4-2021

Evaluation of Two Artifacts related with the duration of RNA-protein immunoprecipitation experiments

Kevin Avanzino

Evaluation of Two Artifacts related with the duration of RNA-protein immunoprecipitation experiments

by

Kenneth Avanzino

Undergraduate honors thesis under the direction of

Dr. Anastasios Vourekas

Department of Biological Science

Submitted to the LSU Roger Hadfield Ogden Honors College in partial fulfillment of

the Upper Division Honors Program.

April 2021

Louisiana State University

& Agricultural and Mechanical College

Baton Rouge, Louisiana

Abstract: This study aimed to create a shortened Ribonucleoprotein Immunoprecipitation (RNP-IP) protocol that could be done in tandem with a traditional CLIP-seq protocol to answer the question if CLIP-seq reads from the Ribonucleoprotein DDX3X represents accurately the full length mRNA boudn by the protein in the cells, or the mRNA population is skewed and or affected by RNases released after cell lysis and present during the immunoprecipitation. Three different RNP-IP incubation durations of 2 minutes, 30 minutes, and 3 hours were utilized and the two artifacts of reassociation and RNA degradation were evaluated via Western blot analysis, Nanodrop spectroscopy, UREA Page Gel electrophoresis and RT-qPCR. The data showed that two minutes was sufficiently long enough to achieve quality protein yield equivalent to a 3-hour incubation, but RNA yield significantly increased at a disproportionate rate with increasing incubation duration. The data showed over a fivefold difference in RNA abundance between 2 minutes 3 hours incubation. This suggests that increasing incubation duration to attain greater protein yield leads to diminishing returns by greatly exacerbating artefactual reassociation. Interestingly, this study was not able to identify degradation fragments that were present using RT-qPCR, and instead observed greatly increased mRNA abundance in the 3-hour incubation durations. In addition, DDX3X in prolonged IPs showed biases for certain genes to such a high degree that needs to be considered for the design of immunoprecipitation experiments under denaturing conditions that prevent RNA reassociation like CLIP-seq.

Rational/Background:

RNA serves as a critical nexus point that ties the information stored in the genome to the protein molecules it will subsequently encode, and the manner in which RNA is metabolized allows for fine-tuned gene expression necessary for the dynamic physiology of a living cell. One group of enzymes that are essential drivers in RNA metabolism are RNA helicases. RNA helicases are crucial to the cellular processing of RNA, i.e., transcription, translation, and mRNA degradation. Their general roles include remodeling of RNA secondary structures and/or RNA protein complexes associated with the biological processes mentioned above. One group of RNA helicases that have been an essential topic of discussion are DEAD-box helicases (Linder & Jankowsky, 2011). These highly conserved enzymes make up the largest RNA helicase group and are vital to RNA metabolism, with many having a variety of functions beyond the simple unraveling of an RNA strand (Caruthers & McKay, 2002).

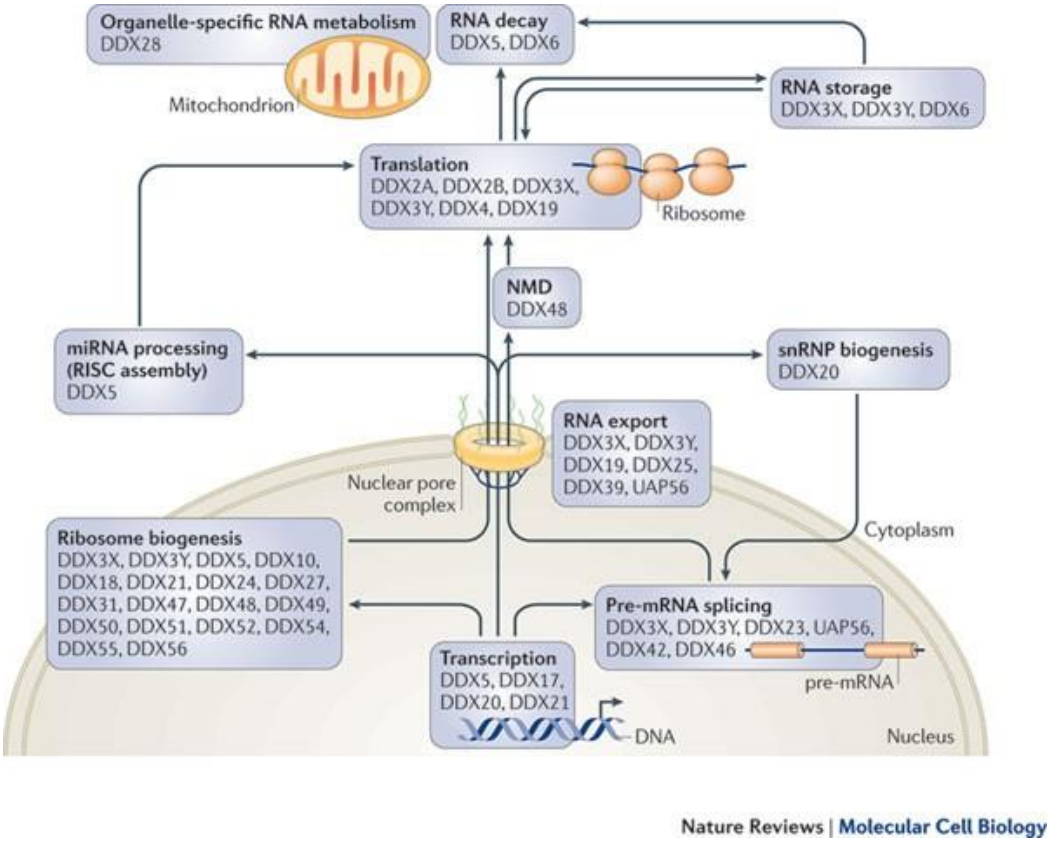

Figure 1: Linder& Jankowsky, Nature Reviews Molecular Cell Biology, 2011

They are characterized by two identical and highly conserved helicase cores that contain at least 12 conserved sequence motifs, most notable of which is Motif II (Linder et al., 1989; Fairman-Williams et al., 2010; Caruthers & McKay, 2002) which consists of an Asp-Glu-Ala-Asp motif abbreviated as DEAD, which is where this class of helicases gets their name (Linder & Jankowsky, 2011). The two helicase domains form a cleft between each other containing an ATP-binding site and an associated RNA binding site that runs across the two-helicase domains (Linder & Jankowsky, 2011). It has been shown that they bind RNA in a highly conserved fashion over five nucleotide motifs through protein to phosphate backbone interactions. In addition to a conserved helicase core, as well as RNA binding sites, and ATP

binding sites, DEAD-box helicases harbor auxiliary domains surrounding the helicase core, which is where the specialized functions of DEAD-box helicases are believed to come from (Linder & Jankowsky, 2011).

DDX3X, a member of the DEAD-box helicases, has proven to be an area of great interest regarding DEAD-box helicases. No other DEAD-box helicase is found over a wider variety of RNA metabolism than DDX3X, but the primary function of DDX3X remains elusive and often contradictory (Figure 1). Its mutation and dysregulation have been associated with a variety of diseases, most notably in cancer etiology. However, researchers cannot agree on whether or not it is a tumor-suppressing gene or an oncogene. DDX3X has been shown to have a role in translation initiation and stress granule formation. Various studies have shown that it can associate with the eIF4F complex to promote translation of a very select group of mRNAs, function as an eIF4E inhibitor to suppress translation, as well as potentially direct stress granule formation by its N-terminal low complexity domain through the biophysical process of liquid-liquid phase separation (Soto-Rifo et al., 2012; Shih et al., 2012; Valentin-Vega et al., 2016). Accurately defining where DDX3X binds mRNA in the transcriptome is crucial to devising a molecular model that will illuminate its proper function. This much is true for any Ribonucleoprotein (RNP). One such technique that has proven to help determine transcriptome binding sites for RNPs is CLIP-seq (Figure 2), a novel technique that combines UV cross-linking, RNA footing printing, immunoprecipitation, and High throughput sequencing to develop a comprehensive library of RNP binding sites within the transcriptome.

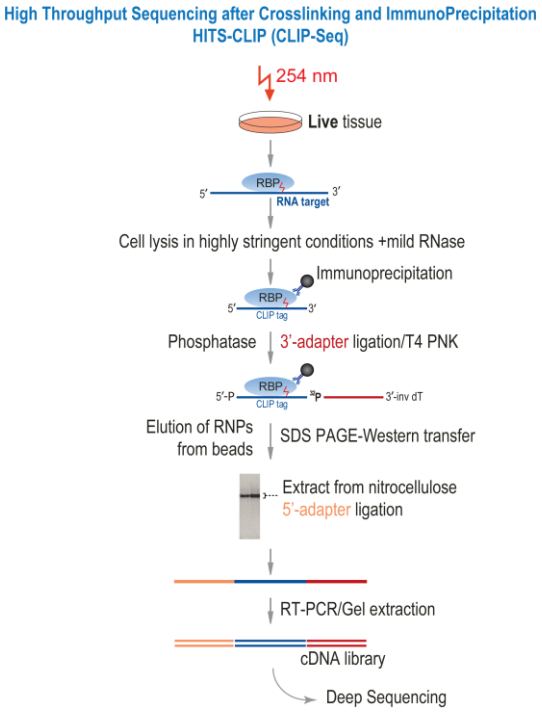

Figure 2: CLIP-Seq schematic.

Ule, Darnell, *Science* 2003, Vourekas and Mourelatos, *Methods Mol Biol* 2014

Researchers have already begun to apply this technique with great success. One study conducted by Valentin-Vega et al. in 2016 utilized CLIP-seq to create a comprehensive mRNA binding profile of DDX3X and found that it binds along the proximal side of 5' UTRs close to the AUG start codon of a specialized set of mRNAs providing valuable insight to the function of DDX3X and supporting its role in translation initiation. However, the CLIP-seq protocol has limitations as it only provides short RNP protected fragments for sequencing, and as a result, it cannot be determined whether or not these protein-protected fragments represent *in vivo* Full-Length mRNAs or Degradation Fragments. This limitation in traditional CLIP-seq protocol

provides the rationale for this thesis: to evaluate the development of shortened timed Ribonucleoprotein Immunoprecipitations (RNP-IP) followed by RT-qPCR analysis of Coimmuno-precipitated RNA to be conducted in tandem with CLIP-seq protocol to inform the nature of *in vivo* mRNA reads. It is hypothesized that if the mRNA reads represent degradation fragments, there will be decreased mRNA abundance detection in longer immunoprecipitation durations. Moreover, when the cells are lysed, it is possible that RNA binding proteins (RBP) bind RNAs with which they were not bound in vivo, which is an artifact called re-association with RNA, within the bulk contents of the cells released by cell lysis to perform the IP. To profile the in vivo bound mRNAs by an RBP and to effectively utilize RNP-IP in conjunction with CLIP-sec, minimization of the reassociation and degradation artifacts must be achieved. Since RNP-IP protocol conditions are not under protein denaturing conditions, nor is UV-crosslinking utilized to immobilize the RNP, we wanted to evaluate whether prolonged immunoprecipitation allowed the protein to bind RNA after cell lysis (which of course misrepresents the *in vivo* function of the protein), during the immunoprecipitation step, which in standard protocols lasts 3-16 hours. In order to test this hypothesis and find the optimal RNP-IP incubation duration to effectively balance between quality protein and RNA yield while maintaining the fidelity of the RNP: three different RNP-IPs for DDX3X were conducted at the following incubation durations: 2 mins, 30 mins, and 3 hrs. The findings of this experiment will not only inform CLIP-seq experiments regarding DDX3X but other RNPs as well, where it is more likely that the protein (or complex) binds RNAs at various time points in a maturation process, and therefore if a fragment of an long RNA is retrieved by CLIP-Seq, there is uncertainty as to what form the fragment represents, the full length or the in vivo processed form.

Materials:

Cell culture:

- 293T cells (Invitrogen R75007)
- 15 cm plates, Falcon™ Standard Tissue Culture Dishes (Fischer Fcientific 353003)
- Full (+) media
  (900 ml DMEM, 100 ml FBS Atlas Biological #F-0500-D, 10 ml L-glutamine 100x, 10 ml non-essential Amino acids, 10 mL Na pyruvate).
- Trypsin (labcorp 010355)
- PBS
- Liquid Nitrogen

RNP-IP:

- Protein A Dyna beads, Invitrogen 10002D
- Antibody binding buffer (1x PBS, 0.1% Nonidet P40)
- RSB 250 (50 mM Tris pH 7.5, 250 mM NaCl, 0.5% NP-40 IGEPAL, 0.1% Triton X-100 1.25 mM EDTA)
- Denaturing SDS-PAGE loading buffer: 1ml 4X LDS (Thermo #NP0007) plus 100 ul beta mercaptoethanol (Gibco #21985-023)
- Denaturing SDS-PAGE loading buffer for beads: two thirds (V/V) of the above buffer, one third V/V of RSB-250
- Cell lysis buffer: 10 ml RSB 250 plus 100 ul of Protease and phosphatase inhibitor Single-Use Cocktail (100x) (ThermoFisher, #78442), and 10 ul of rRnasin (Promega #N2511) per 400 ul of cell lysis.
- Trizol (Thermo Fisher #1559618)
- DDX3X Ab, custom, rabbit polyclonal (acquired from Genscript, preparation D)

RNA Extraction:

- Nuclease free water
- Phenol (invitrogen #1559-031)
- Isoamyl chloroform (Millipore #18F0456069)
- Sodium acetate
- Glycogen
- 70% ethanol
- Trizol (Invitrogen 15596026)

Urea Page Gel:

- 40% poly acrylamide (fisher chemical #BP1406-1)
- miliQ water
- ammonia persulfate (Thermo Fischer #17874)
- TEMED (Thermo Fischer #17919)
- NEB Low size RNA marker (New England Biolabs #B7025)

Western Blot:

- 15 well NuPAGE 4-12% Bis Tris gel (Thermo NP0336BOX)
- TBS
- TBS-0.1%Tween 20 (TBS-T)
- 2.5% milk TBST
- iBright marker (Thermo Fischer LC5615)
- 1/2000 of 1ug/ul solution DDX3X Ab D plus,
- 1/5000 mouse Ab tubulin CST (Cell signalling technology #2146)
- starbright 700nm 1/50,000 (Biorad #120004159)
- Dylight Goat anti mouse 800nm 1/5000 (Thermo Fisher scientific SA5-10176)

RT-QPCR:

- random hexamers DNTPS
- RNasin
- DDT 10 mM
- Superscript III (Invitrogen 18080093)
- RNase H (Promega M4285)
- Power UP SYBR green Mix (Thermo Fisher Scientific #A25741)
- ATP5 HIST1H1E and rsp6 RT-qPCR PRIMERS (ordered from IDT, PrimeTime primers)
- Maxiprep kit Invitrogen Purelink HiPure Plasmid Filter Maxiprep Kit (K210016)
- Xho1 (New England Biolabs MFCD00166586)
- RLUC plasmid pUC57 (addgene plasmid #50562)

Methods:

Cell culture:

Cell line of 293T was acquired from Thermo Scientific and grown to 75-90% confluency 15.0 cm cell plates in "Full +" media. Cells were harvested after 48-96 hours following inoculation, and detachment of cells from tissue plate was accomplished via Trypsinization and light shaking. Once all cells were resuspended, trypsin was neutralized via equal volume of Full + media, cells from each plate were collected 50 mL conical tubes by centrifugation (5 min at 1500 g), washed with PBS and collected finally in in pre-weighed 1.5 mL microtube tubes, and pelleted at 4 °C, at 1500 rpm for 1 minutes on a table-top microtube centrifuge. Pellets were then weighed, and flash frozen in liquid nitrogen and stored at -80 C.

RNP-IP:

Antibody binding

100 μl of Protein A Dynabeads was used per IP sample. Upon collection of appropriate amount of Dynabeads three subsequent washes were conducted using Antibody binding buffer and a magnetic strip. Upon completion of third wash, the wash buffer was discarded and 600 μl fresh antibody binding buffer containing the appropriate amount of perspective antibody at a ration of 8μg/ 100 μl of Dynabeads was added to the beads' suspension and left to incubate at room temperature for 1 hour under rotation. DDX3X Ab, custom, rabbit polyclonal (genscript, preparation D), and Non-immune Rabbit Serum (NRS) was used as the respective antibodies.

Cell lysis

Approximately 70 mg of cell pellets collected from previous cell culture experiment were utilized per IP sample. The frozen cell pellets were thawed/resuspended in Lysis buffer at a ratio of 9 μl per mg of Cell pellet and subject to further lysis via sonication, followed by centrifugation for 15 mins, 17,000 g, and 4 °C. Cell lysate superantant was collected and stored in a master tube, and 40 μl of untreated of lysate was collected at this time to be used as a pre-IP sample in Western Blot analysis.

Immunoprecipitation

Upon successful harvesting of Cell lysates, the supernatant (Antibody Binding buffer) of the fully prepared beads was removed and 600 μl of cell lysates per IP was added to the beads from which the final wash solution was removed, and incubated by rotation at 4 °C for the following duration: 2 mins, 30 mins, and 3 hours. Upon completion of incubation depleted lysates were removed and 50 μl was collected per IP and combined with equal volume of SDS-PAGE buffer for Western Blot analysis. Beads were then washed 3 times with RSB 250 and on last wash using 900 μl of RSB spilt into ¼ (225 μl) for Western Blot analysis and ¾ (675 μl) for RNA analysis.

Bead processing

Beads that were collected for Western Blot analysis were combined with 30 μl of Denaturing SDS-PAGE loading buffer for beads, and incubated on a shaking heat block for 12 min at 70 °C for protein elution. Upon completion of incubation, liquid fraction was isolated from beads via a

magnetic strip to be later used in Western blot analysis. Beads collected for RNA analysis were collected from supernatant and 500 μl of Trizol was added to dry beads followed by a 30 second vortexing. The beads were then left to incubate at room temperature for 5 mins. Upon completion 150 μl of Chloroform was added to solution, vortexed for 30 secs, and left to incubate at room temperature for 2 mins. Following incubation, the sample was centrifuged at 17,000 g and the upper aqueous phase of approximately 300 μl was collected where 3 μl of glycogen (5 μg/ μl) and 350 μl of isopropanol was then added and the aliquot was then subjected to incubation at -20 °C for 20 mins. Up completion of incubation, the sample was then centrifuged at 4 °C, 1700 gs for 30 mins, the supernatant was then removed and pellet was resubmerged in 1 mL of 70% EtOH followed by another round of centrifugation 4 °C, 1700 g's for 5 mins. The pellet was then aspirated, followed by air drying, and resuspended in 20 μl of Nuclease free water. A DNase I treatment of the RNA followed (20 min at 37oC) and another round of Phenol extraction – ethanol precipitation (see below). The amount of extracted RNA was then quantified via Nanodrop Spectrophotometry and used for RNA analysis via UREA PAGE Gel and RT-qPCR.

Western Blot: Western Blot protein analysis was performed to verify DDX3X precipitation based on molecular weight. Following acquisition of protein samples from the RNP-IP the samples run on a gel and transferred to a nitrocellulose membrane via semi-dry transfer. The membranes were blocked in 2.5% non-fat dry milk 1xTBS-T, washed briefly with 1x TBS and incubated with custom antibodies for at least 3 hours at 4 °C, then washed 3 times with TBST for 10 minutes each. Following washing, the membranes were incubated with the secondary

antibody for a minimum of 45 minutes. After incubation, the membranes were washed a total of 3 times for ten minutes each— Twice with TBST, and TBS for the last and final wash. The membranes were then imaged on a Chemidoc MP imaging software, and quantification of protein yield was quantified via antibody dependent fluorescence through Image lab software.

Quantification of RNA yield from RNP-IP: Following the RNA extraction from RNP-IP beads, amount RNA yield was quantified via Nanodrop spectrophometer.

UREA PAGE RNA Gel Electrophoresis: Evaluation of extracted of RNA samples was conducted via UREA PAGE RNA Gel Electrophoresis. Extracted RNA samples were either diluted to the same concentration of 1.4 ng/ µl or kept at their native concentration and run on a 6% Polyacrylamide gel, which was formed through solidification of the following mixture: 2.5 mL 10x TBE, 10.5 g UREA, 3.7 mL 40% polyacrylamide, and filled up to 25 mL with MilliQ water, 15 µl TEMED, and 150 µl of 15% ammonium persulfate.

Q-PCR:

Maxiprep: *Escherichia coli* cells expressing the reporter plasmid RLUC pUC57 was grown overnight in LB Broth Lennox broth containing 100 µg/mL concentration of ampicillin at 37 °C and 225 rpm. 200 ml of liquid medium was used for plasmid extraction following the PureLink™ HiPure Plasmid Maxiprep Kit (K210006) protocol, a type of column anion

exchange chromatography to purify plasmid DNA from cell lysates. Once the RLUC pUC57 plasmid was isolated from the cells they were subjected to Restriction Endonuclease Digestion for linearization with Xho1.

Plasmid digestion: A mixture containing, 25 µL of Xho1, 7.5 µL of nuclease free water containing the RLUC pUC57 plasmid (concentration 1328 µg/ µL), 5 µL Buffer R (10X), and 35 µL of nuclease free water were left to incubate for 1 hr. Upon completion of digestion, the linearized plasmid was isolated via Phenol Extraction and Ethanol precipitation.

Phenol Extraction and ethanol precipitation: Digested sample of linearized RLUC pUC57 was diluted to a volume of 400 µL nuclease free water, equal volume of phenol was then added vortexed for 1 min and centrifuged at 1700 G's for 5 minutes. Top water phase was then collected and equal volume of chloroform isoamyl-alcohol was added. The mixture was then vigorously vortexed and centrifuged at 1700g for 2 mins followed by collection of the water phase again. Upon collection of the water phase 40 µl of sodium acetate, 1.5 µl of glycogen (3.5 ng/ µl), and 1 ml of cold 100% ethanol was added and then incubated in -70 °C for 30 mins. Upon completion of cold incubation, the mixture was then centrifuged at 17,000 g for 30 minutes at 4 °C supernatant was discard followed by addition of 1 ml of 70% ethanol and centrifuged once more at 17,000 G's for 15 minutes at 4 °C. Upon completion of centrifugation,

the newly formed pellet was aspirated and left to air dry on a 37 °C heat block for 10 mins. Dry pelleted was then resuspended in nuclease free water to be used for In vitro transcription.

In vitro transcription: The luciferase gene in the plasmid is under the control of T7 RNA polymerase promoter, so we can transcribe it in vitro by incubating with T7 RNA polymerase. In vitro transcription was conducted via creation of a mixture containing: 10 µl of 5x Transcription Buffer, 10 µl of ATP/GTP/CTP/UTP each at 10 mM concentration, I ng of linearized plasmid template, 1.25 µl of rRNAsin RNase Inhibitor, 1.5 µl of T7 RNA Polymerase, and filled up to a 50 µl volume with nuclease free water. The mixture was left to incubate for 2 hours at 37 °C. Upon completion of incubation period 2 µl of DNase I, was added to mixture and incubated for 15 mins. DNase treatment was stopped via addition of 2 µl of 0.5 M EDTA, pH 8.0 and left to incubate at 65 °C for 10 min. Newly transcribed IVT RNA was then isolated by the previously described Phenol Extraction and ethanol precipitation protocol.

Generation of cDNA: 7 µl of the RNA samples extracted from the immunoprecipitation experiments were combined with the following reagents for generation of cDNA, 1 µL of IVT extracted Rluc mRNA (1 ng/ µL), 1 µL of random hexamers and 1 µL DNTPs and subjected to brief mixing and centrifugation followed by a 5-minute incubation at 65 °C. RLuc mRNA is used as a spike – reference mRNA to quantify relative RNA abundance differences in IP samples (no "housekeeping" gene is present in the IP samples). Upon completion of incubation, the samples

were left on ice for 1-minute to cool down followed by addition of 4 µL of 5x First Strand Synthesis, 1 µL of RNasin, 1 µL DTT (100 mM), 1 µL SS III, and 3 µL nuclease free water. Once the mixture was properly mixed and concocted, the samples were placed in a Miniamp thermocycler to incubate the samples for 10 minutes at 25 °C, 50-minute incubation at 50 °C, and 5-minute incubation at 85 °C for one cycle. Once completion of cDNA synthesis 1 µL of RNase H was added to each sample and incubated at 37 °C for 20 min.

RT-qPCR: Upon synthesis of cDNA containing gene spike. Equal volume of the cDNA was mixed with gene specific primers (ATP5F1E, RSP6, HIST1H1E, (IDT) and SYBR green Mastermix: Renilla Luciferase (rLuc) primers at the following volumes and concentrations: 1 µL RT reaction 2 µL primer mix (100 µM stock, 1 µL from each), 7 µL H20, and 10 µL SYBR Green mix. The samples were then plated and run at default Delta Delta CT qPCR settings in a Quant studio 3 qPCR machine.

Results:

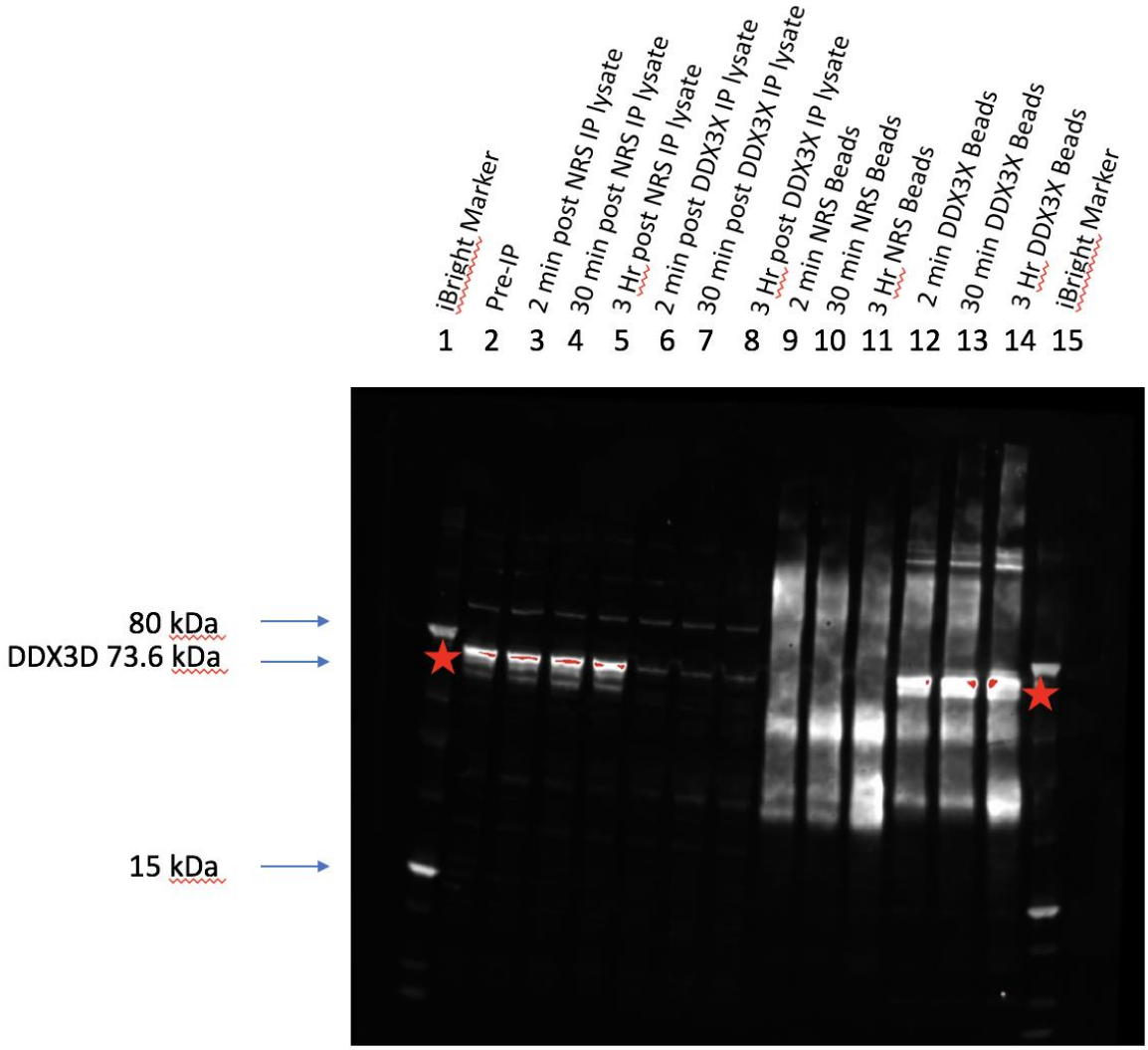

Figure 3: BW Western blot image for DDX3X on a 15 well gell. Each well represents the corresponding Protein sample and IP duration. The protein signal (DDX3X) is indicated with a red star

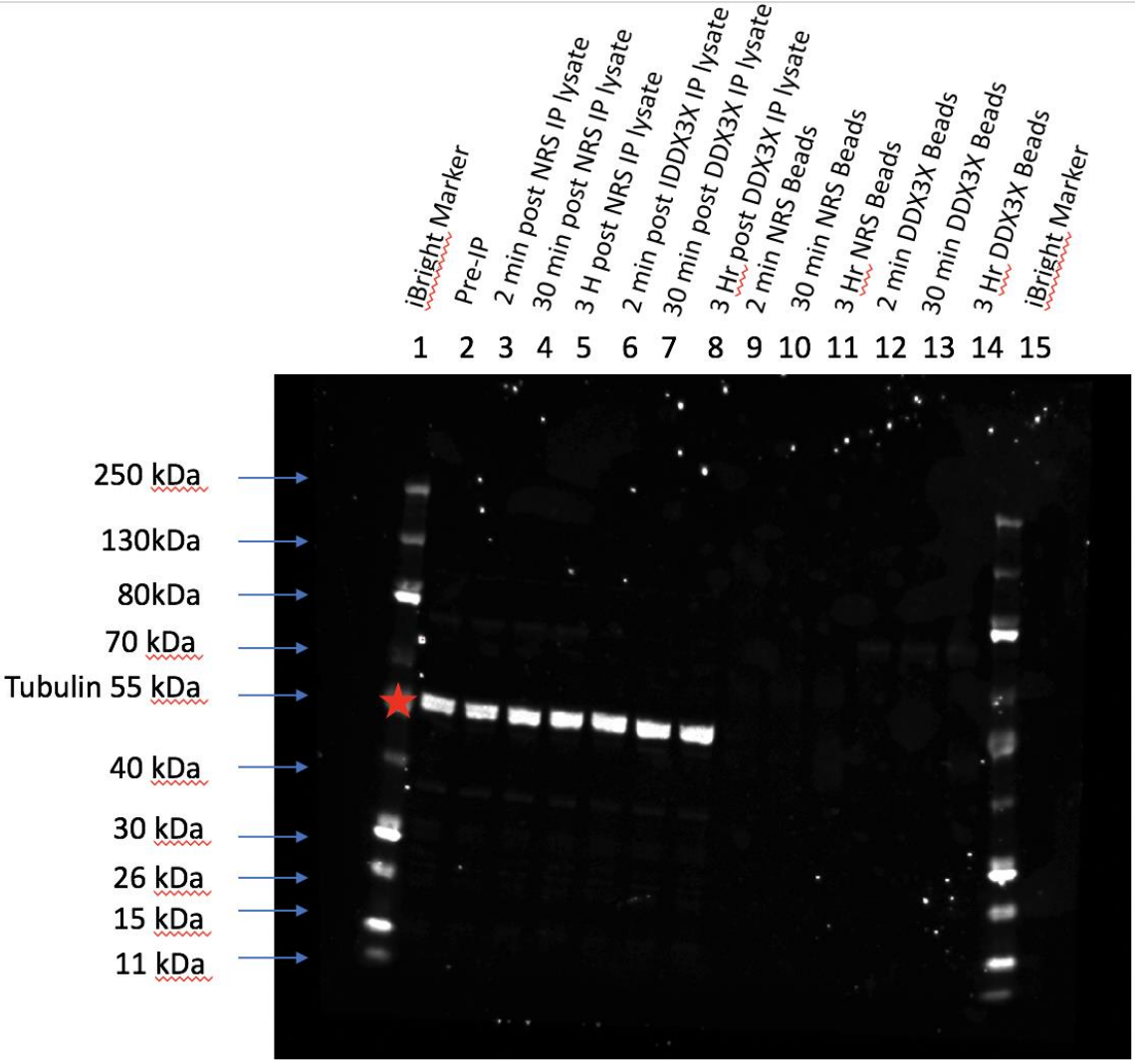

Figure 4: BW blot image for Tubulin (positive control) on a 15 well gell. Each well represents the corresponding Protein sample, and IP duration. The protein signal (tubulin) is indicated with a red star

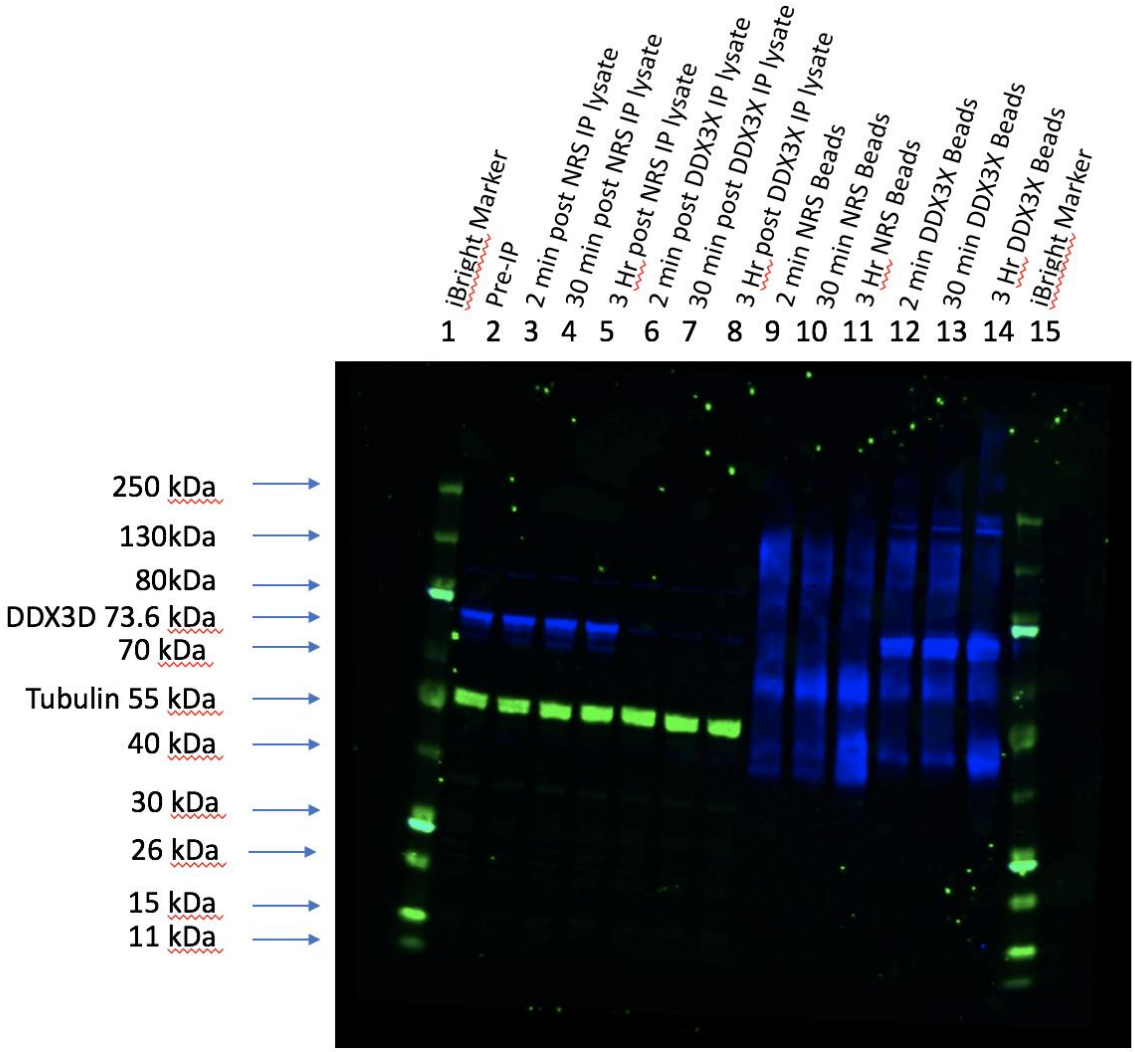

Figure 5: Composite Western blot image for DDX3X and Tubulin (positive control) on a 15 well gell. Each well represents the corresponding Protein sample and IP duration.

Replicate	Protein	IP Duration	Protein Yield (AFU)
1	DDX3X	2 minutes	43,437,352
1	DDX3X	30 minutes	44,121,116
1	DDX3X	3 hours	46,930,000
2	DDX3X	2 minutes	146,339,975
2	DDX3X	3 hours	166,428,422

Table 1: Quantitative analysis data of DDX3X yield at varying IP durations via Antibody dependent fluorescence through Image Lab software. Each row represents the replicate number, protein treatment, corresponding IP duration, and Protein yield measured in Arbitrary Fluorescence Units (AFU)

Western blot and DDX3X Antibody dependent fluorescence:

DDX3X was successfully immunoprecipitated in all RNP-IP incubation durations based on presence of bands in the Bead sample columns and with very minimal detection of DDX3X protein in the depleted lysate samples (Figure 5 -DDX3X WB-, compare lanes 2 with lanes 6-8).

Fluorescent-based detection of DDX3X levels shows minimal differences in protein yield between different incubation durations. In both replicates 1 and 2, the 3-hour RNP-IP incubation treatments showed the highest level of fluorescence for DDX3X at 46,930,000 AFU and 166428422 AFU respectively, but the increase compared to 2 min IP was 7 and 14 % respectively(Table 1).

Replicate	Protein	IP duration	RNA yield (ng/µL)
1	DDX3X	2 minutes	14.6 ng/µl
1	DDX3X	30 minutes	21.6 ng/µl
1	DDX3X	3 hours	42.1 ng/µl
1	NRS	2 minutes	5.7 ng/µl
1	NRS	30 minutes	7.8 ng/µl
1	NRS	3 hours	16.5 ng/µl
2	DDX3X	2 minutes	3.5 ng/µl
2	DDX3X	3 hours	28.6 ng/µl
2	NRS	3 hours	9 ng/µl

Table 2: RNA yield quantification via Nanodrop spectrophotometry post RNA extraction of the subsequent RNP IP treatment. Each row represents the replicate number, protein treatment, corresponding IP duration, and RNA yield measured in ng/µL.

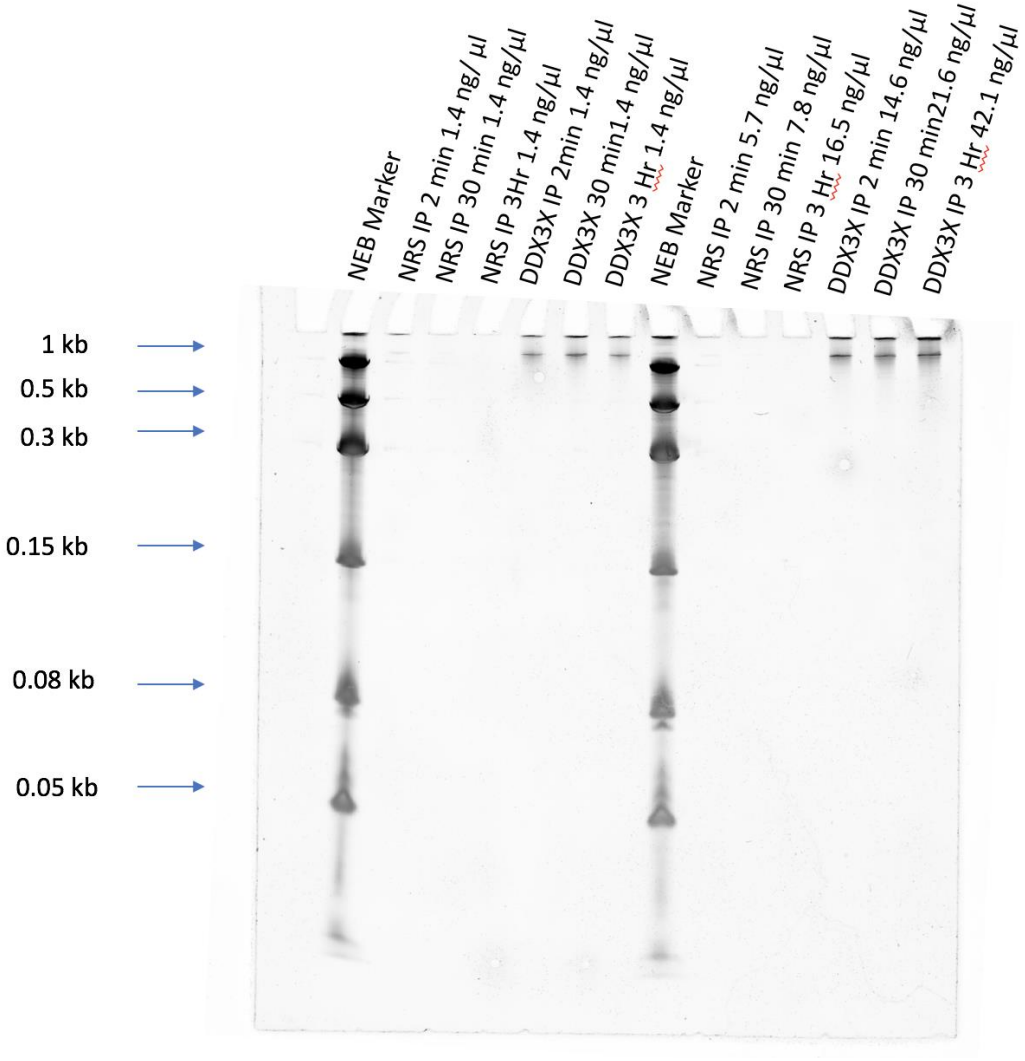

Figure 6: Urea Page Gel of RNA samples from the first Western Blot. Each lane represents the corresponding RNP IP sample, IP duration, and RNA concentration in ng/µL.

Nanodrop spectrophotometry and Urea Page gel electrophoresis:

RNA extraction from each RNP-IP sample was verified by Nanodrop spectrophotometry. UREA Page Gel electrophoresis showed presence of RNA species with differential nucleotide lengths, which is shown as differential bands in the DDX3X RNP-IPs (Figure 6). The data showed that amount of RNA extracted from the RNP-IP significantly increased with increasing incubation

duration (Table 2). The data shows that RNA yield nearly quadrupled from the 2 minutes to 3 hours incubation DDX3X treatment in replicate 1 and nearly increased by a factor of ~8.2 from 2 minutes incubation to 3-hour incubation duration in replicate 2 (Table 2).

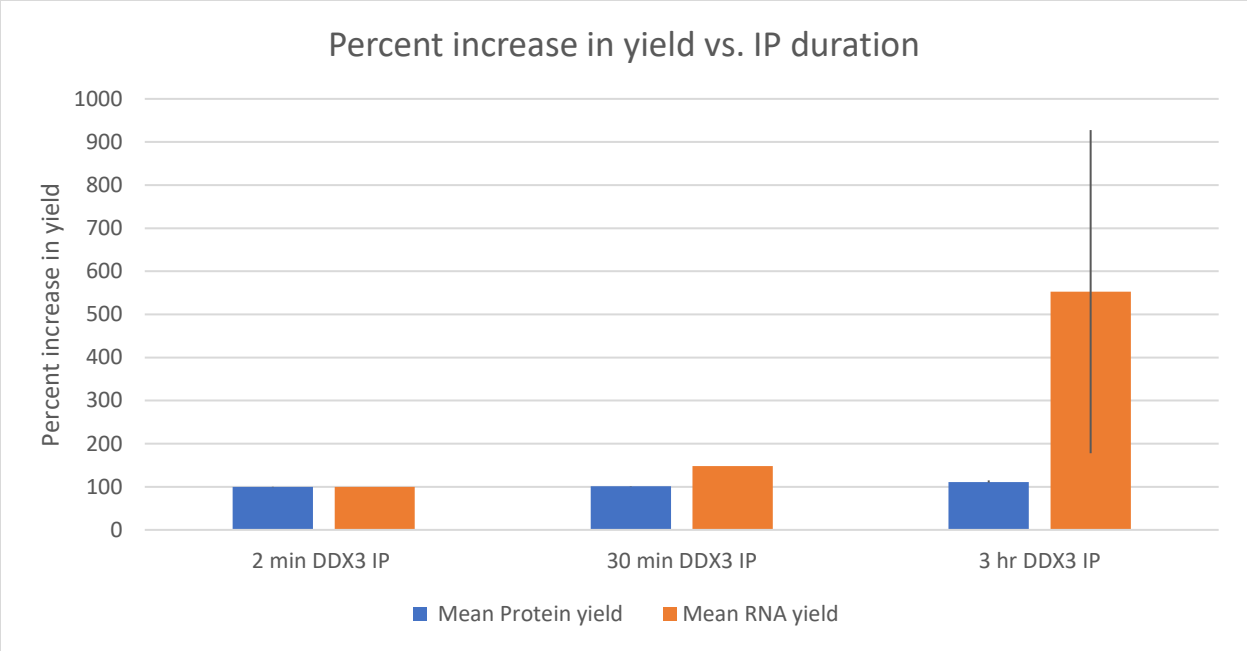

Figure 7: Graphical illustration of the percent increase in yield of Protein and RNA durations for various RNP-IP's for the protein DDX3X at differing durations. Each Bar represents the mean ± min - max (n=2: 3 hrs, and n=1: 30 mins) percent change in yield; 2 min IP durations were treated as the baseline (100%). Replicates not shown in thesis.

Graphical analysis of Percent increase in Protein and RNA yield vs. IP sample duration:

According to Figure 7, it is apparent that RNA yield increases at a disproportional rate compared to Protein yield, which only increased by 110.85 $\pm$ 4.03% when compared to the magnitude at which RNA yield increased 552.75 $\pm$ 374.85% upon reaching a time point of 3-hour incubation (Figure 7).

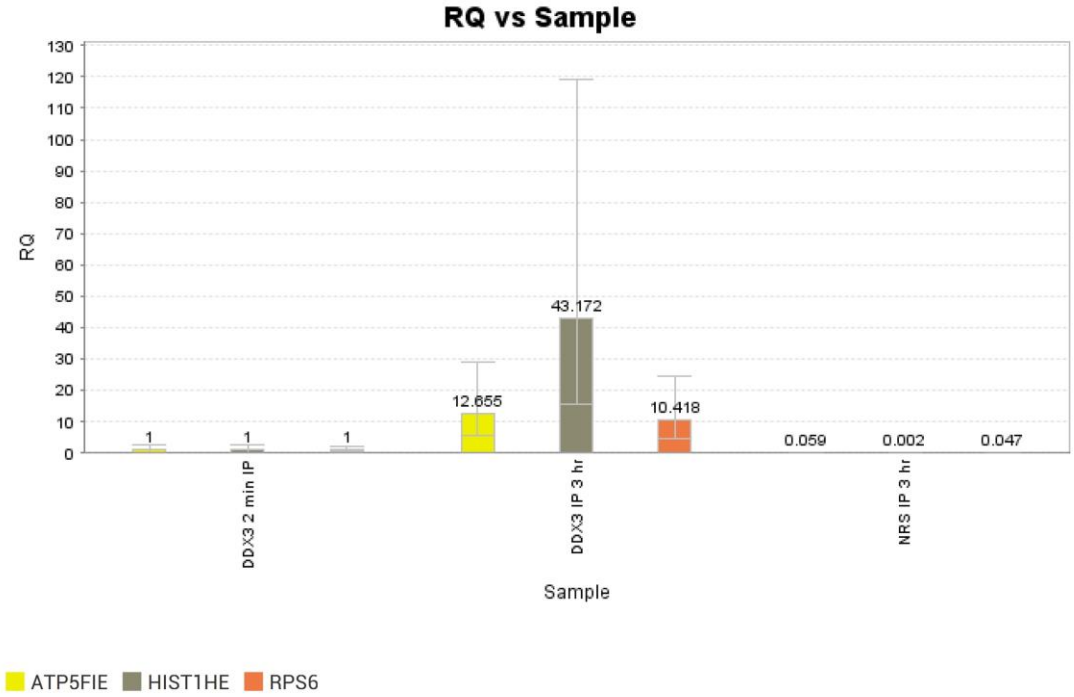

Figure 8: Relative Quantification value (RQ) vs sample treatment. Each bar represents the mean ± std (n=3) RQ for a particular mRNA corresponding to one of the following genes: ATPF1E, HIST1HE, and RPS6. The following IP samples were used: DDX3X 2min IP, DDX3X 3 hr IP, and NRS 3 hr IP (Control).

RT-qPCR analysis:

The data shows that mRNA abundance for all three genes ATPF1E, HIST1HE, and RPS6 significantly increased with RNP-IP incubation duration (Figure 8). Additionally, all three genes were detected in different abundance with HIST1HE being detected at a significantly higher mean RQ value of 43.172 when compared to the other two genes: ATPF1E and RPS6 which were detected to have mean RQ values of 12.665 and 10.418 respectively (Figure 8).

Discussion:

According to the results, the 2 minute IP duration was sufficient to achieve a protein yield equivalent to a 3-hour RNP-IP incubation. It was observed that RNA yield increased at a disproportionate rate compared to the rate protein yield, confirming our hypothesis that the artifact of reassociation becomes an increasing issue with longer RNP-IP durations and further illuminates the importance of decreasing RNP-IP incubation duration. However, what was most startling was the data from the RT-qPCR experiment. The original rationale for using RT-qPCR was to measure the amount of degradation, and one of two outcomes were expected, i.e., decreased amount mRNA abundance in the 3-hour RNP-IP indicating the artifact of degradation is at play, or increased abundance indicating artefactual reassociation. The qPCR results indicated the latter, and to such a high degree that if degradation is occurring, it is likely overshadowed by the reassociation of DDX3X, as detected by RT-qPCR. Interestingly, DDX3X showed biases for reassociation for different mRNA transcripts in the RT-qPCR data, as HIST1HE had a significantly higher RQ value than any other mRNA transcript. Given the relatively constant protein yield and rapid increase of RNA, the data can only be explained by the fact that reassociation is such an enormous artifact that even in other protocols such as CLIP-seq, which perform IP's under denaturing conditions, may be a factor that negatively influences the accuracy of the data. Increasing incubation duration yields diminishing returns as a very small amount of additional protein is yielded, and artefactual association becomes a more significant factor. Further experimentation should be conducted to verify this data and assess the degree of reassociation for proteins like DDX3X under denaturing conditioned protocols like CLIP-seq.

Acknowledgments: I want to thank the Vourekas Laboratory for helping me during my thesis journey. Thank you Dr. Vourekas for taking a chance on me and allowing me to work in your lab. Your mentorship and general life guidance has profoundly affected me in academia and beyond. Thank you also to Claire Fisher, Kali Taylor, Jared Navarre, Monsif Shawky, Mahmoud Dondeti, Eleni Liapi, Allison Scarboro, for the contribution you have made, without your support this thesis wouldn't be a reality.

References

Caruthers, J. M., & McKay, D. B. (2002). Helicase structure and mechanism. *Current Opinion in Structural Biology*, *12*(1), 123–133. https://doi.org/10.1016/s0959-440x(02)00298-1

Fairman-Williams, M. E., Guenther, U.-P., & Jankowsky, E. (2010). SF1 and SF2 helicases: family matters. *Current Opinion in Structural Biology*, *20*(3), 313–324. https://doi.org/10.1016/j.sbi.2010.03.011

Linder, Patrick., Lasko, Paul F, Ashburner, Michael, Leroy, Pascale., Nielson, Peter. J., Nishi, K. Ayoko, Slonimski, Piotrp (1989). Birth of the D-E-A-D box. *Nature*, *337*(6203), 121–122. https://doi.org/10.1038/337121a0

Linder, P., & Jankowsky, E. (2011). From unwinding to clamping — the DEAD box RNA helicase family. *Nature Reviews Molecular Cell Biology*, *12*(8), 505–516. https://doi.org/10.1038/nrm3154

Raman, V. (2015). DDX3, a potential target for cancer treatment. *Molecular Cancer*. https://doi.org/https://doi.org/10.1186/s12943-015-0461-7

Shih, J.-W., Tsai, T.-Y., Chao, C.-H., & Wu Lee, Y.-H. (2007). Candidate tumor suppressor DDX3X RNA helicase specifically represses cap-dependent translation by acting as an eIF4E inhibitory protein. *Oncogene*, *27*(5), 700–714. https://doi.org/10.1038/sj.onc.1210687

Shih, J.-W., Wang, W.-T., Tsai, T.-Y., Kuo, C.-Y., Li, H.-K., & Wu Lee, Y.-H. (2011). Critical roles of RNA helicase DDX3X and its interactions with eIF4E/PABP1 in stress granule assembly and stress response. *Biochemical Journal*, *441*(1), 119–129. https://doi.org/10.1042/bj20110739

Singleton, M. R., Dillingham, M. S., & Wigley, D. B. (2007). Structure and Mechanism of Helicases and Nucleic Acid Translocases. *Annual Review of Biochemistry*, *76*(1), 23–50. https://doi.org/10.1146/annurev.biochem.76.052305.115300

Soto-Rifo, R., Rubilar, P. S., Limousin, T., de Breyne, S., Décimo, D., & Ohlmann, T. (2012). DEAD-box protein DDX3X associates with eIF4F to promote translation of selected mRNAs. *The EMBO Journal*, *31*(18), 3745–3756. https://doi.org/10.1038/emboj.2012.220

Valentin-Vega, Y. A., Wang, Y.-D., Parker, M., Patmore, D. M., Kanagaraj, A., Moore, J., … Taylor, J. P. (2016). Cancer-associated DDX3X mutations drive stress granule assembly and impair global translation. *Scientific Reports*, *6*(1). https://doi.org/10.1038/srep25996