IDENTIFICATION AND FUNCTIONAL CHARACTERIZATION OF PROTIST SWI/SNF AND ISWI COMPLEXES

by

Alejandro Saettone Chipana

Bachiller en Ciencias Biólogicas

Universidad Nacional Mayor de San Marcos, Lima-Perú

October 2008

A thesis presented to Ryerson University

in partial fulfillment of the requirements for the degree of

Master of Science

in the Program of Molecular Sciences

Toronto, Ontario, Canada, 2015

AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A THESIS

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I authorize Ryerson University to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize Ryerson University to reproduce this thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

Abstract

The thesis aims to identify and initiate functional characterization of the SWI/SNF and ISWI complexes in *Tetrahymena thermophila*. Through affinity purification of the conserved subunit Snf5 followed by mass spectrometry (AP-MS), I identified the first SWI/SNF complex in protists. One of the subunits I found is a small bromodomain containing protein named Ibd1. Through AP-MS of Ibd1 I found Ibd1 is versatile and interacts with several additional chromatin remodeling complexes. Bromodomains are known to have affinity for acetylated lysine residues within proteins such as histones. A peptide array experiment suggests that Ibd1 also has affinity for acetylated chromatin. Indirect immunofluorescence (IF) of Ibd1 hints at a role in transcription. My analysis of *Tetrahymena* Iswi1 shows expression during meiosis, vegetative growth and starvation. IF data shows its localization is consistent with Iswi1 function in mitosis/meiosis or maintenance of silent chromatin. AP-MS of ISW1 discovered several interacting proteins of unknown function.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

°C	degrees Celsius
µ	micro
AC	acetylation
ADP	adenosine diphosphate
AP	affinity purification
APS	ammonium persulfate
ATP	adenosine triphosphate
b	base
BD	bromodomain
BLAST	Basic Local Alignment Search Tool
BRCA1	breast cancer type 1 susceptibility protein
C-	carboxy
Da	Dalton
DAPI	4',6-diamidino-2-phenylindole
DCL1	dicer-like protein 1
ddH2O	double distilled water
DOWN	downstream
DNA	deoxyribonucleic acid
dNTP	deoxynucleotides
ECL	enhanced chemiluminescence
EDTA	ethylenediaminetetraacetic acid
FITC	fluorescein isothiocyanate
FZZ	3x FLAG-ZZ epitope tag
g	grams
H	hours
H2A	histone 2A
H2Az	histone 2A variant z
H2B	histone 2B
H3	histone 3

H4	histone 4
HAT	histone acetyltransferase
HMT	histone methyltransferase
HPLC	high pressure liquid chromatography
HRP	horseradish peroxidase
Ibd1	Interactive Bromodomain Protein 1
Ibd1-FZZ	refers to the FZZ strain
Ibd1-FZZ	refers to the tagged Ibd1
IES	internal eliminated sequence
IF	indirect immunofluorescence
IgG	immunoglobulin G
INI1	integrase interactor 1
IP	immunoprecipitation
ISWI	Imitation SWI
ISWI1-FZZ	refers to the FZZ strains
Iswi1-FZZ	refers to the tagged Iswi1
Iswi1	Imitation SWI 1 protein
Iswi 2	Imitation SWI 2 protein
k	kilo
KO	knock-out
L	litres
LC	liquid chromatography
Log	logarithmic
LTQ	Linear Trap Quadrupole
m	milli
M	molar
MAC	macronucleus
MIC	micronucleus
MS	mass spectrometry
MS-AP	mass spectrometry- affinity purification
mt	mating type

N	normal
N-	amino
n	haploid number
ncRNA	non-coding RNA
NURF	nucleosome remodeling factor
Ori	origin of replication
PCR	polymerase chain reaction
p	pico
pH	potential hydrogen
pm-r	paramomycin resistant
PTM	posttranslational modification
rDNA	ribosomal deoxyribonucleic acid
Rho	rhodamine
RNA	ribonucleic acid
RP-HPLC	reversed-phase high-pressure liquid chromatography
rpm	rotations per minute
RSC	remodel the structure of chromatin
SAGA	multi-subunit histone modifying complex
SAINT	significance analysis of interactome
scnRNA	scan RNA
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SNF5-FZZ	refers to the FZZ strains
Snf5-FZZ	refers to the tagged protein
SPP	sequestering protease peptone
SWI/SNF	SWItching/Sucrose Non-Fermenting
SWR	histone variant exchange complex
TCA	trichloroacetic acid
TEV	tobacco edge virus
TGD wiki	*Tetrahymena* genome database wiki
TF	transcription factor
Tt	*Tetrahymena thermophila*

U	unit
UP	upstream
UTR	untranslated region
UV	ultraviolet
V	volts
Veg	vegetative growth
WB	Western Blot
WCE	whole cell extract
WT	wild type
w/v	weight/ volume

1 INTRODUCTION

1.1 Objective and rationale

In multi-cellular eukaryotes, the precise function of ATP-dependent chromatin remodeling complexes such as the SWI/SNF (SWItch/Sucrose Non Fermentable) and the ISWI (Imitation SWI) complexes, remains unclear. In humans, alteration or loss of the SWI/SNF complex function through mutation has been shown to be associated with cancer and there is evidence that relates SWI/SNF and ISWI complexes to tumor suppression (Pal, Vishwanath, Tempst, Sif, & Erdjument-bromage, 2004), (Wong et al., 2000), (Bochar et al., 2000) (Badenhorst, Voas, Rebay, & Wu, 2002). In addition, SWI/SNF miss regulation critically affects cellular differentiation and proliferation (Reisman, Glaros, & Thompson, 2009), (Roberts & Orkin, 2004). In *Drosophila*, the ISWI (also called NURF complex), is required for proper binding of heat-shock transcription factors to their target genes (Badenhorst et al., 2002).

The objective of this research is characterize the composition as well as molecular function and mechanism of the SWI/SNF and ISWI complexes in the unique chromatin environment of a protist model, the Alveolate Tetrahymena thermophila.

This research will further identify molecular functions of SWI/SNF and ISW1 and possibly molecular mechanisms underlying ATP-dependent chromatin remodeling which is a vital process in eukaryotic cells. By analogy, it will also aide in our current understanding of how *SWI/SNF* miss-regulation or altered function is involved in disease. In addition, *T.thermophila* is the most experimentally amenable member of the Alveolates and thus can help us to understand the basic biology of the parasitic and disease-causing Alveolate *Plasmodium* species. *Tetrahymena thermophila* is a proven model in establishing knowledge of relevance as it relates to the improvement of human health. The fundamental knowledge generated from this research will contribute to a greater understanding of human diseases, which include both chronic; such as cancer, and infectious; such as malaria.

1.2 Tetrahymena thermophila as a model organism

Tetrahymena thermophila is the model system for this research. *Tetrahymena thermophila* is part of the superphylum of Alveolates within protists. Protists are eukaryotic unicellular animals and plants. Alveolates has three primary phyla (Figure 1): Ciliates, Dinoflagellates, and the parasitic phylum Apicomplexa. The Apicomplexa phylum contains the parasite *Plasmodium falciparum*, which is responsible for malaria. The ciliate phylum contains parasites such as *Ichthyophthirius multifiliis* that affects aquaculture (Sigh, Lindenstrøm, & Buchmann, 2004). The ciliate *T.thermophila* is the most experimentally amenable member of the Alveolates and of the diverse ciliate species that include free-living organisms. Studies in *T.thermophila* have led directly to two Nobel Prizes: description of catalytic self-splicing RNA in the group I intron of the rDNA gene (Cech, 1990), and cloning of telomeres and identification of telomerase (Blackburn, Greider, & Szostak, 2006). Studies in *Tetrahymena* also led to the identification of the first histone acetyltransferase (HAT) which established the link between histone acetylation and transcription (Brownell et al., 1996). Currently *Tetrahymena* is a leading model for studies of the links between RNAi and chromain formation (Mochizuki, 2012).

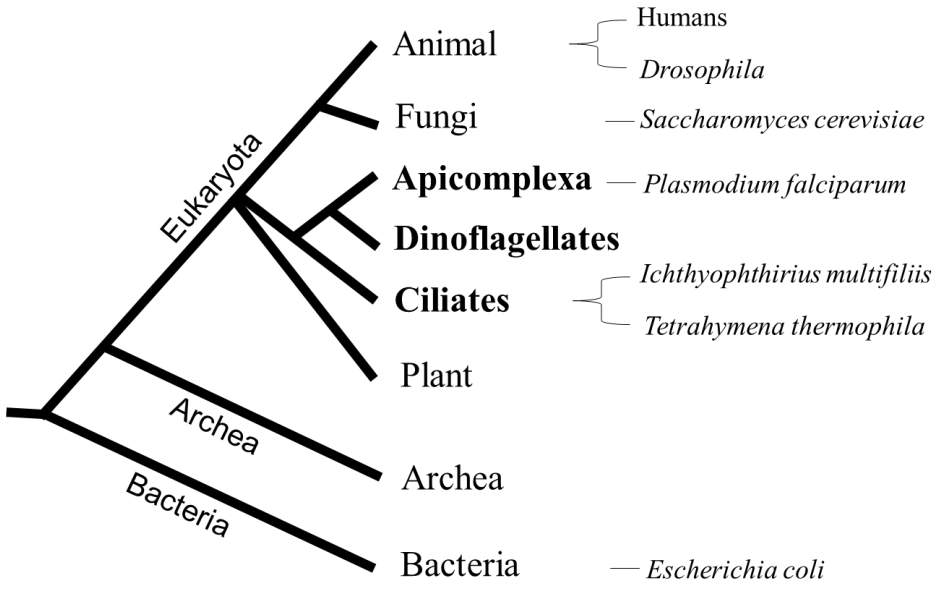


Figure 1. Phylogenetic tree showing bacteria, archea and eukaryota.
The protists branch shows the 3 phyla in bold (ciliate, dinoflagellates and apicomplexa). The phylum ciliate shows *T.thermophila* and the parasite *I.multifiliis* and plasmodium. The phylum apicomplexa shows *P.falciparum.*
Adaptation from (Nash, Nisbet, Barbrook, & Howe, 2008)(Roy & Morse, 2013)

1.2.1 Nuclear dimorphism

The utility of *T.thermophila* as a model system for nuclear structure/function and chromatin studies lies in its nuclear dimorphism, with two distinct nuclei in its single eukaryotic cell. The first of these two nuclei has an amitotic and polyploid (n=250) macronucleus (MAC). Amitosis involves nuclear division without nuclear envelope degradation, chromosome condensation, establishment of clear mitotic spindles and equal chromosome segregation (Wolfe, 1967). Therefore the segregation of alleles to daughter cells during cell division is random (Merriam & Bruns, 1988). The second is a mitotic and diploid (n=5) micronucleus (MIC) (K. Collins, 2012). Gene expression from the MAC is responsible for vegetative growth and mating while the MIC is transcriptionally silent during vegetative growth (K. Collins, 2012) but is transcribed early in conjugation in meiosis (Martindale, Allis, & Bruns, 1985). Much of the utility of *T.thermophila* to basic biomedical research has its origins in investigations of the fundamental differences between the two nuclei (Cech, 1990)(Brownell et al., 1996)(Mochizuki, Fine, Fujisawa, & Gorovsky, 2002) (Blackburn et al., 2006).

1.2.2 *Tetrahymena* life cycles

1.2.2.1 Vegetative growth

Under laboratory conditions the optimal vegetative growth of *T.thermophila* is induced at 30°C and slow shaking (100rpm). During vegetative growth *T.thermophila* cells reproduce by binary fission (Orias, Cervantes, & Hamilton, 2011). The MIC undergoes meiosis and the MAC undergoes amitosis (K. Collins, 2012). Although during conjugation MAC develops from the diploid MIC, for any gene that is heterozygous in the MIC the new MAC will initially contain equal amount of both alleles but the following a number of amitotic divisions will lead to unequal ratios of the alleles in daughter cells. Gene expression from the MAC provides the cell's phenotype and its amitosis during vegetative growth can produce different phenotypes among the progeny. At some point the cells can lose all copies of one of the alleles for a particular gene as long as the lost allele is non-essential for viability. This phenotypic appearance of a homozygous

trait in the MAC from an initially heterozygous situation is called phenotypic assortment (Merriam & Bruns, 1988).

As mentioned above, *Tetrahymena thermophila* vegetative growth is based on asexual reproduction by binary fission (Orias et al., 2011). The MAC undergoes amitosis, therefore chromosomes, and as a result alleles, are not equally segregated to each daughter cell leading to phenotypic assortment. *Tetrahymena* MAC amitotic division naturally leads to phenotypic assortment. Phenotypic assortment can be accelerated by inserting selectable markers into *T.thermophila*'s MAC genome, such as the *neo* cassette, using genetic transformation methods. The *neo* cassette contains a point mutation in the 17S rRNA gene that confers resistance to the antibiotic paromomycin (Yu, Hasson, & Blackburn, 1988). Therefore, acceleration of phenotypic assortment is possible by gradually increasing the concentration of paromomycin into the growth media (Figure 2). During the binary fissions of phenotypic assortment, the MIC undergoes conventional mitosis and therefore equal numbers of chromosomes are portioned to each daughter cell and therefore MIC chromosomes are not subject to phenotypic assortment phenomenon (Orias et al., 2011).

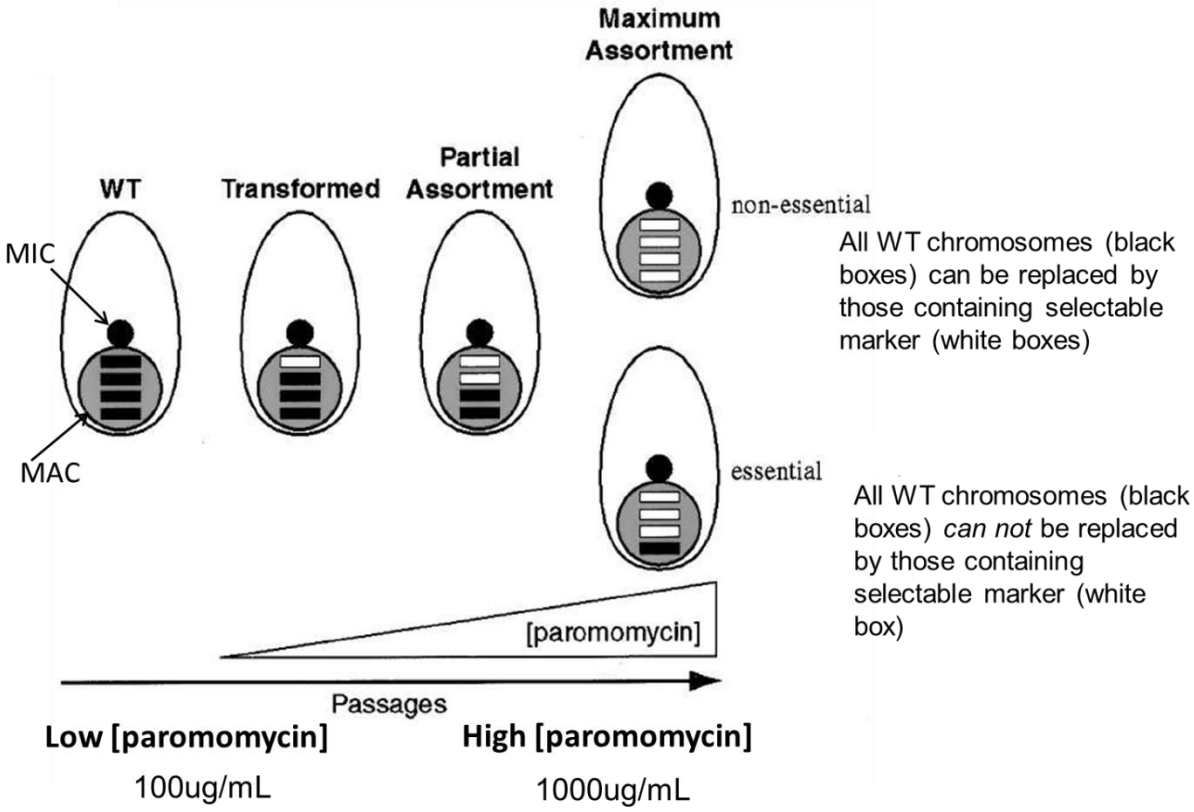


Figure 2: Acceleration of phenotypic assortment of alleles in *T.thermophila.* (Adapted from Miller and Collins, 2000)
In this figure the two cases of maximum assortment are differentiated by whether or not the transgenic MAC chromosome (white box) has some kind of deleterious effect (for example, if the selectable marker is being used to disrupt a gene you get the full replacement if the gene is non-essential for growth – partial replacement if essential for growth

1.2.2.2 Conjugation

Under laboratory conditions mating of *T.thermophila* is induced by starvation. During starvation there is expression of two important genes (*MTA* and *MTB*) that encode for mating type-specific segments and a transmembrane protein. These proteins are essential for starved cells of different mating types to carry out with attachment and conjugation (Cervantes et al., 2013). The first step is cell pairing (Figure 3 a). Meiosis initiates and the MIC becomes transcriptionally active (Martindale, Allis, & Bruns, 1982) for the only time in the cell's life cycle. As part of meiosis, the MIC elongates forming the crescent (prophase I) (Figure 3 b). The end result of meiosis is four haploid gametic nuclei (Figure 3 c), one of which will duplicate mitotically to form two active and identical gametic nuclei (Figure 3 d). The next step involves

reciprocal exchange of one of the gametic nuclei of each cell then fusion of two different gametic nuclei to form diploid zygotic nuclei (Figure 3 e). At this point the zygotic nucleus is present and the degradation of the 3 inactive gametic nuclei occurs (Figure 3 f). A double mitotic duplication of the zygotic nuclei occurs forming four identical nuclei occurs and degradation of parental MAC (OLD MAC) starts (Figure 3 g). Of the four new post-zygotic nuclei, two will develop to form two new MACs (anlagen), one will be degraded and one will mitotically duplicate to form two new MIC (Malone, Anderson, Motl, Rexer, & Chalker, 2005). (Figure 3 h). MAC development occurs in both the MACs in the cell and involves programmed DNA rearrangements which include DNA fragmentation and DNA deletion losing 10-15% from the original zygotic (MIC) genome (K. Collins, 2012). After a cell division with no associated nuclear division, the result of conjugation is four daughter cells with two newly formed nuclei that will have different genomes and different chromatin structures (K. Collins, 2012). (Figure 3 i). In conclusion, *Tetrahymena thermophila* mating, or conjugation, involves MAC development which includes large scale genome rearrangements that are regulated in large part at the level of chromatin structure (Smothers, Mizzen, Tubbert, Cook, & Allis, 1997). Interestingly, these DNA rearrangements are controlled at some level by the cell's RNAi machinery (Mochizuki et al., 2002).

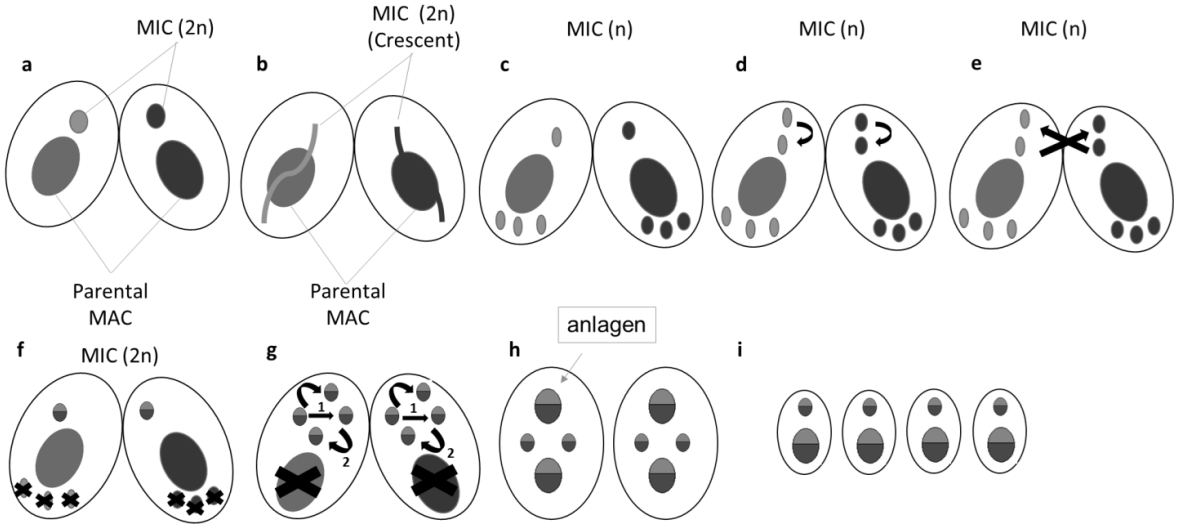


Figure 3: Mating and nuclear development.
a. Two starved cells of two different mating types attached; b. meiosis begins following formation of the crescent (prophase I); c. the result of meiosis is four gametic nuclei; d. one of which will duplicate continue to form two identical haploid gametic nuclei in each of the two cells; reciprocal exchange and fusion of the two active gametic nuclei; f. formation of the zygotic nucleus and degradation of the 3 inactive gametic nuclei; g. double duplication of the post-

zygotic nuclei and degradation of OLD MAC starts; h. swelling of two zygotic nuclei to form anlagen which will become the new MAC; i. the result of conjugation is four daughter cells.

1.2.2.2.1 DNA rearrangements in Tetrahymena thermophila

Post-transcriptional gene silencing (PTGS) is an epigenetic mechanism that, as its name implies, is related to gene silencing (Mochizuki & Gorovsky, 2004b). PAZ and Piwi containing domain proteins (PPD), such as the argonautes AGO1 and AGO2, are involved in PTGS through the formation of complexes with small functional RNAs (sRNA) that after transcription were initially long and were cut by a Dicer protein (DCL1 in *Tetrahymena* (Malone et al., 2005)). It is important to note that in *Tetrahymena* these sRNA are called scnRNA. At this point the protein or complex(es) that regulates the expression of the scnRNA precursor or long RNA is unknown; however, it is known they originate from the meiotic MIC and that RNAPII transcribes them (Mochizuki & Gorovsky, 2004a). Dicer proteins were first described in the botanical model *Arabidopsis thaliana* as the catalytic subunit of the RNA-induced silencing complex (RISC). RISC directs gene silencing by RNA interference (RNAi) (Slotkin & Martienssen, 2007). AGO1 and AGO2 bind many classes of sRNA, such as micro RNAs (miRNA), small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs) (Slotkin & Martienssen, 2007). Another example of PPD proteins are the Piwi-related proteins, which are nucleic acid-binding proteins that bind and cleave RNA, similar to the above described argonautes. One of Piwi-related proteins function (Twip1 in *T.thermophila* (Mochizuki et al., 2002)) is gene silencing (Slotkin & Martienssen, 2007) via heterochromatin formation. In *Tetrahymena*, the heterochromatin is eliminated during IES deletion as part of the large scale DNA rearrangements that occur during MAC development in conjugation (Mochizuki et al., 2002).

During heterochromatin formation and subsequent elimination, the ultimate goal of the RISC-like complex Dicer/sRNA/PPD complex (in *T.thermophila*: DCL1/scnRNA/Twi1p) is to recruit histone lysine methyltransferases (HMT; example in *T.thermophila*: Ezl1p) so this HMT is able to methylate the target histone indicated by the sRNA. During *T.thermophila*'s conjugation, RNAi-mediated heterochromatin formation and elimination starts with the transcription from the meiotic MIC of relatively long unprocessed RNA (Chalker & Yao, 2001). These unprocessed RNA are cut in the cytoplasm by DCL1 into scnRNA (Malone et al., 2005).

These DCL1/scnRNA will bind with argonaute/PIWI homologues termed Twi1p (visualized first in cytoplasm). These complexes (DCL1/scnRNA/Twi1p) join HMT (Ezl1p) forming DCL1/scnRNA/Twi1p/Ezl1p complexes. At this point DCL1 leaves the complexes. The final complexes are small-RNA-guided complexes (scnRNA/Twi1p/Ezl1p) (Mochizuki et al., 2002)(Y. Liu et al., 2007).

First, scnRNA/Twi1p/Ezl1p complexes go to the parental MAC (old MAC) where scnRNAs homologous to parental MAC DNA sequences will be degraded (Mochizuki et al., 2002). Second, the remaining scnRNA/Twi1p/Ezl1p complexes contain the scnRNAs that are homologous to MIC sequences. These complexes will go to the anlagen (developing MAC) and guide heterochromatin formation and subsequent DNA deletion (Mochizuki et al., 2002)(Y. Liu et al., 2007). After this, the chromodomain-1 of the programmed DNA deletion 1 protein (Pdd1p-CD1) recruits proteins involved in DNA elimination (Mochizuki et al., 2002)(Y. Liu et al., 2007).

The suggested scan RNA model for *T.thermophila* implies that scnRNA/Twi1p/Ezl1p will guide Ezl1p to tri-methylate the lysine 27 of H3 (H3K27me3) in the extremes of IES and perhaps the same complex will tri-methylate the lysine 9 (H3K9me3) of the interior sequences of IES (Y. Liu et al., 2007). After this, Pdd1p-CD1 will first recognize the H3K27Me2, then Pdd1p-CD1 and Pdd3p-CD will attach to H3K9Me3 marks (Taverna, Coyne, & Allis, 2002). Pdd1p will assist with recruiting other proteins such as Lia1p which is essential to complete the removal of IES (Rexer & Chalker, 2007). Failure or lack of any of these proteins or events during conjugation will affect heterochromatin formation, DNA elimination and ultimately lead to inviable progeny (Taverna et al., 2002)(Y. Liu et al., 2007).

1.3 Histones and Chromatin

Eukaryotic cells package their genomic DNA in the cell nucleus. Within the nucleus, chromatin enables eukaryotic cells to compactly store a large amount of DNA. The basic unit of chromatin is the nucleosome, which is made of 147 base pairs of DNA tightly wrapped ~1.7 times around a histone octamer. The canonical histone octamer is comprised of eight canonical

histones; two H2A and H2B dimers, and one H3 and H4 tetramer (Luger, Mader, Richmond, Sargent, & Richmond, 1997) (Figure 4). Each individual histone protein has a globular domain and N- and C- termini that are flexible and exposed from the nucleosome core (Luger et al., 1997). The canonical or replicative histones, which are only expressed during S-phase, are deposited into chromatin during DNA replication (Kamakaka & Hiller, 2005). Despite their function in efficient packaging of DNA, cells must have access to DNA within chromatin to express genes or be replicated. In order to allow gene expression or DNA replication or DNA repair , the chromatin-protected region of DNA must be unraveled in a coordinated way and become available to transcription or replication factors (Aalfs & Kingston, 2000). The non-canonical or replacement variants, which are expressed at any time during the cell cycle, are deposited into chromatin to alter nucleosome stability (Biterge & Schneider, 2014).

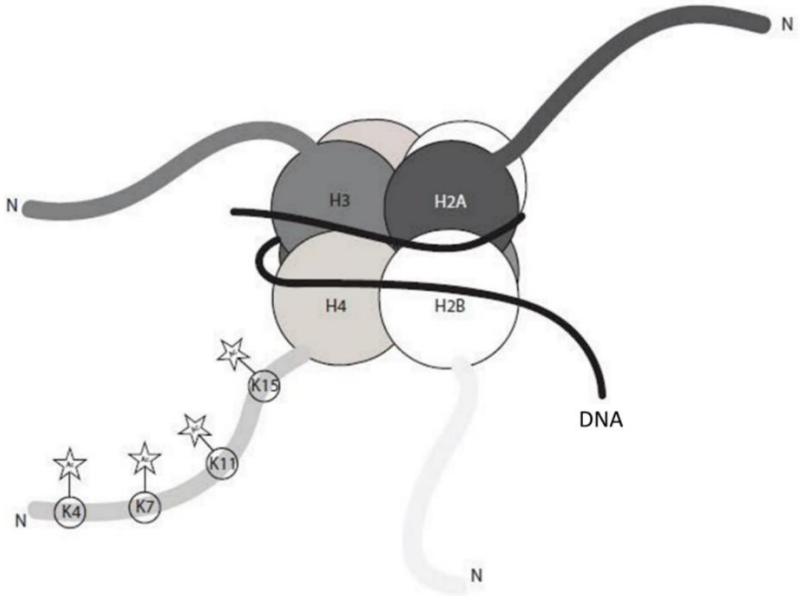


Figure 4: Nucleosome.
This representation of a nucleosome depicts the DNA wrapping the *T.thermophila* histone octamer. It also shows some post-translational modifications in the N-terminal domain of *Tetrahymena* H4. Stars represent acetylation of lysines.

1.4 Epigenetics

The study of epigenetics focuses on the environmental factors that impact phenotypic trait variation. The mechanisms responsible for making these changes work by turning genes on or

off, and by altering DNA sequences; affecting gene expression (Korochkin, 2006). Some examples of epigenetic changes are DNA methylation, as well as histone methylation, acetylation and phosphorylation, histone variant exchange, and the sliding or removing of histones within a nucleosome. Studying these epigenetic changes are important as they may be heritable, and many of these changes lead to cancer and other diseases in humans (Holliday, 2006).

1.5 Chromatin remodeling

Chromatin remodeling is the term used to describe the coordinated process where DNA is unwound from the histone core and exposed to transcription and replication factors (Aalfs & Kingston, 2000). Two general types of chromatin remodeling has been described (Tang, Nogales, & Ciferri, 2010):

1.5.1 Post-translational modification (PTM) of histone proteins

Post-translational modifications of histones are covalent modifications that are made by histone modifying enzymes (Berger, 2007). These post-translational modifications (PTM) are made to specific amino acid residues in a histone N-terminal tail, globular region, or C-terminus. Histone PTMs include 1] acetylation of lysine residues by histone acetyl transferases (HAT) such as Gcn5, and de-acetylation by histone de-acetyl transferases (HDAC). 2] Methylation of lysine or arginine residues by histone methyl transferases (HMT) such as Atxr3/Set1 and de-methylation by histone de-methyl transferases (HDMT). 3] Ubiquitination of lysine by ubiquitin ligases. Ubiquitination is reversible by deubiquitylating enzymes 4] Phosphorylation and de-phosphorylation of serine, threonine or tyrosine by kinases and phosphatases and 5] ribosylation of arginine by arginine adenosine-5′-diphosphoribosylation (ADP-ribosylation)(Strahl & Allis, 2000)(Fischle, Wang, & Allis, 2003). In some cases the PTMs can function by themselves or cis-effect. For example acetylation can negate the positive charge of the lysine side chain and result in looser attachment of the histone to negatively charged DNA. In other cases, histone modifying enzymes add PTMs to histone that are "read" by other proteins or trans-effect. For example, the human Brg1 and *Saccharomyces cerevisiae* (yeast) Sth1 both can recognize, or

“read” acetylated lysines via their bromodomain (BD). Another example of readers are *Drosophila* HP1 or *Tetrahymena* Pdd1p which both read methylated lysines via their chromodomains (CD) (Strahl & Allis, 2000)(Schwope & Chalker, 2014).

It is important to note that histones in *Tetrahymena thermophila* are highly similar but not identical in their primary amino acid sequence to human histones. Based on the analysis presented in Figure 5 (EMBOSS Needle - EMBL-EBI), it can be observed that in humans the positions of some histone residues are different than the histone residues present in *Tetrahymena*. However, the most classic residues related to transcription in humans such as H4K5ac, H4K8ac, H4K12ac and H4K16ac (Turner, Birley, & Lavender, 1992) are analogues to *Tetrahymena* H4K4ac, H4K7ac, H4K11ac and H4K15ac (See Figure 5 for pair alignments).

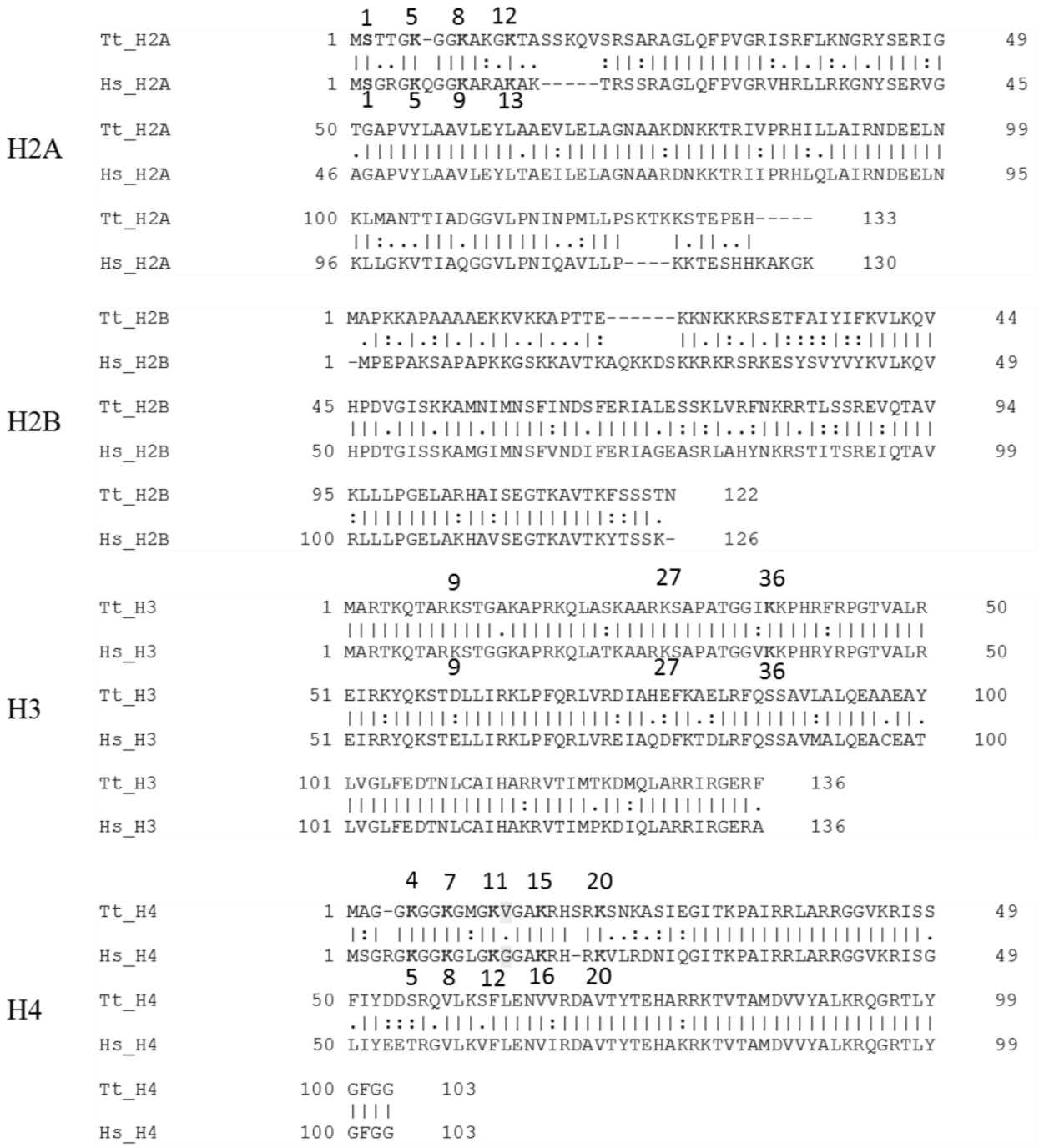


Figure 5: Human and *Tetrahymena* histone alignments
Tt: *Tetrahymena thermophila*. Hs: Human. The numbers represent the position the protein residues. The markup line between residues represent that the residues pair. Note than: in Hs_H2A lysines 9 and 13 are in position 8 and 12 respectively in Tt_H2A. Human H4 has lysines 5, 8, 12 and 16 whereas *Tetrahymen*a has the analogues lysines 4, 7, 11 and 15. (Data retrieved from ciliate.org, NCBI and analyzed with EMBOSS Needle - EMBL-EBI)

1.5.1.1 SAGA complex

A complex that plays a role in transcription is the HAT/transcriptional co-activator complex Spt-Ada-Gcn5-Acetyl-transferase (SAGA). This complex is primarily involved in transcription related processes, such as regulating its activation, elongation, and mRNA export (Baker & Grant, 2007). Additionally, SAGA acetylates numerous lysine residues on the N-terminal tails of H2B and H3 histones. It has been proposed that Gcn5, Ada2 and Ada3 compose the catalytic core of SAGA which allows for the acetylation of histones. Defects in SAGA subunits compromise its ability to acetylate histones and regulate the activation of transcription. Misregulation of this complex in humans may lead to the development of diseases (Baker & Grant, 2007). Abnormal functioning of the SAGA complex in humans could largely affect gene expression and development (Baker & Grant, 2007).

1.5.2 ATP dependent chromatin remodeling complexes

There are two kinds of ATP dependent chromatin remodeling complexes. The first class of ATP dependent chromatin remodeling complexes recognize histones, often due to a specific pattern of histone PTM, and through ATP hydrolysis these complexes will eject or slide histones modulating DNA access (Owen-hughes, 2003)(Vignali, Hassan, Neely, & Workman, 2000). Examples of this class are SWI/SNF and ISWI complexes. The second type of ATP dependent chromatin remodeling complexes, such as SWR, are involved in the exchange of core histones for histone variants within the nucleosome (Krogan et al., 2003). These two examples of chromatin remodeling are performed by a variety of large multi-subunit protein complexes who share the feature of having a central ATPase subunit belonging to the SNF2 superfamily of helicase-like (named after the first member of the class, yeast *Swi2*/*Snf2* gene) (Eisen, Sweder, & Hanawalt, 1995) (Figure 6).

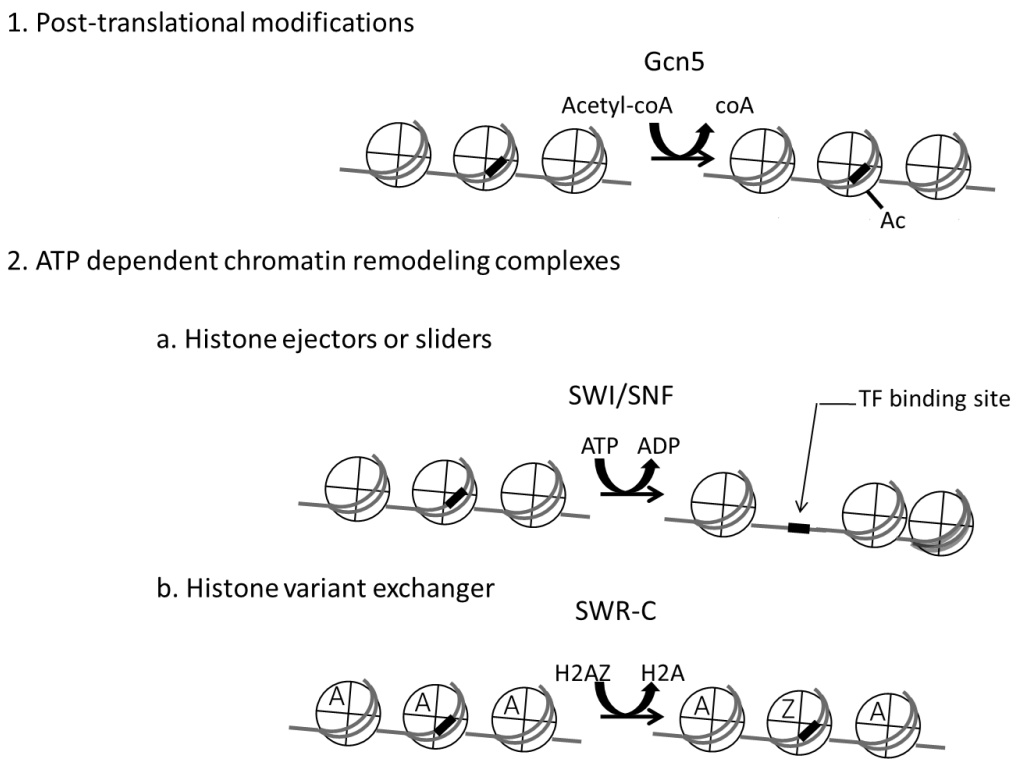


Figure 6: Mechanisms of chromatin remodeling
1; shows the histone acetyltransferase (HAT) Gcn5. 2a; shows the ATP dependent chromatin remodeling complex SWI/SNF sliding a histone and exposing a transcription factor (TF) binding sites. 2b; shows the SWR complex replacing a H2A histones for its variant H2AZ to rapidly expose TF binding sites.

1.5.2.1 SWR complex

A known chromatin remodeling complex in eukaryotes such as humans and yeast is the SWR complex (SWR-C) (Figure 6). This complex is an ATP-dependent complex and works by exchanging variant histone dimers of H2A.Z/H2B and H2A/H2B dimers within nucleosomes (Nguyen et al., 2013). The histone variant H2A.Z has a role in transcription, assists Exo1 nuclease throughout DNA mismatch repair, and enhances double strand break repair in DNA (Van, Williams, Kunkel, & Peterson, 2015). The loss of this histone variant increases the prevalence of mutations in lagging strand synthesis. In *Tetrahymena* H2A.Z is called Hv1. Hv1 localizes to the MAC throughout *T.thermophila* life cycle and to the MIC only during early stages of conjugation when MIC is transcriptionally active. Hv1 is correlated with transcription (Allis et al., 1986)(Stargell et al., 1993) and it was found that loss of Hv1 is lethal (X. Liu, Li, & GorovskyMA, 1996). Inactivation of this complex increases the number of spontaneous

mutations in yeast and a defect in chromatin remodeling performed by the SWR-C reduces overall genome stability. The catalytic subunit of SWR-C is Swr1 (Van et al., 2015).

1.5.2.2 SWI/SNF and ISWI complexes

The SWI/SNF complex was first discovered in yeast in two separate genetic screens. The first was aimed at identifying mutants defective in yeast mating type switching (SWI) (Breeden & Nasmyth, 1987) and the second study aimed to find mutants defective in fermentation of sucrose (SNF) (Neigeborn & Carlson, 1984). The SWI/SNF protein complex is an ATP dependant chromatin remodeling complex which slides and ejects nucleosomes (Figure 6) (Tang et al., 2010). The main purpose of this remodeling is to unravel DNA from histones and expose the DNA to replication or transcription factors (Tang et al., 2010).

In yeast there are two different SWI/SNF-like complexes. The first is SWI/SNF and its catalytic subunit is Swi2/Snf2. This complex has a global role in transcription (Winston & Carlson, 1992). The second yeast SWI/SNF complex is named RSC (**R**emodels **S**tructure of **C**hromatin) and its catalytic subunit is Sth1. Swi2/Snf2 and Sth1 are highly homologous with a basic local alignment search technique for protein sequences (BLASTP) E-value of 0. The closer the E-value is to zero means that the homology across the entire length of the protein is more significant. RSC is essential for mitotic growth and it has the capacity of disturbing nucleosome structure to expose genes to transcription factors. (Cairns et al., 1996). In addition, yeast has two ISWI complexes. Each of these complexes has its own ATP-ase subunit, Iswi1 and Iswi2 respectively, which work independently and cooperatively regulating developmental responses to starvation. Isw2 in association with DNA-binding proteins controls meiosis (Kent, Karabetsou, Politis, & Mellor, 2001).

In humans cells there exist two analogous SWI/SNF complexes, one containing Brg1 (Tang et al., 2010) and the other Brahma (Muchardt, Reyes, Bourachot, Leguoy, & Yaniv, 1996) as catalytic subunits. Brg1 and Brahma are highly homologous (but not identical) with a BLASTP E-value of 0. In addition, there are two ISWI complexes that have as their catalytic subunits two ATPase subunits, Snf2H and Snf2L. The ISWI complex functions are related to

chromatin assembly, nucleosome spacing, replication, and transcriptional repression and activation (Erdel, Schubert, Marth, Längst, & Rippe, 2010).

Tetrahymena thermophila on the other hand encodes a single Brg1/Snf2 yeast homologue. This protein may interact with other proteins that have homology to know members of the SWI/SNF complex in yeast suggesting the presence on a putative SWI/SNF in *Tetrahymena*. The mentioned proteins are describe in Table 1, adapted from Fillingham et al. (2006). This protein was cloned and given the name Brg1 (TTHERM_01245640) by Fillingham et al. (2006). In this paper it was also found that Brg1 localizes to MAC therefore correlates with transcription and expression of Brg1 is essential for growth. TTHERM_ followed by a number represents the gene model identifier code for *T.thermophila* MAC genes. My searches of the sequenced MAC genome for additional Brg1-related genes were unsuccessful suggesting that a single Snf2 related protein exists in *Tetrahymena*. In addition, my searching of the *T.thermophila* sequenced MAC genome suggests it encodes two ISWI related proteins, Iswi1 (TTHERM_00388250) and Iswi2 (TTHERM_00137610). This suggests that the protist contains a single SWI/SNF with Brg1 as the catalytic subunit and two ISWI complexes with catalytic subunits Isw1 and Isw2.

The primary structure of each of catalytic subunits (Swi2/Snf2, Sth1, Brg1 and Brahma) of the SWI/SNF complex in yeast and humans contains both an ATPase domain and a bromodomain (Tang et al., 2010). The bromodomain physically interacts with acetylated histones (Reisman et al., 2009) (Figure 7). Interestingly, despite high similarity to Brg1 through most of its sequence including the ATPase domain, *T.thermophila* Brg1 appears to have lost its bromodomain (Fillingham et al., 2006).

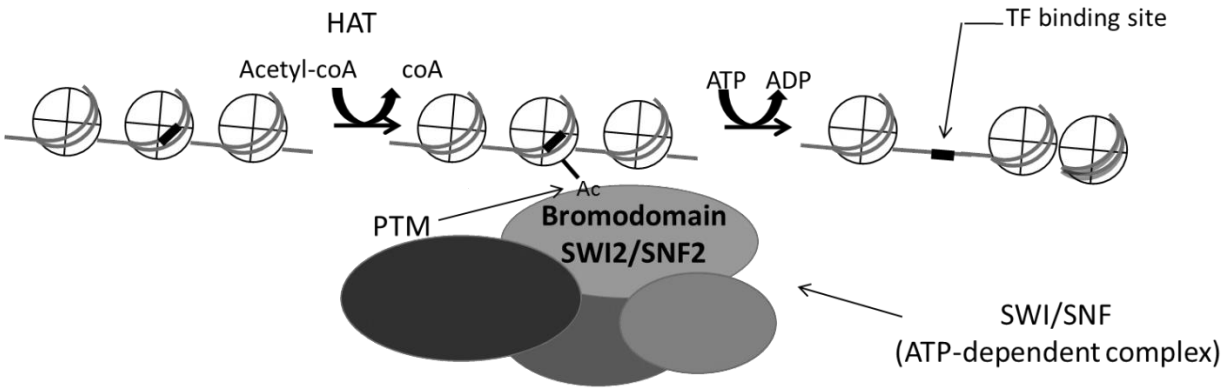


Figure 7: Bromodomains physically interacts with acetylated histones.
This figure describes that the mechanisms of chromatin remodeling are sometimes inter-related. A HAT covalently attached an acetyl group to a lysine located in the N-terminal domain of a histone. This PTM is recognized by a reader, the bromodomain of SWI2/SNF2. SWI2/SNF2 is the catalytic subunit of the SWI/SNF complex in yeast. Through ATP hydrolysis mediated by SWI2/SNF2 the SWI/SNF complex slides the histone octamer exposing the transcription factor (TF) binding site.

1.5.2.2.1 SWI/SNF complex in human health

In humans, mutations of genes encoding several members of the SWI/SNF complex (hSNF5 and BRG1) have been associated with several kinds of cancer (Roberts & Orkin, 2004). In particular, truncating mutations of hSNF5/INI1 is related to extremely aggressive malignant rhabdoid tumors in early childhood since data suggests that hSNF5/INI1 works as a tumor suppressor. (Versteege et al., 1998). The tumor suppressor BRCA1 protein and BRG1 co-purify in the same complex of proteins and a dominat-negative mutation in *hBRG1* will initiate the development of cancer (Bochar et al., 2000). In addition, in multiple tumor cell lines, BRG1 is lost (Reisman et al., 2009), (Wong et al., 2000).

1.6 Starting point to identification of SWI/SNF in *Tetrahymena*

In 2006, molecular genetic analysis of *Terahymena*'s *Brg1* (TtBrg1) suggest that it is a yeast *Brg1/Snf2* homologue; however, differently to the yeast Snf2, TtBrg1 does not have a bromodomain. In this work an antibody against Brg1/Snf2 was generated. A comparative genomic analysis of the recently sequenced *Tetrahymena* MAC genome predicted the existence

of four additional members all of which are part of SWI/SNF in yeast and humans (Table 1, adapted from (Fillingham et al., 2006)). In this paper it was also found that Brg1 localizes to MAC therefore correlates with transcription and expression of Brg1 is essential for growth. The *Tetrahymena* Snf5 (TtSnf5) was chosen for epitope tagging since it is a well-known and conserved protein of the SWI/SNF complex in yeast and humans. Between TtSnf5 and ySnf5 there is a low BLASTP E-value of $2E^{-14}$ and with respect to hSnf5 the value is $4E^{-26}$. This suggests that between TtSnf5 and ySnf5, and TtSnf5 and hSnf5 there is a section that is highly homologous.

Table 1: Predicted subunit of the SWI/SNF complex in *Tetrahymena*

Yeast Protein	***Tetrahymena* potential homologue**
Brg1	TTHERM_01245640 (without bromodomain)
Swi1	TTHERM_00243900
Swi3	TTHERM_00584840
Snf5	TTHERM_00304150
Snf12	TTHERM_00925560

1.7 Summary

To identify the members of the *Tetrahymena* SWI/SNF complex, we engineered a *Tetrahymena* strain which expresses epitope tagged (FZZ) Snf5 (ciliate.org TTHERM_00304150). My four independent Snf5-FZZ affinity purification experiments followed by mass spectrometry have identified the members of the first putative protist SWI/SNF complex. One protein that co-purified with Snf5 is a small bromodomain containing protein (ciliate.org TTHERM_00729230). A bromodomain is a protein domain that has affinity for acetylated lysine residues (Tamkun et al., 1992)(Cairns et al., 1996). I tagged the bromodomain containing protein (BD-FZZ) and through its affinity purification (three replicates) followed by mass spectrometry (MS) confirmed a reciprocal interaction with Snf5 as well as of most of Snf5-interacting proteins, confirming my identification of *Tetrahymena* SWI/SNF complex. Additionally, I found that the bromodomain containing protein is versatile in that it interacts with a set of proteins that likely originate from several additional chromatin remodeling

complexes, specifically *Tetrahymena* SAGA, SWR and a HMT. I have named this Snf5-copurifying bromodomain-containing protein **I**nteractive **B**romo**d**omain protein1 (Ibd1). In order to determine whether the bromodomain of Ibd1 is functional we expressed Ibd1 as a recombinant protein and found it specifically interacted with acetyl-lysine containing histone peptides. Interestingly, our data suggest that Ibd1 may recognize two or more PTM at one time rather than individual PTM. Finally, my indirect immunofluorescence (IF) demonstrates that Ibd1 localizes to the MAC during growth and sexual development consistent with a role in transcription.

The initial rationale for functional analysis of Iswi1 (TTHERM_00388250) was based on its publically available gene expression profile that suggested a role in DNA rearrangements of *Tetrahymena*. In order to begin to get an idea of whether this was the case, I therefore engineered *Tetrahymena* to express epitope tagged Iswi1 (Iswi1-FZZ). My Western blot analysis shows that as expected Iswi1 is expressed when DNA rearrangements occur, but also during vegetative growth and starvation. My indirect IF data shows that Iswi1 localizes to the transcriptionally silent MIC in growth and sexual development consistent with a role in mitosis/meiosis or maintenance of silent chromatin. MS data from two AP experiments of Iswi1-FZZ shows that it interacts with several proteins the significance of which at this point is unknown.

2 MATERIALS AND METHODS

2.1 Strain generation - Construction of gene specific 3xFLAG-TEV-ZZ (FZZ) targeting vectors and transgenic *Tetrahymena*

The predicted coding region of genes encoding proteins epitope tagged in this study were assessed by BLASTP to ensure clear, conserved SWI/SNF and ISWI orthologues with homology across the length of the protein particularly the C-terminus. Any BLASTP E-value that was lower than $1E^{-3}$ was accepted as sufficient molecular identity to proceed with tagging. It is important to know that in *T.thermophila*, chromosomal integration of transgenes predominantly occurs by homologous recombination (Cassidy-Hanley et al., 1997). The approach used in this study was to introduce a DNA sequence encoding the epitope tag itself as well as 3` sequences to ensure proper transcription of the tagged gene linked to a selectable marker.

The vector pBKS-FZZ (Figure 8) was used to engineer the tagging cassette as it possesses DNA sequence encoding the FZZ epitope tag. This vector contains an ampicillin-resistance cassette (for initial selection and subsequent propagation in bacteria), two unique restriction enzyme recognition polylinker sites (to sub-clone in the desired upstream and downstream homologous sequences to guide proper integration of the epitope tag). The tagging cassette contains:

1. DNA sequence encoding an epitope tag FZZ. The epitope tag FZZ contains: A] a three times repeated FLAG sequence (3xFLAG). FLAG is a polypeptide with the following sequence DYKDDDDK. (Knappik & Pluckthun, 1994). The 3xFLAG is recognized by the monoclonal and commercially available M2 antibody and the anti-FLAG M2 affinity gel agarose beads. B] TEV (Tobacco Etch Virus protease cleavage site) sequence. C] ZZ sequence (domain of protein A, recognized by IgG sepharose beads and some secondary IgG antibodies).
2. At the 3' end of the epitope FZZ there is a *RPL29* 3'untranslated region (*RPL29* 3' UTR) which contains the in frame TGA stop codon and a transcriptional terminator signal (polyA signal). This *RPL29* 3' UTR is not technically part of the tag but is

required for the proper transcription of the tagged gene since when the cassette is integrated, the normal 3`UTR of the tagged gene is essentially unlinked from the gene so this replaces it (Kataoka, Schoeberl, & Mochizuki, 2010).

3. A *neo2* chimeric gene that when expressed in *Tetrahymena* confers resistant to paromomycin to successful *T.thermophila* transformants. The *neo2* chimeric gene includes the neomycin resistance gene as the open reading frame (ORF). The ORF is expressed under the control of the *Tetrahymena* H41 promoter and includes a TGA stop codon followed by a transcriptional terminator signal (*BTU1* 3' UTR) (K. Collins, 2012). H41 is an active histone promoter which directs a high level of the neomycin gene transcription ensuring resistance to paromomycin (K. Collins, 2012).

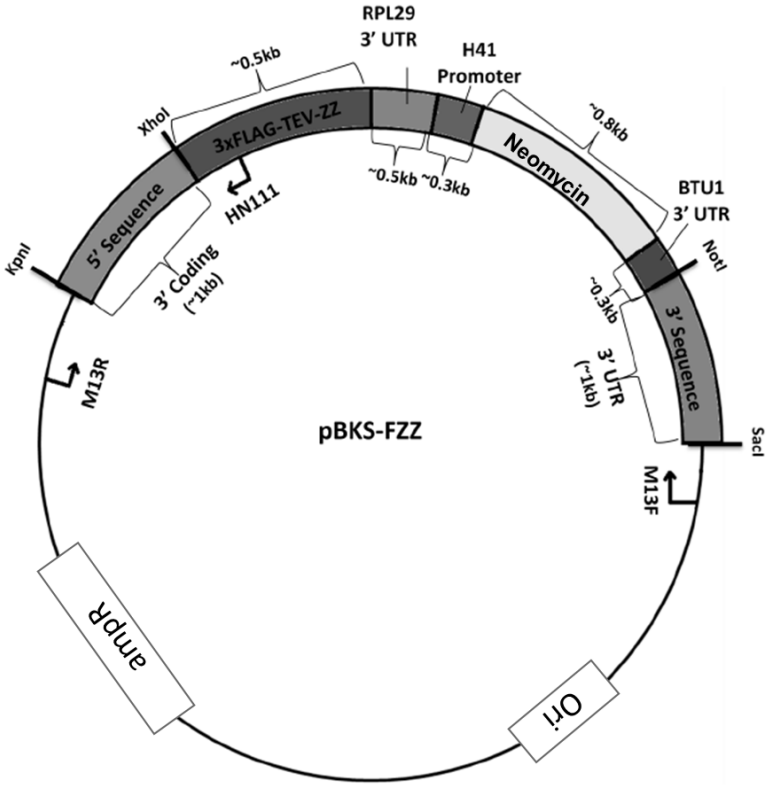


Figure 8: pBKS-FZZ plasmid MAP. (Garg et al., 2013a)

Since the *T.thermophila* MAC can be genetically manipulated and it is possible to fuse the tagging construct to the DNA sequence encoding the C-terminus of the protein of interest using the homologous recombination machinery of the cell. Tagging constructs were introduced into

the *Tetrahymena* MAC using particle bombardment. Particle bombardment is a technique used to transfer genetic material into organisms by shooting coated metal particles at the target cells. This is done using a biolistic device or a gene gun. Ultimately after selection and phenotypic assortment a strain which can express from its own promoter the desired protein fused in-frame to the epitope FZZ is generated. Note the strain can still express the tagged protein before phenotypic assortment which is used to generate a homozygous MAC (Cassidy-Hanley et al., 1997). To accomplish my research goals I used a previously generated SNF5-FZZ strain and I engineered "tagging constructs" and transgenic strains for genes coding for Ibd1 (former BD protein) and Iswi1.

2.1.1 *Tetrahymena thermophila* cell strains, culture, cells collection, starvation and mating

The *T.thermophila* cell strains used for transformation and as wild type controls were B2086 [*Mpr*+/*Mpr*+ (mp-s, mt II)] and CU428 [*Mpr*/*Mpr* (mp-s, mt VII)]. The CU428 strain is homozygous for alleles that confer resistance to 6-methylpurine (6mp) in the diploid MIC; however, in the MAC it carries only the allele that is sensitive to this drug. Thus 6mp selection can be used to select for exconjugants of a mating involving CU428. Both strains have the rDNA gene that is sensitive to paromomycin. (Gaertig, Gu, Hai, & Gorovsky, 1994) and therefore can be transformed with neo2 based plasmids used in this study. These strains were obtained from the *Tetrahymena* stock center of Cornell University, Ithaca NY.

2.1.1.1 Bench culture

These strains were stored in our lab at room temperature as bench cultures, in axenic and sterile conditions, in reusable 25 x 150mm rimless, glass test tubes. Each tube contain 10mL of sterile 1xSPP (sequestrine protease peptone) medium (see Appendix 1) and was covered with reusable 25mm two way plastic test tube cap to allow gas exchange. I will refer to the mentioned system as test tube cultures. The maintenance of bench cultures was performed every month and consisted in transferring 1mL of the previous culture into a new tube containing 9mL of sterile 1xSPP.

2.1.1.2 *Tetrahymena thermophila* vegetative growth (Overnight culture)

For small (below 10mL) and starter cultures, *Tetrahymena thermophila* cells were grown to log phase ($2x10^5$ - $3x10^6$ cells/mL) in test tube cultures in an incubating shaker (MaxQ 4450 Benchtop, Thermo Scientific) at 30^0C at 100rpm for 15hours. For large cultures, I added the starter culture in sterile flasks containing the required volume of 1xSPP media. The cells were grown to log phase using the previous described conditions in a larger incubator with shaker (MaxQ 5000 Incubated/Refrigerated Floor Shaker – Analog, Thermo Scientific). In order to allow efficient aeration, the media volume to flask size ratio was 1:10.

2.1.1.3 Cell collection and washing

For small cultures, 1mL or 10mL of overnight culture were transferred into a micro-centrifuge tube (1.5mL Boilproof Microtube, Frogga Bio) or a conical tube (15mL conical tubes, PP, Frogga Bio) and to keep the cells alive I slowly spun the micro centrifuge tubes (5424 – Eppendorf) or the conical tubes (Allegra X-15R, Beckman Coulter) at room temperature, at 3000 rpm for 4 minutes. To wash the cells, the media was quickly removed by air vacuum aspiration, before cells start swimming, trying not to disturb the cell pellet. The cells were re-suspended and washed with 1mL or 5mL of 10mM Tris-HCl pH 7.4 (starvation buffer) (see Appendix 1) accordingly and the washing and spinning process repeated followed by air vacuum aspiration. This pellet can be stored at -80^0C for later use or re-suspended in starvation buffer for immediate use. For large cultures cells of 250mL, I transferred 250mL of overnight cultures into centrifuge bottles (250mL Centrifuge Bottle PPCO, Fisher Scientific) and spun the bottles (RC5C Plus Centrifuge, Sorvall) at room temperature for 4 minutes at 3000rpm. The media was quickly removed by inversion, before cells start swimming, trying not to disturb the cell pellet. The cells were re-suspended and washed with 100mL of 10mM Tris-HCl pH 7.4 (see Appendix 1). The washing and spinning process was repeated followed by removal of the buffer by inversion. This pellet can be stored at -80^0C for later use or re-suspended in starvation buffer for immediate use.

2.1.1.4 Cells starvation

The collected and washed cells can be starved in the same amount of 10mM Tris-HCl pH 7.4 (starvation buffer) used for growth. The cells must be starved for 24 hours at 30^0C without shaking. Cells can be collected at this stage as described above for further analysis, wash is not required since they are in starvation buffer, or they can be used for mating. This pellet can be stored at -80^0C for later use or re-suspended in starvation buffer for immediate use.

2.1.1.5 Cells mating

The same amount of starved cells (for 24 hours) from different mating types, for example B2086 and CU428, were mixed and incubated at 30^0C without shaking. Cells can be collected at this stage as described above for further analysis, wash is not required since they are in starvation buffer. This pellet can be stored at -80^0C for later use or re-suspended in starvation buffer for immediate use.

2.1.2 Genomic DNA isolation

The used method for *T.themophila* genomic DNA extraction is described by Gaertig et al., 1994. A volume of 10mL of overnight culture of the B2086 strain were grown and collected as described above.

The cell pellet was re-suspended in 500µL of *Tetrahymena* lysis solution (see Appendix 1) using a micro-pipette and this solution was transferred to a 1.5mL micro-centrifuge tube. To remove the proteins and lipids, 600µL of phenol:chloroform (1:1) was added and rotated for one minute at room temperature (Rotorack 343, Fisher Scientific) until a homogeneous opaque colour was reached. The micro-centrifuge tube was spun at room temperature for 1 minute at 13000rpm (5424 – Eppendorf). The top layer was transferred to a new micro-centrifuge tube making sure that only the top layer was transferred and the middle layer was not disturbed. This step was repeated one more time. To the new tube containing the top layer from the previous step, the same amount of chloroform was added, it should be approximately 600µL, and rotated

for one minute at room temperature until a homogeneous opaque colour was reached. The micro-centrifuge tube was spun at room temperature for 1 minute at 13000rpm. The top layer was transferred to a new micro-centrifuge tube. This step was repeated one more time. To the new micro-centrifuge tube with approximately 600µL of the sample, 200µL of 5M NaCl (see Appendix 1) was added. This has a sample and 5M NaCl ratio of 3:1. This solution was mixed by inversion 10 times. To precipitate the DNA, 800µL of isopropanol, corresponding to the same volume as the previous total volume, was added. The sample was rotated for one minute at room temperature. The micro-centrifuge tube was spun at room temperature for 2 minutes at 13000rpm. The supernatant was discarded by inversion. To the same tube and to remove the isopropanol, 200µl of 70% ethanol was added. This solution was mixed by inversion 10 times followed by spinning at room temperature for 2 minutes at 13000rpm. The supernatant was discarded by inversion. This step was repeated one more time. The sample was dried using a Speed Vac System (Speed Vac System, Labconco) at room temperature for 30 minutes to remove ethanol residues. To re-suspend the DNA pellet, 100µL of sterile ddH_2O was added and incubated at 37^0C for 1 hour.

2.1.3 Sequence data retrieval and primers design

In order to design the primers, the gene sequences which code for the Snf5-interacting bromodomain-containing protein (TTHERM_00729230) and Iswi1 (TTHERM_00388250) were obtained from the *Tetrahymena* Genome Database Wiki (TGD Wiki) was used (ciliate.org). The Institute for Genomic Research (TIGR) of Bradly University, Peoria IL, sequenced the *Tetrahymena thermophila* genome and uploaded it onto a database called TGD Wiki. This database can be updated by each user and contains information on *Tetrahymena* genes and proteins from sources such as scientific literature and the research community.

To engineer FZZ-tagged strains, primers were designed to amplify two separate sequences for each desired gene of *T.thermophila* genomic MAC DNA. The first sequence was a 1kb genomic MAC region (1kbUP) located directly upstream and in-frame with, but not including, the wild type stop codon. The second sequence was an additional 1kb genomic MAC region (1kbDOWN) located immediately downstream of the wild type stop codon. The function

of these two sequences is to guide the "tagging construct" to the two complementary MAC gene sequences of interest in order to replace the WT stop codon for the FZZ and the *neo2* chimeric gene by same gene replacement. . The first pair of primers, for the 1kbUP (Appendix 3), have at their 5` end restriction enzyme recognition sites for *KpnI* and *XhoI*. The second pair of primers (Appendix 3), for the 1kbDOWN, have at their 5` ends restriction enzyme recognition sites for *NotI* and *SacI*. These two PCR generated sequences were inserted after appropriate restriction enzyme digestion into the vector (Figure 8) one at the time (tandem) using molecular cloning methods and transformation into *Escherichia coli* DH5-α.

2.1.4 PCR on genomic *Tetrahymena* DNA

The final volume used for PCR reactions was 50μL; H_2O added to complete 50μL, 1U of DNA polymerase and its buffer (Phusion High-Fidelity, Thermo Scientific), 0.2mM of dNTPs, 0.5μM of each primer (Appendix 3) and 1pg-10ng of DNA template. The PCR reaction was prepared in a PCR tube. (0.2mL thin-wall flat cap PCR tubes, MaxyClear Axygen). The PCR reactions were performed in a thermocycler (Perkin-Elmer Geneamp 9600) using the program illustrated in Table 2.

Table 2: PCR program to amplify the gene targets.

Cycles	Program	Temperature (^{0}C)	Time (seconds)
1	Initial denaturation	95	300
35	Denaturation	95	10
	Annealing	50	15
	Elongation	72	60
1	Final elongation	72	600

To assess if the PCR was successful, electrophoresis was employed. The PCR products were run in a mini 1xTBE 0.8% agarose gel (see Appendix 1) at 80V for 30 minutes (PowerPac Universal, BioRad). The gel was visualized by an imaging system (Gel Doc XR System, BioRad).

2.1.5 *Escherichia coli* growth

From our long term storage, cells are in glycerol stocks kept at -80^{0}C, *E.coli* cells containing the plasmid pBKS-FZZ with the cloned in ~800bp UP and ~800bp DOWN sequences for the Asf1 gene were streaked into LB + ampicillin (100μg/mL) plates and incubated overnight at 37^{0}C. A single colony was inoculated in a sterile culture tube (14mL round bottom, polypropylene tube, dual-position snap cap, Falcon) containing 3 mL of LB + ampicillin and shaken overnight at 225rpm at 37^{0}C (MaxQ 4450 Benchtop, Thermo Scientific). The 3mL of culture was used for plasmid isolation.

2.1.6 Plasmid isolation, pBKS-FZZ plasmid and PCR products restriction digestion, ligation, transformation and plasmid sequencing

The two PCR generated sequences, corresponding to the 1kbUP and 1kbDOWN stream of the stop codon from BD (TTHERM_00729230) and ISWI1 (TTHERM_00388250) genes, were inserted after appropriate restriction enzyme digestion into the vector one at the time using molecular cloning methods and transformation into *E.coli* (DH5-α).

2.1.6.1 Plasmid isolation

A 3mL culture of *E.coli*, containing the plasmid pBKS-FZZ, grown overnight was used for plasmid isolation with a miniprep kit (High-Speed Plasmid Mini Kit, Geneaid). To assess if the isolation was successful, the kit eluates were run in a mini 1xTBE 0.8% agarose gel (see Appendix 1) at 80V for 30 minutes (PowerPac Universal, BioRad). The gel was visualized by an imaging system (Gel Doc XR System, BioRad).

2.1.6.2 pBKS-FZZ plasmids and 1kbUP PCR products restriction digestion followed by fragment recovery

The 1kbUP BD (TTHERM_00729230), 1kbUP ISWI1 (TTHERM_00388250) PCR products and the pBKS-FZZ plasmids were digested with *KpnI* and *XhoI* (Life Technologies) following the manufacturer specifications.

The digested PCR products were cleaned (Gel/PCR DNA Fragments Extraction Kit, Geneaid) and the digested plasmid were run in a mini 1xTAE 0.8% agarose gel (see Appendix 1) at 40V for 80 minutes (PowerPac Universal, BioRad). The gel was visualized and cut with a new razor under UV light (Short/long-wave UV lamp, Cole-Parmer). The digested plasmid in the cut gel was extracted (Gel/PCR DNA Fragments Extraction Kit, Geneaid).

2.1.6.3 DNA ligation and transformation

The digested and cleaned 1kbUP BD (TTHERM_00729230) and 1kbUP Iswi1 (TTHERM_00388250) PCR products were separately ligated to the digested and cleaned pBKS-FZZ plasmids following the ligase (T4 DNA Ligase, Life Technologies) manufacturer recommendations. Both ligated products were transformed into competent *E.coli* DH5-α (Subcloning Efficency DH5-α Competent cells, Life Technologies) following the manufacturer instructions and usage of transformation controls for each experiment. After transformation, the *E.coli* cells were spread into LB + ampicillin (100μg/mL) plates and incubated overnight at 37^0C. Six single colonies from each experiment were inoculated in six sterile culture tubes (14mL round bottom, polypropylene tube, dual-position snap cap, Falcon) containing 3.5 mL of LB + ampicillin and shaken overnight at 225rpm at 37^0C (MaxQ 4450 Benchtop, Thermo Scientific). The 3mL of culture was used for plasmid isolation (High-Speed Plasmid Mini Kit, Geneaid) to verify if the colonies contained a plasmid, the rest was stored at 4^0C for further inoculation.

To verify if these plasmids were the pBKS-FZZ plasmids and contained the corresponding 1kbUP PCR product, 2μL of plasmids from each experiment were double digested

with *KpnI* and *XhoI* (Life Technologies) following the manufacturer specifications. The rest was stored at -20^{0}C for sequencing.

To verify if the released product was 100% identical to the required sequence, the isolated plasmids (undigested), corresponding to the digested plasmid, were sequenced by an external provider (ACGT Corp.), by Sanger using the generic primer HN111 (See Appendix 3). The obtained sequence was aligned (Nucleotide EMBOSS Needle, EMBL-EBI) to the gene sequence obtained from ciliate.org. The previously stored cells at 4^{0}C corresponding to this experiment were growth for plasmid isolation for the subsequent addition of the 1kbDOWN stream of the stop codon sequence.

In order to insert the 1kbDOWN BD (TTHERM_00729230), 1kbDOWN Iswi1 (TTHERM_00388250), the same steps described above were followed. The restriction enzymes used were *NotI* and *SacI* (Life Technologies) and the generic primer used for Sanger sequencing was M13F (See Appendix 3). The final arrangement of sequence on the "tagging construct" within the vector was be KpnI-1kbUP-FZZ-RPL29 3'UTR-NEO2-1kbDOWN-SacI. The cells containing the pBKS-FZZ plasmid with the required tagging construct were grown overnight. A volume of 800µL of the overnight culture was mixed with 800µL of 50% glycerol (sterile) and stored at -80^{0}C.

2.1.7 Tagging construct recovery

A volume of 30mL of *E.coli* cells containing the plasmids with the "tagging constructs" for Iswi1 (TTHERM_00388250) and BD (TTHERM_00729230) were grown and the plasmids were isolated using 10 miniprep per gene. The cell input per miniprep was 3mL. The obtained plasmid were combined obtaining 500µL of miniprep DNA [50µg]. The miniprep DNA was digested as follows; in a 1.5 mL micro-centrifuge tube 100µL miniprep DNA [10µg], 40µL of buffer, 8µL of *KpnI* and 8µL of *SacI* H_2O to complete 400µL were added and incubated at 37^{0}C for 16 hours. To assess if the double digestion was successful, 2µL of each digested product was run in a mini 1xTBE 0.8% agarose gel (see Appendix 1) at 80V for 50 minutes (PowerPac Universal, BioRad). The gel was visualized by an imaging system (Gel Doc XR System,

BioRad). The final arrangement of sequence on the "tagging cassette" within the vector were KpnI-1kbUP-FZZ-*RPL29* 3'UTR-NEO2-1kbDOWN-SacI. Prior to *T.thermophila* transformation, the plasmid in *E.coli* with the construct needed to be digested with restriction enzymes (*KpnI* and *SacI*) to free the homologous ends. For transformation into *T.thermophila*, 4-5μg of digested and purified DNA was needed.

The digested genes were cleaned using phenol-chloroform extraction. The digestion (400μL) reaction was re-suspended with the same amount of phenol:chloroform (1:1) and rotated for one minute at room temperature (Rotorack 343, Fisher Scientific) until a homogeneous opaque colour was reached. The micro-centrifuge tube was spun at room temperature for 1 minute at 13000 rpm (5424 – Eppendorf). The top layer was transferred to a new micro-centrifuge tube. To the new tube containing the top layer from the previous step, the same amount of chloroform was added, it should be approximately 400μL, and rotated for one minute at room temperature until a homogeneous opaque colour was reached. The micro-centrifuge tube was spun at room temperature for 1 minute at 13000rpm. The top layer was transferred to a new micro-centrifuge tube. This step was repeated one more time. To the new micro-centrifuge tube with approximately 400μL of the sample, 40μL (1:10) of commercially available 3M NaOAc was added. This solution was mixed by inversion 10 times. To precipitate the DNA, 1mL of 100% cold ethanol, corresponding to 2.5 times the volume than the previous 400μL volume, was added. The sample was inverted 10 times and stored at -20^{0}C overnight. The micro-centrifuge tube was spun at 4^{0}C for 15 minutes at 13000rpm (22R Refrigerated Microcentrifuge, Beckman Coulter). The supernatant was discarded by inversion. To the same tube and to remove the 100% ethanol, 1.4mL of 70% cold ethanol was added. This solution was mixed by inversion 10 times followed by spinning at 4^{0}C for 15 minutes at 13000rpm (22R Refrigerated Microcentrifuge, Beckman Coulter). The supernatant was discarded by inversion. The sample was dried using a Speed Vac System (Speed Vac System, Labconco) at room temperature for 30 minutes to remove ethanol residues. To re-suspend the DNA pellet, 12μL of sterile ddH2O was added and incubated at 37^{0}C for 1 hour. A volume of 1μL was run in an agarose gel to assess if the extraction was successful. The remaining sample was stored at -20^{0}C to use in *T.thermophila* transformation.

2.2 Transformation of *T.thermophila*

Biolistics (Bruns & Cassidy-Hanley, 1999) was used to transform *T.thermophila* MAC by gene replacement. A total volume of 50mL of B2086 and CU428 cells (different mating types, for mating experiments) were grown to log phase and starved as described above.

2.2.1 Biolistics

The gold particles (1µm Gold Microcarriers, BioRad) were coated with 5µg of the linearized DNA (digested plasmid, 1µg/µL). To attach the DNA onto the gold particles, 1M of calcium chloride and 10mM of spermidine (Sigma) were pre-chilled and added to 1µm gold particles and DNA in a microcentrifuge tube. This solution was mixed using a vortex for 15 minutes at 4^0C. The solution was spun for 10 seconds at maximum speed and the supernatant was removed. A volume of 200µL of 70% ice cold ethanol was added and span for 10 seconds at maximum speed, the supernatant was removed. A volume of 200µL of 100% ice cold ethanol was added and span for 10 seconds at maximum speed, the supernatant was removed. The DNA coated gold particles were re-suspended with 20µL of 70% ice cold ethanol. The gene gun (PDS-1000/He Biolistic particle delivery system, BioRad) apparatus and parts were washed with ddH_2O and 70% ethanol and air dried. The 50mL of starved B2086 and CU428 cells were spun at room temperature for 4 minutes at 3000rpm. The starvation buffer was aspirated and re-suspended in 1mL of fresh starvation buffer. The gene gun was assembled and the cells located into the gene gun as described (Bruns & Cassidy-Hanley, 1999). The DNA coated gold particles were shot into the cells. The shot cells were transferred to 50mL of pre-warmed to 30^0C SPP media and incubated at 30^0C at 100rpm for 3 hours for cell recovery.

2.2.2 Cell passages and phenotypic assortment

To the 50mL of recovered cells, paromomycin (Paromycin Sulfate, Calbiochem) to a final concentration of 100µg/mL was added. The 50mL of cells containing paromomycin was transferred to 96 well plates (Tissue culture plate 96-well flat bottom suspension cells, Sarstedt), 200µL per well, and incubated at 30^0C for 3 days. After this period, the cells were observed

using a microscope to locate growing cells. The growing cells are cells where identical gene replacement by homologous recombination occurred. After this initial selection, paromomycin concentration was gradually increased, to accelerate phenotypic assortment, up to 1mg/mL. This process is to replace all endogenous copies of the targeted gene with the FZZ-tagged versions. These cells were stored in bench cultures.

2.2.3 Trichloroacetic acid precipitation and Western blotting (TCA-WB)

Trichloroacetic acid (TCA) precipitation followed by Western blotting analysis was used to assess if paromomycin resistant strains were expressing the tagged protein.

2.2.3.1 TCA precipitation

The cell extracts were prepared using 10% TCA precipitation (Bright, Kambesis, Nelson, Jeong, & Turkewitz, 2010). In a microcentrifuge tube 1mL of log phase, starved or mating cells were collected, as described in cells collection, re-suspended in 90µL of starvation buffer. A volume of 10µL of 100% TCA was added to obtain a final concentration of 10% TCA. The lysis-precipitation solution remained in ice for 20 minutes. The micro-centrifuge tube was spun at room temperature for 2 minutes at 13000rpm (5424 – Eppendorf) and the supernatant was discarded. The pellet containing soluble and insoluble proteins was re-suspended in 90µL of SDS loading dye (see Appendix 1). To neutralize the solution 10µL of 1M Tris (see Appendix 1) was added. The samples were incubated in boiling water for 5 minutes. A volume of 10µL of each sample was loaded into sodium dodecyl sulfate-polyacrylamide gels (SDS gels) and the rest stored at -20^{0}C for further analysis.

2.2.3.2 Western blotting

Western blot is a simple and efficient procedure based on the electrophoretic transfer of proteins from SDS gels to unmodified nitrocellulose and detection by enhanced chemiluminescence (ECL) of specific antibodies (Burnette, 1981). The SDS gels for SDS-PAGE were prepared as follows; the stacking part of the gel was 4% (see Appendix 1) and the running

portion was either 10% or 15%. A volume of 10µL of samples was loaded ($2x10^5$ - $3x10^6$ cells/mL) to be resolved either with PiNK Plus Prestained Protein Ladder (FroggaBio) or BLUeye Prestained Protein Ladder (FroggaBio) (see Appendix 4). The gels were run (Mini-PROTEAN Tetra Handcast Systems, BioRad) in 1xSDS-running buffer (see Appendix 1) at a constant voltage of 100V for 2.5 hours (PowerPac Universal, BioRad). The separated proteins were transferred (Trans-Blot Cell, BioRad) to a 0.45µm nitrocellulose membrane (BioRad) using transfer buffer (see Appendix 1) at a constant voltage of 80V for 2.5 hours (PowerPac Universal, BioRad). The nitrocellulose membrane was blocked in 3% skim milk (BioShop) (see Appendix 1) shaking (The belly dancer, Stovall) for one hour at room temperature or overnight at 4^0C. After blocking, the nitrocellulose membrane was washed 3 times for 5 minutes each wash in 1xTBS (see Appendix 1) shaking at room temperature (The belly dancer, Stovall). The nitrocellulose membrane was incubated with different antibodies as indicated in Table 3 . Following manufacturer specifications, ECL detection (WestarNova 2011, Cyanagen) and visualization were performed (Gel Doc XR System, BioRad).

Table 3: Conditions for incubation of Antibodies

Primary antibody	Washes	Secondary antibody	Washes
1:4000; mouse monoclonal to FLAG (Sigma-Aldrich): Primary antibody buffer (see Appendix 1). 1 hour incubation with shaking	3 times for 5 minutes each wash in 1xTBS with shaking	1:5000; HRP-conjugated goat polyclonal to mouse (Cedarlane):3% skim milk (see Appendix 1). 45 minutes incubation with shaking	3 times for 15 minutes each wash in 1xTBS with shaking
1:1000; rabbit polyclonal to Brg1 (In house): Primary antibody buffer (see Appendix 1). 1 hour incubation with shaking	3 times for 5 minutes each wash in 1xTBS with shaking	1:5000; HRP-conjugated goat polyclonal to rabbit (Cedarlane):3% skim milk (see Appendix 1). 45 minutes incubation with shaking	3 times for 15 minutes each wash in 1xTBS with shaking
1:2500; rabbit polyclonal to Pdd1 (abcam): Primary antibody buffer (see Appendix 1). 1 hour incubation with shaking	3 times for 5 minutes each wash in 1xTBS with shaking	1:5000; HRP-conjugated goat polyclonal to rabbit (Cedarlane):3% skim milk (see Appendix 1). 45 minutes incubation with shaking	3 times for 15 minutes each wash in 1xTBS with shaking

2.3 Affinity purification – Western blotting (AP-WB)

The 3xFLAG sequence is recognized by the monoclonal and commercially available M2 antibody and the anti-FLAG M2 affinity gel agarose beads. The epitope FZZ is normally used for tandem affinity protein purification which involves two sequential steps of affinity purification. First, affinity purification with IgG sepharose beads followed by TVE protease and a second purification using anti-FLAG M2 affinity gel agarose beads. In this research, only one step of affinity purification using ANTI-FLAG M2 affinity gel agarose beads was used.

2.3.1 Whole cell extraction

Nuclear proteins isolation requires relatively high salt concentrations. The concentration of NaCl used was 300mM. This is relatively high, high enough to lyse the nuclei and free genomic DNA which will make the lysed cells extremely viscous. A cell pellet of less than 3mL in a 50mL centrifuge tube from 500mL (2 x 250 mL) of overnight culture or mating cells was used as starting material. Frozen pellets were kept cold during the thawing procedure. Cells pellets thaw relatively slowly. The procedure required room temperature water in a big beaker. The tubes containing the samples were swirled for 30 seconds and then placed on ice, this step was repeated until the cells pellet were totally thawed. On ice, an equal volume to the cells pellet of ice cooled 2x lysis buffer (see Appendix 1) containing protease inhibitors for yeast (Sigma-Aldrich) and 200µl of 100mM PMSF were added and mixed by gently pipetting up and down. At this point the solution became viscous due to the DNA. The solution was adjusted to a total volume of 15mL using 1×lysis buffer (see Appendix 1) containing protease inhibitors for yeast (Sigma-Aldrich) and 200µl of 100mM PMSF and mixed by gently pipetting up and down. A volume of 300µl of 10%NP40 to 0.2% final concentration were added. A volume of 5µl of Benzonase (Sigma-Aldrich) to digest all the released genomic DNA was added. Viscosity of the whole cell extraction (WCE) was tested using a cut blue tip and a P1000. There was viscosity present at this point. The tube was gently inverted 10 times and left to rotate (Rotorack 343, Fisher Scientific) at 4^{o}C for 1 hour to allow benzonase to digest all the released genomic DNA. Viscosity of the WCE was tested as mentioned above. Viscosity disappeared at this point. To clarify, this step gets rid of insoluble proteins, the samples by centrifugation of the 15mL of

WCE were divided into 10 1.5mL pre-chilled micro centrifuge tubes and spun (22R Refrigerated Microcentrifuge, Beckman Coulter) at 4^{o}C at maximum speed for 30 minutes. This step was repeated one more time by transferring supernatant by inversion to fresh pre-chilled micro centrifuge tubes. The clarified WCE (supernatants) were pooled from all tubes into a single 50mL pre-chilled centrifuge tube on ice and spun (Allegra X-15R, Beckman Coulter) at 4^{o}C at 4750rpm for 6min to separate remaining insoluble proteins, carbohydrates and lipids. The supernatant was transferred to a new pre-chilled 50mL centrifuge tube using a 25 mL pipette, on ice, and the pellet was discarded. A volume of 50µl of WCE were saved in a separate pre-chilled microcentrifuge tube at -80^{o}C to be used as input material for Western blotting.

2.3.2 Immunoprecipitation

Affinity purification (AP), affinity chromatography or immunoprecipitation (IP) permits the preparation of adsorbents (agarose beads) containing ligands (FLAG antibody) attached by bonds to the tagged proteins and this tagged protein is attached to its complex. The FLAG antibodies bonds are susceptible to specific chemical breakage using ammonium hydroxide (NH_4OH), thus providing means of removing the intact protein-ligand complex from the affinity adsorbent (Cuatrecasas, 1970). In other words, affinity chromatography using agarose beads with FLAG antibodies is a method which facilitates the separation of FLAG tagged proteins from whole cell extracts (WCE) and because native condition are used during this separation, other proteins that bind the tagged protein will co-purify as well. It is based on the highly specific interaction between the 3x-FLAG antigen, attached to the protein of interest, and the FLAG antibody, linked to the agarose beads plus native conditions during the proteins extraction. The immunoprecipitation (IP) protocol followed was designed specifically for *T.thermophila* which works accordingly (Garg et al., 2013b). In addition, from these samples a small part were used for western blotting to A] confirm and assess quality of AP by amount of bait recovered and B] to determine of TtBrg1 co-purifies as was done for Snf5-FZZ (Figure 1). Affinity chromatography was done at different physiological stages such as vegetative growth and conjugation. The rationale for this is that during vegetative growth (MAC is active) the *T.thermophila* SWI/SNF complex is expected to be interacting with a certain set of proteins. Fillingham et al. (2006) previously showed that TtBrg1 was required for progression through

conjugation. Therefore, I am interested to determine if the *Tetrahymena* SWI/SNF complex interacts with additional or different proteins than yeast/humans during growth and conjugation. Futher analysis of any specific newly identified protein(s) will predict functions.

A volume of 100µl of M2 agarose slurry (ANTI-FLAG M2 Affinity Gel, Sigma Aldrich) was used for each sample. This 100µl of M2 agarose slurry corresponds to 50µl of M2 agarose beads. Using a cut tip, the required slurry containing the beads was added to a pre-chilled 15mL centrifuge tube and washed with 5mL of cold 1xlysis buffer and slowly inverted 5 times to re-suspend beads. The 15mL centrifuge tube containing the slurry and cold 1xlysis buffer was spun at 4^{o}C at 3000rpm for 2 minutes and the supernatant was aspirated. The washing of the beads was repeated two more times. After the last aspiration the beads were equilibrated with an equal amount of 1xlysis buffer to the beads alone. Using a cut tip, 100µl per sample of the new slurry containing the equilibrated beads was transferred to pre-chilled 15mL centrifuge tubes, one per sample. The pooled supernatants were added to the respective 15mL falcon tubes containing the M2 beads and rotated at 4^{o}C for 3 hours. The tubes were spun at 4^{o}C at 3000rpm for 2 minutes. The supernatant was aspirated and discarded. The beads containing the tagged proteins and interacting proteins were washed 1 time with 10mL of IPP300 (see Appendix 1), 2 times with 5mL of IPP100 (see Appendix 1) and 2 times with 5mL of IPP100 without detergent (see Appendix 1). Between every wash the samples were rotated at 4^{o}C for 5 minutes, spun at 4^{o}C at 3000rpm for 2 minutes and the supernatant was aspirated and discarded. For the final wash, 750µl of cold $CaCl_2$/Tris solution (see Appendix 1) was added. Using a cut tip the slurries were transferred to microcentrifuge tubes and mixed by inversion 5 times. The microcentrifuge tubes were spun at room temperature at 5000rpm for 1min and the supernatant carefully removed using a micropipette. To separate the beads from the tagged proteins and interacting proteins, 500µl of freshly made NH_4OH (see Appendix 1) were added, this is a denaturing condition. The microcentrifuge tubes were rotated at room temperature for 20min and spun at room temperature at 5000rpm for 2min. The supernatant contains the proteins were transferred to new pre-chilled microcentrifuge tubes and 50µl (IP) from this tube were transferred to separate pre-chilled microcentrifuge tubes for Western blotting. The remaining 450µl were used for mass spectrometry (MS). Both samples were stored at -80^{o}C until use.

2.3.3 Western blotting

The purpose of this step is to use Western blot analysis to confirm that affinity purification was successful. The used samples are the input from WCE and IP from immunoprecipation. The samples that showed good recovery of bait were used for MS. The employed method and antibodies are described above.

2.4 Mass spectrometry and SAINTexpress analysis

Mass spectrometry was used to identify the bait protein and any co-purifying proteins in our AP experiments. The approach for this step is to use the eluates from the affinity purification step and analyze these samples using mass spectrometry (AP-MS) to identify sets of co-purifying proteins for each bait protein. Mass spectrometry is an analytical method that is used to measure and determine the presence and abundance of specific compounds by separation through mass/charge ratio. For example, based on the unique spectra of molecules we can determine compounds such as peptides. Each spectra is based on the individual and specific molecule's mass present within a sample (Fenn et al., 2013). In other words, mass spectrometry will facilitate the determination of the bait-pray interactions or the presence of any polypeptides that are co-purifying with FZZ-tagged proteins (bait).

FZZ-tagged proteins (bait) and any co-purifying proteins (prey) were digested with trypsin and were loaded to a C18 column to be separated by reversed-phase high-pressure liquid chromatography (RP-HPLC). The RP-HPLC column had C18 (hydrophobic) as static phase and 2% acetonitrile and 0.1% formic acid as mobile phase. This column was attached to the LTQ Ion Trap Mass Spectrometer to create a continuous system. All these mentioned steps were performed by the Gingras laboratory. The Mass Spectrometer outputs are spectra. The obtained spectra was identified and further filtered by a MASCOT (search engine) score to obtain spectral counts. Spectral count sums (SpecSum) represent the number of proteins in a complex mixture by providing the total number of mass spectra that match peptides to a particular protein.

The SAINTexpress (Significance Analysis of INTeractome express) algorithm, developed by the Choi, Nesvizhskii and Gingras laboratories, was used to transform the MASCOT SpecSum to SAINTexpress SpecSum. The obtained SAINTexpress SpecSum represents a given bait-prey interaction into the probability the physical interaction is significant (Skarra et al., 2011). SAINTexpress eliminates background data to obtain a reproducible set of proteins that co-purify with a given bait. Therefore SAINTexpress filters out false positive from non-specific binding proteins since the algorithm incorporates all available purification data from other tagged proteins as well as more than 30 untagged control purifications or mock IP. In other words, mock IP is important for SAINTexpress because it shows that co-purification of the protein is a direct consequence of the tagged protein and not unspecific binding to the chromatography resin. In conclusion, SAINTexpress converts spectral counts for a specific bait-prey interaction into the probability that it represents a true interaction (100%) which must pass a set high cut-off value to be included in a SAINTexpress curated data table. After SAINTexpress, the set of proteins that co-purify with the given bait are expected. Meaning, the proteins found are interacting with the FZZ-tagged (bait) proteins. The AP-MS was repeated two or more times for each bait to have a rigorous idea of the set of co-purifying proteins.

Mass spectrometry and SAINTexpress were performed by Dr. Jean Philippe Lambert in the collaborating laboratory of Dr. Ann-Claude Gingras in the Lunenfield-Tanenbaum Research Institute in the Mount Sinai Hospital in Toronto as described in the literature (Teo et al., 2014). The eluates that have the separated FZZ proteins plus any co-purifying proteins were analyzed by mass spectrometry (Garg et al., 2013) as follows. Affinity purified proteins were eluted in 0.5M NH_4OH and were taken to dryness in a speed-vac without heat. The proteins were digested with trypsin (Kean, Couzens, & Gingras, 2012). The resulting peptides were manually bomb loaded onto a capillary column and placed in the LTQ (Linear ion-Trap Quadrupole) mass spectrometer equipped with an HPLC (High Performance Liquid Chromatography) Agilent 1100 capillary (Garg et al., 2013) followed by MASCOT and SAINTexpress as described in the literature (Teo et al., 2014).

2.5 Indirect immunofluorescence

Indirect immunofluorescence (IF) was used to identify where in the cell these FZZ tagged proteins (bait) are localized. I harvested and fixed the cells for IF using the method of Wenkert and Allis (Wenkert & Allis, 1984). The nuclear counterstaining was with 4,6-diamidino-2-phenylindole dihydrochloride (DAPI) for fluorescence microscopy (Garg et al., 2013b).

In a microcentrifuge tube 1mL of washed cells in different physiological stages were spun at room temperature at 12000rpm for 2 minutes. The starvation buffer was removed by aspiration to leave 500µL of it. A volume of 1.7µL of Shandin (see Appendix 1) was added to the samples to fix the cells. The sample was spun at room temperature at 10000rpm for 1 minute and the supernatant was discarded. A volume of 100µL of cold acetone was added, inverted 10 times and placed on ice for 5 minutes. The sample was spun at room temperature at 10000rpm for 1 minute and the supernatant was discarded. A volume of 500µL of 1xPBS (see Appendix 1) was added, inverted 10 times and spun at room temperature at 10000rpm for 1 minutes and the supernatant was discarded. The primary antibodies were mixed on 1xPBST (see Appendix 1) in a 1:1000 ratio and 100µL was added. The samples were rotated at room temperature for 1 hour. The sample was spun at room temperature at 5000rpm for 2 minutes and the supernatant was discarded. The samples were re-suspended in 500µL of 1xPBS, spun at room temperature at 5000rpm for 2 minutes and the supernatant was discarded. The fluorescently conjugated secondary antibodies were mixed with 1xPBST (see Appendix 1) in a 1:1000 ratio and 100µL was added. The samples were rotated at room temperature for 1 hour in the dark. The sample was spun at room temperature at 5000rpm for 2 minutes and the supernatant was discarded. The samples were re-suspended in 500µLof 1xPBS, spun at room temperature at 5000rpm for 2 minutes and the supernatant was discarded. A volume of 100 µL of DAPI (4',6-diamidino-2-phenylindole) solution was added and incubated in the dark for 5 minutes. A total of 20µL of each sample was mounted on slides with coverslips and epi-fluorescence examined on a compound, upright Leica DM5000B microscope under magnification of 400x.

2.6 Comprehensive analysis of SWI/SNF complex

To characterize the possible function of the SWI/SNF complex during vegetative growth and conjugation, I used the data obtained from the SAINTexpress analysis of the specific newly identified protein(s) and predicted their functions based on functions of orthologous proteins present in other organisms. These proteins were grouped in putative complexes using Venn diagrams.

2.7 Synthetic gene - cloning and expression of *Tetrahymena thermophila* Interactive Bromodomain Protein 1 (Ibd1)

In order to begin to understand the function of Ibd1, it was necessary to determine if its bromodomain recognizes a specific histone post-translational modification. In order to accomplish this, a recombinant protein was engineered.

2.7.1 DNA optimization

Tetrahymena uses a non-universal genetic code where UAA and UAG specify glutamine and not stop codons. Therefore in order to express a full length Ibd1 protein in *E.coli*, we engineered a synthetic gene to use appropriate codons for glutamic acid. In addition, we took the opportunity to optimize codon usage for expression in *E.coli* BL21. The engineering of the synthetic gene was performed by an external provider (BioBasics Inc.) The gene was provided in a pUC57 plasmid.

2.7.2 Strain generation

2.7.2.1 Amplification of Synthetic Ibd1

The Ibd1 synthetic gene was amplified using polymerase chain reaction (PCR). To amplify Ibd1, primers were designed carrying at their 5` end restriction enzyme recognition sites to facilitate its cloning into pET28a. Specifically, the forward primer contained an *NdeI* site, and

the reverse primer a *BamHI* site (Appendix 3). The template was a 1:1000 dilution of the pUC57 plasmid mentioned above. The PCR conditions are outlined in the PCR section.

2.7.2.2 Restriction digestion, ligation into expression vector and transformation into *E.coli* DH5α

The PCR products and an expression pET28 plasmid were double digested with *NdeI* and *BamHI* (Life Technologies) following the manufacturer specifications. To assess if the digestion was successful, 2μL of the digested products were run in a mini 1xTBE 0.8% agarose gel (see Appendix 1) at 80V for 20 minutes (PowerPac Universal, BioRad). The gel was visualized by an imaging system (Gel Doc XR System, BioRad). The obtained image of the gel (Figure X) shows that the digestion was successful. The PCR products were cleaned, the plasmids gel purified, and both parts ligated as described above. The ligated DNA was transformed into *E.coli* DH5α strains and selected for transformants using media containing the antibiotic kanamycin. The transformed cells were streaked onto a plate with media containing kanamycin, allowing for the selection of transformants with kanamycin resistance. The success of the cloning was confirmed by sequencing. Plasmid DNA was purified, analyzed using restriction enzyme digestion and agarose gel electrophoresis, and sequenced by an external provider (M13F and M13R generic primers (See Appendix 3)) to verify that no mutations were introduced by PCR and that the gene was cloned into the correct position of the pET28 vector, in-frame with DNA sequence that when translated will add a N-terminal 6x His-tag to Ibd1. The rest of the plasmids were used for transformation into *E.coli* BL21 (DE3).

2.7.2.3 Transformation into *E.coli* BL21 (DE3)

Next the sequence verified plasmid was transformed into *E.coli* BL21 (DE3) (New England Bio Labs) following the manufacturer specifications. Positive transformants were grown in media containing kanamycin +/- IPTG. Whole cell extracts by sonication, SDS loading dye following by incubation in boiling water were made. The samples were separated by SDS-PAGE, and visualized by staining with Fast SeeBand protein staining solution in order to determine if recombinant 6xHIS-Ibd1 protein expressed at the correct size.

2.7.3 Inducing Expression of Ibd1 in E. coli BL21 (DE3)

The addition of IPTG was required to induce transcription and displace the lac repressor from the operator site in the host *E.coli* BL21 (DE3) cells. The cells were grown overnight at 37°C. The next day, I re-suspended them in liquid culture and incubated them until the OD_{600} reached approximately 0.6. The culture was then induced with IPTG for 4 hours. I took samples at 1 hour time intervals starting from time 0 to 4 hours post induction.

2.7.4 Purification of Expressed 6XHis-Ibd1

Extraction of the 6XHis-Ibd1 was done using Ni-NTA Fast Start Kit made (Qiagen). The recombinant protein was purified under native conditions. First, 3mL cell pellet from *E.coli* BL21 (DE3) culture were obtained and suspended it in native Lysis Buffer. The lysate was centrifuged to ensure cellular debris was contained in the pellet. The supernatant was retained, as this contained the soluble portion of the recombinant protein. A volume of 5µL of the supernatant was aliquoted and kept on ice. The nickel containing resin was re-suspended in a Fast Start Column by inversion. The seal was broken and allowed for the storage buffer to drain out before adding the cell lysate supernatant to the column. The flow through fraction was collected and a 5µL aliquot was obtained. The column was washed twice, each time with 4mL of native wash buffer. A 5µL aliquot was taken after each wash. With the addition of two 1mL aliquots of Native elution buffer, the bound 6XHis-Ibd1 was eluted from the column. The wash and elution buffers used had low concentrations of imidazole added to act as a competitive inhibitor, binding to the nickel and promoting the elution of the 6XHis-Ibd1. I collected 5µL of each flow through, wash and elution fraction in separate tubes and mixed them with 5µL of SDS loading dye and boiled them for 5 minutes. I used SDS PAGE to verify the success of each elution.

2.7.4.1 Buffer Exchange by Size Exclusion Chromatography

Buffer exchange is based on size exclusion chromatography (PD-10 Desalting Columns, GE Healthcare). The objective of this step is to obtain the 6XHis-Ibd1 in a buffer free from imidazole, appropriate for downstream applications. Elutions 1 and 2 from the Ni-NTA purification were passed through the column. The protein of interest, 6XHis-Ibd1, was large enough to enter the pores of the resin and was therefore eluted from the column in the new buffer. This elution was obtained in different fractions. Each of the 8 samples were run in an SDS-PAGE, followed by staining of the gel by Fast SeeBand Protein stain solution (FroggaBio).

2.8 Histone Peptide Array

A commercial human histone peptide array (MODified Histone Peptide Array, Active Motif Catalog number 13005) and its reagents (MODified Protein Domain Binding Kit, Active Motif Catalog number 13007) was used to screen Ibd1 for binding interactions with 384 unique posttranslational modification combinations. The array contained 384 sites, each with a different combination of PTM. The mechanism of the array was comparable to a Western blot. The array was blocked with 5% milk for 1 hour, washed with 1xTBS and then incubated with my protein of interest, 6XHis-Ibd1, for 1 hour. The array was washed with 1xTBS and incubated first with the primary antibody; anti-6XHis, and then with the secondary antibody; anti-rabbit HRP. After washing again with 1X TBS, ECL was used for detection. The array was visualized by an imaging system (Gel Doc XR System, BioRad). The obtained image was processed in the software provided with the kit (Array Analyze Software, Active Motif). The output of the software was used to analyze each of the sites with binding interactions.

3 CHAPTER 1 - MOLECULAR CLONING OF TAGGING CONSTRUCTS AND GENERATION OF TAGGED LINES RESULTS

3.1 Genomic DNA isolation

I purified whole cell genomic DNA from WT *Tetrahymena* in order to use as template for PCR amplification of homologous sequence that were cloned in pBSK-FZZ. Volumes of 2, 5 and 10µL were run in an agarose gel to assess if the extraction was successful. The remaining sample was stored at -20^{0}C for use in PCR (Polymerase Chain Reaction). To assess if the extraction was successful, the DNA was run by electrophoresis. The obtained image of the gel (Figure 9) shows that both extractions were successful and contain DNA.

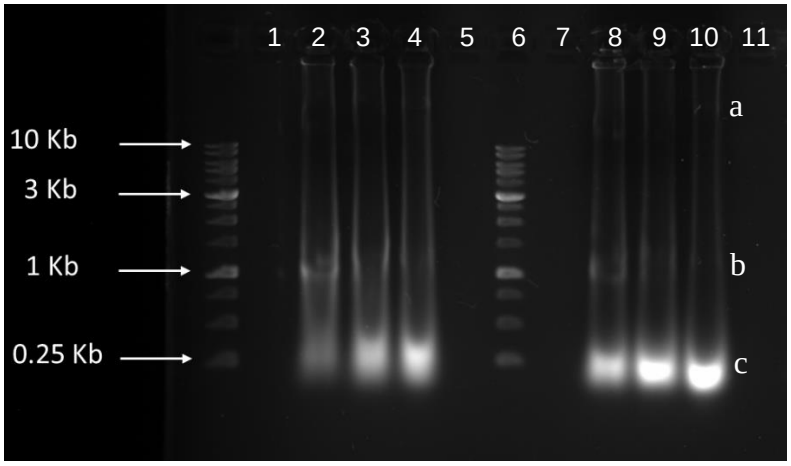


Figure 9: Two DNA isolations from whole cell genomic DNA.
First DNA Isolation: lane 1, 4µL ladder; lane 2, empty; lane 3, 4µL loaded (2µL blue loading dye + 2µL DNA); lane 4, 7µL loaded (2µL blue loading dye + 5µL DNA); lane 5, 7µL loaded (2µL blue loading dye + 10ʎ DNA); lane 6, empty. **Second DNA Isolation:** lane 7, 4µL ladder; lane 8, empty; lane 9, 4µL loaded (2µL blue loading dye + 2µL DNA); lane 10, 7 µL loaded (2µL blue loading dye + 5µL DNA); lane 11, 7 µL loaded (2µL blue loading dye + 10µL DNA); lane 12, empty. **a.** DNA; **b and c.** RNA.

3.2 PCR on genomic *Tetrahymena* DNA

Next I amplified UP and DOWN sequence for Ibd1 and ISW1 genes. The obtained image of the gel (Figure 10) shows that the four reactions; 1 for UP Ibd1, 1 for DOWN Ibd1, 1 for UP ISWI1 and 1 for DOWN ISW1 were successful and contain the expected PCR product of approximately 1kb. The remaining sample was separated from PCR Master Mix (Gel/PCR DNA Fragments Extraction Kit, Geneaid), the PCR products recovery was assessed by electrophoresis. The cleaned and recovered products were stored at -20^{0}C to be used for enzyme digestion.

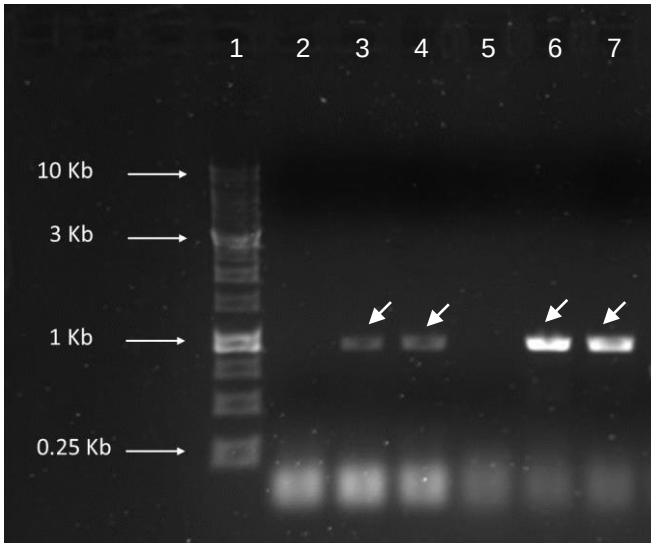


Figure 10: PCR on genomic Tetrahymena DNA for BD (TTHERM_00729230) and Iswi1 (TTHERM_00388250).
BD (TTHERM_00729230) PCR: lane 1, 4µL ladder; lane 2, negative control (water as template); lane 3; 1kbUP; lane 4, 1kbDOWN. **Iswi1 (TTHERM_00388250) PCR:** lane 5, negative control (water as template); lane 6, 1kbUP; lane 7, 1kbDOWN. 5µL of every mixed sample was loaded (2µL blue loading dye + 5µL PCR products). The arrows show the PCR products.

3.3 Plasmid isolation

PCR amplified UP and DOWN sequence for Asf1 were cloned into pBSK-FZZ prior to this thesis work. The obtained image of the gel (Figure 11) shows that the 12 plasmid isolation of pBSK-FZZ were successful since the band corresponding to the plasmids are running at the expected size ~9kb.

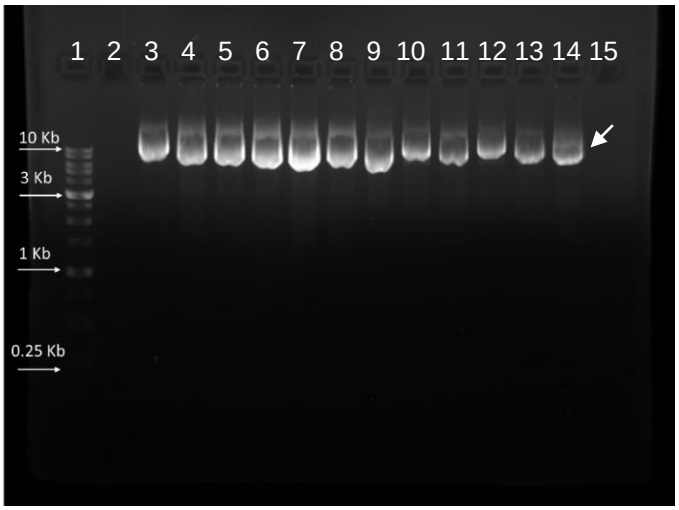


Figure 11: Plasmid pBSK-FZZ isolation.
Lane 1, 4µL ladder; lane 2, empty; lane 3 - 14, plasmids; lane 15, empty. The arrow shows the 12 plasmid isolation of pBSK-FZZ. The bands correspond to the plasmids and are running at the expected size ~9kb.

3.4 pBKS-FZZ plasmids and 1kbUP PCR products restriction digestion followed by fragment recovery

Two plasmids containing the cloned in UP and DOWN (~800bp) sequences for Asf1, UP BD PCR product and UP ISWI1 PCR product were double digested with *KpnI* and *XhoI*. To assess if the digestion was successful, 2µL of the digested products were assessed by electrophoresis The obtained image of the gel (Figure 12) shows that the digestion was successful since we were able to observe that the two plasmids containing the cloned in Asf1 gene released the ~800bp corresponding to the UP region.

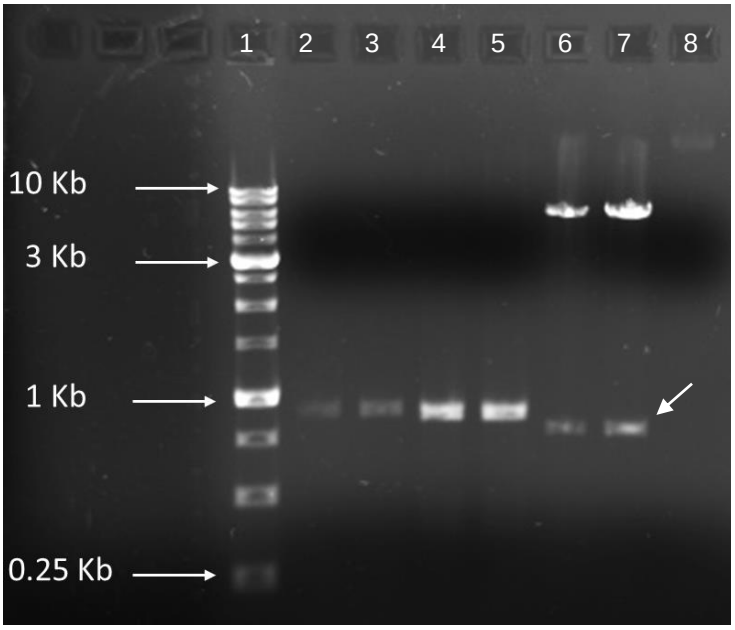


Figure 12: Restriction digestion with *KpnI* and *XhoI*.
Lane 1, 4μL ladder. **UP BD (TTHERM_00729230) PCR products digestion (~1kb):** lane 2, undigested UP PCR product; lane 3; digested UP PCR product. **UP ISWI1 (TTHERM_00388250) PCR products digestion (~1kb):** lane 4, undigested UP PCR product; lane 5; digested UP PCR product. There is no difference between the undigested and the digested PCR product because the eliminated flanking DNA sequence at both ends is of approximately 20 base pairs in total. **pBKS-FZZ plasmids (with Asf1 cloned in) digestion:** lane 6 and 7, two digested plasmids and the released insert (showed with arrow) contained in the plasmid (~800bp); lane 8, undigested plasmid (digestion control).

3.5 DNA ligation and transformation

Prior to ligation and to verify recovery, 2μL of both digested and cleaned PCR products and plasmid (vector) were assessed by electrophoresis. The obtained image of the gel (Figure 13) shows that the fragment recovery was successful since presents the PCR and digested products.

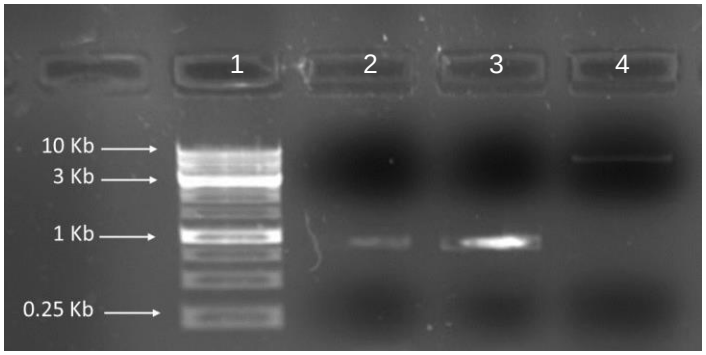


Figure 13: Fragments recovery.
Lane 1, 4µL ladder; Lane 2, UP BD (TTHERM_00729230) PCR products digested and cleaned; lane 3, UP ISWI1 (TTHERM_00388250) PCR products digested and cleaned; lane 4, pBKS-FZZ plasmids digested and cleaned.

I ligated the UP BD to the vector (experiment 1) and the UP ISWI1 to the vector (experiment 2) using T4 ligase. I followed the ligation with transformation of competent *E.coli* (DH5α). From the transformants colonies I selected six colonies from each experiment (1 and 2), growth them in liquid media and isolate the plasmids. From the plasmid isolation, 2µL of each of the twelve plasmid isolations, six corresponding to experiment 1 and six to experiment 2, were assessed by electrophoresis. The obtained image of the gel (Figure 14) shows that the selected colonies contain a plasmid.

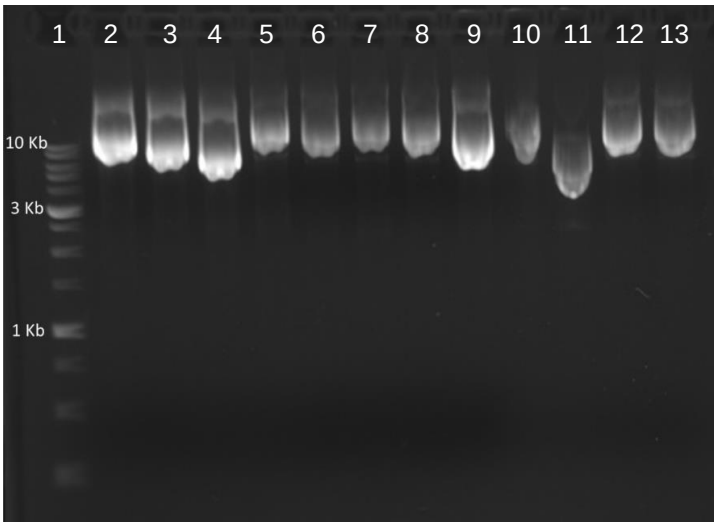


Figure 14: pBKS-FZZ Plasmids after 1kb UP transformation.
Lane 1, 4µL ladder; Lane 2-7, digested UP BD (TTHERM_00729230) PCR products ligated to digested pBKS-FZZ linear plasmids; lane 8-13, digested UP ISWI1 (TTHERM_00388250) PCR products ligated to digested pBKS-FZZ linear plasmids.

In order to assess if the transformed plasmids contain the approximately 1kb in size corresponding to the UP BD and UP ISWI1, I performed double digestion with *KpnI* and *XhoI*. To assess if the released portions correspond to the expected size of approximately 1kb, the digestion products were run in an electrophoresis gel. The obtained image of the gel (Figure 15) shows that the plasmids contain a product of approximately 1kb.

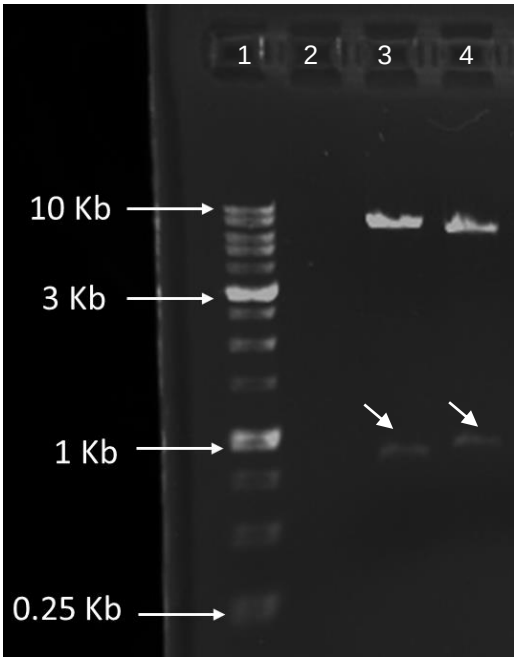


Figure 15: pBKS-FZZ Plasmids digestion.
Lane 1, 4µL ladder; lane 2, empty; lane 3, *KpnI* and *XhoI* double digestion of a candidate plasmid that may contain the 1kbUP ISWI1 (TTHERM_00388250); lane 4, *KpnI* and *XhoI* double digestion of a candidate plasmid that may contain the 1kbUP BD (TTHERM_00729230). The arrows show the approximately 1kb sequences.

After Sanger sequencing, the obtained sequence was aligned (Nucleotide EMBOSS Needle, EMBL-EBI) to the gene sequence obtained from ciliate.org, both sequences corresponding to the 1kbUP stream of the stop codon from Iswi1 (TTHERM_00388250) and BD (TTHERM_00729230) were 100% identical. The previously stored cells at 4^0C corresponding to this experiment were grown for plasmid isolation for the subsequent cloning of the 1kbDOWN stream sequence. The addition of this sequences was successful (data not shown).

3.6 Tagging construct recovery

In order to transform *Tetrahymena* by biolistics, it is necessary to release the entire tagging cassette from the plasmid backbone to free the ends of the homologous sequence. Double digestion was with *KpnI* and *SacI*. Per gene, I used minipreps to extract plasmids from a total of 30mL of grown cells obtaining 50ug of miniprep DNA. I double digested the miniprep DNA in 5 reactions per gene. To assess the digestion, I used 2 μL of each digestion for electrophoresis. The obtained image of the gel (Figure 16) shows that the digested products contain a product of approximately 3kb corresponding to the tagging construct (TC) and .a higher band of approximately 6kb corresponding to the plasmid backbone (PB).

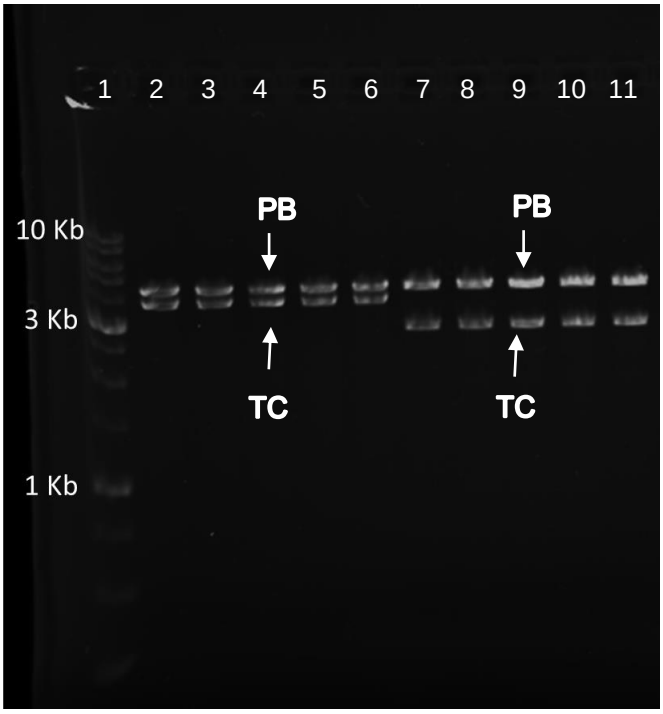


Figure 16: Double digestion to obtain the tagging constructs.
Lane 1, 4μL ladder; lane 2-6, BD (TTHERM_00729230) double digested with *KpnI* and *SacI*; lane 7-11, ISW1 (TTHERM_00388250) double digested with *KpnI* and *SacI*. The ~3kb bands correspond to the tagging constructs (TC) and the higher band of ~6kb correspond to the plasmid backbone (PB).

Each of the 5 digestions per gene were cleaned using phenol-chloroform extraction and recovered by ethanol precipitation. To assess if the extraction was successful the DNA was run

by electrophoresis. The obtained image of the gel (Figure 17) shows that both extractions were successful and contain the tagging construct and plasmid backbone.

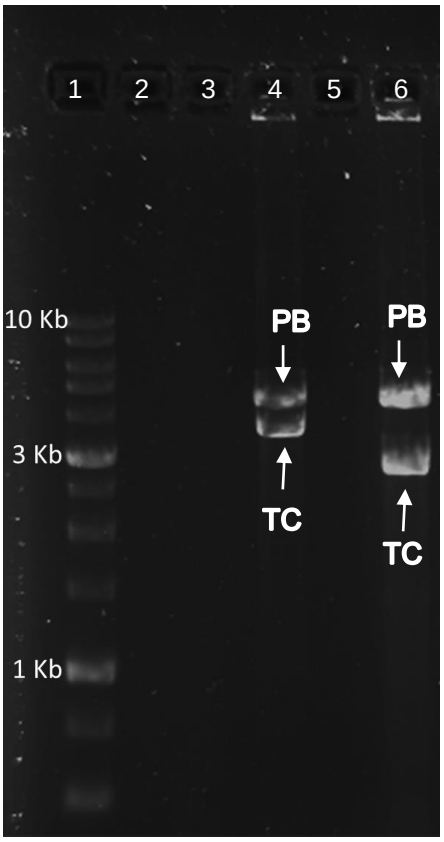


Figure 17: Tagging construct clean up
Lane 1, 4µL ladder; lane 2-3, empty; lane 4 BD (TTHERM_00729230) double digested with *KpnI* and *SacI* and cleaned with phenol:chloroform; lane 5, empty; lane 6, ISWI1 (TTHERM_00388250) double digested with *KpnI* and *SacI* and cleaned with phenol:chloroform. The ~3kb bands correspond to the tagging constructs (TC) and the higher band of ~6kb correspond to the plasmid backbone (PB).

4 CHAPTER 2 – Identification of *Tetrahymena* SWI/SNF and identification and analysis of Ibd1

In multi-cellular eukaryotes, SWI/SNF complex is known to function as an ATP dependent chromatin remodeler (Vignali et al., 2000). The identity of individual SWI/SNF proteins are well known in fungi (Peterson, Dingwall, & Scott, 1994), (Logie & Peterson, 1999), humans (Muchardt et al., 1996) and *Drosophila* (Dingwall et al., 1995). In addition to physical interactions within the complex, interactions exists with other proteins that presumable recruit SWI/SNF to specific locations. In humans, *hBRG* is involved in many protein-protein interactions such as in β-catenin (Barker et al., 2001), the Fanconi anemia protein A (FANCA) (Otsuki et al., 2001), and Hp1α (Nielsen et al., 2002). Although the putative catalytic subunit Brg1 has been clones and functionally analyze (Fillingham et al., 2006), the SWI/SNF complex remains uncharacterized at the experimental level in *T.thermophila*.

4.1 Assessment to determine if Snf5-FZZ pm-r transformants are expressing the tagged protein (Snf5-FZZ) at the expected size

I started my work with a potential SNF5-FZZ that had been generated previously but not confirmed by western blotting. To confirm the identity of the strains, I used TCA precipitation followed by Western blotting analysis to assess if the provided SNF5-FZZ strains were expressing the tagged protein. When I blotted with the M2 antibody, Figure 18 shows that only the SNF5-FZZ and not the untagged WT strain is expressing a protein of about 60kDa which corresponds to 18kDa from the FZZ and 42kDa size predicted for the Snf5 protein. For the Brg1 antibody, Figure 18 shows that both, a WT strain and the SNF5-FZZ strain, are expressing Brg1 (loading control). This confirms that I have the correct strain.

4.2 Expression analysis of Snf5

In order to identify whether Snf5 is member of a *Tetrahymena* SWI/SNF complex, I affinity purified the FZZ-tagged protein and any co-purifying proteins by performing a one-step

affinity chromatography on whole cell extracts made from cells in vegetative growth expressing SNF5-FZZ. My Western blotting using M2 antibody for the input and affinity purified material in Figure 18 shows that only the SNF5-FZZ strain is expressing a protein of about 60kDa corresponding to SNF5-FZZ. If Snf5 is a member of *Tetrahymena* SWI/SNF, then as predicted, it should co-purify with Brg1, the *Tetrahymena* Snf2 homolog and likely ATP-dependent engine of the complex. Since I had an antibody that recognizes Brg1 (Fillingham et al., 2006), I used it in western blotting. Anti-Brg1 for the input and affinity purified material shows that both the WT and the SNF5-FZZ show a similar layout to the TCA extraction (Figure 18), as expected. However, some degradation of Brg1 was observed (Figure 18). Specifically, after the TCA extraction, Brg1 presented 1 band (Figure 18), and after the whole cell extraction with the gentler AP-MS lysis buffer, Brg1 presented 2-3 bands (Figure 18). If a PTM had been present on Brg1, two bands would have been visible after the TCA extraction. However, this statement is not all inclusive since all Brg1 could be modified with this hypothetical PTM but this is unlikely to happen. Brg1 with a PTM such as phosphorylation or ubiquitination would appear heavier than the WT size and degradation would make it smaller than WT size. Because the additional bands seen are smaller than the WT size I think it is more likely that the multiple bands recognized by anti-Brg1 seen in Fig 17 in the AP experiment represent degradation of Brg1 and not the presence of a PTM.

When blotting with the M2 antibody, Figure 18 shows that only the SNF5-FZZ strain is expressing a protein of about 60kDa corresponding to SNF5-FZZ. WT represents the mock IP. When I blotted with the anti- Brg1 antibody Figure 18 shows that only the SNF5-FZZ is co-purifying with Brg1. This is a predicted result based on the fact that Snf5 and Brg1 interact in other systems. Thus as I predicted I have likely purified a *Tetrahymena* specific SWI/SNF complex. From this result, I expect to see Brg1 identified in the SAINTexpress analysis of the affinity purified material (see next section).

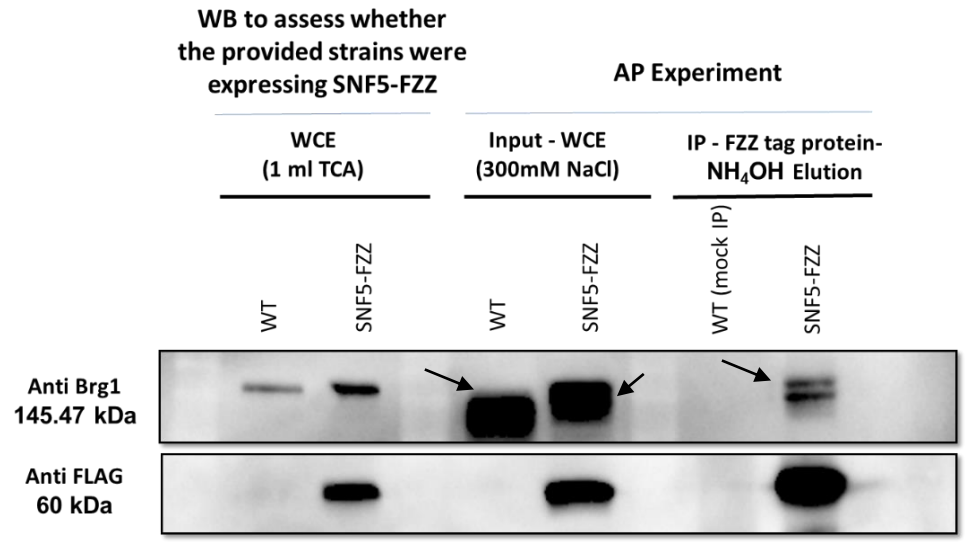


Figure 18. Expression Analysis / Affinity Purification of SNF5-FZZ
Snf5-FZZ is recognized by anti FLAG. The arrows point to where degradation of Brg1 was observed.

4.3 MS and SAINTexpress analysis of affinity purified Snf5-FZZ during vegetative growth

With affinity purified SNf5-FZZ we next used LCMS to identify any co-purifying proteins. The approach for this step is to use the affinity purified Snf5-FZZ eluates from the affinity purification step and analyze them using LCMS (AP-MS) to identify protein-protein interaction profiles. The SAINTexpress (Significance Analysis of INTeractome express) algorithm, developed by the Choi, Nesvizhskii and Gingras laboratories in Toronto, Ontario, was used to transform spectral counts obtained for a given bait-prey interaction into the probability of whether the physical interaction is significant or not (Skarra et al., 2011). Therefore only true interactions that pass a set cut-off value were included in a SAINTexpress curated data table (Table 4). At this final step I found the SNF5-FZZ co-purifying proteins, or, the proteins that are interacting with SNF5-FZZ. In order to eliminate background data and obtain a reproducible set of proteins that co-purify with a given bait, these analyses were performed in conjunction with a mock IP from untagged WT cells and database mock IP. Mock IP are very important for SAINTexpress to work properly because with respect to background proteins there are some of these proteins that show up all the time in every mock IP and there are some background proteins that show up only a percentage of the times. Therefore having multiple mock IPs give

SAINTexpress better confidence when predicting whether a given bait-prey interaction is significant. Similarly, multiple repeats of the given AP-MS is important. AP-MS was repeated four times for each bait to have a rigorous idea of the set of co-purifying proteins. These proteins can be visualized in Table 4.

The SAINTexpress analysis of the MS data for the four replicates of SNF5-FZZ (Table 4) reveals five of the predicted SWI/SNF members by Fillingham et al, 2006 (Fillingham et al., 2006); Brg1 (also as expected from my western blotting of Figure 18), Swi1, Swi3, and Snf12 in addition to the bait protein Snf5 (Table 1) as well as 8 potentially novel members of *Tetrahymena* SWI/SNF; one protein that contains a PHD domain (TTHERM_00241840), tetrin, four Tet-specific proteins, one ciliate-specific and interestingly one that contains a BD. PHD domain containing proteins are not present in yeast or human SWI/SNF complexes. One function of PHD domains is to mediate specific interactions with methylated lysine on histone proteins to positively regulate transcription (Shi et al., 2006). In this specific case the found PHD domain is conserved with that of yeast JHD2 which is a HDMT (Table 4). Four Tet-specific proteins with no predicted functions were found and they may be acting as species-specific members of the SWI/SNF complex in *Tetrahymena thermophila*. A ciliate-specific protein (TTHERM_00092790) was found with no predicted function and it may be a conserved member of SWI/SNF complex in ciliates. Interestingly, tetrin (TTHERM_00006320) is an insoluble cytoskeletal protein unique to cilia (Williams & Hontst, 1995) and was present in our soluble extraction. Another interesting result is that a bromodomain-containing protein (TTHERM_00729230) co-purifies with Snf5-FZZ. Recall that an earlier functional analysis of *Tetrahymena* Brg1 (Fillingham et al. 2006) (Fillingham et al., 2006) revealed that unlike yeast and human version it does not possess a C-terminal bromodomain. The presence of the BD containing protein is consistent with a SWI/SNF complex in *Tetrahymena* interacting with acetylated histones similar to SWI/SNF in yeast and human cells. Next I decided to perform a reciprocal purification with a member of Table 4 with the intent to confirm the identity of members of the SWI/SNF complex in *Tetrahymena*

Of particular interest was the BD-containing protein, TTHERM_00729230. As previously mentioned, bromodomains mediate specific physical interaction with acetyl-lysine on

histones (Dhalluin et al., 1999). In most eukaryotes the catalytic subunit of SWI/SNF itself possesses a bromodomain (Tang et al., 2010). The catalytic subunit in *Tetrahymena* BRG1/SNF2 is unusual in that it does not possess a bromodomain (Fillingham et al., 2006).

Table 4: SAINTexpress-curated MS data for 4 replicates of SNF5-FZZ affinity purification during vegetative growth.
MS data for SNF5-FZZ reveals predicted and novel members of the putative *Tetrahymena* SWI/SNF complex. BOLD: Predicted in Fillingham et al., 2006 (Fillingham et al., 2006). *Italic*: Novel. Only true interactions that pass a set high cut-off value were included in this table. SpecSum represents the abundance of each protein.

Protein	Gene Name	SpecSum	Yeast
Swi3	**TTHERM_00584840**	**338**	**Swi3**
Tet-specific	*TTHERM_00346460*	*284*	---
Brg1	**TTHERM_01245640**	**256**	**Snf2/Sth1**
Swi1	**TTHERM_00243900**	**154**	**Swi1**
Snf5 (BAIT)	**TTHERM_00304150**	**143**	**Snf5**
Tet-specific	*TTHERM_00129650*	*138*	---
PHD domain	*TTHERM_00241840*	*120*	*JHD2*
Snf12	**TTHERM_00925560**	**78**	**Snf12**
tetrin	*TTHERM_00006320*	*76*	---
ciliate-specific	*TTHERM_00092790*	*70*	---
Tet-specific	*TTHERM_00657560*	*56*	---
Tet-specific	*TTHERM_00637690*	*43*	---
BD protein	*TTHERM_00729230*	*40*	---

4.4 Molecular cloning and generation of FZZ tagged TTHERM_00729230 (Ibd1)

I therefore generated transformants of *Tetrahymena* expressing FZZ-tagged BD-containing protein (TTHERM_00729230) from its WT chromosomal locus. The *Tetrahymena* BD protein was chosen for epitope tagging since it is conserved with the Sth1 bromodomain in yeast with a low BLASTP E-value of $3E^{-5}$. This E-value suggests homology between the Sth1 bromodomain and the bromodomain of the BD protein. I used the molecular cloning and transformation methods described in chapter 1 to FZZ tag the BD protein (TTHERM_00729230) in order to

confirm reciprocal AP-MS interactions with SNF5. The rationale behind reciprocal purifications is to validate the identity of the SWI/SNF complex through co-purification via an additional subunit than Snf5, and possibly identify new co-purifying proteins. A Snf5-FZZ co-purifying protein, the bromodomain protein, was selected for FZZ tagging because of our interest in the lost bromodomain of Brg1 as mentioned above. Using the described molecular cloning methods I obtained 5 paromomycin resistant (pm-r) transformants each of which should express FZZ tagged TTHERM_00729230

4.5 Assessment to determine if pm-r transformants are expressing the tagged protein (BD-FZZ) at the expected size

I used TCA precipitation to generate a whole cell extract which I separated on SDS-PAGE and used Western blotting analysis to assess whether the 5 transformed pm-r cells are expressing BD-FZZ. Of the 5 pm-r transformants, three were independent transformants in B2086 strain; BD-FZZ B1, BD-FZZ B2, and BD-FZZ B3, and two from CU428 cells; BD-FZZ C1 and BD-FZZ C2. Western blotting of whole cell extracts using the M2 antibody shows that only the SNF5-FZZ strain is expressing a protein of about 50kDa which corresponds to 18kDa from the FZZ and 32kDa from the BD protein (Figure 19). Western blotting using the Brg1 antibody shows that both the WT strain and the BD-FZZ strains are expressing Brg1 (Figure 19, loading control). Therefore, I successfully engineered two new BD-FZZ strains with two different mating types.

4.6 Analysis of Ibd1 interactions during vegetative growth

In order to determine the expression pattern of BD-FZZ in the *Tetrahymena* life cycle I affinity purified Snf5-FZZ along with an y co-purifying proteins by performing one-step affinity chromatography on whole cell extracts made from cells in vegetative growth expressing BD-FZZ (I used BD-FZZ B3 from (Figure 19).

Western blotting using the Brg1 antibody (Figure 19) shows that both the WT and the BD-FZZ have Brg1; however, like my earlier experiments, degradation of this protein is present

(showed with arrows in Figure 19). Western blotting using the M2 antibody shows that only the BD-FZZ strain is expressing a protein of about 50kDa corresponding to BD-FZZ, as expected (Figure 19). WT represents the mock IP from untagged cells. Western blotting using the Brg1 antibody shows that the BD-FZZ but not the mock IP from untagged cells is co-purifying with Brg1, consistent with the BD-containing protein being a member of *Tetrahymena* SWI/SNF as expected from my Snf5 analysis (Table 1). Degradation of Brg1 is present on the western blot (showed with arrows in Figure 19) as before in Figure 18. From this data I can assume that the BD protein and Brg1 are bona fide members of the SWI/SNF complex and I expect to see Brg1 in SAINTexpress analysis of the affinity purified material (see next section).

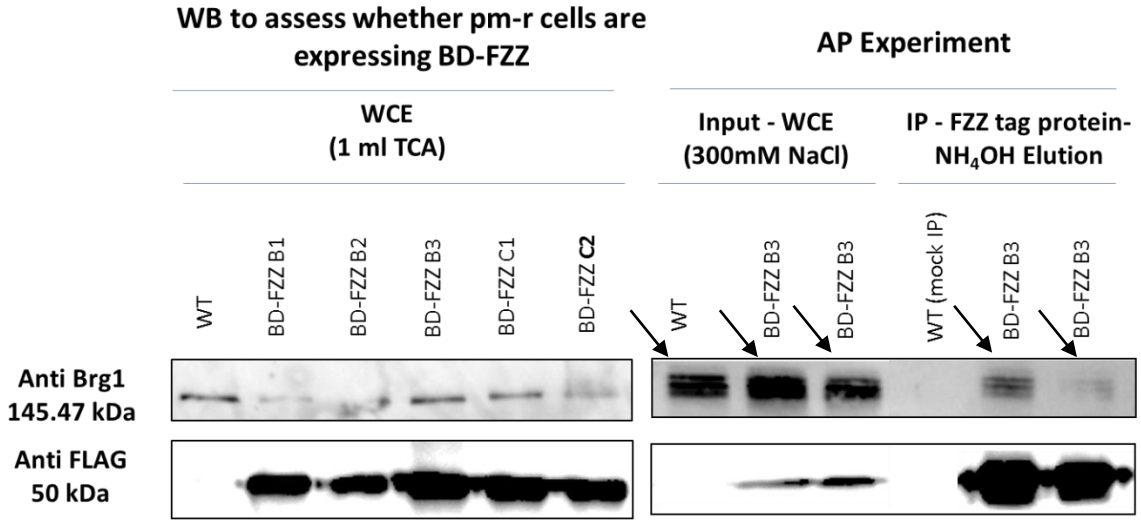


Figure 19: Expression analysis of BD-FZZ during vegetative growth
The BD containing protein is recognized by anti FLAG and co-purifies with Brg1. The arrows point to where degradation of Brg1 was observed. WT represents the mock IP from untagged cells.

4.7 SAINTexpress analysis of MS data from affinity purified BD-FZZ during vegetative growth

The SAINTexpress analysis of MS data from the affinity purified BD protein is shown in Table 5. This data suggests that BD protein is a bona-fide subunit of the SWI/SNF complex, in addition to other chromatin remodeling complexes, such as, a putative *Tetrahymena* SWR and SAGA and a HMT (Table 5, and see below). For this reason I renamed the BD protein

Interactive **B**romo**d**omain Protein **1** (Ibd1) since it interacts with multiple transcription associated complexes. From the SAINTexpress curated MS data, I also found different types of actin. Actin related proteins are found in several ATP dependent chromatin remodeling complexes, although their precise function is unclear (Schubert et al., 2013). The presence of actin during this purification process suggests that the co-purifying subunits are related to chromatin remodeling.

Table 5: SAINTexpress-curated data for 3 replicates of BD-FZZ during vegetative growth
The SpecSum represents the abundance of each protein. The found proteins were labeled as follows (SWI/SNF-C), (Putative SWR-C), (Putative HMT), (Putative SAGA-C), and (undetermined) based on their predicted function.

Protein	Gene Name	SpecSum
BD protein (BAIT)	TTHERM_00729230	1348
Swi3 (SWI/SNF-C)	TTHERM_00584840	495
RvB2 (Putative SWR-C)	TTHERM_00046920	372
RvB1 (Putative SWR-C)	TTHERM_00476820	216
Brg1 (SWI/SNF-C)	TTHERM_01245640	176
Atxr3/Set1 (Putative HMT)	TTHERM_00486440	154
Swr1 (Putative SWR-C)	TTHERM_01546860	134
Actin (Putative SWR-C)	TTHERM_00975380	114
Snf12 (SWI/SNF-C)	TTHERM_00925560	90
Swi1 (SWI/SNF-C)	TTHERM_00243900	80
PHD finger (undetermined)	TTHERM_00444700	67
Actin (Putative SWR-C)	TTHERM_00317000	52
Snf5 (SWI/SNF-C)	TTHERM_00304150	46
Arp6 (Putative SWR-C)	TTHERM_01005190	45
ciliate-specific (SWI/SNF-C)	TTHERM_00092790	44
Gcn5 (Putative SAGA-C)	TTHERM_00248390	44
Hypothetical protein (undetermined)	TTHERM_00170260	34
Tet-specific (SWI/SNF-C)	TTHERM_00346460	33
Swc4 (Putative SWR-C)	TTHERM_00357110	33

ciliate-specific (undetermined)	TTHERM_00136450	29
PHD domain (SWI/SNF-C)	TTHERM_00241840	29
Tet-specific (SWI/SNF-C)	TTHERM_00637690	27
AT-hook containing (undetermined)	TTHERM_00355040	23
Tet-specific (SWI/SNF-C)	TTHERM_00129650	22
Ada2 (Putative SAGA-C)	TTHERM_00790730	21
Swc2-like (Putative SWR-C)	TTHERM_00388500	18
Yaf9 (Putative SWR-C)	TTHERM_00561450	13
Hypothetical protein (undetermined)	TTHERM_00046150	4

4.8 The first ciliate SWI/SNF complex

Through AP-MS of *Tetrahymena* Snf5, and Snf5-interacting Ibd1, I have identified the first protist SWI/SNF complex, based on the intersection of the 2 sets of co-purifying proteins (Refer to Figure 20 for Venn diagram intersections). The *Tetrahymena thermophila* SWI/SNF complex contains 11 subunits (Table 6). It has a catalytic subunit (Brg1) and a standalone bromodomain containing protein (Ibd1). This SWI/SNF complex also has well characterized orthologues subunits present in yeast (Swi1, Swi3, Snf5, and Snf12), a ciliate specific protein (TTHERM_00092790), 3 *Tetrahymena* specific proteins (TTHERM_00346460, TTHERM_00129650 and TTHERM_00637690) and a PHD domain containing protein (TTHERM_00241840). PHD domain containing proteins are not known members of the SWI/SNF complex in other organisms. One function of PHD domains is to mediate specific interactions with methylated lysine on histone proteins to positively regulate transcription (Shi et al., 2006). In this specific case the found PHD domain is conserved with the yeast JHD2 (HDMT). In conclusion, perhaps *Tetrahymena* SWI/SNF has two proteins that recognize PTM on histones, the PHD domain and Ibd1. These two proteins could work in conjunction to positively regulate transcription at an increased rate.

Table 6: *Tetrahymena thermophila* SWI/SNF complex

Protein	Gene	Yeast
Brg1 (catalytic subunit)	TTHERM_01245640	Snf2/Sth1
Ibd1 (Interactive Bromodomain Protein 1)	TTHERM_00729230	Sth1 bromodomain
Snf5	TTHERM_00304150	Snf5
Swi3	TTHERM_00584840	Swi3
Swi1	TTHERM_00243900	Swi1
Snf12	TTHERM_00925560	Snf12
Ciliate-specific	TTHERM_00092790	---
Tet-specific	TTHERM_00346460	---
Tet-specific	TTHERM_00129650	---
Tet-specific	TTHERM_00637690	---
PHD domain	TTHERM_00241840	JHD2

4.9 Putative SWR and SAGA complexes

In addition to the SWI/SNF complex and from data from Table 5, I am beginning to establish members of *Tetrahymena thermophila* putative SWR (Table 7) and SAGA (Table 8) complexes.

Table 7. *Tetrahymena thermophila* putative SWR complex

Protein	Gene
RvB2	TTHERM_00046920
Ibd1	TTHERM_00729230
RvB1	TTHERM_00476820
Swr1	TTHERM_01546860
Actin	TTHERM_00975380
Actin	TTHERM_00317000
Arp6	TTHERM_01005190
Swc4	TTHERM_00357110
Swc2-like	TTHERM_00388500
Yaf9	TTHERM_00561450

Table 8. *Tetrahymena thermophila* putative SAGA complex

Protein	Gene
Gcn5	TTHERM_00248390
Ibd1	TTHERM_00729230
Ada2	TTHERM_00790730

4.10 During vegetative growth Ibd1 interacts with members of chromatin remodeling complexes related to transcription

From the data obtained from the SAINTexpress curated MS analysis of Ibd1-FZZ AP, individual proteins were identified and assigned to a protein complex (SWR or SAGA) based on orthologue proteins present in other organisms. SWI/SNF proteins identified as explained above in section 4.8. I predicted that Ibd1 is involved in several independent protein complexes and all these complexes form a mega complex to increase transcription. The Ibd1 interacting proteins were grouped in predicted complexes using Venn diagrams (Figure 20). I grouped the members of each transcription related complex that interacts with Ibd1. The complexes, SWI/SNF, SWR-C, and SAGA related subunits and a HMT protein, are present in yeast and may be homologous in *Tetrahymena*. Ibd1 also interacts with other proteins that were classified in the group "Ibd1 (other interactions)" showed in Figure 20. These proteins are two hypothetical proteins (TTHERM_00046150 and TTHERM_00170260), a PHD Finger (TTHERM_00444700, not the same as SWI/SNF PHD domain; TTHERM_00241840), an AT-hook containing (TTHERM_00355040) and a ciliate-specific (TTHERM_00136450 not the same as SWI/SNF ciliate specific; TTHERM_00092790).

In conclusion, Ibd1 is shared with multiple protein complexes (SAGA, SWR, and SWI/SNF) and possibly a HMT protein all of which share common function in chromatin remodeling and transcription.

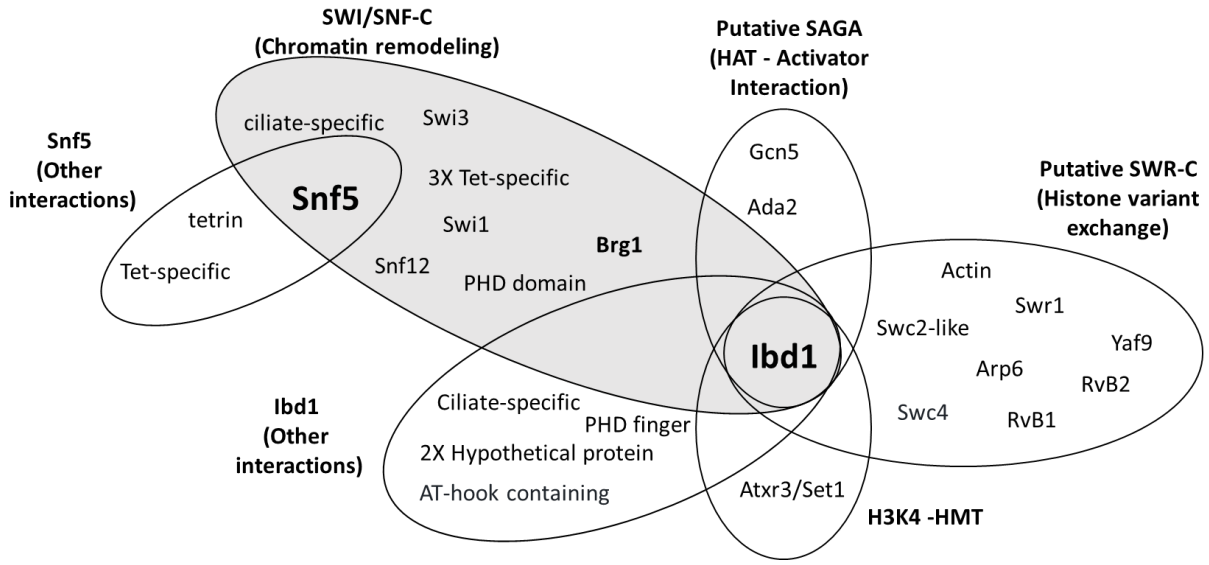


Figure 20: Comprehensive analysis of the Ibd1 interactome during vegetative growth.
The *Tetrahymena* SWI/SNF-C was found based on the intersection of the two sets of co-purifying proteins Snf5 and Ibd1. In addition, Ibd1 interacts with members of other chromatin remodeling complexes, such as, a putative SWR-C and SAGA and a HMT. Ibd1 also interacts with proteins that cannot be grouped in any of the mentioned complexes at this time.

4.11 Expression analysis of BD-FZZ during *T.thermophila*'s life cycle

The purpose of this experiment was to corroborate the Ibd1 protein expression profile with the mRNA expression profile publically available from ciliate.org which indicates that at the mRNA level this gene expresses throughout *Tetrahymena*'s life cycle. It also will give us a good reference whether Ibd1 expresses at different hours related to specific key events during conjugation such as vegetative growth, starvation, meiosis (3H), gametic nuclei fusion (4.5H), zygotic nuclei duplication (6H), anlagen development (8H) and when the daughter cells have formed (24H).

I used western blotting of TCA-generated whole cell extracts to perform a time course expression analysis of Ibd1. I used Brg1 as a loading control for expression throughout *T.thermophila*'s life cycle, and Pdd1 strictly as a developmental loading control. With these results, it was evident that Ibd1 expresses throughout *T.thermophila*'s life cycle, and it might be stable enough to be present at 24H after mating when the new MAC is active.

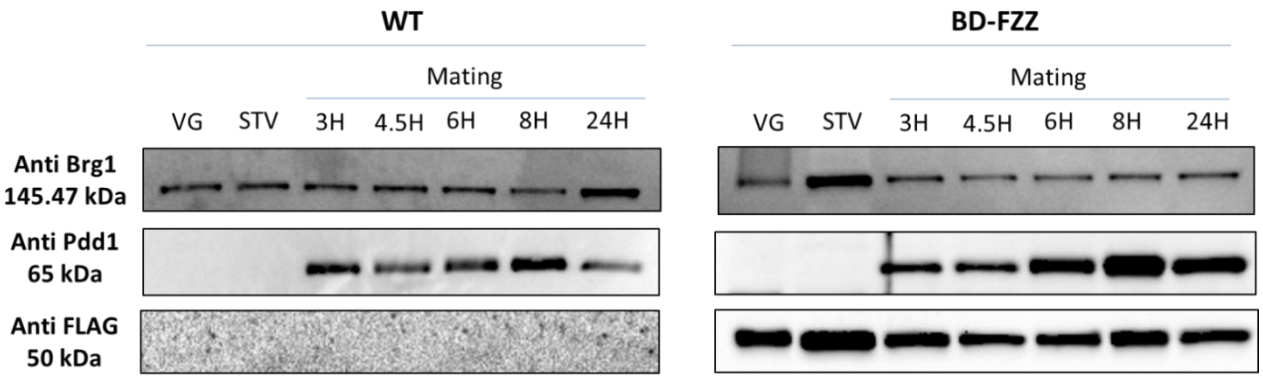


Figure 21: Expression analysis of Ibd1-FZZ during *T.thermophila*'s life cycle
Ibd1 (Anti FLAG) expresses throughout *T.thermophila*'s life cycle. Brg1 is a loading control for expression throughout *T.thermophila*'s life cycle and Pdd1 is strictly a developmental loading control.

4.12 Analysis of Ibd1 interactions during conjugation

In order to identify the set of proteins that co-purify with Ibd1-FZZ I performed one-step affinity chromatography on whole cell extracts made from cells expressing Ibd1-FZZ (I used Ibd1-FZZ B3 x BD-FZZ C1) and harvested 5 hours after the initiation of conjugation.

Western blotting of affinity purified Ibd1-FZZ using M2 antibody shows that only the BD-FZZ cells mating are expressing a protein of about 50kDa corresponding to BD-FZZ (Figure 22). I used Pdd1p antibody as developmental control and its signal was not present for this blot in the WCE, perhaps because it is not soluble (data not shown). Even though I cannot observe Pdd1p in the WB analysis, I know that the cells are conjugating based on microscope visualization. For the Brg1 antibody Figure 22 shows that both the WT cells mating and the BD-FZZ cells mating have degradation of Brg1.

Western blotting of affinity purified material using the M2 antibody shows that only the BD-FZZ cells mating are expressing a protein of about 50kDa corresponding to BD-FZZ (Figure 22). WT represents the mock IP. Western blotting using the Brg1 antibody (Figure 22) shows that BD-FZZ cells mating are not co-purifying with Brg1. This is different than what occurs during vegetative growth (Figure 19). From this result it seems possible that during conjugation

Ibd1 does not co-purify with Brg1. Perhaps during conjugation Idb1 does not interact with SWI/SNF but remains bound to other complexes such as SWR, SAGA and HMT and possibly new ones. Alternatively, it is possible that during conjugation Brg1 co-purifies with Ibd1 but the amount of it is reduced to a level that is difficult to detect using western blotting. To answer this, next I equilibrated the relative amount of Ibd1 to be loaded in SDS-PAGE from the samples taken during vegetative growth and conjugation (see below). In addition, MS will provide further data (see below).

.

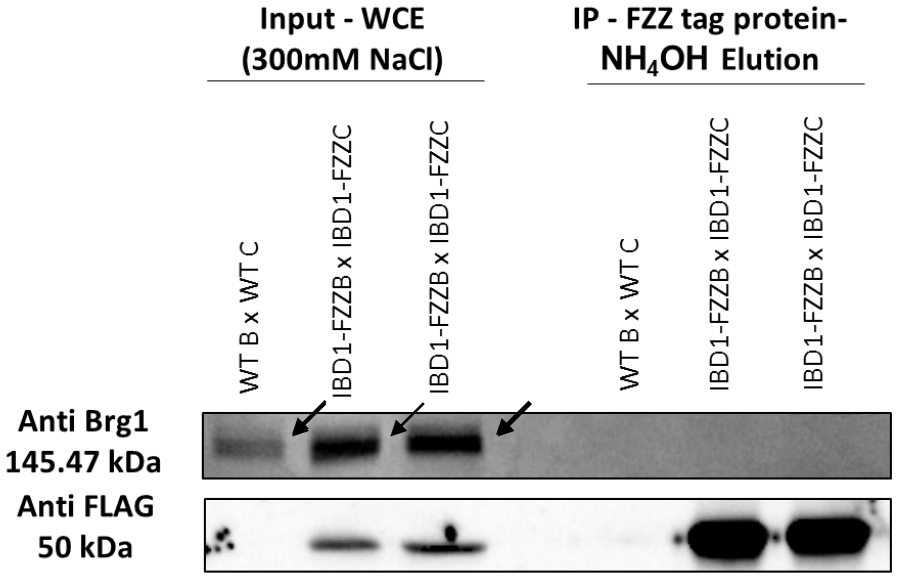


Figure 22: Expression analysis of BD-FZZ 5 hours after conjugation.
In the IP samples taken during conjugation Brg1 cannot be detected. Pdd1p cannot be detected (data not shown). The arrows point to where degradation of Brg1 was observed. Ibd1 is recognized by anti FLAG.

4.13 SAINTexpress curation of MS data from AP of Ibd1-FZZ during conjugation

Conversely to WB analysis, SAINTexpress analysis of BD-FZZ cells 5 hours after initiation of mating shows that Ibd1 still interacts with Brg1. Ibd1 interacts with less proteins than during vegetative growth; however, all the interacting proteins are the same than during vegetative growth and there are not new interacting proteins (Table 9). At this stage Ibd1 still interacts with subunits of SWI/SNF, putative SWR-C, HMT and putative SAGA. With this data,

I can now answer the question, “What is the composition of Ibd1 protein-protein interactions during conjugation?”

Table 9: SAINTexpress-curated data for 2 replicates of BD-FZZ 5 hours after mating.

Protein	Gene Name	SpecSum
BD protein (BAIT)	TTHERM_00729230	738
RvB2 (Putative SWR-C)	TTHERM_00046920	263
RvB1 (Putative SWR-C)	TTHERM_00476820	170
Swr1 (Putative SWR-C)	TTHERM_01546860	137
Actin (Putative SWR-C)	TTHERM_00975380	106
Swi3 (SWI/SNF-C)	TTHERM_00584840	100
PHD Finger (SWI/SNF-C)	TTHERM_00444700	57
Gcn5 (Putative SAGA-C)	TTHERM_00248390	41
Arp6 (Putative SWR-C)	TTHERM_01005190	41
AT-hook containing (undetermined)	TTHERM_00355040	35
Swc4 (Putative SWR-C)	TTHERM_00357110	32
ciliate-specific (undetermined)	TTHERM_00136450	26
Swc2-like (Putative SWR-C)	TTHERM_00388500	22
Atxr3/Set1 (Putative HMT)	TTHERM_00486440	22
Ada2 (Putative SAGA-C)	TTHERM_00790730	19
Hypothetical protein (undetermined)	TTHERM_00046150	16
Snf12 (SWI/SNF-C)	TTHERM_00925560	13
Brg1 (SWI/SNF-C)	TTHERM_01245640	11
ciliate specific (SWI/SNF-C)	TTHERM_00092790	7
Swi1 (SWI/SNF-C)	TTHERM_00243900	6

4.14 Ibd1 modifies its interactions at 5 hours after conjugation and during vegetative growth

From an inspection of Table 7, Ibd1 still interacts with SWI/SNF, putative SWR-C, HMT and putative SAGA sub-units 5 hours after initiation of conjugation. At this stage Ibd1 interacts with less proteins than during vegetative growth; however, all the interacting proteins are the same than during vegetative growth and there are not new interacting proteins To see if there are any changes in the relative amount of Ibd1 interacting with these respective protein complexes, I set the SpecSum values for the baits (BD-FZZ) in Table 5 and Table 9 to 1, allowing me to standardize the remaining values in each table in relation to the corresponding bait. I then compared the trends of the two tables in Figure 23 (data analysis in Table 11, Appendix 2) and observed in both that there is a decrease of all SWI/SNF subunits (-76.51% on average); accounting for why Brg1 (-88.58%) could not be detected in the SDS-PAGE (Figure 24). Another protein that decreases in conjugation in the amount of Ibd1 bound is the HMT (-73.91%).

Not all proteins decrease: there is an increase of all putative SWR-C (65.67% on average) and putative SAGA (67.73% on average) subunits present in this purification. Perhaps Ibd1 has an important role in histone variant exchange and HAT related proteins at 5 hours after initiation of conjugation and not as a member of the SWI/SNF complex or HMT.

Interestingly, this analysis also shows that the proteins that are in the group “Ibd1 (other interactions)” showed in Figure 20 such as hypothetical protein (TTHERM_00046150, not a member of the SWI/SNF complex) drastically increases in percentage bound to Ibd1 (630%) during conjugation. The protein PHD Finger (TTHERM_00444700, not the same as SWI/SNF PHD domain; TTHERM_00241840), an AT-hook containing (TTHERM_00355040) and a ciliate-specific (TTHERM_00136450 not the same as SWI/SNF ciliate specific; TTHERM_00092790) increase 55.39%, 177.95% and 63.76%, respectively (Figure 23). Perhaps all the proteins that are increasing in amount during conjugation are important for this stage. Calculations are in the Appendix 2.

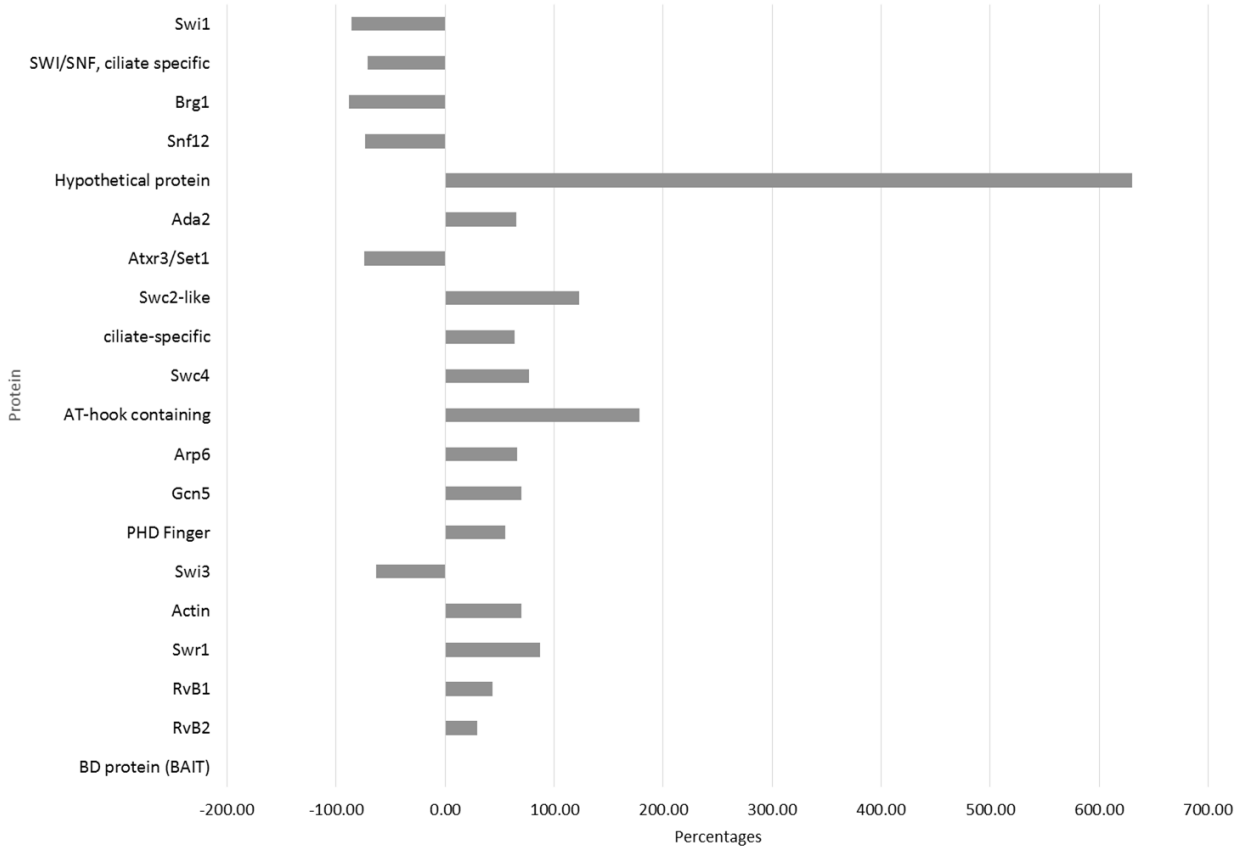


Figure 23: Relative proportion of encountered proteins during vegetative growth and 5 hours after initiation of conjugation
For percentages, if % > 0 there is more protein interacting with Ibd1 in conjugation or if % < 0 there is less protein interacting with Ibd1 during conjugation compared to during vegetative growth.

4.15 SDS-PAGE to equilibrate the relative amount of Ibd1

In Figure 19, we can observe that Ibd1 co-purify with Brg1 during vegetative growth and during conjugation Figure 22 shows that Brg1 does not co-purify with Brg1. I initially hypothesized that at 5 hours after initiation of conjugation there is: 1] a decrease of Brg1 along with a decrease of the SWI/SNF complex subunits; or 2] total absence of the Brg1 subunit alone; or 3] total absence of the SWI/SNF complex. AP-MS showed that there is reduced presence of Brg1 during conjugation compared to vegetative growth, suggesting that during conjugation the SWI/SNF complex decreases in abundance (Figure 23). Thus, probably the decrease of Brg1 influences the sensitivity of WB analysis. We also have to consider that for the previous WB

shown in Figure 19 and Figure 22 the IP loaded samples were not representing equivalent expression of Ibd1. In order to visualize equivalent expression of Ibd1 with respect to Brg1, it was necessary to increase the amount of IP sample added from conjugation and reduce the amount added from vegetative growth. In conclusion, I found that in IP samples the relative amount of interacting Brg1 to Ibd1 is less, as is shown in Figure 24. This is consistent with the previous result in Figure 23.

4.16 Pdd1p is insoluble

As seen in Figure 21, I ran samples belonging to TCA precipitations throughout *Tetrahymena thermophila*'s life cycle. In this WB, Pdd1p can be observed. However, as seen in Figure 22, I ran samples belonging to soluble extracts taken during conjugation. It is evident here that Pdd1p cannot be observed and I hypothesized that Pdd1p is insoluble in our extraction buffer that we use for generation of WCE for AP. I used fresh TCA extracts and compared them to soluble extracts. This can be seen in Figure 24, where anti-Pdd1 recognizes the TCA WCE and not the input WCE. This confirms that Pdd1 is insoluble under our AP conditions.

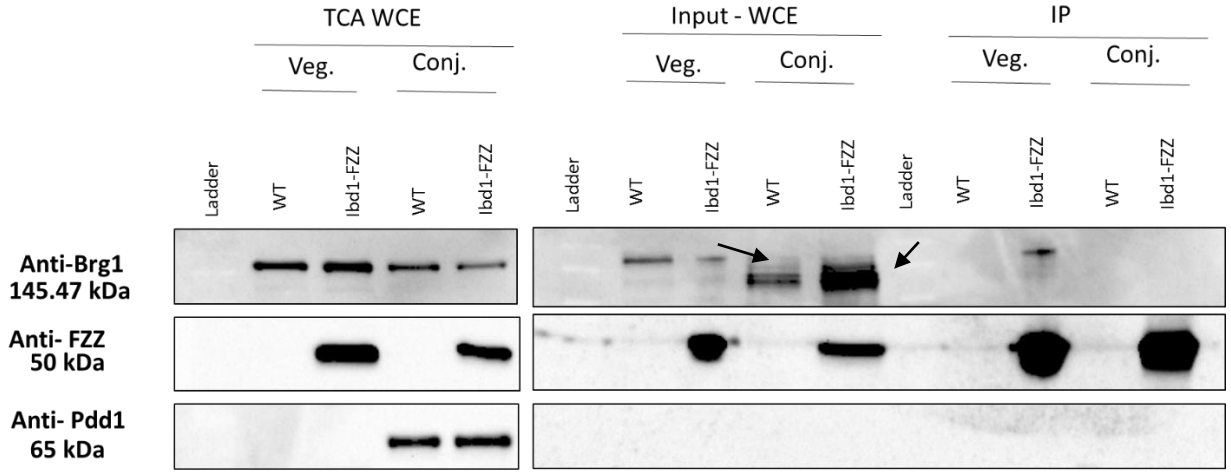


Figure 24: SDS-PAGE to equilibrate the relative amount of Ibd1 and Pdd1p solubility.
In IP samples the relative amount of interacting Brg1 to Ibd1 (Anti-FZZ) is less. The arrows point to where degradation of Brg1 was observed. Anti-Pdd1 recognizes the TCA WCE and not the input WCE confirming that Pdd1 is insoluble.

4.17 Ibd1 recognizes multiple rather than single PTM

An important question raised by the previous analyses is whether Ibd1's bromodomain is functional. My next objective was to determine if the bromodomain of Ibd1 was functional. In order to do so, I engineered a synthetic Ibd1 gene to express a recombinant 6XHis-Ibd1 in *E.coli* BL21 (D3). The purified recombinant protein was used in a commercially available histone peptide array to assess Ibd1's binding interactions with various PTM. The array contained 384 sites with single and combinations of PTM. The principle behind use of this array is described in the materials and methods section.

4.17.1 Strain generation

In order to begin to understand the function of Ibd1 and to determine if its bromodomain recognizes a specific histone post-translational modification I first engineered a synthetic gene to express a 6XHis-Ibd1 recombinant protein in *E.coli* BL21 (D3).

4.17.1.1 DNA optimization

The first steps was DNA optimization, since *Tetrahymena* uses a non-universal genetic code where UAA and UAG specify glutamine and not stop codons as in *E.coli*. Therefore in order to express a full length Ibd1 protein in *E.coli*, I engineered a synthetic gene to use appropriate codons for glutamic acid. In addition, I took the opportunity to optimize codon usage for expression in *E.coli* BL21. The engineering of the synthetic gene was performed by an external provider (BioBasics Inc.) and provided in a pUC57 plasmid.

4.17.1.2 Amplification of synthetic Ibd1

After DNA optimization of Ibd1, PCR was carried out using the pUC57 plasmid containing the Ibd1 synthetic gene as template and the used primers are described in Appendix 3. To assess success of the PCR, the samples were run in an agarose gel electrophoresis following the methods mentioned above (Figure 25).

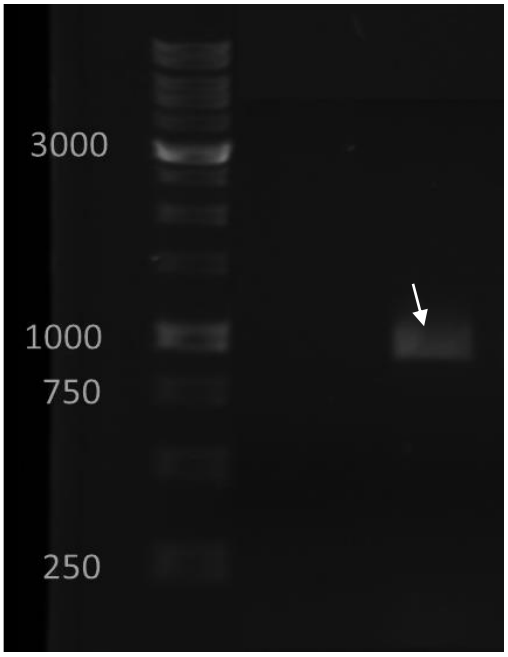


Figure 25: PCR of Synthetic Ibd1.
Arrow shows that the PCR product migrated the expected distance, corresponding to 813bp.

4.17.1.3 Restriction digestion, ligation into expression vector and transformation into *E.coli* DH5α followed by transformation into *E.coli* BL21 (DE3)

The PCR products and a pET28 plasmid (expression plasmid) were double digested with *NdeI* and *BamHI* followed by cleaning and gel purification respectively and ligation. The ligated DNA was transformed into *E.coli* DH5α strains and the success of the cloning was confirmed by Sanger sequencing. The pET28 plasmid containing the Ibd1 synthetic gene was then used for transformation into *E.coli* BL21 (DE3).

4.17.2 Inducing expression of Ibd1 in *E. coli* BL21 (DE3)

The transformed *E.coli* BL21 (DE3) containing were grown in media containing kanamycin +/- IPTG over a period of 4 hours. Whole cell extracts by sonication, SDS loading dye following by incubation in boiling water were made. I took samples at 1 hour time intervals starting from time 0 to 4 hours post induction. I ran the samples against a ladder in an SDS-PAGE. I applied Fast SeeBand Protein stain solution to the gel following electrophoresis (Figure 26). One the stained gel (Figure 25) a band can be seen at the expected size for 6xHIS-Ibd1 that

is not present at Time 0 but is present at the other times indicating that I successfully expressed the protein.

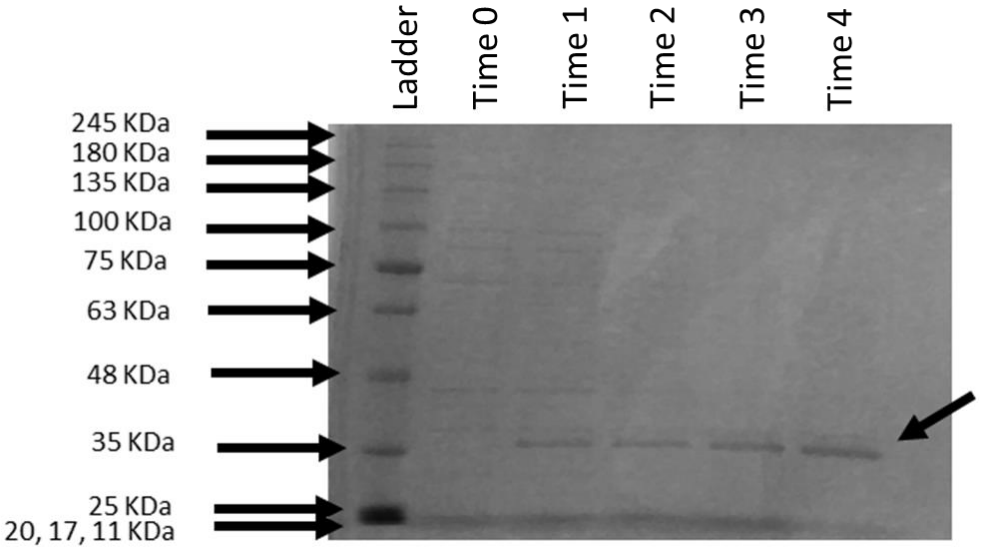


Figure 26: Inducing expression of Ibd1 in *E.coli* BL21 (DE3).
The arrow shows the Ibd1 predicted molecular weight is ~32 kDa (32kDa + 1kDa for 6XHis-Tag)

4.17.3 Affinity Purification of 6XHis-Ibd1 from BL21

Extraction of the 6XHis-Ibd1 was done using Ni-NTA Fast Start Kit (Qiagen). The recombinant protein was purified under native conditions. The input for the kit was 3mL of cell pellet from IPTG induced *E.coli* BL21 (DE3). The extraction kit has many flow through, wash and elution steps which are described in the materials and methods section. I collected 5µL of each flow through, wash and elution fraction in separate tubes. The 5µL aliquots from each fraction were mixed with 5µL of SDS loading dye and stored at -20° C for subsequent analysis with SDS-PAGE. The aliquots from each of the fractions were run in an SDS-PAGE. The gel was then stained with Fast SeeBand Protein stain solution to confirm the presence of 6XHis-Ibd1 in the samples (Figure 27).

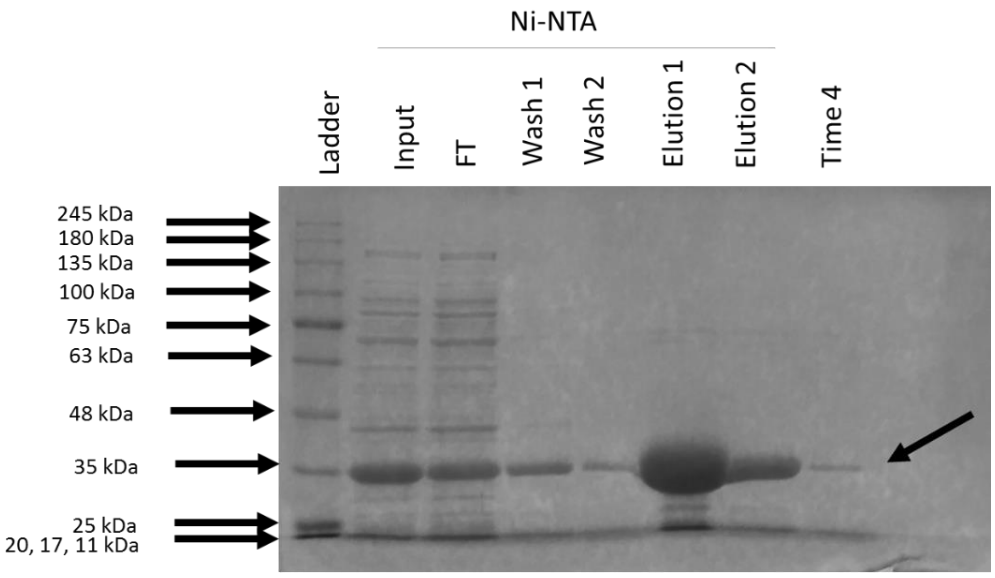


Figure 27: Purification of Expressed 6XHis-Ibd1
The arrow shows the Ibd1 predicted molecular weight is ~32 kDa (32kDa + 1kDa for 6XHis-Tag)

4.17.3.1 Buffer exchange by size exclusion chromatography

To obtain the 6XHis-Ibd1 in a buffer free from imidazole appropriate for downstream applications, I used buffer exchange which is based on size exclusion chromatography. I used the elutions 1 and 2 from the Ni-NTA purification and the new elution was obtained in 8 different fractions. Each of the 8 samples were run in an SDS-PAGE, followed by staining of the gel by Fast SeeBand Protein stain solution.

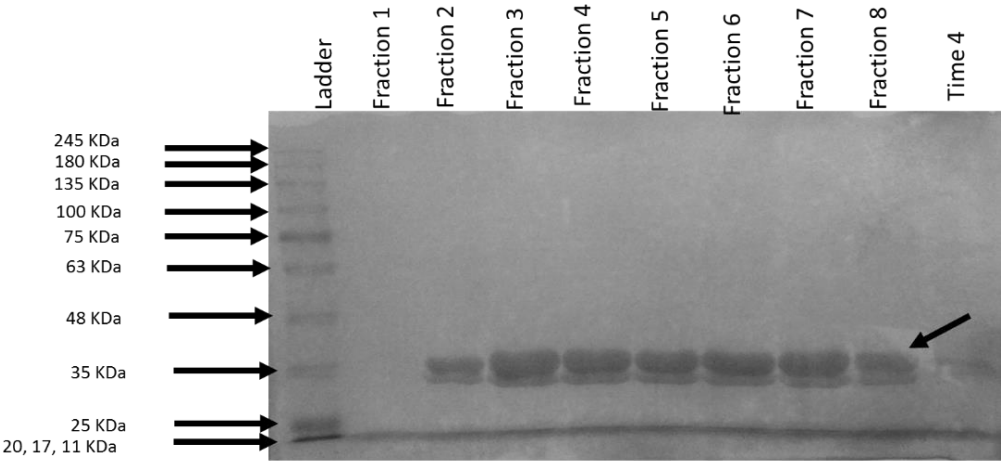


Figure 28: Buffer exchange by size exclusion chromatography
The arrow shows the Ibd1 predicted molecular weight is ~32 kDa (32kDa + 1kDa for 6XHis-Tag)

4.17.4 Functional characterization of Ibd1

I used 200nM final concentration of the 6XHis-Ibd1 recombinant protein and incubated it with the array in the reaction buffer provided by the kit. This was followed by blocking, anti-His incubation, HRP conjugated secondary antibody, ECL, washes and imaging of the array. I then used the Active Motif software provided with the histone peptide array to obtain a detailed analysis of the array sites that delivered the highest signal intensities. A high intensity indicates that Ibd1 has a high affinity for this (these) PTM. The sites in the array contain 384 multiple peptides representing N-terminal domains of human histones with a single or multiple PTM. The software assigned values based on the signal intensities from the raw data (Figure 31) and depending on the specific analysis these values may vary. There are two different analyses provided by the software. First, the specificity analysis of multiple peptide average gives a specificity factor. The specificity factor is calculated by taking the average intensity of all spots containing that PTM divided by the average intensity of all spots not containing that PTM. The ten modifications with the highest specificity factors for sites containing both single and multiple peptides are shown in Figure 32. Second, the single peptide reactivity will measure the impact of neighbouring modifications on binding specificity. The peptide reactivity analyses are shown from Figure 33 to Figure 42. For each of the peptides with highest specificity factors from the specificity analysis of multiple peptide average and the specificity analysis of single peptides, an individual analysis of single peptide reactivity was done to compare the specificity of this peptide alone compared to when it shared a site with other modified peptides. I also used the provided positive control to validate the experiment which expected patterns are in Appendix 5.

4.17.5 Peptide array for the control (His-G9a)

4.17.5.1 Raw data of the peptide array for the control (His-G9a)

The provided control is a recombinant human His-G9a expressed en *E.coli*. His represents a histidine tag. G9a is a histone methyltransferase (via its SET domain) and only recognizes mono- and dimethylated H3K9 sites (via its ankyrin repeats)(R. E. Collins et al.,

2008). As a control for this experiment, I incubated a separate array with the provided His-G9a (10nM final concentration). The His-tag antibody was used to detect the His portion of the His-G9a. G9a only recognizes mono- and dimethylated H3K9 sites. The raw data shows the signal intensities of individual interactions between the His-G9a and the sites (circles) containing mono- and dimethylated H3K9. These sites represent N-terminal domains of human histones with a single or multiple PTM. A visual comparison of signal intensities between the obtained peptide array (Figure 29) and the model control showed in Appendix 5 was made and the layout was virtually identical.

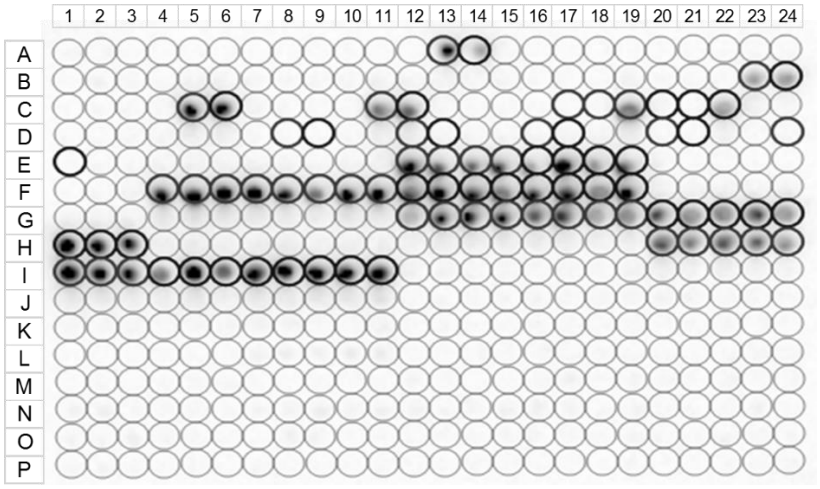


Figure 29. Raw data of the peptide array for the control (His-G9a).
The light circles represent N-terminal domains of human histones with a single or multiple PTM. The dark circles represent mono- and dimethylated H3K9 spots. The array was incubated with His-G9a and detected with the His-tag antibody. The annotations for each site are shown in Appendix 5.

4.17.5.2 Specificity analysis of multiple peptide average for the control (His-G9a)

After visual inspection of the raw data I used the software provided with the kit. The software assigned values based on the signal intensities from the raw data (Figure 28) and depending on the specific analysis these values may vary. The specificity analysis of multiple peptide average gives a specificity factor. The specificity factor is calculated by taking the average intensity of all spots containing that PTM divided by the average intensity of all spots not containing that PTM. The ten modifications with the highest specificity factors for sites

containing both single and multiple peptides are shown in Figure 30. These values cannot be compared to the previous analyses outputted values. This analysis presents two high values: 8.2 for H3K9me2 and 7.6 for H3K9me1. The expected graphical analysis of the control is shown in Appendix 5. Our results match the model graphical analysis for the control thus validating Ibd1 analysis.

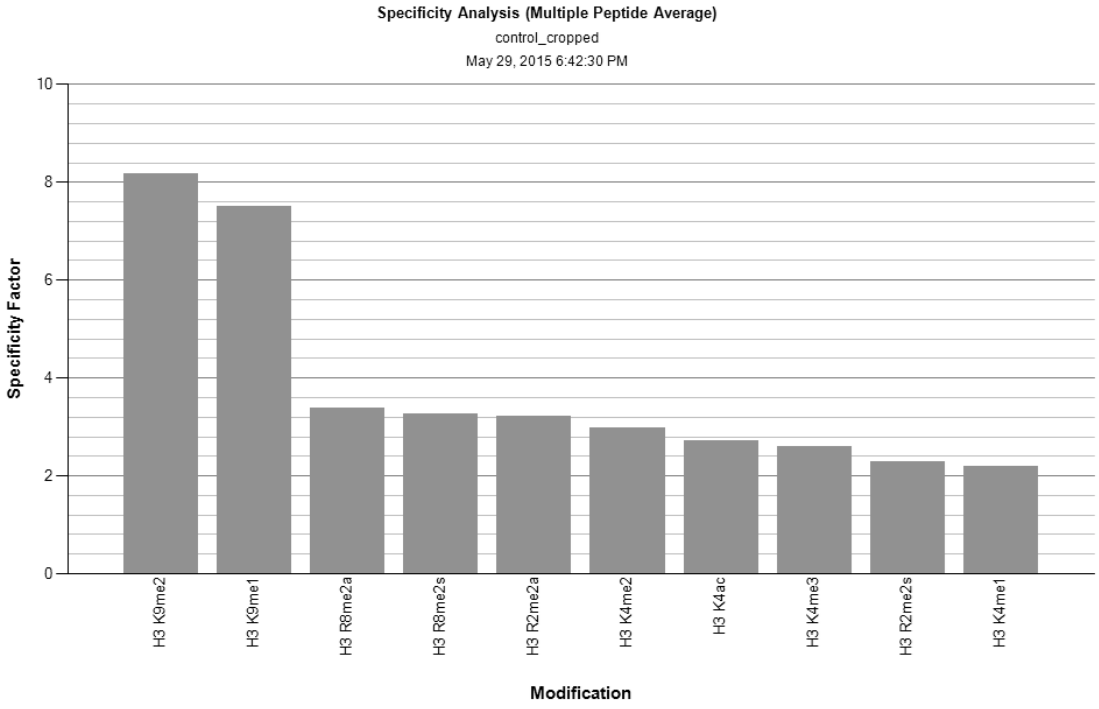


Figure 30: Multiple peptides specificity analysis for the control (His-G9a)
Peptide array used as a control for the array. The array was incubated with His-G9a and detected with the 6xHis-tag antibody.

4.17.6 Peptide array for Ibd1

4.17.6.1 Raw data of the peptide array for Ibd1

Now that we know the assay works, I took the recombinant 6XHis-Ibd1 and tried it in the array. The raw data shows the signal intensities of individual interactions between 6XHis-Ibd1 and the multiple peptides in each site (circles). These sites represent N-terminal domains of human histones with a single or multiple PTM. The software assigned values based on the signal

intensities from the raw data (Figure 31) and depending on the specific analysis these values may vary. Ibd1 did not recognize any methylated peptides like the control protein, rather it recognized almost exclusively acetylated peptides consistent with having a functional BD.

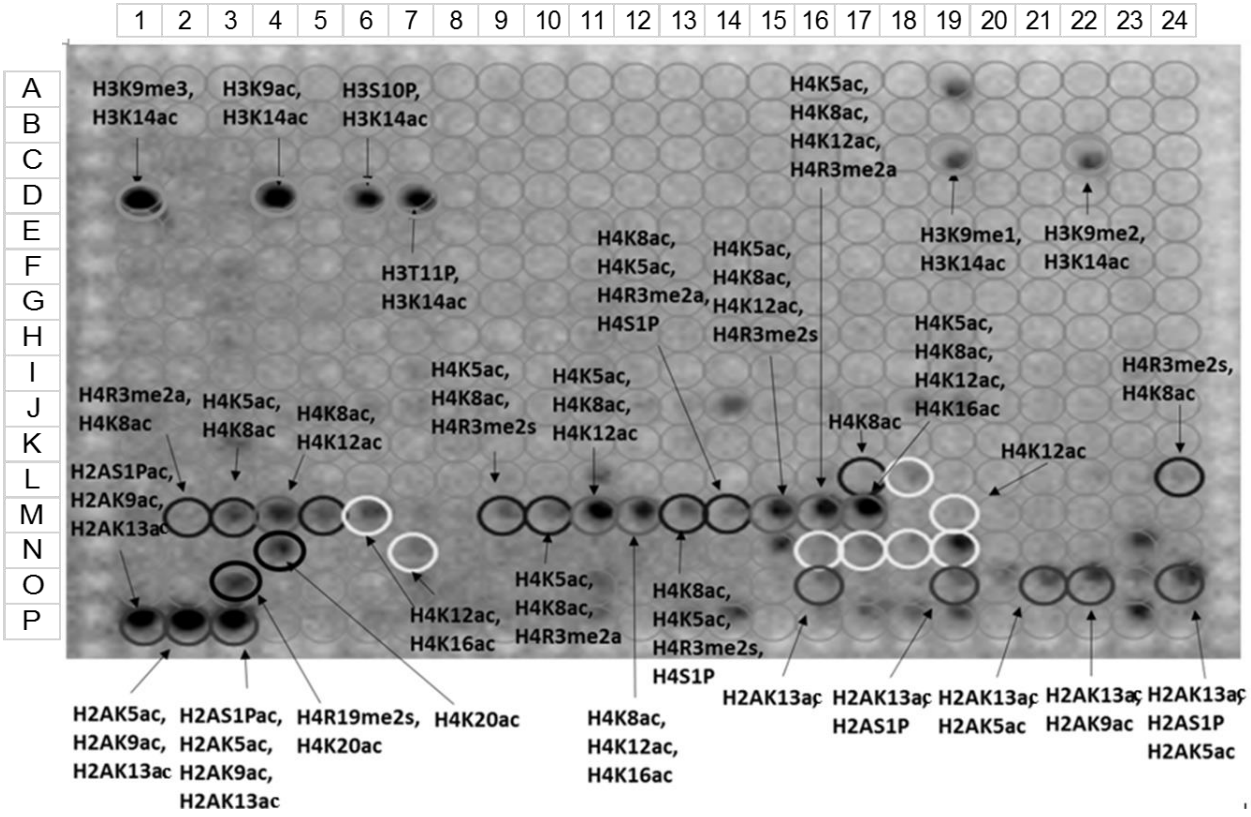


Figure 31: Ibd1 Raw data of the peptide array.
The circles represent N-terminal domains of human histones with a single or multiple PTM. The array was incubated with 6XHis-Ibd1 and detected with the His-tag antibody. The annotations for each site are shown in Appendix 5.

4.17.6.2 Specificity analysis of multiple peptide average for Ibd1

The specificity analysis of multiple peptide average gives a specificity factor. The specificity factor is calculated by taking the average intensity of all spots containing that PTM divided by the average intensity of all spots not containing that PTM. The ten modifications with the highest specificity factors for sites containing both single and multiple peptides are shown in Figure 32. These values cannot be compared to other analyses outputted values since scales are

different. This analysis in Figure 32 presents three high values: 8.8 for hH4K8ac, 8.4 for hH2AK13ac and 8.0 for hH4K12ac. This data suggests that the best candidate combinations to be recognized by Ibd1 include hH4K8ac and hH4K12ac. I am excluding hH2AK13ac since H2A resembles H4 N-terminus. I will next individually analyse each of these modifications using single peptide reactivity analysis.

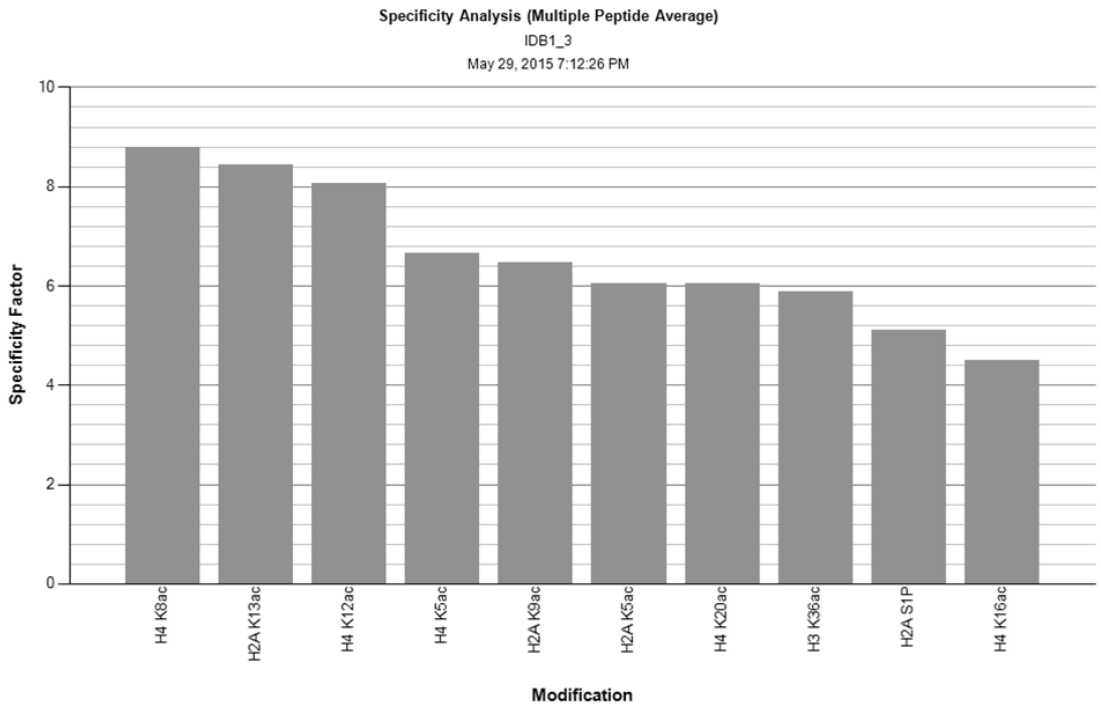


Figure 32: Multiple peptides specificity analysis for Ibd1.
Three high values: 8.8 for hH4K8ac, 8.4 for hH2AK13ac and 8.0 for hH4K12ac. This data suggests that the best candidate combinations to be recognized by Ibd1 include hH4K8ac and hH4K12ac.

4.17.6.3 Single peptide reactivity analysis for Ibd1

I will use the obtained data from the specificity analysis of multiple peptide average to individually analyse each of the outputted PTM using the single peptide reactivity analysis. The single peptide reactivity analysis will measure the impact of neighbouring modifications on binding specificity. This impact can be favorable or not for Ibd1 binding. The peptide reactivity analyses are shown from Figure 33 to Figure 42. The objective is to further understand what marks are being read by the bromodomain of Ibd1.

4.17.6.3.1 Single peptide reactivity analysis for hH4K8ac

The individual analysis of the sites containing only hH4K8ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 33. Ibd1 recognizes the human combination of H4K5ac, H4K8ac, H4K12ac (H4-Triplet) and the addition of human H4K16ac will not improve the recognition (this is depicted in Figure 33-1). (See Figure 5 for alignments). The addition of a human H4R3me2a to the H4-triplet will reduce Ibd1 recognition (*Tetrahymena* does not have this residue) and absence of human H4K12ac will also reduce recognition (Figure 33-2). These data suggest that the absence of any member of the H4-triplet decreases binding specificity (Figure 33-3).

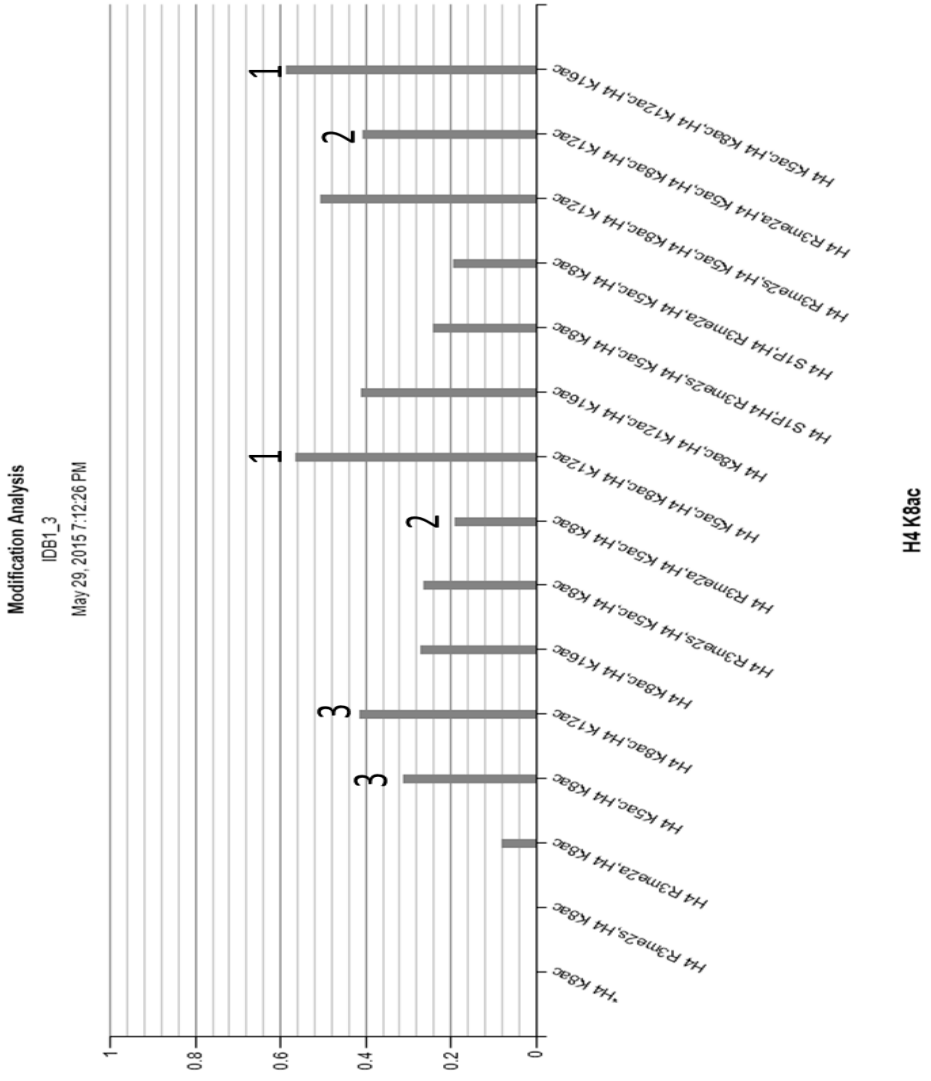


Figure 33: Single peptide reactivity analysis for hH4K8ac. The H4K8ac mark is a posttranslational modification in humans (hH4K8ac). In *Tetrahymena*, this acetylation mark is on the lysine residue at position 7 and is subsequently denoted TtH4K7ac. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

4.17.6.3.2 Single peptide reactivity analysis for hH2AK13ac

The individual analysis of the sites containing only hH2AK13ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 34. Ibd1 also recognizes the human combination of H2AK5ac, H2AK9ac, H2AK13ac with the highest specificity in comparison to other sites with various PTM combinations (Figure 34-1). The addition of H2AS1P slightly decreases the binding specificity of the triplet, however its presence slightly increases the specificity when only 1 or 2 of the marks occupy one site (Figure 34-2).

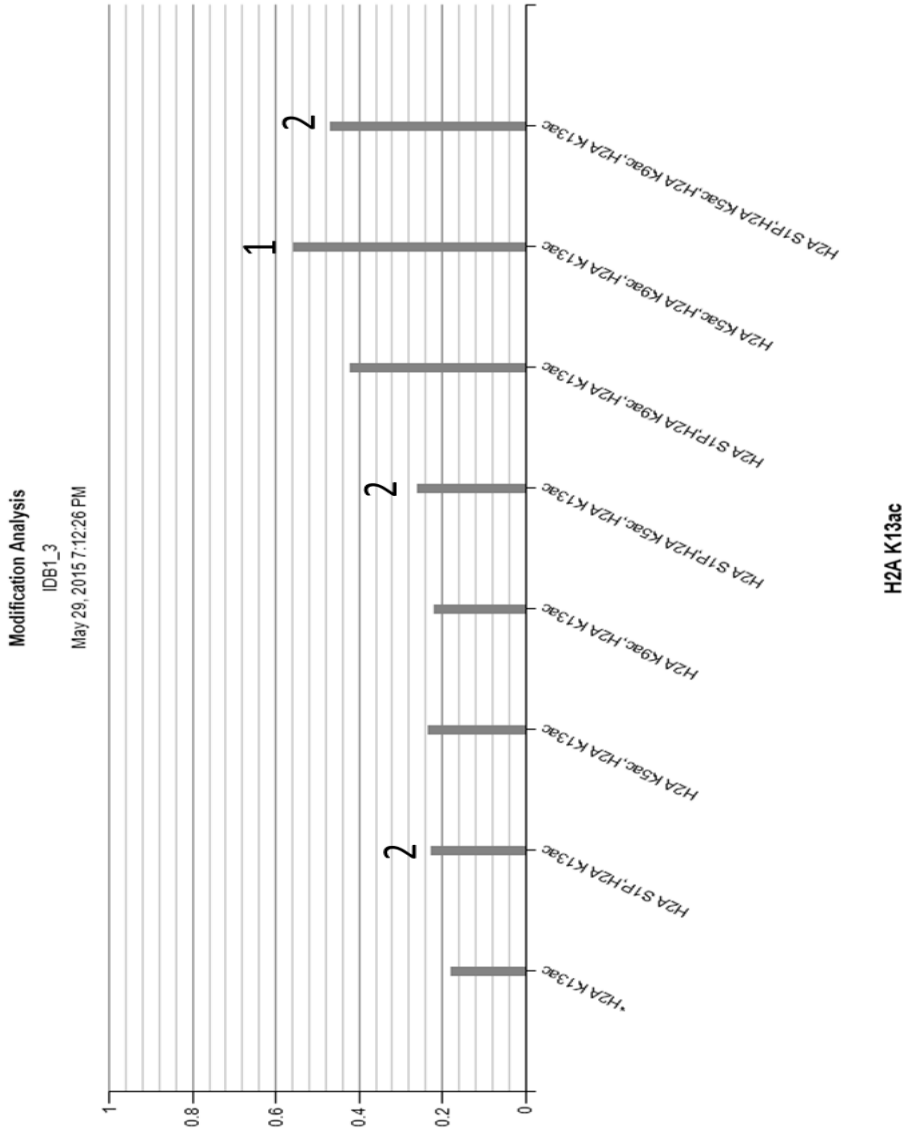


Figure 34: Single peptide reactivity analysis for hH2AK13ac. The H2AK13ac mark is a posttranslational modification in humans (hH2AK13ac). In *Tetrahymena*, this acetylation mark is on the lysine residue at position 12 and is subsequently denoted TtH2AK12ac. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

4.17.6.3.3 Single peptide reactivity analysis for hH4K12ac

The individual analysis of the sites containing only hH4K12ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 35. Ibd1 has the highest binding specificity with triplet and quadruple acetylations (Figure 35-1). Ibd1 has very little recognition of H4K12ac on its own, however its recognition is greatly increased in the presence of another PTM, such as H4K8ac or H4K16ac (Figure 35-2). Mono-, di-, and trimethylations of H4K20 decrease Ibd1's binding specificity to H4K12ac, H4K16ac (Figure 35-3). The H4R3me2 mark seen in humans does not exist in *Tetrahymena*.

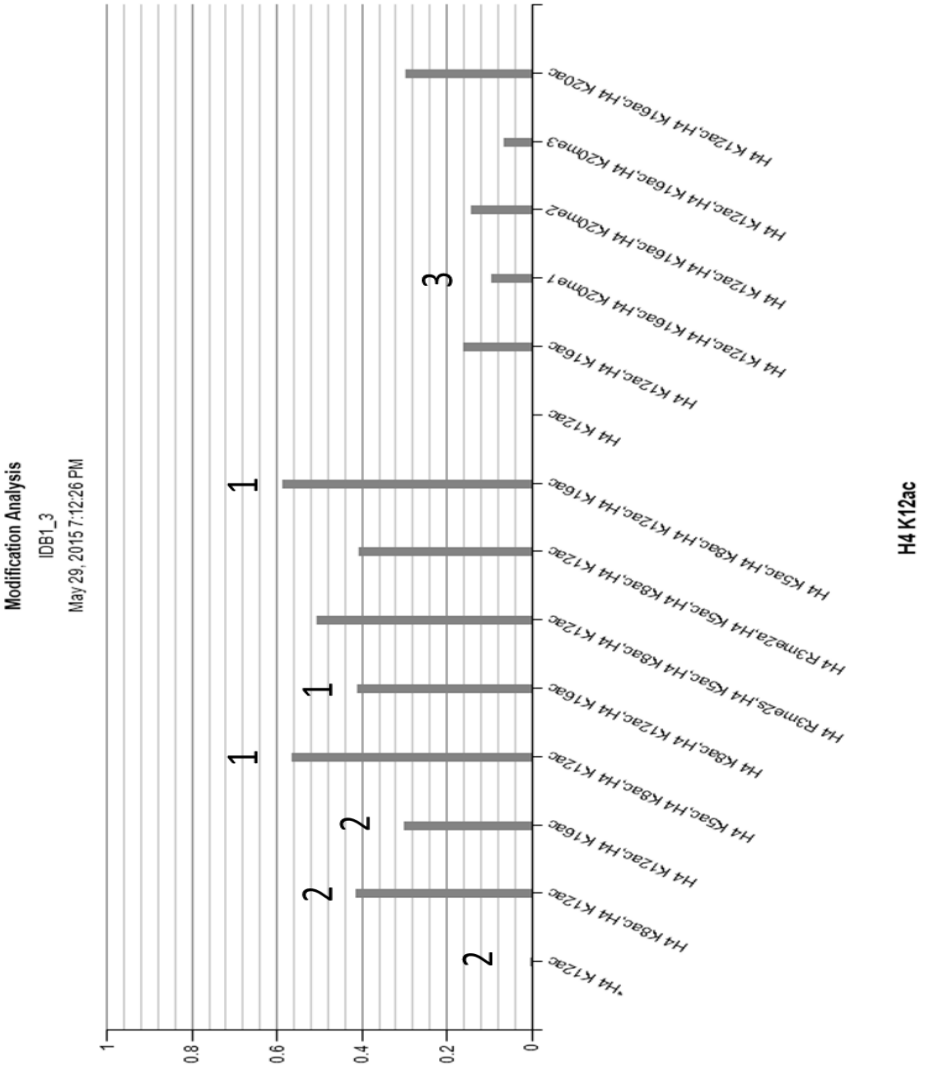


Figure 35: Single peptide reactivity analysis for hH4K12ac. The H4K12ac mark is a posttranslational modification in humans (hH4K12ac). In *Tetrahymena*, the equivalent etylation mark is found on lysine residue at position 11 and is subsequently denoted TtH4K11ac. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

4.17.6.3.4 Single peptide reactivity analysis for hH4K5ac

The individual analysis of the sites containing only hH4K5ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 36. Ibd1 recognizes H4K5ac with a low binding specificity (Figure 36-1). Ibd1 has a higher specificity for sites containing H4K5ac coupled to H4K8ac, and even higher for sites coupled to H4K8ac and H4K12ac (Figure 36-2). The addition of H4K16ac to the triplet only slightly increases Ibd1's binding specificity at this site (Figure 36-3).

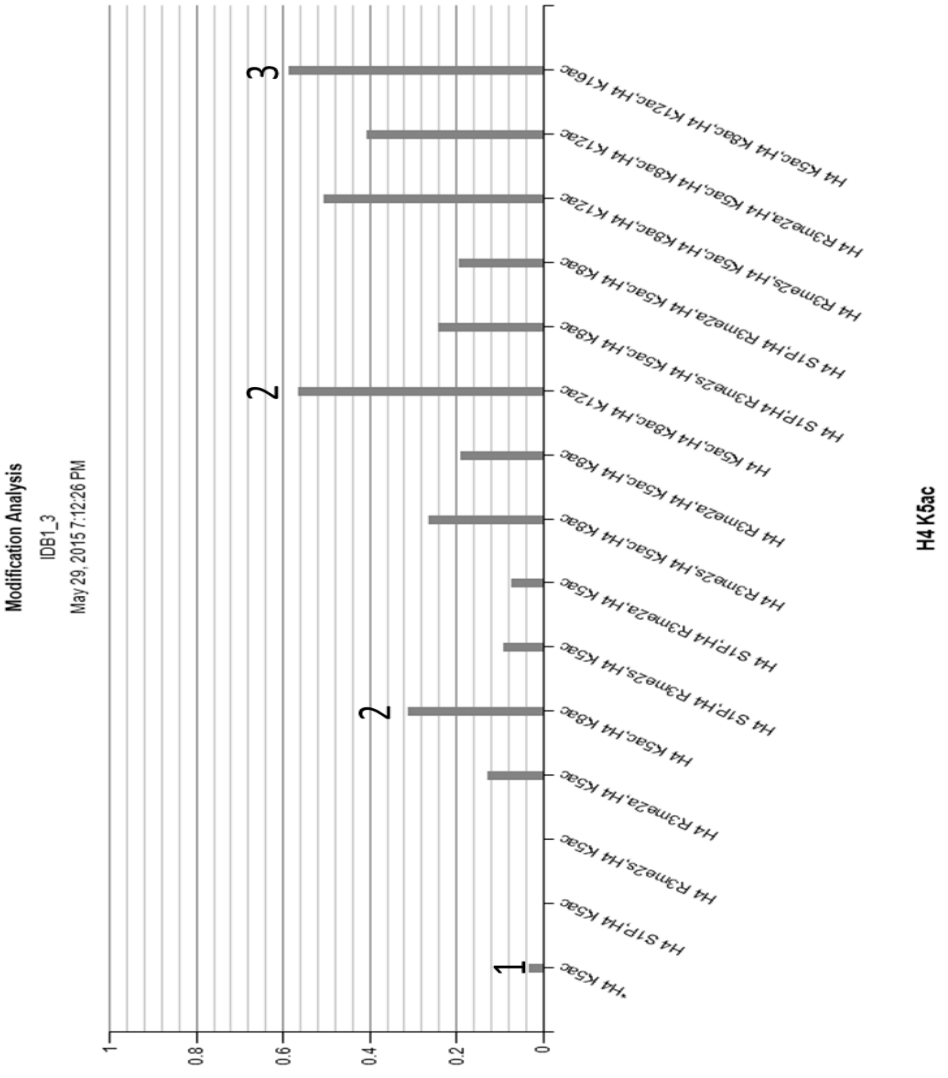


Figure 36: Single peptide reactivity analysis for hH4K5ac.
The H4K5ac mark is a posttranslational modification in humans (hH4K5ac). In *Tetrahymena*, this acetylation mark is on the lysine residue at position 4 and is subsequently denoted TtH4K4ac. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

4.17.6.3.5 Single peptide reactivity analysis for hH2AK9ac

The individual analysis of the sites containing only hH2AK9ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 37. Ibd1 recognizes the human H2AK9ac with a low binding specificity, increased slightly in combination with one other PTM (Figure 37-1). Ibd1's recognition of H2AK9ac increases greatly when in combination with 2 or 3 additional PTM; including H2AK13ac (Figure 37-2).

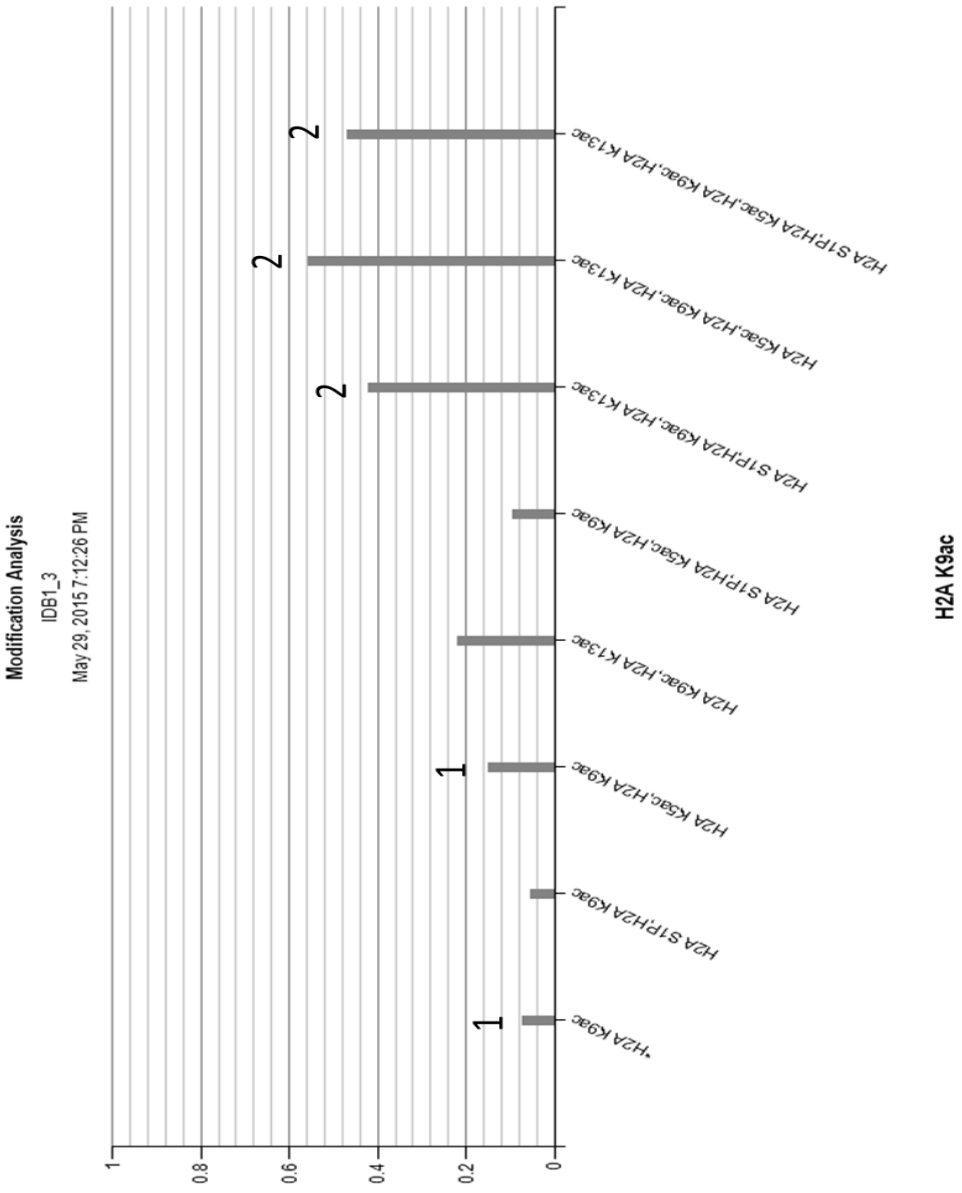


Figure 37: Single peptide reactivity analysis for hH2AK9ac. The H2AK9ac mark is a posttranslational modification in humans (hH2AK8ac). In *Tetrahymena*, this acetylation mark is on the lysine residue at position 8 and is subsequently denoted TtH2AK8ac. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

4.17.6.3.6 Single peptide reactivity analysis for H2AK5ac

The individual analysis of the sites containing only H2AK5ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 38. Ibd1 has slight binding interaction with H2AK5ac, and the addition of H2AS1P decreases the specificity (Figure 38-1). An additional acetylated peptide increases the binding specificity of Ibd1 for H2K5ac Figure 38-2). The combination of H2AK5ac with H2AK9ac and H2AK13ac greatly increases the binding specificity when they share a site (Figure 38-3).

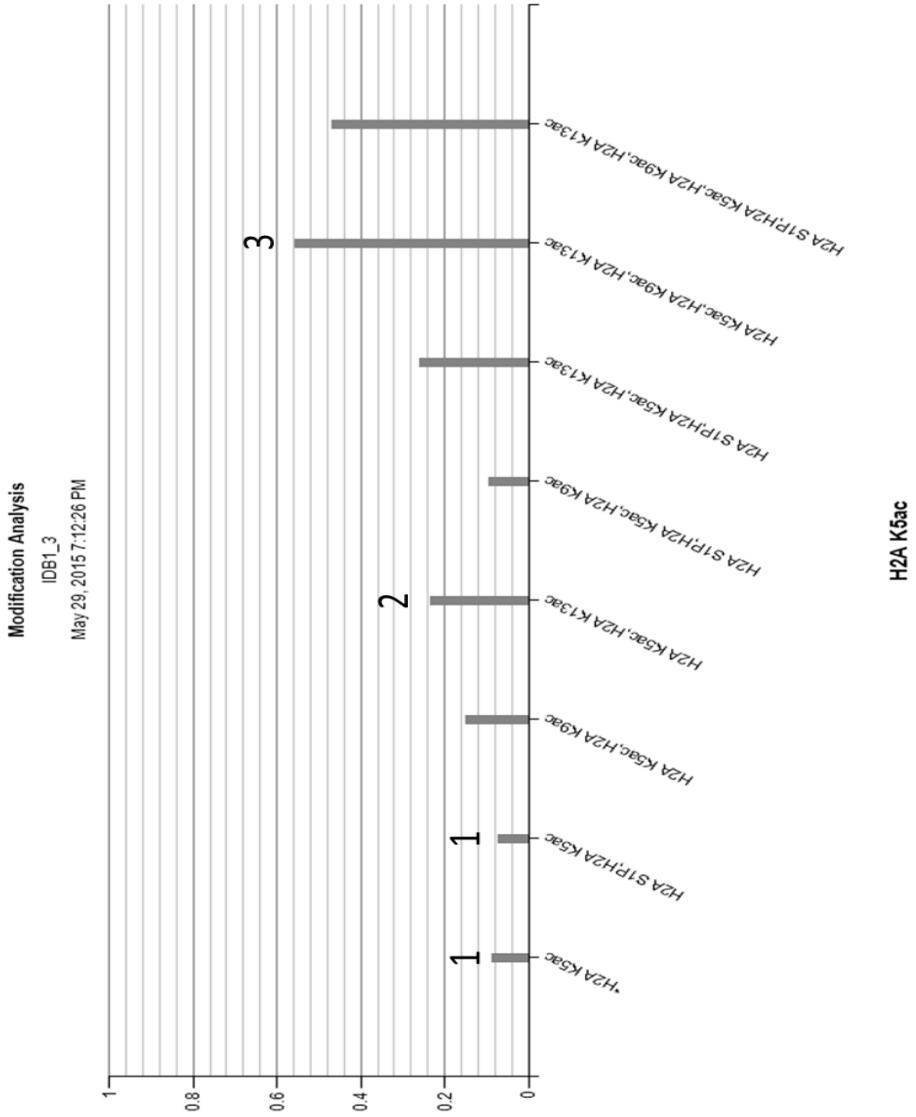


Figure 38: Single peptide reactivity analysis for H2AK5. The H2AK5ac mark is a posttranslational modification in humans and *Tetrahymena.* An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

4.17.6.3.7 Single peptide reactivity analysis for H4K20ac

The individual analysis of the sites containing only H4K20ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 39. Ibd1 recognizes the H4K20ac with higher binding specificity than most of the other acetylated PTM (Figure 39-1). The addition of acetylated peptides to the same site does not significantly affect Ibd1's recognition of H4K20ac (Figure 39-2). Asymmetrical dimethylations decrease Ibd1's recognition for H4K20ac, however symmetrical dimethylations appear to have negligible effects on recognition (Figure 39-3).

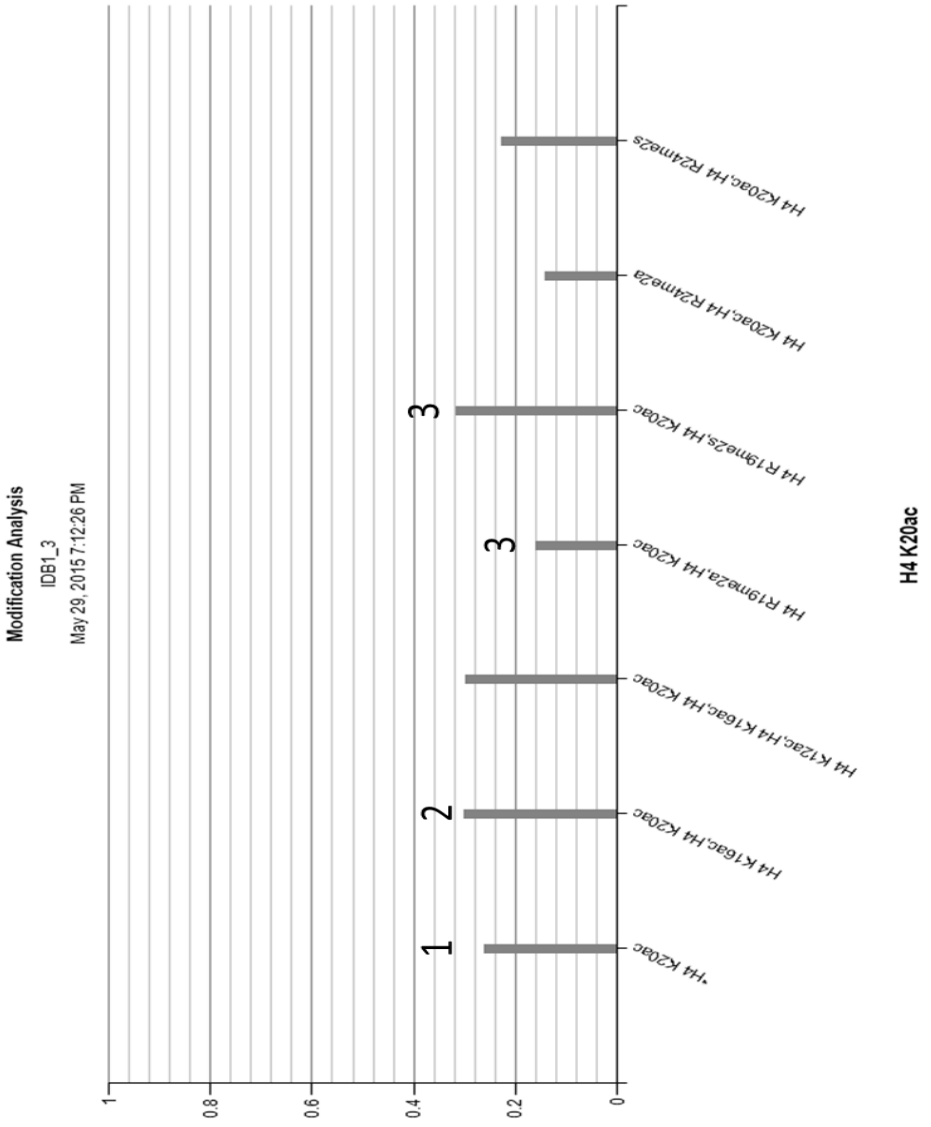


Figure 39: Single peptide reactivity analysis for H4K20ac. The H4K20ac mark is a posttranslational modification in humans and *Tetrahymena*. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above

4.17.6.3.8 Single peptide reactivity analysis for H3K36ac

The individual analysis of the sites containing only H3K36ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 40. Ibd1 recognizes H3K36ac in combination with no other PTM (Figure 40-1).

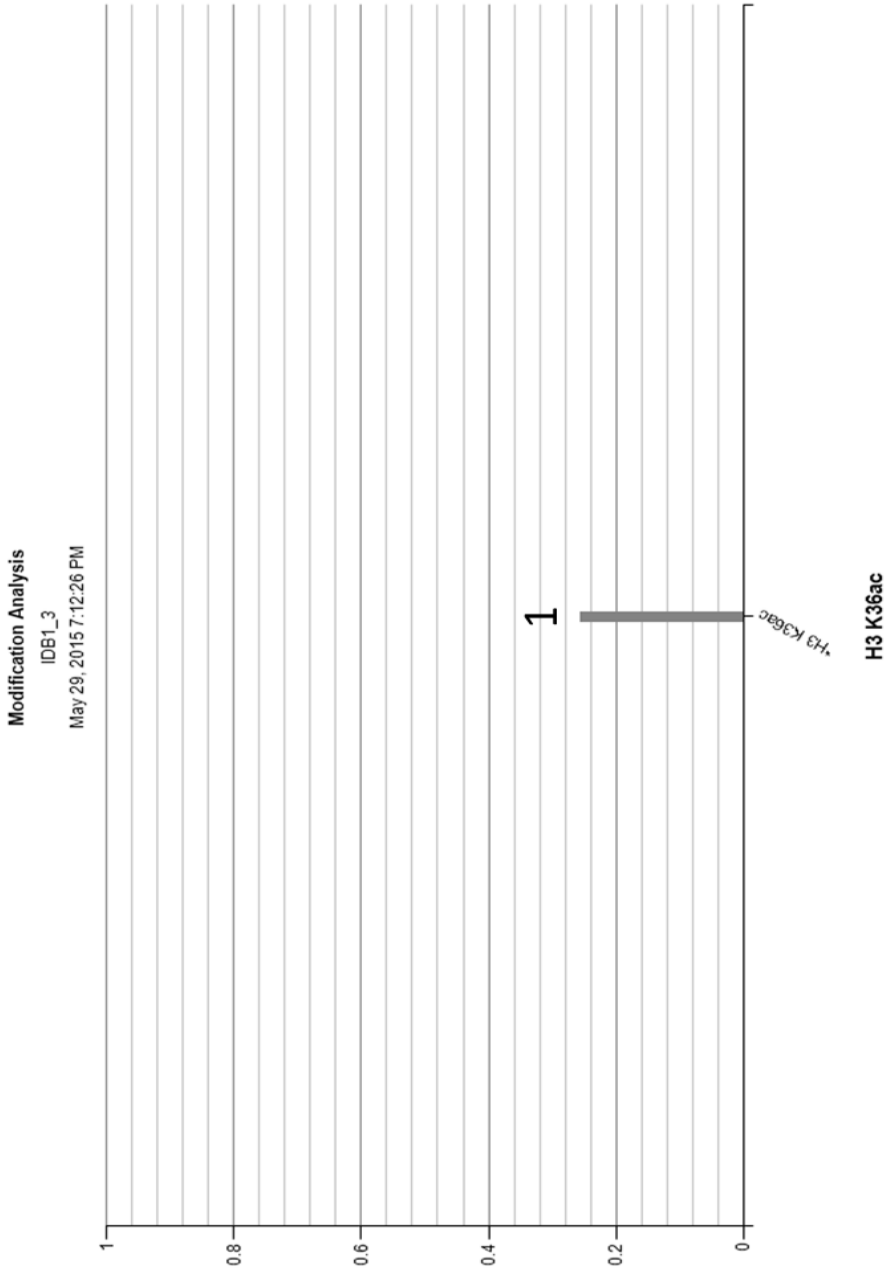


Figure 40: Single peptide reactivity analysis for H3K36ac.
The H3K36ac mark is a posttranslational modification in humans and *Tetrahymena*. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

4.17.6.3.9 Single peptide reactivity analysis for H2AS1P

The individual analysis of the sites containing only H2AS1P compared to sites containing this mark as well as additional modified peptides is seen in Figure 41. Ibd1 has low recognition for H2AS1P on its own and in combination with one or both of H2AK5ac and H2AK9ac (H2AK8ac in *Tetrahymena*) (Figure 41-1). When in combination with H2AK13ac (H2AK12ac in *Tetrahymena*), the recognition increases (Figure 41-2). Ibd1 shows the greatest binding specificity for sites containing H2AS1P and the H2AK5ac, H2AK9ac, H2AK13ac triplet (H2AKac, H2AK8ac, and H2AK12ac in *Tetrahymena*) (Figure 41-3).

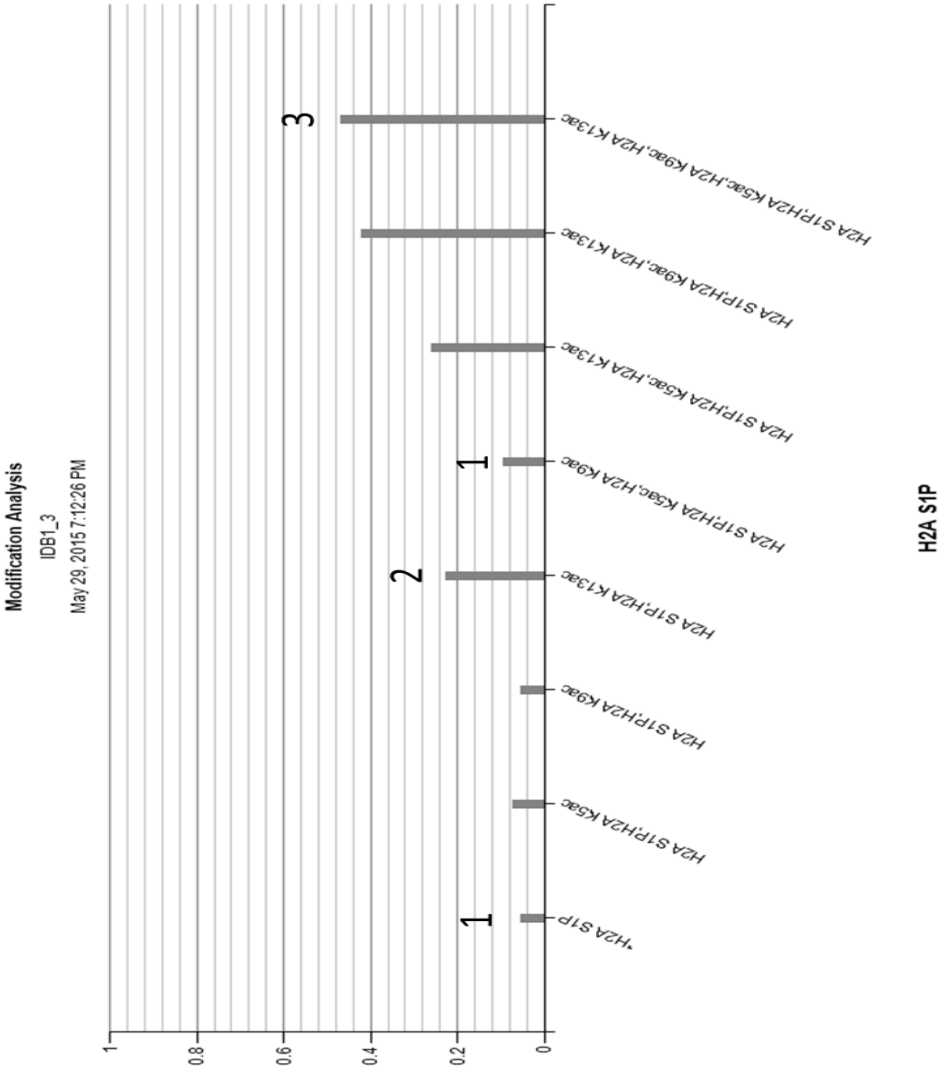


Figure 41: Single peptide reactivity analysis for H2AS1P. The H2AS1P mark is a posttranslational modification in humans and *Tetrahymena*. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

4.17.6.3.10 Single peptide reactivity analysis for hH4K16ac

The individual analysis of the sites containing only hH4K16ac compared to sites containing this mark as well as additional modified peptides is seen in Figure 42. Ibd1 does not recognize H4K16ac on its own (Figure 42-1). The addition of 1, 2 and 3 acetylated peptide(s) increasingly affects the binding specificity of Ibd1 for the site (Figure 42-2). Mono-, di- and trimethylations increase Ibd1's recognition of the site compared to H4K16ac on its own. The symmetry of the methylation does not appear to affect recognition (Figure 42-3).

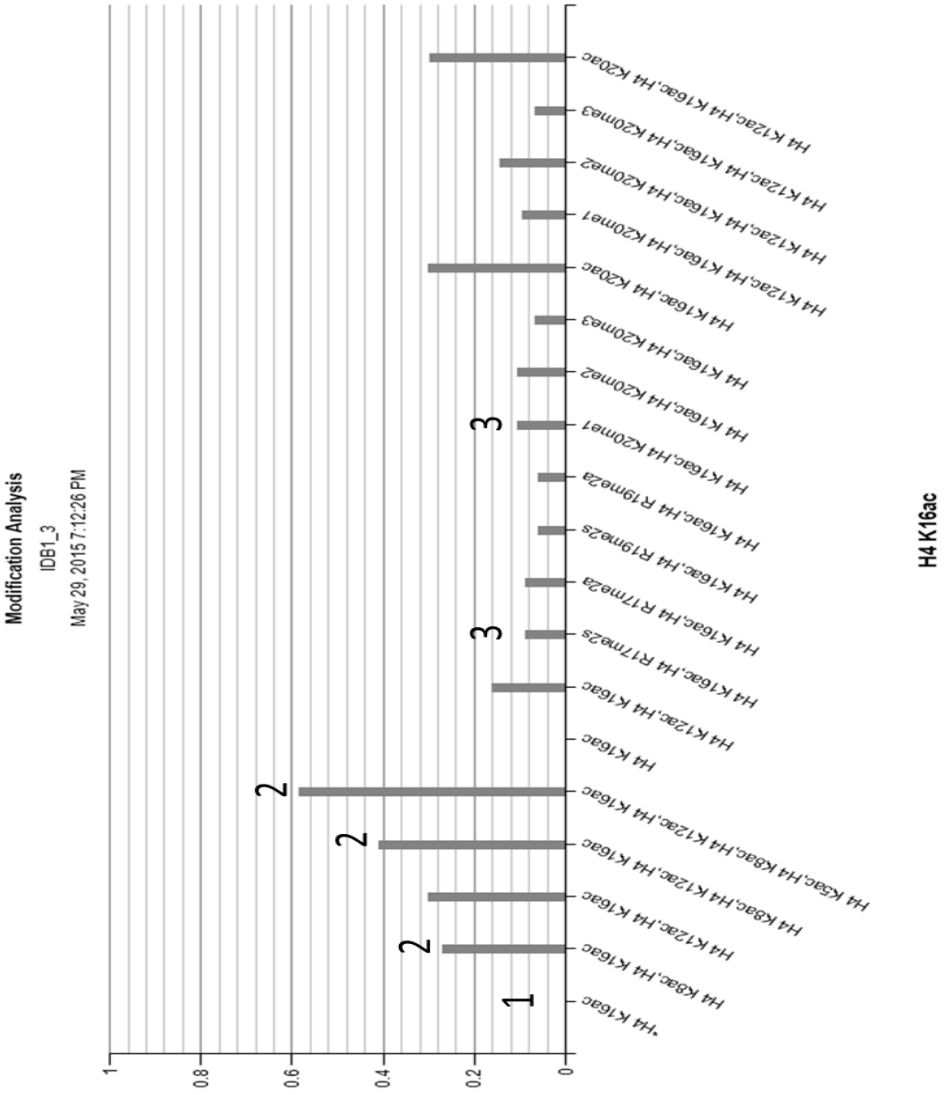


Figure 42: Single peptide reactivity analysis for hH4K16ac.
The H4K16ac mark is a posttranslational modification in humans (hH4K16ac). In *Tetrahymena*, this acetylation mark is on the lysine residue at position 15 and is subsequently denoted TtH4K15ac. An analysis of Ibd1's recognition for this singly acetylated site compared to sites with multiple peptides containing PTM is shown above.

From the single peptide reactivity analysis, it can be seen that the human combination of marks that Ibd1 recognizes with the highest specificity (a value of 0.58) is human H4K5ac, H4K8ac, H4K12ac, H4K16ac. For these human PTM the *Tetrahymena* analogues will be H4K4ac, H4K7ac, H4K11ac, H4K15ac. The next highest value of 0.56 results from the hH4K5ac, hH4K8ac, hH4K12ac combination. For these human PTM the *Tetrahymena* analogues will be H4K4ac, H4K7ac, H4K11. The third highest value of 0.56 results from the combination of hH2AK5ac, hH2AK9ac, hH2AK13ac. For these human PTM the *Tetrahymena* analogues will be H2AK5ac, H2AK8ac, H2AK12ac. Based on this analysis, Ibd1's strongest binding interaction is with the PTM H4K20ac, giving a value of 0.26, and H3K36 with a value of 0.25. There is a clear preference of Ibd1 to multiple PTM.

4.18 Ibd1 localizes to MAC

During growth, I expected to find that SWI/SNF proteins would be present in the MAC but not the MIC, which were consistent with a role in transcription. During conjugation, I expected that cells would show localization again to the MAC and not the meiotic MIC similar to what was shown for TtBrg1 (Fillingham et al., 2006). If the targeted proteins show similar localization to TtBrg1, this is good evidence for common function with TtBrg1 and consistent with function in a SWI/SNF complex. Thus, TtBrg1 does not localize to MIC, which is unusual because several transcription proteins localize to meiotic MIC. Perhaps in *T.thermophila* the SWI/SNF complex does not function in MIC transcription but only in MAC transcription and other functions such as amitosis.

In order to further characterize the function of the TtSWI/SNF complex, I performed IF of Ibd1. Figure 43 shows that Ibd1 localizes to the MAC during growth and sexual development including meiosis, consistent with function in transcription. Fillingham et al. (Fillingham et al., 2006) showed that TtBrg1 localizes to the MAC during growth and development. This is consistent with a role in transcription during growth and development. The targeted protein (Ibd1) shows similar localization to TtBrg1, providing substantial evidence for common function with TtBrg1 and consistent with function in a SWI/SNF complex.

4.19 H4K8ac localizes to MAC

The rationale of doing this experiment is based on the peptide array results that suggest that Ibd1 recognizes modifications where the PTM H4K8ac is present. If the previous statement is correct, H4K8ac should have similar localization to Ibd1. Figure 43 shows that H4K8ac and Ibd1 both only localize to the MAC during growth and sexual development including meiosis, consistent with function in transcription, consistent with a role in transcription.

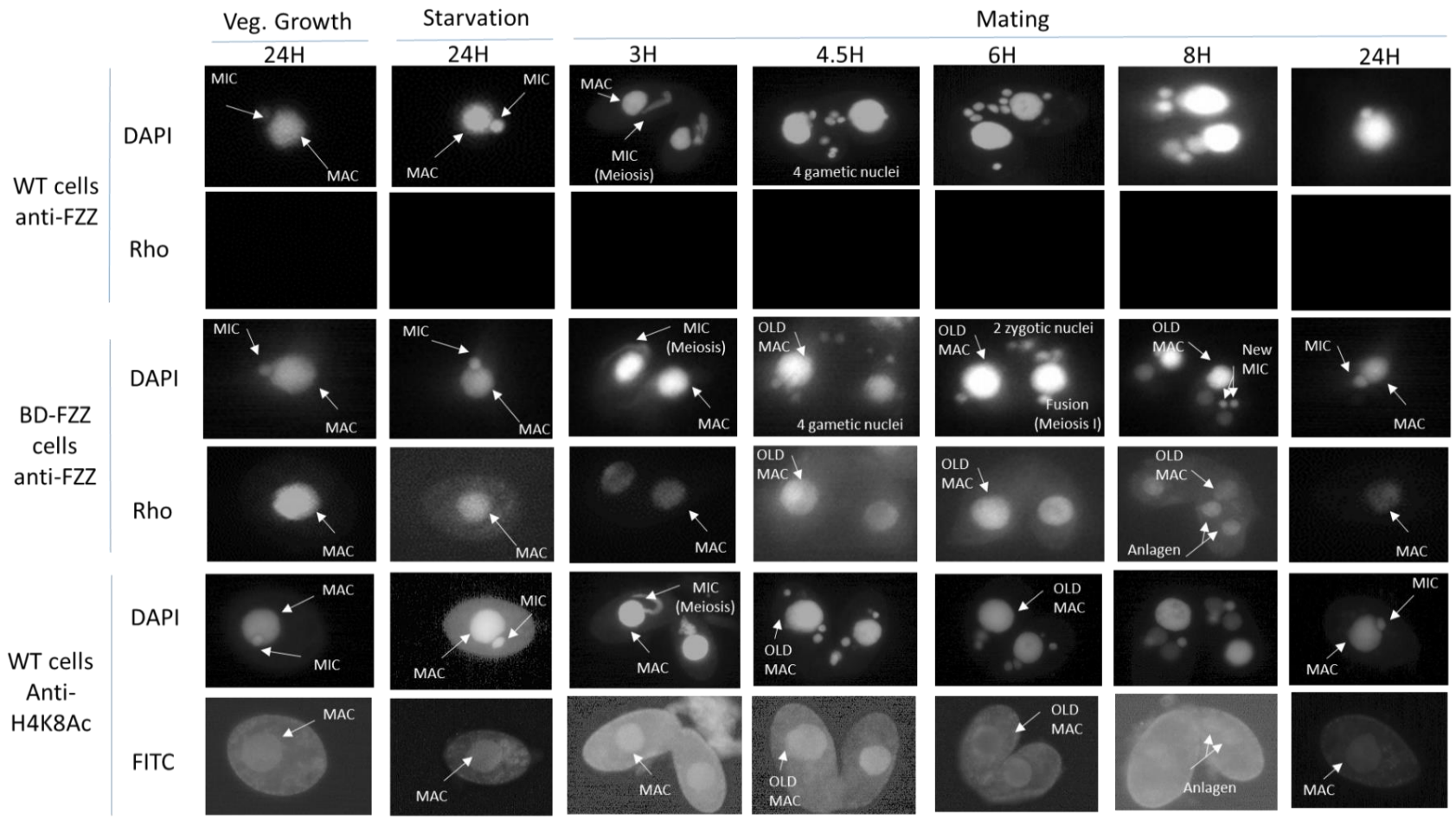


Figure 43: Ibd1 and H4K8Ac localizes to the MAC during growth and sexual development including meiosis

4.20 Model of Ibd1 localization during *T.thermophila's* life cycle

Figure 44 shows that Ibd1 localizes to the MAC during vegetative growth, starvation, meiosis (3H), gametic nuclei fusion (4.5H) and zygotic nuclei duplication (6H). At 8H during anlagen development, Ibd1 localizes to anlagen (new MAC) and old MAC similar to Brg1 (Fillingham et al., 2006). This is consistent with a role in transcription and new MAC development. At 24H when the daughter cells have formed, the expressed Ibd1 from the parental MAC might be stable enough to localize to the MAC of the daughter cells.

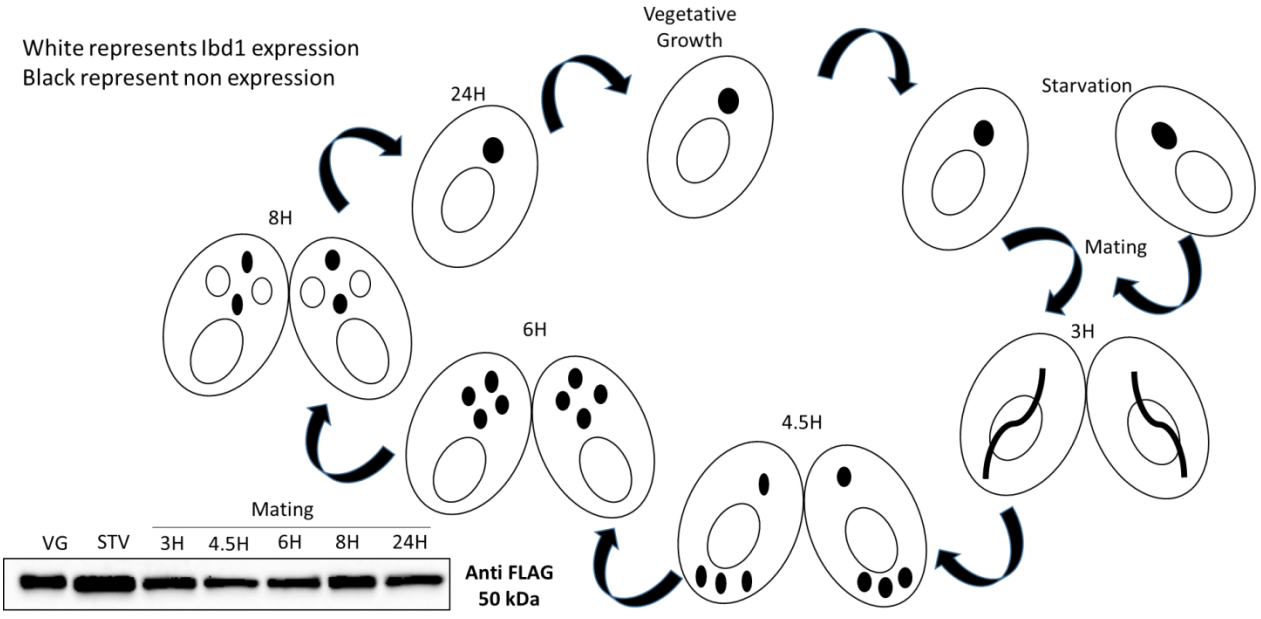


Figure 44: Model of Ibd1 localization during *T.thermophila's* life cycle.

5 CHAPTER 3 - ISWI RESULTS

An important part of *Tetrahymena*'s life cycle are DNA rearrangements (Madireddi et al., 1996)(Smothers et al., 1997)(Chalker et al., 1999)(Coyne, Nikiforov, Smothers, Allis, & Yao, 1999)(Mochizuki et al., 2002)(Mochizuki et al., 2002)(Taverna et al., 2002)(Malone et al., 2005) (Y. Liu et al., 2007)(Chalker, 2008)(Schwope & Chalker, 2014). One initial interest of our lab was to characterize chromatin related proteins involved in DNA rearrangements. I initially hypothesized that Iswi1 is involved in DNA rearrangements. I based this hypothesis on a publically available gene expression profile of mRNA (www.ciliate.org) for Iswi1 (Figure 45), which indicates that Iswi1 is highly expressed at 2 and 7 hours during conjugation. This coincides with the time during conjugation where other proteins (Figure 46) involved in DNA rearrangements are being expressed. At 2 hours, Iswi1 is highly expressed (Figure 45) and the meiotic transcription of long ncRNAs (scnRNA precursors) is occurring (Mochizuki et al., 2002). Therefore Iswi1 may be the protein that regulates the expression of long ncRNA. To answer this, I tagged the Iswi1 protein (TTHERM_00388250).

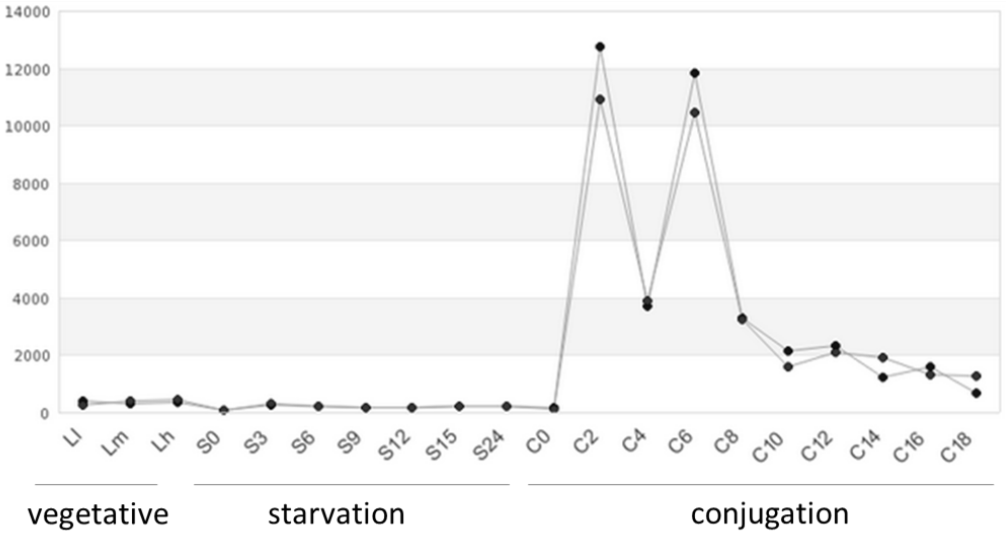


Figure 45: Iswi1 gene expression profile from RNA-seq
Tetrahymena functional genomics database from ciliate.org

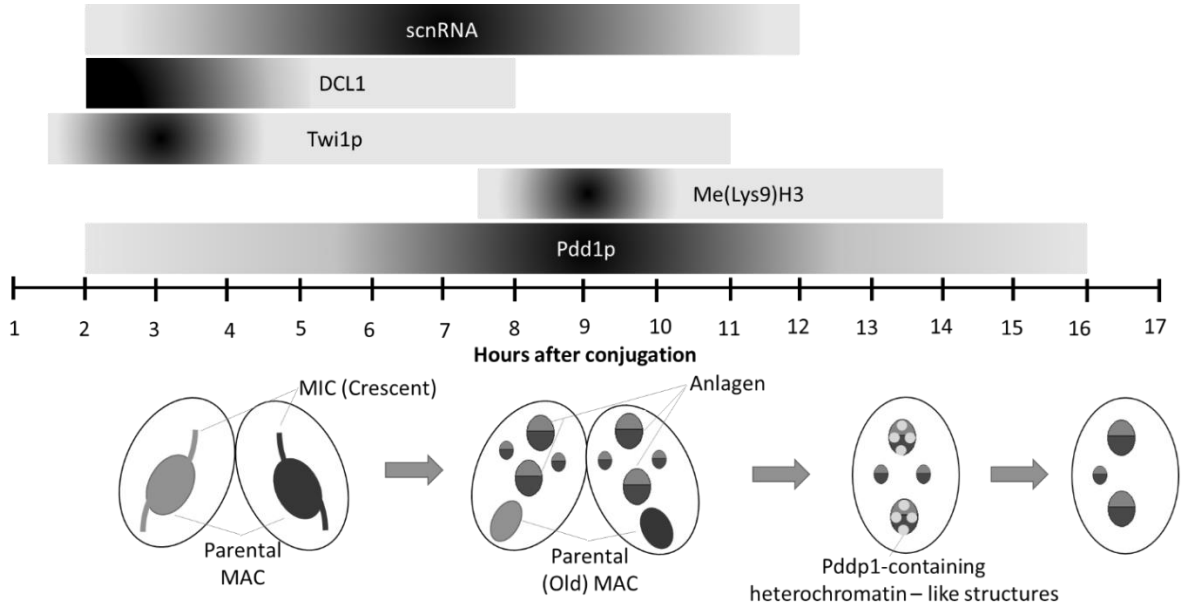


Figure 46: Intermediated molecules involved in *T.thermophila*'s DNA rearrangement.
The darker areas of the bars represent the expression peaks. Ezlp1 tri-methylates H3K9 (Me(Lys9)H3)

5.1 Characterization of potential ISW1-FZZ transformants during development

The publicly available gene expression data (Figure 45) suggests Iswi1 is not expressed during growth. In materials and methods and in chapter 1, I described the steps for cloning of ISW1-FZZ tagging construct and generation of lines expression ISW1-FZZ. I was able to tag one strain from B2086 cells; ISWI1-FZZ B, and three from CU428 cells; ISWI1-FZZ C1, ISWI1-FZZ C2 and ISWI1-FZZ C3. Starved cells of different mating types were mixed and collected at 2 and 3 hours (Figure 47), as these are the approximate times that the expression profile indicated high Iswi1 expression (Figure 45). I used TCA precipitation followed by Western blotting analysis to assess whether the 4 transformed pm-r cells are expressing ISWI1-FZZ during conjugation. Western blotting of the WCEs using the M2 antibody shows that only the ISWI1-FZZ strain is expressing a protein of about 150kDa which corresponds to 18kDa from the FZZ and 132kDa from the Iswi1 protein. As expected from data shown in Figure 47, ISWI1 expresses during conjugation.

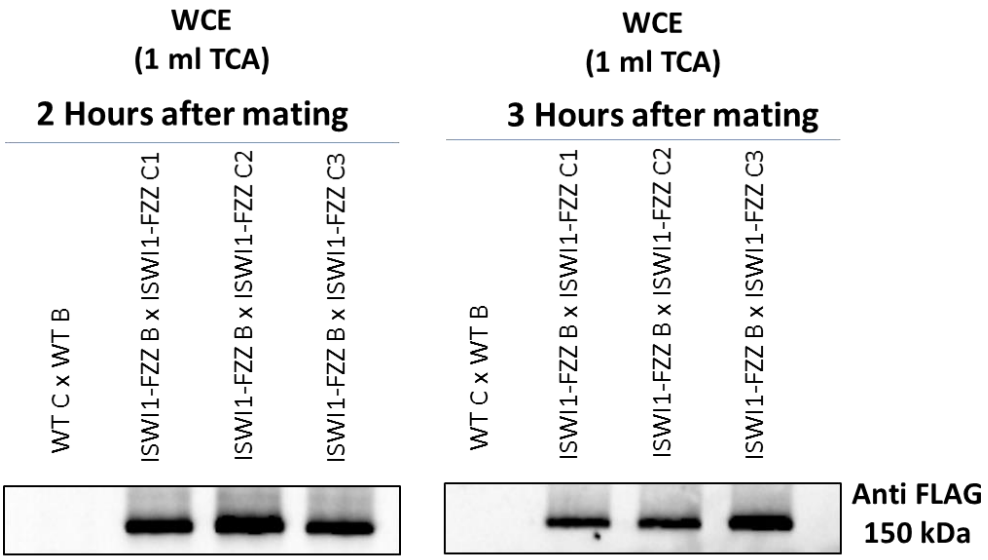


Figure 47: Characterization of potential ISW1-FZZ transformants.
Iswi1 (Anti FLAG) expresses during conjugation.

5.1.1 Expression analysis to assess whether individual strains are expressing ISWI1-FZZ during conjugation

For the conjugations that gave a positive signal for ISW1-FZZ in Figure 47, it is impossible to tell whether one or both of the strains are actually expressing ISW1-FZZ. Therefore I mixed starved cells of each of the obtained transformants with WT cells of a different mating type and then performed a similar experiment to that described above to assess whether individual strains are expressing ISWI1-FZZ. The western blot of Figure 48 shows that every transformed cell is expressing ISWI1-FZZ. Western blotting using the M2 antibody shows that only the ISWI1-FZZ strain is expressing a protein of about 150kDa which corresponds to the ISWI1-FZZ protein. Western blotting using anti-Brg1 is used as a loading control.

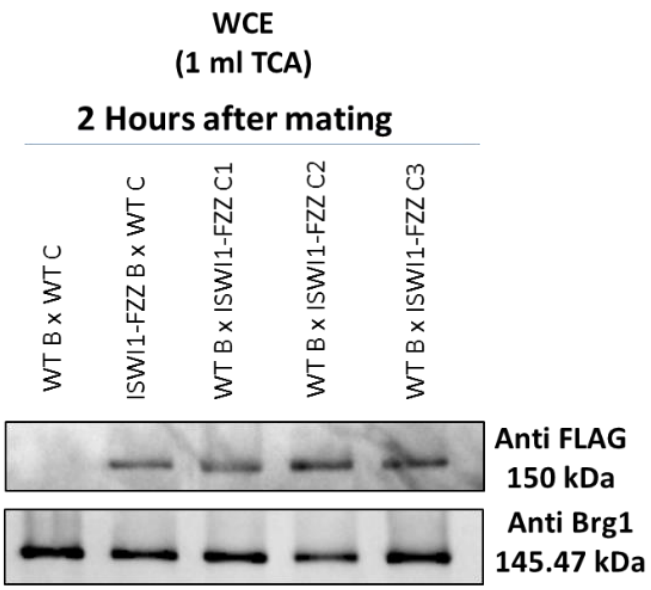


Figure 48: Expression analysis to assess whether individual strains are expressing ISWI1-FZZ during conjugation.
Anti-FLAG shows that all transformants are expressing ISWI1-FZZ.

5.2 Expression analysis of ISW1 during growth

Although the publically available gene expression data indicates it is not expressed during growth (Figure 47), I next used TCA precipitation followed by Western blotting analysis to assess whether the two of the four transformed pm-r cells are expressing Iswi1 during vegetative growth. Unexpectedly, Iswi1 is also expressed during vegetative growth. Western blotting using the M2 antibody shows that the ISWI1-FZZ (Figure 49) strain is expressing a protein of about 150kDa which corresponds to the ISWI1-FZZ protein. To further characterize ISW1 expression at the protein level confirm this, a time course experiment for ISWI1 expression throughout *T.thermophila*'s life cycle was performed.

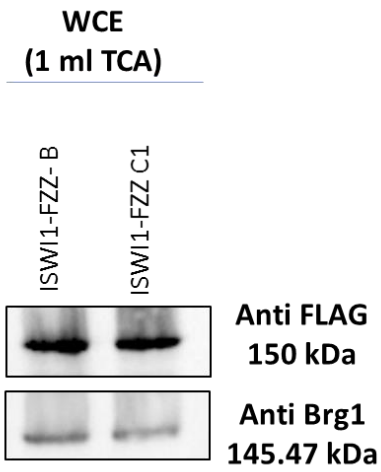


Figure 49: Expression analysis of ISWI1-FZZ during vegetative growth.
Iswi1 (Anti FLAG) expresses during vegetative growth.

5.3 Expression analysis of ISWI1-FZZ during *T.thermophila*'s life cycle

I next performed a time course expression analysis of Iswi1 (Figure 50). I made TCA WCEs from the indicated times in the *Tetrahymena* life cycle and analyzed them using SDS-PAGE and Western blotting. I used anti-Brg1 antibody as a loading control for expression throughout *T.thermophila*'s life cycle, and monitored Pdd1 expression strictly as a developmental loading control. With these results, it was evident that Iswi1 expresses throughout *T.thermophila*'s life cycle.

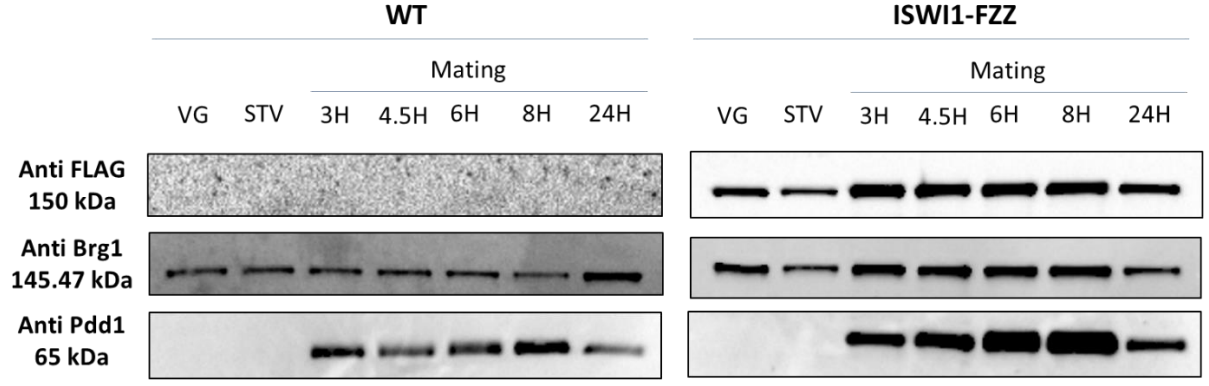


Figure 50: Expression analysis of ISWI1-FZZ during *T.thermophila*'s life cycle
Iswi1 (Anti FLAG) expresses throughout *T.thermophila*'s life cycle. Brg1 is a loading control for expression throughout *T.thermophila*'s life cycle and Pdd1 is strictly a developmental loading control.

5.4 Analysis of ISWI1-FZZ during conjugation

In order to analyze the expression analysis of BD-FZZ during conjugation I affinity purified the FZZ-tagged protein and any co-purifying proteins, I performed one-step affinity chromatography on whole cell extracts made from cells in conjugation expressing ISWI1-FZZ. For the Input, the M2 antibody Figure 51 shows that only the ISWI-FZZ strains are expressing a protein of about 150kDa corresponding to ISWI1-FZZ. For the IP, the M2 antibody Figure 51 shows that only the ISWI1-FZZ strain is expressing a protein of about 150kDa corresponding to ISWI1-FZZ. WT represents the mock IP.

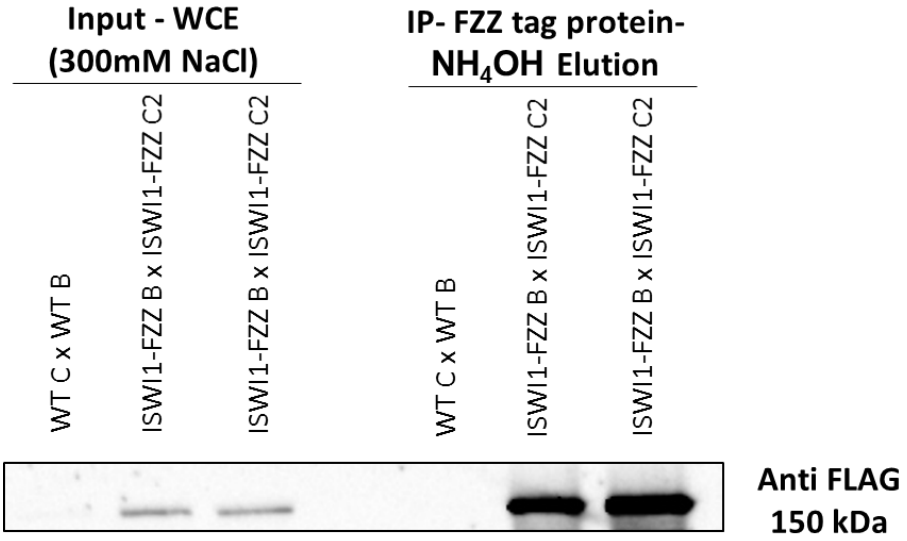


Figure 51: Expression analysis of ISWI1-FZZ during conjugation
Iswi1 (Anti FLAG) expresses during conjugation

5.5 MS and SAINTexpress analysis of ISWI1-FZZ during conjugation

Through SAINTexpress analysis of the mass spectrometry data from the Iswi1 affinity purification, I have identified several potential interacting proteins at this point it is not known if they are related to DNA rearrangements.

Table 10: SAINTexpress-curated data for 2 replicates of ISWI1-FZZ during conjugation

Protein	Gene Name	SpecSum
Iswi1	TTHERM_00388250	57
hypothetical protein	TTHERM_00247110	28
hypothetical protein	TTHERM_00582070	19
Sm protein (spliceosome)	TTHERM_01403810	7
hypothetical protein	TTHERM_00709740	6
RPT6 (26S proteasome regulatory subunit T6)	TTHERM_00551090	4

The hypothetical protein (TTHERM_00247110) has a similar (Figure 45) gene expression profile to Iswi1 (Figure 52) therefore they may be working together. This makes this protein an excellent candidate for tagging and further analysis. In addition, this protein does not have known domains (NCBI).

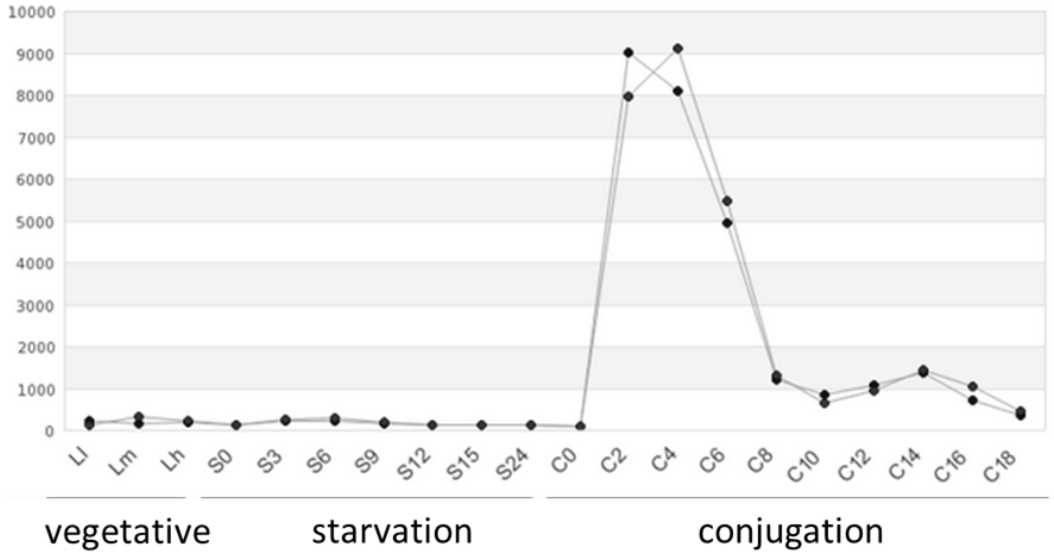


Figure 52: TTHERM_00247110 gene expression profile from RNA-seq
Tetrahymena functional genomics database from ciliate.org

5.6 Iswi1 localizes to the MIC during growth and sexual development including meiosis

In order to further characterize the function of Iswi1, I performed IF. I found that Iswi1 localizes to the MIC during growth and sexual development, including meiosis. Localization of Iswi1 protein to the MIC is consistent with role in mitosis, or maintenance of silent chromatin (Figure 53). Iswi1 does not localize to where other proteins (Pdd1p, Twi1p, Ezlp1 and DCL1) related to DNA rearrangement localize. Therefore Iswi1 localization to MIC might mean that it is involved in transcription of MIC long ncRNAs that are the precursors to scnRNA.

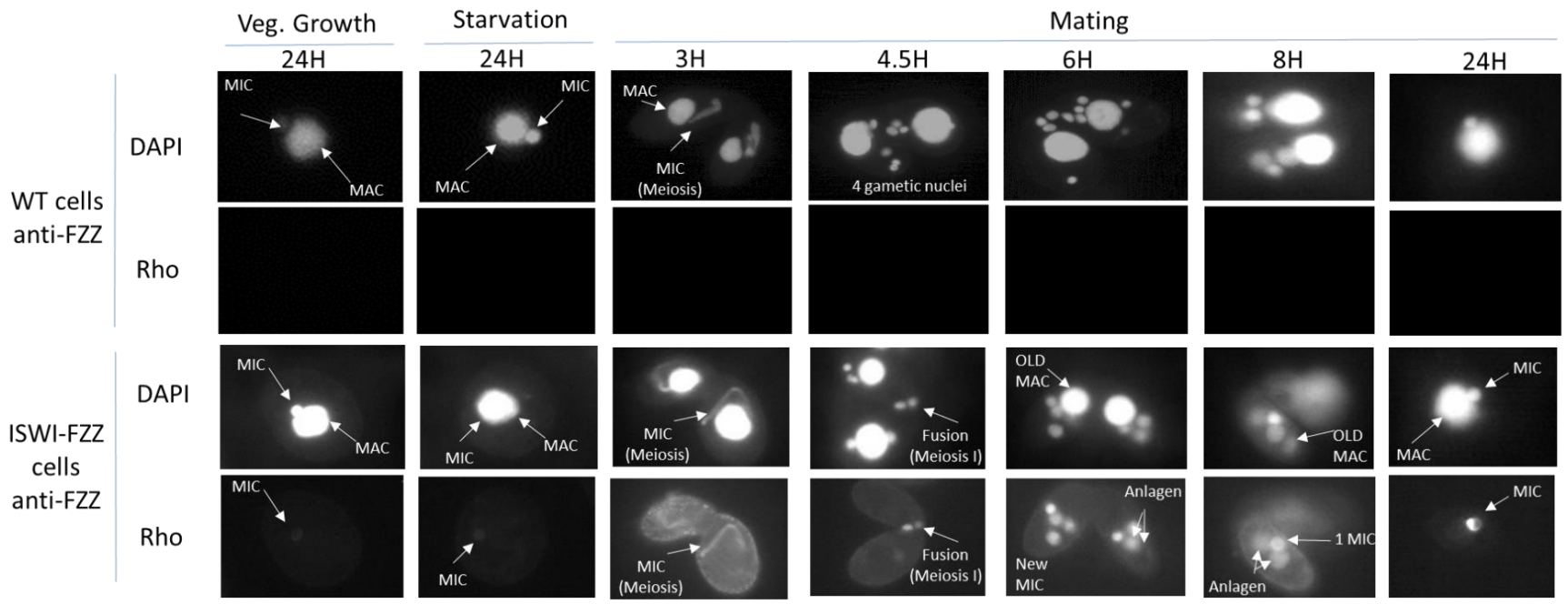


Figure 53: Iswi1 localizes to the MIC during growth and sexual development including meiosis

5.7 Model of ISWI1 localization during *T.thermophila*'s life cycle

Figure 54 shows that Iswi1 localizes to the MIC during vegetative growth, starvation, meiosis (3H), gametic nuclei fusion (4.5H) and zygotic nuclei duplication (6H). At 8H during anlagen development, Iswi1 localizes to anlagen (new MAC) and new MIC. This is consistent with a role in mitosis, maintenance of silent chromatin, perhaps transcription of some specific genes during meiosis, gametic differentiation and fusion, and development of the new MIC and MAC. At 24H when the daughter cells have formed, the expressed Iswi1 from the parental MAC might be stable enough to localize to the MIC of the daughter cells.

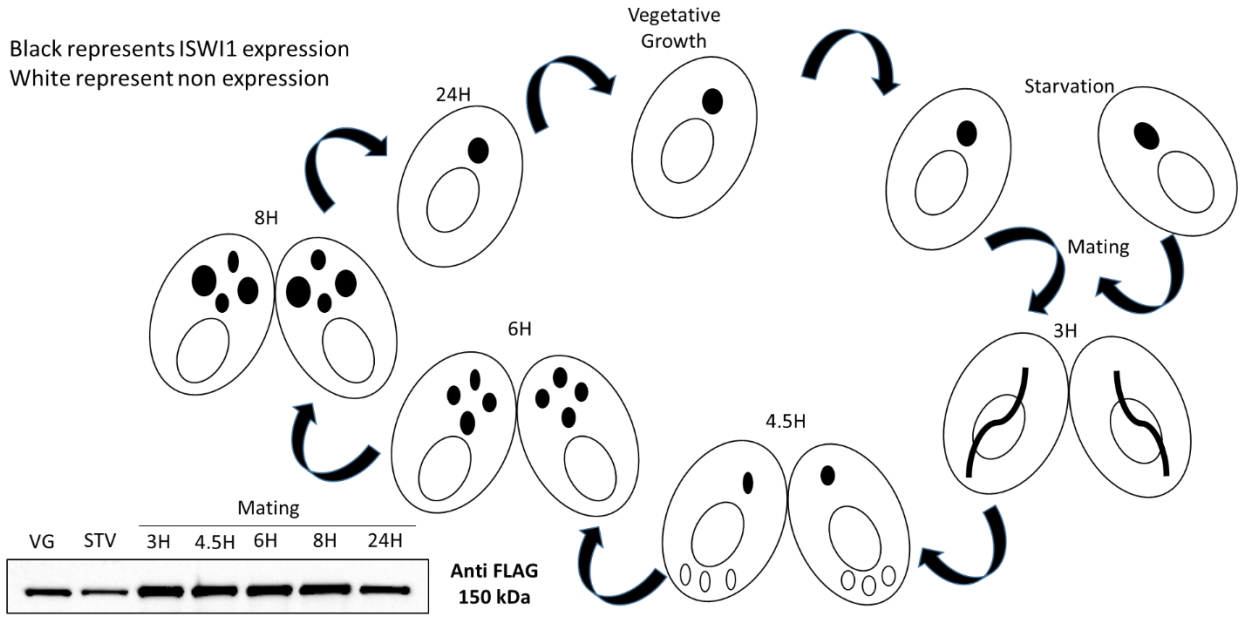


Figure 54: Model of ISWI1 localization during *T.thermophila*'s life cycle

6 SUMMARY AND FUTURE DIRECTIONS

6.1 SWI/SNF complex

AP-MS of tagged TtSnf5 revealed members of a putative TtSWI/SNF complex. One of these revealed members was a bromodomain containing protein (Ibd1). This protein was selected and tagged using molecular cloning methods. The purpose of tagging Ibd1 was to confirm reciprocal AP-MS interactions with Snf5. The rationale behind reciprocal purifications is to validate the identity of the SWI/SNF complex and possibly identify new members. Through AP-MS of *Tetrahymena* Snf5, and Snf5-interacting Ibd1, I have identified the first protist SWI/SNF complex. In addition, perhaps *Tetrahymena* SWI/SNF has two proteins that potentially recognize PTM on histones, the PHD domain (TTHERM_00241840) and Ibd1 (TTHERM_00729230). These two proteins could work in conjunction to positively regulate transcription at an increased rate. One possible double PTM recognized by SWI/SNF could be chromatin tri-methylated at H3K4 (Taverna et al., 2007) and multiple acetylations on H4 which is correlated with highly transcribed genes (Turner et al., 1992). A ciliate-specific protein (TTHERM_00092790) was found and for this protein it could be established if the expression of the gene encoding this protein is essential for growth. If the protein is essential it can be used as a target protein for a rational design drug of ciliate specific antibiotic. For example this drug could be used against the parasitic ciliate *Ichthyophthirius multifiliis* which affects aquiculture (Sigh et al., 2004).

Fillingham et al. (Fillingham et al., 2006) showed that TtBrg1 localizes to the MAC during growth and development. This is consistent with a role in transcription during growth. IF for SWI/SNF showed localization of Ibd1 protein and H4K8ac to MAC consistent with Fillingham's results. To further illustrate the SWI/SNF member's localization, IF of Snf5 can be performed. Our data for Ibd1 localization suggests that Ibd1 localizes to the MAC during vegetative growth, starvation, meiosis (3H), gametic nuclei fusion (4.5H) and zygotic nuclei duplication (6H). At 8H during anlagen development, Ibd1 localizes to the anlagen (new MAC) and old MAC. At this point the anlagen or new MAC takes over transcription from the old MAC which begins to degrade. Since Brg1 (Fillingham et al (2006) and Ibd1 (this study) follow this pattern we hypothesize that Ibd1 and SWI/SNF have a role in transcription.

6.1.1 Ibd1

AP-MS of Ibd1 during vegetative growth suggests that in addition to being a member of SWI/SNF in *Tetrahymena*, (Table 5). Ibd1 interacts with multiple proteins (SWR, HMT, SAGA) that have not been yet studied in *Tetrahymena*. These Ibd1 interacting proteins correspond to ortologues that are part of known transcription-related complexes in yeast and humans. Our data suggests that during conjugation, there is a change in the complexes that Ibd1 is interacting with. However, my analysis shows that 5 hours after the initiation of conjugation, when the zygotic nucleus is being formed by gametic nuclei fusion, Ibd1 still interacts with the same complexes but the composition is different (Table 9). Figure 23 shows that at 5 hours after conjugation there is a decrease in the amount of Ibd1 that is interacting with all SWI/SNF subunits and the HMT (TTHERM_00486440). However, there is an increase relative to SWI/SNF and HMT (TTHERM_00486440) of all putative SWR-C and putative SAGA subunits present in the Ibd1 affinity purification in conjugation. Perhaps Ibd1 has an important role in histone variant exchange and histone acetylation during gametic nuclei fusion. Perhaps Ibd1's function as a member of the SWI/SNF complex or HMT is not required. Interestingly, the hypothetical protein (TTHERM_00046150), along with other proteins present in this purification, such as PHD Finger (TTHERM_00444700), AT-hook containing (TTHERM_00355040) and ciliate-specific (TTHERM_00136450), drastically increase in the percentage interacting with Ibd1 at this stage during conjugation. These last mentioned proteins might be also important for gametic nuclei fusion along with putative SWR-C and putative SAGA.

An interesting question that arose from this research is if during vegetative growth and conjugation all of the Ibd1 related complexes work independently or as a mega complex. This question was based on the fact that *Tetrahymena thermophila* is a large unicellular organism (~50μm) and having a high rate of transcription is of paramount importance for maintenance. The first model supported by this study is during vegetative growth and suggests that highly transcribed regions have multiple acetylations in H4. These regions are recognized by Ibd1 to recruit SWI/SNF, SWR, SAGA and HMT all of which work together to form a mega complex. This mega complex will increase the transcription rate on a subset of genes during vegetative

growth. The second model suggest that during conjugation the complexes work independently. The comparative analysis of both AP-MS data from vegetative growth and 5 hours after mating (Figure 23) shows that the number of subunits of each complex is being drastically modified during conjugation compared to vegetative growth. Perhaps during conjugation these complexes are related to chromatin assembly and not to increased transcription rate. In order to determine if the mentioned complexes work independently or as a mega complex in these two physiological states, AP-MS experiments of SAGA, SWRC, HMT members are needed. Alternatively, a gel filtration of Ibd1 to determine the number of distinct Ibd1-containing protein complexes is needed. For the gel filtration of Ibd1, first I would follow all the mentioned steps to perform an AP in materials and methods. However, instead of using a denaturing elution presented in this work I would use a native elution (3xFLAG peptide) to conserve the complex's composition. The 3xFLAG peptide competes with the FZZ-tagged protein for the M2 matrix. After I have the eluates with the native composition of the complexes I will gel filtrate them to separate them by size and then I will analyse each fraction by MS followed by SAINTexpress as described in materials and methods. With this data I will be able to determine the complexes composition.

Another important question raised by this work is whether Ibd1's bromodomain is functional since Brg1 lacks a BD. Based on this my next objective was to answer if Ibd1 is functional. To determine Ibd1's functionality, I engineered a synthetic Ibd1 gene to express a recombinant 6xHis-Ibd1 in *E.coli* BL21 (D3). The purified recombinant protein was used in an array containing sites with single PTM and combinations of PTM. The peptide array confirms that Ibd1 is functional in that it recognizes the canonical target of bromodomains and suggests a role in transcription. My data is consistent with Ibd1 recognizing combinations of marks rather than singly acetylated residues. As suggested in the literature (Filippakopoulos et al., 2012) bromodomains may recognize acetylations and phosphorylations, suggesting that some bromodomains recognize combinations of marks rather than singly acetylated residues. The data obtained from the peptide array suggest that Ibd1 better recognizes multiple peptides. The best candidate combinations to be recognized by Ibd1 as suggested by the peptide array are TtH4K4ac, TtH4K7ac, TtH4K11ac, TtH4K15ac and TtH4K4ac, TtH4K7ac, TtH4K11ac. These PTM are related to highly transcribed regions (Turner et al., 1992). To determine the exact combination of PTM recognized by Ibd1, a dissociation constant experiment can be performed

since it is possible to measure the binding constants (K_d) that Ibd1 may have with the different candidate PTM as described in the literature (Taverna et al., 2002). Once the PTM are determined we could use ChIP-Seq to identify the Ibd1 binding sites in the genome during growth. This experiment will be followed by ChIP-Seq with the antibodies against the PTM given by the K_d experiment. Thus, with these two results we will determine if Ibd1 and the PTM are found in the same genomic region corresponding to highly transcribed regions.

Some other questions raised by this work are: 1] is the SWI/SNF PHD domain (TTHERM_00241840) functional. In order to answer this question a recombinant protein can be engineered and the peptide array experiment can be used. 2] *Tetrahymena thermophila* has 12 bromodomain containing proteins including Gcn5 and Ibd1. From the 12 bromodomain containing proteins, one protein (TTHERM_00483490) has a very similar structure to Ibd1 and characterizing this protein could be interesting.

6.2 ISWI complex

I initially hypothesized that Iswi1 is involved in meiosis/DNA rearrangements. I based this hypothesis on the publically available gene expression profile for Iswi1. Iswi1 is highly expressed at 2 and 7 hours during conjugation, which coincides with the time during conjugation where other proteins involved in DNA rearrangements are being expressed and I thought Iswi1 may have been the protein that regulates the expression of the scnRNA precursor or long RNA. To test this hypothesis, I tagged the Iswi1 protein (TTHERM_00388250). Unexpectedly, Iswi1 also expresses during vegetative growth. I confirmed this through a time course experiment for ISWI1 expression as assessed by western blotting throughout *Tetrahymena*'s life cycle was performed. Through SAINTexpress analysis of AP-MS data of Iswi1-FZZ I have identified several potential interacting proteins, at this point it is not known if they are related to DNA rearrangements. Significantly, the hypothetical protein (TTHERM_00247110) has a similar gene expression profile to Iswi1 making this protein an excellent candidate for tagging and further analysis.

In order to further characterize the function of Iswi1, I performed IF. I found that Iswi1 localizes to the MIC during growth and sexual development, including meiosis. Localization of Iswi1 protein to the MIC is consistent with role in mitosis, or maintenance of silent chromatin (Figure 53). Iswi1 does not localize to where other proteins (Pdd1p, Twi1p, Ezlp1 and DCL1) related to DNA rearrangement localize. Therefore Iswi1 localization to MIC might mean that it is involved in transcription of MIC long ncRNAs that are the precursors to scnRNA. Our model for Iswi1 localization suggests that Iswi1 localizes to the MIC during vegetative growth, starvation, meiosis (3H), gametic nuclei fusion (4.5H) and zygotic nuclei duplication (6H). At 8H during anlagen development,

Some of the questions raised by this work are: 1] what is the function of ISWI1 in *Tetrahymena*? 2] What is the phenotypic analysis of MIC integrity in ISWI1-KO in growth and meiosis? 3] What is the function of ISWI1 interacting proteins – are they specific for development?

7 APPENDICES

7.1 Appendix 1 – Recipes

Reagent	Composition
0.8% Agarose Gel (w/v) (50 mL)	0.4 g agarose 50 mL 1xTBE or 1xTAE 5 µL RedSafe Nucleic Acid Stain, FroggaBio
0.5M Ammonium Hydroxide (NH_4OH)	1 mL 14.5M NH_4OH 28 mL ddH_2O
10% APS (Ammonium Persulfate) (w/v)	0.1 g ammonium persulfate 1 mL ddH_2O
1 M $CaCl_2$ (1L)	$CaCl_2$ = 110.99 g/mol 1 mole $CaCl_2$:1 L ddH_2O
2mM $CaCl_2$/20mM Tris	100 µL 1M $CaCl_2$ 1 mL 1M Tris pH 8.0 48.9 mL ddH_2O
0.5M EDTA, Iron (III) Sodium Salt pH 8.0 (500mL)	91.78 g Na_2EDTA (367.1 g/mol) ddH2O up to 500 mL, adjust pH to 8.0 using NaOH (1N)
2x Lysis Buffer (50 mL) Note: For 1x Lysis Buffer, prepare equal parts 2x Lysis Buffer and ddH_2O	2 mL 1M Tris pH 8.0 50 µL 1M $MgCl_2$ 42 mL ddH_2O 6 mL 5M NaCl 500 µL Protease Inhibitor Cocktail, Sigma 200 µL phenylmethylsulfonyl fluoride (PMSF)
1M $MgCl_2$ (203.3 g/mol) (100 mL)	20.33 g $MgCl_2$; ddH_2O to 100 mL
5xTBS	48.44 g Tris 584.4 g NaCl 3.5 L ddH_2O Adjust pH to 7.5 using HCl Add ddH_2O up to 4L
3% Milk (50 mL)	1.5 g dried milk powder 50 mL 1xTBS
10x Western Buffer	121.2 g Tris 576 g glycine Add up to 4L of ddH_2O
1x Running Buffer	100 mL 10x Western Buffer 900 mL ddH_2O 10 mL 10% SDS
10% SDS (w/v)	500 mL ddH_2O 50 g SDS
100mM NaCl Wash Buffer (IPP100)	500 µL 1M Tris pH 8.0 1 mL 5M NaCl 500 µL 10% NP-40

	48 mL ddH_2O
300mM NaCl Wash Buffer (IPP300)	500 µL 1M Tris pH 8.0 3 mL 5M NaCl 500 µL 10% NP-40 46 mL ddH_2O
5M NaCl (500 mL)	146.1 g NaCl ddH_2O up to 500 mL
10% NP-40 (v/v)	2.5 mL NP-40 22.5 ddH_2O
10xPBS pH 7.3 (1L)	82 g NaCl 2.64 g NaH_2PO_4 16 g Na_2HPO_4 ddH_2O up to 1 L, pH 7.3
1xPBST (500 mL)	500 mL 1xPBS 250 µL Tween 20
100mM PMSF (10 mL)	0.1742 g PMSF 10 mL isopropanol
Ponceau (0.1% w/v) (1 L)	1 g Ponceau S 50 mL acetic acid ddH_2O up to 1L
2xSDS Loading Dye	1.15 mL 10% SDS 37.5 µL 2M Tris, pH 6.8 62.5 µL beta-mercaptoethanol Pinch of bromophenol blue
SPP Media (4 L)	40 g protease peptone 4 g yeast extract 8 g glucose 0.12 g Sequestrine, Sigma ddH_2O up to 4L
Tetrahymena Lysis Solution (500 mL)	210 g urea 35 mL 5M NaCl 5 mL 1M Tris, pH 7.4 10 mL 0.5M EDTA 50 mL 10% SDS ddH_2O up to 500 mL
1xTEV Cleavage Buffer	500 µL 1M Tris, pH 8.0 1 mL 5M NaCl 500 µL 10% NP-40 50 µL 0.5M EDTA 48 mL ddH_2O
10mM Tris, pH 7.4 (1 L)	1.21 g Tris ddH_2O up to 1 L, pH 7.4
LB Liquid Media	1 L ddH_2O 25 g LB, Miller
LB Media for Plates	1 L ddH2O 25 g LB, Miller

	15 g agar
Running Gel (10%)	10 mL 1.5M Tris, pH 8.8 13 mL 30% Bis/Acrylamide, 37.5:1 17 mL ddH2O 40 µL Temed 200 µL 10% APS
Stacking Gel (4%)	3.8 mL 0.5M Tris, pH 6.8 2.0 mL 30% Bis/Acrylamide, 37.5:1 9.0 mL ddH_2O 15 µL Temed 75 µL 10% APS
Primary Antibody Buffer (50 mL)	1.5 g albumin 50 mL 1xTBS 12.5 µL Tween 20
Shandin	100 mL ethanol 200 mL $HgCl_2$

Table 11: Data analysis for Ibd1 during vegetative growth and 5 hours after conjugation.

			5 hours after conjugation			Vegetative			Comparison	
Gene Name	Protein	Possible Complex	SpecSum	% with respect to BAIT	Fold with respect to the bait	SpecSum	% with respect to BAIT	Fold with respect to the bait	*5 hours after conjugation the compared fold of bait with respect to vegetative growth	% Normalized data
TTHERM_00729230	BD protein (BAIT)	Many	738	1.000	1	1348	1.000	1	1.00	0.00
TTHERM_00046920	RvB2	SWR-C	263	0.356	3	372	0.276	4	1.29	29.14
TTHERM_00476820	RvB1	SWR-C	170	0.230	4	216	0.160	6	1.44	43.76
TTHERM_01546860	Swr1	SWR-C	137	0.186	5	134	0.099	10	1.87	86.75
TTHERM_00975380	Actin	SWR-C	106	0.144	7	114	0.085	12	1.70	69.84
TTHERM_00584840	Swi3	SWI/SNF	100	0.136	7	495	0.367	3	0.37	-63.10
TTHERM_00444700	PHD Finger	Other	57	0.077	13	67	0.050	20	1.55	55.39
TTHERM_00248390	Gcn5	SAGA	41	0.056	18	44	0.033	31	1.70	70.20
TTHERM_01005190	Arp6	SWR-C	41	0.056	18	45	0.033	30	1.66	66.42
TTHERM_00355040	AT-hook containing	Other	35	0.047	21	23	0.017	59	2.78	177.95
TTHERM_00357110	Swc4	SWR-C	32	0.043	23	33	0.024	41	1.77	77.12
TTHERM_00136450	ciliate-specific	Other	26	0.035	28	29	0.022	46	1.64	63.76
TTHERM_00388500	Swc2-like	SWR-C	22	0.030	34	18	0.013	75	2.23	123.25
TTHERM_00486440	Atxr3/Set1	HMT	22	0.030	34	154	0.114	9	0.26	-73.91
TTHERM_00790730	Ada2	SAGA	19	0.026	39	21	0.016	64	1.65	65.26
TTHERM_00046150	Hypothetical protein	Other	16	0.022	46	4	0.003	337	7.31	630.62
TTHERM_00925560	Snf12	SWI/SNF	13	0.018	57	90	0.067	15	0.26	-73.62
TTHERM_01245640	Brg1	SWI/SNF	11	0.015	67	176	0.131	8	0.11	-88.58
TTHERM_00092790	SWI/SNF, ciliate specific	SWI/SNF	7	0.009	105	44	0.033	31	0.29	-70.94
TTHERM_00243900	Swi1	SWI/SNF	6	0.008	123	80	0.059	17	0.14	-86.30

* 5 hours after conjugation the compared fold of bait with respect to vegetative growth: did not change if fold= 1, increased if fold > 1 or decreased if fold < 1

7.3 Appendix 3 – Primers

7.3.1 Sequencing primers

M13F 5'-tgtaaaacgacggccagt-3'
M13R 5'-caggaaacagctatgac-3'
HN111 5'-tatcatcatcatctttgtaatcaaatatc-3'

7.3.2 Bd containing protein (TTHERM_00729230) primers

Restriction sites are underlined

7.3.2.1 Up stream TGA codon

FWR KpnI 5'-ccccGGTACCaaataatttttgaatgtattttggttttat-3'
RVR XhoI 5'-ccccCTCGAGaatcttcactctctttgtagtagtttcagt-3'

7.3.2.2 Down stream TGA codon

FWR NotI 5'-ccccGCGGCCGCagtagaaagaagaaattttaaagaagcttg-3'
RVR SacI 5'-ccccGAGCTCatatttctgctcataatatacaaactaaag-3'

7.3.3 ISWI1 (TTHERM_00388250) primers

7.3.3.1 Up stream TGA codon

FWR KpnI 5'-ccccGGTACCagagctttctggtaaagaggag-3'
RVR XhoI 5'-ccccCTCGAGctttttaactattttatcaattttactaaaat-3'

7.3.3.2 Down stream TGA codon

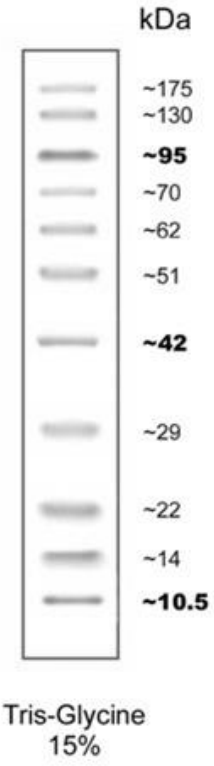
FWR NotI 5'-ccccGCGGCCGCtttattatgctctcaaaaaagtttta-3'
RVR SacI 5'-ccccGAGCTCatagttatagctaaatttcttgtttgatta-3'

7.3.4 Synthetic Ibd1 gene primers

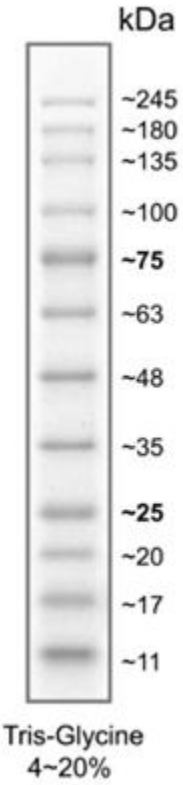
FWR NdeI 5'-ggggCATATGatgtctggtttcaagtcttccctga-3'
RVR BamHI 5'-ggggGGATCCttaaattttaacgcgcttcgtcg-3'

7.4 Appendix 4 – Protein ladder

PiNK Plus Prestained Protein Ladder (FroggaBio)



BLUeye Prestained Protein Ladder (FroggaBio)



7.5 Appendix 5 – Peptide Array

Example analysis of positive control G9a protein

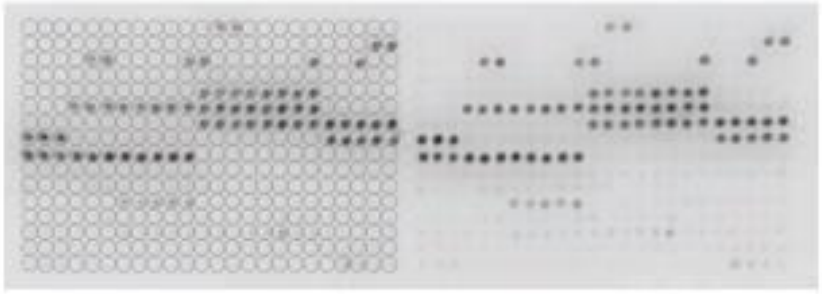
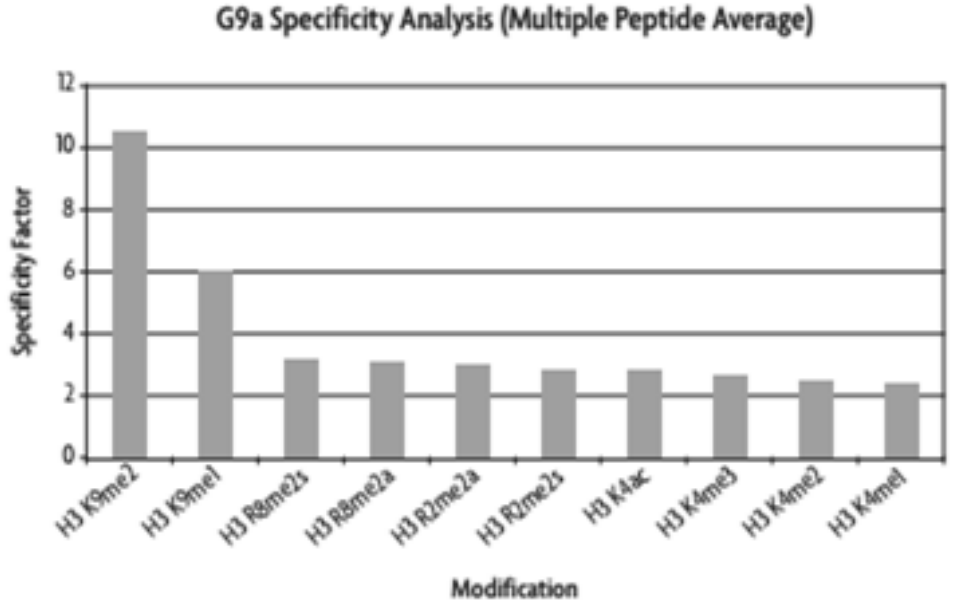


Image of ECL detection of positive control Recombinant His-G9a protein.

The positive control His-G9a protein was incubated with the MODified Histone Peptide array in combination with a c-Myc antibody (Active Motif Catalog No. 39502) for control peptide P21 according to the MODified Histone Protein Domain Binding manual at a concentration of 10 nM for 2 hours. The anti-His6-tag antibody and c-Myc antibody were each used at a 1:6,000 dilution, followed by incubation with the anti-mouse HRP-conjugated antibody at a 1:2,500 dilution and ECL detection. The chemiluminescent signal from the array was captured with a CCD camera. The image shows the grid overlay from the Array Analyze software for spot identification on the left hand side of the array image.



Graphical analysis of the positive control G9a protein.

Active Motif's Array Analyze Software was used to analyze spot intensity from the chemiluminescent image. The results were graphed as a specificity factor, which is the ratio of the average intensity of all spots containing the modification divided by the average intensity of all spots not containing the modification. The graph shows the top ten most reactive histone modifications based on the specificity factor values. G9a is a known histone methyltransferase associated with the specific mono- and di-methylation of lysine 9 on Histone H3. The data shown here confirm the association of G9a with its preferred histone modifications.

Annotations for each site

o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o
o	A 1	A 2	A 3	A 4	A 5	A 6	A 7	A 8	A 9	A 10	A 11	A 12	A 13	A 14	A 15	A 16	A 17	A 18	A 19	A 20	A 21	A 22	A 23	A 24	o
o	B 1	B 2	B 3	B 4	B 5	B 6	B 7	B 8	B 9	B10	B11	B12	B13	B14	B15	B16	B17	B18	B19	B20	B21	B22	B23	B24	o
o	C 1	C 2	C 3	C 4	C 5	C 6	C 7	C 8	C 9	C10	C11	C12	C13	C14	C15	C16	C17	C18	C19	C20	C21	C22	C23	C24	o
o	D 1	D 2	D 3	D 4	D 5	D 6	D 7	D 8	D 9	D10	D11	D12	D13	D14	D15	D16	D17	D18	D19	D20	D21	D22	D23	D24	o
o	E 1	E 2	E 3	E 4	E 5	E 6	E 7	E 8	E 9	E10	E11	E12	E13	E14	E15	E16	E17	E18	E19	E20	E21	E22	E23	E24	o
o	F 1	F 2	F 3	F 4	F 5	F 6	F 7	F 8	F 9	F 10	F 11	F 12	F 13	F 14	F 15	F 16	F 17	F 18	F 19	F 20	F 21	F 22	F 23	F 24	o
o	G 1	G 2	G 3	G 4	G 5	G 6	G 7	G 8	G 9	G10	G11	G12	G13	G14	G15	G16	G17	G18	G19	G20	G21	G22	G23	G24	o
o	H 1	H 2	H 3	H 4	H 5	H 6	H 7	H 8	H 9	H10	H11	H12	H13	H14	H15	H16	H17	H18	H19	H20	H21	H22	H23	H24	o
o	I 1	I 2	I 3	I 4	I 5	I 6	I 7	I 8	I 9	I10	I11	I12	I13	I14	I15	I16	I17	I18	I19	I20	I21	I22	I23	I24	o
o	J 1	J 2	J 3	J 4	J 5	J 6	J 7	J 8	J 9	J10	J11	J12	J13	J14	J15	J16	J17	J18	J19	J20	J21	J22	J23	J24	o
o	K 1	K 2	K 3	K 4	K 5	K 6	K 7	K 8	K 9	K10	K11	K12	K13	K14	K15	K16	K17	K18	K19	K20	K21	K22	K23	K24	o
o	L 1	L 2	L 3	L 4	L 5	L 6	L 7	L 8	L 9	L10	L11	L12	L13	L14	L15	L16	L17	L18	L19	L20	L21	L22	L23	L24	o
o	M 1	M 2	M 3	M 4	M 5	M 6	M 7	M 8	M 9	M10	M11	M12	M13	M14	M15	M16	M17	M18	M19	M20	M21	M22	M23	M24	o
o	N 1	N 2	N 3	N 4	N 5	N 6	N 7	N 8	N 9	N10	N11	N12	N13	N14	N15	N16	N17	N18	N19	N20	N21	N22	N23	N24	o
o	O 1	O 2	O 3	O 4	O 5	O 6	O 7	O 8	O 9	O10	O11	O12	O13	O14	O15	O16	O17	O18	O19	O20	O21	O22	O23	O24	o
o	P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8	P 9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	P21	P22	P23	P24	o
o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o	o

Active Motif MODified™ Histone Peptide Array* Catalog Nos. 13001 & 13005

*CelluSpots™ arrays are manufactured under license by INTAVIS Bioanalytical Instruments AG

Location	*Peptide sequence*	*Name*	*Mod1*	*Mod2*	*Mod 3*	*Mod 4*	*N-terminus*
A 1	A R T K Q T A R K S T G G K A P R K Q	H3 1-19	unmod				free
A 2	A Rme2s T K Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s				free
A 3	A Rme2a T K Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a				free
A 4	A Cit T K Q T A R K S T G G K A P R K Q	H3 1-19	R2Citr				free
A 5	A R pT K Q T A R K S T G G K A P R K Q	H3 1-19	T3P				free
A 6	A R T Kme1 Q T A R K S T G G K A P R K Q	H3 1-19	K4me1				free
A 7	A R T Kme2 Q T A R K S T G G K A P R K Q	H3 1-19	K4me2				free
A 8	A R T Kme3 Q T A R K S T G G K A P R K Q	H3 1-19	K4me3				free
A 9	A R T Kac Q T A R K S T G G K A P R K Q	H3 1-19	K4ac				free
A10	A R T K Q T A Rme2s K S T G G K A P R K Q	H3 1-19	R8me2s				free
A11	A R T K Q T A Rme2a K S T G G K A P R K Q	H3 1-19	R8me2a				free
A12	A R T K Q T A Cit K S T G G K A P R K Q	H3 1-19	R8Citr				free
A13	A R T K Q T A R Kme1 S T G G K A P R K Q	H3 1-19	K9me1				free
A14	A R T K Q T A R Kme2 S T G G K A P R K Q	H3 1-19	K9me2				free
A15	A R T K Q T A R Kme3 S T G G K A P R K Q	H3 1-19	K9me3				free
A16	A R T K Q T A R Kac S T G G K A P R K Q	H3 1-19	K9ac				free
A17	A R T K Q T A R K pS T G G K A P R K Q	H3 1-19	S10P				free
A18	A R T K Q T A R K S pT G G K A P R K Q	H3 1-19	T11P				free
A19	A R T K Q T A R K S T G G Kac A P R K Q	H3 1-19	K14ac				free

A20	A Rme2s pT K Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	T3P			free
A21	A Rme2s T Kme1 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	K4me1			free
A22	A Rme2s T Kme2 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	K4me2			free
A23	A Rme2s T Kme3 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	K4me3			free
A24	A Rme2s T Kac Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	K4ac			free
B 1	A Rme2a pT K Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	T3P			free
B 2	A Rme2a T Kme1 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	K4me1			free
B 3	A Rme2a T Kme2 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	K4me2			free
B 4	A Rme2a T Kme3 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	K4me3			free
B 5	A Rme2a T Kac Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	K4ac			free
B 6	A Cit pT K Q T A R K S T G G K A P R K Q	H3 1-19	R2Citr	T3P			free
B 7	A Cit T Kme1 Q T A R K S T G G K A P R K Q	H3 1-19	R2Citr	K4me1			free
B 8	A Cit T Kme2 Q T A R K S T G G K A P R K Q	H3 1-19	R2Citr	K4me2			free
B 9	A Cit T Kme3 Q T A R K S T G G K A P R K Q	H3 1-19	R2Citr	K4me3			free
B10	A Cit T Kac Q T A R K S T G G K A P R K Q	H3 1-19	R2Citr	K4ac			free
B11	A R pT Kme1 Q T A R K S T G G K A P R K Q	H3 1-19	T3P	K4me1			free
B12	A R pT Kme2 Q T A R K S T G G K A P R K Q	H3 1-19	T3P	K4me2			free
B13	A R pT Kme3 Q T A R K S T G G K A P R K Q	H3 1-19	T3P	K4me3			free
B14	A R pT Kac Q T A R K S T G G K A P R K Q	H3 1-19	T3P	K4ac			free
B15	A Rme2s pT Kme1 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	T3P	K4me1		free
B16	A Rme2s pT Kme2 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	T3P	K4me2		free
B17	A Rme2s pT Kme3 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	T3P	K4me3		free
B18	A Rme2s pT Kac Q T A R K S T G G K A P R K Q	H3 1-19	R2me2s	T3P	K4ac		free
B19	A Rme2a pT Kme1 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	T3P	K4me1		free
B20	A Rme2a pT Kme2 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	T3P	K4me2		free
B21	A Rme2a pT Kme3 Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	T3P	K4me3		free
B22	A Rme2a pT Kac Q T A R K S T G G K A P R K Q	H3 1-19	R2me2a	T3P	K4ac		free
B23	A R T K Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R8me2s	K9me1			free
B24	A R T K Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R8me2s	K9me2			free
C 1	A R T K Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R8me2s	K9me3			free
C 2	A R T K Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R8me2s	K9ac			free
C 3	A R T K Q T A Rme2a K pS T G G K A P R K Q	H3 1-19	R8me2s	S10P			free
C 4	A R T K Q T A Rme2a K S pT G G K A P R K Q	H3 1-19	R8me2s	T11P			free
C 5	A R T K Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R8me2a	K9me1			free
C 6	A R T K Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R8me2a	K9me2			free
C 7	A R T K Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R8me2a	K9me3			free
C 8	A R T K Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R8me2a	K9ac			free
C 9	A R T K Q T A Rme2a K pS T G G K A P R K Q	H3 1-19	R8me2a	S10P			free
C10	A R T K Q T A Rme2a K S pT G G K A P R K Q	H3 1-19	R8me2a	T11P			free
C11	A R T K Q T A Cit Kme1 S T G G K A P R K Q	H3 1-19	R8Citr	K9me1			free
C12	A R T K Q T A Cit Kme2 S T G G K A P R K Q	H3 1-19	R8Citr	K9me2			free
C13	A R T K Q T A Cit Kme3 S T G G K A P R K Q	H3 1-19	R8Citr	K9me3			free
C14	A R T K Q T A Cit Kac S T G G K A P R K Q	H3 1-19	R8Citr	K9ac			free
C15	A R T K Q T A Cit K pS T G G K A P R K Q	H3 1-19	R8Citr	S10P			free
C16	A R T K Q T A Cit K S pT G G K A P R K Q	H3 1-19	R8Citr	T11P			free

C17	A R T K Q T A R Kme1 pS T G G K A P R K Q	H3 1-19	K9me1	S10P			free
C18	A R T K Q T A R Kme1 S pT G G K A P R K Q	H3 1-19	K9me1	T11P			free
C19	A R T K Q T A R Kme1 S T G G Kac A P R K Q	H3 1-19	K9me1	K14ac			free
C20	A R T K Q T A R Kme2 pS T G G K A P R K Q	H3 1-19	K9me2	S10P			free
C21	A R T K Q T A R Kme2 S pT G G K A P R K Q	H3 1-19	K9me2	T11P			free
C22	A R T K Q T A R Kme2 S T G G Kac A P R K Q	H3 1-19	K9me2	K14ac			free
C23	A R T K Q T A R Kme3 pS T G G K A P R K Q	H3 1-19	K9me3	S10P			free
C24	A R T K Q T A R Kme3 S pT G G K A P R K Q	H3 1-19	K9me3	T11P			free
D 1	A R T K Q T A R Kme3 S T G G Kac A P R K Q	H3 1-19	K9me3	K14ac			free
D 2	A R T K Q T A R Kac pS T G G K A P R K Q	H3 1-19	K9ac	S10P			free
D 3	A R T K Q T A R Kac S pT G G K A P R K Q	H3 1-19	K9ac	T11P			free
D 4	A R T K Q T A R Kac S T G G Kac A P R K Q	H3 1-19	K9ac	K14ac			free
D 5	A R T K Q T A R K pS pT G G K A P R K Q	H3 1-19	S10P	T11P			free
D 6	A R T K Q T A R K pS T G G Kac A P R K Q	H3 1-19	S10P	K14ac			free
D 7	A R T K Q T A R K S pT G G Kac A P R K Q	H3 1-19	T11P	K14ac			free
D 8	A R T K Q T A Rme2s Kme1 pS T G G K A P R K Q	H3 1-19	R8me2s	K9me1	S10P		free
D 9	A R T K Q T A Rme2s Kme2 pS T G G K A P R K Q	H3 1-19	R8me2s	K9me2	S10P		free
D10	A R T K Q T A Rme2s Kme3 pS T G G K A P R K Q	H3 1-19	R8me2s	K9me3	S10P		free
D11	A R T K Q T A Rme2s Kac pS T G G K A P R K Q	H3 1-19	R8me2s	K9ac	S10P		free
D12	A R T K Q T A Rme2s Kme1 S pT G G K A P R K Q	H3 1-19	R8me2s	K9me1	T11P		free
D13	A R T K Q T A Rme2s Kme2 S pT G G K A P R K Q	H3 1-19	R8me2s	K9me2	T11P		free
D14	A R T K Q T A Rme2s Kme3 S pT G G K A P R K Q	H3 1-19	R8me2s	K9me3	T11P		free
D15	A R T K Q T A Rme2s Kac S pT G G K A P R K Q	H3 1-19	R8me2s	K9ac	T11P		free
D16	A R T K Q T A Rme2a Kme1 pS T G G K A P R K Q	H3 1-19	R8me2a	K9me1	S10P		free
D17	A R T K Q T A Rme2a Kme2 pS T G G K A P R K Q	H3 1-19	R8me2a	K9me2	S10P		free
D18	A R T K Q T A Rme2a Kme3 pS T G G K A P R K Q	H3 1-19	R8me2a	K9me3	S10P		free
D19	A R T K Q T A Rme2a Kac pS T G G K A P R K Q	H3 1-19	R8me2a	K9ac	S10P		free
D20	A R T K Q T A Rme2a Kme1 S pT G G K A P R K Q	H3 1-19	R8me2a	K9me1	T11P		free
D21	A R T K Q T A Rme2a Kme2 S pT G G K A P R K Q	H3 1-19	R8me2a	K9me2	T11P		free
D22	A R T K Q T A Rme2a Kme3 S pT G G K A P R K Q	H3 1-19	R8me2a	K9me3	T11P		free
D23	A R T K Q T A Rme2a Kac S pT G G K A P R K Q	H3 1-19	R8me2a	K9ac	T11P		free
D24	A R T K Q T A Rme2a Kme1 pS pT G G K A P R K Q	H3 1-19	R8me2a	K9me1	S10P	T11P	free
E 1	A R T K Q T A Rme2a Kme2 pS pT G G K A P R K Q	H3 1-19	R8me2a	K9me2	S10P	T11P	free
E 2	A R T K Q T A Rme2a Kme3 pS pT G G K A P R K Q	H3 1-19	R8me2a	K9me3	S10P	T11P	free
E 3	A R T K Q T A Rme2a Kac pS pT G G K A P R K Q	H3 1-19	R8me2a	K9ac	S10P	T11P	free
E 4	A Rme2s T Kme1 Q T A Rme2s K S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2s		free
E 5	A Rme2s T Kme2 Q T A Rme2s K S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2s		free
E 6	A Rme2s T Kme3 Q T A Rme2s K S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2s		free
E 7	A Rme2s T Kac Q T A Rme2s K S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2s		free
E 8	A Rme2a T Kme1 Q T A Rme2a K S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2a		free
E 9	A Rme2a T Kme2 Q T A Rme2a K S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2a		free
E10	A Rme2a T Kme3 Q T A Rme2a K S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2a		free
E11	A Rme2a T Kac Q T A Rme2a K S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2a		free
E12	A Rme2s T Kme1 Q T A R Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	K9me1		free
E13	A Rme2s T Kme2 Q T A R Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	K9me1		free

E14	A Rme2s T Kme3 Q T A R Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	K9me1		free
E15	A Rme2s T Kac Q T A R Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	K9me1		free
E16	A Rme2a T Kme1 Q T A R Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	K9me2		free
E17	A Rme2a T Kme2 Q T A R Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	K9me2		free
E18	A Rme2a T Kme3 Q T A R Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	K9me2		free
E19	A Rme2a T Kac Q T A R Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	K9me2		free
E20	A Rme2s T Kme1 Q T A R Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	K9me3		free
E21	A Rme2s T Kme2 Q T A R Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	K9me3		free
E22	A Rme2s T Kme3 Q T A R Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	K9me3		free
E23	A Rme2s T Kac Q T A R Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	K9me3		free
E24	A Rme2a T Kme1 Q T A R Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	K9ac		free
F 1	A Rme2a T Kme2 Q T A R Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	K9ac		free
F 2	A Rme2a T Kme3 Q T A R Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	K9ac		free
F 3	A Rme2a T Kac Q T A R Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	K9ac		free
F 4	A R T Kme1 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	K4me1	R8me2s	K9me1		free
F 5	A R T Kme2 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	K4me2	R8me2s	K9me1		free
F 6	A R T Kme3 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	K4me3	R8me2s	K9me1		free
F 7	A R T Kac Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	K4ac	R8me2s	K9me1		free
F 8	A R T Kme1 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	K4me1	R8me2a	K9me1		free
F 9	A R T Kme2 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	K4me2	R8me2a	K9me1		free
F10	A R T Kme3 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	K4me3	R8me2a	K9me1		free
F11	A R T Kac Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	K4ac	R8me2a	K9me1		free
F12	A R T Kme1 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	K4me1	R8me2s	K9me2		free
F13	A R T Kme2 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	K4me2	R8me2s	K9me2		free
F14	A R T Kme3 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	K4me3	R8me2s	K9me2		free
F15	A R T Kac Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	K4ac	R8me2s	K9me2		free
F16	A R T Kme1 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	K4me1	R8me2a	K9me2		free
F17	A R T Kme2 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	K4me2	R8me2a	K9me2		free
F18	A R T Kme3 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	K4me3	R8me2a	K9me2		free
F19	A R T Kac Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	K4ac	R8me2a	K9me2		free
F20	A R T Kme1 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	K4me1	R8me2s	K9me3		free
F21	A R T Kme2 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	K4me2	R8me2s	K9me3		free
F22	A R T Kme3 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	K4me3	R8me2s	K9me3		free
F23	A R T Kac Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	K4ac	R8me2s	K9me3		free
F24	A R T Kme1 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	K4me1	R8me2a	K9me3		free
G 1	A R T Kme2 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	K4me2	R8me2a	K9me3		free
G 2	A R T Kme3 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	K4me3	R8me2a	K9me3		free
G 3	A R T Kac Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	K4ac	R8me2a	K9me3		free
G 4	A R T Kme1 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	K4me1	R8me2s	K9ac		free
G 5	A R T Kme2 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	K4me2	R8me2s	K9ac		free
G 6	A R T Kme3 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	K4me3	R8me2s	K9ac		free
G 7	A R T Kac Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	K4ac	R8me2s	K9ac		free
G 8	A R T Kme1 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	K4me1	R8me2a	K9ac		free
G 9	A R T Kme2 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	K4me2	R8me2a	K9ac		free
G10	A R T Kme3 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	K4me3	R8me2a	K9ac		free

G11	A R T Kac Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	K4ac	R8me2a	K9ac		free
G12	A Rme2s T Kme1 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2s	K9me1	free
G13	A Rme2s T Kme2 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2s	K9me1	free
G14	A Rme2s T Kme3 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2s	K9me1	free
G15	A Rme2s T Kac Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2s	K9me1	free
G16	A Rme2a T Kme1 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2s	K9me1	free
G17	A Rme2a T Kme2 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2s	K9me1	free
G18	A Rme2a T Kme3 Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2s	K9me1	free
G19	A Rme2a T Kac Q T A Rme2s Kme1 S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2s	K9me1	free
G20	A Rme2s T Kme1 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2s	K9me2	free
G21	A Rme2s T Kme2 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2s	K9me2	free
G22	A Rme2s T Kme3 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2s	K9me2	free
G23	A Rme2s T Kac Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2s	K9me2	free
G24	A Rme2a T Kme1 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2s	K9me2	free
H 1	A Rme2a T Kme2 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2s	K9me2	free
H 2	A Rme2a T Kme3 Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2s	K9me2	free
H 3	A Rme2a T Kac Q T A Rme2s Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2s	K9me2	free
H 4	A Rme2s T Kme1 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2s	K9me3	free
H 5	A Rme2s T Kme2 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2s	K9me3	free
H 6	A Rme2s T Kme3 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2s	K9me3	free
H 7	A Rme2s T Kac Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2s	K9me3	free
H 8	A Rme2a T Kme1 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2s	K9me3	free
H 9	A Rme2a T Kme2 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2s	K9me3	free
H10	A Rme2a T Kme3 Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2s	K9me3	free
H11	A Rme2a T Kac Q T A Rme2s Kme3 S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2s	K9me3	free
H12	A Rme2s T Kme1 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2s	K9ac	free
H13	A Rme2s T Kme2 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2s	K9ac	free
H14	A Rme2s T Kme3 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2s	K9ac	free
H15	A Rme2s T Kac Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2s	K9ac	free
H16	A Rme2a T Kme1 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2s	K9ac	free
H17	A Rme2a T Kme2 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2s	K9ac	free
H18	A Rme2a T Kme3 Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2s	K9ac	free
H19	A Rme2a T Kac Q T A Rme2s Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2s	K9ac	free
H20	A Rme2s T Kme1 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2a	K9me1	free
H21	A Rme2s T Kme2 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2a	K9me1	free
H22	A Rme2s T Kme3 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2a	K9me1	free
H23	A Rme2s T Kac Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2a	K9me1	free
H24	A Rme2a T Kme1 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2a	K9me1	free
I 1	A Rme2a T Kme2 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2a	K9me1	free
I 2	A Rme2a T Kme3 Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2a	K9me1	free
I 3	A Rme2a T Kac Q T A Rme2a Kme1 S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2a	K9me1	free
I 4	A Rme2s T Kme1 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2a	K9me2	free
I 5	A Rme2s T Kme2 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2a	K9me2	free
I 6	A Rme2s T Kme3 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2a	K9me2	free
I 7	A Rme2s T Kac Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2a	K9me2	free

I 8	A Rme2a T Kme1 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2a	K9me2	free
I 9	A Rme2a T Kme2 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2a	K9me2	free
I10	A Rme2a T Kme3 Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2a	K9me2	free
I11	A Rme2a T Kac Q T A Rme2a Kme2 S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2a	K9me2	free
I12	A Rme2s T Kme1 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2a	K9me3	free
I13	A Rme2s T Kme2 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2a	K9me3	free
I14	A Rme2s T Kme3 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2a	K9me3	free
I15	A Rme2s T Kac Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2a	K9me3	free
I16	A Rme2a T Kme1 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2a	K9me3	free
I17	A Rme2a T Kme2 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2a	K9me3	free
I18	A Rme2a T Kme3 Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2a	K9me3	free
I19	A Rme2a T Kac Q T A Rme2a Kme3 S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2a	K9me3	free
I20	A Rme2s T Kme1 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R2me2s	K4me1	R8me2a	K9ac	free
I21	A Rme2s T Kme2 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R2me2s	K4me2	R8me2a	K9ac	free
I22	A Rme2s T Kme3 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R2me2s	K4me3	R8me2a	K9ac	free
I23	A Rme2s T Kac Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R2me2s	K4ac	R8me2a	K9ac	free
I24	A Rme2a T Kme1 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me1	R8me2a	K9ac	free
J 1	A Rme2a T Kme2 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me2	R8me2a	K9ac	free
J 2	A Rme2a T Kme3 Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4me3	R8me2a	K9ac	free
J 3	A Rme2a T Kac Q T A Rme2a Kac S T G G K A P R K Q	H3 1-19	R2me2a	K4ac	R8me2a	K9ac	free
J 4	A R K S T G G K A P R K Q L A T K A A R	H3 7-26	unmod				acetylated
J 5	A R K S T G G Kac A P R K Q L A T K A A R	H3 7-26	K14ac				acetylated
J 6	A R K pS T G G Kac A P R K Q L A T K A A R	H3 7-26	K14ac	S10P			acetylated
J 7	A R K S pT G G Kac A P R K Q L A T K A A R	H3 7-26	K14ac	T11P			acetylated
J 8	A R K S T G G K A P Rme2s K Q L A T K A A R	H3 7-26	R17me2s				acetylated
J 9	A R K S T G G K A P Rme2a K Q L A T K A A R	H3 7-26	R17me2a				acetylated
J10	A R K S T G G K A P Cit K Q L A T K A A R	H3 7-26	R17Citr				acetylated
J11	A R K S T G G K A P R Kac Q L A T K A A R	H3 7-26	K18ac				acetylated
J12	A R K S T G G Kac A P Rme2s K Q L A T K A A R	H3 7-26	K14ac	R17me2s			acetylated
J13	A R K S T G G Kac A P Rme2a K Q L A T K A A R	H3 7-26	K14ac	R17me2a			acetylated
J14	A R K S T G G Kac A P R Kac Q L A T K A A R	H3 7-26	K14ac	K18ac			acetylated
J15	A R K S T G G K A P Rme2s Kac Q L A T K A A R	H3 7-26	R17me2s	K18ac			acetylated
J16	A R K S T G G K A P Rme2a Kac Q L A T K A A R	H3 7-26	R17me2a	K18ac			acetylated
J17	A R K S T G G K A P Cit Kac Q L A T K A A R	H3 7-26	R17Citr	K18ac			acetylated
J18	A R K S T G G Kac A P Rme2s Kac Q L A T K A A R	H3 7-26	K14ac	R17me2s	K18ac		acetylated
J19	A R K S T G G Kac A P Rme2a Kac Q L A T K A A R	H3 7-26	K14ac	R17me2a	K18ac		acetylated
J20	P R K Q L A T K A A R K S A P A T G G	H3 16-35	unmod				acetylated
J21	P R K Q L A T K A A Rme2s K S A P A T G G	H3 16-35	R26me2s				acetylated
J22	P R K Q L A T K A A Rme2a K S A P A T G G	H3 16-35	R26me2a				acetylated
J23	P R K Q L A T K A A Cit K S A P A T G G	H3 16-35	R26Citr				acetylated
J24	P R K Q L A T K A A R Kme1 S A P A T G G	H3 16-35	K27me1				acetylated
K 1	P R K Q L A T K A A R Kme2 S A P A T G G	H3 16-35	K27me2				acetylated
K 2	P R K Q L A T K A A R Kme3 S A P A T G G	H3 16-35	K27me3				acetylated
K 3	P R K Q L A T K A A R Kac S A P A T G G	H3 16-35	K27ac				acetylated
K 4	P R K Q L A T K A A R K pS A P A T G G	H3 16-35	S28P				acetylated

K 5	P R K Q L A T K A A Rme2s Kme1 S A P A T G G	H3 16-35	R26me2s	K27me1			acetylated
K 6	P R K Q L A T K A A Rme2s Kme2 S A P A T G G	H3 16-35	R26me2s	K27me2			acetylated
K 7	P R K Q L A T K A A Rme2s Kme3 S A P A T G G	H3 16-35	R26me2s	K27me3			acetylated
K 8	P R K Q L A T K A A Rme2s Kac S A P A T G G	H3 16-35	R26me2s	K27ac			acetylated
K 9	P R K Q L A T K A A Rme2s K pS A P A T G G	H3 16-35	R26me2s	S28P			acetylated
K10	P R K Q L A T K A A Rme2a Kme1 S A P A T G G	H3 16-35	R26me2a	K27me1			acetylated
K11	P R K Q L A T K A A Rme2a Kme2 S A P A T G G	H3 16-35	R26me2a	K27me2			acetylated
K12	P R K Q L A T K A A Rme2a Kme3 S A P A T G G	H3 16-35	R26me2a	K27me3			acetylated
K13	P R K Q L A T K A A Rme2a Kac S A P A T G G	H3 16-35	R26me2a	K27ac			acetylated
K14	P R K Q L A T K A A Rme2a K pS A P A T G G	H3 16-35	R26me2a	S28P			acetylated
K15	P R K Q L A T K A A Cit Kme1 S A P A T G G	H3 16-35	R26Citr	K27me1			acetylated
K16	P R K Q L A T K A A Cit Kme2 S A P A T G G	H3 16-35	R26Citr	K27me2			acetylated
K17	P R K Q L A T K A A Cit Kme3 S A P A T G G	H3 16-35	R26Citr	K27me3			acetylated
K18	P R K Q L A T K A A Cit K pS A P A T G G	H3 16-35	R26Citr	S28P			acetylated
K19	P R K Q L A T K A A R Kme1 pS A P A T G G	H3 16-35	K27me1	S28P			acetylated
K20	P R K Q L A T K A A R Kme2 pS A P A T G G	H3 16-35	K27me2	S28P			acetylated
K21	P R K Q L A T K A A R Kme3 pS A P A T G G	H3 16-35	K27me3	S28P			acetylated
K22	P R K Q L A T K A A R Kac pS A P A T G G	H3 16-35	K27ac	S28P			acetylated
K23	P R K Q L A T K A A Rme2s Kme1 pS A P A T G G	H3 16-35	R26me2s	K27me1	S28P		acetylated
K24	P R K Q L A T K A A Rme2s Kme2 pS A P A T G G	H3 16-35	R26me2s	K27me2	S28P		acetylated
L 1	P R K Q L A T K A A Rme2s Kme3 pS A P A T G G	H3 16-35	R26me2s	K27me3	S28P		acetylated
L 2	P R K Q L A T K A A Rme2s Kac pS A P A T G G	H3 16-35	R26me2s	K27ac	S28P		acetylated
L 3	P R K Q L A T K A A Rme2a Kme1 pS A P A T G G	H3 16-35	R26me2a	K27me1	S28P		acetylated
L 4	P R K Q L A T K A A Rme2a Kme2 pS A P A T G G	H3 16-35	R26me2a	K27me2	S28P		acetylated
L 5	P R K Q L A T K A A Rme2a Kme3 pS A P A T G G	H3 16-35	R26me2a	K27me3	S28P		acetylated
L 6	P R K Q L A T K A A Rme2a Kac pS A P A T G G	H3 16-35	R26me2a	K27ac	S28P		acetylated
L 7	R K S A P A T G G V K K P H R Y R P G	H3 26-45	unmod				acetylated
L 8	R K S A P A T G G V Kme1 K P H R Y R P G	H3 26-45	K36me1				acetylated
L 9	R K S A P A T G G V Kme2 K P H R Y R P G	H3 26-45	K36me2				acetylated
L10	R K S A P A T G G V Kme3 K P H R Y R P G	H3 26-45	K36me3				acetylated
L11	R K S A P A T G G V Kac K P H R Y R P G	H3 26-45	K36ac				acetylated
L12	S G R G K G G K G L G K G G A K R H R	H4 1-19	unmod				free
L13	pS G R G K G G K G L G K G G A K R H R	H4 1-19	S1P				free
L14	S G Rme2s G K G G K G L G K G G A K R H R	H4 1-19	R3me2s				free
L15	S G Rme2a G K G G K G L G K G G A K R H R	H4 1-19	R3me2a				free
L16	S G R G Kac G G K G L G K G G A K R H R	H4 1-19	K5ac				free
L17	S G R G K G G Kac G L G K G G A K R H R	H4 1-19	K8ac				free
L18	S G R G K G G K G L G Kac G G A K R H R	H4 1-19	K12ac				free
L19	S G R G K G G K G L G K G G A Kac R H R	H4 1-19	K16ac				free
L20	pS G Rme2s G K G G K G L G K G G A K R H R	H4 1-19	S1P	R3me2s			free
L21	pS G Rme2a G K G G K G L G K G G A K R H R	H4 1-19	S1P	R3me2a			free
L22	pS G R G Kac G G K G L G K G G A K R H R	H4 1-19	S1P	K5ac			free
L23	S G Rme2s G Kac G G K G L G K G G A K R H R	H4 1-19	R3me2s	K5ac			free
L24	S G Rme2s G K G G Kac G L G K G G A K R H R	H4 1-19	R3me2s	K8ac			free
M 1	S G Rme2a G Kac G G K G L G K G G A K R H R	H4 1-19	R3me2a	K5ac			free

M 2	S G Rme2a G K G G Kac G L G K G G A K R H R	H4 1-19	R3me2a	K8ac			free
M 3	S G R G Kac G G Kac G L G K G G A K R H R	H4 1-19	K5ac	K8ac			free
M 4	S G R G K G G Kac G L G Kac G G A K R H R	H4 1-19	K8ac	K12ac			free
M 5	S G R G K G G Kac G L G K G G A Kac R H R	H4 1-19	K8ac	K16ac			free
M 6	S G R G K G G K G L G Kac G G A Kac R H R	H4 1-19	K12ac	K16ac			free
M 7	pS G Rme2s G Kac G G K G L G K G G A K R H R	H4 1-19	S1P	R3me2s	K5ac		free
M 8	pS G Rme2a G Kac G G K G L G K G G A K R H R	H4 1-19	S1P	R3me2a	K5ac		free
M 9	S G Rme2s G Kac G G Kac G L G K G G A K R H R	H4 1-19	R3me2s	K5ac	K8ac		free
M10	S G Rme2a G Kac G G Kac G L G K G G A K R H R	H4 1-19	R3me2a	K5ac	K8ac		free
M11	S G R G Kac G G Kac G L G Kac G G A K R H R	H4 1-19	K5ac	K8ac	K12ac		free
M12	S G R G K G G Kac G L G Kac G G A Kac R H R	H4 1-19	K8ac	K12ac	K16ac		free
M13	pS G Rme2s G Kac G G Kac G L G K G G A K R H R	H4 1-19	S1P	R3me2s	K5ac	K8ac	free
M14	pS G Rme2a G Kac G G Kac G L G K G G A K R H R	H4 1-19	S1P	R3me2a	K5ac	K8ac	free
M15	S G Rme2s G Kac G G Kac G L G Kac G G A K R H R	H4 1-19	R3me2s	K5ac	K8ac	K12ac	free
M16	S G Rme2a G Kac G G Kac G L G Kac G G A K R H R	H4 1-19	R3me2a	K5ac	K8ac	K12ac	free
M17	S G R G Kac G G Kac G L G Kac G G A Kac R H R	H4 1-19	K5ac	K8ac	K12ac	K16ac	free
M18	G K G G A K R H R K V L R D N I Q G I T	H4 11-30	unmod				acetylated
M19	G Kac G G A K R H R K V L R D N I Q G I T	H4 11-30	K12ac				acetylated
M20	G K G G A Kac R H R K V L R D N I Q G I T	H4 11-30	K16ac				acetylated
M21	G K G G A K Rme2s H R K V L R D N I Q G I T	H4 11-30	R17me2s				acetylated
M22	G K G G A K Rme2a H R K V L R D N I Q G I T	H4 11-30	R17me2a				acetylated
M23	G K G G A K R H Rme2s K V L R D N I Q G I T	H4 11-30	R19me2s				acetylated
M24	G K G G A K R H Rme2a K V L R D N I Q G I T	H4 11-30	R19me2a				acetylated
N 1	G K G G A K R H R Kme1 V L R D N I Q G I T	H4 11-30	K20me1				acetylated
N 2	G K G G A K R H R Kme2 V L R D N I Q G I T	H4 11-30	K20me2				acetylated
N 3	G K G G A K R H R Kme3 V L R D N I Q G I T	H4 11-30	K20me3				acetylated
N 4	G K G G A K R H R Kac V L R D N I Q G I T	H4 11-30	K20ac				acetylated
N 5	G K G G A K R H R K V L Rme2a D N I Q G I T	H4 11-30	R24me2a				acetylated
N 6	G K G G A K R H R K V L Rme2s D N I Q G I T	H4 11-30	R24me2s				acetylated
N 7	G Kac G G A Kac R H R K V L R D N I Q G I T	H4 11-30	K12ac	K16ac			acetylated
N 8	G K G G A Kac Rme2s H R K V L R D N I Q G I T	H4 11-30	K16ac	R17me2s			acetylated
N 9	G K G G A Kac Rme2a H R K V L R D N I Q G I T	H4 11-30	K16ac	R17me2a			acetylated
N10	G K G G A Kac R H Rme2s K V L R D N I Q G I T	H4 11-30	K16ac	R19me2s			acetylated
N11	G K G G A Kac R H Rme2a K V L R D N I Q G I T	H4 11-30	K16ac	R19me2a			acetylated
N12	G K G G A Kac R H R Kme1 V L R D N I Q G I T	H4 11-30	K16ac	K20me1			acetylated
N13	G K G G A Kac R H R Kme2 V L R D N I Q G I T	H4 11-30	K16ac	K20me2			acetylated
N14	G K G G A Kac R H R Kme3 V L R D N I Q G I T	H4 11-30	K16ac	K20me3			acetylated
N15	G K G G A Kac R H R Kac V L R D N I Q G I T	H4 11-30	K16ac	K20ac			acetylated
N16	G Kac G G A Kac R H R Kme1 V L R D N I Q G I T	H4 11-30	K12ac	K16ac	K20me1		acetylated
N17	G Kac G G A Kac R H R Kme2 V L R D N I Q G I T	H4 11-30	K12ac	K16ac	K20me2		acetylated
N18	G Kac G G A Kac R H R Kme3 V L R D N I Q G I T	H4 11-30	K12ac	K16ac	K20me3		acetylated
N19	G Kac G G A Kac R H R Kac V L R D N I Q G I T	H4 11-30	K12ac	K16ac	K20ac		acetylated
N20	G K G G A K R H Rme2a Kme1 V L R D N I Q G I T	H4 11-30	R19me2a	K20me1			acetylated
N21	G K G G A K R H Rme2a Kme2 V L R D N I Q G I T	H4 11-30	R19me2a	K20me2			acetylated
N22	G K G G A K R H Rme2a Kme3 V L R D N I Q G I T	H4 11-30	R19me2a	K20me3			acetylated

N23	G K G G A K R H Rme2a Kac V L R D N I Q G I T	H4 11-30	R19me2a	K20ac			acetylated
N24	G K G G A K R H Rme2s Kme1 V L R D N I Q G I T	H4 11-30	R19me2s	K20me1			acetylated
O 1	G K G G A K R H Rme2s Kme2 V L R D N I Q G I T	H4 11-30	R19me2s	K20me2			acetylated
O 2	G K G G A K R H Rme2s Kme3 V L R D N I Q G I T	H4 11-30	R19me2s	K20me3			acetylated
O 3	G K G G A K R H Rme2s Kac V L R D N I Q G I T	H4 11-30	R19me2s	K20ac			acetylated
O 4	G K G G A K R H R Kme1 V L Rme2a D N I Q G I T	H4 11-30	R24me2a	K20me1			acetylated
O 5	G K G G A K R H R Kme2 V L Rme2a D N I Q G I T	H4 11-30	R24me2a	K20me2			acetylated
O 6	G K G G A K R H R Kme3 V L Rme2a D N I Q G I T	H4 11-30	R24me2a	K20me3			acetylated
O 7	G K G G A K R H R Kac V L Rme2a D N I Q G I T	H4 11-30	R24me2a	K20ac			acetylated
O 8	G K G G A K R H R Kme1 V L Rme2s D N I Q G I T	H4 11-30	R24me2s	K20me1			acetylated
O 9	G K G G A K R H R Kme2 V L Rme2s D N I Q G I T	H4 11-30	R24me2s	K20me2			acetylated
O10	G K G G A K R H R Kme3 V L Rme2s D N I Q G I T	H4 11-30	R24me2s	K20me3			acetylated
O11	G K G G A K R H R Kac V L Rme2s D N I Q G I T	H4 11-30	R24me2s	K20ac			acetylated
O12	S G R G K Q G G K A R A K A K S R S S	H2a 1-19	unmod				free
O13	pS G R G K Q G G K A R A K A K S R S S	H2a 1-19	S1P				free
O14	S G R G Kac Q G G K A R A K A K S R S S	H2a 1-19	K5ac				free
O15	S G R G K Q G G Kac A R A K A K S R S S	H2a 1-19	K9ac				free
O16	S G R G K Q G G K A R A Kac A K S R S S	H2a 1-19	K13ac				free
O17	pS G R G Kac Q G G K A R A K A K S R S S	H2a 1-19	S1P	K5ac			free
O18	pS G R G K Q G G Kac A R A K A K S R S S	H2a 1-19	S1P	K9ac			free
O19	pS G R G K Q G G K A R A Kac A K S R S S	H2a 1-19	S1P	K13ac			free
O20	S G R G Kac Q G G Kac A R A K A K S R S S	H2a 1-19	K5ac	K9ac			free
O21	S G R G Kac Q G G K A R A Kac A K S R S S	H2a 1-19	K5ac	K13ac			free
O22	S G R G K Q G G Kac A R A Kac A K S R S S	H2a 1-19	K9ac	K13ac			free
O23	pS G R G Kac Q G G Kac A R A K A K S R S S	H2a 1-19	S1P	K5ac	K9ac		free
O24	pS G R G Kac Q G G K A R A Kac A K S R S S	H2a 1-19	S1P	K5ac	K13ac		free
P 1	pS G R G K Q G G Kac A R A Kac A K S R S S	H2a 1-19	S1P	K9ac	K13ac		free
P 2	S G R G Kac Q G G Kac A R A Kac A K S R S S	H2a 1-19	K5ac	K9ac	K13ac		free
P 3	pS G R G Kac Q G G Kac A R A Kac A K S R S S	H2a 1-19	S1P	K5ac	K9ac	K13ac	free
P 4	P D P A K S A P A P K K G S K K A V T	H2B 1-19	unmod				free
P 5	P D P A Kac S A P A P K K G S K K A V T	H2B 1-19	K5ac				free
P 6	P D P A K S A P A P K Kac G S K K A V T	H2B 1-19	K12ac				free
P 7	P D P A K S A P A P K K G pS K K A V T	H2B 1-19	S14P				free
P 8	P D P A K S A P A P K K G S Kac K A V T	H2B 1-19	K15ac				free
P 9	P D P A Kac S A P A P K Kac G S K K A V T	H2B 1-19	K5ac	K12ac			free
P10	P D P A Kac S A P A P K K G pS K K A V T	H2B 1-19	K5ac	S14P			free
P11	P D P A Kac S A P A P K K G S Kac K A V T	H2B 1-19	K5ac	K15ac			free
P12	P D P A K S A P A P K Kac G pS K K A V T	H2B 1-19	K12ac	S14P			free
P13	P D P A K S A P A P K Kac G S Kac K A V T	H2B 1-19	K12ac	K15ac			free
P14	P D P A K S A P A P K K G pS Kac K A V T	H2B 1-19	S14P	K15ac			free
P15	P D P A Kac S A P A P K Kac G pS K K A V T	H2B 1-19	K5ac	K12ac	S14P		free
P16	P D P A Kac S A P A P K Kac G S Kac K A V T	H2B 1-19	K5ac	K12ac	K15ac		free
P17	P D P A Kac S A P A P K K G pS Kac K A V T	H2B 1-19	K5ac	S14P	K15ac		free
P18	P D P A K S A P A P K Kac G pS Kac K A V T	H2B 1-19	K12ac	S14P	K15ac		free
P19	P D P A Kac S A P A P K Kac G pS Kac K A V T	H2B 1-19	K5ac	K12ac	S14P	K15ac	free

P20	Bio A A N W S H P Q F E K A A	Biotin, control peptide					biotinylated
P21	E Q K L I S E E D L A	c-myc tag					free
P22	HAc	neg. contol					acetylated
P23	K Kme1 Kme2 Kme3 Kac R Rme2s R Rme2a R Cit K Kme1 Kac Kme3 R K	background 01					acetylated
P24	R Rme2s K Kme1 Kac R Rme2a Kme2 K Kme3 R Kme1 Rme2s K Kac R K	background 02					acetylated

8 REFERENCES

Aalfs, J. D., & Kingston, R. E. (2000). Review: What does “chromatin remodeling” mean? *Trends in Biochemical Sciences*, *25*, 548–555.

Allis, C. D., Richman, R., Gorovsky, M. a, Ziegler, Y. S., Touchstone, B., Bradley, W. a, & Cook, R. G. (1986). hv1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes. *The Journal of Biological Chemistry*, *261*(4), 1941–1948.

Badenhorst, P., Voas, M., Rebay, I., & Wu, C. (2002). Biological functions of the ISWI chromatin remodeling complex NURF. *Genes and Development*, *16*, 3186–3198.

Baker, S. P., & Grant, P. a. (2007). The SAGA continues: expanding the cellular role of a transcriptional co-activator complex. *Oncogene*, *26*, 5329–5340.

Barker, N., Hurlstone, A., Musisi, H., Miles, A., Bienz, M., & Clevers, H. (2001). The chromatin remodelling factor Brg-1 interacts with beta-catenin to promote target gene activation. *The EMBO Journal*, *20*, 4935–4943.

Berger, S. L. (2007). The complex language of chromatin regulation during transcription. *Nature*, *447*, 407–412.

Biterge, B., & Schneider, R. (2014). Histone variants: Key players of chromatin. *Cell and Tissue Research*, *356*, 457–466.

Blackburn, E. H., Greider, C. W., & Szostak, J. W. (2006). Telomeres and telomerase: the path from maize, Tetrahymena and yeast to human cancer and aging. *Nature Medicine*, *12*, 1133–1138.

Bochar, D. a, Wang, L., Beniya, H., Kinev, A., Xue, Y., Lane, W. S., … Shiekhattar, R. (2000). BRCA1 is associated with a human SWI/SNF-related complex: linking chromatin remodeling to breast cancer. *Cell*, *102*, 257–265.

Breeden, L., & Nasmyth, K. (1987). Cell cycle control of the yeast HO gene: cis- and trans-acting regulators. *Cell*, *48*, 389–97.

Bright, L. J., Kambesis, N., Nelson, S. B., Jeong, B., & Turkewitz, A. P. (2010). Comprehensive analysis reveals dynamic and evolutionary plasticity of Rab GTPases and membrane traffic in Tetrahymena thermophila. *PLoS Genetics*, *6*, 1–18.

Brownell, J. E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D. G., Roth, S. Y., & Allis, C. D. (1996). Tetrahymena histone acetyltransferase: a homologue to yeast Gcn5p linking histone acetylation to gene activation. *Cell*, *84*, 843–851.

Bruns, P. J., & Cassidy-Hanley, D. (1999). Biolistic transformation of macro- and micronuclei. *Methods in Molecular Biology*, *62*, 501–512.

Burnette, N. W. (1981). "Western Blotting": Electrophoretic Transfer of Proteins from Sodium Dodecyl Sulfate-Polyacrylamide Gels to Unmodified Nitrocellulose and Radiographic Detection with Antibody and Radioiodinated Protein A. *Analytical Biochemistry*, *112*, 195–203.

Cairns, B. R., Lorch, Y., Li, Y., Zhang, M., Lacomis, L., Erdjument-Bromage, H., … Kornberg, R. D. (1996). RSC, an essential, abundant chromatin-remodeling complex. *Cell*, *87*, 1249–1260.

Cassidy-Hanley, D., Bowen, J., Lee, J. H., Cole, E., Verplank, L. A., Gaertip, J., … Bruns, P. J. (1997). Germline and Somatic Transformation of Mating Tetrahymena thermophila by Particle Bombardment. *Genetics Society of America*, *146*, 135–147.

Cech, T. R. (1990). Self-splicing and enzymatic activity of an intervening sequence RNA from Tetrahymena. *Bioscience Reports*, *24*, 362–385.

Cervantes, M. D., Hamilton, E. P., Xiong, J., Lawson, M. J., Yuan, D., Hadjithomas, M., … Orias, E. (2013). Selecting One of Several Mating Types through Gene Segment Joining and Deletion in Tetrahymena thermophila. *PLoS Biology*, *11*(3), e1001518.

Chalker, D. L. (2008). Dynamic nuclear reorganization during genome remodeling of Tetrahymena. *Biochimica et Biophysica Acta*, *1783*(11), 2130–6.

Chalker, D. L., Terza, A. La, Wilson, A., Kroenke, C. D., Yao, M., & Terza, A. L. A. (1999). Flanking Regulatory Sequences of the Tetrahymena R Deletion Element Determine the Boundaries of DNA Rearrangement Flanking Regulatory Sequences of the Tetrahymena R Deletion Element Determine the Boundaries of DNA Rearrangement. *Molecular and Cellular Biology*, *19*(8), 5631–5641.

Chalker, D. L., & Yao, M. C. (2001). Nongenic, bidirectional transcription precedes and may promote developmental DNA deletion in Tetrahymena thermophila. *Genes and Development*, *15*, 1287–1298.

Collins, K. (2012). Methods in Cell Biology Volume 109 - Tetrahymena thermophila. In *Academic Press* (p. 13).

Collins, R. E., Northrop, J. P., Horton, J. R., Lee, D. Y., Zhang, X., Stallcup, M. R., & Cheng, X. (2008). The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nature Structural & Molecular Biology*, *15*, 245–250.

Coyne, R. S., Nikiforov, M. a, Smothers, J. F., Allis, C. D., & Yao, M. C. (1999). Parental expression of the chromodomain protein Pdd1p is required for completion of programmed DNA elimination and nuclear differentiation. *Molecular Cell*, *4*(5), 865–872.

Cuatrecasas, P. (1970). Protein Purification by Affinity Chromatography. *Journal of Biological Sciences*, *245*, 3050–3065.

Dingwall, a K., Beek, S. J., McCallum, C. M., Tamkun, J. W., Kalpana, G. V, Goff, S. P., & Scott, M. P. (1995). The Drosophila snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex. *Molecular Biology of the Cell*, *6*, 777–791.

Eisen, J. a, Sweder, K. S., & Hanawalt, P. C. (1995). Evolution of the SNF2 family of proteins: subfamilies with distinct sequences and functions. *Nucleic Acids Research*, *23*, 2715–2723.

Erdel, F., Schubert, T., Marth, C., Längst, G., & Rippe, K. (2010). Human ISWI chromatin-remodeling complexes sample nucleosomes via transient binding reactions and become immobilized at active sites. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 19873–19878.

Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Craig, M., Meng, C. K. A. I., … Whitehouse, C. M. (2013). Electrospray Ionization of Large for Mass Spectrometry Biomolecules. *American Association for the Advancement of Science*, *246*(4926), 64–71.

Filippakopoulos, P., Picaud, S., Mangos, M., Keates, T., Lambert, J. P., Barsyte-Lovejoy, D., … Knapp, S. (2012). Histone recognition and large-scale structural analysis of the human bromodomain family. *Cell*, *149*, 214–231.

Fillingham, J. S., Garg, J., Tsao, N., Vythilingum, N., Nishikawa, T., & Pearlman, R. E. (2006). Molecular genetic analysis of an SNF2/brahma-related gene in Tetrahymena thermophila suggests roles in growth and nuclear development. *Eukaryotic Cell*, *5*, 1347–1359.

Fischle, W., Wang, Y., & Allis, C. D. (2003). Histone and chromatin cross-talk. *Current Opinion in Cell Biology*, *15*, 172–183.

Gaertig, J., Gu, L., Hai, B., & Gorovsky, M. a. (1994). High frequency vector-mediated transformation and gene replacement in Tetrahymena. *Nucleic Acids Research*, *22*, 5391–5398.

Garg, J., Lambert, J.-P., Karsou, A., Marquez, S., Nabeel-Shah, S., Bertucci, V., … Fillingham, J. S. (2013a). Conserved Asf1-importin β physical interaction in growth and sexual development in the ciliate Tetrahymena thermophila. *Journal of Proteomics*, *94*, 311–326.

Garg, J., Lambert, J.-P., Karsou, A., Marquez, S., Nabeel-Shah, S., Bertucci, V., … Fillingham, J. S. (2013b). Conserved Asf1-importin β physical interaction in growth and sexual development in the ciliate Tetrahymena thermophila. *Journal of Proteomics*, *94C*, 311–326.

Holliday, R. (2006). Epigenetics: A historical overview. *Epigenetics*, *1*, 76–80.

Kamakaka, R., & Hiller, N. (2005). Histone variants: deviants? *Genes and Development*, *19*, 295–310.

Kataoka, K., Schoeberl, U. E., & Mochizuki, K. (2010). Modules for C-terminal epitope tagging of Tetrahymena genes. *Journal of Microbiological Methods*, *82*, 342–346.

Kent, N. a., Karabetsou, N., Politis, P. K., & Mellor, J. (2001). In vivo chromatin remodeling by yeast ISWI homologs Isw1p and Isw2p. *Genes and Development*, *15*, 619–626.

Knappik, A., & Pluckthun, A. (1994). An Improved Affinity Tag Based on the FLAG Peptide for the Detection and Purification of Recombinant Antibody Fragments. *BioTechniques*, *17*, 754–761.

Korochkin, L. I. (2006). What is epigenetics. *Genetika*, *42*, 1156–1164.

Krogan, N. J., Keogh, M.-C., Datta, N., Sawa, C., Ryan, O. W., Ding, H., … Greenblatt, J. F. (2003). A Snf2 family ATPase complex required for recruitment of the histone H2A variant Htz1. *Molecular Cell*, *12*, 1565–1576.

Liu, X., Li, B., & GorovskyMA. (1996). Essential and nonessential histone H2A variants in Tetrahymena thermophila. *Molecular and Cellular Biology*, *16*, 4305–4311.

Liu, Y., Taverna, S. D., Muratore, T. L., Shabanowitz, J., Hunt, D. F., & Allis, C. D. (2007). RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in Tetrahymena. *Genes & Development*, *21*(12), 1530–45. doi:10.1101/gad.1544207

Logie, C., & Peterson, C. L. (1999). Purification and Biochemical Properties of Yeast SWI/SNF Complex. *Methods in Enzymology*, *304*, 726–741.

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

Madireddi, M. T., Coyne, R. S., Smothers, J. F., Mickey, K. M., Yao, M. C., & Allis, C. D. (1996). Pdd1p, a novel chromodomain-containing protein, links heterochromatin assembly and DNA elimination in Tetrahymena. *Cell*, *87*, 75–84.

Malone, C. D., Anderson, A. M., Motl, J. A., Rexer, C. H., & Chalker, D. L. (2005). Germ Line Transcripts Are Processed by a Dicer-Like Protein That Is Essential for Developmentally Programmed Genome Rearrangements of Tetrahymena thermophila. *Molecular and Cellular Biology*, *25*, 9151–9164.

Martindale, D. W., Allis, C. D., & Bruns, P. J. (1982). Conjugation in Tetrahymena thermophila. A Temporal Analysis of Cytological Stages. *Experimental Cell Research*, *140*, 227–236.

Martindale, D. W., Allis, C. D., & Bruns, P. J. (1985). RNA and protein Synthesis During Meiotic Prophase in Tetrahymena thermophila. *Protozoology, 32*, 345–350.

Merriam, E. V, & Bruns, P. J. (1988). Phenotypic assortment in Tetrahymena thermophila: assortment kinetics of antibiotic-resistance markers, tsA, death, and the highly amplified rDNA locus. *Genetics*, *120*, 389–395.

Mochizuki, K. (2012). Developmentally programmed, RNA-directed genome rearrangement in Tetrahymena. *Development Growth and Differentiation*, *54*, 108–119.

Mochizuki, K., Fine, N. a, Fujisawa, T., & Gorovsky, M. a. (2002). Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in Tetrahymena. *Cell*, *110*, 689–699.

Mochizuki, K., & Gorovsky, M. a. (2004a). RNA polymerase II localizes in Tetrahymena thermophila meiotic micronuclei when micronuclear transcription associated with genome rearrangement occurs. *Eukaryotic Cell*, *3*, 1233–1240.

Mochizuki, K., & Gorovsky, M. a. (2004b). Small RNAs in genome rearrangement in Tetrahymena. *Current Opinion in Genetics and Development*, *14*, 181–187.

Muchardt, C., Reyes, J. C., Bourachot, B., Leguoy, E., & Yaniv, M. (1996). The hbrm and BRG-1 proteins, components of the human SNF/SWI complex, are phosphorylated and excluded from the condensed chromosomes during mitosis. *The EMBO Journal*, *15*, 3394–3402.

Nash, E. a., Nisbet, R. E. R., Barbrook, A. C., & Howe, C. J. (2008). Dinoflagellates: a mitochondrial genome all at sea. *Trends in Genetics*, *24*, 328–335.

Neigeborn, L., & Carlson, M. (1984). Genes Affecting The Regulation Of SUC2 Gene Expression By Glucose Repression In Saccharomyces cerevisiae. *Genetics Society of America*, *108*, 845–858.

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*, 1220–1231.

Nielsen, A. L., Sanchez, C., Ichinose, H., Cerviño, M., Lerouge, T., Chambon, P., & Losson, R. (2002). Selective interaction between the chromatin-remodeling factor BRG1 and the heterochromatin-associated protein HP1alpha. *The EMBO Journal*, *21*, 5797–5806.

Orias, E., Cervantes, M. D., & Hamilton, E. P. (2011). Tetrahymena thermophila, a unicellular eukaryote with separate germline and somatic genomes. *Research in Microbiology*, *162*, 578–586.

Otsuki, T., Furukawa, Y., Ikeda, K., Endo, H., Yamashita, T., Shinohara, A., … Liu, J. M. (2001). Fanconi anemia protein, FANCA, associates with BRG1, a component of the human SWI/SNF complex. *Human Molecular Genetics*, *10*, 2651–2660.

Owen-hughes, T. (2003). Pathways for remodelling chromatin. *Colworth Medal Lecture*, *31*, 893–905.

Pal, S., Vishwanath, S. N., Tempst, P., Sif, S., & Erdjument-bromage, H. (2004). Human SWI/SNF-Associated PRMT5 Methylates Histone H3 Arginine 8 and Negatively Regulates Expression of ST7 and NM23 Tumor Suppressor Genes. *Molecular and Cellular Biology*, *24*, 9630–9645.

Peterson, C. L., Dingwall, A., & Scott, M. P. (1994). Five SWI/SNF gene products are components of a large multisubunit complex required for transcriptional enhancement. *Proceedings of the National Academy of Sciences of the United States of America*, *91*, 2905–2908.

Reisman, D., Glaros, S., & Thompson, E. a. (2009). Review: The SWI/SNF complex and cancer. *Oncogene*, *28*, 1653–1668.

Rexer, C. H., & Chalker, D. L. (2007). Lia1p, a novel protein required during nuclear differentiation for genome-wide DNA rearrangements in Tetrahymena thermophila. *Eukaryotic Cell*, *6*, 1320–1329.

Roberts, C. W. M., & Orkin, S. H. (2004). Review: The SWI/SNF complex--chromatin and cancer. *Nature Reviews. Cancer*, *4*, 133–142.

Roy, S., & Morse, D. (2013). Transcription and Maturation of mRNA in Dinoflagellates. *Microorganisms*, *1*, 71–99.

Schwope, R. M., & Chalker, D. L. (2014). Mutations in Pdd1 reveal distinct requirements for its chromodomain and chromoshadow domain in directing histone methylation and heterochromatin elimination. *Eukaryotic Cell*, *13*, 190–201.

Sigh, J., Lindenstrøm, T., & Buchmann, K. (2004). Expression of pro-inflammatory cytokines in rainbow trout (Oncorhynchus mykiss) during an infection with Ichthyophthirius multifiliis. *Fish & Shellfish Immunology*, *17*, 75–86.

Skarra, D. V, Goudreault, M., Choi, H., Mullin, M., Nesvizhskii, A. I., Gingras, A.-C., & Honkanen, R. E. (2011). Label-free quantitative proteomics and SAINT analysis enable interactome mapping for the human Ser/Thr protein phosphatase 5. *Proteomics*, *11*, 1508–1516.

Slotkin, R. K., & Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nature Reviews. Genetics*, *8*(4), 272–285.

Smothers, J. F., Mizzen, C. A., Tubbert, M. M., Cook, R. G., & Allis, C. D. (1997). Pdd1p associates with germline-restricted chromatin and a second novel anlagen-enriched protein in developmentally programmed DNA elimination structures. *Development (Cambridge, England)*, *124*, 4537–4545.

Stargell, L. a., Bowen, J., Dadd, C. a., Dedon, P. C., Davis, M., Cook, R. G., … Gorovsky, M. a. (1993). Temporal and spatial association of his tone H2A variant hv1 with transcriptionally competent chromatin during nuclear development in Tetrahymena thermophila. *Genes and Development*, *7*, 2641–2651.

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

Tamkun, J. W., Deuring, R., Scott, M. P., Kissinger, M., Pattatucci, a M., Kaufman, T. C., & Kennison, J. a. (1992). brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell*, *68*, 561–572.

Tang, L., Nogales, E., & Ciferri, C. (2010). Review: Structure and function of SWI/SNF chromatin remodeling complexes and mechanistic implications for transcription. *Progress in Biophysics and Molecular Biology*, *102*, 122–128.

Taverna, S. D., Coyne, R. S., & Allis, C. D. (2002). Methylation of Histone H3 at Lysine 9 Targets Programmed DNA Elimination in Tetrahymena. *Cell*, *110*, 701–711.

Taverna, S. D., Ueberheide, B. M., Liu, Y., Tackett, A. J., Diaz, R. L., Shabanowitz, J., … Allis, C. D. (2007). Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proceedings of the National Academy of Sciences of the United States of America*, *104*, 2086–2091.

Teo, G., Liu, G., Zhang, J., Nesvizhskii, A. I., Gingras, A. C., & Choi, H. (2014). SAINTexpress: Improvements and additional features in Significance Analysis of INTeractome software. *Journal of Proteomics*, *100*, 37–43.

Turner, B. M., Birley, a J., & Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. *Cell*, *69*, 375–384.

Van, C., Williams, J. S., Kunkel, T. a., & Peterson, C. L. (2015). Deposition of histone H2A.Z by the SWR-C remodeling enzyme prevents genome instability. *DNA Repair*, *25*, 9–14.

Versteege, I., Sevenet, N., Lange, J., Rousseau-Merck, M.-F., Ambros, P., Handgretinger, R., … Delattre, O. (1998). Truncating mutations of hSNF5/INI1 in aggressive paediatric cancer. *Nature*, *394*, 203–206.

Vignali, M., Hassan, A. H., Neely, K. E., & Workman, J. L. (2000). ATP-Dependent Chromatin-Remodeling Complexes. *Molecular and Cellular Biology*, *20*, 1899–1910.

Wenkert, D., & Allis, C. D. (1984). Timing of the appearance of macronuclear-specific histone variant hv1 and gene expression in developing new macronuclei of Tetrahymena thermophila. *The Journal of Cell Biology*, *98*, 2107–2117.

Williams, N. E., & Hontst, J. E. (1995). Isolation and Fractionation of the Tetrahymena Cytoskeleton and Oral Apparatus. *Methods in Cell Biology*, *47*, 301–306.

Winston, F., & Carlson, M. (1992). Yeast SNF/SWI transcriptional activators and the SPT/SIN chromatin connection. *Trends in Genetics*, *8*, 387–391.

Wolfe, J. (1967). Structural Aspects of Amitosis: A Light and Electron Microscope Study of the Isolated Macronuclei of Parameium aurelia and Tetrahymena pyriformis. *Chromosoma*, *23*, 59–79.

Wong, A. K. C., Shanahan, F., Chen, Y., Lian, L., Ha, P., Hendricks, K., … Lees, E. (2000). BRG1 , a Component of the SWI-SNF Complex , Is Mutated in Multiple Human Tumor Cell Lines. *Cancer Reseach*, *60*, 6171–6177.

Yu, G. L., Hasson, M., & Blackburn, E. H. (1988). Circular ribosomal DNA plasmids transform Tetrahymena thermophila by homologous recombination with endogenous macronuclear ribosomal DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *85*(14), 5151–5155.