2017

Molecular Basis For H3.3/h4 Deposition By The Hira Histone Chaperone Complex

Michael Daniel Ricketts
University of Pennsylvania, md.ricketts@gmail.com

Molecular Basis For H3.3/h4 Deposition By The Hira Histone Chaperone Complex

Abstract
The HIRA histone chaperone complex is composed of the proteins HIRA, UBN1, and CABIN1 that cooperate with ASF1a to specifically mediate deposition of H3.3/H4 into chromatin. While many of the protein/protein contacts that allow for assembly of this multi-protein histone chaperone complex have been elucidated, how these proteins are able to cooperate to specifically bind and deposit H3.3/H4 over H3.1/H4, which differ by only 5 amino acids, has remained unclear. Here we demonstrate the HIRA complex subunit UBN1 specifically binds to H3.3/H4 over H3.1/H4 using a combination of biochemical and structural studies. We show H3.3-specific binding is mediated by highly conserved residues in the UBN1 Hpc2-related domain (HRD) and residue G90 in H3.3. Interestingly the UBN1 HRD and another H3.3-specific histone chaperone, Daxx, bind to H3.3/H4 in the region of G90 with striking similarity. Using sequence conservation and secondary structure prediction we have identified the UBN1 middle domain. We show the middle domain is able to form monomer and dimer populations as well as bind to H3/H4 through interaction with the N-terminal histone tails. We have additionally discovered that a loop region linking the UBN1 HRD and middle domains binds non-specifically to DNA with nanomolar affinity, although UBN1 fragments containing the HRD, DNA binding loop, and middle domain are unable to bind assembled nucleosomes. Finally we demonstrate with analytical ultracentrifugation that the HIRA C-terminal domain forms a stable trimer and interacts with CABIN1 in a HIRA(3)/CABIN1(2) stoichiometry. We show that the HIRA C-terminal domain is organized into separate strand and helical regions that must be linked for trimer formation. Further we show that a replisome protein called Ctf4 is structurally related to HIRA, as it also arranges into a trimer composed of separate strand and helical regions and interacts in a 3:2 stoichiometry with binding partners. From these observations we propose a model for H3.3/H4 deposition by the HIRA complex involving DNA binding and dimer formation by UBN1 to mediate (H3.3/H4)2 tetramer formation organized on a HIRA(3)/UBN1(3)/CABIN1(2) complex where the third UBN1 molecule may harbor ASF1a/H3.3/H4 in reserve.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Biochemistry & Molecular Biophysics

First Advisor
Ronen Marmorstein

Keywords
Histone, Histone Chaperone

MOLECULAR BASIS FOR H3.3/H4 DEPOSITION BY THE HIRA HISTONE CHAPERONE COMPLEX

Michael Daniel Ricketts

A DISSERTATION

in

Biochemistry and Molecular Biophysics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2017

Supervisor of Dissertation

_____

Ronen Marmorstein, PhD

George W. Raiziss Professor

Graduate Group Chairperson

_____

Kim A. Sharp, PhD, Associate Professor of Biochemistry and Biophysics

Dissertation Committee

Shelley L. Berger, PhD, Daniel S. Och University Professor

Ben E. Black, PhD, Associate Professor of Biochemistry and Biophysics (Chair)

Hua-Ying Fan, PhD, Assistant Professor of Biochemistry and Biophysics

F. Bradley Johnson, MD–PhD, Associate Professor of Pathology and Laboratory Medicine

Gregory D. Van Duyne, PhD, Jacob Gershon-Cohen Professor of Medical Science

Peter D. Adams, PhD, Professor (Sanford Burnham Prebys Medical Discovery Institute)

For my brother Johnny

Who was killed on 07/13/2016 at the age of 22 by heroin laced with fentanyl

And my brother Billy

Who is currently in recovery for opioid addiction

ABSTRACT

Molecular basis for H3.3/H4 deposition by the HIRA histone chaperone complex

Michael Daniel Ricketts

Ronen Marmorstein

The HIRA histone chaperone complex is composed of the proteins HIRA, UBN1, and CABIN1 that cooperate with ASF1a to specifically mediate deposition of H3.3/H4 into chromatin. While many of the protein/protein contacts that allow for assembly of this multi-protein histone chaperone complex have been elucidated, how these proteins are able to cooperate to specifically bind and deposit H3.3/H4 over H3.1/H4, which differ by only 5 amino acids, has remained unclear. Here we demonstrate the HIRA complex subunit UBN1 specifically binds to H3.3/H4 over H3.1/H4 using a combination of biochemical and structural studies. We show H3.3-specific binding is mediated by highly conserved residues in the UBN1 Hpc2-related domain (HRD) and residue G90 in H3.3. Interestingly the UBN1 HRD and another H3.3-specific histone chaperone, Daxx, bind to H3.3/H4 in the region of G90 with striking similarity. Using sequence conservation and secondary structure prediction we have identified the UBN1 middle domain. We show the middle domain is able to form monomer and dimer populations as well as bind to H3/H4 through interaction with the N-terminal histone tails. We have additionally discovered that a loop region linking the UBN1 HRD and middle domains binds non-specifically to DNA with nanomolar affinity, although UBN1 fragments containing the HRD, DNA binding loop, and middle domain are unable to bind assembled nucleosomes. Finally we demonstrate with analytical ultracentrifugation that the HIRA C-terminal domain forms a stable trimer and interacts with CABIN1 in a HIRA(3)/CABIN1(2) stoichiometry. We show that the HIRA C-terminal

domain is organized into separate strand and helical regions that must be linked for trimer formation. Further we show that a replisome protein called Ctf4 is structurally related to HIRA, as it also arranges into a trimer composed of separate strand and helical regions and interacts in a 3:2 stoichiometry with binding partners. From these observations we propose a model for H3.3/H4 deposition by the HIRA complex involving DNA binding and dimer formation by UBN1 to mediate $(H3.3/H4)_2$ tetramer formation organized on a HIRA(3)/UBN1(3)/CABIN1(2) complex where the third UBN1 molecule may harbor ASF1a/H3.3/H4 in reserve.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: A molecular prospective for HIRA complex assembly and H3.3-specific histone chaperone function

1.1 – Introduction to the HIRA histone chaperone complex

Incorporation of variant histone sequences, in addition to post-translational modification of histones, serves to modulate the chromatin environment. Different histone chaperone proteins mediate the storage and chromatin deposition of variant histones. Although the two non-centromeric histone H3 variants, H3.1 and H3.3, differ by only five amino acids, replacement of histone H3.1 with H3.3 can modulate transcription for highly expressed and developmentally required genes, lead to the formation of repressive heterochromatin, or aid in DNA and chromatin repair. The human HIRA complex composed of HIRA, UBN1, CABIN1 and transiently ASF1a, forms one of two complexes that binds and deposit H3.3/H4 into chromatin. A number of recent biochemical and structural studies have revealed important details underlying how these proteins assemble and function together as a multi-protein H3.3-specific histone chaperone complex. Here we present a review of existing data and present a new model for assembly of the HIRA complex and for HIRA-mediated incorporation of H3.3/H4 into chromatin.

About 3 billion base pairs of DNA are packaged into the nucleus of the majority of cells in the human body. To allow for the storage and regulation of this great length of genetic material the negatively charged DNA is compacted through binding to positively charged histone proteins for the formation of nucleosomes, the basic repeating units of chromatin(Kornberg, 1974; Olins and Olins, 1974). The nucleosome core particle is built from 147 base pairs of DNA wrapped

around an octamer of core histone proteins composed of two copies each of H2A, H2B, H3, and H4(Luger et al., 1997a; Luger et al., 1997b; Richmond and Davey, 2003). While the majority of nucleosomes are composed of canonical histones, there are also histones with variant amino acid sequences that are incorporated at specific regions of chromatin(Maze et al., 2014). Targeted incorporation of these variant histones(Maze et al., 2014) and post-translation modification to a variety of histone residues(Bannister and Kouzarides, 2011) are the two major ways that histones contribute to modulation and regulation of the chromatin environment to mediate distinct DNA-templated activities such as DNA repair, transcriptional regulation or DNA replication.

Nucleosome assembly is coordinated by histone chaperone proteins, which bind specific histones to mediate their storage, eviction, or deposition from/or into chromatin(Burgess and Zhang, 2013; Laskey et al., 1978; Mattiroli et al., 2015). Histone chaperones are generally dedicated to binding and deposition of either the H3/H4 unit or the H2A/H2B unit although there are examples of proteins capable of chaperoning both H3/H4 and H2A/H2B(Tsunaka et al., 2016; Winkler et al., 2012). Some histone chaperons are able to bind several different variants of a particular histone, while others are dedicated to specifically binding just one variant sequence(Mattiroli et al., 2015). The histone chaperone ASF1a binds to either H3.1/H4 and H3.3/H4 and associates with the CAF-1 complex for H3.1/H4 deposition or the HIRA complex for H3.3/H4 deposition(Tagami et al., 2004;

Tang et al., 2006). Although H3.1 and H3.3 differ by only 5 amino acids, the HIRA complex specifically mediates replication-independent H3.3/H4 deposition, while the CAF-1 histone chaperone carries out H3.1/H4 deposition coupled to DNA replication and damage repair(Gaillard et al., 1996; Tagami et al., 2004).

The HIRA complex is composed of the proteins HIRA, Ubinuclein-1 (UBN1), and CABIN1(Banumathy et al., 2009; Rai et al., 2011; Tagami et al., 2004). These three proteins function with ASF1a to mediate deposition of H3.3/H4 primarily at the bodies of actively transcribed genes(Goldberg et al., 2010; Pchelintsev et al., 2013), gene regulatory regions(Banumathy et al., 2009; Jin et al., 2009; Rai et al., 2011), developmentally regulated genes(Banaszynski et al., 2013; Orsi et al., 2013), and areas of DNA and chromatin damage and repair(Adam et al., 2013; Ray-Gallet et al., 2011; Schneiderman et al., 2012; Zhang et al., 2007). The ATRX/DAXX histone chaperone complex also binds specifically to H3.3/H4 but functions independently of ASF1a and the HIRA complex to deposit H3.3/H4 mainly into areas of heterochromatin at telomeres, pericentromeres, and endogenous retroviral elements(Corpet et al., 2014; Elsaesser and Allis, 2010; Elsasser et al., 2012; Elsasser et al., 2015; Goldberg et al., 2010). While the HIRA and ATRX/DAXX complexes have largely separate functions, they employ similar structural features to bind H3.3/H4(Ricketts et al., 2015), both complexes may have some overlapping function in H3.3/H4 deposition during cellular senescence(Banumathy et al., 2009; Corpet et al., 2014; Rai et al., 2011; Ye et

al., 2007; Zhang et al., 2005), and ATRX/DAXX has been reported to deposit H3.3/H4 to activate gene transcription in neuronal cells.

In recent years, a number of biochemical and structural studies have contributed to a greatly increased understanding of the molecular assembly and function of the HIRA histone chaperone complex. In this review, we focus on detailing these developments in order to present an updated model of HIRA complex assembly for H3.3-specific binding and deposition functions.

1.2 – HIRA complex subunits

The HIRA complex is assembled from HIRA, UBN1 and CABIN1 and coordinates with ASF1a to carry out H3.3-specific deposition. All three subunits and ASF1 were identified and characterized individually before it was discovered that they work in concert to mediate deposition of H3.3/H4(Tagami et al., 2004). In this section we will review the known biological functions of the individual proteins that comprise the HIRA complex.

ASF1a

ASF1 (anti-silencing function 1) was originally identified in *S. cerevisiae* and named for the observed de-repression of silent mating type loci when overexpressed(Le et al., 1997). ASF1 was later characterized as a novel H3/H4 binding protein involved in histone deposition during DNA replication and repair(Tyler et al., 1999). In human cells, ASF1 has two isoforms, ASF1a and

ASF1b(Tagami et al., 2004; Tang et al., 2006). ASF1a/b bind to both H3.1/H4 and H3.3/H4 with similar affinity(Tagami et al., 2004). The CAF-1 complex cooperates with both ASF1a/b for H3.1/H4 deposition, while only ASF1a serves to deliver H3.3/H4 to the HIRA complex(Tagami et al., 2004; Tang et al., 2006; Zhang et al., 2005). Both HIRA and CAF-1 associate with the ASF1 N-terminal core using a small but conserved B-domain(Tang et al., 2006; Zhang et al., 2005). ASF1a/b have also been shown to cooperate with MCM2 to coordinate H3/H4 deposition. While ASF1a and ASF1b are highly conserved in their N-terminal core domains, they are divergent in their C-terminal tails; it has been shown that the C-terminal tail of ASF1b in combination with some region of HIRA outside the B-domain work together to inhibit HIRA binding, leading to a preferential association of HIRA with ASF1a(Tang et al., 2006).

HIRA

Human HIRA was identified within a genomic region lost in chromosome 22q11 deletion syndrome (DiGeorge syndrome)(Lamour et al., 1995). Depletion of HIRA was initially thought to be a possible contributing factor to the symptoms observed in patients with DiGeroge syndrome(Lamour et al., 1995), although deletion of the Tbx1 gene was later identified as responsible for the symptoms of DiGeorge syndrome(Lindsay et al., 2001). Previous to the identification of HIRA within the 22q11 deletion, two *S. cerevisiae* proteins were identified as transcriptional regulators of the histone genes and named Hir1 and Hir2 for their

histone regulatory function(Sherwood et al., 1993; Spector and Osley, 1993). When the open reading frame in the 22q11 deletion was analyzed it was discovered that it was homologous to Hir1 and Hir2 and the protein was named HIRA. The histone chaperone function of HIRA was not discovered until later when the HIRA complex was purified from HeLa cells using a tagged H3.3 to pull down a complex containing ASF1a, HIRA, UBN1, and CABIN1(Tagami et al., 2004). The transcriptional regulation of histone genes observed for *S. cerevisiae* Hir1/Hir2 has not been observed for human HIRA, although histone deposition by the HIRA complex has been linked to transcriptional activation for a number of highly expressed and developmentally regulated genes(Banaszynski et al., 2013; Dilg et al., 2016; Dutta et al., 2010; Goldberg et al., 2010). After the histone chaperone function of the human HIRA complex was characterized, *S. cerevisiae* Hir1/Hir2 were shown to form a complex with Hpc2 and Hir3, the *S. cerevisiae* orthologs of UBN1 and CABIN1, respectively, to carry out replication independent H3/H4 deposition(Green et al., 2005; Prochasson et al., 2005).

UBN1

UBN1 was originally identified as a ubiquitously expressed protein in the nuclei of human cells and shown to associate with endogenous and viral transcription factors to compete with binding to their DNA target sites(Aho et al., 2000; Aho et al., 2009). Although this transcription factor binding activity of human UBN1 is not well characterized at the biochemical level, the HIRA complex member CABIN1

has also been reported to bind a number of transcription factors(Hammond and Udvadia, 2010; Han et al., 2003; Jang et al., 2009; Liu et al., 2001; Youn and Liu, 2000) so it is possible that the original observation was mediated through CABIN1. Similarly to HIRA, UBN1 was identified as a member of the HIRA complex when pulled down with H3.3. HIRA and UBN1 together appear to form a functional core of the HIRA complex as knock down, mutation, and depletion experiments suggest that both ASF1a and CABIN1 are dispensable for proper H3.3 deposition activity of the complex(Bonnefoy et al., 2007; Ray-Gallet et al., 2007; Ray-Gallet et al., 2011; Song et al., 2013). Although the two proteins are weakly conserved, UBN1 was identified as an ortholog of the *S. cerevisiae* Hir complex subunit Hpc2 through alignment of the most conserved domain between the two proteins referred to as the Hpc2-related domain (HRD)(Banumathy et al., 2009), or less commonly as the Hpc2-Ubinuclein-1 domain (HUN)(Balaji et al., 2009). Recently UBN1 was shown to have H3.3-specific binding activity and likely mediates the specific deposition of H3.3/H4 over H3.1/H4 by the HIRA complex(Ricketts et al., 2015).

CABIN1

The largest member of the HIRA complex at 2220 amino acids, CABIN1, is predicted to have around thirty tetratricopeptide repeats (TPR)(Amin et al., 2013; Balaji et al., 2009). TPR domains are built from repeats of bi-helical pairs and often serve to mediate protein-protein interactions in large complexes(Blatch and

Lassle, 1999). Outside of binding to the C-terminal domain of HIRA(Rai et al., 2011), the role of this TPR containing protein in the H3.3 deposition process remains unknown. Several functions of CABIN1 independent of the HIRA complex have been well studied. CABIN1 was identified as a binding partner of calcineurin and acts as a regulator of the calcineurin mediated signaling pathway(Song et al., 2013). CABIN1 has also been shown to physically associate with transcription factors p53(Jang et al., 2009) and MEF2(Esau et al., 2001; Hammond and Udvadia, 2010; Han et al., 2003; Youn and Liu, 2000) to modulate their transcriptional regulatory activity. Further, modulation of MEF2 activity by CABIN1 is also dependent on HIRA and ASF1a histone chaperone activity(Yang et al., 2011); so it is possible that the role of CABIN1 is to bind to these, and potentially other, transcription factors and recruit the HIRA complex to deposit H3.3/H4 at targeted areas of chromatin. This HIRA mediated H3.3/H4 deposition may serve to modulate the local chromatin architecture and alter levels of transcription. Similar to CABIN1, the Hir3 *S. cerevisiae* ortholog of CABIN1 forms a complex with Hir1, Hir2, and Hpc2 to regulate transcription of histone genes and carry out replication independent histone deposition(Green et al., 2005; Prochasson et al., 2005; Spector and Osley, 1993).

1.3 – HIRA complex assembly and structure

In recent years, several studies have focused on determining many of the key protein-protein contacts involved in assembly of the HIRA complex and its specific association with H3.3/H4 over H3.1/H4. Here we present a summary of current biochemical and structural data in an effort to describe the most accurate model of HIRA complex assembly and function as an H3.3-specific histone chaperone.

HIRA as a molecular scaffold

From several early studies, it became clear that the protein HIRA was the central member of this H3.3-specific histone chaperone complex. The histone chaperone activity of HIRA was described independently of ASF1(Loppin et al., 2005; Ray-Gallet et al., 2002), and its direct interaction and cooperation with ASF1a was identified shortly thereafter(Tagami et al., 2004; Tang et al., 2006; Zhang et al., 2005). The obvious functional significance of the HIRA protein in H3.3 deposition led to the current nomenclature of referring to the entire HIRA/UBN1/CABIN1 complex, and transiently ASF1a, as the HIRA complex. The undefined C-terminal domain of HIRA has been reported to bind H3/H4 with phosphorylation at H4S47(Kang et al., 2011), although this data has not been expanded upon and several other studies suggest that UBN1 may be responsible for the H3.3-specific binding and deposition by the HIRA complex(Banumathy et al., 2009; Ray-Gallet et al., 2011; Ricketts et al., 2015; Song et al., 2013). Most of

the existing data suggests that HIRA likely serves as a molecular scaffold for assembly of the complex, as it is known to directly associate with every other member of the HIRA complex (Figure 1.1). Interestingly the stability of the core HIRA/UBN1/CABIN1 complex is dependent on the presence of all three proteins, as knockdown or removal of any individual subunit results in severely decreased protein expression levels for the remaining subunits(Banumathy et al., 2009; Rai et al., 2011; Song et al., 2013).

ASF1a/H3/H4

Binding and deposition of H3/H4 by the ASF1 core domain has been extensively studied using biochemical and structural techniques[35,42,58-60] (Figure 1.1). While there is less structural information for the ASF1a/H3/H4 complex(Agez et al., 2007), the ASF1 core domain is so conserved that information from the *S. cerevisiae* structure is likely applicable to the human ASF1a system(Adkins and Tyler, 2004; English et al., 2006; English et al., 2005; Green et al., 2005; Tyler et al., 1999). It is clear that ASF1 binds to one H3/H4 dimer and has extensive contacts with H3/H4 along the $(H3/H4)_2$ tetramer interface(English et al., 2006; English et al., 2005; Luger et al., 1997a; Natsume et al., 2007) (Figure 1.2a) thus inhibiting the formation of the $(H3/H4)_2$ tetramer, which in the absence of ASF1 exists in dynamic equilibrium with the H3/H4 dimer(Banks and Gloss, 2004; Winkler et al., 2012). ASF1a does not make any contacts to or near H3 residues

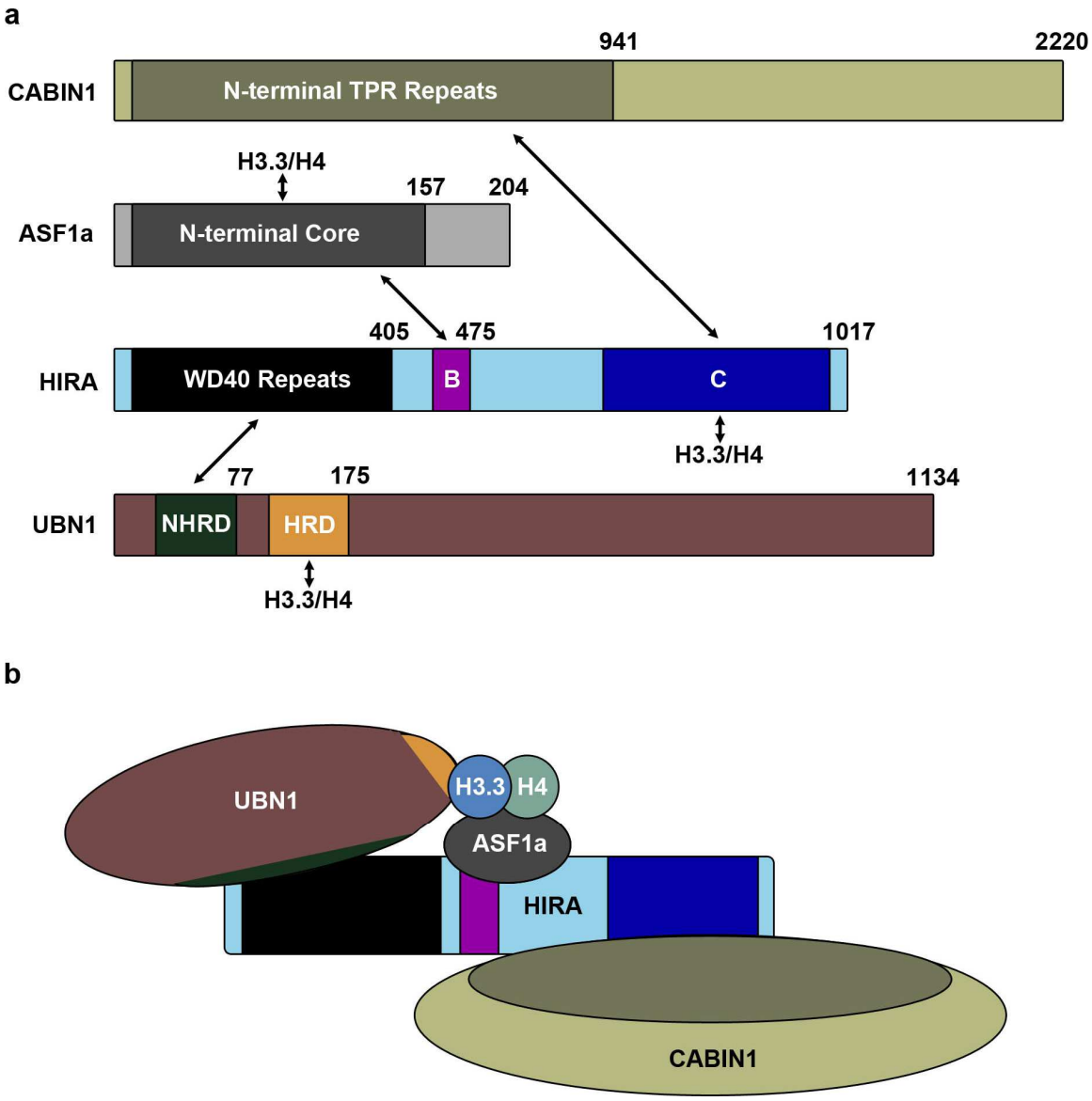


Figure 1.1 – Domain structures and interactions of HIRA complex members and H3.3/H4. (A) Domain structures of HIRA, UBN1, CABIN1 and ASF1a with arrows depicting all reported interactions between the complex members and H3.3/H4. (B) Proposed model for assembly of the HIRA complex based on the reported interacting domains.

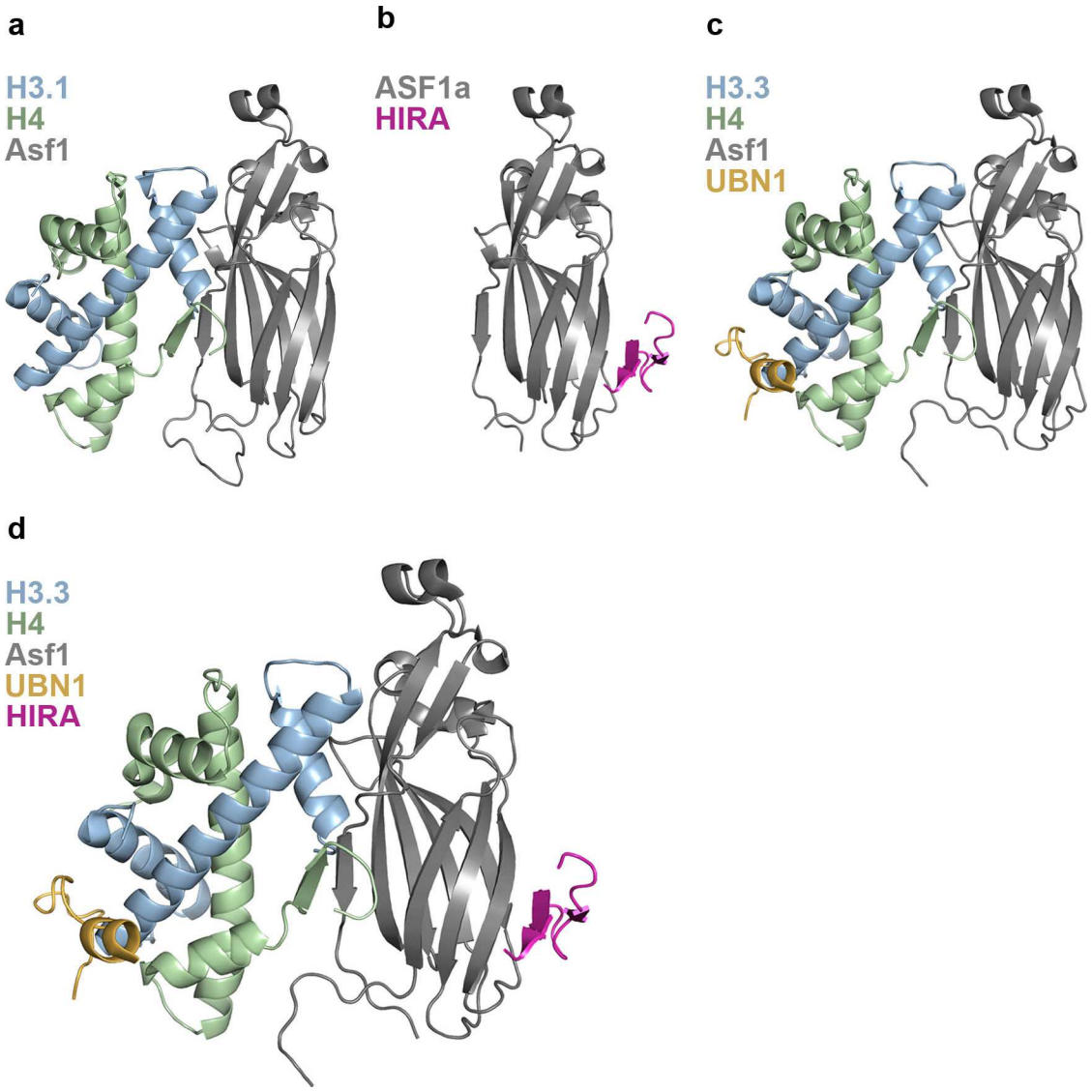


Figure 1.2 – Structural basis for HIRA complex assembly and function. (A) 1.7 Å crystal structure of *S. cerevisiae* Asf1 bound to *X. laevis* H3.1/H4 depicting the interaction of Asf1 with one H3/H4 dimer. (B) 2.7 Å crystal structure of the human HIRA/ASF1a complex showing association of the HIRA B-domain with the ASF1a core on the surface opposite to where ASF1a binds H3/H4. (C) 2.3 Å crystal structure of the UBN1/H3.3/H4/Asf1 complex revealing the molecular contacts between the UBN1 HRD and histone H3.3. (D) Structural model of the HIRA B-domain superimposed onto the UBN1/H3.3/H4/Asf1 complex depicting a hypothetical intermediate in the pathway of HIRA-mediated H3.3/H4 deposition.

87-90, which are the major differentiating amino acids between H3.3 and H3.1, and is capable of supplying H3.1/H4 to the CAF-1 complex or H3.3/H4 to the HIRA complex(Tagami et al., 2004; Tang et al., 2006; Zhang et al., 2005).

HIRA/ASF1a

The interaction between ASF1a and HIRA was discovered upon the initial pull-down of the HIRA complex form HeLa cells with tagged H3.3(Tagami et al., 2004) and the HIRA B-domain spanning residues 439-475 was identified as the region responsible for contacting the ASF1a N-terminal core(Zhang et al., 2005) (Figure 1.1). The specific amino acid contacts between HIRA and ASF1a were then identified through determination of a 2.7 Å crystal structure of HIRA B-domain residues 446-466 bound to the ASF1a core domain(Tang et al., 2006) (Figure 1.2b). Importantly the structure reveals that HIRA associates with the ASF1a core domain on the ASF1 surface opposite to where ASF1 binds the H3/H4 dimer, indicating that another domain of HIRA or member of the complex is likely capable of binding to H3.3/H4 on the outer surface, where the differing residues between H3.3 and H3.1 reside, to potentiate H3.3/H4 specific binding and chromatin deposition (Figure 1.2a,b).

HIRA/UBN1

The interaction between HIRA and UBN1 is essential for the stability of the HIRA complex as endogenous knockdown of UBN1 results in decreased HIRA

expression, and the same is observed for UBN1 expression when HIRA is knocked down(Banumathy et al., 2009). Additionally UBN1 is essential for H3.3-specific deposition as depletion of UBN1 leads to decreased H3.3/H4 incorporation into chromatin(Ray-Gallet et al., 2011). As the interaction between HIRA and UBN1 is essential for the stability and H3.3/H4 deposition function of the HIRA complex(Banumathy et al., 2009; Ray-Gallet et al., 2011), understanding the molecular basis the HIRA/UBN1 interaction is of great interest. It was initially reported that the N-terminal WD40 repeat domain in HIRA interacted with the HRD in UBN1(Banumathy et al., 2009). As the HRD is the most evolutionarily conserved domain in UBN1 this result initially made biological sense(Balaji et al., 2009; Banumathy et al., 2009). Although upon further biochemical dissection of the HIRA/UBN1 complex it was discovered that a weakly conserved domain slightly N-terminal to the HRD was essential for HIRA binding activity by UBN1, while the HRD was dispensable, this new UBN1 domain N-terminal to the HRD was called the NHRD(Tang et al., 2012) (Figure 1.1).

HIRA/CABIN1

CABIN1 was shown to be dispensable for H3.3/H4 deposition by the HIRA complex(Ray-Gallet et al., 2011), although similarly to UBN1, it appears to be important for the structural integrity of the HIRA complex. Depletion of endogenous CABIN1 levels in cells results in decreased expression of

endogenous HIRA, and depletion of HIRA has the same effect on CABIN1 expression(Rai et al., 2011). Additionally, in the yeast Hir complex, removal of Hir1, Hir2 or Hpc2 results in decreased expression of Hir3(Song et al., 2013). Together these data suggest an important role for CABIN1 in the stability of the HIRA complex. While the entire 2220 amino acids of CABIN1 are spanned by about 30 TPR repeats(Amin et al., 2013), deletion experiments have mapped the N-terminal TPR segment of CABIN1, encompassing residues 1-941, as the minimal region required for HIRA binding(Rai et al., 2011) (Figure 1.1). Deletions probing HIRA have identified the structurally undefined C-terminal domain of HIRA composed from residues 736-963 as responsible for contacting CABIN1(Rai et al., 2011) (Figure 1.1).

UBN1/H3.3/H4

After the NHRD of UBN1 was identified as required for complex formation with the HIRA WD40 repeats, two studies implicated the more strongly conserved UBN1 HRD in H3.3-specific binding and chromatin deposition(Ricketts et al., 2015; Song et al., 2013). Experiments probing function of the yeast Hir complex in H3.3-specific deposition in a human cell system showed that deletion of the HRD in Hpc2 abrogates H3.3/H4 deposition specificity(Song et al., 2013). Following this, a 2.3 Å crystal structure of UBN1 HRD residues 122-142 bound to H3.3/H4/Asf1 was determined and together with biochemical studies revealed that the highly conserved UBN1 HRD binds to H3.3 in the proximity of residues

87-90 to mediate specificity for binding to H3.3/H4 over H3.1/H4(Ricketts et al., 2015) (Figure 1.2c). The crucial H3.3 residue for UBN1 binding specificity appears to be G90, as swapping M90 from H3.1 into H3.3 strongly decreases binding affinity(Ricketts et al., 2015).

Structural model for the HIRA complex

Based on the existing structural data for the HIRA sub-complexes of HIRA/ASF1a and UBN1/H3.3/H4/Asf1 we can infer a structural model where ASF1a associates with histones H3.3/H4 and delivers them to the HIRA complex for deposition through association of ASF1a with the HIRA B-domain, and H3.3/H4 with the UBN1 HRD (Figure 1.2d). As the order of assembly for the HIRA complex is not known, we propose that the UBN1 HRD binds to H3.3/H4 in complex with ASF1a at some point during ASF1a/HIRA complex formation. As the UBN1 HRD has greatly reduced binding affinity for H3.1/H4 in comparison with H3.3/H4(Ricketts et al., 2015), UBN1 likely functions to reject any ASF1a/H3.1/H4 complexes and ensure that only ASF1a bound to H3.3/H4 will be fully accepted for chromatin deposition by the HIRA complex.

1.4 – Model for histone deposition by the HIRA complex

A compilation of recent findings on HIRA complex assembly and H3.3-specific binding and deposition has led to our proposed structural model for the HIRA complex represented by a HIRA/UBN1/ASF1a/H3.3/H4 sub-complex (Figure

1.2d). While all existing data currently point to our proposed molecular assembly, this complex may represent an intermediate in the formation of the $(H3/H4)_2$ tetramer to be deposited into nucleosomes. ASF1a binds to a dimer of H3/H4 along the tetramerization interface and prevents the formation of the $(H3/H4)_2$ tetramer(English et al., 2006). In the absence of ASF1, H3/H4 dimers exist in equilibrium with $(H3/H4)_2$ tetramers(Banks and Gloss, 2004; Bowman et al., 2010; Winkler et al., 2012) (Figure 1.3a). Stable isotope labeling with amino acids in cell culture (SILAC) mass spectrometry data indicates that the majority of H3/H4 is deposited into nucleosomes in the $(H3/H4)_2$ tetramer form and tetramer splitting between nucleosomes is relatively rare(Xu et al., 2010) (Figure 1.3a).

Many histone chaperones (CAF-1(Winkler et al., 2012), MCM2(Huang et al., 2015; Richet et al., 2015), SPT2(Chen et al., 2015), Rtt106(Fazly et al., 2012), Vps75(Bowman et al., 2011), Nap1(Bowman et al., 2011), and SPT16(Tsunaka et al., 2016)) have been shown to stably bind an H3/H4 tetramer. MCM2 has been shown to associate with H3/H4/ASF1 or an $(H3/H4)_2$ tetramer during DNA replication(Huang et al., 2015; Winkler et al., 2012), and similarly to the UBN1 HRD, a loop of MCM2 interacts with the histone surface between the □1 and □2 helices of H3(Huang et al., 2015; Ricketts et al., 2015). One molecule of MCM2 is capable of binding to an H3/H4 dimer associated with ASF1 while two MCM2 molecules can also bind to an $(H3/H4)_2$ tetramer; indicating that MCM2 may act as an intermediate in the formation of $(H3.1/H4)_2$ tetramers for deposition and

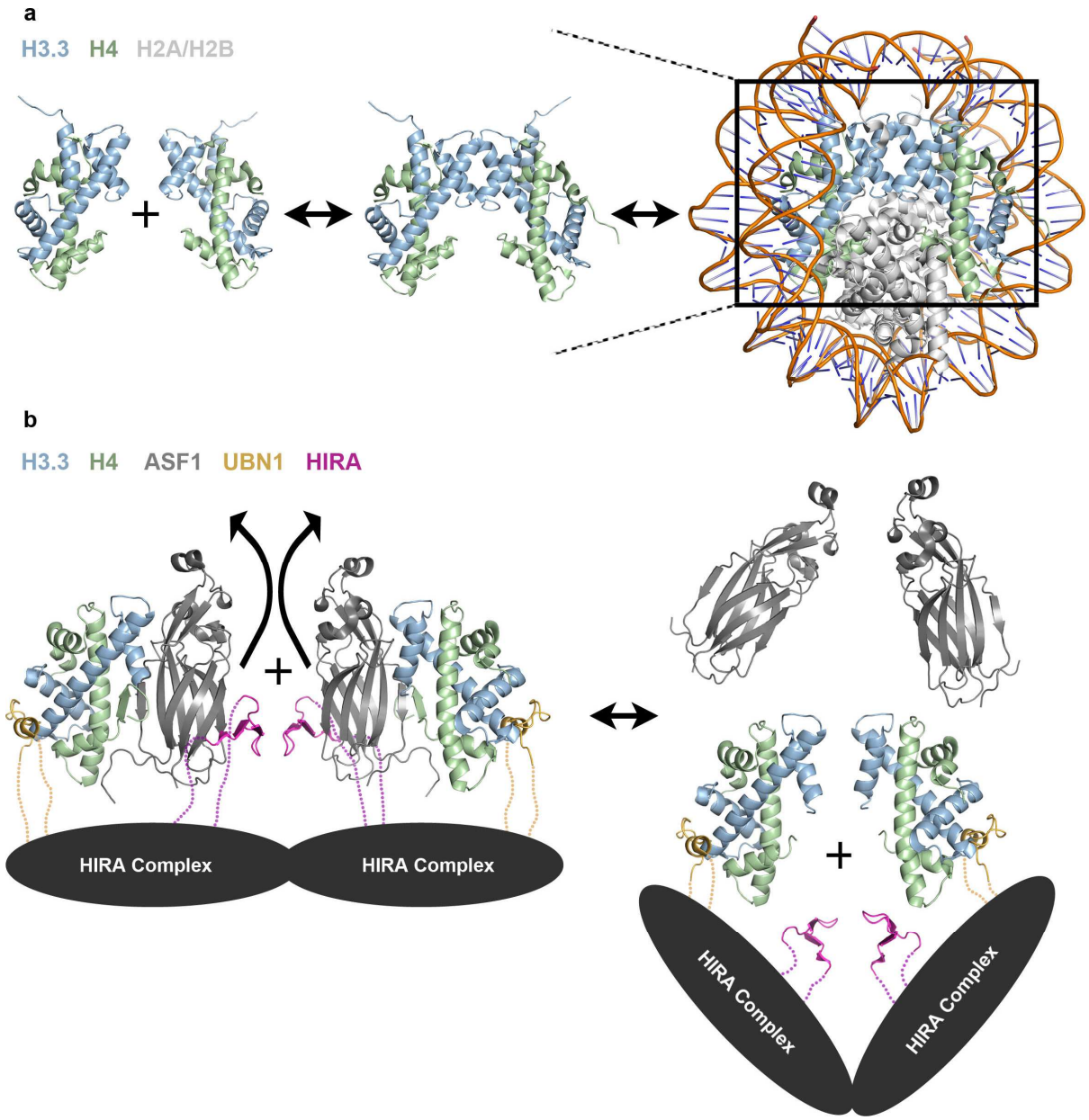


Figure 1.3 – Model for incorporation of H3.3/H4 into nucleosomes by the HIRA complex. (A) In solution, histones H3/H4 exist in equilibrium between dimer and tetramer populations, although strong evidence suggests that they are deposited into the nucleosome in the tetramer form. (B) Model for HIRA complex mediated formation of an $(H3.3/H4)_2$ tetramer prior to nucleosome deposition based on existing structural and biochemical data. The dotted lines represent presumed contacts lacking structural insight.

stabilization of $(H3.1/H4)_2$ tetramers for recycling during DNA replication(Huang et al., 2015). Additionally, several studies indicate that the HIRA complex may be able to still deposit H3.3/H4 even in the absence of ASF1a(Bonnefoy et al., 2007; Ray-Gallet et al., 2007; Ray-Gallet et al., 2011). These data indicate that the ASF1a is likely not a permanent member of the complex and may serve to bind and protect H3.3/H4 before delivering them to the HIRA complex for tetramer formation and chromatin deposition. Based on these observations we propose a model for H3.3/H4 deposition by the HIRA complex where H3.3/H4/ASF1a associates with HIRA/UBN1/CABIN1 complex followed by eviction of ASF1a and UBN1-mediated formation of an $(H3.3/H4)_2$ tetramer prior to deposition onto DNA (Figure 1.3b).

Although crystallographic evidence shows that residues 122-148 of the UBN1 HRD bound to an H3.3/H4 dimer in complex with ASF1(Ricketts et al., 2015), in solution ITC data in the absence of ASF1 indicates that fragments of UBN1 as large as 92-175 can bind an $(H3.3/H4)_2$ tetramer with a stoichiometry of 2 molecules or UBN1 to one $(H3.3/H4)_2$ tetramer(Ricketts et al., 2015). As the UBN1 ITC experiments were conducted with H3/H4 concentrations where the histones should exist almost entirely in the tetramer form(Ricketts et al., 2015; Winkler et al., 2012), it can be concluded that the data represents two molecules of UBN1 associating with identical binding sites on one $(H3.3/H4)_2$ tetramer. Considering both the ITC and crystallographic evidence, it seems possible that

UBN1 may function, similarly to MCM2, to bind H3/H4 dimers associated with ASF1 and mediate tetramer formation (Figure 1.3b).

The molecular mechanism of how two UBN1 molecules cooperate to form the $(H3.3/H4)_2$ tetramer is not known. Although there is no existing evidence for oligomerization of the human HIRA complex, the *S. cerevisiae* Hir complex has been reported to form an assembly of one copy each of Hir1 and Hir3 and two copies each of Hir2 and Hpc2(Prochasson et al., 2005). With two copies of the UBN1 ortholog Hpc2 found in the *S. cerevisiae* complex, it is possible that two HIRA complex molecules may come together for the formation of the $(H3/H4)_2$ tetramer (Figure 1.3b).

Additionally, it is unclear if the HIRA complex cooperates with ATP-dependent chromatin remodelers to mediate deposition of H3.3/H4 onto DNA similarly to the complex formed by the histone chaperone DAXX and chromatin remodeler ATRX. It has been reported that the *S. cerevisiae* Hir complex likely functions to recruit the SWI/SNF complex to histone promoters(Dimova et al., 1999; Prochasson et al., 2005) but a direct interaction between the Hir and SWI/SNF complexes has not been conclusively shown. In mammalian cells the HIRA and CAF-1 complexes have been shown to work in concert with the chromatin remodeler INO80 during DNA double-strand break repair, although there is no evidence for complex formation as INO80 is essential for histone eviction during non-homologous end joining prior to nucleosome reassembly mediated by the

HIRA or CAF-1 complexes independent of INO80(Li and Tyler, 2016). While the possibility remains open, more comprehensive studies are needed to identify how the HIRA complex functions, whether alone or in complex with an ATP-dependent chromatin remodeler.

1.5 – Conclusion

The HIRA complex is an evolutionarily conserved histone chaperone complex that operates to mediate H3.3-specific deposition in a wide variety of chromatin environments. While this review is not meant to be an exhaustive summary of HIRA complex biological function, a more comprehensive review is available(Amin et al., 2013) that features detailed analysis of the biological functions of the *S. cerevisiae* Hir complex and the HIRA complex in a number of species. Through summarizing the existing biochemical and structural data, this review has aimed to improve the current understanding of the architecture of the HIRA complex and to provide a testable model for H3.3-specific histone deposition by the complex. Nonetheless, it is clear that our model would benefit from more structural information, and significant work still needs to be done to uncover the detailed molecular mechanism for the H3.3-specific histone chaperone function of the human HIRA complex.

CHAPTER 2: Ubinuclein-1 confers histone H3.3 binding specificity by the HIRA histone chaperone complex

2.1 – Overview

Histone chaperones bind specific histones to mediate their storage, eviction or deposition from/or into chromatin. The HIRA histone chaperone complex, composed of HIRA, Ubinuclein-1 (UBN1), and CABIN1, cooperates with the histone chaperone ASF1a to mediate H3.3-specific binding and chromatin deposition. Here we demonstrate that the conserved UBN1 Hpc2-related domain (HRD) is a novel H3.3-specific binding domain. Biochemical and biophysical studies show the UBN1-HRD preferentially binds H3.3/H4 over H3.1/H4. X-ray crystallographic and mutational studies reveal that conserved residues within the UBN1-HRD and H3.3 G90 as key determinants of UBN1-H3.3 binding specificity. Comparison of the structure with the unrelated H3.3-specific chaperone DAXX reveals nearly identical points of contact between the chaperone and histone in the proximity of H3.3 G90, although the mechanism for H3.3 G90 recognition appears to be distinct. This study points to UBN1 as the determinant of H3.3-specific binding and deposition by the HIRA complex.

In metazoan cells, there are two major H3 variants: H3.1 is deposited by the trimeric CAF-1 complex during DNA replication(Tagami et al., 2004) and repair of UV-induced DNA damage(Gaillard et al., 1996), while histone H3.3 is deposited in a replication-independent manner(Ahmad and Henikoff, 2002; Tagami et al., 2004) by either DAXX/ATRX largely at heterochromatin, including pericentromeres, telomeres and endogenous retroviral elements (Elsaesser and

Allis, 2010; Elsasser et al., 2015; Goldberg et al., 2010; Wong et al., 2010), or by the HIRA complex predominantly at gene regulatory regions, gene bodies, developmentally regulated genes(Banaszynski et al., 2013; Banumathy et al., 2009; Goldberg et al., 2010; Jin et al., 2009; Orsi et al., 2013; Pchelintsev et al., 2013; Zhang et al., 2007), and sites of DNA and chromatin damage and repair (Adam et al., 2013; Ray-Gallet et al., 2011; Schneiderman et al., 2012; Zhang et al., 2007). However, DAXX/ATRX and HIRA likely also have some overlapping functions. For example, DAXX promotes H3.3 deposition at gene regulatory regions of activated neurons and both HIRA and DAXX are involved in H3.3 deposition in non-proliferating senescent cells(Corpet et al., 2014). While the H3/H4 histone chaperone ASF1a interacts with both CAF-1 and HIRA complexes and binds to both H3.3/H4 and H3.1/H4 for histone deposition(English et al., 2006; Galvani et al., 2008; Natsume et al., 2007; Tagami et al., 2004), CABIN1 and UBN1 are unique members of the HIRA complex(Tagami et al., 2004). CABIN1 and UBN1 were initially identified as a negative regulator for calcineurin signaling in T lymphocytes(Sun et al., 1998) and as a ubiquitously expressed nuclear protein that interacts with cellular and viral transcription factors(Aho et al., 2000), respectively, and later shown to be functional members of the HIRA histone chaperone complex(Banumathy et al., 2009; Orsi et al., 2013; Rai et al., 2011; Tagami et al., 2004).

The human HIRA complex is orthologous to the *S. cerevisiae* Hir histone

chaperone complex, which regulates histone gene transcription(Sherwood et al., 1993; Spector et al., 1997) and mediates replication-independent deposition of H3/H4 in yeast(Green et al., 2005). HIRA is homologous to Hir1 and Hir2(Lamour et al., 1995), CABIN1 is homologous to Hir3(Balaji et al., 2009; Rai et al., 2011; Tagami et al., 2004) and UBN1 is homologous to Hpc2 with strong sequence conservation in the HRD(Balaji et al., 2009; Banumathy et al., 2009). Although the UBN1-HRD was initially proposed to mediate association with the WD repeats of HIRA(Banumathy et al., 2009), we more recently demonstrated that a less conserved region of UBN1 that resides sixty residues N-terminal to the HRD is, in fact, responsible for HIRA-binding activity in UBN1; we named this domain the NHRD(Tang et al., 2012) (Figure 2.1a). These data in combination with the observations of others(Balaji et al., 2009) led us to the hypothesis that that the highly conserved UBN1-HRD may function to confer histone H3.3 binding specificity.

The molecular basis for how the HIRA complex selectively binds to H3.3, which differs from H3.1 by only 5 amino acids, is unknown. In this study we report that H3.3-specific binding by the HIRA complex is mediated by the UBN1-HRD. We demonstrate that the UBN1-HRD specifically binds to H3.3/H4 over H3.1/H4 in a manner that is independent of the protein HIRA. Determination of a crystal structure containing a UBN1-HRD peptide bound to H3.3/H4/Asf1 and mutational studies aided us in identification of residues within the UBN1-HRD and H3.3 that

contribute to the specificity and strength of the UBN1/H3.3 interaction. Additionally the UBN1/H3.3/H4/Asf1 crystal structure reveals striking structural similarity between UBN1/H3.3/H4 and DAXX/H3.3/H4. Together, our results demonstrate that the UBN1-HRD mediates H3.3-specific binding by the HIRA complex and exemplify the evolutionary power of the H3.3 surface in encouraging the disparate histone chaperones UBN1 and DAXX to adopt similar structures.

2.2 – A region of UBN1 containing the HRD specifically binds to H3.3

We initially compared H3.3/H4 and H3.1/H4 binding activity of the HIRA(1-405)/GST-UBN1(41-175) complex to that of GST-UBN1(92-175) (Figure 2.1b). HIRA(1-405) has predicted WD repeats and is the minimum UBN1-binding domain, UBN(41-175) has both the NHRD and HRD while UBN1(92-175) harbors only the HRD. GST-tagged proteins were incubated with H3.1/H4 or H3.3/H4 alone or in competition with the reciprocal His-H3/His-H4 (H3.1 or H3.3) complex and subjected to GST pull-down. We observed H3.3/H4 selective binding over H3.1/H4 for both HIRA(1-405)/GST-UBN1(41-175) and GST-UBN (92-175), indicating that a UBN1 HRD containing protein region can specifically bind H3.3 without contribution from HIRA (Figure 2.1b). In a control pull-down study, we demonstrated that GST-ASF1a binds to H3.1/H4 and H3.3/H4 with equal affinity while GST-UBN1(92-175) specifically binds to H3.3/H4 (Figure 2.2).

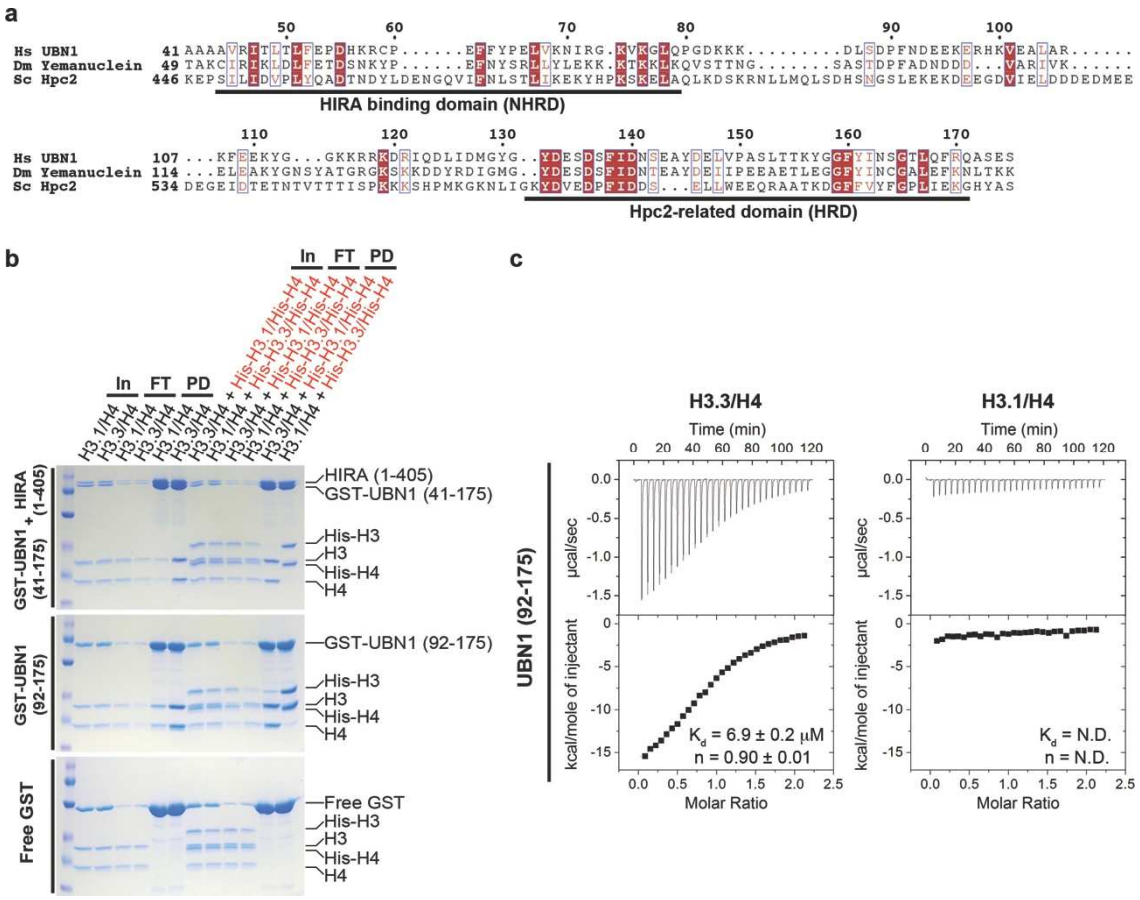


Figure 2.1 – A region of UBN1 containing the Hpc2-related domain (HRD) is able to specifically bind to histone H3.3. (a) Sequence alignment comparing two highly conserved domains in *Homo sapiens* UBN1, *Drosophila melanogaster* Yemanuclein, and *Saccharomyces cerevisiae* Hpc2. (b) GST pull-down histone binding assay. Binding to H3.3/H4 and H3.1/H4 is tested separately in a single input H3/H4 pull-down (left), and in competition by subjecting an equimolar mixture of H3/H4 and His-H3/His-H4 to pull-down (right). The H3.3-specific binding activity of GST-UBN1(41-175)/HIRA(1-405) (top), GST-UBN1(92-175) (middle), and free GST (bottom) is compared. (c) Isothermal titration calorimetry was used to quantitatively compare UBN1(92-175) binding to H3.3/H4 (left) and H3.1/H4 (right), +/- values represent the standard error of the ITC fit using Origin 7.0.

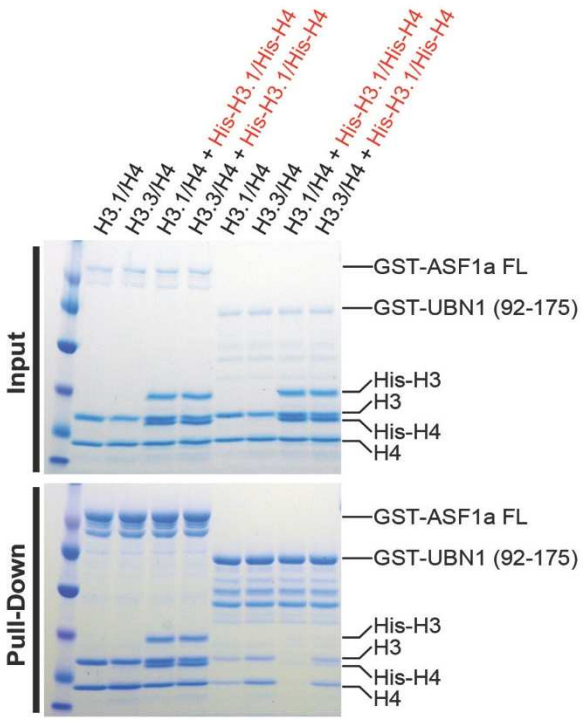


Figure 2.2 – ASF1a control pull-down. UBN1(92-175) (right) is able to selectively bind H3.3/H4 while ASF1a FL (left) is not. Both chaperones are tested for binding to H3.1/H3 and H3.3/H4 alone as well as in competition with His-H3.1/His-H4.

Specificity for H3.3/H4 over H3.1/H4 was especially marked when UBN1 was presented with both H3/H4 variant complexes simultaneously. Specificity for H3.3 was apparent in this assay regardless of whether H3.3 was His-tagged or untagged. Interaction of UBN1(92-175) with H3.3/H4 and H3.1/H4 was monitored

directly with isothermal titration calorimetry (ITC) and indirectly through competition binding against a FITC-labeled UBN1 peptide (Figure 2.1c, Figure 2.3d, Figure 2.5a). Both assays demonstrated that UBN1(92-175) binds to H3.3/H4 with a dissociation constant in the low micromolar range (~ 7-17 □M depending on the assay used), while significant interaction with H3.1/H4 could not be detected with either assay (Figure 2.1c, Figure 2.3d).

2.3 – Conserved HRD residues are required for binding to H3.3/H4

UBN1 residues involved in the interaction with H3.3/H4 were investigated with an alanine scanning mutagenesis experiment targeted at the UBN1-HRD. A survey of the HRD was conducted by testing eighteen GST-UBN1(92-175) triple alanine mutants in the context of GST-UBN1(92-175), spanning residues 122-175, for their ability to pull-down with H3.3/H4 (Figure 2.4). The strongest ablation of binding was observed for mutants of UBN1 residues 135-142 (Figure 2.4) containing some of the most highly conserved residues of the HRD (Figure 2.1a, Figure 2.3). We then focused on this region by generating 27 alanine point mutations spanning UBN1 residues 122-148 (Figure 2.3a). UBN1 point mutants

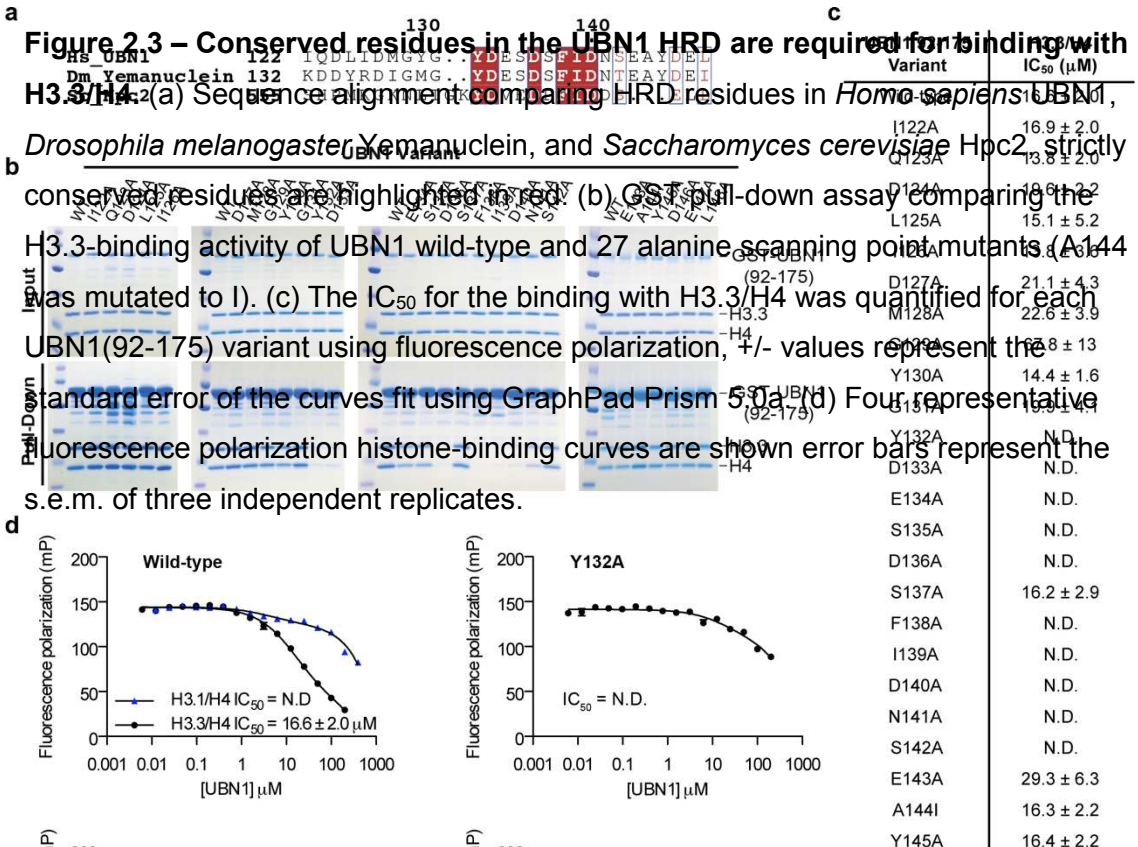


Figure 2.3 – Conserved residues in the UBN1 HRD are required for binding with H3.3/H4. (a) Sequence alignment comparing HRD residues in *Homo sapiens* UBN1, *Drosophila melanogaster* Yemanuclein, and *Saccharomyces cerevisiae* Hpc2, strictly conserved residues are highlighted in red. (b) GST pull-down assay comparing the H3.3-binding activity of UBN1 wild-type and 27 alanine scanning point mutants (A144 was mutated to I). (c) The IC_{50} for the binding with H3.3/H4 was quantified for each UBN1(92-175) variant using fluorescence polarization, +/- values represent the standard error of the curves fit using GraphPad Prism 5.0a. (d) Four representative fluorescence polarization histone-binding curves are shown error bars represent the s.e.m. of three independent replicates.

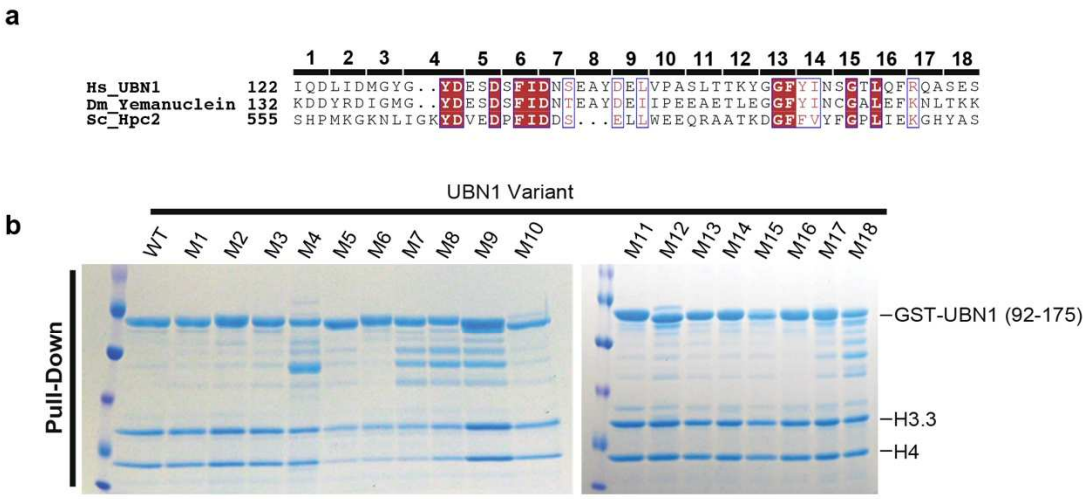


Figure 2.4 – Scanning alanine mutagenesis targeted at the UBN1 HRD. (a) A group of 18 triple alanine mutants were designed across the UBN1 HRD. (b) The gel representing a GST-pull down comparing the H3.3/H4 binding ability of the mutants to wild-type UBN1.

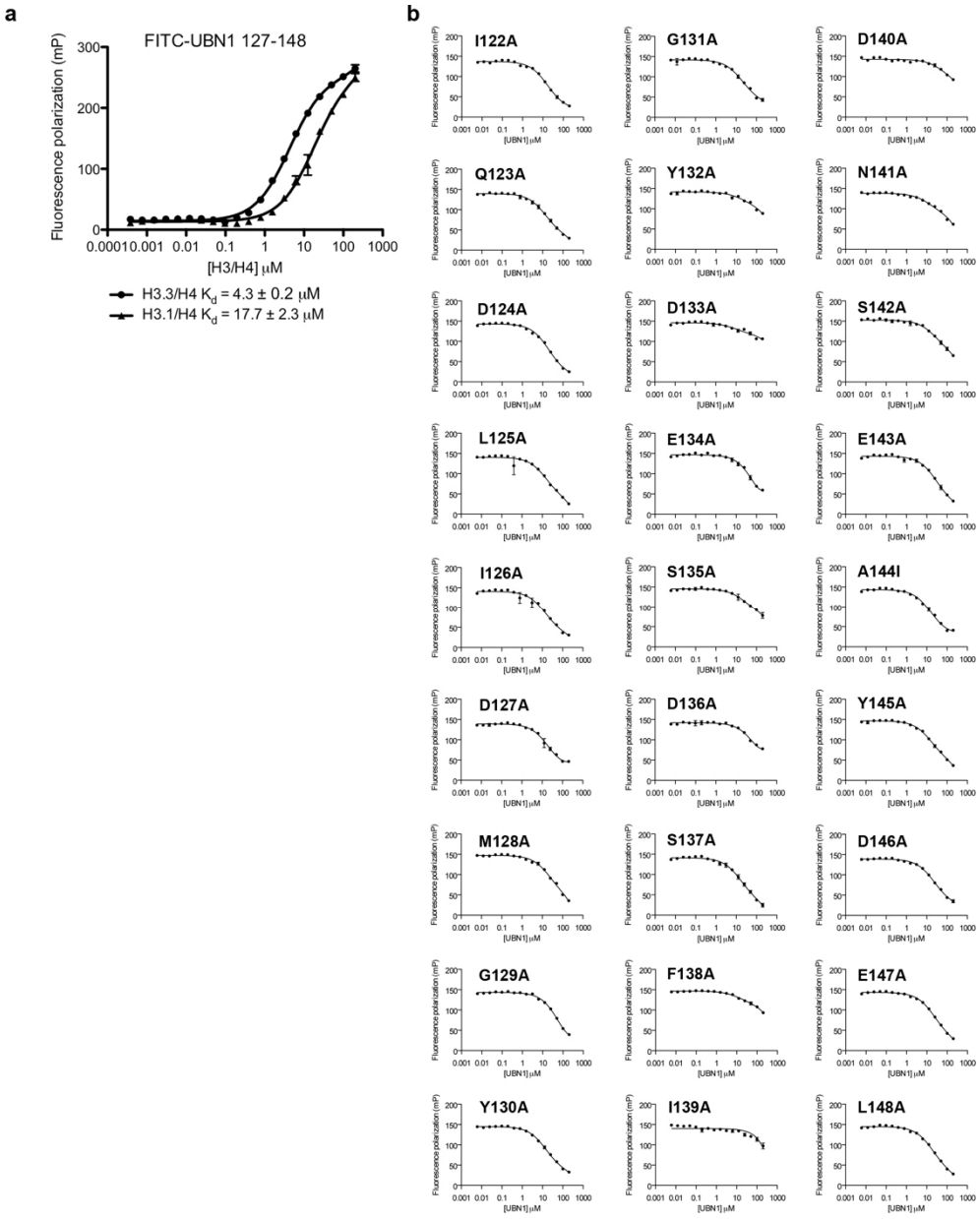


Figure 2.5 – Fluorescence polarization histone-binding assay. (a) FITC-UBN1(127-148) fluorescence polarization binding curves with both H3.3/H3 and H3.1/H4. (b) A group of 27 alanine point mutants of UBN1 (92-175) were analyzed for binding to H3.3/H4 with a competition binding fluorescence polarization assay. Error bars represent the s.e.m. of three independent replicates, +/- values represent the standard error of the curves fit using GraphPad Prism 5.0a

were tested for interaction with H3.3/H4 using GST pull-down assays and fluorescence polarization histone-binding assays. Both experiments showed that the UBN1 Y132A, D133A, D136A, F138A, I139A, D140A, and N141A mutants exhibited reduced histone binding relative to wild-type UBN1 (Figures 2.3b-d, Figure 2.5). All of these residues are strictly conserved from yeast to human (Figure 2.3a), with the exception of N141 (which is sometimes conservatively substitute to a D). Alanine mutants of UBN1 residues 122-148 were additionally subjected to pull-down with H3.1/H4, and these mutants showed either no effect or a significantly smaller effect on UBN1 association with H3.1/H4, suggesting that the most conserved residues within the UBN1-HRD play a predominant role in H3.3 binding specificity (Figure 2.6).

A minimal UBN1-HRD fragment, consisting of residues 122-148, was then investigated to determine if this conserved core region is sufficient for H3.3 binding-specificity. In a pull-down assay GST-UBN1(122-148) exhibited specificity for binding to H3.3/H4 in comparison with H3.1/H4 (Figure 2.7b). ITC and FP histone-binding assays were employed to quantitatively confirm the H3.3-specificity of the UBN1(122-148) peptide (Figure 2.7c, Figure 2.8b). While UBN1(122-148) maintains binding specificity for H3.3/H4 in comparison with H3.1/H4, UBN1(92-175) binds to H3.3/H4 about 3 to 6-fold more strongly than UBN1(122-148). This suggests that residues flanking 122-148 of the UBN1 HRD also contribute to H3.3-binding affinity. UBN1(92-148) and UBN1(122-175)

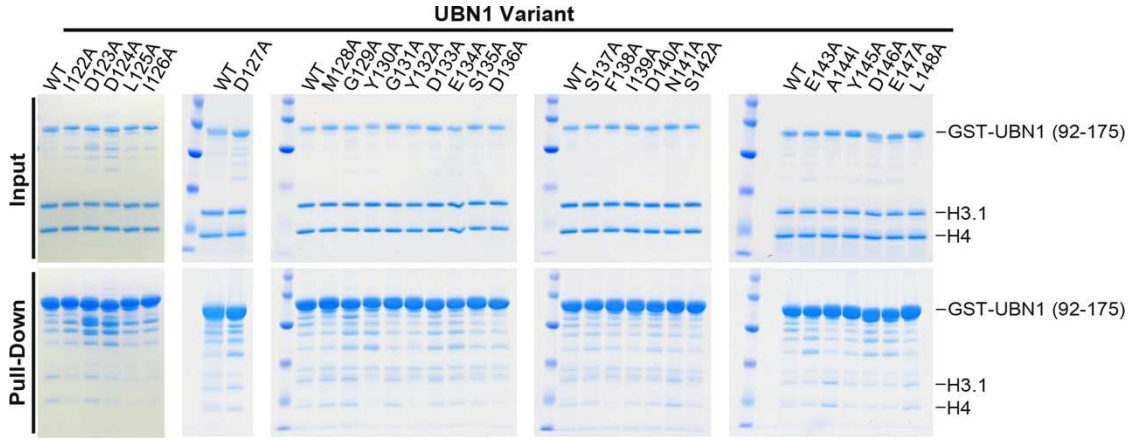


Figure 2.6 – GST pull-down to test the H3.1/H4 binding ability of alanine mutants of UBN1(92-175).

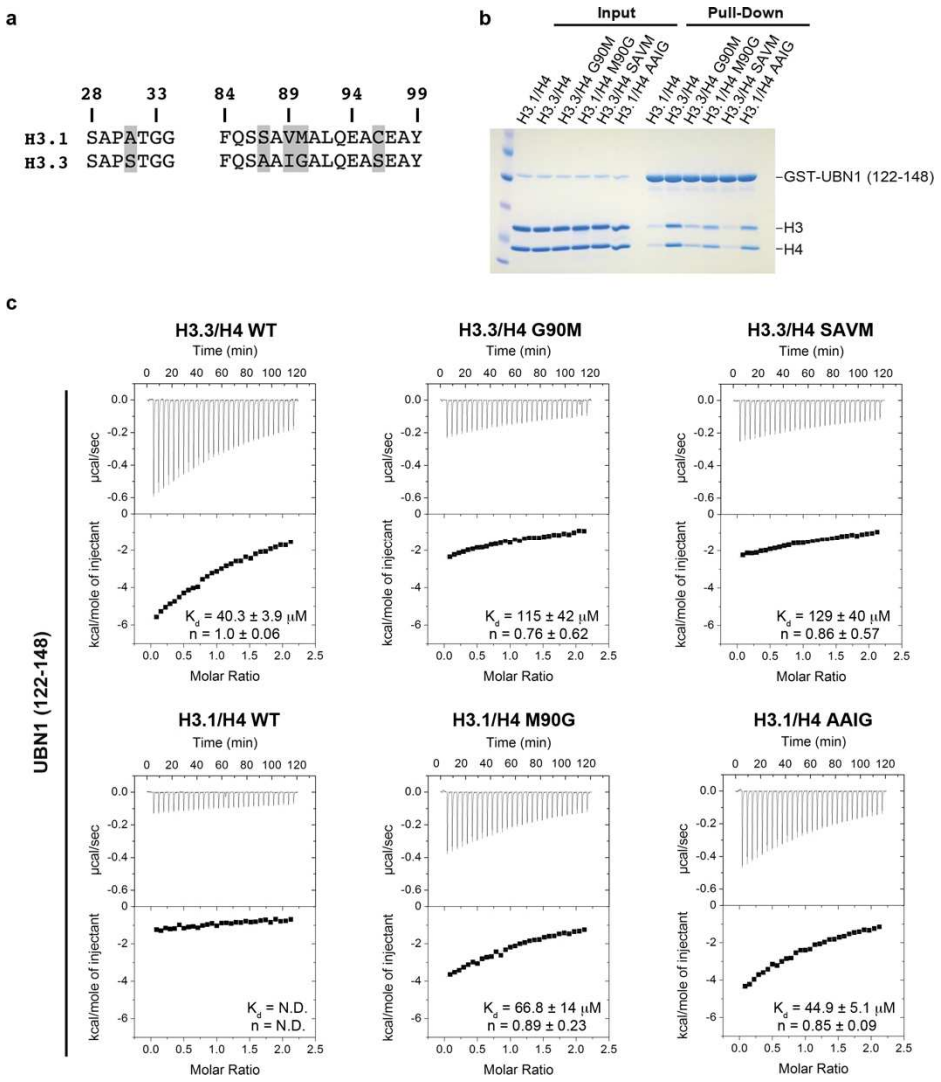


Figure 2.7 – H3.3 residue G90 mediates UBN1 specificity. (a) Sequence alignment comparing human histones H3.1 and H3.3 with the five amino acid differences highlighted in grey (b) GST pull-down assay comparing the ability of UBN1(122-148) to select between histone mutants where critical residues have been swapped between H3.1 and H3.3. UBN1 binding with wild-type H3.3/H4 and H3.1/H4 was analyzed in addition to H3/H4 carrying several H3 mutants: H3.3(G90M), H3.1(M90G), H3.3(SAVM), and H3.1(AAIG). (c-h) Isothermal titration calorimetry quantification of the UBN1(122-148) interaction with wild-type and mutant H3/H4 complexes. Dissociation constant (K_d) and stoichiometry of UBN1 binding to H3/H4 are reported, +/- values represent the standard error of the ITC fit using Origin 7.0.

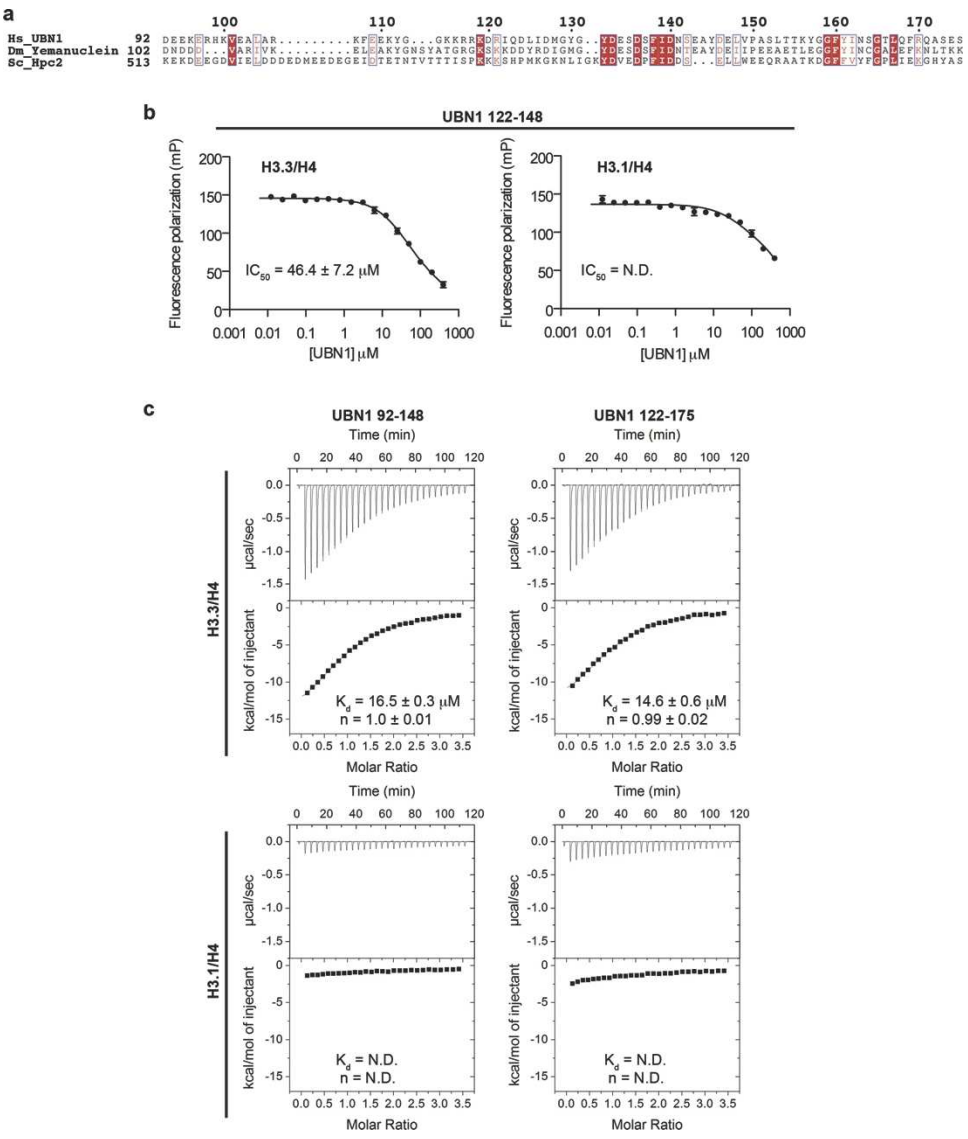

a
100
110
120
130
140
150
160
170
Hs_UBN1 92
Dm_Yemanuclein 102
Sc_Hpc2 513
b
UBN1 122-148
H3.3/H4
IC50 = 46.4 ± 7.2 μM
H3.1/H4
IC50 = N.D.
Fluorescence polarization (mP)
[UBN1] μM
c
UBN1 92-148
UBN1 122-175
Time (min)
H3.3/H4
H3.1/H4
μcal/sec
kcal/mol of injectant
Kd = 16.5 ± 0.3 μM
n = 1.0 ± 0.01
Kd = 14.6 ± 0.6 μM
n = 0.99 ± 0.02
Kd = N.D.
n = N.D.
Molar Ratio

Figure 2.8 – Different HRD-containing UBN1 constructs binding with both H3.3/H4 and H3.1/H4. (a) Sequence alignment comparing residues in *Homo sapiens* UBN1, *Drosophila melanogaster* Yemanuclein, and *Saccharomyces cerevisiae* Hpc2, strictly conserved residues are highlighted in red. (b) Fluorescence polarization competition binding assay to quantify the binding of UBN1(122-148) to both H3.3/H4 and H3.1/H4. (c) Isothermal titration calorimetry was used to confirm that both UBN1 (92-148) and (122-175) fragments specifically bind to an H3.3 containing H3/H4 complex, +/- values represent the standard error of the ITC curves fit using Origin 7.0.

were analyzed for binding to H3.3/H4 and H3.1/H4 with ITC to address the contributions of these flanking residues. We found that both UBN1 fragments bind H3.3/H4 with similar dissociation constants (K_D = ~14-16 □M) while binding with H3.1/H4 was undetectable (Figure 2.8c). This binding capacity is stronger than UBN1(122-148) but weaker than UBN1(92-175), further suggesting that UBN1 residues 92-121 and 149-175 contribute to the binding affinity of H3.3/H4.

2.4 – H3.3 residue G90 mediates UBN1 specificity

To establish the amino acids in H3.3 and H3.1 that confer specific binding of the former to UBN1-HRD, we analyzed the five residues that differ between H3.3 and H3.1 with a residue-swap mutagenesis experiment. The five differing residues are clustered in three patches along the sequence of H3 (S31A, SAVM87-90AAIG and C96S) (Figure 2.7a, Figure 2.9a). We first tested H3.3(S31A), H3.3(SAVM) and H3.3(S96C) for binding to GST-UBN1(92-175) in pull down assays, in competition with wild-type His-H3.1/His-H4 (Figure 2.9b). Only the H3.3(SAVM) mutant abolished H3.3-specific binding from UBN1 (Figure 2.9b).

To assess the individual residue effects from this patch, H3.3 A87S, I89V and G90M mutants were applied in the same pull-down assays; only H3.3(G90M) prevented binding to UBN1 (Figure 2.9c). The reverse experiment, with H3.3 residues introduced into H3.1, showed that the H3.1(M90G) mutant uniquely acquired binding to UBN1 (Figure 2.9d). GST-ASF1a FL and GST-UBN1(92-175) were

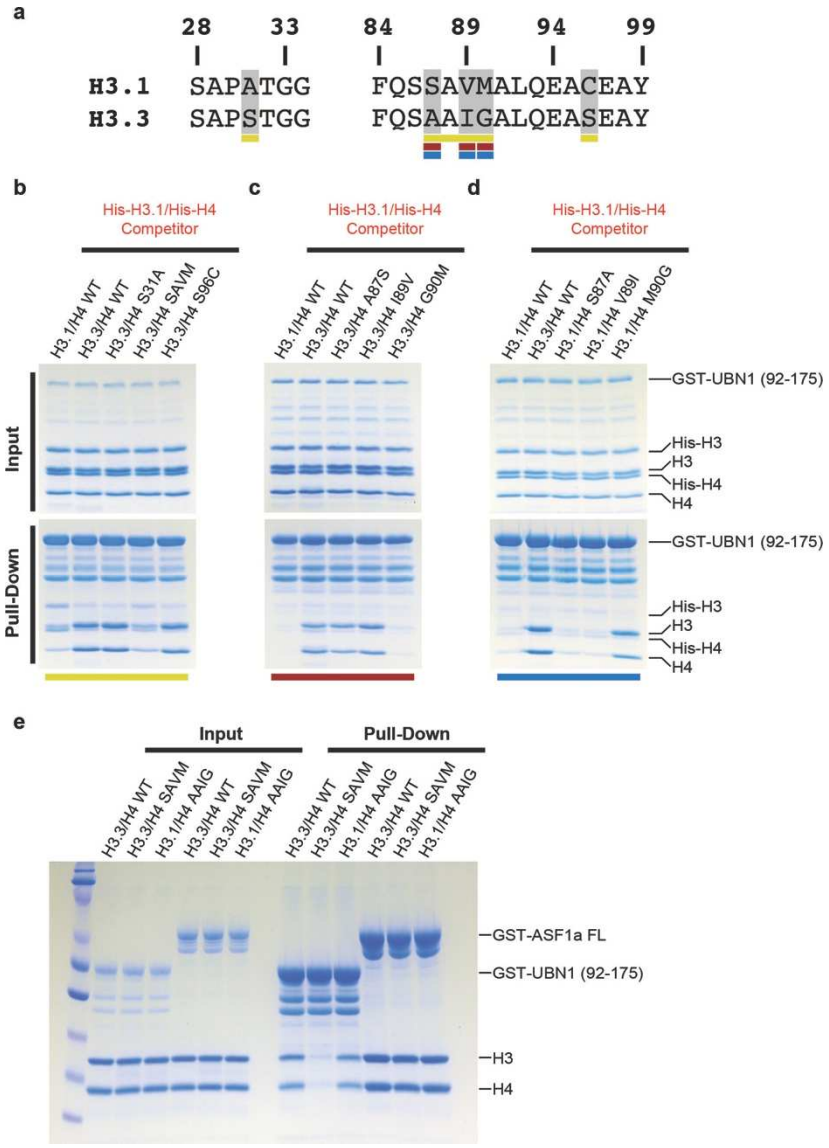


Figure 2.9 – Binding of UBN1 to H3.3 wild-type and mutants. (a) Sequence alignment highlighting the 5 differing residues between H3.3 and H3.1. Residues involved in pull-downs b, c and d are respectively underlined in yellow, red and blue. (b) GST pull-down indicating that the H3.3 residues AAIG mediate UBN1 specificity. (c) GST pull-down indicating that the H3.3 residue G90 is the crucial residue in the AAIG cluster responsible for UBN1 binding specificity. (d) Reverse GST pull-down experiment that shows that the lack of specificity in UBN1 for H3.1 can be rescued with a M90G mutant.

compared for selectivity between H3.3(SAVM) and H3.1(AAIG) mutants, and only GST-UBN1(92-175) exhibited binding specificity for H3.1(AAIG) (Figure 2.9e). GST-UBN1(122-148) was analyzed for selectivity between the crucial histone residue-swap mutants and showed decreased binding to H3.3(SAVM) and H3.3(G90M), but increased binding to H3.1(AAIG) and H3.1(M90G) (Figure 2.7b). A more quantitative analysis of the effect of H3.1 and H3.3 mutants on UBN1(122-148) binding was performed with ITC and demonstrated that the H3.3(G90M) and H3.3(SAVM) mutants had about a 3-fold decrease in UBN1(122-148) binding relative to H3.3 WT, and that the H3.1(M90G) and H3.1(AAIG) swap mutants bound UBN1(122-148) nearly as well as H3.3 WT (Figs. 2.7c).

2.5 – UBN1/H3.3/H4/Asf1 structure reveals mode of H3.3 binding

Having determined that the UBN1 HRD is necessary and sufficient for selective binding of H3.3 over H3.1 to better understand this at the molecular level, we set out to determine a crystal structure of the UBN1/H3.3 interaction. We were unable to obtain crystals of UBN1 in complex with H3.3/H4 alone. Previously a structure of the Asf1 core domain bound to H3.1/H4 lacking N-terminal tails was determined. We obtained this construct and modified it to generate a Asf1/H3.3/H4 complex that we successfully used for cocrystallization with a UBN1 peptide harboring the HRD region. From those crystals we determined the 2.3 Å X-ray crystal structure of a complex containing human UBN1(122-142)

bound to *X. laevis* H3.3(61-134), *X. laevis* H4(24-101), and *S. cerevisiae* Asf1(1-154) (Figure 2.10a, Table 2.1). Not surprisingly, the Asf1/H3.3/H4 component of the structure superimposes perfectly with the previously reported Asf1/H3.1/H4 structure (RMSD = 0.821 Å for all identical atoms) since the differences between H3.3 and H3.1 do not reside at the Asf1-histone interface. An omit map confirms electron density for UBN1 residues 122-142 (Figure 2.11). UBN1 residues 122-127 form a short helix and residues 128-142 a coil, with the coil wrapping over a surface of H3.3 formed by helices 1 and 2 and the intervening turn. The UBN1 helix (residues 122-127) projects away into solvent (Figure 2.10b). Confirming the biochemical studies, the UBN1 HRD makes intimate interactions with H3.3 centered around G90 (Figure 2.10). The UBN1-H3.3 interface is nucleated by UBN1 residues F138 and Y132 that protrude into largely hydrophobic pockets on H3.3 and position the coil of residues 128-132 in a saddle that sits over the top of H3.3 G90 (Figure 2.10b). The hydroxyl group of UBN1 Y132 forms a hydrogen bond with the side chain of H3.3 Q93 (Figure 2.10c). The aromatic ring of UBN1 Y132 makes extensive van der Waals interactions with the aliphatic side chain of H3.3 K64; a hydrogen bond with the UBN1 Y130 backbone carbonyl locks H3.3 K64 into place above UBN1 Y132 (Figure 2.10c). A van der Waals interaction between the beta carbon of UBN1 D133 and the side chain of H3.3 L65 may also help anchor UBN1 Y132 into position (Figure 2.10c). UBN1 F138 sits in a hydrophobic pocket where it forms an extensive network of van der Waals

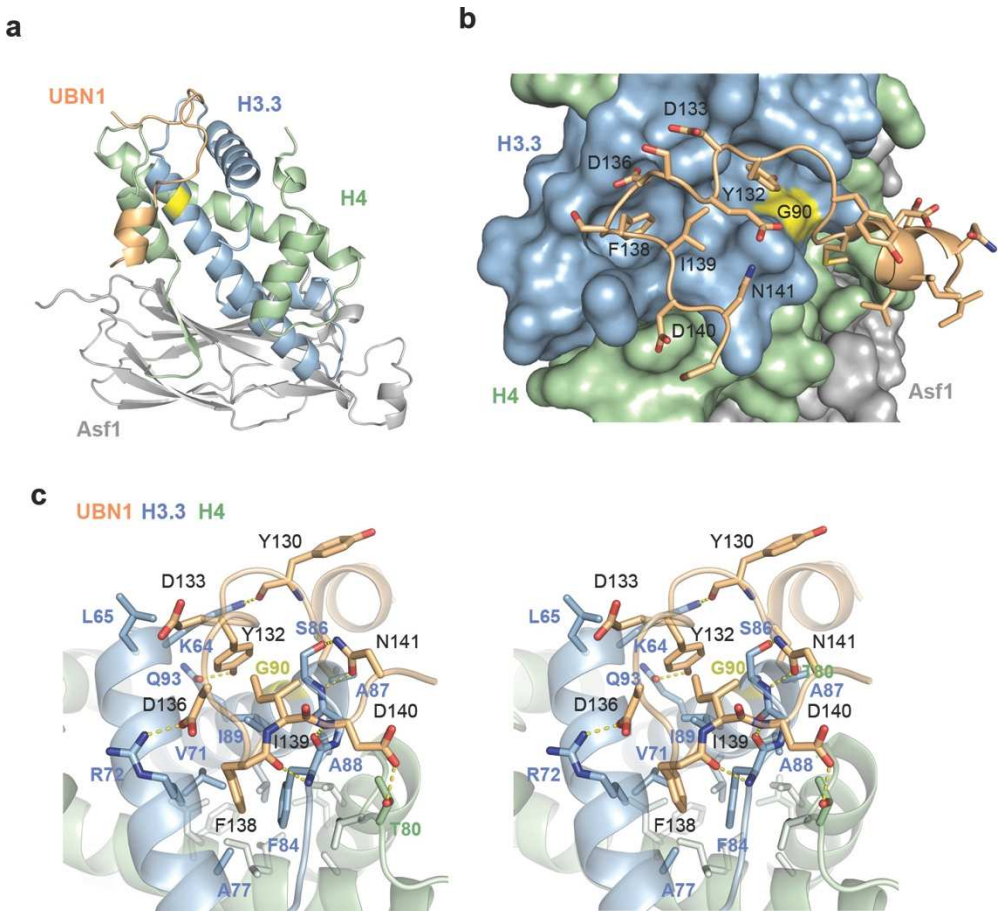


Figure 2.10 – Crystal structure of the UBN1/H3.3/H4/Asf1 complex at 2.3 Å resolution. (a) Overall architecture of the UBN1/H3.3/H4/Asf1 complex with the position of H3.3 G90 highlighted in yellow. The structure is composed of UBN1 122-148, H3.3 61-134, H4 24-102, and Asf1 1-154. (b) UBN1 has an intimate association with the H3.3/H4 surface with close proximity to H3.3 G90 (yellow). UBN1 residues with sensitivity to alanine mutation are labeled. (c) Detailed stereo view of UBN1 interactions with H3.3/H4.

Table 2.1. X-ray data collection and refinement statistics

Data collection[a]	
Space Group	$P3_221$
Cell Dimensions	
a, b, c (Å)	90.024, 90.024, 120.734
α, β, γ (°)	90.00, 90.00, 120.00
Wavelength (Å)	1.5418
Resolution (Å)	25.41 - 2.25 (2.33 - 2.25)[b]
R_{sym}	5.4 (44.0)
$I / \sigma I$	26.4 (2.7)
Completeness (%)	99.1 (93.3)
Redundancy	6.0 (3.8)
Refinement statistics	
Resolution (Å)	25.41 - 2.25 (2.33 - 2.25)
No. reflections (total / unique)	725,259 / 26,431
R_{work}/R_{free} (%)	20.82 / 23.48
No. Atoms	2866
Protein	2692
Ligand/ion	16
Water	158
B-factors (Å^2)	
Protein (Asf1/H3.3/H4)	54.7
Protein (UBN1)	67.2
Ligand/ion	58.6
Water	53.4
R.m.s. deviations	
Bong lengths (Å)	0.003
Bond angles (°)	0.637
Ramachandran statistics	
Favored (%)	99.1
Allowed (%)	0.90

[a]One crystal was used for data collection and refinement.

[b]Values in parentheses are for the highest resolution shell.

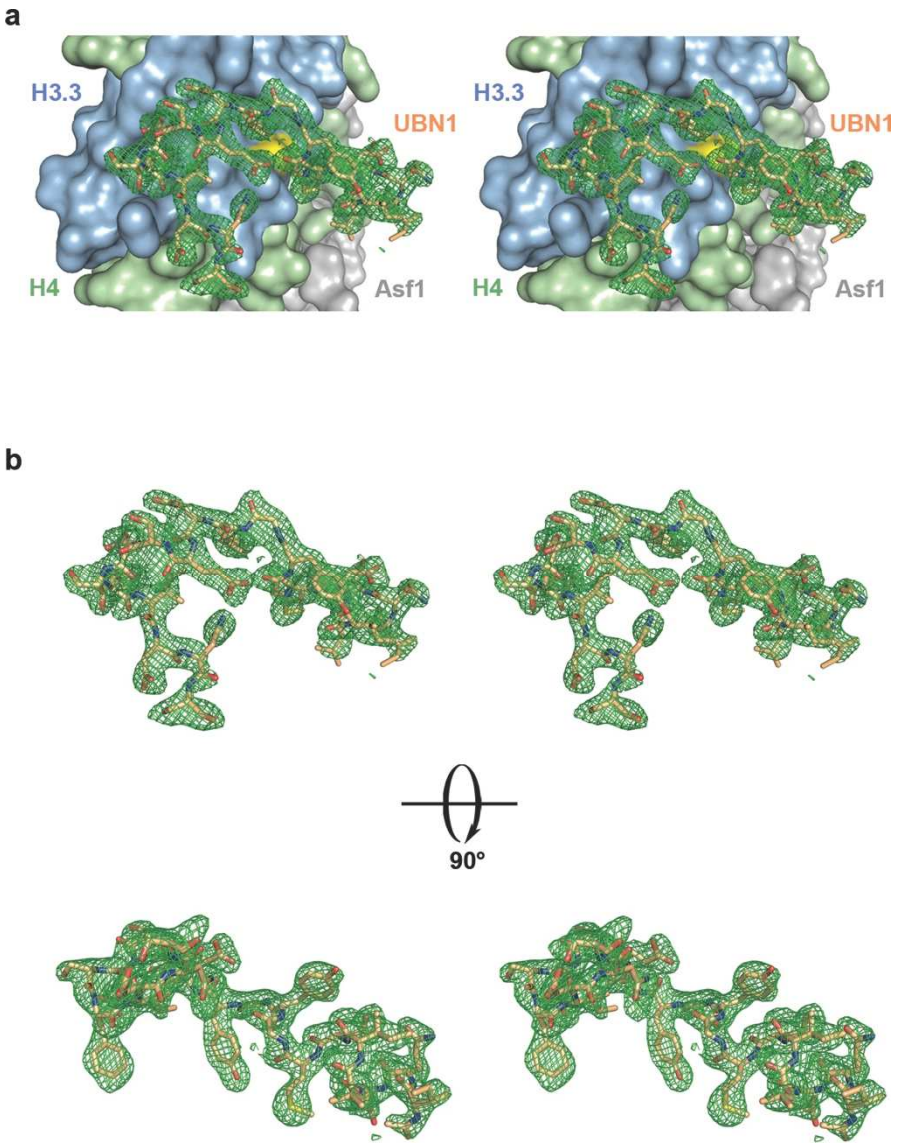


Figure 2.11 – Simulated annealing F_o - F_c omit map (contoured at 2σ) presented in stereo illustrating the high confidence of the UBN1 structural model. (a) Omit map was overlaid with the UBN1 structure (stick model) in the context of the bound H3.3/H4/Asf1 that is shown in surface representation in distinct colors. (b) View of the UBN1 omit map/structure overlay alone, with the same orientation as in (a) (top panel), or rotated 90 degree (bottom panel) for clarity of the various side chains.

interactions with the aliphatic chain of H3.3 R72 and the aromatic ring of H3.3 F84 ,as well as supporting interactions with H3.3 V71, A77 and I89 (Figure

2.10c). UBN1 I139 makes van der Waals interactions with the H3.3 I89 side chain and helps to cap the hydrophobic pocket in which UBN1 F138 binds (Figure 2.10c). UBN1 D136 and D140 form hydrogen bonds with H3.3 R72 and H4 T80 respectively, while the UBN1 N141 side chain makes hydrogen-bonding interactions with both the backbone amide and the side chain hydroxyl of H3.3 S86 (Figure 2.10c). The backbone carbonyl from UBN1 F138 and backbone amide from UBN1 D140 form hydrogen bonds with the backbone amide and carbonyl from H3.3 F84 to further solidify the UBN1/H3.3 interaction (Figure 2.10c). Notably, the structure reveals that evolutionarily conserved UBN1 residues whose mutation disrupts H3.3 binding form hydrogen bond or van der Waals interaction with H3.3/H4 (Figure 2.10c). UBN1 residues that are not sensitive to mutation (E134, S135, and S137) do not make contact with H3.3/H4 (Figure 2.12).

2.6 – Asf1 and histones tails do not alter UBN1/H3.3 interaction

To confirm the veracity of the interactions between UBN1 and H3.3/H4 observed in our crystal structure, we investigated how the presence of Asf1 in complex with H3/H4 affects the interaction between UBN1 and H3/H4. We conducted ITC to determine the affinity of UBN1(122-148) for complexes assembled with Asf1 and

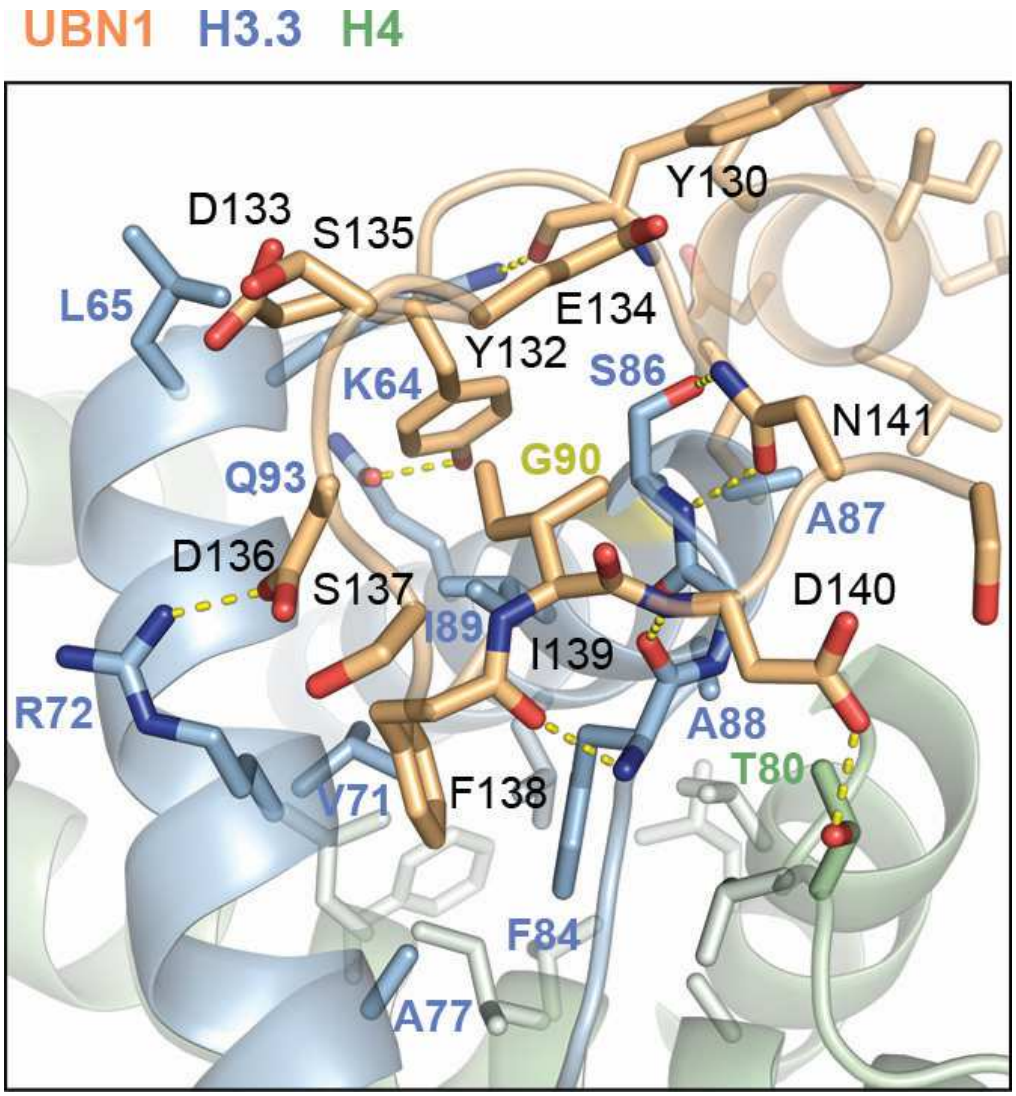


Figure 2.12 – Detailed view of the UBN1/H3.3/H4 interaction including UBN1 side chains E134, S135, and S137, which have no interaction with H3.3/H4.

FL H3/H4 or H3(60-135)/H4(20-102) lacking the N-terminal tails (Supplementary Figure9). We found that UBN1(122-148) interacts with the Asf1/H3.3(FL)/H4(FL) complex with a dissociation constant (~49 □M) comparable to that of UBN1(122-148) binding to H3.3/H4 alone (~40-46 □M depending on the assay) ; no significant interaction between UBN1(122-148) and Asf1/H3.1(FL)/H4(FL) was detected (Figure 2.13a). We then tested UBN1 (122-148) binding with Asf1/H3.3(60-135)/H4(20-102). The dissociation constant determined for this interaction (~110 □M) is roughly 2 fold weaker than for UBN1(122-148) binding to the Asf1/H3.3(FL)/H4(FL) complex, although specificity for H3.3/H4 over H3.1/H4 was still evident (Figure 2.13b). These data suggest that the presence of Asf1 in complex with the histones does not significantly influence the binding and specificity of the UBN1-HRD for H3.3/H4 over H3.1/H4.

2.7 – UBN1 and DAXX bind to H3.3/H4 with structural similarity

A comparison of the structure reported here with the DAXX/H3.3/H4 complex(Elsasser et al., 2012) reveals that while the histone chaperones UBN1 and DAXX share no sequence conservation, they bind the region of H3.3/H4 surrounding G90 with striking similarity (Figure 2.14 a-b, Table 2.2). The structural similarity between how UBN1 and DAXX bind H3.3/H4 is evident although the two chaperones have opposite orientations of N to C-terminal sequence when bound to H3.3/H4 and no conservation when aligned in either orientation (Figure 2.14 b).

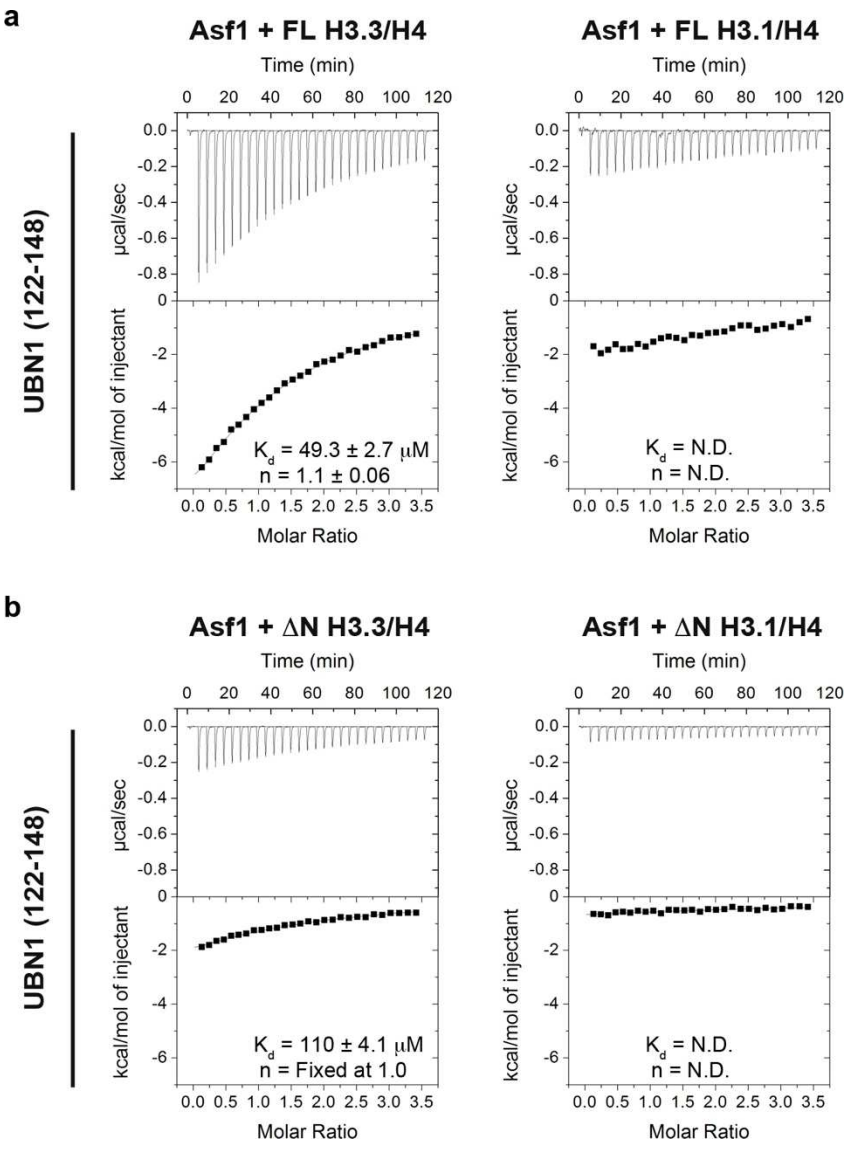


Figure 2.13 – ITC analysis of UBN1 binding to Asf1/H3/H4 complexes containing FL and N-terminally truncated histones. (a) ITC analysis of UBN1(122-148) interaction with FL H3.3/H4 or FL H3.1/H4 in complex with Asf1(1-169). (b) ITC analysis of UBN1(122-148) interaction with H3.3(60-135)/H4(20-102) or H3.1(60-135)/H4(20-102) in complex with Asf1(1-169), +/- values represent the standard error of the ITC curves fit using Origin 7.0.

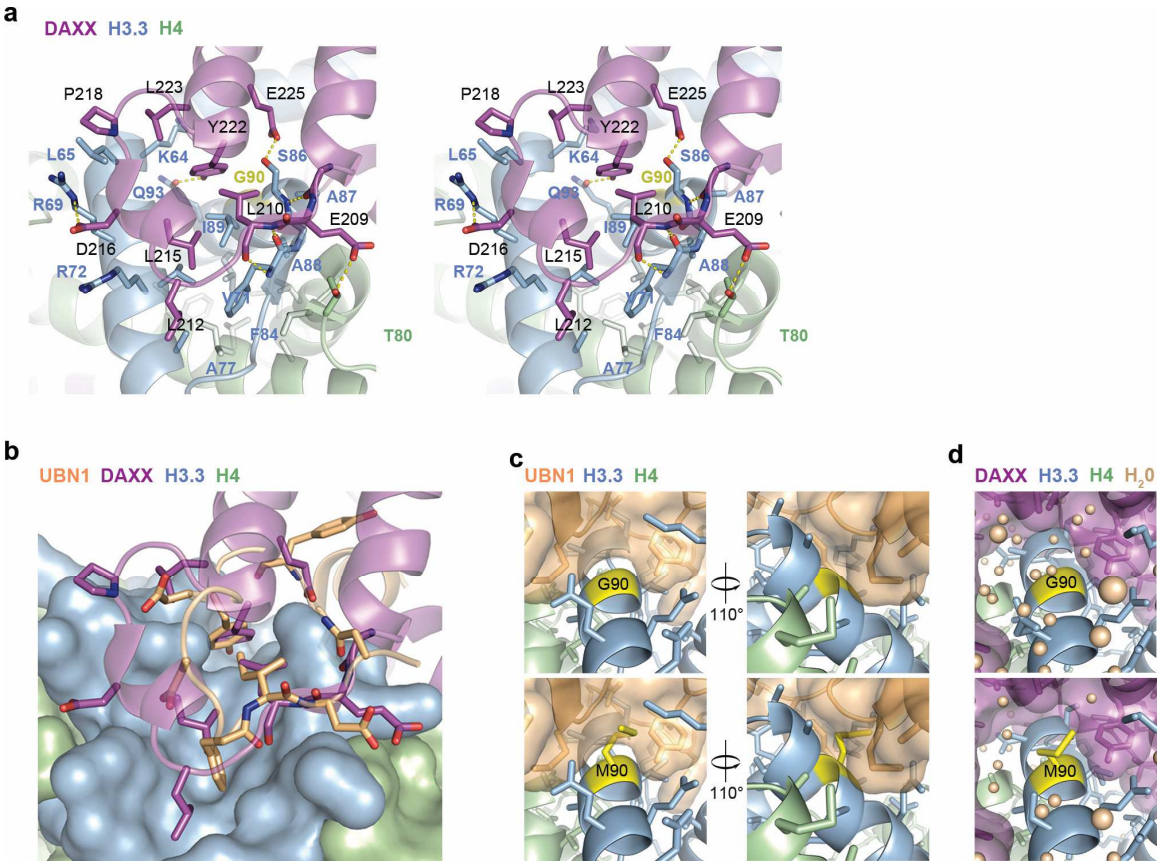


Figure 2.14 – Comparison of UBN1 and Daxx interactions with H3.3/H4. (a) Detailed stereo view of DAXX interactions with H3.3/H4 showing conserved interaction surfaces. (b) Alignment of the UBN1/H3.3/H4/Asf1 and DAXX/H3.3/H4 structures illustrating structural similarity between UBN1 and DAXX. (c) Surface representation of UBN1 shows that residues 128-132 rest closely over H3.3 G90 (top) and would likely sterically occlude binding with H3.1 M90 (bottom). M90 is represented by substitution of G90 with the highest percentage methionine rotamer. (d) The same view of the DAXX/H3.3/H4 structure shows that DAXX has a cavity where UBN1 128-131 resides. The comparable DAXX site fills with water molecules to sterically occlude binding with H3.1/H4.

Table 2.2 – Comparison of Daxx and UBN1 residues

H3.3 Residue	H4 Residue	UBN1 Residue	DAXX Residue
K64		Y130, Y132	Y222, L223
L65		D133	P218, L223
R69		X	D216
V71		F138	L215
R72		D136, F138	L212, L215
A77		F138	L212
F84		F138	L215
S86		N141	E225
I89		F138, I139	L210, L215
	T80	D140	E209
F84 - S86*		F138 - N141	K208 - L210

*Backbone contacts are formed within this group of residues

DAXX residues Y222, L210, and E209 are in essentially the same positions and mediate the same interactions as UBN1 residues Y132, I139 and D140, respectively (Figures 2.10c, 2.14a, and 2.15) In addition, L212 and L215 from DAXX maintain the same van der Waals interaction network that UBN1 F138 forms with H3.3 (Figures 2.10c and 2.14a). Similar backbone interactions are also maintained between the DAXX and UBN1 chaperones with H3.3 (Figures 2.10c, 2.14a). Table 2.2 summarizes the contacts that UBN1 and DAXX make with key H3.3/H4 residues.

While DAXX and UBN1 contact H3.3/H4 in much the same way, they do not use the same mechanism to select for H3.3 over H3.1. UBN1 residues M128-Y132 sit in close proximity above H3.3 G90 (Figures 2.10b and 2.14c). A modeling of H3.3 G90M bound to UBN1 predicts that UBN1 would directly clash with a methionine at this position, suggesting that UBN1 employs a mechanism of steric occlusion to prevent unfavorable binding with H3.1 (Figure 2.14c). In contrast, the corresponding region within the DAXX/H3.3/H4 complex is solvent accessible and filled with a network of water molecules that are proposed to provide steric bulk to occlude binding with H3.1[24] (Figure 2.14d). Although G90 is the primary H3.3 residue responsible for UBN1 specificity other H3.3-specific residues may contribute to the UBN1 interaction. The delta carbon from H3.3 I89 forms a van der Waals interaction with UBN1 I139 and the H3.3 A87 beta carbon forms van

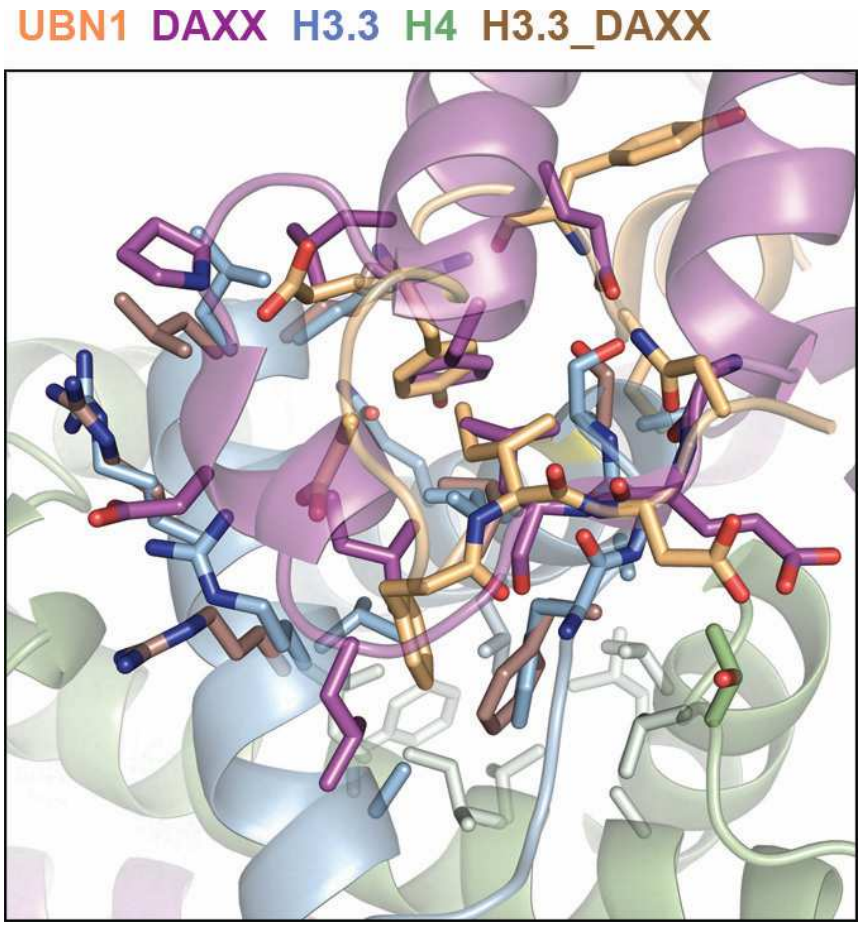


Figure 2.15 – Overlay of UBN1 and DAXX residues that contact the H3.3/H4 surface with close proximity to H3.3G G90. H3.3 residues from the DAXX/H3.3/H4 crystal structure that differ significantly from those in the UBN1/H3.3/H4/Asf1 structure are shown in brown.

der Waals interactions with the beta and gamma carbons from UBN1 M128

(Figure 2.10c). Both of these residues are unique to H3.3 and may be responsible for the increased affinity that UBN1 has for the SAVM(87-90)AAIG mutation in H3.1 in comparison with a loan M90G mutation (Figure 2.7c). Taken together, while UBN1 and DAXX bind to the same region of histone H3.3, they achieve specificity for H3.3 over H3.1 by distinct mechanisms.

2.8 – H3.3-specific binding occurs within HIRA/UBN1/CABIN1 complex

To determine the contributions of the UBN1-HRD to H3.3 specificity in more complete HIRA and CABIN1-containing complexes, we prepared complexes containing full-length HIRA, UBN1 and CABIN1 proteins in wild-type form or bearing a UBN1 FID138-140AAA mutant, and assayed the ability of these complexes to pull down H3.3/H4 and H3.1/H4. This experiment demonstrated that HIRA/UBN1 and HIRA/UBN1/CABIN1 complexes assembled from full-length proteins preferentially bind to H3.3/H4 over H3.1/H4, and the preference is abolished in complexes containing the UBN1 FID138-140AAA mutant (Figure 2.16).

2.9 – Discussion

Our data demonstrate that the UBN1 HRD is an evolutionarily conserved H3.3-specific binding domain that exploits residue G90 in H3.3 to mediate specific binding and so deposition by the HIRA histone chaperone complex. Many studies

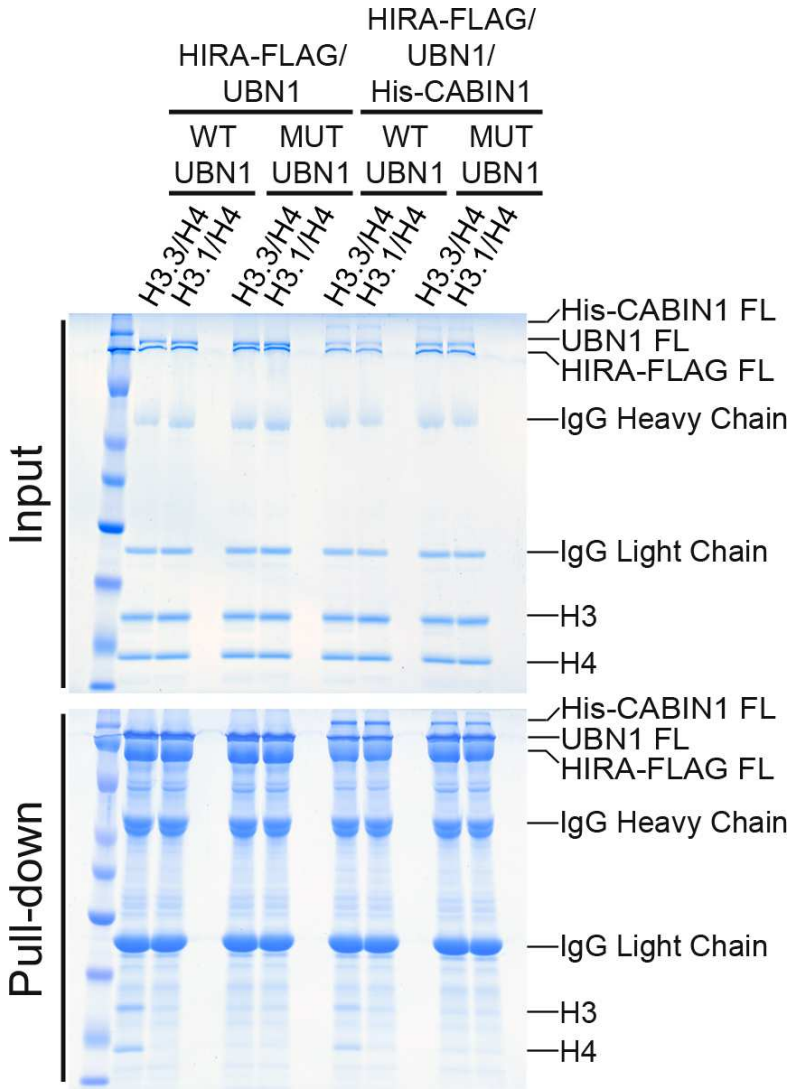


Figure 2.16 – UBN1-HRD/H3.3 interface is required for H.3.3 binding specificity in intact complexes. FLAG pull-down histone binding assay. Recombinant complexes of HIRA-FLAG(FL)/UBN1(FL) and HIRA-FLAG(FL)/UBN1(FL)/His-CABIN1(FL) with wild-type and FID138-140AAA mutant UBN1 were analyzed for binding to both H3.3/H4 and H3.1/H4.

have shown that HIRA complex specifically binds to and deposits H3.3 and is required for deposition of H3.3 at selected genomic regions (Banaszynski et al., 2013; Corpet et al., 2014; Elsasser et al., 2015; Goldberg et al., 2010; Loppin et al., 2005; Tagami et al., 2004; Zhang et al., 2007). Our structural analysis shows this specificity for H3.3 comes primarily from HIRA-bound UBN1, not HIRA itself.

The specificity of the UBN1 HRD for G90 may be evolutionarily conserved with the Hpc2 subunit of the yeast Hir complex. Recent data have shown that the yeast Hir complex can mediate H3.3-specific deposition when ectopically expressed in human cells and this specificity is dependent on the HRD in Hpc2(Song et al., 2013). In human cells, the UBN1 HRD functions to specifically bind H3.3 over H3.1 and presumably other variants containing M90 (H3.2, H3.4), and bulky L90 (CENP-A). H3.X, H3.Y and H3.5 represent primate-specific H3 variants that also contain G90(Schenk et al., 2011; Wiedemann et al., 2010). Our data suggest that deposition of these H3 variants could also be coordinated by UBN1 (HIRA complex) or alternatively by DAXX.

We show that the structures of the UBN1 and DAXX interactions with H3.3 have surprising similarity, although the two chaperones use different mechanisms to mediate H3.3 G90 specificity. The structures are broadly similar in the sense that both chaperones bind proximal to H3.3 G90 of H3.3/H4 with nearly identical molecular contacts. However, unlike the UBN1-HRD, DAXX wraps around the entire H3/H4 heterodimer to make more extensive histone contacts that likely

contribute to its enhanced general H3/H4 binding affinity relative to UBN1-HRD. Accordingly, it is possible that other UBN1 regions and/or other proteins within the HIRA complex also contribute to general histone binding affinity(Kang et al., 2011; Lorain et al., 1998). In support of the participation of additional UBN1 regions in H3/H4 binding is our observation that UBN1(92-175) binds H3.3/H4 about 3 to 6-fold more strongly than UBN1(122-148) containing only the HRD (Supplementary Figs. 5b and 5c).

With respect to G90 specificity, the steric bulk provided by UBN1 residues 128-131 is perhaps more stringent than the water molecules that DAXX positions over G90. These water molecules can potentially be evacuated from the cavity to allow association with histone H3 variants with a bulkier amino acid at residue 90. This idea was demonstrated by a crystal structure of DAXX/H3.3(G90M)/H4(Elsasser et al., 2012), and supported by a recent report(Lacoste et al., 2014) that DAXX can mediate promiscuous deposition of the H3 variant CENP-A that contains an L in the corresponding position 90 of H3.

In addition, although we have demonstrated that UBN1 is primarily responsible for H3.3 binding specificity of the HIRA complex, the molecular basis for how the HIRA complex mediates deposition requires further studies. Our UBN1/H3.3/H4/Asf1 crystal structure may represent an intermediate in a potential hand off of histones between ASF1a/H3.3/H4 and HIRA/UBN1/CABIN1 complexes prior to deposition into chromatin, as several studies suggest that

deposition function of the HIRA complex is not dependent on the presence of ASF1a(Bonnefoy et al., 2007; Ray-Gallet et al., 2007; Ray-Gallet et al., 2011). Nonetheless, the studies presented here provide important new molecular insights into how the HIRA histone chaperone complex specifically recognizes the histone H3.3 variant to regulate diverse genomic activities.

2.10 – Materials and methods

Multiple sequence alignments

Multiple sequence alignments were generated using ClustalW2(Goujon et al., 2010; Larkin et al., 2007). Alignments were further formatted for publication using ESPript version 3.0 to better visualize sequence conservation(Gouet et al., 2003), with the exception of those in Figure 2.9a and supplementary Figure 2.7a which were manually formatted for publication.

Expression and purification of GST-tagged proteins

To generate the plasmid DNA construct encoding GST-UBN1(41-175), the cDNA encoding UBN1 residues 41-175 was PCR amplified from a previously discribed UBN1 construct (Banumathy et al., 2009) and ligated into the *Bam*HI/XhoI sites of pFastBacGST (Invitrogen). The plasmid DNA construct encoding HIRA(1-405) was generated though PCR amplification and ligation of the cDNA endoding HIRA(1-405) into the BamHI/XhoI sites of pFastBacHTB (Invitrogen), removal of the 6xHistidine tag was facilitated through mutagenesis using the Stratagene

QuikChange protocol(Weiner and Costa, 1994). GST-UBN1(41-175) and HIRA(1-405) were co-expressed in baculovirus infected Sf9 cells. GST-ASF1a FL, GST-UBN1(92-175) and GST-UBN1(122-148) were generated though PCR by amplification of the proper cDNA sequence and ligation into the BamHI/XhoI sites of a custom engineered pCDFduet-1 (Novagen) *Escherichia coli* protein expression vector carrying an N-terminal GST tag that is removable through TEV protease cleavage. GST-UBN1(92-175) mutants were generated with the Stratagene QuikChange protocol(Weiner and Costa, 1994). GST-tagged proteins produced in *Escherichia coli* were induced with 0.8 mM IPTG added to BL21-Gold(DE3) cells (Agilent) containing the desired expression plasmid and cells were grown overnight at 18°C. To purify GST-tagged proteins expressed in *E. coli* and Sf9, isolated cell pellets were suspended in 50 mL lysis buffer (phosphate buffered saline with 1 mM PMSF and 5 mM BME) per liter of cells and lysed with sonication. Lysate was clarified by centrifugation and the supernatant was incubated with glutathione agarose resin (Gold Biotechnology) for 1 hour prior to washing the resin with 60 column volumes of wash buffer (phosphate buffered saline with 5 mM BME). GST-tagged proteins were either eluted from the glutathione agarose with 5 column volumes of elution buffer (phosphate buffered saline with 5 mM BME and 20 mM reduced glutathione) or TEV protease was added to the protein-bound glutathione agarose and incubated overnight to cleave the target protein off the GST fusion. Eluted GST-UBN1(92-175), GST-UBN1(122-148), GST-ASF1a and GST-UBN1(41-

175)/HIRA(1-405) were then dialyzed overnight into buffer (20 mM Tris pH 8.0, 25 mM NaCl, 5 mM BME) for ion exchange purification using a HiTrap™ Q HP column (GE Healthcare). Protein released from GST by TEV cleavage was eluted with 6 column volumes of buffer (Hepes pH 7.5, 750 mM, 1 mM TCEP) and concentrated using a 3K cutoff centrifugal filter device (Millipore) prior to a final gel filtration purification step with a Superdex S75 10/300 GL column (GE Healthcare).

Expression and purification of recombinant H3/H4

Human histones H3.3 and H3.2 were amplified by PCR from a human cDNA library (ATCC 77430) and ligated into the *Bam*HI/XhoI sites of a custom engineered pETduet-1 (Novagen) *Escherichia coli* protein expression vector with an N-terminal 6xHistidine tag that is removable through TEV protease cleavage. A construct to express the amino acid sequence for H3.1 was generated by site-directed mutagenesis of H3.2. Human histone H4 has a sequence that carries several rare codons which are not conducive to expression in *Escherichia coli* and to overcome this issue we obtained a previously described codon-optimized construct encoding H4 from *X. laevis* which has the same amino acid sequence as human H4(Luger et al., 1997b). The codon-optimized *X. laevis* H4 cDNA was PCR amplified and ligated into the BamHI/XhoI sites of the same *Escherichia coli* protein expression vector as the H3 constructs. H3 and H4 proteins were expressed as inclusion bodies in BL21-Gold(DE3) cells (Agilent) induced with 0.8

mM IPTG, protein was expressed for 4 hours at 37°C. To purify H3 and H4 cells were lysed with sonication in 50 mL of lysis buffer (20 mM Tris pH 7.5, 500 mM NaCl, 5 mM BME) per liter of cells, lysate was clarified with centrifugation. Cell pellets containing insoluble H3 or H4 were then resuspended in denaturing buffer (20 mM Tris pH 7.5, 500 mM NaCl, 7.0 M GdHCl, 5 mM BME) and clarified with centrifugation. Denatured H3 or H4 were incubated with Ni-NTA agarose resin for 1 hour prior to washing the resin with 60 column volumes of denaturing buffer. Proteins were either eluted from the Ni-NTA agarose with 5 column volumes of elution buffer (300 mM Imidazole, 20 mM Tris pH 7.5, 500 mM NaCL, 7.0 M GdHCl, 5 mM BME). Eluted proteins were the dialyzed into water and lyophilized. Refolding was conduced through resuspending the lyophilized H3 and H4 in denaturing buffer and mixing the two histones at 1:1 prior to refolding by dialysis against two changes of non-denaturing buffer (20 mM Tris pH 7.5, 1.0 M NaCl, 5 mM BME, 1 mM EDTA) followed by dialysis into a low salt buffer (20 mM Tris pH 7.5, 100 mM NaCl, 5 mM BME, 1 mM EDTA) prior to ion exchange purification using a HiTrap™ SP HP column (GE Healthcare). The 6xHistidine tags on H3 and H4 were either left on the protein or removed by treatment with TEV protease during the refolding reaction. Following ion exchange the refolded H3/H4 complexes were further purified with gel filtration using a Superdex 200 10/300 GL column (GE Healthcare). H3.3 and H3.1 residue-swap point mutants were generated using the Stratagene QuikChange protocol(Weiner and Costa, 1994).

Synthetic petides

Large scale peptide synthesis of UBN1(122-148) and FITC-UBN1(127-148) was commissioned by a commercial vendor (Genscript).

GST pull-down histone binding assay

GST-tagged UBN1 proteins and H3/H4 complexes were prepared as described above. The pull-down experiments represented in Figure 2.1 and Figure 2.2 were conducted by incubating 2 μM GST-tagged protein with 2 μM H3/H4 (or 2 μM H3/H4 in competition with 2 μM His-H3/His-H4) in a buffer containing 20 mM Hepes pH 7.5, 500 mM NaCl, and 1 mM TCEP at 4°C for 30 minutes. Proteins were then subjected to pull-down by incubation with glutathione agarose resin (Gold Biotechnology) for 30 minutes. Resin was washed with 120 column volumes of buffer (20 mM Hepes pH 7.5, 500 mM NaCl, and 1 mM TCEP) prior to elution of bound proteins by boiling resin in SDS gel-loading buffer. All other GST pull-down experiments were conducted by incubating 2 μM GST-tagged protein with 4 μM H3/H4 (or 4 μM H3/H4 in competition with 4 μM His-H3/His-H4) in a buffer containing 20 mM Hepes pH 7.5, 750 mM NaCl, and 1 mM TCEP at 4°C for 30 minutes. Proteins were then subjected to pull-down by incubation with glutathione agarose resin (Gold Biotechnology) for 30 minutes. Resin was washed with 120 column volumes of buffer (20 mM Hepes pH 7.5, 750 mM NaCl, and 1 mM TCEP) prior to elution of bound proteins by boiling resin in SDS gel-loading buffer. Results of the pull-down assays were analyzed though

visualization of input and pull-down samples with SDS-polyacrylamide gel electrophoresis. Gels were stained using Coomassie Brilliant Blue G-250.

Isothermal titration calorimetry

Quantitative analysis of the UBN1/histone interaction was conducted with a VP-ITC microcalorimeter (MicroCal). All proteins/peptides were prepared in a buffer with 20 mM Hepes pH 7.5, 750 mM NaCl and 1 mM TCEP prior to ITC analysis. Experiments for Figures 2.1c and 2.7c were conduced by injecting 0.450 mM UBN1 into a calorimetry cell containing 30 μM H3/H4, where UBN1 was injected in 7 μL increments every 3 minutes at 22°C and ITC reference power set to 30. Experiments for supplementary figure 5c were conducted with the same cell and syringe concentrations but UBN1 was injected in 10 uL increments every 3 minutes at 22°C and ITC reference power was set to 20. Data were analyzed using Origin version 7.0 and corrected for heat of mixing and dilution.

Fluorescence polarization histone-binding assay

Quantitative analysis of the UBN1/histone interaction was conducted with a FP-based competition-binding assay; all data were collected using a PerkinElmer EnVision Xcite Multilabel plate reader. All proteins and peptides for this assay were prepared in a buffer containing 20 mM Hepes pH 7.5, 750 mM NaCl and 1mM TCEP. We commissioned commercial synthesis of a UBN1(127-148) peptide with an N-terminal FITC moiety (Genscript). The K_d of FITC-UBN1(127-

148) for binding to H3.3/H4 (K_d = 4.3 ± 0.2 μM) and H3.1/H4 (K_d = 17.7 ± 2.3 μM) was measured by monitoring the change in fluorescence polarization when increasing H3/H4 is titrated into a constant concentration of FITC-UBN1(127-148) (10 nM) (Figure 2.5a). The curves were fit using a one site – total binding model with GraphPad Prism (version 5.0a) (Y=Bmax*X/(Kd+X) + NS*X + Background). We found that the FITC moiety or linker used to attach it to UBN1 artificially enhances the UBN1/histone interaction in comparison with both UBN1(92-175) and UBN1(122-148) (Figure 2.5a). Despite this, we were still able to use FITC-UBN1(127-148) for a competition-binding assay. To obtain a suitable fluorescence polarization range in the competition binding assay, the concentration of H3/H4 was held at or slightly above the relative dissociation constant for FITC-UBN1(127-148). Increasing concentration of UBN1 competitor was titrated into solutions containing constant concentrations of FITC-UBN1(127-148) (10 nM) and either H3.3/H4 (6.5 μM) or H3.1/H4 (25 μM) to estimate IC_{50} values by competition binding. UBN1 wild-type and all 27 point mutants were also titrated into a solution containing only the FITC-UBN1(127-148) to generate curves used to subtract background polarization, which is possibly caused by UBN1 self-association. Curves for D136A, S137A, F138A, I139A, and D140A had a slightly different scale in comparison with others and were normalized to the average baseline; normalization did not alter data interpretation. Competition binding curves were fit using the one site – total binding model with GraphPad Prism (version 5.0a).

Recombinant H3/H4 are known to form a strong $(H3/H4)_2$ homotetramer at concentrations above 1 μM(Winkler et al., 2012). Due to this H3/H4 conformation, two molecules of UBN1 are necessary to completely compete both FITC-UBN1 molecules away from the histone tetramer. Taking the binding stoichiometry of $(UBN1)_2/(H3/H4)_2$ into account, the concentration of UBN1 reported in the fluorescence polarization titration curves actually represents half of the molecular concentration of UBN1.

Crystallization and structure determination

We obtained as a generous gift from Mair Churchill (University of Colorado School of Medicine) the GST-Asf1/H3.1/H4 polycistronic *Escherichia coli* co-expression construct that was previously used to produce Asf1/H3.1/H4 for crystallization and structure determination(English et al., 2006). We used mutagenesis(Weiner and Costa, 1994) to generate a construct encoding Asf1/H3.3/H4 for our purposes. The GST-Asf1/H3/H4 complex was expressed in BL21-Gold(DE3) cells (Agilent) induced with 0.8 mM IPTG, protein was expressed for 8 hours at 28 °C. Cell pellets were resuspended in 50 mL of lysis buffer (20 mM Tris pH 8.0, 500 mM NaCl, 1 mM PMSF, 5 mM BME) per liter of cells and lysed with sonication. Lysate was clarified by centrifugation and the supernatant was incubated with glutathione agarose resin (Gold Biotechnology) for 1 hour prior to washing the resin with 60 column volumes of wash buffer (20 mM Tris pH 8.0, 500 mM NaCl, 5 mM BME). The Asf1/H3/H4 complex was then

released from the resin with on-column cleavage using prescission protease. The cleaved Asf1/H3/H4 complex was eluted with 6 column volumes of elution buffer (20 mM Tris pH 8.0, 500 mM NaCl, 1 mM TCEP). The free Asf1/H3/H4 complex was then concentrated using a 3K cutoff centrifugal filter device (Millipore) prior to a final gel filtration purification step with a HiLoad 16/600 Superdex S75 pg column (GE Healthcare). The UBN1(122-148)/Asf1/H3.3/H4 complex was formed by incubating 0.2 mM Asf1/H3.3/H4 with 0.8 mM UBN1(122-148) in a buffer containing 20 mM Tris pH 8.0, 500 mM NaCl and 1 mM TCEP prior to crystallization. Hanging drops for vapor diffusion were formed by mixing two volumes of protein solution with one volume of crystallization solution (0.1 M Sodium Cacodylate pH 6.0, 8% PEG 8K, 0.2 M NaCl) and crystals appeared within 48 hours. X-ray diffraction data were collected using an in-house Rigaku MicroMax-007HF microfocus rotating anode x-ray generator and a Saturn 944+ CCD detector. The diffraction data were indexed and scaled using HLK3000, and a Molecular Replacement solution was obtained using Phaser in the Phenix software suite with Asf1/H3.1/H4 (PDB ID, 2HUE) used as a search model. The solution was refined using phenix.refine in the Phenix software suite and model building was carried out using the molecular-graphics program Coot. The solution model contains UBN1(122-142), H3.3(61-134), H4(24-102), and Asf1(1-154). Electron density was not observed for Asf1 residues 155-169, H3.3 residue 60, and H4 residues 19-23. Table 2.1 shows a summary of the refinement statistics. Simulated annealing F_o - F_c omit map for the UBN1 peptide was generated using

phenix.refine in the Phenix software suite. Figures for publication were generated using the program PyMOL. Structure comparison of UBN1 and DAXX were carried out using the published crystal structures of DAXX/H3.3/H4 (PDB ID, 4H9N) and DAXX/H3.3(G90M)/H4 (PDB ID, 4H9O)(Elsasser et al., 2012).

Assembly of Asf1/H3(FL)/H4(FL) complexes

As described previously(English et al., 2005), expression and purification of the heterotrimeric Asf1(1-169)/H3(60-135)/H4(20-102) complex from *Escherichia coli* yields a large excess of Asf1 not in complex with H3/H4 that can be separated from the Asf1/H3/H4 complex on a HiLoad 16/600 Superdex S75 pg column (GE Healthcare). We collected this excess Asf1 to be used for complex formation with either FL H3.3/H4 or FL H3.1/H4. Complexes were formed through mixing a 2-fold excess of Asf1 with H3.3/H4 or H3.1/H4 in a buffer containing 20 mM Hepes pH 7.5, 2.0 M NaCl, and 1 mM TCEP, this mixture was then dialyzed overnight into a buffer containing 20 mM Hepes pH 7.5, 750 mM NaCl, and 1 mM TCEP. The mixture was then injected onto a HiLoad 16/600 Superdex S75 pg column (GE Healthcare) to separate the Asf1/H3/H4 complex from the excess Asf1. The resulting Asf1/H3/H4 was used for ITC studies.

pFastBacDual HIRA/UBN1

A C-terminally FLAG tagged open reading frame of full-length human HIRA encoded by nucleotides 221 to 3271 of NM_003325 were PCR amplified and

subcloned into the XmaI and SphI restriction sites 3' to the P10 promoter in pFastBacDual. Nucleotides 843 to 4247 of NM_016936 encoding full-length human UBN1 were PCR amplified and subcloned into the SalI and XbaI restriction sites 3' to the PH promoter in pFastBacDual. The FID(138-140)AAA mutation was introduced into UBN1 in the pFastBacDual vector by the Stratagene QuikChange protocol(Weiner and Costa, 1994).

Expression of recombinant proteins in insect cells

Sequence-confirmed pFastBac transfer vectors containing sequences for human HIRA, UBN1 and CABIN1 were transformed into DH10Bac cells. Proper recombination of the HIRA, UBN1 and CABIN1 sequences into the baculovirus genome was determined by PCR and positive bacmid DNAs were transfected into Sf9 cells(Gillette et al., 2011). High-titer passage 1 (P1) virus stocks were recovered 120 hours post-transfection. A high-titer P2 virus stock was generated by infecting Sf9 at an MOI (Multiplicity Of Infection) of ~0.1, followed by incubation for 120 hours. For productions, 1 x 10^6 Sf9 cells/ml in Sf900-III medium (Invitrogen) were infected with virus at an MOI of 1. Infected cells were harvested 48 hours post-infection.

Histone pull-downs with the intact HIRA/UBN1/CABIN1 complex

The HIRA-FLAG/UBN1 and HIRA-FLAG/UBN1/His-CABIN1 complexes containing wild-type UBN1 or UBN1(FID138-140AAA) mutant was purified by

suspending isolated cell pellets in 50 mL lysis buffer (20 mM Tris pH 8.0, 500 mM NaCl, 1 mM PMSF, 1 mM BME and a complete EDTA-free protease inhibitor tablet (Roche)) per liter of cells, and lyses with sonication. Lysate was clarified by centrifugation and the supernatant was incubated with ANTI-FLAG M2 affinity gel (Sigma-Aldrich) for 2 hours prior to washing the gel with 60 column volumes of wash buffer (20 mM Tris pH 8.0, 500 mM NaCl, 1 mM BME). The FLAG gel was bound to capacity for both the wild-type and mutant complexes. To conduct the FLAG pull-down assay, 50 □L of bound gel was incubated for 30 minutes with 4 μM H3/H4 in a buffer with 20 mM Tris pH 8.0, 750 mM NaCl and 1 mM BME. Then the gel was washed with 120 column volumes of buffer (20 mM Tris pH 8.0, 750 mM NaCl, and 1 mM BME) prior to elution of bound proteins by boiling the gel in SDS gel-loading buffer. Results of the pull-down assays were analyzed though visualization of input and pull-down samples with SDS-polyacrylamide gel electrophoresis. Gels were stained using Coomassie Brilliant Blue G-250.

2.11 – Acknowledgements

We would like to thank John Domsic, Adam Olia, and Jasna Maksimoska for helpful discussions regarding the crystal structure determination and refinement. The construct for expression of the Asf1/H3.1/H4 complex was a gift from Mair Churchill (University of Colorado School of Medicine), and the construct for expression of H4 was a gift from Karolin Luger (Colorado State University). This work was supported by the National Institutes of Health grant AG031862 to D.S.,

P.D.A and R.M and an American Heart Association predoctoral fellowship 12PRE12030157 to M.D.R. We acknowledge the use of the Wistar Protein Expression facility, which is supported in part by NIH grant CA010815. Coordinates and structure factors for the Asf1/H3.3/H4/UBN1 complex have been deposited in the Protein Data Bank under accession code 4ZBJ.

CHAPTER 3: The HIRA histone chaperone complex subunit UBN1 harbors H3/H4 and DNA binding activity

3.1 – Overview

The deposition of histones onto DNA for nucleosome formation and chromatin maintenance is a hallmark of eukaryotic cells and serves to package and protect DNA for the efficient regulation of gene expression. Histone deposition is coordinated by histone chaperone proteins, which serve to bind and protect the basic histone proteins from non-specific interactions prior to deposition into nucleosomes (Burgess and Zhang, 2013; Laskey et al., 1978). Additionally, histones have many variant sequences that are incorporated into nucleosomes at specific genomic locations to effect changes on the local chromatin environment, and many distinct histone chaperones have specifically evolved to bind and deposit a single variant histone into chromatin (Mattiroli et al., 2015; Maze et al., 2014). The human HIRA histone chaperone complex specifically binds to and deposits H3.3/H4 into chromatin in a replication-independent manner (Tagami et al., 2004). The complex is composed of the three core subunits, HIRA, UBN1, and CABIN1, with the transient addition of ASF1a, which serves to bring an H3.3/H4 dimer to the complex but is not required for ultimate H3.3/H4 deposition into nucleosomes (Bonnefoy et al., 2007; Ray-Gallet et al., 2007; Tagami et al., 2004). Much work has been done to elucidate how the HIRA complex assembles (Banumathy et al., 2009; Rai et al., 2011; Tang et al., 2006; Tang et al., 2012), and we have recently reported that UBN1 specifically binds an H3.3/H4 dimer bound to ASF1a (Ricketts et al., 2015); but the mechanism of how the HIRA

complex ejects ASF1a and transitions the H3/H4 dimer to an $(H3.3/H4)_2$ tetramer for deposition is not clear. The study of $(H3/H4)_2$ tetramer deposition has been the focus of work for several other histone chaperones system (Chen et al., 2015; Huang et al., 2015; Mattiroli et al., 2017; Richet et al., 2015; Sauer et al., 2017; Wang et al., 2015; Zasadzinska et al., 2013) since the discovery that the vast majority of H3/H4 is deposited onto DNA in the tetramer form for eventual nucleosome formation upon the addition of two H2A/H2B dimers (Xu et al., 2010).

The HIRA complex subunit UBN1 has been reported to bind to the WD40 domain of HIRA though a small domain of about 30 amino acids called the NHRD (Tang et al., 2012), named after discovery of it's HIRA binding activity because it lies just N-terminal to a strongly conserved UBN1 domain called the Hpc2-related domain (HRD), named for it's conservation with Hpc2, the S. cerevisiae homolog of UBN1. After identification of the NHRD as a HIRA-binding domain we were able to show that the HRD specifically binds to H3.3/H4 and determined a crystal structure of the UBN1/H3.3/H4/Asf1 complex (Ricketts et al., 2015). While, the crystal structure is informative about the exact nature of the H3.3-specific interactions that are mediated by UBN1, it only contains 20 amino acids from UBN1. So with roughly 50 amino acids from UBN1 being well characterized for function out of 1134 residues, we decided to attempt to characterize additional

regions of UBN1 to identify previously unknown functions to potentially learn more about how the HIRA complex binds and deposits H3.3/H4 into chromatin.

Here we characterize two UBN1 domains of previously unknown function. Through secondary structure prediction and sequence conservation we have identified a roughly centrally located region of UBN1, which we call the UBN1 middle domain, spanning residues ~300-600, predicted to have helical secondary structure. Though size-exclusion chromatography and analytical ultracentrifugation experiments we have shown that the C-terminal region of the UBN1 middle domain, residues 504-584, are responsible for dimer formation. Through pull-down and deletion analysis we have also shown that the N-terminal region of the UBN1 middle domain, residues 296-341, is required for H3/H4 binding. Additionally we have identified several vertebrate-specific conserved lysine residues in a very basic loop region of low predicted structure residing between the UBN1 HRD and middle domain. Using fluorescence polarization and size-exclusion chromatography analysis of UBN1/DNA complex formation we have determined that these conserved lysine residues are essential for non-specific DNA binding by UBN1. Together, our observations of H3/H4 binding and dimer formation by the UBN1 middle domain indicate that UBN1 may be capable of binding two H3.3/H4 molecules to induce $(H3.3/H4)_2$ tetramer formation in a pathway to nucleosome deposition that may also involve UBN1-mediated DNA binding by the HIRA complex.

3.2 – The UBN1 middle domain forms stable monomer and dimer populations

To identify regions of UBN1 that contribute to the biological function of the human HIRA histone chaperone complex, beyond the well characterized HRD and NHRD domains, we performed a multiple sequence alignment of metazoan UBN1 homologs to hone in on additional regions of conservation to probe for functional significance (Figure 3.1). Hpc2, the UBN1 homolog from S. cerevisiae, only shares conservation in the HRD and NHRD domains, to identify nuanced features of human UBN1 outside of these domains we omitted UBN1 homologs from yeast from our alignment. Through alignment of Homo sapiens UBN1, Mus musculus UBN1, Gallus gallus UBN1, Scleropages formosus UBN1, and Drosophila melanogaster Yemanuclein it is evident that the NHRD (residues 41-77) and HRD (119-175) are very highly conserved, but there are additional regions of conservation, most notably the region spanning residues ~300-600 (Figure 3.1).

To compliment the multiple sequence alignment, we also performed secondary structure prediction for the human UBN1 protein sequence using the PSIPRED server (Buchan et al., 2013) (Figure 3.2). While the majority of human UBN1 has little predicted structure or low confidence prediction, UBN1 residues ~300-600 are predicted to form helical secondary structure with high confidence of prediction. This, in combination with the metazoan sequence conservation,

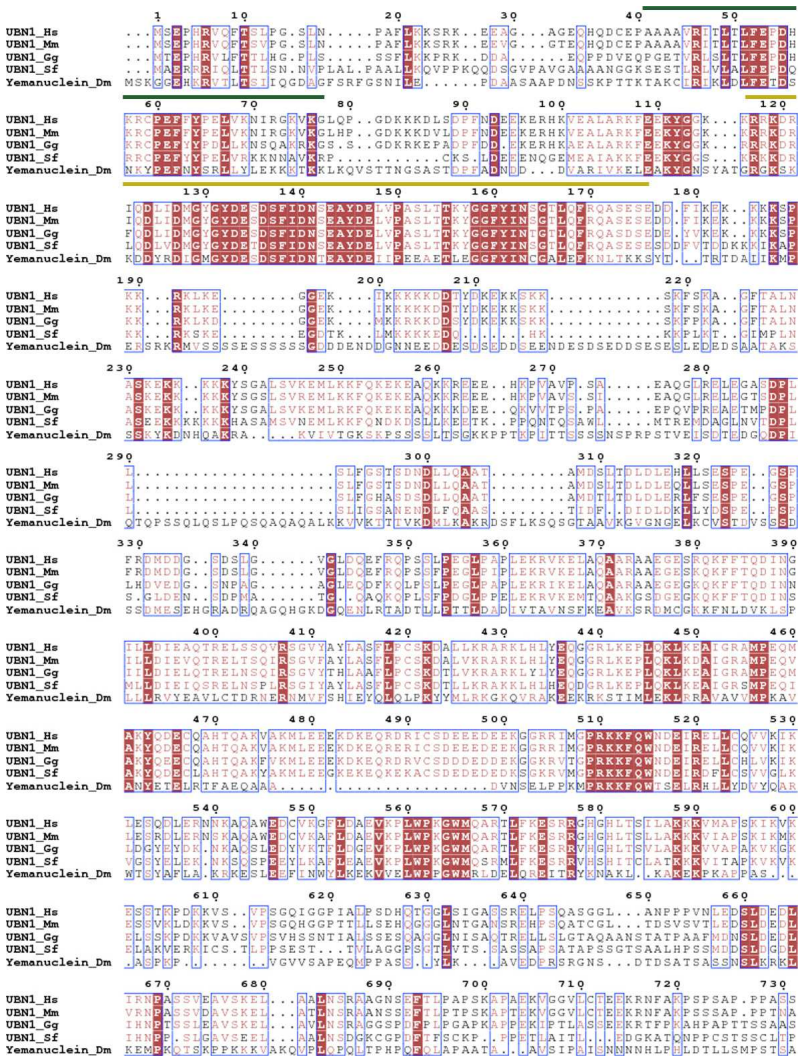


Figure 3.1 – Alignment of UBN1 metazoan homologs and secondary structure prediction. (A) Alignment depicts sequence conservation among *Homo sapiens* UBN1, *Mus musculus* UBN1, *Gallus gallus* UBN1, *Scleropages formosus* UBN1, and *Drosophila melanogaster* Yemanuclein. The conserved UBN1 HRD (residues 120-175) (yellow line over sequence) and NHRD (residues 41-71) (green line over sequence) domains have been well characterized for their functions. UBN1 residues 176-700 are well conserved among metazoan species, although the biological function of this UBN1 region remains unclear. Note that *Scleropages formosus* represents UBN1 from a fish (Asian bonytongue), as there is not an annotated full-length sequence for UBN1 from *Danio rerio*.

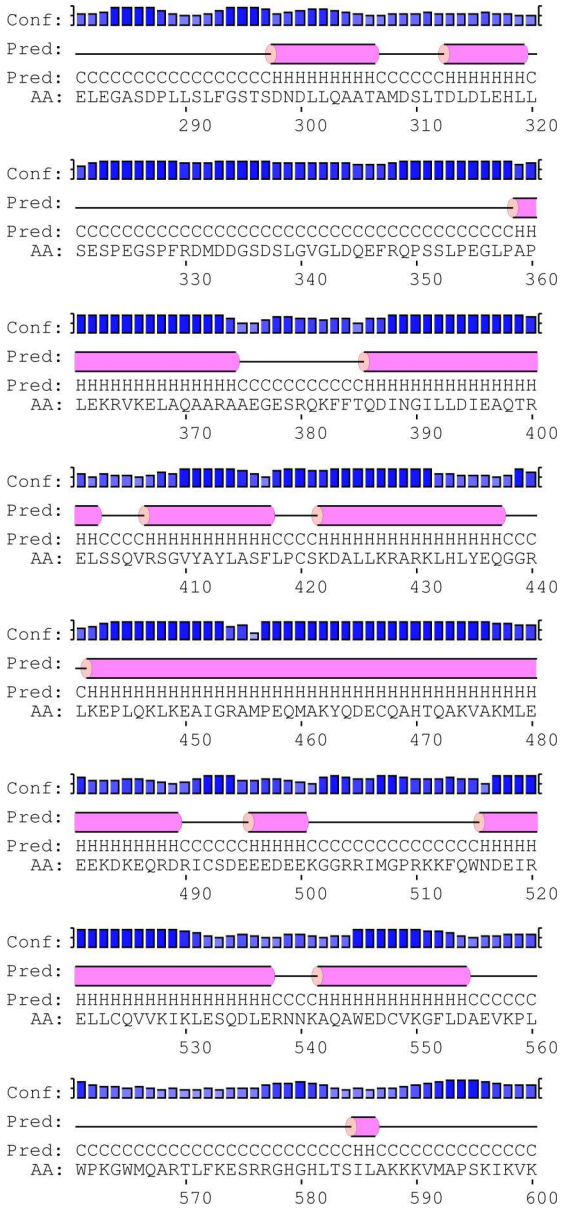


Figure 3.2 – Secondary structure prediction for UBN1 residues 291-600. Secondary structure for UBN1(291-600) was predicted based on the amino acid sequence by the PSIPRED protein sequence analysis server.

identifies this domain of UBN1 as a good target for functional characterization with biochemical and structural studies. As this domain has been uncharacterized for function in human UBN1, we preliminarily name this region the UBN1 middle domain.

A construct encoding UBN1 middle domain residues 296-584, bearing an N-terminal 6xHis that is removable with TEV protease, was generated and expressed in E. coli cells (Figure 3.3a). This UBN1(296-584) fragment was easily purified and stable in solution. Upon size-exclusion chromatography analysis using an s200 16/600 column (GE Healthcare), we observed this fragment of UBN1 forms two monodisperse populations that we call “peak 1” and “peak 2” (Figure 3.3b). Upon pooling and re-injecting the isolated populations back onto the size-exclusion column, it was observed that these populations are relatively stable and do not readily re-distribute back into the same two populations (Figure 3.3b-c). To further analyze these stable UBN1(296-584) peaks, we then injected the isolated populations onto an analytical 10/300 s200 column (GE Healthcare) and compared the elution profiles to a set of molecular-weight standards (Biorad) (Figure 3.4a-b). Comparison with the standards indicates that peak 1 has an elution profile similar to that of the 158 kDa standard and peak 2 elutes just before the 44 kDa standard. To further verify our observations that these populations are non-exchanging, we performed a pull-down experiment where GST-UBN1(296-584) was incubated with untagged UBN1(296-584) from the

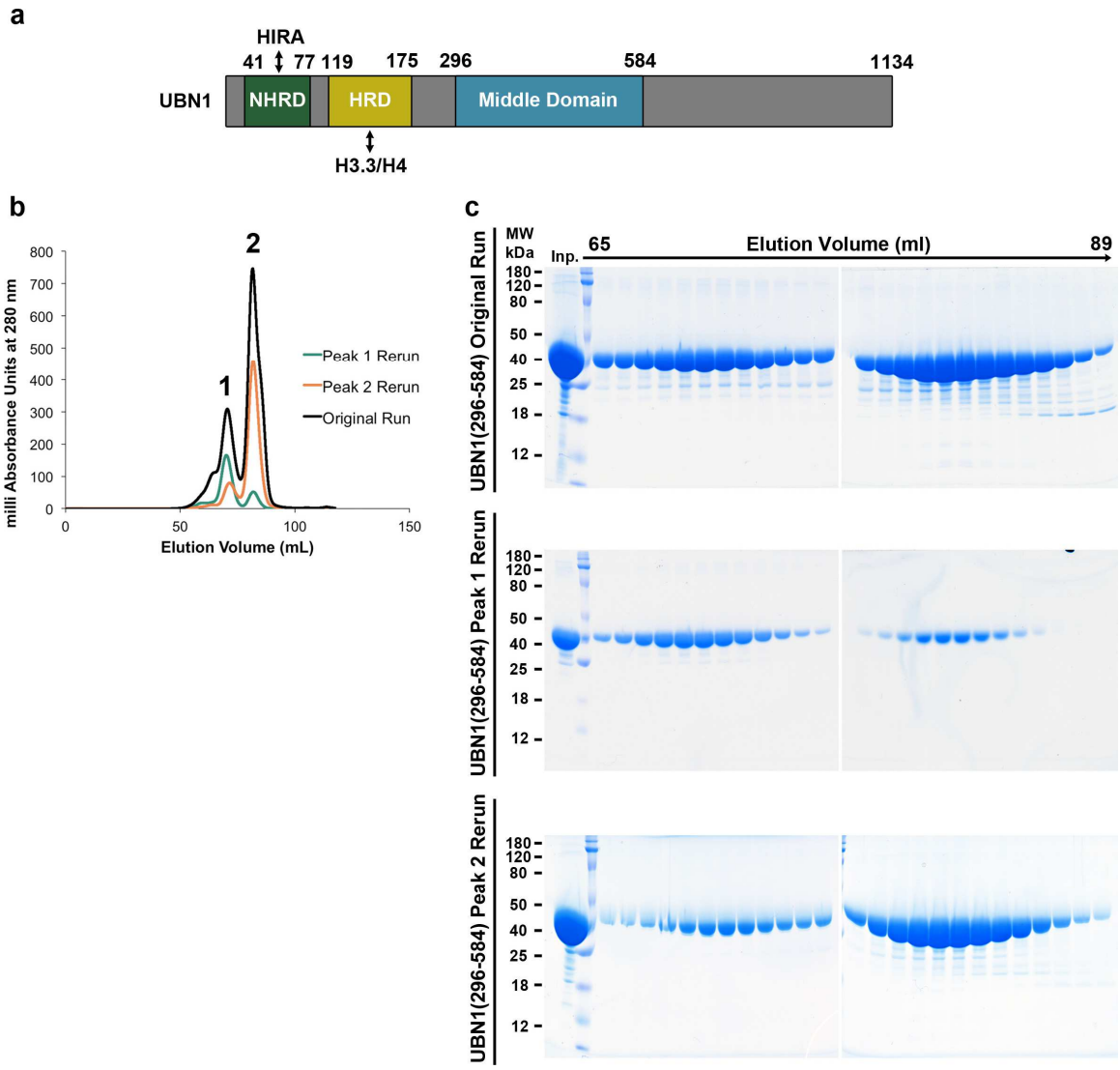


Figure 3.3 – Analysis of UBN1(296-584) with size-exclusion chromatography. (A) Domain architecture diagram depicting the characterized UBN1 NHRD, HRD and Middle domains. (B) Chromatograph of UBN1(296-584) resolved on a HiLoad s200 16/600 size-exclusion column. Fractions from 65-75 mL were pooled and re-loaded onto the column for "Peak 1 Rerun", while fractions from 76-89 mL were pooled and re-loaded onto the column for "Peak 2 Rerun". (C) SDS-PAGE gel analysis of 1 mL fractions spanning 65-89 mL for all three runs.

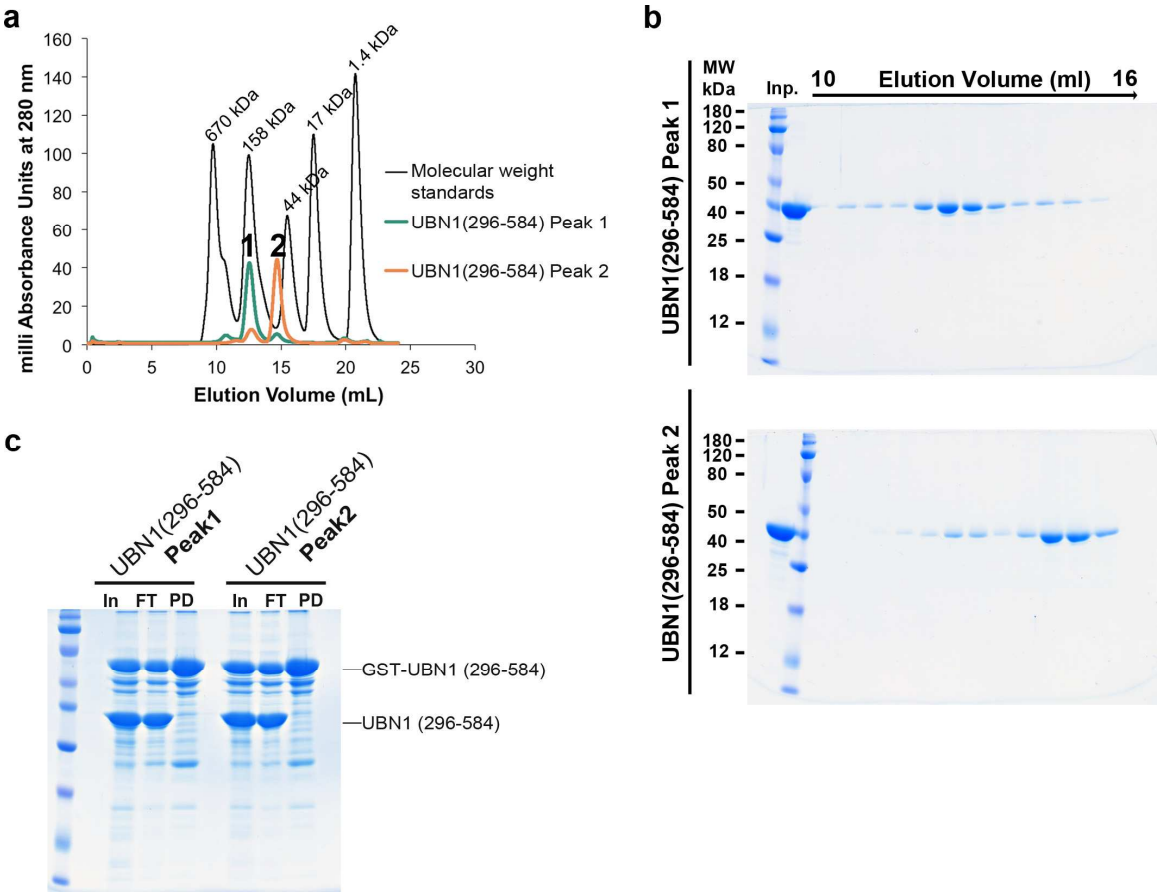


Figure 3.4 – UBN1(296-584) forms two non-exchanging monodisperse populations. (A) Chromatographs from UBN1(296-584) peak 1 and peak 2 populations resolved on an analytical s200 10/300 size-exclusion column in comparison with molecular-weight standards. (B) SDS-PAGE gel analysis of 0.5 mL fractions spanning 10-16 mL for both runs. (C) GST pull-down confirming that both peak 1 and peak 2 are unable to form a complex with GST-UBN1(296-584), further indicating that these are stable and non-exchanging populations.

peak 1 or peak 2 populations (Figure 3.4c). We observed that both peak 1 and peak 2 are unable to form complexes with GST-UBN1(296-584), indicating that these populations do not exchange after they are initially formed.

A theoretical monomer of UBN1(296-584) has a molecular weight of 33 kDa. To determine if these two UBN1 populations represent different oligomerization states we preformed a sedimentation equilibrium analytical-ultracentrifugation analysis of the two populations. Both peaks were monitored for the equilibrium distribution of three different protein concentrations (0.3, 0.5, 0.7 OD A280) at three centrifugation speeds (12000, 18000, 26000 RPM). The combined equilibrium data were globally fit with an ideal model to determine the average molecular weight of the protein in solution. Analysis of peak 1 determined an experimental molecular weight of 56.7 kDa (Figure 3.5), while analysis of peak 2 determined an experimental molecular weight of 34.6 kDa (Figure 3.6). Comparison with the molecular weight of a theoretical UBN1(296-584) monomer (33 kDa) and dimer (66 kDa) indicates that peak 1 forms a majority dimer population while peak 2 forms a majority monomer population. For better visualization of peak 1 and peak 2 oligomerization states, representative curves from each population were plotted on a linear scale and compared with predicted curves for a theoretical UBN1(296-584) monomer, dimer, and trimer. In comparison with the predicted curves it appears that peak 2 aligns very well with

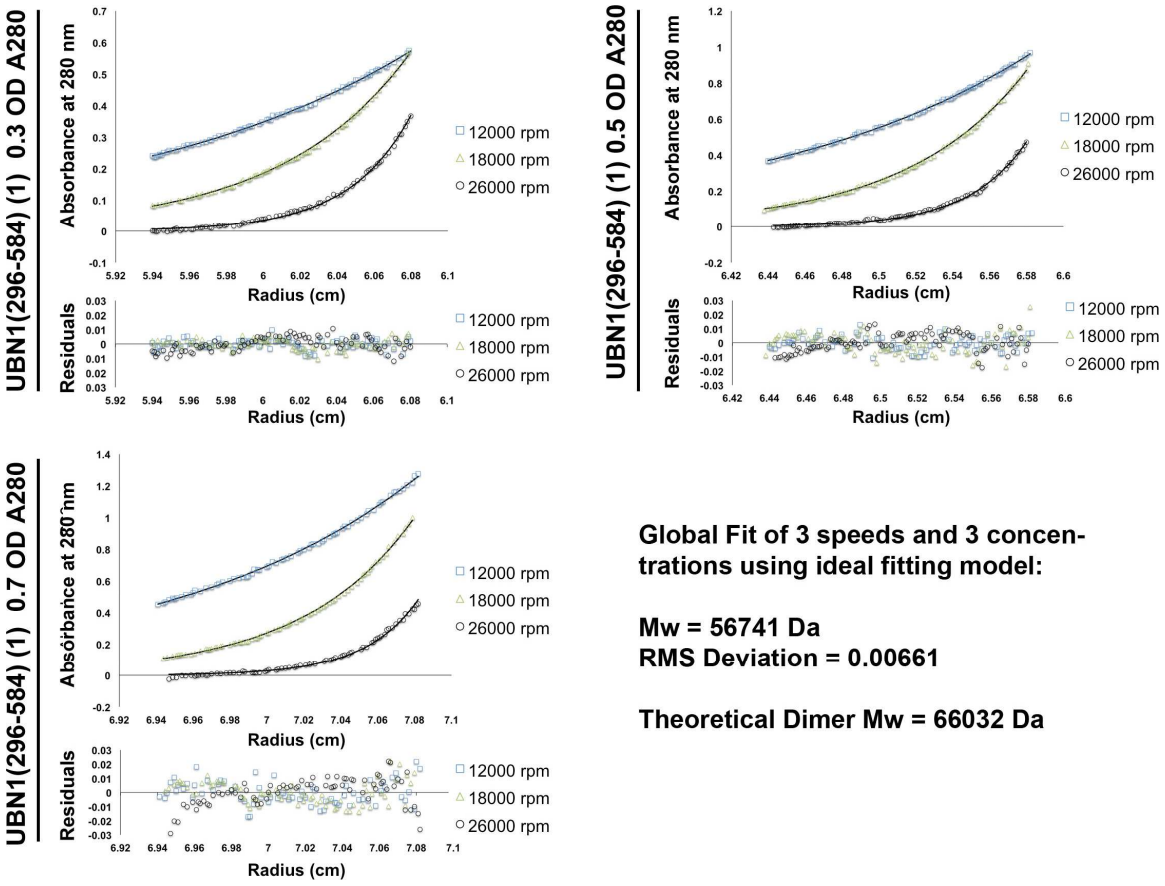


Figure 3.5 – Analytical ultracentrifugation of UBN1(296-584) peak 1 dimer.
Sedimentation equilibrium of UBN1(296-584) peak 1 population, three different protein concentrations (0.3, 0.5, 0.7 OD A280) were analyzed for equilibrium distribution at three different speeds (12000, 18000, and 26000 RPM). These data were fit to the ideal fitting model using the program HeteroAnalysis. The ideal fit yielded an experimental molecular weight of 56741 Da, indicating this is a majority dimer population.

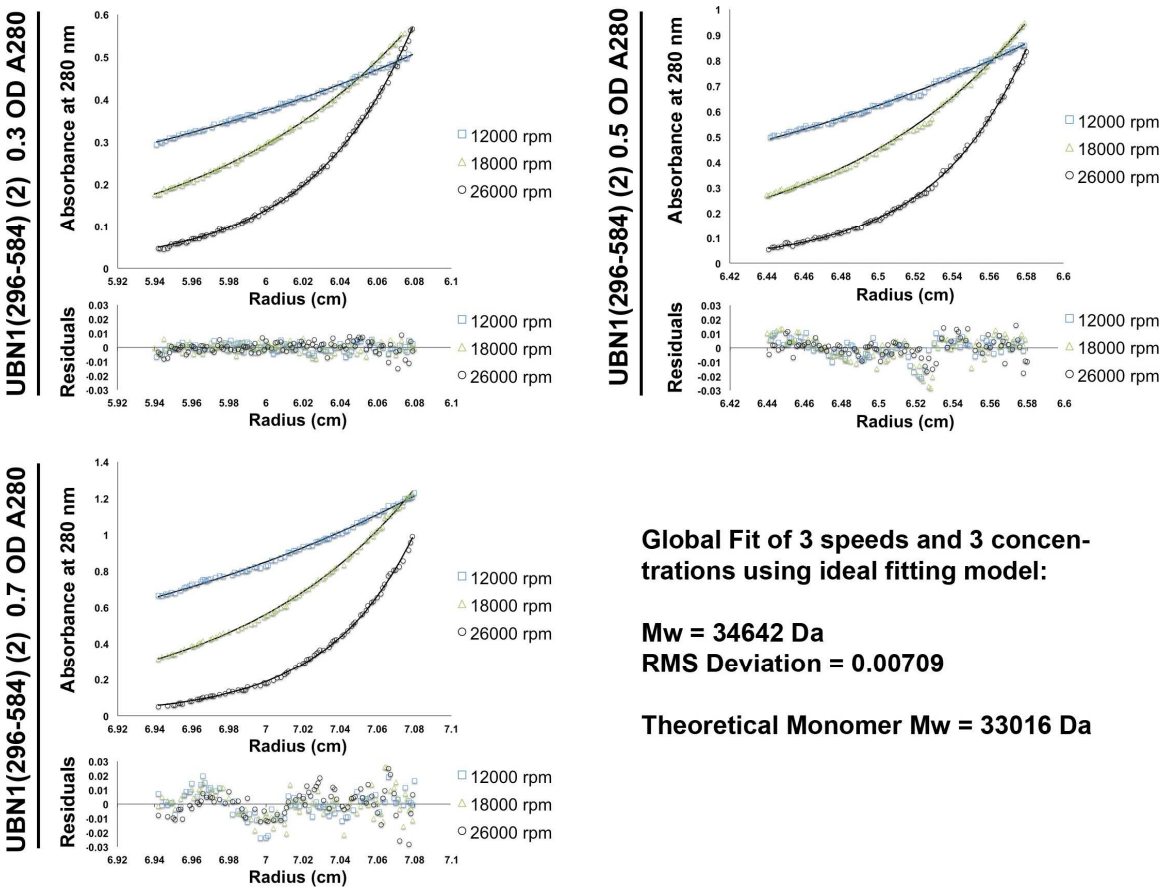


Figure 3.6 – Analytical ultracentrifugation of UBN1(296-584) peak 2 monomer. Sedimentation equilibrium of the UBN1(296-584) peak 2 population, three different protein concentrations (0.3, 0.5, 0.7 OD A280) were analyzed for equilibrium distribution at three different speeds (12000, 18000, and 26000 RPM). These data were fit to the ideal fitting model using the program HeteroAnalysis. The ideal fit yielded an experimental molecular-weight of 34642 Da, indicating this is a majority monomer population.

the predicted monomer, and peak 1 aligns best with the predicted dimer (Figure 3.7).

We then expressed and purified UBN1(341-584) in the same manner as UBN1(296-584) to monitor the effect of truncating the predicted N-terminal helices and loop (Figure 3.8a). UBN1(341-584) resolves into the same peak 1 and peak 2 populations observed for UBN1(296-584) on an s200 16/600 column, the only major difference for this fragment was that we observed the formation of a specific degradation product that runs between the 18 kDa and 12 kDa molecular-weight standards on an SDS-PAGE gel (Figure 3.8b-c). We decided to preform LC-MS analysis on the degradation product and determined that the degradation product represented UBN1 residues 341-503.

We then prepared UBN1(296-503) in the same manner as UBN1 (296-584) and UBN1(341-584) to identify the contribution of UBN1 residues 504-584 to the formation of the monomer and dimer populations (Figure 3.9a). When UBN1(296-503) was resolved on an s200 16/600 column, only one major peak with a small shoulder was observed (Figure 3.9b-c). We did not see the same distinctly separate populations observed for UBN1(296-584) and UBN1(341-584), but we decided to treat the shoulder as a potentially separate population. We pooled and re-injected the shoulder and the main peak as peak 1 and peak 2, respectively. We observed that both rerun of peak 1 and peak 2 resulted in formation of the same distribution as the original run, indicating that the peak

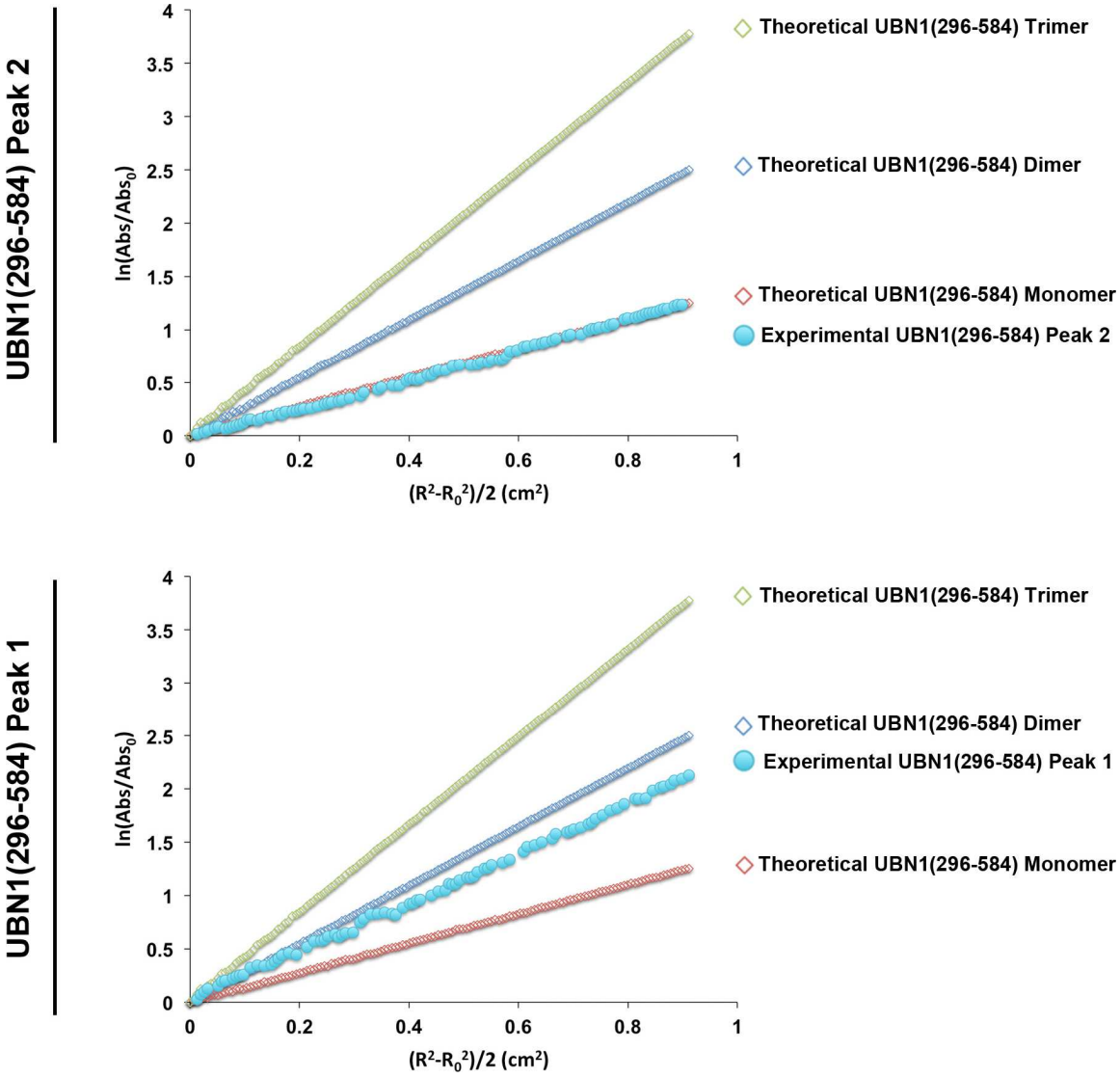


Figure 3.7 – Linear plot of UBN1(296-584) peak 1 / peak 2 sedimentation equilibrium data. Representative data for UBN1 peak 1 (bottom) and peak 2 (top) at 0.5 OD and 18000 RPM were compared to predicted curves for a theoretical UBN1(296-584) monomer, dimer, and trimer. Theoretical curves were generated using the program HeteroAnalysis. To represent the data on a linear scale the curves were plotted using a logarithmic normalization.

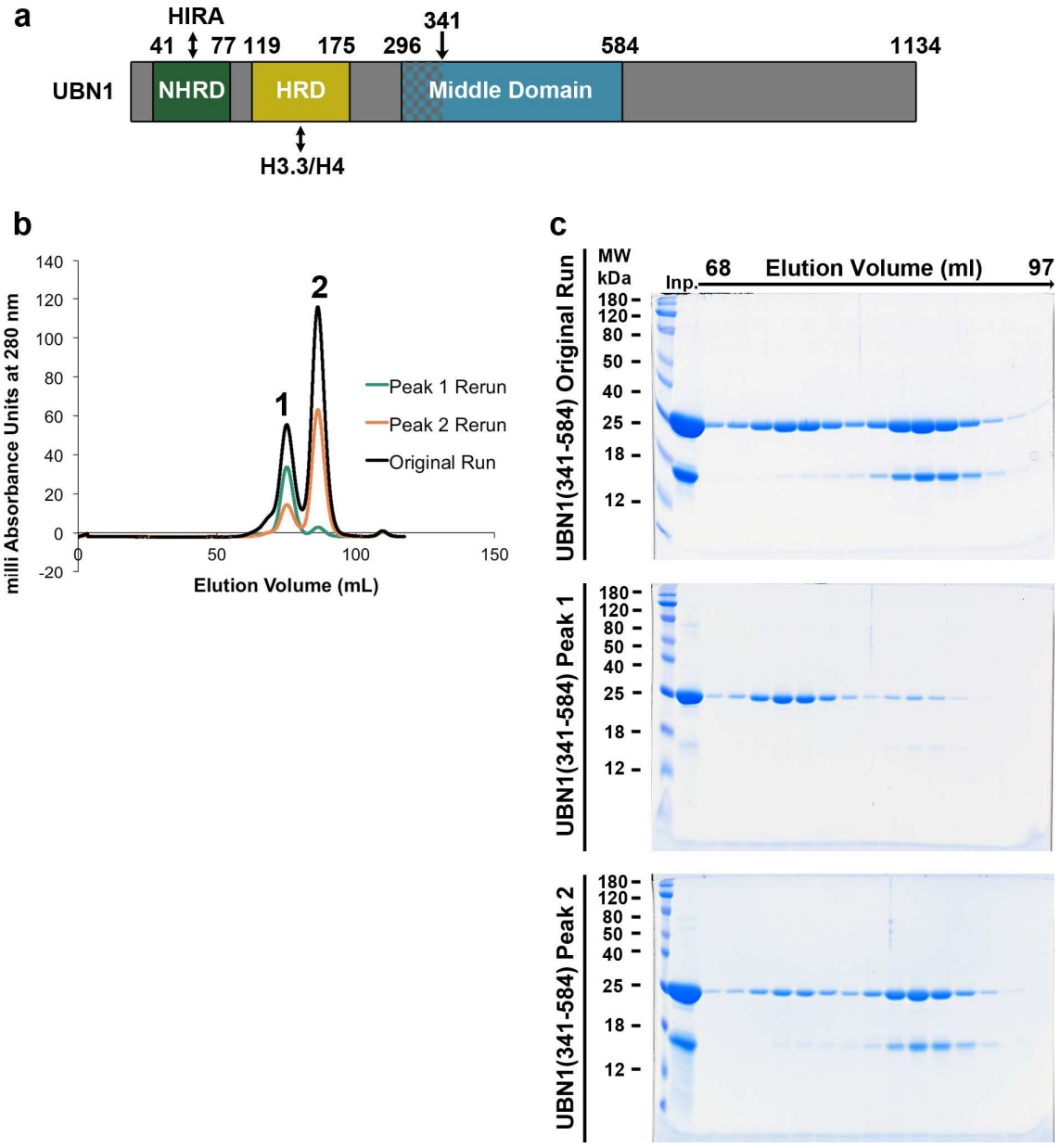


Figure 3.8 – Analysis of UBN1(341-584) with size-exclusion chromatography. (A) Domain architecture diagram for UBN1 depicting middle domain fragments (B). Chromatograph of UBN1(341-584) resolved on a HiLoad s200 16/600 size-exclusion column. Fractions from 68-78 mL were pooled and re-loaded onto the column for "Peak 1 Rerun", while fractions from 80-94 mL were pooled and re-loaded onto the column for "Peak 2 Rerun". (C) SDS-PAGE gel analysis of every other 1 mL fraction spanning 68-97 mL for all three runs.

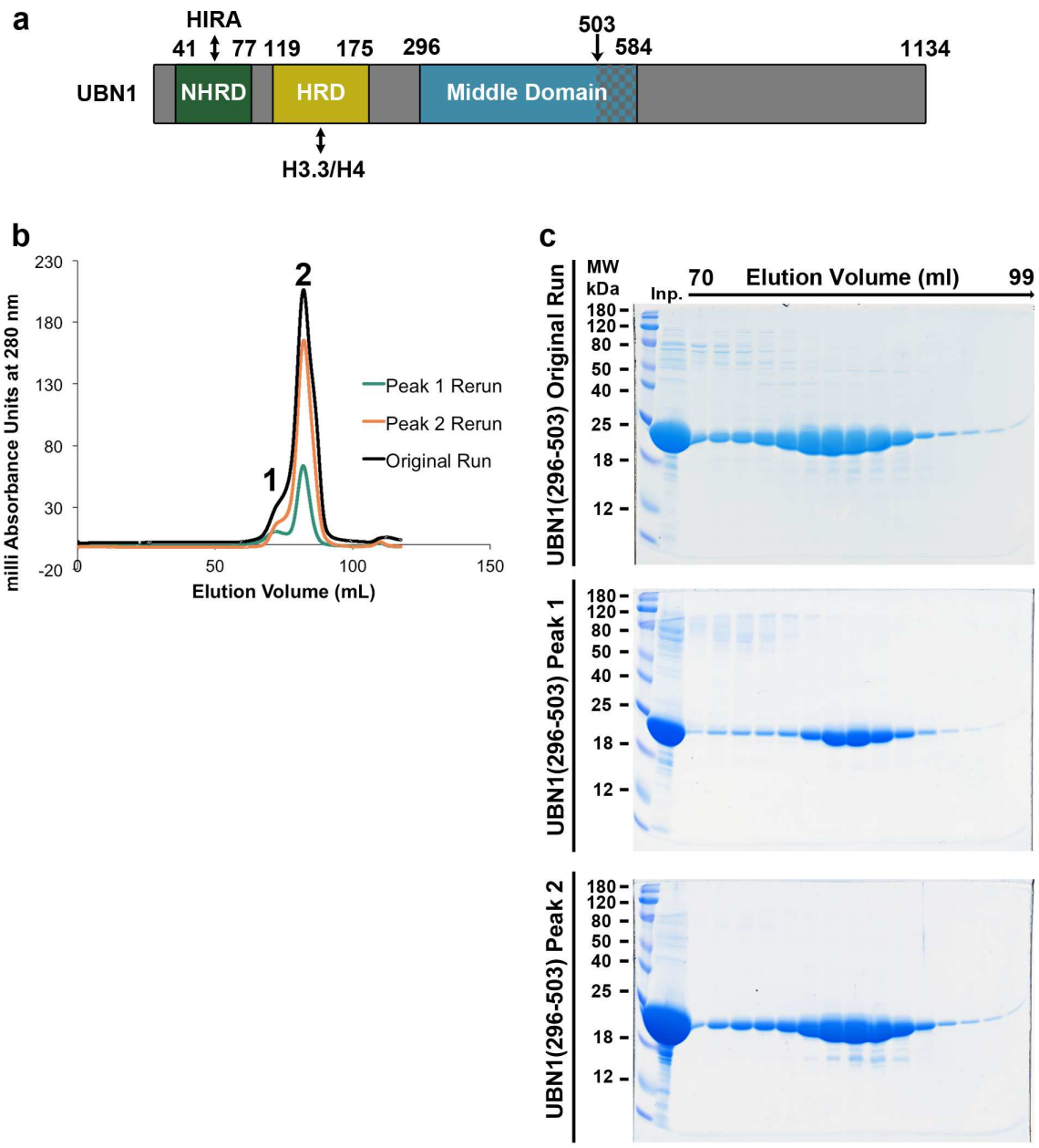


Figure 3.9 – Analysis of UBN1(296-503) with size-exclusion chromatography. (A) Domain architecture diagram for UBN1 depicting middle domain fragments. (B) Chromatograph of UBN1(296-503) resolved on a HiLoad s200 16/600 size-exclusion column. Fractions from 70-78 mL were pooled and re-loaded onto the column for "Peak 1 Rerun", while fractions from 80-90 mL were pooled and re-loaded onto the column for "Peak 2 Rerun". (B) SDS-PAGE gel analysis of every other 1 mL fraction spanning 70-99 mL for all three runs.

represents only the monomer population with a shoulder. This observation leads to the conclusion that residues 504-584 of UBN1 are required for dimer formation. We attempted to purify UBN1(504-584) as an MBP-UBN1(504-584) fusion, but this fragment formed very large soluble aggregates. We were unable to isolate soluble UBN1(504-584) with TEV protease cleavage and when MBP-UBN1(504-584) was resolved on a superose 6 10/300 column the majority eluted in the void volume, indicating that this fragment of UBN1 is aggregated and not suitable for biophysical analysis in isolation (Figure 3.10).

3.3 – The UBN1 middle domain specifically binds H3/H4

Given our observation here that the UBN1 middle domain forms stable monomer and dimer populations, together with our previous finding that the UBN-HRD domain (residues 119-175) binds to H3.3/H4, we decided to test if the UBN1 middle domain also binds to histones, potentially to contribute to $(H3.3.H4)_2$ deposition into chromatin. We prepared GST-UBN1(296-584) for a pull-down based histone-binding assay. GST-UBN1(296-584) was mixed with either H3.1/H4 or H3.3/H4 and binding was analyzed in buffer containing either 300 mM or 750 mM NaCl. We observed that UBN1(296-584) was able to bind H3.1/H4 and H3.3/H4 with equal affinity, and the interaction was only evident in the 300 mM NaCl buffer (Figure 3.11a), indicating that the interaction between H3/H4 and the UBN1 middle domain is electrostatic in nature and not specific for H3.3/H4. We then tested if the monomer or dimer population of UBN1(296-584) had any

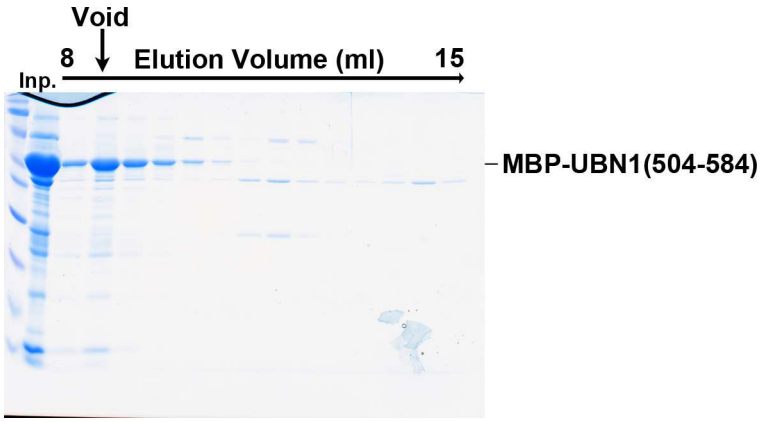


Figure 3.10 – Analysis of MBP-UBN1 with size-exclusion chromatography
MBP-UBN1(504-584) was resolved on a Superose 6 10/300 size-exclusion column. The majority of MBP-UBN1(504-584) eluted in the void volumn.

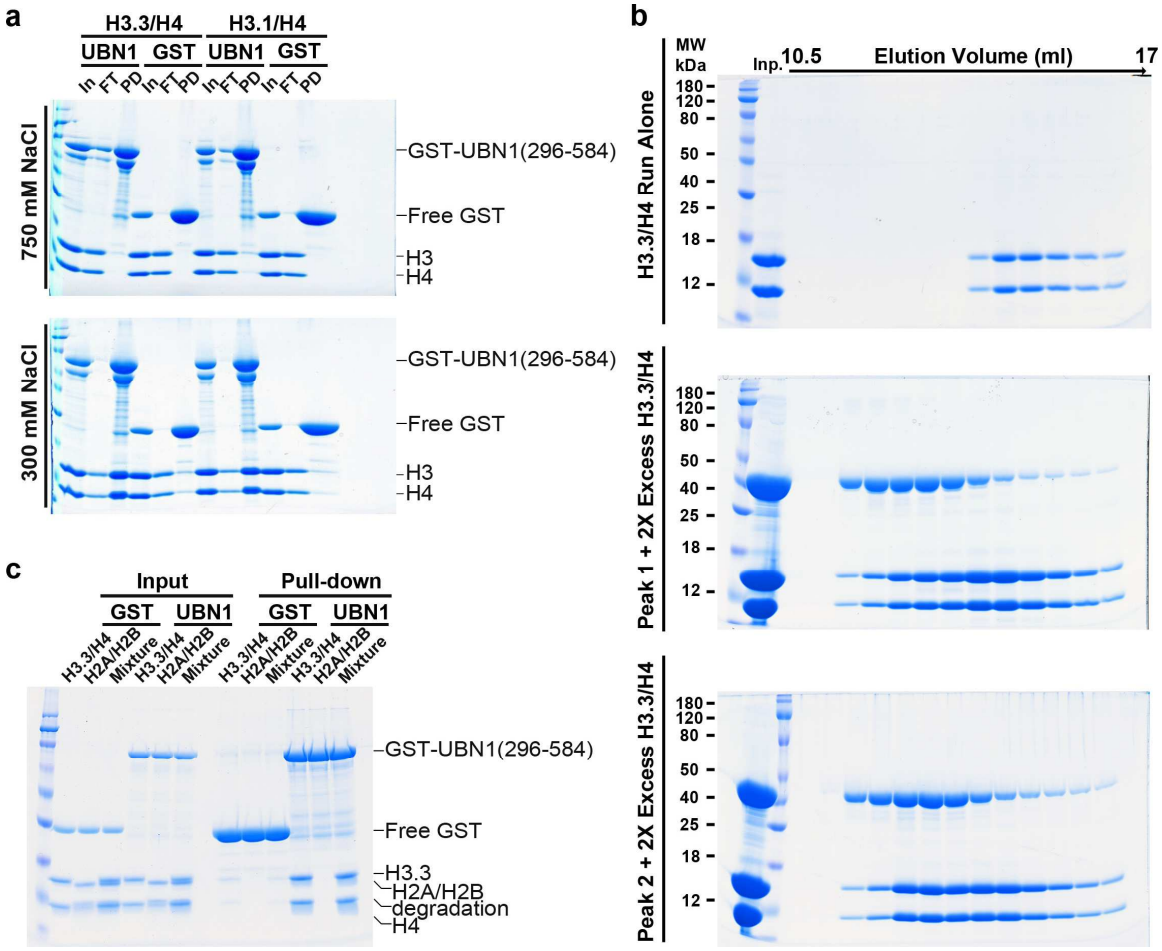


Figure 3.11 – UBN1(296-584) binds histones H3/H4 with no specificity for H3.3 vs. H3.1. (A) GST pull-down conducted with GST-UBN1(296-584) and H3.3/H4 or H3.1/H4 at 300 mM and 700 mM NaCl. (B) Size-exclusion histone binding experiment where UBN1(296-584) peak 1 and peak 2 were mixed with excess H3.3/H4 and run on an analytical s200 10/300 column to analyze complex formation. (C) GST pull-down conducted using GST-UBN1(296-584) with H3.3/H4 or H2A/H2B alone or in combination to analyze binding specificity.

influence on this middle domain/histone interaction. We isolated UBN1(296-584) peak 1 or peak 2 and incubated the individual populations with excess H3.3/H4 prior to resolving the mixtures over an s200 10/300 column, both the monomer and dimer populations were capable of forming a similar complex with H3.3/H4 on the s200 column (Figure 3.11b). To determine if the middle domain interaction with H3/H4 is significant, or just a non-specific electrostatic interaction, we conducted a pull-down histone-histone binding assay to compare the middle domain interaction with H3/H4 and H2A/H2B. Both H3/H4 and H2A/H2B have a strong positive charge, but we hypothesized that UBN1 would likely show specificity towards H3/H4 if the interaction is biologically relevant. GST-UBN1 (296-584) was incubated with H3.3/H4 or H2A/H2B individually as well as a 1:1 pool of H3.3/H4 and H2A/H2B in competition. The pull-down show that GST-UBN1(296-584) is able to bind H3.3/H4 but not H2A/H2B when the histones are assayed individually, and GST-UBN1(296-584) only binds H3.3/H4 from the competitive mixture (Figure 3.11c). These studies demonstrate that the UBN1 middle domain specifically binds H3/H4 but with no preference between H3.1 and H3.3.

Following the discovery that the UBN1 middle domain interacts specifically with H3/H4 in an electrostatic dependent manner, we then assayed the significance of the flanking middle domain regions, 296-340 and 504-584, to the H3/H4 interaction. GST-UBN1(296-503) and GST-UBN1 (341-503) were incubated with

H3.3/H4 and subjected to pull-down. While GST-UBN1(296-503) was able to pull-down H3.3/H4, no interaction above background was observed for GST-UBN1(341-503) (Figure 3.12a). Suggesting that UBN1 residues 504-584 are dispensable for H3/H4 binding, while the interaction is dependent on UBN1 residues 296-340. There is some mild evolutionary conservation in the region of UBN1 residues 296-340, but nothing striking to target for individual point mutations (Figure 3.1). We employed a deletion strategy in this region instead and generated several N-terminal deletions to compare for H3/H4 binding; GST-UBN1(296-503), GST-UBN1(316-503), GST-UBN1(331-503), and GST-UBN1(341-503). Each UBN1 fragment was incubated with H3.3/H4 and subjected to GST pull-down (Figure 3.12b). We observed that GST-UBN1(296-503) was able to form a stable complex with H3.3/H4, but with each successive N-terminal deletion we observed less H3.3/H4 being pulled-down, with GST-UBN1 341-503 having no interaction above background. This result indicates that many amino acids within UBN1(296-340) contribute to H3/H4 binding.

We then investigated if the middle domain binds to the folded histone core or the lysine-rich N-terminal tails. We previously characterized the UBN1-HRD/H3.3/H4 interaction and showed that the HRD binds primarily to the folded histone core with a smaller contribution from the histone tails (Ricketts et al., 2015). GST-UBN1(296-584) was incubated with either FL H3.3/H4 or N-terminally truncated H3.3(45-135)/H4(20-102) and subjected to pull-down. The results reveal that

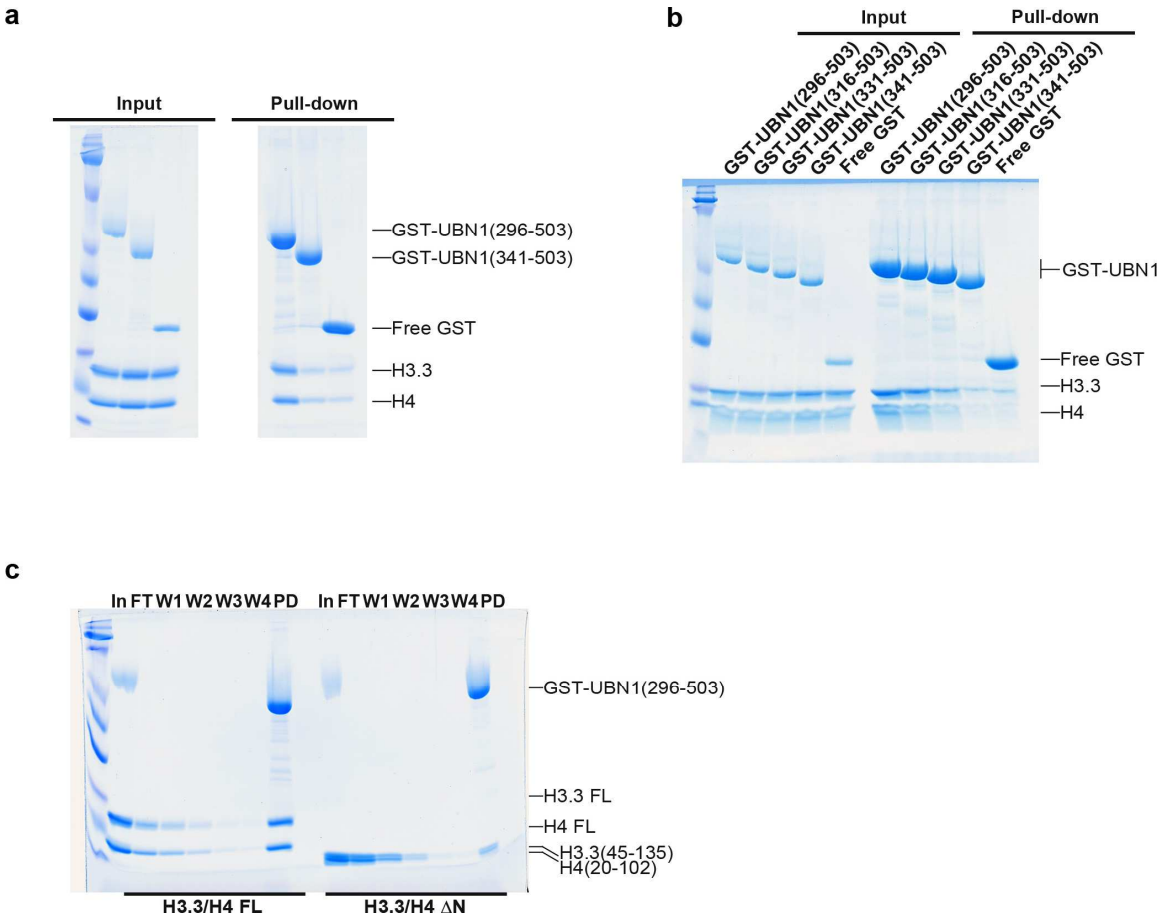


Figure 3.12 – A region of UBN1 spanning residues 296-341 is involved in binding to the H3/H4 N-terminal tails. (A) GST pull-down conducted with H3.3/H4 and GST-UBN1(296-503) or GST-UBN1(341-503). (B) GST pull-down conducted with H3.3/H4 and various UBN1 fragments illustrating that residues 296-341 contribute to H3/H4 binding. (C) GST pull-down conducted with GST-UBN1(296-503) and FL or ΔN H3.3/H4.

GST-UBN1(296-584) is able to pull-down significantly more FL H3.3/H4 in comparison with the N-terminally truncated complex (Figures 3.12c). This suggests that the UBN1 middle domain binds primarily to the N-terminal H3/H4 tails.

3.4 – UBN1 residues 176-295 non-specifically bind to DNA

The UBN1 region of 176-295, located between the HRD and middle domain, has little predicted secondary structure and is extremely lysine rich. There is no significant conservation in this region with the UBN1 homolog Yemanuclein from D. melanogaster, but if Yemanuclein is excluded from the multiple sequence alignment, it becomes clear that the lysine rich loop is well conserved among vertebrate UBN1 homologs (Figure 3.13a-b). This observation in combination with the report that UBN1 has been shown to have stronger non-specific interaction with DNA in comparison with other members of the HIRA complex (Ray-Gallet et al., 2011) lead to the hypothesis that this region of UBN1 may be involved in non-specific binding to DNA. To address this hypothesis, we generated MBP-UBN1(122-584); harboring the HRD, lysing rich loop, and middle domain. Within MBP-UBN1(122-584) we made several deletions including the entire 175-296 loop as well as smaller stretches of conserved lysine residues. Additionally we generated MBP-UBN1(176-195) to address if the loop alone would be sufficient for DNA binding. We employed a fluorescence polarization based DNA binding assay to test interaction with DNA. A FAM labeled 46 bp

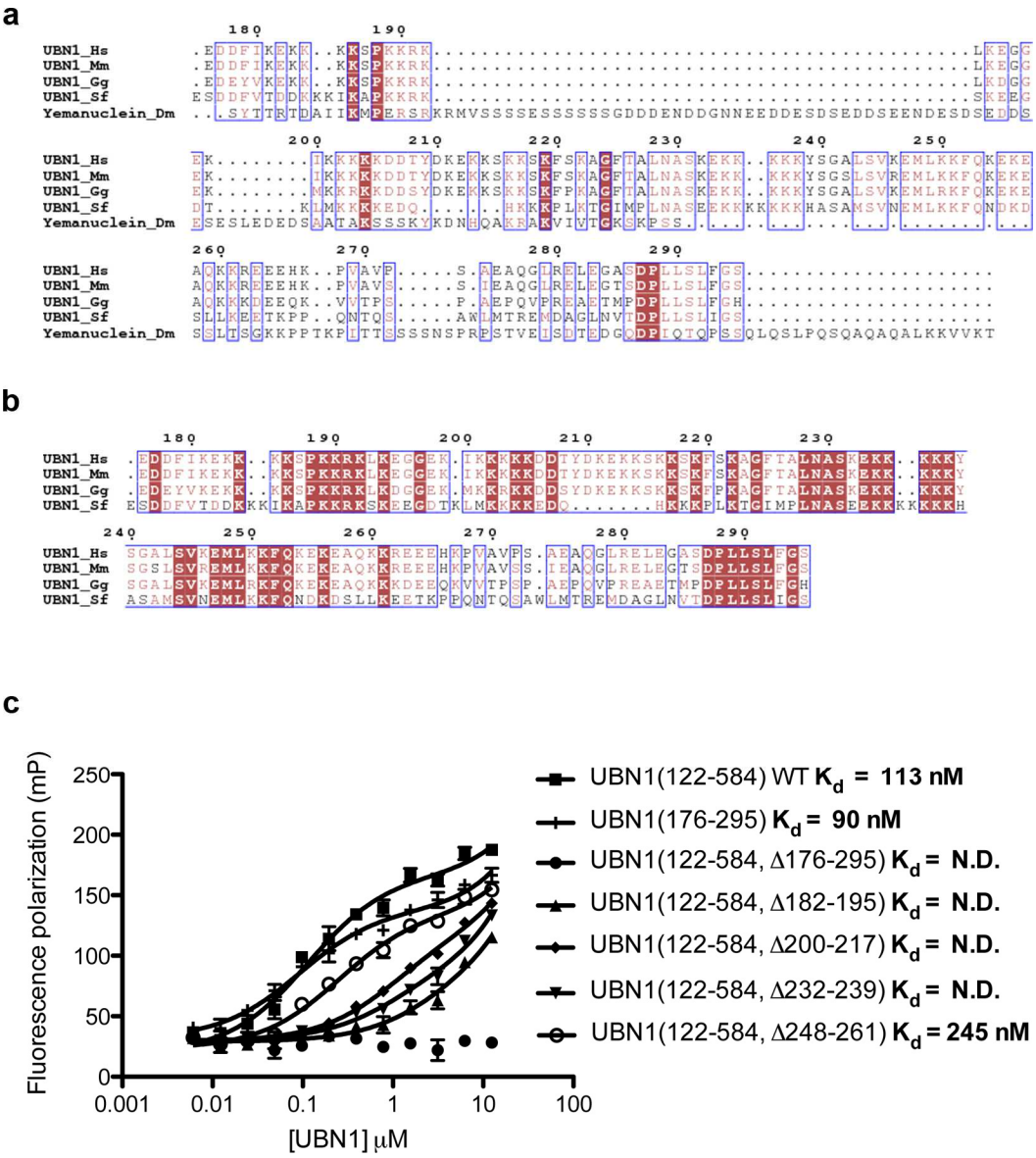


Figure 3.13 – UBN1 residues 176-295 bind to DNA with nanomolar affinity. (A) Alignment of UBN1 residues 176-295 with metazoan homologs. (B) Alignment of UBN1 176-295 with vertebrate homologs, lysine residues are clearly conserved among vertebrates. (C) Fluorescence polarization binding experiment conducted with MBP-UBN1 fragments and a double stranded LBS 1/2 FP probe. Deletion analysis of UBN1 shows that conserved lysine residues harbored in the 176-295 region contribute to DNA binding affinity.

double-stranded LBS1/2 DNA, originally generated to study binding with the Kaposi's sarcoma associated herpesvirus DNA-binding protein LANA (Domsic et al., 2013) was repurposed for our study as a non-specific DNA binding FP probe. We incubated the probe with increasing MBP-UBN1(122-584) to monitor the change in fluorescence polarization as the UBN1/DNA complex formed, and from the resulting FP curve we were able to determine the Kd of the UBN1/DNA interaction to be 113 nM (Figure 3.13c). Using the same FP assay we were able to show that MBP-UBN1(176-295) alone is able to bind DNA with a very comparable Kd of 90 nM, while MBP-UBN1(122-584, □176-295) did not have any detectable binding to DNA (Figure 3.13c). Additionally deletion of conserved lysine patches 182-195, 200-217, and 232-239 resulted in incomplete binding curves that could not be fit to determine Kd values, while deletion of 248-261 resulted in reduced binding with a Kd of 245 nM (Figure 3.13c). These data combined indicate that the lysine rich UBN1 region of 176-295 is essential for non-specific DNA binding.

To further confirm the veracity of our findings from the FP DNA binding assay, we analyzed MBP-UBN1(122-584) WT and □176-295 mutant for binding with LBS1/2 double stranded DNA without a FAM label on size-exclusion chromatography. MBP-UBN1(122-584) WT was incubated with 2-fold molar excess LBS1/2 double stranded DNA and resolved the complex on an s200 10/300 column superose 6 10/300 column. In this experiment, we observed a

significant shift in the elution volume of the MBP-UBN1(122-584) upon addition of DNA, indicating a higher molecular weight complex is formed with DNA (Figure 6d). While MBP-UBN1(122-584, □176-295) had no observable complex formation with the addition of DNA (Figure 3.14).

We then generated a new double stranded DNA fragment to compare with the LBS1/2 one we have been using to confirm that the DNA interaction is not dependent on local sequence. The double stranded DNA was 40 base pairs long, composed of 10 repeats of “ATGC”, which we named 10xATGC. Excess 10xATGC was incubated with MBP-UBN1(176-295) and the mixture was resolved on both the s200 and superose 6 10/300 size-exclusion column. In the same manner as LBS1/2 with MBP-UBN1(122-584), the 10xATGC formed a complex with MBP-UBN1(176-295), resulting in shift for the larger molecular weight species on the size-exclusion column in comparison with MBP-UBN1(176-295) resolved alone (Figure 3.15). These data indicate that UBN1 is able to bind both DNA sequences in a similar manner.

To quantitatively compare the 10xATGC and LBS1/2 interactions as well as probe the DNA specificity of the interaction, a competition FP experiment was preformed to compare the ability of 10xATGC DNA, Heparin, and tRNA to compete UBN1 from the FAM-LBS1/2 DNA. Heparin and tRNA represent non-DNA molecules with strong negative charge. 10xATGC DNA was able to

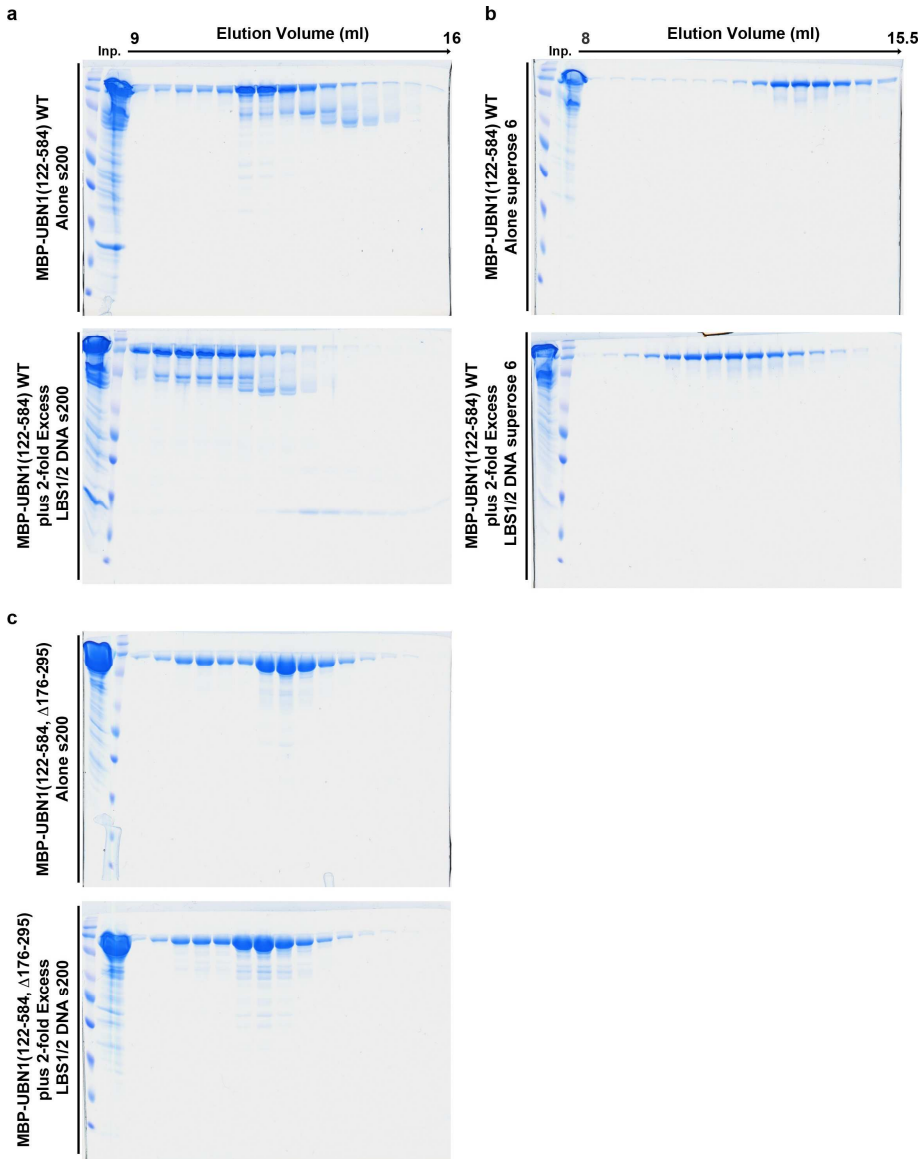


Figure 3.14 – Size-exclusion analysis of UBN1 / DNA complex formation. (A) MBP-UBN1(122-584) resolved on an s200 10/300 alone (top) or with 2-fold excess LBS1/2 double stranded DNA (bottom). (B) MBP-UBN1(122-584) resolved on a superpose 6 10/300 alone (top) or with 2-fold excess LBS1/2 double stranded DNA (bottom). (C) MBP-UBN1(122-584, Δ176-295) resolved on an s200 10/300 alone (top) or with 2-fold excess LBS1/2 double stranded DNA (bottom).

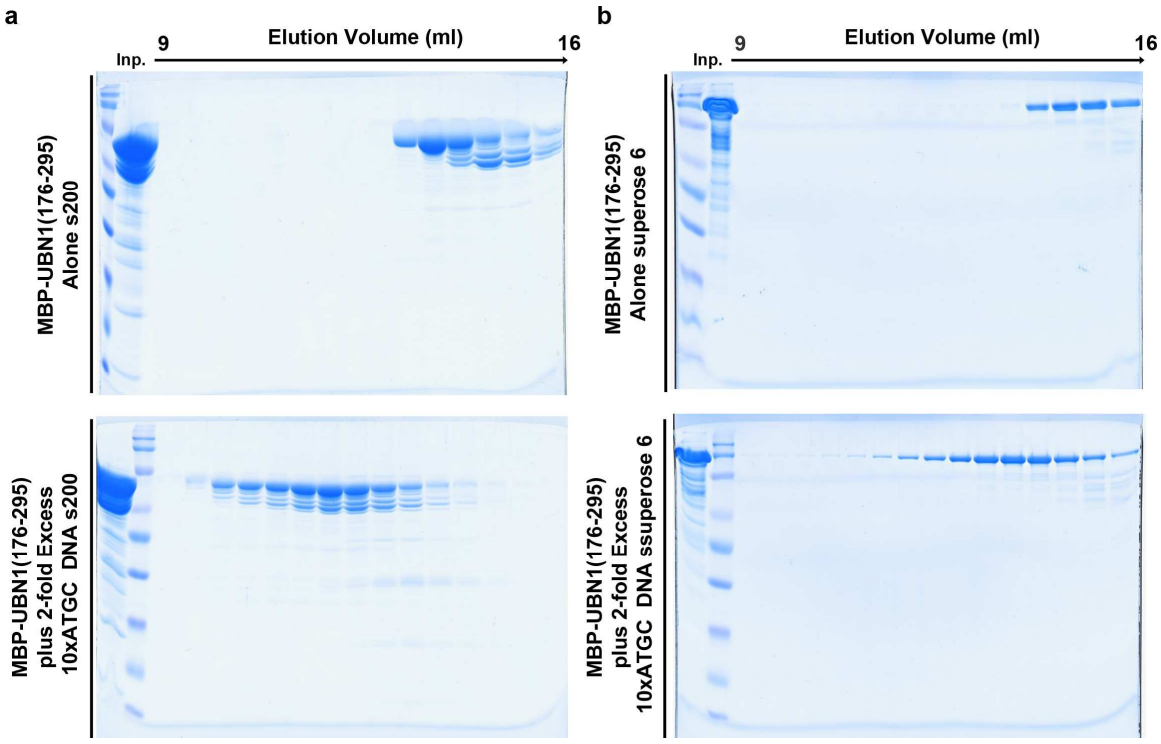


Figure 3.15 – Size-exclusion analysis of UBN1(176-295) / DNA complex formation. (A) MBP-UBN1(176-295) resolved on an s200 10/300 alone (top) or with 2-fold excess 10xATGC double stranded DNA (bottom). (B) MBP-UBN1(176-295) resolved on a superpose 6 10/300 alone (top) or with 2-fold excess 10xATGC double stranded DNA (bottom).

compete UBN1(122-584) away from FAM-LBS1/2 with an IC50 or 182 nM, while heparin had an IC50 of 80 nM and tRNA had an IC50 of 430 nM (Figure 3.16). While there is some variability among the IC50 value for the three competitors, all interactions are on the same order of magnitude, indicating that UBN1 residues 176-295 have a propensity to non-specifically bind negatively charged species. Although the interaction is based on non-specific charge, these results, together with the known biological activity of UBN1 to participate in histone deposition into chromatin is most consistent with a biologically significant role of UBN1 in DNA binding.

3.5 – The UBN1 HRD and middle domain bind H3/H4 in a mutually exclusive manner

As both the UBN1 HRD and middle domain have histone bindings activity, we investigated if they represent partial interactions of a larger UBN1/H3.3/H4 complex. First we mixed UBN1(296-584), UBN1(92-175), and H3.3/H4 in a 2:2:1 ratio and resolved the mixture over an s200 10/300 size-exclusion column. We observed that UBN1(92-175) appears to form complex with H3.3/H4 but UBN1 (296-584) does not appear incorporated in the complex (Figure 3.17a top). To further verify this result we ran the same column with just the 2:1 ratio of UBN1 (92-175) and H3.3/H4 and observed that the UBN1(92-175)/H3.3/H4 complex elutes in the same fractions, showing that there is no shift upon addition of UBN1(296-584) (Figure 3.17a bottom). Following this result, we analyzed if either

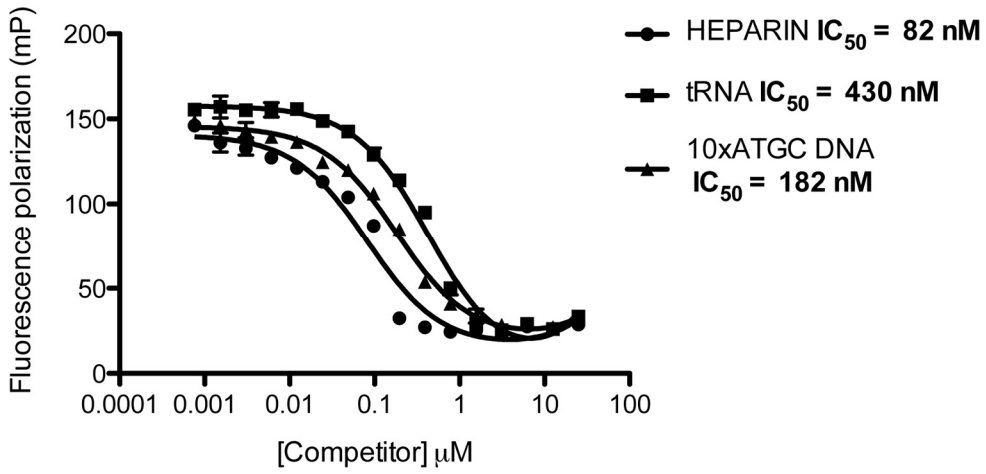


Figure 3.16 – UBN1 binds DNA though a non-specific electrostatic mechanism. A fluorescence polarization competition experiment was conducted by pre-forming a complex of UBN1 with FAM tagged LBS1/2 DNA. This complex was then titrated with heparin, tRNA, or double stranded 10xATGC DNA to determine the IC_{50} for the competitors.

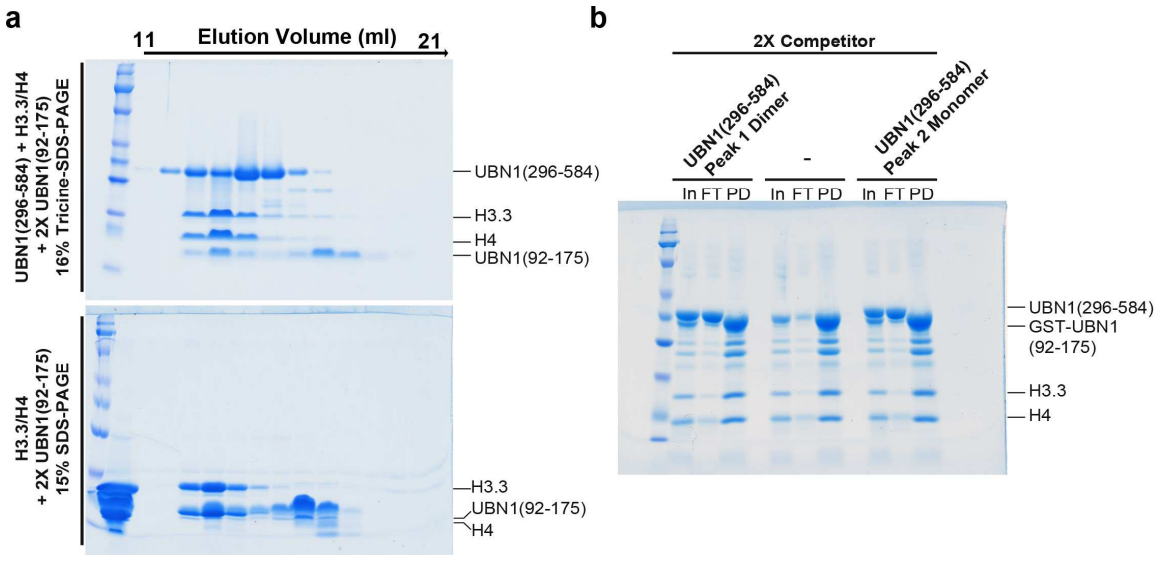


Figure 3.17 – The UBN1 HRD and middle domain bind histones in a mutually exclusive manner. (A) Size-exclusion experiment performed by adding 2-fold excess UBN1(296-584) and UBN1(92-175) to H3.3/H4 and resolving the mixture on an s200 10/300 column (top), control was run by adding 2-fold 92-175 to H3.3/H4 without UBN1(296-584) (bottom). (B) GST pull-down conducted by mixing GST-UBN1(92-175), H3.3/H4, and either UBN1(296-584) peak 1 or peak 2.

the UBN1(296-584) monomer or dimer populations could alter the mutually exclusive nature of the UBN1 HRD and middle domain histone interactions. A pull-down histone-binding assay was preformed though incubating either UBN1(296-584) monomer or dimer population with GST-UBN1(92-175) and H3.3/H4 prior to GST pull-down. Neither the UBN1(296-584) monomer or dimer were able to be incorporated into the GST-UBN1(92-175)/H3.3/H4 complex, and the interaction level for GST-UBN1(92-175) with H3.3/H4 was not altered by the addition of a UBN1(296-584) monomer or dimer competitor (Figure 3.17b). These observations combined lead to the conclusion that the UBN1-HRD and middle domain interact with H3.3/H4 in a mutually exclusive manner. Although this is somewhat confounding, the observations may be explained by a hand-off mechanism. While we observe preferential interaction of H3.3/H4 with the HRD over the middle domain, there may be some sort of biological switch, such as a post-translational modification or allosteric process, that allows for handoff from one domain to the other, although further investigation is necessary to confirm this hypothesis.

3.6 – UBN1 binds free H3/H4 but does not bind nucleosomes

Following our observations that UBN1 has an additional histone tail binding region as well as a DNA binding region, we investigated if these newly characterized domains in UBN1 could be involved in nucleosome binding. Nucleosomes were prepared with histones H3.1, H4, H2A, and H2B and a 147

bp double stranded DNA fragment bearing the 601 nucleosome positioning (Luger et al., 1999). Although UBN1-HRD binding is selective for H3.3/H4 (Ricketts et al., 2015), the UBN1 middle domain binds H3.1/H4 and H3.3/H4 with similar affinity and the H3.3-specific contact surface of the UBN1-HRD is not accessible in the nucleosome core particle (Figure 3.12c). We compared binding of nucleosomes versus H3.3/H4 alone with several different UBN1 fragments in a pull-down based binding assay. MBP-UBN1(122-584, WT), MBP-UBN1(122-584, □176-295), GST-UBN1(296-584), and GST-UBN1(122-175) were incubated with nucleosomes or H3.3/H4 and subjected to pull-down. We observed that while all four UBN1 fragments were capable of binding free H3.3/H4, none showed any detectable binding with nucleosomes (Figure 3.18). We also preformed the nucleosome pull-down for MBP-UBN1(122-584, WT) and MBP-UBN1(122-584, □176-295) at 150 mM NaCl in addition to 300 mM, we still observed no nucleosome binding at the lower salt concentration (Figure 3.18). Based on these results, we conclude that UBN1 does not participate in nucleosome binding.

3.7 – Discussion

UBN1 is an essential member of the HIRA complex, but only two very small domains have been studied in any depth. In our attempt to characterize regions of UBN1 outside of the well studied HRD and NHRD domains, we have identified the middle domain as a structured region of UBN1 that has dimerization and

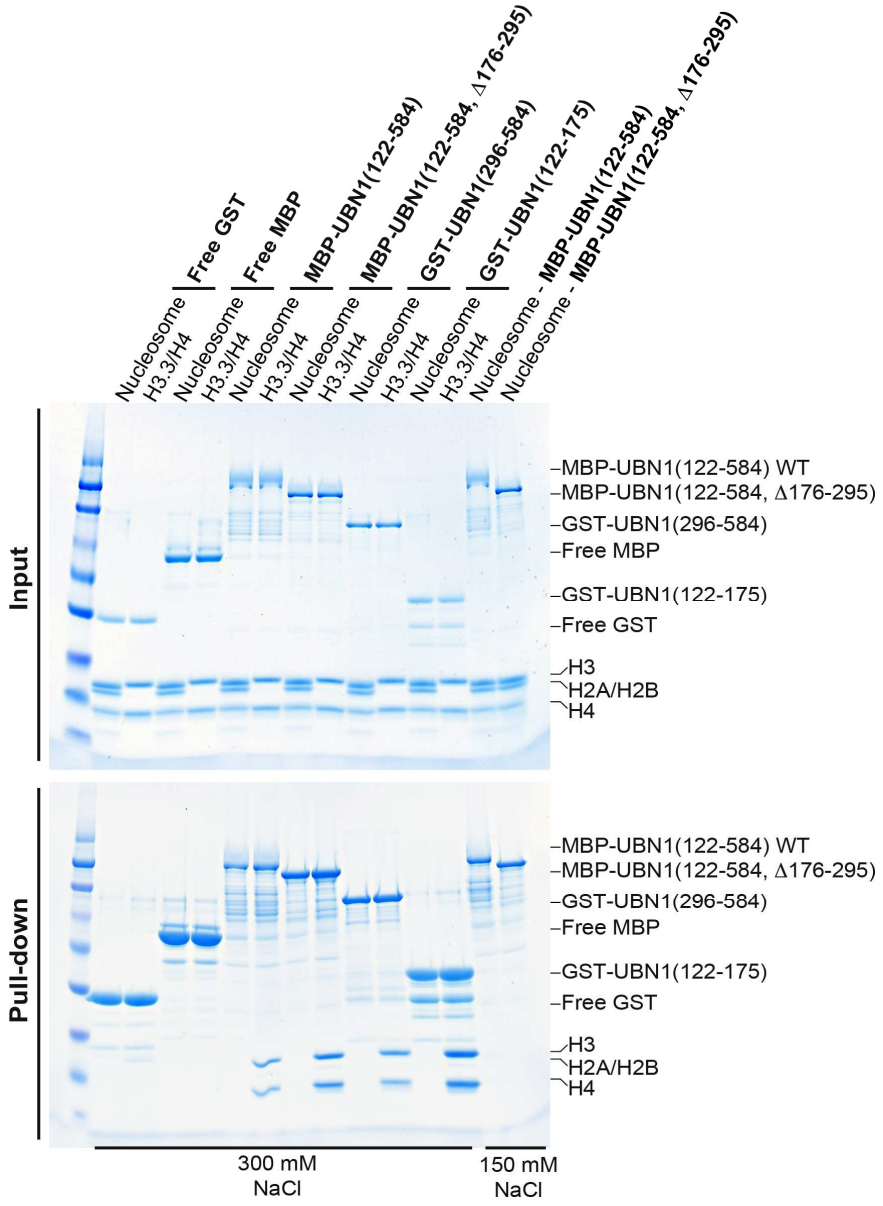


Figure 3.18 – UBN1 does not bind Nucleosomes. Pull-down experiment comparing ability of MBP-UBN1(122-584, WT), MBP-UBN1(122-584, Δ176-295), GST-UBN1(296-584), and GST-UBN1(122-175) to bind free H3.3/H4 or nucleosomes. As UBN1 residues 176-295 have been shown to contribute to DNA binding, the pull-down with MBP-UBN1(122-584, WT) and MBP-UBN1(122-584, Δ176-295) was attempted at 150 mM NaCl as well as 300 mM NaCl.

histone H3/H4 binding activity. Additionally, we also identified that the loop separating the HRD and middle domain harbors many conserved lysine residues that harbor non-specific DNA binding activity. Finally, we have shown that although UBN1 is capable of binding H3/H4 and DNA in isolation, it does not associate with assembled nucleosomes. Several studies have identified dimer-formation in other histone chaperones that bind to an $(H3.3/H4)_2$ tetramer for nucleosome deposition (Fazly et al., 2012; Huang et al., 2015; Richet et al., 2015; Su et al., 2012; Winkler et al., 2012; Zasadzinska et al., 2013), and it was previously demonstrated that the majority of H3.3/H4 is deposited into nucleosomes in the tetramer form (Xu et al., 2010). Furthermore, human UBN1 has been previously shown to exhibit DNA binding activity (Ray-Gallet et al., 2011), as have several other histone chaperones such as FACT (Safina et al., 2017) and CAF-1 (Sauer et al., 2017; Zhang et al., 2016). From these observations in combination with ours, we propose a model for HIRA complex function in which UBN1 binds and stabilizes an $(H3.3/H4)_2$ tetramer in concert with DNA binding, prior to H3.3/H4 deposition and final nucleosome assembly with the aid of additional histone chaperones and cellular factors (Figure 3.19).

Experiments with the full-length HIRA complex and different HIRA sub-complexes are needed to rigorously confirm our model, as we have only observed a lack of nucleosome binding for UBN1 in isolation. It has recently been reported that the WD40 domain of HIRA is essential for nucleosome assembly

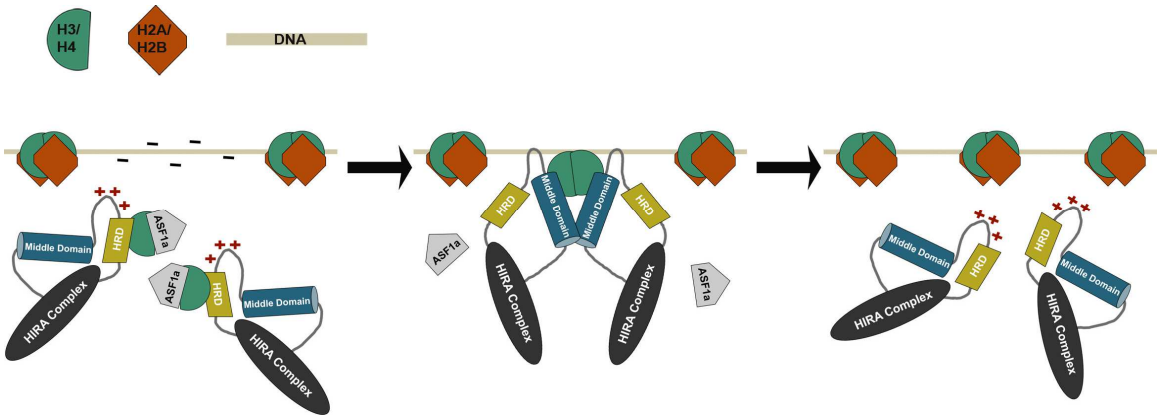


Figure 3.19 – Proposed model for H3.3/H4 deposition by the HIRA complex

Proposed model for H3.3/H4 deposition coordinated by the HIRA complex based on our observation. The three-step model first shows an H3.3/H4 dimer bound by ASF1 and the UBN1-HRD. Next, UBN1 binds DNA and middle domain forms a dimer and associates with H3.3/H4 to displace ASF1a and the UBN1-HRD to allow for $(H3.3/H4)_2$ tetramer formation. Finally the HIRA complex disassociates from the $(H3.3/H4)_2$ tetramer and DNA to allow additional chaperones to deposit H2A/H2B and coordinate final nucleosome formation.

(Zhu et al., 2017), although this observation could be due to the function of HIRA as a molecular scaffold for the HIRA complex, as the UBN1-NHRD binds to the HIRA-WD40 domain (Tang et al., 2012). Additionally, our observation of mutually exclusive H3.3/H4 binding by the UBN1-HRD and middle domain is somewhat confounding, further investigation will be required to determine if there is a potential hand-off or other allosteric mechanism that could be driven by a yet to be determined factor. Of particular interest to our studies is that the CAF-1 complex has recently been reported to form an H3/H4 tetramer only upon binding to an H3/H4 dimer together with DNA, which allows for the formation of a CAF-1 dimer to drive $(H3/H4)_2$ tetramer formation prior to deposition (Mattiroli et al., 2017). Perhaps UBN1 could operate through a similarly triggered mechanism. While there is still much that remains unknown about how the HIRA complex and other histone chaperones mediate histone deposition, our data suggests that UBN1 may function similarly to other H3/H4 histone chaperones though a mechanism that employs both dimerization and DNA binding activities to efficiently deposit $(H3.3/H4)_2$ onto DNA for nucleosome formation.

3.8 – Materials and Methods

Multiple sequence alignments

Alignments for Figures 3.1 and 3.13 were generated using Clustal Omega (Sievers et al., 2011), and formatted for better visualization using ESPript version 3.0 (Gouet et al., 2003).

Protein secondary structure prediction

Predicted secondary structure of UBN1 depicted in Figure 1b was generated using the PSIPRED secondary structure prediction server (Buchan et al., 2013).

Generation of expression plasmids

The plasmids encoding His-UBN1(296-584), His-UBN1(341-584), and His-UBN1(296-503) were generated by PCR amplification from a FL UBN1 construct (Banumathy et al., 2009) and ligation into the BamHI/XhoI sites of a custom engineered pCDFduet-1 (Novagen) Escherichia coli expression vector carrying a N-terminal 6xHis tag that is removable by TEV protease cleavage. The plasmids encoding GST-UBN1(296-584), GST-UBN1(296-503), GST-UBN1(316-503), GST-UBN1(330-503), GST-UBN1(341-503), GST-UBN1(92-175), and GST-UBN1(122-175) were generated by PCR amplification and ligation into the BamHI/XhoI sites of a custom engineered pCDFduet-1 (Novagen) Escherichia coli expression vector carrying a N-terminal GST tag that is removable by TEV protease cleavage. The plasmid encoding His-MBP-UBN1(122-584) was generated by PCR amplification and ligation into the BamHI/XhoI sites of a pDB.His.MBP Escherichia coli expression vector carrying a N-terminal His-MBP tag that is removable by TEV protease cleavage (Seiler et al., 2014). The plasmids encoding His-MBP-UBN1(122-584, □176-295), His-MBP-UBN1(122-584, □182-195), His-MBP-UBN1(122-584, □200-217), His-MBP-UBN1(122-584, □232-239), and His-MBP-UBN1(122-584, □□□□□□□□)□were generated by

site-directed mutagenesis(Weiner and Costa, 1994). The plasmids encoding FL H3.1, H3.3, and H4 were generated as described previously (Ricketts et al., 2015). The plasmid encoding H2B was a gift from Don Cleveland at UCSD. The plasmids encoding H3.3(45-135) and H4(20-102) were generated by PCR amplification of the H3.3 and H4 sequences and ligation into the BamHI/XhoI sites of custom engineered pETduet-1 (Novagen) Escherichia coli expression vector carrying a N-terminal His tag that is removable by TEV protease cleavage.

Protein expression and purification

H2A, H2B, H3, and H4 individual histone proteins were expressed to the inclusion bodies in BL21-Gold(DE3) cells (Agilent). The individual histone proteins were purified and refolded to form H3/H4 and H2A/H2B as previously described (Luger et al., 1999; Ricketts et al., 2015). These proteins were additionally used to form nucleosomes with the addition of 601 nucleosome positioning sequence double stranded DNA according to a previously reported protocol (Luger et al., 1999). His-UBN1(296-584), His-UBN1(341-584), and His-UBN1(296-503) were generated in BL21-Gold(DE3) cells (Agilent) induced with 0.8 mM IPTG and expressed overnight at 18°C. Cells were re-suspended in buffer containing 20 mM Tris pH 8.0, 500 mM NaCl, 5 mM BME, and 10 mM imidazole and lysed with sonication. Lysate was clarified with centrifugation and the supernatant was subjected to nickel affinity chromatography to isolate the His tagged protein. Proteins were eluted with 200 mM Imidazole and dialyzed

overnight with the addition of TEV protease to remove the His tag, dialysis buffer contains 20 mM Tris pH 7.5, 50 mM NaCl, and 5 mM BME. Proteins were then subjected to ion exchange chromatography using a HiTrap SP column (GE Healthcare). Following ion exchange, proteins were concentrated using a spin concentrator (Millipore) and loaded either onto a Superdex 200 10/300 or 16/600 column (GE Healthcare) in buffer with 20 mM Tris pH 7.5, 300 mM NaCl, and 1 mM TCEP for further purification and analysis by gel filtration. GST-UBN1(296-584), GST-UBN1(296-503), GST-UBN1(316-503), GST-UBN1(330-503), GST-UBN1(341-503), GST-UBN1(92-175), and GST-UBN1(122-175) were generated in BL21-Gold(DE3) cells (Agilent) induced with 0.8 mM IPTG and expressed overnight at 18°C. Cells were re-suspended in 1XPBS supplemented with 5 mM BME and lysed by sonication. Lysate was clarified by centrifugation and supernatant was subjected to GST affinity chromatography. Protein was eluted with 20 mM reduced glutathione and dialyzed into a buffer containing 20 mM Tris pH 7.5, 300 mM NaCl and 5 mM BME for storage. UBN1 92-175 was generated by adding TEV protease to GST-UBN1 92-175 bound to GST affinity resin followed by incubation at 4°C overnight. Cleaved protein was eluted with a buffer containing 20 mM Tris pH 7.5, 300 mM NaCl, and 5 mM BME. Eluted protein was concentrated using a spin concentrator (Millipore) and loaded onto a Superdex 75 10/300 (GE Healthcare) for further purification by gel filtration in buffer with 20 mM Tris pH 7.5, 300 mM NaCl, and 1 mM TCEP. His-MBP-UBN1(122-584, □176-295), His-MBP-UBN1(122-584, □182-195), His-MBP-UBN1(122-584,

□200-217), His-MBP-UBN1(122-584, □232-239), and His-MBP-UBN1(122-584, □□□□□□□□)□were expressed and purified in the same method as His-UBN1(296-584), His-UBN1 (341-584), and His-UBN1(296-503) with omission of the TEV cleavage step.

Size-exclusion analysis of UBN1 peak 1 and peak 2 distribution

UBN1(296-584) was purified though the ion exchange step, protein eluted from the ion exchange column was pooled and loaded onto a Superdex 200 16/600 column (GE Healthcare) in buffer containing 20 mM Tris pH 7.5, 300 mM NaCl, and 1 mM TCEP. Protein eluting in the range of 65-75 mL was pooled as peak 1 and protein eluting in the range of 76-89 mL was pooled as peak 2. Both peak 1 and peak 1 were re-loaded back onto the column for further analysis of distribution. UBN1(341-584) was analyzed in the same method as UBN1(296-584) with the exception that peak 1 was defined as 68-78 mL and peak 2 was defined as 80-94 mL. UBN1(296-503) was analyzed in the same method as UBN1(296-584) with the exception that peak 1 was defined as 70-78 mL and peak 2 was defined as 80-90 mL. Analysis of UBN1(296-584) peak 1 and peak 2 on the analytical Superdex 200 10/300 column was conducted by loading protein isolated from the first run on the Superdex 200 16/600 onto the analytical column for comparison with molecular-weight standards (Biorad).

Analytical ultracentrifugation – sedimentation equilibrium

UBN1(296-584) peak 1 was prepared for sedimentation equilibrium analysis by conducting an initial gel filtration run on a Superdex 200 16/600. Protein eluting in the region of 65-75 mL was pooled and re-loaded onto the Superdex 200 16/600, the protein eluting in the same region for the second run was then pooled and analyzed for concentration. Samples were prepared at 0.3, 0.5, and 0.7 OD 280 nm, loaded into a 6-channel ultracentrifugation cell with buffer blanks, and analyzed for equilibrium distribution analysis at 12000, 18000, and 26000 RPM using an Optima XL-I analytical ultracentrifuge (Beckman). Equilibrium distribution data were fit with the program HeteroAnalysis using the ideal fitting model (Cole, 2004). UBN1(296-584) peak 2 was prepared for sedimentation equilibrium analysis by conducting an initial gel filtration run on a Superdex 200 16/600. Protein eluting in the region of 76-89 mL was pooled and re-loaded onto the Superdex 200 16/600. The protein eluting in the same region for the second run was then pooled and analyzed for concentration. Protein was analyzed for equilibrium distribution similarly to peak 1.

LC-MS analysis of protein samples

LC-MS digestion and peptide analysis was preformed by the proteomics and metabolomics facility at The Wistar Institute in Philadelphia, Pennsylvania. We provided the facility with an SDS-PAGE gel stained with coomassie brilliant blue and indicated the protein bands for analysis.

GST/MBP pull-down binding assay

Pull-down experiments were conducted by incubating 2 μM tagged bait protein with 4 μM untagged prey protein for 30 minutes at 4°C. Proteins were then subjected to pull-down by incubation with glutathione agarose (Fisher Scientific) or amalyose (NEB) resin for 30 minutes. Bound resin was washed with 120 column volumes of buffer before elution of bound proteins by boiling resin in SDS loading buffer. Pull-downs were analyzed by visualization using SDS-PAGE stained with coomassie brilliant blue G-250. All pull-downs were conducted in buffer with 20 mM Tris pH 7.5, 300 mM NaCl, and 1 mM TCEP with the exception of Figure 3.11a which was conducted in both this buffer and one with 750 mM NaCl, and Figure 3.18 which was conducted in both this buffer and one with 150 mM NaCl. Pull-down conducted in Figure 7b was conducted by mixing 2 μM GST-UBN1(92-175), 2 μM H3.3/H4 and 4 μM UBN1(296-584) peak 1 or peak 2.

Formation of protein/protein complexes using size-exclusion chromatography

To form complexes with UBN1(296-584) peak 1 or peak 2 and H3.3/H4 for Figure 3.11c, 200 μM UBN1 (296-584) was incubated with 400 μM H3.3/H4 and resolved on a Superdex 200 10/300 column (GE Healthcare) in a buffer with 20 mM Tris pH 7.5, 300 mM NaCl, and 1 mM TCEP. H3.3/H4 was run alone at a concentration of 100 μM. For analysis of UBN1 92-175 and UBN1 296-584 H3.3/H4 binding cooperation for Figure 3.17; 200 μM UBN1 296-584, 200 μM

UBN1 92-175 and 100 μM H3.3/H4 were incubated in a buffer containing 20 mM Tris pH7.5, 300 mM NaCl, and 1 mM TCEP and resolved on a Superdex 200 10/300 column (GE Healthcare). Analysis of UBN1 92-175 and H3.3/H4 alone was conducted by incubating 200 μM UBN1 92-175 with 100 μM H3.3/H4 and resolving the mixture on the column in the same 20 mM Tris pH 7.5, 300 mM NaCl, 1 mM TCEP buffer.

DNA probes

The FAM labeled LBS1/2 (LANA binding sites 1 and 2) double stranded DNA used for the fluorescence polarization DNA binding assay was assembled from synthesized DNA fragments (IDT). This FAM labeled probe contains a DNA sequence recognized by the KSHV protein LANA and was originally used to study this interaction(Domsic et al., 2013), it was repurposed for our study as UBN1 has non-specific DNA binding activity. Unlabeled LBS1/2 and 10XATGC double stranded DNA was assembled from synthesized sense and anti-sense fragments. Sense fragment sequences:

LBS1/2:

AGCGGCCCCATGCCCGGGCGGGAGGCGCCGCAGGCCCCGGCGGCGT

10XATGC:

ATGCATGCATGCATGCATGCATGCATGCATGCATGCATGCATGCATGC

Sense and anti-sense fragments were dissolved in a buffer with 10 mM Tris pH 7.5, 10 mM NaCl, and 1 mM EDTA, mixed in a 1:1 ratio and boiled followed by slow cooling to room temperature to anneal and form double stranded DNA.

Fluorescence polarization DNA-binding assay

To quantitatively analyze the UBN1/DNA interaction, a fluorescence polarization based binding assay was performed. Protein and DNA used for this assay were prepared in buffer containing 20 mM Tris pH 7.5, 150 mM NaCl, 1 mM TCEP, and 2 mg/mL BSA. The change in fluorescence polarization was monitored as increasing concentration of UBN1 protein was titrated into FAM-LBS1/2 DNA (10 nM) to determine the Kd of the UBN1/DNA interaction. All data were collected using a Perkin Elmer EnVision Xcite Multilabel plate reader, and curves were fit using a one site – total binding model with GraphPad Prism (version 5.0a). A competition-binding assay was also performed to determine the IC50 for the interaction between UBN1 and: Heperin (Fisher), tRNA, and 10XATGC DNA. To monitor the ability of each competitor to compete UBN1 from LBS1/2 DNA, increasing concentration of competitor was titrated into solution containing 10 nM FAM-LBS1/2 and 300 nM MBP-UBN1(122-584). The tRNA used for this study was a pooled tRNA purified from yeast (Invitrogen).

3.9 – Acknowledgments

We would like to thank John Domsic for use of the LBS1/2 probe that he designed. The plasmid encoding H2B was a gift from Don Cleveland (University of California San Diego), and the plasmid encoding H4 was a gift from Karolin Luger (University of Colorado Boulder). This work was supported by the National Institutes of Health grant AG031862 to R.M and an American Heart Association predoctoral fellowship 12PRE12030157 to M.D.R. We acknowledge the Proteomics and Metabolomics Facility at The Wistar Institute for preparing Gel/LC-MS/MS data, and the Molecular Screening and Protein Expression Facility at The Wistar Institute for use of the Perkin Elmer EnVision Xcite Multilabel plate reader.

CHAPTER 4: Molecular characterization of HIRA oligomer formation and interaction with CABIN1

4.1 – Overview

The human HIRA histone chaperone complex carries out deposition of H3.3/H4 into chromatin in a replication independent manner(Tagami et al., 2004); H3.3/H4 deposition by the HIRA complex is necessary for proper regulation of gene expression (Banaszynski et al., 2013; Goldberg et al., 2010; Jullien et al., 2012), as well as DNA and chromatin repair and maintenance (Adam et al., 2013; Ray-Gallet et al., 2011; Schneiderman et al., 2012; Zhang et al., 2007). The complex is composed of the proteins HIRA, UBN1, and CABIN1, which are supplied with H3.3/H4 by the transient subunit ASF1a. The assembly and function of the HIRA complex has been relatively well studied; the individual protein/protein contacts responsible for complex formation for HIRA/UBN1(Tang et al., 2012), HIRA/ASF1a(Tang et al., 2006; Zhang et al., 2005), and UBN1/H3.3/H4(Ricketts et al., 2015) have been the focus of in depth biochemical and structural studies. However, the molecular contacts responsible for formation of the HIRA/CABIN1 complex remain relatively unclear. It has been reported that CABIN1 residues 1-941 are sufficient and necessary for complex formation with HIRA, and that this interaction is also dependent on HIRA residues 520-1017. It has been reported that CABIN1 is dispensable for H3.3/H4 deposition by the HIRA complex, while HIRA and UBN1 are essential for H3.3/H4 function(Ray-Gallet et al., 2011). Despite not being required for H3.3/H4 deposition, CABIN1 is an important member of the HIRA complex as knockdown of CABIN1 has been reported to

decrease cellular expression of HIRA, and knockdown/microarray studies have shown HIRA and CABIN1 regulate expression for many of the same genes(Rai et al., 2011). CABIN1 has been reported to interact with the transcription factors Mef2 and p53 (Choi et al., 2013; Esau et al., 2001; Jang et al., 2009; Youn and Liu, 2000), and a crystal structure of a CABIN1 peptide bound to Mef2 has been reported(Han et al., 2003). The existing evidence leads to the hypothesis that CABIN1 could play a role in the structural stability of the HIRA complex and may bind with some transcription factors to recruit the HIRA complex to specific areas of chromatin. To address this hypothesis, additional investigation into the protein/protein contacts responsible for HIRA/CABIN1 complex formation are necessary in order to design rational experiments to probe the biological significance of the HIRA/CABIN1 interaction.

Based on the gap in knowledge outlined above, we decided to carry out experiments to probe the region of HIRA that interacts with CABIN1 as well as the HIRA/CABIN1 interaction. We anticipated that these studies would provide insights into the roles played HIRA and CABIN1 H3.3/H4 deposition by the HIRA complex. Our studies were facilitated by unpublished cellular data from a collaborator demonstrating that HIRA residues 600-1017 are responsible for HIRA oligomer formation as well as CABIN1 binding. Specifically, our collaborators demonstrated that deletion within this HIRA domain abrogated co-IP of differentially tagged HIRA as well as CABIN1 (data not shown), suggesting

that oligomer formation and CABIN1 binding by this HIRA domain may be functionally linked. Here we present biochemical data demonstrating that HIRA(661-1017) exists a stable trimer in solution and is able to form a complex with CABIN1. Sedimentation equilibrium experiments show that the HIRA/CABIN1 complex has a stoichiometry of HIRA(3)/CABIN1(2). We demonstrate that HIRA trimer formation is dependent on the presence of two linked domains within this HIRA region; HIRA(661-872) likely forms a strand domain and HIRA(905-1017) a helical domain, the two are linked by a loop spanning residues 873-904. We show that replacing the HIRA 873-904 loop with a glycine/serine linker still allows for trimer formation while abrogating CABIN1 binding. Further we show that a HIRA F870A_R871A_L874A triple mutant decreases the HIRA/CABIN1 interaction. Together these data suggest that HIRA/CABIN1 complex formation is dependent on HIRA trimer formation, while specific contacts in the region of the HIRA(873-904) loop are additionally necessary for CABIN1 interaction.

4.2 – The HIRA C-terminal domain forms a trimer and interacts with CABIN1

We first set out to make a C-terminal fragment of HIRA that could be stably expressed and purified from *E. coli*. While the N-terminal region of HIRA has been well characterized as a predicted WD40 repeat domain, the C-terminal region of HIRA has been relatively understudied. We used the PSIPRED server (Buchan et al., 2013) to predict the secondary structure for HIRA residues 651-

1017 (Figure 4.1 top), which shows that this HIRA region is largely structured and organized into a strand region (~residues 670-860) and helical region (~residues 885-1017). We have assigned these previously uncharacterized domains on a new HIRA domain architecture schematic (Figure 4.1 bottom). We decided to generate a plasmid for expression of HIRA(661-1017), and were able to successfully purify HIRA(661-1017) from *E. coli*. We showed that this HIRA fragment forms a stable monodisperse population eluting between the 158 kD and 44 kD standards on a Superdex 200 10/300 GL column (Figure 4.2a). This domain of HIRA has previously been reported to interact with H3.3/H4, with increased interaction for H3.3/H4 bearing phosphorylation on H4 S47(Kang et al., 2011). To test this interaction we incubated HIRA(661-1017) with 2-fold molar excess H3.3/H4 and resolved the mixture on a Superdex 200 10/300 GL column. We observed no complex formation (Figure 4.2b), although this experiment still needs to be repeated with H3.3/H4 harboring the S47 phosphorylation on H4.

To study complex formation with CABIN1 FL we incubated 5-fold molar excess HIRA(661-1017) with CABIN1 FL purified from Sf9 cells and resolved the mixture on a Superdex 200 10/300 GL column in comparison with HIRA(661-1017) and CABIN1 FL run alone on the same column (Figure 4.3a-b). We observed an increased molecular weight shift and co-migration of the HIRA and CABIN1 bands on an SDS-PAGE gel, indicating that HIRA(661-1017) forms a stable complex with CABIN1 FL in this assay.

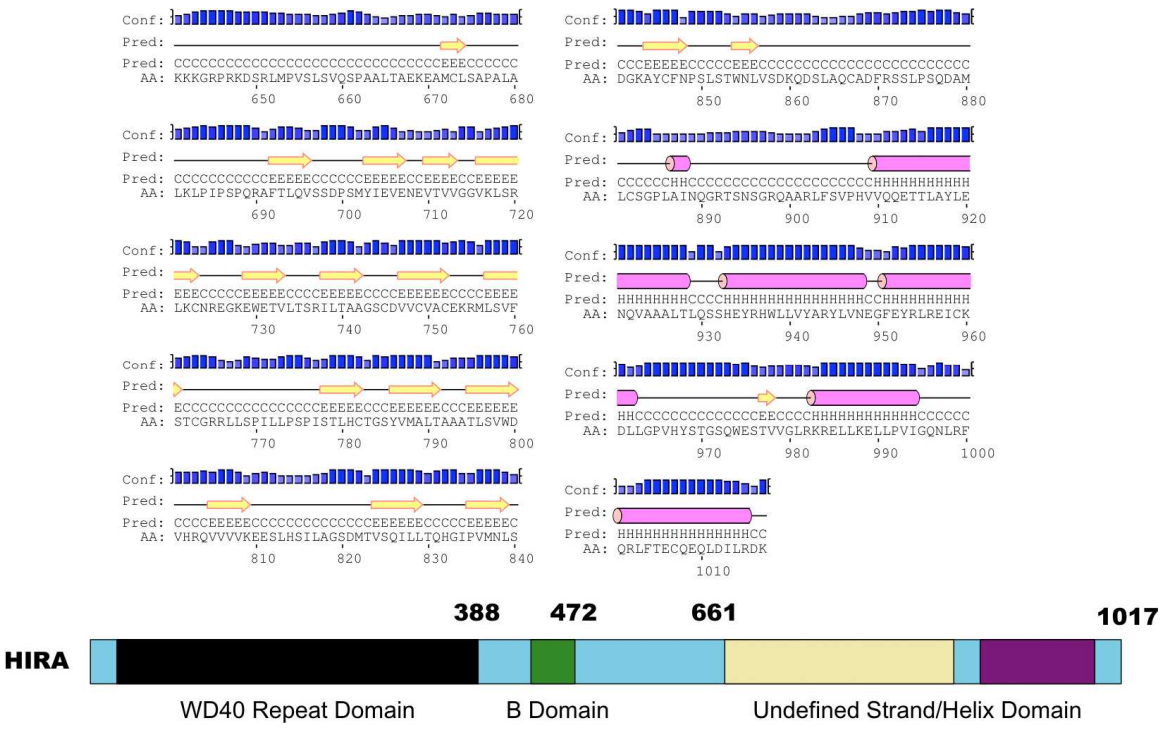


Figure 4.1 – Secondary structure prediction and domain architecture for HIRA. (Top) Secondary structure prediction for HIRA(651-1017) depicting the predicted strand domain spanning residues 670-860 and the predicted helical domain spanning residues 885-1017. (Bottom) Domain architecture diagram for full-length human HIRA.

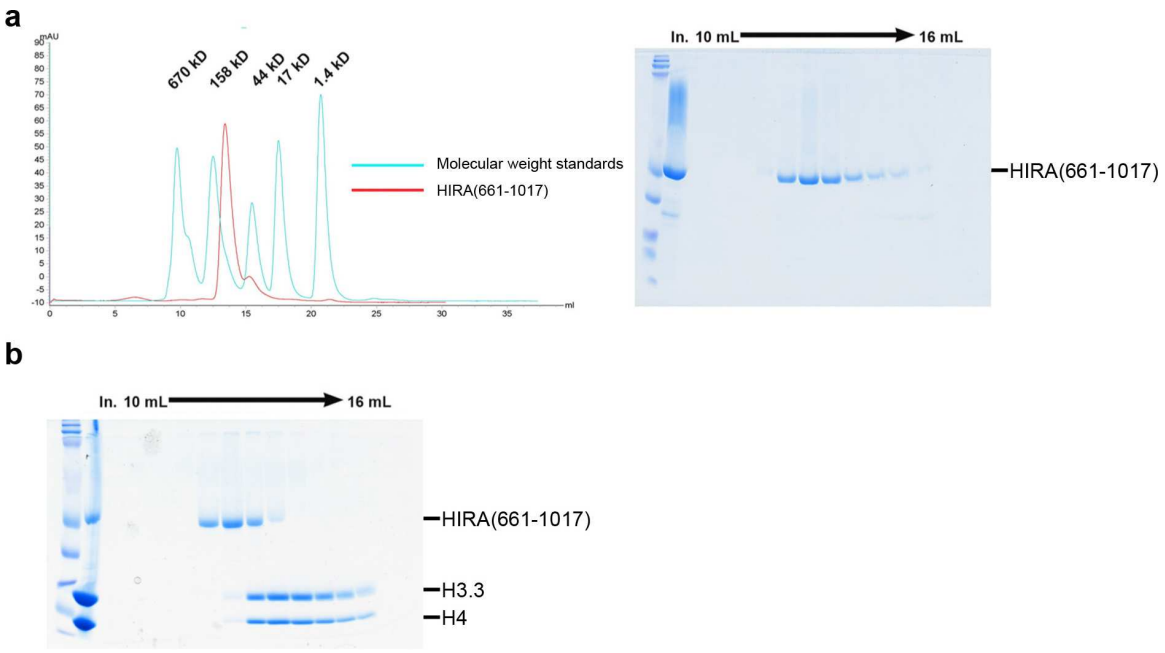


Figure 4.2 – Size-exclusion analysis of HIRA(661-1017) alone and with H3.3/H4. (A) HIRA(661-1017) elutes from an s200 10/300 GL size exclusion column in a monodisperse peak which sits between the 158 and 44 kD standard peaks. (B) HIRA(661-1017) was mixed with 2-fold excess H3.3/H4 and resolved on the s200 10/300 GL column, it appears as if no complex is formed between the proteins.

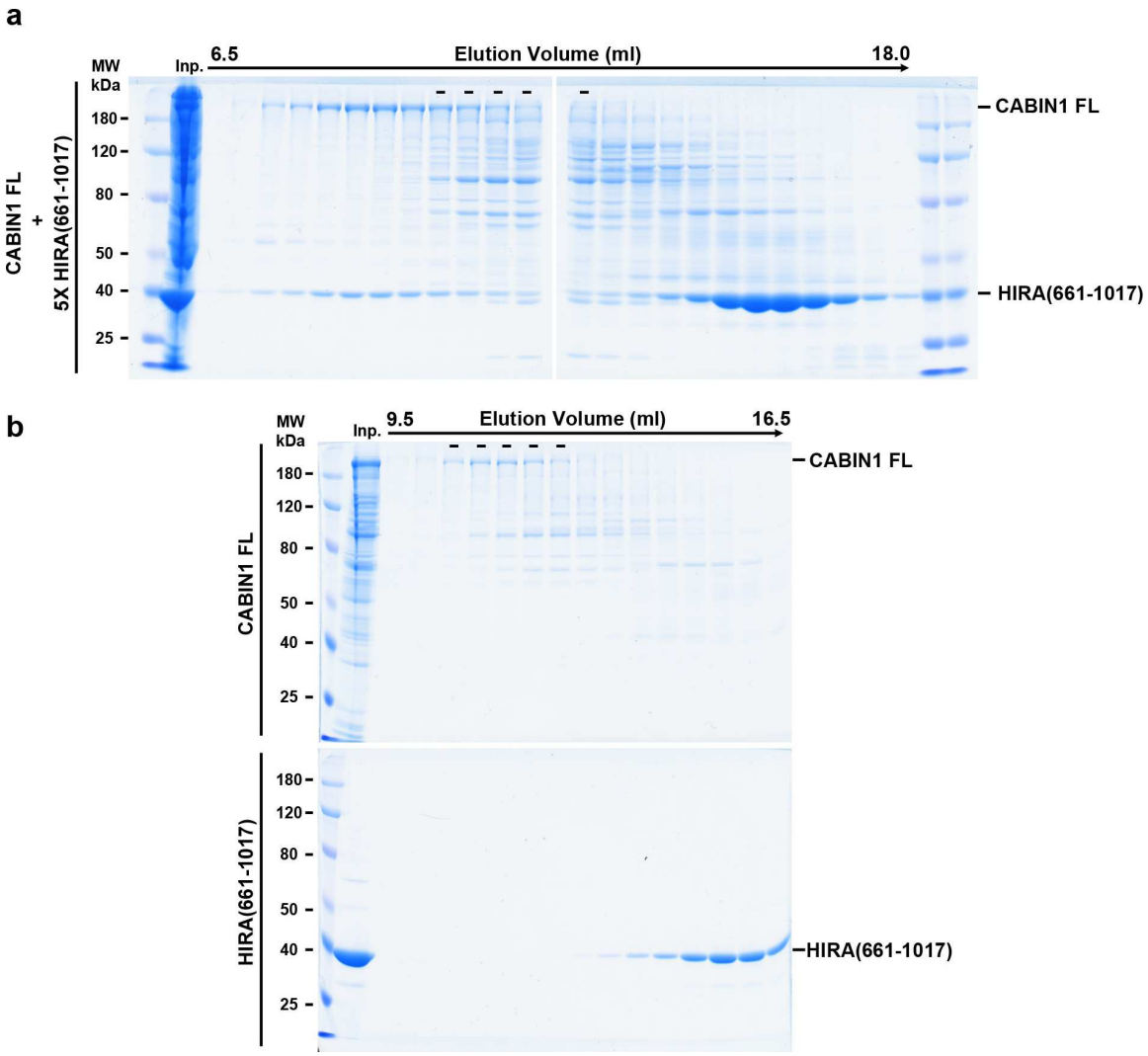
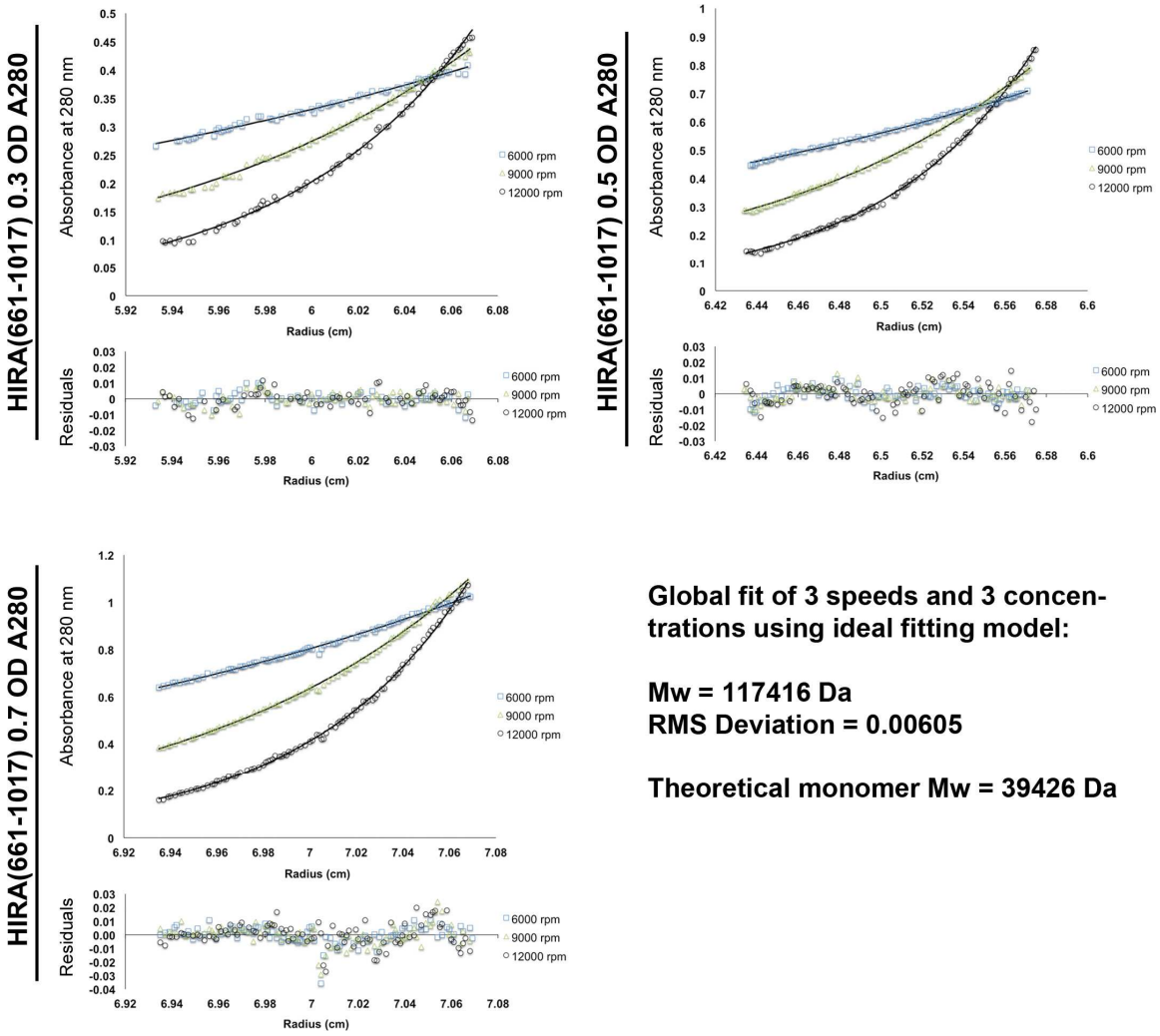


Figure 4.3 – Size-exclusion analysis of HIRA(661-1017) and CABIN1. (A) Excess HIRA(661-1017) was mixed with CABIN1 FL and resolved on a Superose 6 10/300 GL column, a HIRA/CABIN1 complex appears be co-migrating on the column. (B) CABIN1 FL and HIRA(661-1017) were resolved individually on the same Superose 6 column. Dashed lines represent where CABIN1 FL elutes individually.

We then sought to better understand the biophysical nature of the HIRA/CABIN1 complex. HIRA(661-1017), CABIN1 FL, and HIRA/CABIN1 complex that were resolved on size-exclusion chromatography. To do this, each of the resolved size-exclusion peaks were subjected to analytical ultracentrifugation to determine the molecular assembly of the proteins alone as well as the complex. Sedimentation equilibrium analysis of HIRA(661-1017) alone indicates that the species has a solution molecular weight of 117416 Da. A comparison with the theoretical molecular weight of a HIRA(661-1017) monomer indicates that this HIRA fragment exists as a stable trimer in solution (Figure 4.4). While analysis of CABIN1 alone indicates a molecular weight of 244736 Da, which in comparison with a theoretical molecular weight of 246352 Da for a monomer suggests that CABIN1 alone exists as a monomer in solution (Figure 4.5). Interestingly when we analyzed the equilibrium distribution of the HIRA/CABIN1 complex it revealed an experimental molecular weight of 605989 Da (Figure 4.6). This indicates that the complex is likely a HIRA(3)/CABIN1(2) complex, which would have a theoretical molecular weight of 610982 Da. For clear visualization of these data we have plotted representative curves for HIRA(661-1017), CABIN1 FL, and the HIRA/CABIN1 complex on a linear scale in comparison with predicted data for various molecular assemblies (Figure 4.7).



Global fit of 3 speeds and 3 concentrations using ideal fitting model:

Mw = 117416 Da
RMS Deviation = 0.00605

Theoretical monomer Mw = 39426 Da

Figure 4.4 – Analytical ultracentrifugation of HIRA(661-1017). (A) Sedimentation equilibrium was performed with HIRA(661-1017) alone, three different protein concentrations (0.3, 0.5, 0.7 OD A280) were analyzed for equilibrium distribution at three different speeds (6000, 9000, and 12000 RPM). These data were fit to the ideal fitting model using the program HeteroAnalysis. The ideal fit yielded an experimental molecular weight of 117416 Da, indicating this is population may represent a HIRA trimer.

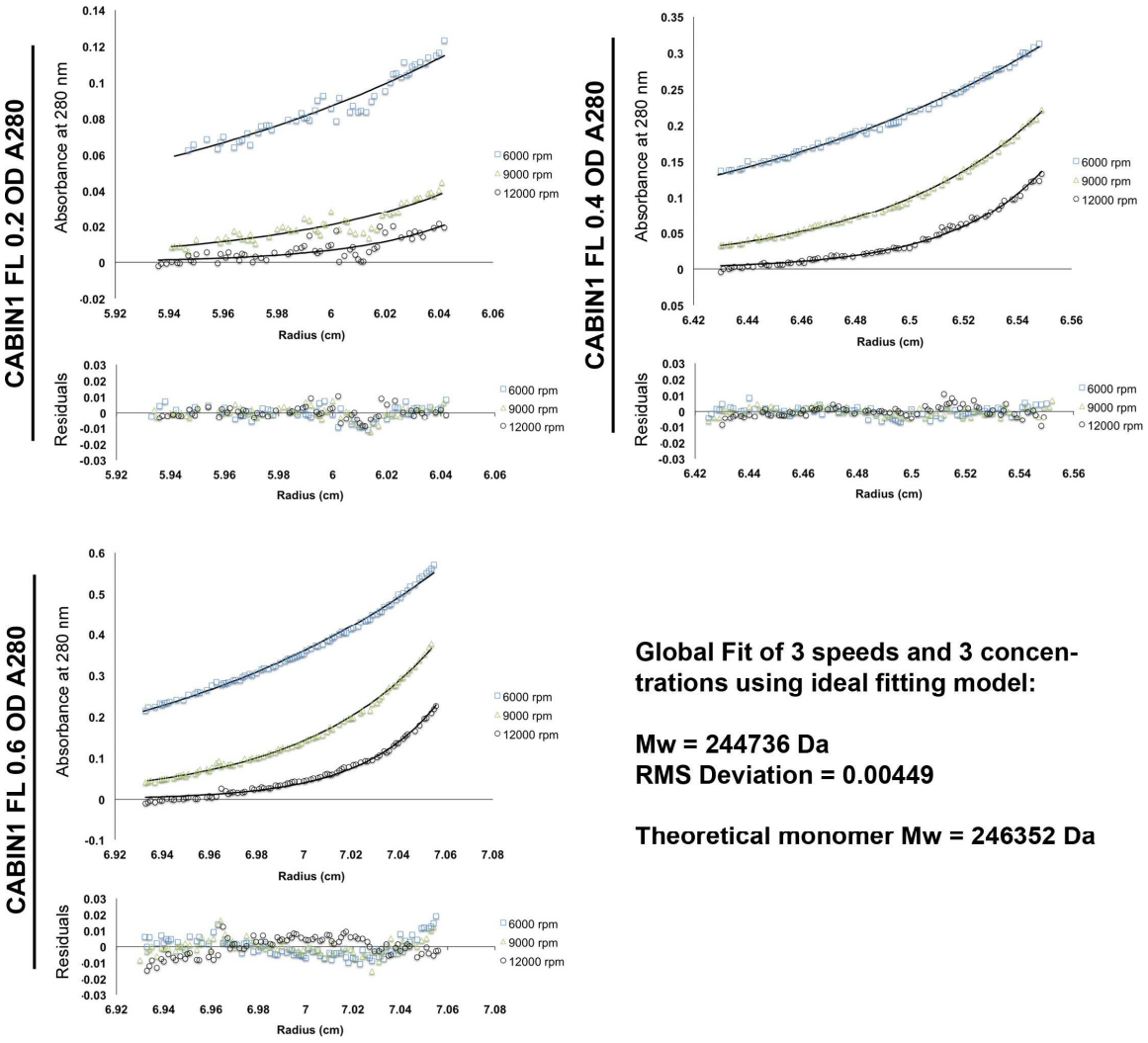


Figure 4.5 – Analytical ultracentrifugation of CABIN1. (A) Sedimentation equilibrium was performed with CABIN1 alone, three different protein concentrations (0.2, 0.4, 0.6 OD A280) were analyzed for equilibrium distribution at three different speeds (6000, 9000, and 12000 RPM). These data were fit to the ideal fitting model using the program HeteroAnalysis. The ideal fit yielded an experimental molecular weight of 244736 Da, indicating this is population may represent a CABIN1 monomer.

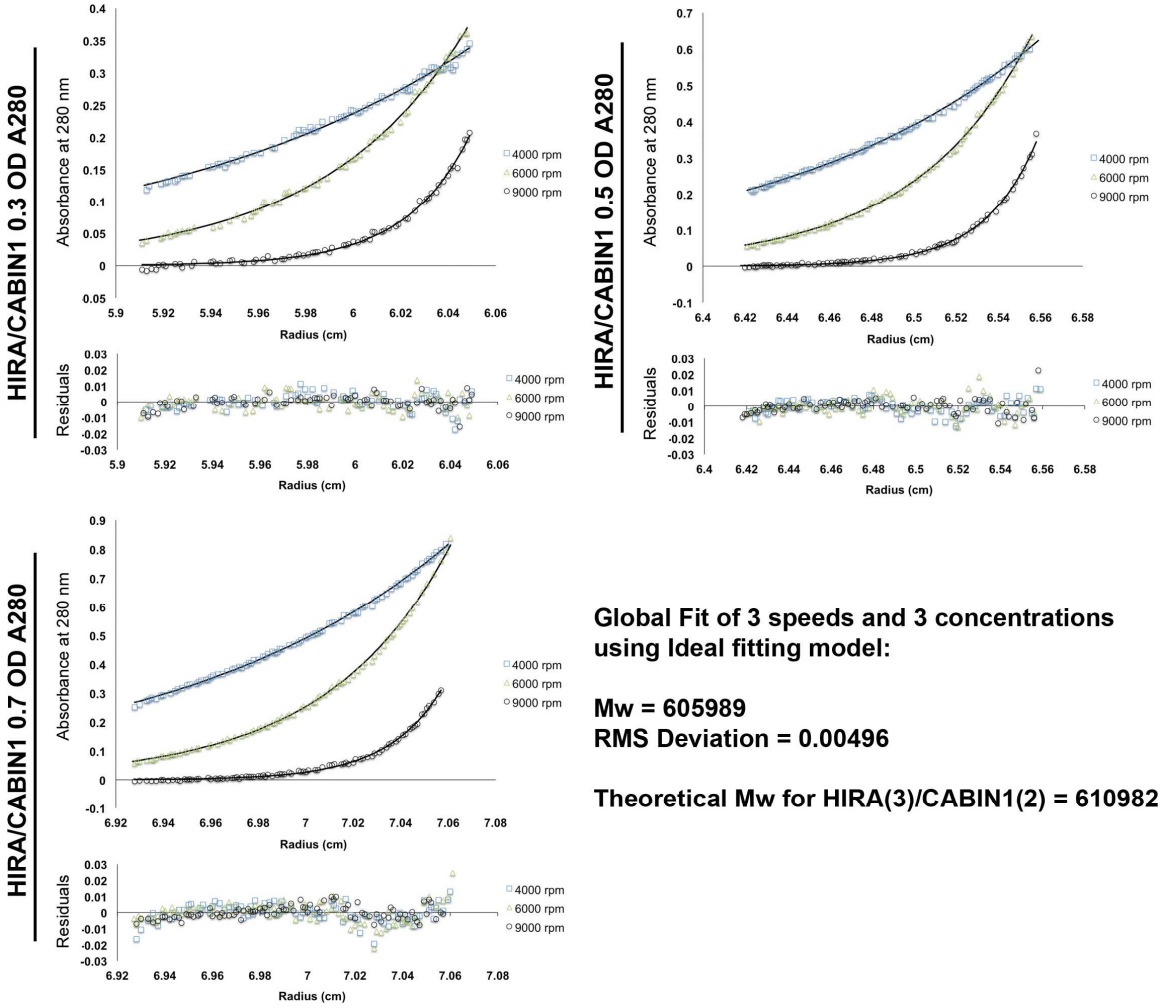


Figure 4.6 – Analytical ultracentrifugation of the HIRA/CABIN1 complex. (A) Sedimentation equilibrium was performed with the HIRA/CABIN1 complex, three different protein concentrations (0.3, 0.5, 0.7 OD A280) were analyzed for equilibrium distribution at three different speeds (4000, 6000, and 9000 RPM). These data were fit to the ideal fitting model using the program HeteroAnalysis. The ideal fit yielded an experimental molecular weight of 605989 Da, indicating this is population may represent a $HIRA_3/CABIN1_2$ complex.

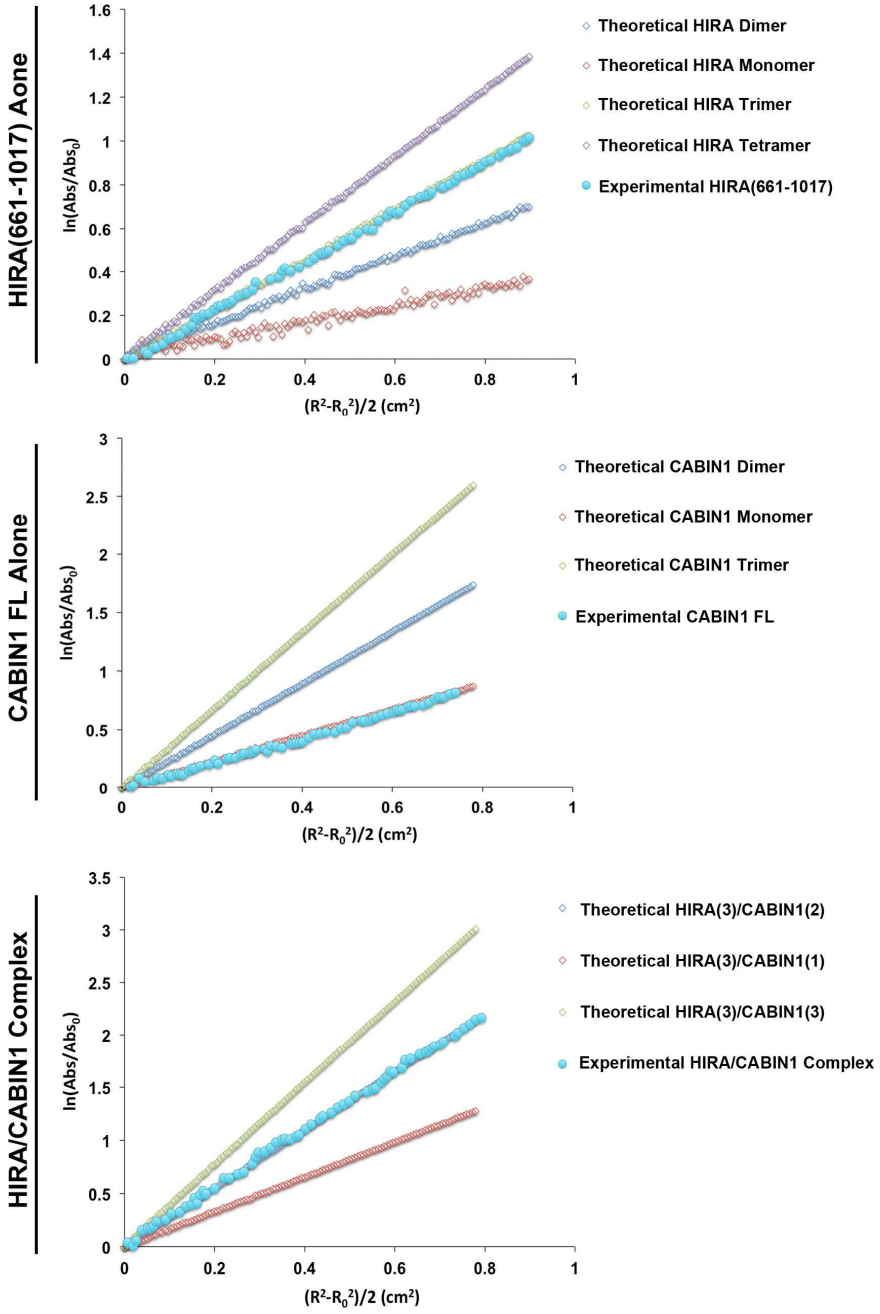


Figure 4.7 – Linear plot of HIRA/CABIN1 sedimentation equilibrium data. Representative data for HIRA, CABIN1, and the HIRA/CABIN1 complex at 0.5 OD and 9000 RPM were compared to predicted curves for threoretical monomer, dimer, and trimer for HIRA and CABIN1 alone, as well as $HIRA_3/CABIN1_1$, $HIRA_3/CABIN1_2$, $HIRA_3/CABIN1_3$. Theoretical curves were generated using the program HeteroAnalysis. To represent the data on a linear scale the curves were plotted using a logarithmic normalization.

4.3 – HIRA(661-872) and HIRA(905-1017) are tethered for trimer formation

As HIRA(661-1017) is predicted to have a strand region and a helical region, which are separated by a relatively unstructured loop (Figure 4.1), we decided to perform limited proteolysis in an attempt to potentially narrow down functional domains within residues 661-1017. HIRA(661-1017) was incubated with a 1:50 dilution of 1 mg/ml or 0.1 mg/ml Trypsin, Chymotrypsin, or Glu-C protease for either 15 minutes or 1 hour and resolved on a SDS-PAGE gel. The resulting gel indicated that both Trypsin and Chymotrypsin were able to cleave HIRA(661-1017) into two smaller fragments (Figure 4.8). Bands from the gel were analyzed by LC/MS-MS and the proteolysis products were shown to represent HIRA(661-872) and HIRA(905-1017), indicating that the loop spanning residues 873-904 is unstructured and susceptible to protease cleavage. Following this observation we decided to monitor the interaction between the separate HIRA(661-872) strand domain and HIRA(905-1017) helix domain, as HIRA(661-1017) forms a stable trimer and these two regions may potentially interact for trimer formation. HIRA(661-1017) was incubated with Trypsin and resolved on a Superdex 200 10/300 GL column in comparison with HIRA(661-872) mixed in a 1:1 ratio with HIRA(905-1017) (Figure 4.9). Both experiments show that HIRA(661-872) and HIRA(905-1017) do not co-migrate on the column, indicating that the.HIRA(661-872) strand domain and HIRA(905-1017) helix domain do not interact with each other to mediate trimer formation.

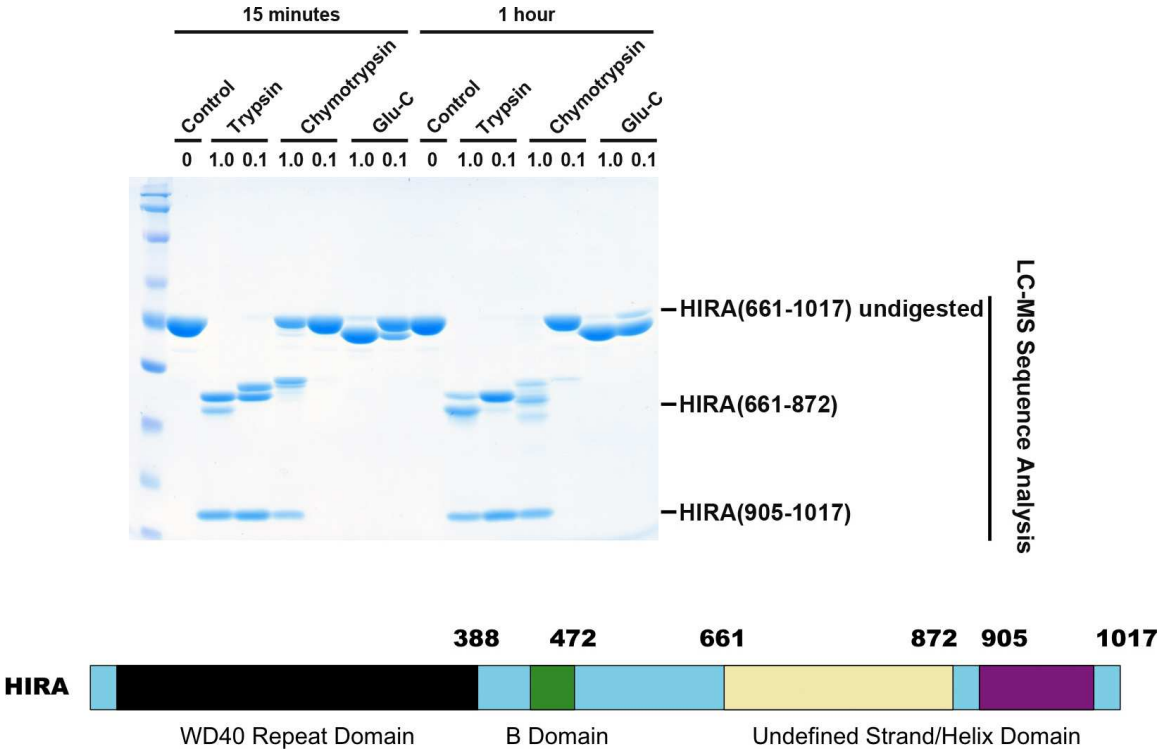


Figure 4.8 – Limited proteolysis and LC-MS analysis of HIRA(661-1017). (Top) HIRA(661-1017) was incubated with protease from either a 1.0 mg/ml or 0.1 mg/ml stock solution. Sensitivity of HIRA(661-1017) to Trypsin, Chymotrypsin, and Glu-C proteases were screened in this way, time points were taken at 15 minutes and 1 hour of incubation. LC-MS analysis was performed on the digested HIRA(661-1017) fragments, results of analysis are indicated on the proteolysis gel. (Bottom) Domain architecture for HIRA updated with results from LC-MS analysis.

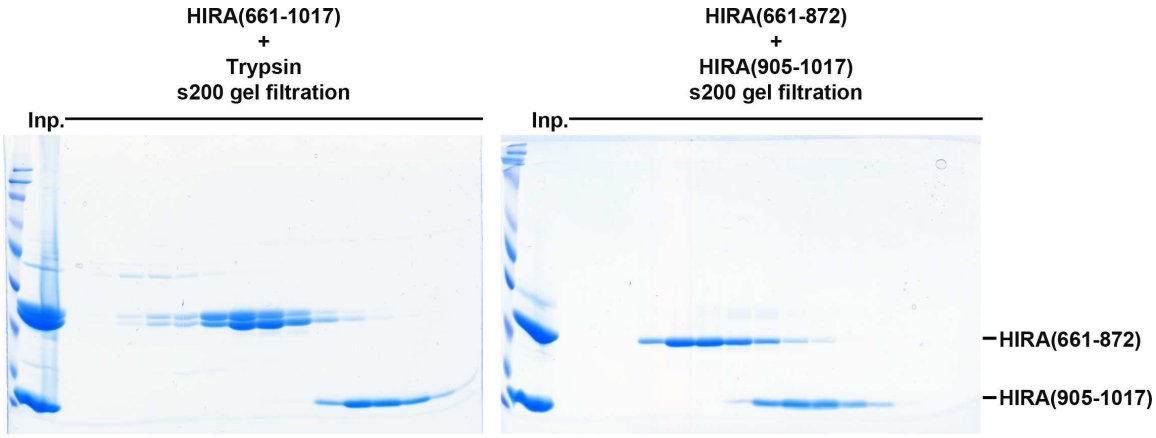


Figure 4.9 – Size-exclusion analysis of digested and mixed HIRA C-terminal fragments (Left) HIRA(661-1017) was incubated with Trypsin and resolved on an s200 10/300 GL column. (Right) HIRA(661-872) and HIRA(905-1017) were purified separately and mixed in a 1:1 ratio before resolving on an s200 10/300 GL column.

We performed sedimentation equilibrium analysis on HIRA(661-872), HIRA(905-1017), as well as an HIRA(661-1017, ∆873-904) where we replaced the HIRA 873-904 loop with a GSGSGSGSGSGS linker. The resulting data indicate that both HIRA(661-872) (Figure 4.10) and HIRA(905-1017) (Figure 4.11) are monomers in solution, while HIRA(661-1017, ∆873-904) still forms a trimer (Figure 4.12). These data suggest that the individual strand and helical domains need to be linked by the 873-904 loop to be able to form the trimer, and replacing this loop with another amino acid sequence still allows for trimer formation.

4.4 – The HIRA/CABIN1 complex is dependent on HIRA trimer formation and specific contacts from the HIRA(873-904) loop region

To further elucidate the molecular contacts underlying the HIRA/CABIN1 complex, we monitored complex formation between CABIN1 FL and HIRA(661-1017, ∆873-904), HIRA(661-872), and HIRA(905-1017) in comparison with HIRA(661-1017). CABIN1 FL was incubated with 5-fold excess HIRA and resolved on a Superdex 200 10/300 GL column. We observed that when CABIN1 FL was incubated with HIRA(661-1017, ∆873-904), HIRA(661-872), and HIRA(905-1017) there was no detectable complex formation in comparison with HIRA(661-1017) (Figure 4.13). Although it does appear as if some HIRA is migrating in the region of CABIN1 for HIRA(661-1017, ∆873-904) and HIRA(661-872), there does not appear to be complex formation, as the CABIN1 FL band had not undergone an increased molecular weight shift. The bands for

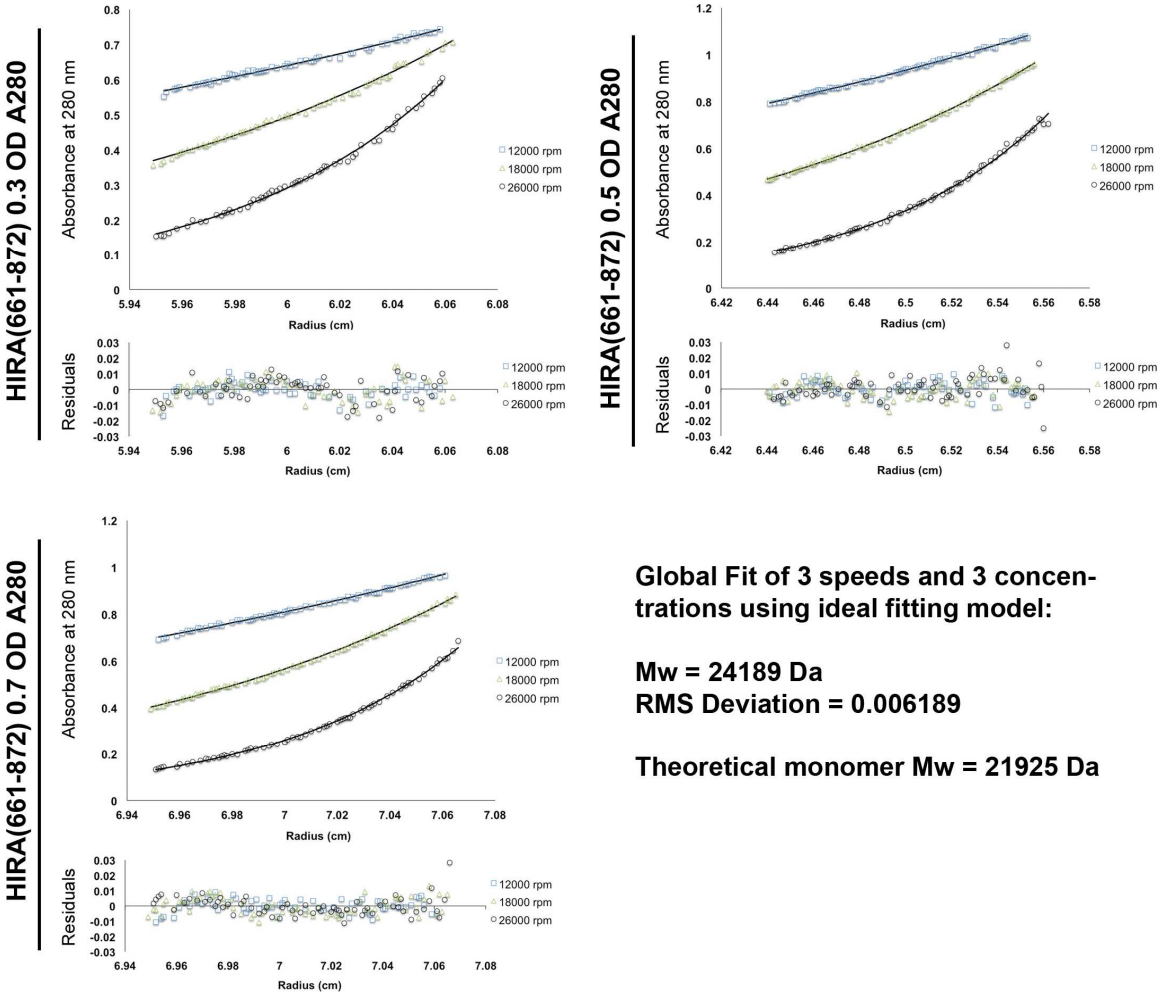


Figure 4.10 – Analytical ultracentrifugation of HIRA(661-872). (A) Sedimentation equilibrium was performed with HIRA(661-872) alone, three different protein concentrations (0.3, 0.5, 0.7 OD A280) were analyzed for equilibrium distribution at three different speeds (12000, 18000, and 26000 RPM). These data were fit to the ideal fitting model using the program HeteroAnalysis. The ideal fit yielded an experimental molecular weight of 24189 Da, indicating this is population may represent a HIRA(661-872) monomer.

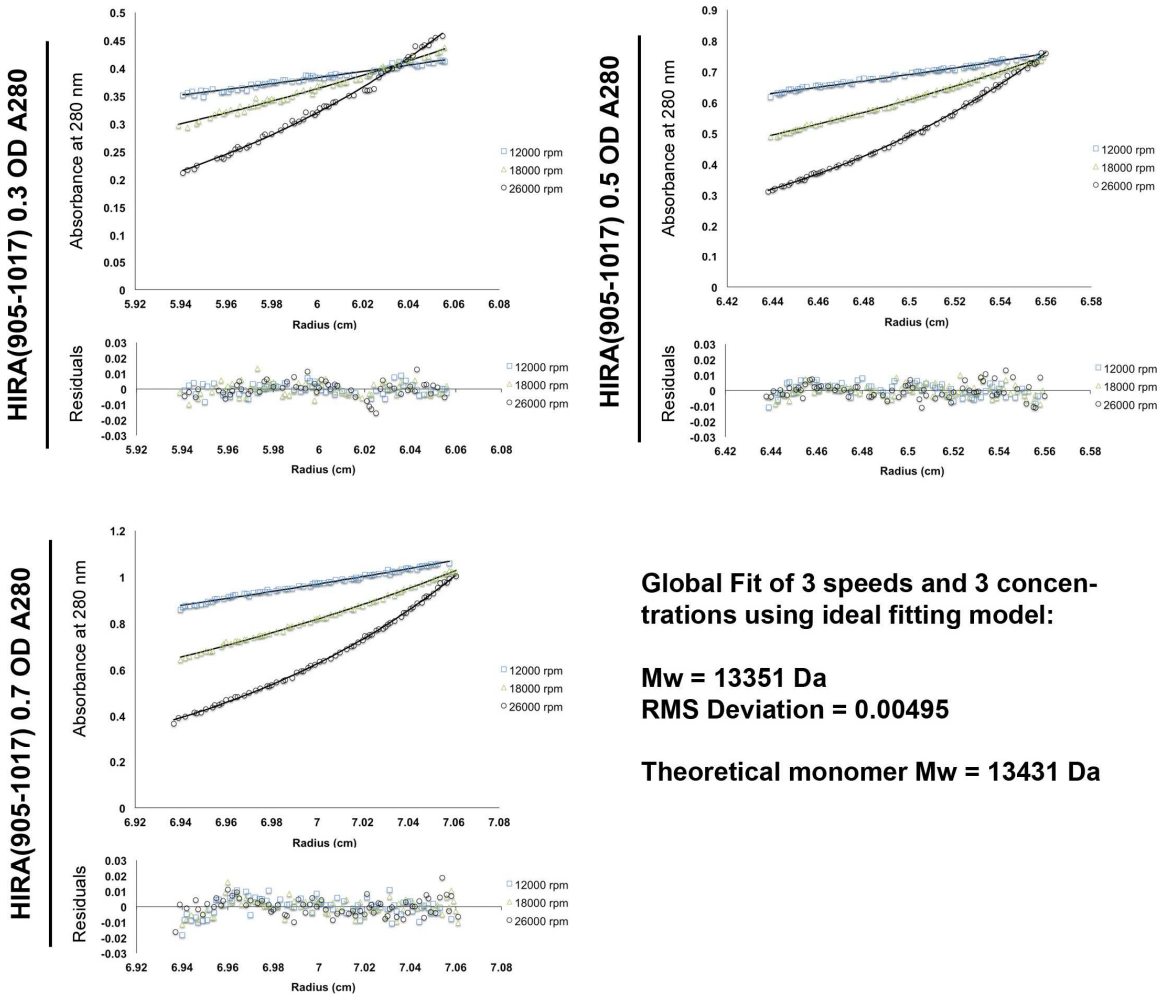


Figure 4.11 – Analytical ultracentrifugation of HIRA(905-1017). (A) Sedimentation equilibrium was performed with HIRA(905-1017) alone, three different protein concentrations (0.3, 0.5, 0.7 OD A280) were analyzed for equilibrium distribution at three different speeds (12000, 18000, and 26000 RPM). These data were fit to the ideal fitting model using the program HeteroAnalysis. The ideal fit yielded an experimental molecular weight of 13351 Da, indicating this is population may represent a HIRA(905-1017) monomer.

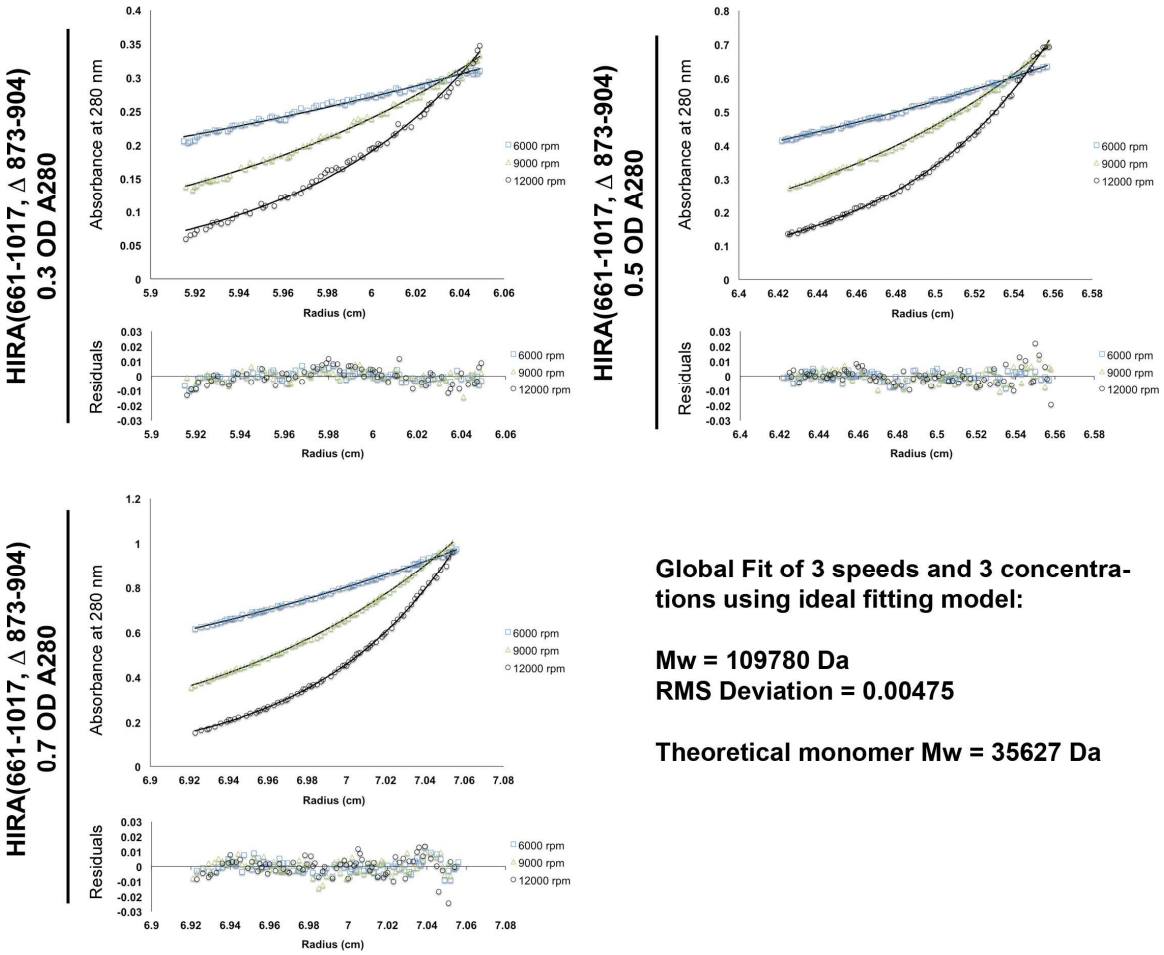


Figure 4.12 – Analytical ultracentrifugation of HIRA(661-1017, Δ873-904). (A) Sedimentation equilibrium was performed with HIRA(661-1017, Δ873-904) alone, three different protein concentrations (0.3, 0.5, 0.7 OD A280) were analyzed for equilibrium distribution at three different speeds (6000, 9000, and 12000 RPM). These data were fit to the ideal fitting model using the program HeteroAnalysis. The ideal fit yielded an experimental molecular weight of 109780 Da, indicating this is population may represent a HIRA(661-1017, Δ873-904) trimer.

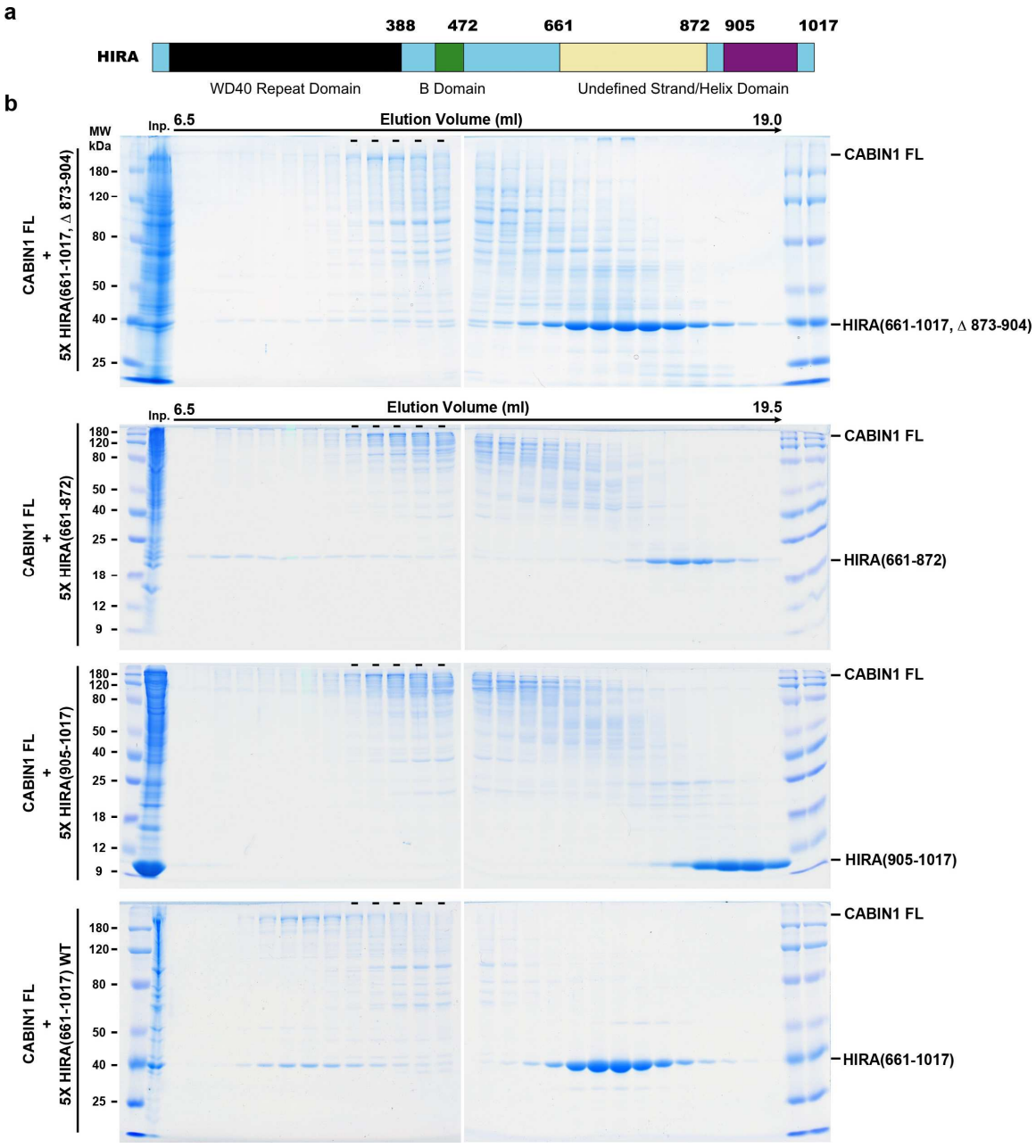


Figure 4.13 – Size-exclusion analysis of HIRA C-terminal fragments with CABIN1. (A) Domain architecture for HIRA. (B) Size-exclusion analysis of complex formation between CABIN1 FL and HIRA(661-1017) WT, HIRA(661-1017, Δ873-904), HIRA(661-872), and HIRA(905-1017). For all runs 5X excess HIRA was incubated with CABIN1 FL and resolved on a Superose 6 10/300 GL column. Dashed lines represent where CABIN1 FL elutes individually

HIRA(661-1017, Δ873-904) and HIRA(661-872) in the vicinity of CABIN1 are also present when these fragments are resolved on the column in absence of CABIN1 FL, indicating that these bands represent a small amount of HIRA aggregation, not CABIN1 interaction (Figure 4.14).

As we observed no complex formation for CABIN1 FL with HIRA(661-1017, Δ873-904), HIRA(661-872), and HIRA(905-1017), we hypothesized that CABIN1 may be binding to HIRA in the region of the 873-904 loop. To test this we decided to monitor CABIN1 interaction with HIRA(873-1017) to see if the 873-904 loop alone is sufficient for CABIN1 binding. We also attempted to make HIRA(661-904) for this experiment but this HIRA fragment turned out to be insoluble and not suitable for biophysical studies (data not shown). We incubated HIRA(873-1017) with CABIN1 FL and resolved the mixture on a Superdex 200 10/300 GL column in comparison with HIRA(661-1017) (Figure 4.15). We observed that HIRA(873-1017) was not able to form a complex with CABIN1. This indicates that while the presence of the HIRA(873-904) loop is important for CABIN1 binding, the interaction is likely dependent on the HIRA trimer in combination with the 873-904 loop.

Based on our observation that HIRA(661-1017, Δ873-904) does not bind CABIN1, our collaborators decided to generate some combined alanine point mutants in the loop region. They tested interaction of HIRA FL bearing the F870A_R871A_L874A mutations with CABIN1 FL and observed that this triple

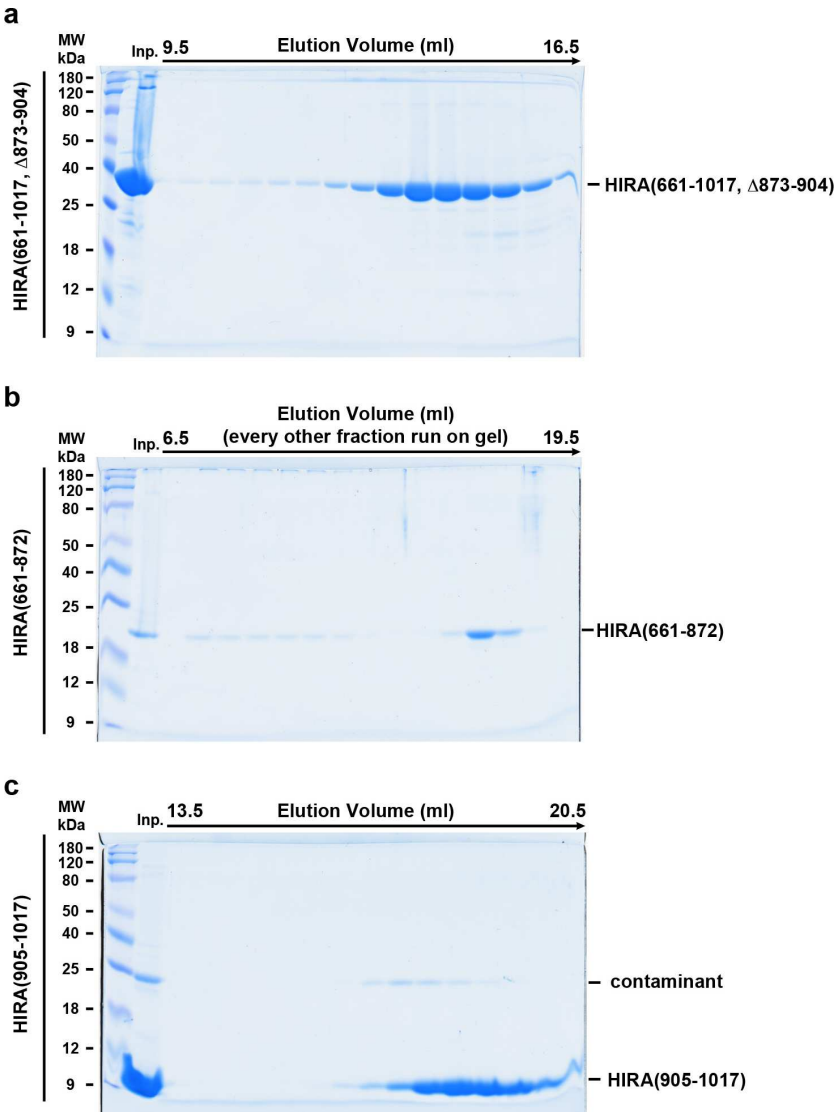


Figure 4.14 – Size-exclusion analysis of HIRA C-terminal fragments. (A) Size-exclusion analysis of HIRA(661-1017, Δ873-904) alone. (B) Size-exclusion analysis of HIRA(661-872) alone. (C) Size-exclusion analysis HIRA(905-1017) alone. All runs were performed on a Superose 6 10/300 GL column.

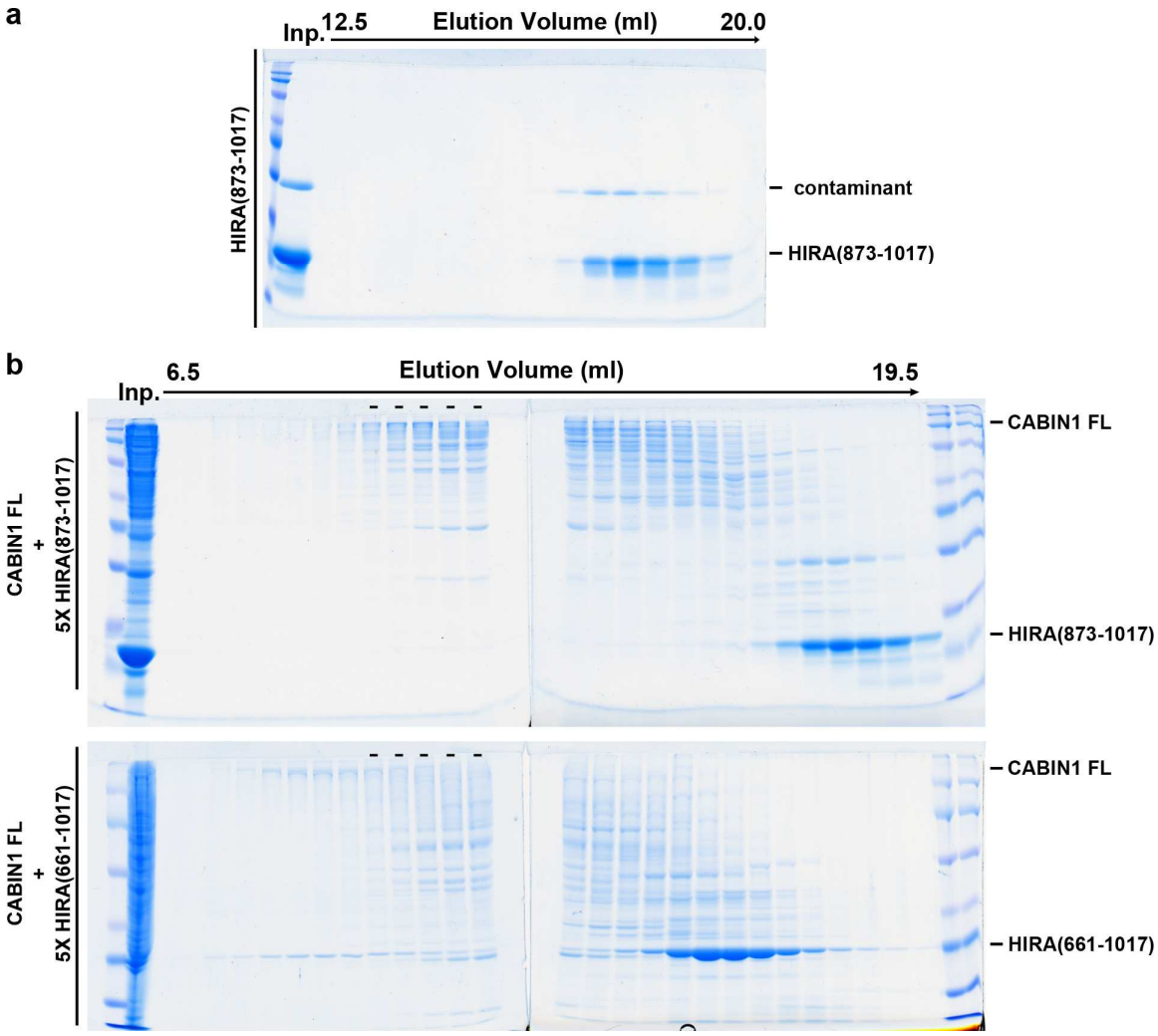


Figure 4.15 – Size-exclusion analysis of HIRA(873-1017) with CABIN1. (A) Size-exclusion analysis of HIRA(873-1017) alone. (B) Size-exclusion analysis of HIRA(873-1017) with CABIN1 (Top), in comparison with the complex formed between HIRA(661-1017) and CABIN1 (Bottom). All runs were performed on a Superose 6 10/300 GL column. Dashed lines represent where CABIN1 FL elutes individually

mutant shows decreased CABIN1 binding similarly to HIRA FL with the 873-904 loop replaced with the GSGSGSGSGSGS linker (data not shown). We then introduced the same F870A_R871A_L874A triple mutant into HIRA(661-1017) and monitored complex formation with CABIN in comparison with HIRA(661-1017) WT (Figure 4.16). In our hands the HIRA F870A_R871A_L874A triple mutant showed reduced complex formation with CABIN1, but not a complete abrogation of the HIRA/CABIN1 interaction. In comparison with CABIN1 FL run alone we observe a very significant larger molecular weight shift with the addition of HIRA(661-1017). This shift is not as evident when CABIN1 FL is mixed with HIRA(661-1017) F870A_R871A_L874A, although there is still complex formation, indicating that the triple mutant decreases but does not ablate CABIN1 interaction.

4.5 – HIRA may be structurally homologous to Ctf4

Concurrently with our biochemical investigation into the HIRA/CABIN1 interaction, we also pursued the X-ray crystal structure of the HIRA trimer alone and in complex withCABIN1. Several rounds of crystal screens were set up for HIRA(661-1017), HIRA(661-872), HIRA(905-1017), HIRA(661-1017, Δ873-904), and the CABIN1/HIRA(661-1017) complex. From all of our attempts we have only been able to prepare crystals of HIRA(661-872) (Figure 4.17a). These crystals diffract to about 3.5Å (Figure 4.17b), but are sensitive to radiation and have relatively low symmetry with a P21 space group. We have collected some

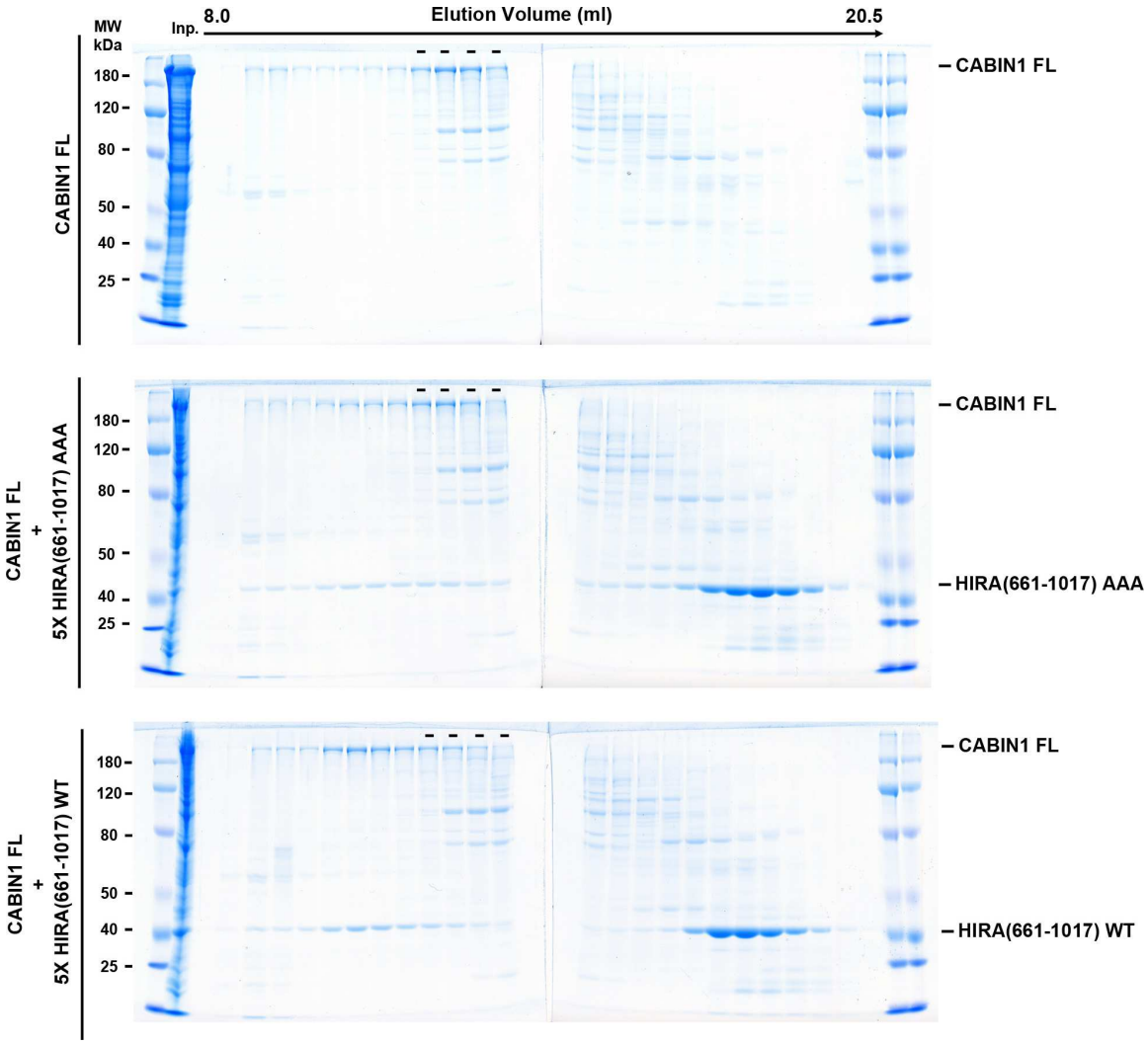


Figure 4.16 – Size-exclusion analysis of HIRA(661-1017, F870A_R871A_L874A) with CABIN1. Size-exclusion analysis of CABIN1 FL alone (Top), HIRA(661-1017, F870A_R871A_L874A) with CABIN1, and HIRA(661-1017) WT with CABIN1 (bottom). All runs were performed on a Superose 6 10/300 GL column. Dashed lines represent where CABIN1 FL elutes individually.

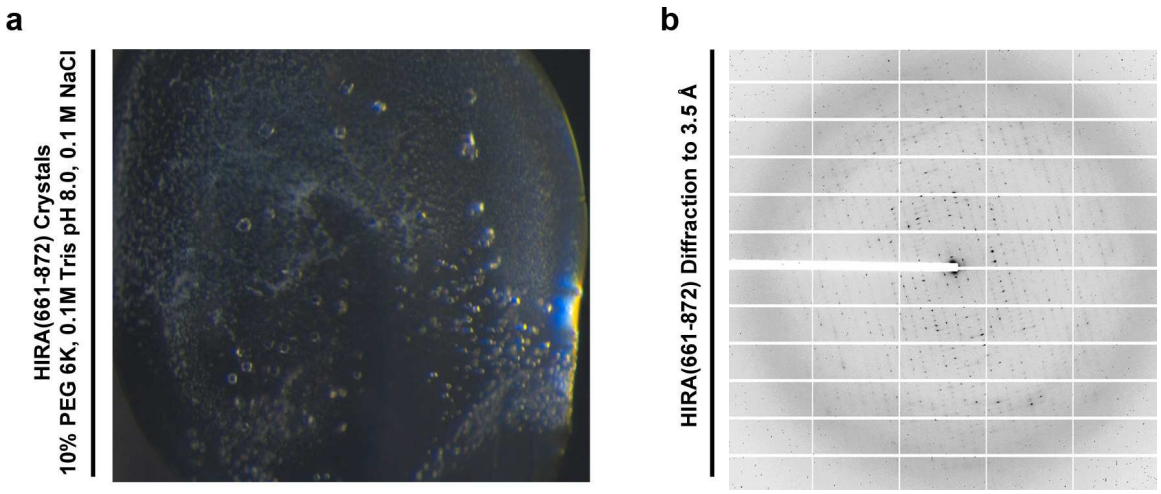


Figure 4.17 – Crystallization and X-ray diffraction of HIRA(661-872). (A) Crystals formed for HIRA(661-872) in a solution of 10% PEG 6K, 0.1M Tris pH 8.0, and 0.1M NaCl. (B) Representative diffraction frame for HIRA(661-872) crystal, data goes to roughly 3.5 Å, diffraction data collected at the Advanced Photon Source (APS).

datasets for selenomethionine and platinum derived crystals that diffract to about 3.5Å or 4.0Å, but have yet to determine the structure.

While attempting to determine the HIRA(661-872) structure, we submitted the sequence for HIRA(661-1017) to the Protein Homology/analogY Recognition Engine (Phyre) for structure prediction in attempt to find a molecular replacement search model for the crystallographic data. While the Phyre model did not work for molecular replacement, it did help us identify a previously uncharacterized structural homolog for HIRA. The top hit in the Phyre search was a protein called Ctf4, which has function in linking the Cdc45-MCM-GINS (CMG) helicase to DNA polymerase alpha, along with other proteins that have been found to harbor a CIP-box (Ctf4-interacting-peptide) motif(Villa et al., 2016). Interestingly the C-terminal region of Ctf4 is also separated into strand and helix domains which interact to form a homotrimer (Figure 4.18a) (Simon et al., 2014; Villa et al., 2016). This is noteworthy, as we have identified another strand/helix trimer forming domain, which may be similar to HIRA(661-1017). Ctf4 also harbors an N-terminal WD40 domain similar to HIRA, further indicating that these two proteins may be distantly related structural homologs (Figure 4.18b). Ctf4 is involved in linking several proteins to the CMG complex, containing the histone chaperone MCM2, which is involved in H3/H4 shuttling during DNA replication. Based on this relation to Ctf4 we hypothesize that HIRA may interact with RNA

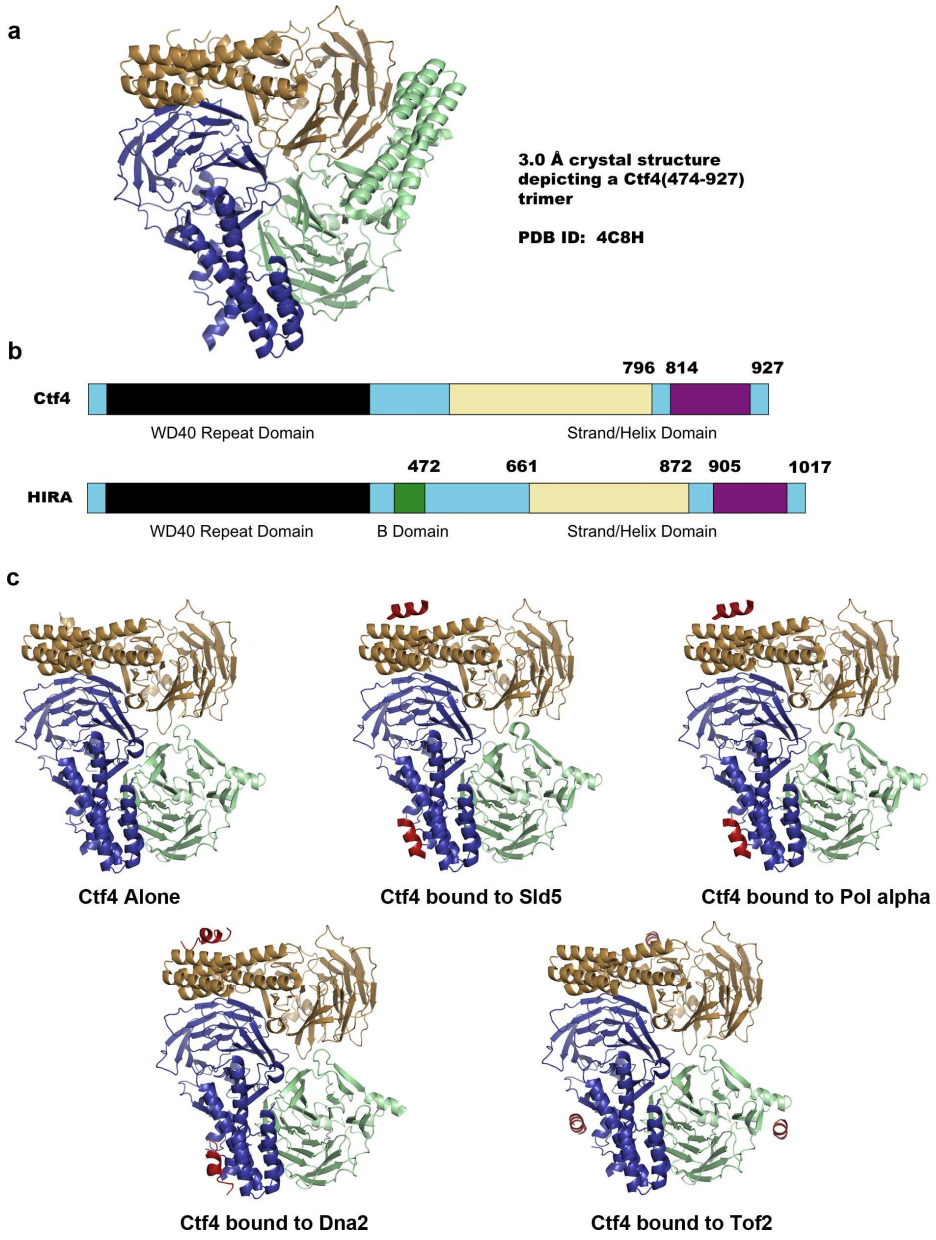


Figure 4.18 – The HIRA C-terminal strand/helix domain may form a trimer similar to Ctf4. (A) 3.0 Å crystal structure for Ctf4(474-927) depicting the formation of a homotrimer. (B) Comparison of domain architecture for HIRA and Ctf4 representing the similarity in the domain organization between the two proteins. (C) Remaining Ctf4 crystal structures representing data for the three strand but only two helical domain conformation, and Ctf4(3)/Peptide(2) binding stoichiometry for Sld5, Pol alpha, and Dna2 peptides.

polymerase and potentially other proteins in a similar manner during gene transcription, although additional studies are needed to confirm this prediction.

Of particular interest is that of the six crystal structures deposited into the PDB for the Ctf4 C-terminal trimer, only one structure has density for all three helical domains (4C8H) (Figure 4.18a). All five remaining structures have three strand domains, but only two helical domains resolved (5HOG, 5HOI, 4C8S, 4C93, 4C95) (Figure 4.18c). One of these is of the Ctf4 trimer alone (4C8S), while the other four depict a Ctf4 trimer associated with Ctf4-interacting peptides from either Pol alpha, Sld5, Dna2, or Tof2. The structures for Ctf4 bound to Pol alpha, Sld5, and Dna2 have a Ctf4(3)/Peptide(2) stoichiometry with the peptide associated with the helical region of Ctf4, while the structure for Ctf4 bound to Tof2 has a Ctf4(3)/Peptide(3) stoichiometry and the Tof2 peptide is bound to the cleft between the strand and helical regions of Ctf4. We have identified that the loop separating the stand and helical domains for HIRA is essential for CABIN1 interaction, the HIRA(3)/CABIN1(2) stoichiometry we have observed may be related to how Ctf4 interacts with some of its binding partners using a 3:2 stoichiometry.

4.6 – Hir1/Hir2 from *S. cerevisiae* associate to form a heterotrimer

An interesting note on the observed HIRA(3)/CABIN1(2) stoichiometry may be gleaned from the Hir complex in *S. cerevisiae*, where two proteins called Hir1 and Hir2 work together to form a functional HIRA unit. It has been reported that

the Hir complex harbors two copies of Hir2 and one copy of Hir1(Prochasson et al., 2005). Both Hir1 and Hir2 are homologous to HIRA residues 1-858; the Hir1 C-terminus corresponds to HIRA residue 1-858, while only Hir2 harbors sequence homologous to HIRA residues 859-1017 harboring both the loop and the helical regions (Figure 4.19). It is possible that the HIRA(3)/CABIN1(2) stoichiometry is an evolutionary holdover from yeast where the complex did not have three binding sites for the CABIN1 homolog Hir3. To address this hypothesis we prepared the representative C-terminal regions for Hir2 alone and a complex of Hir1/Hir2 from co-expression in *E. coli* (Figure 4.20a). We were unable to produce soluble Hir1 when it was expressed alone. We used sedimentation equilibrium to analyze the stoichiometry of our Hir2 and Hir1/Hir2 complex and showed that Hir2 alone exists as a dimer, and forms a heterotrimer when co-expressed with Hir1, presumably with a stoichiometry of Hir2(2)/Hir1(1) (Figure 4.20b). We attempted to prepare a complex between Hir1/Hir2 and Hir3, although we were unable to prepare soluble Hir3 in bacteria and insect cells. With evidence from Ctf4 and the yeast Hir complex, the HIRA trimer and the HIRA(3)/CABIN1(2) stoichiometry likely represent biologically relevant confirmations.

4.7 – Discussion

Although our studies into HIRA timer formation and the HIRA/CABIN1 complex are incomplete, we have been able to characterize several previously unknown

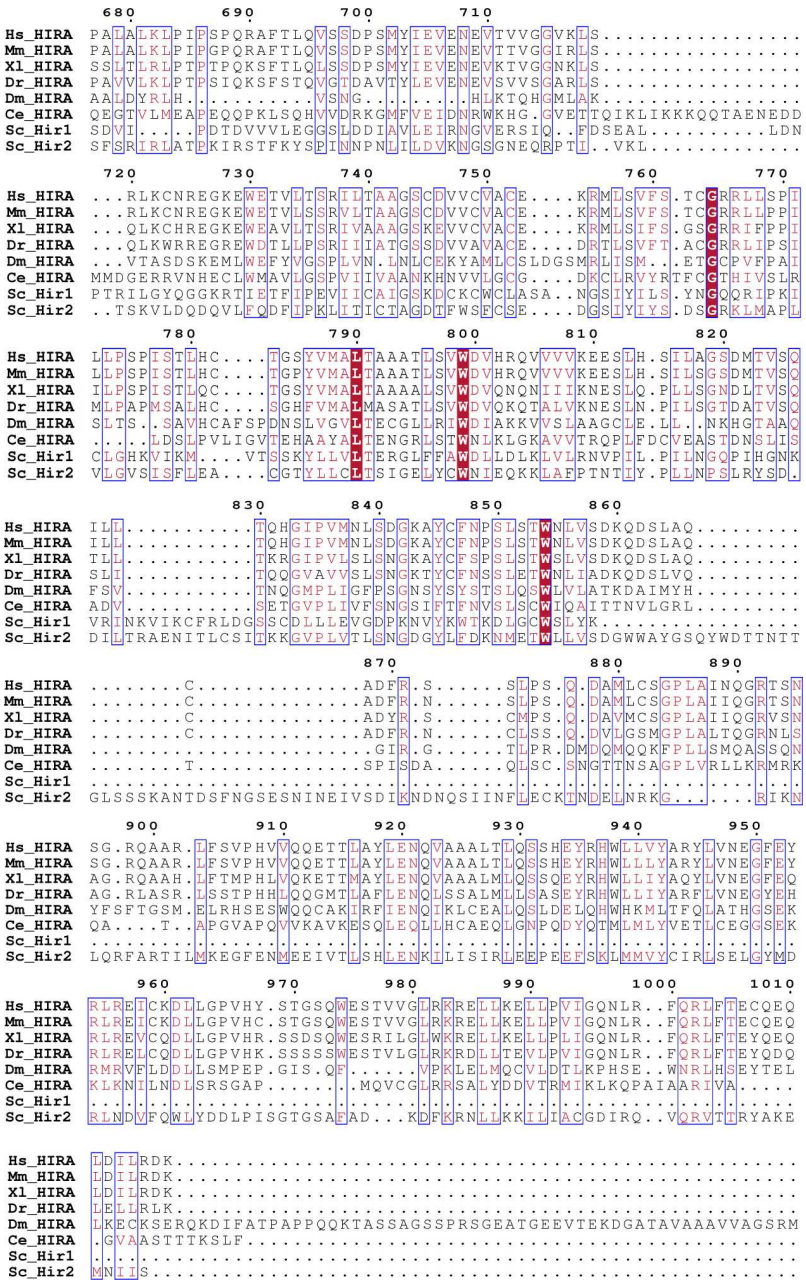


Figure 4.19 – Alignment of HIRA homologs. Multiple sequence alignment showing conservation among *Homo sapiens* HIRA, *Mus musculus* HIRA, *Xenopus laevis* HIRA, *Danio rerio* HIRA, *Drosophila melanogaster* HIRA, *Caenorhabditis elegans* HIRA, *Saccharomyces cerevisiae* Hir1, and *Saccharomyces cerevisiae* Hir2. Alignment contains residues corresponding to human HIRA residues 677-1017.

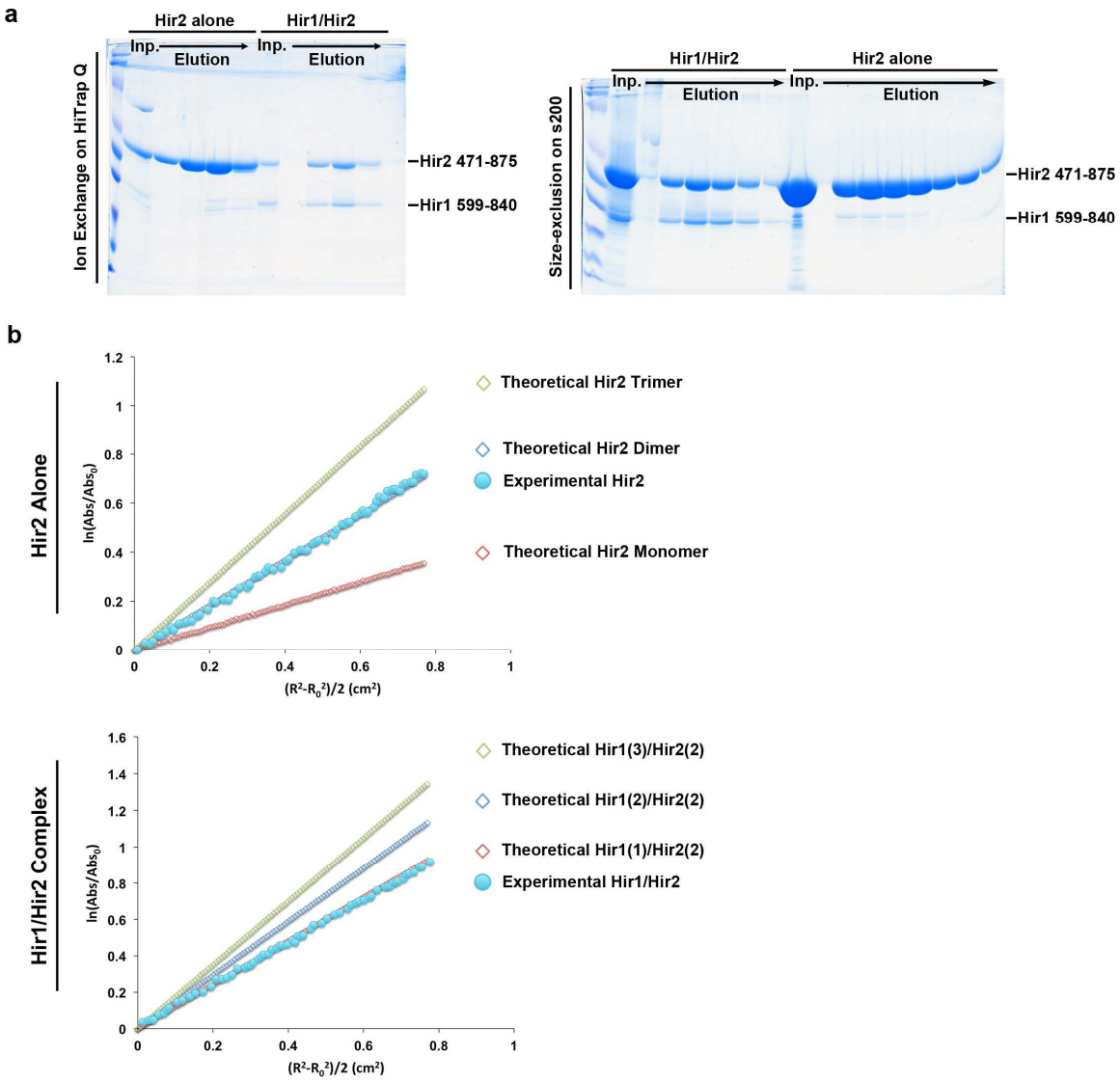


Figure 4.20 – Preparation of the Hir1/Hir2 complex. (A) Ion exchange and size-exclusion purification of the C-terminal regions of Hir2 alone and the Hir1/Hir2 complex. (B) Sedimentation equilibrium analysis of Hir2 alone and the Hir1/Hir2 complex, experimental data in comparison with predicted theoretical data.

aspects of how the HIRA C-terminal domain functions to form a homotrimer which is essential for binding with CABIN1. We have shown that the HIRA(661-1017) is organized into separate strand (HIRA(661-872)) and helical (HIRA(905-1017)) domains which are tethered by a loop spanning HIRA residues 873-904. We have shown that the strand and helical domains must be tethered by a linker for proper trimer formation, although the sequence of the linker is not important for formation of the trimer. Additionally we have shown that formation of the HIRA/CABIN1 complex is dependent on the HIRA trimer as well as the proper sequence of the HIRA(873-904) loop, as deletion or mutagenesis of the loop decreases formation of the complex. Finally we have shown that HIRA may be structurally homologous to the protein Ctf4, which has very similar domain organization to HIRA and has also been shown to form a homotrimer.

A crystal structure for the HIRA C-terminal region would help us further understand the nature of the HIRA trimer and to determine if it does indeed form a structure similar to Ctf4. While we have focused on HIRA(661-872) as we have been able to produce crystals for this fragment, it may not yield structural information into the molecular contacts required for trimer formation, as HIRA(661-872) exists as a monomer in solution. We plan to focus on further crystal screening for HIRA(661-1017) as well as HIRA(661-1017, Δ873-904) and potentially other variants of the loop deletion as there is no density for the homologous loop spanning residues 796-814 in Ctf4 (Simon et al., 2014). We

have reinforced our data demonstrating the HIRA trimer and HIRA(3)/CABIN1(2) stoichiometry with biological evidence from Ctf4 and the yeast Hir complex, although it is unclear why this complex would need to form a trimer to carry out the function of depositing an $(H3.3/H4)_2$ tetramer into chromatin. As we have discussed extensively in chapter 3, UBN1 may form a dimer to allow formation of an $(H3.3/H4)_2$ heterotetramer from two ASF1a bound H3.3/H4 heterodimers. If we combine the idea of two UBN1 molecules bound to a $(H3.3/H4)_2$ tetramer, associated with a trimer of HIRA, it is possible that a HIRA complex with a third UBN1 molecule harboring an ASF1a/H3.3/H4 unit is held in reserve to be transferred into the next $(H3.3/H4)_2$ tetramer to be formed for deposition (Figure 4.21). This idea of a histone reservoir may be difficult to test biochemically, although we plan to use an electron microscopy based approach to investigate the structural confirmation of the full-length HIRA complex. We believe these structural studies may help confirm the molecular assembly of the full-length HIRA complex associated with histones and ASF1a. Although further studies are needed to elaborate on the observations we have presented here. Our investigation has characterized the previously unknown HIRA trimer, and further elucidated the molecular contacts underlying the HIRA/CABIN1 complex.

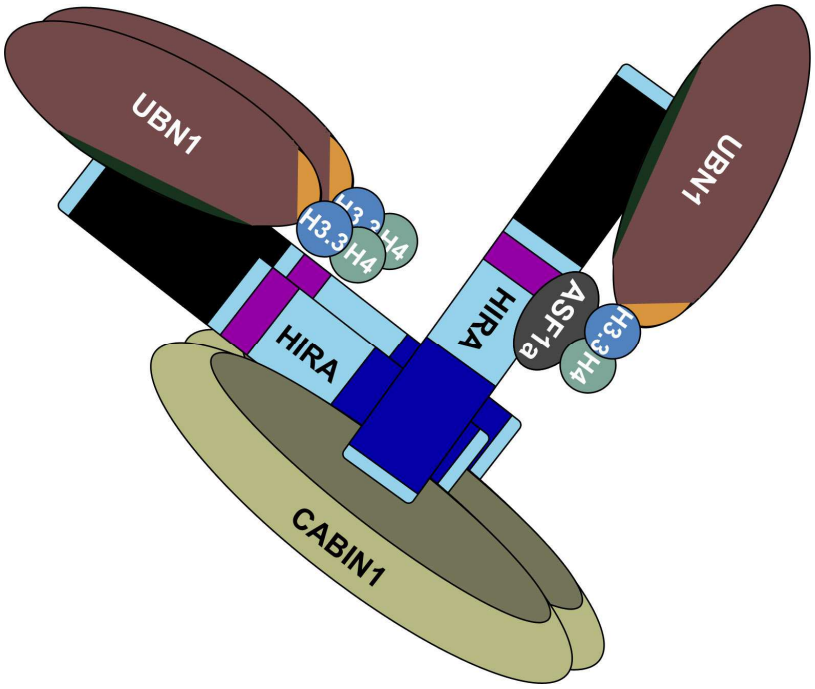


Figure 4.21 – Model for molecular assembly of the HIRA complex. Depiction of the HIRA complex showing a HIRA(3)/UBN1(3)/CABIN1(2) stoichiometry. The UBN1 NHRD is shown in green while the orange domain represents the UBN1 HRD and middle domain as they both may contribute to $(H3.3/H4)_2$ tetramer binding.

4.8 – Materials and methods

Multiple sequence alignments. Alignment for Figures 4.19 was generated using Clustal Omega (Sievers et al., 2011), and formatted for better visualization using ESPript version 3.0 (Gouet et al., 2003).

Protein secondary structure prediction. Predicted secondary structure of HIRA depicted in Figure 4.1 was generated using the PSIPRED secondary structure prediction server (Buchan et al., 2013).

Generation of expression plasmids. The plasmids encoding GST-HIRA(661-1017), GST-HIRA(661-872), GST-HIRA(905-1017), and GST-HIRA(873-1017) and were generated by PCR amplification from a FL HIRA construct (Banumathy et al., 2009; Tang et al., 2012) and ligation into the EcoRI/XhoI sites of a custom engineered pRSFduet-1 (Novagen) *Escherichia coli* expression vector carrying a N-terminal GST tag that is removable by TEV protease cleavage. The plasmids encoding GST-HIRA(661-1017, D873-904) and GST-HIRA(661-1017) F870A_R871A_L874A□were generated by site-directed mutagenesis(Weiner and Costa, 1994). The plasmids encoding FL H3.1, H3.3, and H4 were generated as described previously (Ricketts et al., 2015). The plasmid encoding FL His-CABIN1 was generated as described previously (Rai et al., 2011; Ricketts et al., 2015; Tang et al., 2012). The plasmid encoding GST-Hir1(599-840) was generated by PCR amplification and ligation into the EcoRI/XhoI sites of a custom engineered pRSFduet-1 (Novagen) *Escherichia coli* expression vector

carrying a N-terminal GST tag that is removable by TEV protease cleavage. The plasmid encoding His-Hir2(471-875) was generated by PCR amplification and ligation into the EcoRI/XhoI sites of a custom engineered pETduet-1 (Novagen) *Escherichia coli* expression vector carrying a N-terminal His tag that is removable by TEV protease cleavage.

Protein expression and purification. H3.3 and H4 individual histone proteins were expressed to the inclusion bodies in BL21-Gold(DE3) cells (Agilent). The individual histone proteins were purified and refolded to form H3.3/H4 as previously described (Luger et al., 1999; Ricketts et al., 2015). GST-HIRA(661-1017), GST-HIRA(661-872), GST-HIRA(905-1017), GST-HIRA(873-1017), GST-HIRA(661-1017, D873-904) and GST-HIRA(661-1017) F870A_R871A_L874A were generated in BL21-Gold(DE3) cells (Agilent) induced with 0.8 mM IPTG and expressed overnight at 18°C. Cells were re-suspended in 1XPBS supplemented with 5 mM BME and lysed by sonication. Lysate was clarified by centrifugation and supernatant was subjected to GST affinity chromatography. The GST-Hir1(599-840)/His-Hir2(471-875) complex was co-expressed and subjected to GST affinity chromatography similar to the GST-HIRA fragments. GST fusion proteins were then subjected to on-column cleavage overnight though incubation with TEV protease. Proteins were then eluted into buffer containing 20 mM Tris pH 8.0, 50 mM NaCl, and 5 mM BME. Proteins were then subjected to ion exchange chromatography using a HiTrap Q column (GE Healthcare). Following

ion exchange, proteins were concentrated using a spin concentrator (Millipore) and loaded onto a Superdex 200 10/300 GL or Superose 6 10/300 GL (GE Healthcare) column for further purification by size-exclusion in a buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, and 1 mM TCEP. His-Hir2(471-875) alone was generated in BL21-Gold(DE3) cells (Agilent) induced with 0.8 mM IPTG and expressed overnight at 18°C. Cells were re-suspended in buffer containing 20 mM Tris pH 8.0, 500 mM NaCl, 5 mM BME, and 10 mM imidazole and lysed with sonication. Lysate was clarified with centrifugation and the supernatant was subjected to nickel affinity chromatography to isolate the His tagged protein. Protein was eluted with 200 mM Imidazole and dialyzed overnight with the addition of TEV protease to remove the His tag, dialysis buffer contains 20 mM Tris pH 8.0, 50 mM NaCl, and 5 mM BME. Protein was then subjected to ion exchange chromatography using a HiTrap Q column (GE Healthcare) followed by size-exclusion on an s200 10/300 GL or Superose 6 10/300 GL column in buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, and 1 mM TCEP. His-CABIN1 was generated through expression in Sf9 cells as described previously (Rai et al., 2011; Ricketts et al., 2015). Cells were re-suspended in buffer containing 20 mM Tris pH 8.0, 500 mM NaCl, 5 mM BME, and 10 mM imidazole and lysed with sonication. Lysate was clarified with centrifugation and the supernatant was subjected to nickel affinity chromatography to isolate the His tagged protein. Protein was eluted with 200 mM Imidazole and dialyzed overnight into buffer with 20 mM Tris pH 8.0, 300 mM NaCl, 1 mM TCEP. Protein was then concentrated

using proteins were concentrated using a spin concentrator (Millipore) and resolved on a Superose 6 10/300 GL column in a buffer with 20 mM Tris pH 8.0, 300 mM NaCl, 1 mM TCEP.

Assembly of HIRA/CABIN1 complexes using size-exclusion chromatography. To form complexes of HIRA and CABIN1, HIRA was incubated in a five-fold molar excess with CABIN1, generally about 100 μM HIRA incubated with 20 μM CABIN1. The complex was allowed to incubate at 4°C for 30 minutes prior to resolving on a Superose 6 10/300 GL column in a buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, and 1 mM TCEP. For the experiment where HIRA(661-872) was resolved with HIRA(905-1017) on a s200 10/300 GL column, both fragments were incubated together at 50 μM prior to resolving on the column in a buffer containing 20 mM Tris pH 8.0, 300 mM NaCl, and 1 mM TCEP.

Analytical ultracentrifugation – sedimentation equilibrium. HIRA, CABIN1, Hir2, Hir1/Hir2, or the HIRA/CAIBN1 complex eluting from the size-exclusion column was then pooled and analyzed for concentration. Samples were prepared at roughly 0.3, 0.5, and 0.7 OD 280 nm, loaded into a 6-channel ultracentrifugation cell with buffer blanks, and analyzed for equilibrium distribution analysis several speeds using an Optima XL-I analytical ultracentrifuge (Beckman). Equilibrium distribution data were fit with the program HeteroAnalysis

using the ideal fitting model (Cole, 2004). Global fits were conducted using three centrifugation speeds and three protein concentrations.

Limited Proteolysis. HIRA 661-1017 was held at a concentration of 1 mg/ml and incubated with a 1:50 dilution of either 1 mg/ml or 0.1 mg/ml Trypsin, Chymotrypsin, or Glu-C protease. The proteolysis was monitored by SDS-PAGE gel visualization at 15 minutes and 1 hour of incubation. The resulting proteolysis products observed on the gel were then analysis by LC/MS-MS with the aid of the Proteomics and Metabolomics Facility. For size-exclusion analysis of HIRA(661-1017) incubated with Trypsin, 500 µL of HIRA(661-1017) at 1 mg/ml was incubated with 10 µl of 0.1 mg/ml Trypsin for 30 minutes before resolving the mixture on an s200 10/300 GL column in a buffer with 20 mM Tris pH 8.0, 300 mM NaCl, and 1 mM TCEP.

Crystallization and X-ray diffraction. Crystals of HIRA(661-872) were initially grown by hanging drop vapor diffusion with a solution of 0.1 M HEPES pH 7.5, 10% PEG 4K, and 0.1 $MgCl_2$. Although these crystals did not show any evidence of X-ray diffraction we used them for microseeding trials. Crystals were harvested in the well solution and crushed into microseeds with the use of a Seed Bead (Hampton Research). Microseeds were then added to crystal drops across a new series of sparse matrix seeding in attempt to identify new crystal conditions. The largest seeded crystals of HIRA(661-872) grew with a well solution containing 0.1

M Tris pH 8.0, 10% PEG 6K, and 0.1 M NaCl. We were able to collect diffraction data from these seeded crystals.

Ctf4 structural models. Representative models of the Ctf4 trimer alone and bound to peptide were formatted for display using The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.

CHAPTER 5: Conclusions and Future Directions

5.1 – Overview

Through these studies we have identified several previously unknown functions of the HIRA complex and proposed how our observations are likely involved in the mechanism of H3.3/H4 deposition by the HIRA complex. In this chapter we summarize our findings and propose additional studies to expand upon these discoveries.

5.2 – Ubinuclein-1 confers histone H3.3 binding specificity by the HIRA histone chaperone complex

When we initially set out to identify the molecular basis for H3.3/H4 deposition by the multi-protein HIRA histone chaperone complex it was clear that the three proteins HIRA, UBN1, and CABIN1 cooperated with ASF1a to facilitate deposition of H3.3/H4 into nucleosomes, although it was unclear how this complex was able to specifically binds to and deposit H3.3/H4 over the highly homologous H3.1/H4 where H3.3 and H3.1 differ by only five amino acids.

The general feeling in the field at the time that we started this project was that the protein HIRA likely mediated H3.3/H4 binding and deposition with the aid of UBN1, CABIN1, and ASF1a, as HIRA had been the focus of many of the initial studies into this multi-subunit histone chaperone complex ((Loppin et al., 2005; Ray-Gallet et al., 2002; Ray-Gallet et al., 2007; Tang et al., 2006; Zhang et al., 2005). More recently the protein UBN1 has been identified as a homolog of the yeast Hir complex subunit Hpc2 though conservation in the highly acidic HRD

domain. It was initially suspected that the HRD was responsible for binding with HIRA(Banumathy et al., 2009), although we later identified a domain N-terminal to the HRD that was essential and sufficient for HIRA binding which we named the NHRD(Tang et al., 2012).

Following the discovery of the less conserved NHRD domain, we then focused our efforts on identifying if the acidic UBN1 HRD was able to bind H3.3/H4, as histone chaperones have long been known to harbor acidic regions to interact with the highly basic histone proteins(Laskey et al., 1978). We were able to express and purify the UBN1 HRD and show that it bound to H3.3/H4 with specificity over H3.1/H4 in the presence or absence of the HIRA N-terminal WD40 domain in a GST pull-down histone-binding assay. The UBN1/H3.3/H4 interaction was then quantified by ITC and a fluorescence polarization binding assay, UBN1 bound to H3.3/H4 with a Kd in the low micromolar range while we were not able to quantify interaction with H3.1/H4. We then went on to show that highly conserved residues in the UBN1 HRD were essential for H3.3/H4 binding as mutation of these residues to alanine abrogated the UBN1/H3.3/H4 interaction by GST pull-down and in a fluorescence polarization histone-binding assay. We then used mutagenesis to probe the five differing amino acids between H3.1 and H3.3 and demonstrated with GST pull-down and ITC that UBN1 recognizes the AAIG patch spanning residues 87-90 in H3.3 with G90 contributing the most to the specificity of UBN1 binding.

At this point we had spent significant time and effort trying to crystallize the UBN1 bound to H3.3/H4 alone, so we decided to us the previously crystalized H3/H4/Asf1 complex as a staring point to help us determine a 2.3Å crystal structure of UBN1 HRD residues 122-142 bound to H3.3/H4/Asf1. Our structure revealed that conserved UBN1 HRD residues sensitive to alanine mutagenesis are involved in key contacts between UBN1 and H3.3/H4 in the vicinity of the 87-90 AAIG patch on the H3.3 surface, confirming the veracity of our biochemical results. From our structure we also observed that UBN1 and the unrelated H3.3/H4 chaperone Daxx bind to the H3.3/H4 surface in the region of the AAIG patch with surprising similarity of structure, even though there not sequence conservation between UBN1 and Daxx and they actually bind to the histones with differing N to C-terminal orientation. The similarity in how UBN1 and Daxx bind to the H3.3/H4 surface represents the evolutionary importance of H3.3/H4; as these are two seemingly un-related histone chaperones that target differing areas of chromatin for H3.3/H4 deposition.

Finally we have demonstrated that mutation of critical HRD residues abrogates the ability of the full-length HIRA complex to bind to H3.3/H4 with specificity over H3.1/H4, indicating that binding by the UBN1 HRD is essential for H3.3/H4 specific binding. It is worth noting that while the region of Daxx we have discussed shows similarity to how the UBN1 HRD associates with H3.3/H4, the determined crystal structures of Daxx/H3.3/H4 contain a much larger fragment of

Daxx (residues 182-386) which form a large helical structure that full encompasses an H3.3/H4 dimer with many points of contact on the H3.3/H4 surface (Elsasser et al., 2012; Liu et al., 2012). From this, evidence form the full-length HIRA complex, and our observations that flanking regions of the UBN1 HRD (residues 92-121 and 149-175) also contribute to histone binding, we hypothesize that additional regions of UBN1, with potential contribution from other HIRA complex subunits, likely associate to form a larger complex with additional H3.3/H4 contacts in comparison with what we have observed in our UBN1(122-142)/H3.3/H4/Asf1 structure. Although our pull-down results with the full-length HIRA complex clearly demonstrate that the UBN1 HRD is essential for H3.3/H4 binding specificity.

Although, the mechanism for how the HIRA complex actually incorporates H3.3/H4 into nucleosomes likely involves additional steps beyond the specific binding to H3.3/H4 that we have demonstrated. To expand upon current understanding of how the HIRA complex specifically deposits H3.3/H4 into chromatin we set to characterize additional regions of the UBN1.

5.3 – The HIRA histone chaperone complex subunit UBN1 harbors H3/H4 and DNA binding activity

Following our discovery of H3.3/H4 specific binding activity by the UBN1 HRD, we then decided to characterize additional regions of UBN1 as it is a very large protein with only the two small NHRD and HRD domains studied in any detail. As

discussed above we were primarily interested in investigating if UBN1 harbored any additional regions with H3.3/H4 binding activity. We also suspected based on previous evidence that UBN1 may be able to bind DNA (Ray-Gallet et al., 2011). We performed alignment and secondary structure predication and identified a conserved and structured region roughly spanning residues 300-600 as a starting point for our functional characterization. Due to the roughly central location of this region within the UBN1 sequence we have titled this the UBN1 middle domain.

We initially showed that the middle domain fragment UBN(296-584) resolves into two peaks on a size-exclusion column and employed sedimentation equilibrium analysis to show that these peaks represent static monomer and dimer populations. We found this of great interest as it has been reported that the majority of H3.3/H4 is deposited into nucleosomes in the $(H3.3/H4)_2$ tetramer confirmation. Additionally we have previously observed with ITC conduced with the UBN1 HRD peptide and H3.3/H4 in the absence of Asf1, that UBN1 interacts with H3.3/H4 in a 1:1 ratio. This result could be interpreted as a 1:1:1 UBN1/H3.3/H4 complex, although it is more likely that this represents a 2:2:2 UBN1/H3.3/H4 complex as H3.3/H4 have been reported to exist as a tetramer at the concentrations at which we conduced ITC(Winkler et al., 2012). If we combine our observation of UBN1 dimer formation with these existing data a hypothesis for a model of HIRA complex function emerges where UBN1 forms a dimer to associate with a H3.3/H4 tetramer. We then showed that the C-terminal

region of the middle domain (residues 504-584) is essential for formation of the dimer population, while residues 296-341 are dispensable.

Next we identified that the UBN1 middle domain also has histone binding activity with no specificity towards H3.3/H4 over H3.1/H4, although it does appear to be specific to H3/H4 and does not bind H2A/H2B. This interaction seems to be driven by electrostatics as the middle domain binds to H3/H4 well at 300 mM NaCl, but the interaction is abated when the pull-down was conducted with 750 mM NaCl. We then showed that UBN1 residues 504-584 are dispensable for H3/H4 binding while deletion of residues 296-341 abolishes the interaction between the UBN1 middle domain and H3/H4. We then made a series of deletions across residues 296-341 each of which showed a gradual reduction in H3/H4 binding indicating that several residues in this region are involved in binding to the histones. Finally we tested the interaction of the UBN1 middle domain with full-length H3/H4 and H3/H4 with the N-terminal tails removed. We demonstrated that the middle domain binds to full-length H3/H4 well but shows greatly reduced binding with the tailless H3/H4 indicating that the UBN1 middle domain interacts with the charged N-terminal H3/H4 tails.

The UBN1 HRD and middle domain are separated by a highly basic linker region, which harbors many lysine residues that are conserved among vertebrate UBN1 homologs. It had been previously reported that UBN1 was able to nonspecifically bind DNA with greater affinity than HIRA or CABIN1. We hypothesized that

UBN1 DNA binding activity may be mediated by this highly basic linker region spanning UBN1 residues 176-295. To address this hypothesis we tested the ability of UBN1 to bind DNA using a fluorescence polarization assay and demonstrated that WT UBN1 is able to bind DNA with nanomolar affinity, and if we delete UBN1 residues 176-295 and replace them with a poly glycine serine linker UBN1 is no longer able to form complex with DNA. We then made a series of targeted deletions of individual patches of lysine residues within this UBN1 region and showed that removal of these residues decreases the binding affinity, from which we can conclude that these conserved lysine residues are involved in a non-specific electrostatic DNA binding activity.

As UBN1 needs to deposit H3.3/H4 into chromatin it is highly intriguing that UBN1 harbors a DNA binding domain directly between adjacent histone binding domains. From this discovery we developed a hypothesis for H3.3/H4 deposition that involved histone binding, dimer formation to induce formation of an $(H3.3/H4)_2$ tetramer, and DNA binding activity by UBN1. We then investigated if the UBN1 HRD and middle domain could cooperate to bind H3.3/H4 together in formation of a larger chaperone/histone complex. We analyzed the contribution of both the UBN1 middle domain dimer and monomer populations to binding with H3.3/H4 and observed that the UBN1 HRD and middle domain compete for H3.3/H4 binding and associate with the histones in a mutually exclusive manner. Finally we investigated the ability of UBN1 to bind free H3.3/H4 in comparison

with assembled nucleosomes. We discovered that UBN1 fragments as large as UBN1(122-584) were capable of binding to free H3.3/H4 with no detectable interaction with the nucleosome. Indicating that UBN1 is likely involved in chaperoning free H3.3/H4 for nucleosome assembly.

From our observations we proposed a multi-step model for H3.3/H4 deposition by the HIRA complex involving a handoff of H3.3/H4 between the HRD and middle domain and DNA binding. So far it is unclear what the exact order of these events may be and our proposed model represents just one possible scenario. It is likely that yet to be determined cellular events also contribute to this multi-step process, one outstanding detail is how ASF1a disassociates to allow for the proposed formation of the $(H3.3/H4)_2$ tetramer. We have yet to be able to demonstrate the handoff of histones between ASF1a and the HIRA complex biochemically. It is possible that a yet to be determined cellular factor allows for the release of ASF1a to allow for formation of the $(H3.3/H4)_2$ tetramer. ASF1a binds very tightly to the HIRA B-domain, and we have yet to find conditions that allow it to dissociate from HIRA or H3.3/H4 (data not shown). To address the remaining questions associated with our proposed model for H3.3/H4 deposition by the HIRA complex, we plan to pursue electron microscopy studies with the full-length HIRA complex with the addition of ASF1/H3.3/H4, DNA, or both to identify potentially allosteric structural changes that occur upon binding of these factors. Although it is possible that cellular factors such as yet to

be discovered PTMs to the HIRA complex, ASF1a, or H3.3/H4 allow for transitions between these pre-deposition conformations, which would be more difficult to pin down without additional cellular studies to identify potentially important PTMs.

5.4 – Molecular characterization of HIRA oligomer formation and interaction with CABIN1

Concurrently with our studies into the UBN1 middle domain and DNA binding loop, we began to investigate a link between HIRA oligomer formation and binding with the subunit CABIN1. Another group observed that deletions within the HIRA C-terminal region reduced the ability of HIRA to both bind CABIN1 and pull-down a differentially tagged HIRA. We were invited to collaborate on biochemical and biophysical analysis of their findings. We were highly intrigued by this as we were already working on the UBN1 middle domain dimerization and found it interesting that HIRA may also have oligomer formation activity, which would contribute well to our proposed model for HIRA complex function.

We were able produce a soluble fragment of the HIRA C-terminal domain, HIRA(661-1017) which is predicted to be organized into a strand and helical region separated by a short linker region. This region of HIRA has been previously reported to associate with H3.3/H4 with increased affinity for H3.3/H4 bearing an S47 phosphorylation(Kang et al., 2011), although we were unable to observe any binding with our un-phosphorylated H3.3/H4. From this result it is

worth noting that the only histone biding activity we have observed from the HIRA-complex subunits we have been able to produce comes from the UBN1 HRD and middle domain. From our studies both this HIRA C-terminal region and the N-terminal WD40 domain of HIRA appear to be dispensable for binding with H3.3/H4.

Using size-exclusion chromatography we were able to show that HIRA(661-1017) readily forms a complex with full-length CABIN1. To analyze the molecular assembly of HIRA alone in comparison with the HIRA/CABIN1 complex we performed sedimentation equilibrium analysis on both HIRA(661-1017) and CABIN1 alone as well as the complex. We found that HIRA(661-1017) forms a stable trimer in isolation while CABIN1 exists as a monomer. Analysis of the HIRA/CABIN1 complex revealed that it has a HIRA(3)/CABIN1(2) stoichiometry.

Limited-proteolysis experiments revealed that the HIRA(661-1017) region is likely organized into a HIRA(661-872) strand domain and a HIRA(905-1017) helical domain which exist as monomers individually and do not form a complex when mixed. Although we were able to replace the 873-905 loop with a GS linker to generate HIRA(661-1017, Δ873-905) which still formed a stable trimer. From this we concluded that the HIRA(873-904) linker is necessary to bring these domains together in the proper conformation for trimer formation.

Interestingly neither HIRA(661-872), HIRA(905-1017), HIRA(661-1017, Δ873-905), or even HIRA(873-1017) were able to bind CABIN1, which suggests that

the HIRA(873-904) loop is critical for CABIN1 binding but not sufficient without HIRA trimer formation. We then went on to show that a triple mutant in the loop region, HIRA(661-1017) F870A_R871A_L874A, displayed decreased complex formation with CABIN1 further verifying the CABIN1 likely binds in the region of the HIRA(873-904) loop.

We are particularly interested in identifying the amino acid contacts which allow for HIRA trimer formation and have extensively attempted crystal screening for both HIRA(661-1017) and HIRA(661-1017, ∆873-905) although we have not had any hits for either. Although it forms a monomer we have also pursued crystallization of HIRA(661-872) which has yielded several hits. We collected multiple 3.5-4.0Å datasets for these crystals but have so far been unable to solve the structure. Future directions for crystallization include continued optimization of these HIRA(661-872) crystals as well as continued attempts with HIRA(661-1017, ∆873-904) and potentially variations on the length of the GS linker which we have replaced the loop with as well as variations on the HIRA residues we are removing in this region.

In attempt to identify a good molecular replacement search model we used the Phyre server to predict a model for HIRA(661-872), and to our surprise the top hit was a protein called Ctf4 which has remarkably similar domain architecture to HIRA. The C-terminal strand/helix domain in Ctf4 also forms a trimer, of which there have been 6 different crystal structures reported. Of the six structures only

one has density for all three helical domains, and of 4 structures in complex with binding partners, 3 have a Ctf4(3)/Peptide(2) stoichiometry. These observations about the arrangement of the Ctf4 trimer and the peptide binding stoichiometry are potentially crystallographic artifacts, but the similarity to the HIRA trimer and HIRA/CABIN1 interaction is very interesting. It is possible that HIRA and Ctf4 are distantly related structural homologs and could function in the same way to associate with binding partners in a 3:2 stoichiometry.

An interesting note in comparison of our HIRA(661-1017) trimer is that the Ctf4 trimer that crystallized is composed of Ctf(474-927) with no loop deletion (although there is not density for the loop in any of the published structures). The Ctf4 trimer is composed of an additional 97 residues that reside in the N-terminal region of the Ctf4 strand domain. Although HIRA has little predicted structure in the region of HIRA(564-661), if would be interesting to generate a HIRA(564-1017) to be of comparable length to the Ctf4 trimer crystallized. It is possible that these residues, which are predicted to be unstructured, could contribute to the strand domain fold and allow for a more stable HIRA trimer for crystallization trials.

We then turned our focus onto the yeast Hir complex where functional HIRA is composed of two homologous proteins called Hir1 and Hir2. Hir1 and Hir2 have been reported to associate in a Hir1(1)/Hir2(2) ratio, and only Hir2 has sequence homologous to HIRA(873-1017). From this we hypothesized that the HIRA trimer

and the odd HIRA(3)/CABIN1(2) binding ratio could also be represented in the Hir complex by a Hir1(1)/Hir2(2)/Hir3(2) complex. We were able to produce a complex of Hir1 and Hir2 and showed that they do indeed for a Hir1(1)/Hir2(2) trimer although we still need to investigate the assembly of the Hir1/Hir2/Hir3 complex. Although we have had difficulty producing soluble Hir3 this remains an important future experiment to further validate our findings about how the HIRA/CABIN1 complex assembles and is possibly related to Ctf4 structure and function. Additionally we could attempt to see if human CABIN1 will associate with the Hir1/Hir2 complex we have produced, as our human CABIN1 expresses well in Sf9 cells. Although this may not work, while Hir3 and CABIN1 are organized into similar TPR repeat structures, there is little sequence conservation between the two homologs(Rai et al., 2011).

Finally we have combined our observations of a HIRA(3)/CABIN1(2) complex with our previously presented hypothesis that UBN1 dimerization may facilitate formation of a $(H3.3/H4)_2$ tetramer to propose a new model for assembly of the HIRA complex. This model has a HIRA trimer associated with two molecules of CABIN1 while each HIRA trimer is bound by UBN1. Of the three UBN1 molecules we propose that two may form a dimer that associates with an (H3.3/H4)2 tetramer, while the third is waiting in the wings with H3.3/H4/ASF1a to be incorporated into the next tetramer to be formed for deposition.

To verify this model we are optimizing production of the full-length HIRA complex in Sf9 insect cells to be used in AUC experiments to study the molecular assembly of the full-length complex to compliment our studies with the UBN1 and HIRA fragments. In addition to AUC studies with the full length HIRA complex we also plan to study the full-length HIRA complex using an electron microscopy based approach to further understand the molecular assembly of the full complex as well as any structural rearrangement that may occur upon the addition of ASF1a/H3.3/H4 and DNA.

BIBLIOGRAPHY

Adam, S., Polo, S.E., and Almouzni, G. (2013). Transcription recovery after DNA damage requires chromatin priming by the H3.3 histone chaperone HIRA. Cell *155*, 94-106.

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

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. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Molecular cell *9*, 1191-1200.

Aho, S., Buisson, M., Pajunen, T., Ryoo, Y.W., Giot, J.F., Gruffat, H., Sergeant, A., and Uitto, J. (2000). Ubinuclein, a novel nuclear protein interacting with cellular and viral transcription factors. The Journal of cell biology *148*, 1165-1176.

Aho, S., Lupo, J., Coly, P.A., Sabine, A., Castellazzi, M., Morand, P., Sergeant, A., Manet, E., Boyer, V., and Gruffat, H. (2009). Characterization of the ubinuclein protein as a new member of the nuclear and adhesion complex components (NACos). Biology of the cell / under the auspices of the European Cell Biology Organization *101*, 319-334.

Amin, A.D., Vishnoi, N., and Prochasson, P. (2013). A global requirement for the HIR complex in the assembly of chromatin. Biochimica et biophysica acta *1819*, 264-276.

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

Banaszynski, L.A., Wen, D., Dewell, S., Whitcomb, S.J., Lin, M., Diaz, N., Elsasser, S.J., Chapgier, A., Goldberg, A.D., Canaani, E.*, et al.* (2013). Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. Cell *155*, 107-120.

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

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell research *21*, 381-395.

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. Molecular and cellular biology *29*, 758-770.

Blatch, G.L., and Lassle, M. (1999). The tetratricopeptide repeat: a structural motif mediating protein-protein interactions. BioEssays : news and reviews in molecular, cellular and developmental biology *21*, 932-939.

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 genetics *3*, 1991-2006.

Bowman, A., Ward, R., El-Mkami, H., Owen-Hughes, T., and Norman, D.G. (2010). Probing the (H3-H4)2 histone tetramer structure using pulsed EPR spectroscopy combined with site-directed spin labelling. Nucleic acids research *38*, 695-707.

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. Molecular cell *41*, 398-408.

Buchan, D.W., Minneci, F., Nugent, T.C., Bryson, K., and Jones, D.T. (2013). Scalable web services for the PSIPRED Protein Analysis Workbench. Nucleic acids research *41*, W349-357.

Burgess, R.J., and Zhang, Z. (2013). Histone chaperones in nucleosome assembly and human disease. Nature structural & molecular biology *20*, 14-22.

Chen, S., Rufiange, A., Huang, H., Rajashankar, K.R., Nourani, A., and Patel, D.J. (2015). Structure-function studies of histone H3/H4 tetramer maintenance during transcription by chaperone Spt2. Genes & development *29*, 1326-1340.

Choi, S.Y., Jang, H., Roe, J.S., Kim, S.T., Cho, E.J., and Youn, H.D. (2013). Phosphorylation and ubiquitination-dependent degradation of CABIN1 releases p53 for transactivation upon genotoxic stress. Nucleic acids research *41*, 2180-2190.

Cole, J.L. (2004). Analysis of heterogeneous interactions. Methods in enzymology *384*, 212-232.

Corpet, A., Olbrich, T., Gwerder, M., Fink, D., and Stucki, M. (2014). Dynamics of histone H3.3 deposition in proliferating and senescent cells reveals a DAXX-dependent targeting to PML-NBs important for pericentromeric heterochromatin organization. Cell Cycle *13*, 249-267.

Dilg, D., Saleh, R.N., Phelps, S.E., Rose, Y., Dupays, L., Murphy, C., Mohun, T., Anderson, R.H., Scambler, P.J., and Chapgier, A.L. (2016). HIRA Is Required for Heart Development and Directly Regulates Tnni2 and Tnnt3. PloS one *11*, e0161096.

Dimova, D., Nackerdien, Z., Furgeson, S., Eguchi, S., and Osley, M.A. (1999). A role for transcriptional repressors in targeting the yeast Swi/Snf complex. Molecular cell *4*, 75-83.

Domsic, J.F., Chen, H.S., Lu, F., Marmorstein, R., and Lieberman, P.M. (2013). Molecular basis for oligomeric-DNA binding and episome maintenance by KSHV LANA. PLoS pathogens *9*, e1003672.

Dutta, D., Ray, S., Home, P., Saha, B., Wang, S., Sheibani, N., Tawfik, O., Cheng, N., and Paul, S. (2010). Regulation of angiogenesis by histone chaperone HIRA-mediated incorporation of lysine 56-acetylated histone H3.3 at chromatin domains of endothelial genes. The Journal of biological chemistry *285*, 41567-41577.

Elsaesser, S.J., and Allis, C.D. (2010). HIRA and Daxx constitute two independent histone H3.3-containing predeposition complexes. Cold Spring Harbor symposia on quantitative biology *75*, 27-34.

Elsasser, S.J., Huang, H., Lewis, P.W., Chin, J.W., Allis, C.D., and Patel, D.J. (2012). DAXX envelops a histone H3.3-H4 dimer for H3.3-specific recognition. Nature *491*, 560-565.

Elsasser, S.J., Noh, K.M., Diaz, N., Allis, C.D., and Banaszynski, L.A. (2015). Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. Nature.

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

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

Esau, C., Boes, M., Youn, H.D., Tatterson, L., Liu, J.O., and Chen, J. (2001). Deletion of calcineurin and myocyte enhancer factor 2 (MEF2) binding domain of Cabin1 results in enhanced cytokine gene expression in T cells. The Journal of experimental medicine *194*, 1449-1459.

Fazly, A., Li, Q., Hu, Q., Mer, G., Horazdovsky, B., and Zhang, Z. (2012). Histone chaperone Rtt106 promotes nucleosome formation using (H3-H4)2 tetramers. The Journal of biological chemistry *287*, 10753-10760.

Gaillard, P.H., Martini, E.M., Kaufman, P.D., Stillman, B., Moustacchi, E., and Almouzni, G. (1996). Chromatin assembly coupled to DNA repair: a new role for chromatin assembly factor I. Cell *86*, 887-896.

Galvani, A., Courbeyrette, R., Agez, M., Ochsenbein, F., Mann, C., and Thuret, J.Y. (2008). In vivo study of the nucleosome assembly functions of ASF1 histone chaperones in human cells. Molecular and cellular biology *28*, 3672-3685.

Gillette, W.K., Esposito, D., Taylor, T.E., Hopkins, R.F., Bagni, R.K., and Hartley, J.L. (2011). Purify First: rapid expression and purification of proteins from XMRV. Protein expression and purification *76*, 238-247.

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.

Gouet, P., Robert, X., and Courcelle, E. (2003). ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic acids research *31*, 3320-3323.

Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J., and Lopez, R. (2010). A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic acids research *38*, W695-699.

Green, E.M., Antczak, A.J., Bailey, A.O., Franco, A.A., Wu, K.J., Yates, J.R., 3rd, and Kaufman, P.D. (2005). Replication-independent histone deposition by the HIR complex and Asf1. Current biology : CB *15*, 2044-2049.

Hammond, D.R., and Udvadia, A.J. (2010). Cabin1 expression suggests roles in neuronal development. Developmental dynamics : an official publication of the American Association of Anatomists *239*, 2443-2451.

Han, A., Pan, F., Stroud, J.C., Youn, H.D., Liu, J.O., and Chen, L. (2003). Sequence-specific recruitment of transcriptional co-repressor Cabin1 by myocyte enhancer factor-2. Nature *422*, 730-734.

Huang, H., Stromme, C.B., Saredi, G., Hodl, M., Strandsby, A., Gonzalez-Aguilera, C., Chen, S., Groth, A., and Patel, D.J. (2015). A unique binding mode enables MCM2 to chaperone histones H3-H4 at replication forks. Nature structural & molecular biology *22*, 618-626.

Jang, H., Choi, S.Y., Cho, E.J., and Youn, H.D. (2009). Cabin1 restrains p53 activity on chromatin. Nature structural & molecular biology *16*, 910-915.

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. Nature genetics *41*, 941-945.

Jullien, J., Astrand, C., Szenker, E., Garrett, N., Almouzni, G., and Gurdon, J.B. (2012). HIRA dependent H3.3 deposition is required for transcriptional reprogramming following nuclear transfer to Xenopus oocytes. Epigenetics & chromatin *5*, 17.

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 & development *25*, 1359-1364.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Lacoste, N., Woolfe, A., Tachiwana, H., Garea, A.V., Barth, T., Cantaloube, S., Kurumizaka, H., Imhof, A., and Almouzni, G. (2014). Mislocalization of the Centromeric Histone Variant CenH3/CENP-A in Human Cells Depends on the Chaperone DAXX. Molecular cell *53*, 631-644.

Lamour, V., Lecluse, 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. Human molecular genetics *4*, 791-799.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R.*, et al.* (2007). Clustal W and Clustal X version 2.0. Bioinformatics *23*, 2947-2948.

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

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

Li, X., and Tyler, J.K. (2016). Nucleosome disassembly during human non-homologous end joining followed by concerted HIRA- and CAF-1-dependent reassembly. eLife *5*.

Lindsay, E.A., Vitelli, F., Su, H., Morishima, M., Huynh, T., Pramparo, T., Jurecic, V., Ogunrinu, G., Sutherland, H.F., Scambler, P.J.*, et al.* (2001). Tbx1 haploinsufficieny in the DiGeorge syndrome region causes aortic arch defects in mice. Nature *410*, 97-101.

Liu, C.P., Xiong, C., Wang, M., Yu, Z., Yang, N., Chen, P., Zhang, Z., Li, G., and Xu, R.M. (2012). Structure of the variant histone H3.3-H4 heterodimer in complex with its chaperone DAXX. Nature structural & molecular biology *19*, 1287-1292.

Liu, W., Youn, H.D., and Liu, J.O. (2001). Thapsigargin-induced apoptosis involves Cabin1-MEF2-mediated induction of Nur77. European journal of immunology *31*, 1757-1764.

Loppin, B., Bonnefoy, E., Anselme, C., Laurencon, 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., Lecluse, Y., Almouzni, G., and Lipinski, M. (1998). Core histones and HIRIP3, a novel histone-binding protein, directly interact with WD repeat protein HIRA. Molecular and cellular biology *18*, 5546-5556.

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

Luger, K., Rechsteiner, T.J., Flaus, A.J., Waye, M.M., and Richmond, T.J. (1997b). Characterization of nucleosome core particles containing histone proteins made in bacteria. Journal of molecular biology *272*, 301-311.

Luger, K., Rechsteiner, T.J., and Richmond, T.J. (1999). Preparation of nucleosome core particle from recombinant histones. Methods in enzymology *304*, 3-19.

Mattiroli, F., D'Arcy, S., and Luger, K. (2015). The right place at the right time: chaperoning core histone variants. EMBO reports *16*, 1454-1466.

Mattiroli, F., Gu, Y., Yadav, T., Balsbaugh, J.L., Harris, M.R., Findlay, E.S., Liu, Y., Radebaugh, C.A., Stargell, L.A., Ahn, N.G.*, et al.* (2017). DNA-mediated association of two histone-bound complexes of yeast Chromatin Assembly Factor-1 (CAF-1) drives tetrasome assembly in the wake of DNA replication. eLife *6*.

Maze, I., Noh, K.M., Soshnev, A.A., and Allis, C.D. (2014). Every amino acid matters: essential contributions of histone variants to mammalian development and disease. Nature reviews Genetics *15*, 259-271.

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.

Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). Science *183*, 330-332.

Orsi, G.A., Algazeery, A., Meyer, R.E., Capri, M., Sapey-Triomphe, L.M., Horard, B., Gruffat, H., Couble, P., Ait-Ahmed, O., and Loppin, B. (2013). Drosophila Yemanuclein and HIRA cooperate for de novo assembly of H3.3-containing nucleosomes in the male pronucleus. PLoS genetics *9*, e1003285.

Pchelintsev, N.A., McBryan, T., Rai, T.S., van Tuyn, J., Ray-Gallet, D., Almouzni, G., and Adams, P.D. (2013). Placing the HIRA histone chaperone complex in the chromatin landscape. Cell reports *3*, 1012-1019.

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 & development *19*, 2534-2539.

Rai, T.S., Puri, A., McBryan, T., Hoffman, J., Tang, Y., Pchelintsev, N.A., 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. Molecular and cellular biology *31*, 4107-4118.

Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. Molecular cell *9*, 1091-1100.

Ray-Gallet, D., Quivy, J.P., Sillje, H.W., Nigg, E.A., and Almouzni, G. (2007). The histone chaperone Asf1 is dispensable for direct de novo histone deposition in Xenopus egg extracts. Chromosoma *116*, 487-496.

Ray-Gallet, D., Woolfe, A., Vassias, I., Pellentz, C., Lacoste, N., Puri, A., Schultz, D.C., Pchelintsev, N.A., Adams, P.D., Jansen, L.E.*, et al.* (2011). Dynamics of histone H3 deposition in vivo reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. Molecular cell *44*, 928-941.

Richet, N., Liu, D., Legrand, P., Velours, C., Corpet, A., Gaubert, A., Bakail, M., Moal-Raisin, G., Guerois, R., Compper, C.*, et al.* (2015). Structural insight into how the human helicase subunit MCM2 may act as a histone chaperone together with ASF1 at the replication fork. Nucleic acids research *43*, 1905-1917.

Richmond, T.J., and Davey, C.A. (2003). The structure of DNA in the nucleosome core. Nature *423*, 145-150.

Ricketts, M.D., Frederick, B., Hoff, H., Tang, Y., Schultz, D.C., Singh Rai, T., Grazia Vizioli, M., Adams, P.D., and Marmorstein, R. (2015). Ubinuclein-1

confers histone H3.3-specific-binding by the HIRA histone chaperone complex. Nature communications *6*, 7711.

Safina, A., Cheney, P., Pal, M., Brodsky, L., Ivanov, A., Kirsanov, K., Lesovaya, E., Naberezhnov, D., Nesher, E., Koman, I.*, et al.* (2017). FACT is a sensor of DNA torsional stress in eukaryotic cells. Nucleic acids research *45*, 1925-1945.

Sauer, P.V., Timm, J., Liu, D., Sitbon, D., Boeri-Erba, E., Velours, C., Mucke, N., Langowski, J., Ochsenbein, F., Almouzni, G.*, et al.* (2017). Insights into the molecular architecture and histone H3-H4 deposition mechanism of yeast Chromatin assembly factor 1. eLife *6*.

Schenk, R., Jenke, A., Zilbauer, M., Wirth, S., and Postberg, J. (2011). H3.5 is a novel hominid-specific histone H3 variant that is specifically expressed in the seminiferous tubules of human testes. Chromosoma *120*, 275-285.

Schneiderman, J.I., Orsi, G.A., Hughes, K.T., Loppin, B., and Ahmad, K. (2012). Nucleosome-depleted chromatin gaps recruit assembly factors for the H3.3 histone variant. Proceedings of the National Academy of Sciences of the United States of America *109*, 19721-19726.

Seiler, C.Y., Park, J.G., Sharma, A., Hunter, P., Surapaneni, P., Sedillo, C., Field, J., Algar, R., Price, A., Steel, J.*, et al.* (2014). DNASU plasmid and PSI:Biology-Materials repositories: resources to accelerate biological research. Nucleic acids research *42*, D1253-1260.

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. Molecular and cellular biology *13*, 28-38.

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

Simon, A.C., Zhou, J.C., Perera, R.L., van Deursen, F., Evrin, C., Ivanova, M.E., Kilkenny, M.L., Renault, L., Kjaer, S., Matak-Vinkovic, D.*, et al.* (2014). A Ctf4

trimer couples the CMG helicase to DNA polymerase alpha in the eukaryotic replisome. Nature *510*, 293-297.

Song, Y., Seol, J.H., Yang, J.H., Kim, H.J., Han, J.W., Youn, H.D., and Cho, E.J. (2013). Dissecting the roles of the histone chaperones reveals the evolutionary conserved mechanism of transcription-coupled deposition of H3.3. Nucleic acids research *41*, 5199-5209.

Spector, M.S., and Osley, M.A. (1993). The HIR4-1 mutation defines a new class of histone regulatory genes in Saccharomyces cerevisiae. Genetics *135*, 25-34.

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. Molecular and cellular biology *17*, 545-552.

Su, D., Hu, Q., Li, Q., Thompson, J.R., Cui, G., Fazly, A., Davies, B.A., Botuyan, M.V., Zhang, Z., and Mer, G. (2012). Structural basis for recognition of H3K56-acetylated histone H3-H4 by the chaperone Rtt106. Nature *483*, 104-107.

Sun, L., Youn, H.D., Loh, C., Stolow, M., He, W., and Liu, J.O. (1998). Cabin 1, a negative regulator for calcineurin signaling in T lymphocytes. Immunity *8*, 703-711.

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.

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

Tang, Y., Puri, A., Ricketts, M.D., Rai, T.S., Hoffmann, J., Hoi, E., Adams, P.D., Schultz, D.C., and Marmorstein, R. (2012). Identification of an ubinuclein 1 region required for stability and function of the human HIRA/UBN1/CABIN1/ASF1a histone H3.3 chaperone complex. Biochemistry *51*, 2366-2377.

Tsunaka, Y., Fujiwara, Y., Oyama, T., Hirose, S., and Morikawa, K. (2016). Integrated molecular mechanism directing nucleosome reorganization by human FACT. Genes & development *30*, 673-686.

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

Villa, F., Simon, A.C., Ortiz Bazan, M.A., Kilkenny, M.L., Wirthensohn, D., Wightman, M., Matak-Vinkovic, D., Pellegrini, L., and Labib, K. (2016). Ctf4 Is a Hub in the Eukaryotic Replisome that Links Multiple CIP-Box Proteins to the CMG Helicase. Molecular cell *63*, 385-396.

Wang, H., Wang, M., Yang, N., and Xu, R.M. (2015). Structure of the quaternary complex of histone H3-H4 heterodimer with chaperone ASF1 and the replicative helicase subunit MCM2. Protein & cell *6*, 693-697.

Weiner, M.P., and Costa, G.L. (1994). Rapid PCR site-directed mutagenesis. PCR methods and applications *4*, S131-136.

Wiedemann, S.M., Mildner, S.N., Bonisch, C., Israel, L., Maiser, A., Matheisl, S., Straub, T., Merkl, R., Leonhardt, H., Kremmer, E.*, et al.* (2010). Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. The Journal of cell biology *190*, 777-791.

Winkler, D.D., Zhou, H., Dar, M.A., Zhang, Z., and Luger, K. (2012). Yeast CAF-1 assembles histone (H3-H4)2 tetramers prior to DNA deposition. Nucleic acids research *40*, 10139-10149.

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. (2010). ATRX interacts with H3.3 in maintaining telomere structural integrity in pluripotent embryonic stem cells. Genome research *20*, 351-360.

Xu, M., Long, C., Chen, X., Huang, C., Chen, S., and Zhu, B. (2010). Partitioning of histone H3-H4 tetramers during DNA replication-dependent chromatin assembly. Science *328*, 94-98.

Yang, J.H., Choi, J.H., Jang, H., Park, J.Y., Han, J.W., Youn, H.D., and Cho, E.J. (2011). Histone chaperones cooperate to mediate Mef2-targeted transcriptional regulation during skeletal myogenesis. Biochemical and biophysical research communications *407*, 541-547.

Ye, X., Zerlanko, B., Zhang, R., Somaiah, N., Lipinski, M., Salomoni, P., and Adams, P.D. (2007). Definition of pRB- and p53-dependent and -independent steps in HIRA/ASF1a-mediated formation of senescence-associated heterochromatin foci. Molecular and cellular biology *27*, 2452-2465.

Youn, H.D., and Liu, J.O. (2000). Cabin1 represses MEF2-dependent Nur77 expression and T cell apoptosis by controlling association of histone deacetylases and acetylases with MEF2. Immunity *13*, 85-94.

Zasadzinska, E., Barnhart-Dailey, M.C., Kuich, P.H., and Foltz, D.R. (2013). Dimerization of the CENP-A assembly factor HJURP is required for centromeric nucleosome deposition. The EMBO journal *32*, 2113-2124.

Zhang, K., Gao, Y., Li, J., Burgess, R., Han, J., Liang, H., Zhang, Z., and Liu, Y. (2016). A DNA binding winged helix domain in CAF-1 functions with PCNA to stabilize CAF-1 at replication forks. Nucleic acids research *44*, 5083-5094.

Zhang, R., Liu, S.T., Chen, W., Bonner, M., Pehrson, J., Yen, T.J., and Adams, P.D. (2007). HP1 proteins are essential for a dynamic nuclear response that rescues the function of perturbed heterochromatin in primary human cells. Molecular and cellular biology *27*, 949-962.

Zhang, R., Poustovoitov, M.V., Ye, X., Santos, H.A., Chen, W., Daganzo, S.M., Erzberger, J.P., Serebriiskii, I.G., Canutescu, A.A., Dunbrack, R.L.*, et al.* (2005). Formation of MacroH2A-containing senescence-associated heterochromatin foci and senescence driven by ASF1a and HIRA. Developmental cell *8*, 19-30.

Zhu, R., Iwabuchi, M., and Ohsumi, K. (2017). The WD40 Domain of HIRA Is Essential for RI-nucleosome Assembly in Xenopus Egg Extracts. Cell structure and function *42*, 37-48.