2012

# Dissecting Mammalian Replication-Independent Chromatin Assembly â€“ Biochemical and Structural Studies on the H3.3-Specific Histone Chaperones Hira and Daxx

Simon Elsässer


# DISSECTING MAMMALIAN REPLICATION-INDEPENDENT CHROMATIN ASSEMBLY – BIOCHEMICAL AND STRUCTURAL STUDIES ON THE H3.3-SPECIFIC HISTONE CHAPERONES HIRA AND DAXX

A Thesis Presented to the Faculty of

The Rockefeller University

in Partial Fulfillment of the Requirements for

the degree of Doctor of Philosophy

by

Simon Elsässer

June 2012

# DISSECTING MAMMALIAN REPLICATION-INDEPENDENT CHROMATIN ASSEMBLY – BIOCHEMICAL AND STRUCTURAL STUDIES ON THE H3.3-SPECIFIC HISTONE CHAPERONES HIRA AND DAXX

Simon Elsässer, Ph.D.

The Rockefeller University

Defended October 4, 2011

Histones are architectural proteins that wrap approximately two turns of DNA around their octameric core structure, constituting the fundamental packaging unit of eukaryotic chromatin – the nucleosome. Beyond their structural role, they regulate virtually all processes that act on or depend on DNA, such as replication, gene expression and maintenance of centromeres and telomeres. Despite the high conservation of the four core histones, H3, H4, H2A and H2B, throughout all eukaryotes, histone variants have emerged with variable degree of divergence from their canonical counterparts. These variants are thought to expand the regulatory repertoire of chromatin.

The histone replacement variant H3.3 is implicated in the formation and maintenance of specialized chromatin structure in metazoan cells. H3.3-containing nucleosomes are assembled in a replication-independent manner by means of dedicated histone chaperone proteins. By purifying H3.3-containing nuclear complexes from HeLa cells, I have identified proteins that mediate this replication-independent deposition of H3.3. I have been able to single out the death-associated protein Daxx as an *bona fide* histone H3.3 chaperone. Furthermore, I show that Daxx

cooperates with an ATP-dependent chromatin remodeler, ATRX, in depositing histone H3.3 at telomeres *in vivo* (Chapter 2). I show that an evolutionary conserved histone binding domain of Daxx interacts directly with variant-specific residues of the H3.3 core. I further identify homologies with the centromere-specific histone chaperone HJURP that hint to a conserved mechanism used across replication-independent chromatin assembly pathways (Chapter 3). I report crystallization of a complex of the Daxx histone binding domain and histones H3.3-H4 with the aim of solving the cocrystal structure to reveal how Daxx can tell the subtle differences in H3 variants apart. (Chapter 4).

Through a more detailed biochemical fractionation and genetic analysis, I found that Daxx constitutes a histone deposition pathway independent of the previously described HIRA histone chaperone complex (Chapter 5), corroborating the notion that distinct replication-independent chromatin assembly pathways maintain specialized chromatin at distinct genomic locations. To facilitate the study of histone H3 variants across biological model organisms, I have developed universal, variant-specific antibodies (Chapter 6). I showcase their application in deriving a semi-quantitative histone variant landscape in murine embryonic stem cells, on the basis of which the functions of H3.3 and replication-independent chromatin assembly can be studied in new detail.

In the light of recent findings that place Daxx and ATRX in the roles of potent tumor suppressors for a subset of neuroendocrine cancers, my studies will help to elucidate the molecular etiology of replication-independent chromatin assembly pathways in human disease.


Daxx! Thank goodness you came!!
We can always use more chaperones!!

## Acknowledgements

I would like to thank Dr. David Allis for his mentorship. He has giving me guidance but also a tremendous amount of freedom to pursue ideas or follow up on results that I found interesting. His creativity, enthusiasm and passion for science have infused my scientific thinking and have been a great source of motivation. I am very grateful for his fantastic support in developing my career.

I would like to thank the members of the Allis laboratory for making it an inspiring environment for learning and doing science, to discuss ideas, to get ample advice and also constructive criticism, or just have a good laugh. I am indebted to my mentor Dr. Alex Ruthenburg, who has made a huge commitment to helping my project to take off and has been never tired to give advice and share his thoughts. I am also very thankful to Dr. Peter Lewis, who taught me important lab skills and has been a great collaborator. I would also like to thank Dr. Laura Banaszynski for much experimental advice and generous reagent gifts, as well as Dr. Kyung-Min Noh, Dr. Sonja Stadler and Dr. Aaron Goldberg for experimental help. I am also grateful for advice from past and present lab members including Drs. Christina Hughes, Jung-Ae Kim, David Shechter, Ronen Sadeh, Greg Wang, Tom Milne, as well as my fellow graduate students Lindsey Baker, Sarah Whitcomb, Ying Liu, Fabio Casadio and Dr. Holger Dormann. I would also like to thank Dr. Ariane Chapgier for the ongoing collaboration and her helpfulness during her visit in the Allis lab. Jamie Winshell has been invaluable in making the lab run smoothly in all aspects.

I am indebted to the support from Dr. Dinshaw Patel in pursuing my x-ray crystallography studies. In his lab, I have enjoyed helpful advice from Drs. Jiamu Du, Jikui Song and Zhanxin Wang. I am very grateful for Dr. Hongda Huang's collaboration and help in crystallizing the Daxx complex. I would also like to thank Dr. Deena Oren for her commitment to providing excellent equipment and know-how at the structural biology resource center. Her help has greatly facilitated my structural work. I have also profited from other resource centers at Rockefeller, including the proteomics, spectroscopy, bio-imaging and genomics facilities.

I would like to thank my committee members, Drs. Tom Muir, Seth Darst and David Cowburn for their advice in shaping my project, as well as Dr. Kami Ahmad for his inspiring work that seeded my interest in histone variants and for serving as my external referee.

I am indebted the support by the David Rockefeller Graduate Program and the Dean's office staff. The Boehringer Ingelheim Fonds have also provided financial support and I am grateful of being part of their fellows community. I would like to thank the German National Academic Foundation and the German Academic Exchange service for travel fellowships.

I cannot thank my family enough, in particular my parents Helga and Wolfgang Elsässer, for shaping my personal and professional development. I would like to thank Birthe Meineke for critical reading of my thesis, just one example of her love and support.

# TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| ATP | Adenosinetriphosphat |
| cenH3 | centromeric H3 |
| ChIP | Chromatin immunoprecipitation |
| CRD | Chaperone recognition domain |
| DNAP | DNA Polymerase |
| ESC | embryonic stem cell |
| GST | Glutathion-S-transferase |
| HBD | Histone binding domain |
| HBH | Histone binding helix |
| HMM | Hidden Markov Model |
| LIC | Ligation-independent cloning |
| MALS | Multiangle Light Scattering |
| MNase | Micrococcal nuclease |
| MS/MS | Tandem mass spectrometry |
| NMR | Nuclear Magnetic Resonance |
| ORF | Open reading frame |
| PAGE | Polyacrylamide gel electrophoresis |
| PanNET | Pancreatic neuroendocrine tumors |
| PEV | Position effect variegation |
| PHD | Plant homeodomain |
| PTMs | posttranslational modifications |
| qPCR | quantitative PCR (real-time PCR) |
| RD | Replication-dependent |
| RE | Regulatory Element |
| RI | Replication-independent |
| RNAP | RNA Polymerase |
| SEC | Size exclusion chromatography (Gel filtration) |
| TAU | Triton-acid-urea gel |
| TES | Transcription end site |
| TFBS | Transcription factor binding site |
| TSS | Transcription start site |
| XNP | X-linked nuclear protein |

# 1 Introduction

## 1.1 Chromatin is the physiologic state of eukaryotic genomes

The nuclear genome of every eukaryotic cell is organized into chromosomes, large nucleoprotein – chromatin – entities that contain a single linear, double-stranded deoxyribonucleic acid (DNA) molecule. The extraordinary dynamic compaction of few to many megabases of DNA into the limited nuclear volume is achieved through an intricate interplay of fractal globule folding and ordered repetitive spooling. The fundamental repeating unit of chromatin is the nucleosome, in which histones proteins create an octameric core structure that wraps approximately two superhelical turns of DNA (Figure 1.1A) (Luger et al., 1997). A wealth of discoveries in the last decades has transformed the image of histones from static scaffolding proteins to dynamic modulators of virtually all processes that act on or depend on DNA, such as replication, gene expression and maintenance of centromeres and telomeres. Histones are subject to a wide variety of posttranslational modifications (PTMs) such as lysine acetylation and methylation many of which are characteristic of active or repressed gene states, as well as DNA replication, recombination, damage or repair events. The resulting variety of modification patterns constitute a 'histone code' that has the potential for cellular memory and transmission of epigenetic information through cell divisions (Bernstein et al., 2007; Ruthenburg et al., 2007; Strahl and Allis, 2000).

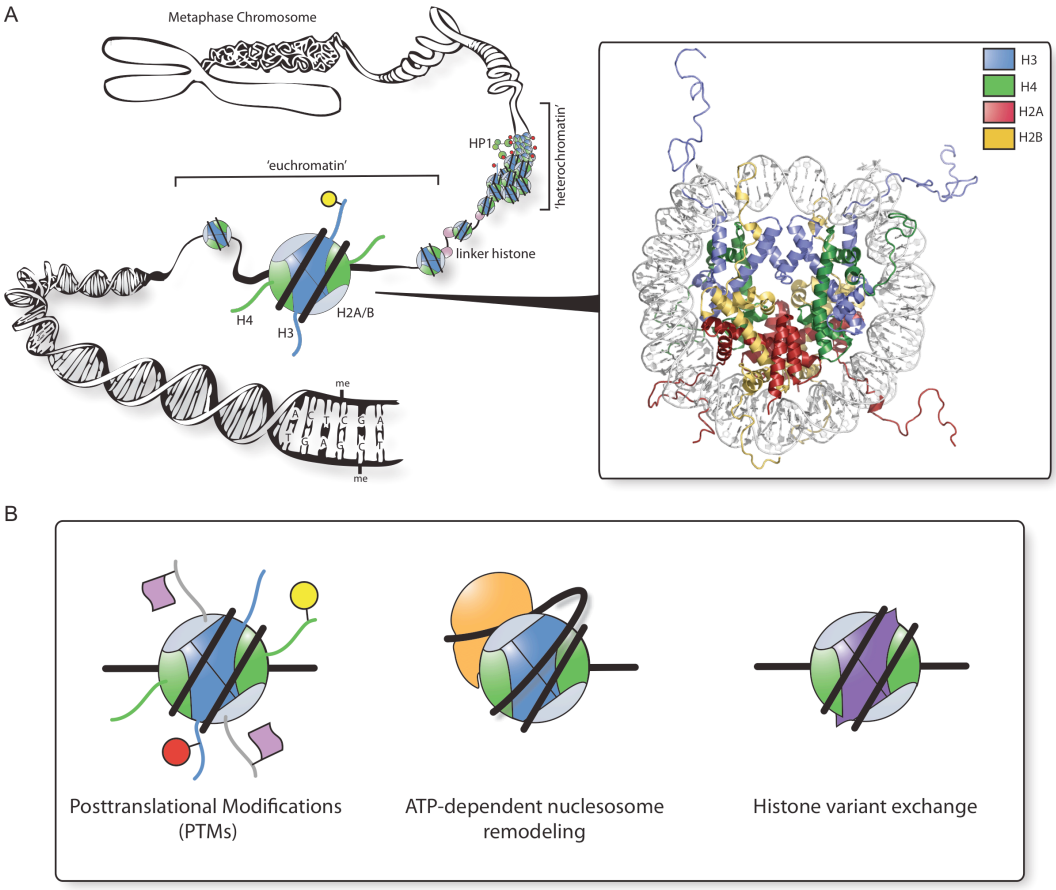



Figure 1.1: The eukaryotic genome is packaged in chromatin.

(A) The fundamental repeating unit of chromatin is the nucleosome, in which ~147 base pairs of DNA are wrapped around a histone octamer core (a H3-H4 tetramer, blue/green, flanked by two H2A/H2B dimers, gray). Perhaps unexpectedly given their seemingly strict structure (see crystal structure inset, PDB 1KX5) and monotonous repetitive appearance along the chromatin fiber, the arrangement and properties of nucleosomes allow for highly variable chromatin states. 'Euchromatin' describes an open, transcriptionally active chromatin state, whereas silent, condensed 'heterochromatin' is formed through nucleosome packing and the help of nucleosome-associated protein such as the linker histones and heterochromatin protein 1 (HP1). (B) Such and more subtle variation in chromatin states are thought to be dependent on posttranslational modification (PTMs) of the histones, the activity of ATP-dependent chromatin remodelers and the switching of the histone identity through introduction of histone variants. These mechanism are further integrated with modification of the DNA itself (methylation and hydroxymethylation of cytosine) and non-coding RNAs to create potentially heritable, or 'epigenetic', information.

Apart from the four major-type histones (H3, H4, H2A and H2B), metazoans have a number of histone variants. Like histone posttranslational modifications and nucleosome positioning and remodeling, the use of histone variants, such as centromeric H3, H3.3, H2A.X and H2A.Z, contributes to the regulatory repertoire of chromatin (Figure 1.1B).

### 1.1.1 Organization of eukaryotic chromosomes

The overall chromosome architecture can be observed most clearly during mitosis, when the highly compacted duplicate DNA molecules (sister chromatids) line up along the metaphase plate (Figure 1.1, A). The sister chromatids touch at the so-called centromere, which serves as kinetochore attachment for their subsequent separation. Telomeres, the tips of each sister chromatid, are specialized structures that protect the ends of the linear DNA molecule (Blackburn, 1991; Cleveland et al., 2003).

> *"Euchromatin is genicly active, heterochromatin genicly passive. Heterochromatic chromosomes or pieces of chromosomes contain no genes or somehow passive genes" – Emil Heitz, 1929*

Two fundamentally distinct compartments of eukaryotic chromatin, euchromatin and heterochromatin, were recognized not long after the discovery of chromosomes themselves in the 1880s (Zacharias, 1995). Heterochromatin has been defined as those parts of chromosomes that do not decondense in interphase and had been traced to the centers and distal ends of the chromosomes by Emil Heitz (Figure 1.2A, heterochromatin stains dark). From the linear arrangement of genes on chromosomes that explained genetic linkage

(first described by Thomas Hunt Morgan), Emil Heitz concluded that genes must be concentrated in the euchromatic 'arms' of chromosomes where the majority of linkage phenomena had been observed. The constitutively condensed centromeres and telomeres, on the other hand, would consequently be devoid of genes.

This notion has been corroborated and dissected in much more detail through cytogenetic studies, for example on the *Drosophila* polytene chromosomes. Gene activity and silencing has been found to be much more dynamic than the clear-cut distinction of euchromatin and heterochromatin could explain. One seminal finding was the facultative silencing of genes depending on their distance from centromeric or telomeric heterochromatin domains, so-called position-effect variegation (PEV), originally described by Hermann J Muller (Muller, 1930). The variegated or mosaic 'ON' or 'OFF' switching of genes placed in the neighborhood of heterochromatin is a pivotal example of a heritable phenotype that does not involve DNA mutation, a phenomenon therefore termed *epi*genetic.

It should be noted that, while histones and their modifications are often loosely attributed to carry 'epigenetic' information, there is consent in the field that this term should be reserved to describe proven *heritable* variation beyond DNA sequence changes that can rely on – or work independent of – chromosome inheritance (Bonasio et al., 2010; Henikoff and Shilatifard, 2011; Jenuwein and Allis, 2001; Strahl and Allis, 2000).

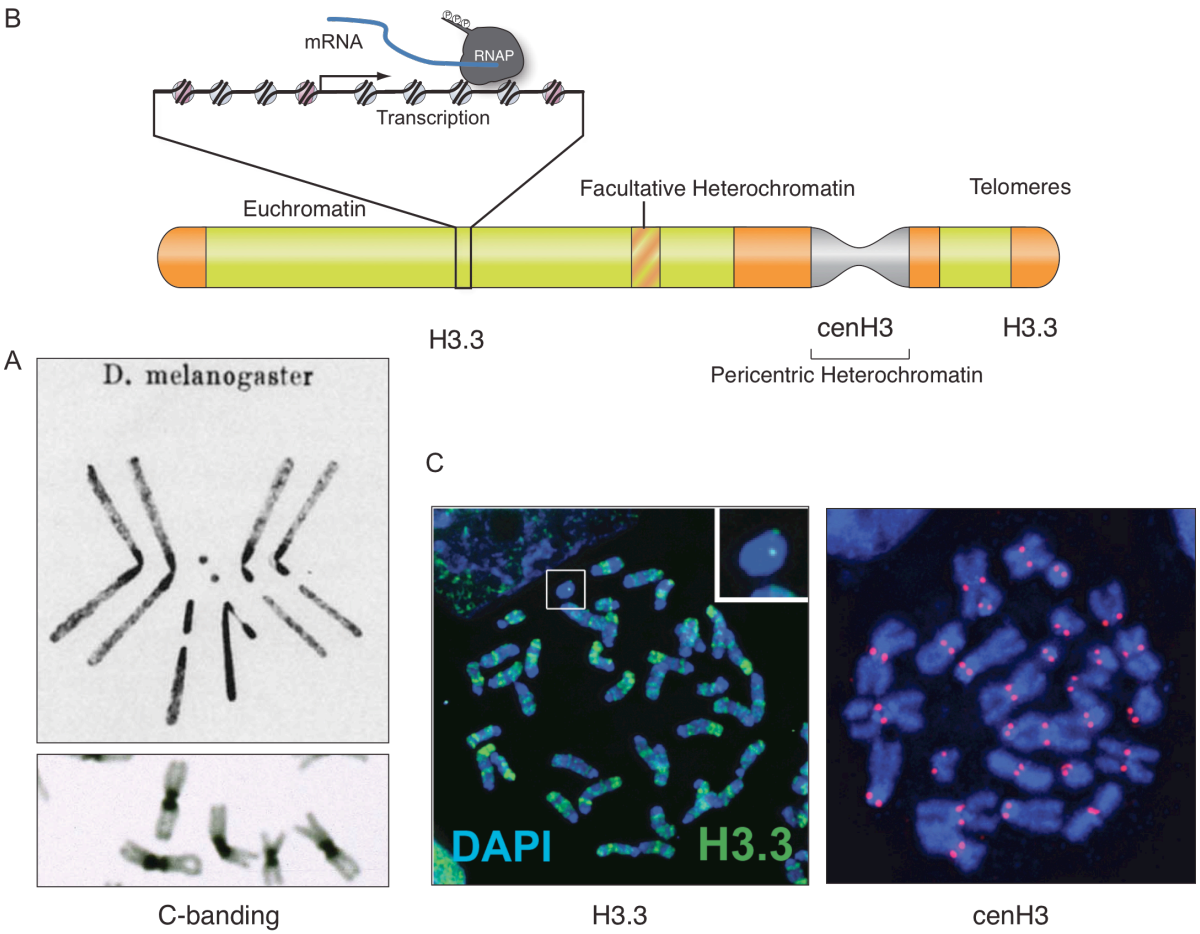



Figure 1.2: Chromosome structure and compartements as exemplified by histone variants. (A) Carmine staining of chromosomes as used to visualize heterochromatin (dense) and euchromatin (light stain), also called C banding. An original drawing by Heintz (top) and a contemporary photograph of fish chromosome carmine stain (bottom; C Ozouf-Costaz), with pericentric heterochromatin stained most darkly (B) Schematic of a chromosome. Heterochromatin locates to the region around the centromere (pericentromere) and to the ends of chromosomes (telomeres). Euchromatin on the 'arms' of chromosomes is gene-rich and transcriptionally active. However large regions of euchromatin can be silenced, giving rise to facultative heterochromatin. (C) H3 variants H3.3 and centromeric H3 (cenH3) localize to characteristic regions of chromosomes (murine ESC metaphase spread, DAPI-stained DNA blue, H3.3 green; A Goldberg). H3.3 is enriched at euchromatic arms of chromosomes (i.e., to one side of acrocentric murine ESC chromosomes shown here). Inset shows telomeric enrichment of H3.3 on the transcriptionally inactive Y-chromosomes. CenH3 defines centromeres and is absent from other regions of the chromosome (S2 cell metaphase spread, DAPI stained DNA blue, cenH3 purple; S. Erhardt).

Nevertheless, many chromatin components have been discovered in screens for modifiers of epigenetic states (Henikoff, 1990). Today, we know much about the identity of these modifiers and how they reversibly affect chromatin through histone PTMs, ATP-dependent remodeling or the use of histone variants. They therefore are implicated in encoding chromosome-bound heritable information (Figure 1.1B).

Of note, the two histone H3 variants that I will discuss in great detail throughout my thesis, H3.3 and cenH3, visibly localize to specific chromosomal regions: H3.3 localizes to the active arms of chromosomes and is broadly absent from heterochromatin except for discrete foci at telomeres in murine ESC (Figure 1.1C, left). CenH3, on the other hand, marks the centromeres but not adjacent pericentric heterochromatin.

Modern genome-wide mapping technologies, in particular next-gen sequencing, have changed our view from a birds-eye perspective of chromosomes to the macro level of individual base pair resolution. Those techniques now refine and redefine chromosomal domains through the analysis of the landscape of histone variants and PTMs, transcription machinery and other chromatin-associated proteins along the linear chromosomal arrangement. In particular, it has become clear that euchromatin and heterochromatin are not two uniform entities but can be subdivided into distinct classes as a function of their characteristic chromatin properties. Consequently, two pioneering publications recently proposed a Technicolor update of Heitz' black-and-white C-bands, delineating 5 (Filion et al., 2010) or 9 (Kharchenko et al., 2011) principal states or 'colors' of chromatin that capture fundamentally different subtypes of euchromatin and heterochromatin. In keeping with this metaphor, we can expect

the need for more technological advances before we will be able to see chromatin in its full color.

In my thesis work, I have been exploring the contribution of a particular histone variant, H3.3, to the properties of both eu- and hetero- chromatin. I have discovered a novel pathway for H3.3 to be incorporated into chromatin (Chapter 2) and made progress towards understanding how it's specificity for this variant is achieved (Chapter 3 and 4), as well as why animals might require distinct pathways for the same variant (Chapter 5). With novel immunological tools I gained a more detailed insights how histone H3 variants are distributed within euchromatin and heterochromatin, introducing a new nuance to the color spectrum (Chapter 6).

## 1.2 Replacement histone H3.3: new functions for an old variant[‡]

### 1.2.1 Preface

The description of the identity of the four core histones and their oligomeric arrangement in the nucleosome dates to the 1960s and 1970s (Kornberg and Thomas, 1974; Lewis, 1976; van Holde, 1988). Intriguingly, even in purified nucleosome core particle more than four non-identical protein species could be discerned if electrophoresis was carried out in the presence of non-ionic detergents (Zweidler, 1978). These species turned out to be related to, but not identical to the major-type histones H3, H2A, H2B and were therefore termed

[‡] Parts of this chapter have been published in Elsaesser, S. J., Goldberg, A. D., and Allis, C. D. (2010). New functions for an old variant: no substitute for histone H3.3. *Curr Opin Genet Dev* 20, 110–117.

variants (Franklin and Zweidler, 1977). Histone variants have subsequently been isolated from many organisms including the protozoan *Tetrahymena*, that features unconventional separation of euchromatin in macronuclei from heterochromatin retained in micronuclei by large-scale somatic genomic rearrangements. This nuclear dimorphism facilitated biochemical fractionation of chromatin and lead to the striking finding that specific histone variants were exclusively associated with the transcriptionally active macronucleus (Allis et al., 1980). Only the cloning of histones genes from a number of species revealed that the histone variants hv1 and hv2 associated with euchromatin in *Tetrahymena* (Allis et al., 1986; Allis et al., 1982) have functional counterparts in metazoans. Metazoan and yeast H2A.Z is a true ortholog of *Tetrahymena* hv1, whereas the functional ortholog of hv2 was found to be H3.3 (Ahmad and Henikoff, 2002c).

### 1.2.2 Histone H3 variants have distinct sequences and expression patterns

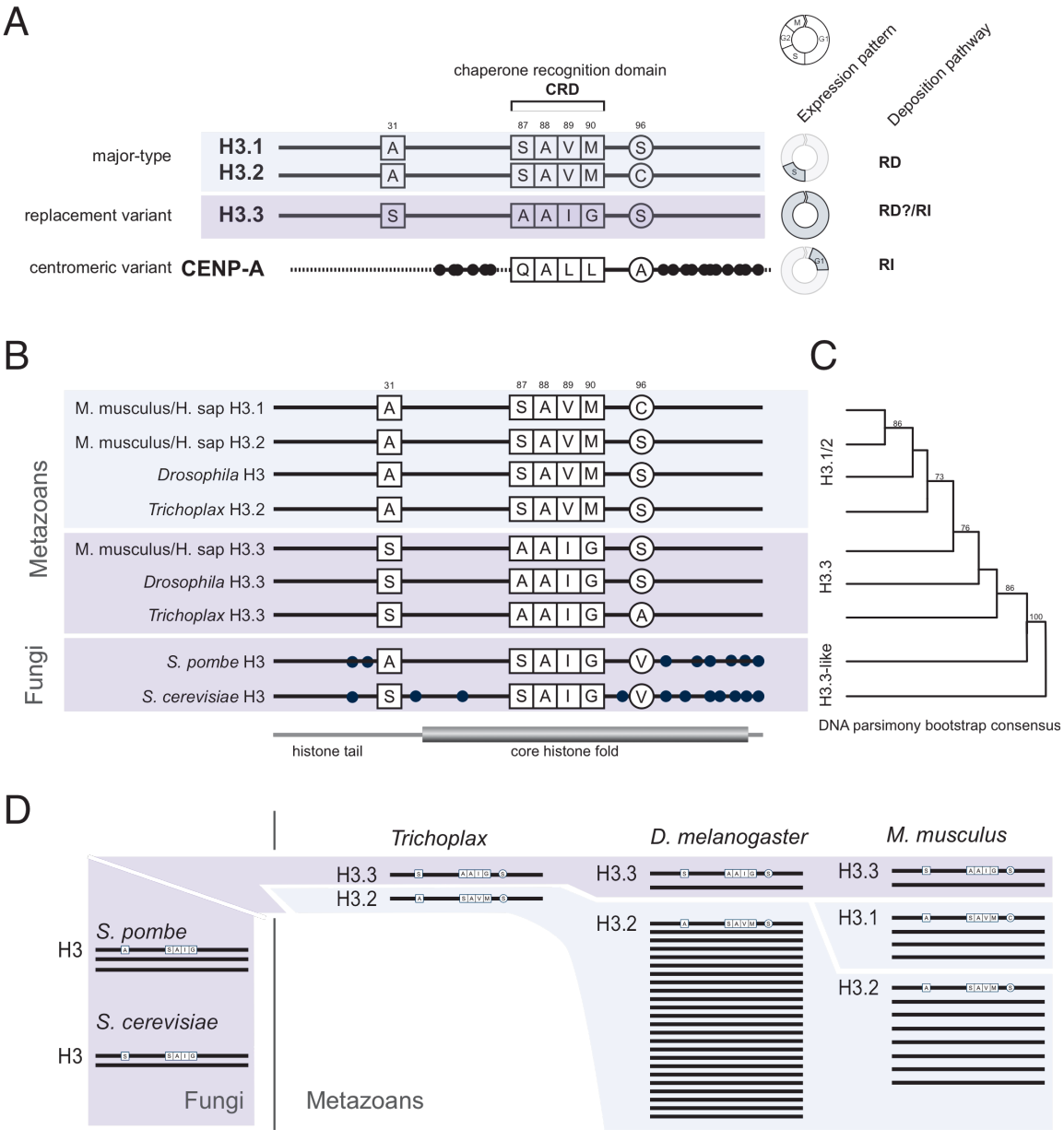
In metazoans, three main classes of histone H3 genes encode distinct H3 proteins: the major-type, replication-dependent (RD) histone H3, the replication-independent (RI) variant H3.3, and the centromeric H3 (cenH3 hereafter) variant CENP-A (Figure 1.3A). Apart from the major type histone H3 shared by all metazoans (systematically called H3.2), mammals possess another RD variant, H3.1, with a single amino acid substitution at position 96.

Two major, highly conserved distinctions account for unique functions of H3.3: differential expression during the cell cycle and amino acid variation in residues 87-90 of the histone core region (Figure 1.3A). Most higher eukaryotes organize their genes for all four major-type histones in repeats with a total of 10-

50 intronless copies of each histone gene (Marzluff et al., 2002; Rooney et al., 2002). Organisms at the base of the metazoan tree, such as *Trichoplax adhearens*, have only one or few copies of H3.2 (Figure 1.3D). H3.1/2 transcripts from histone clusters lack a poly(A) tail but share a conserved 3' stem loop (Dominski and Marzluff, 1999). These atypical mRNA features are thought to be responsible for the tight restriction of replication-dependent histone gene expression to S phase (Harris et al., 1991).

In contrast, H3.3 genes are present in single copies, often contain introns, and give rise to classical polyadenylated mRNAs. Unlike H3.1/2, the expression of H3.3 genes is replication-independent, and H3.3 has long been established as the predominant H3 variant in quiescent, G1, and G2 cells (Wu et al., 1982) Consequently, its cell cycle-independent expression enables H3.3 to serve as a substrate for histone replacement processes that occur outside of S phase.

Figure 1.3: Protein sequences and gene complements of the non-centromeric histone H3 variants in fungi and metazoans.
(A) Schematic representation of the major human histone H3 variants, their expression timing during cell cycle and their mode of incorporation (RD replication-coupled, RI replication-independent) (B) Schematic representation of the major non-centromeric histone H3 protein sequences from human, mouse and *Trichoplax adhearens* (one of the most basal metazoan species), as well as budding and fission yeast. Amino acids that distinguish variants are highlighted with residue numbers, additional differences are indicated as dots. H3.1 only exists in mammals and only differs in position 96 from H3.2 present in all metazoans. H3.2 and H3.3 are distinguished by one amino acid difference at position 31 in the histone tail and three in amino acids 87-90 in the core histone fold. (C) Phylogenetic relationship of the respective histone H3 genes. An unrooted parsimony tree was constructed based on representative coding sequences (consensus tree of 100 bootstraps, excluding the wobble bases). The H3 genes of *S. pombe* and *S. cerevisiae* cluster with metazoan H3.3. (D) Schematic overview of the major non-centromeric gene complements of the indicated species. The placozoan *Trichoplax adhearens* is has only one gene for H3.2 and H3.3 each, while higher metazoans have greatly expanded H3.1/2 gene complements.


A
chaperone recognition domain
CRD
Expression pattern
Deposition pathway
major-type
H3.1
H3.2
replacement variant
H3.3
centromeric variant
CENP-A
RD
RD?/RI
RI
B
Metazoans
Fungi
M. musculus/H. sap H3.1
M. musculus/H. sap H3.2
Drosophila H3
Trichoplax H3.2
M. musculus/H. sap H3.3
Drosophila H3.3
Trichoplax H3.3
S. pombe H3
S. cerevisiae H3
histone tail
core histone fold
C
H3.1/2
H3.3
H3.3-like
DNA parsimony bootstrap consensus
D
Trichoplax
D. melanogaster
M. musculus
H3.3
H3.2
H3.1
S. pombe
S. cerevisiae
H3
Fungi
Metazoans

### 1.2.3 H3.3 is the evolutionary conserved, universal histone H3

Given the overwhelming number of synonymous RD histone genes in higher metazoans, one might be tempted to explain the generation of minor variants by transfer of one or more histone gene from the genomic clusters into an orphan location and subsequent evolution of a variant sequence. Such case can be made for the testis-specific, H3.2-like variant H3t in humans (Franklin and Zweidler, 1977; Witt et al., 1996). For the relationship of H3.1/2 and H3.3, phylogenetic analysis with lower metazoans draw a different picture of variant evolution that places histone H3.3 at the root of the evolutionary tree: Unlike metazoans and multicellular fungi, the yeasts *S. cerevisiae* and *S. pombe* only have a single non-centromeric H3 variant. Protein and DNA sequence analysis suggests that yeast H3 is related to the metazoan H3.3. As yeasts share some of the same replication-dependent and independent chromatin assembly pathways (see Figure 1.5), H3.3 might be considered as an ancestral histone H3 that is compatible with both major assembly pathways. This is in contrast to the metazoan H3.1/2, whose function seems to be largely restricted to the replication-dependent pathway.

Nucleotide sequence alignments suggest that H3 variants have arisen or have already been present at the root of metazoans – RD and RI genes form two separate clades in a maximum-likelihood tree, with RI genes sharing significant homology to the yeast H3 (Figure 1.3B,C). As expected, the simplest metazoan organism identified to date, the placozoan *Trichoplax adhearens*, branches off at the root of this tree, most closely related to the histone genes of the fungi lineage.

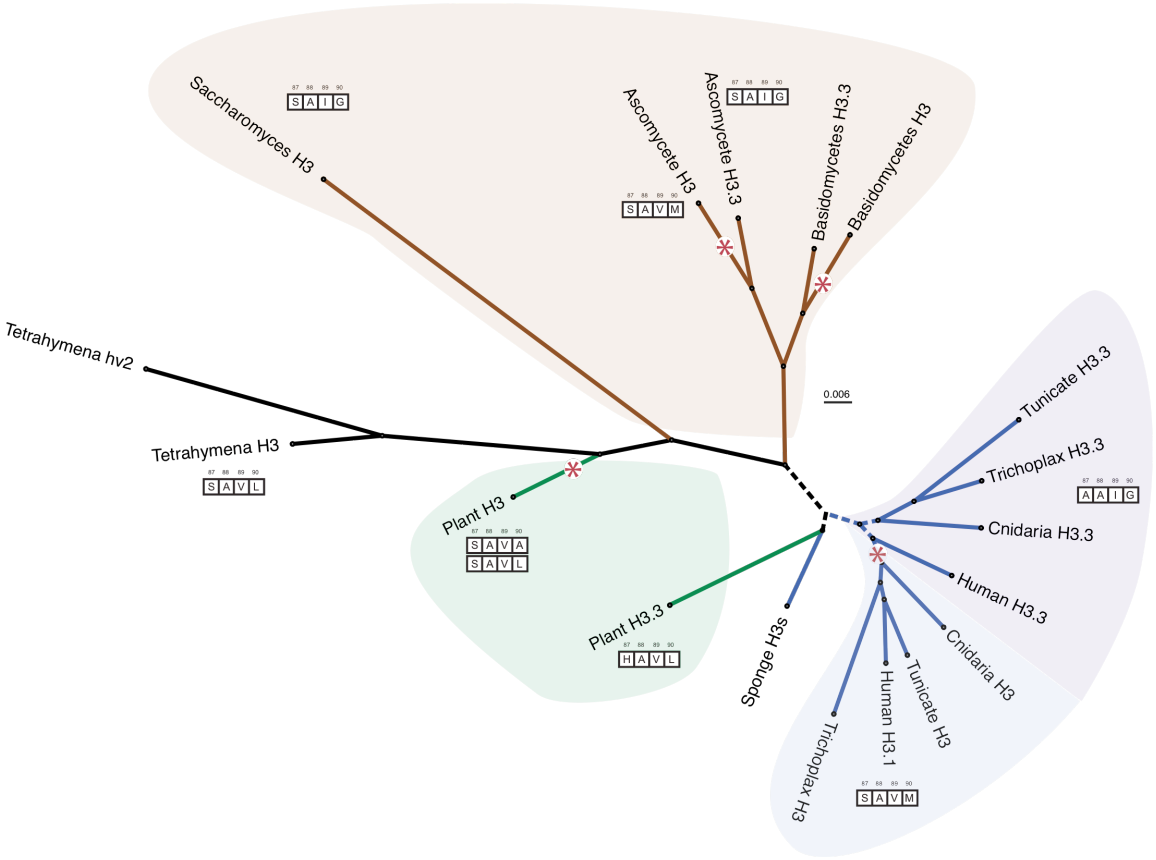



Figure 1.4: Convergent evolution of non-centromeric histone H3 variants in all kingdoms of life.

A phylogenetic tree was created based on a Kimura distance matrix of representative coding sequences (excluding the wobble base) with the PHYLIP package (Felsenstein, 1989). Length of the branches corresponds to evolutionary distance. Dotted lines represent branches that are not supported by bootstrap values >= 0.8 (therefore, branches at the base of the tree cannot be accurately positioned). Branch color by kingdom; fungi brown, plants green, animals blue. Amino acid sequences in the chaperone recognition domain (CRD) is shown for selected branches (all other divergent positions are not shown but included in the phylogenetic analysis). Red stars represent switches in the CRD to the RD variant (all switches are supported by bootstrapping values > 0.8). The isolated appearance of RD variants within each kingdom suggests convergent evolution from the H3.3-like RI ancestor.

Interestingly, *T. adhearens* possesses only one gene per variant, arguing for a later expansion of the RD genes into the canonical histone gene clusters that we observe in higher animals (Figure 1.3D). It is tempting to speculate that it is the dramatically expanding genome size in metazoan evolution that triggered the multiplication of histone genes and potentially also led to the evolution and fixation of the RD variant (H3.1/2) sequence that restricts its use to packaging the newly replicated genome.

Extending my phylogenetic analysis to all three kingdoms – animals, plants and fungi, I found evidence that RD variants have appeared independently in all three kingdoms by convergent evolution (Figure 1.4A). While evolutionary pressure on the protein sequence has restricted amino acid variation to a level that does not allow inference of evolutionary relationships, the mostly synonymous mutations occurring in the coding sequence of histones permit meaningful phylogenetic analysis: A distance matrix-based tree (Felsenstein, 1989) constructed from the coding sequence of a representative set of RD and RI histones in fungi, plants and animals places each respective RI species (H3.3) at a common root. RD species branch off at the root of each multicellular speciation. In the case of multicellular fungi, residues 87-90 of the RD histone H3 are synonymous with the animal H3.1/2, despite the substantially divergent nucleotide sequences. It can therefore be concluded that the particular 'SAVM' motif has been selected for by convergent evolution.

### 1.2.4 H3.3 is enriched at active genes, promoters, and regulatory elements

As discussed above, major-type histones are incorporated in bulk during DNA replication. Once assembled into nucleosomes, the vast majority of H3-H4 tetramers has been observed to remain chromatin-bound over several cell generations, whereas a substantial amount of H2A-H2B dimers is exchanged within hours (Kimura, 2005). In the absence of replication, histone H3 has been reported to extremely stable with a half-life of ~150 days (Piña and Suau, 1987).

Despite the slow turnover of bulk histones, pioneering cytological studies of H3 variant deposition in *Drosophila* provided evidence for rapid H3-H4 exchange at specific loci in euchromatin (Ahmad and Henikoff, 2002c). While low levels of H3.3 are likely deposited together with H3.2 during replication, H3.3 was specifically enriched within actively transcribed genes by a replication-independent replacement process dependent on active transcription (Ahmad and Henikoff, 2002c; Schwartz and Ahmad, 2005). Conversion of the H3.3 variant region 87-90 ('AAIG') to the H3.1/2 sequence 'SAVM' abolished replication-independent incorporation (Ahmad and Henikoff, 2002c). These findings underscore the importance of the variant H3.3 sequence in addition to its cell-cycle independent expression. Interestingly, the single replacement of a Ser with an Ala at position 31 of the histone H3 tail did not have any influence on the deposition pathway, suggesting that H3.3 S31 and its phosphorylation do not play a role in H3.3 deposition (Ahmad and Henikoff, 2002c; Hake et al., 2005).

Recent advances in chromatin immunoprecipitation (ChIP) technologies have allowed a more detailed map of H3.3 deposition, revealing specific H3.3

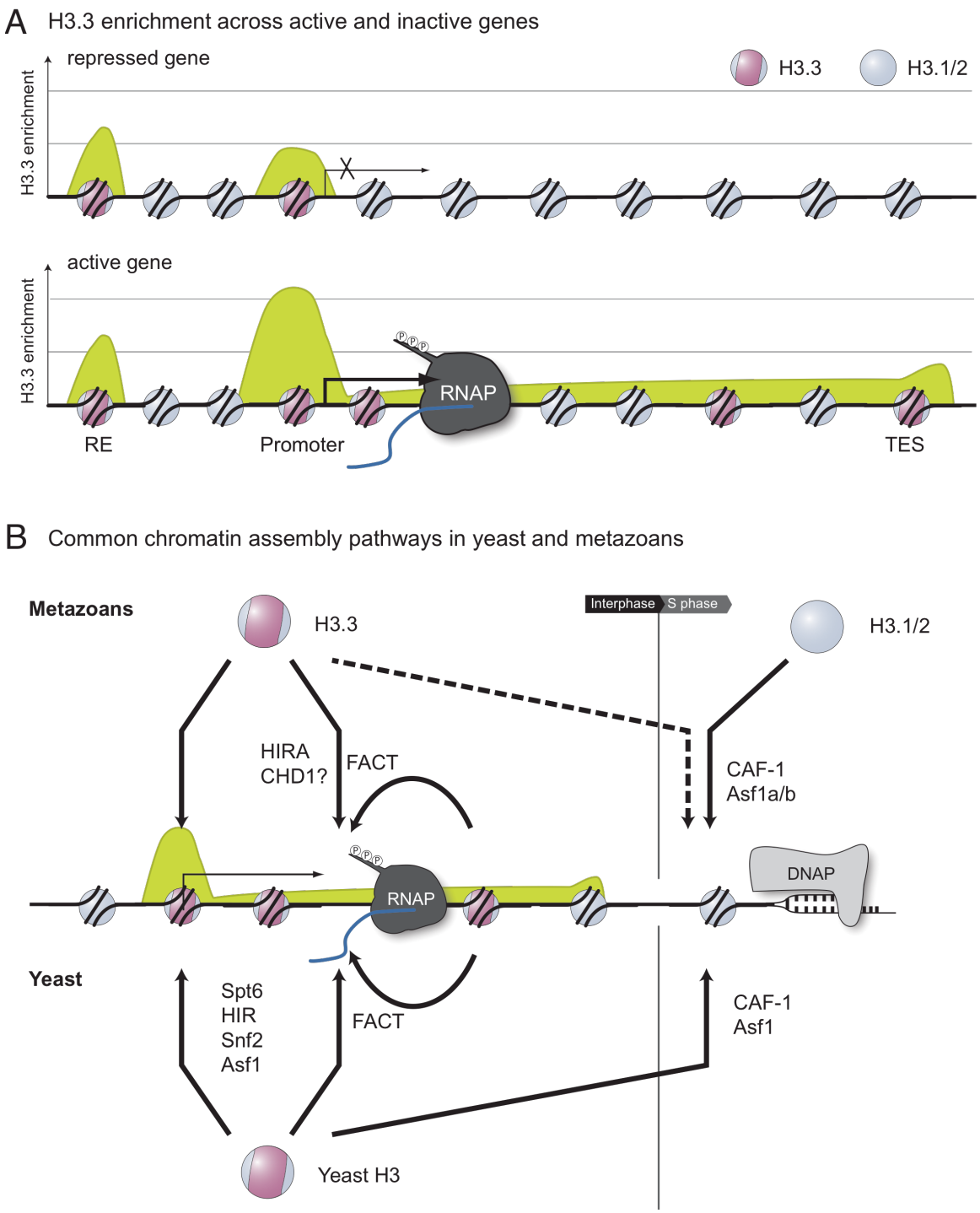
incorporation throughout the gene body of transcribed genes as well as highly enriched foci at the promoter region in *Drosophila* and mammalian cells (Figure 1.5A) (Chow et al., 2005; Daury et al., 2006; Janicki et al., 2004; Jin et al., 2009; Mito et al., 2005, 2007; Nakayama et al., 2007; Sutcliffe et al., 2009; Tamura et al., 2009; Wirbelauer et al., 2005). H3.3 enrichment at promoter regions has been observed not only at active genes but also at inactive genes, possibly accounting for a 'poised' state of these genes (Mito et al., 2007; Tamura et al., 2009). Furthermore, H3 replacement occurs at genic and intergenic regulatory regions in various metazoans (Figure 1.5A) (Goldberg et al., 2010; Jin et al., 2009; Mito et al., 2007; Nakayama et al., 2007).

### 1.2.5 H3 replacement in yeast

Despite relying on a single H3.3-like species, both replication-dependent and replication-independent H3 deposition pathways are found in yeast: in *S. pombe*, H3 expressed outside of S phase is preferentially incorporated in euchromatin (Choi et al., 2005; Takayama and Takahashi, 2007). A number of studies in *S. cerevisiae* detected H3 replacement at active (Jamai et al., 2007) and also inactive (Dion et al., 2007; Rufiange et al., 2007) promoters, but only to a small extent throughout transcribed gene bodies. Genetic studies in yeast delineated a pathway comprising the chromatin remodeler Snf2 and the histone chaperone Asf1, as well as Hir1 or Spt6 for H3 exchange at the promoter region (Adkins and Tyler, 2004, 2006; Gkikopoulos et al., 2009; Rufiange et al., 2007; Schermer et al., 2005). H3 deposition at the gene body required active transcription, Hir1, and Asf1 (Dion et al., 2007; Kim et al., 2009; Rufiange et al., 2007). Hir1, Hir2, Hir3 and Hpc2 constitute the yeast HIR repressor complex (an overview of histone

chaperones is found in Section 1.3, Table 1.1) that has been shown to catalyze replication-independent histone deposition together with the H3-H4 chaperone Asf1 *in vitro* (Green et al., 2005; Prochasson et al., 2005). Spt6 has also been shown to facilitate nucleosome assembly *in vitro* (Bortvin and Winston, 1996). Therefore, chromatin remodelers, histone chaperones and deposition factors cooperate in the eviction of old and deposition of new histones in yeast (Figure 1.5B). Interestingly, the elongation complex FACT (Spt16/ Pob3) redeposits H3-H4 units in the wake of RNAPII, favoring recycling of 'old' histones over exchanging them with 'new' H3-H4 (Jamai et al., 2009). When Spt16 is deleted, a Hir1-dependent pathway takes over to deposit more 'new' H3 (Formosa et al., 2002; Jamai et al., 2009). In conclusion, yeast genetics of replication-independent histone exchange processes might yield clues to yet undiscovered components of metazoan H3.3-deposition pathways (Figure 1.5B).

Figure 1.5: Transcription-associated histone dynamics in metazoan and yeast.
A Schematic map of an active and inactive gene locus comprising an upstream regulatory element (RE), transcription start site (Promoter) and transcription end site (TES). The distribution of histone H3.3 across the locus is shown in green, with representative H3.3 and H3.1/2 nucleosomes, as well as RNA polymerase II (RNAP). (b) Summary of the known factors involved in replication-dependent (right, S phase) and replication-independent (left, Interphase) chromatin assembly pathways, in metazoans and yeast. Replication-coupled assembly is thought to be mediated by the CAF1 complex and Asf1 proteins in the wake of DNA polymerase (DNAP). H3.3-enrichment at telomeres is dependent on ATRX. Replication-independent deposition at promoters, regulatory elements and genic regions in metazoans requires HIRA, CHD1, and/or other factors, analogous to pathways in yeast mediated by Snf2, Asf1, HIR complex and/or Spt6. The FACT complex (Spt16 and Pob3/Ssrp1) might contribute to incorporation of new or recycling of old histones.


A
H3.3 enrichment across active and inactive genes
repressed gene
H3.3
H3.1/2
H3.3 enrichment
active gene
H3.3 enrichment
RNAP
RE
Promoter
TES
B
Common chromatin assembly pathways in yeast and metazoans
Metazoans
H3.3
Interphase
S phase
H3.1/2
HIRA
CHD1?
FACT
CAF-1
Asf1a/b
RNAP
DNAP
Yeast
Spt6
HIR
Snf2
Asf1
FACT
CAF-1
Asf1
Yeast H3

### 1.2.6 Mechanism of H3.3-specific deposition in metazoans

HIRA, the homolog of yeast Hir1 in higher eukaryotes, has been shown to assemble chromatin independent of replication and to interact specifically with ASF1a/b in a multisubunit complex specific for H3.3 (Ray-Gallet et al., 2002; Tagami et al., 2004). HIRA and the SWI/SNF family chromatin remodeler CHD1 have also been implicated in H3.3 deposition *in vivo* (Konev et al., 2007). Interestingly, H3.3 continues to be incorporated into chromatin even in the absence of HIRA or CHD1 (Bonnefoy et al., 2007; Konev et al., 2007). Therefore, an imminent question at the time I started my thesis work was if these identified factors were only mediating H3.3-deposition at a subset of genomic locations or if yet unknown pathways would compensate for their loss. Indeed, studies by my colleagues subsequently showed HIRA mediated H3.3 deposition at transcribed regions and a subset of regulatory elements (Goldberg et al., 2010). Furthermore, H3.3-enrichment at murine ESCs was not dependent on HIRA but not on the ATP-dependent chromatin remodeler ATRX (Goldberg et al., 2010; Wong et al., 2010). Despite these first insights into locus-specific chromatin assembly pathways, many questions remained about the precise mechanisms of H3.3 replacement: Are HIRA and ATRX *bona fide* deposition factors for histone H3.3 or do they orchestrate the incorporation of H3.3 by indirect means. What are the histone chaperones that can distinguish the subtle differences between H3.3 and H3.1/2, and how do they achieve site-specific deposition? Importantly, neither HIRA nor ATRX had been shown to interact directly with histone H3.3 and structural details on the molecular recognition of H3.3 were lacking.

### 1.2.7 H3.3 function: a balancing act between facilitating and repressing transcription?

As a highly conserved replacement variant, does H3.3 have a conserved function at promoters, coding regions, and regulatory elements? Two recent studies assessing inducible gene expression suggest that incorporation of H3.3 promotes initial gene activation (Placek et al., 2008; Tamura et al., 2009). One possibility is that nucleosome eviction and H3.3 deposition may serve as a mechanism for the rapid removal of inhibitory histone posttranslational modifications and/or replacement with activating marks (Ahmad and Henikoff, 2002c). However, even though nucleosomal H3.3 is enriched in activating modifications such as H3K4me3, these modifications in particular seem to be established only after nucleosomal deposition (Loyola et al., 2006; McKittrick et al., 2004). Rather than introducing a particular set of PTMs, ongoing histone exchange could therefore contribute to a highly dynamic steady state of establishment and removal of histone PTMs at specific genomic locations (Ahmad and Henikoff, 2002c; Deal et al., 2010). Continuous histone exchange and H3.3-incorporation at boundaries of chromatin domains has therefore been proposed to limit the spreading of certain histone modifications (Dion et al., 2007; Mito et al., 2007).

Based on the apparent lability of H3.3 nucleosomes in chromatin extracts, it has been proposed that nucleosome-destabilizing properties could help promote and propagate an active chromatin state (Jin and Felsenfeld, 2007; Jin et al., 2009). As *in vitro* studies found little stability difference in recombinant H3.1/2 and H3.3 nucleosome (Flaus et al., 2004; Thakar et al., 2009), this effect

might be potentiated by histone PTMs or inherent to CG-rich promoter DNA sequences that often coincide with H3.3-enrichment (Ramirez-Carrozzi et al., 2009). Furthermore, cooperative effects with H2A.Z and exclusion of the linker histone H1 could account for some of the properties of H3.3-containing nucleosomes (Braunschweig et al., 2009; Kim et al., 2007).

Is H3.3 a general marker of active chromatin? Notably, the HIR complex has been shown to have a repressive role on transcription in yeast (Anderson et al., 2009; Blackwell et al., 2004; Fillingham et al., 2009; Kim et al., 2007; Sharp et al., 2001; Sherwood et al., 1993). Hir1 was first identified as a potent repressor of the canonical histone genes in *S. cerevisiae* (Sherwood et al., 1993), and recently its repressor function in *S. pombe* has been mapped to a large number of promoters and also to suppression of cryptic transcripts from within coding regions (Anderson et al., 2009), likely by repopulating nucleosome-free regions (Schermer et al., 2005). It is tempting to speculate that replication-independent H3.3 deposition in metazoans is similarly used to replenish nucleosome-free regions. Indeed, H3.3 knockdown leads to a slight decrease in nucleosome density (Braunschweig et al., 2009).

Apart from its unresolved effect on gene expression from euchromatin, cumulative evidence from others and our studies supported a role for H3.3 in facultative and constitutive, highly repetitive heterochromatic regions. HIRA, ASF1a and the mammalian Hir2 homolog Ubinuclein-1 have been implicated in the formation of facultative heterochromatin (Banumathy et al., 2009; Zhang et al., 2007a), and H3.3 has been observed in pericentric heterochromatin and at

telomeres (Drané et al., 2010; Goldberg et al., 2010; Hake et al., 2005; Lewis et al., 2010; Wong et al., 2009).

### 1.2.8 Biological significance of replication-independent H3.3 deposition

Clues for the *in vivo* functional significance of H3.3 come from genetic studies in flies and mice. Loss of both genes of H3.3 in flies leads to complete sterility, mild transcriptional defects, particularly at highly expressed genes, and partial but incomplete lethality (~42% viability) (Sakai et al., 2009). Intriguingly, the grossly normal development to adulthood of the surviving H3.3-deficient flies indicates that H3.3 is not absolutely required for transcription and development (Hödl and Basler, 2009; Sakai et al., 2009). Indeed, while gene expression of a subset of genes in adults was perturbed, the precisely timed and localized expression of developmental key factors was not affected in H3.3-deficient flies. Similarly, although HIRA is required for fertility, adult HIRA null flies have no phenotypic abnormalities (Bonnefoy et al., 2007). In mice, targeted mutagenesis of HIRA resulted in gastrulation defects and patterning abnormalities of mesendodermal derivatives prior to early embryonic lethality (Roberts et al., 2002), suggesting a more prominent role for replication-independent chromatin assembly during mammalian development. Although HIRA may have various H3.3-independent functions, H3.3 itself is also important for mammalian development: a retroviral gene trap insertion into the murine H3.3A gene generated an H3.3 hypomorph that caused developmental defects and neonatal lethality (Couldrey et al., 1999).

How do flies compensate so well for the loss of H3.3 or HIRA? Intriguingly, an unknown mechanism seems to allow the cells to sense overall histone levels, as

replication-dependent histone H3.2 genes are upregulated in H3.3-deficient flies. Furthermore, upregulated endogenous replication-dependent histone gene transcripts were found to be polyadenylated to some extent, likely achieved by a previously described alternative histone mRNA processing mechanism (Akhmanova et al., 1997; Sakai et al., 2009). Importantly, viability and wild type expression of most genes are fully restored when an additional H3.2 transgene is introduced (Sakai et al., 2009). Thus, elevated levels of H3.2 can largely rescue the transcriptional phenotype in adult H3.3 null flies. In rapidly dividing cells, replication-dependent deposition of H3.1/2 could compensate for loss of nucleosomes during transcription (Figure 1.6A). Alternatively, replication-independent pathways could tolerate H3.1/2 as substrates in the absence of H3.3 (Figure 1.6B). Interestingly, global H3K4me3 levels in flies lacking H3.3 were comparable to wild-type but drastically reduced in flies with a H3.3K4A transgene (Sakai et al., 2009), indicating that only in the absence of H3.3, H3.2 becomes the major carrier for this mark. I therefore speculate that in the absence of a replacement variant, re-deposition of histones in *cis* partially substitutes for replication-independent incorporation of new histones (Bell and Schübeler, 2009), which is analogous to the competing pathways for 'new' and 'old' histone observed in yeast (Figure 1.6C)(Jamai et al., 2009). Thus, if no 'new' histones are available, more 'old' histones with 'old' marks might be retained.

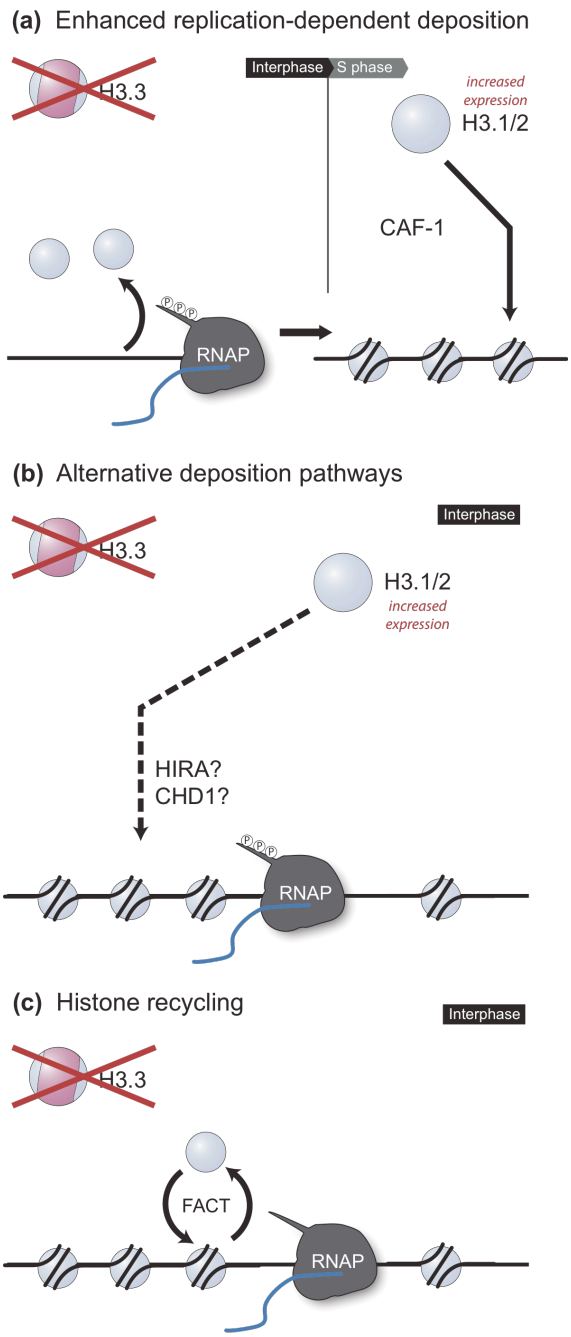



Figure 1.6: Putative compensatory mechanisms for the loss of H3.3.
(A) Absence of H3.3 as substrate for replication-independent chromatin assembly could create nucleosome-free regions. In rapidly dividing cells, these gaps could be filled during the next S phase via canonical replication-dependent chromatin assembly. (B) Elevated levels of H3.1/2 throughout the cell cycle could provide substrate for replication-independent chromatin assembly factors that are not restricted to H3.3. (C) In the absence of de-novo chromatin assembly, the FACT complex could favor transient eviction and redeposition of histone units in *cis*.

### 1.2.9 Histone replacement by H3.3 is essential for reproduction in metazoans

Despite potential compensatory mechanisms, HIRA and H3.3 play critical roles in sexual reproduction in all studied metazoans (Bonnefoy et al., 2007; Hödl and Basler, 2009; Sakai et al., 2009). Both male and female H3.3 null flies are sterile. In mammals, the hypomorphic gene trap of H3.3A described above also led to male sub-fertility (Couldrey et al., 1999). Strikingly, H3.3 is substrate for several large-scale chromatin remodeling events during metazoan reproduction, in particular gametogenesis and fertilization (Figure 1.7, see also (Banaszynski et al., 2010)). With the discovery of the metazoan-specific histone chaperone Daxx, another potential player in these chromatin remodeling events will need to be considered.

Meiotic sex-chromosome inactivation in mammalian male germ cells also involves massive incorporation of H3.3 into the X and Y chromosomes and subsequent silencing (van der Heijden et al., 2007), a process that could be mechanistically related to the formation of facultative heterochromatin (Banumathy et al., 2009; Zhang et al., 2007a). Meiosis is partly impaired in H3.3 null flies due to a defect in chromosome segregation (Sakai et al., 2009). After meiosis, the condensation of sperm chromatin requires removal of most histones and replacement with protamines, although some pool of H3.3 is retained in mammalian and *C. elegans* sperm chromatin (Ooi et al., 2006; van der Heijden et al. 2007).

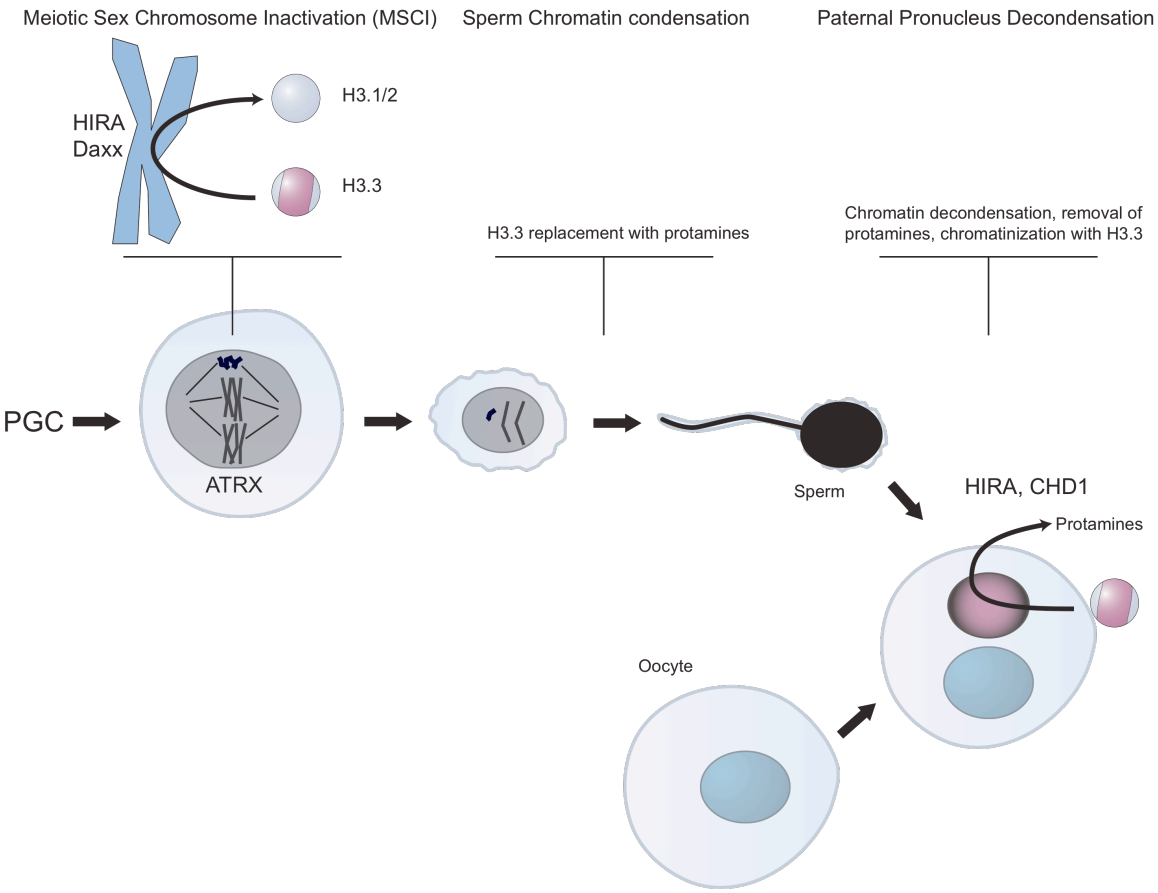



Figure 1.7: Large-scale replication-independent histone-exchange in the germ line.
Early in mammalian spermatogonia differentiation, both X and Y chromosomes are condensed and remain silent throughout spermatogenesis. Condensation appears to be brought about by chromosome-wide remodeling that includes H3.3 incorporation (van der Heijden et al., 2007). Both HIRA and Daxx localize to the XY body at this stage (Rogers et al., 2004; van der Heijden et al., 2007). ATRX is required for faithful meiotic chromosome segregation (de La Fuente et al., 2004). Upon fertilization, H3.3 is incorporated into the paternal pronucleus concomitant with protamine removal whereas no H3.3 is detected in the maternal pronucleus (Bonnefoy et al., 2007; Konev et al., 2007; Ooi et al., 2006; Torres-Padilla et al., 2006; van der Heijden et al., 2005). This asymmetric histone variant incorporation correlates with an earlier onset of transcription on the paternal pronucleus. Both the histone chaperone HIRA and the chromatin remodeler CHD1 have been shown to be required for protamine removal, decondensation and H3.3 incorporation in flies (Bonnefoy et al., 2007; Konev et al., 2007).

After fertilization, a maternal pool of H3.3 is used to rechromatinize the paternal genome in the male pronucleus (Bonnefoy et al., 2007; Konev et al., 2007; Ooi et al., 2006; Torres-Padilla et al., 2006). This asymmetrical distribution of H3 variants could be important in epigenetic distinction of maternal and paternal information.

Critically, all remodeling events in the germline seem to be exquisitely specific to H3.3, as a H3.2 transgene under the H3.3 promoter cannot rescue the fly's sterility (Hödl and Basler, 2009; Sakai et al., 2009). It is therefore likely that the phenotype is a direct consequence of impaired large-scale chromatin remodeling rather than a secondary effect due to gene expression changes related to transcriptional defects in the absence of H3.3. Consistent with this notion, H3.3 incorporation in the male pronucleus precedes onset of transcription and relies on HIRA and CHD1 activity (Bonnefoy et al., 2007; Konev et al., 2007). The essential germline functions of H3.3 therefore likely created the strong evolutionary pressure that drove the exceptional conservation of the H3.3 protein in higher eukaryotes.

### 1.2.10 Conclusions

Why is the use of H3.3 so diverse and widespread while not all of its functions are essential in metazoans? I speculate that both germline and somatic functions of H3.3 have evolved from the single H3.3-like ancestor present in unicellular organisms. Differential timing of H3 gene expression might have allowed some tailoring of H3 variants for replication-dependent and independent functions, but ultimately the diversification of H3.1/2 amino acid sequence efficiently excluded these replication-dependent histones from H3.3-

specific pathways. The separation of replication-dependent and -independent pools of H3-H4 might have allowed subsequent multiplication of the replication-dependent histone genes to fuel the growing need for bulk histones during replication of larger genomes without affecting fine-tuned histone replacement processes. We will need more detailed studies on how these H3.3-specific pathways affect chromatin structure and function to ultimately understand why metazoans evolved this exquisite specificity. Of note, wherever H3 variants have been studied in non-metazoan organisms, plants and the protozoans *Ustilago*, *Tetrahymena* and *Trypanosoma brucei*, similar patterns of RD and RI incorporation have been observed (Allis et al., 1980; Anju et al., 2011; Siegel et al., 2009; Waterborg, 1993). This strongly suggests an overarching theme in chromatin regulation by histone variants.

## 1.3 Histone chaperones[‡]

### 1.3.1 The structure of the nucleosome

Core histone proteins are structurally defined by the presence of a histone fold, composed of approximately 65 amino acids (Luger et al., 1997). These amino acids adopt a three ⟨-helix structure in the presence of a partner histone, yielding the heterodimers H3-H4 and H2A-H2B. Two H3-H4 dimers further associate via a four ⟨-helix bundle to form a $(H3\text{-}H4)_2$ tetramer (Figure 1.8B,C). In the nucleosome, this $(H3\text{-}H4)_2$ tetramer is flanked by an H2A-H2B dimer on either side (Figure 1.8A). DNA is held around this histone octamer through multiple direct contacts (Richmond and Davey, 2003). These include at least one salt-bridge between a backbone phosphate of each DNA strand and each histone (Richmond and Davey, 2003). The disordered histone tails protrude between or in close proximity to the DNA gyres from the somewhat circular nucleosome core (Figure 1.8A). All four core histones belong to the most conserved proteins amongst mono- and multicellular eukaryotic organisms. Conservation of H2A, H2B, H3 and H4 extends beyond the core histone fold to peripheral ⟨-helices and disordered N-terminal tails (also C-terminal in H2A). The variants of higher eukaryotes substitute only one subunit of the heterodimer, H3 and H2A, while their hetero-dimerization partners H4 and H2B are invariant with few exceptions (reviewed in Malik and Henikoff, 2003).

[‡] Parts of this chapter have been published as: Elsässer, S. J., and D'Arcy, S. (2011). Towards a mechanism for histone chaperones. *Biochim Biophys Acta*.

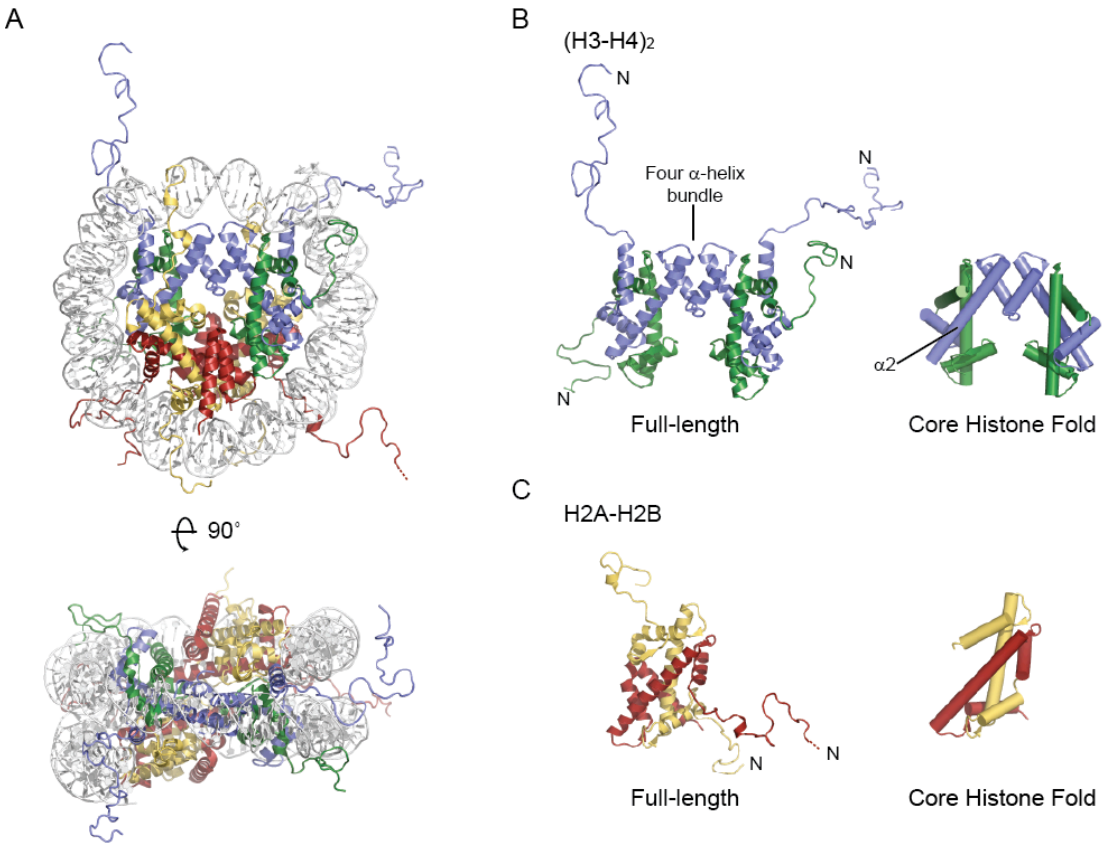



Figure 1.8 Nucleosome structure.

(A), $(H3\text{-}H4)_2$ tetramer (B) and H2A-H2B dimer (C). Cartoons were rendered from PDB 1KX5. DNA is white, H3 is blue, H4 is green, H2A is yellow and H2B is red. Missing residues at the N-terminus of H2B are shown by a dotted line. In *B* and *C*, an N indicates the proteins N-terminus, and the representation on the right shows only the core histone fold. The four ⟨-helix bundle that mediates H3-H4 tetramer formation and the second ⟨-helix of H3 are indicated.

These variants have high sequence similarity to their major-type counterpart in the histone fold domain, containing few amino acid substitutions or short insertions confined to specific histone surfaces. Consequently, the overall structure of the histone fold is almost invariant, but contacts between the H3-H4 tetramer and the H2A-H2B dimer or DNA interactions can be affected (Luger et al., 1997; Suto et al., 2000; Tachiwana et al., 2011).Conservation of histone protein

family members alludes to heavy structural constraints imposed by nucleosome formation, and is evidence that histone sequence, structure and function are highly inter-dependent and major contributors to organism fitness.

### 1.3.2 Histones are post-translationally modified

Histones are extensively post-translationally modified by diverse enzymatic activities including phosphorylation, acetylation, methylation and ubiquitination (Kouzarides, 2007). Histones can be a substrate as part of the nucleosome or when in a non-nucleosomal complex, depending on the enzyme and the location of the modified residue (Loyola and Almouzni, 2007). Post-translational modifications (PTMs) occur predominantly in the lysine/arginine-rich histone tails, but also in the core histone fold (Cosgrove et al., 2004). A number of specific lysine acetylations directly influence histone biochemistry, particularly in the context of nucleosome assembly and stability (Andrews et al., 2010; Cosgrove, 2007). Many studies have been dedicated to identifying the enzymes involved in histone PTM and PTM-removal, as well as to deciphering the function of histone PTMs alone and in combinations (Kharchenko et al., 2011; Kouzarides, 2007; modENCODE Consortium et al., 2010; Ruthenburg et al., 2007; Strahl and Allis, 2000; Taverna et al., 2007).

### 1.3.3 Prerequisites of nucleosome formation

The structure of the nucleosome with its octameric histone core is in keeping with the many biochemical studies on recombinant histone proteins. H3-H4 heterodimer exists in equilibrium with the $(H3\text{-}H4)_2$ tetramer at physiological ionic strength in the absence of other protein or DNA factors (Baxevanis et al.,

1991; Donham et al., 2011). Contacts between the H2A-H2B dimer and $(H3\text{-}H4)_2$ tetramer to form a stoichiometric octamer, also occur in solution, albeit only at high ionic strength which, like the vicinity of polyanionic DNA, neutralizes some of the positive charges of the histones. Nucleosome assembly is a thermodynamically favored process due to the many strong electrostatic interactions that can be formed between the histone octamer and DNA.

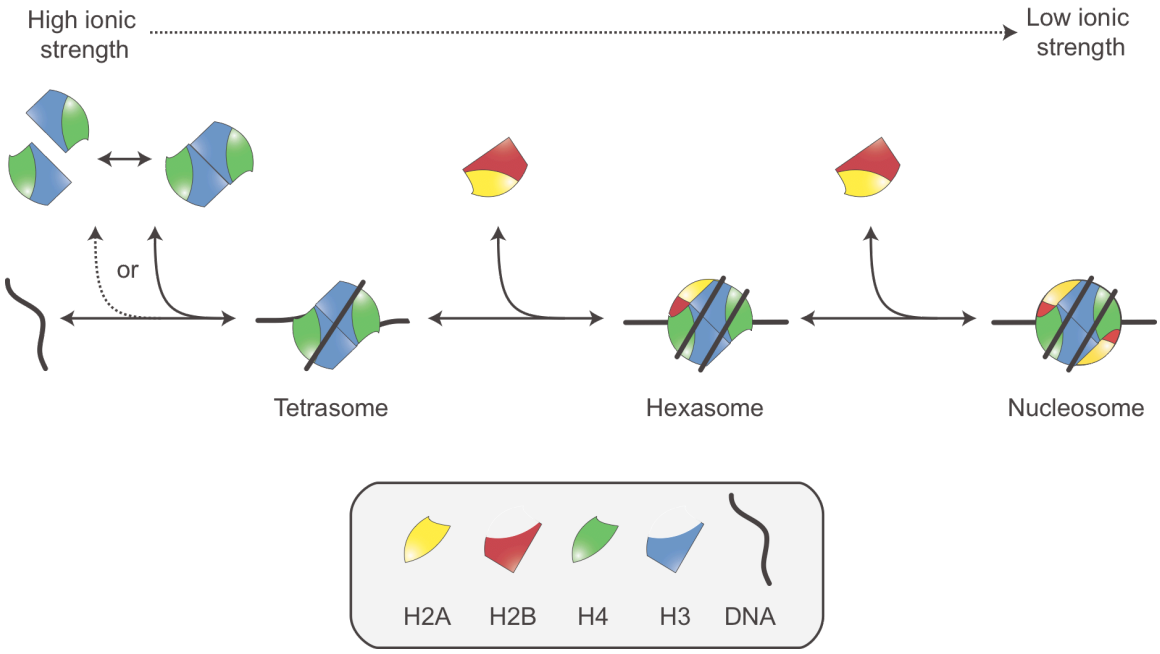



Figure 1.9: Model of sequential nucleosome assembly.
The $(H3\text{-}H4)_2$ tetramer (or two H3-H4 dimers) binds to DNA to form the tetrasome. This is followed by the addition of two H2A-H2B dimers. The intermediate, a tetrasome plus a single H2A-H2B dimer, is described as a hexasome. This process is reversible in salt gradient experiments as indicated. A similar sequence of events has been delineated for Nap1 nucleosome assembly *in vitro*.

Perhaps the integral observation in histone biochemistry is that salt gradients (high to low) of equimolar histone-DNA mixtures, induce nucleosome formation (Thomas and Butler, 1977) and circumvent the observed precipitation of histone-containing mixtures at physiological ionic strength (Laskey et al.,

1978a). Analysis of salt-dependent *in vitro* nucleosome reconstitution lead to a model of sequential nucleosome assembly (Figure 1.9). The sequential model postulates that the $(H3\text{-}H4)_2$ tetramer (or two H3-H4 dimers) is first deposited onto DNA, followed by the addition of two H2A-H2B dimers.

The ability of $(H3\text{-}H4)_2$ tetramer to bind DNA at higher ionic strength than H2A-H2B dimer supports this model (Oohara and Wada, 1987; Wilhelm et al., 1978). While this model has been confirmed *in vitro* and is reversible (Bohm et al., 2011; Mazurkiewicz et al., 2006), evidence for its relevance *in vivo* is somewhat indirect (Smith and Stillman, 1991; Worcel et al., 1978). The distinct exchange rates of H2A-H2B (fast) and H3-H4 (slow) in chromatin are consistent with the model, although these may be a result of specific cellular variables and not generally applicable (Jamai et al., 2007; Kimura and Cook, 2001). It can also be argued that the superhelical arrangement of the DNA makes it difficult to envision deposition of an intact histone octamer, the only other stable histone complex detected *in vitro* (Andrews et al., 2010). Cross-linked octamers, however, do assembly into nucleosomes with the same efficiency as uncross-linked histones *in vitro* (Stein et al., 1977; Stein et al., 1979). In the context of the sequential model, exchange of the central $(H3\text{-}H4)_2$ tetramer, for example for incorporation of replication-independent H3 variants, would require the complete disassembly of the nucleosome.

In the cell, multiple steps have to precede the productive formation of nucleosomes. Histones are synthesized and folded in the cytoplasm, before being imported into the nucleus and recruited to subnuclear territories and ultimately the sites of deposition at the DNA (Figure 1.10). This flow of histones has to be

highly facilitated and regulated to meet the supply and demand of DNA-templated processes. Proteins that catalyze one or more of these steps are termed 'histone chaperones'.

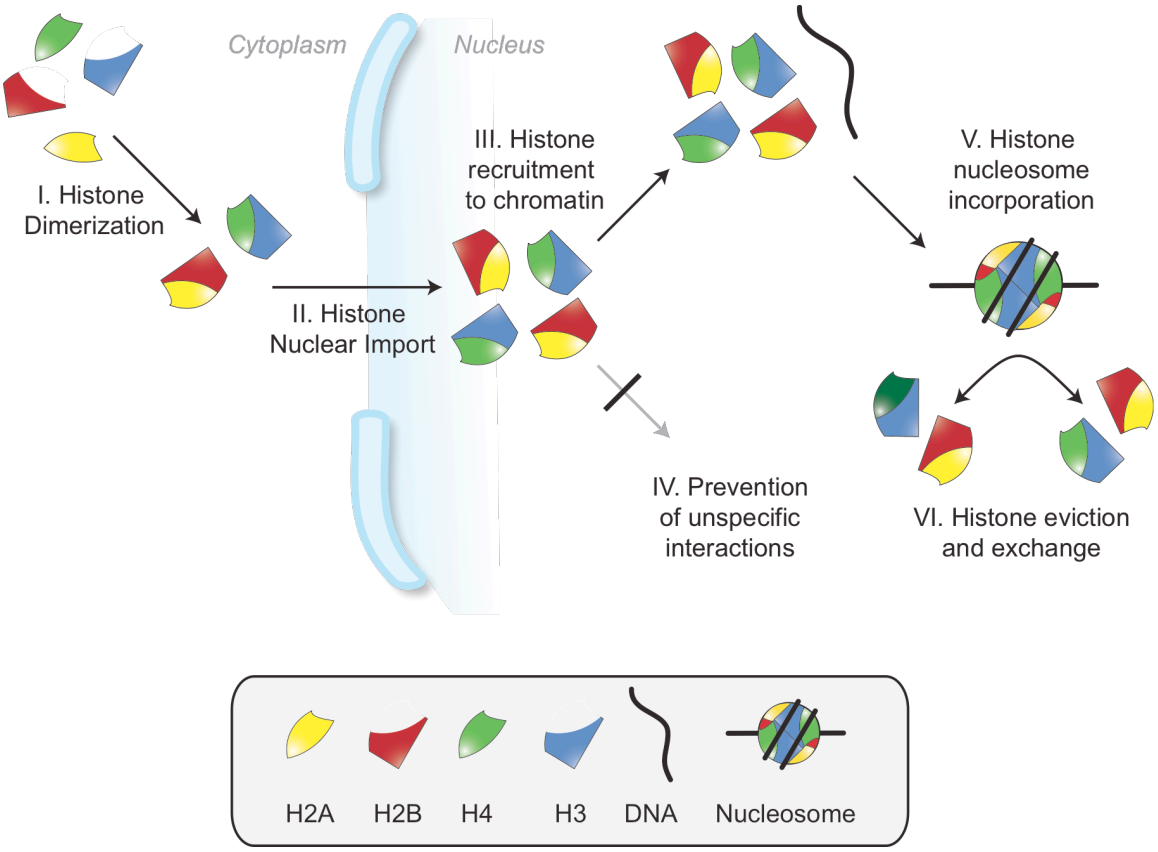



Figure 1.10: Prerequisites of nucleosome formation and contexts of histone chaperone activity.

Histone chaperones may facilitate nucleosome formation by being involved in some or all of the processes I-VI indicated. Histones may be transferred between chaperones to complete all these processes in a regulated manner.

### 1.3.4 A framework of histone chaperone function

Even before the definition of molecular folding 'helpers', the term 'chaperone' was applied to a protein involved in nucleosome formation in the late seventies (Laskey et al., 1978a): Laskey recognized that the complete instruction manual for nucleosome assembly *in vitro* is contained in the histones and DNA. He thus proposed that histone chaperones function by preventing promiscuous histone-DNA interactions. Studies by Laskey and colleagues identified the acidic protein NPM in *Xenopus laevis* egg extracts able to bind histones and facilitate nucleosome formation in an ATP-independent manner (Laskey et al., 1978a; Laskey et al., 1977). It is remarkable that these two properties remain the only shared and thus defining characteristics of histone chaperones today (Figure 1.11). It should be reinforced that histone chaperones possess *both* histone-binding and ATP-independent nucleosome assembly activity, but are not stable constituents of the final nucleosome product. This distinguishes them from proteins with histone PTM-recognition domains, such as bromodomains or PHD fingers, as well as from ATP-dependent nucleosome remodelers. Histone association is routinely assayed both *in vitro* and *in vivo* using standard binding assays, while nucleosome assembly is frequently only assessed by *in vitro* experiments on linear or circular DNA. A protein involved in any one of the aforementioned prerequisites is likely to have *in vitro* nucleosome assembly activity. Experiments involving yeast genetics or cell-free extracts are more stringent tests of so-called direct nucleosome assembly *in vivo*.

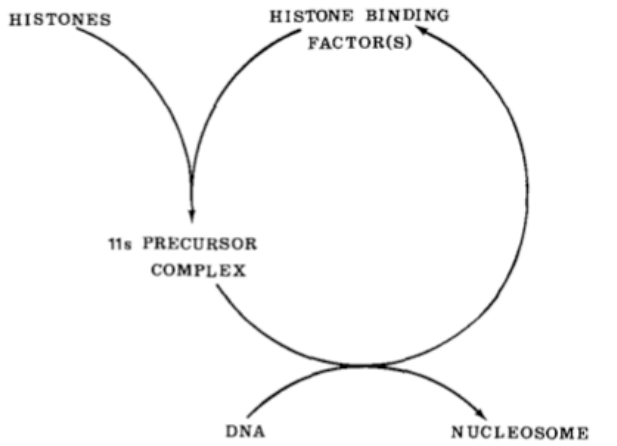



**Figure 6.** A proposed scheme for nucleosome assembly, in which a catalytic factor binds histones to form nucleosome precursor complexes which then interact with DNA to form nucleosomes.

Figure 1.11: Laskey's histone chaperone definition.
The founding definition of a molecular chaperone as a catalytic factor in chromatin assembly by Ron, 1997 (Laskey et al., 1978b)

Efforts by many laboratories employing diverse experimental strategies have identified conserved histone chaperones in yeast and metazoans (see Table 1.1 and (Koning et al., 2007) for a more thorough description of individual chaperones). Cytosolic H2A-H2B and H3-H4, for example, are bound by Nap1 and NASP, Asf1 respectively, and might remain bound during translocation to the nucleus (Campos et al., 2010; Drané et al., 2010; Mosammaparast et al., 2002; Osakabe et al., 2010; Tagami et al., 2004). Nap1 and Asf1 also function in the nucleus, where the histones are channeled into distinct pathways associated with DNA metabolic events. These events are typically classified as DNA replication-dependent or independent. CAF1 is thought to be a replication-dependent chaperone, while HIRA activity is replication-independent (Goldberg et al., 2010; Kaufman et al., 1995; Tagami et al.). Many chaperones are functionally linked to transcriptional regulation, such as FACT, which facilitates eviction of H2A-H2B

from the nucleosome (Bortvin and Winston, 1996; Han et al., 2010; Kundu et al., 2011; Orphanides et al., 1999). NPM represents a storage chaperone insofar as it binds the large pool of histones that will be used for packaging newly replicated DNA in the *X. laevis* egg (Laskey et al., 1977). Homologous NPM chaperones also act as sinks for histones removed during sperm development and fertilization (Philpott and Leno, 1992; Philpott et al., 1991). Linker histone chaperone activity has been described as additional feature of NAP1, sNASP and NPM1, potentially regulating higher order chromatin structure *in vivo* (Finn et al., 2008; Gadad et al., 2011; Kepert et al., 2005; Shintomi et al., 2005).

The sheer number of histone chaperones and histone-binding proteins identified to date supports the long-held view of there being little or no free histones at any given time or locus in the cell. Implicit in the idea that most histones are nucleosomal or chaperone-bound, is the transfer of histones from one chaperone to another (hereafter a 'hand-off'). The recent description of distinct, ordered complexes associated with newly synthesized histones suggests that this notion is applicable to H3-H4 (Campos et al., 2010). Biochemical analysis of H3-H4 purified from either the cytoplasm or nucleus, failed to detect free or unbound histones (Campos et al., 2010; Drané et al., 2010; Elsaesser and Allis, 2010; Lewis et al., 2010; Tagami et al., 2004). Rather, H3-H4 was found in different histone chaperone-containing complexes. Details of the temporal hierarchy of these complexes, as well as a similar study for H2A-H2B, are eagerly awaited. The localization of *bona fide* chaperones in both the nucleus and the cytoplasm (e.g. (Mosammaparast et al., 2002)), as well as the often stoichiometric amounts of histone chaperones and histones (Elsaesser and Allis, 2010; Tagami et al., 2004), also attest to a small or absent population of free histones in the cell.

Table 1.1:

Brief overview of major histone chaperones and their biological functions.

Homologs in selected species are given where confirmed (blue background) or putative (blue diagonal stripes). Described functions are highlighted ('main' function purple, additional observed or implied functions in purple stripes). For references see main text.

| | Homologues | | | | | Putative Function | | | |
|---|---|---|---|---|---|---|---|---|---|
| Name | *S. cerevisiae* | *Drosophila* | *Xenopus* | Mammals | Specificity | Replication/Repair | Replication-independent | Transcription | Shuttling-Storage |
| Asf1 | Asf1 | Asf1 | Asf1 | ASF1a<br>ASF1b | H3-H4 | | | | |
| H3-H4 dimer chaperone with low nM affinity | | | | | | | | | |
| RBBP4-7 | (Cac3) | p55 | p46<br>p48 | p46<br>p48 | H4<br>(H3-H4) | | | | |
| H3-H4 chaperone, also cenH3-H4 chaperone | | | | | | | | | |
| CAF1 | Cac1<br>Cac2<br>Cac3 | p180<br>p105<br>p55 | p150<br>p60<br>p46-48 | p150<br>p60<br>p46-48 | H3.1/2-H4<br>(yeast H3-H4) | | | | |
| Replication-dependent H3.1-H4 histone chaperone | | | | | | | | | |
| HIRA | Hir1<br>Hir2<br>Hir3<br>Hpc2 | HIRA<br>HIRA<br>?<br>Yema | HIRA<br>HIRA<br>?<br>Ubn1 | HIRA<br>HIRA<br>Cabin1<br>Ubn1 | H3.3-H4<br>(yeast H3-H4) | | | | |
| Replication-independent H3.3-H4 histone chaperone, associated with transcription | | | | | | | | | |
| Daxx | - | DLP | Daxx | Daxx | H3.3-H4 | | | | |
| Replication-independent histone chaperone | | | | | | | | | |
| HJURP | Scm3 | HJURP | HJURP | HJURP | CENPA-H4 | | | | |
| Centromere-specific, replication-independent histone chaperone | | | | | | | | | |
| FACT | Spt16<br>Pob3 | Spt16<br>SSRP1 | Spt16<br>SSRP1 | Spt16<br>SSRP1 | H3-H4<br>H2A-H2B | | | | |
| Facilitates transcription through chromatin by H2A/H2B removal, Ssrp1 and Spt16 Peptidase domain also bind H3-H4 | | | | | | | | | |
| Spt6 | | | | | H3-H4 | | | | |
| Transcriptional elongation factor, also shown to contain histone chaperone activity | | | | | | | | | |
| NASP | (Hif1) | NASP | N1/N2 | sNASP<br>tNASP | H3-H4 | | | | |
| Associated with nuclear and cytosolic H3-H4, also *in vitro* chromatin assembly activity | | | | | | | | | |
| Rtt106 | Rtt106 | - | | | H3-H4 | | | | |
| H3-H4 histone chaperone in yeast | | | | | | | | | |
| Nucleoplasmin/Nucleophosmin/NPMs | | | | | H2A-H2B, H3-H4 | | | | |
| Histone storage chaperone for sperm and oocyte | | | | | | | | | |
| NAP1 | Nap1<br>Vps75 | Nap1 | Nap1 | NAP1L<br>SET<br>(APLF) | H2A-H2B, H3-H4 | | | | |
| Histone H3-H4 and H2A-H2B chaperone | | | | | | | | | |
| Chz1 | | | | | H2A.Z-H2B | | | | |
| Histone chaperone associated with replication-independent variant H2A.Z | | | | | | | | | |

#### 1.3.4.1 Applicability of *in vitro* results *in vivo*

Another challenge in determining histone chaperone mechanism is the interpretation of factor-dependent *in vitro* nucleosome reconstitution assays. These assays determine if a particular factor promotes nucleosome formation in histone-DNA mixtures at less than 250 mM salt and pH 7.5 to 8.0. These ionic and pH conditions allow recapitulation of general chromatin properties observed *in vivo*. If precipitation is avoided and nucleosomes are formed, then the factor is said to have nucleosome assembly activity and consequently thought to deposit histones *in vivo*. Such assembly activity is often assumed to stem from direct deposition of histones onto DNA by the histone chaperone. A recent thermodynamic study of Nap1-mediated nucleosome assembly however, reinforces that such a direct mechanism is not required *in vitro* or *in vivo* (Andrews et al., 2010). The results of this study and others (e.g. (Mazurkiewicz et al., 2006)) indicate that *any* histone-binding protein (or other chemical factor) with an affinity in the range of the respective histone-DNA complex product can potentially facilitate nucleosome assembly *in vitro*. *In vivo*, these histone-binding proteins may however function in steps that precede ultimate deposition on DNA. Asf1 is an example of a histone chaperone with nucleosome assembly activity only under certain *in vitro* conditions (Donham et al., 2011; Munakata et al., 2000; Umehara et al., 2002) that does not directly assemble nucleosomes in a *X. laevis* extract system (Koning et al., 2007; Ray-Gallet et al., 2007). Thus the ability to assemble nucleosomes *in vitro* may simply reflect an ability to bind histones tightly and avoid non-nucleosomal interactions.

The somewhat artificial scenario of the *in vitro* reconstitution assay is further indicated by experiments showing that poly-glutamic acid or poly-aspartic acid successfully assembly nucleosomes (Stein et al., 1979). Unlike Nap1, however, they do this by stabilizing the histone octamer even in the absence of DNA (Stein et al., 1979). This alludes to a purpose for the acidic stretches often found in histone chaperones and highlights that *in vitro* nucleosome reconstitution can in fact be informative. In particular, it is useful to identify potential pathways of nucleosome formation and to see how they are influenced by particular factors.

Another difficulty in interpreting *in vitro* mechanistic data for *in vivo* function relates to the specificity of a histone chaperone for particular histones. Nap1, for example, binds H2A-H2B and H3-H4 equally well *in vitro* (Andrews et al., 2008). Chz1 is also thought to be H2A.Z-H2B-specifc, and yet the co-structure fails to identify any H2A.Z-specific contacts (Zhou et al., 2008). It is thus more likely that specificity is achieved through direct competition or temporal and spatial regulation of histone chaperone activity *in vivo*. The fact that histone chaperones are often found in complex with other factors suggests that spatial and temporal regulation could be achieved by accessory subunits. Furthermore, it is still uncertain to what extent activities from biochemically independent, non-homologous histone chaperones are redundant in the cell.

### 1.3.5 Models of histone chaperone function

An important principle emerges from these introductory remarks: histone chaperones are mechanistically diverse. Mechanistic diversity reflects the many types of histone proteins (e.g. major-type, centromeric, replacement histones), the many contexts of 'chaperoning' activity, and the structural heterogeneity of

histone chaperone proteins. These physical and contextual variances ultimately make it extremely difficult to define histone chaperones from a mechanistic standpoint. I will therefore outline models for histone chaperone function based on known structural and biochemical data.

### 1.3.6 Histone chaperone structural diversity

Given the common task, a surprising number of unrelated structural motifs are found throughout known histone chaperones. While many full-length or multimeric histone chaperone structures are outstanding, several have been solved (e.g. (Daganzo et al., 2003; Dutta et al., 2001; Park and Luger, 2006; VanDemark et al., 2006; Zhou et al., 2008)). Structures of Nap1, Asf1 and NPM are illustrative examples. Nap1 adopts a homo-dimeric earmuff fold that is characterized by a non-coiled coil motif and two composite ⟨/® domains (Park and Luger, 2006). Asf1 and NPM, however, have ®-sandwich architectures with immunoglobulin-like and jelly roll topologies, respectively (Daganzo et al., 2003; Dutta et al., 2001). The Nap1 fold has only been found in functionally similar homologues such as Vps75 and SET (Muto et al., 2007; Park and Luger, 2008), while Asf1 and NPM folds have been observed in proteins with varied functions. Further, in both solution and in the crystal, Nap1 is an obligate homo-dimer, while Asf1 is a monomer and NPM is a homo-pentamer or decamer (Daganzo et al., 2003; Dutta et al., 2001). These oligomeric states are thought to persist *in vivo*. Hetero-oligomeric chaperones also exist with examples including FACT, composed of SSRP1 and SPT16; and CAF1, composed of p150, p60 and RbAp48, in humans. The latter two subunits adopt yet another structural motif, a ®-

propeller fold (Lejon et al., 2011; Song et al., 2008). This diversity in basic structure and oligomeric state suggests that each histone chaperone evolved independently to serve a specialized function in the cell.

Although histone chaperone structure is varied, one commonality somewhat widespread is the presence of regions rich in glutamic and aspartic acids. These acidic stretches are often found near the chaperones C-terminus and are presumably disordered, at least in the absence of a binding partner. A shortlist of chaperones containing such acidic regions includes Nap1, Asf1, Rtt106, Chz1, NPM, DAXX (not at the C-terminus) and the SPT16 subunit of FACT (Belotserkovskaya, 2003). Unlike the folded domains of histone chaperones, these acidic stretches are not necessarily conserved. Fungal Asf1, for examples, contains these stretches, while human Asf1 does not (Daganzo et al., 2003). Similarly, the length and composition of these stretches varies between the many Nap1-like proteins in humans, as well as between yeast and *Drosophila melanogaster* Nap1. The latter is intriguing as the acidic stretches in *D. melanogaster* Nap1 contain sites of polyglutamylation (Regnard et al., 2000). This PTM may compensate for a less acidic composition.

#### 1.3.6.1 Interactions between histones and their chaperones

For many histone chaperones, a missing link in the elucidation of mechanism is the absence of structural detail regarding their interaction with histones. While numerous chaperone-only structures are available, they often lend little insight into the mode of histone binding. Major-type co-structures solved to date include human and yeast Asf1 with H3-H4 (Agez et al., 2007; English et al., 2006; Natsume et al., 2007), and the homologous p46/p55 subunits

of human and *D. melanogaster* CAF1 respectively with a H4 peptide (Murzina et al., 2008; Song et al., 2008). Even though p46 paralog p48 assembles centromeric H3 nucleosomes *in vitro* (Furuyama and Henikoff, 2009), the significance of the latter in terms of histone chaperone mechanism is vague. This is because p46/p55 is a member of several complexes involved in histone biochemistry (Suganuma et al., 2008). It is interesting to note however that the insertion of the H4 histone fold ⟨-helix into the binding groove of p46/p55 requires a conformational rearrangement in the H3-H4 heterodimer. In the Asf1-H3-H4 co-structure, Asf1 binds at the H3 interface required for $(H3\text{-}H4)_2$ tetramer formation (Figure 1.12C). The C-terminal ®-strand of H4 also extends away from the H3-H4 dimer and inserts into a hydrophobic pocket in Asf1. In the nucleosome or histone octamer, this strand folds back and aligns with the C-terminal ®-strand of H2A. Both large and small conformational differences in the co-structures attest to a degree of plasticity in the arrangement of histone protein secondary structures, particularly when the histones are non-nucleosomal. Such plasticity, as well as the involvement of intrinsically unstructured regions, may account for the lack of co-structures available for the many other histone chaperones.

The insight provided by the Asf1-H3-H4 structure is testimony to the information that may be gleaned from co-structures. Asf1 binding to an obligate H3-H4 dimer has implications for the nucleosome assembly and disassembly mechanism, histone PTMs (as Asf1, like Vps75, activates histone acetyltransferase Rtt109 (Tsubota et al., 2007)) and a potential H3-H4 hand-off (see (Koning et al., 2007) for a more complete discussion). Overall, the Asf1-H3-

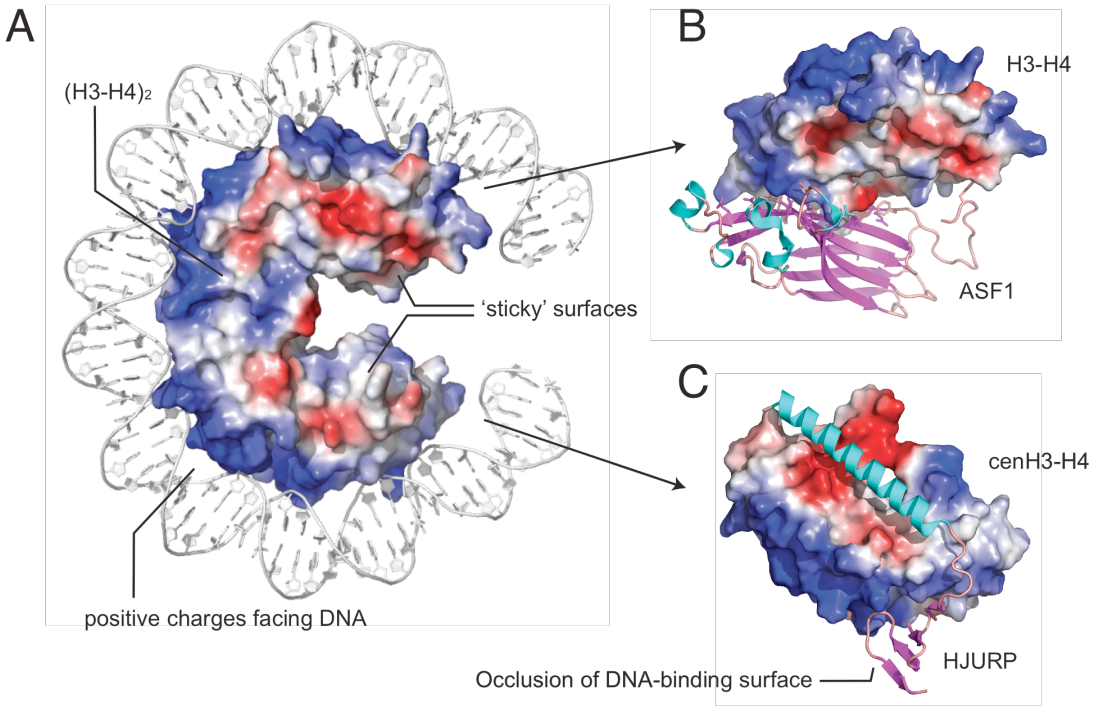
H4 co-structure clearly indicates that binding to histone chaperones and nucleosome formation, are mutually exclusive. This idea is echoed in other co-structures available involving variant histones. These include Chz1 with H2A.Z-H2B (Zhou et al., 2008), and HJURP/Scm3 with cenH3-H4 (Cho and Harrison, 2011; Hu et al., 2011; Zhou et al., 2011). HJURP and Scm3 are cenH3-H4-specific chaperones in human and yeast respectively. Histone-binding by HJURP/Scm3 is incompatible with the histone-DNA interactions observed in the nucleosome (Figure 1.12C).

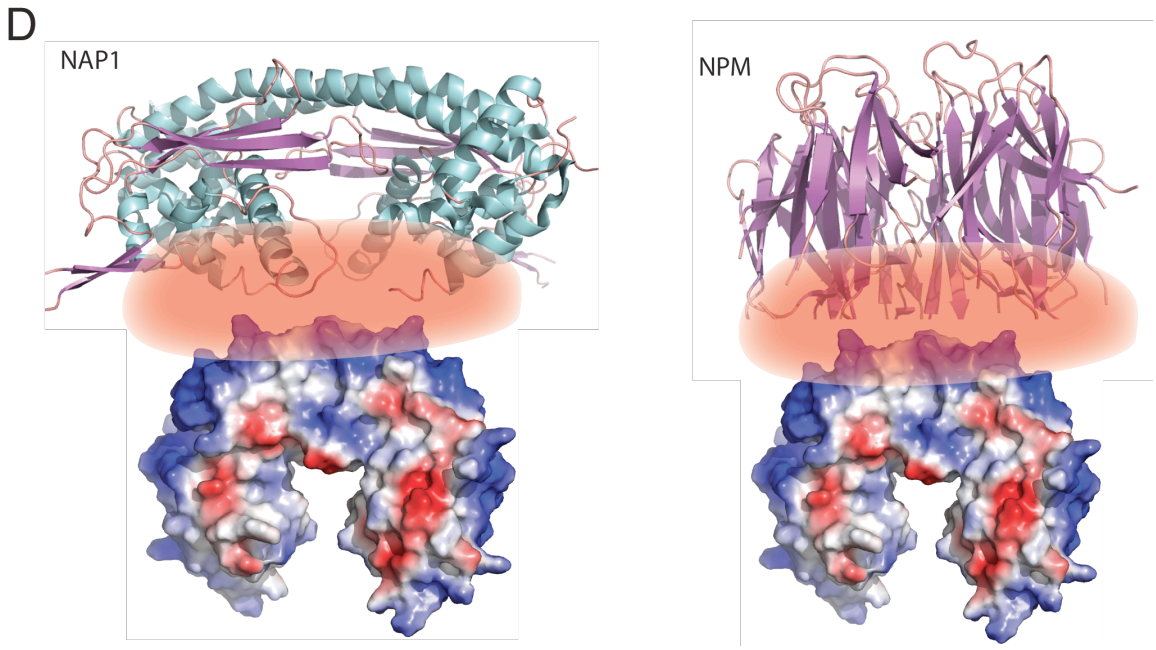
Like Asf1, HJURP and Scm3 preclude $(cenH3\text{-}H4)_2$ tetramer formation, albeit through a different mechanism.In HJURP-cenH3-H4, the interface required for $(cenH3\text{-}H4)_2$ tetramer formation is sterically occluded, while in Scm3-cenH3-H4 it is also distorted. Another common feature of these co-structures is the presence of extended hydrophobic interactions at the histone interface. This reinforces that histone chaperones are likely to be doing more than just mimicking salt.

The co-structures of HJURP and Scm3 with cenH3-H4 have also yielded insight into histone chaperone discrimination of H3 variants. Both HJURP and Scm3 contain a long ⟨-helix that runs anti-parallel to the central ⟨-helix of the cenH3 histone fold. Overall affinity is achieved through interlinking hydrophobic residues flanked by salt bridges, reminiscent of a coiled-coil interaction. Variant-specificity is conferred through small highly conserved alterations in the hydrophobic surface, but also through neighboring loop-loop contacts, spanning the region of cenH3 disparate in sequence to major-type H3. It is interesting to note that variations in the same hydrophobic surface occur in

mammalian variants H3.1, H3.2 and H3.3. As this region does not make crucial contacts in the histone octamer, it may have evolved novel functions more easily.

Figure 1.12: Nature of histone-DNA and histone-histone chaperone interactions. (**A**) A vacuum electrostatic surface charge map of $(H3\text{-}H4)_2$ with 72 base pairs of DNA from PDB 1KX5. Basic surfaces (blue) accompany the path of DNA. The top, bottom and inner surfaces are partly involved in inter-histone interactions and are more hydrophobic. The four ⟨-helix bundle of the $(H3\text{-}H4)_2$ tetramer is buried in (**A**). H3-H4 dimers in (**B**) and (**C**) are shown in an identical orientation to (**A**). (**B**) Asf1 (cartoon) binds along the four ⟨-helix bundle interface and prevents $(H3\text{-}H4)_2$ tetramer formation (PDB 2HUE). (**C**) HJURP (cartoon) binds along ⟨2 of CENP-A, disrupting $(cenH3\text{-}H4)_2$ tetramer formation. HJURP also blocks DNA binding through a C-terminal ®-sheet domain (PDB 3R45). (**D**) Schematics of NAP1- and NPM-histone interaction, basic surfaces are thought to mediate electrostatic interactions (indicated in red).


A
(H3-H4)2
'sticky' surfaces
positive charges facing DNA
B
H3-H4
ASF1
C
cenH3-H4
HJURP
Occlusion of DNA-binding surface


D
NAP1
NPM

#### 1.3.6.2 Mechanistic implications of distinct histone-binding modes

Based on *in vitro* experiments, unbound or free H3-H4 and H2A-H2B units are presumed to be 'sticky' in the cell. The excess positive surface charges of free histones limits the formation of the histone octamer, as well as the $(H3\text{-}H4)_2$ tetramer at physiological conditions (Stein et al., 1979). As a consequence, normally protected, hydrophobic histone-histone interfaces become solvent-exposed outside of the nucleosomal octameric arrangement (Banks and Gloss, 2003, 2004; Baxevanis et al., 1991; Eickbush and Moudrianakis, 1978; Karantza et al., 1996). In this state of high potential energy, the histones readily engage in non-specific interactions , electrostatic or hydrophobic in nature, with nucleic acids, proteins or other cellular components (Andrews et al., 2010; Stein, 1979) (see also Figure 1.13). The putative detriment of these non-nucleosomal complexes to nucleosome formation has been illustrated by a study with Nap1 (Andrews et al., 2010). Based on available structural and biochemical data, we propose that histone chaperones exhibit two alternative histone-binding modes that, by distinct mechanisms, can both prevent promiscuous histone interactions. It remains to be seen if both modes can be employed by a single histone chaperone.

The first mode involves *neutralizing excess positive charge* by providing acidic residues to alleviate charge repulsion, as exemplified by the well-studied histone chaperones NAP-1 and NPM (Figure 1.12D) This mode seems to provide little selectivity for H2A-H2B or H3-H4 (and in some cases the linker histone family), and alludes to a purpose of the large negatively charged surfaces found in numerous histone chaperones. The acidic residues may be

structured or unstructured in the apo-histone chaperone, and may form either tight or transient interactions with the histones. Phosphorylation of non-acidic residues will also contribute to this mode, as observed for NPM (Cotten et al., 1986; Taneva et al., 2008). The involvement of PTMs that regulate chaperone-histone affinity might be crucial to ensure that chaperone activity occurs in the applicable biological context. While alleviation of electrostatic repulsion favors self-association of histone hetero-dimers into $(H3\text{-}H4)_2$ tetramers and potentially histone octamers, steric constraints imposed by chaperone architecture would confer additional specificity. For example, the symmetric geometry of Nap1 and Vps75 facilitates binding of an obligate (H3-H4)2 tetramer (Bowman et al., 2011). The mechanistic benefits of electrostatic shielding stem from the histone chaperones ability to directly compete with non-specific electrostatic interactions, particularly with DNA. Under favorable thermodynamic conditions, pre-formed $(H3\text{-}H4)_2$ tetramers, H2A-H2B dimers or histone octamers can be 'presented' to the recipient DNA or other histone chaperones.

The second, fundamentally different mode is to *block histone hydrophobic surfaces.* This mode is evident in the co-structures of Asf1-H3-H4 (Figure 1.12B), HJURP-cenH3-H4 (Figure 1.12C), Scm3-cenH3-H4 and Chz1-H2A.Z-H2B (English et al., 2006; Hu et al., 2011; Natsume et al., 2007; Zhou et al., 2008; Zhou et al., 2011). This mode is likely to be more specific as surface-exposed amino acids are divergent between H2A-H2B and H3-H4. In fact, subtle variations in the hydrophobic surfaces of the H3-H4 dimer may form the basis of the histone variant specificity observed for a number of histone chaperones. Histone chaperones that employ this mode are expected to retain histone-binding at high ionic strength. Indeed, Asf1, DAXX, HJURP and Scm3 histone complexes are

stable at high salt, with hydrophobic contacts outweighing salt bridges and hydrogen bonds (Drané et al., 2010; Hu et al., 2011; Lewis et al.; Natsume et al., 2007; Zhou et al., 2011). While the most immediate effect of this mode counteracts aggregation via hydrophobic histone surfaces, the chaperone can also prevent nucleosome formation through steric occlusion. As described previously, Asf1 and HJURP/Scm3 prevent $(H3\text{-}H4)_2$ tetramer formation and occlude part of the DNA-binding site of the H3-H4 dimer (Figure 1.12) (English et al., 2006; Hu et al., 2011; Natsume et al., 2007; Zhou et al., 2011). Promiscuous histone interactions are thus efficiently prevented, despite retaining some of the excess positive charge in the chaperone-histone complex.

The use of diverse and somewhat complementary histone-binding modes reinforces the variety of mechanisms employed by histone chaperones to accomplish the seemingly simple task of preventing aggregation. This diversity, however, sets the stage for the formation of multi-chaperone complexes. Histones may be passed between chaperones within these complexes before eventually being made available for nucleosome assembly. Other components of these complexes may be integral in fine-tuning histone-binding and -release onto DNA or other pre-nucleosomal complexes. This resolves the conundrum of histone chaperones essentially rendering the histones inert for any (specific or non-specific) interactions with other histones or DNA.

#### 1.3.6.3 Thermodynamic and kinetic components of histone chaperone activity

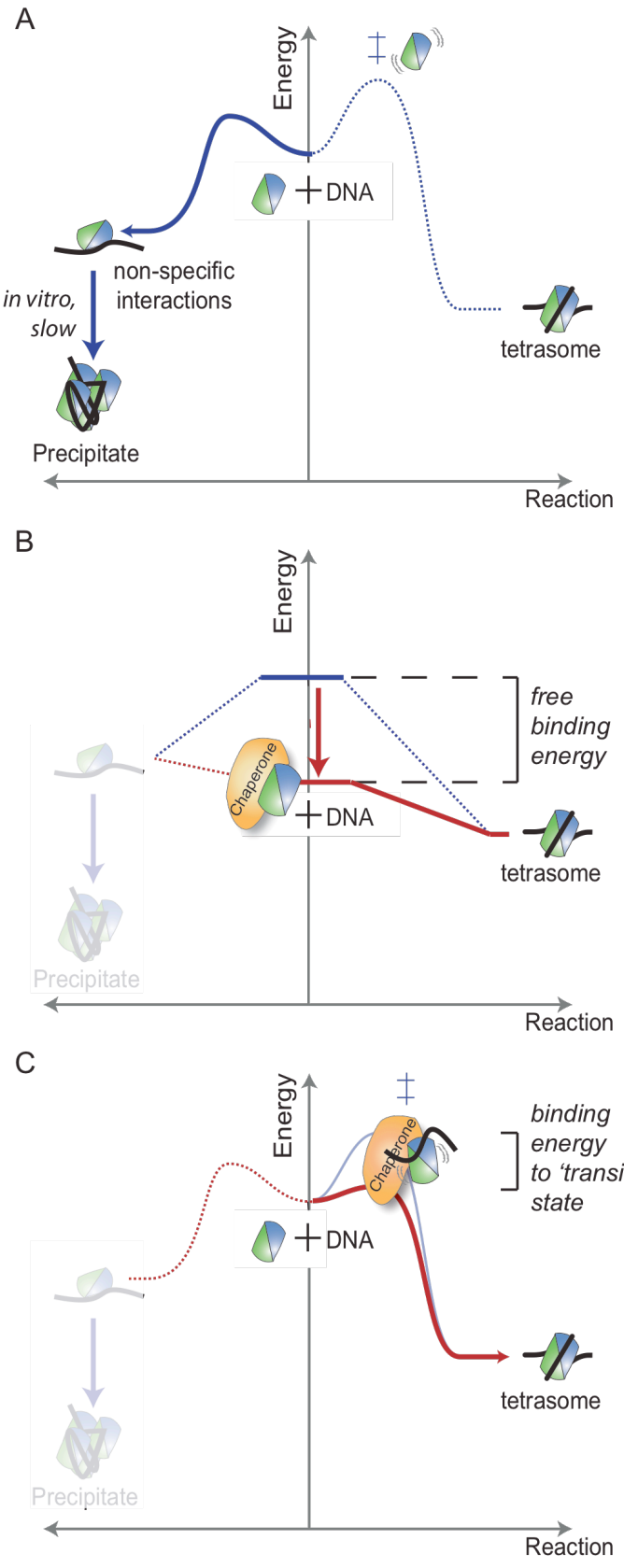
Potential mechanisms of histone chaperones can be illustrated using simplified reaction diagrams such as those shown in (Figure 1.13). These diagrams are based on thermodynamic constants of complexes involving histone

chaperones, histones and/or DNA. As described earlier, free histones readily engage in electrostatic interactions with DNA. These interactions can be specific and associated with nucleosome formation (Figure 1.13, right) or so-called non-specific resulting in irreversible precipitation *in vitro* (Figure 1.13, left). The specific complexes may be $(H3\text{-}H4)_2$ tetramer bound to DNA (tetrasome), tetrasome plus one copy of H2A-H2B dimer (hexasome), or the nucleosome. In the absence of a histone chaperone, the predominant reaction product will be the non-specific complexes because the specific complexes involve only a small subset of all potential histone-DNA interactions (Figure 1.13A).

A histone chaperone may promote specific complex formation by lowering the free energy of the histones (Figure 1.13B). Free energy is lowered by the formation of a chaperone-histone complex that is thermodynamically favored over non-specific histone-DNA complexes. This mechanism has been proposed for Nap1, which binds either H3-H4 or H2A-H2B with nano-molar affinity *in vitro* (Andrews and Luger, 2011). Through these interactions, Nap1 assembles nucleosomes via tetrasome and hexasome intermediates (Figure 1.9) (Mazurkiewicz et al., 2006). The two key measurements are the affinities of H2A-H2B for DNA (non-specific), and H2A-H2B for the tetrasome (specific) (Andrews et al., 2010). The absolute affinities vary depending on the DNA sequence, with a high-affinity sequence (the so-called Widom 601) measuring 44 ± 5 nM and 13 ± 3 nM correspondingly (Andrews et al., 2010). These numbers fit well with the Nap1 affinity for H2A-H2B of 7.8 ± 0.4 nM insofar as H2A-H2B will prefer to bind Nap1 over DNA (7.8 < 44 nM), but binds to Nap1 or tetrasome at similar affinity (7.8 ~ 13 nM).

Figure 1.13: Models of histone chaperone function illustrated as free-energy reaction diagrams.

The change in free-energy is plotted along the reaction coordinate as the substrate (DNA and histones, middle) becomes the product (precipitate, left, or tetrasome, right). Analogous to a standard chemical reaction, substrates are postulated to go through a 'transition state' of highest free-energy (‡) before forming the product. The transition state represents a kinetic barrier for the reaction. (A) shows that in the absence of histone chaperones, non-specific interactions of histones (here H3-H4) with DNA are favored (left) over the tetrasome formation (right). Tetrasome formation requires histones and DNA to encounter in a non-random orientation. (B) shows that a Nap1-type chaperone binds histones with a free-energy similar to the final tetrasome product. In the chaperone-bound state, non-specific interaction with DNA (left) is costly, and thus the tetrasome formation (right) is favored. (C) shows an alternative mode of chaperone action. The chaperone binds to form a 'transition state' that is competent to directly deposit histones onto DNA. This state might represent a chaperone-histone-DNA trimeric complex, or a conformationally activated histone moiety that is primed to form nucleosomal DNA contacts (right).




In the absence of a histone chaperone, non-specific histone-DNA interactions are favored over the non-random interactions of the tetrasome.

In the presence of a histone chaperone, non-specific histone-DNA interactions are efficiently competed by the more stable histone chaperone-histone complex.

The histone chaperone-histone complex is a 'transistion' state able to directly deposit histones into a tetrasome arrangement.

When generalized to other histone chaperones with nano-molar histone affinities, it seems that chaperones may make histones available for nucleosome formation (Figure 1.13B, right) by disfavoring non-specific histone-DNA complexes (Figure 1.13B, left). Unfortunately, as little is known about the kinetics of chaperone-histone complexes, it is unclear if the substrates of nucleosome assembly are the chaperone-histone complexes or the small proportion of free histones in dynamic equilibrium.

Histone chaperones may also guide nucleosome assembly in a more directed fashion by presenting the histones in the 'correct' orientation (Figure 1.13C, right). In contrast to the case above (Figure 1.13B), the chaperone does not necessarily lower the free energy of the histones. Instead, it is involved in a transient chaperone-histone-DNA complex that is resolved by DNA wrapping and chaperone dissociation. By increasing the likelihood of specific interactions between histone and DNA, the chaperone skews the reaction kinetics towards a specific histone-DNA complex (i.e. tetrasome, hexasome or nucleosome) (Figure 1.13C, right). The formation of a transient ternary complex causes an apparent decrease in activation energy to be surmounted for nucleosome formation, favoring it over non-specific histone-DNA interactions (Figure 1.13, left). While such ternary complexes have not been observed directly, as expected by their transient nature, they can be inferred in a few *in vitro* cases. Asf1, for example, facilitates formation of a disome [$(H3\text{-}H4)_1$-DNA] opposed to the tetrasome formed in the absence of chaperone (Donham et al., 2011). Similarly, for the reverse, nucleosome disassembly reaction, chaperone-nucleosome complexes may occur for FACT (Orphanides et al., 1999) and Nap1 (Park et al., 2005). It

should also be noted that a transient, ternary complex may also be involved in preventing non-specific histone-DNA interactions. If histone chaperones do directly deposit histones, how is the transition made kinetically and thermodynamically more favorable? One possible answer is that histone chaperones exploit the conformational flexibility of the histone fold. Co-structures involving p46/p55 and Scm3 suggest that the histone fold can be distorted upon chaperone binding (Murzina et al., 2008; Song et al., 2008; Zhou et al., 2011). These conformational changes may represent high-energy 'transition' states resolved by nucleosome formation upon encounter of DNA.

#### 1.3.6.4 Nucleosome assembly versus disassembly

One might wonder if nucleosome disassembly is a direct reversal of a given assembly pathway. Chaperone-histone affinities range from micromolar to nanomolar and are thus equivalent of the nanomolar free energy of a nucleosome (Andrews et al., 2010). As such, it has been speculated that histone chaperones can directly facilitate histone eviction. This notion is supported by excess Nap1 removing H2A-H2B dimers from a preformed nucleosome *in vitro* (Park et al., 2005). As discussed previously, affinities of H2A-H2B for tetrasome and Nap1 are matched to allow assembly reactions to be driven in both directions. The actual direction will depend on the abundance of Nap1, histones and DNA. FACT has also been shown to be competent to remove H2A-H2B from a nucleosome in the context of transcription (Belotserkovskaya, 2003). It is important to note however that the energy required to evict H2A-H2B (and in fact H3-H4) depends on nucleosome conformation and stability, with histone PTMs, DNA sequence and buffer all being influential variables (Andrews et al., 2010). These variables most likely contribute to the likelihood of histone chaperone-mediated nucleosome disassembly *in vivo*.

In contrast to H2A-H2B, the $(H3\text{-}H4)_2$ tetramer has higher dissociation energy and more potential to be kinetically trapped. In the context of the nucleosome, kinetically trapped simply means that the DNA does not visit an 'unwrapped' conformation for long enough to allow the histones to dissociate. This is indicated by excess Asf1 failing to disassemble nucleosomes or tetrasomes at physiological salt concentration (Donham et al., 2011). Thermodynamics alone would predict that Asf1 would disassembly nucleosomes, given its nanomolar

affinity that is comparable to the tetrasome energy (Donham et al., 2011). Interestingly, thermal unwinding of the DNA still does not allow Asf1 to compete for H3-H4 dimer, likely because at least half of the tetrasome DNA would need to be unwrapped to uncover its binding site (Donham et al., 2011). In light of nucleosomes being kinetically trapped, it seems probable that histone chaperones alone cannot directly disassemble nucleosomes *in vivo*. However, they may facilitate disassembly alongside ATP-dependent remodelers, as hydrolysis of two to three ATP molecules would theoretically provide sufficient energy to completely dissociate all histone-DNA interactions within the nucleosome (Saha et al., 2006). Furthermore, mechanical stress imposed by DNA-templated processes such as the positive super-coils induced by traveling RNA polymerase, might destabilize nucleosomes sufficiently to allow direct disassembly (Schwabish and Struhl, 2006). Therefore, the dynamic environment *in vivo* will alleviate some of the kinetic barriers observed *in vitro* and allow a direct competition of histone chaperones and DNA for histones. Nucleosome density could therefore be controlled by mass action of histone chaperones (with high affinity such as Asf1 for H3-H4 and NAP1 for H2A-H2B).

The acidic stretches found in many histone chaperones may also function in nucleosome disassembly. That is, they may initiate the unwrapping of the nucleosomal DNA (or stabilize it in an 'open' state) to allow ATP-dependent chromatin remodeling activity or histone eviction (Park et al., 2005). This is consistent with acidic stretches also being found in other chromatin modifying enzymes, such as the histone methyltransferase SET1, as well as ATP-dependent remodelers such as SWR1 and ATRX. An accessory function accounts for the fact

that these acidic stretches are sometimes dispensable and not necessarily conserved.

#### 1.3.6.5 Channeling histones along a histone chaperone pathway

The regulation of histone chaperone activity is also at the core of their *in vivo* mechanism. How do histone chaperones ensure that the thermodynamic flow of histones is regulated in an ordered fashion? This is particularly important as there is evidence that chromatin assembly pathways compete for histone supply (Lewis et al., 2010). A theme emerging from recent literature is the regulation of histone hand-offs by histone PTMs (Figure 1.14). For example, the acetylation of H4-lysine 5 and H4-lysine 12 is well-established and promotes histone association with nuclear import complexes (Alvarez et al., 2011).

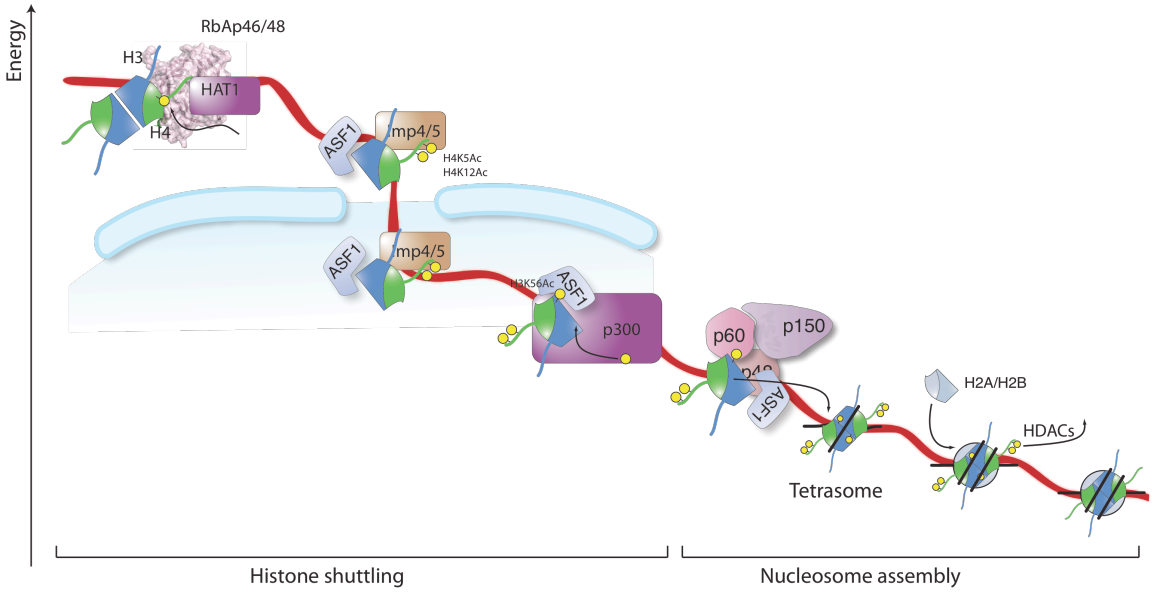



Figure 1.14: Ordered hand-offs in a chromatin assembly pathway.
A series of histone chaperones mediates the shuttling of histones from their synthesis to incorporation into the nucleosome. Histone acetyl transferases (such as HAT1 and p300) transiently acetylate H3 and H4. These acetylations are thought to facilitate the ordered hand-off or creating checkpoints along the pathway. For details see main text.

H3 is further acetylated at lysine 56 before chromatin assembly (Li et al., 2008; Masumoto et al., 2005; Recht et al., 2006). The histone chaperone Asf1 is a required cofactor in the acetylation of H3 at lysine 56 (Das et al., 2009; Tang et al., 2011). The multiple acetylation of H3-H4 in turn facilitates nuclear import and may enhance H3 affinity for CAF1 thereby promoting chromatin assembly (Barman et al., 2008; Das et al., 2009; Glowczewski et al., 2004; Li et al., 2008). Following incorporation, H3-lysine 56 is deacetylated, potentially creating a more stable nucleosome (Andrews et al., 2010). Similarly, a recent study implicates phosphorylation of H3-serine 47 in modulating association with histone chaperone complexes and chromatin assembly (Kang et al., 2011). A series of histone hand-offs regulated by histone PTMs can thus be delineated. Such series would fine-tune histone chaperone activity in the context of chromatin homeostasis, essentially creating a series of checkpoints that keep the histone traffic ordered (Barman et al., 2008; Glowczewski et al., 2004; Groth et al., 2007; Rufiange et al., 2007; Sharp et al., 2005).

#### 1.3.6.6 Outlook

Since their discovery, the unusual and unique biochemistry of the histone proteins has continued to astound and amaze (Kornberg and Thomas, 1974; Lewis, 1976; van Holde, 1988). A wealth of advances over the last decade has transformed them from static to dynamic modulators of virtually all processes that act or depend on DNA. Closely associated both figuratively and literally, with the histones, is a network of histone chaperones. Through structural and functional diversity, the histone chaperones facilitate nucleosome formation. They are involved in all facets of histone biology, from folding in the cytoplasm

to nuclear import, storage and deposition into chromatin. The precise mechanisms of histone chaperone function, however, have remained elusive. The lack of chaperone-histone co-structures, difficulties interpreting *in vitro* results in an *in vivo* context, and a poor understanding of the biologically relevant chaperone substrate and product, have hindered mechanism determination. Even with these limitations, we can postulate several models of histone chaperone mechanism by considering the thermodynamics of nucleosome assembly and disassembly. These models relate to how histone chaperones bind to their substrates, prevent promiscuous histone interactions and balance nucleosome assembly and disassembly, the fate of chaperone-bound histones, and the importance of histone hand-offs between different histone chaperones. Testing these hypotheses will be the seed for future studies aimed at determining histone chaperone mechanism.

# 2 The H3.3-specific histone chaperone Daxx cooperates with ATRX in replication-independent chromatin assembly at telomeres[‡]

## 2.1 Introduction

The assembly of chromosomal DNA into nucleosomes represents the most fundamental step in the formation of eukaryotic chromatin structure. The assembly, remodeling and eviction of nucleosomes have been shown to be important for a variety of DNA-templated processes such as replication, repair and transcription. Histone deposition pathways are thought to play a critical role in the establishment and maintenance of epigenetic information encoded by histone modifications, nucleosome positioning and higher-order chromatin structure (see also Section 1.1). As discussed in Section 1.2.6, higher eukaryotes utilize separate chaperones and deposition pathways for the different histone H3 variants. Previous work identified two major pathways: replication coupled deposition of H3.1/H3.2 by the CAF1 complex, and replication independent

[‡] This chapter has been published as: Lewis, P. W.*, Elsaesser, S. J.*, Noh, K.-M., Stadler, S. C., and Allis, C. D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proc Natl Acad Sci USA*, 107, 14075–14080. Where primary data was contributed by others, the authorship is indicated in the figure legend.

(* equal contribution)

deposition of H3.3 by the HIRA complex (Ray-Gallet et al., 2002; Tagami et al., 2004; Verreault et al., 1996).

While originally associated with euchromatic sites of active transcription, H3.3 has recently been found associated with regulatory elements and constitutive heterochromatin at telomeres (Goldberg et al.; Jin et al., 2009; Mito et al., 2007; Wong et al., 2009). We found that HIRA is required for localization of H3.3 to actively transcribed regions, while the α-thalassemia X-linked mental retardation protein ATRX is essential for H3.3 incorporation at telomeres. Apart from ATRX, we also identified the death domain associated protein Daxx in H3.3 immunoprecipitations (IP) (Goldberg et al.). Daxx and ATRX have been shown to form a complex (Tang et al., 2004) and colocalize at pericentric heterochromatin and promyelocytic leukemia bodies (PML bodies) (Nan et al., 2007; Xue et al., 2003). Loss of ATRX leads to epigenetic alterations, including abnormal levels of DNA methylation at repetitive elements such as telomeres in murine cells (Gibbons et al., 2000). Moreover, ATRX and H3.3 have essential roles in maintaining telomere chromatin (Goldberg et al.; Wong et al., 2010).

To gain further insight into H3.3-specific deposition pathways, we sought to identify the direct binding partner of H3.3. Here we show that Daxx binds directly to H3.3, and importantly, this binding is specific for H3.3 and not H3.1. We find that Daxx alone, or when present in the ATRX-Daxx complex, can effectively assemble H3.3-containing nucleosomes. Additionally, we show that ATRX recruits Daxx to telomeres and that both complex subunits are required for H3.3 deposition at telomeric chromatin in murine ESCs.

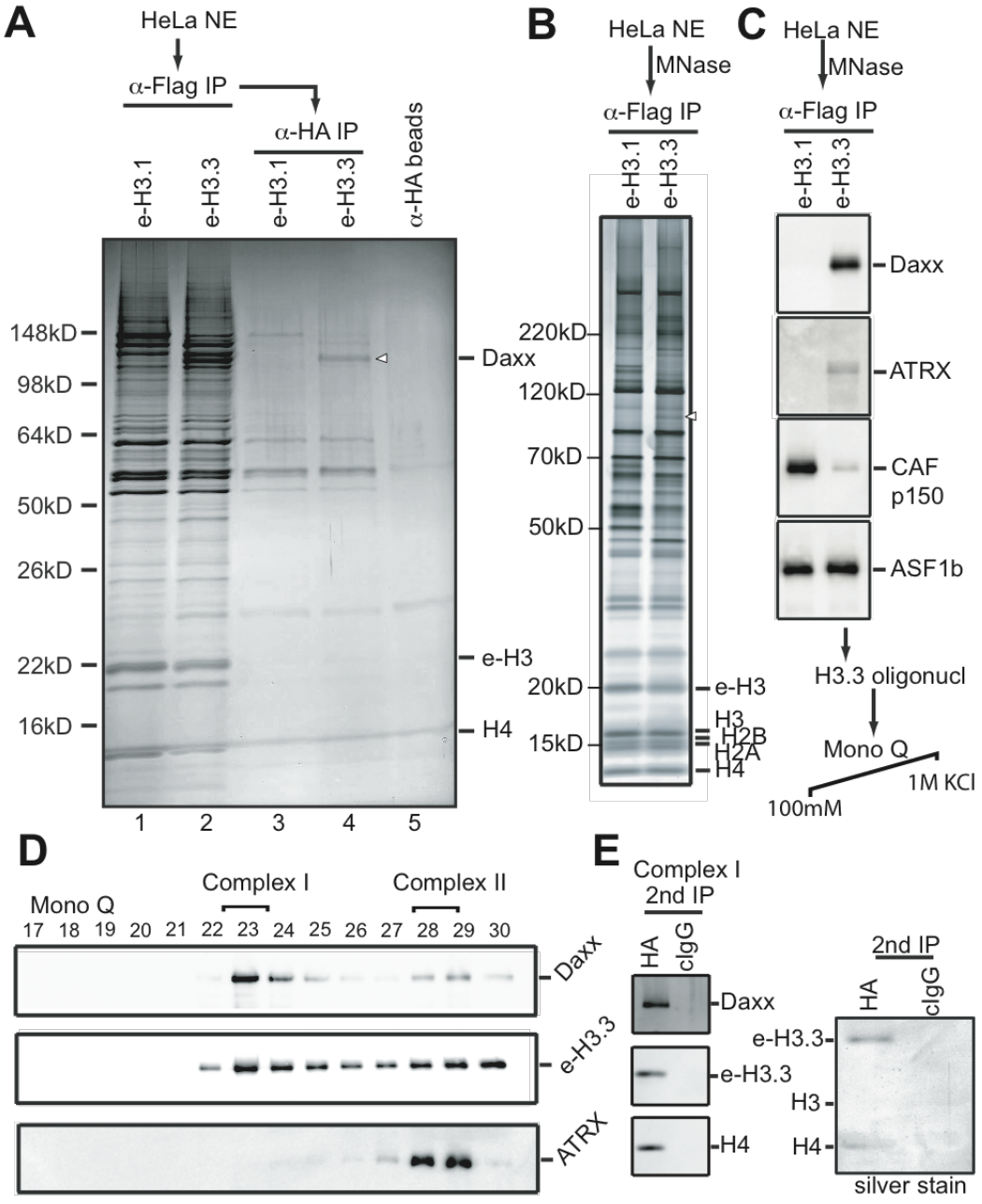
## 2.2 Results

### 2.2.1 Daxx associates specifically with H3.3

Having identified Daxx and ATRX as novel H3.3-associated factors, we wanted to determine if either of these polypeptides binds directly to H3.3, but not to H3.1. Using HeLa cells lines that stably express FLAG-HA-tagged H3.1 or H3.3 (hereafter e-H3.1 and e-H3.3) (Tagami et al., 2004), I isolated minimal H3.1 and H3.3-containing complexes by tandem-affinity purification from soluble nuclear extracts with a series of stringent salt washes of up to 1 M NaCl (Figure 2.1A). After sequential anti-FLAG and anti-HA IP, I consistently recovered a single polypeptide band of ~100 kD apparent molecular weight that co-purified with H3.3, but not H3.1 (see arrowhead next to lane 4); mass spectrometric analysis of this band yielded a significant number of peptides matching Daxx. Concurrent with previous results (Tagami et al., 2004), I also detected HIRA in the first H3.3 co-IP step (Figure 2.2). While we recently found ATRX associated with H3.3 in whole-cell preparations (Goldberg et al., 2010), it was only weakly associated with H3.3 in soluble nuclear extracts (Figure 2.2).

As the tandem-affinity chromatography was performed on nuclear extracts that did not contain chromatin, I reasoned that the H3.3 and H3.1 complexes purified from these soluble nuclear extracts likely represent a pool of pre-deposition histones. We also sought to identify proteins that were enriched with H3.3 in the chromatin fraction. To this end, my colleague Dr. Peter Lewis solubilized oligonucleosomes from the tagged cell lines by brief digestion with micrococcal nuclease (MNase) and isolated H3.1- and H3.3-enriched chromatin by FLAG IP (Figure 2.1B).

Figure 2.1: Identification of Daxx as an H3.3-specific binding protein.
(A) Silver stain of H3.1 and H3.3 associated proteins. Soluble nuclear extracts were prepared from Hela cells that express e-H3.1 or e-H3.3. FLAG-M2 immunoprecipitation was performed on the nuclear extract (lanes 1,2). The eluate from the FLAG immunoprecipitation was then subjected to HA immunoprecipitation (lanes 3, 4). HA immunoprecipitation was performed on untagged Hela nuclear extract as a control (lane 5). Protein bands from the HA affinity column were then subjected to identification by mass spectrometry. (B) Silver stain of FLAG eluate from H3.1 and H3.3 chromatin-associated fractions. Nuclei from the e-H3.1 and e-H3.3 Hela cells were subjected to brief MNase treatment. Solublized chromatin was clarified and FLAG M2 immunoprecipitation was performed. (C) Immunoblot was performed on the FLAG eluate with anti-Daxx, anti-ATRX, anti-CAF150, and anti-ASF1a/b sera. (D) Fractionation scheme for H3.3 chromatin-associated proteins. (E) Eluate from chromatin associated H3.3 was resolved by Mono Q chromatography. Immunoblotting analysis was performed on the fractions with anti-Daxx, anti-FLAG and anti-ATRX. (F) Immunoblot of anti-HA and control rabbit IgG immunoprecipitation of e-H3.3 from pooled Mono Q complex I fractions (left panel). Silver stain of immunoprecipitated material (right panel). (B-E) contributed by P. Lewis.


A
HeLa NE
α-Flag IP
α-HA IP
e-H3.1
e-H3.3
e-H3.1
e-H3.3
α-HA beads
148kD
98kD
64kD
50kD
26kD
22kD
16kD
Daxx
e-H3
H4
1
2
3
4
5
B
HeLa NE
MNase
α-Flag IP
e-H3.1
e-H3.3
220kD
120kD
70kD
50kD
20kD
15kD
e-H3
H3
H2B
H2A
H4
C
HeLa NE
MNase
α-Flag IP
e-H3.1
e-H3.3
Daxx
ATRX
CAF
p150
ASF1b
H3.3 oligonucl
Mono Q
1M KCl
100mM
D
Mono Q
Complex I
Complex II
17 18 19 20 21 22 23 24 25 26 27 28 29 30
Daxx
e-H3.3
ATRX
E
Complex I
2nd IP
HA
cIgG
Daxx
e-H3.3
H4
2nd IP
HA
cIgG
e-H3.3
H3
H4
silver stain

By immunoblot analysis, he found that both Daxx and ATRX were associated with H3.3, but not H3.1 chromatin (Figure 2.1C). The p150 subunit of the replication-coupled Chromatin Assembly Complex 1 (CAF1) was found primarily but not exclusively associated with the H3.1-associated fraction. The histone chaperone ASF1 was found in both e-H3.1 and e-H3.3 eluates. We further fractionated proteins associated with H3.3 chromatin using Mono Q anion-exchange chromatography (Figure 2.1D). Immunoblot analysis of the column fractions revealed two biochemically distinct populations of Daxx. The majority of Daxx eluted from the Mono Q column at ~0.4 M KCl, (complex I). While this fraction contained H3.3 he did not detect any ATRX by immunoblot. The second population of H3.3-bound Daxx eluted from the Mono Q column at higher salt (0.75 M KCl). These fractions contained Daxx, ATRX and H3.3 (complex II).

The isolation of two distinct populations of Daxx provided further evidence that Daxx is the direct binding partner of H3.3. In order to determine if Daxx was bound to H3.3 in complex I, he performed HA-IP with pooled fractions and immunoblotted for Daxx (Figure 2.1E). he found that Daxx co-immunoprecipitated H3.3 in these fractions. We failed to detect H2A, H2B, or endogenous untagged histone H3 in the Mono Q fractions that contained complex I. Immunoblot and silver stain analysis indicated that only epitope tagged-H3.3 and endogenous H4 are present in complex I (Figure 2.1E). As untagged H3 was not associated with Daxx, we concluded that the H3.3-Daxx pre-deposition complex includes a heterodimer of H3.3-H4. Consistent with this result, others had previously described purified pre-deposition complexes (e-H3.1.com and e-H3.3.com) that contained H3-H4 heterodimers (Tagami et al., 2004).

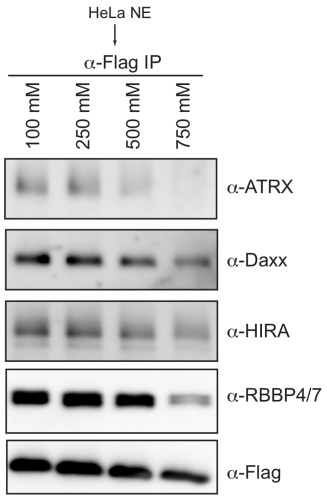


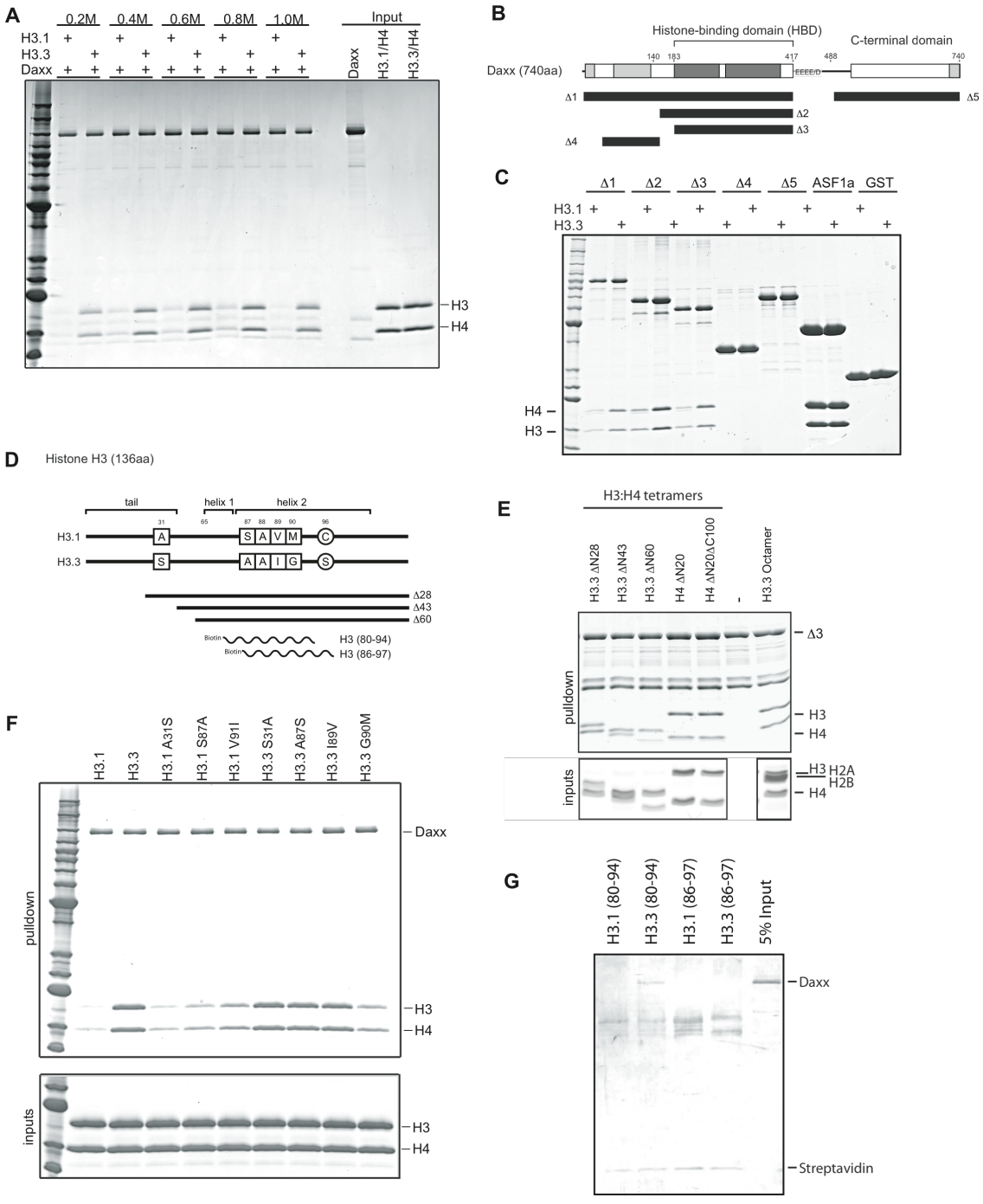

Figure 2.2: ATRX association with H3.3 in soluble nuclear extract is salt-dependent.
FLAG-M2 immunoprecipitation from e-H3.3 cell lines were performed as in Figure 1A, washes were performed with indicated NaCl concentrations. The FLAG eluates were immunoblotted for ATRX, Daxx, HIRA, RBBP4/7, FLAG.

### 2.2.2 Daxx directly binds the globular core of H3.3

We next carried out *in vitro* pull-downs to confirm the direct and variant-specific interaction with Daxx and H3.3. Recombinant $His_6$-tagged human histone proteins were prepared from *E. coli* and recombinant human Daxx from insect cells. Interestingly, a small fraction of the purified Daxx isolated from insect cells was stably bound to endogenous histone H3-H4 (Figure 2.3A). Recombinant Daxx protein bound our reconstituted H3.3-H4 tetramers over a wide range of salt concentrations, demonstrating that Daxx is a novel histone-binding protein. Notably, Daxx bound H3.1-H4 much less avidly at physiological, as well as high salt concentrations (Figure 2.3A).

The 740 amino acid human Daxx is comprised of an N-terminal, overall basic (pI 9.3), alpha-helical domain and an S/P/T-rich C-terminal fold connected by a likely unstructured highly acidic linker region (Figure 2.3B). The N-terminal domain contains a highly conserved region found in all metazoan Daxx proteins (amino acids 183-417) with no homolog in yeast (Figure 2.4A).

Figure 2.3: Daxx directly interacts with histone H3.3.
(A) Co-immunoprecipitation of H3.1 and H3.3 with recombinant FLAG-Daxx. Recombinant H3.1-H4 or H3.3-H4 were mixed with Daxx and resulting complexes captured on M2-FLAG agarose. Agarose bead were washed extensively with buffers of indicated KCl concentrations. Eluates and inputs were resolved by SDS-PAGE gel and stained with Coomassie Blue. (B) Domain structure of human Daxx. Shaded boxes indicate highly conserved regions (Dark gray: highly conserved in all metazoan Daxx, Light gray: highly conserved in vertebrates). GST-fusion constructs corresponding to different domains of Daxx are shown in black. (C) GST-pulldown experiment with Daxx fragments shown in B) with H3.3-H4 or H3.1-H4. (D) Schematic showing amino acid differences between H3.1 and H3.3. Relevant secondary structure features are indicated. Solid lines represent histone deletion constructs use in E), biotinylated peptides used in G) are shown below. (E) GST-Pulldown with the Daxx HBD (Δ4) and H3.3-H4 tetramers with histone tail deletions and the H3.3-H4 octamer with H2A/H2B. (F) Co-immunoprecipitation was performed with single H3 point mutations as described in A). (G) Peptide pulldown with biotinylated peptides of residues 80-94 or 86-97 of H3.1 and H3.3. SDS Elution of the streptavidin beads and 5% input shown. (A) contr. by P. Lewis.


A
0.2M
0.4M
0.6M
0.8M
1.0M
Input
H3.1
H3.3
Daxx
Daxx
H3.1/H4
H3.3/H4
H3
H4
B
Histone-binding domain (HBD)
C-terminal domain
Daxx (740aa)
140
183
417
488
740
EEEE/D
Δ1
Δ2
Δ3
Δ4
Δ5
C
Δ1
Δ2
Δ3
Δ4
Δ5
ASF1a
GST
H3.1
H3.3
H4
H3
D
Histone H3 (136aa)
tail
helix 1
helix 2
H3.1
H3.3
Δ28
Δ43
Δ60
Biotin
H3 (80-94)
H3 (86-97)
E
H3:H4 tetramers
H3.3 ΔN28
H3.3 ΔN43
H3.3 ΔN60
H4 ΔN20
H4 ΔN20ΔC100
-
H3.3 Octamer
pulldown
Δ3
H3
H4
inputs
H3
H2A
H2B
H4
F
H3.1
H3.3
H3.1 A31S
H3.1 S87A
H3.1 V91I
H3.3 S31A
H3.3 A87S
H3.3 I89V
H3.3 G90M
pulldown
Daxx
H3
H4
inputs
H3
H4
G
H3.1 (80-94)
H3.3 (80-94)
H3.1 (86-97)
H3.3 (86-97)
5% Input
Daxx
Streptavidin

The Glu/Asp-rich linker between N-and C-terminal domain is resembling similar acidic stretches found in many histone chaperones including NAP1, Asf1p, Rtt106, Vsp75, FACT, CAF1p150. Additionally, two hydrophobic motives at the far N-and C-termini are conserved in vertebrates (Figure 2.4A). Guided by conserved domains and secondary structure prediction, I designed a set of GST-fusion proteins that covered the N- and C-terminal domains. GST-pulldowns with the Daxx fusion proteins and recombinant H3.3-H4 or H3.1-H4 were performed (Figure 2.3C). I found that the fusion proteins containing the highly conserved region in the N-terminus of Daxx all bound specifically to H3.3-H4. I termed this region the 'histone binding domain' (HBD) of Daxx, necessary and sufficient for histone binding (Figure 2.3C). The HBD is exceptionally conserved throughout metazoans (Figure 2.4B). As a control, I found that a recombinant GST-Asf1a protein bound equally well to both H3.1-H4 and H3.3-H4 (Figure 2.3C). I further investigated the Daxx-H3.3 interaction by making a series of 'tailless' recombinant H3.3-H4 (Figure 2.3D). I found that the Daxx HBD fusion protein bound effectively to all H3.3-H4 tail deletions constructs (Figure 2.3E), indicating that Daxx contacts H3.3-H4 via the histone globular domain. Furthermore, the Daxx HBD did not bind to the H2A/H2B dimer (Figure 2.3E).

H3.1 and H3.3 differ in sequence at five residues; four residues found within the histone-fold domain and an A31S substitution in the unstructured N-terminal histone tail. Three substitutions are clustered at the base of helix 2 of the histone fold, and these residues were previously shown to confer binding specificity for particular histone deposition pathways (Ahmad and Henikoff, 2002c; Tagami et al., 2004) (Figure 2.3B).

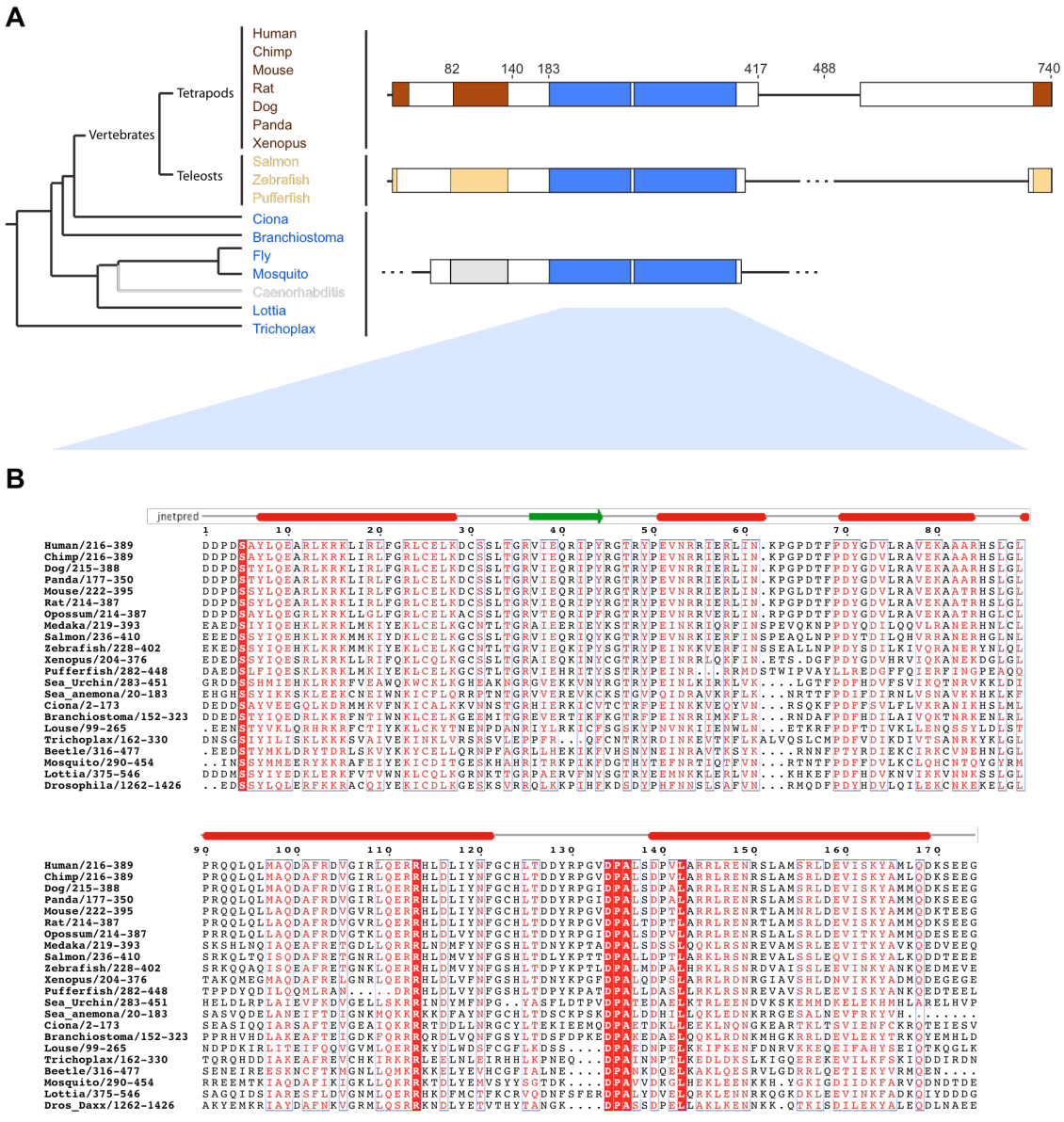



Figure 2.4: Domain structure and conservation of Daxx homologs.

(A) Domain structure of Daxx homologs were analyzed for selected metazoan species from the JGI Tree of Life (left schematic tree, not drawn to scale). The two blue central regions forming the histone binding domain (HBD) are highly conserved in all Daxx homologs. Red/orange regions are highly conserved among tetrapods and teleosts, but less conserved or missing in non-vertebrates. The *Caenorhabditis* genus does not identifiable Daxx homologs (light gray). (B) A multiple sequence alignment spanning residues 216-389 of human Daxx and corresponding regions of other species as listed show exceptionally conserved residues within histone binding domain.

In order to identify which of the five residues are responsible for differential binding by Daxx, I constructed a series of single point mutations in the unique H3.1 or H3.3 residues. Each of these mutant H3-H4 complexes was subject to co-IP with recombinant full length Daxx (Figure 2.3F). We found that no single point mutation in H3.1 conferred markedly increased binding to Daxx. Similarly, we found that no single mutation in H3.3 abrogated binding to Daxx, demonstrating that single point mutations had little effect on the overall interaction.

Mutants in position 31 had no effect on Daxx binding to H3.3, signifying that the primary mode of recognition occurs via the globular domain. Mutation of H3.3 Gly 90 to Met (the residue present in H3.1) had the largest single effect on Daxx binding. To directly test if the interaction between Daxx and H3.3 occurs via the unique '**A**A**IG**' motif at the base of H3.3 helix 2, I performed pull-downs with peptides that corresponded to the residues 80-94 and from 86-97 of either H3.1 or H3.3 (Figure 2.3D). Peptides corresponding to residues 80-94 of H3.3, but not H3.1, bound effectively, indicating that this region is necessary and sufficient for interaction with Daxx (Figure 2.3G). H3 peptides spanning residues 86-97 failed to interact with Daxx in the pull-down assay.

### 2.2.3 Daxx is a histone H3.3-specific chaperone

The purification of Daxx in a nucleoplasmic complex with H3.3-H4 indicated that Daxx might act as a histone chaperone for the deposition of newly synthesized H3.3. We saturated Daxx with excess H3.3-H4 and purified a stable stoichiometric complex over a Mono Q ion exchange column (Figure 2.5A).

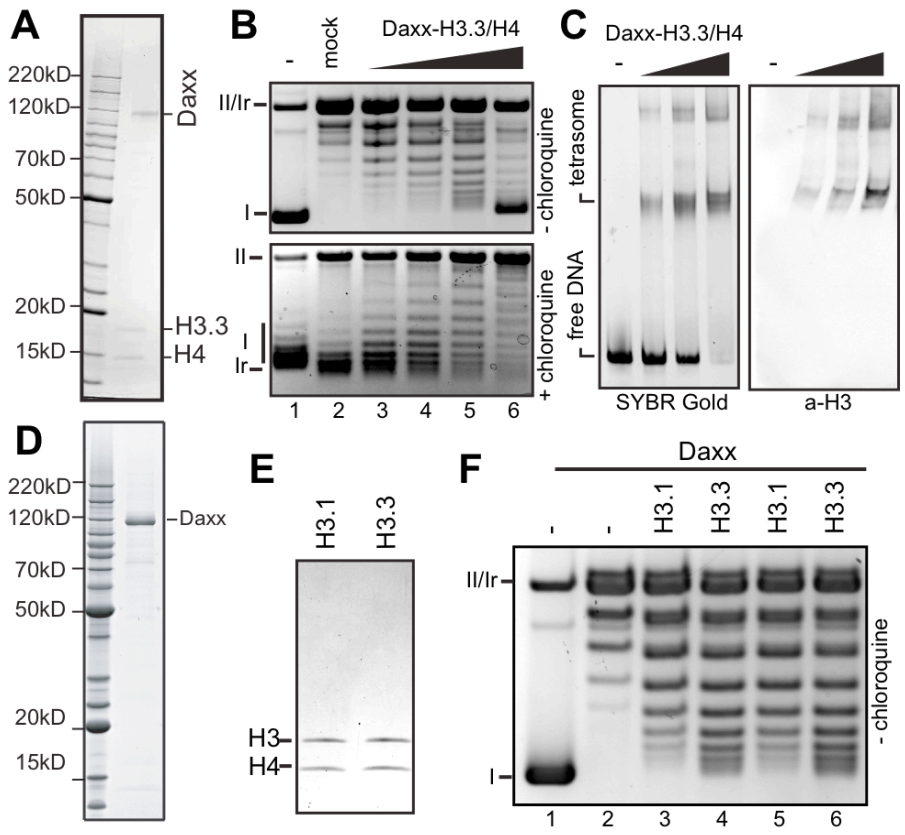



Figure 2.5: Daxx is a histone H3.3 chaperone.

(A-C) A recombinant Daxx-H3.3-H4 complex assembles chromatin *in vitro*. Histone deposition activity of an *in vitro* reconstituted and purified Daxx-H3.3-H4 complex (A) was assessed in two assays: (B) chromatin-induced supercoiling was analyzed by agarose gel electrophoresis in the absence (upper panel) and presence (lower panel) of chloroquine; 0.25-2.5 $\mu$g Daxx-H3.3-H4 was mixed with 500 ng Topo-1 relaxed plasmid DNA (lanes 3-6). Supercoiled input plasmid (lane 1) and Topo-1 relaxed plasmid (lane 2, mock) were run as controls. (C) Tetrasome deposition by the Daxx-H3-H4 complex was assayed on a 148 bp DNA fragment. The electrophoretic mobility shift was analyzed on a native TBE gel. (D-F) Daxx preferentially assembles H3.3 chromatin. 2 $\mu$g recombinant Daxx (D) or NAP-1 was preincubated with H3.1-H4 or H3.3-H4 tetramers (E) and added to 300 ng Topo-1 relaxed plasmid. The resulting plasmid supercoiling was analyzed by agarose gel electrophoresis (F). Histone tetramers or the chaperones alone did not induce supercoiling (lanes 3-5, 8) While NAP-1 induced supercoiling with 250 ng H3.1-H4 or H3.3-H4 to a similar extant (lanes 6, 7), Daxx assembled chromatin more efficiently with H3.3-H4 than with H3.1-H4 at all tested concentrations (150, 200, 250 ng; lanes 9-14). (A,D) contributed by P. Lewis.

Next, I added this complex to a relaxed plasmid template and found that it induced supercoiling in a concentration-dependent manner. The decreased mobility of the products in the presence of chloroquine indicated negative supercoiling characteristic for H3-H4 deposition onto the plasmid (Figure 2.5B). Furthermore, a 148bp DNA fragment formed a tetrasome upon addition of the Daxx-H3.3-H4 complex as indicated by a characteristic shift in electrophoretic mobility of the DNA and co-migrating H3.3 (Figure 2.5C). These results show that Daxx has intrinsic histone chaperone activity.Given the marked specificity of Daxx for H3.3 in our *in vitro* binding experiments, I wondered if its histone chaperone activity would be specific to H3.3, as well. I performed plasmid supercoiling assays as described above, but with purified Daxx (Figure 2.5D) and H3.1 or H3.3 (Figure 2.5E). While Daxx promoted some chromatin assembly in the presence of both H3.1-H4 and H3.3-H4, it proved more efficient in depositing H3.3-H4 tetramers (Figure 2.5F). In contrast, NAP-1-mediated assembly was equally efficient with H3.1-H4 (Figure 2.6). I therefore identified Daxx as a *bona fide* histone chaperone with intrinsic preference for H3.3.

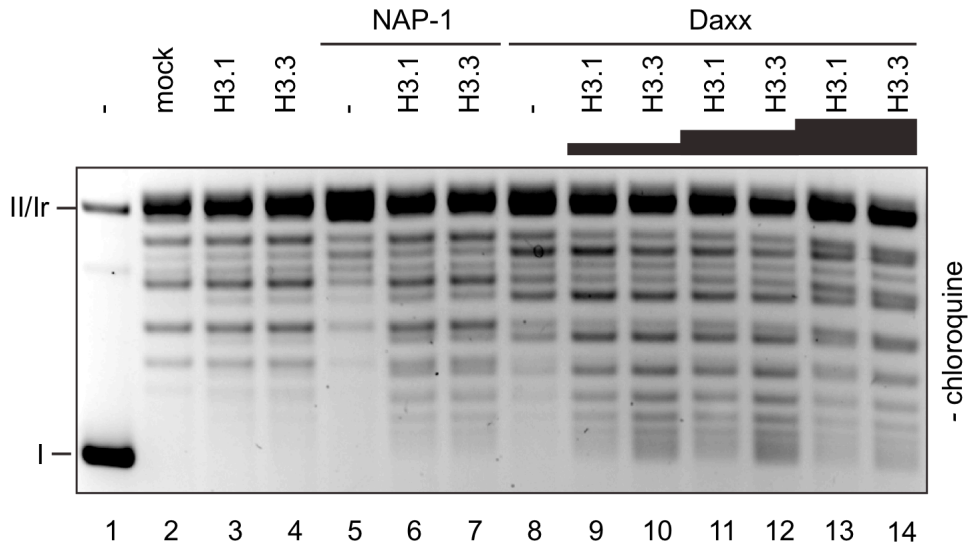



Figure 2.6: Daxx preferentially assembles H3.3 chromatin.
2 $\mu$g recombinant Daxx or NAP-1 was preincubated with H3.1-H4 or H3.3-H4 tetramers and added to 300 ng Topo-1 relaxed plasmid. The resulting plasmid supercoiling was analyzed by agarose gel electrophoresis. Histone tetramers or the chaperones alone did not induce supercoiling (lanes 3-5, 8) While NAP-1 induced supercoiling with 250 ng H3.1-H4 or H3.3-H4 to a similar extant (lanes 6, 7), Daxx assembled chromatin more efficiently with H3.3-H4 than with H3.1-H4 at all tested amounts (150, 200, 250 ng; lanes 9-14).

### 2.2.4 The ATRX-Daxx complex has chromatin remodeling and H3.3-deposition activity

ATRX localizes to telomeric chromatin, and ATRX-/- ESCs fail to incorporate H3.3 into telomeres, suggesting that ATRX plays a direct role in incorporating H3.3 into chromatin (Goldberg et al.). SNF2-family chromatin remodeling proteins such as ATRX use ATP hydrolysis to translocate nucleosomes along the DNA (Saha et al., 2006). Additionally, CHD1 and ISWI family remodelers function alone or in the context of protein complexes to

facilitate the assembly of extended, periodic nucleosome arrays (Ito et al., 1997; Lusser et al., 2005). To investigate the histone deposition properties of the ATRX-Daxx complex, we performed chromatin assembly assays with a completely purified system containing recombinant ATRX-Daxx, recombinant Daxx and recombinant H3.1-H4, H3.3-H4 and H2A/H2B and a DNA template that contained a series of tandem *Xenopus* 5S rDNA nucleosome positioning sequences. Micrococcal nuclease (MNase) digestion was used to analyze the reaction products for nuclease-protected mono- or oligonucleosomal fragments (Figure 2.7C, D). While we demonstrated above that Daxx has intrinsic histone deposition activity (Figure 2.5), this analysis revealed that the ATRX-Daxx complex catalyzed the formation of extended nucleosome arrays (Figure 2.7D).

We compared the assembly activity of the ATRX-Daxx complex to recombinant human ACF complex and human Nap1. We found that the ATRX-Daxx complex assembled H3.3-containing nucleosomes to a greater extent than H3.1 nucleosomes, whereas ACF and NAP1 assembled both H3.1 and H3.3 nucleosomes equally. ATRX assembly in the absence of Daxx could not be assessed, as we were unable to purify recombinant full-length ATRX alone. The enhanced deposition of nucleosome arrays suggested that the ATRX-Daxx complex contains remodeling activity in addition to its histone deposition activity. Previously, purified ATRX-containing complexes have been shown to have DNA-stimulated ATPase, DNA translocase and chromatin remodeling activities (Emelyanov et al.; Tang et al., 2004; Xue et al., 2003). The p185 isoform of *Drosophila* ATRX homolog (XNP) was found in a complex with HP1, and exhibited *in vitro* chromatin-remodeling activity (Emelyanov et al.).

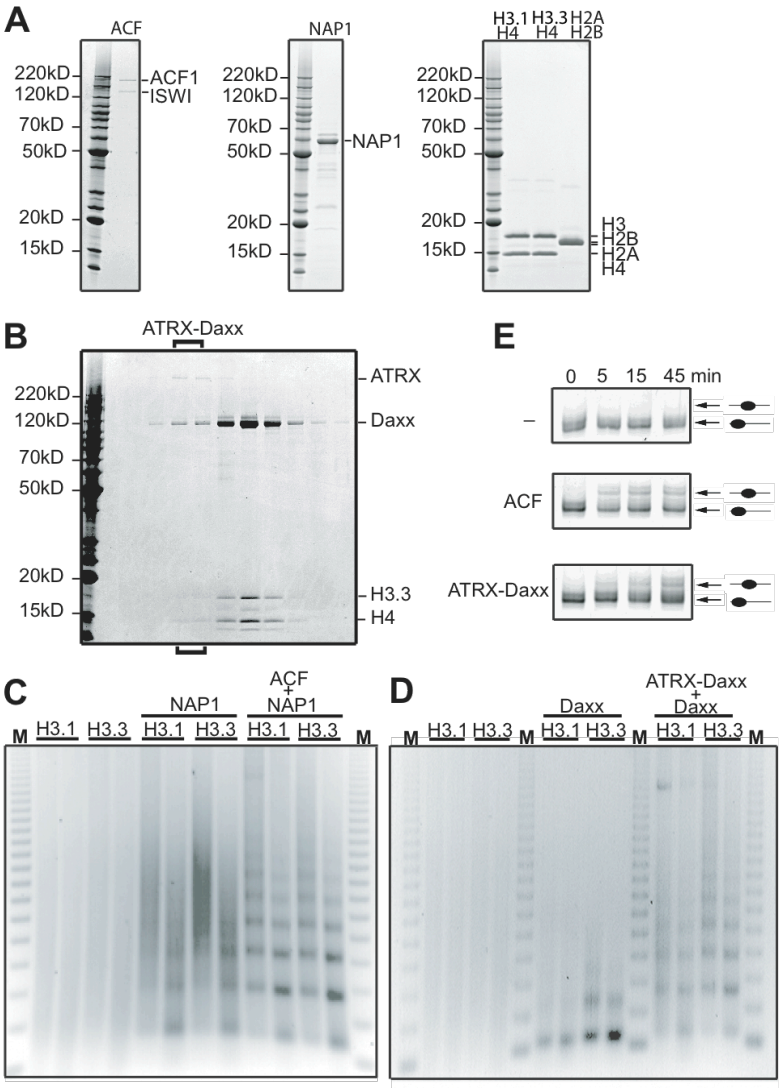



Figure 2.7: ATRX-Daxx is a histone deposition and remodeling complex.
(A) Purification of recombinant ACF complex (ACF1 and SNF2H), recombinant hNAP1, recombinant Daxx, recombinant H3.1-H4, H3.3-H4 and H2A/H2B. (B) SMART Superdex 200 gel filtration fractionation of ATRX-Daxx complex. ATRX-Daxx complex used in C)-E) was purified from free Daxx and Daxx-H3.3-H4 complex. (C) MNase digestion of chromatin assembly reactions with H3.1-H4 or H3.3-H4 alone, or in combination with NAP1 and ACF complex. A 123bp ladder is loaded in the marker lane (D) MNase digestion of chromatin assembly reactions with H3.1-H4 or H3.3-H4 alone, or in combination with Daxx and ATRX-Daxx complex. (E) Analysis of remodeling activity of ACF and ATRX-Daxx complexes on H3.3 mononucleosomes by native polyacrylamide gel electrophoresis. (B-D) contributed by P. Lewis.

We investigated whether the ATRX-Daxx complex could mobilize a positioned H3.3-containing nucleosome. Remodeling assays were carried out on a 194-bp DNA template that contained a nucleosome positioning sequence. An H3.3-histone octamer was assembled onto the DNA and incubated with ATRX-Daxx and ATP. Analysis of the remodeled substrate by native gel electrophoresis indicated that the ATRX-Daxx complex could effectively mobilize the nucleosome along the DNA template (Figure 2.7E).

### 2.2.5 Daxx is required for H3.3 deposition at telomeres

Previously, we found that ATRX-/- ESCs exhibited a dramatic loss of H3.3 found at telomeres, indicating that ATRX may have a direct role in the incorporation or maintenance of H3.3-containing nucleosomes (Goldberg et al., 2010). We sought to determine if Daxx, like ATRX, is important for H3.3 deposition at telomeres.

We have earlier described the use of zinc finger nucleases to generate murine ESC lines that carry one H3.3B allele tagged with a C-terminal HA epitope (Goldberg et al.). We targeted the H3.3B locus in 129/SvEv ESCs and Daxx -/- ESCs to generate cell lines that contained one HA-tagged allele of H3.3B. The Daxx-/- ESCs were generated from the 129/SvEv cell line (Michaelson et al., 1999). Immunoblot analysis of extract from the H3.3-HA tagged ATRX-/-, Daxx-/- and C6 ESCs indicated that levels of H3.3-HA were similar between the cell lines. Interestingly, we found that ATRX levels were reduced in Daxx -/- ESCs, suggesting that Daxx may be required for ATRX protein stability or expression (Figure 2.8A).

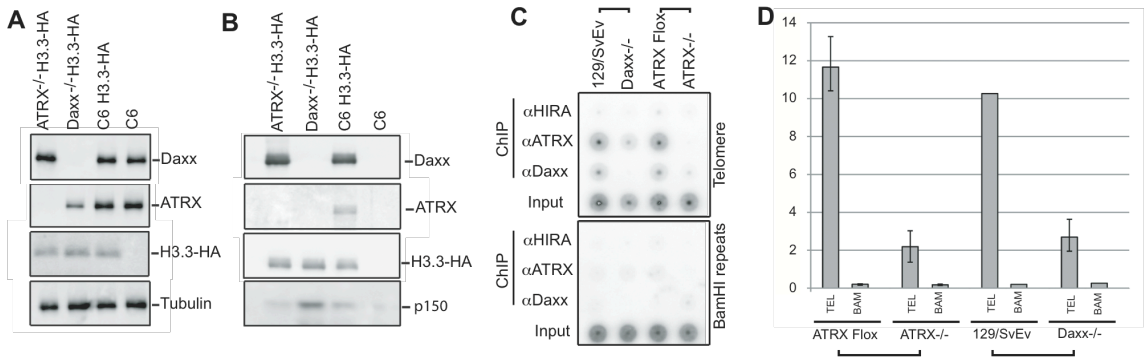



Figure 2.8: ATRX-Daxx complex is required for H3.3-deposition at telomeres.
(A) Nuclear extract was prepared from wild type C6 ESCs, and from ATRX-/-, Daxx-/- and C6 ESCs with HA-epitope tagged endogenous H3.3B. Immunoblots were performed on extracts for Daxx, ATRX and H3.3-HA. Tubulin immunoblot served as a loading control (B) HA-Immunoprecipitation from the extracts in A). Immunoprecipitated material was immunoblotted for Daxx, ATRX and H3.3-HA. (C) Chromatin immunoprecipitation of HIRA, ATRX and Daxx proteins was on 129/SvEv, Daxx-/-, ATRX-Flox, and ATRX-/- ESCs. Dot blot analysis was performed on the immunoprecipitated DNA using telomere and BamH1 repeat probes. (D) HA-H3.3 chromatin immunoprecipitation was performed on 129/SvEv, Daxx-/-, ATRX-Flox, and ATRX-/- ESCs. Bar graph showing the input-normalized signal from the dot blot on telomere and BamHI repeats with standard errors. (A,B) contributed by P. Lewis, (C,D) contributed by K.M. Noh.

---

Immunoprecipitation of H3.3-HA was performed from nuclear extract from these cell lines (Figure 2.8B). We found that H3.3-HA could efficiently co-IP Daxx in ATRX-/- cells, suggesting that ATRX is not required for Daxx interaction with H3.3. ATRX also failed to co-IP with H3.3-HA in Daxx -/- ESC. These data are consistent with our previous experiments that indicated a direct contact between Daxx and H3.3. Interestingly, the CAF1 p150 signal increased in Daxx -/- cells, suggesting that in the absence of Daxx, H3.3 may associate with other histone chaperone complexes.

My colleague Dr. Kyung Min Noh performed ChIP on H3.3-HA in the Daxx-/- ESCs to determine if Daxx, like ATRX, is required for H3.3 deposition at telomeres (Figure 2.8C). The H3.3 signal was decreased to a similar degree in both ATRX-/- and Daxx-/- ESCs. We performed ChIP with Daxx antiserum and found that Daxx was bound to telomeric chromatin in wild type murine ESCs (Figure 2.8D). Daxx failed to ChIP to telomeric chromatin in ATRX-/- ESCs, suggesting that ATRX is required for targeting Daxx and H3.3 to telomeric chromatin. We therefore conclude that the ATRX-Daxx complex localizes to telomeres and directly mediates H3.3 deposition at telomeric chromatin.

## 2.3 Discussion

### 2.3.1 Daxx is a H3.3 chaperone

We demonstrated that Daxx forms stable complexes with H3.3-H4, but not with H3.1-H4. We assessed Daxx histone chaperone activity alone and coupled to ATP-dependent remodeling by ATRX and have observed preference for H3.3 in both deposition assays. Similar H3.3-specificity has been observed in a minicircle assembly assay (Drané et al.).

I identified a minimal 234 amino acid segment of Daxx that specifically interacts with H3.3-H4. The histone-binding domain also represents the most highly conserved segment of Daxx, suggesting that the H3.3-specific binding and chaperone activity may be conserved among Daxx homologs. Besides a poly-Glu/Asp acidic stretch, the Daxx histone-binding domain lacks sequence homology with any known histone binding proteins, suggesting that the Daxx contains a novel histone chaperone domain (Figure 2.4).

I have shown that Daxx distinguishes histone H3 variants through direct interaction with the variant-specific residues 87-90 in the core histone fold of H3.3 (Figure 2.3F, G). The three unique residues in the H3.3 'AAIG' motif cooperatively confer specificity for binding as single point mutants only have a modest effect. My *in vitro* binding data closely resembles the composite specificity for replication-independent H3.3 deposition observed in *Drosophila* cells (Ahmad and Henikoff, 2002c). While specific binding was achieved with a 15-residue peptide, it is conceivable that other surfaces of H3.3-H4 contribute to the overall binding affinity. These shared regions likely account for the residual binding to H3.1-H4 observed in all *in vitro* assays. We speculate that selectivity is enhanced *in vivo* by the existence of competing chaperone complexes and deposition pathways. In agreement with such a model, we found increased levels of H3.3 association with CAF1 in Daxx-/- ESCs as compared to wild type (Figure 2.8B).

Our previous ChIP-Seq studies indicated that H3.3-nucleosomes at telomeres and transcription factor binding sites (TFBS) were deposited independently of HIRA (Goldberg et al.). While the ATRX-Daxx complex may account for the HIRA-independent H3.3 deposition at telomeres, the factors involved in H3.3 deposition at TFBS remain unknown. Our purification of H3.3-bound proteins from cell extracts yielded two biochemically distinct Daxx populations. We speculate that Daxx may be recruited to TFBS and other genomic loci for assembly of H3.3-nucleosomes.

Although Daxx has been studied extensively in many systems, its function in transcriptional regulation is not well understood. In light of our findings, the

mechanism by which Daxx influences transcription may be directly linked to H3.3 deposition activity at TFBS or other regulatory elements.

### 2.3.2 ATRX has nucleosome remodeling activity

We demonstrated that recombinant ATRX-Daxx complex has chromatin remodeling activity and can assist Daxx in the assembly of H3.3-nucleosomes. Previously, immunoprecipitated human ATRX was shown to exhibit both DNA translocase and mononucleosome disruption activity focused near the nucleosome entry/exit site (Xue et al., 2003). The deposition of H3.3-H4 by Daxx, followed by the mobilization of H3.3-nucleosomes by the ATP-dependent remodeling activity of ATRX may serve an important role in the formation of higher-order chromatin structure. Human ATRX is a homolog of the repair and recombination protein, RAD54. *In vitro* studies found that other human RAD54 homologs displayed similar remodeling activities to the ISWI-type ATP-dependent remodeler proteins near the entry and exit sites (Längst and Becker, 2001; Zhang et al., 2007b). These data are consistent with a model whereby RAD54 family remodeling proteins mobilize the histone octamer, as opposed to catalyzing the formation of DNA loops on the nucleosome surface.

### 2.3.3 ATRX-Daxx complex may assemble specialized heterochromatin

Our ChIP results indicate that the ATRX-Daxx complex is present at telomeric chromatin in murine ESCs (Figure 2.8C). Moreover, we show that Daxx localization to telomeres is dependent on ATRX, suggesting that ATRX may be involved in recruitment of the complex to telomeres. While Daxx histone chaperone activity is sufficient to assemble chromatin *in vitro*, H3.3-deposition at

telomeres is additionally dependent on ATRX (Figure 2.8D), suggesting a dual role in both complex recruitment and ATP-dependent chromatin remodeling.

In differentiated cells, H3.3 has been found to localize to pericentric chromatin (Hake et al., 2005; Wong et al., 2009). Recent work performed in murine embryonic fibroblasts (MEFs) found H3.3 deposited at pericentric DNA repeats in an ATRX-Daxx-dependent manner (Drané et al.). ATRX-Daxx also localizes to PML bodies, and ATRX has been shown to co-localize with heterochromatin on both the inactive X chromosome and Y chromosome in mice (Baumann and de La Fuente, 2009; Baumann et al., 2008). These findings raise the intriguing possibility that the ATRX-Daxx complex may serve as a specialized chromatin assembly pathway for repetitive regions such as telomeres, centromeres and other regions of constitutive heterochromatin. In agreement with this model, ATRX-depleted cells display centromere dysfunction indicated by defective sister chromatid cohesion at the metaphase plate, as well as abnormal chromosome alignment (de La Fuente et al., 2004; Ritchie et al., 2008). Cells from ATR-X syndrome patients also display altered pericentric DNA methylation (Nan et al.). Also, ESCs depleted for ATRX or H3.3 exhibited signs of telomere dysfunction (Goldberg et al.; Wong et al., 2010). *Drosophila xnp* mutants display a position-effect phenotype (PEV) and defects in pericentric heterochromatin (Bassett et al., 2008; Emelyanov et al.; Schneiderman et al., 2009).

Replication-independent histone deposition complexes help to fill in nucleosome-free regions created by RNA polymerase passage and ATP-dependent remodeling at transcribed genes and enhancer elements. Both telomeric and centromeric sequences are transcribed to produce long non-coding

RNAs (Luke and Lingner, 2009; Zaratiegui et al., 2007). Telomere repeat-containing RNAs (TERRA) appear to affect telomere structure and replication (Luke and Lingner, 2009), and centromeric RNA is important for pericentric heterochromatin formation (Maison et al., 2002). While heterochromatic regions are generally considered to have slow histone exchange rates, the intrinsic repetitive features of telomeric and centromeric DNA may promote nucleosome instability. Indeed, telomere repeats have a low propensity to form nucleosomes *in vitro* (Cacchione et al., 1997), and deposition of H3.3 outside of S-phase may help maintain a proper nucleosome density for heterochromatin formation.

### 2.3.4 Recruitment of the ATRX-Daxx complex

In addition to the chromatin remodeling SNF2-like ATPase domain, ATRX contains an N-terminal ATRX-DMNT3L-DNMT3A (ADD) domain that consists of an N-terminal GATA-like zinc finger and a PHD finger. The ADD domain of DNMT3A and DNMT3L preferentially interact with the unmodified extreme N-terminus of histone H3 (Ooi et al., 2007; Otani et al., 2009). Very recent reports now showed that the ATRX ADD domain, like the heterochromatin protein-1 (HP1) (Jacobs and Khorasanizadeh, 2002), prefers H3K9me3, leading to an attractive model of ADD-dependent recruitment or stabilization of the ATRX-Daxx complex to heterochromatin for remodeling and histone deposition (Dhayalan et al., 2011; Eustermann et al., 2011; Iwase et al., 2011). Both telomeres and pericentric regions are enriched in DNA methylation and heterochromatin histone modifications such as H3K9me3 and H4K20me3 (Luke and Lingner, 2009; Schoeftner and Blasco, 2009), and these modifications may recruit the ATRX-Daxx complex to facilitate H3.3 deposition. In support, loss of the

H3K9me3 methyltransferase, Suv39h1, had similar phenotypes as a reduction of ATRX, Daxx or H3.3: abnormal telomeres, aberrant chromosome segregation and premature sister chromatid separation (García-Cao et al., 2004; Lehnertz et al., 2003). In addition to histone modifications, DNA methylation may be a means to recruit ATRX-Daxx to specific genomic loci through an interaction with the methyl-CpG binding protein 2 (MeCP2) (Nan et al., 2007).

### 2.3.5 Functional significance of H3.3 incorporation

Viewed in the broader context, our results raise the question of whether newly incorporated H3.3 serves a specialized role beyond simply replacing the 'old' histone H3. While it has been proposed that H3.3 nucleosomes are intrinsically less stable (Jin and Felsenfeld, 2007), and promote transcription (Tamura et al., 2009), the exact role of transcription in heterochromatin still remains unclear and poorly defined. Others and we have observed mitotic phosphorylation of H3.3 S31 at ESC telomeres, as well as pericentric heterochromatin of differentiated cells (Hake et al., 2005; Wong et al., 2009). Notably, S31 is unique to H3.3 as substituted by Ala in H3.1/2 and makes this heterochromatin-associated mark strictly dependent on H3.3 incorporation. Future studies will elucidate these unexpected heterochromatic functions of histone variant H3.3, initially discovered as a marker of 'active' chromatin.

# 3 Identification of a common histone binding helix in the histone chaperones HJURP and Daxx

## 3.1 Introduction

### 3.1.1 Centromeric histone H3 variants define the eukaryotic centromeres

As introduced in Section 1.1, centromeres serve as attachment point on chromosomes for the mitotic spindle. The cenH3 variant (CENP-A in humans, CID in flies, Cse4 in yeast) is the defining protein component of functional centromeres (Earnshaw and Migeon, 1985) that by yet unknown mechanisms recruits the remaining CENP proteins (Howman et al., 2000). While the N-terminal, unstructured tail of CENP-A is quite divergent from H3, the histone fold domain is conserved with the exception of a CENP-A specific chaperone recognition domain (CRD) and preceding loop between helices α1 and α2 (Figure 3.1). Instead of the CRD region of cenH3 ,the loop and potentially other amino acid substitutions in the histone fold is not used for variant discrimination (Black et al., 2007). Like H3.3, cenH3 is incorporated independent of replication (Ahmad and Henikoff, 2002b).

The overall architecture of centromeres and the underlying centromere-defining DNA sequences have diverged considerably between metazoans and fungi.

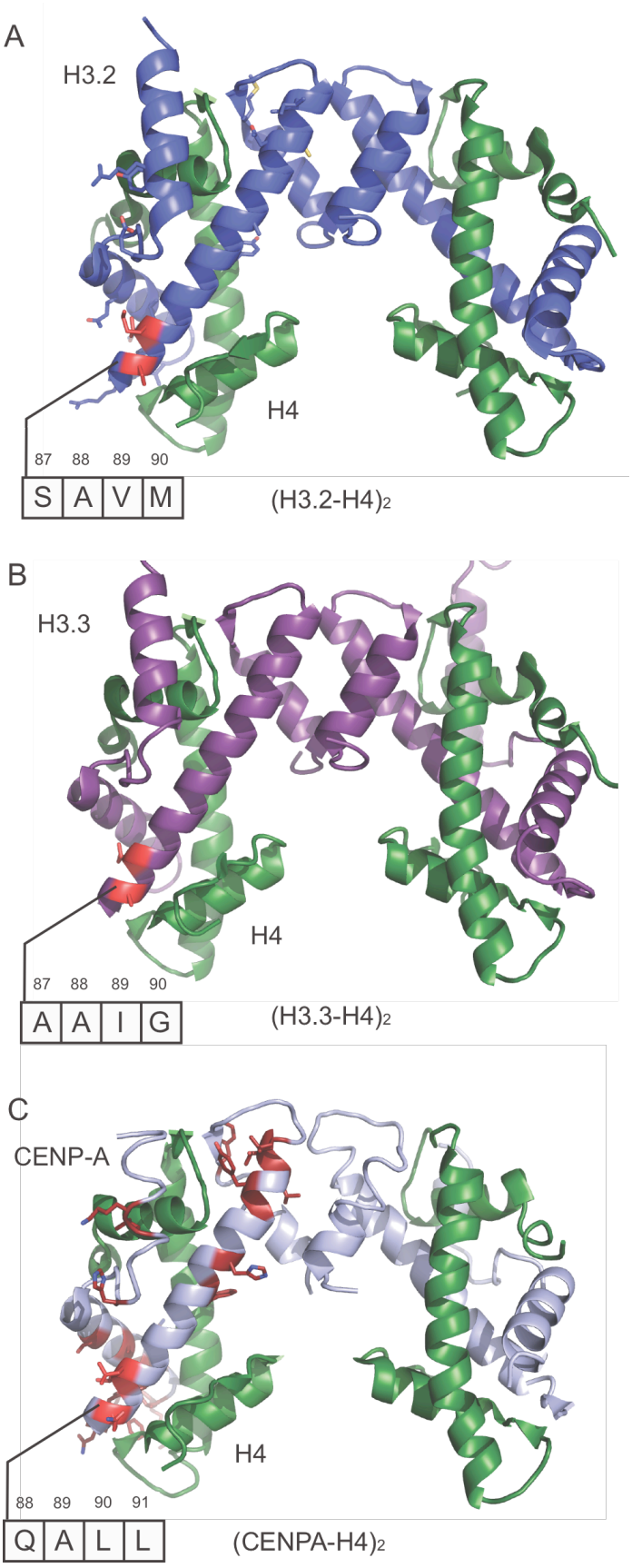



Figure 3.1: The chaperone recognition domain (CRD) distinguishes histone H3 variants.

(**A**) *Xenopus* H3.2 (blue), (**B**) human H3.3 (purple) and (**C**) human CENP-A (grayblue) tetramers with H4 (green) from respective nucleosome crystal structures (PDB 1KX5, 3AV2, 3AN2) are shown. The CRD at the base of helix α2 is highlighted in bright red with variant-specific amino acids shown in boxes. Further amino acid differences are colored in dark red.

The histone chaperones HJURP (human) and Scm3 (fungal) form predeposition complexes with cenH3-H4 (Dunleavy et al., 2009; Foltz et al., 2009; Shuaib et al., 2010) and mediate their direct deposition onto DNA (Bergmann et al., 2011; Dechassa et al., 2011; Shivaraju et al., 2011).

Interestingly, targeting HJURP to a specific genomic region is sufficient to create functional centromeres (Barnhart et al., 2011; Bergmann et al., 2011). Recent reports on the structure of metazoan and fungal cenH3 variant chaperones, HJURP and Scm3, respectively, have highlighted an evolutionarily conserved binding mode shared amongst distant eukaryotes (Cho and Harrison, 2011; Hu et al., 2011; Zhou et al., 2011). This suggests a feedback loop of CENP-A deposition by HJURP and recruitment of HJURP to CENP-A containing regions. Centromeres can be propagated independent of the presence of a centromeric DNA sequence. Therefore, establishment and inheritance of centromeres is therefore a to date unique example of DNA sequence-independent, epigenetic propagation (Black and Cleveland, 2011). Striking examples of this epigenetic phenomenon are the formation and generational inheritance of neocentromers following certain human chromosomal rearrangements (Choo, 2001).

The exact mechanism creating such a feedback loop is not yet well understood. In particular, CENP-A is not incorporated into centromeres during S-phase but in the next G1 phase, raising the question of how faithful propagation of centromers is achieved through G2/M phase (Dunleavy, 2011) and how HJURP is later recruited to the centromeres to replenish CENP-A. The structural and biochemical insight gained from the HJURP-CENPA-H4 complex helped to understand one part of the propagation mechanism – how CENP-A can be enriched at specific genomic regions by recruitment of its cognate

chaperone (Barnhart et al., 2011; Bergmann et al., 2011). These findings might now also have implications for the mechanism of Daxx-dependent H3.3 enrichment at telomeric and centromeric repeat sequences (see Chapter 2 and (Drané et al., 2010)).

Traditionally, centromeric and replacement histone chaperones have been viewed to be unrelated, each highly specialized to its cognate histone variant. However, the striking homology between both chaperone families discussed in this chapter suggests common functional principles. Through amino acid motif analysis and homology modeling, I have gathered evidence that the overall binding mode observed for the cenH3 variant is also applicable to Daxx and potentially HIRA, both chaperones specific to the histone replacement variant H3.3. This suggests a fundamental role and common evolutionary origin of a specific interaction surface of the H3-H4 dimer in replication-independent chromatin assembly, irrespective of the H3 variant subspecification. Through more detailed analysis of available crystal structures and homology models, I find that this class of histone chaperones mimics inter-histone interactions, allowing them to bind a H3-H4 dimer in its native conformation.

## 3.2 Results

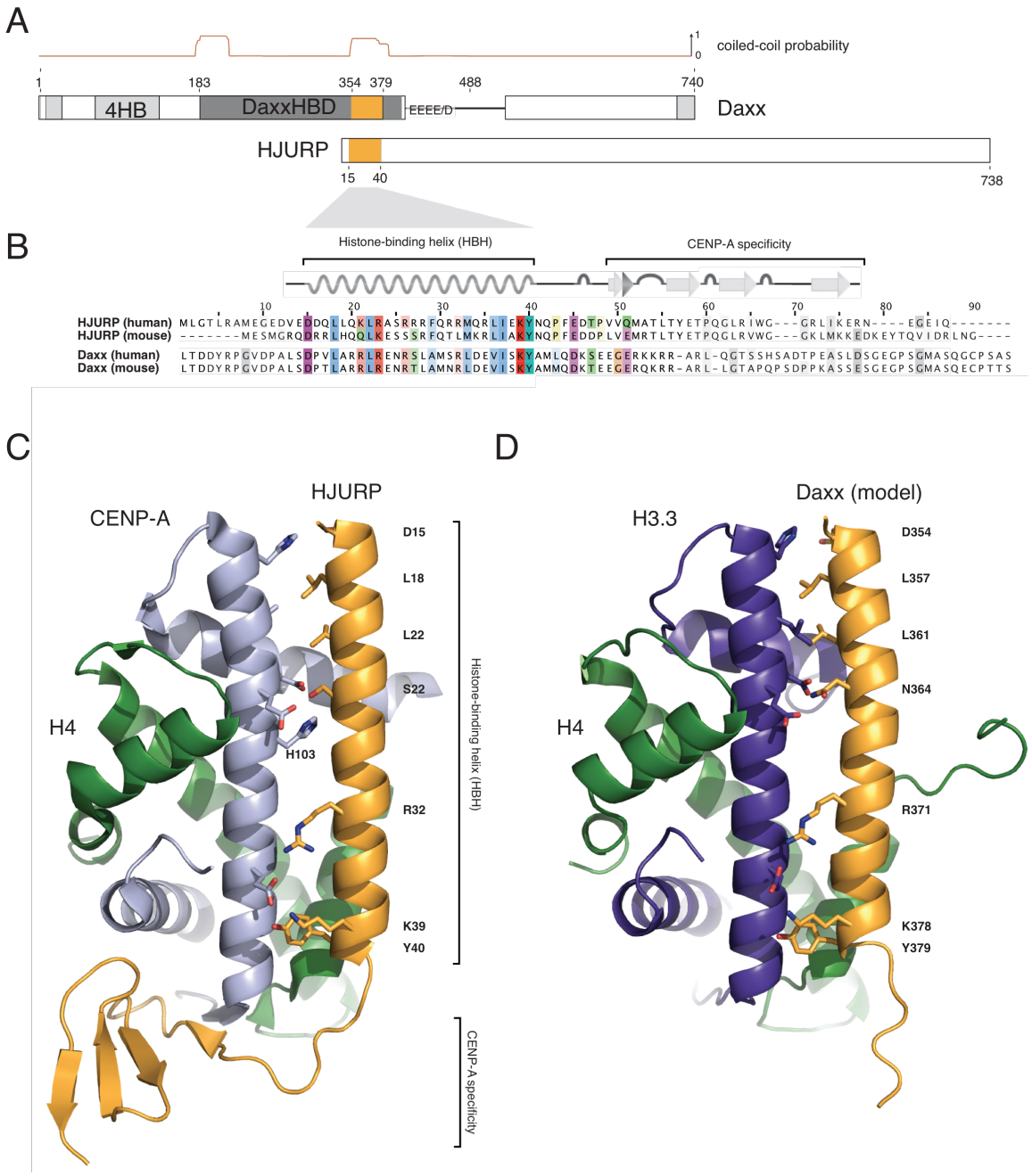
### 3.2.1 A C-terminal helix of the Daxx histone binding domain is homologous to the histone-binding helix of HJURP.

Detailed inspection of the long ⟨-helix of HJURP (histone-binding helix, HBH hereafter), which contacts the ⟨2-helix of CENP-A in the HJURP-CENPA-H4 complex structure (Hu et al., 2011), revealed patterns of evenly spaced large hydrophobic (leucine, isoleucin, valine) residues (every 3 to 4 amino acids, corresponding to one helical turn). Such pattern is characteristic of a coiled-coil structure between two adjacent helices (also called leucine zipper). Indeed, all hydrophobic residues of the HJURP HBH interdigitate with corresponding hydrophobic residues in the CENP-A ⟨2-helix, forming the characteristic leucine zipper (Hu et al., 2011). As coiled-coil interactions are predicted for at least two helices in the Daxx histone binding domain (HBD) by the sequence-based Lupas algorithm (Lupas et al., 1991), I aligned the HJURP HBH with all candidate Daxx helices and found a significant match in the most C-terminal predicted helix of the Daxx HBD (Figure 3.2A, aa 354-379). Importantly, residues involved in histone contacts are almost completely conserved between the HJURP HBH and this putative Daxx HBH, whereas residues with shifted register, pointing away from the helix interface, are quite divergent (Figure 3.2B). The β–sheet fold ensuing the HJURP HBH is not conserved in Daxx, consistent with its proposed role in conferring specificity to the CENP-A variant (Figure 3.2B, C) (Hu et al., 2011). Based on the sequence alignment, I constructed a homology model for the Daxx HBH in complex with H3.3 and H4 (Figure 3.2D). I replaced the CENPA-

H4 moiety of the HJURP-CENPA-H4 complex with a H3-H4 dimer from the nucleosome structure (PDB 1KX5) yielding a hybrid HJURP-H3-H4 structure. The subsequent conversion of HJURP to Daxx (and H3 to H3.3) was carried out *in silico* by automated amino acid mutagenesis and energy minimization by the Rosetta Backrub routine (Das and Baker, 2008; Smith and Kortemme, 2008). In the resulting model of the DaxxHBH-H3.3-H4 complex, all major contacts between the Daxx HBH and H3.3-H4 are analogous to the HJURP–CENPA interactions. The N-termini of both HJURP and Daxx HBH helices are anchored via an aspartate residue and two adjacent leucine residues to the H3 α2-helix.

Approaching the C-terminal end of the HBH, two subsequent hydrophobic residues contact H3 α2– and H4 α3–helices, followed by a salt bridge between a lysine and glutamate and a tyrosine that packs against a hydrophobic patch near the N-terminus of the H3 α2–helix. The overall binding mode of Daxx could therefore be very similar to HJURP, despite the different variants they associate with. As discussed below, other parts of Daxx and HJURP likely contribute to specificity and affinity to their respective histone.

Figure 3.2: Histone-binding helix (HBH) of HJURP and Daxx.
(A) Location of the homologous regions within the Daxx and HJURP proteins. Graph illustrates coiled-coil prediction based on amino acid sequence (Lupas et al., 1991). Gray areas designate conserved regions of Daxx including the 4-helix bundle (4HB) and the histone binding domain (HBD). Orange area designates the histone binding helices (HBH). No homology is observed outside of this region. (B) Sequence alignment of HJURP N-terminus (1-80) and corresponding residues of Daxx (340-431). Secondary structure derived from crystal structure (C) is indicated. (C) Crystal structure of the HJURP-CENPA-H4 complex (PDB 3R45), featuring residues involved in histone binding as stick models. (D) Homology model of the Daxx-H3.3-H4 complex derived from (C).


A
coiled-coil probability
1
0
1
183
354
379
488
740
4HB
DaxxHBD
EEEE/D
Daxx
HJURP
15
40
738
B
Histone-binding helix (HBH)
CENP-A specificity
10
20
30
40
50
60
70
80
90
HJURP (human) MLGTLRAMEGEDVEDDQLLQKLRASRRRFQRRMQRLIEKYNQPFEDTPVVQMATLTYETPQGLRIWG---GRLIKERN----EGEIQ-------
HJURP (mouse) -------MESMGRQDRRLHQQLKESSSRFQTLMKRLIAKYNQPFEDDPLVEMRTLTYETPQGLRVWG---GKLMKKEDKEYTQVIDRLNG----
Daxx (human) LTDDYRPGVDPALSDPVLARRLRENRSLAMSRLDEVISKYAMLQDKSEEGERKKRR-ARL-QGTSSHSADTPEASLDSGEGPSGMASQGCPSAS
Daxx (mouse) LTDDYRPGVDPALSDPTLARRLRENRTLAMNRLDEVISKYAMMQDKTEEGERQKRR-ARL-LGTAPQPSDPPKASSESGEGPSGMASQECPTTS
C
CENP-A
HJURP
H4
H103
D15
L18
L22
S22
R32
K39
Y40
Histone-binding helix (HBH)
CENP-A specificity
D
H3.3
Daxx (model)
H4
D354
L357
L361
N364
R371
K378
Y379

### 3.2.2 HJURP and Daxx histone binding helices mimick inter-histone contacts

A comparison of the modeled DaxxHBH-H3.3-H4 complex with the nucleosomal $(H3\text{-}H4)_2$ tetramer suggests that the base of the HBH occupies the same position as the C-terminal α3-helix of the second H3 in the tetramer in the H3-H3' homodimerization interface (a 4-helix bundle). This explains why the binding of Daxx or HJURP competes with tetramer formation (Figure 4.7) (Hu et al., 2010)). Surprisingly, Daxx and HJURP HBH not only occupy the same space, but also contact the H3 α2-helix with a set of residues identical to the H3' α3-helix in H3-H3 4-helix bundle. Indeed, sequence alignment reveals a shared motif between the H3 α3-helix, HJURP and Daxx (Figure 3.3A, B). Another striking feature of the HBH is the tyrosine at the far C-terminal end packing against the H3 α2-helix. This aromatic residue has a functional counterpart in the H4 C-terminus that, at least in the context of the nucleosome, folds back onto H3 in a similar manner (Figure 3.3B, bottom right). It is surprising how closely both histone chaperones have adopted existing motives to interact with histones. As discussed in Section 1.3.1, most residues of the core histone fold are under high evolutionary pressure as judged from the high identity of primary sequence even between remote species and histone variants. Histone chaperone mechanism therefore might have evolved on a mostly fixed histone scaffold, limiting the potential for variation and skewing towards reutilizing existing structural themes. A counter-example, however, is the histone chaperone Asf1, which also interacts with the 4-helix bundle interface of the H3-H4 dimer, but with a unique β-sheet fold (Adkins et al., 2007; Natsume et al., 2007).

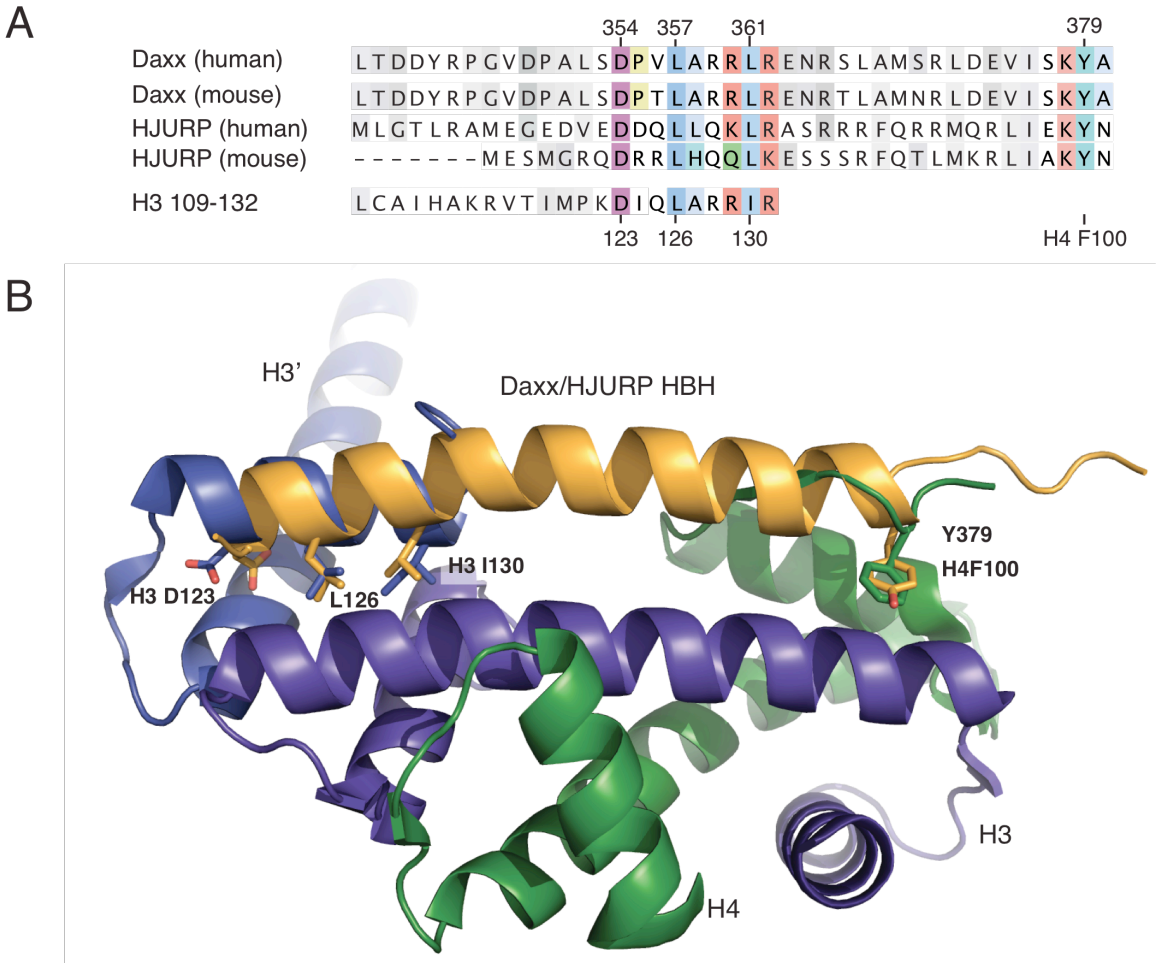



Figure 3.3: HJURP and Daxx mimick inter-histone interactions.

(**A**) Alignment of Daxx and HJURP with C-terminus of H3 (109-132). (**B**) Overlay of the H3-H3′ inter-histone contacts in the $(H3\text{-}H4)_2$ tetramer (PDB 1KX5) and the Daxx HBH model. Polar and hydrophobic residues have similar arrangements (H3 blue, Daxx HBH orange), the HBH therefore substitutes for the second H3-H4 dimer. Furthermore, Daxx HBH Y379 occupies a hydrophobic patch analogous to F100 of the H4 C-terminus (green).

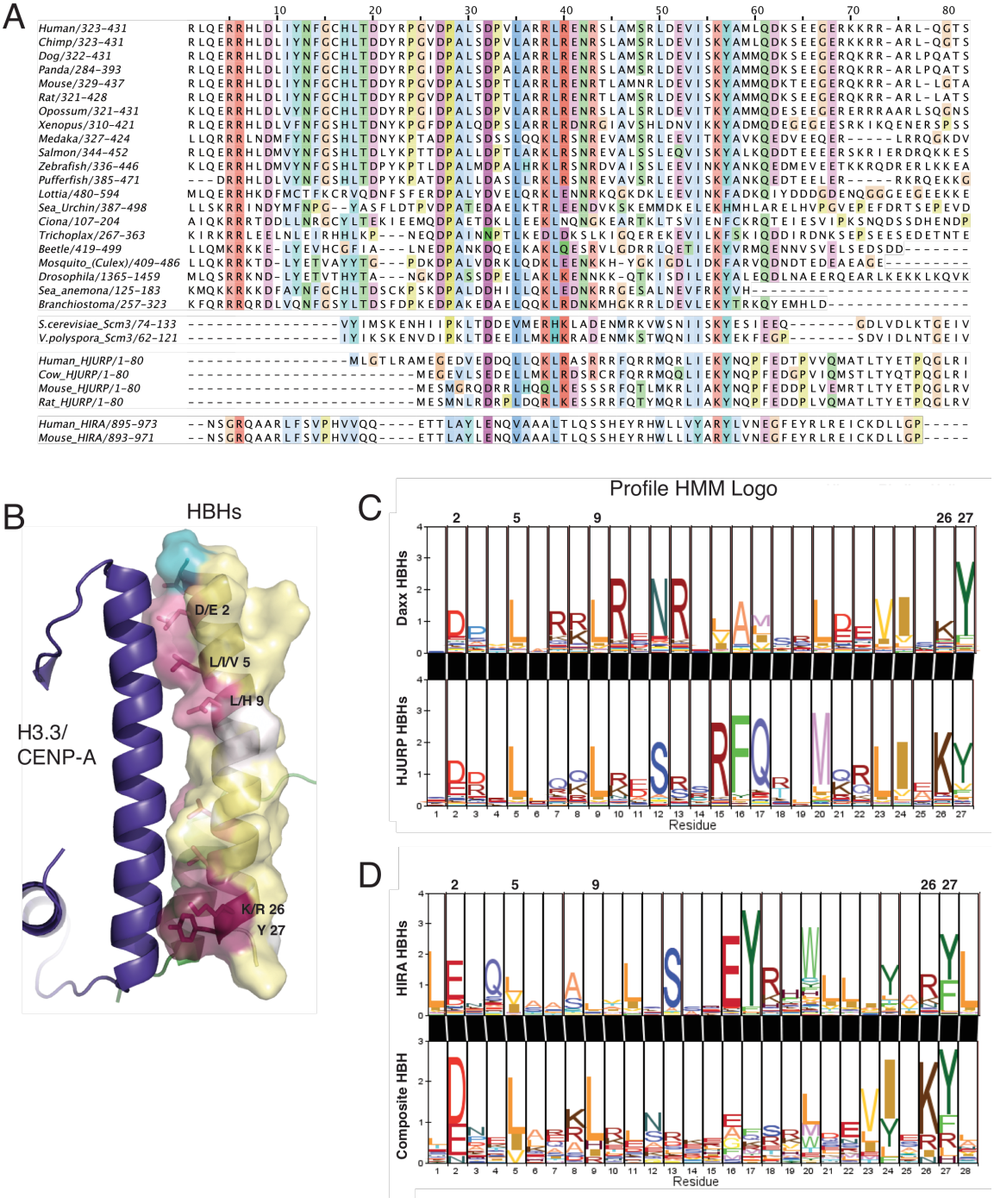
### 3.2.3 Widespread conservation of the histone binding helix motif in centromeric and H3.3-specific histone chaperones in metazoans and fungi

Despite widespread appearance in metazoans, no Daxx orthologs have been identified in lower eukaryotes, fungi or plants (Lewis et al., 2010). HJURP is found in many metazoan lineages but surprisingly it is missing from some well-studied organisms, such as *C. elegans* and *D. melanogaster*. However, a functional similar protein Scm3 is found in fungi (Mizuguchi et al., 2007). Homology between HJURP and Scm3 is confined to a limited number of motives that includes a highly conserved HBH (Sanchez-Pulido et al., 2009). Crystal and NMR structures of Scm3HBH-H3-H4 complexes show overall similar complex architecture despite somewhat divergent amino acid sequences (Cho and Harrison, 2011; Hu et al., 2011; Zhou et al., 2011). Considering that distant Daxx orthologs in fungi might be more similar to Scm3, I extended my HBH alignment to fungal Scm3 species. Interestingly, the 'middle' region of the HBH is quite diverse between Daxx, HJURP and Scm3, while the previously discussed hydrophobic side chains that make direct histone contacts are mostly identical in all available sequences (Figure 3.4A). The color-coded HBH model further illustrates that only residues with direct histone contacts are conserved (Figure 3.4B).

Traditional sequence search engines (such as BLAST) perform poorly with this apparent position-specific conservation, which might have left analogous HBH motives in other histone chaperones unrecognized. I therefore turned to a hidden markov model (HMM)-based homology search that takes position-

specific weights into account (Eddy, 1998). HMMs constructed from Daxx HBH alignments returned HJURP and Scm3 proteins and vice versa with high significance scores, as expected from the highly similar position scores (Figure 3.4B). A search with a composite HMM from all known HBH-containing proteins returned the H3.3-specific histone chaperone HIRA in addition to the input set, albeit with low significance.

Figure 3.4: Bioinformatic search for more distant HBH-containing proteins.
(**A**) Alignment of a selection of HBHs from Daxx (top), HJURP and Scm3 proteins, as well as the novel candidate HIRA. (**B**) A HBH model colored with conservation scores (purple=identity, yellow/white=intermediate, green=highly variable) from Consurf (Ashkenazy et al., 2010) illustrates that only histone-interacting residues are conserved across Daxx, HJURP and Scm3. (**C**) HMM for individual alignments of all Daxx and HJURP proteins visualized with LogoMat-P (Schuster-Böckler et al., 2004). Width of the column represents weight of the position, height of each amino acid its individual weight. Amino acids are color-coded by chemical property groups. (**D**) Comparison of the composite HMM of all known Daxx, HJURP, Scm3 proteins, and the putative HBH in the HIRAB domain of metazoan HIRA.


A
B
HBHs
D/E 2
L/I/V 5
L/H 9
H3.3/
CENP-A
K/R 26
Y 27
C
Profile HMM Logo
Daxx HBHs
HJURP HBHs
Residue
D
HIRA HBHs
Composite HBH
Residue

A HMM based on an alignment of corresponding HIRA fragments (mapping into the C-terminal HIRAB domain) did not return Daxx, HJURP or Scm3, but comparison of position scores shows common patterns, such as the E/D-L-L zipper and the N-terminal K/R-Y (Figure 3.4C). While attractive for its common biology with Daxx, bioinformatic inference alone does not warrant the conclusions that HIRA contains a bona fide HBH. A molecular model can be derived from the HJURP-CENPA-H4 structure analogous to the in silico procedure described for Daxx, however its accuracy is difficult to judge due to the large number of mutations that need to be introduced. I am therefore in the process of experimentally testing this hypothesis.

### 3.2.4 Biochemical and mutational analysis of the HBH-histone H3 interface

The detailed bioinformatic analysis of the Daxx HBD suggested that at least part of its binding to histones is mediated through an HJURP-like HBH. To test this hypothesis, I have designed experiments to characterize the putative Daxx HBH *in vitro* and conduct mutagenesis, coimmunoprecipitation and crosslinking studies of exogenous H3.3 and Daxx in the human 293T cell line. These studies are ongoing and I will discuss preliminary *in vitro* and *in vivo* results in this section.

To test if Daxx or HIRA bind via the predicted coiled-coil with the α2-helix of H3.3, I introduced point mutants in FLAG-tagged H3.3, transiently expressed in 293T cells. Based on the DaxxHBH-H3.3-H4 model, several classes of point mutations might disrupt the coiled-coil interface (Figure 3.5A): (i) where the backbones of both helices are predicted to be in close proximity, introduction

of bulky, polar amino acids in place of small hydrophobic residues might disrupt binding significantly. To this end, I generated H3.3 A91N and G102Q. (ii) The leucine zipper motive relies on interdigitating large, uncharged residues. I either mutated them to alanine (L109A H113A, data not shown), or introduced charge (L109K). (iii) Flanking the hydrophobic core, several salt bridges and hydrogen bonds seem to 'staple' the two helices together, adding additional affinity. I reversed the charge of the hydrogen-bonded E94 (E94K) or neutralized two subsequent negative charges at positions 105 and 106, (ED105QN).

FLAG-immunoprecipitation of transfected cells yielded H3-H4 and associated proteins (Figure 3.5B, upper panel). Some of the H3.3 mutants and associated H4 consistently shifted toward higher molecular weights. However, the chemical nature of this shift has not been identified. Endogenous copurifying Daxx or HIRA were detected by western blot. While insertion of bulky hydrophobic amino acids into the putative coiled-coil interface (A91N and G102Q) did not disrupt (and in the case of A91N slightly enhance) binding of Daxx, mutants relying on the neutralization or introduction of repulsive charge (ED105QN and E94K, L109K, respectively) markedly decreased association with Daxx. These mutants are properly folded as judged by quantitative association with H4 (Figure 3.5B, upper panel). As they are expressed at wild type level and Daxx protein levels are unchanged (Figure 3.5B, bottom panel), I conclude that these residues specifically abrogate association with Daxx. Although I gathered *in vitro* evidence that the recombinant Daxx HBH directly binds to H3-H4 (Figure 3.5C), more indirect effects on the Daxx interaction evoked by this mutation *in vivo* cannot be ruled out.

Notably, mutation of the H3.3 CRD (87-AAIG-90) to the respective H3.2 residues (SAVM) completely abrogate Daxx binding. This is consistent with previous results (Drané et al., 2010; Goldberg et al., 2010; Lewis et al., 2010). The terminal tyrosine of the Daxx HBH, Y379, is predicted to contact H3.3 A91. However, this interaction does not appear to be affected by the identity of the adjacent amino acids in the model (Figure 3.5C).

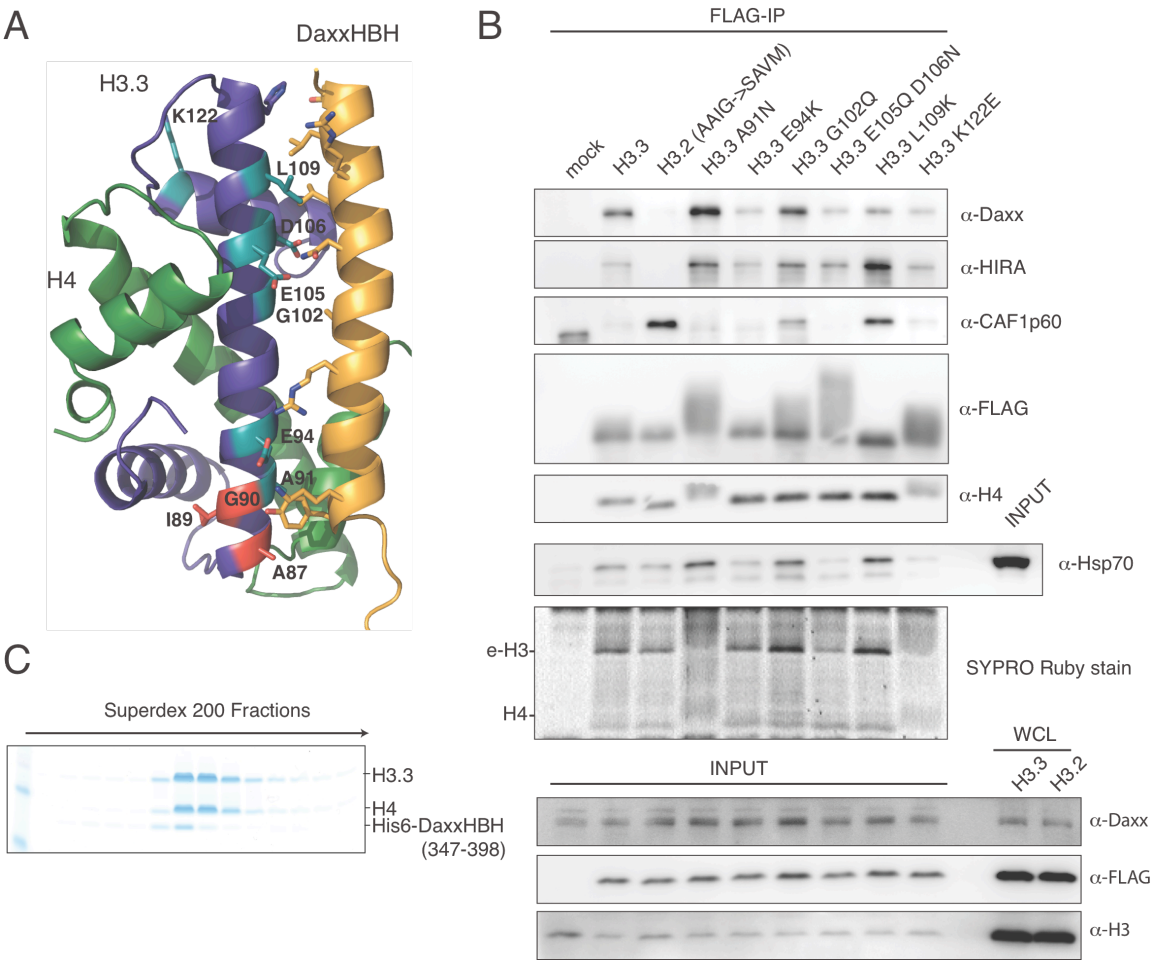



Figure 3.5: H3.3 Mutagenesis to map Daxx and HIRA binding.
(A) Model of the DaxxHBH-H3.3-H4 complex (from Figure 3.2). Mutated residues in H3.3 along the HBH shown in cyan, H3.3 CRD residues (AAIG) shown in red. (B) Coimmunoprecipitation of H3.3 mutants from transiently transfected 293T cells (upper panel), Pre-IP soluble extracts and whole-cell lysates (bottom panel). (C) Size exclusion chromatography of a recombinant, *in vitro* refolded H3.3-H4-DaxxHBH complex.

Therefore, the model fails to explain why the CRD swapping has such a large effect on Daxx binding. CRD recognition must be attributed to other parts of the Daxx HBD as discussed in Section 3.3.1. H3 K120 is surface-exposed in the H3-H4 dimer and involved in the binding of the H3-H4 dimer chaperone Asf1 (Adkins et al., 2007; Natsume et al., 2007). K120E decreases binding of Daxx despite its somewhat remote location from the HBH. This region, however, could also be contacted by additional residues of the Daxx HBD, in particular the ones immediately preceding the HBH.

Broadly, HIRA binding is affected by the mutations in H3.3 in a similar manner as Daxx (Figure 3.5B): H3.3 A91N and G102Q still interact with HIRA, whereas E94K and K120E mutants associate with HIRA to a lesser extent. Curiously, wild type H3.3 associates with HIRA relatively weakly, potentially because Daxx has a considerably higher affinity for wild-type H3.3 (see Figure 2.1). Given the potential competition of HIRA and Daxx for newly synthesized H3.3, it is difficult to dissect the effects of point mutations that decrease Daxx association simultaneously. I therefore plan to introduce reciprocal mutations for H3.3 E94K in Daxx or HIRA by swapping the opposing charged residues that form the salt bridge. The double mutation (E94 to K in H3.3, K102 to E in Daxx/HIRA) should rescue at least some of the binding.

One point mutation, H3.3 L109K shows preferred association with HIRA over wild-type H3.3. Whereas the Daxx HBH positions an arginine (R360) in proximity of H3.3 L109, HIRA does not have a basic residue at the corresponding position (Figure 3.4A) and might therefore tolerate the H3.3 L109K mutation.

One remaining question is whether other histone chaperones compensate when interaction with Daxx or HIRA is specifically impaired. I have therefore

assessed association of above mutants. With the CAF1 complex. As expected, CAF1 subunit p60 does not associate with wild-type H3.3 and instead uniquely binds H3.2 (Figure 3.5B), suggesting that the H3.2 CRD (SAVM) is necessary for CAF1 binding. However, H3.3 L109K and to a lesser extent H3.3 G102Q interact with CAF1, as well. It remains to be determined if these mutations stimulate CAF1 interaction directly. Interestingly, H3.3 E94K, which loses significant association with both Daxx and HIRA, does not interact with CAF1, either. Further candidates for buffering predeposition histones such as Asf1, NASP or NPM will need to be considered in future experiments.

In summary, Daxx and HIRA are sensitive to H3.3 mutations in the interaction surface derived from sequence homology to the CENPA-HJURP interaction. This finding corroborates bioinformatic evidence that three major replication-independent histone chaperones for H3.3 and CENP-A, respectively, use a similar binding mode and are functionally related.

## 3.3 Discussion

### 3.3.1 Additional affinity and histone variant specificity is conferred by other parts of the Daxx and HJURP histone chaperones

It is self-evident that faithful discrimination of histone variants by their cognate chaperones is intimately tied to their relevance in crucial biological processes, such as the control of chromosome segregation in the case of CENP-A. It can easily be imagined how misincorporation of CENP-A outside of centromeres could interfere with proper kinetochore formation. Indeed, overexpression of CENP-A prompts its chromosome-wide localization and

subsequent chromosome segregation defects (Heun et al., 2006; Van Hooser et al., 2001). Deletion of Scm3 in *S. cerevisiae* leads to cell cycle arrest in metaphase (Camahort et al., 2007) and levels of cenH3 are tightly regulated by ubiquitin-mediated proteolysis (Hewawasam et al., 2010; Ranjitkar et al., 2010).

It is therefore somewhat puzzling and even counter-intuitive to observe such striking homologies between a centromeric and at least one unrelated RI histone chaperone. High-resolution ChIP studies even corroborate the notion that misincorporation of CENP-A relies on RI deposition pathways cognate to H3.3 in yeast (Lefrançois et al., 2009) and flies (unpublished data, S. Henikoff, AECOM Epigenomics Symposium, 2010). It would seem sensible that histone variants or their chaperones would have diverged far enough to achieve exquisite specificity if not forced to retain some common motives due to rigid constrains imposed by nucleosome structure and assembly mechanism. It remains elusive if the common HBH stems from a single ancestral histone chaperone or if it evolved convergently. In either case, we can conclude that CENP-A and H3.3 histone chaperones have a common central HBH for the purpose of chaperone activity and evolved variant specificity in separate domains.

To date, we only have insight into this histone H3 variant discrimination from the HJURP-CENPA-H4 cocrystal structure (Hu et al., 2011): In this case, the HJURP HBH is followed by a β-sheet domain that contacts the CENP-A α1-helix (Figure 3.2C). Mutating the CENP-A specific Gln68 to Ser (present at this position in all other H3s) abrogates binding of the entire HJURP fragment under moderate *in vitro* conditions (300 mM NaCl). Similarly, partial deletion of the

HJURP β-sheet domain disrupts histone binding (Shuaib, 2010). This suggests that, despite the extended hydrophobic surface, the overall affinity of the HBH to histone-chaperone interactions is limited. This might be explained with the fact that the HBH directly competes with inter-histone contacts (as discussed in Section 3.2.2), which might dominate at high histone concentrations. Nevertheless, complexes with close to minimal Scm3 HBH have been amenable for both solution NMR and crystal structure determination (Cho and Harrison, 2011; Zhou et al., 2011).

Similar 'division of labor' might hold true for the Daxx HBD. While I observed some association of the Daxx HBH with H3.3-H4 on a size exclusion column (Figure 3.5C), I have not been able to purify a biochemically stable complex with this fragment. This is in stark contrast to the full Daxx HBD, which forms an exquisitely stable complex with H3.3-H4 over a wider range of salt concentration (see Chapter 4). Which regions of the Daxx HBD contribute this additional affinity and is this additional binding sensitive to the H3.3 CRD?

Given the location of the H3.3 CRD ('AAIG') at the base of the ⟨2-helix, residues immediately following the Daxx HBH (aa 380-400) could make contacts to this motive (compare Figure 3.5A and Figure 3.2). However, the Daxx HBH itself is located at the C-terminal extreme of the highly conserved Daxx HBD (aa 180-400) and more C-terminal regions are poorly conserved and predicted to be unstructured (Figure 3.2). The HBH could therefore alternatively be part of a larger folded domain with additional histone contacts in the N-terminal part of unknown structure (aa 180-350). In either case, only a molecular structure could resolve the long-standing question of how specificity for H3.3 is achieved.

# 4 Biochemical characterization and crystallization trials of the DaxxHBD-H3.3-H4 complex

## 4.1 Introduction

Human Daxx is a 740 amino acid protein that divides into a structured N-terminal ~400 amino acid domain, an acidic disordered linker region and a C-terminal domain scarcely defined by predictable secondary structure (Figure 4.1). The extreme N- and C-termini each contain a highly conserved SUMO-interacting motive, that might help to recruit it to specific subnuclear localizations (Lin et al., 2006; Santiago et al., 2009). Close to the N-terminus, a 4-helix bundle is formed by amino acids 59-129 as determined by solution NMR (Escobar-Cabrera et al., 2010).

Since discovering that the highly conserved histone binding domain (HBD) of Daxx is necessary and sufficient for interaction with histone H3.3-H4 (see section 0), I have focused my research on understanding its role in deposition of histone H3.3-H4 as well as discriminating against other histone variants. The approximately 234 amino acid HBD is small enough to be recombinantely expressed in *E. coli*, allowing purification of milligram quantities for biochemical and structural studies. Interestingly, the HBD carries an overall positive charge (pI 9.2) that is opposed by a highly acidic, Glu/Asp-rich stretch of ~100 amino acids C-terminal of the HBD (Figure 4.1). Within the HBD boundaries, evolutionary conservation correlates well with predicted secondary structure content, a total of 8 ⟨-helices connected by short linker sequences (Figure 4.1). As described below, domain boundaries were also confirmed and

refined experimentally. The central questions guiding my biochemical characterization were how this domain interacts with histones, how its chromatin assembly activity can be explained biochemically and how variant-specificity is integrated into the chaperoning mechanism.

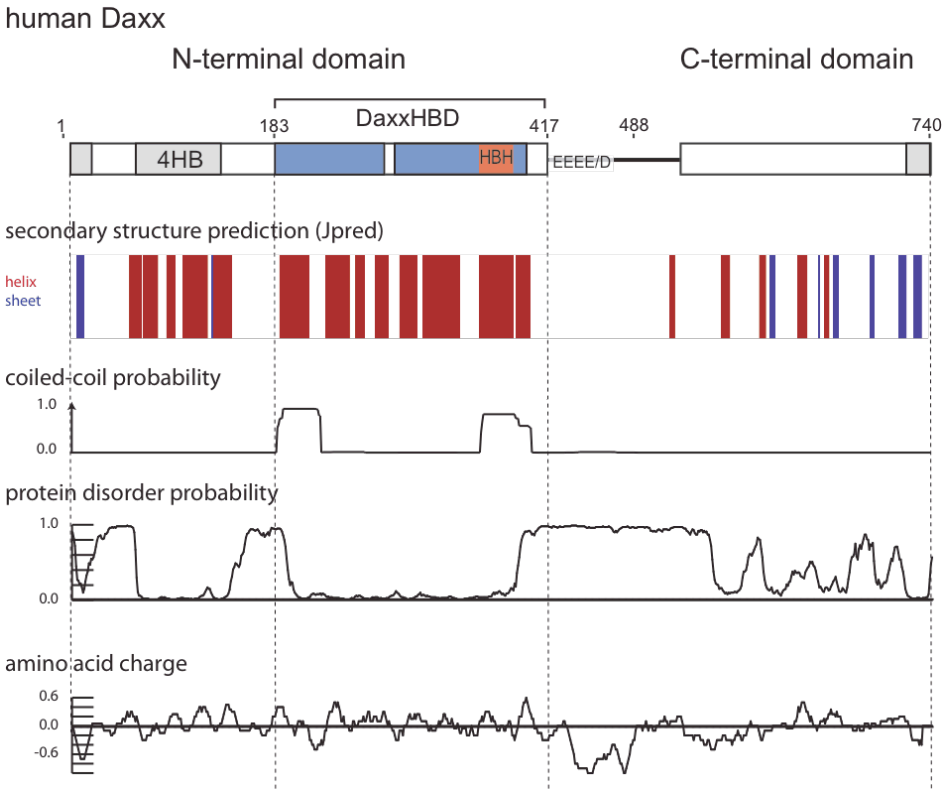



Figure 4.1: Domain structure of human Daxx.
Domain boundaries are based on evolutionary conservation, primary sequence analysis and experimental evidence. Secondary structure, coiled-coil, disorder predictions from JPred (Cole et al., 2008), are shown, as well as a moving window charge average.

## 4.2 Results

### 4.2.1 Coexpression of Daxx and H3.1/3-H4 in *E. coli* delineates a stable proteolytic fragment that specifically binds H3.3.

The first evidence for a direct and variant-specific interaction of an N-terminal subdomain of Daxx with H3.3-H4 came from an initial coexpression trial (Figure 4.2); I cloned full-length, N-terminally His-tagged Daxx into a bacterial expression vector with a second, bicistronic ORF expressing H3.3 and H4. Daxx was heavily proteolyzed in *E. coli*, and Ni-affinity purification from lysates yielded a heterogeneous mix of His-tagged Daxx fragments - therefore proteolyzed from the C-terminus (Figure 4.2A). Histones were predominantly detected when H3.3 and not H3.1 was coexpressed. Subsequent removal of aggregated protein and purification on a cation exchange column yielded three major Daxx species. Importantly, the two large N-terminal Daxx fragments elute at high salt concentration (~1 M NaCl) with H3.3, explained by the high pI of the observed complex (Figure 4.2B), whereas no interaction with H3.1 was observed on the cation exchange column as evident from the broad elution profile (Figure 4.2C). The apparent size of the largest Daxx fragment detected via its N-terminal His-tag fit well with the approximate 81 kDa full-length protein (Figure 4.2B). The smaller fragment coeluting with H3.3 was estimated to 48 kDa, suggesting a proteolysis site ~420 amino acids from the His-tagged N-terminus. Tryptic digestion and tandem mass spectrometric analysis of this band further refined the C-terminus to residue 417. A smaller N-terminal fragment of about 20 kDa eluted early on the Mono S column, suggesting that the N-terminal 150 amino

acids do not bind histones. Design of subsequent Daxx fragments was therefore centered on the region between residues 150 and 417.

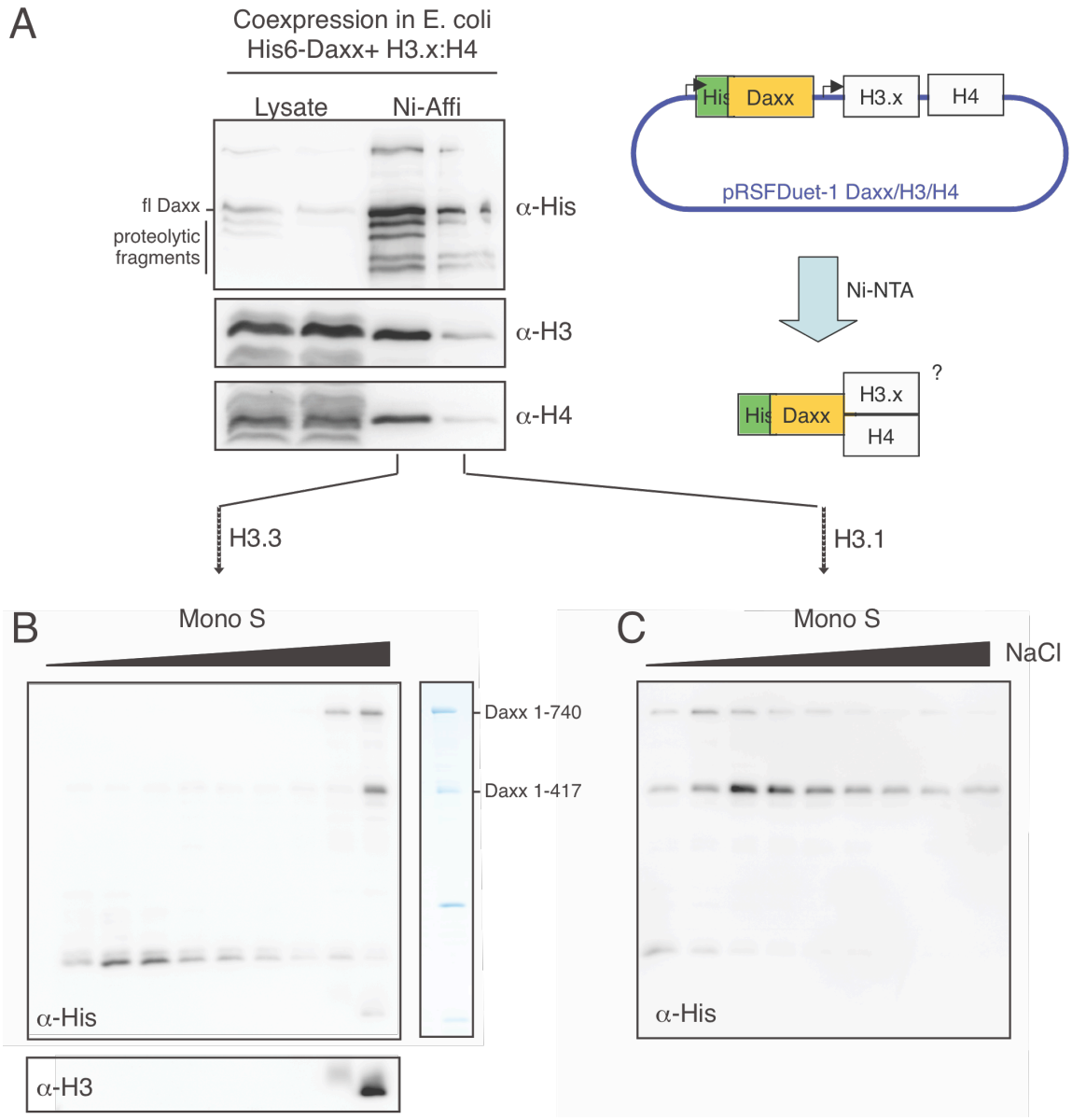



Figure 4.2: Coexpression of full-length Daxx and H3-H4 in *E. coli*.
(A) Coexpression of Daxx-H3.3-H4 and Daxx-H3.1-H4 from a single plasmid (right), and Ni-affinity purification via an N-terminal His-tag. (B) Cation exchange chromatography of eluted Daxx coexpressed with H3.3-H4 and (C) H3.1-H4. Fractions from a ~250 mM – 1 M NaCl gradient are shown. The Coomassie-stained lane in (B) corresponds to the right-most fraction shown in the western blot.

### 4.2.2 *In vitro* refolding of DaxxHBD-H3.3-H4 yields large quantities of homogenous, pure and stoichiometric complex.

In search of a reliable and robust method to express the DaxxHBD-H3.3-H4 complex, I evaluated several methods to purify either Daxx HBD alone or coexpressed with histones:

(i) His-tagged expression as exemplified in the previous section had generally low yield and purity. In the absence of histones, any constructs that included the Daxx HBD or parts of it were completely insoluble, even though the N-terminal 4-helix bundle and the C-terminal domain were soluble when expressed separately. Coexpression of histones improved solubility, but yields of DaxxHBD-H3.3-H4 were small, likely due to unbalanced expression of the three components.

(ii) GST-tagged Daxx HBD could be expressed solubly at low temperature, albeit with low yield, and could be partially purified. However the fusion protein aggregated upon further purification and, after cleavage of the GST-tag, the Daxx HBD precipitated quantitatively. Furthermore, the *E. coli* folding chaperone DnaJ was copurifying to significant amounts with GST-DaxxHBD. While including ATP in the wash buffer efficiently removed DnaJ during glutathione (GSH) affinity purification, its presence suggested that the Daxx HBD might be partially unfolded in the fusion protein. Only saturating amounts of histone H3.3-H4 stabilized the GST-DaxxHBD fusion protein, allowing cleavage of the tag and subsequent purification of a stoichiometric complex of DaxxHBD-H3.3-H4. At least when recombinantly produced in *E. coli*, the Daxx HBD therefore is partially unfolded but can form a native complex upon H3.3-H4 addition. This method

allowed me to test a number of Daxx constructs containing the HBD for their biochemical stability (Figure 4.3) and to conduct initial crystal trials (see also Table 4.1). I included constructs that were shorter than the originally defined HBD (Figure 4.2). However they did not bind H3-H4 with similar affinity as the full HBD (Figure 4.3B) In general, this method yielded only limited amounts of protein complex for crystal screening due to the low soluble expression of the GST-Daxx HBD and a large amount of impurities (Figure 4.3B).

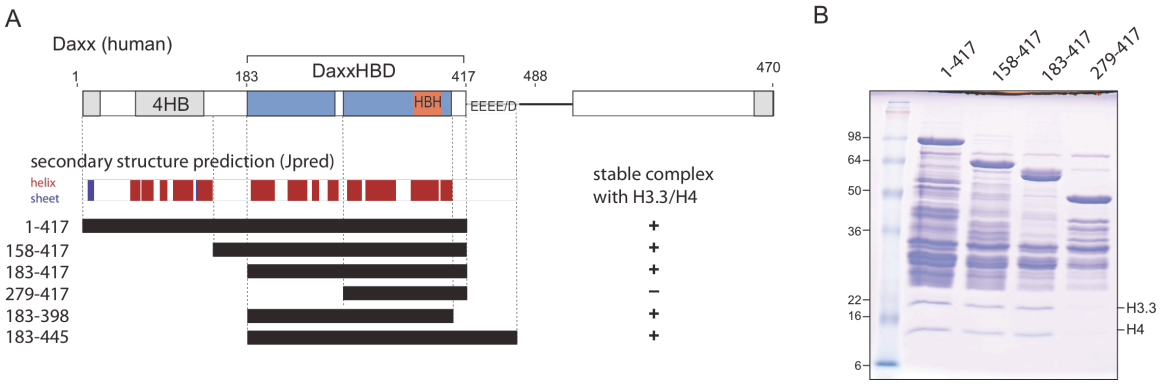



Figure 4.3: GST-tagged Daxx HBD constructs can be expressed soluble in *E. coli* and reconstituted with histones to form a stable DaxxHBD-H3.3-H4 complex.
(A) Initial set of constructs used for DaxxHBD-H3.3-H4 complex reconstitution and crystal trials. (B) Example of DaccHBD-H3.3-H4 complex reconstitutions: GST-tagged Daxx HBD constructs were expressed in *E. coli* at 20°C for 12h. Lysate was applied to glutathione resin. After washing, folded, recombinant H3.3-H4 was added in excess at 500 mM NaCl salt concentration. After further washing with 500mM NaCl, complexes were eluted.

(iii) Given the abundant production of inclusion bodies from His-tagged Daxx HBD, I considered purifying the unfolded protein in 6 M Guanidine·HCl and refolding it *in vitro* by slow dialysis to a native buffer. A typical purification is shown in Figure 4.4. His-tagged Daxx HBD was expressed and purified in inclusion bodies, solubilized with 6 M Guanidine·HCl and further purified on Ni

NTA resin (Figure 4.4A). As expected from the previous results, the Daxx HBD did not fold into a well-behaved structure on its own. Size exclusion chromatography shows that after dialyzing out the Guanidine·HCl, Daxx HBD quantitatively forms aggregates (Figure 4.4B, C).

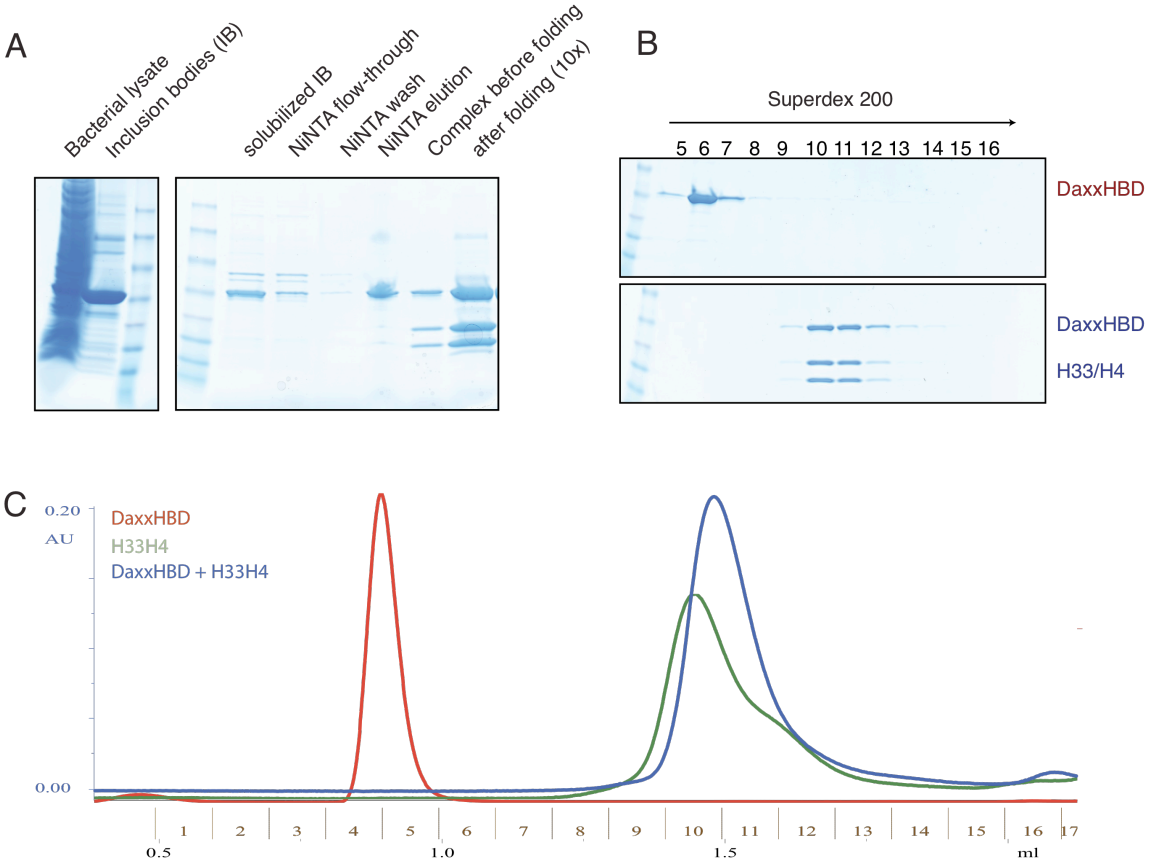



Figure 4.4: *In vitro* folding of the DaxxHBD-H3.3-H4 complex.
(A) Purification of His-tagged Daxx HBD from inclusion bodies. Inclusion bodies were purified from bacterial lysate by differential centrifugation (left panel) and solubilized in 6 M Guanidine·HCl. Daxx HBD was further purified on Ni NTA resin under denaturing conditions (right panel). Equimolar amounts of Daxx HBD, H3.3 and H4 were mixed in 6 M Guanidine·HCl and dialyzed to 50 mM MOPS pH7, 500 mM NaCl (rightmost lanes). (B, C) Daxx HBD refolds into a stoichiometric complex with histones, whereas it runs as an aggregate on a size exclusion column in the absence of histones. Lanes in (B) correspond to the respective fractions taken from (C).

However, when mixing equimolar amounts of recombinant H3.3 and H4 (Section 9.1) and Daxx HBD prior to dialysis, I was able to refold a stoichiometric, trimeric DaxxHBD-H3.3-H4 complex *in vitro*, that eluted as a well-defined peak from a size exclusion column (Figure 4.4B, C). The individual purification of all three components under denaturing conditions yielded pure starting material, which allowed the production of highly pure complexes in a single size exclusion chromatography step after refolding. Furthermore, the method proved highly versatile for reconstituting and purifying multiple combinations of Daxx HBD and histone constructs (i.e. histone tail deletions and mutants) and became my method of choice for producing mg amounts of DaxxHBD-H3.3-H4 for crystallization trials.

### 4.2.3 DaxxHBD-H3.3-H4 is a constitutive heterotetrameric complex

With recent discoveries of obligatory H3-H4 dimer and tetramer chaperones (Adkins et al., 2007; Bowman et al., 2011), it has been appreciated that histone chaperone mechanisms must vary fundamentally between those binding modes (see Section 1.3.6.2). In particular, chaperoning H3-H4 dimers requires either sequential or concerted deposition of two independent H3-H4 units to form the $(H3\text{-}H4)_2$ tetramer that is thought to be prerequisite for subsequent H2A-H2B addition. Coomassie staining of the DaxxHBD-H3.3-H4 complex suggested an even stoichiometry. I sought to confirm this notion with quantitative measurements. To this end, I separated the components of the DaxxHBD-H3.3-H4 complex on a reversed-phase HPLC column and integrated the 280 nm trace to calculate the molar ratio of the three peaks (Figure 4.5A).

While the ratio is not exactly 1:1:1, it certainly does not support a 1:2:2 ratio indicative of a complex with a $(H3\text{-}H4)_2$ tetramer. An alternative quantification can be derived from amino acid analysis. The complex was chemically degraded into single amino acids and their abundance was quantified by HPLC. I compared the reported amino acid frequencies to the hypothetical frequencies in a 1:1:1 or 1:2:2 complex. Experimental values consistently matched a 1:1:1 stoichiometry. As both methods can only assess the relative stoichiometry, I further wanted to investigate the oligomerization state of the DaxxHBD-H3.3-H4 complex. It is formally possible, for example, that two DaxxHBD bind to one $(H3.3\text{-}H4)_2$ tetramer or that the DaxxHBD homodimerizes to carry two separated H3.3-H4 dimers. Both modes would have implications in how nucleosomes can be assembled from the predeposition histones chaperoned by Daxx.

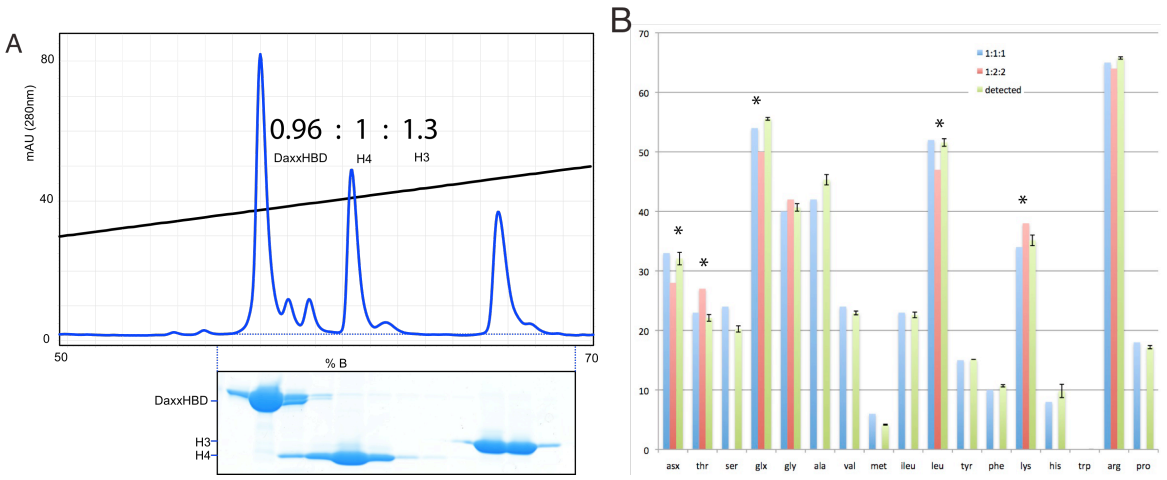



Figure 4.5: The Daxx HBD forms a stoichiometric complex with H3.3-H4.
(A) Reversed-Phase Chromatography of the Daxx HBD-H3.3-H4 complex. Molar ratio was calculated by dividing the area under each peak by the extinction coefficient of the protein. (B) Amino acid analysis of the complex. The experimentally determined number of amino acids per complex was compared to the calculated composition of a 1:1:1 and 1:2:2 complex (where these are significantly different). Stars indicate experimental data matching the 1:1:1 better than 1:2:2 with $p < 0.05$.

There are very few methods that can accurately determine the native composition of protein complexes. one of those is dynamic multi-angle light scattering (MALS). Through measuring the intensity of scattered light at a range of angles, the mass of the scattering object can be accurately determined. MALS in combination with size exclusion chromatography (SEC-MALS) determines the molecular mass of the eluting peak. This is particularly useful because the retention on the size exclusion column is dependent on the hydrodynamic radius of the complex rather than its molecular mass. I confirmed the accuracy of the SEC-MALS by running a number of standard histone complexes (Figure 4.6A). As expected, H3.3-H4 is tetrameric at 500 mM NaCl, dimeric in complex with Asf1 and tetrameric when bound by NAP1. The experimentally determined mass of 61 kDa for the DaxxHBD-H3.3-H4 complex matches the expected mass of a 54 kDa 1:1:1 heterotrimer reasonably well (Figure 4.6A). The Daxx HBD therefore induces constitutive H3.3-H4 dimers. In the light of the HBH model discussed in the previous chapter (see Section 3.2.2), this can be interpreted as a consequence of Daxx binding within the H3-H3 homodimerization interface, disrupting the tetramer. Comparison of the Asf1-H3-H4 and DaxxHBD-H3.3-H4 (Figure 4.6, bottom panels) shows that the Daxx HBD complex runs much larger than the ASF1 complex despite its smaller molecular weight. This indicates that the Daxx HBD has a more extended conformation that occupies a larger solvent volume. A non-globular shape might also explain the slight deviation of the measured 61 kDa mass from the calculated 54 kDa. In line with an extended shape, dynamic light scattering (DLS, which is sensitive to hydrodynamic radius in

solution rather than mass) of the complex yielded an apparent weight almost twice the true mass of the heterotrimer.

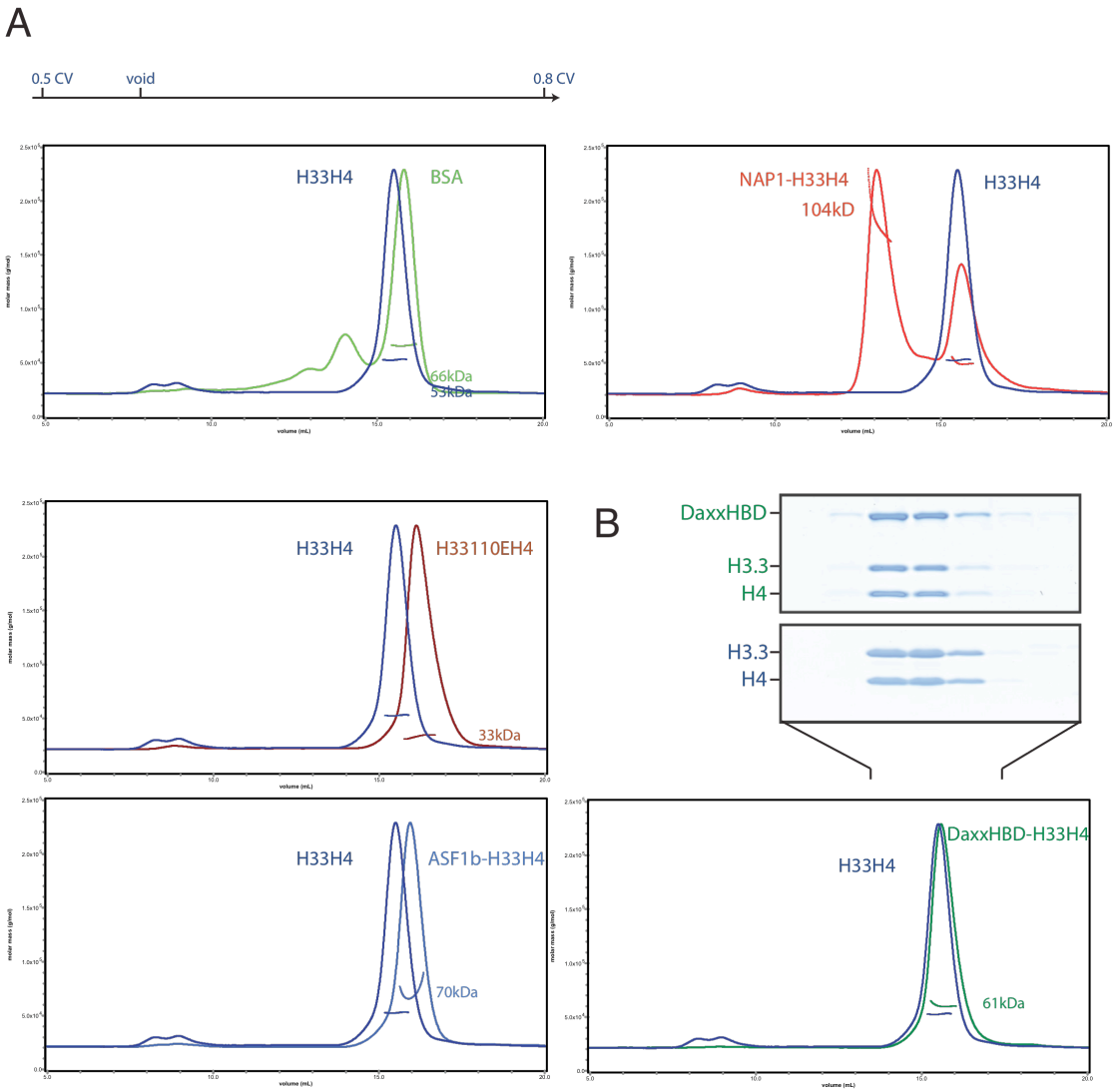



Figure 4.6: Size exclusion chromatography-coupled multiangle dynamic light scattering. SEC-MALS shows that the DaxxHBD-H3.3-H4 complex is a heterotrimer. (A) Reference elution profiles and MALS-derived molecular masses of a number of histone and histone-chaperone complexes. All experiments were carried out in 10 mM MOPS pH7, 1 M NaCl. H33-H4 is a tetramer of 53 kDa although its retention volume suggests a size of 70-80 kDa (as compared to the 66 kDa BSA monomer, top left). A mutation in the H3-H3 homodimerization interface, H3C110E, induces constitutive H3-H4 dimers (middle left). ASF1 forms a heterotrimer of 70 kDa with H3-H4 (bottom left). NAP1 binds one $(H3\text{-}H4)_2$ tetramer. (B) DaxxHBD forms a 61 kDa heterotrimer with H3.3-H4.

### 4.2.4 DaxxHBD blocks the formation of $(H3.3\text{-}H4)_2$ tetramers

All size exclusion chromatography experiments above were carried out at salt concentrations of 1 M NaCl as I have observed non-specific binding of the complex to the column resin at lower salt concentrations. Furthermore, the overall positive surface charge of both Daxx HBD (pI 9.2) and H3.3-H4 (pI 11.3) suggested that higher salt concentrations would stabilize the complex in the absence of counteracting acidic stretches found in the full-length Daxx protein (Figure 4.1). As both inter-histone and histone-chaperone interactions might be salt- and protein-concentration dependent (Banks and Gloss, 2003; Donham et al., 2011), it was important to also assess the stoichiometry of the DaxxHBD-H3.3-H4 complex at physiological salt concentration. I therefore directly tested the oligomeric state of H3-H4 by chemical crosslinking.

In the center of the H3-H3 homodimerization interface, two cysteins (C110 on both H3 units) symmetrically face towards each other, bringing their sulfhydryls sufficiently close to form a disulfide bond or to be bridged by a bivalent Cys-reactive crosslinker. Conveniently, there are no other cysteins in H3.3 or H4, minimizing background crosslinking. Crosslinking therefore leads to the specific formation of stable H3 homodimers associated with a shift in electrophoretic mobility on a SDS-PAGE gel. I incubated either H3.3-H4 tetramers or equimolar amounts of DaxxHBD-H3.3-H4 complex at physiological salt concentration (150 mM NaCl) with either a mild oxidizing reagent (Figure 4.7A) or one of two bis-maleimide crosslinkers (Figure 4.7B). H3.3 homodimers could be observed in the presence of oxidizing or crosslinking reagents for H3.3-

H4 alone. However, the presence of the Daxx HBD efficiently inhibited the formation of H3-H3' homodimers.

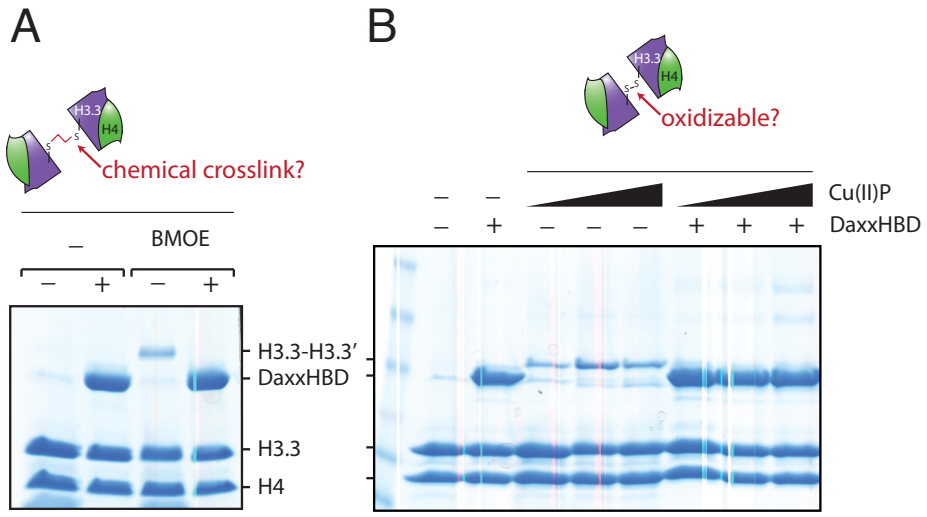



Figure 4.7: The Daxx HBD prevents formation of H3.3-H4 tetramers at physiological salt concentrations.
(A) Crosslinking of H3C110 by bivalent bismaleimide crosslinkers of different lengths. In the presence of Daxx HBD, H3C110 cannot be oxidized or crosslinked. (B) $(H3\text{-}H4)_2$ Tetramer formation probed by oxidation of adjacent H3C110 residues located in the H3-H3' homodimerization interface with excess Cu(II)-Phenanthroline. Resulting H3 dimers are shifted on a non-reducing SDS-PAGE.

This result corroborates the notion that Daxx HBD forms a trimeric complex with H3.3-H4 and furthermore directly proves that the Daxx HBD blocks the H3 homodimerization interface.

### 4.2.5 Crystallization of the DaxxHBD-H3.3-H4 complex

Despite the interesting insights that derive from the homology in the HBHs of Daxx and HJURP (see Chapter 3), major questions about the overall fold and histone contacts of the Daxx HBD, in particular the H3.3 specificity, can only be elucidated by a *de novo* structure. Both X-ray crystallography and, to a lesser

extent, solution NMR spectroscopy have been employed to gather structural information on histone-chaperone interactions (Adkins et al., 2007; Black et al., 2007; Cho and Harrison, 2011; Hu et al., 2011; Natsume et al., 2007; Zhou et al., 2008 2008; Zhou et al., 2011 2008). While solution NMR has been elegantly used to capture a solution state of the Scm3 HBH bound to cenH3-H4 (Zhou et al., 2011), a complex of the full DaxxHBD with H3.3-H4 exceeds the typical limit for NMR structures. Furthermore, major discrepancies between the NMR and crystal structures of Scm3-CenH3-H4 complexes raised questions about the validity of the single-chain fusion approach that was used to reduce the number of residues and stabilize the complex for NMR measurements (Cho and Harrison, 2011; Zhou et al., 2011). As I aimed to elucidate the structure and function of the entire Daxx HBD, I decided to attempt crystallization of the DaxxHBD-H3.3-H4 complex.

#### 4.2.5.1 Pre-crystallization tests and construct considerations

The cumulative biophysical data described above, in particular the homogeneity of the complex as determined by size-exclusion chromatography, MALS and DLS, suggested that the complex would be suited for crystallization trials. Furthermore, while DLS measurements at salt concentrations lower than 500 mM NaCl indicated a tendency to form aggregates (data not shown), at 500 mM NaCl or above, the complex remained intact for weeks at 4˚C or flash-frozen at -80˚C, as determined by analytic size-exclusion chromatography. Therefore, the protein complex was always purified and concentrated at 500 mM to 1 M NaCl. To preserve a constant surface charge, pH 7 was used in most

purifications, well below the pI of the complex (9-10.5, depending on the histone and Daxx HBD constructs used)

Another often mentioned prerequisite for successful crystallization is the absence of unstructured, highly mobile stretches within a protein, such as extended loops or tails. Various segments of the N-terminal H3 and/or H4 tails were therefore removed in some crystallization trials. However, it should be noted that the presence of the histone tails did not impede high-resolution crystal structure determination for several histone-chaperone complexes (Hu et al., 2011; Natsume et al., 2007).

The first ~43 residues of H3 are unstructured in the nucleosome (Figure 1.8). While constituting the first (⟨N)-helix of H3 in the nucleosome core, the following residues 43 to 60 appear to be disordered in the context of H3-H4 tetramers (Adkins et al., 2007; Natsume et al., 2007; Sekulic et al., 2010). In an earlier experiment, which initially characterized the Daxx-histone interaction, I have found the first 60 residues of H3 and 20 of H4 to be dispensable for Daxx HBD binding (Figure 2.3). Therefore, deletions of the first 28, 43, 57 and 60 amino acids of H3 were tested in various crystal trials. Biophysical properties of the DaxxHBD-H3-H4 complex were affected by H3 tail deletions beyond residue 40, as evident from a greatly increased tendency to aggregate during complex refolding and concentration. All deletions of the H4 tail, up to residue 20, behaved normal.

In order to identify possible unstructured regions in the Daxx HBD, I conducted limited proteolysis of the DaxxHBD-H3.3-H4 complex with a set of

diverse proteases (Figure 4.8). Trypsin and subtilisin were the only proteases that significantly digested the Daxx HBD.

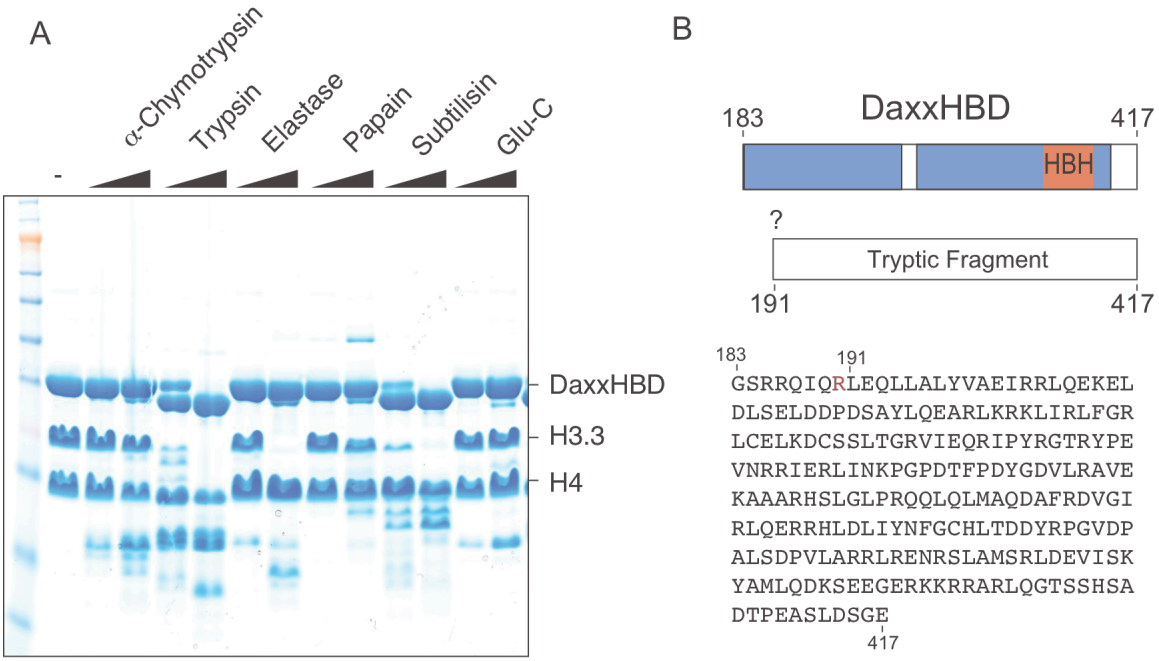



Figure 4.8: Limited proteolysis of DaxxHBD-H3.3-H4 complex for crystallography.
(A) Incubation of the DaxxHBD-H3.3-H4 complex with indicated proteases for 4h at 37°C at a 1:1000 or 1:100 (w/w) ratio. (B) Potential fragment corresponding to the slightly faster migrating band in the Trypsin and Subtilisin digest. Tryptic digest and MS/MS peptide fingerprinting of this band suggests the presence of the complete C-terminal sequence (peptides detected ending in LDSGE, data not shown) but the lack of at least 8 amino acids on the N-terminus.

---

The resulting fragment was analyzed by MS/MS peptide mapping. Interestingly, peptides from the far C-terminus (400-417) were found abundantly, suggesting that it was not cut despite the fact that it was predicted to be unstructured (Figure 4.1). Notably, the fragment found after full-length Daxx expression in *E. coli* showed similar protease resistance up to residue 417 (Figure 4.2). This might indicate that the far C-terminus of the Daxx HBD folds back onto the central structured part. Given the intact C-terminus, the proteolytic trypsin/subtilisin

fragment in Figure 4.8 must be cleaved near the N-terminus, which was further corroborated by the absence of peptides mapping to this region in the lower, proteolyzed band. Consequently, I included Daxx HBD constructs that started after potential trypsin cleavage sites (i.e. 191-417) in my screening efforts.

#### 4.2.5.2 The wild type DaxxHBD-H3.3-H4 complex could not be crystallized

Using the protocols outlined in Section 4.2.2, I reconstituted a variety of DaxxHBD-H3.3-H4 complexes for crystallization trials (Table 4.1). All complexes were purified and salt-exchanged into crystallization buffer (10 mM MOPS pH 7, 0.5–1 M NaCl, 0.5 mM EDTA, 5 mM DTT or 1 mM TCEP) over a final Superdex 200 size exclusion column and concentrated up to 30–50 mg/mL immediately before manual or robotic screen set up. Despite prior assessment of biochemical and biophysical homogeneity of the complex (see previous section), initial trials did not yield any crystals. At high protein concentrations, many conditions showed phase separation phenomena or abundant precipitation. Anecdotal crystals or crystal-like conditions discovered in the primary screens proved to be salt crystals or did not reproduce in subsequent refinement screens. Given the large number of conditions that were sampled in the sum of the screens (the 183-417 fragment of Daxx HBD was screened in at least 1000 unique conditions), I had to consider that some intrinsic property of the complex prohibited ordered crystal packing. This did not appear to stem from unstructured tails as trimming of the constructs *per se* did not facilitate crystallization (Table 4.1). Therefore, methods to address other potential obstacles intrinsic to the complex needed to be evaluated.

One detriment that I suspected was the large number of surface lysines and argingines of the complex which generally add to a high surface entropy through their charge and side chain flexibility (Derewenda, 2004).

Table 4.1: Summary of crystallization trials.
3-20 mg of purified and concentrated protein were produced in each trial and usually screened against a set of 4-10 non-redundant primary 96-well plate screens. Crystals are only indicated where they could be reproduced in a subsequent refinement screen. The best diffraction values are given where available.

| **GST-TEV-HBD** | **H3.3** | **H4** | **crystals** | **diffraction** |
|---|---|---|---|---|
| 183-417 | full length 1-135 | full length 1-102 | - | |
| 183-417 lysine-methylated | full length | full length | - | |
| 186-417 | 28-135 | 15-102 | - | |
| 191-417 | 28-135 | 20-100 | - | |
| **His-TEV-DaxxHBD** | **H3.3** | **H4** | | |
| 183-417 | 60-135 | 20-100 | - | |
| 183-417 | 43-135 | 20-100 | - | |
| 183-417 | 1-135 | 1-94 | - | |
| **His-TEV-DaxxHBD** | **H3.3$^{7CENPA}$** | **H4** | | |
| 183-417 | full length | full length | + | 9Å |
| 183-417 lysine-methylated | full length | full length | - | |
| 183-417 | full length | 20-100 | + | 9Å |
| 183-417 | full length | 1-94 | - | |
| 183-445 | full length | full length | + | 9Å |
| HHHHHH-183-417 noncleavable tag | full length | full length | + | 9Å |
| 158-417 | full length | full length | + | ? |
| 183-398 | full length | 20-102 | + | 3Å |

While targeted mutations of surface residues could potentially interfere with native contacts made between histones and chaperones, reductive methylation of the preformed complex provided a non-destructive method to increase the hydrophobicity or 'stickiness' of exposed lysines to facilitate crystal packing (Kim et al., 2008). However, no improvement in crystallization behavior was observed by this method (in combination with mutations described below, it even inhibited crystal formation).

#### 4.2.5.3 Stabilizating mutations in the histone fold allow crystallization of the DaxxHBD-H3.3-H4 complex

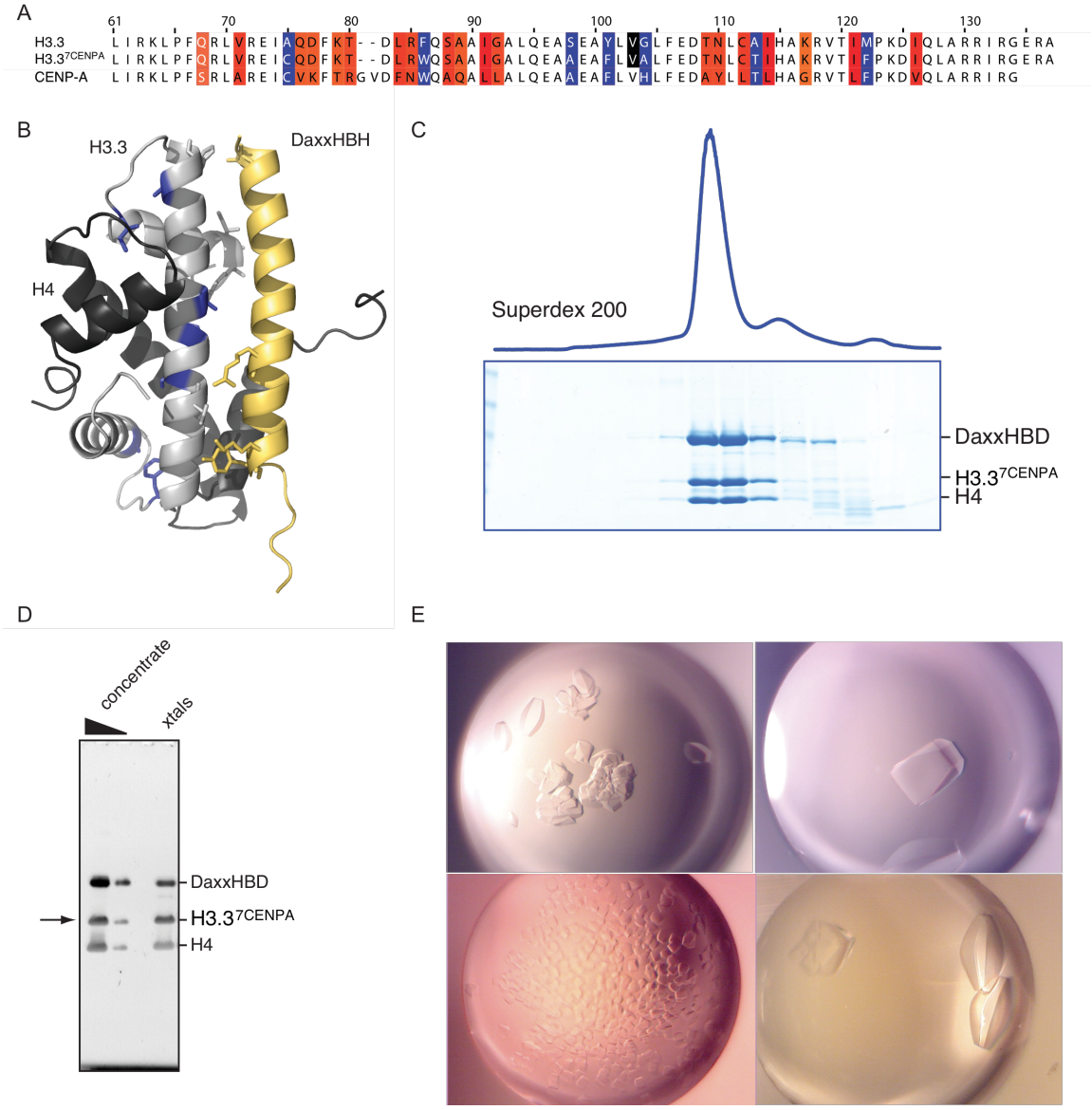
I considered another source of entropy that prevented ordered crystal packing to come from the histone fold itself. This hypothesis is based on an interesting observation made by Dr. Ben Black and colleagues (Black et al., 2004; Sekulic et al., 2010): using hydrogen/deuterium exchange-mass spectrometry, they detected considerable proton exchange with solvent at side chains of H3 and H4 that are predicted to lie in the hydrophobic (solvent-inaccessible) core of the histone fold. He therefore concluded that the solution state of the H3-H4 histone fold is quite dynamic, in contrary to the rigid organization observed in all available H3-H4-containing crystal structures. Interestingly, the corresponding internal residues of the CENPA-H4 heterodimer showed much less hydrogen/deuterium exchange and therefore suggested a more rigid architecture (Black et al., 2004), which could later be attributed to increased hydrophobicity of a number of CENP-A specific side chains extending into the core of the histone fold (Sekulic et al., 2010). This difference in rigidity furthermore lends an explanation to the curious observation that CENPA-H4

readily crystallized as a tetramer whereas no crystals could be derived from free H3-H4 tetramers despite intense efforts (B. Black, personal communication).

I reasoned that it would be possible to separate the mutations that stabilize the CENPA-H4 core from mutations that are involved in CENPA-specific chaperone recognition. Unfortunately, no extensive single point mutagenesis between H3 and CENP-A residues has been carried out to date. Nevertheless, I identified a number of unique CENPA-H4 side chains in the CENPA-H4 tetramer that had virtually no contact with its surface in the crystal structure (Sekulic et al., 2010) (Figure 4.9B, blue; also see Figure 3.1). Their counterparts in the H3-H4 tetramer had highly correlated positions, suggesting that they contributed to the internal architecture of both heterodimers without affecting their exposed surfaces. Importantly, H3.1/2 and H3.3 shared 100% identify in those internal residues, arguing that they are not involved in variant discrimination. Consequently, it seemed possible that introducing stabilizing mutations into H3.3 would lower the H3.3-H4 heterodimer entropy without affecting variant-specific properties.

I mutated a total of 7 residues in H3.3 towards their CENP-A counterparts (creating a hybrid termed $H3.3^{7CENPA}$, Figure 4.9A). The resulting DaxxHBD- $H3.3^{7CENPA}$-H4 complex behaved as wild type and could be purified to homogeneity (Figure 4.9C). Within a primary crystallization screen of ~500 conditions, several hits were obtained, all containing sodium formate as a precipitant. Crystals contained the stoichiometric complex (Figure 4.9D).

Figure 4.9: Engineering of a more rigid H3.3-H4. Mutations were guided by sequence comparison of H3.3 and the more rigid CENP-A.
(A) Alignment of H3.3, CENP-A and the engineered H3.3$^{7CENPA}$. Nonsynonymous residues between H3.3 and CENP-A are highlighted in color; orange residues are surface-exposed and thought to facilitate specific recognition of CENP-A by chaperones; blue residues are completely buried in the hydrophobic core of the CENPA-H4 fold. The latter 7 residues were mutated in the H3.3 sequence towards the CENP-A counterpart in the H3.3$^{7CENPA}$ (middle lane). (B) Location of the mutated residues in the DaxxHBD-H3.3-H4 models; none of the mutations affect the histone-chaperone interface. (C) Purification and (D) crystallization of the DaxxHBD-H3.3$^{7CENPA}$-H4 complex with formate. (E) Optimized crystals derived from this complex (top row, bottom right) and DaxxHBD$^{183-445}$-H3.3$^{7CENPA}$-H4 (bottom left).


A
61
70
80
90
100
110
120
130
H3.3 L I R K L P F Q R L V R E I A Q D F K T - - D L R F Q S A A I G A L Q E A S E A Y L V G L F E D T N L C A I H A K R V T I M P K D I Q L A R R I R G E R A
H3.3^7CENPA L I R K L P F Q R L V R E I C Q D F K T - - D L R W Q S A A I G A L Q E A A E A F L V A L F E D T N L C T I H A K R V T I F P K D I Q L A R R I R G E R A
CENP-A L I R K L P F S R L A R E I C V K F T R G V D F N W Q A Q A L L A L Q E A A E A F L V H L F E D A Y L L T L H A G R V T L F P K D V Q L A R R I R G
B
H3.3
DaxxHBH
H4
C
Superdex 200
DaxxHBD
H3.3^7CENPA
H4
D
concentrate
xtals
DaxxHBD
H3.3^7CENPA
H4
E

Formate conditions could be optimized to yield few large singular crystals, which looked extremely uniform by eye (Figure 4.9E). Subsequent crystallization of a number of constructs revealed that in the presence of the 7 mutations in H3.3, the extent of the Daxx HBD fragment or the presence of the histone tails did not majorly influence crystal packing (Figure 4.9E, Table 4.1).

#### 4.2.5.4 Alternate crystal forms of the DaxxHBD-H3.3$^{7CENPA}$-H4 complex with greatly different diffraction limits.

Initial crystals of the DaxxHBD-H3.3$^{7CENPA}$-H4 diffracted to 10Å. Crystallization with formate as a precipitant proved to be exquisitely robust to a pH range between 6 and 8, and to additives such as salts, detergents and alcohols. However, none of those refinements and additives improved diffraction beyond 9Å. In addition, Varying the length of the Daxx HBD fragment or the histone tails did not yield higher diffracting crystals either under any tested formate condition (Table 4.1). Close distance of diffraction spots suggested a large asymmetric unit with high disordered water content. Consequently, crystal dehydration was attempted by raising the formate concentration slowly to 5 M in a hanging drop, but diffraction was still limited to 9Å.

I therefore rescreened a more minimal complex (DaxxHBD$^{183-398}$-H3.3$^{7CENPA}$-H4$^{20-102}$, see Figure 4.10A) under ~500 new conditions (including the full JSCG Core Suite) and found a second condition that contained 0.1 M Na/K phosphate (pH 6.2) and 2.5 M NaCl to yield a single, large crystal. This crystal diffracted to 3.5Å, with some imperfections (Figure 4.10C). Crystals could be reproduced and are now under optimization (Figure 4.10B). Preliminary rescreening of other complexes (e.g. with DaxxHBD$^{183-417}$) under the phosphate

condition did not produce crystals, suggesting that this condition might require the truncated Daxx HBD.

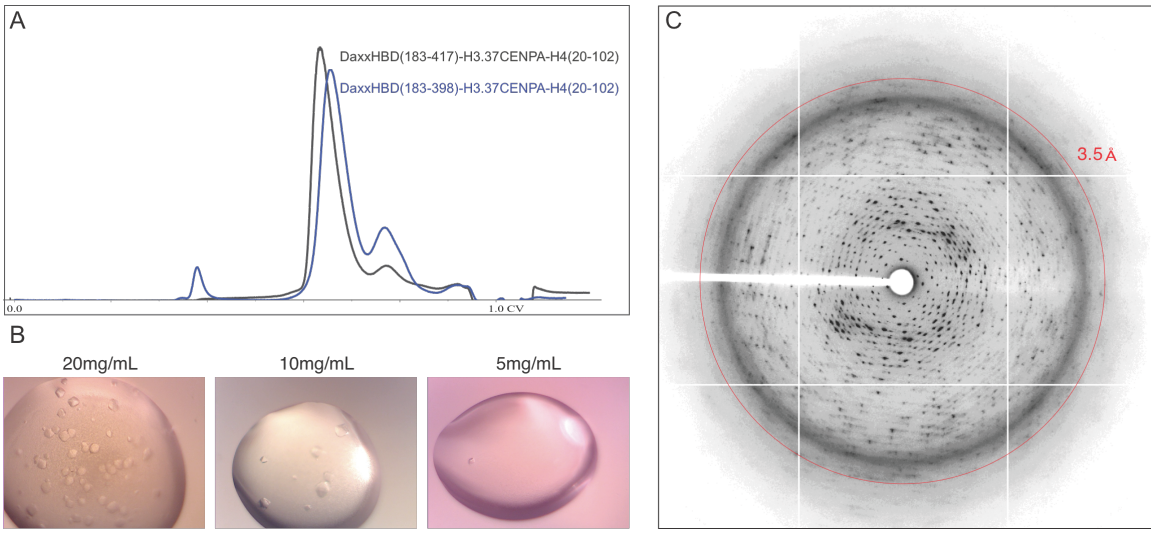



Figure 4.10: A C-terminally truncated Daxx HBD yields crystals under an alternative salt condition.

(A) Purification of the DaxxHBD$^{183\text{-}398}$-H3.3$^{7CENPA}$-H4$^{20\text{-}102}$ complex over a Superdex 200 size exclusion column. (B) crystal growth at indicated protein concentrations in 0.1 M Na/K phosphate pH6.2, 2.5 M NaCl. (C) Diffraction pattern of the single crystal in the initial sitting drop screen.

## 4.3 Outlook

With the current crystals diffracting to at least 3.5Å, it should be possible to solve an initial structure of the DaxxHBD-H3.3-H4 complex. Phases may be solved by molecular replacement with the histone moiety of Asf1 or HJURP costructures, as the histone fold architecture is expected to be preserved in the complex. As an alternative, the simple and efficient expression of the Daxx HBD in inclusion bodies makes production of a Se-Met derivative feasible. The resulting structure will be particularly valuable for understanding the specific recognition of the H3.3 CRD by Daxx, given the subtle amino acid differences to H3.1/2 (Figure 3.1).

Point mutants in the Daxx HBD will be designed that make contacts to the CRD to see if altering the recognition interface interferes with the H3.3 specificity. In addition, different parts of the interface could be mutated in Daxx *in vivo* to study their impacts on discriminating and chaperoning H3.3-H4 dimers.

# 5 HIRA and Daxx constitute two independent histone H3.3-containing predeposition complexes[‡]

## 5.1 Introduction

Despite the inherent physical stability of the nucleosome particle (e.g. as discussed in Section 1.3.3), the resulting chromatin structure must be highly dynamic to allow access to the DNA-encoded information. H3-H4 and H2A/H2B units are deposited and evicted from nucleosomes by different histone chaperones and ATP-dependent chromatin remodelers. Histone chaperones thereby serve two major functions: first, to buffer the pool of non-chromatin associated histones in predeposition complexes, and second, to facilitate favorable histone-DNA interactions during the step-wise assembly of the histone octamer-DNA complex from H3-H4 and H2A/H2B units (see section 1.3.5). Mutants of histone chaperones in *S. cerevisiae*, for example, cannot form proper chromatin structure (Adkins and Tyler, 2004; Andrews et al., 2010; Sharp et al., 2002). Metazoans have evolved histone variants to meet the diverse needs for establishing and maintaining complex epigenetic information. Histone H3.3 is a minor variant of the 'canonical' S-phase histones H3.1 and H3.2 in metazoans.

[‡] This chapter has been published as: Elsaesser, S. J., and Allis, C. D. (2010). HIRA and Daxx constitute two independent histone H3.3-containing predeposition complexes. Cold Spring Harb Symp Quant Biol 75, 27–34.

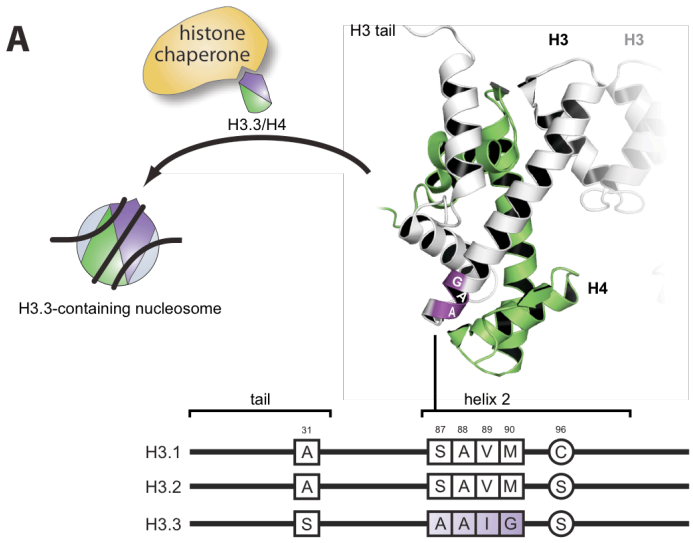



Figure 5.1: A highly conserved 'AAIG' motif in the core histone fold of metazoan H3.3. Schematic representation of non-centromeric histone H3 variants in mammals. H3.1 and H3.2 are commonly referred to as 'canonical' or S phase histone H3. Variable amino acids are shown with residue number. Three unique residues of H3.3 from the 'AAIG' motif at the base of helix 2 of the histone H3 core (purple, note that the isoleucine is hidden in the helical depiction). The molecular structure of a representative H3-H4 dimer (H3 white, H4 green) has been adapted from the nucleosome structure (PDB:1KX5), the H3:H3 interface of the H3-H4 tetramer is indicated.

However, its role as a replacement histone outside of S-phase is conserved to lower eukaryotes, and accumulating evidence suggests that by virtue of a unique amino acid motif, H3.3 engages in several variant-specific chromatin assembly pathways. While the amino acid sequence of H3.3 is over 95% identical to H3.1/2, the small 'AAIG' motif in the core histone fold (Figure 5.1) allows targeting of the variant to specific genomic loci by a replication-independent mechanism (Ahmad and Henikoff, 2002c; Goldberg et al., 2010). The existence of identical H3 variants throughout the metazoan lineage underscores their

importance as carriers of epigenetic information in somatic cells and the germline (see Section 1.2.3). Reading out the subtle differences between H3.3 and H3.1/2 poses an intricate structural problem to variant-specific chaperones. Hence, a number of recent studies have aimed to elucidate which proteins confer specificity in chromatin assembly. By purifying H3.1 and H3.3-associated complexes from HeLa nuclear extracts, the replication-independent histone deposition factor HIRA was found to associate specifically with H3.3 (Tagami et al., 2004). Using the same approach, others and we identified a different polypeptide, the Fas death-domain associated protein Daxx, as an H3.3-specific binding protein (Drané et al., 2010; Goldberg et al., 2010; Lewis et al., 2010). I found that Daxx specifically interacts with the 'AAIG' motif of H3.3 *in vitro* and is a bona fide H3.3-specific histone chaperone (Figure 2.3). The ATP-dependent chromatin remodeler and ⟨-thalassemia X-linked mental retardation protein ATRX forms a complex with Daxx-H3.3-H4 that was essential for H3.3 chromatin assembly *in vivo* (Goldberg et al., 2010; Lewis et al., 2010).

While I was able to detect HIRA in H3.3-associated fractions, this association was unstable upon exposure to the more stringent biochemical conditions used to identify Daxx (

Figure 2.2). Preceding my characterization of the putative histone-binding motif (Section 3.2.3), this raised the question of whether HIRA and its associated proteins Ubinuclein-1 and Cabin-1 were recruited indirectly, potentially via a secondary interaction with Daxx or another as yet unidentified protein. I addressed this issue by biochemically fractionating H3.3-containing sub-complexes from HeLa cells and purifying H3.3-associated proteins from mouse

embryonic stem cells (ESCs) devoid of HIRA, Daxx or ATRX. I found that the HIRA/Ubinuclein-1/Cabin-1 and ATRX/Daxx complexes are two separate biochemical entities that independently associate with H3.3. We have previously shown that HIRA is responsible for incorporation of H3.3 at genic regions whereas ATRX is required for telomeric deposition of H3.3 (Goldberg et al., 2010). HIRA and Daxx/ATRX therefore constitute two separate histone H3.3-deposition machineries that act on distinct genomic loci.

## 5.2 Results

### 5.2.1 Biochemical fractionation of H3.3-containing complexes

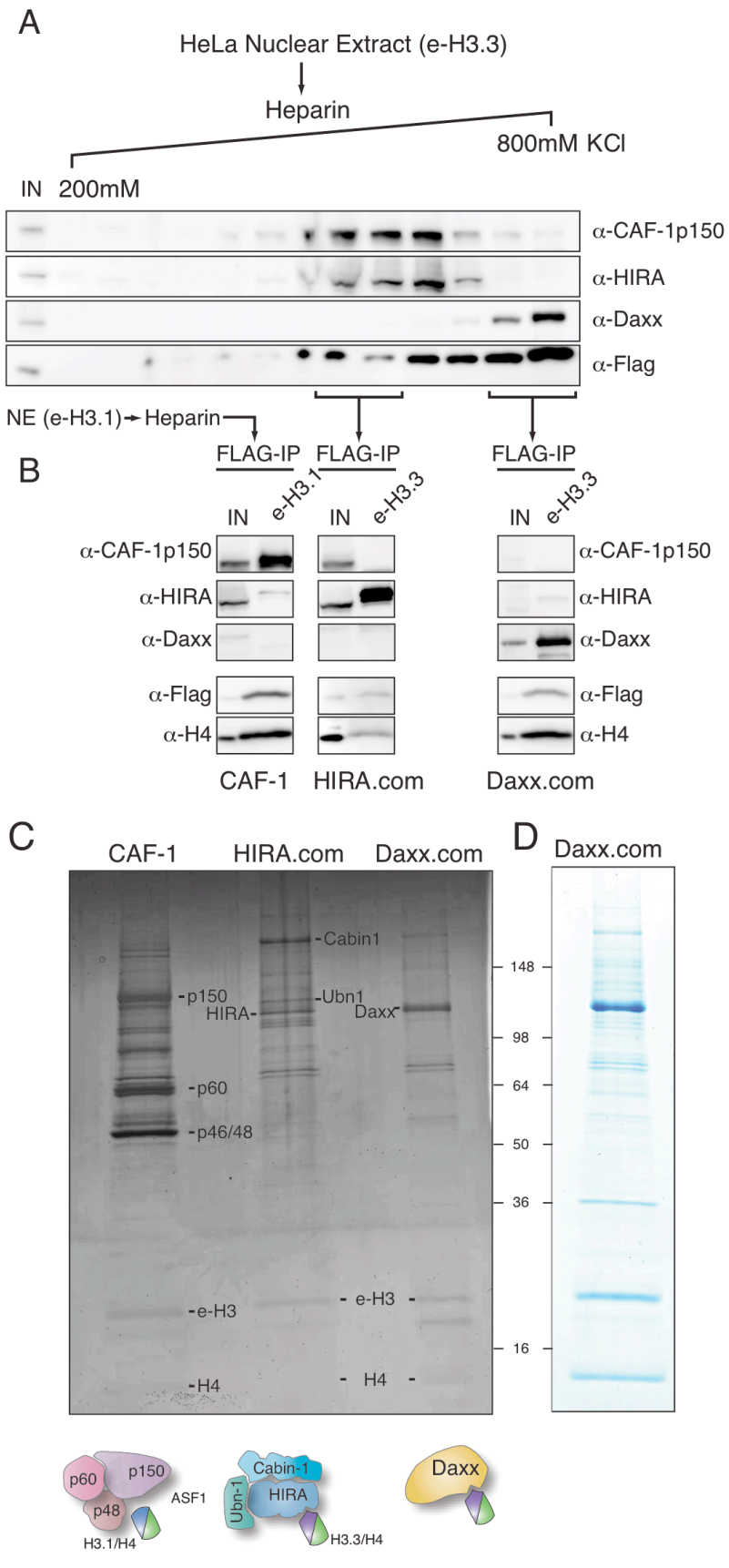
Previous studies used double-affinity purification of FLAG-HA-tagged H3.3 (e-H3.3) from nuclear extracts to isolate H3.3-associated factors (Drané et al., 2010; Lewis et al., 2010; Tagami et al., 2004). However, this strategy did not allow me to identify putative sub-complexes that independently co-immunoprecipitate with H3.3. Thus, I developed an alternative biochemical approach to fractionate H3.3-containing complexes from nuclear extracts based on their affinity to a negatively charged heparin matrix. I noted that e-H3.3 eluted over a broad range of salt concentrations from immobilized heparin (Figure 5.2A). While HIRA and CAF1 p150 complexes eluted with the bulk of nuclear proteins at medium salt, Daxx only dissociated at higher salt concentrations from the heparin column (Figure 5.2A). This approach therefore provided me with a means to separate distinct subpopulation of e-H3.3 and associated proteins. I combined fractions containing almost exclusively

HIRA/CAF1 p150 or Daxx and immunoprecipitated e-H3.3 (Figure 5.2B). HIRA, but not CAF1 p150 copurified with H3.3. As Daxx was absent from the immunoprecipitated material, I conclude that HIRA associates specifically with H3.3, independent of Daxx. As a control, I heparin-fractionated e-H3.1 nuclear extracts (data not shown) and performed immuoprecipitation from the corresponding HIRA/CAF1 p150 fractions; this e-H3.1 immunoprecipitation yielded CAF1 p150 (Figure 5.2B). Daxx coimmunoprecipitated with H3.3 from the higher salt fractions in the absence of HIRA (Figure 5.2B). Thus, I was able to isolate non-overlapping H3.3 sub-complexes from heparin-fractionated nuclear extract containing either HIRA or Daxx.

In addition to HIRA, I found Cabin-1 and Ubinuclein-1 associated with the first pool of e-H3.3-H4 (Figure 5.2C). HIRA, Cabin-1 and Ubinuclein-1 have been previously found to form a complex in the presence and absence of H3.3-H4 (Drané et al., 2010; Tagami et al., 2004). Iterative alignment searches identified Cabin-1 and Ubinuclein-1 as mammalian homologs of the yeast proteins Hir3 and Hcp2p, respectively, suggesting a HIRA complex analogous to the yeast Hir complex. The control H3.1 immunoprecipitation yielded all three subunits of the CAF1 complex, p150, p60 and RbAp46/48. Daxx co-precipitated with H3.3 from the higher salt fractions without apparent stoichiometric partners (Figure 5.2C, D), in agreement with our previous observation that the direct interaction of Daxx with H3.3-H4 is not dependent on other proteins such as ATRX (Lewis et al., 2010). In conclusion, I find that the known H3.3-associated proteins fractionate into two biochemically distinct subcomplexes constituted by HIRA and Daxx.

Figure 5.2: HIRA and Daxx constitute two biochemically distinct H3.3-containing subcomplexes.

(A) Biochemical fractionation of nuclear H3.3 complexes: 200-800 mM KCl fractions from a heparin column were probed for e-H3.3, HIRA, CAF1 and Daxx by western blotting. (B) Affinity purification of e-H3.3 and e-H3.1 complexes. e-H3.3/H3.1 complexes were immunoprecipitated from indicated pools and analyzed by western blotting. Immunoprecipitation from the medium-salt e-H3.3 and e-H3.1 fractions yielded predominantly HIRA and CAF1, respectively; Daxx coprecipitated with high-salt e-H3.3. (C) Silver staining of eluted CAF1, HIRA and Daxx complexes from above. Indicated bands were all identified by MS and/or verified by western blotting. (D) Coomassie stain of the Daxx complex to estimate protein abundance.


A
HeLa Nuclear Extract (e-H3.3)
Heparin
800mM KCl
IN 200mM
α-CAF-1p150
α-HIRA
α-Daxx
α-Flag
NE (e-H3.1)→Heparin
FLAG-IP
FLAG-IP
FLAG-IP
B
IN
e-H3.1
IN
e-H3.3
IN
e-H3.3
α-CAF-1p150
α-HIRA
α-Daxx
α-Flag
α-H4
CAF-1
HIRA.com
Daxx.com
C
CAF-1
HIRA.com
Daxx.com
Cabin1
p150
HIRA
Ubn1
Daxx
p60
p46/48
e-H3
H4
e-H3
H4
148
98
64
50
36
16
D
Daxx.com
p60
p150
p48
ASF1
H3.1/H4
Cabin-1
Ubn-1
HIRA
H3.3/H4
Daxx

### 5.2.2 Genetic dissection of H3.3-containing complexes

Next, I assessed how the deletion of either HIRA or Daxx histone chaperones affects H3.3 and its associated complexes in respective knockout mouse embryonic stem cell (ESC) lines. I used a zinc-finger based genome editing strategy to specifically tag one allele of the endogenous H3.3B gene in wild type W9.5, HIRA-/-, Daxx-/- and HIRA-/- ESCs with a C-terminal HA tag (H3.3-HA) (Goldberg et al., 2010), assuring endogenous expression levels of the tagged histone. I found elevated levels of H3.3-HA in nuclear extracts of HIRA-/- ESCs as previously reported (Meshorer et al., 2006), while the soluble pool of H3.3-HA was reduced in Daxx-/- ESCs (Figure 5.3, left panel). I speculate that Daxx buffers a considerable pool of H3.3 in the nucleus by forming a stable complex with H3-H4 units. Overall H3 and H4 levels (accounting for the untagged H3.3, as well as H3.1/2) in the nucleoplasm were unchanged, indicating that reduced levels of H3.3 might be compensated by an increase in soluble H3.1/2. Levels of chromatin-bound H3.3-HA were comparable in all cell lines (Figure 5.3, bottom). As previously reported, ATRX levels were reduced in Daxx-/- cells (Lewis et al., 2010). Unexpectedly, Cabin-1 was depleted in HIRA-/- nuclear extracts (Figure 5.3, left panel) and whole cell lysates (data not shown). Cabin-1 mRNA levels are not perturbed in HIRA-/- cells (Goldberg et al., 2010), suggesting that the Cabin-1 protein might be unstable and degraded in the absence of HIRA. A recent study confirmed that Cabin-1, like Ubinuclein-1 contributes directly to the function of the HIRA complex (Banumathy et al., 2009; Rai et al., 2011).

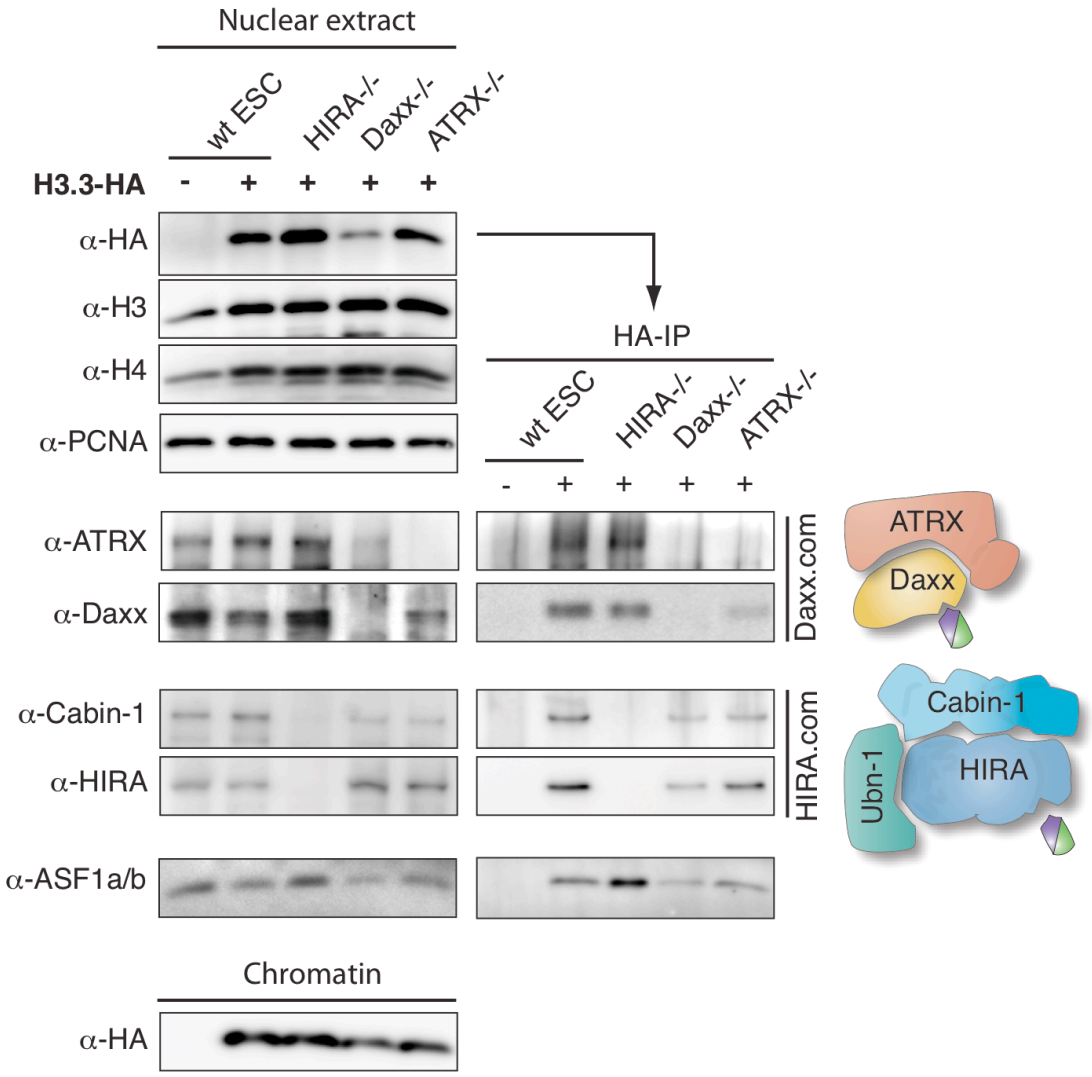



Figure 5.3: HIRA and Daxx independently associate with nucleoplasmic H3.3.
Nuclear extracts were prepared as above from indicated mouse embryonic stem cell lines (Lewis et al., 2010; Roberts et al., 2002), either parental or targeted heterozygous H3.3B-HA knock-in (left panel). Protein levels in nucleoplasm (top) and insoluble chromatin (bottom) were immunoblotted. e-H3.3 was bound to ⟨-HA-Agarose, washed and eluted with 1.5% SDS. Copurifying proteins were analyzed by western blotting (right panel).

### 5.2.3 HIRA and Daxx complexes independently bind H3.3

Based on our biochemical fractionation of H3.3-containing complexes, I hypothesized that association of HIRA and Daxx with H3.3 are independent of one another. I scrutinized this notion by immunoprecipitating H3.3-HA from the above ESC nuclear extract. Indeed, HIRA copurified with H3.3-HA in Daxx-/- cells and vice versa (Figure 3, right panel). ATRX coprecipitated with Daxx but ATRX deletion did not influence binding of either histone chaperone, as expected from our earlier studies (Lewis et al., 2010). ASF1 was constitutively associated with H3.3-HA.

I therefore conclude that the HIRA and Daxx complexes independently interact with H3.3 and likely maintain distinct pools of H3.3-H4 units in the nucleus. This is in line with our recent observation that HIRA and ATRX/Daxx deposit histones at distinct genomic locations: HIRA at genic regions (Goldberg et al., 2010) and ATRX/Daxx at telomeres as well as pericentromeric heterochromatin (Drané et al., 2010; Lewis et al., 2010). While my *in vitro* studies confirmed that Daxx directly interacts with the H3.3 'AAIG' motif (Figure 2.3), similar direct evidence is thus far missing for HIRA. As discussed in Section 3.2.3, the presence of a histone binding helix motif in the C-terminus is a first indication of a direct binding element in HIRA. However, my finding that HIRA consistently copurifies with H3.3 in a complex with Cabin-1 and Ubinuclein-1 (Lewis et al., 2010; Tagami et al., 2004) raises the possibility that these components also contribute to binding and specificity.

### 5.2.4 Do predeposition complexes represent independent histone pools?

While H3.3 was originally characterized to be associated with actively transcribed genes (Ahmad and Henikoff, 2002c), the discovery of new H3.3 chaperone pathways and heterochromatic target regions (Drané et al., 2010; Hake et al., 2005; Lewis et al., 2010; Santenard et al., 2010; van der Heijden et al., 2007) poses some intriguing questions: Does H3.3 carry out unique functions at distinct genomic locations? How can a single replacement variant act activating or repressing in a context-dependent manner? How do the specialized deposition machineries contribute to the functional outcome of H3.3 incorporation?

The functional diversity of H3.3 could be explained if the distinct predeposition complexes would represent independent pools of H3.3-H4 units. Assuming that the exchange between those pools is slow due to the biochemical stability of the complexes or their distinct spatial or temporal distribution during the cell cycle, the associated histones could acquire specific posttranslational modifications (PTMs) that might prime them for their respective functions at the genomic target sites.

### 5.2.5 Distinct histone posttranslational modifications are present in predeposition complexes

Histone variants have been found enriched in PTMs that correlate with the local chromatin environment at their target sites (Hake et al., 2006). A set of specific modifications has been identified for the soluble pool of histones composed of newly synthesized histones but possibly also evicted histones 'recycled' from chromatin: ubiquitous H4K5/K12 diacetylation (H4K5/K12diAc)

(Sobel et al., 1995) and H3K9 monomethylation (H3K9me1) on both H3.1/2 and H3.3, as well as H3K9 dimethylation (H3K9me2) and H3K9 acetylation (H3K9Ac) enriched on H3.3 (Loyola et al., 2006). Given the overall heterogeneity of predeposition PTMs on the single H3.3K9 residue, I wondered if any of these marks might be characteristic of a subset of predeposition H3.3 associated with a specific chaperone complex.

Using heparin fractionation to separate e-H3.3-containing complexes found in HeLa cells, we identified a subpopulation of e-H3.3 that was enriched in H3K9Ac (Figure 5.4A), a modification that is thought to be established on non-nucleosomal H3 by the histone acetyltransferase Gcn5 (Adkins et al., 2007). As the enrichment of H3K9Ac tracked well with the HIRA complex, we further analyzed the PTMs present on histones within the CAF1, HIRA and Daxx complexes (Figure 5.4B). We found that H3K9Ac on e-H3.3 associated with the HIRA complexes but was undetectable in the Daxx complex. Notably, a mass spectrometric analysis of the H3.3/Daxx-associated proteins in (Figure 5.2D) yielded the H3K9 methyl transferases G9a and GLP, suggesting that H3K9 is mono- or dimethylated (Wen et al., 2009) in this complex (see Appendix 9.1). However, I have not followed up on this recent finding. Both Daxx and HIRA complexes contained low levels of acetylated H4. e-H3.1 in the CAF1 complex was enriched in H4 acetylation, with H3K9Ac present at a lower level than on e-H3.3. With the differential H3K9 acetylation as a first indication, we conclude that unique H3.3-H4 PTMs can be associated with predeposition complexes.

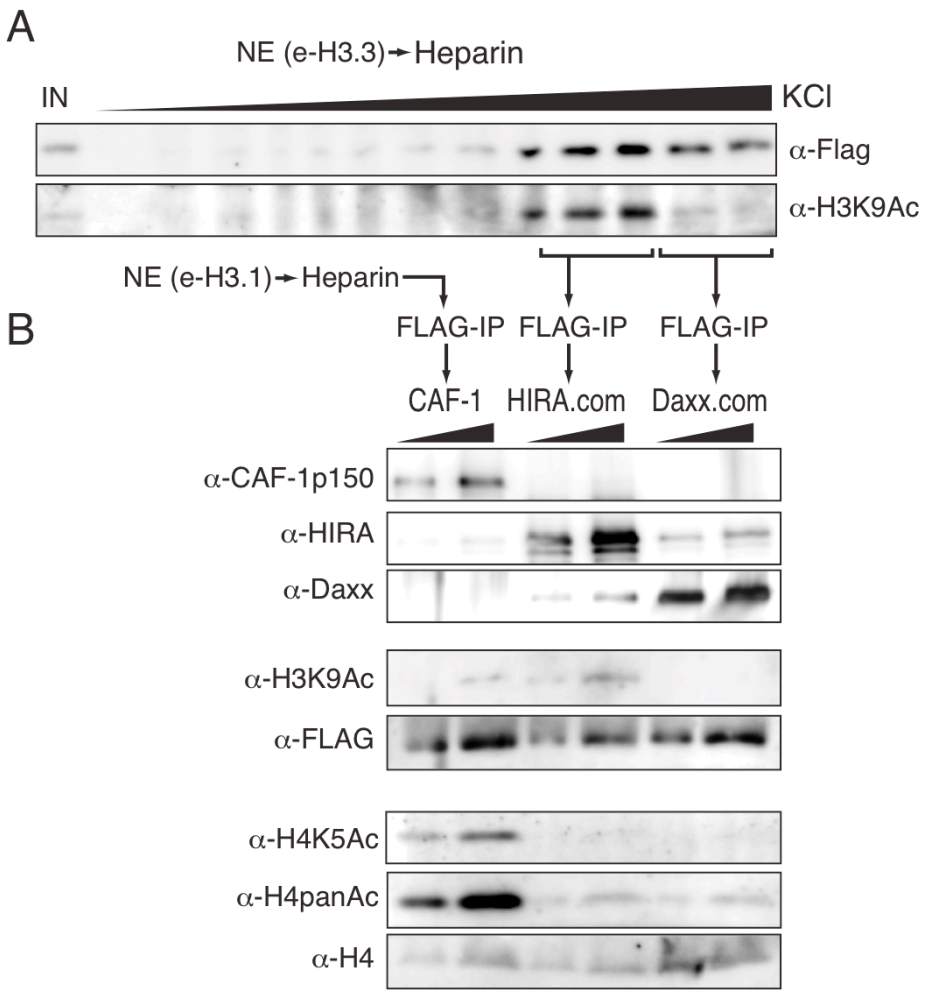



Figure 5.4: H3.3 predeposition complexes are associated with distinct histone posttranslational modifications.

(A) Western blot analysis of e-H3.3 lysine 9 acetylation (K9Ac) in the heparin salt gradient. K9Ac on e-H3.3 is enriched in fractions that correspond to the HIRA complex. (B) The CAF1, HIRA and Daxx complexes were analyzed for selected H3 and H4 posttranslational modifications. K9Ac is enriched on e-H3.3 within the HIRA complex but undetectable in the Daxx complex.

# 6 H3 variant-specific antibodies for genome-wide profiling of histone variants in primary cells or tissues

## 6.1 Introduction

The previous chapter has highlighted the diligence with which chromatin assembly is controlled by histone chaperones. It is easy to imagine how the cell could employ the site-specific installation of histone variants, carrying specific predeposition PTMs, to modulate chromatin states. However, such model relies on the assumption that the respective variant is sufficiently abundant at the site of incorporation to dominate the chromatin state. Global abundance of H3 variants is thought to vary greatly amongst cell types and age (Hake et al., 2006; Piña and Suau, 1987; Wu et al., 1982), likely impacting their enrichment at specific loci. It has been difficult however, to quantitatively assess the abundance of H3 variants in many cell types or primary tissues, as available methods to separate and detect variants, such as triton-acid urea (TAU) gel electrophoresis, HPLC and tandem mass spectrometry, require substantial amounts of protein and have limited resolution. While specific antibodies for immunological detection and quantification of a number of histone variants (i.e. CENP-A, H2A.Z, H2A.X) are well established, no such reagents were available for the non-centromeric H3 variants when I started my thesis research.

Once quantified globally, can the relative abundance of histone variants be determined genome-wide? While this seems a trivial question, current methodologies are intrinsically blind to the abundance of histone variants or PTMs relative to each other at a given genomic locus. The absolute abundance of

histone marks can be quantified by mass spectrometry (Villar-Garea et al., 2008) and next-generation ChIP-Seq technology allows assessment of the relative distribution of a given mark along a genomic region. However, neither method nor their combination can be used to infer what percentage of nucleosomes at a fixed locus have a variant identity or carry a certain mark. In particular comparison of ChIP-Seq read density merely yields enrichment scores between distinct loci that cannot be quantitatively compared across different antibodies. As introduced in Section 1.2.4, a wealth of studies has determined the H3.3 landscape in a number of different organisms and cell types (Chow et al., 2005; Daury et al., 2006; Janicki et al., 2004; Jin et al., 2009; Mito et al., 2007; Nakayama et al., 2007; Sutcliffe et al., 2009; Tamura et al., 2009; Wirbelauer et al., 2005). Given the lack of specific immunologic reagents discussed above, prerequisite of all these studies was the ability to genetically modify the cell line or organism to introduce an epitope-tagged version of H3.3. While most studies relied on the addition of an exogenous H3.3 locus, work by my colleague Dr. Aaron Goldberg allowed tagging of one of two endogenous H3.3 loci by zinc-finger mediated recombination in mouse embryonic stem cells (Goldberg et al., 2010). In either case the caveat remained that total H3.3 profiles were inferred from the distribution of the measurable, epitope-tagged subpopulation. Given the cell-cycle dependent chromatin dynamics, this inference is however only valid if expression patterns and cell-cycle dependent abundance between the tagged and untagged H3.3 population are sufficiently similar. Furthermore, no analogous genomic tagging of H3.1/2 species seems feasible in mammalian cells due to the number of H3.1/2 genes.

It was therefore desirable to develop immunologic reagents that would allow the specific detection and discrimination of H3.1/2 and H3.3. In collaboration with Trinette Chuang/EMD Millipore, I devised variant-specific peptide epitopes for the in-house generation of rabbit polyclonal antibodies at EMD Millipore. I further tested resulting sera, characterized candidate antibody fractions, leading to two new highly specific rabbit polyclonal products for H3.1/2 and H3.3.

## 6.2 Results

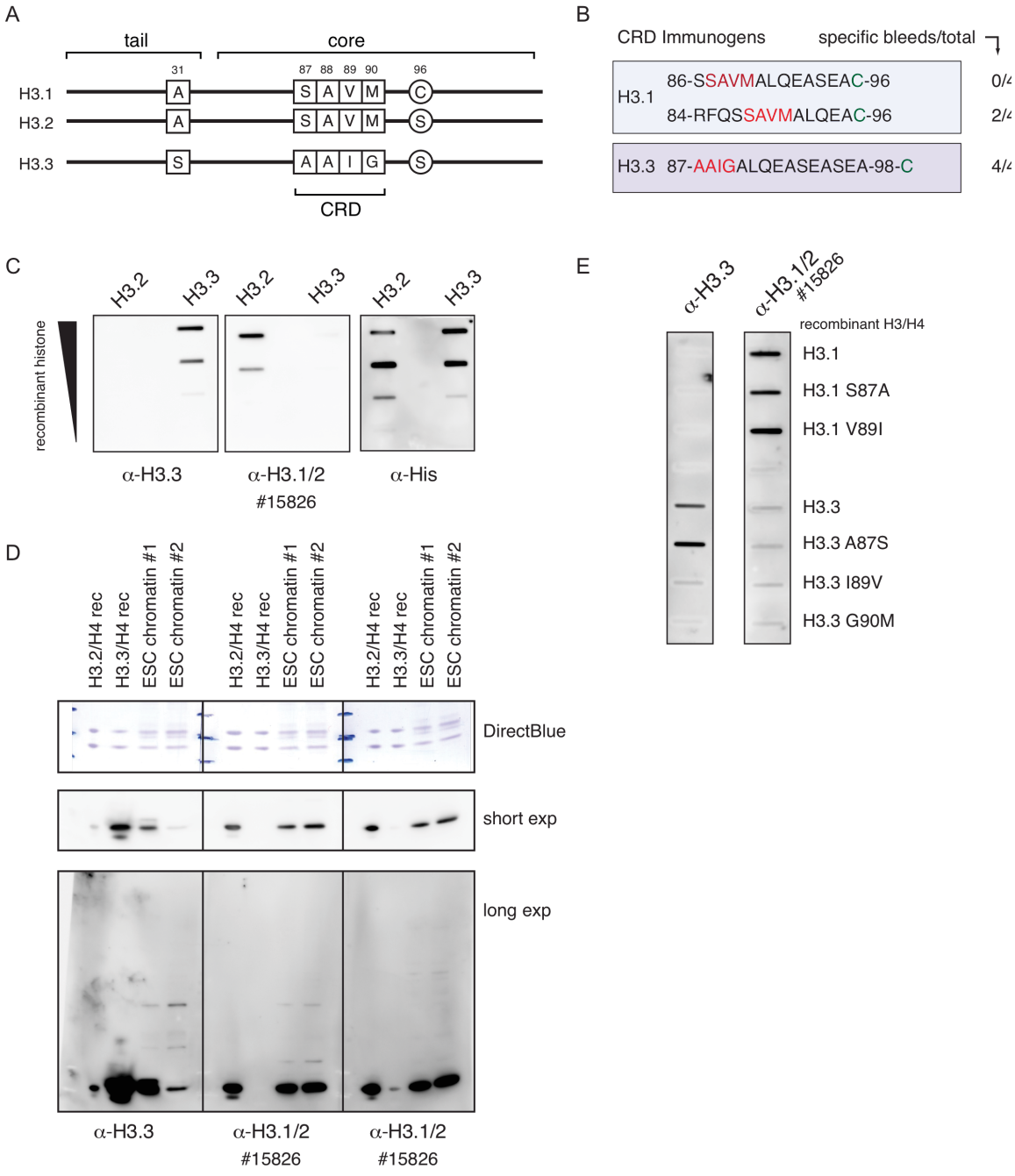
### 6.2.1 Generation of H3 variant-specific rabbit sera and purified antibodies

The amino acid sequences of major-type and replacement histone H3.3 (Figure 6.1A) are identical in all metazoans with very few exceptions. We intended to create antibodies that recognize a specific epitope in either variant. Two candidate regions for the design of immunogen peptides were considered, one in the tail (aa 31) and one around the chaperone recognition domain CRD (aa 87-90). The latter has three unique amino acids between the variants and therefore seemed more suited for the generation of a highly variant-specific antibody. As the only difference between mammalian H3.1 and H3.2 is located 6 residues downstream of the CRD, we expected to generate an antibody that would recognize both subtypes in addition to the major-type H3 in most other metazoans (identical to mammalian H3.2).

Peptides were kept short to avoid the presence of too identical sequences that would promote cross-reactivity. As the CRD is at the base of an ⟨-helix in H3, we included some of the downstream residues to promote a native conformation

of the peptide, which could facilitate the generation of antibodies that can immunoprecipitate native, folded protein (Figure 6.1A). Bleeds from all of the rabbits immunized with the H3.3 peptide tested positive against the epitope. However, none of the H3.1/2 bleeds showed any reactivity against the H3.1/2 immunizing peptide (Figure 6.1B, top). In a second trial, with a N-terminally elongated peptide, two bleeds contained H3.1/2-specific antibodies. I further tested positive bleeds on human and mouse whole cell lysates to determine the extent of cross-reactivity with other cellular proteins. The cleanest bleeds were further affinity-purified with the immunizing peptide and, if necessary, depleted with a corresponding peptide from the other variant. The purified antibodies showed no cross-reactivity against the respective other histone variant or other histones (two examples shown in Figure 6.1C,D). No major crossreacting bands were observed when tested on nuclear lysates (Figure 6.1D). Preliminary analysis against histone CRD point mutants further suggested that the H3.3 antibody recognizes the unique residue G90, whereas the tested H3.1/2 antibody has a more composite specificity against the H3.1/2 CRD (Figure 6.1E).

Figure 6.1: Development of H3 variant-specific, rabbit polyclonal antibodies.
(A) Schematic of the three non-centromeric H3 variants in mammals. Only regions that vary between those variants are shown. Despite the difference in position 96, major-type histones H3.1 and H3.2 are thought function in the same deposition pathways. Position 31 in the tail and the chaperone recognition domain (CRD) 87-90 distinguish them from the replacement variant H3.3. (B) Peptide immunogen design for the production of antibodies specific to the H3.1/2 and H3.3 CRD. Unique region is shown in red, the cysteine for adjuvant conjugation is shown in green. Number of rabbits that produced antibodies against the unique (red) region is given, as well as total number of immunized animals. (C, D) Specificity of purified antibodies against His-tagged full-length recombinant histones H3.1, H3.2 or H3.3 and point mutants thereof. (E) H3 variant-specific antibodies tested for specificity and cross-reactivity with other nuclear proteins. Western blots with the indicated purified antibodies against recombinant histone H3-H4 tetramers and two mouse embryonic stem cell (ESC) nuclei (control, treatment; kindly provided by L. Banaszynski) are shown.


A
tail
core
31
87 88 89 90
96
H3.1
H3.2
H3.3
CRD
B
CRD Immunogens
specific bleeds/total
H3.1
86-SSAVMALQEASEAC-96
84-RFQSSAVMALQEAC-96
H3.3 87-AAIGALQEASEASEA-98-C
C
H3.2
H3.3
recombinant histone
α-H3.3
α-H3.1/2
#15826
α-His
D
H3.2/H4 rec
H3.3/H4 rec
ESC chromatin #1
ESC chromatin #2
DirectBlue
short exp
long exp
α-H3.3
α-H3.1/2
#15826
E
α-H3.3
α-H3.1/2
#15826
recombinant H3/H4
H3.1
H3.1 S87A
H3.1 V89I
H3.3
H3.3 A87S
H3.3 I89V
H3.3 G90M

### 6.2.2 Antibody-based estimation of relative H3 abundance

As discussed above, determination of histone variant protein levels in primary tissues has been technically challenging. It is therefore unknown or uncertain for many interesting cell types, how much of the cellular histone H3 is contributed by which variant. In particular, post-mitotic cells such as neurons and germ cells do not use replication-dependent assembly and produce exclusively H3.3 (Wu et al., 1982). It has, for example, been observed that rat brains accumulate H3.3 (Piña and Suau, 1987). However, accurate determination was not possible, and despite advances in methodology this question has not been addressed, to date.

I therefore aimed to quantitatively assess H3.3 levels in a representative, post-mitotic cell type, the mature spermatozoan (in a collaboration with Dr. Shahin Rafii and Ying Liu). I established a co-western blot of H3.3 and H4, as well as H3.1/2 and H4 to serve as a per-lane loading control (Figure 6.2A). A comparison between human spermatozoa and a highly proliferative murine ESC line that contains only 10-20% H3.3 as determined by HPLC (L. Banaszynski, unpublished data) suggests an opposite ratio in the spermatozoan (Figure 6.2A). A more accurate determination required an external standard curve. To this end, I prepared recombinant H3.2-H4 and H3.3-H4 tetramers and also mixed them in 25%/50%/75% ratios. This panel should allow quantification of the western blot signal in relation to the stoichiometric H4 loading control. A preliminary quantification experiment is shown in Figure 6.2B, suggesting that H3.3 levels are in the range of 75% of total H3. More precise quantification was, however,

not possible, due to the weak staining of the H3.1/2 antibody and imperfect match in total histone between sample and controls.

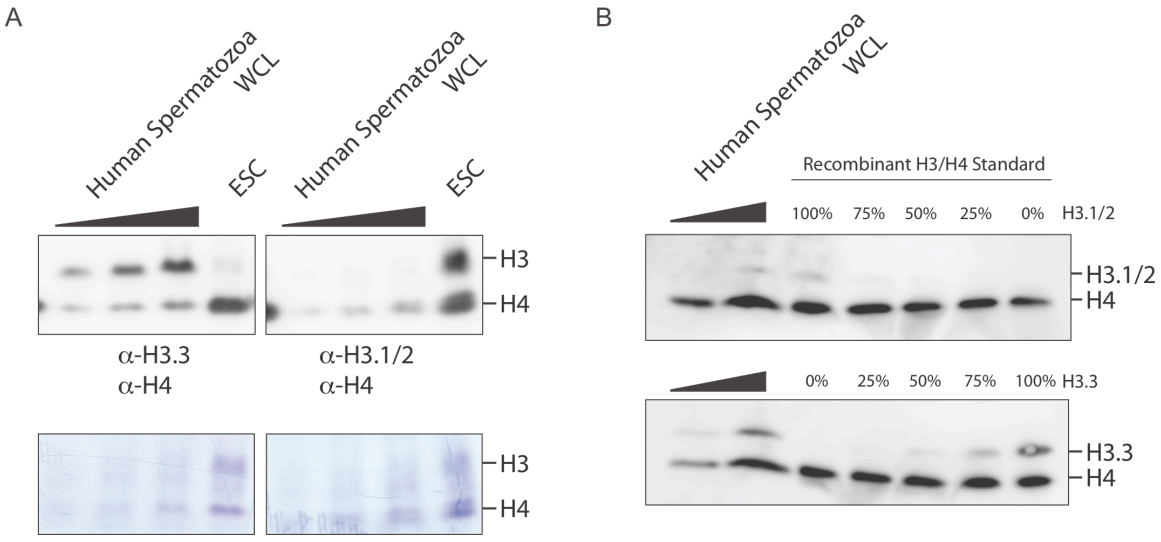



Figure 6.2: Estimating the relative abundance of H3 variants using western blot.
(A) Simultaneous detection of histone H3 variants and H4 as a loading control. A comparison of H3 variant content of a postmitotic (human spermatozoa) and fast dividing (mouse embryonic stem) cell type shows that the postmitotic cells contain much more H3.3 and less H3.1/2 than the embryonic stem cells. (B) A recombinant H3-H4 standard can be used to estimate the relative abundance of H3 variants. Recombinant H3-H4 tetramers were mixed at the indicated ratios. Comparison with the standard curve suggests that human spermatozoa contain more than 50% H3.3.

### 6.2.3 Chromatin immunoprecipitation with H3 variant-specific antibodies

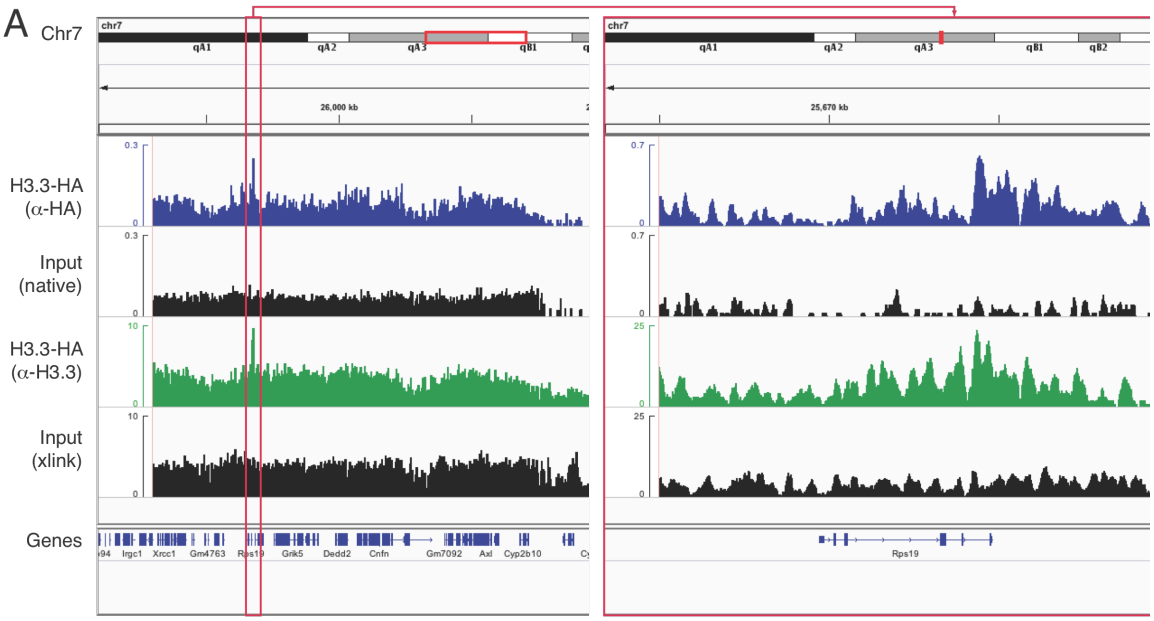
Given the excellent antibody specificity observed in western blotting, especially the absence of crossreacting nuclear proteins, the antibodies proved potential for a wider application range such as ChIP (Egelhofer et al., 2010). One concern for immunoprecipitation of chromatin was the accessibility of the antibody epitope, as the CRD is covered by DNA in the nucleosome structure.

Native conditions (mononucleosomes generated by MNase) could be too mild to unwind or relax DNA sufficiently to allow antibody binding. Crosslinking ChIP (formaldehyde treatment and subsequent sheering of the DNA by sonication) allows higher salt and detergent concentrations to reveal buried epitopes on the nucleosome but, vice versa, crosslinking could trap the epitope inside the core. Therefore, I initially carried out trials for both methods.

For assessment of the ⟨-H3.3 antibody, I was able to use our published H3.3 data from murine ESCs as a benchmark (Goldberg et al., 2010). As this study relied on a genetic knock-in of a HA-tag on the endogenous H3.3B locus, the distribution of the HA-tagged H3.3 was expected to closely reflect the distribution of total endogenous H3.3. I was therefore able to test and compare immunoprecipitation efficiency of the ⟨-H3.3 antibody in published conditions for native and crosslinking (xlink) ChIP (Goldberg et al., 2010) by probing known H3.3-enriched loci by qPCR. No enrichment was observed in native ChIP conditions as determined by qPCR, whereas ⟨-H3.3 ChIP after crosslinking and sonication consistently reproduced enrichment previously observed with ⟨-HA ChIP, albeit with ~50% loss in signal over background (data not shown). Therefore, the CRD epitope was likely buried under native conditions but accessible in the crosslinking protocol. The loss in signal over background in comparison to the HA signal could either be due to incomplete epitope accessibility, difference in affinity (HA antibodies bind extremely tightly), or it could reflect a biologically significant difference of looking at a tagged subpopulation versus the total H3.3.

Next, I sequenced a representative ⟨-H3.3 xlink ChIP sample using the Illumina GAII platform. Genome-wide correlations between ⟨-H3.3 and ⟨-HA H3.3 ChIP were excellent (data not shown), and patterns at individual loci matched well (Figure 6.3A). Characteristic patterns of H3.3 incorporation, such as the enrichment of H3.3 at transcription start sites (TSS) and towards the 5′ end of active genes were reproduced (Figure 6.3B, compare to (Goldberg et al., 2010)). As suggested by qPCR, the peak height over background was consistently reduced in the ⟨-H3.3 ChIP compared to native ⟨-HA H3.3 ChIP. Again, technical or biological effects could explain this phenomenon. However, comparison of native and xlink ⟨-HA H3.3 suggests that the crosslinking method adds some intrinsic noise (Goldberg et al., 2010). Nevertheless, the H3.3 antibody proved to reproduce data that previously has been dependent on the use of a tagged species and therefore represents a valuable tool to study H3.3 distribution in untagged cell lines, primary cell lines and tissues. The conservation of the epitope makes it universal for studying H3.3 deposition across metazoans.

Validation of the H3.1/2 antibody represented a more difficult task. Due to the repetitiveness and large number of major-type histone genes, no genomic-tagging strategy has been tackled to date. Instead, H3.1/2 has been either expressed exogenously from an inducible promoter (Mito et al., 2005), or the replication-independent H3.3B gene has been converted into a tagged H3.1/2 species (Goldberg et al., 2010). In both cases, the tagged H3.1/2 expression was not cell-cycle regulated and therefore did not mirror the endogenous H3.1/2.


A
Chr7
chr7
qA1
qA2
qA3
qB1
26,000 kb
25,670 kb
H3.3-HA
(α-HA)
Input
(native)
H3.3-HA
(α-H3.3)
Input
(xlink)
Genes
Irgc1
Xrcc1
Gm4763
Grik5
Dedd2
Cnfn
Gm7092
Axl
Cyp2b10
Rps19

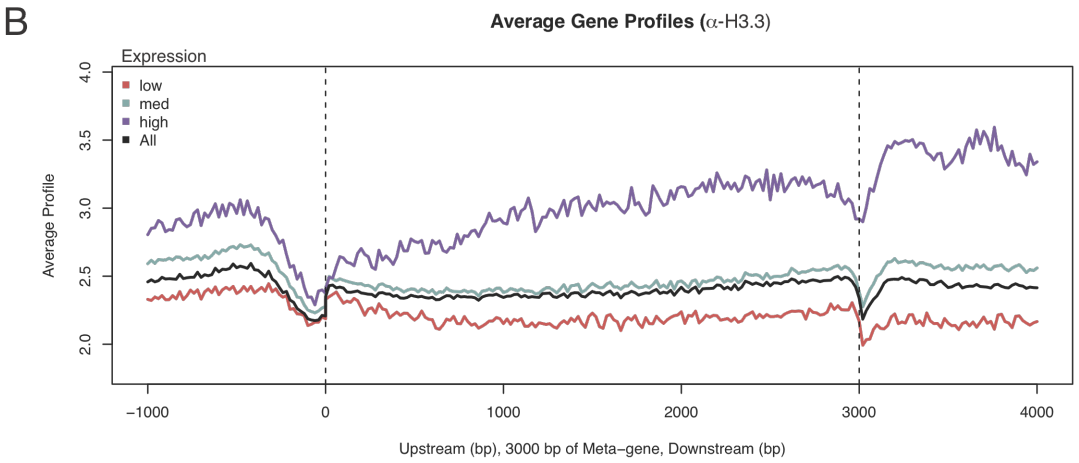

B
Average Gene Profiles (α-H3.3)
Expression
low
med
high
All
Average Profile
4.0
3.5
3.0
2.5
2.0
−1000
0
1000
2000
3000
4000
Upstream (bp), 3000 bp of Meta−gene, Downstream (bp)

Figure 6.3: ChIP-Seq of endogenous H3.3 in mouse embryonic stem cells with H3.3 specific antibody.

(A) Comparison of H3.3 profiles derived from ⟨-HA ChIP in H3.3-HA-knockin cell lines (Goldberg et al., 2010) and ⟨-H3.3 ChIP. ⟨-HA ChIP was carried out from MNased chromatin (native ChIP), ⟨-H3.3 ChIP from formaldehyde treated and sonicated chromatin (xlink ChIP) inputs. The chromosomal neighborhood of ribosomal protein gene Rps19 (left) and a close-up view on its gene locus (right) is shown. (B) ⟨-H3.3 antibody yields characteristic profiles of transcription-dependent genic enrichment comparable to ⟨-HA H3.3 ChIP (Goldberg et al., 2010). Profiles of high, medium, low expressed genes (as in (Goldberg et al., 2010)) were averaged over transcription start site, gene body and end. ChIP-Seq libraries in (A) were prepared by K.M. Noh.

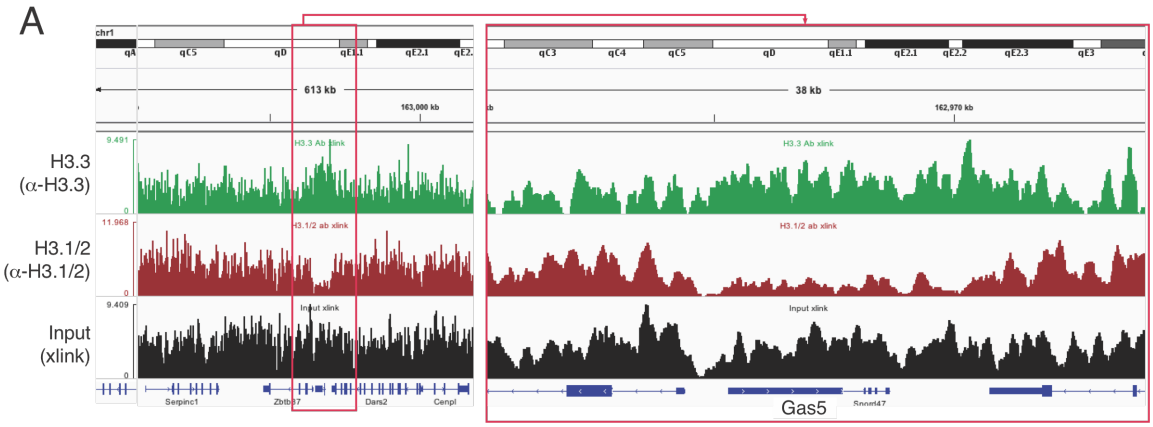
It could only be speculated that as a consequence of this exogenous expression, the tagged H3.1/2 species was found enriched at typical H3.3-rich regions, albeit much lower than H3.3 (Goldberg et al., 2010; Mito et al., 2005).

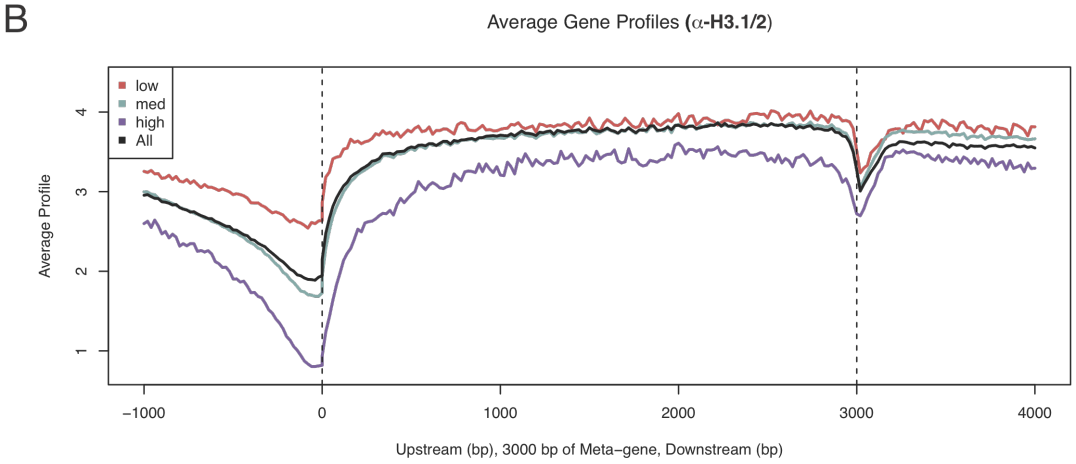
The intuitive expectation for H3.3-rich regions would instead be a depletion of H3.1/2. As discussed above, the extent of this depletion would depend on the relative abundance of H3.3 versus H3.1/2. Given the small fraction of H3.3 in murine ESCs (10-20%), it could be expected that the remaining 80-90% of H3.1/2 would be broadly distributed, mostly resembling input. It was therefore not possible to confirm H3.1/2 depletion *a priori* by qPCR measurements. The same epitope accessibility considerations applied to the H3.1/2 antibody as discussed for H3.3. Therefore, an xlink ⟨-H3.1/2 ChIP was carried out in parallel with a pan-H3 ChIP under the same conditions as a the ⟨-H3.3 ChIP. ChIP-Seq libraries were prepared for the input, H3.1/2 and pan-H3 ChIP (sequencing data for pan-H3 ChIP not yet available). Sequencing was carried out on a Illumina HiSeq platform that registers more than 100 million

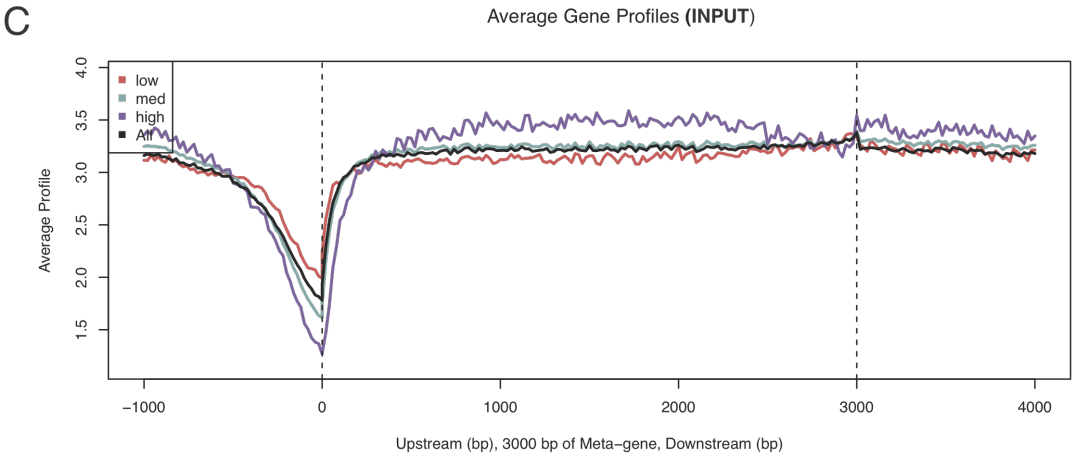
reads per lane and allows reasonably dense read coverage even of large mammalian genomes.

Inspection of the H3.1/2 ChIP profile by eye confirmed a broad overlap with Input signal. Intriguingly, regions of high H3.3 enrichment did show a discernable depletion in the H3.1/2 profile (Figure 6.4A). To acquire a more unbiased view of the H3.1/2 landscape, I averaged the reads over lists of high, medium and low expressed genes as above (Figure 6.3B). In contrast to the activity-dependent enrichment for H3.3, I was able to observe an activity-dependent depletion of H3.1/2 over the TSS as well as the whole gene body (Figure 6.4B). This was not a consequence of a decrease in DNA recovery from those regions as the input profiles were not dependent on gene activity (Figure 6.4C). In conclusion, these data suggest that the ⟨-H3.1/2 antibody indeed specifically ChIPs endogenous H3.1/2 and that, despite low global abundance of H3.3, some regions of H3.3 enrichment are effectively depleted in H3.1/2. Even in fast cycling stem cells, some genomic regions are therefore predominantly occupied with H3.3, suggesting that its incorporation could imply a mechanism to profoundly affect the local chromatin environment.

Figure 6.4: ChIP-Seq of endogenous H3.1/2 in mouse embryonic stem cells with H3.1/2 specific antibody.
(A) Example of a transcribed region (the ncRNA gene Gas5) where H3.3 enrichment coincides with a significant depletion of H3.1/2 over input. (B) ⟨-H3.1/2 antibody is depleted from highly transcribed gene bodies and promoters. Profiles of ⟨-H3.1/2 ChIP and crosslinked input at high, medium, low expressed genes were averaged over transcription start site, gene body and end. Compare depletion of H3.1/2 (top graph) with input (bottom graph) and H3.3 enrichment (Figure 6.4). ChIP was carried out by L. Banaszynski.


A
H3.3
(α-H3.3)
H3.1/2
(α-H3.1/2)
Input
(xlink)
613 kb
38 kb
Gas5


B
Average Gene Profiles (α-H3.1/2)
low
med
high
All
Average Profile
Upstream (bp), 3000 bp of Meta−gene, Downstream (bp)


C
Average Gene Profiles (INPUT)
low
med
high
All
Average Profile
Upstream (bp), 3000 bp of Meta−gene, Downstream (bp)

### 6.2.4 Estimation of locus-specific H3 variant content

Given the potential importance of H3.3 'hotspots' in the genome, I wanted to identify regions that contain predominantly or exclusively H3.3 in an unbiased manner. From Figure 6.4A (in particular from the left zoomed tracks) it is apparent that most H3.3 'peaks', even the highest ones, coincide with considerable H3.1/2 occupancy, therefore a typical peak-finding algorithm applied to the H3.3 track does not necessarily yield regions with concurrent H3.1/2 depletion. Instead, the fraction of the number of H3.3 reads of total H3 reads is an indicator for the relative enrichment of H3.3. To yield the total H3 reads, I calculated the sum of H3.3 and H3.1/2 reads (weighted 20:80 to reflect the global ration of the variants) over fixed 128bp windows. I then expressed H3.3 and H3.1/2 as a fraction of this total. The fraction value lies between 0 and 1, with 1 (=100%) meaning that *every* nucleosome at the given position has the variant identity.

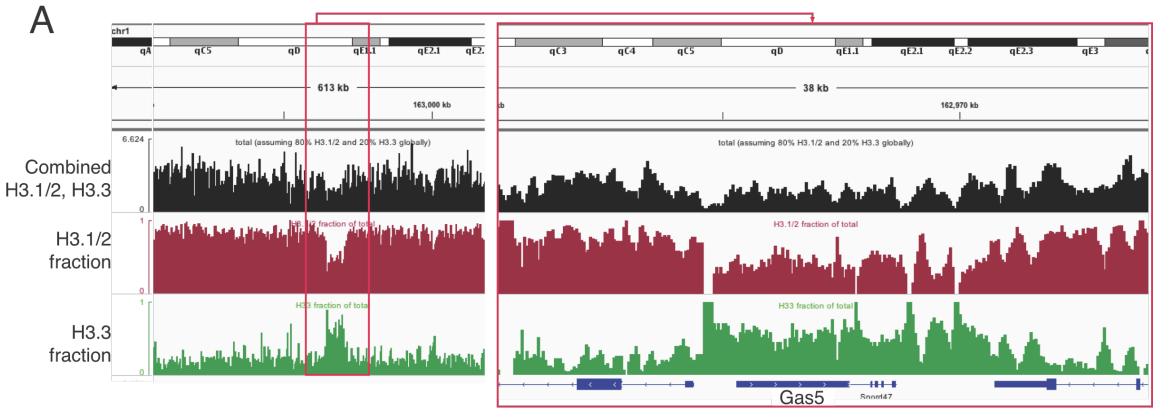
Continuous profiles were drawn from these 128 bp windows for the H3.3 and H3.1/2 fractions (as ratios, they add up to 1 at any given window) in Figure 6.5A. The top track represents the sum of H3.3 and H3.1/2 reads, i.e. the total *H3* occupancy. This track should match the pan-H3 ChIP track (not yet sequenced). As apparent from the H3.3 ratio profile, transcription start site (TSS) and transcription end site of *Gas5* feature regions of 100% H3.3 content (i.e. no H3.1/2 reads). Such H3.3 'hotspots' might appear by chance if the H3.1/2 track does not contain any reads in a 128 bp window due to insufficient sequencing depth and therefore can only be considered significant if they appear recurrently. Therefore, I averaged the H3.3 profile over the gene categories described above. An activity-

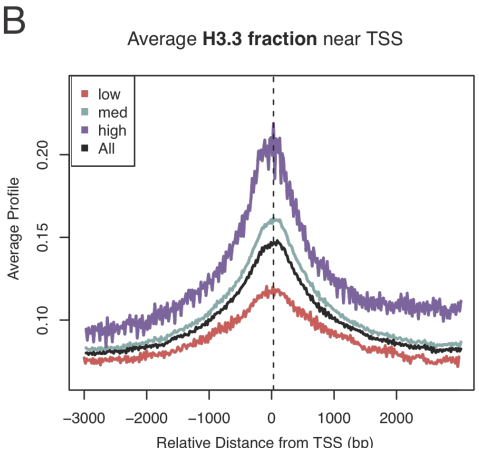
dependent H3.3-enrichment centered over the TSS was apparent from those averaged plots (Figure 6.5B). However, the average ratio for H3.3 did not exceed 0.25, therefore TSSs of the high expression gene set contain on average less than ¼ H3.3. H3.3 has been shown to be enriched in CpG islands (Goldberg et al., 2010; Mito et al., 2005), I therefore plotted the average H3.3 ratio over all annotated CpG islands (Figure 6.5C). Interestingly, while the nucleosome occupancy dips (CpG islands disfavor nucleosomes (Ramirez-Carrozzi et al., 2009)), the H3.3 ratio peaks at ~50%, suggesting that H3.3 is an abundant component of CpG islands in murine ESCs.

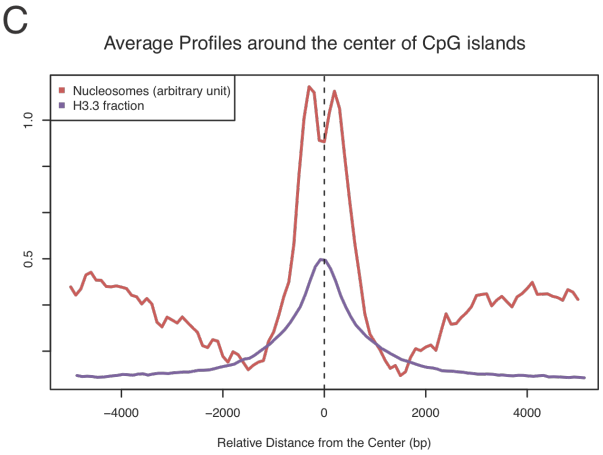
In a more unbiased approach, ~10 000 peaks were called from the H3.3 ratio track (Cistrome, (Liu et al., 2011)) with a significance cut-off of $p=10^{-6}$. The called peaks had an average H3.3 ratio of 0.9. Therefore, the corresponding regions were almost exclusively occupied by H3.3. 20% of these peaks fell in the TSS (Figure 6.5D, top), but not in regions immediately adjacent to the TSS (Figure 6.5D, bottom). Therefore, those H3.3 hotspots are highly localized to TSSs. About 50% of the H3.3 hotspots were found in intergenic regions and await further analysis. So far, we have only considered annotated regions of the genome. As H3.3 has been detected at repetitive regions, i.e. telomeres and centromeres (Drané et al., 2010; Goldberg et al., 2010), I extended my analysis of H3.1/2 and H3.3 density to a number of repetitive regions. I calculated the enrichment over input of repetitive sequences by counting the occurrence of ~20 bp motifs unique to the respective repeat element in ChIP as well as input samples. The resulting ratio (ChIP/input) is given in Figure 6.5E. Repetitive elements were moderately enriched in H3.1/2 and depleted in H3.3, except for telomeres, which were enriched in H3.3 (~3-fold over input). Immunofluorescence in murine ESC

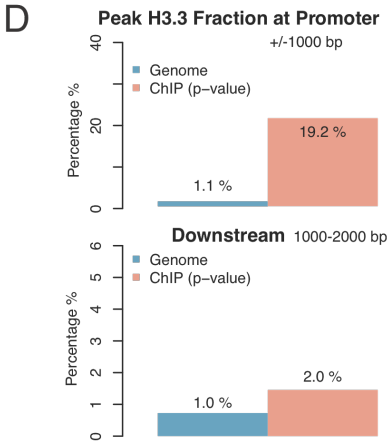
revealed H3.3 foci at telomeres, suggesting that H3.3 might constitute the major H3 variant in this region. However, the ChIP-Seq data does not support this notion, as no depletion of telomeric sequences in the H3.1/2 reads is observed. An H3.3 fraction calculated analogous to the method described above for the annotated regions of the genome yields a value of ~25% at telomeres.

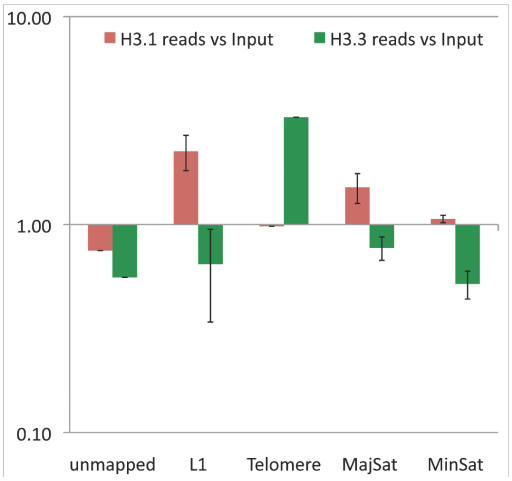
Figure 6.5: Genome-wide integration of H3.1/2 and H3.3 signals yields an estimate of relative H3.3 abundance at specific genomic loci.
(A) Calculation of a continuous fraction value across the genome for H3.1/2 and H3.3. Reads from H3.3 and H3.1/2 ChIP-Seq runs were binned and divided by the sum thereof to yield a fraction (0-1) value that corresponds to the relative abundance of the respective variant. (B) H3.3 fraction at transcription start site (TSS) corresponds to gene activity. Average H3.3 fraction does not exceed ¼ over ~10,000 most highly expressed genes. (C) Nucleosome occupancy (as measured by sequencing MNase-treated chromatin input) dips at the center of CpG islands, however the remaining nucleosomes are highly enriched in H3.3 (averaging to 50% of total H3). (D) The 10,000 most significant maxima in H3.3 ratio (>= 90%) are highly enriched over promoter regions but not the adjacent gene body. Percentage overlap with annotated promoters (top) and adjacent 1kb regions (bottom) is compared to their total occurrence in the annotated genome. (E) H3.3 enrichment but no H3.1 depletion from telomeric repeats. Enrichment of repetitive sequences in ⟨-H3.1/2 and ⟨-H3.3 ChIP. Occurrences of one ore more (where standard deviation is given as error bars) of ~20 bp sequences unique to the given repeat class counted in the unmapped portions of the respective ⟨-H3.1/2 and ⟨-H3.3 ChIP-Seq dataset. Ratio of reads counted in ChIP dataset versus input is shown on a log scale.


A
Combined H3.1/2, H3.3
H3.1/2 fraction
H3.3 fraction
613 kb
38 kb
Gas5


B
Average H3.3 fraction near TSS
low
med
high
All
Average Profile
Relative Distance from TSS (bp)


C
Average Profiles around the center of CpG islands
Nucleosomes (arbitrary unit)
H3.3 fraction
Relative Distance from the Center (bp)


D
Peak H3.3 Fraction at Promoter
+/-1000 bp
Genome
ChIP (p–value)
Percentage %
19.2 %
1.1 %
Downstream 1000-2000 bp
2.0 %
1.0 %


E
H3.1 reads vs Input
H3.3 reads vs Input
unmapped
L1
Telomere
MajSat
MinSat

## 6.3 Outlook

Based on my preliminary results described in this chapter, the H3.3 and H3.1/2 antibodies developed in collaboration with EMD/Millipore will provide powerful tools to profile histone variants in a multitude of cell types across metazoans. Their compatibility with standard formaldehyde-crosslinking ChIP protocols will allow their use for primary tissues and samples of small quantities.

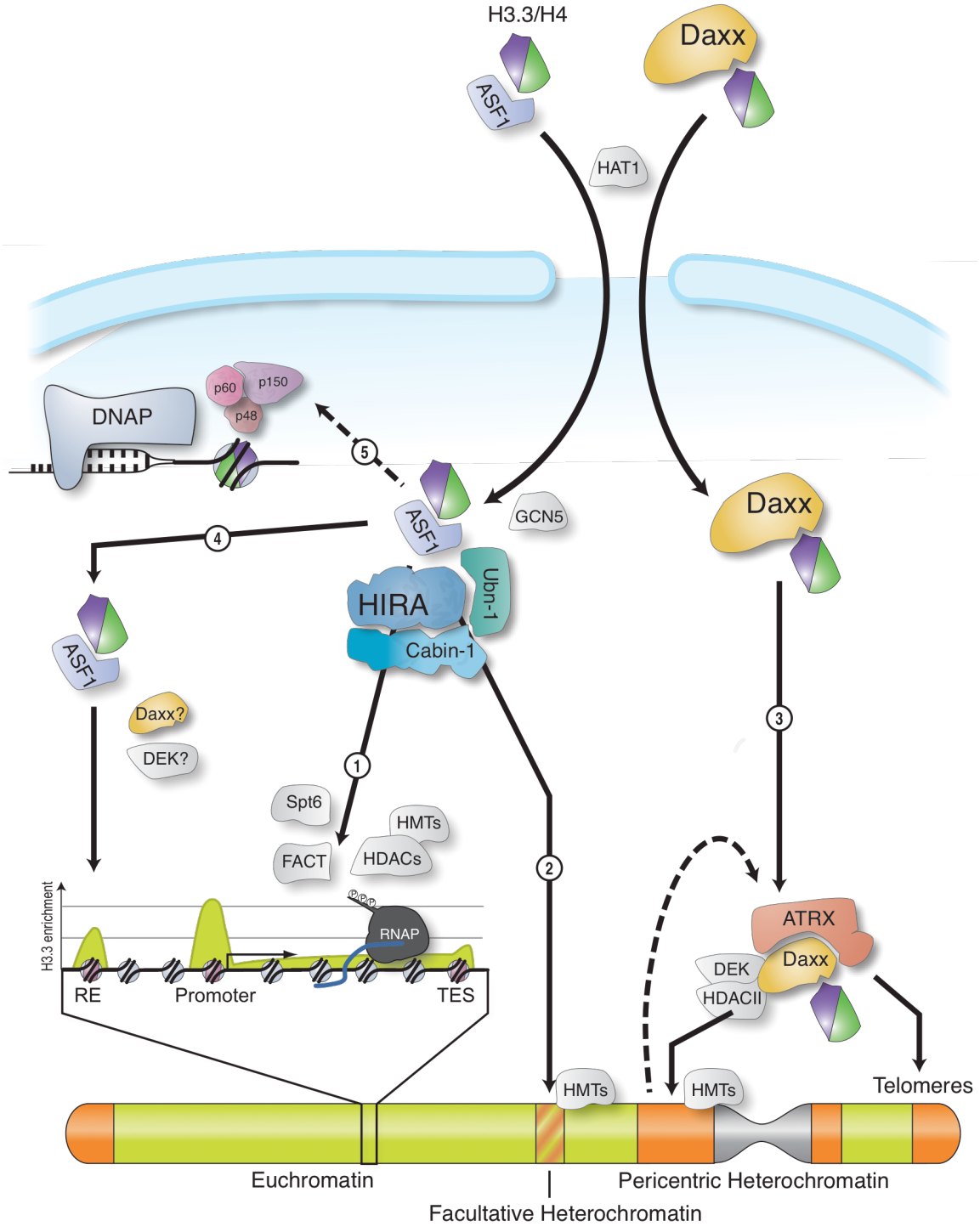
In the long term, it will be interesting to find out if H3.3-deposition mechanisms are conserved in all metazoans. So far, data from *Drosophila* and *C. elegans* and mouse suggest overlapping themes in H3.3 deposition (Ahmad and Henikoff, 2002c; Goldberg et al., 2010; Mito et al., 2005; Ooi et al., 2006), however the absence of a Daxx homolog in *C. elegans* also suggests significant differences. Going down the evolutionary tree, it will be particularly interesting if genome-wide patterns of H3 variants are conserved even to the simplest eukaryotes, implying a vital role for multicellular organisms.

# 7 Discussion

## 7.1 A model for the different pathways of H3.3 chromatin assembly

The biochemical characterization of distinct H3.3 deposition machineries discussed in Chapters 2 and 5 (as well as in (Goldberg et al., 2010)), allows us to construct a model for the journey of histones from synthesis to incorporation at specific genomic locations. Newly synthesized histones in the cytosol are likely immediately chaperoned by abundant 'general' histone chaperones such as ASF1 and/or NASP (Figure 7.1, top). Furthermore, H4 is diacetylated on K5 and K12. ASF1-bound H3.3-H4 dimers then serve as substrates for various assembly pathways. While a Daxx-H3.3-H4 complex has been reported in the cytosol (Drané et al., 2010), HIRA seems to be exclusively localized to the nucleus and has not been found in a cytosolic complex with H3.3-H4 (Lorain et al., 1998; Tagami et al., 2004). Although ASF1 is ubiquitously detectable with H3.3-H4 (Drané et al., 2010; Tagami et al., 2004), our biochemical data suggest that it is not a required component of HIRA and Daxx predeposition complexes (Figure 5.2C,D), (Lewis et al., 2010).

Figure 7.1: Predeposition complexes mediate H3.3 chromatin assembly.
A model for the flow of histones from their synthesis in the cytosol (top) to their target genomic regions in the nucleus (bottom). Newly synthesized histones H3.3-H4 are chaperoned by ASF1 and a cytosolic pool of Daxx. HAT1 acetylates H4K5/K12 in the cytosol. Predeposition complexes HIRA.com and Daxx.com chaperone distinct pools of H3.3-H4. (1) HIRA.com mediates deposition of H3.3 at promoters and transcribed regions together with RNA polymerase (RNAP)-associated FACT and Spt6 histone chaperones (2) HIRA.com deposits histones at facultative heterochromatin in senescent cells. Histone methyltransferases (HMTs) establish repressive chromatin state (3) Daxx.com cooperates with ATRX and possibly other factors in incorporating H3.3 into pericentric heterochromatin and telomeres. The chromatin-bound deposition complex might also contain the histone chaperone Dek and the histone deacetylase HDACII (4). An unknown factor delivers H3.3 to regulatory elements (RE). (5) In the absence of Daxx.com or HIRA.com, H3.3 associates with the replication-dependent CAF1 complex that mainly deposits histones after the DNA polymerase (DNAP) replication fork.


H3.3/H4
Daxx
ASF1
HAT1
DNAP
p60
p150
p48
5
ASF1
GCN5
Daxx
4
HIRA
Ubn-1
Cabin-1
ASF1
Daxx?
DEK?
3
1
Spt6
HMTs
FACT
HDACs
2
H3.3 enrichment
RNAP
RE
Promoter
TES
ATRX
Daxx
DEK
HDACII
HMTs
HMTs
Telomeres
Euchromatin
Pericentric Heterochromatin
Facultative Heterochromatin

### 7.1.1 The HIRA complex and histone deposition at transcribed gene bodies and promoters

HIRA histone chaperone activity has been well studied in Xenopus egg extracts, where it is necessary and sufficient for replication-independent chromatin assembly on exogenous DNA (Ray-Gallet et al., 2007; Ray-Gallet et al., 2002; Tagami et al., 2004). However, *in vivo* it accounts for only a subset of replication-independent chromatin assembly (Figure 7.1, pathway 1), namely within transcribed regions, at promoters and some regulatory elements (RE) (Goldberg et al., 2010).

Evidence from yeast showing that the Hir complex and Asf1p genetically and physically interact with the Set2 methyltransferase suggests that histone deposition might be functionally linked to H3K36 trimethylation. Histone deacetylases remove predeposition acetylation marks from nucleosomes after RNA polymerase passage.

H3.3 enrichment in gene bodies was consistently lost in HIRA-/- ESCs, but gene expression patterns and the transcription elongation mark H3K36me3 were only mildly affected (Goldberg et al., 2010). Several other histone chaperones including Spt6 and FACT have also been implicated in histone exchange during transcription. Thus, the purpose of the transcription-associated H3.3 deposition and possible compensatory mechanisms in the absence of HIRA have yet to be defined.

The homology of HIRA to the yeast Hir corepressor complex subunits Hir1/2p suggests that its histone deposition activity might suppress basal transcription (Anderson et al., 2009; Spector et al., 1997). Furthermore, it will be

necessary to define the roles of the remaining complex members Cabin-1 and Ubinuclein-1. Cabin-1 has recently been shown to act as a corepressor at p53 target genes and it will be interesting to see if this activity requires HIRA (Jang et al., 2009). Ubinuclein-1 cooperates with HIRA in a chromatin assembly pathway specific to the establishment of facultative heterochromatin domains in senescent cells (Figure 7.1, pathway 2) (Banumathy et al., 2009).

### 7.1.2 Daxx and ATRX assemble H3.3 chromatin at pericentric heterochromatin and telomeres

A considerable fraction of H3.3 is localized to telomeres in murine ESCs, and we have recently shown that this enrichment is dependent on the Daxx/ATRX complex, but not HIRA (Goldberg et al., 2010; Lewis et al., 2010). While Daxx readily assembles H3.3-H4 chromatin *in vitro*, ATRX is required *in vivo* to target Daxx and H3.3-H4 to the telomeres (Lewis et al., 2010), as well as to pericentric heterochromatin (Drané et al., 2010). A diffusible Daxx-H3.3-H4 complex therefore likely delivers histones to an ATRX-containing complex associated with telomeric chromatin (Figure 7.1, pathway 3) (Wong et al., 2010).

The chromatin-bound complex might furthermore contain the histone chaperone Dek and histone deacetylase HDACII as suggested by biochemical studies (Hollenbach et al., 2002), to prime a repressive chromatin state. This notion is supported by our finding that H3.3 does not carry the 'activating' K9Ac mark in the Daxx complex, but potentially acquires K9me1 or K9m2 methylation mediated by the G9a/GLP methyltransferase (Figure 5.4B, Figure 9.1). Recent data suggested two alternative, but not exclusive, models of how the ATRX/Daxx complex is recruited to telomeres. First, an intrinsic affinity of the

ATRX protein for a unconventional DNA hairpin structure, called G-quadruplex, that is prone to occur at repetitive elements, in particular the G-rich telomeric repeat (Law et al., 2010); second, the recognition of the heterochromatic H3K9me3 mark that decorates most repetitive elements in heterochromatin, but also euchromatin, through the ATRX ADD domain (Dhayalan et al., 2011; Eustermann et al., 2011; Iwase et al., 2011). Failure of H3K9me3 establishment could therefore lead to similar consequences as the genetic knockouts of ATRX and Daxx that we have been studying. Importantly, through stimulation of H3K9me3 at sites of H3.3 incorporation and subsequent recognition of the mark by the ATRX ADD domain (Figure 7.2), this could establish a feedback mechanism in a similar manner to the epigenetic propagation of centromeres discussed in Section 3.1.1. Such a mechanism might contribute to maintenance of heterochromatin as suggested by the increase in TERRA transcription upon ATRX deletion (Goldberg et al., 2010) that could ultimately lead to genomic instability exemplified by the occurrence of DNA damage and γ-H2A.X repair foci (Wong et al., 2010)) in the absence of ATRX. I will detail possible connections between epigenetic instability and genomic instability in Section 7.4. Notably, Daxx is a highly conserved histone chaperone that appeared with the emergence of metazoans. Unlike the H3.3-chaperone activity of HIRA that is shared in yeast, Daxx therefore likely evolved to diversify the functionality of H3.3 within the more complex chromatin structure of higher eukaryotes. A specific example is the asymmetric incorporation of histone H3.3 between male and female pronucleus in the fertilized egg (Loppin et al., 2005; Torres-Padilla et al., 2006). Specific incorporation of H3.3 into the paternal genome is required for

establishing heterochromatin in the decondensed sperm nucleus (Santenard et al., 2010). It is tempting to speculate that Daxx has taken on this and other metazoan-specific challenges.

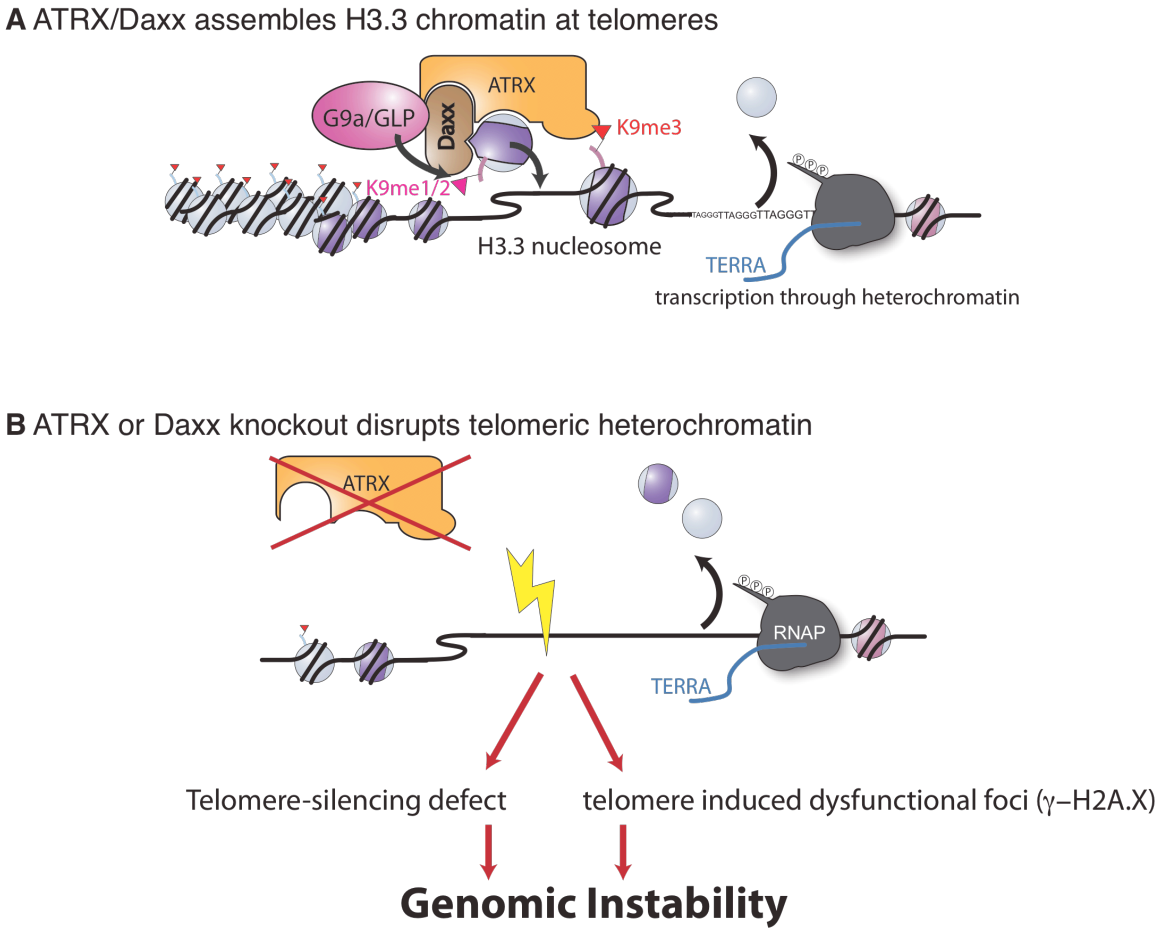



Figure 7.2: Model for ATRX/Daxx function at telomeres.
(A) The ATRX/Daxx complex is recruited to heterochromatin via H3K9me3 recognition by the ADD domain of ATRX or through recognition of DNA structures that occur in nucleosome-depleted repetitive elements. Deposited H3.3 nucleosomes prime H3K9me3 through pre- or codeposition methylation of H3K9 to the mono- or dimethylated state. (B) Consequence of Daxx or ATRX knockout include a desilencing of transcripts from repetitive elements, as well as genomic lesions, both associated with genomic instability.

### 7.1.3 H3.3 deposition at regulatory elements

Many DNA sequences with regulatory functions, including transcription factor binding sites (TFBSs), insulator elements and origins of replication (ORFs), are characterized by fast histone exchange dynamics. These epigenetic regulatory elements acquire newly synthesized histones and are constitutively enriched in H3.3 (Deal et al., 2010; Goldberg et al., 2010; Jin et al., 2009; Mito et al., 2007; Nakayama et al., 2007). Neither HIRA nor ATRX deletion significantly altered the global incorporation of H3.3 at these sites (Goldberg et al., 2010), suggesting that a yet unknown assembly pathway is used. While 'general' histone chaperones such as ASF1 or NASP could provide the histones for ATP-dependent chromatin remodelers that occupy these locations, it is important to note that this pathway is also specific to H3.3 (Goldberg et al., 2010). We therefore speculate that either Daxx, Dek or yet unknown factors mediate the deposition of H3.3 at regulatory elements (Figure 7.1, pathway 4). Dek has recently been found to cooperate with ATP-dependent chromatin remodelers and act as a coactivator for nuclear receptors (Sawatsubashi et al., 2010).

Intriguingly, H3K4 monomethylation is another hallmark of regulatory elements and it correlates well with exchange hotspots and H3.3 enrichment in genome wide analyses. It remains elusive which histone methyltransferase activity is responsible for establishing the H3K4me1 mark, but we speculate that methylation and histone exchange are functionally linked at these genomic regions.

### 7.1.4 Alternate pathways for H3.3 deposition

As apparent from the imperfect discrimination that can be observed in pure *in vitro* assays with variant-specific chaperones (Drané et al., 2010; Lewis et al., 2010), only the interplay and competition of variant-specific histone chaperones can explain the exquisite specificity of H3.3 incorporation observed *in vivo*. Whereas in unperturbed systems, H3.1 and H3.3 associate exclusively with their cognate chaperone systems (CAF1, HIRA and Daxx, respectively), we did observe CAF1 copurifying with e-H3.3 in the absence of Daxx and HIRA (Figure 7.1, pathway 5) (Lewis et. al. 2010, Drane et. al., 2010). A more comprehensive set of histone mutations along the lines of Figure 3.5 and molecular structures will help to define the overlapping or distinct binding surfaces used by various pathways. Furthermore, there is evidence that low levels of H3.3 are broadly incorporated into chromosomes by the replication-dependent machinery in *Drosophila* cells (Ahmad and Henikoff, 2002c; Schwartz and Ahmad, 2005). Therefore, similar to the single yeast H3.3-like histone, H3.3 might represent a universal variant for replication-dependent and independent deposition pathways that is simply outnumbered by the large pool of H3.1/2 present during S phase. Additionally, studies in flies suggest that H3.1/2 can take over some, but not all, functions of H3.3 (Hödl and Basler, 2009; Sakai et al., 2009).

### 7.1.5 Conclusion

As illustrated by the replacement histone H3.3 and the centromeric H3, a rapidly emerging literature points to remarkable complexity for the use of relatively minor histone variants in punctuating epigenomes in organisms ranging from yeast to man. Multiple predeposition and assembly systems are called upon to deposit histone variants in select genomic locations. While these variants often differ in only a small number of amino acids, mounting evidence suggests that, once assembled, they carry out specialized functions that remain poorly defined. It has been proposed that selective use of histone variants may contribute to a 'nucleosome code' (Bernstein and Hake, 2006), providing additional variation to that provided by chromatin remodeling and posttranslational modifications. We look forward to future experiments identifying and characterizing the predeposition complexes that engage these variants, giving the field more mechanistic insights into how these variants make their complicated journeys from synthesis to distinct chromatin functions.

## 7.2 Nucleosome dynamics, histone variants and cellular memory

> *"Every cell contains within its chromosomes all the information necessary for the manufacture of every protein in the entire organism. […] Each could be doing what the other does, but it doesn't. What determines what information gets used? What switches on certain genes in the DNA? What is the difference in the chromosomes of those two cells? That's what we want to find out." - Vincent G Allfrey, 1980 (Bardossi, 1980)*

How do all the pathways of replication-independent chromatin assembly outlined above integrate with other mechanisms known to modulate chromatin states? Prerequisite for the incorporation of new histones is the eviction of old ones. The ATRX/Daxx complex has chromatin remodeling activity but it has not been shown yet that the complex can 'switch' nucleosomes as has been shown for the SWR1 remodeler that exchanges canonical H2A for the histone variant H2A.Z (Luk et al., 2010). Rather, a multitude of DNA transactions directly or indirectly create dynamic chromatin. Transcription destabilizes or evicts nucleosomes through DNA torsion stress and transcription factors recruit ATP-dependent chromatin remodelers to promoters or regulatory elements. Therefore one might speculate that H3.3 deposition is merely a consequence of dynamic chromatin (Ahmad and Henikoff, 2002a; Deal et al., 2010; Henikoff and Shilatifard, 2011). As part of dynamic chromatin, H3.3, like typically 'activating' histone PTMs such as H3 and H4 acetylation, could however contribute to a cellular memory that would predispose transcriptionally active regions for a new transcription cycle after replication (Henikoff et al., 2004). However it will be important to dissect if, in replenishing nucleosome free regions, H3.3 acts predominantly as a substrate to the assembly pathway or in addition can adopt a unique role once incorporated into chromatin. An argument has been made that H3.3 could create intrinsically unstable nucleosomes (Jin and Felsenfeld, 2007; Jin et al., 2009), however *in vitro* data does not support such a model (Flaus et al., 2004; Thakar et al., 2009). Instead, the observed instability could reflect the properties of dynamic chromatin that, in part, are determined by DNA sequence (Deal et al., 2010; Ramirez-Carrozzi et al., 2009). It should be pointed out, however, that the unique serine in position 31 of H3.3, phosphorylated or not,

could affect a number of 'readers' and 'writers' in the vicinity (such as for H3K27 or H3K36).

It has been argued that chromatin states in general do not determine gene expression outcome and solely reflect the consequence of DNA-templated processes (Bryant et al., 2008 2011). In particular, when chromatin states are implicated in the maintenance of cell identity, it has to be considered if cell-type specific 'master' transcription factors ultimately set the cell identity. The reprogramming of somatic cells to pluripotent stem cells by four core stem cell transcription factors is a pivotal example of this concept (Takahashi and Yamanaka, 2006). An important consideration in this respect is how chromatin affects transcription factor binding. There is a conundrum in transcription factor biology – DNA sequence cannot accurately predict the occupancy of a transcription factor binding site with its cognate transcription factor, owing to the fact that there are orders of magnitude more nonspecific binding sites in the genome than the estimated number of transcription factor molecules (Stratmann and Schibler, 2011). A mechanism to 'assist' a transcription factor to find its cell type-specific binding sites would solve this problem, and only recently such a mechanism has been shown to rely on chromatin changes (Voss et al., 2011). In this particular example, the hyperdynamic chromatin state around a transcription factor binding site, once primed through a binding event, maintains accessibility to subsequent binding events of the same or different transcription factors. It is therefore warranted to postulate that some properties of chromatin, which will likely include histone PTMs and histone variants or a combination thereof, contribute to cell identity through an epigenetic propagation mechanism (Moazed, 2011).

## 7.3 Defects in H3.3 deposition pathways are associated with a wide spectrum of human diseases

Given the dynamic state of chromatin, replication-independent chromatin assembly pathways might be expected to have an important role in faithful maintenance of gene expression and developmental programs (Ahmad and Henikoff, 2002a). Indeed, mutations or deletions of histone chaperones are associated with a number of diverse developmental disorders as detailed below.

The *HIRA* gene falls in the 22q11.21 locus deleted in the DiGeorge/CATCH22 syndrome, a craniofacial developmental disorder (Lamour et al., 1995). As a number of candidate genes in the locus might contribute to the syndrome's etiology, it has not been possible to dissect individual gene functions. However, *HIRA* deletion in mice is embryonic lethal with early defects in mesodermal and cardiac development, establishing potential links with the defects observed in CATCH22 patients (Roberts et al., 2002).

Familial mutations in the *ATRX* gene give rise to a blood disease and neurological disorder called X-linked ⟨-thalassemia mental retardation (ATR-X) syndrome (Gibbons, 2006; Picketts et al., 1996; Xue et al., 2003). The causal genetic alterations ultimately affect the expression of genes (such as the ⟨-globin locus) that are megabases apart from the mapped locus or even on different chromosomes. ATRX therefore acts as an epigenetic modifier of remotely located factors that control blood cell and neuronal development, plasticity and excitability. Strikingly, conditional deletion of the *ATRX* gene in mice impairs the

development in a similar fashion (Ritchie et al., 2008; Seah et al., 2008). ATRX has been identified as a Snf2-family ATP-dependent chromatin remodeler that localizes to pericentromeric and telomeric heterochromatin (McDowell et al., 1999; Picketts et al., 1996). Patient mutations are found both in its catalytic core as well as the so-called ADD domain that directly binds the heterochromatic trimethylated H3K9me3 (Argentaro et al., 2007; Dhayalan et al., 2011; Eustermann et al., 2011; Gibbons, 2006; Iwase et al., 2011; Ooi et al., 2007). Mutations in the ADD H3K9me3 recognition domain are associated with impaired DNA cytosine-methylation in heterochromatin but also with gene expression changes in the vicinity of heterochromatin domains (Gibbons et al., 2000). Considering its ubiquitous expression and broad nuclear distribution along heterochromatin, it is surprising that major symptoms of familiar *ATRX* mutations are confined to the brain and the hematopoietic system. It is interesting to speculate that in the absence of replication-dependent mechanisms in postmitotic neurons, replication-independent chromatin assembly becomes crucial for maintaining chromatin structure and in particular the dense packaging of heterochromatin (Ahmad and Henikoff, 2002a).

Daxx has been identified as a mandatory subunit of the ATRX chromatin remodeling complex (Drané et al., 2010; Lewis et al., 2010; Xue et al., 2003) and therefore likely plays an essential role in ATRX biology. Daxx and ATRX knockout mice show phenotypically similar embryonic lethality (Ishov et al., 2004; Michaelson et al., 1999; Xue et al., 2003). Our recent findings that Daxx links ATRX remodeling activity to the deposition of the histone variant H3.3 first shed light into the mechanism of the ATRX/Daxx complex as an epigenetic modifier (Goldberg et al., 2010; Lewis et al., 2010). No Daxx or histone H3.3 mutations

have been identified thus far in ATR-X patients, however at least one known mutation in ATRX might compromise its ability to bind Daxx (Fichera et al., 1998; Gibbons et al., 2008; Tang et al., 2004). Germline polymorphisms likely exist in the *DAXX* gene, as well, and future targeted studies might reveal mutations that associate with syndromes similar to ATR-X.

Daxx has been originally described as a component of the proapoptotic death receptor pathway (Yang et al., 1997) and since then has been implicated in a variety of pro- and antiapoptotic functions in the nucleus and cytoplasm (Salomoni and Khelifi, 2006). From the variety of functions and molecular interactions associated with the Daxx proteins it is apparent that future research on the role of Daxx will require careful testing and a mechanistic dissection. The functions of individual domains of the Daxx protein will have to be dissected to learn if phenotypes can be attributed to a specific domain or rely on cooperation of multiple domains. In particular, it will be interesting to see if the apoptotic regulation is mechanistically related to the histone H3.3 chaperone function.

## 7.4 A sound genome in sound chromatin – Epigenetic stability is important for genomic stability.

Constitutive heterochromatin encompasses all major repetitive elements in eukaryotic genomes, in particular the abundant centromeric α-satellite and telomeric repeat elements. Aberrant transcription in repeat regions is thought to cause genome instability through the reactivation of retrotransposons, but likely also through to other less well-understood mechanisms. Therefore, defects in heterochromatin silencing have been attributed to genomic instability (Fry et al.,

2006; Mostoslavsky et al., 2006; Oberdoerffer and Sinclair, 2007). However, direct links between such epigenetic defects and human cancers have been rare. This has dramatically changed in very recent years, namely through novel technologies, in particular exome sequencing – the targeted sequencing of all protein coding regions in the human genome. The systematic discovery of somatic mutations has redefined a number of well-known chromatin components as potent tumor suppressors in specific types of tumors (Dalgliesh et al., 2010; Jones et al., 2010; Parsons et al., 2011; van Haaften et al., 2009; Varela et al., 2011; Wiegand et al., 2010). Inactivating mutations in those epigenetic factors seem to occur frequently early in neoplasia, facilitating subsequent genomic alterations. We have termed these early, highly tissue-specific, epigenetic tumor suppressors 'backseat drivers' of cancer, as opposed to the well-studied oncogenes 'drivers' that are activated later and common to many diverse types of tumors (Elsässer et al., 2011)

One particularly intriguing study used exome sequencing to explore the genetic etiology of pancreatic neuroendocrine tumors (PanNET) - a class of tumors that frequently arise from islet cells of the pancreas (Jiao et al., 2011). In addition to well-studied tumor suppressors *PTEN* and *TSC2* in the mTOR pathway, the authors found that more than 60% of PanNETs are mutated in either *MEN1* (a histone methyltransferase), *ATRX* or *DAXX*. Simultaneous mutations of *ATRX* and *DAXX* were never observed in the same tumor, suggesting that inactivation of one was sufficient to drive those tumors. Subsequent work by the Meeker and Papadopoulos groups found that *ATRX*/*DAXX* mutations correlated with vastly elongated telomeres, indicative of an activation of the alternative telomere lengthening (ALT) pathway (Heaphy et

al., 2011), affirming our speculation that ATRX and Daxx are essential for the maintenance of telomere integrity (see Section 7.1.2). Interestingly, other recent reports highlight a role for heterochromatin-mediated silencing of transcription through repetitive elements in tumor suppression: aberrant overexpression of centromeric α-satellite repeats and LINE-1 retrotransposons are found in a variety of epithelial cancers (Ting et al., 2011). Most surprisingly, the potent tumor suppressor and E3 ubiquitin ligase BRCA1, well studied for its detrimental heritable mutations leading to breast and ovarian cancer, has now been shown to organize heterochromatin (Zhu et al., 2011).

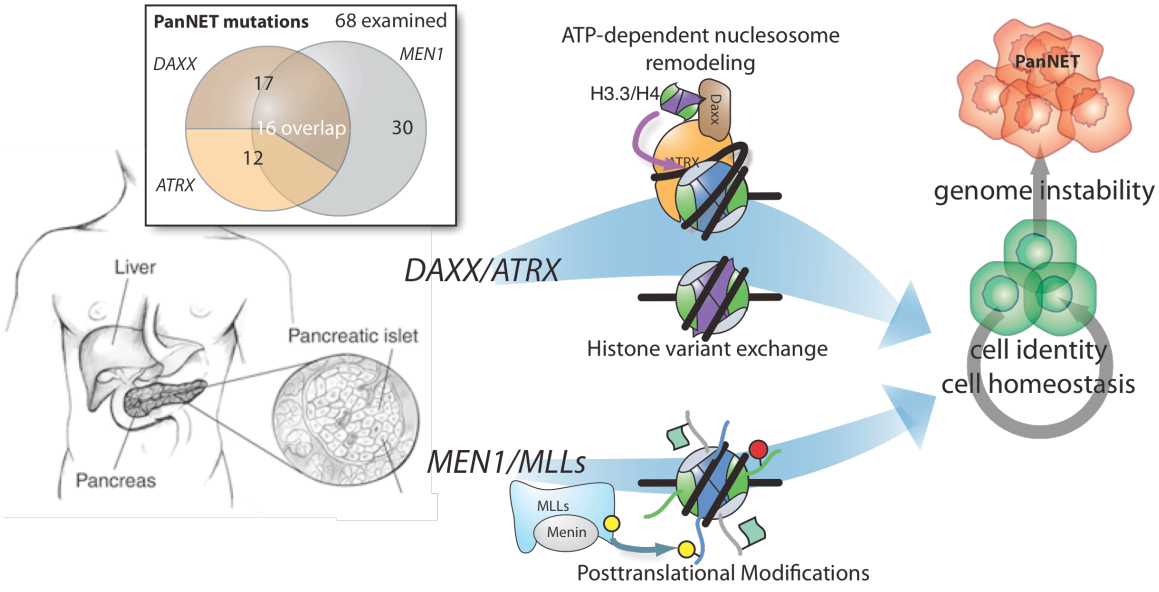



Figure 7.3: Chromatin regulators mutated in pancreatic neuroendocrine tumors.
Epigenetic regulators Menin and ATRX-Daxx promote genome integrity and maintain cell identity through the modification of chromatin structure by chromatin remodeling, variant histone nucleosome assembly and histone posttranslational modifications. Loss of these pathways may play a critical role in the development of neuroendocrine tumors arising from the pancreatic islet.

*BRCA1* mutations lead to desilencing of centromeric satellite repeats, likely through a defect in histone H2A ubiquitinylation. Cumulative evidence from these diverse studies suggests that proper maintenance of chromatin states is absolutely essential for genome maintenance. In light of these discoveries, I am looking forward to future studies on how the H3.3-deposition activity of ATRX/Daxx cooperates with other well-studied components of heterochromatin such as the H3K9 methyltransferase SUVH39 and the H3K9me3 binder HP1. The study of PanNET tumors will yield unprecedented insight into the *in vivo* function of the ATRX/Daxx complex. Molecular and structural details of these proteins will help to mechanistically explain these findings and guide further experiments *in vivo*.

# 8 Methods

## 8.1 Molecular Biology

### 8.1.1 Ligation-independent cloning for *E. coli* expression

For generation of all expression constructs, ligation-independent cloning (LIC) strategies were used. LIC vectors for N-terminal GST-$His_6$-TEV (pRUTH2) and $His_6$-MBP-TEV (pRUTH3) fusions were derived from pMCSG7 (Stols et al., 2007) and novel LIC site for an N-terminal $His_6$-TEV tag was designed by Dr. Alexander Ruthenburg in the pET-based pRUTH5 vector (Figure 8.1). 1 $\mu$g vector was linearized at its unique LIC restriction site (SspI for pRUTH2/3, BseRI for pRUTH5/pRSFDuet-LIC) and purified by agarose gel electrophoresis and gel extraction (Qiagen). The insert was created by amplifying the coding region of interest with primers containing the LIC overhangs in a 20-50 $\mu$L PCR reaction (NEB Phusion), followed by a PCR purification column (Qiagen). Long 5' overhangs for LIC were created by limited 3' resection with T4 Polymerase in the following reaction:

T4 Pol digest (*20-30 min at RT, 10 min at 80°C*)
~28 $\mu$l vector or insert (column-purified)
0.8 $\mu$l of 100 mM dNTP as follows
pRUTH2/3 vector: dGTP; PCR insert: dCTP
pRUTH5/pRSF vector: dCTP; PCR insert: dGTP
0.3 $\mu$l BSA (NEB)
3.3 $\mu$l T4 Pol buffer (NEB 2)
1 $\mu$l of T4 DNA Polymerase (NEB).

For transformation, 0.3-1 $\mu$L vector was mixed with 1-3 $\mu$L insert (both components at RT) and incubated at RT for 10-20 min. 1 $\mu$L of 50 mM EDTA was added and after 5 min, the mix was transferred onto 25-50 $\mu$L competent cells (Invitrogen BL21(DE3)Star or Rosetta(DE3)pLysS).

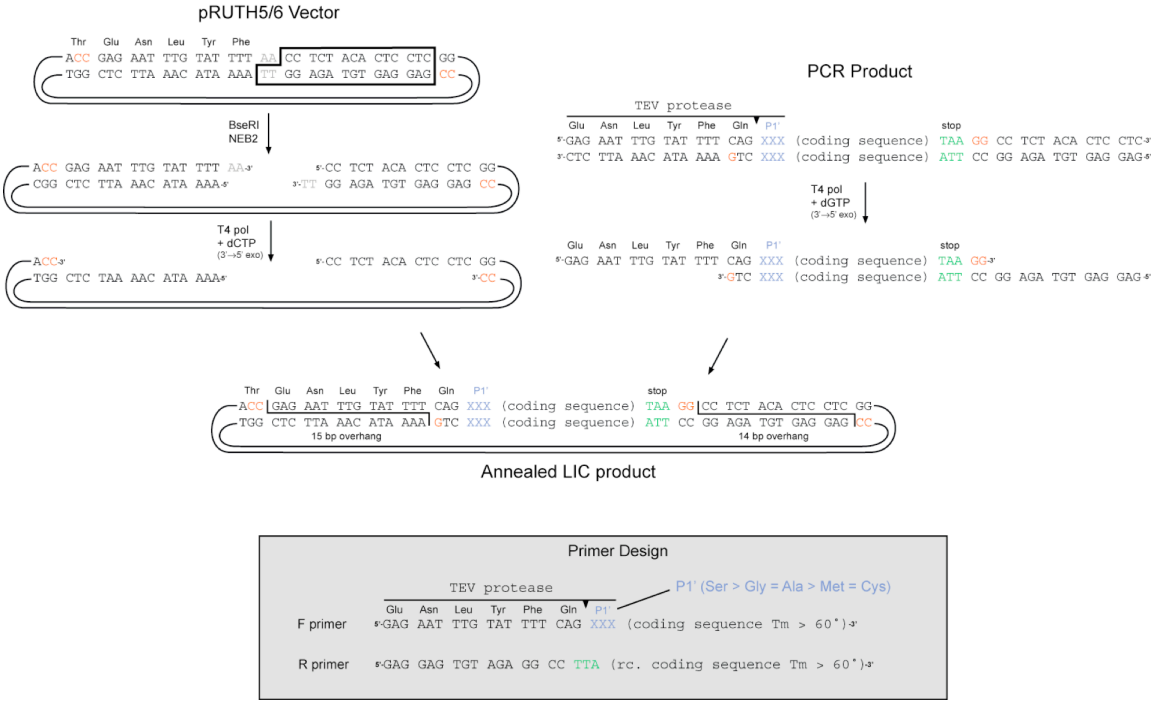



Figure 8.1: Ligation-independent cloning strategy for pRUTH5 and pRSFDuet-LIC (courtesy of Dr. A Ruthenburg)

### 8.1.2 LIC-compatible coexpression vectors for *E. coli*

For coexpression in *E. coli*, I created a pRSF-Duet derivate, transferring the pRUTH5 LIC site into the first ORF by restriction cloning, yielding pRSFDuet-LIC. A bicistronic insert for H3 and H4 expression was generated by 2-step PCR and cloned into the second ORF of pRSFDuet-LIC via restriction sites, yielding pRSFDuet-LIC/H4/H3.3. This vector allows expression of a His-tagged protein in the first ORF with the two untagged histones. Cloning protocol for the LIC site is analogous to pRUTH5 described above. Both pRSFDuet-LIC and pRSFDuet-

LIC/H4/H3.3 are further compatible for coexpression with pET-based vectors, including pMCSG7 and pRUTH2/3.

## 8.2 Protein analysis

### 8.2.1 SDS-PAGE

One-dimensional polyacrylamide gel electrophoresis (PAGE) was carried out as described with either 15% ProtoGel (Molecular Diagnostics, EC-890) (Gallagher, 2006) gels or 4-20% Precast TGX Gels (Bio-Rad 456-1091). Further, proteins were either transferred onto a PVDF membrane or stained using 0.02% Coomassie G-250, 10% Ethanol, 5% Acetic acid.

### 8.2.2 Protein transfer to PVDF membrane

Proteins were blotted onto a polyvinylidene difluoride (PVDF) membrane with a Trans-Blot® SD Semi-Dry Transfer Cell system (BioRad 170-3940). Transfer times of 30 min to 3 hours were used depending on the size of proteins that were to be transferred.

### 8.2.3 Western blotting

PVDF Membranes were blocked with 5% fat-free milk in TBS-T (TBS, 0.05% Tween) under agitated movement on a rocking platform for 60 min. Primary antibodies were prepared in the respective dilution in 5% non-fat milk in TBS-T. Membranes were incubated on a rocking platform with primary antibody over night at 4°C and washed 4 x 5 min with TBS-T. Then, membranes were incubated with secondary HRP-conjugated antibody 1:2000 in TBS-T for 1 h at RT on a

rocking platform. Membranes were washed 5 x 5 min with TBS-T before ECL detection.

### 8.2.4 Antibodies

Commercial antibodies used in this study were as follows: α-FLAG-HRP (SIGMA), α-HA-HRP A190-107P (Bethyl), α-HIRA WC119 and WC15 clones (Peter Scambler), α-HIRA (H300) sc-48774, α-p150 sc-10772 (Santa Cruz), α-Daxx sc-7152 (Santa Cruz), α-p60 NB500-212 (Novus), α-Cabin1 ab3349 (Abcam), α-ATRX H-300 (Santa Cruz), α-ASF1a/b 4A1/3 (Santa Cruz), α-HDAC1 2E10 #05-614 (Millipore).

Commercial histone (modification) antibodies: α-H3 ab1791 (Abcam), α-H3K9Ac #07-352 (Millipore), α-acetyl-H4 #06-598 (Millipore), α-H4K5Ac #06-759 (Millipore).

For ECL detections, Goat Anti-Rabbit IgG (H+L)-HRP Conjugate (BioRad #172-1019) and Goat Anti-Mouse IgG (H+L)-HRP Conjugate (BioRad #172-1011) were used.

### 8.2.5 ECL Detection

Immunoblotted bands were visualized by enhanced chemiluminescent (ECL) detection (Amersham) as described in the manufacturer's handbook. Incremental luminescent images were recorded with a CCD camera (Fujifilm LAS-3000).

## 8.3 Cell culture

### 8.3.1 Mammalian cell culture

293T cells were grown in DMEM supplemented with 10% FBS, 1x Penicillin/Streptomycin (Invitrogen 15140-122) at 37˚C, 5% $CO_2$ and 100% Humidity. Murine embryonic stem cells (ESCs) were cultured on gelatin (flasks precoated with 0.1% Gelatin/PBS and dried) in KO DMEM (Invitrogen 1082-9018) supplemented with 15% ES-grade FBS (Gibco 10439-024), 2 mM L-Glutamine (Gibco 25030-081), 0.1 $\mu$M 2-mercaptoethanol, 1x Penicillin/Streptomycin (Invitrogen 15140-122) and LIF (homemade) at 37˚C and 5% $CO_2$.

### 8.3.2 Insect cell culture

Sf9 cells were grown in Grace's Insect Cell Medium (Invitrogen 11605-102) supplemented with 1x Penicillin/Streptomycin (Invitrogen 15140-122) in a humidified 30˚C incubator. For large-scale expression, Sf9 cells were grown in baffled 0.5-2L flasks shaking at 80 rpm at 30˚C.

## 8.4 Recombinant protein production

### 8.4.1 Recombinant histones

Individual histones were expressed and purified as described in the point-by-point protocol in Appendix 9.3. Equimolar ratios of H3/H4 or H3/H4/H2A/H2B were mixed in refolding buffer (4 M Guanidine, 50 mM MOPS pH 7, 2 M NaCl, 5 mM EDTA, 15% Glycerol) to reconstitute histone tetramers or octamers, respectively. Tetramers were dialyzed against 500 mM

NaCl, Octamers against 2 M NaCl, MOPS pH 7, 1 mM EDTA. Octamers were purified on a Superdex 200 gel filtration column.

### 8.4.2 GST-Fusion proteins

GST-Fusion proteins were expressed from pRUTH2 plasmids in Rosetta(DE3)pLysS cells in either LB or TBA (Appendix 9.3) medium at 20˚C overnight. The cells were lysed in 50 mM Tris pH 7.5, 500 mM NaCl, 1 mM EDTA, 5 mM DTT, 0.8 mM PMSF in an EmulsiFlex-C5 (Avestin) and spun for 30 min at 30 000 g. Soluble fusion proteins were bound to Glutathion Sepharose (Amersham), washed with 20 mM Tris pH 7.5, 1 M NaCl, 1 mM EDTA, 5 mM DTT. GST fusion protein eluted in 20 mM Tris pH 7.5, 25 mM Glutathione, 500 mM NaCl, 0.5 mM EDTA, 5 mM DTT (prepared fresh and titrated).

### 8.4.3 Recombinant Daxx-H3.3-H4 complex

For coexpression of Daxx and histones H3.3-H4, the full ORF of Daxx was cloned into pRSFDuet-LIC/H3.3/H4 to yield pRSFDuet-Daxx/H3.3/H4. The vector encodes a N-terminal His-tag for Daxx, whereas H3.3 and H4 are untagged. This construct was transformed into Rosetta(DE3)pLysS. The complex was expressed in TBA (Appendix 9.3) medium at 20˚C overnight. The cells were lysed in 50 mM Tris pH 7.5, 500 mM NaCl, 10 mM 2-mercaptoethanol, 0.8 mM PMSF in an EmulsiFlex-C5 (Avestin). Lysate was cleared by centrifugation for 30 min at 30 000 g. Supernatant was bound to Ni-NTA resin (Qiagen). Resin was washed with 20 CV lysis buffer and bound complex was eluted with 50 mM Tris pH 7.5, 500 mM NaCl, 250 mM Imidazol.

For subsequent cation exchange chromatography on a Mono S 5/50 GL column (Amersham), eluate was diluted 1:5 with 50 mM Tris pH 7.5 and filtered through a 0.45 $\mu$m filter immediately before loading onto the column. Bound proteins were eluted with a gradient of 100 mM – 2 M NaCl in the same buffer.

### 8.4.4 Recombinant DaxxHBD-H3.3-H4 complex for crystallography

Various DaxxHBD constructs were cloned into the pRUTH5 or pRSFDuet-LIC vectors for N-terminal His6-TEV tag. Proteins were expressed in inclusion bodies for 4-6 hours at 37˚C in TBA and purified as described in the point-by-point protocol in Appendix 9.3. Equimolar ratios of DaxxHBD, H3 and H4 were mixed in refolding buffer (4 M Guanidine, 50 mM MOPS pH 7, 2 M NaCl, 5 mM EDTA, 15% Glycerol) to reconstitute the DaxxHBD-H3.3-H4 complex. The mixture was renatured by dialysis against 50 mM MOPS pH 7, 1 M NaCl, 1 mM EDTA, 10% Glycerol. After 24-48 hours, insoluble precipitate was removed by centrifugation (30 min at 30 000 g) or filtering through a 0.45 $\mu$m filter. Soluble complex concentration (Amicon Ultra, Millipore) was determined by measuring the absorption at 280 nm against dialysis buffer on a Nanodrop 2000 (Thermo Scientific) using the extinction coefficient calculated by ProtPram. Recombinant TEV protease was added to the complex in a 1:100 mass ratio to cleave the His6-tag. If cleavage did not proceed to completion after 6 hours at 4˚C, another 1:100 (w/w) aliquot of TEV was added followed by incubation for 6 hours at 10˚C. Complex was then concentrated to 10-20 mg/mL and applied to a size exclusion chromatography column (16/60 or GL 10/300 Superdex 200 for more or less than 5 mg total protein, respectively. The size exclusion column was equilibrated and run in crystal screening buffer (10 mM MOPS pH 7, 500 mM NaCl, 0.1 mM

EDTA, 0.1 mM PMSF, 5 mM DTT or 1 mM TCEP) DaxxHBD-H3.3-H4 peak fractions were pooled and concentrated (Amicon Ultra, Millipore).

### 8.4.5 Recombinant full-length Daxx and ATRX–Daxx

Full length human ATRX (isoform 2) and full length human Daxx were cloned into pFastbac1. ATRX was cloned with a C-terminal FLAG epitope, and Daxx was cloned with an N-terminal FLAG epitope. Sf9 cells were used to generate recombinant ATRX–Daxx complex and Daxx. Proteins were purified on M2 agarose followed by SMART Mono Q and Superdex 200 chromatography.

## 8.5 Crystallography

### 8.5.1 Commercial and custom 96-well screens

Commercial screens were manually set up in 2 mL deep well blocks. Screens used in this study were

> Qiagen JCSG+ Suite (130720), Qiagen PACT Suite (130718), Qiagen JCSG Core I-IV Suites (130924, 130925, 130926, 130927), Qiagen pHClear (130709), Qiagen Anions (130707), Qiagen Cations (130708), Hampton SaltRx (HR2-107 + HR2-109), Hampton Index (HR2-144), Hampton Crystal Screen (HR2-110 + HR2-112)

For custom screens, a liquid handling robot (Formulator, Formulatrix) was used to set up 2 mL deep well blocks or 80 $\mu$L well solutions directly in a screening plate.

### 8.5.2 Robotic setup for sitting drops vapor diffusion

96-well format sitting drops were set up on MRC 2 Well Crystallization Plates (Swissci) using a mosquito® liquid handling robot (TTP Labtech). 80 $\mu$L

precipitant solution was transferred to each well using a multichannel pipette. In primary screens one or two drops were set up with 0.15 $\mu$L protein + 0.15 $\mu$L well solution. Volumes for refinement screens were increased to 0.25 $\mu$L protein + 0.25 $\mu$L well.

### 8.5.3 Hanging drop vapor diffusion setup

Hanging drops were set up in 24-well EasyXtal plates (Qiagen), using 0.5 – 2 $\mu$L of protein and well solutions.

### 8.5.4 Cryocrystallography and X-ray data acquisition

Crystal cryoprotection conditions were optimized by raising the precipitant concentration or adding 15-20% glycerol compared to the original well solution or mother liquor composition. Cryoprotection solutions were tested for the appearance of salt rings in the x-ray diffraction of a flash-frozen drop. Crystals were harvested from sitting or hanging drop and equilibrated in cryoprotection solution for 1 minute before flash freezing in liquid nitrogen or directly in the cryostream. Crystal quality was assessed on a Rigaku HighFlux HomeLab x-ray diffraction system.

### 8.5.5 Dynamic and Multiangle Light Scattering

Dynamic light scattering (DLS) of protein solutions were measured in a 10 $\mu$L cuvette in a DynaPro instrument (Wyatt).

Multiangle light scattering (MALS) was measured in line with a HPLC size-exclusion column setup (SEC-MALS). A DAWN® HELEOS™ DLS (Wyatt) instrument was directly connected to an Agilent 1100 HPLC system equipped with a Superdex 200 GL 10/300 column (Amersham). 100 $\mu$L of sample (5

mg/mL) was injected and DLS signals recorded for 25 mL at 0.5 mL/min flow rate.

### 8.5.6 Reversed-phase analysis of complex

The complex was dialyzed against 0.1% acetic acid, 5 mM 2-mercaptoethanol and lyophilized before loading on the reversed phase column. A C8 column (220 x 4.6 mm Aquapore RP-300 (PerkinElmer) was used on an Beckman Coulter System Gold 126 HPLC system. A linear gradient of 30-70% B (solvent A: 5% acetonitrile, 0.1% trifluoroacetic acid; solvent B: 90% acetonitrile, 0.1% trifluoroacetic acid) was applied over 30 min at 1.0 ml/ min flow-rate.

## 8.6 Biochemical purification of native complexes

### 8.6.1 Double-affinity purification from nuclear extracts

Nuclear extract was prepared from 2x109 cells (Dignam et al., 1983) and diluted to 150 mM KCl. The first immunoprecipitation was carried out with 300 $\mu$l anti-Flag resin (Sigma A2220) for 4h at 4ºC. The resin was transferred to a disposable column, washed with 30 volumes of wash buffer (20 mM Hepes•HCl pH7.9, 750 mM NaCl, 0.2 mM EDTA , 5 mM 2-mercaptoethanol, 10% Glycerol, 0.01% NP-40, 0.4 mM PMSF) by gravity flow. Histone complexes were eluted in Flag-elution buffer (0.25 mg/mL Flag3 peptide, 20 mM Hepes•HCl pH7.9, 150 mM NaCl, 0.2 mM EDTA, 5 mM 2-mercaptoethanol, 10% Glycerol, 0.01% NP-40, 0.4 mM PMSF) and incubated with 50 $\mu$L goat anti-HA agarose (Bethyl 190-107) for 4h at 4ºC. The immunoprecipitated material was washed with 10 volumes wash buffer (see above, 1 M NaCl) and eluted in 50uL 20 mM Glycine pH2.5, 150 mM NaCl. Samples were subsequently run on a SDS PAGE.

### 8.6.2 Affinity-purification of H3.3 complexes from murine ESCs

Nuclear extracts were prepared as described (Lewis et al., 2010) from murine ESCs (Goldberg et al., 2010 2010). Immunoprecipitations were carried out in 20 mM Hepes•HCl pH7.5, 150 mM NaCl, 10% Glycerol, 0.2 mM EDTA, 2 mM MgCl2 in the presence of DNAse and RNAse, washed 5x with wash buffer (20 mM Hepes•HCl pH7.5, 250 mM NaCl, 10% Glycerol, 0.2 mM EDTA, 0.01% NP-40) and either eluted with FLAG peptide or 1% SDS. Eluted material was separated by SDS PAGE and analyzed by western blotting.

### 8.6.3 Transient transfection and pulldown from 293T cells

A pTRIP vector (containing lentiviral integration sites and large T antigen replication origin for episomal propagation) containing the H3.2 or H3.3 gene with C-terminal Flag-HA-tag (courtesy of Dr. Beth Duncan) were used. Point mutations were introduced using the AccuPrimer Pfx Polymerase Mastermix (Invitrogen) with mutagenic complementary primers as used in standard QuikChange™ reactions (designed with PrimerX, http://www.bioinformatics.org/primerx).

For transfection, 293T cells were seeded in 6-well plates at $2x10^5$ cell/well. After 12-24 hours, cell were transfected with 2.5 $\mu$g plasmid with Transit LT1 reagent (Mirus) according to the manufacturers manual. Cell were harvested after 48 hours by scraping and tituration in warm DMEM medium and pelleted. Cell pellet was washed with PBS and subsequently with hypotonic buffer (15 mM HEPES pH7.5, 30 mM KCl, 5 mM $MgCl_2$). For lysis, cell pellet was resuspended in 250 $\mu$L hypotonic buffer supplemented with 0.02% NonidentP-40, 0.8 mM PMSF, 1x EDTA-free Protease Inhibitor Cocktail (Roche), 10 $\mu$g/mL RNAse I.

Lysis was allowed to proceed for 15 with rotation at 4˚C. 22 $\mu$L 5 M NaCl was added (final 400 mM NaCl) under agitation to extract nuclear proteins. Extract was rotated for 15 min at 4˚C and spun 30 min at 4˚C at 20 000 g. Supernatant was diluted with 300 $\mu$L hypotonic buffer and was incubated with 25 $\mu$L Anti-HA Affinity Gel (SigmaAldrich E6779) for 4 hours at 4˚C. Beads were washed 2x 5 minutes with hypotonic buffer + 250 mM NaCl. Bound proteins were eluted with 1x laemmli SDS sample buffer (63 mM TrisHCl pH 6.8, 10% Glycerol, 2% SDS). DTT was added to eluate to 1% before boiling and running on an SDS PAGE.

## 8.7 *In vitro* binding and functional assays

### 8.7.1 GST-Pulldown assays

10 $\mu$g of histone tetramer was mixed with 10 $\mu$g GST fusion proteins in binding buffer (20 mM Hepes•HCl pH7.5, 200 mM NaCl, 0.1% (w/v) BSA, 5 mM 2-mercaptoethanol, 0.01% NP-40, 10% Glycerol) and incubated for 30 min at 4ºC. 25 $\mu$L Glutathione Sepharose 4B resin (GE) were added an rotated for 30 min. The resin was washed 3x with 1 mL W500 (20 mM Hepes•HCl pH7.5, 500 mM NaCl, 5 mM 2-mercaptoethanol, 0.1% NP-40, 10% Glycerol), 3x 1 mL W200 and eluted in 20 mM Hepes•HCl pH7.5, 200 mM NaCl, 30 mM Glutathione.

### 8.7.2 Peptide pulldown

An excess of N-biotinylated peptide was bound to 20 uL High Capacity Streptavidin Agarose (Pierce). The resin was washed extensively with PBS and binding buffer (20 mM Hepes•HCl pH7.5, 200 mM NaCl, 0.1% (w/v) BSA, 5 mM 2-mercaptoethanol, 0.01% NP-40, 10% Glycerol). 10 $\mu$g of recombinant Daxx was

incubated in 500 μL binding buffer with 20 uL resin for 4 h at 4°C. The resin was washed 6x 5 min with wash buffer (20 mM Hepes•HCl pH7.5, 200 mM NaCl, 5 mM 2-mercaptoethanol, 0.01% NP-40, 10% Glycerol). Peptide and bound protein were eluted in 1x SDS Loading Buffer and separated on a 10-20% Tricine gel (Invitrogen). Proteins were transferred onto a PVDF membrane and stained with Colloidal Gold Total Protein Stain (Bio-Rad).

### 8.7.3 Cystein chemical crosslinking

Bismaleimidoethane (BMOE, Pierce 22323) and 1,4-Bismaleimidobutane (BMB, Pierce 22331) were dissolved at 20 mM concentration in DMSO. In each reaction, 10 μM H3.3-H4 dimers or DaxxHBD-H3.3-H4 trimeric complex were present in 100 μl reaction volume. The reaction buffer was 20 mM MOPS pH7, 150 mM NaCl, 1 mM EDTA, 0.5 mM TCEP. BMOE or BMB was rapidly added from stock solution to a final of 20, 50 or 100 μM concentration. Reactions were incubated for 30 min at room temperature, and stopped by adding 6x laemmli sample buffer including 1% DTT, boiled and run on a SDS PAGE.

### 8.7.4 Copper(II)-based oxidation of disulfide bonds

A 100 mM Cu(II)-Phenanthroline$_3$ (CuP) solution was made from $CuCl_2$ and three molar equivalents of 1,10-Phenanthroline (Sigma 131377). H3.3-H4 dimers or DaxxHBD-H3.3-H4 trimeric complex were reduced with 50 μM TCEP and diluted at least 1:10 to 10 μM in the final reaction buffer (20 mM MOPS pH7, 150 mM NaCl). 0.1 mM CuP was added and the reaction was incubated for 30 min at 4˚C. 6x laemmli sample buffer without reducing agents was added and sampled were loaded on an SDS PAGE without boiling.

### 8.7.5 Plasmid supercoiling assay

Histones and histone chaperones were mixed and incubated on ice for 30 min in 20 $\mu$L assembly buffer AB1 (10 mM HEPES pH 7.5, 40 mM KCl, 60 mM NaCl, 2 mM $MgCl_2$, 0.5 mM EGTA). 300 ng pBluescript II KS+ was prerelaxed with recombinant human topoisomerase I (Promega) in 10 $\mu$L AB1 for 90 min at 37°C and added to the protein mix followed by 90 min incubation at 37°C. The assembly reaction was stopped by addition of 0.5 $\mu$g Proteinase K and 0.3% SDS final and subsequent incubation for 10 min at 55°C. 150 ng of purified plasmid was run on a 1.2% Agarose gel in either 1x TAE or 1x TPE, 25 $\mu$g/mL Chloroquine for 10h at 3V/cm. Gels were stained with SYBR Gold (Invitrogen)

### 8.7.6 Mononucleosome assembly assay

Recombinant Daxx-H3.3/H4 complex was reconstituted and purified from free histones by ion exchange chromatography. 50-200 $\mu$g Daxx-H3.3/H4 was mixed with 100 ng DNA in 20uL assembly buffer AB1. Reaction was incubated 30 min at room temperature and analyzed on a 6% PAGE in 0.5x TBE. Gel was stained with SYBR Gold, subsequently transferred onto a PVDF membrane and probed for H3 by standard western blotting procedures.

### 8.7.7 Mononucleosome assembly for remodeling assay

Equimolar amounts of H3.3 histone octamer and a 194 bp PCR product containing the 601 positioning sequence (1) were mixed in R2000 (20 mM Tris pH7.5, 2 M NaCl, 1 mM EDTA, 10% Glycerol, 1 mM DTT, 0.4 mM PMSF) and dialyzed against 100 mL R2000 diluted with 900 mL R0 (20 mM Tris pH7.5, 0 M NaCl, 1 mM EDTA, 10% Glycerol, 1 mM DTT, 0.4 mM PMSF) gradually over 24

h. Mononucleosomes were further dialyzed against R50 and stored at 4 °C with 1× Protease Inhibitor Cocktail (Roche). Mononucleosome mobilization assays were performed by adding ATRX-Daxx (150 ng) or ACF (50 ng) to the assembled 194 bp (200 ng) template with ATP. Reactions were performed at 37C for the time indicated. Reactions were quenched with a Stop solution (12% sucrose, 5 mM EDTA, 0.5 μg dA⁄dT). Reactions were analyzed by native 5% acrylamide, 0.5× TBE at 4 C.

### 8.7.8 Plasmid histone deposition assay

Sf9 cells were infected with baculoviruses encoding for human Daxx–FLAG and human ATRX– FLAG. Proteins were purified by M2 affinity chromatography and loaded onto a SMART Superdex 200. Fractions containing ATRX–Daxx complex were pooled and mixed with either H3.1/ H4 or H3.3/H4 tetramer. Daxx-mediated histone deposition assays: Daxx (9 μg) and H3.1/H4 or H3.3/H4 (550 ng) tetramers, and ATRX–Daxx (2 μg) with H2A/H2B (550 ng) dimers were slowly added to the reaction mixture during the 30˚C incubation (0.1 μg × 5 points during 4 hr incubation).

NAP1 (7.5 μg) and H3.1/H4 or H3.3/H4 tetramers (550 ng) were mixed together. ATP (3.25 mM), MgCl2 (4.6 mM), phosphocreatine (30 mM), creatine phosphokinase (2.5 ng), ACF (0.5 μg) and a circularized PCR fragment from pG5ML (1 μg) were subsequently added to the reaction. The reaction was incubated at 30˚C for four hours. H2A/H2B (550 ng) were added to the reaction during incubation.

### 8.7.9 ChIP and DNA dot blot

Crosslinking ChIP was performed as described (2). Daxx (M-112, sc-7152) and Atrx (H-300, sc-15408) antibodies were obtained from Santa Cruz Biotechnology, Hira antibody (ab20655) from Abcam. To examine telomere enrichment, ChIP and input DNA from 129⁄SvEv, Daxx-/-, Atrxflox and Atrx-/- ESCs were blotted with Minifold-I Dot-Blot System (Schleicher & Schuell). Blotting membrane was probed with a TTAGGG repeat probe or BamHI repeat probe as described (3).

## 8.8 Bioinformatics

### 8.8.1 Homology Modeling

Homology Modeling was carried out using a local installation of Rosetta 3.3 (http://www.rosettacommons.org/). The Backrub module was used with following parameters:

```
-backrub:ntrials 50000
-resfile resfiles/mutations.resfile
-nstruct 5
-initial_pack
```

PDB files were prepared with PyMol (http://www.pymol.org/) The residue file mutations.refile contained a list of all residues in the chains to mutate with wither the NATAA flag to keep the original residue or the PIKAA flag followed by the new amino acid (1 letter code) in the following format:

```
NATRO
start
ResNumber Chain NATAA
ResNumber Chain PIKAA AminoAcid
```

### 8.8.2 ChIP-Seq analysis

ChIP-Seq data was essentially processed as described {Ruthenburg, 2011}. Reads were mapped with Bowtie (http://bowtie-bio.sourceforge.net/). Bam files were sorted and cleaned of PCR duplicates with Picard (http://picard.sourceforge.net/). Continuous wiggle profiles were created using IGVTools (http://www.broadinstitute.org/igv/igvtools) count function extending reads to the 150bp average fragment length. Wiggle files were normalized to 50 Mio reads with a perl script:

```
#!/usr/bin/perl –w
open(IN,"$ARGV[1]");
open(OUT,">$ARGV[2]");
while(<IN>) {
   if (/^\d+/){
      chomp;
      @_=split;
      printf OUT "%d\t%.4f\n",$_[0],$_[1]/$ARGV[0];
   }
   else{
      print OUT $_;
   }
}
```

To create tracks for IGV (http://www.broadinstitute.org/igv/home) visualization, normalized wig files were tiled with igvtools tile command. Normalized wig files were also used for accumulation plots with the Cis-regulatory element annotation system (CEAS, http://liulab.dfci.harvard.edu/CEAS/). Cumulative and ratio tracks were created from wig files by concatenating columns in an interval file (wigBed)

through the cistrome online analysis portal (http://cistrome.org/ap/) and afterwards converted back to wig with the following unix shell script:

```
#!/bin/sh
SPAN=$3
S=$1
R=$2
rm -f ${R}
head -45 ${S} | egrep "^browser|^track" > ${R}
#grep "^chr" ${S} | cut -f1 | sort -u > chr.list
cat chr.list | while read C
do
   echo "variableStep chrom=${C} span=${SPAN}" >> ${R}
   awk '{if (match($1,"^"${C}"$")) { print } }' ${S} | sort -k2n awk
'
{
   printf "%d\t%g\n", $2+1, $8/$5
}
' >> ${R}
done
```

For repeat analysis, nonmapped reads were extracted from bam files with bam2fastq (http://www.hudsonalpha.org/gsl/software/bam2fastq.php). Occurrences of the repeat sequences were counted with the unix grep –c command.

### 8.8.3 Other bioinformatic software packages/online services

| | |
|---|---|
| Extinction coefficient and pI | http://web.expasy.org/protparam/ |
| Secondary structure prediction | http://www.compbio.dundee.ac.uk/www-jpred/ |
| Protein charge plot | http://mobyle.pasteur.fr/ |
| Protein alignments | http://www.jalview.org/ |
| DNA alignments/HYLIP | http://ugene.unipro.ru/ |
| Maximum Likelihood Parsimony | http://mobyle.pasteur.fr/ |

# 9 Appendix

## 9.1 Mass Spectrometry of native Daxx-H3.3-H4 complex

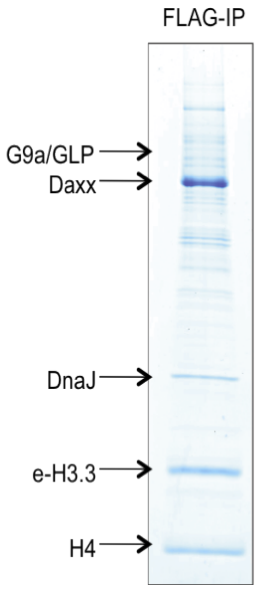



Figure 9.1: MS/MS identification of proteins associated with the Daxx-H3.3-H4 predeposition complex.
Daxx complex purification as described in Figure 5.2. FLAG-tagged H3.3 (e-H3.3) was immunoprecipitated from the high salt fraction of the heparin gradient (see Section 5.3.1). The methyltransferase complex G9a/GLP was found in the H3.3-associated material.

## 9.2 Biochemical characterization of the HIRA complex

HIRA shares homology with the *S. cerevisiae* Hir1 (N-terminal beta-propeller) and Hir2 (C-terminal Tup1-like domain) proteins. As Hir1 and Hir2 participate in a quaternary complex with Hir3 and Hpc2, it is likely that HIRA forms a complex as well. By sequence homology to Hir3 and Hpc2, Cabin1 and

Ubinuclein-1 (UBN1), respectively, are likely candidates (Balaji et al., 2009) and both proteins have been described to interact with HIRA (Banumathy et al., 2009; Tagami et al., 2004). However, the putative HIRA complex has not been empirically defined. Therefore, an affinity purification of the HIRA complex from HeLa cells was carried out with subsequent Mass Spectrometric analysis.

A mixture of two monoclonal α-HIRA antibodies (WC119 and WC15) was immobilized on a 1 mL Protein G Sepharose column with DMP. HIRA and associated proteins were coimmunoprecipitated from HeLa nuclear extracts (20 mL, from 6L HeLa cells). Nuclear extracts (100 mM NaCl, 25 mM HEPES pH7.9) were first run over an empty Protein G Sepharose column (negative control) and then over the α -HIRA column with subsequent extensive washing (Figure 9.2A). Loosely bound proteins were dissociated with a NaCl gradient (100-500 mM). HIRA and strongly associated proteins were eluted at low pH (100 mM Glycine pH2.5). Aliquots of a-HIRA washes and elution, as well as the control Protein G elution, were separated by SDS PAGE and analyzed by tryptic digestion and MS/MS fingerprinting (Figure 9.2B). Selected tandem mass spectrometry data were confirmed western blots (Figure 9.2C). Complete complexes represented by significant peptides of all subunits in the MS/MS data are shown in Figure 9.2D.

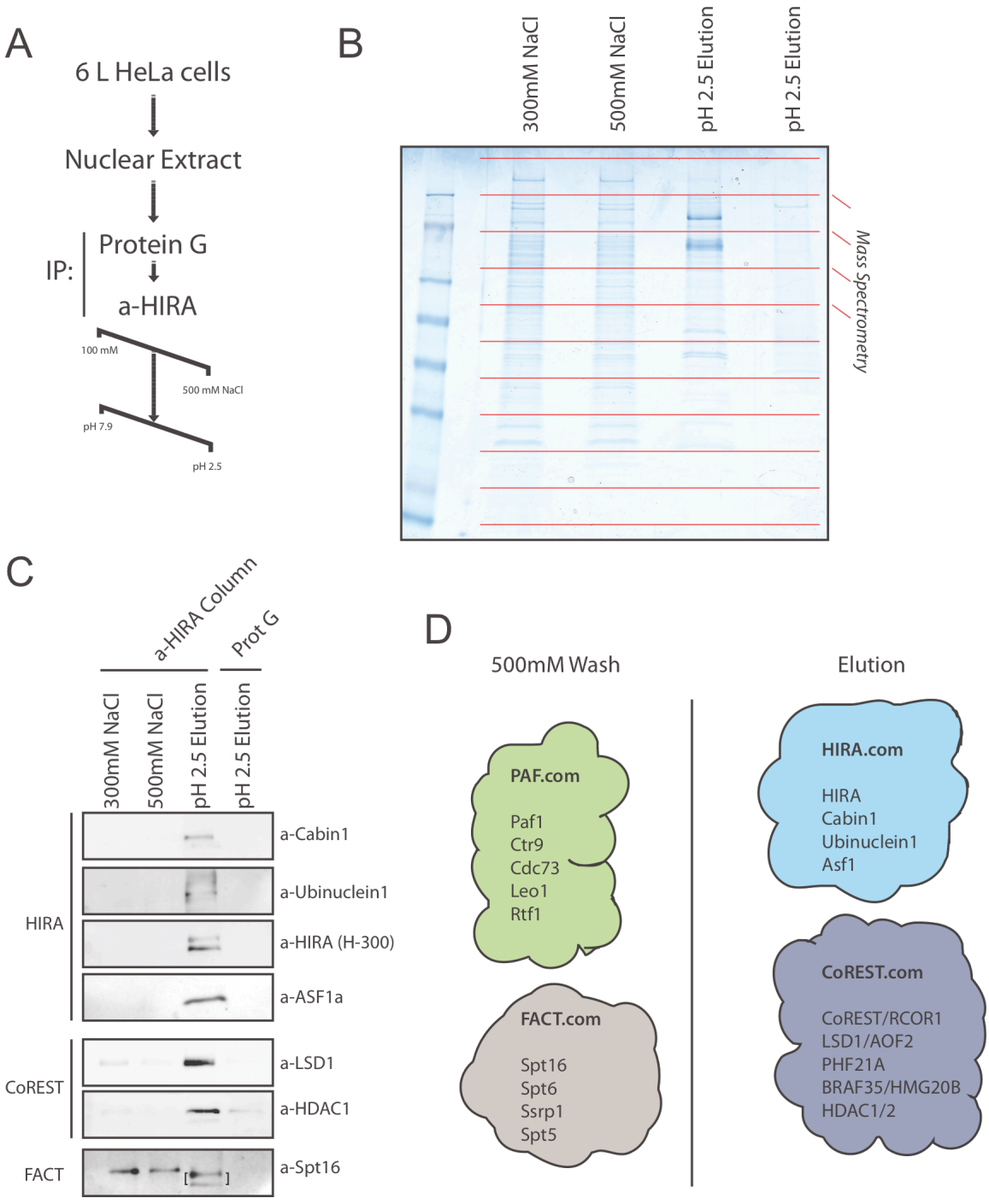



Figure 9.2: HIRA complex coimmunoprecipitation from HeLa nuclear extract.

The total length of each lane from the SDS PAGE in Figure 9.2B was divided into 10 segments that were analyzed separately by tryptic digest and tandem mass spectrometry. Peptides were matched with the human proteome database by the MASCOT search engine. Hits were further analyzed individually by the following criteria:

- Proteins with less then three matching peptides were discarded

- The molecular weights of each identified protein was verified against the SDS PAGE segment it was identified in. Only proteins within a reasonable molecular weight range were further analyzed.

- Proteins that were also identified in the eluate of the Protein G column (negative control) were disregarded.

Known interactions of candidate proteins were retrieved from databases (BioGRID, STRING) to identify known complexes the might interact with HIRA By the criteria above, the most significant 50 proteins identified in the 500 mM wash fraction and the pH 2.5 elution are listed in Figure 9.3.

The remaining eluate from the α-HIRA affinity column was subjected to further fractionation by ion exchange chromatography (Figure 9.4B) and gel filtration (Figure 9.4C). Not the separation of a higher migrating species of HIRA on both columns that does not correlate with the remaining complex members. I speculate that it is a cell cycle-dependent hyperphosphorylated form of HIRA (Hall et al., 2001). The phosphorylated form does no longer interact with the other complex members, which explains why it is no longer localized to

chromosomes (Hall et al., 2001). The remaining proteins could stay as a complex on chromatin and define genomic HIRA binding sites.

Figure 9.3: Mass spectrometry of HIRA complex and associated proteins.

**Table 1:** Top50 Proteins from MS/MS Analysis of HIRA IP (Elution pH 2.5)

| prot_acc | complex | Protein Name | Mass (kDa) | MassSpec Score (HIRA IP) | MassSpec Score (Protein G) |
|---|---|---|---|---|---|
| IPI00217560 | HIRA | HIRA Isoform Long of Protein HIRA | 113 | 1241 | 0 |
| IPI00002355 | HIRA | CABIN1 Calcineurin-binding protein cabin-1 | 249 | 1096 | 0 |
| IPI00099834 | HIRA | UBN1 Ubinuclein | 122 | 521 | 0 |
| IPI00292168 | HIRA | ASF1A Histone chaperone ASF1A | 23 | 125 | 0 |
| IPI00041959 | CoREST | PHF21A Isoform 2 of PHD finger protein 21A | 71 | 992 | 0 |
| IPI00217540 | CoREST | AOF2 Isoform 2 of Lysine-specific histone demethylase 1 | 96 | 590 | 0 |
| IPI00008531 | CoREST | RCOR1 REST corepressor 1 | 53 | 495 | 0 |
| IPI00464951 | CoREST | HMG20B SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamil | 36 | 492 | 0 |
| IPI00289601 | CoREST | HDAC2 histone deacetylase 2 | 66 | 128 | 0 |
| IPI00013774 | CoREST | HDAC1 Histone deacetylase 1 | 56 | 124 | 0 |
| IPI00477535 | | ERCC5 DNA-repair protein complementing XP-G cells | 134 | 15221 | 0 |
| IPI00014952 | | GDAP2 Isoform 1 of Ganglioside-induced differentiation-associated protein 2 | 57 | 3907 | 0 |
| IPI00180730 | | - Similar to Elongation factor 1-alpha 1 | 50 | 2496 | 0 |
| IPI00024235 | | BACH1 Transcription regulator protein BACH1 | 84 | 2187 | 0 |
| IPI00339353 | | ARHGEF12 Isoform 2 of Rho guanine nucleotide exchange factor 12 | 172 | 1552 | 0 |
| IPI00023098 | | FOXP1 Isoform 1 of Forkhead box protein P1 | 76 | 970 | 0 |
| IPI00031820 | | FARSA Phenylalanyl-tRNA synthetase alpha chain | 58 | 915 | 0 |
| IPI00020928 | | TFAM Transcription factor A, mitochondrial precursor | 29 | 622 | 0 |
| IPI00013683 | | TUBB3 Tubulin beta-3 chain | 51 | 566 | 0 |
| IPI00031410 | | FRAP1 FKBP12-rapamycin complex-associated protein | 291 | 350 | 0 |
| IPI00031018 | | MAFK V-maf musculoaponeurotic fibrosarcoma oncogene homolog K (Fragment) | 18 | 335 | 0 |
| IPI00785015 | | tcag7.1228 Isoform 1 of Uncharacterized protein KIAA2030 | 147 | 331 | 0 |
| IPI00016786 | | CDC42 Isoform 2 of Cell division control protein 42 homolog precursor | 22 | 257 | 0 |
| IPI00003362 | | HSPA5 HSPA5 protein | 72 | 243 | 0 |
| IPI00030247 | | CCNT1 Cyclin-T1 | 81 | 240 | 0 |
| IPI00166680 | | MINK1 Isoform 3 of Misshapen-like kinase 1 | 150 | 235 | 0 |
| IPI00031115 | | GOLGA1 Golgin subfamily A member 1 | 88 | 217 | 0 |
| IPI00743194 | | - Kappa light chain variable region (Fragment) | 13 | 211 | 0 |
| IPI00103525 | | PSPC1 Isoform 1 of Paraspeckle component 1 | 59 | 203 | 0 |
| IPI00007311 | | MAFG Transcription factor MafG | 18 | 201 | 0 |
| IPI00018971 | | TRIM21 52 kDa Ro protein | 55 | 195 | 0 |
| IPI00017454 | | TUBA4B Putative tubulin-like protein alpha-4B | 28 | 164 | 155 |
| IPI00554788 | | KRT18 Keratin, type I cytoskeletal 18 | 48 | 159 | 212 |
| IPI00102926 | | GBL Isoform 1 of Target of rapamycin complex subunit LST8 | 37 | 149 | 0 |
| IPI00386277 | | FOXP4 FOXP4 protein (Fragment) | 42 | 142 | 0 |
| IPI00006752 | | MAP4K4 Isoform 1 of Mitogen-activated protein kinase kinase kinase kinase 4 | 143 | 135 | 0 |
| IPI00218905 | | TNIK Isoform 6 of TRAF2 and NCK-interacting protein kinase | 153 | 135 | 0 |
| IPI00026272 | | HIST1H2AB;HIST1H2AE Histone H2A type 1-B | 14 | 126 | 331 |
| IPI00787362 | | LOC730839 similar to Hornerin | 191 | 95 | 0 |
| IPI00157442 | | ARHGEF11 Rho guanine nucleotide exchange factor 11 | 168 | 91 | 0 |
| IPI00888063 | | KRT18P33 similar to hCG21219 | 106 | 86 | 0 |
| IPI00293464 | | DDB1 DNA damage-binding protein 1 | 128 | 82 | 56 |
| IPI00030243 | | PSME3 Isoform 1 of Proteasome activator complex subunit 3 | 30 | 81 | 0 |
| IPI00418262 | | ALDOC Fructose-bisphosphate aldolase C | 40 | 72 | 0 |
| IPI00045917 | | CRBN Isoform 1 of Protein cereblon | 51 | 70 | 0 |
| IPI00006510 | | TUBB1 Tubulin beta-1 chain | 51 | 66 | 55 |
| IPI00740019 | | CCDC88C Isoform 1 of Protein Daple | 229 | 66 | 0 |
| IPI00015148 | | RAP1B Ras-related protein Rap-1b precursor | 21 | 63 | 0 |

**Table 2:** Top50 Proteins from MS/MS Analysis of HIRA IP (Wash 500mM NaCl)

| prot_acc | complex | Protein Name | Mass (kDa) | MassSpec Score (500mM Wash) | MassSpec Score (HIRA IP) | MassSpec Score (Protein G) | Unique B |
|---|---|---|---|---|---|---|---|
| IPI00026970 | FACT | SUPT16H FACT complex subunit SPT16 | 0 | 393 | 0 | 0 | |
| IPI00298058 | FACT | SUPT5H Isoform 1 of Transcription elongation factor SPT5 | 121 | 258 | 0 | 0 | |
| IPI00005154 | FACT | SSRP1 FACT complex subunit SSRP1 | 81 | 220 | 0 | 0 | |
| IPI00784161 | FACT | SUPT6H Isoform 1 of Transcription elongation factor SPT6 | 200 | 34 | 0 | 0 | |
| IPI00384593 | Paf | PAF1 Isoform 2 of RNA polymerase II-associated factor 1 homolog | 55 | 249 | 0 | 0 | |
| IPI00477468 | Paf | CTR9 RNA polymerase-associated protein CTR9 homolog | 134 | 194 | 0 | 0 | |
| IPI00300659 | Paf | CDC73 Parafibromin | 61 | 138 | 0 | 0 | |
| IPI00103090 | Paf | LEO1 Isoform 1 of RNA polymerase-associated protein LEO1 | 75 | 90 | 0 | 0 | |
| IPI00303832 | Paf | RTF1 RNA polymerase-associated protein RTF1 homolog | 80 | 57 | 0 | 0 | |
| IPI00296337 | | PRKDC Isoform 1 of DNA-dependent protein kinase catalytic subunit | 474 | 2345 | 408 | 0 | |
| IPI00786995 | | LOC731751 similar to DNA dependent protein kinase catalytic subunit | 470 | 2218 | 0 | 0 | |
| IPI00644712 | | XRCC6 ATP-dependent DNA helicase 2 subunit 1 | 70 | 1696 | 0 | 0 | |
| IPI00008359 | | KRT76 Keratin, type II cytoskeletal 2 oral | 66 | 1293 | 0 | 0 | |
| IPI00174775 | | KRT73 Isoform 1 of Keratin, type II cytoskeletal 73 | 59 | 1148 | 680 | 0 | |
| IPI00293665 | | KRT6B Keratin, type II cytoskeletal 6B | 60 | 1129 | 643 | 0 | |
| IPI00220740 | | NPM1 Isoform 2 of Nucleophosmin | 30 | 1027 | 0 | 0 | |
| IPI00888013 | | LOC389901 hypothetical LOC389901 | 97 | 1019 | 0 | 0 | |
| IPI00604620 | | NCL Isoform 1 of Nucleolin | 77 | 992 | 0 | 0 | |
| IPI00398625 | | HRNR Hornerin | 283 | 920 | 0 | 0 | |
| IPI00027569 | | HNRNPCL1 Heterogeneous nuclear ribonucleoprotein C-like 1 | 32 | 817 | 0 | 0 | |
| IPI00220834 | | XRCC5 ATP-dependent DNA helicase 2 subunit 2 | 83 | 767 | 0 | 0 | |
| IPI00386854 | | HNRNPA2B1 HNRPA2B1 protein | 28 | 704 | 0 | 0 | |
| IPI00886929 | | LOC100131813 similar to hCG2041218 | 27 | 657 | 0 | 0 | |
| IPI00003519 | | EFTUD2 116 kDa U5 small nuclear ribonucleoprotein component | 110 | 631 | 0 | 0 | |
| IPI00007752 | | TUBB2C Tubulin beta-2C chain | 50 | 615 | 685 | 0 | |
| IPI00025491 | | EIF4A1 Eukaryotic initiation factor 4A-I | 46 | 597 | 0 | 0 | |
| IPI00465233 | | EIF3EIP Eukaryotic translation initiation factor 3, subunit E interacting protein | 71 | 582 | 0 | 0 | |
| IPI00215965 | | HNRNPA1 Isoform A1-B of Heterogeneous nuclear ribonucleoprotein A1 | 39 | 569 | 0 | 0 | |
| IPI00013475 | | TUBB2A Tubulin beta-2A chain | 50 | 555 | 656 | 0 | |
| IPI00646779 | | TUBB6 TUBB6 protein | 51 | 478 | 341 | 0 | |
| IPI00420014 | | ASCC3L1 Isoform 1 of U5 small nuclear ribonucleoprotein 200 kDa helicase | 246 | 477 | 0 | 0 | |
| IPI00215637 | | DDX3X ATP-dependent RNA helicase DDX3X | 74 | 465 | 0 | 0 | |
| IPI00449049 | | PARP1 Poly [ADP-ribose] polymerase 1 | 114 | 445 | 0 | 0 | |
| IPI00000875 | | EEF1G Elongation factor 1-gamma | 50 | 417 | 2294 | 0 | |
| IPI00023048 | | EEF1D Elongation factor 1-delta | 31 | 415 | 1460 | 0 | |
| IPI00413108 | | RPSA 33 kDa protein | 33 | 413 | 0 | 0 | |
| IPI00017617 | | DDX5 Probable ATP-dependent RNA helicase DDX5 | 70 | 399 | 0 | 0 | |
| IPI00031812 | | YBX1 Nuclease-sensitive element-binding protein 1 | 36 | 374 | 0 | 0 | |
| IPI00240675 | | PDCD4 programmed cell death 4 isoform 2 | 51 | 360 | 0 | 0 | |
| IPI00009328 | | EIF4A3 Eukaryotic initiation factor 4A-III | 47 | 355 | 0 | 0 | |
| IPI00008530 | | RPLP0 60S acidic ribosomal protein P0 | 34 | 352 | 0 | 0 | |
| IPI00216049 | | HNRNPK Isoform 1 of Heterogeneous nuclear ribonucleoprotein K | 51 | 344 | 0 | 0 | |
| IPI00019269 | | WDR61 WD repeat-containing protein 61 | 34 | 339 | 0 | 0 | |
| IPI00748303 | | ZFR 115 kDa protein | 116 | 321 | 290 | 0 | |
| IPI00216587 | | RPS8 40S ribosomal protein S8 | 24 | 317 | 0 | 0 | |
| IPI00647650 | | EIF3H Eukaryotic translation initiation factor 3 subunit 3 | 42 | 310 | 0 | 0 | |
| IPI00013788 | | HTATSF1 HIV Tat-specific factor 1 | 86 | 293 | 0 | 0 | |
| IPI00217030 | | RPS4X 40S ribosomal protein S4, X isoform | 30 | 288 | 0 | 0 | |
| IPI00654777 | | EIF3F Eukaryotic translation initiation factor 3 subunit 5 | 39 | 277 | 0 | 0 | |
| IPI00023598 | | TUBB4 Tubulin beta-4 chain | 50 | 276 | 0 | 0 | |
| IPI00029012 | | EIF3A Eukaryotic translation initiation factor 3 subunit A | 167 | 270 | 0 | 0 | |
| IPI00872047 | | SMU1 Uncharacterized protein SMU1 (Fragment) | 58 | 266 | 0 | 0 | |
| IPI00013068 | | EIF3E Eukaryotic translation initiation factor 3 subunit E | 53 | 263 | 0 | 0 | |
| IPI00029400 | | ZRANB2 Isoform 1 of Zinc finger Ran-binding domain-containing protein 2 | 38 | 256 | 0 | 0 | |
| IPI00741317 | | CSNK2A1 Casein kinase 2 alpha isoform | 45 | 256 | 0 | 0 | |
| IPI00013881 | | HNRNPH1 Heterogeneous nuclear ribonucleoprotein H | 49 | 247 | 0 | 0 | |
| IPI00301503 | | SFRS10 Isoform 1 of Splicing factor, arginine/serine-rich 10 | 34 | 245 | 0 | 0 | |
| IPI00016910 | | EIF3CL;EIF3C Eukaryotic translation initiation factor 3 subunit C | 106 | 242 | 0 | 0 | |
| IPI00396370 | | EIF3B Isoform 1 of Eukaryotic translation initiation factor 3 subunit B | 93 | 235 | 0 | 0 | |
| IPI00015953 | | DDX21 Isoform 1 of Nucleolar RNA helicase 2 | 88 | 228 | 0 | 0 | |
| IPI00024933 | | RPL12 60S ribosomal protein L12 | 18 | 222 | 0 | 0 | |
| IPI00479217 | | HNRNPU Isoform Short of Heterogeneous nuclear ribonucleoprotein U | 90 | 211 | 0 | 0 | |
| IPI00023785 | | DDX17 DEAD box polypeptide 17 isoform 1 | 81 | 206 | 0 | 0 | |
| IPI00844578 | | DHX9 ATP-dependent RNA helicase A | 142 | 205 | 0 | 0 | |
| IPI00641706 | | TUBB6 46 kDa protein | 46 | 202 | 0 | 0 | |

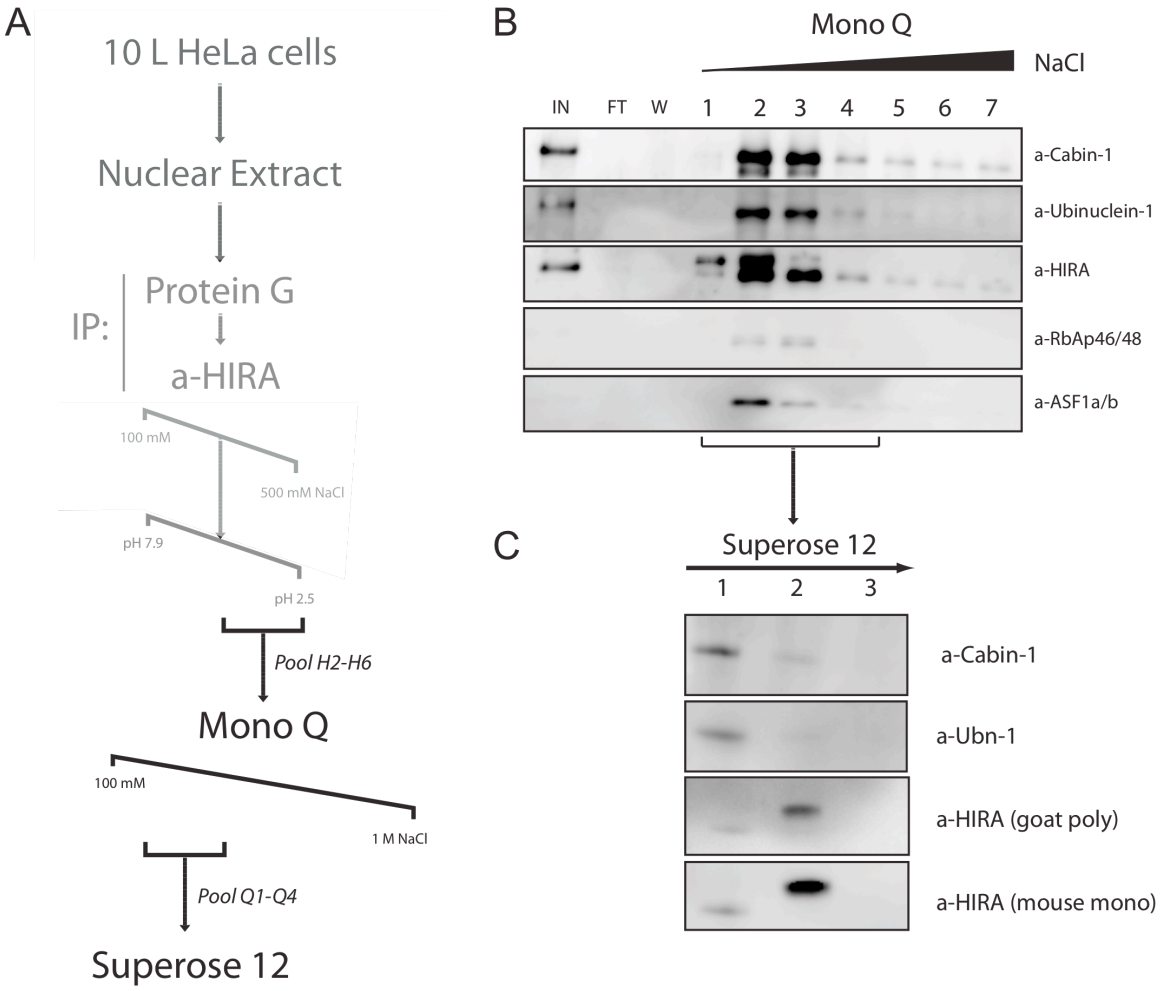



Figure 9.4: Biochemical fractionation of the HIRA complex.

## 9.3 A simple method for purifying recombinant histones[‡]

### 9.3.1 Introduction

Purified recombinant histones expressed in *E. Coli*, and nucleosomes assembled with them, have proven to be extraordinarily powerful reagents for studying chromatin biology and biochemistry. The production of purified recombinant histones was an essential prerequisite to high-resolution crystal structures of the nucleosome core particle (Davey et al., 2002; Luger et al., 1997), as well as crucial reagents in landmark histone modification enzyme (Wang et al., 2001) and chromatin remodelling biochemistry (Hamiche et al., 2001). Although much has been gained from the reductionist approach employing histone peptides as surrogates for full-length histones in their native contexts, there are countless examples of chromatin modification activities that require full-length histones, chaperone-complexes, nucleosomes, or nucleosomal arrays. The explosion of interest in epigenetics and chromatin biology has drawn laboratories outside of the field into working with histone and nucleosome reagents. Here we report improved methods for preparation of purified recombinant core histones of *S. cerevisiae*, *X. Laevis*, and *H. sapiens* in *E. coli*, and modest simplifications of the nucleosome assembly process to make this protocol more accessible to non-specialists.

---

[‡] This chapter has been prepared for submission as: Simon J Elsässer, Laura A. Banaszynski[*], Christina M. Hughes[*], Sarah J. Whitcomb[*], C. David Allis, Alexander J. Ruthenburg[§], "A simple method for expression and purification of recombinant histone"

[*]Equal contribution, [§]Corresponding author

The chromatin field has a checkered history of mismatching the organism from which a given protein or complex is derived with the source of recombinant histones employed—indeed the source of histones in many cases is still the original *Xenopus laevis* histone constructs. Although amenable to large-scale protein production necessitated by crystallographic material demands, this protocol requires both HPLC and FPLC instruments and columns and yields protein of only modest purity. Indeed this histone preparation may co-purify with unwanted nucleic acid and minor protein impurities from the *E. coli* expression host, and the use of 7 M urea in the course of this purification, even when freshly deionized, may lead to protein carbamylation adducts. Extreme care must be taken to avoid protein carbamylation that occurs as urea containing buffers spontaneously decompose into ammonium cyanate, which in turn, reacts with the amino terminus and lysine residues of histones.

A salient improvement on this procedure in terms of technical ease was afforded by Tanaka and colleagues, by expression of the human core histones as protease cleavable hexahistidine fusion proteins (Tanaka et al., 2004). However, significant drawbacks to this approach remain: as this preparation continues the use of urea, the potential carbamylation issue was not resolved, more importantly, the cleaved purified histones retain additional vector encoded amino acid "scars" and thereby differ at the N-terminus relative to native histones. A number of recognition modules appear to recognize the N-terminus itself (Ruthenburg et al., 2007), consequently, these additional amino acids may

compromise the utility of these histones in termini-sensitive downstream applications.

The expression and purification conditions described here were designed to circumvent the above conditions and allow scalable and streamlined production of all four core histones from various species. We make use of a series of pET-based vectors that encode a cleavable N-terminal His-tag. The use of TEV and R3C protease sites allows site-specific processing to yield a scarless, native N-terminus for all major histone species.

As in previous described methods, we purify histones from inclusion bodies. Thorough wash conditions help to remove nucleic acids and proteinacious impurities from the histone pellet. A subsequent nickel-affinity purification step in 6 M Guanidine hydrochloride yields pure and concentrated denatured histone. We provide alternate protocols for direct octamer assembly from those histones with subsequent tag cleavage, or individual processing. Either method yields highly pure, native histone octamers that can be used for nucleosome assembly. We also describe optimized workflows to produce H3-H4 tetramers and H2A-H2B dimers.

### 9.3.2 Reagent Setup

**Terrific Broth Autoinduction Medium (TBA)**

Make a 50x Autoinduction Stock '5052' by dissolving 250 g glycerol, 25 g glucose and 100 g a-lactose in 1L ddH2O. Sterile filter solution. Autoclave a 1x TB-solution (12 g Tryptone, 24 g Yeast extract, no glycerol). Add 20 mL '5052' per liter autoclaved TB- just before inoculation. Add 1 mL of kanamycin stock (mg/mL) per liter TBA. Add a single drop of Antifoam A concentrate (viscous

solution, mix well before pipetting and use 1 mL pipette tip) to avoid excessive foam formation during shaking.

**Resuspension buffer (R)**

Make a solution containing 50 mM Tris•HCl pH 8, 5 mM EDTA. Add 0.8 mM PMSF to prechilled buffer R just before use. You will need 40 mL per 1L culture.

**Lysis Buffer (LW)**

Make a solution containing 50 mM Tris•HCl pH 8, 250 mM NaCl, 5 mM MgCl2. You will need 40 mL lysis buffer per 1L culture. Add 0.8 mM PMSF to prechilled buffer L just before use.

**Inclusion Body Wash Buffer (IW)**

Make a solution containing 50 mM Tris · HCl pH 8, 250 mM NaCl You will need ~100 mL buffer IW per 1L TBA culture. Add 0.8 mM PMSF to prechilled buffer IW just before use.

**10% (w/v) Triton X-100 Stock solution**

Weigh in 5 g Triton X-100 detergent in a 50 mL Falcon Tube, add ddH2O to 50 mL, mix by slow rotation at room temperature. Store protected from light.

**Denaturing Buffer DB**

Always make a fresh solution of this buffer, containing 50 mM Tris•HCl pH 8.0, 500 mM NaCl, 6.3 M Guanidine•HCl, 5 mM 2-mercaptoethanol. As proteins are efficiently unfolded in this buffer, protease inhibitors are now obsolete.

To make 100 mL of D500 buffer, weigh in 60 g Guanidine•HCl in a beaker. Add 5 mL of 1 M Tris•HCl pH 8.0 and 10 mL of 5 M NaCl. Add water to 80 mL total volume and stir at room temperature until Guanidine is fully dissolved. Use pH meter do check that pH is between 7.6 and 8.0. There are batch-to-batch variations in residual HCl in various commercial Guanidine•HCl that are often

sufficient to perturb the final pH - adjust with concentrated HCl or NaOH if necessary. Add ddH2O to 100 mL and 35uL 2-mercaptoethanol.

**Denaturing Wash Buffer DWB**

Always make a fresh solution of this buffer, containing 50 mM Tris•HCl pH 8.0, 1 M NaCl, 4.2 M Guanidine•HCl, 5 mM 2-mercaptoethanol, as described above.

**Denaturing Elution Buffer DEB**

Always make a fresh solution of this buffer, containing 50 mM Tris•HCl pH 8.0, 1 M NaCl, 4.2 M Guanidine•HCl, 300 mM imidazole, 5 mM 2-mercaptoethanol, as described above.

**Octamer refolding buffer ORB**

Make a 20 mM Tris pH 8, 2 M NaCl, 1 mM EDTA solution.

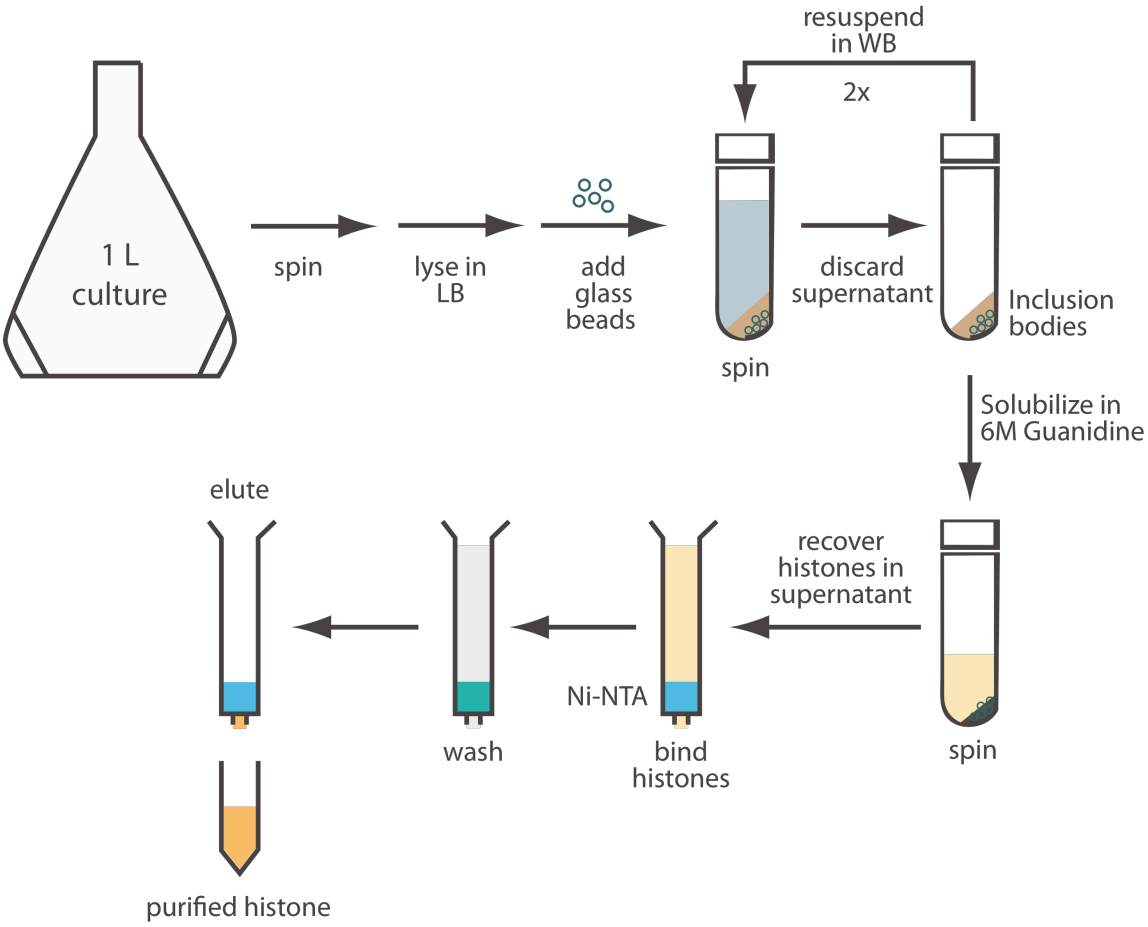
Supplement with fresh reagents to: 10 mM 2-mercaptoethanol, 1 mM PMSF

**Tetramer refolding buffer TRB**

20 mM Tris pH 6.5, 500 mM NaCl, 1 mM EDTA, 10% (v/v) glycerol

Supplement with fresh reagents to: 10 mM 2-mercaptoethanol, 1 mM PMSF

### 9.3.3 Procedure




Schematic of general procedure for purifying individual histones

**Expression of histones ● TIMING 1 d**

1) Inoculate 5-50 mL overnight culture in LB from fresh transformation plate.

2) Inoculate 1 L TBA expression culture with 5 mL volume of the dense LB culture. Shake at 37 ˚C until the culture reaches OD600 ~0.7. Add 1 mL IPTG Stock solution. Shake for another 4-6 hours at 37 ˚C.

3) Harvest cells by spinning down TBA culture 15 min at 4000 rpm.

4) Resuspend pellet in 30-40 mL resuspension buffer R, transfer to 50 mL Falcon tube and spin down for 15 min at 4000rpm. Discard supernatant (can be turbid).

5) If not immediately proceeding to lysis, resuspend pellet in 5 mL lysis buffer, divide into two 50 mL Falcon tubes and flash-freeze in liquid nitrogen. Store at -80˚C. ■ **PAUSE POINT**

**Bacterial lysis with emulsiflex (1+ liters culture) ● TIMING 1h**

6) Resuspend fresh or frozen cell pellet in buffer L (freshly supplemented with 0.8 mM PMSF, final volume 40 mL per 1 L culture)

7) Run bacterial suspension twice through Emulsiflex or French press. If lysate remains viscous, repeat. High inclusion body content should look milky or opaque but not viscous. 'Residual sample viscosity signifies either incomplete lysis or a significant unsheared genomic DNA in the sample— both pitfalls will impinge upon downstream processing.

**Inclusion Body Purification ● TIMING 3 h**

▲ **CRITICAL STEP** Glass beads will greatly facilitate resuspension of the inclusion body pellet and allow efficient washing steps.

8) Add ~1 mL glass beads (0.5-1 mm diameter) to each 50 mL tube of lysate.

9) Centrifuge the lysate (15 min x 15,000 g, 4°C) in oak ridge SS-34 centrifuge tubes, or others rated to 30,000 g with screw caps—one tube per liter of original culture volume greatly facilitates resuspension and downstream processing.

10) Set aside supernatant lysate, histones will pellet quantitatively in the inclusion bodies. You should recover a rubbery to brittle pellet.

11) Add 5 mL of LW, supplemented with 1% Triton X-100 and 0.8 mM PMSF. Resuspend inclusion bodies by vigorous vortexing.

12) Add another 30 mL of LW to inclusion body suspension, rotate for 10 min at 4°C.

13) Pellet the inclusion bodies (15 min x 15,000 g, 4°C).

14) Add 5 mL of IW, supplemented with 0.8 mM PMSF. Resuspend inclusion bodies by vigorous vortexing.

15) Add another 30 mL of IW to inclusion body suspension, rotate for 10 min at 4°C.

16) Pellet the inclusion bodies (15 min x 15,000 g, 4°C).

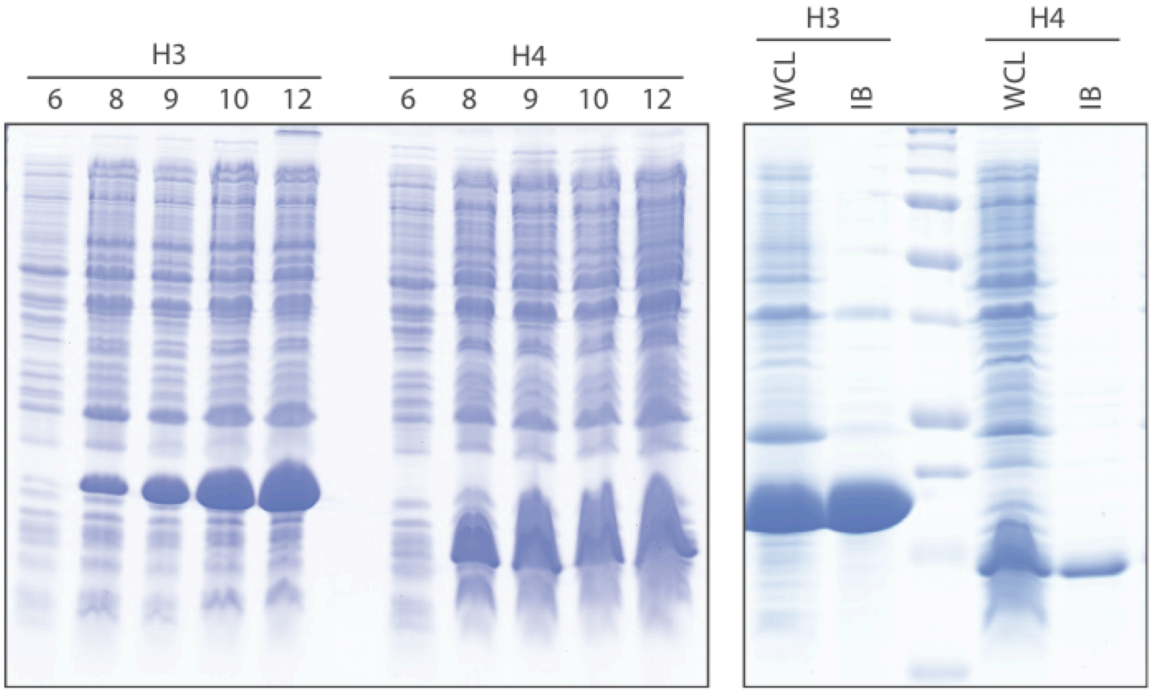
■ **PAUSE POINT** Inclusion body pellet can be transferred to 50 mL Falcon Tube before last spin, pelleted by centrifugation for 30 min x 4000rpm, flash-frozen in liquid nitrogen and stored indefinitely at -80°C.

**ALTERNATE PROTOCOL: Autoinduction and enzymatic lysis for small-scale, parallel expression (1L or less) ● TIMING overnight + 4 hours**

1) Inoculate 5-50 mL TBA from fresh transformation plate (evening).

2) Shake at 37 ˚C overnight.

3) Harvest cells by spinning down TBA culture 10 min at 4000 rpm in 50 mL Falcon tubes.




Example of autoinduction expression of H3 and H4 (left, hours after inoculation given). Inclusion bodies contain predominantly pure histone (right)

4) Supplement lysis buffer LW with 2% Triton, 0.8 mM PMSF, 500 $\mu$g/mL Lysozyme, 10 $\mu$g/mL DNase.

5) Resuspend fresh cell pellet in 25 mL supplemented buffer LW at room temperature. Lysate will become viscous due to genomic DNA released from the cells.

6) Rotate bacterial suspension at room temperature until lysate does not contain any cell fragments and becomes evenly milky with low viscosity.

▲ **CRITICAL STEP** Complete lysis and mechanical or enzymatic breakdown of DNA before pelleting insoluble material is essential to remove nucleic acids and contaminating proteins from inclusion bodies. The lysis process can be assisted by tip sonication on ice.

7) Add ~500 uL glass beads (0.5-1 mm diameter) to each 50 mL tube of lysate.

8) Centrifuge lysate for 30' at 3220 g (Eppendorf Centrifuge 5810, or at maximum speed with similar swinging bucket tabletop centrifuge). Carefully take off supernatant without disturbing white pellet and resuspend pellet in 25 mL IW by vortexing.

9) Pellet the inclusion bodies for 30' at 3220 g (Eppendorf Centrifuge 5810). Discard supernatant and resuspend pellet in 25 mL IW by vortexing.

10) Pellet the inclusion bodies for 30' at 3220 g (Eppendorf Centrifuge 5810) and discard supernatant.

■ **PAUSE POINT** Inclusion body pellet can flash-frozen in liquid nitrogen and stored indefinitely at -80°C. Continue with step 17).

**Inclusion body extraction and denaturing Nickel Affinity Purification ● TIMING 4 h**

17) Soak inclusion bodies in 1 mL DMSO and loosen pellet by vigorous vortexing.

18) Add 25 mL D500 to the suspension, vortex, then and slowly rotate at 4°C for ~1 hour.

19) Spin down insoluble material (15 min x 30,000 g, 4°C), continue with supernatant.

▲ **CRITICAL STEP** Residual insoluble material should be clear and gel-like, a white firm pellet hints to incomplete resolubilization of the histones. In this case, repeat extraction with another 25 mL D500.

20) Equilibrate 2-5 mL Ni-NTA resin with 10 mL D500 in a 50 mL chromatography column and drain buffer.

▲ **CRITICAL STEP** Adapt Ni-NTA volume to expected yield. The capacity for histones under denaturing conditions is usually close to the maximum capacity indicated by the manufacturer (typically 20 mg/mL resin bed volume). Histone H2A, H2B and H3 typically yield > 100 mg per liter TBA/LB culture, H4 between 10 and 50 mg.

21) Add D500 inclusion body supernatant to Ni-NTA and slowly rotate 0.5-1 hour at 4°C.

22) Drain column, apply 25-50 mL D500. (optional: verify that you are not saturating the affinity resin by saving the flow through for gel analysis. In some cases, there is sufficient residual histone in the flow though for a second round of Ni-chromatography, in which case, the flow through may be flash frozen and stored at -80° until needed.)

23) Wash with 2 x 15-20 mL D1000.

24) Elute 3x with 2 bed volumes of DEB. Ni-NTA should change color to deep blue and can be reused or regenerated according to manufactures protocols. Quantify eluate by 280 nm extinction measurement against elution buffer (see box). Add 5 mM EDTA to the eluate to prevent leached Ni-mediated his-tag aggregation. ■ **PAUSE POINT** Eluate can be stored at -20°C or used for refolding directly.

25) Dialyze histones in a 3.5k MWCO membrane against ddH2O, 1% (v/v) acetic acid, 0.2 mM PMSF, 5 mM 2-mercaptoethanol for 24 hours.

26) Quantify histones against dialysis solution, add Tris pH7 to 100 mM for tag cleavage or lyophilize.

▲ **CRITICAL STEP** Quantify protein concentration in eluate and check purity on gel (See Box). Check for DNA contamination by OD260/280 reading.

**Octamer preparation ● TIMING 3 d**

For determining the appropriate mixing ratios, you will need all four core histones H2A, H2B, H3 and H4 in a D1000 or DEB solution. You can use the individual Ni-NTA elution directly (concentrate in spin concentrator if the concentration is <1 mg/mL) or reconstitute solutions from individual lyophilized core histones.

27) Accurately determine the protein concentration by measuring the 280 nm absorption against an identical buffer reference in triplicate (or the concentrator flowthrough).

28) Mix histones in a ratio 1.1:1.1:1:1 H2A:H2B:H3:H4. Your target protein concentration should be within 0.5 and 5 mg/mL.

29) Add glycerol to 15% final and NaCl to 2 M final, transfer the histone mix to a 3.5k MWCO membrane dialysis bag.

30) Dialyze for >12 h against 1-2L of ORB

31) If the reconstitution is performed with tagged histone, take out dialysis bag and add a 20 $\mu$g TEV, 5 $\mu$g R3C per mg total protein.

32) Dialyze against fresh 1-2L of ORB for 24 h.

33) Check cleavage of all four tags on gel. If there is uncleaved histone left, add another 20 $\mu$g TEV and/or 5 $\mu$g R3C per mg total protein.

34) Purify octamer on Superdex 200 column with ORB or desired final buffer.

**H3-H4 Tetramer preparation ● TIMING 3 d**

27) Mix histones in a ratio 1:1 H3:H4. Your target protein concentration should be within 0.5 and 5 mg/mL.

28) Add glycerol to 15%, transfer the histone mix to a 3.5k MWCO membrane dialysis bag.

29) Dialyze for >12 h against 1-2L of TRB

30) Take our dialysis bag and add a 10 $\mu$g TEV per mg total protein.

31) Dialyze against fresh 1-2L of TRB for 12 h.

32) Check cleavage of tags on gel. If there is uncleaved histone left, add another 10 $\mu$g TEV and incubate for 12h at 4¬C.

33) Purify octamer on Superdex 200 column with TRB or desired final buffer.

# BIBLIOGRAPHY


Adkins, M.W., Carson, J.J., English, C.M., Ramey, C.J., and Tyler, J.K. (2007). The histone chaperone anti-silencing function 1 stimulates the acetylation of newly synthesized histone H3 in S-phase. J Biol Chem *282*, 1334-1340.

Adkins, M.W., and Tyler, J.K. (2004). The histone chaperone Asf1p mediates global chromatin disassembly in vivo. J Biol Chem *279*, 52069-52074.

Adkins, M.W., and Tyler, J.K. (2006). Transcriptional activators are dispensable for transcription in the absence of Spt6-mediated chromatin reassembly of promoter regions. Mol Cell *21*, 405-416.

Agez, M., Chen, J., Guerois, R., van Heijenoort, C., Thuret, J.-Y., Mann, C., and Ochsenbein, F. (2007). Structure of the histone chaperone ASF1 bound to the histone H3 C-terminal helix and functional insights. Structure *15*, 191-199.

Ahmad, K., and Henikoff, S. (2002a). Epigenetic consequences of nucleosome dynamics. Cell *111*, 281-284.

Ahmad, K., and Henikoff, S. (2002b). Histone H3 variants specify modes of chromatin assembly. Proc Natl Acad Sci USA *99 Suppl 4*, 16477-16484.

Ahmad, K., and Henikoff, S. (2002c). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell *9*, 1191-1200.

Akhmanova, A., Miedema, K., Kremer, H., and Hennig, W. (1997). Two types of polyadenated mRNAs are synthesized from Drosophila replication-dependent histone genes. Eur J Biochem *244*, 294-300.

Allis, C.D., Glover, C.V., Bowen, J.K., and Gorovsky, M.A. (1980). Histone variants specific to the transcriptionally active, amitotically dividing macronucleus of the unicellular eucaryote, Tetrahymena thermophila. Cell *20*, 609-617.

Allis, C.D., Richman, R., Gorovsky, M.A., Ziegler, Y.S., Touchstone, B., Bradley, W.A., and Cook, R.G. (1986). hv1 is an evolutionarily conserved H2A variant that is preferentially associated with active genes. J Biol Chem *261*, 1941-1948.

Allis, C.D., Ziegler, Y.S., Gorovsky, M.A., and Olmsted, J.B. (1982). A conserved histone variant enriched in nucleoli of mammalian cells. Cell *31*, 131-136.

Alvarez, F., Munoz, F., Schilcher, P., Imhof, A., Almozuni, G., and Loyola, A. (2011). Sequential establishment of marks on soluble histones H3 and H4. Journal of Biological Chemistry.

Anderson, H., Wardle, J., Korkut, S., Murton, H., López-Maury, L., Bähler, J., and Whitehall, S. (2009). The fission yeast HIRA histone chaperone is required for promoter silencing and the suppression of cryptic antisense transcripts. Mol Cell Biol.

Andrews, A.J., Chen, X., Zevin, A., Stargell, L.A., and Luger, K. (2010). The Histone Chaperone Nap1 Promotes Nucleosome Assembly by Eliminating Nonnucleosomal Histone DNA Interactions. Mol Cell *37*, 834-842.

Andrews, A.J., Downing, G., Brown, K., Park, Y.-J., and Luger, K. (2008). A thermodynamic model for Nap1-histone interactions. J Biol Chem *283*, 32412-32418.

Andrews, A.J., and Luger, K. (2011). A coupled equilibrium approach to study nucleosome thermodynamics. Meth Enzymol *488*, 265-285.

Anju, V., Kapros, T., and Waterborg, J.H. (2011). Identification of a replication-independent replacement histone H3 in the basidiomycete Ustilago maydis. Journal of Biological Chemistry *286*, 25790-25800.

Argentaro, A., Yang, J.-C., Chapman, L., Kowalczyk, M.S., Gibbons, R.J., Higgs, D.R., Neuhaus, D., and Rhodes, D. (2007). Structural consequences of disease-causing mutations in the ATRX-DNMT3-DNMT3L (ADD) domain of the chromatin-associated protein ATRX. Proc Natl Acad Sci USA *104*, 11939-11944.

Ashkenazy, H., Erez, E., Martz, E., Pupko, T., and Ben-Tal, N. (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res *38*, W529-533.

Balaji, S., Iyer, L.M., and Aravind, L. (2009). HPC2 and ubinuclein define a novel family of histone chaperones conserved throughout eukaryotes. Mol BioSyst *5*, 269-275.

Banaszynski, L.A., Allis, C.D., and Lewis, P.W. (2010). Histone variants in metazoan development. Dev Cell *19*, 662-674.

Banks, D.D., and Gloss, L.M. (2003). Equilibrium folding of the core histones: the H3-H4 tetramer is less stable than the H2A-H2B dimer. Biochemistry *42*, 6827-6839.

Banks, D.D., and Gloss, L.M. (2004). Folding mechanism of the (H3-H4)2 histone tetramer of the core nucleosome. Protein Sci *13*, 1304-1316.

Banumathy, G., Somaiah, N., Zhang, R., Tang, Y., Hoffmann, J., Andrake, M., Ceulemans, H., Schultz, D., Marmorstein, R., and Adams, P.D. (2009). Human UBN1 is an ortholog of yeast Hpc2p and has an essential role in the HIRA/ASF1a chromatin-remodeling pathway in senescent cells. Mol Cell Biol *29*, 758-770.

Bardossi, F. (1980). Gene Control and Enzymatic Zippers. In Research Profiles (New York, The Rockefeller University).

Barman, H.K., Takami, Y., Nishijima, H., Shibahara, K.-i., Sanematsu, F., and Nakayama, T. (2008). Histone acetyltransferase-1 regulates integrity of cytosolic histone H3-H4 containing complex. Biochem Biophys Res Commun *373*, 624-630.

Barnhart, M.C., Kuich, P.H.J.L., Stellfox, M.E., Ward, J.A., Bassett, E.A., Black, B.E., and Foltz, D.R. (2011). HJURP is a CENP-A chromatin assembly factor sufficient to form a functional de novo kinetochore. J Cell Biol.

Bassett, A.R., Cooper, S.E., Ragab, A., and Travers, A.A. (2008). The chromatin remodelling factor dATRX is involved in heterochromatin formation. PLoS ONE *3*, e2099.

Baumann, C., and de La Fuente, R. (2009). ATRX marks the inactive X chromosome (Xi) in somatic cells and during imprinted X chromosome inactivation in trophoblast stem cells. Chromosoma *118*, 209-222.

Baumann, C., Schmidtmann, A., Muegge, K., and de La Fuente, R. (2008). Association of ATRX with pericentric heterochromatin and the Y chromosome of neonatal mouse spermatogonia. BMC Mol Biol *9*, 29.

Baxevanis, A., Godfrey, J., and Moudrianakis, E. (1991). Associative behavior of the histone (H3-H4) 2 tetramer: dependence on ionic environment. Biochemistry.

Bell, O., and Schübeler, D. (2009). Chromatin: sub out the replacement. Curr Biol *19*, R545-547.

Belotserkovskaya, R. (2003). FACT Facilitates Transcription-Dependent Nucleosome Alteration. Science *301*, 1090-1093.

Bergmann, J.H., Rodríguez, M.G., Martins, N.M.C., Kimura, H., Kelly, D.A., Masumoto, H., Larionov, V., Jansen, L.E.T., and Earnshaw, W.C. (2011). Epigenetic engineering shows H3K4me2 is required for HJURP targeting and CENP-A assembly on a synthetic human kinetochore. EMBO J *30*, 328-340.

Bernstein, B.E., Meissner, A., and Lander, E.S. (2007). The mammalian epigenome. Cell *128*, 669-681.

Bernstein, E., and Hake, S.B. (2006). The nucleosome: a little variation goes a long way. Biochem Cell Biol *84*, 505-517.

Black, B.E., and Cleveland, D.W. (2011). Epigenetic centromere propagation and the nature of CENP-a nucleosomes. Cell *144*, 471-479.

Black, B.E., Foltz, D.R., Chakravarthy, S., Luger, K., Woods, V.L., and Cleveland, D.W. (2004). Structural determinants for generating centromeric chromatin. Nature *430*, 578-582.

Black, B.E., Jansen, L.E.T., Maddox, P.S., Foltz, D.R., Desai, A.B., Shah, J.V., and Cleveland, D.W. (2007). Centromere identity maintained by nucleosomes assembled with histone H3 containing the CENP-A targeting domain. Mol Cell *25*, 309-322.

Blackburn, E.H. (1991). Telomeres. Trends Biochem Sci *16*, 378-381.

Blackwell, C., Martin, K.A., Greenall, A., Pidoux, A., Allshire, R.C., and Whitehall, S.K. (2004). The Schizosaccharomyces pombe HIRA-like protein Hip1 is required for the periodic expression of histone genes and contributes to the function of complex centromeres. Mol Cell Biol *24*, 4309-4320.

Bohm, V., Hieb, A.R., Andrews, A.J., Gansen, A., Rocker, A., Toth, K., Luger, K., and Langowski, J. (2011). Nucleosome accessibility governed by the dimer/tetramer interface. Nucleic Acids Res *39*, 3093-3102.

Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular signals of epigenetic states. Science *330*, 612-616.

Bonnefoy, E., Orsi, G.A., Couble, P., and Loppin, B. (2007). The essential role of Drosophila HIRA for de novo assembly of paternal chromatin at fertilization. PLoS Genet *3*, 1991-2006.

Bortvin, A., and Winston, F. (1996). Evidence that Spt6p controls chromatin structure by a direct interaction with histones. Science *272*, 1473-1476.

Bowman, A., Ward, R., Wiechens, N., Singh, V., El-Mkami, H., Norman, D.G., and Owen-Hughes, T. (2011). The Histone Chaperones Nap1 and Vps75 Bind Histones H3 and H4 in a Tetrameric Conformation. Mol Cell *41*, 398-408.

Braunschweig, U., Hogan, G.J., Pagie, L., and van Steensel, B. (2009). Histone H1 binding is inhibited by histone variant H3.3. EMBO J *28*, 3635-3645.

Bryant, G.O., Prabhu, V., Floer, M., Wang, X., Spagna, D., Schreiber, D., and Ptashne, M. (2008). Activator Control of Nucleosome Occupancy in Activation and Repression of Transcription. PLoS Biol *6*, e317.

Cacchione, S., Cerone, M.A., and Savino, M. (1997). In vitro low propensity to form nucleosomes of four telomeric sequences. FEBS Lett *400*, 37-41.

Camahort, R., Li, B., Florens, L., Swanson, S.K., Washburn, M.P., and Gerton, J.L. (2007). Scm3 is essential to recruit the histone h3 variant cse4 to centromeres and to maintain a functional kinetochore. Mol Cell *26*, 853-865.

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

Cho, U.-S., and Harrison, S.C. (2011). Recognition of the centromere-specific histone Cse4 by the chaperone Scm3. Proc Natl Acad Sci USA *108*, 9367-9371.

Choi, E.S., Shin, J.A., Kim, H.S., and Jang, Y.K. (2005). Dynamic regulation of replication independent deposition of histone H3 in fission yeast. Nucleic Acids Res *33*, 7102-7110.

Choo, K.H. (2001). Domain organization at the centromere and neocentromere. Dev Cell *1*, 165-177.

Chow, C.-M., Georgiou, A., Szutorisz, H., Maia e Silva, A., Pombo, A., Barahona, I., Dargelos, E., Canzonetta, C., and Dillon, N. (2005). Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. EMBO Rep *6*, 354-360.

Cleveland, D.W., Mao, Y., and Sullivan, K.F. (2003). Centromeres and kinetochores: from epigenetics to mitotic checkpoint signaling. Cell *112*, 407-421.

Cole, C., Barber, J.D., and Barton, G.J. (2008). The Jpred 3 secondary structure prediction server. Nucleic Acids Res *36*, W197-201.

Cosgrove, M.S. (2007). Histone proteomics and the epigenetic regulation of nucleosome mobility. Expert review of proteomics *4*, 465-478.

Cosgrove, M.S., Boeke, J.D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. Nat Struct Mol Biol *11*, 1037-1043.

Cotten, M., Sealy, L., and Chalkley, R. (1986). Massive phosphorylation distinguishes Xenopus laevis nucleoplasmin isolated from oocytes or unfertilized eggs. Biochemistry *25*, 5063-5069.

Couldrey, C., Carlton, M.B., Nolan, P.M., Colledge, W.H., and Evans, M.J. (1999). A retroviral gene trap insertion into the histone 3.3A gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice. Hum Mol Genet *8*, 2489-2495.

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

Dalgliesh, G.L., Furge, K., Greenman, C., Chen, L., Bignell, G., Butler, A., Davies, H., Edkins, S., Hardy, C., Latimer, C., *et al.* (2010). Systematic sequencing of renal carcinoma reveals inactivation of histone modifying genes. Nature *463*, 360-363.

Das, C., Lucia, M.S., Hansen, K.C., and Tyler, J.K. (2009). CBP/p300-mediated acetylation of histone H3 on lysine 56. Nature *459*, 113-117.

Das, R., and Baker, D. (2008). Macromolecular modeling with rosetta. Annu Rev Biochem *77*, 363-382.

Daury, L., Chailleux, C., Bonvallet, J., and Trouche, D. (2006). Histone H3.3 deposition at E2F-regulated genes is linked to transcription. EMBO Rep *7*, 66-71.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 a resolution. J Mol Biol *319*, 1097-1113.

de La Fuente, R., Viveiros, M.M., Wigglesworth, K., and Eppig, J.J. (2004). ATRX, a member of the SNF2 family of helicase/ATPases, is required for chromosome alignment and meiotic spindle organization in metaphase II stage mouse oocytes. Developmental Biology *272*, 1-14.

Deal, R.B., Henikoff, J.G., and Henikoff, S. (2010). Genome-wide kinetics of nucleosome turnover determined by metabolic labeling of histones. Science *328*, 1161-1164.

Dechassa, M.L., Wyns, K., Li, M., Hall, M.A., Wang, M.D., and Luger, K. (2011). Structure and Scm3-mediated assembly of budding yeast centromeric nucleosomes. Nat Commun *2*, 313.

Derewenda, Z.S. (2004). Rational protein crystallization by mutational surface engineering. Structure/Folding and Design *12*, 529-535.

Dhayalan, A., Tamas, R., Bock, I., Tattermusch, A., Dimitrova, E., Kudithipudi, S., Ragozin, S., and Jeltsch, A. (2011). The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. Hum Mol Genet *20*, 2195-2203.

Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983). Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res *11*, 1475-1489.

Dion, M.F., Kaplan, T., Kim, M., Buratowski, S., Friedman, N., and Rando, O.J. (2007). Dynamics of replication-independent histone turnover in budding yeast. Science *315*, 1405-1408.

Dominski, Z., and Marzluff, W.F. (1999). Formation of the 3' end of histone mRNA. Gene *239*, 1-14.

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

Drané, P., Ouararhni, K., Depaux, A., Shuaib, M., and Hamiche, A. (2010). The death-associated protein DAXX is a novel histone chaperone involved in the replication-independent deposition of H3.3. Genes Dev *24*, 1253-1265.

Dunleavy, E. (2011). H3.3 is deposited at centromeres in S phase as a placeholder for newly assembled CENP-A in G1 phase. 1-12.

Dunleavy, E.M., Roche, D., Tagami, H., Lacoste, N., Ray-Gallet, D., Nakamura, Y., Daigo, Y., Nakatani, Y., and Almouzni-Pettinotti, G. (2009). HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. Cell *137*, 485-497.

Dutta, S., Akey, I.V., Dingwall, C., Hartman, K.L., Laue, T., Nolte, R.T., Head, J.F., and Akey, C.W. (2001). The crystal structure of nucleoplasmin-core: implications for histone binding and nucleosome assembly. Mol Cell *8*, 841-853.

Earnshaw, W.C., and Migeon, B.R. (1985). Three related centromere proteins are absent from the inactive centromere of a stable isodicentric chromosome. Chromosoma *92*, 290-296.

Eddy, S.R. (1998). Profile hidden Markov models. Bioinformatics *14*, 755-763.

Egelhofer, T.A., Minoda, A., Klugman, S., Lee, K., Kolasinska-Zwierz, P., Alekseyenko, A.A., Cheung, M.-S., Day, D.S., Gadel, S., Gorchakov, A.A., *et al.* (2010). An assessment of histone-modification antibody quality. Nat Struct Mol Biol *18*, 91-93.

Eickbush, T.H., and Moudrianakis, E.N. (1978). The histone core complex: an octamer assembled by two sets of protein-protein interactions. Biochemistry *17*, 4955-4964.

Elsaesser, S.J., and Allis, C.D. (2010). HIRA and Daxx constitute two independent histone H3.3-containing predeposition complexes. Cold Spring Harb Symp Quant Biol *75*, 27-34.

Elsässer, S.J., Allis, C.D., and Lewis, P.W. (2011). Cancer. New epigenetic drivers of cancers. Science *331*, 1145-1146.

Emelyanov, A.V., Konev, A.Y., Vershilova, E., and Fyodorov, D.V. (2010). Protein complex of Drosophila ATRX/XNP and HP1a is required for the formation of pericentric beta-heterochromatin in vivo. J Biol Chem *285*, 15027-15037.

English, C.M., Adkins, M.W., Carson, J.J., Churchill, M.E.A., and Tyler, J.K. (2006). Structural basis for the histone chaperone activity of Asf1. Cell *127*, 495-508.

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

Escobar-Cabrera, E., Lau, D.K.W., Giovinazzi, S., Ishov, A.M., and Mcintosh, L.P. (2010). Structural characterization of the DAXX N-terminal helical bundle domain and its complex with Rassf1C. Structure *18*, 1642-1653.

Eustermann, S., Yang, J.-C., Law, M.J., Amos, R., Chapman, L.M., Jelinska, C., Garrick, D., Clynes, D., Gibbons, R.J., Rhodes, D., *et al.* (2011). Combinatorial readout of histone H3 modifications specifies localization of ATRX to heterochromatin. Nat Struct Mol Biol *18*, 777-782.

Felsenstein, J. (1989). PHYLIP (phylogeny inference package) (Cladistics).

Fichera, M., Romano, C., Castiglia, L., Failla, P., Ruberto, C., Amata, S., Greco, D., Cardoso, C., Fontés, M., and Ragusa, A. (1998). New mutations in XNP/ATR-X gene: a further contribution to genotype/phenotype relationship in ATR/X syndrome. Mutations in brief no. 176. Online. Hum Mutat *12*, 214.

Filion, G.J., van Bemmel, J.G., Braunschweig, U., Talhout, W., Kind, J., Ward, L.D., Brugman, W., de Castro, I.J., Kerkhoven, R.M., Bussemaker, H.J., *et al.* (2010). Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. Cell *143*, 212-224.

Fillingham, J., Kainth, P., Lambert, J.-P., Bakel, H.v., Tsui, K., Pena-Castillo, L., Nislow, C., Figeys, D., Hughes, T.R., Greenblatt, J., *et al.* (2009). Two-Color Cell Array Screen Reveals Interdependent Roles for Histone Chaperones and a Chromatin Boundary Regulator in Histone Gene Repression. Mol Cell *35*, 340-351.

Finn, R., Browne, K., Hodgson, K., and Ausio, J. (2008). sNASP, a histone H1-specific eukaryotic chaperone dimer that facilitates chromatin assembly. Biophys J.

Flaus, A., Rencurel, C., Ferreira, H., Wiechens, N., and Owen-Hughes, T. (2004). Sin mutations alter inherent nucleosome mobility. EMBO J *23*, 343-353.

Foltz, D.R., Jansen, L.E.T., Bailey, A.O., Yates, J.R., Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. Cell *137*, 472-484.

Formosa, T., Ruone, S., Adams, M.D., Olsen, A.E., Eriksson, P., Yu, Y., Rhoades, A.R., Kaufman, P.D., and Stillman, D.J. (2002). Defects in SPT16 or POB3 (yFACT) in Saccharomyces cerevisiae cause dependence on the Hir/Hpc pathway: polymerase passage may degrade chromatin structure. Genetics *162*, 1557-1571.

Franklin, S.G., and Zweidler, A. (1977). Non-allelic variants of histones 2a, 2b and 3 in mammals. Nature *266*, 273-275.

Fry, C.J., Norris, A., Cosgrove, M., Boeke, J.D., and Peterson, C.L. (2006). The LRS and SIN domains: two structurally equivalent but functionally distinct nucleosomal surfaces required for transcriptional silencing. Mol Cell Biol *26*, 9045-9059.

Furuyama, T., Dalal, Y., and Henikoff, S. (2006). Chaperone-mediated assembly of centromeric chromatin in vitro. Proc Natl Acad Sci USA *103*, 6172-6177.

Furuyama, T., and Henikoff, S. (2009). Centromeric nucleosomes induce positive DNA supercoils. Cell *138*, 104-113.

Gadad, S.S., Senapati, P., Syed, S.H., Rajan, R.E., Shandilya, J., Swaminathan, V., Chatterjee, S., Colombo, E., Dimitrov, S., Pelicci, P.G., *et al.* (2011). The Multifunctional Protein Nucleophosmin (NPM1) Is a Human Linker Histone H1 Chaperone. Biochemistry.

Gallagher, S.R. (2006). One-Dimensional SDS Gel Electrophoresis of Proteins. In Current Protocols in Immunology (John Wiley & Sons, Inc.).

García-Cao, M., O'Sullivan, R., Peters, A.H.F.M., Jenuwein, T., and Blasco, M.A. (2004). Epigenetic regulation of telomere length in mammalian cells by the Suv39h1 and Suv39h2 histone methyltransferases. Nat Genet *36*, 94-99.

Gibbons, R. (2006). Alpha thalassaemia-mental retardation, X linked. Orphanet journal of rare diseases *1*, 15.

Gibbons, R.J., McDowell, T.L., Raman, S., O'Rourke, D.M., Garrick, D., Ayyub, H., and Higgs, D.R. (2000). Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. Nat Genet *24*, 368-371.

Gibbons, R.J., Wada, T., Fisher, C.A., Malik, N., Mitson, M.J., Steensma, D.P., Fryer, A., Goudie, D.R., Krantz, I.D., and Traeger-Synodinos, J. (2008). Mutations in the chromatin-associated protein ATRX. Hum Mutat *29*, 796-802.

Gkikopoulos, T., Havas, K.M., Dewar, H., and Owen-Hughes, T. (2009). SWI/SNF and Asf1p cooperate to displace histones during induction of the saccharomyces cerevisiae HO promoter. Mol Cell Biol *29*, 4057-4066.

Glowczewski, L., Waterborg, J.H., and Berman, J.G. (2004). Yeast chromatin assembly complex 1 protein excludes nonacetylatable forms of histone H4 from chromatin and the nucleus. Mol Cell Biol *24*, 10180-10192.

Goldberg, A.D., Banaszynski, L.A., Noh, K.-M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., *et al.* (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell *140*, 678-691.

Green, E.M., Antczak, A.J., Bailey, A.O., Franco, A.A., Wu, K.J., Yates, J.R., and Kaufman, P.D. (2005). Replication-independent histone deposition by the HIR complex and Asf1. Curr Biol *15*, 2044-2049.

Groth, A., Corpet, A., Cook, A.J.L., Roche, D., Bartek, J., Lukas, J., and Almouzni, G. (2007). Regulation of replication fork progression through histone supply and demand. Science *318*, 1928-1931.

Hake, S.B., Garcia, B.A., Duncan, E.M., Kauer, M., Dellaire, G., Shabanowitz, J., Bazett-Jones, D.P., Allis, C.D., and Hunt, D.F. (2006). Expression patterns and post-translational modifications associated with mammalian histone H3 variants. J Biol Chem *281*, 559-568.

Hake, S.B., Garcia, B.A., Kauer, M., Baker, S.P., Shabanowitz, J., Hunt, D.F., and Allis, C.D. (2005). Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. Proc Natl Acad Sci USA *102*, 6344-6349.

Hall, C., Nelson, D.M., Ye, X., Baker, K., DeCaprio, J.A., Seeholzer, S., Lipinski, M., and Adams, P.D. (2001). HIRA, the human homologue of yeast Hir1p and Hir2p, is a novel

cyclin-cdk2 substrate whose expression blocks S-phase progression. Mol Cell Biol *21*, 1854-1865.

Hamiche, A., Kang, J.G., Dennis, C., Xiao, H., and Wu, C. (2001). Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. Proc Natl Acad Sci USA *98*, 14316-14321.

Han, J., Li, Q., Mccullough, L., Kettelkamp, C., Formosa, T., and Zhang, Z. (2010). Ubiquitylation of FACT by the Cullin-E3 ligase Rtt101 connects FACT to DNA replication. Genes Dev *24*, 1485-1490.

Harris, M.E., Böhni, R., Schneiderman, M.H., Ramamurthy, L., Schümperli, D., and Marzluff, W.F. (1991). Regulation of histone mRNA in the unperturbed cell cycle: evidence suggesting control at two posttranscriptional steps. Mol Cell Biol *11*, 2416-2424.

Heaphy, C.M., de Wilde, R.F., Jiao, Y., Klein, A.P., Edil, B.H., Shi, C., Bettegowda, C., Rodriguez, F.J., Eberhart, C.G., Hebbar, S., *et al.* (2011). Altered telomeres in tumors with ATRX and DAXX mutations. Science *333*, 425.

Henikoff, S. (1990). Position-effect variegation after 60 years. Trends Genet *6*, 422-426.

Henikoff, S., McKittrick, E., and Ahmad, K. (2004). Epigenetics, histone H3 variants, and the inheritance of chromatin states. Cold Spring Harb Symp Quant Biol *69*, 235-243.

Henikoff, S., and Shilatifard, A. (2011). Histone modification: cause or cog? Trends Genet.

Heun, P., Erhardt, S., Blower, M.D., Weiss, S., Skora, A.D., and Karpen, G.H. (2006). Mislocalization of the Drosophila centromere-specific histone CID promotes formation of functional ectopic kinetochores. Dev Cell *10*, 303-315.

Hewawasam, G., Shivaraju, M., Mattingly, M., Venkatesh, S., Martin-Brown, S., Florens, L., Workman, J.L., and Gerton, J.L. (2010). Psh1 is an E3 ubiquitin ligase that targets the centromeric histone variant Cse4. Mol Cell *40*, 444-454.

Hödl, M., and Basler, K. (2009). Transcription in the absence of histone H3.3. Curr Biol *19*, 1221-1226.

Hollenbach, A.D., McPherson, C.J., Mientjes, E.J., Iyengar, R., and Grosveld, G. (2002). Daxx and histone deacetylase II associate with chromatin through an interaction with core histones and the chromatin-associated protein Dek. J Cell Sci *115*, 3319-3330.

Howman, E.V., Fowler, K.J., Newson, A.J., Redward, S., MacDonald, A.C., Kalitsis, P., and Choo, K.H. (2000). Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. Proc Natl Acad Sci USA *97*, 1148-1153.

Hu, H., Liu, Y., Wang, M., Fang, J., Huang, H., Yang, N., Li, Y., Wang, J., Yao, X., Shi, Y., *et al.* (2011). Structure of a CENP-A-histone H4 heterodimer in complex with chaperone HJURP. Genes Dev.

Hu, W., Feng, Z., Modica, I., Klimstra, D.S., Song, L., Allen, P.J., Brennan, M.F., Levine, A.J., and Tang, L.H. (2010). Gene Amplifications in Well-Differentiated Pancreatic Neuroendocrine Tumors Inactivate the p53 Pathway. Genes Cancer *1*, 360-368.

Ishimi, Y., Yasuda, H., Hirosumi, J., Hanaoka, F., and Yamada, M. (1983). A protein which facilitates assembly of nucleosome-like structures in vitro in mammalian cells. J Biochem *94*, 735-744.

Ishov, A.M., Vladimirova, O.V., and Maul, G.G. (2004). Heterochromatin and ND10 are cell-cycle regulated and phosphorylation-dependent alternate nuclear sites of the transcription repressor Daxx and SWI/SNF protein ATRX. J Cell Sci *117*, 3807-3820.

Ito, T., Tyler, J.K., and Kadonaga, J.T. (1997). Chromatin assembly factors: a dual function in nucleosome formation and mobilization? Genes Cells *2*, 593-600.

Iwase, S., Xiang, B., Ghosh, S., Ren, T., Lewis, P.W., Cochrane, J.C., Allis, C.D., Picketts, D.J., Patel, D.J., Li, H., *et al.* (2011). ATRX ADD domain links an atypical histone methylation recognition mechanism to human mental-retardation syndrome. Nat Struct Mol Biol *18*, 769-776.

Jacobs, S.A., and Khorasanizadeh, S. (2002). Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. Science *295*, 2080-2083.

Jamai, A., Imoberdorf, R.M., and Strubin, M. (2007). Continuous histone H2B and transcription-dependent histone H3 exchange in yeast cells outside of replication. Mol Cell *25*, 345-355.

Jamai, A., Puglisi, A., and Strubin, M. (2009). Histone Chaperone Spt16 Promotes Redeposition of the Original H3-H4 Histones Evicted by Elongating RNA Polymerase. Mol Cell *35*, 377-383.

Jang, H., Choi, S.-Y., Cho, E.-J., and Youn, H.-D. (2009). Cabin1 restrains p53 activity on chromatin. Nat Struct Mol Biol *16*, 910-915.

Janicki, S.M., Tsukamoto, T., Salghetti, S.E., Tansey, W.P., Sachidanandam, R., Prasanth, K.V., Ried, T., Shav-Tal, Y., Bertrand, E., Singer, R.H., *et al.* (2004). From silencing to gene expression: real-time analysis in single cells. Cell *116*, 683-698.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. Science *293*, 1074-1080.

Jiao, Y., Shi, C., Edil, B.H., de Wilde, R.F., Klimstra, D.S., Maitra, A., Schulick, R.D., Tang, L.H., Wolfgang, C.L., Choti, M.A., *et al.* (2011). DAXX/ATRX, MEN1, and mTOR Pathway Genes Are Frequently Altered in Pancreatic Neuroendocrine Tumors. Science.

Jin, C., and Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes Dev *21*, 1519-1529.

Jin, C., Zang, C., Wei, G., Cui, K., Peng, W., Zhao, K., and Felsenfeld, G. (2009). H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. Nat Genet *41*, 941-945.

Jones, S., Wang, T.-L., Shih, I.-M., Mao, T.-L., Nakayama, K., Roden, R., Glas, R., Slamon, D., Diaz, L.A., Vogelstein, B., *et al.* (2010). Frequent mutations of chromatin remodeling gene ARID1A in ovarian clear cell carcinoma. Science *330*, 228-231.

Kang, B., Pu, M., Hu, G., Wen, W., Dong, Z., Zhao, K., Stillman, B., and Zhang, Z. (2011). Phosphorylation of H4 Ser 47 promotes HIRA-mediated nucleosome assembly. Genes Dev *25*, 1359-1364.

Karantza, V., Freire, E., and Moudrianakis, E.N. (1996). Thermodynamic studies of the core histones: pH and ionic strength effects on the stability of the (H3-H4)/(H3-H4)2 system. Biochemistry *35*, 2037-2046.

Kaufman, P.D., Kobayashi, R., Kessler, N., and Stillman, B. (1995). The p150 and p60 subunits of chromatin assembly factor I: a molecular link between newly synthesized histones and DNA replication. Cell *81*, 1105-1114.

Kepert, J.F., Mazurkiewicz, J., Heuvelman, G.L., Tóth, K.F., and Rippe, K. (2005). NAP1 modulates binding of linker histone H1 to chromatin and induces an extended chromatin fiber conformation. J Biol Chem *280*, 34063-34072.

Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., Minoda, A., Riddle, N.C., Ernst, J., Sabo, P.J., Larschan, E., Gorchakov, A.A., Gu, T., *et al.* (2011). Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. Nature *471*, 480-485.

Kim, H.-J., Seol, J.-H., and Cho, E.-J. (2009). Potential role of the histone chaperone, CAF-1, in transcription. BMB Rep *42*, 227-231.

Kim, H.-J., Seol, J.-H., Han, J.-W., Youn, H.-D., and Cho, E.-J. (2007). Histone chaperones regulate histone exchange during transcription. EMBO J *26*, 4467-4474.

Kim, Y., Quartey, P., Li, H., Volkart, L., Hatzos, C., Chang, C., Nocek, B., Cuff, M., Osipiuk, J., Tan, K., *et al.* (2008). Large-scale evaluation of protein reductive methylation for improving protein crystallization. Nat Methods *5*, 853-854.

Kimura, H. (2005). Histone dynamics in living cells revealed by photobleaching. DNA Repair (Amst) *4*, 939-950.

Kimura, H., and Cook, P.R. (2001). Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. J Cell Biol *153*, 1341-1353.

Konev, A.Y., Tribus, M., Park, S.Y., Podhraski, V., Lim, C.Y., Emelyanov, A.V., Vershilova, E., Pirrotta, V., Kadonaga, J.T., Lusser, A., *et al.* (2007). CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. Science *317*, 1087-1090.

Koning, d., Corpet, Haber, and Almouzni (2007). Histone chaperones: an escort network regulating histone traffic. Nat Struct Mol Biol *14*, 997-1007.

Kornberg, R.D., and Thomas, J.O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865-868.

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

Kundu, L.R., Seki, M., Watanabe, N., Murofushi, H., Furukohri, A., Waga, S., Score, A.J., Blow, J.J., Horikoshi, M., Enomoto, T., *et al.* (2011). Biphasic chromatin binding of histone chaperone FACT during eukaryotic chromatin DNA replication. Biochim Biophys Acta.

Lamour, V., Lécluse, Y., Desmaze, C., Spector, M., Bodescot, M., Aurias, A., Osley, M.A., and Lipinski, M. (1995). A human homolog of the S. cerevisiae HIR1 and HIR2 transcriptional repressors cloned from the DiGeorge syndrome critical region. Hum Mol Genet *4*, 791-799.

Längst, G., and Becker, P.B. (2001). ISWI induces nucleosome sliding on nicked DNA. Mol Cell *8*, 1085-1092.

Laskey, R.A., Honda, B.M., Mills, A.D., and Finch, J.T. (1978a). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. Nature *275*, 416-420.

Laskey, R.A., Honda, B.M., Mills, A.D., Morris, N.R., Wyllie, A.H., Mertz, J.E., De Roberts, E.M., and Gurdon, J.B. (1978b). Chromatin assembly and transcription in eggs and oocytes of Xenopus laevis. Cold Spring Harb Symp Quant Biol *42 Pt 1*, 171-178.

Laskey, R.A., Mills, A.D., and Morris, N.R. (1977). Assembly of SV40 chromatin in a cell-free system from Xenopus eggs. Cell *10*, 237-243.

Law, M.J., Lower, K.M., Voon, H.P.J., Hughes, J.R., Garrick, D., Viprakasit, V., Mitson, M., de Gobbi, M., Marra, M., Morris, A., *et al.* (2010). ATR-X syndrome protein targets tandem repeats and influences allele-specific expression in a size-dependent manner. Cell *143*, 367-378.

Lefrançois, P., Euskirchen, G.M., Auerbach, R.K., Rozowsky, J., Gibson, T., Yellman, C.M., Gerstein, M., and Snyder, M. (2009). Efficient yeast ChIP-Seq using multiplex short-read DNA sequencing. BMC Genomics *10*, 37.

Lehnertz, B., Ueda, Y., Derijck, A.A.H.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H.F.M. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol *13*, 1192-1200.

Lejon, S., Thong, S.Y., Murthy, A., AlQarni, S., Murzina, N.V., Blobel, G.A., Laue, E.D., and Mackay, J.P. (2011). Insights into association of the NuRD complex with FOG-1 from the crystal structure of an RbAp48.FOG-1 complex. J Biol Chem *286*, 1196-1203.

Lewis, P.N. (1976). Histone-histone interactions. I. An electrophoretic study. Can J Biochem *54*, 641-649.

Lewis, P.W., Elsaesser, S.J., Noh, K.-M., Stadler, S.C., and Allis, C.D. (2010). Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. Proc Natl Acad Sci USA *107*, 14075-14080.

Li, Q., Zhou, H., Wurtele, H., Davies, B., Horazdovsky, B., Verreault, A., and Zhang, Z. (2008). Acetylation of histone H3 lysine 56 regulates replication-coupled nucleosome assembly. Cell *134*, 244-255.

Lin, D.-Y., Huang, Y.-S., Jeng, J.-C., Kuo, H.-Y., Chang, C.-C., Chao, T.-T., Ho, C.-C., Chen, Y.-C., Lin, T.-P., Fang, H.-I., *et al.* (2006). Role of SUMO-interacting motif in Daxx SUMO modification, subnuclear localization, and repression of sumoylated transcription factors. Mol Cell *24*, 341-354.

Liu, T., Ortiz, J.A., Taing, L., Meyer, C.A., Lee, B., Zhang, Y., Shin, H., Wong, S.S., Ma, J., Lei, Y., *et al.* (2011). Cistrome: an integrative platform for transcriptional regulation studies. Genome Biol *12*, R83.

Loppin, B., Bonnefoy, E., Anselme, C., Laurençon, A., Karr, T.L., and Couble, P. (2005). The histone H3.3 chaperone HIRA is essential for chromatin assembly in the male pronucleus. Nature *437*, 1386-1390.

Lorain, S., Quivy, J.P., Monier-Gavelle, F., Scamps, C., Lécluse, Y., Almouzni, G., and Lipinski, M. (1998). Core histones and HIRIP3, a novel histone-binding protein, directly interact with WD repeat protein HIRA. Mol Cell Biol *18*, 5546-5556.

Loyola, A., and Almouzni, G. (2007). Marking histone H3 variants: how, when and why? Trends Biochem Sci *32*, 425-433.

Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. Mol Cell *24*, 309-316.

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

Luk, E., Ranjan, A., Fitzgerald, P.C., Mizuguchi, G., Huang, Y., Wei, D., and Wu, C. (2010). Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. Cell *143*, 725-736.

Luk, E., Vu, N.-D., Patteson, K., Mizuguchi, G., Wu, W.-H., Ranjan, A., Backus, J., Sen, S., Lewis, M., Bai, Y., *et al.* (2007). Chz1, a nuclear chaperone for histone H2AZ. Mol Cell *25*, 357-368.

Luke, B., and Lingner, J. (2009). TERRA: telomeric repeat-containing RNA. EMBO J, 1-8.

Lupas, A., Van Dyke, M., and Stock, J. (1991). Predicting coiled coils from protein sequences. Science *252*, 1162-1164.

Lusser, A., Urwin, D.L., and Kadonaga, J.T. (2005). Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. Nature Publishing Group *12*, 160-166.

Maison, C., Bailly, D., Peters, A.H.F.M., Quivy, J.-P., Roche, D., Taddei, A., Lachner, M., Jenuwein, T., and Almouzni, G. (2002). Higher-order structure in pericentric heterochromatin involves a distinct pattern of histone modification and an RNA component. Nat Genet *30*, 329-334.

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

Marzluff, W.F., Gongidi, P., Woods, K.R., Jin, J., and Maltais, L.J. (2002). The human and mouse replication-dependent histone genes. Genomics *80*, 487-498.

Masumoto, H., Hawke, D., Kobayashi, R., and Verreault, A. (2005). A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. Nature *436*, 294-298.

Mazurkiewicz, J., Kepert, J.F., and Rippe, K. (2006). On the mechanism of nucleosome assembly by histone chaperone NAP1. J Biol Chem *281*, 16462-16472.

McDowell, T.L., Gibbons, R.J., Sutherland, H., O'Rourke, D.M., Bickmore, W.A., Pombo, A., Turley, H., Gatter, K., Picketts, D.J., Buckle, V.J., *et al.* (1999). Localization of a

putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. Proc Natl Acad Sci USA *96*, 13983-13988.

McKittrick, E., Gafken, P.R., Ahmad, K., and Henikoff, S. (2004). Histone H3.3 is enriched in covalent modifications associated with active chromatin. Proc Natl Acad Sci USA *101*, 1525-1530.

Mehrotra, P.V., Ahel, D., Ryan, D.P., Weston, R., Wiechens, N., Kraehenbuehl, R., Owen-Hughes, T., and Ahel, I. (2011). DNA repair factor APLF is a histone chaperone. Mol Cell *41*, 46-55.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell *10*, 105-116.

Michaelson, J.S., Bader, D., Kuo, F., Kozak, C., and Leder, P. (1999). Loss of Daxx, a promiscuously interacting protein, results in extensive apoptosis in early mouse development. Genes Dev *13*, 1918-1923.

Mito, Y., Henikoff, J.G., and Henikoff, S. (2005). Genome-scale profiling of histone H3.3 replacement patterns. Nat Genet *37*, 1090-1097.

Mito, Y., Henikoff, J.G., and Henikoff, S. (2007). Histone replacement marks the boundaries of cis-regulatory domains. Science *315*, 1408-1411.

Mizuguchi, G., Xiao, H., Wisniewski, J., Smith, M.M., and Wu, C. (2007). Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. Cell *129*, 1153-1164.

Moazed, D. (2011). Mechanisms for the inheritance of chromatin States. Cell *146*, 510-518.

modENCODE Consortium, Roy, S., Ernst, J., Kharchenko, P.V., Kheradpour, P., Negre, N., Eaton, M.L., Landolin, J.M., Bristow, C.A., Ma, L., *et al.* (2010). Identification of functional elements and regulatory circuits by Drosophila modENCODE. Science *330*, 1787-1797.

Mosammaparast, N., Ewart, C.S., and Pemberton, L.F. (2002). A role for nucleosome assembly protein 1 in the nuclear transport of histones H2A and H2B. EMBO J *21*, 6527-6538.

Mostoslavsky, R., Chua, K., Lombard, D., and Pang, W. (2006). Downtime | Sciverse ScienceDirect. Cell.

Muller, H.J. (1930). Types of visible variations induced by X-rays inDrosophila. Journ of Gen *22*, 299-334.

Munakata, T., Adachi, N., Yokoyama, N., Kuzuhara, T., and Horikoshi, M. (2000). A human homologue of yeast anti-silencing factor has histone chaperone activity. Genes Cells *5*, 221-233.

Murzina, N.V., Pei, X.-Y., Zhang, W., Sparkes, M., Vicente-Garcia, J., Pratap, J.V., McLaughlin, S.H., Ben-Shahar, T.R., Verreault, A., Luisi, B.F., *et al.* (2008). Structural

basis for the recognition of histone H4 by the histone-chaperone RbAp46. Structure/Folding and Design *16*, 1077-1085.

Muto, S., Senda, M., Akai, Y., Sato, L., Suzuki, T., Nagai, R., Senda, T., and Horikoshi, M. (2007). Relationship between the structure of SET/TAF-Ibeta/INHAT and its histone chaperone activity. Proc Natl Acad Sci USA *104*, 4285-4290.

Nakayama, T., Nishioka, K., Dong, Y.-X., Shimojima, T., and Hirose, S. (2007). Drosophila GAGA factor directs histone H3.3 replacement that prevents the heterochromatin spreading. Genes Dev *21*, 552-561.

Nan, X., Hou, J., Maclean, A., Nasir, J., Lafuente, M.J., Shu, X., Kriaucionis, S., and Bird, A. (2007). Interaction between chromatin proteins MECP2 and ATRX is disrupted by mutations that cause inherited mental retardation. Proc Natl Acad Sci USA *104*, 2709-2714.

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

Oberdoerffer, P., and Sinclair, D.A. (2007). The role of nuclear architecture in genomic instability and ageing. Nat Rev Mol Cell Biol *8*, 692-702.

Oohara, I., and Wada, A. (1987). Spectroscopic studies on histone-DNA interactions. I. The interaction of histone (H2A, H2B) dimer with DNA: DNA sequence dependence. J Mol Biol *196*, 389-397.

Ooi, S.K.T., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.-P., Allis, C.D., *et al.* (2007). DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. Nature *448*, 714-717.

Ooi, S.L., Priess, J.R., and Henikoff, S. (2006). Histone H3.3 variant dynamics in the germline of Caenorhabditis elegans. PLoS Genet *2*, e97.

Orphanides, G., Wu, W.H., Lane, W.S., Hampsey, M., and Reinberg, D. (1999). The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. Nature *400*, 284-288.

Osakabe, A., Tachiwana, H., Matsunaga, T., Shiga, T., Nozawa, R.-S., Obuse, C., and Kurumizaka, H. (2010). Nucleosome formation activity of human sNASP. J Biol Chem *285*, 11913-11921.

Otani, J., Nankumo, T., Arita, K., Inamoto, S., Ariyoshi, M., and Shirakawa, M. (2009). Structural basis for recognition of H3K4 methylation status by the DNA methyltransferase 3A ATRX-DNMT3-DNMT3L domain. EMBO Rep *10*, 1235-1241.

Park, Y.-J., Chodaparambil, J.V., Bao, Y., McBryant, S.J., and Luger, K. (2005). Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. J Biol Chem *280*, 1817-1825.

Park, Y.-J., and Luger, K. (2008). Histone chaperones in nucleosome eviction and histone exchange. Curr Opin Struct Biol *18*, 282-289.

Park, Y.J., and Luger, K. (2006). The structure of nucleosome assembly protein 1. Proc Natl Acad Sci USA *103*, 1248-1253.

Parsons, D.W., Li, M., Zhang, X., Jones, S., Leary, R.J., Lin, J.C.-H., Boca, S.M., Carter, H., Samayoa, J., Bettegowda, C., *et al.* (2011). The genetic landscape of the childhood cancer medulloblastoma. Science *331*, 435-439.

Philpott, A., and Leno, G.H. (1992). Nucleoplasmin remodels sperm chromatin in Xenopus egg extracts. Cell *69*, 759-767.

Philpott, A., Leno, G.H., and Laskey, R.A. (1991). Sperm decondensation in Xenopus egg cytoplasm is mediated by nucleoplasmin. Cell *65*, 569-578.

Picketts, D.J., Higgs, D.R., Bachoo, S., Blake, D.J., Quarrell, O.W., and Gibbons, R.J. (1996). ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. Hum Mol Genet *5*, 1899-1907.

Pidoux, A.L., Choi, E.S., Abbott, J.K.R., Liu, X., Kagansky, A., Castillo, A.G., Hamilton, G.L., Richardson, W., Rappsilber, J., He, X., *et al.* (2009). Fission yeast Scm3: A CENP-A receptor required for integrity of subkinetochore chromatin. Mol Cell *33*, 299-311.

Piña, B., and Suau, P. (1987). Changes in histones H2A and H3 variant composition in differentiating and mature rat brain cortical neurons. Developmental Biology *123*, 51-58.

Placek, B., Huang, J., Kent, J., Dorsey, J., Rice, L., Fraser, N., and Berger, S. (2008). The histone variant H3.3 regulates gene expression during lytic infection by Herpes Simplex Virus, HSV-1. J Virol.

Prochasson, P., Florens, L., Swanson, S.K., Washburn, M.P., and Workman, J.L. (2005). The HIR corepressor complex binds to nucleosomes generating a distinct protein/DNA complex resistant to remodeling by SWI/SNF. Genes Dev *19*, 2534-2539.

Rai, T.S., Puri, A., McBryan, T., Hoffman, J., Tang, Y., Pchelintsev, N., van Tuyn, J., Marmorstein, R., Schultz, D.C., and Adams, P.D. (2011). Human CABIN1 is a functional member of the human HIRA/UBN1/ASF1a histone H3.3 chaperone complex. Mol Cell Biol.

Ramirez-Carrozzi, V.R., Braas, D., Bhatt, D.M., Cheng, C.S., Hong, C., Doty, K.R., Black, J.C., Hoffmann, A., Carey, M., and Smale, S.T. (2009). A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. Cell *138*, 114-128.

Ranjitkar, P., Press, M.O., Yi, X., Baker, R., MacCoss, M.J., and Biggins, S. (2010). An E3 ubiquitin ligase prevents ectopic localization of the centromeric histone H3 variant via the centromere targeting domain. Mol Cell *40*, 455-464.

Ray-Gallet, Quivy, Silljé, Nigg, and Almouzni (2007). The histone chaperone Asf1 is dispensable for direct de novo histone deposition in Xenopus egg extracts. Chromosoma *116*, 487-496.

Ray-Gallet, D., Quivy, J.-P., Scamps, C., Martini, E.M.-D., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. Mol Cell *9*, 1091-1100.

Recht, J., Tsubota, T., Tanny, J.C., Diaz, R.L., Berger, J.M., Zhang, X., Garcia, B.A., Shabanowitz, J., Burlingame, A.L., Hunt, D.F., *et al.* (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. Proc Natl Acad Sci USA *103*, 6988-6993.

Regnard, C., Desbruyeres, E., Huet, J.C., Beauvallet, C., Pernollet, J.C., and Edde, B. (2000). Polyglutamylation of nucleosome assembly proteins. J Biol Chem *275*, 15969-15976.

Richmond, T.J., and Davey, C.A. (2003). The structure of DNA in the nucleosome core. Nature *423*, 145-150.

Ritchie, K., Seah, C., Moulin, J., Isaac, C., Dick, F., and Bérubé, N.G. (2008). Loss of ATRX leads to chromosome cohesion and congression defects. J Cell Biol *180*, 315-324.

Roberts, C., Sutherland, H.F., Farmer, H., Kimber, W., Halford, S., Carey, A., Brickman, J.M., Wynshaw-Boris, A., and Scambler, P.J. (2002). Targeted mutagenesis of the Hira gene results in gastrulation defects and patterning abnormalities of mesoendodermal derivatives prior to early embryonic lethality. Mol Cell Biol *22*, 2318-2328.

Rogers, R.S., Inselman, A., Handel, M.A., and Matunis, M.J. (2004). SUMO modified proteins localize to the XY body of pachytene spermatocytes. Chromosoma *113*, 233-243.

Rooney, A.P., Piontkivska, H., and Nei, M. (2002). Molecular evolution of the nontandemly repeated genes of the histone 3 multigene family. Molecular Biology and Evolution *19*, 68-75.

Rufiange, A., Jacques, P.-E., Bhat, W., Robert, F., and Nourani, A. (2007). Genome-wide replication-independent histone H3 exchange occurs predominantly at promoters and implicates H3 K56 acetylation and Asf1. Mol Cell *27*, 393-405.

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

Saha, A., Wittmeyer, J., and Cairns, B.R. (2006). Chromatin remodelling: the industrial revolution of DNA around histones. Nat Rev Mol Cell Biol *7*, 437-447.

Sakai, A., Schwartz, B.E., Goldstein, S., and Ahmad, K. (2009). Transcriptional and developmental functions of the H3.3 histone variant in Drosophila. Curr Biol *19*, 1816-1820.

Salomoni, P., and Khelifi, A.F. (2006). Daxx: death or survival protein? Trends Cell Biol *16*, 97-104.

Sanchez-Pulido, L., Pidoux, A.L., Ponting, C.P., and Allshire, R.C. (2009). Common ancestry of the CENP-A chaperones Scm3 and HJURP. Cell *137*, 1173-1174.

Santenard, A., Ziegler-Birling, C., Koch, M., Tora, L., Bannister, A.J., and Torres-Padilla, M.E. (2010). Heterochromatin formation in the mouse embryo requires critical residues of the histone variant H3.3. Nat Cell Biol *12*, 853-862.

Santiago, A., Godsey, A.C., Hossain, J., Zhao, L.Y., and Liao, D. (2009). Identification of two independent SUMO-interacting motifs in Daxx: evolutionary conservation from Drosophila to humans and their biochemical functions. Cell Cycle *8*, 76-87.

Sawatsubashi, S., Murata, T., Lim, J., Fujiki, R., Ito, S., Suzuki, E., Tanabe, M., Zhao, Y., Kimura, S., Fujiyama, S., *et al.* (2010). A histone chaperone, DEK, transcriptionally coactivates a nuclear receptor. Genes Dev *24*, 159-170.

Schermer, U.J., Korber, P., and Hörz, W. (2005). Histones are incorporated in trans during reassembly of the yeast PHO5 promoter. Mol Cell *19*, 279-285.

Schneiderman, J.I., Sakai, A., Goldstein, S., and Ahmad, K. (2009). The XNP remodeler targets dynamic chromatin in Drosophila. Proc Natl Acad Sci USA *106*, 14472-14477.

Schoeftner, S., and Blasco, M.A. (2009). A 'higher order' of telomere regulation: telomere heterochromatin and telomeric RNAs. EMBO J *28*, 2323-2336.

Schuster-Böckler, B., Schultz, J., and Rahmann, S. (2004). HMM Logos for visualization of protein families. BMC Bioinformatics *5*, 7.

Schwabish, M.A., and Struhl, K. (2006). Asf1 mediates histone eviction and deposition during elongation by RNA polymerase II. Mol Cell *22*, 415-422.

Schwartz, B.E., and Ahmad, K. (2005). Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev *19*, 804-814.

Seah, C., Levy, M.A., Jiang, Y., Mokhtarzada, S., Higgs, D.R., Gibbons, R.J., and Bérubé, N.G. (2008). Neuronal death resulting from targeted disruption of the Snf2 protein ATRX is mediated by p53. J Neurosci *28*, 12570-12580.

Sekulic, N., Bassett, E.A., Rogers, D.J., and Black, B.E. (2010). The structure of (CENP-A-H4)(2) reveals physical features that mark centromeres. Nature *467*, 347-351.

Sharp, J.A., Fouts, E.T., Krawitz, D.C., and Kaufman, P.D. (2001). Yeast histone deposition protein Asf1p requires Hir proteins and PCNA for heterochromatic silencing. Curr Biol *11*, 463-473.

Sharp, J.A., Franco, A.A., Osley, M.A., and Kaufman, P.D. (2002). Chromatin assembly factor I and Hir proteins contribute to building functional kinetochores in S. cerevisiae. Genes Dev *16*, 85-100.

Sharp, J.A., Rizki, G., and Kaufman, P.D. (2005). Regulation of histone deposition proteins Asf1/Hir1 by multiple DNA damage checkpoint kinases in Saccharomyces cerevisiae. Genetics *171*, 885-899.

Sherwood, P.W., Tsang, S.V., and Osley, M.A. (1993). Characterization of HIR1 and HIR2, two genes required for regulation of histone gene transcription in Saccharomyces cerevisiae. Mol Cell Biol *13*, 28-38.

Shintomi, K., Iwabuchi, M., Saeki, H., Ura, K., Kishimoto, T., and Ohsumi, K. (2005). Nucleosome assembly protein-1 is a linker histone chaperone in Xenopus eggs. Proc Natl Acad Sci USA *102*, 8210-8215.

Shivaraju, M., Camahort, R., Mattingly, M., and Gerton, J.L. (2011). Scm3 is a centromeric nucleosome assembly factor. Journal of Biological Chemistry *286*, 12016-12023.

Shuaib, M., Ouararhni, K., Dimitrov, S., and Hamiche, A. (2010). HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. Proc Natl Acad Sci USA *107*, 1349-1354.

Siegel, T.N., Hekstra, D.R., Kemp, L.E., Figueiredo, L.M., Lowell, J.E., Fenyo, D., Wang, X., Dewell, S., and Cross, G.A. (2009). Four histone variants mark the boundaries of polycistronic transcription units in Trypanosoma brucei. Genes Dev, 1-15.

Smith, C.A., and Kortemme, T. (2008). Backrub-like backbone simulation recapitulates natural protein conformational variability and improves mutant side-chain prediction. J Mol Biol *380*, 742-756.

Smith, S., and Stillman, B. (1991). Stepwise assembly of chromatin during DNA replication in vitro. EMBO J *10*, 971-980.

Sobel, R.E., Cook, R.G., Perry, C.A., Annunziato, A.T., and Allis, C.D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. Proc Natl Acad Sci USA *92*, 1237-1241.

Song, J., Garlick, J., and Kingston, R. (2008). Structural basis of histone H4 recognition by p55. Genes Dev.

Spector, M.S., Raff, A., DeSilva, H., Lee, K., and Osley, M.A. (1997). Hir1p and Hir2p function as transcriptional corepressors to regulate histone gene transcription in the Saccharomyces cerevisiae cell cycle. Mol Cell Biol *17*, 545-552.

Stein, A. (1979). DNA folding by histones: the kinetics of chromatin core particle reassembly and the interaction of nucleosomes with histones. J Mol Biol *130*, 103-134.

Stein, A., Bina-Stein, M., and Simpson, R.T. (1977). Crosslinked histone octamer as a model of the nucleosome core. Proc Natl Acad Sci USA *74*, 2780-2784.

Stein, A., Whitlock, J.P., and Bina, M. (1979). Acidic polypeptides can assemble both histones and chromatin in vitro at physiological ionic strength. Proc Natl Acad Sci USA *76*, 5000-5004.

Stols, L., Zhou, M., Eschenfeldt, W.H., Millard, C.S., Abdullah, J., Collart, F.R., Kim, Y., and Donnelly, M.I. (2007). New vectors for co-expression of proteins: structure of Bacillus subtilis ScoAB obtained by high-throughput protocols. Protein Expression and Purification *53*, 396-403.

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

Stratmann, M., and Schibler, U. (2011). Transcription factor loading: please take my place! Cell *146*, 497-499.

Stuwe, T., Hothorn, M., Lejeune, E., Rybin, V., Bortfeld, M., Scheffzek, K., and Ladurner, A. (2008). The FACT Spt16 "peptidase" domain is a histone H3-H4 binding module. Proc Natl Acad Sci USA.

Suganuma, T., Pattenden, S.G., and Workman, J.L. (2008). Diverse functions of WD40 repeat proteins in histone recognition. Genes Dev *22*, 1265-1268.

Sutcliffe, E.L., Parish, I.A., He, Y.Q., Juelich, T., Tierney, M.L., Rangasamy, D., Milburn, P.J., Parish, C.R., Tremethick, D.J., and Rao, S. (2009). Dynamic histone variant exchange accompanies gene induction in T cells. Mol Cell Biol *29*, 1972-1986.

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

Tachiwana, H., Kagawa, W., Shiga, T., Osakabe, A., Miya, Y., Saito, K., Hayashi-Takanaka, Y., Oda, T., Sato, M., Park, S.-Y., *et al.* (2011). Crystal structure of the human centromeric nucleosome containing CENP-A. Nature *476*, 232-235.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell *116*, 51-61.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Takayama, Y., and Takahashi, K. (2007). Differential regulation of repeated histone genes during the fission yeast cell cycle. Nucleic Acids Res *35*, 3223-3237.

Tamura, T., Smith, M., Kanno, T., Dasenbrock, H., Nishiyama, A., and Ozato, K. (2009). Inducible deposition of the histone variant H3.3 in interferon-stimulated genes. J Biol Chem *284*, 12217-12225.

Tanaka, Y., Tawaramoto-Sasanuma, M., Kawaguchi, S., Ohta, T., Yoda, K., Kurumizaka, H., and Yokoyama, S. (2004). Expression and purification of recombinant human histones. Methods *33*, 3-11.

Taneva, S.G., Muñoz, I.G., Franco, G., Falces, J., Arregi, I., Muga, A., Montoya, G., Urbaneja, M.A., and Bañuelos, S. (2008). Activation of nucleoplasmin, an oligomeric histone chaperone, challenges its stability. Biochemistry *47*, 13897-13906.

Tang, J., Wu, S., Liu, H., Stratt, R., Barak, O.G., Shiekhattar, R., Picketts, D.J., and Yang, X. (2004). A novel transcription regulatory complex containing death domain-associated protein and the ATR-X syndrome protein. J Biol Chem *279*, 20369-20377.

Tang, Y., Holbert, M.A., Delgoshaie, N., Wurtele, H., Guillemette, B., Meeth, K., Yuan, H., Drogaris, P., Lee, E.-H., Durette, C., *et al.* (2011). Structure of the Rtt109-AcCoA/Vps75 complex and implications for chaperone-mediated histone acetylation. Structure *19*, 221-231.

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

Thakar, A., Gupta, P., Ishibashi, T., Finn, R., Silva-Moreno, B., Uchiyama, S., Fukui, K., Tomschik, M., Ausió, J., and Zlatanova, J. (2009). H2A.Z and H3.3 histone variants affect nucleosome structure: biochemical and biophysical studies. Biochemistry.

Thomas, J.O., and Butler, P.J. (1977). Characterization of the octamer of histones free in solution. J Mol Biol *116*, 769-781.

Ting, D.T., Lipson, D., Paul, S., Brannigan, B.W., Akhavanfard, S., Coffman, E.J., Contino, G., Deshpande, V., Iafrate, A.J., Letovsky, S., *et al.* (2011). Aberrant Overexpression of Satellite Repeats in Pancreatic and Other Epithelial Cancers. Science, 1-40.

Torres-Padilla, M.E., Bannister, A.J., Hurd, P.J., Kouzarides, T., and Zernicka-Goetz, M. (2006). Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. Int J Dev Biol *50*, 455-461.

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

Umehara, T., Chimura, T., Ichikawa, N., and Horikoshi, M. (2002). Polyanionic stretch-deleted histone chaperone cia1/Asf1p is functional both in vivo and in vitro. Genes Cells *7*, 59-73.

van der Heijden, G.W., Derijck, A.A.H.A., Pósfai, E., Giele, M., Pelczar, P., Ramos, L., Wansink, D.G., van der Vlag, J., Peters, A.H.F.M., and de Boer, P. (2007). Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. Nat Genet *39*, 251-258.

van der Heijden, G.W., Derijck, A.A.H.A., Ramos, L., Giele, M., van der Vlag, J., and de Boer, P. (2006). Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. Developmental Biology *298*, 458-469.

van der Heijden, G.W., Dieker, J.W., Derijck, A.A.H.A., Muller, S., Berden, J.H.M., Braat, D.D.M., van der Vlag, J., and de Boer, P. (2005). Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. Mech Dev *122*, 1008-1022.

van Haaften, G., Dalgliesh, G.L., Davies, H., Chen, L., Bignell, G., Greenman, C., Edkins, S., Hardy, C., O'meara, S., Teague, J., *et al.* (2009). Somatic mutations of the histone H3K27 demethylase gene UTX in human cancer. Nat Genet *41*, 521-523.

van Holde, K. (1988). Chromatin (New York, Springer).

Van Hooser, A.A., Ouspenski, I.I., Gregson, H.C., Starr, D.A., Yen, T.J., Goldberg, M.L., Yokomori, K., Earnshaw, W.C., Sullivan, K.F., and Brinkley, B.R. (2001). Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A. J Cell Sci *114*, 3529-3542.

VanDemark, A.P., Blanksma, M., Ferris, E., Heroux, A., Hill, C.P., and Formosa, T. (2006). The structure of the yFACT Pob3-M domain, its interaction with the DNA replication factor RPA, and a potential role in nucleosome deposition. Mol Cell *22*, 363-374.

Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C.K., Stephens, P., Davies, H., Jones, D., Lin, M.-L., Teague, J., *et al.* (2011). Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. Nature.

Verreault, A., Kaufman, P.D., Kobayashi, R., and Stillman, B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. Cell *87*, 95-104.

Villar-Garea, A., Israel, L., and Imhof, A. (2008). Analysis of histone modifications by mass spectrometry. Current protocols in protein science *Chapter 14*, Unit 14.10.

Voss, T.C., Schiltz, R.L., Sung, M.-H., Yen, P.M., Stamatoyannopoulos, J.A., Biddie, S.C., Johnson, T.A., Miranda, T.B., John, S., and Hager, G.L. (2011). Dynamic Exchange at Regulatory Elements during Chromatin Remodeling Underlies Assisted Loading Mechanism. Cell *146*, 544-554.

Wang, H., Cao, R., Xia, L., Erdjument-Bromage, H., Borchers, C., Tempst, P., and Zhang, Y. (2001). Purification and functional characterization of a histone H3-lysine 4-specific methyltransferase. Mol Cell *8*, 1207-1217.

Waterborg, J.H. (1993). Histone synthesis and turnover in alfalfa. Fast loss of highly acetylated replacement histone variant H3.2. J Biol Chem *268*, 4912-4917.

Wen, B., Wu, H., Shinkai, Y., Irizarry, R.A., and Feinberg, A.P. (2009). Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. Nat Genet *41*, 246-250.

Wiegand, K.C., Shah, S.P., Al-Agha, O.M., Zhao, Y., Tse, K., Zeng, T., Senz, J., McConechy, M.K., Anglesio, M.S., Kalloger, S.E., *et al.* (2010). ARID1A mutations in endometriosis-associated ovarian carcinomas. N Engl J Med *363*, 1532-1543.

Wilhelm, F.X., Wilhelm, M.L., Erard, M., and Duane, M.P. (1978). Reconstitution of chromatin: assembly of the nucleosome. Nucleic Acids Res *5*, 505-521.

Wirbelauer, C., Bell, O., and Schübeler, D. (2005). Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. Genes Dev *19*, 1761-1766.

Witt, O., Albig, W., and Doenecke, D. (1996). Testis-specific expression of a novel human H3 histone gene. Exp Cell Res *229*, 301-306.

Wong, L., Ren, H., Williams, E., McGhie, J., Ahn, S., Sim, M., Tam, A., Earle, E., Anderson, M., Mann, J., *et al.* (2009). Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. Genome Res.

Wong, L.H., McGhie, J.D., Sim, M., Anderson, M.A., Ahn, S., Hannan, R.D., George, A.J., Morgan, K.A., Mann, J.R., and Choo, K.H.A. (2010). ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. Genome Res *20*, 351-360.

Worcel, A., Han, S., and Wong, M.L. (1978). Assembly of newly replicated chromatin. Cell *15*, 969-977.

Wu, R.S., Tsai, S., and Bonner, W.M. (1982). Patterns of histone variant synthesis can distinguish G0 from G1 cells. Cell *31*, 367-374.

Xue, Y., Gibbons, R., Yan, Z., Yang, D., McDowell, T.L., Sechi, S., Qin, J., Zhou, S., Higgs, D., and Wang, W. (2003). The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. Proc Natl Acad Sci USA *100*, 10635-10640.

Yang, X., Khosravi-Far, R., Chang, H.Y., and Baltimore, D. (1997). Daxx, a novel Fas-binding protein that activates JNK and apoptosis. Cell *89*, 1067-1076.

Zacharias, H. (1995). Emil Heitz (1892-1965): chloroplasts, heterochromatin, and polytene chromosomes., Vol 141.

Zaratiegui, M., Irvine, D.V., and Martienssen, R.A. (2007). Noncoding RNAs and gene silencing. Cell *128*, 763-776.

Zhang, R., Chen, W., and Adams, P.D. (2007a). Molecular dissection of formation of senescence-associated heterochromatin foci. Mol Cell Biol *27*, 2343-2358.

Zhang, Z., Fan, H.-Y., Goldman, J.A., and Kingston, R.E. (2007b). Homology-driven chromatin remodeling by human RAD54. Nat Struct Mol Biol *14*, 397-405.

Zhou, Z., Feng, H., Hansen, D.F., Kato, H., Luk, E., Freedberg, D.I., Kay, L.E., Wu, C., and Bai, Y. (2008). NMR structure of chaperone Chz1 complexed with histones H2A.Z-H2B. Nat Struct Mol Biol *15*, 868-869.

Zhou, Z., Feng, H., Zhou, B.-R., Ghirlando, R., Hu, K., Zwolak, A., Miller Jenkins, L.M., Xiao, H., Tjandra, N., Wu, C., *et al.* (2011). Structural basis for recognition of centromere histone variant CenH3 by the chaperone Scm3. Nature.

Zhu, Q., Pao, G.M., Huynh, A.M., Suh, H., Tonnu, N., Nederlof, P.M., Gage, F.H., and Verma, I.M. (2011). BRCA1 tumour suppression occurs via heterochromatin-mediated silencing. Nature *477*, 179-184.

Zweidler, A. (1978). Resolution of histones by polyacrylamide gel electrophoresis in presence of nonionic detergents. Methods Cell Biol *17*, 223-233.