Methylated arginine nucleoplasmin tail interacts with at least three extended Tudor domains of TDRD6 *in vitro*

Presented to the S. Daniel Abraham Honors Program
in Partial Fulfillment of the
Requirements for Completion of the Program

Stern College for Women
Yeshiva University
August 23rd, 2019

Neda Shokrian

Mentor: Dr. Alyssa Schuck, Biology

Table of Contents:

Abstract

The Central Dogma of biology states that in a cell, DNA must undergo transcription in order to produce RNA, which is then translated into protein. During early embryogenesis in *Xenopus laevis*, however, newly fertilized egg cells rapidly replicate their tightly wound DNA, using proteins not produced by standard method, and the genome is transcriptionally silent. *Xenopus* egg cells contain maternal messenger ribonucleoprotein (mRNP) particles that function to store maternal mRNA in a translationally repressed state. After fertilization, the mRNP particles are somehow remodeled to trigger translation of their stored mRNA transcripts until zygotic genome activation (ZGA). Amongst a handful of proteins contained in the mRNP particles is Tudor domain containing 6 (TDRD6). Various Tudor domain proteins have been shown to interact with binding partners containing arginine-methylated motifs. Nucleoplasmin (Npm2) is the predominant storage chaperone for histones H2A and H2B in *Xenopus* oocytes. The C-terminus tail has a methylated arginine residue whose function is unknown. We hypothesized that the arginine-methylated Npm2 C-terminal tail modulates interaction with TDRD6 to de-repress the translational block of maternal mRNPs during early embryogenesis. Purification of several extended Tudor domains (eTUDs) of TDRD6, followed by an *in vitro* interaction assay with Npm2 constructs, showed that Npm2-C19 with arginine methylation pulled down eTUDs 2, 5, and 6, while Npm2-C19 without arginine methylation did not pull down any eTUDs. While further experiments must be conducted, a connection is now established between the binding of the eTUDs and the methylated arginine residue of the Npm2 tail.

Introduction

(i) *Translation during early embryogenesis*

Cells must manufacture proteins in order to survive and reproduce. The Central Dogma of biology states that DNA must undergo transcription in order to produce RNA, which is then translated into protein. Thus, newly produced daughter cells must produce proteins in order to continue to propagate. The cleavage stage of early embryogenesis presents an exception to this rule. Since a newly fertilized zygote undergoes rapid rounds of mitosis, its genome is kept in a highly condensed state (Amodeo *et al*., 2015). Thus, transcription machinery is blocked from accessing the zygotic DNA needed to initiate protein production. Therefore, the zygotic genome is transcriptionally silent (Laskey, 1985).

Yet, the newly fertilized egg cells still undergo numerous rapid rounds of DNA replication and division in the complete absence of zygotic gene products. This is possible because newly fertilized eggs rely on the maternal factors, such as RNA, proteins, and metabolites, that were deposited into the egg during oogenesis (Sun *et al*., 2014). This period of development is termed "maternal control".

After multiple rounds of division, organisms reach zygotic genome activation (ZGA), the point at which the zygote's genome becomes transcriptionally active and also marks the end of maternal control (Tadros and Lipshitz, 2009). ZGA is concomitant with the mid-blastula transition (MBT) (O'Farrell *et al*., 2004). The point of ZGA differs between organisms; in *Xenopus laevis* embryos, ZGA occurs after the 12th round of division, when there are approximately 4,000 cells in the embryo (Schier, 2007).

(ii) *Maternal mRNPs*

During oogenesis, over 80% of synthesized mRNA is sequestered into storage messenger ribonucleoprotein particles (mRNPs) and is translationally repressed (Davidson, 1986). Maternal mRNPs are germ granules that regulate the release of the maternal factors deposited into the egg during oogenesis (Kimble and Crittenden, 2007). Maternal mRNPs have been studied in depth and have been found to contain multiple proteins and mRNAs essential for primordial germ cell (PGC) development, such as *XDead end*, which assists in PGC migration (Horvay *et al*., 2006), and *Xdazl*, which contains an RNA binding function and is localized to the germ plasm (Houston *et al*., 1998). Maternal mRNPs establish a translational block of stored maternal transcripts in eggs (Richter and Smith, 1984). They are activated upon fertilization by a presently unknown mechanism, and therefore regulate early embryonic gene expression.

(iii) *Tudor domain proteins and TDRD6*

Tudor domain (TDRD) proteins are a family of proteins known to recognize methylated ligands (Maurer-Stroh *et al*., 2003). Much has been elucidated about their structure and function: a Tudor domain contains a ~60 amino acid core structure composed of three to five antiparallel B-strands, which form a barrel-like structure with an aromatic binding pocket at the surface to accommodate methylated ligands (Taverna *et al*., 2007) (Figure 1). Tudor proteins are involved in many cellular processes, such as RNA metabolism. For example, human survival motor neuron (SMN), a member of the Tudor domain family, is implicated in mRNA splicing (Buhler *et al*., 1999). Mammalian Tudor proteins can contain a single Tudor domain alone, multiple tandem Tudor domain repeats, or one or more Tudor

domains in conjunction with other types of domains. Many TDRD proteins have additional 180 residue conserved structural elements flanking the canonical 60-amino-acid Tudor domain units; these elements are known as extended Tudor domains (eTUDs) (Chen *et al*., 2011).

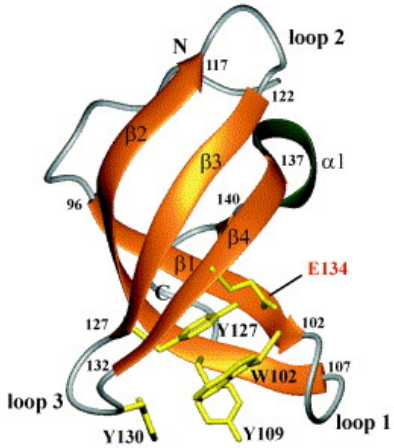


Figure 1. Crystal structure of SMN Tudor domain (Sprangers *et al*., 2003)

One of the proteins associated with maternal mRNPs is Tudor domain containing 6 (TDRD6), a mammalian protein which is present exclusively in early embryonic and germline cells (Mostafa *et al*., 2009). The protein contains six eTUDs and has an intrinsically disordered C-terminal tail (Figure 2). Its transcriptional and translational products occur in germline and early embryonic cells as maternal factors until ZGA but are not found in adult somatic cells (Ikema *et al*., 2002; Hiyoshi *et al.*, 2005). This implies that TDRD6 must serve some function during early embryogenesis but is no longer needed following ZGA. Loss of function of TDRD6 in fertilized eggs shows abnormal microtubule assembly and chromosome condensation during the cleavage of the embryos, resulting in cleavage arrest (Hiyoshi *et al*., 2005). Thus, TDRD6 is necessary in the numerous rounds of mitosis during the cleavage stage. Studies have shown the association of TDRD6 with FRGY2 protein,

another component of maternal mRNPs, as well as maternal RNA, thus indicating that TDRD6 is a component of maternal mRNPs (Mostafa *et al*., 2009).

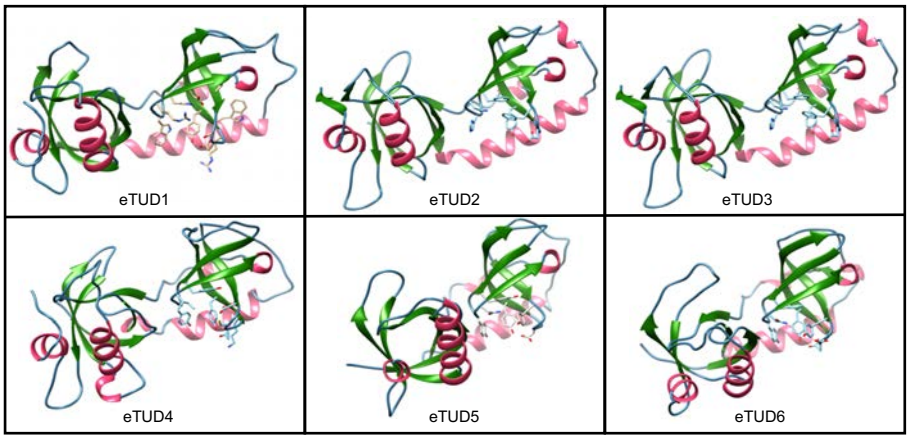


Figure 2. Phyre2 Protein Modeling of TDRD6 eTUD Domains

(iv) *Histone chaperones and PTMs*

During the cleavage stage of embryogenesis, DNA is rapidly replicated and divided (Laskey, 1985). In order for DNA to be condensed during mitosis, histones are required to maintain the DNA in its tightly wound chromatin form (Amodeo *et al*., 2015, Collart *et al*., 2013). Histones are a family of core (H2A, H2B, H3, and H4) and linker (H1 and H5) proteins that package DNA into its chromatin structure (Bannister and Kouzarides, 2011). 147 base pairs of negatively charged DNA wrap around a positively charged nucleosome. A nucleosome is an octamer comprised of two of each of the four core histone proteins, and it forms a "beads on a string" model, where double stranded DNA is wrapped around a nucleosome. Each nucleosome is flanked by linker histones, H1 and H5 (Luger *et al*., 1997). Like any protein, histones are translated by ribosomes in the cytoplasm, but since their primary role is to structurally organize DNA, they must be transported to the nucleus to serve their function. Furthermore, histones are extremely basic, and thus extremely reactive

proteins, so they must somehow be sequestered during transport from the cytoplasm to the nucleus to prevent nonspecific binding until they reach their target, DNA.

Histone chaperones serve both of these functions, as well as many others. Histone chaperones are negatively charged proteins that may bind, store, deposit, or transport histones (Laskey e*t al*., 1978). After a histone is translated, it is stored in a storage chaperone until it is ready to be transported to the nucleus (Finn *et al*., 2012). Histone chaperones may contain nuclear localization sequences, which aid in histone-chaperone transport to the nucleus (Falces *et al*., 2010). Histone chaperones also shield charge and prevent aggregation (Andrews *et al*., 2008). Most importantly, histone chaperones regulate the deposition of histones onto DNA (Jackson and Chalkley, 1981).

One method by which histone chaperones regulate the deposition and sequestration of histones, as well as modify their own functioning and transport, is by post translational modifications (PTMs) of the histone chaperones. These PTMs include, but are not limited to, acetylation (addition of an acetyl group), methylation (addition of a methyl group), and glutamylation (addition of a glutamate residue). These PTMs have been shown to enhance or depress interaction between histones and their chaperones (Onikubo *et al*., 2015, Calvert *et al*., 2008).

(v) *Npm2*

Nucleoplasmin (Npm2) is the predominant histone storage chaperone for histones H2A and H2B found exclusively in *Xenopus laevis* oocytes and through early stages of

embryogenesis (Bouleau *et al*., 2014). Like most histone chaperones, it has an acidic nature to bind histones and neutralize charge (Laskey *et al.*, 1978). The protein adopts a pentameric form, and it contains intrinsically disordered C and N termini (Dutta *et al*., 2001, Bañuelos *et al*., 2003) (Figure 3). Its tail contains three acidic stretches (A1, A2, and A3), each of which contains many negatively charged residues (Dutta *et al*., 2001). Studies have shown that the core is sufficient to bind histones, but the tail also engages in histone binding (Ramos *et al*., 2014).

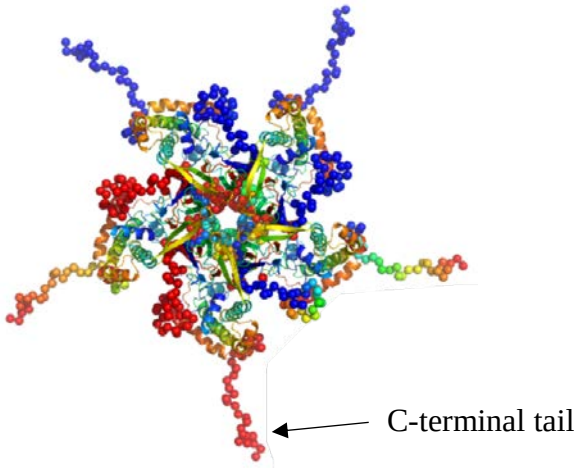


Figure 3. Npm2 adopts a pentameric form and has intrinsically disordered N and C termini (Warren *et al*., 2017)

Nucleoplasmin is extensively post-translationally modified. Both its C and N termini tails accumulate glutamylations and phosphorylations, both of which contribute, either directly or indirectly, to histone deposition and sequestration in the nucleus (Onikubo *et al*., 2015). Npm2 also notably contains a dimethylated arginine (R187me2) residue on its C-terminal tail (Wilczek *et al*., 2011). While studies have identified that this PTM contributes to enhanced deposition at higher histone mass to nucleoplasmin ratios, it does not cause a conformational change in Npm2 (Onikubo *et al*., 2015). We therefore hypothesized that there must be another function for R187me2.

(vi) *Hypothesis: TDRD6-Npm2 Interaction*

E2F proteins are a family of transcription factors best known for their ability to regulate the G1–S transition, and they also contain arginine methylation (Blais and Dynlacht, 2004). Comparing the sequences of the arginine-methylated motifs of both Npm2 and E2F revealed that Npm2 shares 100% identity with the binding motif of E2F in the TDRD1 structure, thus leading to the hypothesis that TDRD6 may also bind to Npm2. Furthermore, TDRD6 presence is highest in the oocyte and egg during early embryogenesis, but significantly decreases after ZGA (Ikema *et al*., 2002, Hiyoshi *et al*., 2005). Therefore, we hypothesized that after release of its histone cargo, the Npm2 Rme C-terminal tail can interact with maternal mRNPs by binding one or more of the TDRD6 extended Tudor domains. This would effectively de-repress the translational block of maternal mRNPs during early embryogenesis (Figure 4).

In this capacity, Npm2 would function as a developmental timer:

(1) Release of histones from Npm2 marks the completion of S-phase
(2) Npm2 binding TDRD6 triggers release of maternal RNA from maternal mRNPs
(3) Maternal RNA is translated, and leads to cell cleavage and progression to the next cell cycle

Our approach to test this hypothesis was to purify each individual TDRD6 extended Tudor domain and perform protein pulldown assays with the C-terminal 19 residues of the Npm2 tail with and without arginine methylation, in order to detect a potential interaction between the two.

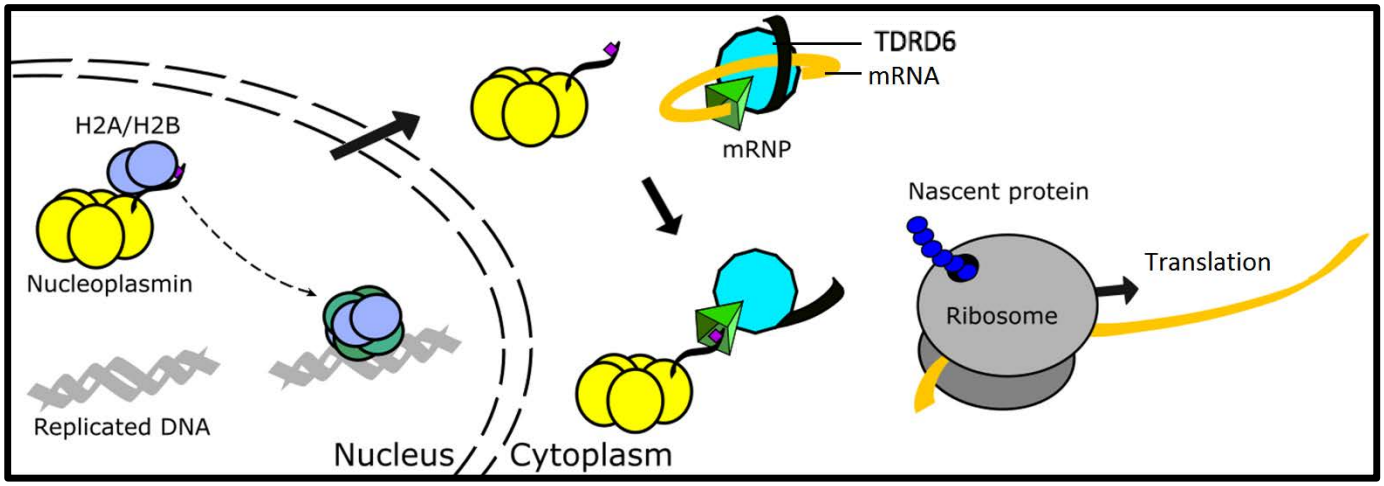


Figure 4. Schematic of hypothesis: as the nuclear envelope breakdown occurs, Npm2, now devoid of histones, can mix with cytoplasmic factors, such as translationally-repressed maternal mRNPs. One protein component of maternal mRNPs is the TDRD6 protein, whose aromatic binding cage can interact with the methylated arginine residue on the C-terminal tail of Npm2.

Materials and Methods

Protein Constructs

Three synthetic cDNA constructs, "eTUD12", "eTUD34", and "eTUD56" with optimized codons for expression in *Escherichia coli* were used as templates for PCR to amplify each individual eTUD domain (1 through 6). A StrepII-tag was added to the C-terminus of each eTUD to assist with purification. The inserts were gel purified and cloned into pRUTH5 using In-Fusion Cloning Kit (Clontech), resulting in plasmids 1.S2-6.S2 (Table 1). Chemically competent DH5α *E. coli* were transformed with plasmid, plated on Luria agar with 5mM Kanamycin, and incubated overnight at 37^{O}C. Successful transformants, verified by sequencing, were cultured overnight in 10mL Luria Broth containing 5mM Kanamycin. Plasmid DNA was isolated by Mini-prep kit (Qiagen).

Table 1. Plasmid constructs and corresponding TDRD6 protein

Name of Plasmid Construct	Name of Protein Produced by Construct
1.S2	eTUD1
2.S2	eTUD2
3.S2	eTUD3
4.S2	eTUD4
5.S2	eTUD5
6.S2	eTUD6

Protein Purification

eTUD 1.S2-6.S2 plasmids encoding protein constructs eTUDs 1-6 were transformed into chemically competent BL21 *E. coli*. Six 1L cultures of *E. coli* cells were grown at 37^{O}C until OD600 = 0.6, and then expression was induced with 1mM IPTG. The cells were grown overnight in Terrific Broth (TB) containing 5mM Chloramphenicol and 5mM Kanamycin. The cultures were spun down at 4,000rpm at 10^{O}C for 20 minutes. Cell pellets were resuspended in 30mL of lysis buffer (50mM Tris-Cl pH 8.0, 150mM NaCl, 5mM β-Mercaptoethanol, 2mM PMSF). Because the 4S.2 pellet was highly concentrated, the 30mL sample was divided between two tubes. 10mL of lysis buffer and 10mL of 1X Buffer (50mM Tris-Cl pH 8.0, 150mM NaCl) were added to each of the two tubes.

The cells were sonicated at 4^{O}C at 30%, 35%, 40%, and 45% amplitude. The tubes were sonicated for 30 seconds at 30% amplitude, then left in ice until the rest of the tubes were sonicated at 30%. This was repeated for the remaining amplitudes. The cells were then centrifuged at 14,000rpm for 45 minutes at 4^{O}C. The supernatant was transferred to a separate tube. (The supernatants from the two 4S.2 tubes were combined into a single sample.)

The insoluble pellets were each resuspended in denaturing buffer (6M Guanidine-HCl, 50mM Tris-Cl pH 8.0, 150mM NaCl, 5mM β-Mercaptoethanol, 5mM Imidazole, 2mM PMSF) to a final volume of 30mL. The samples were still viscous, so they were sonicated again at 30% amplitude for 30 seconds at 4^{O}C. The cells were centrifuged at 14,000rpm for 45 minutes at 4^{O}C. The supernatant was still very viscous, so the cells were sonicated again at 35% amplitude for 30 seconds.

4mL of Ni-NTA resin (2mL Ni-NTA) was added to each of the six samples. The tubes were incubated for 2 hours at 4^{O}C with rotation. The resin was spun at 700g for 1 minute. The supernatant was collected. The resin was washed in batch with 50mL of 5mM imidazole wash (5mM Imidazole, 50mM Tris-Cl pH 8.0, 150mM NaCl, 2mM PMSF, 8M Urea, 5mM β-Mercaptoethanol), then spun again at 700g for 1 minute. The supernatant was collected. The resin was washed in batch with 25mL of 15mM imidazole wash (15mM Imidazole, 50mM Tris-Cl pH 8.0, 150mM NaCl, 2mM PMSF, 8M Urea, 5mM β-Mercaptoethanol) and then poured through a plastic column. The flow through was collected. 2.5mL of the 150mM imidazole wash (150mM Imidazole, 50mM Tris-Cl pH 8.0, 150mM NaCl, 2mM PMSF, 8M Urea, 5mM β-Mercaptoethanol) was added to each column and allowed to flow through completely, repeated a second time, and the flow through was combined. 2.5mL of the 350mM imidazole wash (350mM Imidazole, 50mM Tris-Cl pH 8.0, 150mM NaCl, 2mM PMSF, 8M Urea, 5mM β-Mercaptoethanol) was added to each column and allowed to flow through completely, repeated a second time, and the flow through was combined. 8mL of chase buffer (50mM Tris-Cl pH 8.0, 150mM NaCl, 5mM β-

Mercaptoethanol) was added to each column to empty the resin of any remaining protein. The flow through was collected.

SDS-PAGE

Three 15% SDS-PAGE gels were run to check for presence of soluble protein. 1µL of each sample was added to 9µL of 1X SDS SB (300mM Tris-Cl, pH 6.8, 5mM EDTA, 70mM SDS, 42% Glycerol, 860mM β-Mercaptoethanol, 0.1µg/mL bromphenol blue) and heated at 90^{O}C for 15 minutes. The gels were run for 30 minutes, and then stained in Coomassie Blue stain and microwaved for 12 seconds. The gels were rocked for 5 minutes and then destained in destain buffer overnight.

Protein Dialysis

The 150mM imidazole wash, 350mM imidazole wash, and chase flow through from the 2.S2 construct were pooled. For the 5.S2 and 6.S2 constructs, the 150mM and 350mM washes were pooled. Each protein solution was placed in a dialysis bag. The dialysis bags were placed in 1L of 4M urea buffer (4M urea, 50mM Tris-Cl pH 8.0, 150mM NaCl, 5mM β-Mercaptoethanol) for 3 hours at 4^{O}C, and a stir bar was added. (7mL of the 2.S2 sample was held for a separate dialysis. See below.) This was repeated for 2M Urea buffer and 0M Urea buffer. The dialysis bags were then placed in a second beaker of 1L of 0M urea buffer overnight at 4^{O}C.

7mL of the 2.S2 construct soluble protein solution was placed in a dialysis bag and placed directly into the 0M Urea solution, skipping the stepwise dialysis. The stepwise dialysis was used instead since the protein was not soluble in 0M urea.

Following dialysis, the soluble protein solutions were spun down at 4,000rpm for 10 minutes. The solutions were concentrated to 5mL.

Size Exclusion Chromatography

Proteins eTUD2, eTUD5, and eTUD6 were each run through a Superdex 75 Increase column. 1mL was injected at about 15mg/mL. The flow was 0.5mL/min and 0.5mL fractions were obtained. The sizes of the proteins that were eluted were confirmed on 15% SDS-PAGE.

NDSB Refolding

10mL of denaturing buffer (50mM HEPES-KOH pH 7.5, 6M Guanidine-HCl, 25mM DTT) was added to 5mL of each construct's column flow through. The solutions were spun down in a Millipore tube to concentrate it down to 1mL. The solutions were then diluted to 1mg/mL using the denaturing buffer.

Under vigorous magnetic stirring, 1mL of each construct was quickly injected into 10mL of folding buffer (50mM HEPES pH 7.5, 200mM NaCl, 1mM DTT, 1M NDSB201). The solution continued to stir for 2 minutes, and was then incubated at $4^{O}C$ for 1 hour. The solution was then transferred to a dialysis bag and placed in 500mL of 0M urea buffer

(50mM Tris-Cl pH 8.0, 150mM NaCl, 5mM β-Mercaptoethanol). The soluble protein was spun down at 4,000rpm for 20 minutes in a Millipore tube to concentrate the protein down to less than 500uL.

The protein samples were run on a size exclusion chromatographer, and the sizes of the proteins that were eluted were confirmed on a 15% SDS-PAGE gel.

Protein Interaction Assay

1.8mL of GST-Resin (50% slurry) was spun down at 700g and the supernatant was discarded. 10mL of 1X Buffer (50mM Tris-Cl pH 8.0, 150mM NaCl, 5mM β-Mercaptoethanol) was added to the resin and inverted multiple times. The resin was spun down at 700g for 1 minute and the supernatant was discarded. The resin was resuspended in 900µL of 1X Buffer.

0.1mg/mL of the GST tagged proteins (GST.Npm2 A2, GST.Npm2 Tail, GST.Npm2 Tail.Rme, GST.Npm2 Tail.C19, GST.Npm2 Tail.C19.Rme) was added to 60µL of slurry and rocked overnight at 4^{O}C. All tubes were spun down at 700g for 1 minute and the supernatant was discarded. The resin was washed 3 times with 200µL 1X buffer and the supernatant was discarded each time.

10µM of binding protein (H2A/H2B, eTUD5, eTUD6) or 2.5µM eTUD2 was added to the respective resin tubes. The tubes were rotated at 4^{O}C for 28 hours. The samples were spun down at 700g for 1 minute and the supernatant was discarded. The resin was washed 8

times with 200µL of 1X buffer, spun down at 700g, and the supernatant was discarded each time. The resin was transferred to a new tube after the 7th wash. The protein was eluted with 30µL of 2X SDS SB (300mM Tris-Cl, pH 6.8, 5mM EDTA, 140mM SDS, 42% Glycerol, 860mM β-Mercaptoethanol, 0.1µg/mL bromphenol blue) and heated at 90^{O}C for 5 minutes. The tubes were spun down at 700g for 1 minute and the supernatant was collected. Protein interaction was confirmed on 15% SDS-PAGE gels.

Western Blot

Following SDS-PAGE, the proteins were transferred to a PVDF membrane. The membranes were rocked in 50mL of blocking buffer [1X PBST (Hyclone), 1% ECL Prime Blocking Reagent] for one hour. They were incubated in Mouse StrepII Tag Antibody (1:5,000) at 4^{O}C overnight, and then washed three times in Phosphate Buffered Saline Tween 20 (PBST) for 15 minutes. The membranes were then incubated in HRP conjugated anti-mouse secondary antibody (1:10,000) for 1 hour and then washed three times in PBST for 15 minutes. The membranes were stained in Enhanced Chemiluminescence (ECL) buffer and photographs were taken.

Results

TDRD6 Constructs

In order to study the *in vitro* interaction between TDRD6 and Npm2, purification of both proteins was required. The Npm2 protein had been previously purified in this lab (Onikubo *et* al., 2015), so we proceeded with purifying TDRD6. The complete TDRD6 protein contains 6 extended Tudor domains, and both its C and N termini are intrinsically

disordered (Chen *et al*., 2011) (Figure 5). The enormous size of the full-length protein made it incompatible for production in *E. coli* cells. Since this experiment was particularly focused on identifying the interaction between the Npm2 tail and a specific Tudor domain, each of the six eTUDs were produced separately. Each of the six constructs contained one specific extended Tudor domain.

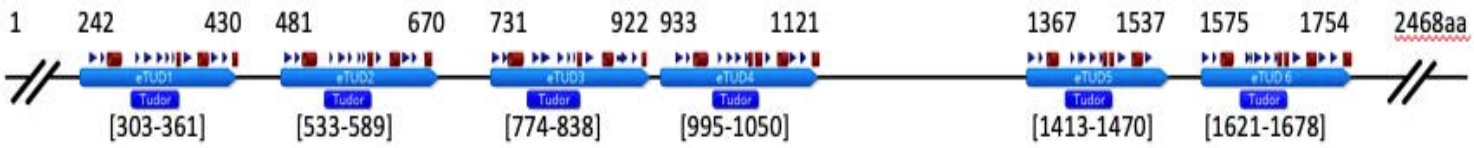


Figure 5. Amino acid location of core Tudor domains 1-6 and eTUDs 1-6 in TDRD6 protein

Solubilization of TDRD6 Protein

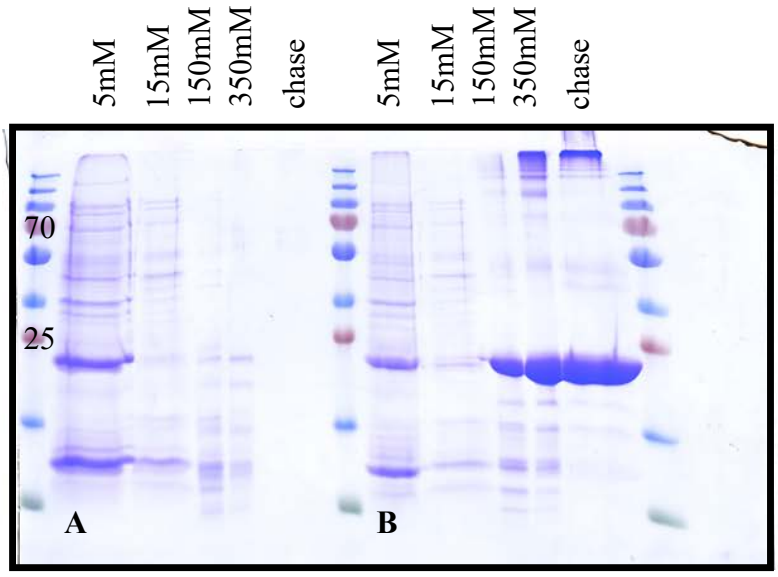
In this study, the goal was to obtain solubilized, purified, and folded eTUD protein to be tested for its interaction with Npm2. L-arginine and sarkosyl were first used to solubilize the protein. However, no protein was purified (data not shown). Instead, complete denaturation of all cellular proteins with Guanidine-HCl, followed by refolding using a non-detergent sulfobetaines (NDSB201) buffer, was used to form properly folded soluble eTUD protein for interaction testing.

Purification of TDRD6 Protein

After all cellular proteins were solubilized using Guanidine-HCl, the TDRD6 protein constructs had to be purified to be used for interaction testing. The DNA sequence coding for the eTUD protein was engineered with a histidine tag, allowing the protein to be purified on a nickel column. The solubilized cellular proteins were run through the column, washed with imidazole, and then eluted with 150mM and 350mM imidazole. The purity and quantity of the protein was checked by SDS-PAGE (Figure 6). Construct 2.S2, which encoded protein

eTUD2, yielded a high volume of protein from the 150mM, 350mM, and chase columns. The protein eluted was approximately 21kDa. The 5.S2 construct (encoding protein eTUD5) and 6.S2 construct (encoding protein eTUD6) each yielded a high volume of protein from the 150mM and 350mM columns, and the protein eluted was also approximately 20kDa. This is consistent with the calculated values of the size of each protein construct (Tables 2, 3).

The protein eluted for each of these three constructs appeared to be extremely pure, suggesting that the imidazole washes were successful in eluting the protein of interest. eTUD constructs 1.S2, 3.S2, and 4.S2 each yielded a small quantity of protein at approximately 21kDa. However, the quantity was considered insignificant for further testing, and thus the protocol continued with only three of the six constructs.



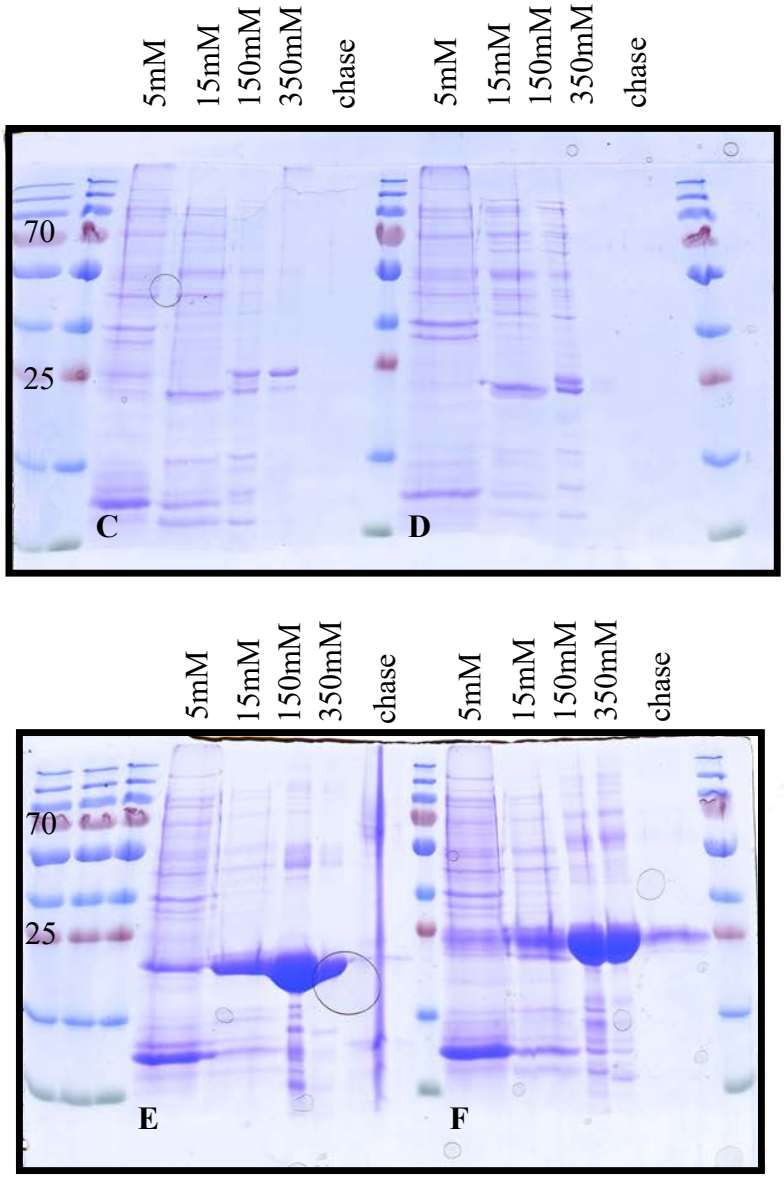


Figure 6. Verification of purified TDRD6 protein for constructs eTUD1 (**A**), eTUD2 (**B**), eTUD3 (**C**), eTUD4 (**D**), eTUD5 (**E**), and eTUD6 (**F**)

Table 2. Summary of eTUD construct sizes

Domain	Amino Acids	Molecular Weight (kDa)
eTUD1	189	21.76
eTUD2	190	21.59
eTUD3	192	21.48
eTUD4	189	21.56
eTUD5	171	19.22
eTUD6	180	20.31

Table 3. Summary of Tudor protein sizes

Domain	Amino Acids	Molecular Weight (kDa)
Tudor1	59	6.83
Tudor2	57	6.69
Tudor3	65	7.40
Tudor4	56	6.41
Tudor5	58	6.52
Tudor6	58	6.72

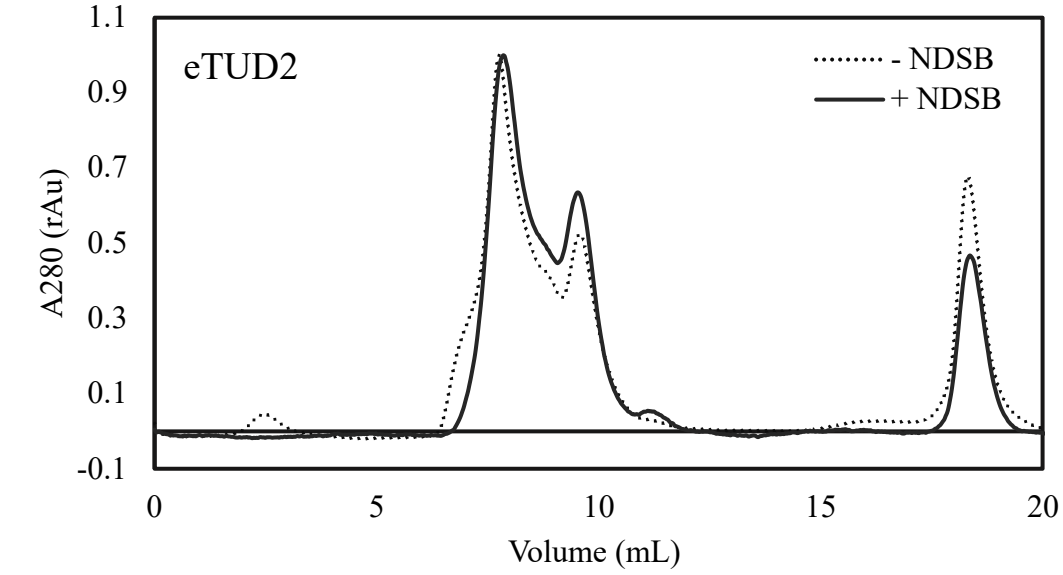
Refolding of TDRD6 Protein and Verification

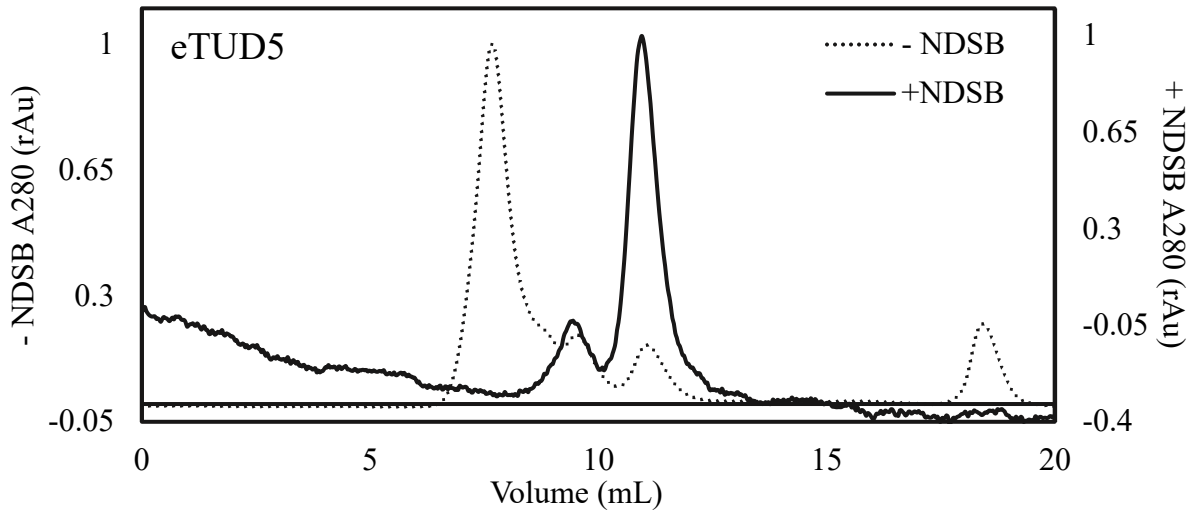
Since Guanidine-HCl is a denaturing agent, much of it had to be removed from the samples' buffer in order to refold the protein. Placing the protein directly into a dialysis buffer without any denaturing reagent caused the protein to crash, as demonstrated by the production of large amounts of precipitate, so stepwise dialysis in urea buffer was used instead to remove the denaturing agent. During stepwise dialysis, the protein was transferred from a Guanidine-HCl buffer to a 4M urea buffer, and then continuously moved to buffers with lower concentrations of urea until it was in a 0M urea buffer. At this point, the protein's buffer contained no denaturing agent.

We hypothesized that the protein may spontaneously refold to its native form after overnight stepwise dialysis in 0M urea buffer. To test this, the samples were concentrated and run on a Superdex 75 Increase column. While a small quantity of protein had spontaneously refolded and eluted at approximately 21kDa, the majority of the protein eluted at a higher molecular weight, indicating that the unfolded protein had aggregated.

The remaining protein that had not refolded spontaneously was refolded with the NDSB protocol. Following this NDSB refolding protocol, the samples were run on a

Superdex 75 Increase column again to check for presence of properly folded protein. While much of the protein was still unfolded and aggregated, as shown by a large peak at 7mL (corresponding to 30kDa), a much larger quantity eluted at approximately 11mL (corresponding to 20kDa), indicating that some of the protein had successfully refolded (Figure 7). While the NDSB protocol was most successful in refolding eTUD5 protein compared to the other protein constructs, sufficient eTUD2 and eTUD6 protein was refolded to continue with the interaction assay.





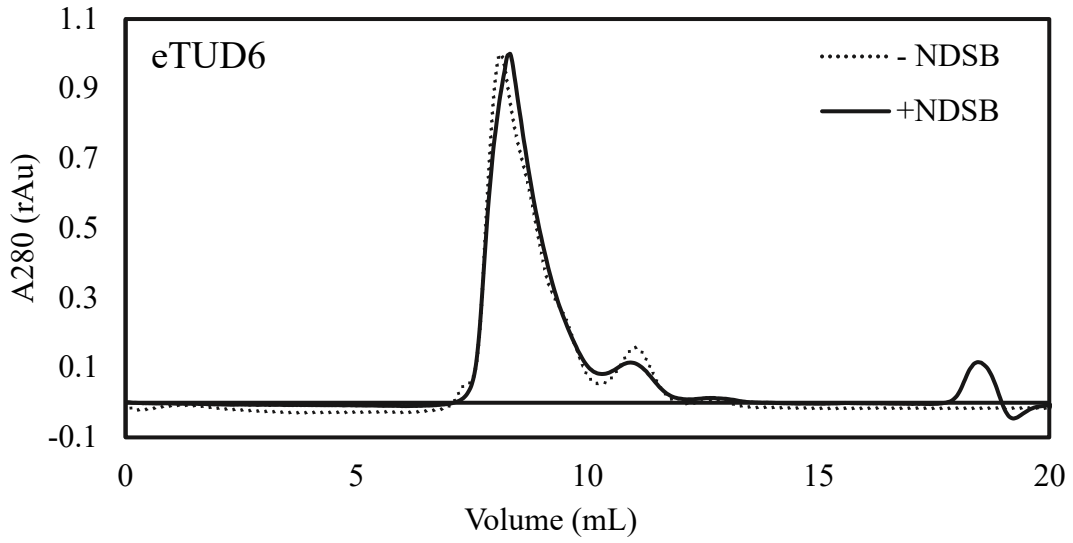


Figure 7. Verification of increase in refolded TDRD6 protein following NDSB refolding protocol for eTUD2, eTUD5, and eTUD6. Graphs are normalized to maximum absorbance by relative absorbance units (rAu)

The fractions that contributed to the 20kDa peak were checked for purity and size of the protein by SDS-PAGE. Results indicated that the protein was pure and had properly refolded. (Figure 8). The stepwise dialysis and NDSB refolding protocols were repeated until sufficient protein was collected for eTUD2, eTUD5, and eTUD6 to be used for an interaction assay.

Figure 8. Verification of Refolded TDRD6 eTUD5

Interaction Between eTUDs 2, 5, 6 and Npm2

An interaction assay, where Npm2 and eTUD protein were incubated together and pulled down using GST resin, was conducted to test if Npm2 tail pulls down one or more of the TDRD6 constructs. All Npm2 constructs were previously purified by this lab. Following the interaction assay, the samples were analyzed by SDS-PAGE gel to check the purity of the inputs and quality of the resin washes (Figure 10). “Resin only” controls were run to ensure that only GST fused proteins attached to the resin, while inputs without a GST tag did not bind to the resin. “GST.Npm2 A2” (Lane 9), “GST.Npm2 Tail” (Lane 10), “GST.Npm2 Tail.Rme” (Lane 11), “GST.Npm2 Tail.C19” (Lane 27), and “GST.Npm2 Tail.C19.Rme” (Lane 28) each showed a positive band in the “resin only” control lanes, indicating that the GST tagged protein properly attached to the resin. In contrast, each of the TDRD6 “resin only” lanes were empty (Lanes 12, 13, 14), indicating that the proteins lacking a GST tag did not attach to the resin. While H2A/H2B did not have a GST tag, a small band appeared in the “resin only” lane (Lane 8). However, it was substantially smaller than the H2A/H2B input control lane (Lane 1), so this small quantity was considered background for future H2A/H2B experiments.

Five models of the Npm2 tail were tested: A2, Tail, Tail.Rme, Tail.C19, and Tail.C19.Rme (Figure 9). A2 refers to the length of tail that begins at the core domain and ends at the second acidic (A2) patch (Lanes 15-18). Each acidic patch contains a high volume of acidic residues in order to bind to charged histones and fulfill nucleoplasmin’s role as a histone chaperone. The GST fused Npm2 A2 sample is known to bind to H2A/H2B histones

and was therefore used as a positive control (Wedlich and Dreyer, 1988). GST.Npm2 A2 pulled down H2A/H2B (Lane 15).

Tail refers to the full length nucleoplasmin tail (Lanes 19-21, 22-24). Previous literature has shown that in its native form, the tail folds back and attaches to the A2 patch of C-terminal tail (Onikubo *et al*., 2015). Therefore, a cut section of the tail, C19, was used in this study to avoid binding competition between TDRD6 protein and the A2 patch of the Npm2 protein. Tail.C19 refers to the 19 C-terminal residues on the Npm2 tail (Lanes 29-31, 32-34). The methylated arginine residue of interest is located on this part of the protein. In order to specifically identify whether the methylated arginine residue is responsible for interaction between Npm2 and TDRD6, the full-length Tail and C19 were used in the assay with and without the methylated arginine residue.

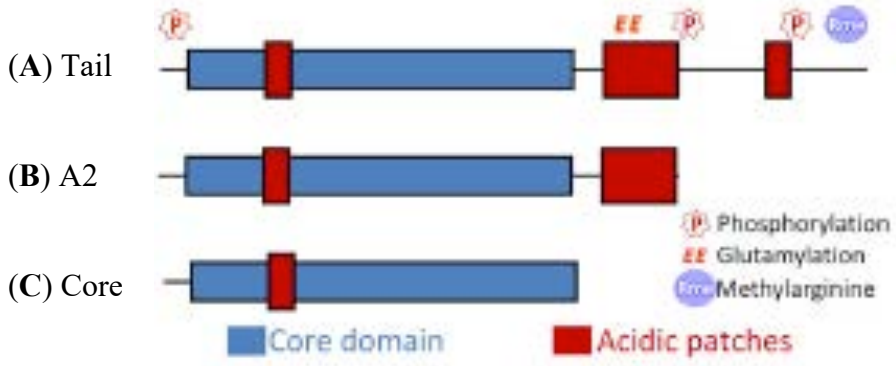


Figure 9. Schematic of nucleoplasmin tail constructs used in interaction assay (Onikubo *et al.*, 2015)

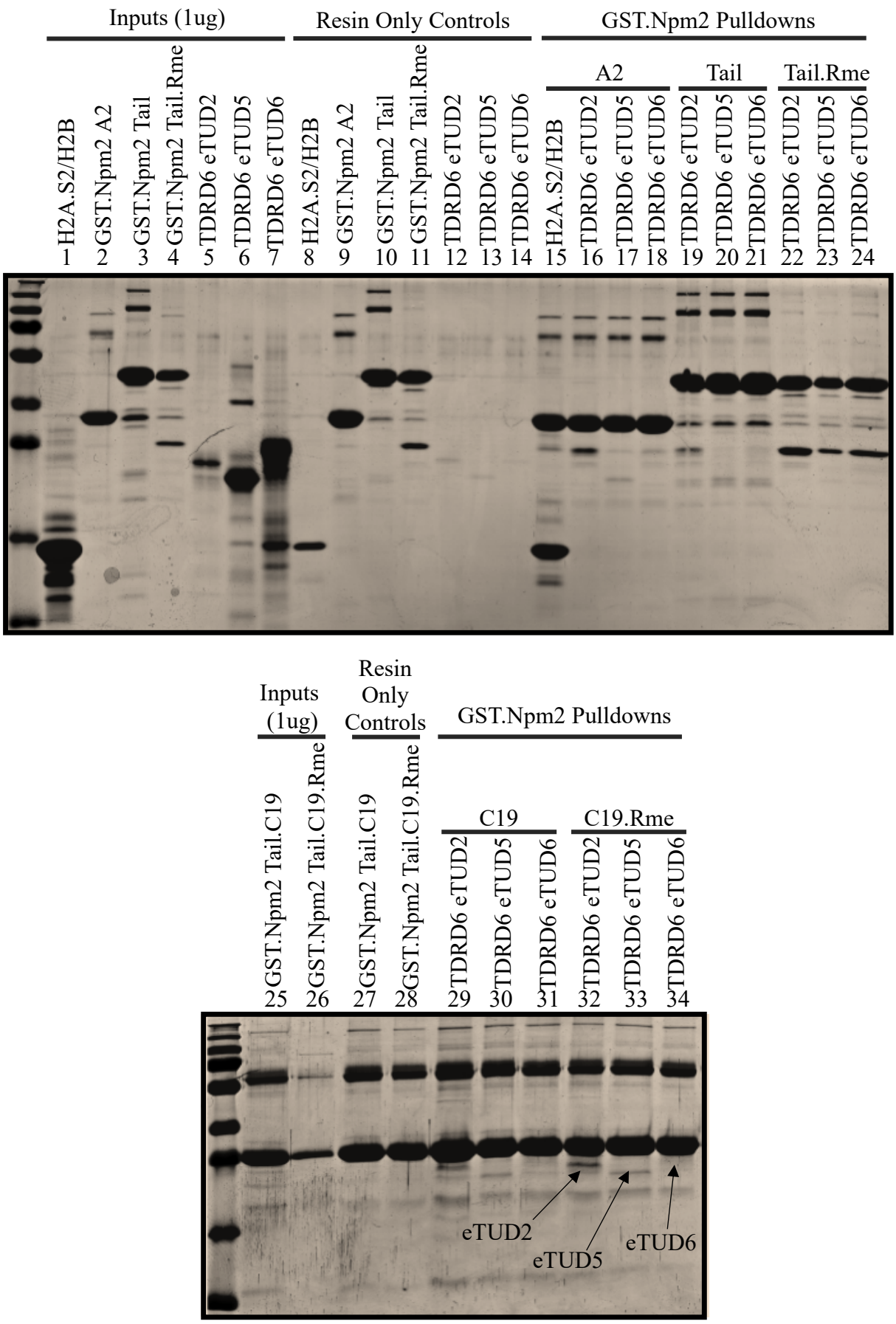


Figure 10. Assay of TDRD6-Npm2 Protein Interaction

Determination of Rme Necessity in TDRD6-Npm2 Interaction

The only interactions observed in the GST pulldown were between the TDRD6 eTUDs and C19, both with and without the methylated arginine residue. All other Npm2 constructs did not appear to have any interaction with TDRD6. Coomassie staining alone was not reliable enough to determine if there was an interaction, so we continued with a Western Blot. Only the C19 construct inputs (Lanes 25, 26), "resin only" controls (Lanes 27, 28), and "GST.Npm2 pulldown" lanes (Lanes 29-34) were blotted to test for the presence of TDRD6 protein.

The TDRD6 constructs all had a Strep II tag, and therefore the primary antibody was targeted against the Strep II tag. A visibly dark band at approximately 20kDa appeared in the "GST.Npm2 Tail.C19.Rme"- "TDRD6 eTUD2" lane. Lighter bands also appeared in the "GST.Npm2 Tail.C19.Rme"- "TDRD6 eTUD5" and "GST.Npm2 Tail.C19.Rme"- "TDRD6 eTUD6" lanes. No bands appeared in the C19 lanes without the methylated arginine residue (Figure 11).

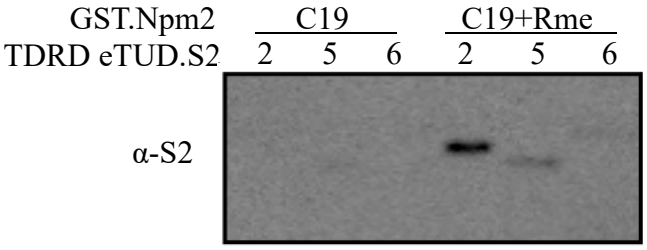


Figure 11. Verification of Npm2 Pulldown of eTUDs 2, 5, and 6 with and without Rme

Discussion

Methylated arginine residue necessary for Npm2-TDRD6 interaction

Histones tend to have many basic residues and are thus positively charged. Histone chaperones, in contrast, are thought to rely on their acidic nature to bind histones and neutralize charge (Laskey *et al.*, 1978). The Npm2 tail has three acidic stretches, as well as many negatively charged post translational modifications (Ramos *et al*., 2014).

As the oocyte continues through development to an egg, nucleoplasmin continues to be post-translationally modified. Studies have found an increase in glutamylation and phosphorylation in the nucleoplasmin tail from oocyte to egg, which is predicted to contribute to its function as a histone storage chaperone (Onikubo *et al*., 2015). The methylation of an arginine residue at the end of the Npm2 tail is not essential in assisting the protein in its function as a histone chaperone. Whereas phosphorylation is essential for histone sequestration, arginine methylation is merely responsible for enhanced deposition at higher histone mass to nucleoplasmin ratios (Onikubo *et al.*, 2015). It is therefore likely that the methylated arginine residue serves another purpose, which was tested in this study.

Our results show that, *in vitro*, Npm2 and eTUD domains of TDRD6 interact. Moreover, the Npm2 C19.Rme tail, but not the Npm2 C19 tail, binds to proteins eTUD2, eTUD5, and eTUD6 (Figures 10, 11). Thus, the methylated arginine residue is likely responsible for the interaction. Our model proposes that upon depositing histones onto newly replicated DNA, the nucleoplasmin protein leaves the nucleus, and the methylated arginine residue on its C-terminal tail interacts with at least one Tudor domain of TDRD6 (Figure 4).

Post-fertilization embryonic *X. laevis* cells are transcriptionally silenced, thus requiring them to rely on maternal stores for proteins and mRNAs (Laskey, 1985). Because TDRD6 protein has been identified as a component of maternal mRNPs, it is likely that the Npm2-TDRD6 interaction is related to the release of maternal mRNA from the maternal mRNPs (Mostafa *et al*., 2009).

Npm2 has highest affinity for eTUD2

Tudor domain proteins are categorically known for binding methylated proteins and methylated DNA (Taverna *et al.*, 2007). Npm2 C19.Rme binds most strongly to eTUD2, even though its concentration in the assay was 2.5μM (due to insufficient protein purification), whereas the concentration of the remaining eTUD proteins in the assay were 10μM (Figure 11). All of the eTUD constructs composing TDRD6 have structurally similar antiparallel β-barrel cores with slight modifications in their primary sequence (Chen *et al*., 2011). It is therefore likely that *in vivo*, Npm2 binds most strongly to eTUD2, but maintains a low binding affinity for the other Tudor domains.

Hyperphosphorylation of Npm2 tail regulates TDRD6 release

Studies show that the C-terminal tail of Npm2 interacts with the A2 patch and blocks histone accessibility, thus causing a mass-ratio-dependent histone deposition pattern where the C-terminal tail and histones compete for A2 binding. Phosphorylation of the A2 patch disrupts this interaction and leaves the A2 patch free to sequester or deposit histones (Onikubo *et al*., 2015). Hyperphosphorylation begins at fertilization and lasts until MBT, which occurs simultaneously with ZGA (Laskey *et al*., 1978). Our results indicated that only

C19.Rme, and not the arginine methylated full-length tail, interacted with the eTUD protein. Combined with this study, we predict that *in vivo*, upon fertilization, hyperphosphorylation would release the C-terminal tail from A2, thus allowing for (1) the Npm2 tail to interact with histones, and (2) the methylated arginine residue to interact with TDRD6. Furthermore, we hypothesize that at ZGA, the Npm2-TDRD6 interaction would end, since maternal RNA would no longer be required. This is consistent with the finding that at ZGA, Npm2 hyperphosphorylation is lost, but other modifications, such as glutamylation and arginine methylation, are retained (Onikubo and Shechter, 2016). Upon losing hyperphosphorylation, the C-terminal tail would interact with the A2 patch, and thus block the methylated arginine residue from further interaction with TDRD6.

Future Studies

To further study the interaction between Npm2 and TDRD6, the three remaining eTUD constructs (1.S2, 4.S2, and 5.S2) must be produced, purified, solubilized, and refolded. While denaturation with Guanidine-HCl was sufficient for eTUDs 2, 5, and 6, this method did not yield a high volume of soluble protein for all 6 constructs (Figure 6). The three remaining eTUD constructs should then be used in a pulldown assay, similar to the one outlined in this study, and blotted to test for interaction between C19.Rme and the Tudor domain proteins.

Furthermore, the complete TDRD6 protein should be grown, preferably in eukaryotic cells that have the capacity for such large proteins. An x-ray crystallography experiment using full length TDRD6 with full length nucleoplasmin should be conducted to see a crystal

structure of the proteins complexed together and to test which specific Tudor domains are most attractive to Npm2. The crystallography experiment can also be conducted with different post translational modifications to nucleoplasmin, such as with and without the methylated arginine residue. This may elucidate whether the methyl group is directly related to the interaction, or whether it changes the protein's secondary structure and causes a binding interaction.

Finally, maternal mRNPs should be isolated from *X. laevis* extract and blotted for Npm2 to test if Npm2 complexes with TDRD6 as part of a maternal mRNP.

Limitations of this study

While these results confirm our hypothesis that there is an interaction between Npm2 and TDRD6, we must also recognize that the presence of a methylated residue on a different protein may have yielded the same result. Since Tudor domain proteins have such a high affinity for methylated side groups, our results may not necessarily be an indication of an *in vivo* interaction. In order to determine if there is an *in vivo* interaction, a co-immunoprecipitation assay utilizing cleavage stage *X. laevis* cell lysate and an anti-TDRD6 antibody is required.

Conclusion

The cleavage stage of early embryogenesis in *Xenopus* presents a fascinating model for testing the activation of translationally repressed maternal mRNPs. TDRD6 is a protein component of maternal mRNPs. Npm2 is the predominant H2A/H2B histone storage

chaperone in *Xenopus* eggs. Since Tudor domain proteins, such as TDRD6, are known for binding methylated ligands, we predicted that Npm2, which is post translationally modified with a methylated arginine residue, may bind to TDRD6 in maternal mRNPs upon fertilization. We have demonstrated an *in vitro* interaction between the C-terminal tail of Npm2 and eTUD2, eTUD5, and eTUD6. We have shown that the tail's methylated arginine residue is responsible for the interaction. Furthermore, we have demonstrated that the Npm2 tail has the highest binding affinity to eTUD2 as compared to eTUD5 and eTUD6. Based on previous research, it is likely that hyperphosphorylation of the A2 patch on the Npm2 tail regulates the interaction of Npm2 and TDRD6. This regulation would effectively remove the translational block of the maternal transcripts in maternal mRNPs upon fertilization, but then end the interaction upon ZGA. Further studies using the full length TDRD6 protein, as well as a co-immunoprecipitation assay utilizing cleavage stage *X. laevis* cell lysate and an anti-TDRD6 antibody, may present the broader picture of how TDRD6 and Npm2 interact *in vivo*.

Acknowledgements

I would like to thank Dr. Cynthia Wachtell for my informative and exciting experience in the S. Daniel Abraham Honors Program at Stern College for Women. Thank you to my mentor, Dr. Alyssa Schuck, for her constant guidance and invaluable insights throughout the honors thesis process. Thank you to Dr. Harvey Babich for reading my work, as well as his help and encouragement throughout Stern college. I would also like to thank Benjamin Lorton, Dr. David Shechter, and the entire Shechter lab at Albert Einstein College of Medicine. Thank you for challenging me to think like a scientist, giving me the opportunity to take on this project, and for providing me with data and information to compose my honors thesis. Finally, I would like to express my endless appreciation to my parents, Mojgan and Iraj Shokrian. You have always supported my endeavors with complete devotion, encouraged me to be my best, and have pushed me to achieve more when I did not think it was possible.

References

Amodeo, Amanda A., et al. “Histone Titration against the Genome Sets the DNA-to-Cytoplasm Threshold for the Xenopus Midblastula Transition.” *Proceedings of the National Academy of Sciences*, vol. 112, no. 10, 2015, doi:10.1073/pnas.1413990112.

Andrews, Andrew J., et al. “A Thermodynamic Model for Nap1-Histone Interactions.” *Journal of Biological Chemistry*, vol. 283, no. 47, 2008, pp. 32412–32418., doi:10.1074/jbc.m805918200.

Bannister, Andrew J, and Tony Kouzarides. “Regulation of Chromatin by Histone Modifications.” *Cell Research*, vol. 21, no. 3, 2011, pp. 381–395., doi:10.1038/cr.2011.22.

Bañuelos, Sonia, et al. “Phosphorylation of Both Nucleoplasmin Domains Is Required for Activation of Its Chromatin Decondensation Activity.” *Journal of Biological Chemistry*, vol. 282, no. 29, 17 May 2007, pp. 21213–21221., doi:10.1074/jbc.m702842200.

Blais, Alexandre, and Brian David Dynlacht. “Hitting Their Targets: an Emerging Picture of E2F and Cell Cycle Control.” *Current Opinion in Genetics & Development*, vol. 14, no. 5, 2004, pp. 527–532., doi:10.1016/j.gde.2004.07.003.

Bouleau, Aurélien, et al. “Maternally Inherited npm2 MRNA Is Crucial for Egg Developmental Competence in Zebrafish1.” *Biology of Reproduction*, vol. 91, no. 2, 1 Aug. 2014, doi:10.1095/biolreprod.114.119925.

Buhler, D., et al. “Essential Role for the Tudor Domain of SMN in Spliceosomal U SnRNP Assembly: Implications for Spinal Muscular Atrophy.” *Human Molecular Genetics*, vol. 8, no. 13, 1999, pp. 2351–2357., doi:10.1093/hmg/8.13.2351.

Calvert, M. E. K., et al. “Phosphorylation by Casein Kinase 2 Regulates Nap1 Localization and Function.” *Molecular and Cellular Biology*, vol. 28, no. 4, 2007, pp. 1313–1325., doi:10.1128/mcb.01035-07.

Chen, Chen, et al. “Deciphering Arginine Methylation: Tudor Tells the Tale.” *Nature Reviews Molecular Cell Biology*, vol. 12, no. 10, 2011, pp. 629–642., doi:10.1038/nrm3185.

Collart, C., et al. “Titration of Four Replication Factors Is Essential for the Xenopus Laevis Midblastula Transition.” *Science*, vol. 341, no. 6148, 1 Aug. 2013, pp. 893–896., doi:10.1126/science.1241530.

Davidson, Eric H. “From Genome to Embryo: The Regulation of Gene Activity in Early Development.” *Gene Activity in Early Development*, 1986, pp. 1–44., doi:10.1016/b978-0-12-205161-6.50005-2.

Dutta, S., et al. “The Crystal Structure of Nucleoplasmin-Core.” *Molecular Cell*, vol. 8, no. 4, 1 Oct. 2001, pp. 841–853., doi:10.2210/pdb1k5j/pdb.

Falces, Jorge, et al. “Recognition of Nucleoplasmin by Its Nuclear Transport Receptor Importin α/β: Insights into a Complete Import Complex.” *Biochemistry*, vol. 49, no. 45, 16 Nov. 2010, pp. 9756–9769., doi:10.1021/bi101179g.

Finn, Ron M., et al. “Vertebrate Nucleoplasmin and NASP: Egg Histone Storage Proteins with Multiple Chaperone Activities.” *The FASEB Journal*, vol. 26, no. 12, 2012, pp. 4788–4804., doi:10.1096/fj.12-216663.

Hiyoshi, Masateru, et al. “Involvement of Xtr (Xenopus Tudor Repeat) in Microtubule Assembly around Nucleus and Karyokinesis during Cleavage in Xenopus Laevis.” *Development, Growth and Differentiation*, vol. 47, no. 2, 2005, pp. 109–117., doi:10.1111/j.1440-169x.2005.00787.x.

Horvay, Katja, et al. “Xenopus Dead End MRNA Is a Localized Maternal Determinant That Serves a Conserved Function in Germ Cell Development.” *Developmental Biology*, vol. 291, no. 1, 2006, pp. 1–11., doi:10.1016/j.ydbio.2005.06.013.

Houston, DW, et al. “A Xenopus DAZ-like Gene Encodes an RNA Component of Germ Plasm and Is a Functional Homologue of Drosophila Boule.” *Development*, vol. 125, no. 2, 1998, pp. 171–180.

Ikema, Yasuko, et al. “Two Novel Genes Expressed InXenopus Germ Line: Characteristic Features of Putative Protein Structures, Their Gene Expression Profiles and Their Possible Roles in Gametogenesis and Embryogenesis.” *Molecular Reproduction and Development*, vol. 62, no. 4, 2002, pp. 421–430., doi:10.1002/mrd.90003.

Jackson, Vaughn, and Roger Chalkley. “Histone Segregation of Replicating Chromatin.” *Biochemistry*, vol. 24, no. 24, 1985, pp. 6930–6938., doi:10.1021/bi00345a027.

Kimble, Judith, and Sarah L. Crittenden. “Controls of Germline Stem Cells, Entry into Meiosis, and the Sperm/Oocyte Decision InCaenorhabditis Elegans.” *Annual Review of Cell and Developmental Biology*, vol. 23, no. 1, 2007, pp. 405–433., doi:10.1146/annurev.cellbio.23.090506.123326.

Laskey, R. A., et al. “Analysis of Chromosome Replication with Eggs of Xenopus Laevis.” *Genetic Engineering: Principles and Methods*, 1985, pp. 135–148., doi:10.1007/978-1-4684-4973-0_6.

Laskey, R. A., et al. “Nucleosomes Are Assembled by an Acidic Protein Which Binds Histones and Transfers Them to DNA.” *Nature*, vol. 275, no. 5679, 1978, pp. 416–420., doi:10.1038/275416a0.

Luger, Karolin, et al. “Crystal Structure of the Nucleosome Core Particle at 2.8 Å Resolution.” *Nature*, no. 389, 18 Sept. 1997, pp. 251–260.

Maurer-Stroh, Sebastian, et al. “The Tudor Domain ‘Royal Family’: Tudor, Plant Agenet, Chromo, PWWP and MBT Domains.” *Trends in Biochemical Sciences*, vol. 28, no. 2, 2003, pp. 69–74., doi:10.1016/s0968-0004(03)00004-5.

Mostafa, Md. Golam, et al. “Xtr, a Plural Tudor Domain-Containing Protein, Coexists with FRGY2 Both in Cytoplasmic MRNP Particle and Germ Plasm in Xenopus Embryo: Its Possible Role in Translational Regulation of Maternal MRNAs.” *Development, Growth & Differentiation*, vol. 51, no. 6, 2009, pp. 595–605., doi:10.1111/j.1440-169x.2009.01121.x.

O'farrell, Patrick H., et al. “Embryonic Cleavage Cycles: How Is a Mouse Like a Fly?” *Current Biology*, vol. 14, no. 1, 2004, doi:10.1016/j.cub.2003.12.022.

Onikubo, Takashi, and David Shechter. “Chaperone-Mediated Chromatin Assembly and Transcriptional Regulation in Xenopus Laevis.” *The International Journal of Developmental Biology*, vol. 60, no. 7-8-9, 13 June 2016, pp. 271–276., doi:10.1387/ijdb.130188ds.

Onikubo, Takashi, et al. “Developmentally Regulated Post-Translational Modification of Nucleoplasmin Controls Histone Sequestration and Deposition.” *Cell Reports*, vol. 10, no. 10, 17 Mar. 2015, pp. 1735–1748., doi:10.1016/j.celrep.2015.02.038.

Ramos, I., et al. “The Intrinsically Disordered Distal Face of Nucleoplasmin Recognizes Distinct Oligomerization States of Histones.” *Nucleic Acids Research*, vol. 42, no. 2, 2013, pp. 1311–1325., doi:10.1093/nar/gkt899.

Richter, Joel D., and L. Dennis Smith. “Reversible Inhibition of Translation by Xenopus Oocyte-Specific Proteins.” *Nature*, vol. 309, no. 5966, May 1984, pp. 378–380., doi:10.1038/309378a0.

Schier, A. F. “The Maternal-Zygotic Transition: Death and Birth of RNAs.” *Science*, vol. 316, no. 5823, 20 Apr. 2007, pp. 406–407., doi:10.1126/science.1140693.

Sprangers, Remco, et al. “High-Resolution X-Ray and NMR Structures of the SMN Tudor Domain: Conformational Variation in the Binding Site for Symmetrically Dimethylated

Arginine Residues." *Journal of Molecular Biology*, vol. 327, no. 2, 21 Mar. 2003, pp. 507–520., doi:10.1016/s0022-2836(03)00148-7.

Sun, Liangliang, et al. "Quantitative Proteomics of Xenopus Laevis Embryos: Expression Kinetics of Nearly 4000 Proteins during Early Development." *Scientific Reports*, vol. 4, no. 1, 14 Mar. 2014, doi:10.1038/srep04365.

Tadros, W., and H. D. Lipshitz. "The Maternal-to-Zygotic Transition: a Play in Two Acts." *Development*, vol. 136, no. 18, 2009, pp. 3033–3042., doi:10.1242/dev.033183.

Taverna, Sean D, et al. "How Chromatin-Binding Modules Interpret Histone Modifications: Lessons from Professional Pocket Pickers." *Nature Structural & Molecular Biology*, vol. 14, no. 11, 2007, pp. 1025–1040., doi:10.1038/nsmb1338.

Warren, Christopher, et al. "Dynamic Intramolecular Regulation of the Histone Chaperone Nucleoplasmin Controls Histone Binding and Release." *Nature Communications*, vol. 8, no. 1, 20 Dec. 2017, doi:10.1038/s41467-017-02308-3.

Wedlich, Doris, and Christine Dreyer. "The Distribution of Nucleoplasmin in Early Development and Organogenesis of Xenopus Laevis." *Cell and Tissue Research*, vol. 254, no. 2, 1988, doi:10.1007/bf00225802.

Wilczek, Carola, et al. "Protein Arginine Methyltransferase Prmt5-Mep50 Methylates Histones H2A and H4 and the Histone Chaperone Nucleoplasmin InXenopus LaevisEggs." *Journal of Biological Chemistry*, vol. 286, no. 49, 2011, pp. 42221–42231., doi:10.1074/jbc.m111.303677.