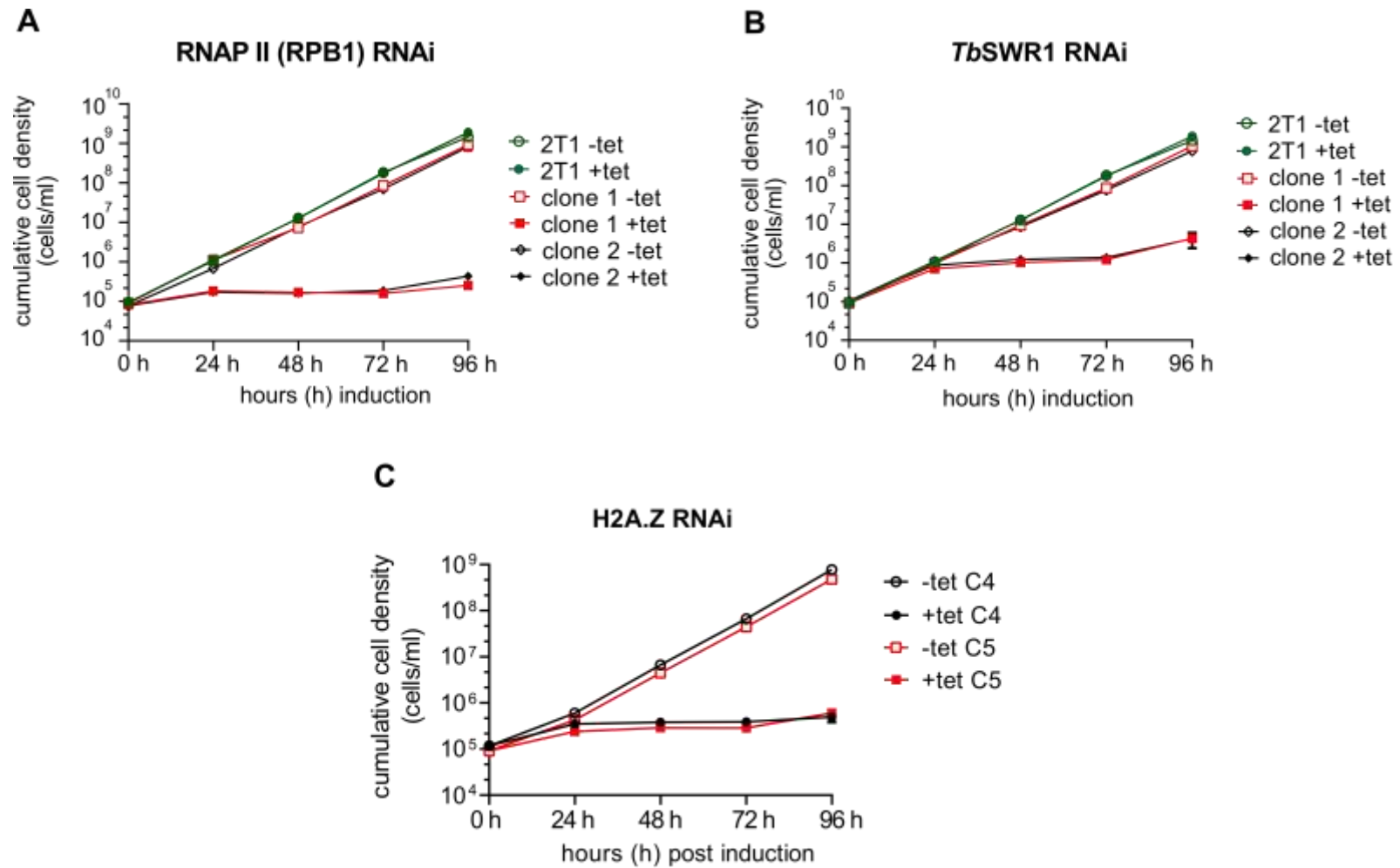
6. Appendix



**Figure S10: Depletion of the histone acetyltransferase HAT1 and HAT2 caused a decrease of reporter luciferase activity within a PTU**

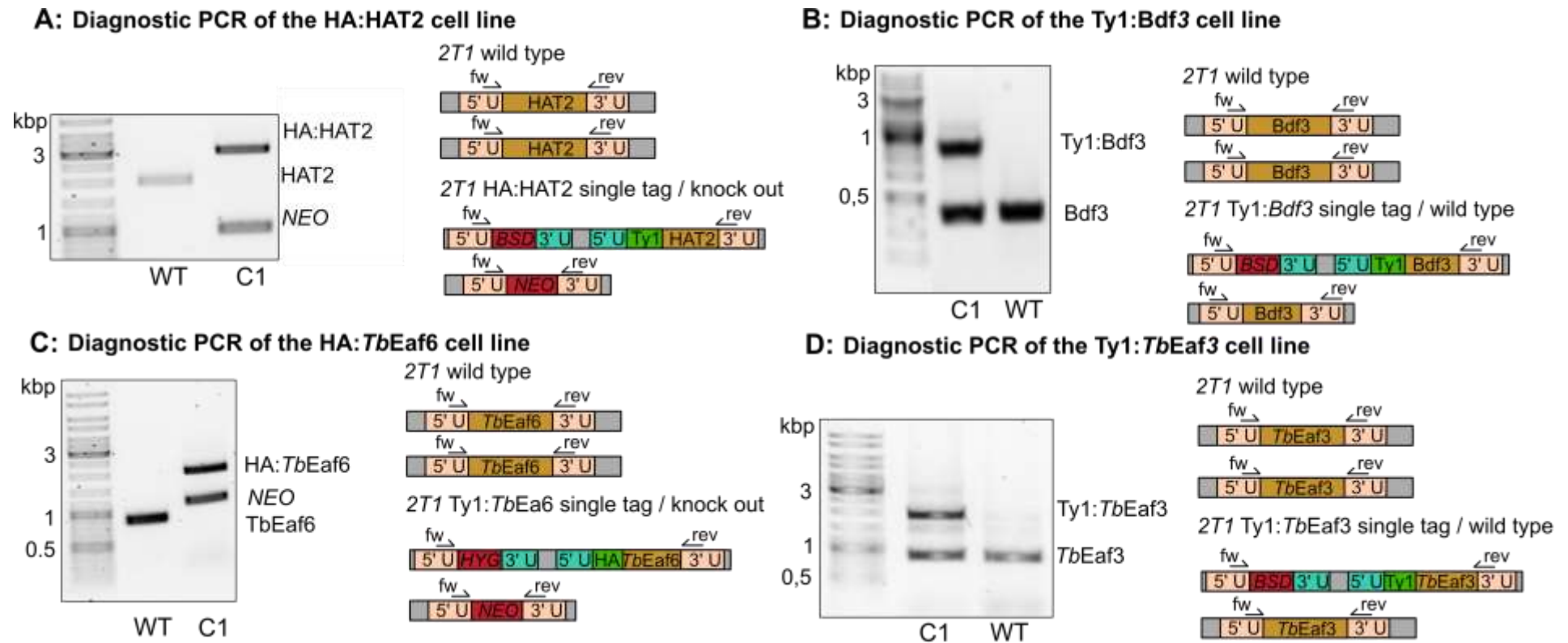
A single luciferase reporter construct was integrated into the tubulin array of an HAT1 (Tb927.7.4560) and an HAT2 (Tb927.11.11530) RNAi cell line. Samples for the luciferase assay were normalised to cell numbers. **(A+D)** Luciferase activity was monitored for 48 h after induction of RNAi in two independent clones. Values of non-induced cells were set to 1. **(B+E)** Live/dead staining of each RNAi cell line was performed in triplicates at the same time points. **(C+F)** Growth of parasites was monitored for 96 hours after RNAi-mediated depletion of HAT1 **(C)** and HAT2 **(F)** using tetracycline (tet) induction. Growth of the parental 2T1 cell line was measured for 96h as a control. (N=3 for all depicted experiments; \*\*\* = p-value <0.001; \*\* = p-value 0.001-0.01; \* = p-value 0.01-0.05). Vellmer *et al.*, 2021 in revision

6. Appendix



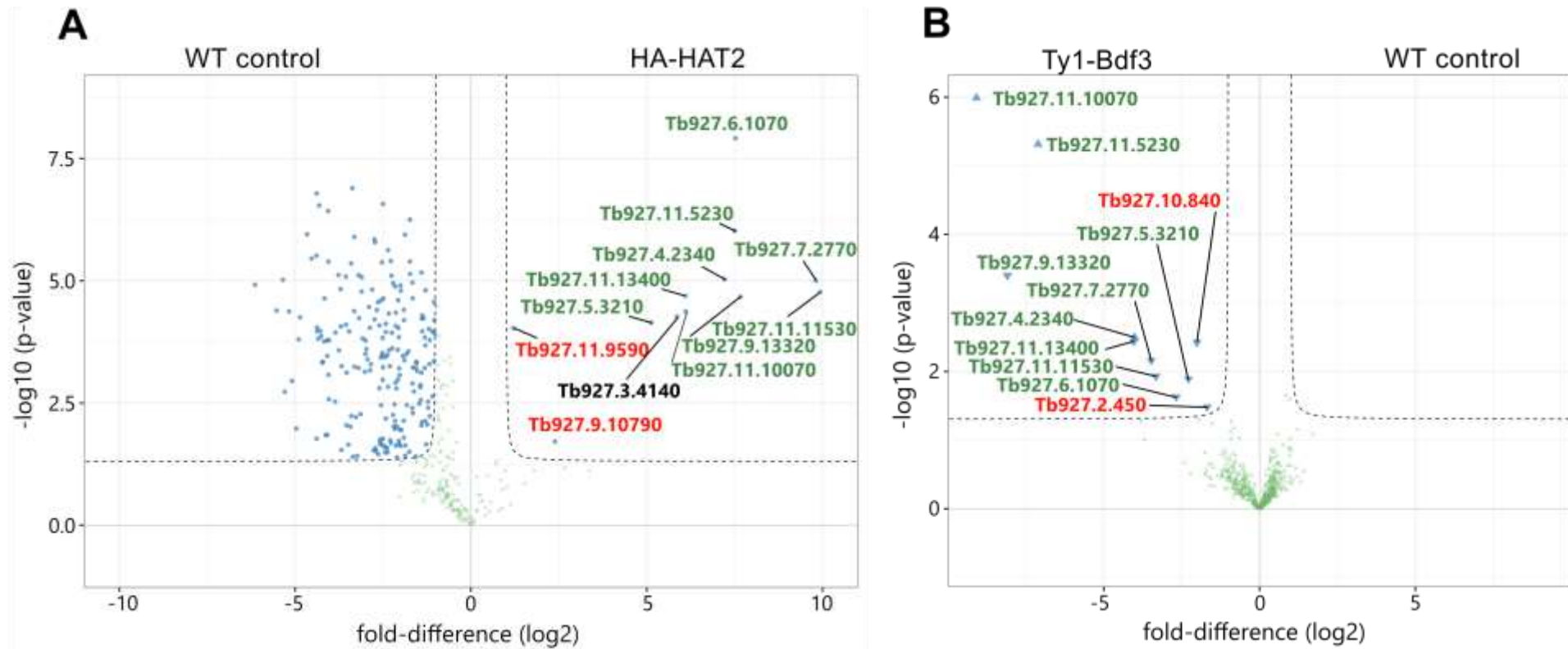
**Figure S11: Growth curves of RNAi cell lines used for the luciferase assay**

Growth of parasites was monitored for 96 hours after RNAi-mediated depletion of **(A)** RPB1, **(B)** *TbSWR1* and **(C)** H2A.Z using tetracycline (tet) induction. Growth of the parental 2T1 cell line was measured for 96h as a control. (N=3 for all depicted experiments. Vellmer *et al.*, 2021 in revision)



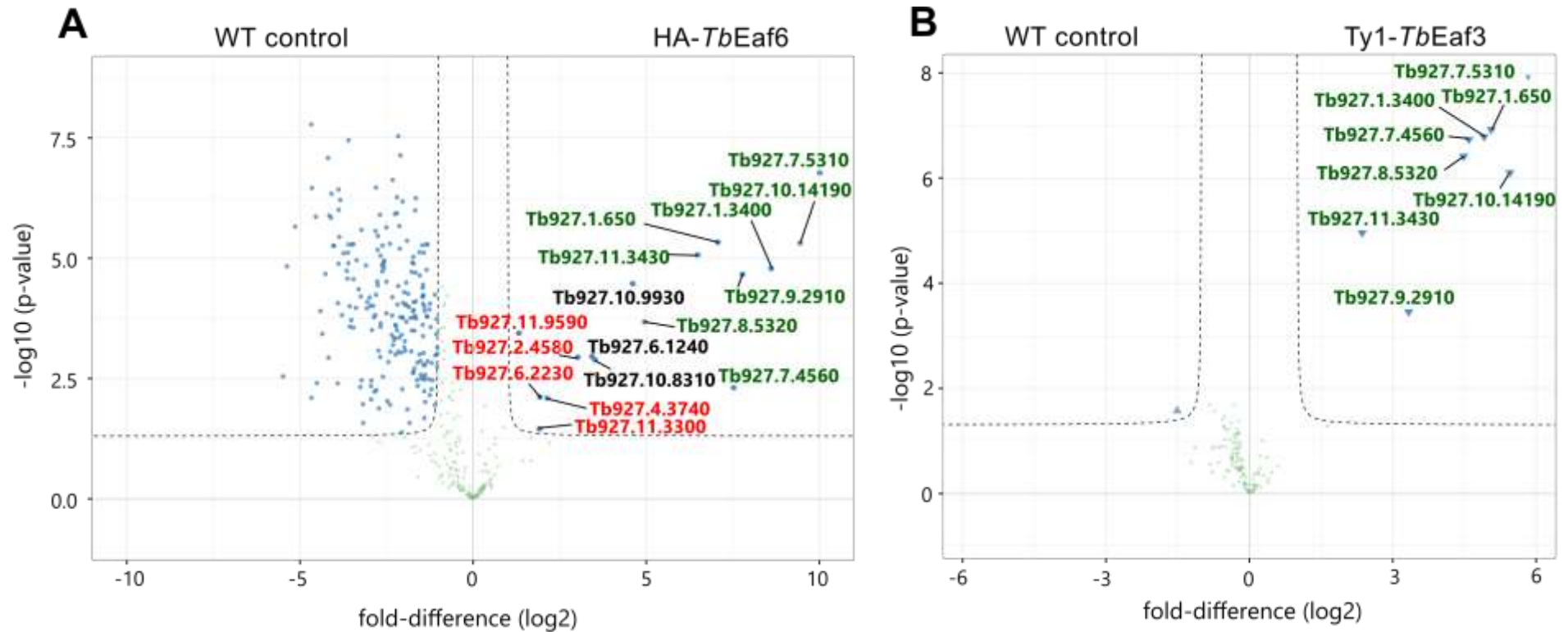
**Figure S12: Evaluation of the cell lines used for the co-IP experiments that identified the HAT1 and HAT2 complex**

(A) Diagnostic PCR of the 2T1 HA:HAT2 cell line, (B) Diagnostic PCR of the SM Ty1:Bdf3 cell line, (C) Diagnostic PCR of the 2T1 HA:TbEaf6 cell line, (D) Diagnostic PCR of the SM Ty1:TbEaf3 cell line; fw indicates the relative binding site of the forward primer, rev indicates the relative binding site of the reverse primer; 5'U indicates the 5' UTR and 3'U indicates the 3' UTR, NEO indicates the location of the neomycin resistance cassette, HYG indicates the location of the neomycin resistance cassette, BSD indicates the location of the blasticidin resistance cassette, HA represents the location of the HA-tag and Ty1 represents the localisation of the Ty1-tag. Vellmer *et al.*, 2021 in revision



**Figure S13:Volcano plots of the co-IPs that identified the HAT2 complex**

Volcano blot of co-purified proteins after **(A)** WT control vs. HA-HAT2 (Tb927.11.11530), **(B)** WT control vs. Ty1-Bdf3 (Tb927.11.10070) co-IPs obtained by MS analysis of four biological replicates. Green dots represent purified proteins with a p-value of  $> 0.01$  or with a fold-enrichment of  $= / < 1$ . Blue dots represent the proteins with a p-value  $= / < 0.01$  or with a fold-enrichment of  $> 1$  that could be identified in at least three of the four co-IPs. Protein IDs in green represent proteins that were identified in both co-IPs, Protein IDs in red represent proteins with a negative NES, Black protein IDs represent proteins that were only identified in the corresponding co-IP. Vellmer *et al.*, 2021 in revision

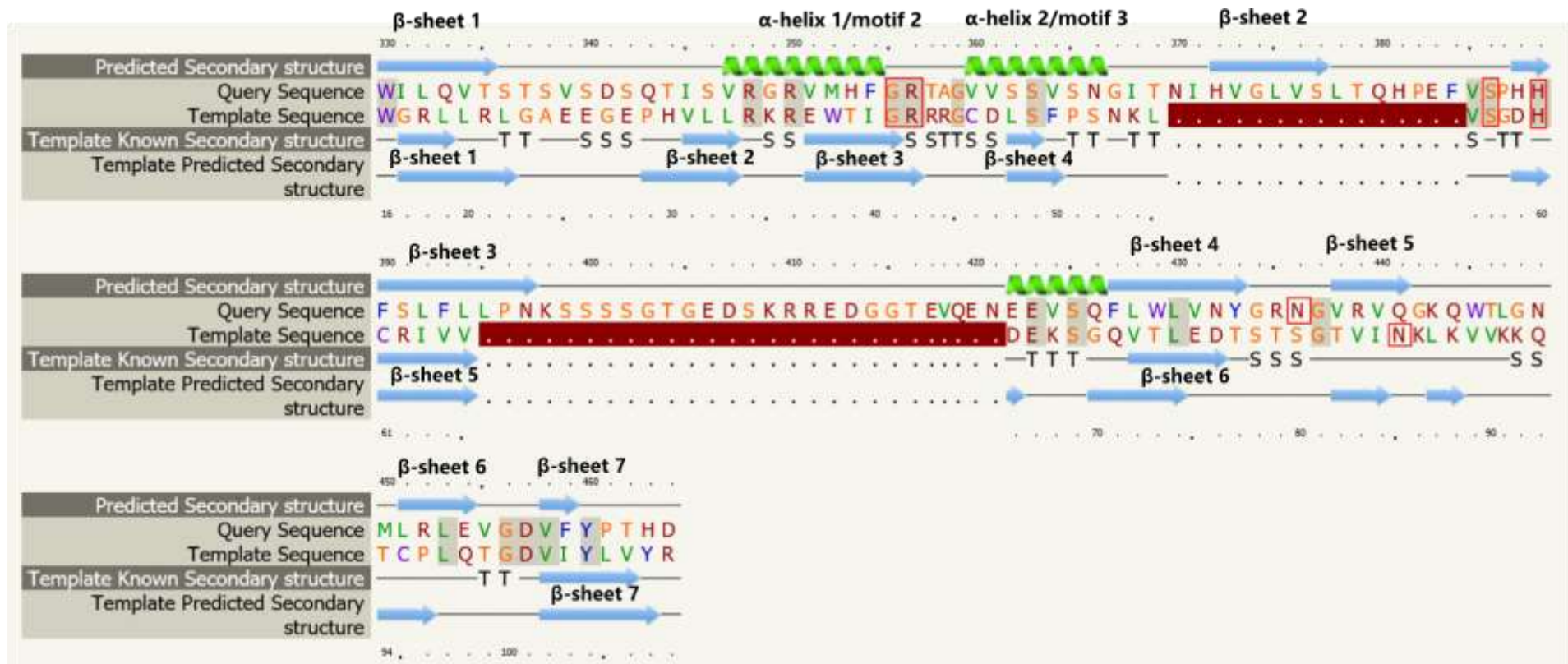


**Figure S14: Volcano plots of the co-IPs that identified the HAT1 complex**

Volcano blot of co-purified proteins after **(A)** WT control vs. HA-*TbEaf6* (Tb927.9.2910), **(B)** WT control vs. HA-*TbEaf3* (Tb927.1.650) co-IPs obtained by MS analysis of four biological replicates. Green dots represent purified proteins with a p-value of  $> 0.01$  or with a fold-enrichment of  $= / < 1$ . Blue dots represent purified proteins with a p-value  $= / < 0.01$  or with a fold-enrichment of  $> 1$ . Protein IDs in green represent proteins that were identified in both co-IPs, Protein IDs in red represent proteins with a negative NES, Black protein IDs represent proteins that were only identified in the corresponding co-IP. Vellmer *et al.*, 2021 in revision

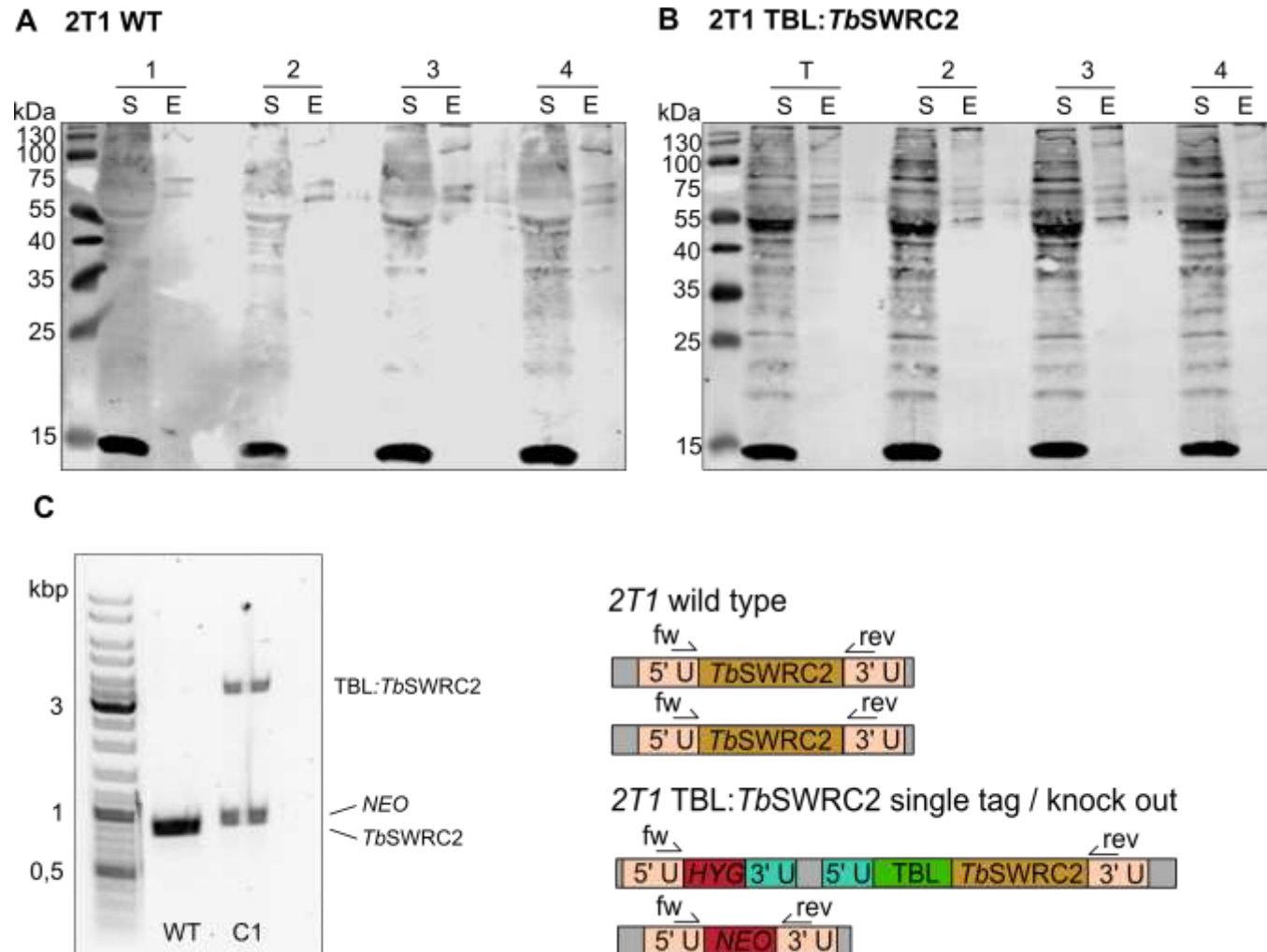


6. Appendix



**Figure S15: Predicted secondary structure of the C-terminal end of Tb927.9.13320 is homologous to an FHA-Domain**

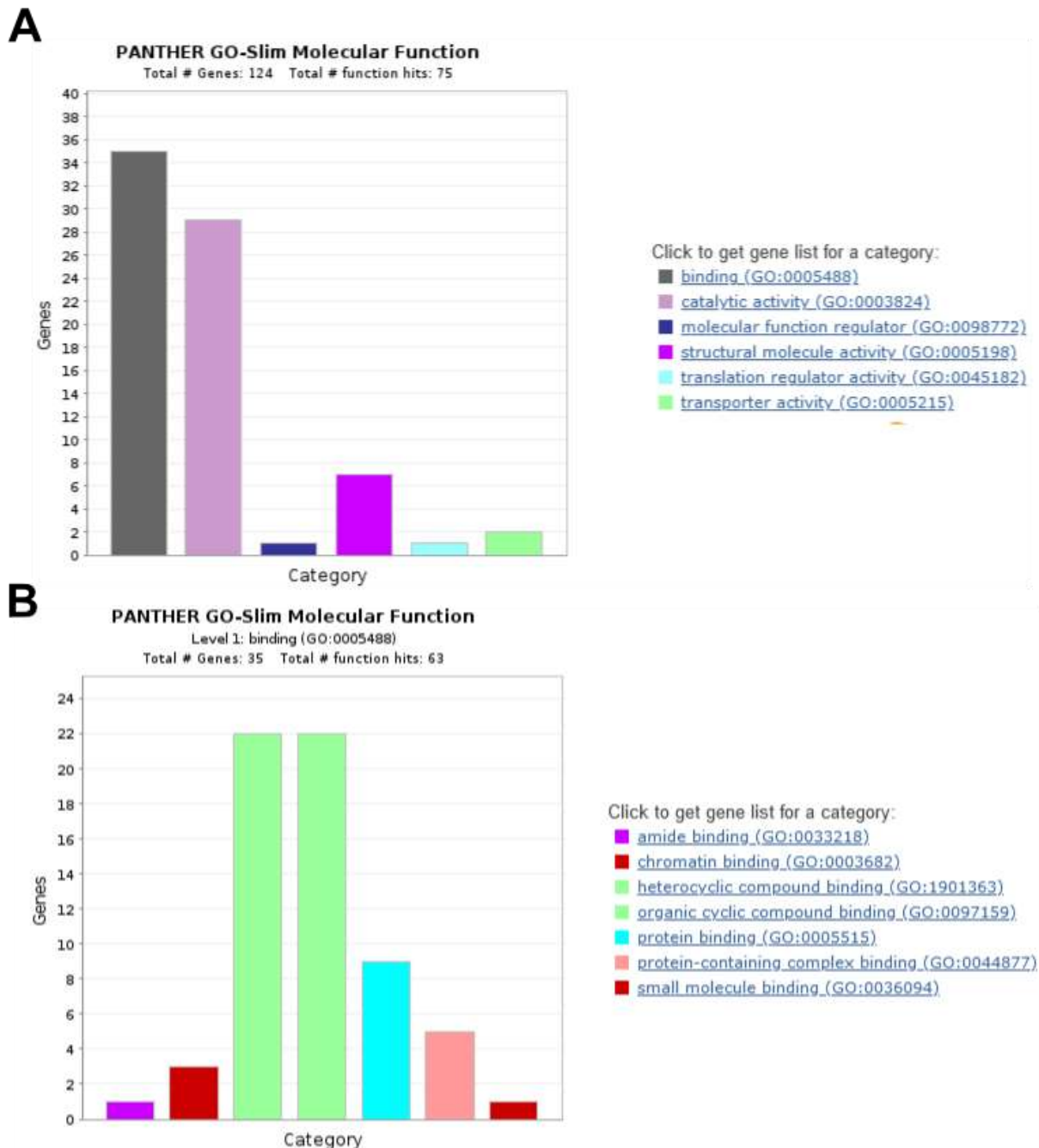
Phyre2 modelling of Tb927.9.13320 could link the C-terminus of the protein to the FHA-domain of the human cell cycle checkpoint protein 'checkpoint with FHA and RING domains' (CHFR; RCSB/PDB ID: 1LGP). In the alignment of Tb927.9.13320 (Query) with CHFR (template) conserved amino acids required for binding of the phosphorylation 11106755 are highlighted with red boxes. Dark-red areas represent sequence gaps in the template. Predicted and known  $\beta$ -sheets are depicted in light-blue. Predicted and known  $\alpha$ -helices are depicted in light-green. The alignment was created by the Phyre2 online modelling platform (Kelley *et al.*, 2015).



**Figure S16: Representative western blots of the *TbSWRC2* BioID and its control**

(A) Protein extracts of 2T1 WT cells were (four replicates) and (B) protein extracts of 2T1 TBL:*TbSWRC2* cells (four replicates) were used for co-IPs with subsequent MS analysis. S (soluble supernatant) and E (eluate: 10-fold compared to the other samples). Blots were probed with IRDye 800CW Streptavidin and an anti-H3 antibody. (C) Diagnostic PCR of one clone of the TBL:*TbSWRC2* cell line that was used for co-IP experiments. *NEO* indicates the location of the neomycin resistance cassette, *HYG* indicates the location of the hygromycin resistance cassette, *BSD* indicates the location of the blasticidin resistance cassette, TBL represents the location of the Turbo Biotinligase.

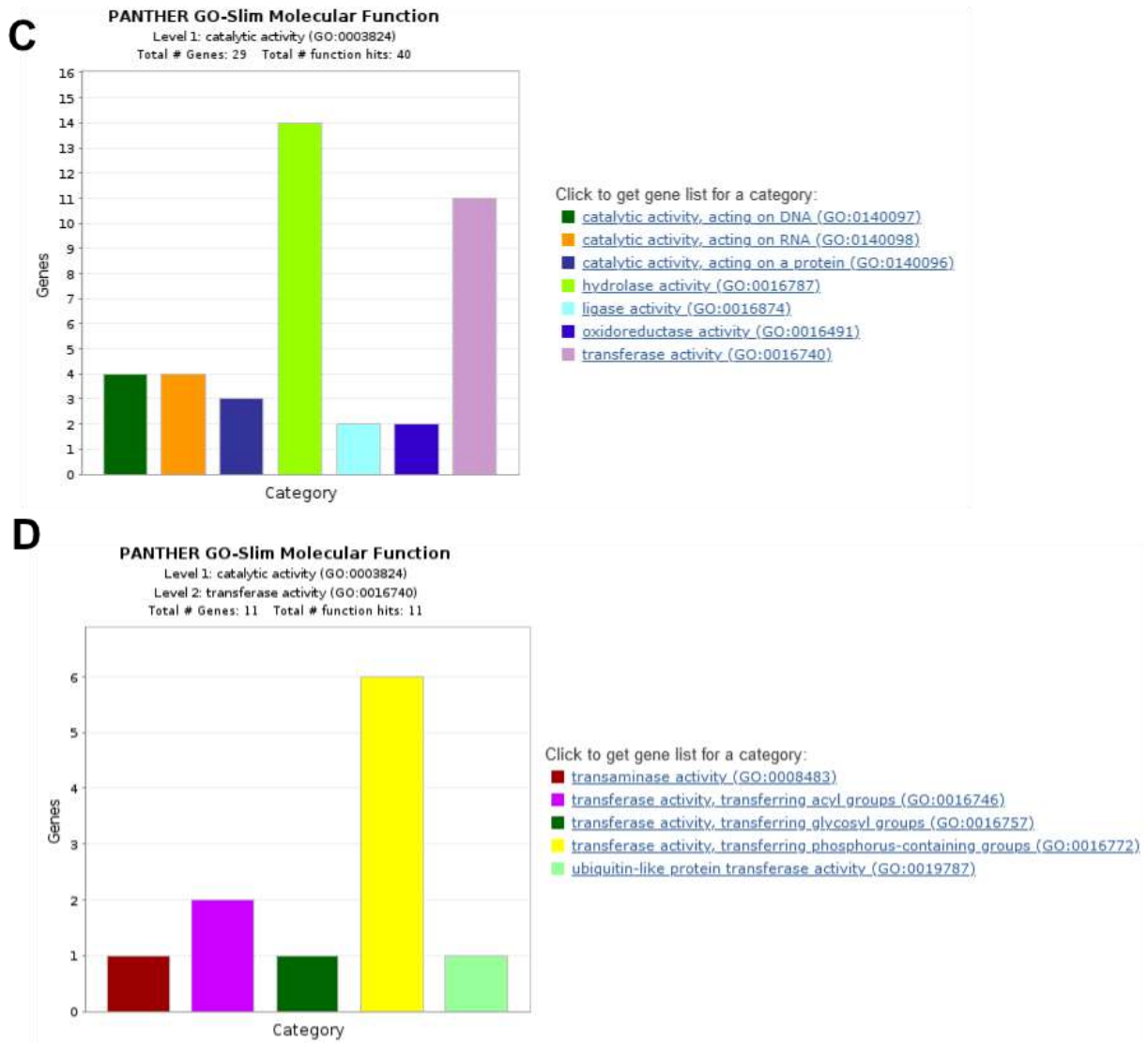




**Figure S17: PANTHER analysis of the proteins identified in the *Tb*SWRC2 BioID, part 1**

**(A)** Summary of the PANTHER GO-term analysis that depicts the function **(B)** breakdown of the binding column (grey) depicted in **(A)**.

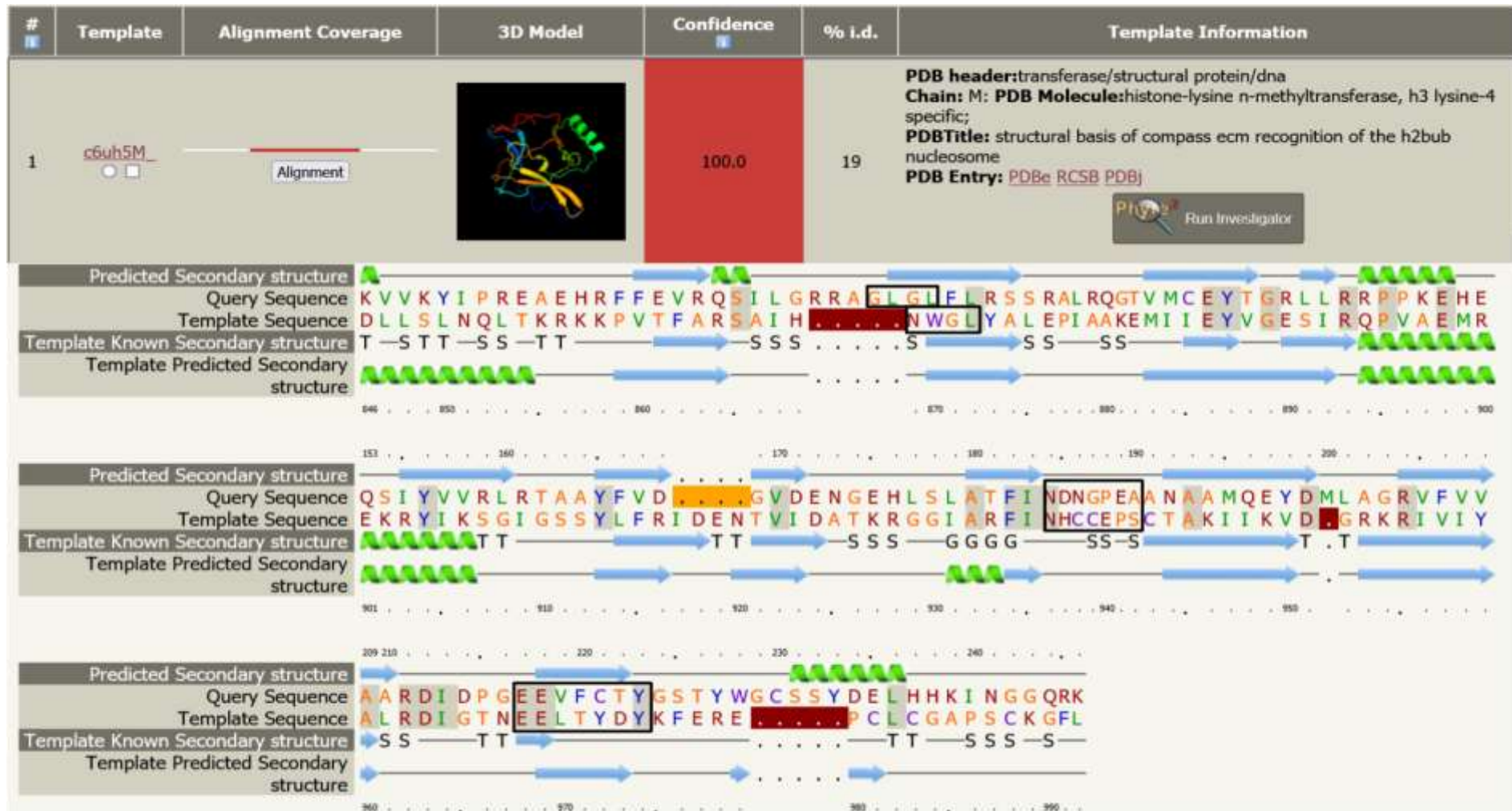
6. Appendix



**Figure S18: PANTHER analysis of the proteins identified in the TbSWRC2 BioID, part 2**

**(C)** breakdown of the catalytic activity column (pink) depicted in (A). **(D)** breakdown of the transferase activity column (pink) depicted in (C).

6. Appendix



**Figure S19: Core region of Tb927.9.8100 is homologous to the SET-domain of the H3K4 specific COMPASS SET-methyltransferase**  
 Conserved SET- methyltransferase domains 17013555 of the *S.cerevisiae* SET1 (Template, PDB ID: 6UH5) and their corresponding areas in Tb927.9.8100 (Query) are highlighted with black boxes. Dark-red areas represent sequence gaps in the template. Yellow areas represent sequence gaps in the query. Predicted and known  $\beta$ -sheets are depicted in light-blue. Predicted and known  $\alpha$ -helices are depicted in light-green. The alignment was created by the Phyre2 online modelling platform (Kelley *et al.*, 2015).



## Contact details

### BORN:

22.12.1987 in Holzminden  
Lower Saxony  
Personal status: unmarried
















### ADRESS:

Steinheilstraße 4 R  
97080 Würzburg

### E-MAIL:

Timvellmer@t-online.de

### study based skills:

- PCR-techniques including qRT-PCR 
- cell culture S2 
- genome editing:
  - cloning 
  - CRISPR-Cas 
- flow cytometry 
- molecular analytics:
  - affinity/Ion-exchange chromatography 
  - dialysis (DNA and Protein) 
  - western blot 
  - ELISA 
  - iso-electric separation 
  - protein expression 
- mass spectrometry:
  - MALDI-MS 
  - data analysis 
- microscopy:
  - light & fluorescence microscopy 
  - EM microscopy sample preparation 

# Curriculum Vitae

## Tim Vellmer

### Profile

- curious person with a broad scientific interest and a proactive working attitude
- passion for teaching – voluntary contribution to educational training at my former school in course of my PhD
- highly self-motivated and very dedicated – data-dependent restart of my doctoral project at the end of my third PhD year

### Education & Studies

- Sep. 2016 - **PhD program in the GSLS Graduate School at the University Würzburg**  
 Sep. 2021  
 Department of Cell and Developmental Biology  
 research group of Prof. Dr. Janzen  
 Topic: "Histone H2A.Z-dependent chromatin dynamics in *Trypanosoma brucei*"
- Establishment and optimization of new methods
  - Supervision of Bachelor and Master students
- Jan. 2015 – **Voluntary Internship at the Australian National University**  
 Jul. 2015  
 Research School Biology; Biomedical and Biochemistry Science  
 research group of Dr. Kevin Saliba  
 "Vitamin B1 utilization in the apicomplexan parasite *Plasmodium falciparum*"
- Oct. 2012 – **M. Sc. Molecular Biology at the University Marburg**  
 Apr. 2015 **(12.3 of 15 points)**  
 majoring in molecular and cellular infection biology  
 Master thesis in the research group of Prof. Dr. Klaus Lingelbach  
 "Investigation of potential phosphorylation of the glutamate transporter EAAT3" (12 von 15 Punkten)
- Oct. 2008 – **B. Sc. Biology at the University Marburg (final grade 2.4)**  
 Apr. 2012  
 majoring in molecular Genetics  
 Bachelor thesis in the research group of Prof. Dr. Regine Kahmann  
 at the MPI for Terrestrial Microbiology  
 "Identification of genes involved in host specificity in smut fungi"
- Apr. 1998 – **Gertrud-Bäumer-Gymnasium Remscheid**  
 May. 2007 Abitur



### additional skills:

- MS Office
- Adobe Illustrator
- Adobe Photoshop



### languages:

- German
- English
- French



### attended courses:

- Cover Letter & CV
- Supervising BSc and MSc students
- Time- and Selfmanagement
- Analysing your market potential as a scientist
- Good Scientific Practice
- Academic Writing
- Vektorgrafiken erstellen mit Adobe Illustrator
- Project Management
- Marketing your skills



### Hobbies:

sports, music, cooking, reading



### Interests:

aerospace, molecular biology, science, the world of tomorrow

## Publications

- Eisenhuth N, Vellmer T, Butter F, Janzen CJ. 2020. A DOT1B/Ribonuclease H2 protein complex is involved in R-loop processing, genomic integrity and antigenic variation in *Trypanosoma brucei*.

Prepublished in bioRxiv doi: <https://doi.org/10.1101/2020.03.02.969337>, submitted In MBio, in revision

- Vellmer T, Hartleb L, Fradera-Sola A, 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

Prepublished in bioRxiv doi: <https://doi.org/10.1101/2021.04.06.438560> , submitted in PLoS Pathogen, in revision

## Conferences

- 2020 31 Annual Molecular Parasitology Meeting, Woods Hole, USA
- 2019 13th International Symposium organized by the doctoral researchers of the Graduate School of Life Sciences "Eureka", Würzburg, Germany

## Relevant experiences & skills

- capacity in teamwork – coordination of research work with my colleagues
- good communication skills – lively technical exchange with Prof. Dr. Janzen and my colleagues in course of my PhD
- analytical and holistic thinking – evaluation and rendition of large datasets in course of my PhD project

## Study related work

- Sep. 2012 – student assistant Max-Planck-Institute for Terrestrial Microbiology
- Dec. 2012 in the research group of Prof. Dr. Kahmann

- Oct.- Nov. 2010/12/14 student assistant: support in Bachelor student supervision

- May 2012 student assist.: support in Master student supervision

- Since Jan. 2018 responsible for the development and implementation of job & orientation courses, as well as self-management courses for the NRW KAOA program at the Gertrud-Bäumer grammar school Remscheid