Biochemical and Structural Features of Phase Separating Proteins

By

Victoria Johnson

B.S, University of Rhode Island, 2020

Thesis

Submitted in partial fulfillment of the requirements for the Degree of Master of Science in the Graduate Program of Biotechnology at Brown University

PROVIDENCE, RHODE ISLAND

MAY 2025

AUTHORIZATION TO LEND AND REPRODUCE THE THESIS

As the sole author of this thesis, I authorize Brown University to lend it to other institutions for the purpose of scholarly research.

Date_____ _____

Victoria Johnson, Author

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

Date_____ _____

Victoria Johnson, Author

This thesis by (Victoria Johnson) is accepted in its present form by the Graduate Program in Biotechnology as satisfying the thesis requirements for the degree of Master of Science

Date_____ Signature: _____

Dr. Nicolas Fawzi, Advisor

Date_____ Signature: _____

Dr. George Lisi, Reader

Date_____ Signature: _____

Dr. Mandar Naik, Reader

Approved by the Graduate Council

Date_____ Signature: _____

Dr. Thomas Lewis, Dean of the Graduate School

CURRICULUM VITAE

Victoria Johnson

9 Newburg Street
Johnston, RI 02919

Phone: +1.401.528.9215
Email: Victoria_johnson@brown.edu

Education

Brown University (2022-2025) Providence, RI
M.S. Biotechnology
Thesis: Biochemical and Structural Features of Phase Separating Proteins

University of Rhode Island, College of Pharmacy (2016-2020) Kingston, Rhode Island
B.S. Pharmaceutical Sciences
Biology Minor
GPA: 3.7/4.0, summa cum laude

Research

Brown University, Research Assistant **July 2020 – present**
Dr. Nicolas Fawzi's Laboratory, Department of Molecular Biology, Cell Biology and Biochemistry

- Performed bacterial expression and purification of proteins (with ^{13}C, ^{15}N, ^{2}H stable isotopic labeling) associated with neurodegenerative diseases, such as TDP-43, FUS, SMN and hnRNPA2 to examine the role of mutations in protein-protein interactions.
- Assisted incoming students on electrophoresis and chromatography techniques (Ni-NTA, ion-exchange, size exclusion) including Bio-Rad chromatography systems.
- Prepared NMR samples with bio-macromolecules and small molecules for solution NMR.
- Performed biochemical and biophysical assays testing protein phase separation and aggregation.
- Maintained laboratory inventories and ordered reagents as necessary for completion of structural biology experiments.

University of Rhode Island, Undergraduate Research **January 2019 – May 2020**
Dr. David Rowley's Laboratory, College of Pharmacy

- Cultivated a marine bacterium (*Pseudoalteromonas piscicida*) and isolated and analyzed bacterial secondary metabolites using HPLC, LC-MS and NMR.
- Isolated and purified peptidoglycan from uropathogenic *E. coli* CFT073 using enzymatic treatments and ultracentrifugation and characterized fragments using HPLC and LC-MS.

Other Experience

Pezzuco Construction, Cranston, RI 02907 **June 2015 – January 2019**

- Managed invoices, payroll information and credit card statements for the entire company
- Created spreadsheets in Microsoft Excel and electronic documents for company record keeping

Publications

Izzy Owen, Debra Yee, Hala Wyne, Theodora Myrto Perdikari, **Victoria Johnson**, Jeremy Smith, Robert Kortum, Nicolas L. Fawzi and Frank Shewmaker. The oncogenic transcription factor FUS-CHOP can undergo nuclear liquid-liquid phase separation. J Cell Sci. (2021).

Liam C. O'Connell, **Victoria Johnson,** Jessica P. Otis, Anika K. Hutton, Anastasia C. Murthy, Mark C. Liang, Szu-Huan Wang, Nicolas L. Fawzi & Kimberly L. Mowry. Intrinsically disordered regions and RNA binding domains contribute to protein enrichment in biomolecular condensates in *Xenopus oocytes*. Sci Rep 14(1) (2024).

Mukulika Ray, Julia Zaborowsky, Pranav Mahableshwarkar, Smriti Vaidyanathan, Jasmine Shum, Renjith Viswanathan, Annie Huang, Szu-Huan Wang, **Victoria Johnson**, Noah Wake, Ashley M. Conrad, Alexander E. Conicella, Ryan Puterbaugh, Nicolas L. Fawzi, and Erica Larschan. Dual DNA/RNA-binding factor regulates dynamics of hnRNP splicing condensates. bioRxiv (2024). *Under review.*

Achievements/Certificates

Completion of the University of Rhode Island's Honors Program 2016 – 2020
Honors Thesis: Metabolomic Investigation of a Marine *Pseudoalteromonas piscicida*

MOOC: Basics of Nuclear Magnetic Resonance, University of Lille 2021

References

Nicolas Fawzi
Professor, Department of Molecular Biology, Cell Biology and Biochemistry
Brown University
Nicolas_fawzi@brown.edu

David Rowley
Professor, Department of Biomedical and Pharmaceutical Sciences
University of Rhode Island
drowley@uri.edu

Roberta King

Professor, Director of the BSPS Program
University of Rhode Island
rking@uri.edu

Deborah Martin
Controller
Pezzuco Construction
401-942-2244 x207
deborah@pezzcon.com

Acknowledgments

There are so many people who helped me get to where I am today, and I am eternally grateful for all of them.

I would first like to thank my thesis advisor, Dr. Nicolas Fawzi. You welcomed me into your lab during a time when the future was completely uncertain, not to mention in the middle of a pandemic. Over the last five years, you have pushed me to be a better scientist and think about those harder questions, while also giving me the freedom to be independent and pursue what I find interesting. I have grown immensely while being under your mentorship and will carry all that you've taught me throughout my research career, wherever that may lead me.

I would like to thank my thesis committee, Dr. George Lisi and Dr. Mandar Naik for serving on my committee and for their support of my career. You both are always willing to answer any questions I may have, and I thank you for that. I would also like to thank the facility managers at Brown, specifically Mandar Naik, Christoph Schorl, and Geoff Williams for training and access to instrumentation.

I would like to thank my program director, Dr. Jacqueline Schell. Your guidance over the past three years has not gone unnoticed. Thank you for always answering all my questions (I know I had lots of them), especially when I was applying into the program. You made the transition into the master's program seamless, and your advice never steered me wrong. Thank you for being so supportive and encouraging throughout my time in the program.

I would like to thank my undergraduate mentor, Dr. David Rowley. You were the first person to give me a chance to really start my career in research and provide me with an encouraging and supportive environment to do so. Thank you for making my first experience so great, and for

teaching me the basic skills I would need to carry on a career in this field. I'm so lucky that the beginning of my career started with you at the University of Rhode Island.

To all the current and former members of the Fawzi lab, thank you for making the lab such a fun place to be. Your support and guidance along the way has been so appreciated, and I thank you for always being willing to help me when I need it. It's been so refreshing to not only be able to talk about science, but also about life, as your friendships have meant so much to me. You have all inspired me every day to be the best scientist I can possibly be. I have no doubts that you all will be successful no matter where you go after Brown.

To my family and friends, thank you for always supporting and encouraging me. This journey has not been easy, but you have motivated me to keep going and achieve my goals. The little distractions whether it's just a quick phone call, a night out or a mini trip we've planned have meant so much. I couldn't have done it without you all.

To my mom and dad, this would not be possible without the both of you. Thank you for all the sacrifices you have made to get me to this point. I appreciate you both more than you know.

And lastly, the biggest thank you to my fiancé. Anthony, thank you for being my biggest supporter throughout this process. Thank you for believing in me when sometimes I didn't even believe in myself. They say as one chapter closes, another one opens, and I can't wait to start this next one with you.

Table of Contents

Table of Figures

Abstract of (Biochemical and Structural Features of Phase Separating Proteins), by (Victoria Johnson), ScM, Brown University, May 2025.

Liquid-liquid phase separation (LLPS) has been rapidly gaining acceptance as a powerful mechanism to explain the formation of membrane-less organelles and their function, playing a vital role in physiology and disease. Different types of interactions can contribute to phase separation, such as protein-protein, protein-RNA and RNA-RNA interactions. In this thesis, biochemical assays were used to determine how different structural domains and features of RNA-binding proteins contribute to phase separation properties for two different multidomain proteins. Biomolecular condensates often contain intrinsically disordered regions (IDRs) and RNA binding domains (RBDs), both of which the *Xenopus* heterogeneous nuclear ribonucleoprotein AB (hnRNPAB) X2 and Negative Elongation Factor E (NELFE) proteins possess. I found that full-length hnRNPAB X2 can form assemblies *in vitro*, while its IDR and RBD domains are able to form liquid condensates both in the presence and absence of *Xenopus* β-globin RNA. Droplets formed from the RBD exhibit liquid behavior with wetting and fusing properties, while X2 full length and the IDR domains form amorphous assemblies that cluster together and do not fuse. Additionally, NELFE full length and a deletion variant removing the RNA recognition motif (ΔRRM) result in phase separation only in the presence of polyadenylic acid (Poly A RNA). Deletion of the negatively charged low complexity sequence 1 (ΔLCS1) phase separates regardless of the presence of RNA. However, removing the RD repeat low complexity sequence two (ΔLCS2) completely abolishes droplet formation. A turbidity assay performed with full length NELFE showed that it contains properties of RNA mediated re-entrant phase separation. These results suggest that different proteins can possess a range of different biochemical properties as well as structural features that influence and regulate the

dynamics of biomolecular condensates. These insights will provide a deeper understanding of how RNA-binding proteins and IDR containing proteins can regulate dynamics of biomolecular condensates in disease states.

Chapter 1: Introduction

The basis of liquid-liquid phase separation

Eukaryotic cells are the basis for every multicellular organism, in which there is a clearly defined nucleus where chromosomes are located. These cells contain various organelles that harbor distinct chemical microenvironments which segregate molecules and proteins to direct the biochemical reactions underlying cellular processes. Eukaryotic cells contain two types of organelles: membrane-bound and membrane-less organelles. Membrane-bound organelles are more commonly known, such as the endoplasmic reticulum that contain the components necessary to synthesize, process and transport secreted proteins. Membrane-less organelles such as Cajal bodies play a vital role in small nuclear ribonucleoprotein assembly and ribosome biogenesis[1]. The formation of membrane-less organelles and their functions have been recently attributed to liquid-liquid phase separation (LLPS)[1].

Liquid-liquid phase separation (LLPS) is a phenomenon that occurs when a homogeneous mixture containing multiple components spontaneously separates into two distinct liquid phases with distinct concentrations of components[2]. Before phase separation occurs in the homogeneous mixture, the molecules remain dispersed. However, when phase separation occurs, they assemble into a concentrated phase, separating themselves from the surrounding dilute phase[3]. Membrane-less organelles can form rapidly and can reversibly concentrate components in discrete loci in cells[4]. Data suggests that proteins that contain intrinsically disordered, low complexity sequence domains (LCDs) can mediate this process. There are many different types of interactions between different domains which can include protein-protein, protein-RNA and RNA-RNA interactions[1].

LLPS implications in neurodegenerative disease

Phase separation has been implicated in neurodegenerative diseases. Under physiological conditions, phase separation is important for many biological processes. However, there are some proteins that although they can undergo phase separation at physiological conditions, they are also found in pathological aggregates. This indicates that pathological protein aggregates may originate from aberrant phase separation.

Protein aggregates have been shown to be the main hallmark of several neurodegenerative diseases, including amyotrophic lateral sclerosis (ALS), Alzheimer's disease (AD), frontotemporal dementia (FTD), and Parkinson's disease (PD). The transition from reversible dynamic LLPS to an irreversible aggregation state has been shown for TAR DNA-binding protein 43 (TDP-43), fused in sarcoma (FUS), tau, and α-synuclein, which are found aggregated in affected neurons of patients with ALS, FTD, AD and PD, respectively[1]. These processes are also altered by disease-associated mutations and regulated by post-translational modifications (PTMs)[1].

The importance of biomolecular condensates in cells

Biomolecular condensates have recently emerged as a key feature of cellular architecture[5]. They are found in the cytoplasm and nucleus, possessing a range of biological functions including ribosome biogenesis, RNA metabolism, signal transduction and DNA damage response[6]. The properties of a phase are influenced by the number and strength of the intermolecular interactions as seen in Figure 1[7]. When intermolecular interactions are weak, biomolecules remain dispersed. However, when the number and strength of the interactions are greater and exceed a threshold,

biomolecules separate into two phases: a dense phase with a high local concentration and a dilute phase with low concentration. This dense phase, or biomolecular condensate, can be liquid-like when the interactions are weak and transient. This allows biomolecules to move within as well as to enter and exit the phase. On the other hand, the dense phase is solid-like when the interactions are strong and long-lived, restricting biomolecular mobility. Many studies have demonstrated that biomolecular condensates can be transformed into materials in different states, such as viscous liquids, gels, and even solid aggregates[8]. Biomolecular condensates that are found in cells are often composed of protein and RNA.

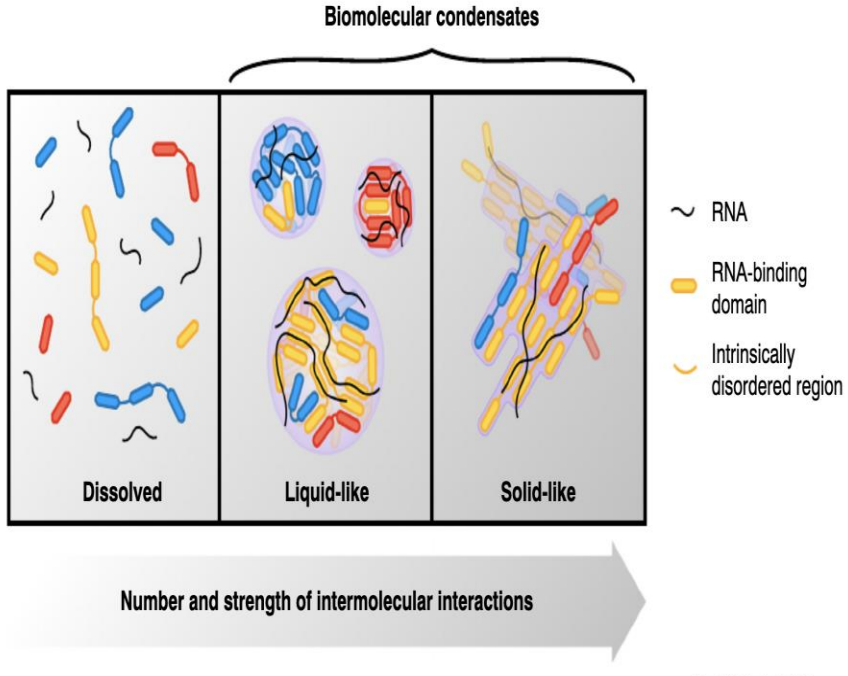


Figure 1: Phases of intracellular biomolecules. Proteins and nucleic acids can form a spectrum of phases with distinct composition, density, and dynamics. The properties of a particular phase are determined by the number and strength of intermolecular interactions within that phase. Biomolecules can remain dispersed, liquid-like, or solid-like[7].

The formation of biomolecular condensates requires their components to undergo a liquid–liquid phase separation or aggregation process, which changes them from a state of dissolution to a

state of separation from their surroundings[9]. Despite phase separation relying on different steps and ionic strength compared to aggregation, which can form assemblies even without phase separation, these two processes are mechanistically connected. There is a direct link between phase separation and aggregation in many disease-associated proteins. However, the formation of condensates by phase separation is often followed by aggregation to a less dynamic state.

In vitro studies performed with biomolecular condensate constituent proteins have shown that intrinsically disordered domains (IDRs) and RNA binding domains (RBDs) can drive phase separation[5]. IDRs are specific regions of a protein that have low degrees of sequence complexity and are not able to form stable higher order structures (such as secondary or tertiary structures). These IDRs remain unfolded, and studies have shown that proteins containing IDRs are able to phase separate *in vitro* by protein-protein interactions, forming bodies containing characteristics similar to biomolecular condensates[5].

Many IDR containing proteins also have RNA binding domains (RBDs). These RBDs can bind RNA with low specificity and affinity[10]. RNA binding proteins (RBPs) participate in the formation of ribonucleoprotein (RNP) complexes that are principally involved in gene expression[11]. The RBP is able to do this by binding to sequence and/or structural motifs in RNA via modular combinations of a limited set of structurally well-defined RNA-binding domains such as the RNA recognition motif (RRM), hnRNP K homology domain (KH) or DEAD box helicase domain. Proteins that contain both IDRs and RBDs have a high propensity to phase separate, suggesting that they function cooperatively in the condensation process, as seen by the protein-RNA interaction shown in Figure 2A[7].

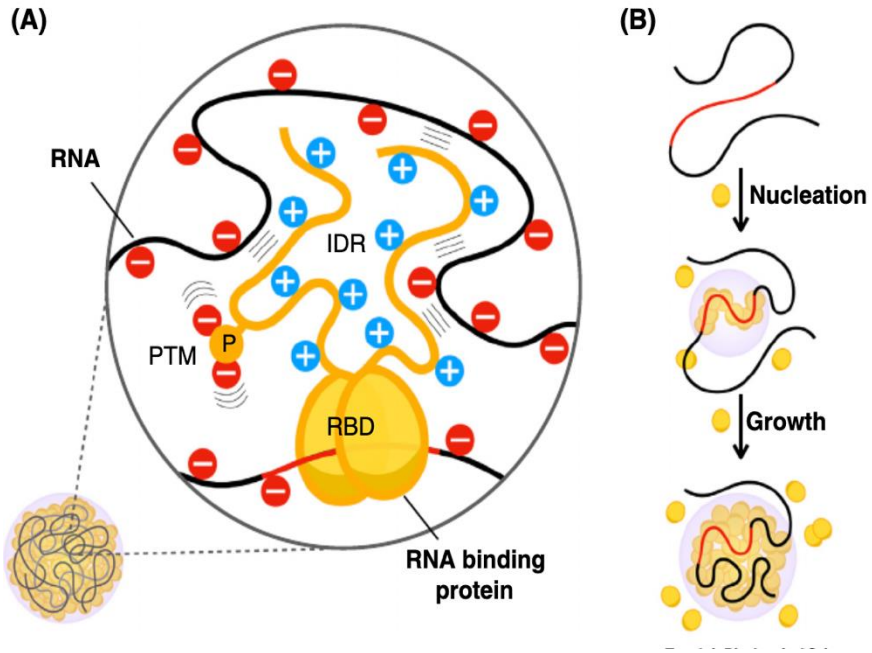


Figure 2: Structure and formation of biomolecular condensates. (A) RNA-binding proteins have a high propensity to undergo phase separation with their substrate RNAs. Multivalency is achieved through both specific and nonspecific interactions with substrate RNAs, mediated by both RNA-binding domains (RBDs) and intrinsically disordered regions (IDRs). These interactions are also modulated by post-translational modifications (PTMs), commonly phosphorylation (P). Straight lines indicate attraction, while curved lines indicate repulsion. **(B)** Biomolecular condensates are often nucleated by high-affinity interactions between RNA-binding proteins and specific sequences or structures on the substrate RNA (red). Increasing valency is provided through protein–protein (e.g., dimerization or higher-ordered oligomerization) and low-affinity protein–RNA interactions (e.g., electrostatic, hydrophobic, π–cation, and π–π interactions), resulting in the formation of a biomolecular condensate[7].

Additionally, there are RGG regions composed of arginine-glycine-glycine repeats of varying lengths and are commonly found in RNA binding proteins[12]. RGG regions can bind RNA, and the variability in the number of RGG repeats can influence the affinity of interactions formed by these RGG regions[13]. An individual RGG region in a protein will have weak RNA-binding

affinities, but are enhanced when all RGG regions are present, such as with RBP fused in sarcoma (FUS) protein[12,14].

Heterogenous nuclear ribonucleoprotein AB (hnRNPAB) X2 contains two RBDs and an IDR domain

Xenopus laevis oocyte biomolecular condensates contain both an intrinsically disordered domain and RNA binding domains that are highly enriched in localization bodies, or L-bodies [5]. L-bodies are defined as ribonucleoprotein (RNP) granules composed of a nondynamic mRNA containing phase enmeshed in a dynamic protein containing layer[15]. They are an important model system for the study of RNA localization. Specifically, *Xenopus* RBP heterogenous nuclear ribonucleoprotein AB (hnRNPAB) X2, a cytoplasmic isoform of hnRNPAB, contains two RBD regions and an IDR domain at its C-terminus. It is highly enriched in L-bodies. RNA binding proteins can regulate all aspects of the RNA lifecycle, such as splicing, modification, transcription, translation, intramolecular trafficking and decay[16]. Studies have shown that adding RNA to RNA binding proteins (RBPs) *in vitro* facilitates the formation of phase separated bodies, suggesting a role for RNA-protein interactions in their formation in addition to IDR interactions[5]. Although the mechanics of *in vitro* phase separated bodies have been studied, it is unclear how this translates in cells with much more complex heterogeneous biomolecular condensates.

In this study, we aim to determine how hnRNPAB X2's intrinsically disordered domain and RNA binding domains affect phase separation *in vitro*. Since it has been shown that the addition of RNA can enable phase separation, we sought to test if the same principle can be applied to hnRNPAB X2. To take this even further, we want to quantitate the data collected, examining

how the size and shape of the condensates formed vary between the different domains, the protein concentration used as well as the addition of RNA. Our findings will demonstrate how the IDR and RBD domains behave in regulating dynamics in biomolecular condensates.

Negative Elongation Factor E (NELFE) is an RNA binding protein

An RNA binding protein that contains many phase separation features is Negative Elongation Factor E (NELFE). It is a vital part of the NELF complex that is responsible for repressing RNA Polymerase II transcript elongation. The NELF complex is a hetero tetramer composed of the subunits NELFA, B, C/D and E[17]. Under stress conditions such as heat shock and arsenic exposure, it can cause an abrupt arrest of ongoing translation of mRNAs encoding metabolic, cell cycle and housekeeping proteins. This process is known to involve the formation of cytosolic stress granules, which are biomolecular condensates that form by phase separation. Recent techniques have confirmed that heat shock specifically causes global transcriptomic downregulation of housekeeping genes in fly, mouse, and human cells[17]. Transcriptional downregulation occurs at the level of promoter-proximal pausing of RNA Polymerase II, implicating the negative elongation factor (NELF) complex[17]. Under stress conditions, NELF binding to promoters of housekeeping genes is rapidly enhanced, coincident with a decreasing of elongating Pol II.

Implications in Cancer

Not only can liquid-liquid phase separation be implicated in neurodegenerative disease, but it can also be implicated in cancers. Studies have indicated that RBPs can mediate cancer-associated transcriptomic changes in hepatocellular carcinoma (HCC). Hepatocellular carcinoma is the most

common type of primary liver cancer, being the 2nd most common cause of cancer related deaths worldwide. Cancer cells, including those in hepatocellular carcinoma, often contain alterations of thousands of seemingly unrelated coding and non-coding RNA transcripts[18]. It has been identified that the RNA binding protein NELFE is oncogenic and can selectively regulate MYC-associated genes in HCC with poor prognosis, since they can regulate the functions and abundances of RNA transcripts at multiple levels[18]. Several studies have shown that RBPs play an essential role in modulating gene expression in homeostasis, cellular development, and disease state[18]. However, it is still largely unexplored and unknown as to how and to what extent RNA binding proteins can regulate the cancer transcriptome.

Structural Features of NELFE

As mentioned previously, the majority of phase separating proteins contain an intrinsically disordered domain. However, NELFE is unique in that it does not possess a typical IDR domain, but instead contains a C-terminal RNA recognition motif (RRM) and two low complexity sequences[19]. The RRM is able to bind RNA. Low complexity sequences (LCS) frequently repeat the same amino acid sequence or contain fragments with low diversity in amino acid composition[20].

The full-length sequence of NELFE is primarily composed of polar, uncharged amino acids such as glutamine (Q), serine (S), threonine (T), and asparagine (N) residues. It has a net charge of positive nine (+9). Additionally, it contains many negatively charged aspartic acid (D) and glutamic acid (E) residues and positively charged arginine (R) residues. In the first low complexity sequence (LCS1), it is a stretch of seventeen amino acid residues, mainly consisting

of aspartic acid, glutamic acid, glycine, and proline residues. Due to the negatively charged residues, if this segment were to be removed from the overall sequence, it changes the overall charge of NELFE, making it more positive.

NELFE contains a long stretch of tandem repeat arginine and aspartic acid (RD) residues in the second low complexity sequence that it has (LCS2), as well as arginine/glutamic acid (RE) and arginine/serine (RS) repeats. Removal of this segment takes most of the charged residues with it away, altering the net charge. The RNA recognition motif is similar to LCS1 in that it contains several aspartic acid and glutamic acid residues. Similarly, it changes the net charge of the sequence to be more positive when this segment is removed.

This project aims to determine whether the RNA binding protein NELFE is capable of undergoing liquid-liquid phase separation on its own without the entire complex as well as how the different domains and deletions within NELFE affect its phase separation. We also want to determine whether NELFE full length phase separation is RNA dependent. These insights will provide a better understanding of how RNA binding proteins induce phase separation as well as modulating cancer disease states.

Chapter 2: Methods

hnRNPAB X2 methods:

In vitro expression and purification of hnRNPAB X2 full length isoform, IDR and RBD domains

N-terminally MBP tagged (pTHMT) hnRNPAB X2 full-length isoform, IDR and RBD domains, and MBP (control) were expressed in *E. coli* BL21 star (DE3) cells (C600003; Thermofisher Scientific). Bacterial cultures were grown to an optical density at 600 nm of 0.7–0.9 before induction with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 4 h at 37° C. Cell pellets were harvested by centrifugation and stored at -80 °C. Cells were resuspended in ~ 40 mL of 20 mM Tris, 500 mM NaCl, 10 mM Imidazole (pH 8.0), with one Roche EDTA-free protease inhibitor tablet (11697498001; Sigma Aldrich) per ~ 4 g cell pellet, and lysed using the Avestin Emulsiflex C3 (Ottawa, Ontario, Canada). The lysate was cleared by centrifugation at 47,850 × g (20,000 rpm) for 50 min at 4 °C, filtered using a 0.2 µm syringe filter, and loaded onto a HisTrap HP 5 ml column (17524701; Cytiva, Marlborough, MA, USA). The protein was eluted with a gradient from 10 to 300 mM imidazole in 20 mM Tris, 500 mM NaCl, pH 8.0. Fractions containing MBP-tagged hnRNPAB X2 full length isoform were loaded onto a HiLoad 26/600 Superdex 200 pg column (28-9893-36; Cytiva) equilibrated in 20 mM Tris, 150 mM NaCl, and 1 mM dithiothreitol (DTT). Fractions containing MBP-tagged hnRNPAB X2-IDR, hnRNPAB X2-RBD, or MBP alone were loaded onto a HiLoad 26/600 Superdex 75 pg (28-9893-34; Cytiva) equilibrated in 20 mM Tris and 150 mM NaCl. Fractions with high purity were identified by SDS-PAGE and concentrated using a centrifugation filter with a 10 kDa cutoff (ACS501024; MilliporeSigma, Burlington, MA, USA). MBP-tagged hnRNPAB X2 full length isoform, the

IDR and RBD domains, and MBP were flash frozen in liquid nitrogen in 20 mM Tris and 150 mM NaCl and stored at -80 °C.

In-vitro phase separation

Purified MBP-tagged hnRNPAB X2 full length isoform, and the IDR and RBD domains or, as a control, MBP alone (including TEV cleavage artifacts and all residues present in MBP-tagged hnRNPAB X2 proteins) in 20 mM NaPi, pH 7.4, 150 mM NaCl were incubated with 0.03 mg/mL TEV protease for 20 min to cleave the MBP solubility tag from the protein. Protein concentrations of 5 µM, 12.5 µM. 25 µM and 50 µM were tested in the presence and absence of 0.25 mg/mL *Xenopus* β-globin RNA, with the addition of 10% polyethylene glycol (PEG) (25322-68-3; Sigma Aldrich) to induce phase separation, analogous to recent protocols for other RNA-binding proteins. After inducing phase separation in microcentrifuge tubes, the samples were incubated at room temperature for ~ 20 min before imaging. Images were taken using a 60x objective; scale bar is 200 µm.

Condensate Imaging

All condensate imaging was performed on a Nikon Ti2-E Fluorescence microscope using a 60× objective. Samples were spotted onto a glass coverslip and condensate formation was evaluated by imaging with differential interference contrast (DIC), acquired with a Prime 95B Mono camera. For DIC time-lapse videos, frames were collected every 0.6 s for 20 s. For fluorescence imaging, *in vitro* phase separation was carried out in the presence of 25 µM Thioflavin T (2390-54-7, Sigma Aldrich). Fluorescence images were quantified with ImageJ software. Images were subjected to background subtraction with a 50.0 pixel rolling ball radius and manual

thresholding. The Analyze Particles plugin was then used to measure condensate number per field of view (294.8 × 294.8 μm), size (in μm), and circularity (on a scale of 0–1.0, with 1.0 representing a perfect circle), excluding particles on the edges. Six to nine images were analyzed per group for each of 3 replicates. Results were analyzed by ordinary two-way ANOVA followed by Tukey's multiple comparison tests for hnRNPAB X2 and the IDR and RBD domains.

NELFE Methods

In vitro expression and purification of NELFE full length protein, ΔLCS1, ΔLCS2 and ΔRRM domains

C-terminally MBP tagged (pJ4M) NELFE full-length isoform, ΔLCS1, ΔLCS2 and ΔRRM domains were expressed in *E. coli* BL21 star (DE3) cells (C600003; Thermo Fisher Scientific). Bacterial cultures were grown to an optical density at 600 nm of 0.7-0.9 before induction with 1 mM isopropyl-β-D-1-thiogalactopyranoside (IPTG) for 4 hours at 37° C. Cell pellets were harvested by centrifugation and stored at -80°C. Cells were resuspended in ~40 mL of 20 mM Tris, 1 M NaCl, 10 mM Imidazole, 1 mM DTT pH 8.0, with one Roche EDTA-free protease inhibitor tablet (11697498001; Sigma Aldrich) for ~4 g cell pellet, and lysed using the Avestin Emulsiflex C3 (Ottawa, Ontario, Canada). The lysate was cleared by centrifugation at 47,850 x g (20,000 rpm) for 50 min at 4°C, filtered using a 0.2 μm syringe filter, and loaded onto a HisTrap HP 5 ml column (17524701; Cytiva, Marlborough, MA, USA). The protein was eluted with a gradient from 10 to 500 mM imidazole in 20 mM Tris, 1 M NaCl, 1 mM DTT pH 8.0. Fractions containing MBP-tagged NELFE full length isoform were loaded onto a HiLoad 26/600 Superdex 200 pg column

(28-9893-36; Cytiva) equilibrated in 20 mM Tris, 300 mM NaCl, and 1 mM dithiothreitol (DTT). Fractions containing MBP-tagged ΔLCS1, ΔLCS2, or ΔRRM were loaded onto a HiLoad 26/600 Superdex 200 pg (28-9893-36; Cytiva) equilibrated in 20 mM Tris, 300 mM NaCl and 1 mM DTT. Fractions with high purity were identified by SDS-PAGE and concentrated using a centrifugation filter with a 10 kDa cutoff (ACS501024; MilliporeSigma, Burlington, MA, USA). MBP-tagged NELFE full length isoform, ΔLCS1, ΔLCS2 and ΔRRM domains were flash frozen in liquid nitrogen in 20 mM Tris, 300 mM NaCl and 1 mM DTT and stored at -80 °C.

Differential interference contrast microscopy

Droplets were prepared in the purification buffer at 50 µM protein concentration for MBP-tagged NELFE full length isoform, and the ΔLCS1, ΔLCS2 and ΔRRM domains in 20 mM NaPi, pH 7.4, 150 mM NaCl, in the presence and absence of 0.5 mg/mL polyadenylic acid (Poly A) potassium salt. The samples were incubated with 0.03 mg ml^{-1} TEV protease for ~20 minutes before visualization. Samples were spotted onto a glass coverslip and droplet formation was evaluated by imaging with differential interference contrast (DIC) on a Ti2-E Fluorescence microscope (Nikon). Images were taken using a 40x objective; scale bar is 20 µm.

Turbidity Measurements

Turbidity was used to evaluate phase separation of 50 µM NELFE full length isoform in the presence of 0.06 mg mL^{-1} TEV protease (~0.3 mg mL^{-1} in 50 mM Tris-HCl pH 7.5, 1 mM EDTA, 5 mM dithiothreitol, 50% glycerol and 0.1% Triton X-100). The experiment was performed in 20 mM NaPi pH 7.4 and 150 mM NaCl with the indicated poly adenylic Poly(A) potassium salt (26763-19-9; Sigma Aldrich) concentrations. Turbidity experiments were performed in a 96-well clear plate with 100 µL samples sealed with optical adhesive film to prevent evaporation (4311971; Thermo Fisher Scientific). The absorbance at 600 nm (A_{600}) was

monitored over time using a Biotek Synergy H1 Plate Reader (BioTek Instruments, Winooski, Vt, USA) at 5 min time intervals for up to sixty minutes.

Chapter 3: Results

hnRNPAB X2 along with the IDR and RBD domains phase separate and self-assemble *in vitro*

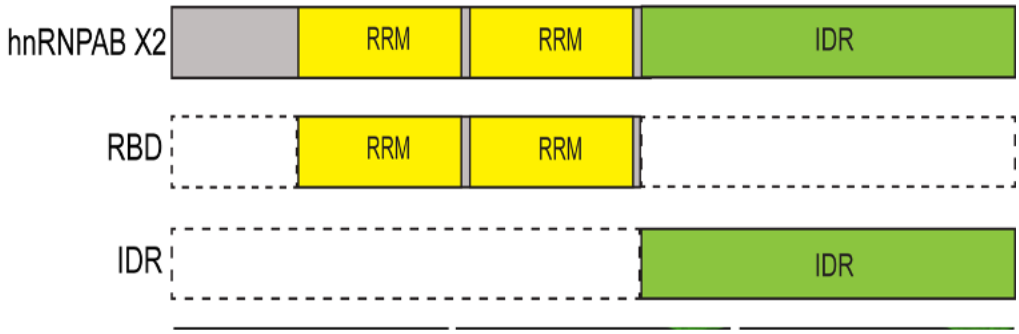


Figure 3: hnRNPAB X2 has two RNA binding domains and one intrinsically disordered region. Schematics of the domain constructs used in this study.

To investigate how RNA binding domains and intrinsically disordered regions regulate protein dynamics in biomolecular condensates we focused on hnRNPAB X2. It is a constituent of L-bodies that contains two RBDs and an IDR (Figure 3). Biomolecular condensates have been characterized to form through phase separation[21]. We tested if hnRNPAB X2 or its domains are capable of undergoing liquid-liquid phase separation using an *in vitro* approach. Phase separation of samples of recombinant hnRNPAB X2, and the IDR and RBD domains were analyzed by differential interference contract (DIC) microscopy. The proteins were purified as fusions with maltose binding protein (MBP) solubility tags to increase their solubility. Samples were prepared in 20 mM sodium phosphate (pH 7.4) and 150 mM sodium chloride buffer conditions. After cleavage of the MBP tag using TEV protease and in the presence of 10% polyethylene glycol (PEG), a crowding agent, full-length hnRNPAB X2, along with the IDR domain formed assemblies in the absence (Figure 4A-B) and in the presence (Figure 4E-F) of *Xenopus* β-globin

RNA. Additionally, the RBD domain formed round liquid droplets both in the presence (Figure 4G) and in the absence (Figure 4C) of RNA. A control protein (MBP alone) was also imaged and did not exhibit phase separation properties (Figure 4D, 4H).

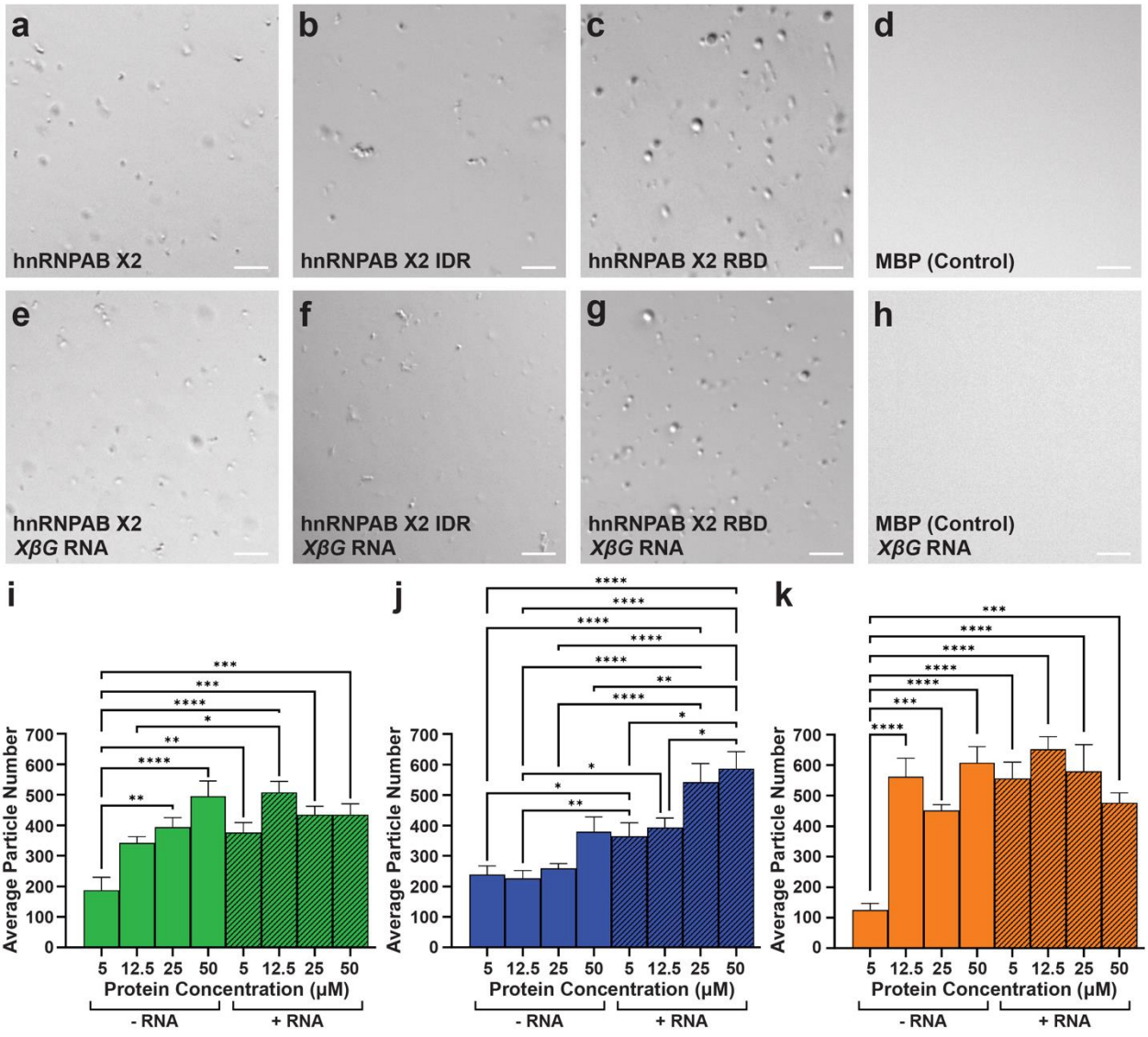


Figure 4: hnRNPAB X2 and its domains phase separate and self-assemble *in-vitro*
(a-h) DIC micrographs of **(a,e)** hnRNPAB X2 full length, **(b,f)** hnRNPAB X2 IDR, **(c,g)** hnRNPAB X2 RBD, and **(d,h)** MBP (control) proteins at 12.5 μM concentration in 20 mM NaPi, pH 7.4, 150 mM NaCl, 10% PEG and 0.03 mg/ml TEV protease in the presence (e-h) or absence (a-d) of 0.25 mg/mL *Xenopus* β-globin (XβG) RNA . Scale bar = 10 μm. (**i**–**k**) Condensate

formation was quantitated for fluorescently labeled condensates by determining the average particle number per field of view for (**i**) full-length hnRNPA2 X2 (green), (**j**) hnRNPAB X2 IDR (blue), and (**k**) hnRNPAB X2 RBD (orange) at 5, 12.5, 25, and 50 µM protein concentrations in the presence (+, hatched bars) and absence (-, open bars) of 0.25 mg/mL *Xenopus* β-globin RNA. Statistics shown are an ordinary two-way ANOVA followed by Tukey's multiple comparisons. Error bars represent standard error of the mean, **** indicates $p < 0.0001$, *** indicates $p < 0.001$, ** indicates $p < 0.01$, * indicates $p < 0.05$, and all brackets not shown are not significant ($p > 0.05$). All quantification analysis and bar graphs were completed by Jessica Otis and come from O'Connell et. al 2024[5].

In collaboration with Dr. Kim Mowry's laboratory at Brown University, we sought to quantitate the condensate numbers for hnRNPAB X2 and each of its domains. Jessica Otis, a member of Kim Mowry's lab, was able to quantitate and analyze the data presented in Figures 4I-K and 5G-L using the ImageJ Analyze Particles plugin and statistical analyzes that were obtained and shown as an ordinary two-way ANOVA followed by Tukey's multiple comparisons. Condensate numbers for hnRNPAB X2 and its IDR and RBD domains were quantitated at 5 µM, 12.5 µM, 25 µM and 50 µM protein concentrations both with and without *Xenopus* β-globin RNA (Figure 4I-K). The full-length protein and the RBD domain form fewer condensates at 5 µM protein concentration in the absence of RNA, while RNA had little to no effect at higher protein concentrations analyzed (25-50 µM) (Figure 4I, 4K). For the IDR domain (Figure 4J), the average number of condensates formed at all protein concentrations increased in the presence of RNA, while the individual IDR condensates were smaller in size in the presence of RNA (Figure 5K). The size of the condensates formed by the RBD domain increased with increasing protein concentration (Figure 5L), and RNA exhibited an effect on condensate size only at the lowest protein concentration of 5 µM, where the average size was higher with RNA present (Figure 5L). hnRNPAB X2 condensate sizes were not strongly affected by the addition of RNA or varying protein concentrations (Figure 5J).

We assessed the circularity of the condensates for each of the three domains (Figure 5G-I). When assessing circularity, a value of 1.0 is defined as a perfect circle. We found that the RBD and full-length domains have high degrees of circularity when RNA is present, as the values average from 0.6-0.8 (Figure 5G, 5I), but when RNA is absent, the average circularity ranges from 0.5-0.5. For the IDR region, circularity is not drastically impacted by the presence of RNA or protein concentration, as the average circularity values remain in the same range (Figure 5H).

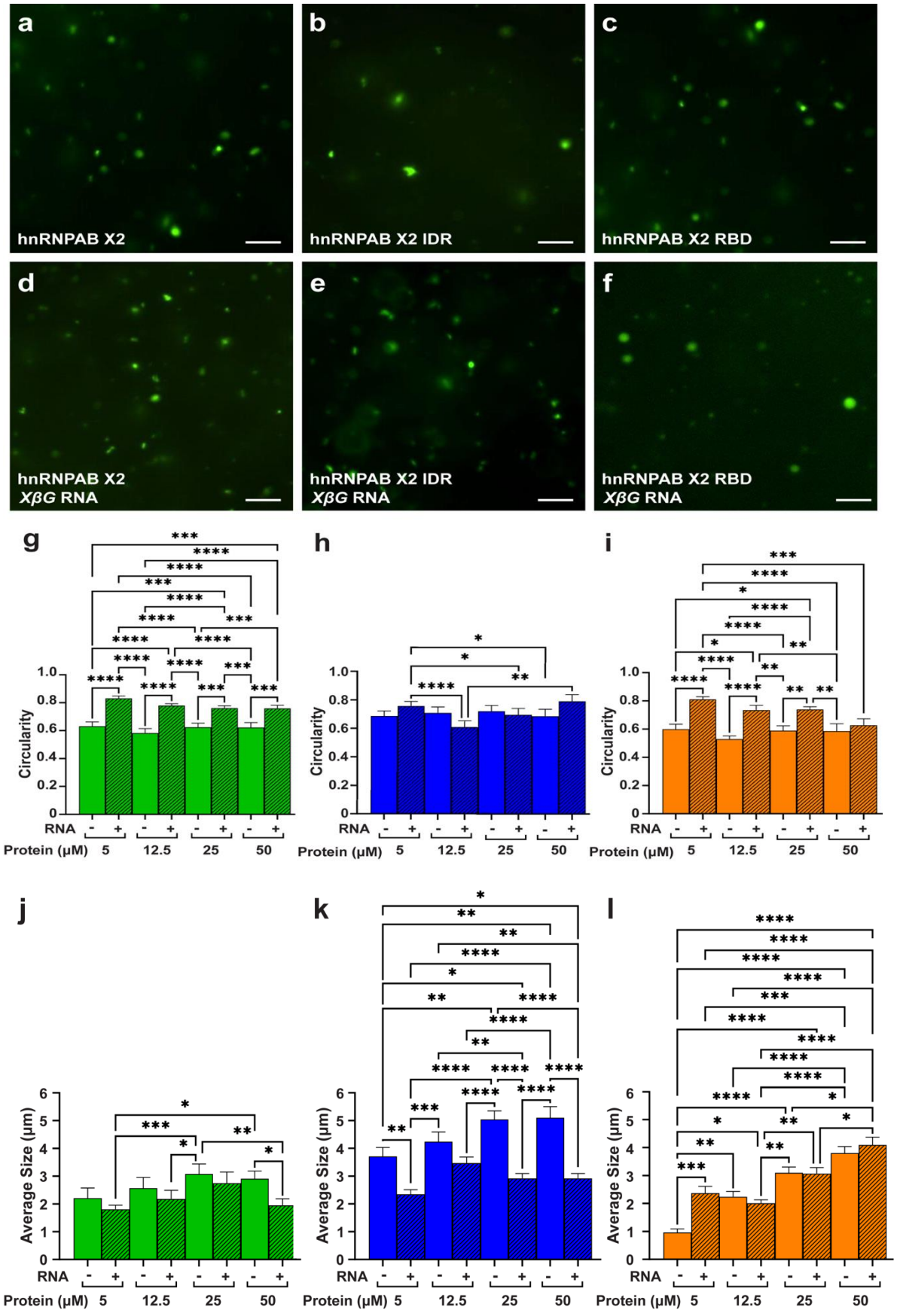
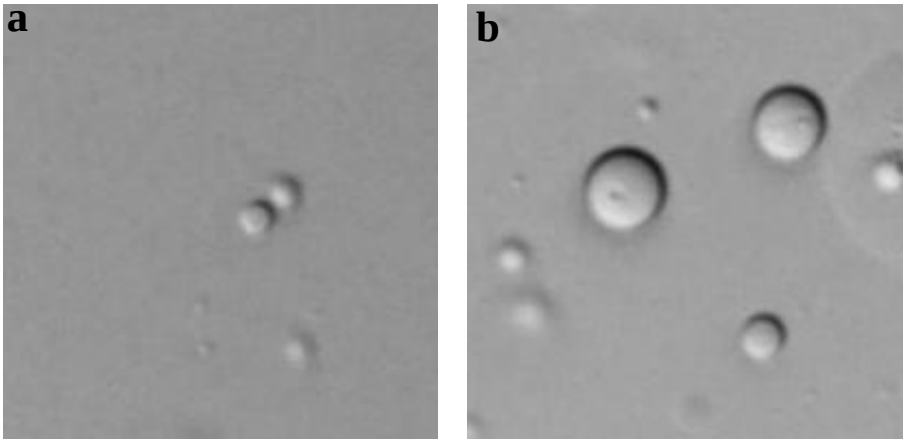

a
hnRNPAB X2
b
hnRNPAB X2 IDR
c
hnRNPAB X2 RBD
d
hnRNPAB X2
XβG RNA
e
hnRNPAB X2 IDR
XβG RNA
f
hnRNPAB X2 RBD
XβG RNA
g
h
i
Circularity
RNA
Protein (μM) 5 12.5 25 50
j
k
l
Average Size (μm)
RNA
Protein (μM) 5 12.5 25 50

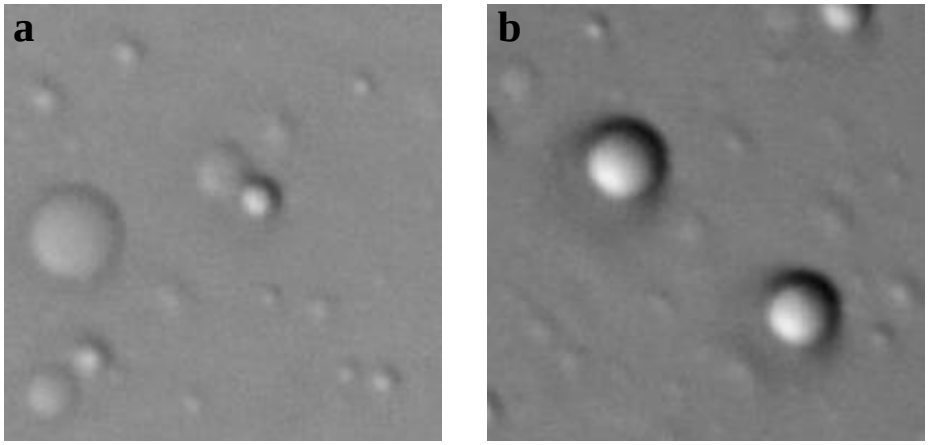
Figure 5: Quantitation of hnRNPAB X2 condensates formed *in-vitro*
(a-f) Fluorescent micrographs of **(a, d)** full-length hnRNPA2 X2, (b. e) hnRNPAB X2 IDR, and **(c, f)** hnRNPAB X2 RBD proteins at 12.5 μM in 20 mM NaPi, pH 7.4, 150 mM NaCl, 10% PEG, and 25 μM of Thioflavin T for fluorescent labeling in the absence (a-c) or the presence (d-f) of 0.25 mg/mL *Xenopus* β-globin (*X*β*G*) RNA. After cleavage of the MBP tags with TEV protease, phase separation was carried out. Images are representative from three or more biological replicates (with independently expressed and purified protein) and three or more technical replicates. Scale bars = 10 μm. **(g-i)** Circularity of the condensates was analyzed using the ImageJ Analyze Particles plugin on a scale of 0-1.0 arbitrary units, with 1.0 representing a perfect circle, for **(g)** full-length hnRNPA2 X2 (green), **(h)** hnRNPAB X2 IDR (blue), and **(i)** hnRNPAB X2 RBD (orange) at 5, 12.5, 25, and 50 μM protein concentrations in the presence (+RNA, hatched bars) and absence (- RNA, open bars) of 0.25 mg/mL *Xenopus* β-globin RNA. **(j-l)** Average size (in μm) of the condensates was determined using the ImageJ Analyze Particles plugin for (j) full-length hnRNPA2 X2 (green), (k) hnRNPAB X2 IDR (blue), and (l) hnRNPAB X2 RBD (orange) at 5, 12.5, 25, and 50 μM protein concentrations in the presence (+RNA, hatched bars) and absence (- RNA, open bars) of 0.25 mg/mL *Xenopus* β-globin RNA. (g-l) Error bars represent standard error of the mean, **** indicates $p<0.0001$, *** indicates $p<0.001$, ** indicates $p<0.01$, * indicates $p<0.05$, and all brackets not shown are not significant ($p>0.05$). Statistics shown are an ordinary two-way ANOVA followed by Tukey's multiple comparisons. All quantification analysis and graphs were created by Jessica Otis and come from O'Connell et. al 2024[5].

We hypothesized that when observing and measuring the circularity of the condensates that each of the three domains formed that they would appear the most circular in the presence of RNA. Even though the condensates formed by hnRNPAB X2 and its IDR and RBD domains appear as round structures (Figure 4A-C, E-G) that display some level of circularity (Figure 5G-I), only the condensates that are formed by the RBD domain can be observed to fuse and wet the slide (Videos 1A-B and 2A-B). Fusing and wetting of the slide are hallmarks that are consistent with liquid-like condensates. Alternatively, condensates formed by the full-length protein (Videos 3 and 4) along with the IDR domain (Videos 5 and 6), cluster together and do not fuse. This feature is an indication of the proteins having more static-like structures. Collectively, these results suggest that with and without RNA both the RBD and IDR domains alone are sufficient

to form phase separated condensates *in vitro*, while full-length hnRNPAB X2 forms amorphous assembly structures *in vitro*.

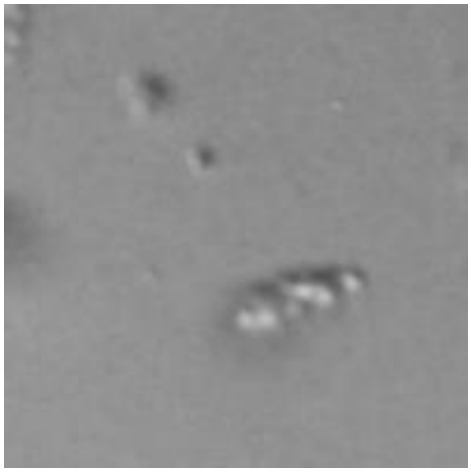


Videos 1: Time lapse videos of hnRNPAB X2 RBD condensates
hnRNPAB X2 RBD condensates fusing **(a)** and wetting **(b)** the slide at room temperature. Condensates were incubated at 50 µM concentration in 20 mM NaPi (pH 7.4), 150 mM NaCl, and 10% PEG, followed by treatment with 0.03 mg/ml TEV protease for 20 minutes. Condensates were imaged on a Nikon Ti2-E Fluorescence microscope, using a 20´ objective. Videos were collected in real time for a total of 20 seconds with frames collected every 0.6 seconds. The display rate is 7 frames per second.

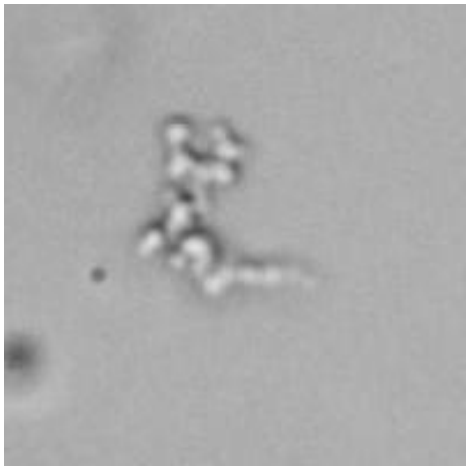


Videos 2: Time-lapse video of hnRNPAB X2 RBD condensates in the presence of RNA.
Condensates formed at room temperature **(a)** fusing and **(b)** wetting the slide. 50 µM hnRNPAB

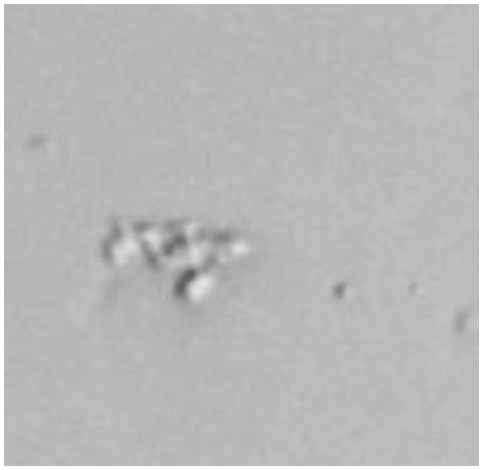
X2 RBD was incubated with 0.25 mg/mL *Xenopus* β-globin RNA in 20 mM NaPi (pH 7.4), 150 mM NaCl, and 10% PEG followed by treatment with 0.03 mg/ml TEV protease. Condensates were imaged on a Nikon Ti2-E Fluorescence microscope, using a 20′ objective. Shown are videos of hnRNPAB X2 RBD condensates captured in real time 20 min. after TEV addition. Videos were captured for 20 seconds total, with frames collected every 0.6 seconds. The display rate is 7 frames per second.



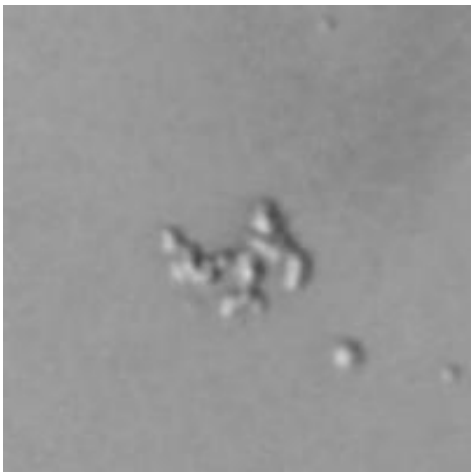
Video 3: Time-lapse video of hnRNPAB X2 full length condensates.
hnRNPAB X2 full length condensates were formed by incubation of 50 µM protein in 20 mM NaPi pH 7.4, 150 mM NaCl, and 10% PEG, followed by treatment with 0.03 mg/ml TEV protease. Condensates were imaged at room temperature on a Nikon Ti2- E Fluorescence microscope using a 20’ objective. Video was taken in real time for a total of 20 seconds with frames collected every 0.6 seconds. Videos begin 20 min. after TEV addition, and the display rate is 7 frames per second.



Video 4: Time-lapse video of hnRNPAB X2 full length condensates in the presence of RNA. Condensates of 50 µM hnRNPAB X2 full length were formed by incubation with 0.25 mg/mL *Xenopus* β-globin RNA in 20 mM NaPi pH 7.4, 150 mM NaCl, and 10% PEG, followed by treatment with 0.03 mg/mL TEV protease. The condensates were imaged at room temperature 20 min. after TEV addition on a Nikon Ti2- E Fluorescence microscope using a 20' objective. Video was taken in real time for a total of 20 seconds, with frames being collected every 0.6 seconds. The display rate is 7 frames per second.



Video 5: Time-lapse video of hnRNPX2 IDR condensates
IDR condensates were formed at room temperature by incubation of 50 µM protein in 20 mM NaPi (pH 7.4), 150 mM NaCl, and 10% PEG followed by 0.03 mg/ml TEV protease treatment. 20 minutes after TEV addition, condensates were imaged on a Nikon Ti2- E Fluorescence microscope using a 20' objective. Video was taken in real time for a total of 20 seconds, with frames collected every 0.6 seconds, and the display rate as 7 frames per second.



Video 6: Time-lapse video of hnRNPAB X2 IDR condensates in the presence of RNA
Condensates were formed at room temperature by incubation of 50 μM hnRNPAB IDR with 0.25 mg/mL *Xenopus* β-globin RNA in 20 mM NaPi, pH 7.4, 150 mM NaCl, followed by treatment with 0.03 mg/ml TEV protease. Condensates were imaged on a Nikon Ti2-E Fluorescence microscope, using a 20′ objective. Shown are videos of hnRNPAB X2 IDR condensates, beginning 20 min. after TEV addition. The video was taken for a total of 20 seconds, with frames collected every 0.6 seconds; the display rate is 7 frames per second.

NELFE undergoes liquid-liquid phase separation

To determine whether NELFE full-length and each of its deletion domains can undergo phase separation *in vitro*, we purified full-length NELFE along with each of its deletion domains. NELFE can form distinct foci in the nucleus which resemble condensates or nuclear bodies that are known to regulate chromatin assembly[22]. It also possesses two low complexity sequences (LCS) and one RNA Recognition Motif (RRM) but does not have a canonical intrinsically disordered region (IDR) that is known to drive LLPS. DISOPRED3 predictions show that NELFE low complexity sequences are highly disordered (Figure 6), suggesting that NELFE may have the capability to undergo phase separation.

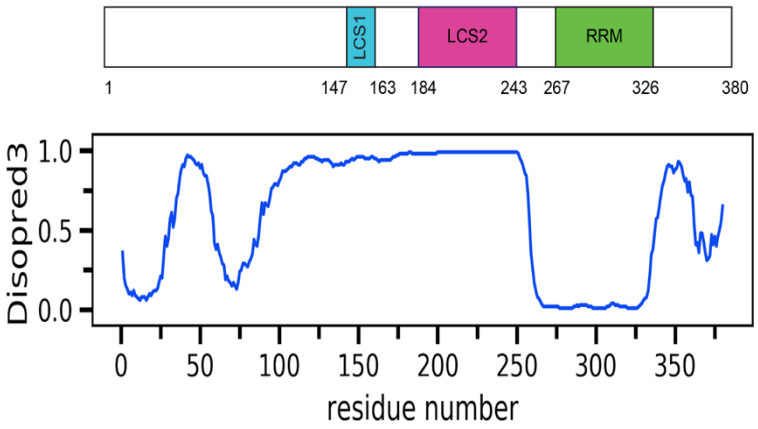


Figure 6: Negative Elongation Factor E (NELFE) has two low complexity regions (LCS) and an RNA recognition motif (RRM)
Disopred3 schematic of NELFE protein along with the domain structures of the constructs.

Using the C-terminal maltose binding protein (MBP) that we have previously used for hnRNPA2 RRMs[23] (Ryan et. al 2020), we purified full length NELFE full and each of its domain deletion variants from *Escherichia coli*. MBP is one of the oldest fusion partners used for producing recombinant proteins in bacterial cells and enhances both the production and solubility of its fusion partner[24]. Each sample was prepared in 20 mM sodium phosphate (pH 7.4) and 150 mM sodium chloride buffer conditions. Upon cleavage of the MBP tag using TEV protease, we found that NELFE full length phase separates only in the presence of polyadenylic acid (Poly A) RNA (Figure 7A). Without RNA, there is no condensate formation (Figure 7E). We then deleted the LCS1, LCS2 or RRM and observed how those deletions impacted phase separation. Deletion of the RRM behaves similarly to the full length NELFE in that it forms liquid droplets only in the presence of Poly A RNA (Figure 7D, 7H). Alternatively, we observed liquid droplet formation of the LCS1 deletion variant both in the presence and absence of RNA (Figure 7B, 7F), while the LCS2 deletion abolished phase separation (Figure 7C, 7G).

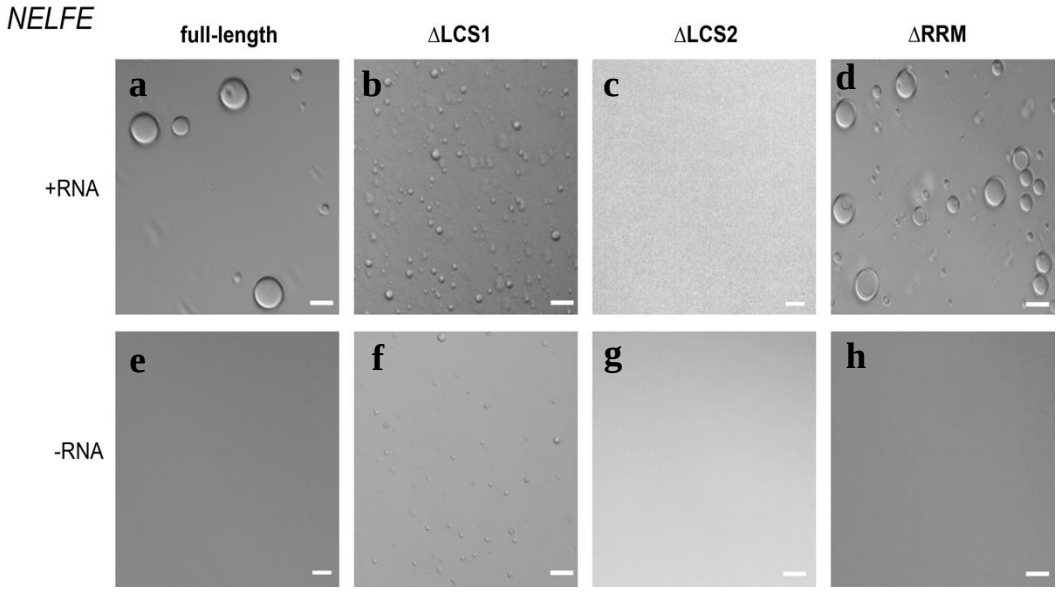


Figure 7: NELFE full length and some of its domains undergo liquid-liquid phase separation
DIC micrographs of **(a,d)** NELFE full length, **(b,f)** ΔLCS1, **(c,g)** ΔLCS2 and **(d,h)** ΔRRM at 50 µM concentration in 20 mM NaPi, pH 7.4, 150 mM NaCl and 0.03 mg/ml TEV protease in the presence (a-d) or absence (e-h) of 0.5 mg/mL Polyadenylic acid (Poly A) RNA. Scale bar = 20 µm.

To complement the phase separation data, we performed a turbidity assay with purified recombinant NELFE to investigate its ability to form liquid droplets. Turbidity assays involve the quantification of the difference in ultraviolet spectra of protein solutions upon phase separation[25]. We assessed NELFE's phase separation via turbidity after 60 minutes with varying Poly A RNA concentrations (Figure 8). The RNA concentrations used range from 0 mg/mL up to 2 mg/mL of RNA. After sixty minutes, at lower RNA concentrations NELFE has higher turbidity values, and then slowly drops off as higher concentrations of RNA are used. As the concentration of Poly A RNA increased, NELFE's turbidity decreased. The data suggests that NELFE displays properties of RNA mediated re-entrant phase separation[26,27].

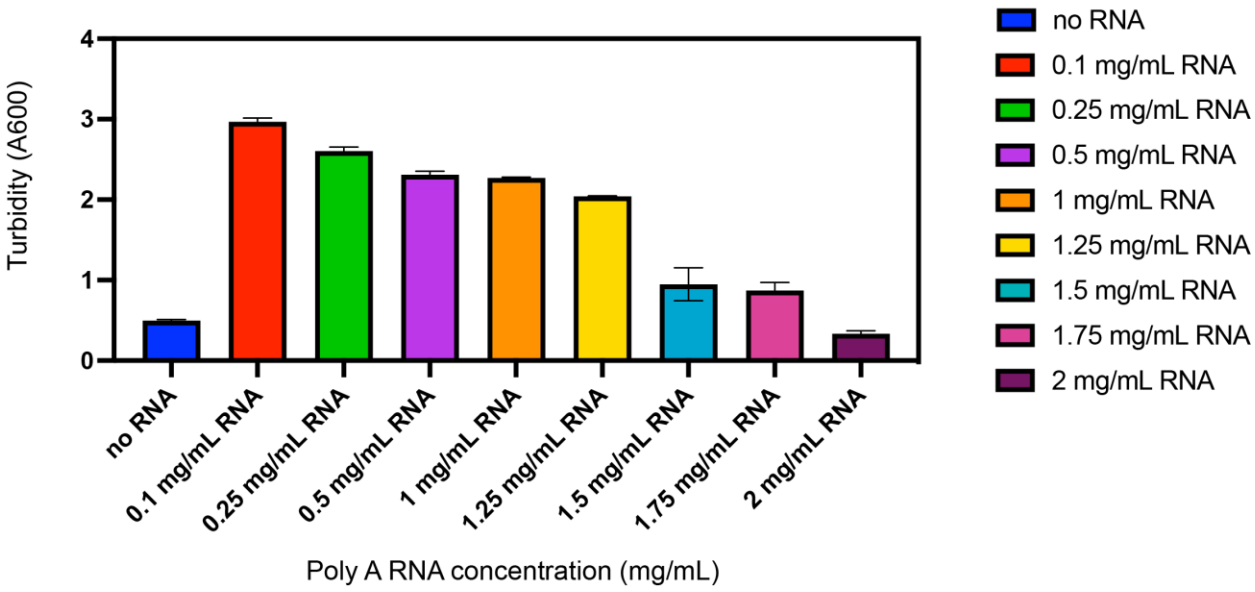


Figure 8: NELFE turbidity decreases with increasing RNA concentration
Turbidity assay of NELFE full length with 50 µM protein in 20 mM NaPi, pH 7.4, 150 mM NaCl, 0.03 mg/ml TEV and increasing concentrations of Poly A RNA up to 2 mg/ml. The absorbance (A_{600}) values are shown after 60 min.

Chapter 4: Discussion and Conclusions

In this work, we used biochemical assays to determine how different structural features of RNA-binding proteins contribute to phase separation properties for two different multidomain proteins. We examined how the multidomain proteins *Xenopus* RBP heterogeneous nuclear ribonucleoprotein AB (hnRNPAB X2) and human Negative Elongation Factor E (NELFE) can phase separate *in vitro*. These proteins contain different structural features such as RNA binding domains, intrinsically disordered regions, and low complexity sequences. Liquid-liquid phase separation is a key mechanism that regulates many aspects of cellular activities that are essential for dynamic chromatin organization and gene regulation[6,28]. Proteins that can undergo LLPS usually contain an IDR that allows for reversible and dynamic multivalent interactions with RNA and/or proteins.

Xenopus RBP hnRNPAB X2 is a protein that contains two RBDs and an IDR and we investigated how it localizes to a biomolecular condensate. This protein is highly enriched in L-body biomolecular condensates and is highly dynamic with them in cells. *In vitro*, full-length hnRNPAB X2 forms condensates, along with the IDR and RBD domains. The IDR and RBD domains are able to phase separate both in the presence and absence of β-globin RNA. Therefore, both the IDR and RBD domains are able to contribute to hnRNPAB X2 self-interactions and phase separation *in vitro*. As previously mentioned, this self-interaction does not require RNA, similar to other RNA binding proteins such as FUS, which can readily phase separate with RNA and whose disordered domain phase separates but does not interact with RNA[29]. The RBD domain exhibited liquid-like properties, as it was able to wet and fuse on the cover slide. Fusing is when droplets are able to "morph" together, while wetting occurs when

adhesive forces cause the droplet to spread across the surface of the solid at the solid-liquid interface. In this case, the solid would be the cover slide. Meanwhile, the full-length and IDR domains seemed to only form more static and amorphous structures, as they clustered together and did not exhibit the properties that are hallmarks of being liquid-like. This indicates that each domain may be interacting differently with the components of RNA and in the protein. The RBD and IDR domains are essential for phase separation to occur, and each domain demonstrates different morphologies.

Through collaboration with Dr. Kim Mowry's laboratory at Brown University, we were able to quantitate specific properties of the condensates that were formed for the full length, IDR and RBD domains of the entire protein. To accomplish this, we stained the protein droplets with Thioflavin T (ThT). Thioflavin T is is a commonly used dye for the study of amyloid pathologies and protein aggregation both *in vitro* and *in vivo*[30]. Although it is most commonly known as a small molecule that gives strong fluorescence upon binding to amyloids and is a commonly used probe to monitor *in vitro* amyloid fibril formation[31], it will fluoresce when portioned into droplets[32] even without amyloid formation and we used it here to mark droplets regardless of amyloid state. For our study purposes we used ThT to quantitate the circularity, average particle number, and average size of the condensates, as this dye is able to go into the condensates.

Dr. Jessica Otis, a scientist in Dr. Mowry's lab, analyzed and created the bar graphs seen in Figures 4I-K and 5G-L to quantify the average number of particles formed, the circularity, and particle size for each individual domain. We found that particle formation for each domain varies depending on whether the RNA is present or not. At higher protein concentrations, RNA has

little to no effect on the number of particles formed. The circularity of the droplets is also not affected by the presence of RNA, as each of the domains relatively exhibited the same degree of circularity of the droplets formed. However, the average size of the particles tended to be larger when RNA was absent from the samples. This indicates that RNA may not be necessary for these interactions to occur or alter any physical properties of the droplets.

Negative Elongation Factor E (NELFE) is a part of a larger complex (NELF) and has been shown to be able to bind to various RNA elements[33]. Recent work shows that the NELF complex can phase separate and condense in response to stress through the IDR of NELFA[17,34]. NELFE does not have a traditional IDR but instead has two low complexity sequences, both of which are predicted to be disordered. Although the LCS is not a canonical IDR, recent work provides evidence that they are able to facilitate LLPS[35]. In this work, we found that full length NELFE can phase separate with the addition of Poly A RNA to form liquid droplets independent of the NELF complex *in vitro*. The deletion of the RNA recognition motif also forms droplets in the presence of RNA, but without RNA no droplets are observed. We found that the deletion of the LCS2 domain completely inhibits droplet formation, whereas the deletion of the LCS1 domain still forms smaller sized droplets with and without RNA. It is likely that the LCS2 domain, which consists of tandem RDRD repeats and is 69 amino acids in length, is required for phase separation to occur and that the LCS1 region is important for specificity of condensate formation.

To further quantify the microscopy data collected, we performed a turbidity assay to examine the role that RNA has on NELFE's phase separation. A turbidity assay utilizes an ultraviolet

absorbance signal to quantify the extent of condensed and aggregated proteins during phase separation that are in the visible range of light[25]. It provides information describing the protein phase separation process, including time-dependent aggregation kinetics. We found that for NELFE after a period of sixty minutes, using low concentrations of RNA resulted in high levels of turbidity, while lower turbidity values were observed for high concentrations of RNA used. This indicates NELFE forms droplets in an RNA dependent manner and suggests NELFE has properties of RNA mediated re-entrant phase separation. In this re-entrant phase separation mechanism, low concentrations of RNA aid in the formation of droplets, while higher concentrations of RNA drive dissolution of these droplets by charge inversion[26]. Taken together, these results show that RBPs are able to induce phase separation.

References

1. Wang, B., Zhang, L., Dai, T., Qin, Z., Lu, H., Zhang, L., & Zhou, F. (2021). Liquid–liquid phase separation in human health and diseases. *Signal Transduction and Targeted Therapy*, *6*(1).
2. Xu, Z., Wang, W., Cao, Y., & Xue, B. (2023). Liquid-liquid phase separation: Fundamental physical principles, biological implications, and applications in supramolecular materials engineering. *Supramolecular Materials*, *2*, 100049.
3. Cai, Z., Mei, S., Zhou, L., Ma, X., Wuyun, Q., Yan, J., & Ding, H. (2023). Liquid–liquid phase separation sheds new light upon cardiovascular diseases. *International Journal of Molecular Sciences*, *24*(20), 15418.
4. Molliex, A., Temirov, J., Lee, J., Coughlin, M., Kanagaraj, A. P., Kim, H. J., Mittag, T., & Taylor, J. P. (2015). Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization. *Cell*, *163*(1), 123–133.
5. O'Connell, L. C., **Johnson, V.,** Otis, J. P., Hutton, A. K., Murthy, A. C., Liang, M. C., Wang, S.-H., Fawzi, N. L., & Mowry, K. L. (2024). Intrinsically disordered regions and RNA binding domains contribute to protein enrichment in biomolecular condensates in xenopus oocytes. *Scientific Reports*, *14*(1).
6. Banani, S., Lee, H., Hyman, A. *et al.* Biomolecular condensates: organizers of cellular biochemistry. *Nat Rev Mol Cell Biol* 18, 285–298 (2017).
7. Sagan, S. M., & Weber, S. C. (2023). Let's phase it: Viruses are master architects of Biomolecular Condensates. *Trends in Biochemical Sciences*, *48*(3), 229–243.

8. Tong, X., Tang, R., J., Wang, W., Zhao, Y., Yu, X., & Shi, S. (2022). Liquid–liquid phase separation in tumor biology. *Signal Transduction and Targeted Therapy*, *7*(1).
9. Niu, X., Zhang, L., Wu, Y., Zong, Z., Wang, B., Liu, J., Zhang, L., & Zhou, F. (2023). Biomolecular condensates: Formation mechanisms, biological functions, and therapeutic targets. *MedComm*, *4*(2).
10. Zigdon, I., Carmi, M., Brodsky, S., Rosenwaser, Z., Barkai, N., & Jonas, F. (2024). Beyond RNA-binding domains: Determinants of protein–RNA binding. *RNA*, *30*(12), 1620–1633.
11. Hentze, M. W., Castello, A., Schwarzl, T., & Preiss, T. (2018). A brave new world of RNA-binding proteins. *Nature Reviews Molecular Cell Biology*, *19*(5), 327–341.
12. 11. Chowdhury, M. N., & Jin, H. (2022). The RGG motif proteins: Interactions, functions, and regulations. *WIREs RNA*, *14*(1).
13. Chong, P. A., Vernon, R. M., & Forman-Kay, J. D. (2018). RGG/RG motif regions in RNA binding and phase separation. *Journal of Molecular Biology*, *430*(23), 4650–4665.
14. Ozdilek, B. A., Thompson, V. F., Ahmed, N. S., White, C. I., Batey, R. T., & Schwartz, J. C. (2017). Intrinsically disordered RGG/RG domains mediate degenerate specificity in RNA binding. Nucleic Acids Research, 45(13), 7984–7996.
15. Neil, C. R., Jeschonek, S. P., Cabral, S. E., O'Connell, L. C., Powrie, E. A., Otis, J. P., Wood, T. R., & Mowry, K. L. (2021). L-bodies are RNA–protein condensates driving RNA localization in *xenopus* oocytes. *Molecular Biology of the Cell*, *32*(22).
16. Gebauer, F., Schwarzi, T., Valcárcel, J., & Hentze, M. W. (2020). RNA-binding proteins in human genetic disease. *Nature Reviews Genetics*, *22*(3), 185–198.

17. Rawat, P., Boehning, M., Hummel, B., Aprile-Garcia, F., Pandit, A. S., Eisenhardt, N., Khavaran, A., Niskanen, E., Vos, S. M., Palvimo, J. J., Pichler, A., Cramer, P., & Sawarkar, R. (2021). Stress-induced nuclear condensation of NELF drives transcriptional downregulation. *Molecular Cell*, *81*(5), 1013–1026.
18. Dang, H., Takai, A., Forgues, M., Pomyen, Y., Mou, H., Xue, W., Ray, D., Ha, K. C. H., Morris, Q. D., Hughes, T. R., & Wang, X. W. (2017). Oncogenic activation of the RNA binding protein NELFE and Myc signaling in hepatocellular carcinoma. *Cancer Cell*, *32*(1), 101–114.
19. Cao, Y., Qin, Y., Zhang, W., Tian, W., Ren, Y., Ren, J., Wang, J., Wang, M., Jiang, J., & Wang, Z. (2023). Structural basis of the human negative elongation factor NELF-b/c/E ternary complex. *Biochemical and Biophysical Research Communications*, *677*, 155–161.
20. Jarnot, P., Ziemska-Legiecka, J., Grynberg, M., & Gruca, A. (2022). Insights from analyses of low complexity regions with canonical methods for protein sequence comparison. *Briefings in Bioinformatics*, *23*(6).
21. Boeynaems, S., S., Alberti, S., Fawzi, N. L., Mittag, T., Polymenidou, M., Rousseau, F., Schymkowitz, J., Shorter, J., Wolozin, B., Van Den Bosch, L., Tompa, P., & Fuxreiter, M. (2018). Protein phase separation: A new phase in Cell Biology. *Trends in Cell Biology*, *28*(6), 420–435.
22. Platani, M., Goldberg, I., Lamond, A.I., and Swedlow, J.R. (2002). Cajal body dynamics and association with chromatin are ATP-dependent. Nat Cell Biol *4*, 502-508.
23. Ryan, V.H., Watters, S., Amaya, J., Khatiwada, B., Venditti, V., Naik, M. T., & Fawzi, N. L. (2020). Weak binding to the A2RE RNA rigidifies HNRNPA2 rrms and reduces

liquid–liquid phase separation and aggregation. *Nucleic Acids Research*, *48*(18), 10542–10554.

24. Lebendiker, M., & Danieli, T. (2010). Purification of proteins fused to maltose-binding protein. *Methods in Molecular Biology*, 281–293.
25. Huang, Y., Bai, Y., Jin, W., Shen, D., Lyu, H., Zeng, L., Wang, M., & Liu, Y. (2021). Common pitfalls and recommendations for using a turbidity assay to study protein phase separation. *Biochemistry*, *60*(32), 2447–2456.
26. Milin, A. N., & Deniz, A. A. (2018). Reentrant phase transitions and non-equilibrium dynamics in membraneless organelles. *Biochemistry*, *57*(17), 2470–2477.
27. Banerjee, P. R., Milin, A. N., Moosa, M. M., Onuchic, P. L., & Deniz, A. A. (2017). Reentrant phase transition drives dynamic substructure formation in ribonucleoprotein droplets. *Angewandte Chemie*, *129*(38), 11512–11517.
28. Jiang, S., Fagman, J. B., Chen, C., Alberti, S., & Liu, B. (2020). Protein phase separation and its role in tumorigenesis. *eLife*, *9*.
29. Burke, K. A., Janke, A. M., Rhine, C. L. & Fawzi, N. L. (2015). Residue-by-residue view of in vitro FUS granules that bind the C-terminal domain of RNA polymerase II. *Mol. Cell.* 60, 231–241.
30. Min, J.-H., Sarlus, H., Oasa, S., & Harris, R. A. (2024). Thioflavin-T: Application as a neuronal body and nucleolar stain and the blue light photo enhancement effect. *Scientific Reports*, *14*(1).
31. Xue, C., Lin, T. Y., Chang, D., & Guo, Z. (2017). Thioflavin T as an amyloid dye: Fibril quantification, optimal concentration, and effect on aggregation. *Royal Society Open Science*, *4*(1), 160696.

32. Wake, N., Weng, S.-L., Zheng, T., Wang, S.-H., Kirilenko, V., Mittal, J., & Fawzi, N. L. (2025). Expanding the molecular grammar of polar residues and arginine in FUS phase separation. *Nature Chemical Biology*.

33. Rao, J. N., Neumann, L., Wenzel, S., Schweimer, K., Rösch, P., & Wöhrl, B. M. (2006). Structural studies on the RNA-recognition motif of NELF E, a cellular negative transcription elongation factor involved in the regulation of HIV transcription. *Biochemical Journal*, *400*(3), 449–456.

34. Xu, W., Ding, M., and Li, P. Struggle for survival: new insights into NELF condensation for adaptive transcriptional reprogramming. (2021). Mol Cell *81*, 896-898.

35. Lee, J., Cho, H., & Kwon, I. (2022). Phase separation of low-complexity domains in cellular function and disease. *Experimental & Molecular Medicine*, *54*(9), 1412-1422.

Appendix

Table 1

Primer	Sequence
X2AB_FL_Fwd	TCCGACACCGAGCAGCAGTGTCTAGAA
X2AB_FL_Rev	GTTCTTACTGTAGTCATAGCCTGGTCCATATCCATAATAG
X2AB_RBD_Fwd	AAAATGTTTGTTGGTGGCTTGAGCTGG
X2AB_RBD_Rev	TGGTTGTGCAATCTTTATCTCACACTTG
X2AB_IDR_Fwd	AAAGAAGTGTATCAGCAACAGTATGGCGG

Table 1: Primers used in this study. All Forward (Fwd) and Reverse (Rev) primers are listed from 5′ to 3′. The following abbreviations are used: X2AB (hnRNPAB X2), FL (full-length). IDR (intrinsically disordered region), and RBD (RNA binding domain).