Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Travis Loya Date

Investigating the structure and function of a prion-like domain in the Nrd1-Nab3-Sen1 transcription termination system

By

Travis Jose Loya
Doctor of Philosophy

Graduate Division of Biological and Biomedical Science
Biochemistry, Cell, and Developmental Biology

Daniel Reines, Ph.D.
Advisor

William Kelly, Ph.D.
Committee Member

Graeme Conn, Ph.D.
Committee Member

Nicholas Seyfried, Ph.D.
Committee Member

Accepted:

Roger Deal, Ph.D.
Committee Member

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Investigating the structure and function of a prion-like domain in the Nrd1-Nab3-Sen1 transcription termination system

By

Travis Jose Loya

B.S. Virginia Tech 2010

Advisor: Daniel Reines, Ph.D.

An abstract of

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Biochemistry, Cell, and Developmental Biology

2018

Abstract

Investigating the structure and function of prion-like domains in the Nrd1-Nab3-Sen1 transcription termination system

By Travis Jose Loya

From a single genome, yeast must adapt their transcriptional profile to respond to or initiate changes in the cell. A key step in this process is transcriptional termination, which in yeast is carried out by two systems for RNA polymerase II (pol II) transcripts. This work focuses on the Nrd1-Nab3-Sen1 (NNS) complex responsible for regulating several classes of small, non-coding RNAs. The complex consists of two RNA recognition motif (RRM)-containing RNA-binding proteins, Nrd1 and Nab3, and a helicase Sen1. Nrd1 and Nab3 interact with pol II, nascent transcripts, and processing factors to coordinate termination and processing of transcripts. While the RNA recognition motifs and other structured interaction domains have been studied in these proteins, we recently discovered mutations in a previously uncharacterized region of Nab3 that cause termination defects and in some cases death. This thesis describes the analysis of this domain.

Interestingly, outside of a small α-helix at the very C-terminus, much of the domain has no predicted structure, as expected from its low complexity sequence. Biochemical characterization of the region revealed it contains a prion-like domain enriched for glutamine and proline that forms amyloid-like fibers *in vitro.* This finding is consistent with a growing body of research showing that prion-like domains (PrLDs) are overrepresented within RNA-binding proteins and are responsible for mediating homo- and heterotypic protein-protein interactions. From this finding, we explored the ability of heterologous prion-like domains to complement Nab3's prion-like domain and found subsets that rescue function.

Prion-like domains are also implicated in facilitating formation of various RNA granules in the cell. These granules are thought to be membrane free compartments where machinery involved in RNA metabolism is sequestered to perform its cellular function. Examples include the nucleolus, stress granules, P-bodies, intranuclear quality control/ juxta nuclear quality control (INQ/JUNQ), and insoluble protein deposit (IPOD). In response to glucose starvation, Nab3 and Nrd1 are localized to a novel nuclear granule. This dissertation describes how Nab3's localization is dependent upon the amyloid forming properties of the prion-like domain, is reversible upon refeeding the yeast, and function in mediating granule assembly. This work deepens our understanding of the regulation and function of prion-like domains in RNA metabolism across eukarya.

Investigating the structure and function of a prion-like domain in the Nrd1-Nab3-Sen1 transcription termination system

By

Travis Jose Loya

B.S. Virginia Tech 2010

Advisor: Daniel Reines, Ph.D.

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Biochemistry, Cell, and Developmental Biology

2018

Table of Contents

Figure and Tables Table of Contents

Tables

Abbreviations

Aβ - amyloid beta

BrU - 5-bromouridine

BSA - bovine serum albumin

CPF/CF - cleavage/polyadenylation factor 1A and B

CPEB - cytoplasmic polyadenylation element binding protein

CID – C-terminal interacting domain

CTD - carboxy terminal domain

CUT – cryptic unstable transcript

DHFR - dihydrofolate reductase

DMSO - dimethyl sulfoxide

EDTA - ethylenediaminetetraacetic acid

GFP - green fluorescent protein

HFIP - hexafluoroisopropanol

hnRNP - heterogeneous nuclear ribonucleoprotein

IDR - intrinsically disordered region

INQ – Intranuclear quality control

IPOD – insoluble protein deposit

IPTG - isopropyl β-D-1-thiogalactopyranoside

JUNQ – juxtanuclear quality control

LCD – low complexity domain

LLPS - liquid-liquid phase separation

NNS - Nrd1-Nab3-Sen1

PCR - polymerase chain reaction

PrLD - Prion-like domain

Pol II - RNA Polymerase II

PTM – post-translational modification

RNP - ribonucleoprotein

RRM - RNA recognition motif

SC – synthetic complete

SDD-AGE - semi-denaturing detergent agarose gel electrophoresis

SDS - sodium dodecyl sulfate

snRNA -small nuclear RNA

snoRNA – small nucleolar RNA

SUT – stable uncharacterized transcript

TEV - tobacco etch virus

TRAMP – Trf4/5-Air1/2-Mtr4

TSS – transcriptional start site

YPD – Yeast peptone dextrose

Chapter 1: Background and significance

1.1 Transcription regulation in *Saccharomyces cerevisiae*

From a single genome, yeast must find a way to adapt their gene expression profiles to efficiently utilize limited resources, deal with changing environmental conditions, accelerate growth when resources are abundant, conserve cellular resources in times of stress, adjust to varying nutrient sources, and coordinate decisions about reproduction, both sexual and asexual. Yeast achieve these outcomes through various modes of transcriptional regulation.

Transcription can be divided broadly into three stages. First, transcription initiation occurs at accessible promoter regions of genes. Typically, a set of gene-specific regulatory factors bind near the site of transcription initiation. When necessary, these factors facilitate making the DNA accessible for the transcriptional machinery by recruiting chromatin remodeling machinery. Once accessible, a set of proteins termed the general transcription factors will bind to core promoter elements in DNA to facilitate RNA polymerase II (pol II) recruitment. With pol II bound to the general transcription factors, the complex melts the DNA at the transcriptional start site (TSS), which allows for RNA synthesis to occur. In recent years it has been shown that these transcriptional initiation events are promiscuous and widespread throughout the genome; even in regions between genes. These abundant transcription initiation events have been termed "pervasive transcription" [1]. Left unregulated, these events would result in cellular dysfunction if initiation complexes compete across promoter regions resulting in over or under expression of needed transcripts, accumulation of these RNAs could result in inefficient use of metabolic resources by competing with the machinery used in RNA metabolism. To prevent this from occurring, yeast regulate transcription by employing two modes of transcription termination of pol II transcripts.

After initiation, a short stretch of RNA is polymerized, pol II disengages from many of the general transcription factors, at which time transcription moves into elongation. In yeast, specific residues in the C-terminal domain (CTD) of the large subunit, Rpb1, of pol II signals the movement from initiation to elongation and results in a shift in factors associated with pol II. Mediating the change in associated factors are post-translational modification (PTMs) made to a conserved heptapeptide repeat (Y-S-P-T-S-P-S) on the C-terminus of the enzyme. The number of heptapeptide repeats varies by organism, yeast have 26 repeats while humans have 52. However, across eukarya, each residue within the repeat can be modified by either phosphorylation (tyrosine, serine, threonine) or isomerization (proline). Of these modifications, the roles of Ser-2 and Ser-5 phosphorylation are best understood with others showing more limited or gene-specific effects. During initiation, Kin28, a subunit of TFIIH, phosphorylates the Ser-5 position of the CTD repeats. Ser-5 phosphorylated CTD is a target for the poly adenylation site (PAS)-independent Nrd1-Nab3-Sen1 (NNS) complex[2] (Fig 1-1). NNS is capable of termination early in a transcriptional unit when transcripts contain Nrd1 and/or Nab3 binding sites. After moving further into elongation, the CTD is dephosphorylated on Ser-5 by Rtr1 and Ssu72 and phosphorylated at Ser-2 by Ctk1 and Bur1 [3-6]. As elongation continues, Ser-2 phosphorylated CTD is a target for another termination system in yeast, the PAS-dependent system. This system, composed of the cleavage/polyadenylation factor 1A and B (CPF/CF) complex, is responsible for processing transcripts terminated late in the transcriptional unit, that are cleaved at a PAS resulting in the addition of a polyadenylate tail by Pap1. The tail prevents degradation, promotes export to the cytoplasm, and recruits translation factors to the mRNA. This dissertation will focus on the NNS complex, its role in transcriptional termination in yeast, and specifically the role of a previously

uncharacterized domain of Nab3 that has unusual and important properties. Several excellent reviews cover initiation, elongation, and termination in detail [7-10].

1.1.1 An overview of Nrd1-Nab3-Sen1 termination

In yeast, a three-protein complex, NNS, is involved in short noncoding transcript termination. The identities of the three proteins originated from distinct investigations [11-13]. Only later was it determined that they form a complex and that the complex regulates termination and coordinates processing of small noncoding RNAs[14]. NNS regulates a large number of short noncoding RNAs that can be divided into several classes. One class is small nucleolar RNAs (snoRNAs), which, upon termination, are 3'-end trimmed to create mature snoRNAs [15]. Mature snoRNAs are involved in directing the modification of other cellular RNAs. A second well-studied class is small nuclear RNAs (snRNAs), which are involved in regulation of splicing. All but snRNA U6 are pol II transcripts. After termination by NNS, snRNA is cleaved at its 3' end then processed by the exosome to create a mature snRNA. Other noncoding RNAs terminated by NNS are a large and diverse class of transcripts. As mentioned above, recent advances have shown that eukaryotic genomes have pervasive transcription of non-protein coding RNAs [1]. These pervasive transcripts are found throughout the genome, are subdivided based on their location in the genome, their stability, and how they are processed by the post transcriptional machinery. One class of unknown function is stable uncharacterized transcripts (SUTs) [16], others called cryptic unstable transcripts (CUTs) are degraded rapidly and only detected when the exosome is disabled, [16-19]. Another group called Xrn1-sensitive unstable transcripts, are enriched in strains lacking Xrn1, a cytoplasmic 5' to 3' exonuclease. Finally, a class of antisense transcripts have been found with some evidence that they may function to repress transcripts analogous to RNA interference in higher eukaryotes [20-24].

1.1.2 Factors involved in NNS termination

Nab3

Nab3 was discovered in a biochemical assay to identify nuclear polyadenylated RNA-binding proteins [25]. Nab3 has no direct homologues in higher eukaryotes but does have similarity to a eukaryotic RNA-recognition motif (RRM) containing protein, hnRNP-C. Apart from the RRM, Nab3 contains two low complexity domains (LCDs), stretches of primary sequence biased towards a small subset of amino acids. The N-terminus has a non-essential aspartic acid/glutamic acid-rich domain (D/E) of unknown function. The C-terminus contains a proline/glutamine-rich domain which was shown to be essential but prior to our work had no established function in the cell[26, 27]. The RRM of Nab3 was initially thought to preferentially bind UCUU but subsequent work has expanded the consensus to UCUUG, CUUG, and UCUU sequences [28-31]. Nab3 contains domains which are known to interact with Nrd1, another RRM-containing RNA-binding protein in the NNS complex, and Sen1, a helicase responsible for terminating chain elongation by releasing pol II from transcripts by an undetermined mechanism [14, 32]. NNS-terminated transcripts are often targeted by the nuclear exosome for degradation, and Nab3 has been shown to interact with Rrp6 in a Nrd1and RNA-independent manner [33].

The C-terminal LCD of Nab3 has been characterized as a prion-like domain (PrLD), a subset of LCDs that have sequence similarity and biochemical features of known prions [34]. Nab3's PrLD forms amyloid-like filaments in vitro, and in response to glucose starvation, Nab3 localizes to a unique nuclear granule in a PrLD dependent manner [35-37]. Interestingly, mutations that destroy the amyloid forming properties of the PrLD are lethal, implying a role for self-assembly of Nab3 in its essential cellular function [36]. Nab3 also contains a α-helical domain at

the C-terminus of the protein adjoining the PrLD. This helix is similar to the helical domain of human hnRNP-C and is necessary for full termination competence of the NNS complex [38].

Nrd1

Nrd1 was first identified in a screen for transcription termination factors [13]. Nrd1, like Nab3 is an RRM-containing protein with a preferred recognition sequence of (A/U)GUAAA [30]. It also has a glutamine/asparagine (Q/N)-rich PrLD [34]. Nrd1 contains a pol II C-terminal domain (CTD)-interacting domain (CID) similar to those found in splicing factors in mammals [13, 39]. The CID has been shown to preferentially bind to the Ser-5 phosphorylated CTD of pol II, helping facilitate recruitment of NNS to pol II during transcription initiation and early elongation [40-42]. Surprisingly, the CID can be deleted with no effect on viability [43]. This tolerance of the loss of the CTD is due to a network of interactions, both protein-protein and protein-nucleic acid, that form a robust termination complex [27]. Nrd1 also directly binds two factors responsible for processing NNS-terminated transcripts, Rrp6 and Trf4, components of the nuclear exosome and the exosome cofactor TRAMP complex, respectively [44].

Nrd1 binding to RNA is mediated *via* its RRM with binding enhanced while heterodimerized to its RRM-containing partner, Nab3 [26, 28, 43]. Similar to its NNS binding partner, Nab3, Nrd1 has been identified as a critical factor for proper termination of a subset of small noncoding RNAs, including snoRNAs, snRNAs, and CUTs [45-47]. Beyond regulation of these small noncoding RNAs, studies looking at transcriptome-wide binding of NNS have shown to be associated with mRNA encoding genes, often when the size of the open reading frame (ORF) falls within the range of NNS activity, approximately 450 nt from the TSS [29, 31, 48-50]. Like

Nab3, Nrd1 also contains a Q/N rich PrLD that assembles into higher ordered filaments *in vitro.* The Nrd1 PrLD also forms high molecular weight complexes in SDD-AGE analysis [36].

Sen1

SEN1 and the family of *SEN* genes were discovered in a genetic screen for loss of tRNA splicing endonuclease activity [12]. *SEN1* is essential and encodes a large nuclear protein with characteristic motifs for nucleic acid helicase and nucleotide binding of Superfamily I [13, 51]. It is a low abundance protein, with roughly 125 copies per cell, a markedly lower prevalence than the other two members of the NNS complex, Nab3 and Nrd1, which are present at 5,800 and 19,600 copies per cell, and lower relative to the number of genes regulated by NNS [52]. This raises the question of the stoichiometry of NNS complexes during termination and of the limiting factors in NNS function. As observed for Nab3 and Nrd1, Sen1 is involved in the termination of a range of short noncoding RNAs [13]. Nuclear run-on assays found that NNS was a distinct mechanism of termination from mRNA generating, polyadenylation-coupled termination[14]. A two-hybrid assay showed that Sen1 plays a role outside of transcription termination. It was found to interact with Rad2, a deoxyribonuclease, and Rnt1, a ribonuclease [32]. This suggests a role for Sen1 in DNA damage repair. This finding is in line with the role of Sen1's human homolog, senataxin, which is not involved in termination as no NNS complex exits in humans [53-55].

The precise mechanism by which Sen1 terminates transcription is unclear. The prevailing model is that Sen1 pulls the transcript from the active site similar to Rho-dependent termination in bacteria, but no helicase activity was detected during this particular study [56]. The *Saccharomyces pombe* homolog however has been shown to exhibit helicase activity [57]. More recently, helicase activity has been confirmed for a purified helicase domain of Sen1 and the domain alone is sufficient to dislodge an elongation complex from RNA *in vitro*. Both the domain

and full-length Sen1 can translocate on nucleic acid templates and dislodge elongation complexes *in vitro.* Interestingly, it has been reported that Sen1 seems to move more efficiently on a single-stranded DNA (ssDNA) template than a double stranded template, suggesting there may be interaction with ssDNA in the transcription bubble [58].

Trf4/5-Air1/2-Mtr4 complex and the exosome

Short noncoding RNAs undergo one of two fates post termination. snRNAs and snoRNAs are 3' end trimmed to create the mature RNA, while unstable RNAs such as CUTs and antisense transcripts are targeted for complete degradation by the nuclear exosome. Two complexes mediate the post-transcriptional fate: the TRAMP polyadenylation complex (Trf4/5, Air1/2, and Mtr4), and the nuclear exosome [17, 59, 60]. The nuclear exosome, consisting of the 9 subunit core exosome, Rrp44, and the nuclear exonuclease Rrp6, is necessary for trimming the 3'-end of snRNAs and snoRNAs to create the mature RNA as well as degrading unstable transcripts [16, 18, 19, 61, 62].

TRAMP adds short oligoadenylate tails to RNA, these short oligoadenylate tails target RNAs for degradation. TRAMP complexes consist of a Trf4/5 subunit, a poly(A) polymerase which adds the oligoadenylate tail and Mtr4, a helicase which governs tail length and manages end processing by nucleases [63, 64]. Finally, Air1/2 are zinc knuckle proteins that facilitate RNA binding by the complex. The exact mechanism by which NNS, TRAMP, and the exosome assemble to coordinate processing is not clear, but some aspects of it are known. Using its CID, Nrd1 interacts with the CTD of pol II and with Trf4/5 of TRAMP. Since both of these interactions are mediated by the same region on Nrd1, the interactions are mutually exclusive making Nrd1 a key regulator in the sequence of events between termination and degradation. TRAMP has been

shown to stimulate exosomal degradation by feeding transcripts into the exosome pore, again linking termination and degradation [65, 66].

1.1.3 NNS recruitment to target transcripts

How does the transcriptional machinery decide between NNS vs CPF/CF mediated termination? NNS termination seems to act within a window near the promoter of transcriptional units, as the transcriptional machinery moves downstream of the TSS the likelihood of NNS termination decreases. This phenomenon was first elucidated in studying the regulation of *IMD2* [41, 67]. The distance between the TSS and Nab3 binding sites affected NNS termination. As the distance grows, termination activity decreases. In order to accomplish this position dependent regulation selection, the transcriptional machinery utilizes an array of interactions, both protein-protein and protein-nucleic acid, all of which contribute to recruitment of termination complexes to sites of active transcription. For NNS, Nab3 and Nrd1 both interact with RNA, Nrd1 preferentially interacts with the Ser-5 phosphorylated CTD of pol II, and Nab3 and Nrd1 interact with each other to recruit Sen1 to terminate transcription [26]. This vast array of interactions helps NNS to target transcripts and is also responsible for the robustness of NNS activity, as loss of a single point of interaction is well tolerated by the cell but cumulative losses become harmful [27].

In spite of these redundant interactions, there seems to be occasions where NNS fails to terminate properly and the polyadenylation-coupled termination complex acts as a fail-safe mechanism [15]. The inverse is also true, where polyadenylation-coupled termination fails and the NNS complex functions as a fail-safe mechanism [29, 31, 68]. These distinct termination complexes provide the cell a layer of protection from limited readthrough errors of the other termination machinery.

1.1.4 NNS regulation at sites of transcriptional attenuation

As mentioned above, there is overlap between the termination targets of the two systems in yeast. A study of the genome wide distribution of Sen1 binding sites showed it at a subset of mRNA encoding ORFs. These ORFs tend to be short (~200-550bp), barely longer than the 450nt window associated with NNS termination and much shorter than the average yeast ORF of ~1.2kb [40, 69]. Supporting this model of NNS regulation at mRNA ORFs was an increase in expression of a small number of genes in response to Nrd1 depletion [47].

A well-characterized example of NNS regulation of an mRNA generating ORF is at *NRD1* itself. An NNS recognition sequence is contained near the 5'end of the *NRD1* transcript auto-regulating its expression [70]. Another example is the regulation of *IMD2*. *IMD2* encodes the rate limiting step for guanine nucleotide biosynthesis and TSS selection of *IMD2* is regulated by cellular GTP levels. When GTP levels are high, transcription starts at G-rich upstream TSSs which are 5' to an NNS-binding site. This results in NNS recruitment to the transcript, early termination and degradation. At low GTP levels, pol II cannot use the G-rich upstream start sites as elongation cannot occur due to sub-optimal GTP levels and pol II instead scans the promoter to a TSS located 3' to the NNS binding site. This allows for expression of the *IMD2* ORF and termination by the polyadenylation-coupled machinery [41, 71, 72].

1.2 Prions, Prion-like domains (PrLDs), and RNA-binding proteins

Prions are self-propagating infectious protein isoforms with the ability to shift conformation from a soluble form to a heritable aggregated amyloid-like state in cells. This conformational shift is dependent upon domains that are often biased towards a small subset of amino acids thought to be needed for amyloid formation, however amyloid formation alone is not sufficient to be a prion. When in an amyloid-like state, proteins assume a characteristic highly ordered, in-register anti-parallel cross β-sheet structure [73, 74]. This state can be confirmed by an assortment of assays, including Thioflavin T binding, Congo Red binding, X-ray fiber diffraction, semi-denaturing detergent agarose electrophoresis (SDD-AGE) and electron microscopy (EM).

In mammals, prions are disease causing agents, notably Creutzfeldt-Jakob disease, variant Creutzfeldt-Jakob disease, fatal familial insomnia, and Kuru in humans as well as Scrapie and Bovine Spongiform Encephalopathy (Mad Cow disease) in sheep and cattle, respectively. In yeast, prions are often benign, and in some cases, beneficial [75, 76]. To explore the depth of prions in the yeast proteome, the Lindquist lab developed a hidden Markov model to scan for potential prions by training it with four known prions, Sup35, Ure2, and Rnq1, and New1 under the assumption that the sequence properties of confirmed prions would be predictive [34]. The algorithm found nearly 200 proteins with domains similar to the training set and these domains were called prion-like domains or PrLDs. Only 19 could form genuine heritable prions and only one of these PrLD containing proteins, Mot3, did so naturally, but many exhibited a subset of the properties of prions. They found that within the set of predicted PrLD containing proteins were a large number of nucleic acid-binding proteins, including Nab3.

Two other papers, this time studied in mammalian cells, showed that many proteins in higher eukaryotes also contain PrLDs and that PrLD-containing proteins involved in nucleic acid metabolism are overrepresented [77, 78]. In a case of serendipity, a biotinylated isoxazole compound developed to precipitate factors involved in stimulating differentiation of neuronal progenitor cells caused spontaneous precipitation of a large number of proteins in cell lysates. Upon mass spectrometry analysis there was found to be an overrepresentation of RNA-binding proteins, especially those containing RRM or hnRNP-K homology (KH)-domains. Many of those proteins also contained long stretches of low complexity sequences, which could be potential PrLDs. RRM and KH-domains would confer RNA binding specificity while PrLDs are potential assembly domains. From these two properties it is possible that these RNA binding proteins would be found in various RNA-containing granules, which had been studied for years but no clear mechanism for their formation was known. From the mass spec data, 162 proteins were precipitated. These proteins were cross referenced with a data set of all proteins known to be associated with granules and 53 of the precipitated proteins were on the list of known granule-associated proteins. Further biochemical analysis showed that *in vitro* many of the proteins, when purified, were able to form liquid droplets in a PrLD-dependent manner that were de-mixed from the surrounding solution. They were amyloid-like, but instead of being irreversibly converted to amyloid they could dynamically switch states. Heterotypic interactions were also possible, opening up the possibility that within a cell PrLDs regulate protein solubility and a variety of proteins may be localized to a granule in a PrLD-dependent manner. These findings offered a potential mechanism for regulating the formation of various non-membrane delimited granules within the cell.

1.3 RNP Granules, RNA and PrLDs

Known granules in the cell

RNP granules are present across eukaryotic cells. Nuclear granules include the nucleolus, paraspeckles, Cajal bodies, and INQ [79-81]. The cytoplasm is home to stress granules, which hold mRNA, RNA-binding proteins, and translational factors, P-bodies containing mRNA and degradation machinery, and IPOD, a storage site for insoluble amyloid [82]. Some cell types contain unique granules. For example, oocytes contain p-granules (a.k.a., germ granules), which sequester maternal RNAs to provide the template to drive protein synthesis in offspring. Neurons contain mRNP granules which help move mRNPs along axons to the synapse where they play a role in synaptic plasticity [83]. Awareness of some of these granules date back nearly two centuries, but it is only recently that we have started to understand the biochemical underpinnings of their formation.

Interactions regulating granule formation

An array of intermolecular interactions participate in RNP granule formation, and a given type of interaction may be pre-eminent for the formation of specific granules, though the order of assembly is typically not clear for *in vivo* formation. Evidence exists for RNA-RNA, RNA-protein, and protein-protein interactions driving granule formation and *in vivo* it is likely that all three interactions contribute to their biogenesis.

RNA-RNA interactions have recently been demonstrated *in vitro* using purified RNAs from yeast which spontaneously assemble into RNA granules. An example of this is the *in vivo* assembly of nuclear paraspeckles driven by expression of Neat1 lncRNA and that depletion of Neat1 results in loss of paraspeckles [84, 85].

The bulk of research has focused on protein-mediated formation of RNP granules for several reasons. First, a common finding is that proteins involved in RNP granules contain some form of intrinsically disordered region (IDR), regions which have no predicted secondary structure but may be able to assume a defined structure under certain conditions. IDRs are polypeptide chains that do not adopt a unique fold but instead can interconvert between multiple conformations. A subset of these IDRs would include the PrLDs discussed above. An example of PrLD-dependent stress granule formation is TIA1 in human cells. Overexpression of TIA1 drives granule formation and deletion of its PrLD abrogates granule formation. Substituting the NM-domain of Sup35 rescues stress granule formation when replacing the PrLD of TIA1 [86].

Second, within a framework where the ability to self-assemble is key for granule formation, it has been repeatedly observed that PrLD-containing proteins can self-assemble *in vitro* by various mechanisms including liquid-liquid phase separation (LLPS), a process in which two liquids de-mix from each other. In biologically relevant phase separation processes, macromolecules in solution condense into dense liquid droplets. The droplets contain high concentrations of the macromolecule and co-exist with a light phase, which contains a low concentration of the macromolecule [87]. This *in vitro* assembly may not be identical to the protein's *in vivo* behavior, but the fact that purified proteins can self-assemble by a similar process, LLPS, makes them a candidate for driving RNA granule formation [88-91].

Finally, many disease states involving proteins which localize to various RNA granules contain mutations which affect their ability to assemble. Mutations in FUS, TDP-43, TIA1, hnRNPA1, and hnRNPA2B1 have all been found in various neurological and muscular diseases

with patients exhibiting aberrant biology related to RNA granules, typically higher aggregation loads or fibrillization at higher rates than healthy individuals [89-92].

Protein mediated assembly is not necessarily driven solely by disordered domains of proteins. Structured domains can also contribute to higher ordered assemblies of proteins. In fact, many PrLDs are flanked by structured α-helices which contribute to assembly as well as function[38, 91, 93]. Post-translation modifications of proteins also help regulate the assembly states of proteins in and out of granules. For example, the MEG (maternal-effect germline defective) proteins are required for fertility. Wang et al. demonstrated that MEG-1 and MEG-3 are substrates of a kinase, MBK-2/DYRK and a phosphatase, PP2A. Phosphorylation and dephosphorylation drive granule disassembly and assembly respectively [94].

RNP Granules as a stress response

As implied by their name, stress granules are a cellular response to stress. In the case of mammalian cells, many forms of stress induce their formation: trapping mRNAs in polysomes with cycloheximide, glucose starvation, oxidative stress, hypo-osmotic stress, and hyper-osmotic stress in mammals. Interestingly, in yeast only glucose starvation and cycloheximide treatment induced granule formation[95]. A finding that was critical to the development of the final chapter of this dissertation was the response of Nab3 and Nrd1 to glucose deprivation described by the Corden laboratory[96]. While investigating potential interactions of NNS termination on other cellular pathways these authors observed a shift in the nuclear localization of Nab3 from the nucleoplasm to a discrete, single puncta in approximately 25% of cells. The Nab3-containing granule also contained Nrd1, its binding partner in the NNS complex, but did not contain markers

of any other yeast RNP granules tested for. Nrd1 was also dephosphorylated under glucose stress suggesting a PTM may have a role in Nrd1 localization to the granule.

1.4 NNS, PrLDs, and a novel granule

My interest in studying NNS termination began after performing a genetic screen using a galactose inducible green fluorescent protein (GFP) reporters containing an NNS-binding site between the promoter and the TSS. In this system mutations affecting NNS-mediated termination would be detected as an increase in GFP fluorescence. Fluorescence activated cell sorting (FACS) was used to identify GFP positive cells which were sorted from the GFP negative population and whole genome sequenced to identify mutant genes responsible for termination defects. A set of Nab3 mutants were identified with nonsense mutations in an uncharacterized C-terminal domain. The smallest mutation resulted in a loss of 19 amino acids from the end of the protein and yielded clear termination defects by Northern blot analysis of SNR33, a snoRNA regulated by NNS. The role of this region was poorly understood other than the observation that truncating roughly 130 amino acids from the C-terminus was lethal, implying some essential function exists within the domain [26, 35].

Subsequent studies showed that a point mutant located two amino acids from the C-terminus caused termination defects. The mutation corresponded to a putative α-helix with similarity to hnRNP-C, a mammalian hnRNP whose α-helix is key to tetramerization of the hnRNP in its role as an RNA packaging factor to deliver RNA for export [93]. Similarly, the α-helix of Nab3 was able to facilitate tetramerization of Nab3 based on crosslinking analysis and size exclusion chromatography of a thioredoxin fusion of the C-terminal 40 amino acids of Nab3. Mutational analysis of the C-terminal domain also confirmed a lethal phenotype upon removal of

134 amino acids from the C-terminus [38]. This essential region of Nab3 was particularly interesting since Alberti et al.'s [34] *in silico* screen for prions found a putative PrLD in Nab3 that overlapped the essential C-terminal domain. That finding, along with Kato et al., [97] led to a model in which Nab3 participates in assembly of a nuclear granule. Previous work on the mammalian protein TDP-43 showed its PrLD was flanked by an α-helix, similar to the arrangement in Nab3[91]. Assimilation of these three reports with our lab's findings about the essential Nab3 PrLD, is the starting point for this dissertation.

Fig 1-1

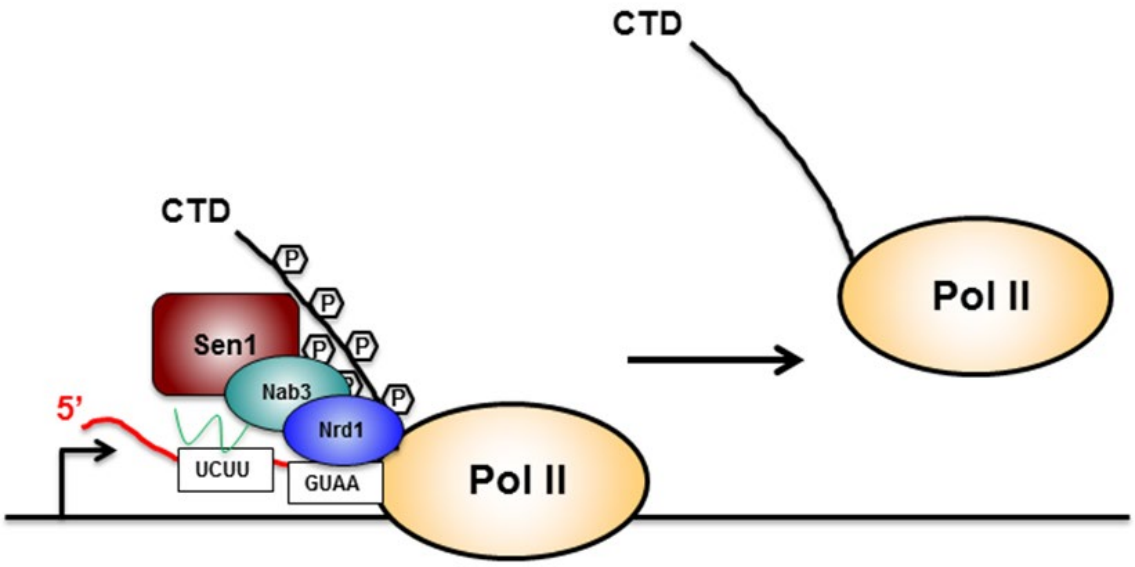


Fig 1-1. Nrd1-Nab3-Sen1 mediated termination. Nrd1 Nab3 Sen1 termination depends on the recognition of specific sequences in nascent RNA's recognized by Nrd1 and Nab3. Upon binding, they recruit Sen1 helicase which causes RNA Polymerase II release from the transcript.

Chapter 2

Amyloid-like Assembly of the Low Complexity Domain of Yeast Nab3

Travis Loya contributed to Fig. 2-3, and was responsible for Figs. 2-5 and 2-6

This is an Accepted Manuscript of an article published by Taylor & Francis in

Prion on 19 February, 2015 , available online:

https://www.tandfonline.com/doi/full/10.1080/19336896.2014.997618

2.1 ABSTRACT

Termination of transcription of short non-coding RNAs is carried out in yeast by the Nab3-Nrd1-Sen1 complex. Nab3 and Nrd1 are hnRNP-like proteins that dimerize and bind RNA with sequence specificity. We show here that an essential region of Nab3 that is predicted to be prion-like based upon its sequence bias, formed amyloid-like filaments. A similar region from Nrd1 also assembled into filaments *in vitro*. The purified Nab3 domain formed a macroscopic gel whose lattice organization was observed by X-ray fiber diffraction. Filaments were resistant to dissociation in anionic detergent, bound the fluorescent dye thioflavin T, and showed a β-sheet rich structure by circular dichroism spectroscopy, similar to human amyloid-β which served as a reference amyloid. A version of the Nab3 domain with a mutation that impairs its termination function, also formed fibers as observed by electron microscopy. Using a protein fragment interaction assay, the purified Nab3 domain was seen to interact with itself in living yeast. A similar observation was made for full length Nab3. These results suggest that the Nab3 and Nrd1 RNA-binding proteins can attain a complex polymeric form and raise the possibility that this property is important for organizing their functional state during termination. These findings are congruent with recent work showing that RNA binding proteins with low complexity domains form a dynamic subcellular matrix in which RNA metabolism takes place but can also aberrantly yield pathological aggregated particles.

2.2 INTRODUCTION

In yeast, a specialized reaction insures the correct termination of transcription by RNA polymerase II for short, non-coding RNAs [98]. This machinery employs the Nrd1-Nab3-Sen1 complex. The former two proteins are hnRNP-like proteins and the third is a putative helicase [11, 13]. Nrd1 and Nab3 dimerize to bind RNA sequence elements through their RRMs[26]. Using a genetic screen, we previously identified a 134 amino acid carboxy-terminal region in Nab3 ("Nab3^{134}") that is essential for cell viability and important for termination at an intergenic terminator involved in the regulation of *IMD2* [35, 41, 99]. This domain of Nab3 is made up of a low complexity amino acid sequence that is intrinsically unstructured and prion-like based on prediction algorithms and its demonstrated biochemical properties [27, 34]. Nab3's very C-terminal residues also contain a short region of structural homology to human hnRNP-C that is predicted to be α-helical and can crosslink to itself (Fig. 2-1)[38]. This biochemical data, combined with genetic data indicating that two different mutant versions of Nab3 can complement cell viability in *trans*, suggests that multiple copies of this protein work together during termination [27]. This extends a model postulating that multiple copies of Nab3 and Nrd1 operate together on a short nascent transcript bearing consensus sequences, while interacting with RNA polymerase II to provoke termination [26].

Some proteins have regions that are considered intrinsically disordered, meaning the polypeptide backbone populates a dynamic set of positions in space; generally lacking fixed secondary structural elements [100, 101]. Intrinsically disordered domains contain amino acid signatures showing a bias toward repeated sequences that use only a few different amino acids. Low complexity sequences enriched in glutamine, proline, and serine, and poor in aromatic amino acids, have been observed as unstructured domains that can nevertheless interact with each other to form polymeric aggregates including highly ordered amyloid-like structures with a stereospecific packing of β-sheets [102]. In some cases, this feature is associated with prion proteins in which the aggregated conformation can recruit monomers for assembly thereby self-perpetuating the state *in vivo* [75].

A number of proteins in yeast display low complexity, intrinsically unstructured regions. Using a genome wide screen starting with a prediction algorithm, the Lindquist lab showed that many of these aggregate into amyloid-like fibers and some are prionogenic [34]. This set of proteins is enriched for those involved in gene expression with a particular over-representation of RNA binding proteins. A number of mammalian proteins involved in RNA metabolism also possess low complexity regions that can spontaneously form fibers and hydrogels [77, 78]. Filamentation and gel formation are thought to reflect a natural function of these proteins and may be part of their ability to form dynamic cellular compartments in which RNA metabolism takes place [103]. Some of the proteins with low complexity domains that self-assemble, associate with RNA polymerase II to impact transcriptional activation [104]. This property also has consequences for human disease as exemplified by heritable mutations in hnRNPA1 and hnRNPA2B1 that result in amyloid-like deposits integral to the pathology of specific neurological diseases [105].

Here we show that the low complexity, prion-like domain of the termination proteins Nab3 and Nrd1 formed amyloid-like filaments *in vitro*. Nab3 formed a hydrogel and bears many of the hallmarks of amyloid aggregates as gauged by biophysical measures. Both the Nab3 low complexity domain and full-length Nab3, made homotypic interactions with themselves *in vivo*. Taken together with prior genetic and biochemical evidence, the data raise the possibility that these proteins function in a polymeric form during transcription termination.

2.3 MATERIALS AND METHODS

Plasmid construction- A PCR product encoding Nab3 amino acids 668-802 ($Nab3^{134}$) was amplified using primers 5'-atatagatctatctcaaactccaatggaccagc-3' and 5'-atatctcgagctatttttgtagttttgctaaac-3', digested with *Bgl*II and *Xho*I, and inserted into similarly cut pET32a (Novagen) to yield pET32-Nab3-134aa. This yields a fusion protein with thioredoxin followed by a His_6 tag, a thrombin cleavage site, and an S-tag to the amino terminal side of Nab3 sequence. A PCR product encoding Nrd1 amino acids 516-575

was amplified using primers 5'-atatagatctagaaaacctgtattttcagggccagcaatatgtgcaacctat-3' and 5'-atatctcgagttagctttgttgttgttgc-3', digested with BglII and *Xho*I, and inserted into similarly cut pET32a to generate pET32a60aaNrd1. This yields a fusion protein with thioredoxin followed by a His_6 tag, a thrombin cleavage site, the S-tag, and a TEV protease cleavage site to the amino terminal side of Nrd1 sequence. A PCR product encoding $Nab3^{134}$ was made using oligos 5'-ggtctagaactagtaccatgtctcaaactccaatgg-3' and 5'-ccatcgatttttttgtagttttgctaaactatc-3' and digested with *Cla*I and *Xba*I and inserted into similarly cut p416-ADH1-ZIP-DHFR[1,2] and p413-ADH1-ZIP-DHFR[3] which replaces the DNA encoding the bZIP domain of yeast Gcn4 [106]. The resulting plasmids (pBatman and pRobin, respectively) were confirmed by sequencing. Full length Nab3 was inserted similarly into each backbone using PCR primers 5'-atatcgatttttttgtagttttgctaaactatctaatagact-3' and 5'-attctagaatgtcagatgaaaaccataacagtg-3' to yield the plasmids pAvengers and pAssemble, respectively.

Protein purification- $Nrd1^{60}$ and $Nab3^{134}$ were purified from *E. coli* following transformation of plasmids into BL21(DE3) cells. Cultures were grown at 37°C to an optical density of 0.5 and induced with 1mM IPTG. Cells were resuspended in lysis buffer (50 mM Tris, pH 7.5, 500mM NaCl, 10 mM imidazole) containing a protease inhibitor tablet (EasyPack Catalog No. 04693132001 Roche), lysed with 10 □g/ml lysozyme, sonicated, and centrifuged at 27,000xg for 30 min. Supernatants were applied to 1 ml HisTrap HP nickel columns (GE Healthcare) equilibrated in lysis buffer, washed with 25ml of lysis buffer, and eluted with lysis buffer containing 250mM imidazole. Fractions were pooled and digested with 0.5 units thrombin (EMD/Novagen), for 16 hrs at 22°C. Digested proteins were exchanged into lysis buffer by centrifugal filtration in Vivaspin units (3 or 5 kDa molecular weight cutoff, GE Healthcare) and loaded onto HisTrap nickel columns as described above. Flow through was collected and the protein was concentrated by centrifugal filtration while exchanging into 20 mM Tris, pH 7.5, 0.2 M NaCl.

Yeast strains and growth- Yeast strain BY4742 was transformed with pBatman and pRobin to generate DY2214, p416-ADH1-ZIP-DHFR[1,2] and p413-ADH1-ZIP-DHFR[3] to generate DY2315

(positive control), pAvengers and pAssemble to generate DY2217, or p416-ADH1-ZIP-DHFR[1,2] and pAssemble to generate DY2219 (negative control). Where indicated, cells were grown in the presence of 200 µg/ml methotrexate and compared to dimethyl sulfoxide (DMSO vehicle) on solid SC ura^- his^- dropout media.

X-ray fiber diffraction- Nab3 hydrogel was dialyzed against water and harvested using a mounted cryoloop (Hampton Research), inserted in a CrystalCap Magnetic Vial (Hampton Research), and dried in an evacuated desiccator overnight. These loops were inserted into a nitrogen cryostream (100K) and exposed to X-rays for 45 sec from a Rigaku Micromax 007HF X-ray generator with a copper anode running at maximum power with Varimax optics and 0.3 mm collimator. Diffraction images were recorded using a Saturn 944+ CCD detector approximately 100 mm from the specimen.

Semi-denaturing detergent agarose electrophoresis- Purified protein samples were adjusted to 2% SDS, 140 mM β-mercaptoethanol, 10% glycerol, 0.002% bromphenol blue, 80 mM Tris, pH 6.8 and separated by electrophoresis in agarose gels (1.5% w/v) in 40 mM Tris-acetate, pH 7.8; 1 mM EDTA, 0.1% SDS run at 4°C. Bio-Rad Precision Plus Protein Kaleidoscope (Cat. No. 161-0375) molecular weight markers were run as standards. Proteins were blotted to Protran nitrocellulose transfer membrane (Whatman) by capillary action for 18 hrs. Filters were blocked in 5% (w/v) nonfat dry milk in Tris-buffered saline with 0.1% Tween-20 and probed with except. Rabbit anti-S tag antibody (catalog no. PM021; MBL International Corp., Woburn, MA) was used as a primary antibody.

Electron microscopy- Five µl of sample suspension were placed on a 400-mesh carbon coated copper grid that had been made hydrophilic by glow discharge. After 5 minutes, the grid was rinsed by briefly touching the sample side to a drop of distilled water. The residual water was then removed by blotting to filter paper. For negative staining, 5µl 1% aqueous phosphotungstic acid (pH 6.5) was applied to the grid immediately after water removal, and excess liquid was removed by blotting after 30 seconds. The grid was air dried before viewing on a JEOL (Tokyo, Japan) IEM-1400 transmission electron

microscope equipped with a Gatan (Pleasanton CA) 2k x 2k US1000 CCD camera. Images were captured digitally as .tif files.

Aβ peptide synthesis- The amyloid β peptide (1-40: DAEFRHDSGYEVHHQKLVFFAEDVGSNKGAIIGLMVGGVV) was synthesized on a Liberty CEM microwave automated peptide synthesizer as described previously [107] with the following modifications. After ether precipitation, the peptide was dissolved in 40% acetonitrile and 0.1% Formic acid. HPLC was performed with a water-acetonitrile-0.1% formic acid gradient. Trifluoroacetic acid was replaced with formic acid. Guanidinium hydrochloride was not used, and the peptide was purified twice by C18-reverse phase chromatography for improved purity. For filament formation, lyophilized powder was dissolved in minimal hexafluoroisopropanol (HFIP, Aldrich Chemical Co.) for 2 h, dried under argon gas to form a clear film which was dissolved in 10 m*M* phosphate buffer, pH 7.4, 0.02% sodium azide and incubated at 37°C for 25 days.

Thioflavin T binding- Assays were carried out as described by LeVine [108]. Thioflavin T (Sigma Chemical, St. Louis, MO) was dissolved in 50mM glycine-NaOH , pH 8.5 and diluted to 10μM for use. The fluorescence from the dye alone, or the dye mixed with protein (2 μM each, except for BSA which was 20 μM), was read in a Shimadzu RF-5301PC spectrofluorophotometer in spectrum mode, with the excitation filter at 450+/-5 nm and the emission filter at 485+/-5 nm. Triplicate readings were averaged and plotted.

Circular dichroism- Proteins were diluted into 50mm NaF, 17mM sodium phosphate, pH 7.25 and spectra were recorded on a JASCO J-810 Spectropolarimeter from 190nm to 300nm with a 100nm/min scan speed, 0.2nm scan pitch at room temperature with either a 1mm or 0.2mm path length quartz cuvette for amyloid beta or Nab3^{134} respectively. Scans were repeated three times. Raw ellipticity data was transformed for each protein to obtain a mean residue ellipticity value at every 0.2nm of wavelength between 190 and 300nm.

2.4 RESULTS

Purified Nab3 and Nrd1 low complexity domains form polymeric filaments in vitro. Nab3's carboxy-terminal 134 residues ($Nab3^{134}$) show a skewed sequence composition typical of intrinsically unstructured domains containing 32% glutamine (Q), 12% proline (P), and 12% serine (S) (Fig. 2-1). This segment is poor in aromatic residues, lacking phenylalanine and tryptophan and containing only a single tyrosine. These features, and the heat stability and protease sensitivity of the recombinant protein purified from *E. coli*, suggest it is intrinsically unstructured and prion-like [27, 34]. Although unstructured, prion-like domains can change conformation and polymerize to form ordered aggregates *in vitro* and *in vivo* [109]. Although essential, the function of this portion of the protein has remained obscure. The properties of Nab3's carboxy-terminal Q/P rich domain suggested it could form amyloid in a manner similar to yeast prions and aggregates containing Nab3 have been observed microscopically in yeast cells when expressed as a GFP-fusion [34]. We previously showed that this region of Nab3, which is distant from the RNA binding portion of the protein (Fig. 2-1), can be robustly expressed in a soluble form in *E. coli* as a fusion to the carboxy-terminus of His_6-tagged *E. coli* thioredoxin. Following affinity chromatography, the piece liberated by thrombin cleavage yielded the $Nab3^{134}$ polypeptide with an S-tag at its amino terminus [27] (Fig. 2-1). Purification and incubation of this piece of Nab3 at 4°C for a number of days at high concentration yielded a translucent macroscopic gel (Fig. 2-2). Our ability to detect polymerization may be due to the choice of boundaries of the domain, the absence of denaturants during the purification from *E. coli*, and/or the fusion of the domain to a distinct protein tag, as compared to a prior study [34]. In any case, the protein polymerizes over a period of hours to days and the rate of formation of fibers increased with protein concentration, incubation at cold temperatures, and with agitation. The dimensions of the fibers seem to be characteristic of each protein. Whether this is an intrinsic feature of their polymerization geometry or a physical property such as their brittleness or flexibility during manipulation, remains to be determined. In any case, the assembled state appears essentially irreversible.

To learn if the Nab3^{134} gel was composed of amyloid-like fibers, we subjected it to X-ray fiber diffraction as described for the low complexity domains of FUS and hnRNPA2, for example [77]. The pattern for Nab3^{134} revealed diffractions centered at 4.7 and 8.9 Å indicative of periodicity in the gel as seen for these two mammalian RNA binding proteins (Fig. 2-2). Examination of a sample of the hydrogel by negative staining and transmission electron microscopy, showed clearly that the gel is composed of filaments averaging 75 nm in length (Fig. 2-3A).

Many amyloid fibers are resistant to the ionic detergent sodium dodecyl sulfate (SDS) (24). In contrast, amorphous, non-amyloid precipitates are solubilized by SDS (25,26). To examine Nab3134 aggregates for this property, we subjected them to semi-denaturing detergent agarose gel electrophoresis (SDD-AGE) which separates monomers from large multimers following their exposure to SDS. Proteins were transferred to nitrocellulose and probed with antibodies against the S-tag epitope (KETAAAKFERQHMDSSTSAA). The Nab3^{134} gel was highly SDS-resistant, running as a complex high molecular weight species as observed for a number of yeast prion-like domains [34]. In contrast, the control protein *E. coli* thioredoxin, runs as a single species, indicating this is not a property of all proteins fused to the S-tag (Fig. 2-4). Boiling in SDS resulted in a subtle decrease in the overall size of aggregated Nab3 but it was largely resistant to this treatment as well (Fig. 2-4). This observation is consistent with the idea that the Nab3 low complexity domain can assume an amyloid form (Fig. 2-4).

Like Nab3, Nrd1 is an essential protein that possesses an RRM and is important for termination of small non-coding RNAs. It dimerizes with Nab3 and also possesses a low complexity domain [32% Q, 17% P, 10% alanine (A)], that is predicted to be prion-like and intrinsically unstructured [34, 100] (Fig. 2-1). However, direct evidence supporting this has not yet been reported. Like the Nab3 domain, the carboxy-terminal 60 residues of Nrd1 could be expressed in *E. coli* and obtained in high yield as a thioredoxin-fusion. Soluble protein with an S-tag was purified following thrombin cleavage (Fig. 2-1). When incubated at 22° or 4°C, the domain formed filaments observable by electron microscopy (Fig. 2-3B), although they

did not form a translucent hydrogel as Nab3^{134} did. Measurements indicated the Nrd1^{60} filaments were on average 590 nm in length. Unexpectedly, monomeric Nrd1^{60} could not be detected by SDD-AGE even though its polymeric form was readily detectable (Fig. 2-4, middle panel). Presumably, it binds poorly to filters (either nitrocellulose or polyvinylidene fluoride membranes). To aid in the detection of Nrd1^{60}-monomers, we studied the behavior of the purified Nrd1^{60}-thioredoxin fusion protein without separating the two by digestion with thrombin. In this case, both monomeric and polymerized forms were detectable and boiling in SDS reduced, but did not completely dissolve, the polymer (Fig. 2-4, right panels). Brief centrifugation of the high molecular weight Nrd1^{60} yielded a more discrete and smaller polymeric fraction in which integral units of polymer could be seen in a ladder-like pattern (Fig. 2-4, right panel). Neither the Nab3 nor Nrd1 polymers could be dissolved by dilution, warming, or 8M urea, although they were readily soluble in dimethyl sulfoxide, as described for human amyloid-β [110]. We conclude that the low complexity domains of these two proteins form amyloid-like polymers *in vitro*.

The Nab3 polymer bears physical hallmarks of amyloid- A characteristic of the β-sheet rich structure of amyloid polymers is their ability to bind thioflavin T which fluoresces with a characteristic emission maximum at 482 nm upon binding to the polymerized protein [111]. To test if Nab3^{134} filaments possessed this property, we incubated the polymer with thioflavin T and scored for fluorescence using spectrofluorometry. We compared Nab3^{134} to a positive control, polymeric human amyloid-β peptide extending from residues 1-40 (Aβ). Nab3^{134} demonstrated a strong response from thioflavin T with a profile comparable to that seen for Aβ and distinct from the lack of fluorescence when ten-fold more bovine serum albumin was mixed with thioflavin T or the dye was excited in the absence of any protein (Fig. 2-5A). Titration of Nab3^{134} showed that the fluorescent signal was linear with respect to protein concentration (Fig. 2-5B).

The structure of the canonical amyloid, human amyloid β has been examined by circular dichroism spectroscopy in which its β sheet content can be observed [112]. We found that the far ultraviolet spectrum

for Nab3^{134} fibers was characteristic of a β-rich structure with a profile similar to the Aβ control, consistent with the model that Nab3^{134} attains an amyloid form (compare Fig. 2-6A and 6B).

A termination defective mutant of Nab3 forms filaments. Mutations in the low complexity domain of Nab3 have been identified that compromise its ability to support termination of short, non-coding transcripts [27, 35, 38]. One of these is a Leu to Ala substitution at position 800 (Fig. 2-1) that shows a strong terminator override phenotype [38]. A recombinant version of Nab3^{134} bearing this substitution was expressed, cleaved off thioredoxin, and purified. The protein readily formed a semi-solid gel and more interestingly, presented as a distinct filamentous form when negative stained and analyzed by electron microscopy (Fig. 2-3A). The filaments were many microns in length and their extreme size made it difficult to measure them precisely. They also tended to form side-by-side bundles in this preparation. When examined by SDD-AGE, the Nab3$^{134\text{-}L800A}$ protein demonstrated resistance to boiling in SDS with no monomeric form detectable (Fig. 2-4).

In vivo analysis of the Nab3 low-complexity domains. As a further test of Nab3's ability to interact with itself *in vivo*, we employed the protein fragment complementation assay developed by Michnick and co-workers [106, 113]. In this assay, a methotrexate resistant dihydrofolate reductase (DHFR) enzyme is recombinantly split into an amino and carboxy terminal portion each of which is fused to candidate proteins. If the candidate proteins interact, the pieces of DHFR are brought together generating active enzyme that renders the growth of yeast resistant to methotrexate. As a positive control, the bZIP domain of the yeast transcription factor GCN4 was shown to interact with itself when fused to both pieces of DHFR (DHFR$^{N\text{-}term}$ and DHFR$^{C\text{-}term}$), as seen previously [106] (Fig. 2-7, top quadrants). We adapted this assay to test if Nab3^{134} interacts with itself by fusing it independently to both the amino- and carboxy-terminal fragments of DHFR and introducing both plasmids into yeast. Growth of this strain in the presence of methotrexate indicated that the Nab3^{134}-DHFR$^{N\text{-}term}$ and Nab3^{134}-DHFR$^{C\text{-}term}$ fusion proteins interact *in vivo* (Fig. 2-7, right quadrants). Pairing a bZIP-DHFR plasmid with a Nab3^{134}-DHFR plasmid served as a

negative control showing that the mere coexistence of both halves of DHFR in the cell was insufficient to provide drug resistance unless they were brought together by fused binding partners. Finally, we extended this approach to full length Nab3 by fusing the full 802 amino acids of this protein to the $DHFR^{N\text{-}term}$ and $DHFR^{C\text{-}term}$ fragments and introducing both plasmids into cells. Again, an interaction was detectable as seen by methotrexate resistance (Fig. 2-7, bottom quadrant). Although this assay does not demonstrate filament formation, it does provide additional independent evidence that Nab3 interacts with itself in living yeast.

2.5 DISCUSSION

In this report we provide new evidence that the low complexity domains of two yeast hnRNP-like proteins, Nab3 and Nrd1, can polymerize into a fibrillar form. Nab3 readily assembles into filamentous structures which organize into a semi-solid gel *in vitro*. The evidence that the low complexity domain of Nab3 forms amyloid-like fibers includes: 1) its ability to form stable, regular filaments, 2) its ability to assemble into a high molecular weight form resistant to disruption by the anionic detergent SDS that is separable from unassembled monomer, similar to yeast prions, 3) its propensity to spontaneously gel and show a periodic organization of fibers as seen by X-ray fiber diffraction that is similar to that seen for other RNA binding proteins with low complexity sequence, 4) its capacity to bind the amyloid-specific fluorescent dye, thioflavin T in a manner indistinguishable from amyloid-β, 5) its circular dichroism properties which closely parallel those of the well-studied amyloid-β fibrils. Together, we take this as compelling evidence that this domain of the yeast RNA binding protein and termination factor Nab3 attains an authentic amyloid form.

The obvious question is what is the biological role of this feature? The following findings make it likely that filament formation is important for Nab3's function as an RNA-binding protein in transcription termination: 1) prior genetic analysis has supported a model that the assembly of multiple copies of Nab3

onto transcripts in living cells is needed for cells to survive [27], 2) evidence presented here using an independent assay indicates that the Nab3 low complexity domain interacts with itself in living cells, as does intact Nab3, 3) the region of Nab3 that assembles into filaments is essential for cell viability [27], the small size of the region limits the number of functions it can contain, increasing the likelihood that its essential function is related to its polymerization potential, and finally, 4) recent findings across species have revealed that RNA-binding proteins, many of which are hnRNPs and many of which contain low complexity regions, undergo reversible associations important for their functions and yield filaments and semi-solid gels *in vitro* [77, 78]. As well, amyloid formation by prion-like proteins is becoming recognized as a functionally important state and it forms the basis of regulatory switches in diverse organisms. Recent examples include neuronal CPEB [114], yeast Mot3 [115], and yeast Mod5 [116]. Based upon these studies, we infer that the assembly of Nab3 we observe *in vitro*, represents the potential for this protein to reversibly cycle between protomers and assembled forms *in vivo*. The biological signals that induce and dissolve these structures remain to be identified. One good candidate, however, is phosphorylation of Nab3 and Nrd1 both of which have been shown to be phospho-proteins [25, 117, 118]. Nrd1's phosphorylation state has been shown to change in conjunction with the availability of a carbon source [96]. Interestingly, dephosphorylation during glucose starvation is concurrent with rearrangement of Nrd1 into nuclear puncta. This process could be related to the assembly potential of Nab3 and Nrd1 observed here. Furthermore, the low complexity domains of other transcriptionally important proteins, such as Taf15, form filaments and are recruited to the carboxy terminal repeat domain of the large subunit of RNA polymerase II [104]. This part of pol II is an intrinsically disordered, simple repeated sequence, and it contains a signature [G/S]Y[G/S] amino acid motif shown to be important for polymerization of some low complexity domains. Filaments made from the FUS and Taf15 domains associate with the pol II repeat in its hypo-phosphorylated form, but not when it is phosphorylated [104]. While these low complexity proteins are transcriptional activators, we propose that a similar situation may exist for termination-related proteins such as Nab3 and Nrd1 which may polymerize, perhaps reversibly, while transiently interacting with pol II.

Association of filamentous Nrd1 or Nab3 with the CTD through their low complexity domain would be a second means of interaction with the CTD in addition to the well-characterized CTD-interaction domain of Nrd1 [40].

A feature of Nab3 and Nrd1 thought to be important for their function is the ability to co-assemble into a stable heterodimer [25]. A careful study of dimerization was done with truncated recombinant Nab3 and Nrd1 derivatives that lacked the self-assembly domains described here. Thus, they can form a heterodimer without these low complexity regions (28). Are heterodimerization and the self-assembly property observed here two-different associations, or can they co-exist? One possibility is that Nrd1-Nab3 heterodimerization is mutually exclusive with the filament assembly ability of each protein. Alternatively, the heterodimer may be a protomeric unit of polymerization, either constitutively or inducibly, which would result in a Nrd1-Nab3 heteropolymer. Further probing of this possibility must await the reconstitution of significant quantities of recombinant full-length Nrd1 and Nab3 that contain both their heterodimerization regions and their low complexity self-assembly domains. Yet a third possibility is that there are distinct pools of Nrd1-Nab3 dimers independent of a filamentous form of each protein.

The findings of filament formation for the low complexity domains of these RNA-binding proteins are in accordance with comparable data obtained for mammalian hnRNPs and other RNA-binding proteins, such as FUS and TDP-43, that have low complexity domains [77, 78]. A growing body of evidence suggests that RNA-binding proteins can assume a polymerized state which manifests itself as a hydrogel *in vitro*, and whose *in vivo* form is a subcellular, non-membranous compartment such as a P-body or stress granule [103]. An emerging concept is that formation of amyloid-like fibers and hydrogels by RNA-interacting proteins reflects their capacity to dynamically and reversibly assemble and disassemble *in vivo* to form localized complexes that handle, metabolize, and modify RNA. The abnormal persistence of such aggregates can lead to a class of diseases known as ‘proteinopathies’, some of which are due to mutations in hnRNPs that potentiate their filamentation [105, 119, 120]. We propose that the assembly of Nrd1 and

Nab3 seen *in vitro*, and that of Nab3 seen *in vivo*, reflects a role for self-assembly in their function during termination of transcription in the nucleus. This would be consistent with prior genetic data indicating that multiple copies of Nab3 are important for function [27], and accumulating evidence that prion-like assemblages are functionally significant [114, 121, 122].

Acknowledgments

The authors thank Drs. Graeme Conn, Tanya Chernova, Yury Chernoff, Richard Cummings, and Xiaodong Cheng for helpful discussions and sharing of reagents and equipment, and Alissandra K. Mowles, Noel X. Li, and Dr. David Lynn for advice on CD and fibrillized Aβ. We also thank Dr. Stephen Michnick for materials and advice. The authors also acknowledge the assistance of Hong Yi in the Emory Electron Microscopy Core. This work was supported, in whole or in part, by National Institutes of Health, NIGMS, Grant R01GM46331, NIH training grant T32GM008367, and NIH grant S10RR025679.

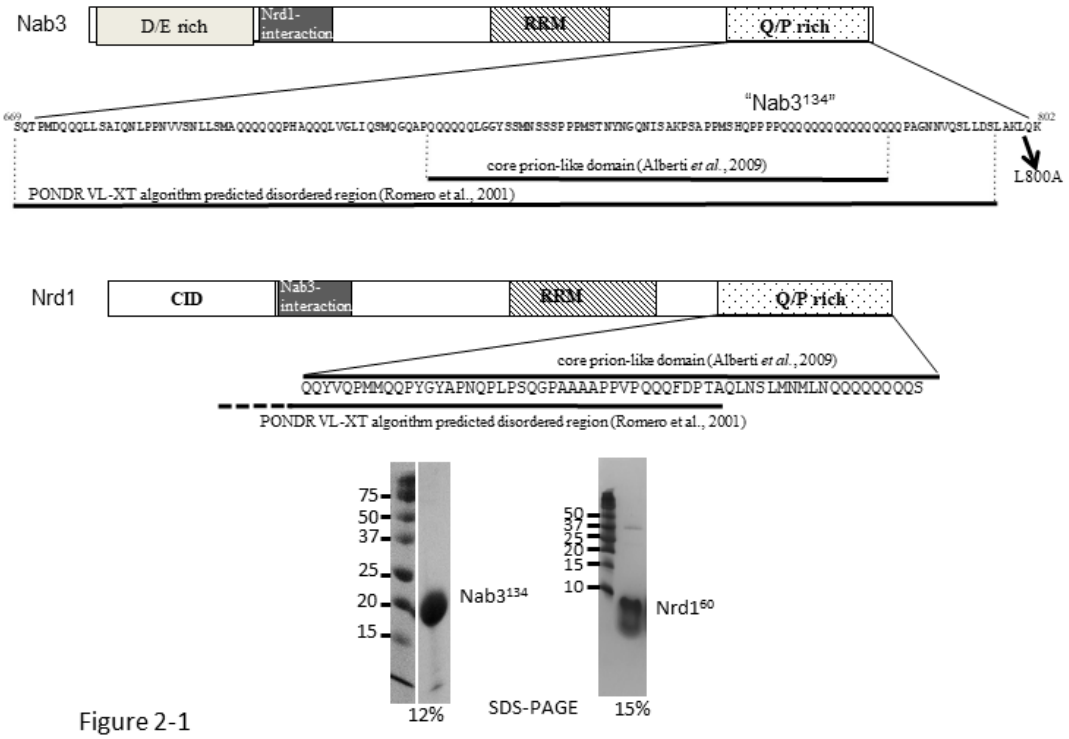
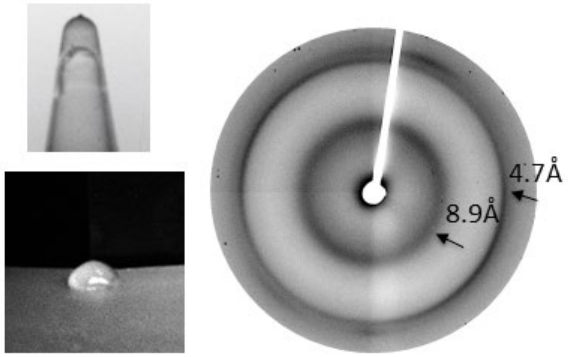


Figure 2-1

Figure 2-1- Schematic of the yeast Nab3 and Nrd1 proteins. A schematic with an expanded view of the low complexity region studied here (not to scale). For Nab3 the Asp-Glu (D/E) rich domain, the Nrd1-interaction domain, the RRM domain, and the Q/P rich domains are indicated with patterned boxes. The computationally predicted core prion-like domain and disordered region are indicated on the expanded sequence. The L800A mutation is shown with an arrow. The Nrd1 schematic shows its CTD-interaction domain, Nab3-interaction domain, RRM, and Q/P rich region with patterned boxes. The expanded sequence shows that C-terminal 60 amino acids studied here and the portions of it that score as prion-like and intrinsically unstructured based on computer algorithms. The purity of purified $Nab3^{134}$ and $Nrd1^{60}$ based upon SDS-PAGE analysis is shown with the acrylamide percentages shown below each image. Both lanes of the $Nab3^{134}$

gel were run on two ends of the same slab. A trace of added thrombin is evident at 36kD in the Nrd1^{60} lane.

Figure 2-2



Nab3[134]

Figure 2-2- Gel formation and X-ray fiber diffraction analysis of Nab3[134]. Photographs of a hydrogel formed from Nab3[134] in an upside-down tube and extruded onto a surface are shown at left. X-ray diffraction of dried fibers of Nab3[134] is shown at right.

Fig 2-3

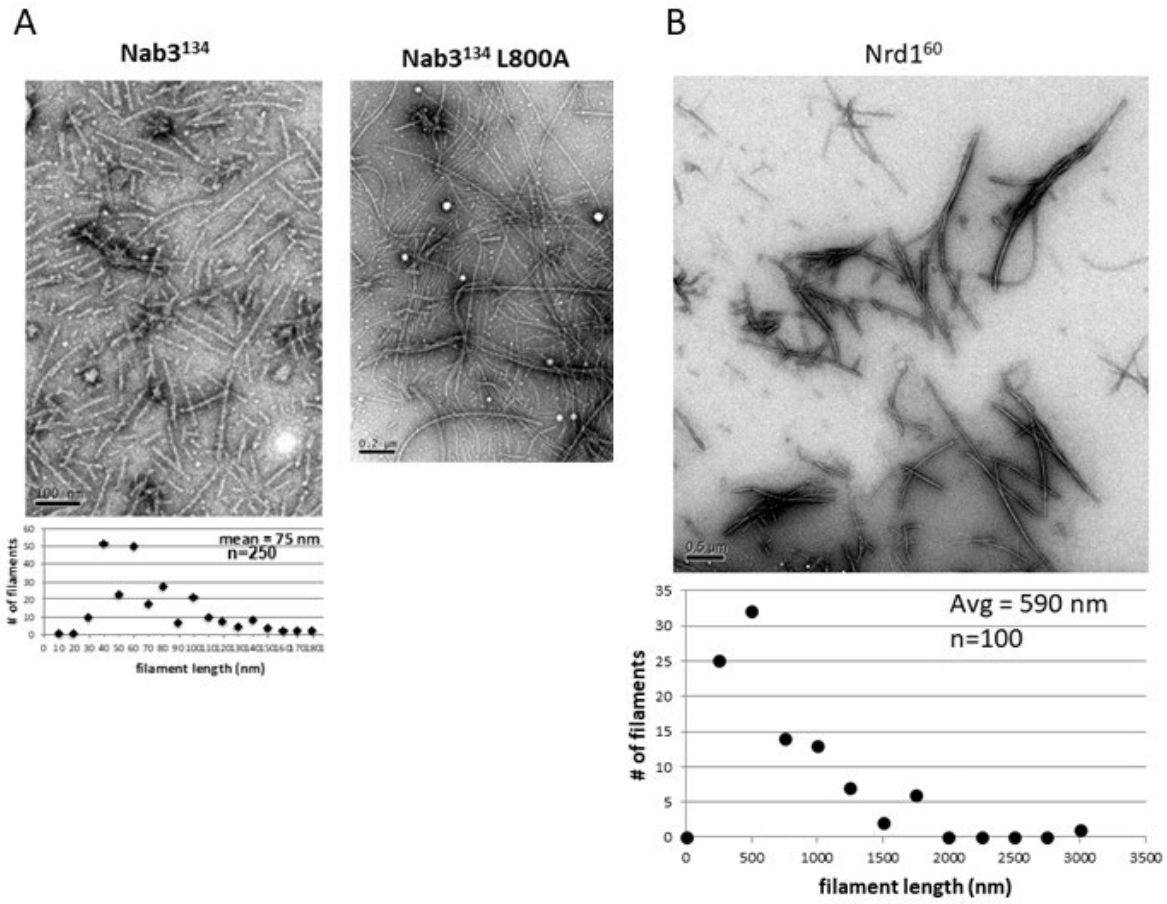


Figure 2-3-Amyoid-like fiber formation by Nab3's and Nrd1's low complexity domains. Transmission electron microscopy on filaments formed from (A) purified Nab3^{134} and its L800A derivative or purified Nrd1^{60} (B). Scale bars are shown at the lower left of each image. A quantification of the size distribution of Nab3^{134} and Nrd1^{60} are plotted below their images.

Figure 2-4

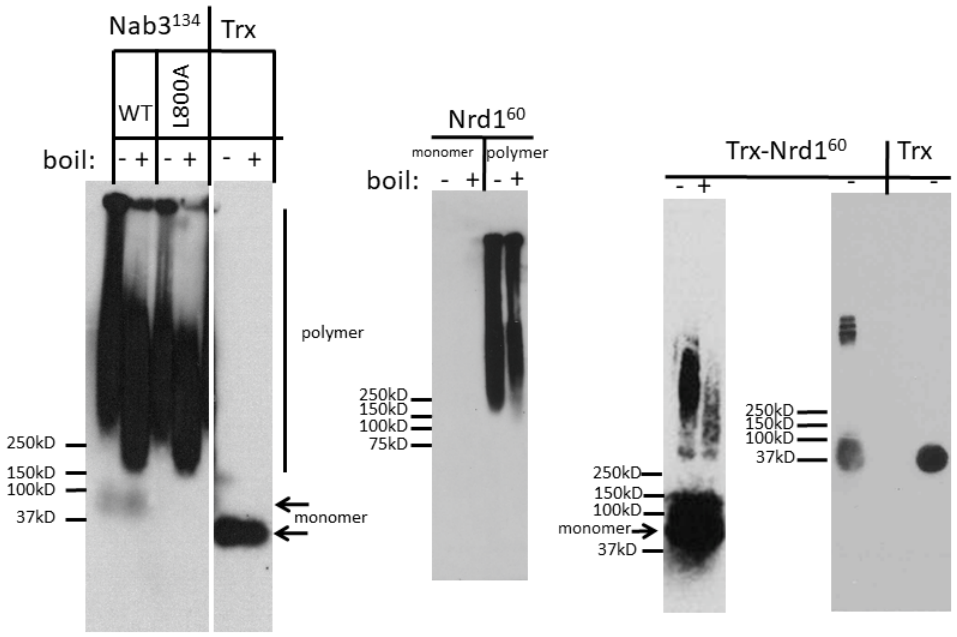


Figure 2-4- Semi-denaturing agarose gel electrophoresis on polymers formed *in vitro.* Purified Nab3^{134}, the L800A mutant derivative of Nab3^{134}, purified Nrd1^{60} (unpolymerized monomer and polymerized), thioredoxin-Nrd1^{60} fusion, and thioredoxin (TRX) were incubated with SDS at 22° (-) or 100°C (+), subjected to electrophoresis, and transferred to nitrocellulose. Membranes were probed with anti-S-tag antibody and developed with chemiluminescent reagents. The positions of monomeric (arrowheads) and polymeric (vertical bar) versions of the proteins are indicated, as are the size (kDa) and migration positions of Precision Plus Protein Kaleidoscope molecular weight standards (Bio-Rad, Inc.). TRX and Nab3^{134} samples from the composite panel on the left were run on different ends of the same gel. Note that Nrd1^{60} monomers were not detectable but the monomeric TRX-Nrd1^{60} fusion protein was.

Figure 2-5

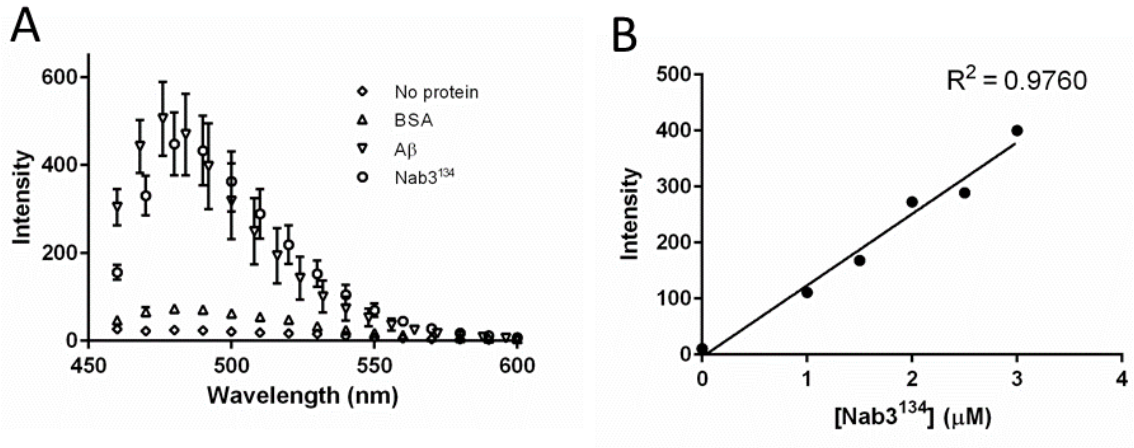


Figure 2-5- Fluorescence spectroscopy of thioflavin T binding to Nab3^{134} and Aβ. A) Thioflavin T was mixed with the indicated proteins (2 μmolar each, except BSA which was 20 μM) and the spectra of fluorescence was scored in a spectrofluorometer in triplicate and averaged. Error bars show the standard deviation around those averages. B) Increasing amounts of Nab3^{134} protein was added to thioflavin and absorbance at 485 nm was quantified and plotted. The correlation coefficient (R^2) was calculated with Prism 6.02 software (GraphPad).

Figure 2-6

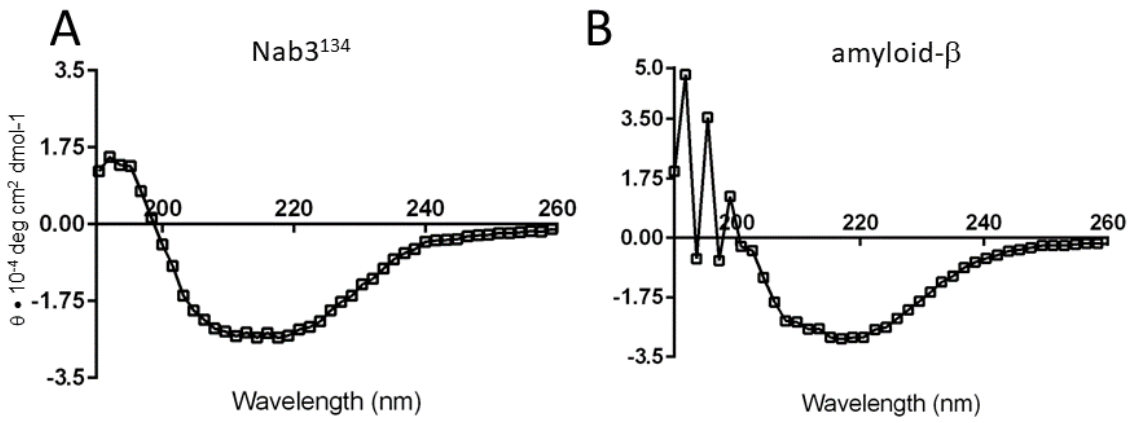


Figure 2-6- Circular dichroism spectroscopy of Nab3^{134} and Aβ. Suspensions of Nab3^{134} and Aβ (1-40) were independently analyzed by CD spectroscopy as described in Materials and Methods. A representative spectrum is shown for each. θ was calculated and plotted at 1.6 nm intervals. Lines connect the average values. Nab3^{134} was read in a 0.2mm path length cuvette and Aβ was read in a 1 cm path length cuvette.

Figure 2-7

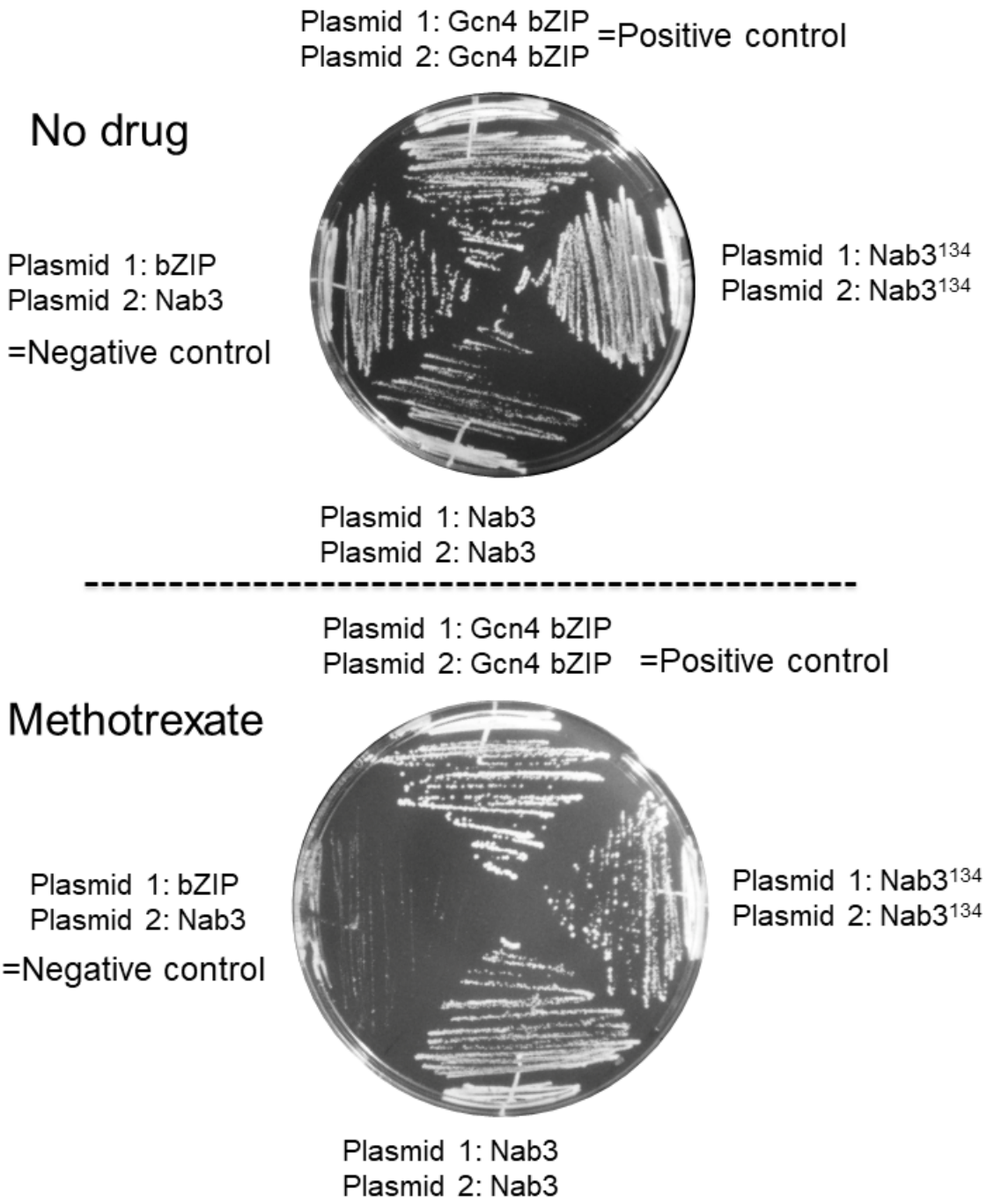

Plasmid 1: Gcn4 bZIP
Plasmid 2: Gcn4 bZIP
=Positive control
No drug
Plasmid 1: bZIP
Plasmid 2: Nab3
=Negative control
Plasmid 1: Nab3134
Plasmid 2: Nab3134
Plasmid 1: Nab3
Plasmid 2: Nab3
Plasmid 1: Gcn4 bZIP
Plasmid 2: Gcn4 bZIP
=Positive control
Methotrexate
Plasmid 1: bZIP
Plasmid 2: Nab3
=Negative control
Plasmid 1: Nab3134
Plasmid 2: Nab3134
Plasmid 1: Nab3
Plasmid 2: Nab3

Figure 2-7- Protein-Protein interaction *in vivo* by Nab3^{134}. Yeast strains containing the indicated plasmids were struck to solid media containing 200 μg/ml methotrexate or its solvent, DMSO, as indicated, and were grown at 30°. The negative control strain was DY2219, the positive control strain was DY2315, the strain with two Nab3^{134}-fusion proteins was DY2214, and that with two full-length Nab3-fusion proteins was DY2217.

Chapter 3

The hnRNP-like Nab3 termination factor can employ heterologous prion-like domains in place of its own essential low complexity domain

Thomas W. O'Rourke contributed to 3-2 and 3-5. Dr. Daniel Reines was responsible for 3-4 and 3-S1, and Dr. Natalya Degtyareva was responsible for 3-S2

This chapter was derived from the published manuscript:

The hnRNP-like Nab3 termination factor can employ heterologous prion-like domains in place of its own essential low complexity domain

Travis J. Loya, Thomas W. O'Rourke, Daniel Reines

Published: October 12, 2017https://doi.org/10.1371/journal.pone.018618

3.1 ABSTRACT

Many RNA-binding proteins possess domains with a biased amino acid content. A common property of these low complexity domains is that they can assemble into an ordered amyloid form, juxtaposing RNA recognition motifs in a subcellular compartment in which RNA metabolism is focused. Yeast Nab3 is one such protein that contains RNA-binding domains and a low complexity, glutamine/proline rich, prion-like domain that can self-assemble. Nab3 also contains a region of structural homology to hnRNP-C that resembles a leucine zipper which can oligomerize. Here we show that the LCD and the hnRNP-C homology domains of Nab3 were experimentally separable, as cells were viable with either segment, but not when both were missing. In exploiting the lethality of deleting these regions of Nab3, we were able to test if heterologous prion-like domains known to assemble into amyloid, could substitute for the native sequence. Those from the hnRNP-like protein Hrp1, the canonical prion Sup35, or the epsin-related protein Ent2, could rescue viability and enable the new Nab3 chimeric protein to support transcription termination. Other low complexity domains from RNA-binding, termination-related proteins or a yeast prion, could not. As well, an unbiased genetic selection revealed a new protein sequence that could rescue the loss of Nab3's essential domain. This new sequence and Sup35's prion domain could also rescue the lethal loss of Hrp1's prion-like domain when substituted for it. This suggests there are different cross-functional classes of amyloid-forming LCDs but that merely appending any assembly-competent LCD to Nab3 does not restore function or rescue viability. The analysis has revealed the functional complexity of LCDs and provides a means by which the differing classes of LCD can be dissected and understood.

3.2 INTRODUCTION

In *Saccharomyces cerevisiae*, a specialized termination machinery ensures the proper termination of short, non-coding transcripts by RNA polymerase II (pol II) [67]. This complex depends upon the NNS complex [123]. Nrd1 and Nab3 are hnRNP-like proteins that dimerize to bind specific RNA sequence elements *via* their RNA recognition motifs (RRMs) and associate with pol II through Nrd1's C-terminal interaction domain (CID) [11, 26]. Sen1, a putative helicase, is thought to join this ribonucleoprotein complex during pol II elongation, thereby terminating transcription [69]. Prior work in our lab established that a C-terminal segment of Nab3 is essential for viability, and is important for proper termination in both a reporter assay as well as at an endogenous target of NNS, the *IMD2* gene [35]. This Nab3 sequence is made up of a low complexity amino acid sequence that is rich in glutamine (Q) and proline (P), intrinsically disordered, and amyloid-like based on computational models and biochemical data. Combinations of electron microscopy, Thioflavin T binding, circular dichroism and semi-denaturing detergent agarose electrophoresis (SDD-AGE) have shown Nab3 and the low complexity domain of Nab3 form amyloid. Furthermore, Q->E mutations in the Nab3 LCD abrogate amyloid formation[36]. There is a noted overrepresentation of LCDs in RNA-binding proteins which is thought to reflect their functional importance [77, 78]. A previously characterized genetic interaction has shown that different mutants of *NAB3* can complement each other in *trans* to restore viability, suggesting multiple copies of Nab3 are present at termination complexes. Nab3 was also shown to interact with itself in living yeast in a protein fragment complementation assay [36, 106]. As well, multiple Nrd1 and Nab3 recognition sites are found in transcripts terminated by the NNS system, again

suggesting multiple copies of the proteins work together *in vivo* [29]. These findings suggest that the LCD contributes to the functionally important assembly of the Nab3 termination factor with itself or other partners. Through its role in RNA biogenesis, the NNS complex is a key regulator of transcriptome changes seen during nutrient stress, showing increased recruitment of Nab3 to genes involved in energy usage and cell growth as well as a decrease in binding to genes regulating stress response [96, 124, 125]. Concomitant with this shift is the movement of Nab3 and Nrd1 to a perinuclear subcellular structure [96]. How this particle assembles, dissociates, and functions in RNA metabolism is poorly understood.

Low complexity domains contain stretches of amino acids with limited sequence diversity that are often intrinsically disordered. However, under some conditions, LCDs can attain very stable ordered secondary structures, often β-sheets, which assemble into intracellular amyloid aggregates. LCDs in prions are important for their ability to aggregate and become heritable particles. The biology of prion proteins is under active investigation and domains of low complexity which often appear prion-like in their composition have received considerable interest due to their implication in numerous disease states [105, 126], their frequent occurrence in RNA-binding proteins [77, 78], and their participation in the formation of non-membrane bound organelles in both the nucleus and cytoplasm [77, 90, 127, 128]. The hallmark of many neurodegenerative disorders is the formation of insoluble protein aggregates in neurons, many of which contain proteins with low complexity domains [129].

Here we have analyzed the functional role of the prion-like LCD of yeast Nab3, which can assemble into a robust amyloid form, and is essential for cell viability. (For

convenience, we use the terms low complexity and prion-like domain interchangeably in this report.) We show that the Nab3 LCD can be exchanged for the low complexity regions of other proteins including those from a bona fide prion, Sup35, or another RNA-binding protein, Hrp1. Although they substitute functionally, these domains lack strong primary sequence homology to the natural Nab3 LCD they replace. Furthermore, amino acid content does not necessarily correlate with their effectiveness. Using the essential nature of the LCD as a genetic tool, we also selected from a pool of bovine DNA, sequences that can functionally replace the natural domain. Taken together, these findings indicate a remarkable degree of versatility between LCDs, however, not all LCDs are capable of complementing function showing that there is specificity within the larger family of low complexity regions. We also show that Nab3's C-terminus is more complex than previously appreciated as it can be subdivided into two regions, a Q/P-rich segment and a small piece with structural homology to an α-helix found in human hnRNP-C. The essential nature of Nab3 and its C-terminal region have provided a genetic means of probing the role of low complexity domains in RNA-binding proteins.

3.3 MATERIALS AND METHODS

Plasmid construction

Nab3 chimeras were generated by cutting pRS315-Nab3 [27] with *Nde*I and *Xho*I and inserting similarly cut PCR products encoding heterologous LCD's amplified with the primers listed in Table 1. The Nab3-Ent2 chimera was made by inserting synthetic DNA (GenScript, Piscataway, NJ) encoding the 60 amino acid core prion-like domain [34] into

the *Nde*I and *Xho*I sites of pRS315-Nab3. The pRS315-Nab3hrp1scr plasmid was made by inserting synthetic DNA (GenScript) encoding a randomization of the 60 amino acid core prion-like domain of Hrp1 (YGNGMMGQTGNYGDGGMNFNNDNGQQNYMNGPEQQGYMGMNYMRPMQN QFYANQKMNPRNQY) into *Nde*I and *Xho*I-cut pRS315-Nab3. pRS315Nab3Δ134α was made by deleting DNA encoding Nab3 amino acids 669-784 in pRS315Nab3 using 5'-aggtggttgaggaggcggacc-3' and 5'-cctgctggcaataatgttcaaagtctatta-3'. pRS315Nab3Δ134αL800A was made from pRS315Nab3Δ134α by Phusion mutagenesis using 5'-gatagtttagcaaaagcacaaaaatagactc-3' and 5'-taatagactttgaacattattgccagc-3'. pRS315-Hrp1 was made by inserting a PCR product encoding *HRP1* (5'-ggatcctaggattgaacagtttcgac-3' and 5'-ctcgagtacagctacttcaagtagag-3') into pRS315 after both were cut with *Bam*HI and *Xho*I. pRS315-HRP1coredel was made by Phusion deletion mutagenesis of pRS315-HRP1 using 5'-aagcaaactggtatggattatactc-3' and 5'-tgatgatttttgttgcatatgtc-3'. pRS315-Hrp1Sup35 and pRS315-Hrp1CT25 were made by inserting synthetic DNA (GenScript) encoding either the Sup35 N-domain residues 2-114, or amino acids VCTCLCMSLCLKKGVKDTRQVVDDHLEG, respectively, in place of the 60 amino acid core prion-like domain of Hrp1 (residues 329-388).

A library of plasmids with bovine genomic DNA fused to the Nab3 sequence was made by digesting pRS315-Nab3 with *Nde*I and *Xho*I. Calf thymus DNA (Worthington) digested with *Nde*I and *Xho*I was gel-purified, ligated to the plasmid DNA and transformed into *E. coli* that was plated on solid medium. Colonies were pooled, and plasmid DNA was prepared and used to transform yeast strain DY3111 (which contains a deletion of

chromosomal *NAB3* and a plasmid bearing wildtype *NAB3* on a *URA3*-marked plasmid) to leu^+. Transformants were selected on SC ura^- leu^- and moved to SC leu^- to allow for loss of the *URA3*-marked plasmid with wildtype *NAB3*. Cells were plated on FOA-containing medium and colonies were re-streaked to FOA for purification. The resulting plasmid was rescued from yeast and sequenced. A number of parental plasmids containing full length *NAB3* were identified and excluded. Bovine DNA from clone CT25 was mapped as an exact match to chromosome 19 from 25435293-25436099 on the Btau4.6.1 assembly at www.bovinegenome.org using BLAST [130]

pRS306-nab3Δ134 was made by inserting the nab3Δ134 sequence released from pRS315-nab3Δ134 [27] by digestion with *Bam*HI and *Xho*I into similarly cut pRS306 [131].

Yeast strains

Yeast strains used in this study are presented in Table 2. Cells were grown in rich medium (YPD), synthetic medium (SC), or standard selective drop-out medium (SC ura^-, SC leu^- or SC ura^- leu^-; [132]) at 30°C unless otherwise indicated. Plasmids were transformed into yeast strains using the lithium acetate method [133]. DY3111 was transformed with pRS315-Nab3Rat1, pRS315-Nab3Pcf11, pRS315-Nab3Rnq1, pRS315-Nab3Sup35, pRS315-Nab3Hrp1, pRS315-Nab3-CT25, pRS315-Nab3Δ134, pRS315-Nab3Δ134α, pRS315-Nab3Δ134αL800A to generate strains DY3184, DY3185, DY3186, DY4002, DY3193, DY4001, DY3134, DY353, and DY377, respectively. The resulting strains were grown at 30˚C on SC leu^- ura^-. For plasmid shuffling, these strains were grown

on SC leu$^-$ followed by selection on FOA. GFP reporter plasmids were transformed into DY3182, DY3196, DY359, DY379, to yield strains DY3187 and DY3204, DY3197, DY3912, DY3913, respectively.

For tetrad analysis, the diploid YH990 [134] served as a starting strain for the 'pop-in/pop-out' [132] introduction of *nab3Δ134* into one of the two *NAB3* loci, generating strain DY3210. Conversion was confirmed by PCR for the two *NAB3* alleles. The diploids were sporulated and dissected.

The strain PSY818 (a gift from Dr. P. Silver, [135]), disrupted for *HRP1* contains the essential *HRP1* gene on a *LEU2*-based plasmid. The *LEU2*-based plasmid was transformed with a *URA3* based plasmid by transformation with pRS316-HRP1 to yield DY307 and screened for the loss of the *LEU2*-marked plasmid, to yield strain DY308. This strain was used as a parent to transform pRS315-Hrp1coredel, pRS315-Hrp1coredel-CT25, pRS315-HRP1, pRS315-Hrp1coredel, pRS315, pRS315-Hrp1Sup35LCD to yield strains DY318, DY387, DY388, DY389, DY1638, and DY3242, respectively. Loss of the *URA3*-marked copy of *HRP1* from these strains *via* shuffling yielded the cognate viable strains DY387F and DY3243.

Flow cytometry

Cells were grown in SC ura$^-$ leu$^-$ in which glucose was replaced by galactose as the carbon source and grown overnight. Overnight cultures were diluted to an OD_{600} of 0.2 and allowed to grow to OD_{600} of approximately 1.0. Flow cytometry was performed using

an LSRII (Becton-Dickson) with an excitation wavelength of 488nm and data was collected using FACSDiva software (Becton-Dickson). Analysis of results was performed using FlowJo ver. 10.1.

Western blotting

Cells were collected, washed with water, and boiled for 5 min in loading buffer [136] before being resolved on a 6% polyacrylamide gel. Proteins were electrophoretically transferred to nitrocellulose which was blocked with 5% nonfat dry milk [137], and probed with antibodies against Hrp1 (gift from Dr. Michael Henry, [135]) or Nab3 (2F12-2; Dr. M. Swanson [11]). Signal was detected using horse radish peroxidase-conjugated anti-mouse IgG (Sigma Chemical Co.) and enhanced chemiluminescence performed in 100mM Tris pH 8.5, 1.25 mM luminol, 200nM p-coumaric acid, and 0.01% H_2O_2.

3.4 RESULTS

Some chimeric Nab3 proteins support viability.

The Nab3 RNA-binding protein contains an essential, carboxy-terminal segment that includes a low complexity domain that is Q/P-rich (Figs. 1 and S1) [11, 26, 27]. The domain scores as prion-like using a prediction algorithm trained with authentic yeast prions (Fig. 3-S1) [34]. Nab3 functions as a multimer in binding nascent RNA, leading to the hypothesis that the essential function of the LCD is to enable the assembly of Nab3, and possibly other proteins, as they work to terminate transcription across the genome [26, 27]. To further test this idea and learn if simply any interaction domain could substitute for Nab3's LCD, we prepared a lethal truncated version of Nab3 (Fig. 3-1, Nab3Δ191) that is

deleted for most of its LCD. This derivative, expressed from the *NAB3* promoter, was used as a stem onto which we grafted other prion-like domains in place of the natural sequence (Fig. 3-1). Using as a starting point the Alberti *et al* [34] catalog of yeast proteins with prion-like domains, we tested: 1) prion-like LCDs from three termination-related yeast RNA-binding proteins Hrp1, Pcf11, and Rat1, 2) LCDs from two well-established cytoplasmic prion proteins, Rnq1 and Sup35, that are unrelated to transcription, and 3) an LCD from a membrane-interacting protein (Ent2) that has a high score as prion-like [34]. These chimeric Nab3 proteins were encoded on *LEU2*-marked plasmids which were individually introduced into a yeast strain in which chromosomal *NAB3* was disrupted, and which contained a covering *URA3*-marked plasmid with wildtype *NAB3*. The ability of different isolates to survive by virtue of their use of the resulting chimeric Nab3 proteins, following the loss of the wildtype *NAB3*-*URA3* plasmid, was tested on media containing fluoroorotic acid (FOA). Surprisingly, some, but not all, of these alternative prion-like sequences, supported cell viability when grafted onto Nab3 (Fig. 3-2). These included the prion-like region from the transcription termination factor Hrp1, the prion domain from the well-studied Sup35 protein, and the core prion-like domain from the membrane-associated cytoplasmic protein Ent2 (Figs. 2, S1). Fusions to Rat1, Pcf11, or Rnq1's prion-like domains, all of which have been previously observed to autonomously form amyloid filaments and one of which, Rnq1, is a *bona fide* prion, did not substitute for Nab3's endogenous LCD in this plasmid shuffle regimen. Western blots using an anti-Nab3 antibody, confirmed that the resulting strains bearing the LCD-fusion partners from Sup35 (Nab3Sup35), Hrp1 (Nab3Hrp1), or Ent2 (Nab3Ent2) contained only the chimeric proteins as their sole source of Nab3 (Fig. 3-2). The chimeric proteins that did not support viability

accumulated to levels similar to wildtype Nab3 in the strains that bore both the wildtype and chimeric plasmid (Rat1, Pcf11, Rnq1; Fig. 3-2), eliminating the possibility that these Nab3-derivatives failed to support growth because they were malformed and degraded. It is also noteworthy, while the 60 amino acid Hrp1 prion-like domain can substitute, that the large (282 amino acid) extended prion domain of Rnq1 could not reconstitute Nab3 function. Thus, the length of the added LCD is not positively correlated with function in this assay.

Work on yeast prion proteins has shown that the primary amino acid sequence of the glutamine- or asparagine-rich prion domains is less important for prion formation than their overall composition [138, 139]. This is seen in experiments in which the primary sequence can be scrambled and yet it remains capable of supporting prion formation [138, 139]. To test if the Hrp1 LCD, which substituted for Nab3's LCD, could be scrambled and still retain function, we subjected the order of the amino acids to randomization which generated a version with an overall amino acid composition identical to the wildtype sequence (Fig. 3-S1), and appended it to the Nab3 stem. As seen previously for other yeast prion-domains, Hrp1's LCD could be scrambled and support cell survival (Fig. 3-2).

To test the ability of Nab3Hrp1, Nab3Sup35, and Nab3Ent2 to support transcription termination, we introduced a termination-reporter plasmid into the strains in which each chimera was the only source of Nab3. This reporter contains a strong terminator from the *IMD2* locus inserted between the *GAL1* promoter and the green fluorescent protein (GFP) reading frame. By using flow cytometry to score GFP abundance in living cells, this plasmid forms the basis of a sensitive assay for the termination of transcription by RNA polymerase II [35]. Strongly fluorescent cells are indicative of terminator readthrough

when Nab3 activity is defective, whereas when termination is operative, GFP fluorescence is not detected (Fig. 3-3A, "-" and "+", respectively). In this assay, all three chimeras displayed nearly full terminator function. These results indicate that some heterologous LCDs restored termination factor activity to Nab3 lacking its C-terminal domain.

The Nab3Sup35 chimera shows an epigenetic inheritance pattern

The strain with Nab3Sup35 showed a curious ability to generate infrequent variants that lost terminator function and started expressing GFP from the reporter plasmid, as seen by direct illumination of colonies on solid medium (Fig 3B). Over time, a plate of termination-competent, GFP-negative colonies will produce sporadic, termination-defective GFP-positive colonies. Re-streaking cells from GFP-negative colonies led to the emergence of GFP-positive colonies at a rate similar to that of the original colony (~$3x10^{-3}$). However, re-streaking of cells from colonies that were GFP-positive led to progeny colonies all of which were GFP-positive (Nab3Sup35*), demonstrating the heritability of this feature as would be expected for prion formation (Figs. 3C and 3D). Since native Sup35 is known to form cytoplasmic aggregates in a prion domain-dependent manner, and this is concomitant with loss of Sup35's function, we hypothesized that the Nab3Sup35 chimera was recruited into an inactive form, presumably as a cytoplasmic amyloid. In so doing, transcription termination on the reporter plasmid would be compromised, and GFP would be expressed. Consistent with the idea that this was a manifestation of Sup35's prion properties, the strain with the Nab3Hrp1 chimera did not generate this type of spontaneous fluorescent colony. Taken together, these data show that some, but not all, heterologous prion-like domains can substitute for the endogenous region of Nab3.

Furthermore, the Nab3-fused Sup35 prion domain appears to retain its capacity to aggregate in living cells and display heritability while attached to Nab3. Interestingly, the rescuing pieces of protein share little or no primary sequence homology with Nab3's LCD (Fig. 3-S1), aside from low complexity, suggesting that rescue is achieved by another feature such as polymer formation, and not simply recognition of a specific primary sequence in the domain.

An unbiased selection for sequences that substitute for Nab3's LCD

Given the apparent lack of sequence specificity of the polypeptides that could rescue Nab3 function, we asked whether random DNA could provide essential sequences to Nab3Δ191 using yeast survival as the basis for an unbiased selection. A library of plasmids containing fragments of calf thymus DNA were fused to the stem of Nab3 used to generate the aforementioned chimera, and the plasmids were introduced into yeast with wildtype *NAB3* on a plasmid. Strains that could survive the loss of wildtype *NAB3* by living off of the Nab3 fusion protein that had acquired calf thymus sequences, were selected on FOA-containing medium. A plasmid that supported growth was recovered from this strain and sequenced showing that the rescuing piece of bovine DNA was an 817bp insert derived from a non-coding portion of bovine chromosome 19 that adds 28 amino acids to the Nab3 stem before a stop codon is encountered (Figs. 2and S1). The resulting fused peptide is not glutamine- or asparagine-rich, a feature of many prion-like low complexity domains, although it does have a somewhat biased content of leucine, valine, and cysteine (14% each). This finding suggests there is a remarkable tolerance in the primary sequence needed for the essential function of this RNA-binding protein, and its role can be fulfilled

by many alternative sequences. However, not every fusion partner sufficed, indicating that simply 'capping' the end with any sequence is not what provides the essential function.

Substitution of LCDs into Hrp1

From a gene ontology perspective, Hrp1 appears similar to Nab3. Both are RNA-binding hnRNPs with LCDs that score highly as prion-like. Their LCDs form filaments *in vitro* [34, 36]. Both exist in a complex of proteins that interact with RNA polymerase II and both are involved in the termination of transcription by the enzyme [67, 69, 140]. First, we tested if Hrp1's natural prion-like core domain (60 amino acids) was essential for cell viability as seen for the Nab3 piece removed above. While wildtype *HRP1* on a plasmid could cover a deletion of a chromosomal copy of the gene, a plasmid encoding Hrp1 lacking its core LCD was not viable, (Fig. 3-4, "FOA") even though the deletion-derivative could be expressed in yeast with wildtype *HRP1* (Fig. 3-4, lane 4). Since the Hrp1 LCD, and the 28-amino acid sequence derived from bovine DNA could be swapped for Nab3's essential domain, we tested if the bovine peptide could substitute for Hrp1's core prion-like domain. Surprisingly, this sequence could also substitute for Hrp1's core prion-like domain, where it was well expressed and enabled cell viability (Fig. 3-4, "FOA" and lane 3), although cells grew more slowly compared to those with wildtype *HRP1*. Since 113-amino acids of Sup35's low complexity prion-domain could be added to the Nab3 stem where it provided an essential function (Figs. 1 and 2), we asked if the Sup35 prion-domain could do the same if exchanged for Hrp1's core prion-like domain. Indeed, this construct also rescued viability (Fig. 3-4, "FOA"), demonstrating the remarkable cross-functionality of prion and prion-like domains, as well as the random sequence selected from mammalian

DNA, when added to two different RNA-binding proteins in place of their native LCDs. (Sup35's prion-domain is not essential for cell viability, so the viability test of exchanging native for heterologous sequences could not be applied to it.) As a negative control, we inserted 60 amino acids of the His_6 and S-Tag-containing linker from pET32a in place of the same length of Hrp1's prion-like domain. This 'Hrp1-linker' derivative, while expressed in cells with wildtype *HRP1*, could not alone support viability (Fig. 3-4 "FOA"). This controls for the possibility that the native Hrp1 sequence serves a 'spacing' function which might mean that replacing it with virtually any sequence would work. Hence, the LCD sequences we added are providing a specific property to the rest of the Hrp1 scaffold to cover function. It should be noted that these grafting experiments, in which the heterologous exchange for Hrp1 was internal, differed from those for Nab3 where additions were made to its C-terminus.

Nab3's essential C-terminus can be subdivided

The smallest portion of Nab3's C-terminal region needed for viability was previously delimited to 134 amino acids as shown by plasmid shuffling or integration of the deleted allele into the chromosome of a diploid followed by tetrad dissection ([27] and Fig. 3-S2). The sequence contains two discernible components, the Q/P-rich region and the C-terminal-most 18 amino acids which has features of a leucine zipper and structural homology to a self-assembling segment of human hnRNP-C [38]. Deletion of this small peptide from Nab3's terminus, or a subtle leucine to alanine substitution within it, are not lethal, but do slow growth and compromise Nab3's termination function [35, 38]. To determine the contribution that these two elements provide to the region's essential

function, we deleted the Q/P-rich segment leaving only the hnRNP-C-homology. Surprisingly, this derivative was viable (Fig. 3-5 "FOA"), indicating that neither loss of a substantial part of the Q/P-rich domain, nor the C-terminal α-helix, are required for cell viability, however, both cannot be lost for the protein to remain functional. Hence, the Nab3 C-terminus has a bipartite nature based upon structural predictions and functional assays. We asked if the hnRNP-C-homology domain, with the previously defined L800A mutation that impacts termination, could support the loss of the Q/P domain. The mutation compromised but did not abolish the homology region's ability to protect against loss of the Q/P sequences. These results suggest that the rescuing effect is related to the role of the domain in termination and that the LCD-chimeras described above appear to provide this dual property of the native sequence when fused to the Nab3 stem.

3.5 DISCUSSION

Numerous cellular RNA-binding proteins contain LCDs that can self-assemble. This enables them to form dynamic compartments in cells in which RNA metabolism is sequestered. In this report, we set out to determine the interoperability of LCDs in the yeast proteome to gain insight into their role in the context of a well-studied RNA-binding protein. We exploited the fact that Nab3 contains an essential C-terminal region, which permitted us to test heterologous LCDs whose sequence and potential for self-assembly stem from their relatedness to yeast prions. The ability to rescue viability by the chimeras was followed by functional tests of termination. A starting set of six LCDs were derived from proteins of varying functions, four RNA-binding proteins, two canonical yeast prions, and one membrane associated epsin-like protein. Three of these resulted in fusion proteins that could supply function to live yeast. The chimeric proteins showed near normal termination activity in our assays. Interestingly, the viable fusions came from proteins with disparate biological roles. Hrp1 is a component of CF1 and is a nuclear RNA-binding protein necessary for the hydrolysis and polyadenylation of pre-mRNA's 3' ends. It is also involved in termination of short non-coding RNAs and autoregulation of its own mRNA. Thus, it is similar to Nab3's biochemical and biological functions. In contrast, Sup35 is a cytoplasmic translation factor that associates with ribosomes and is a canonical yeast prion and Ent2 is a cytoplasmic epsin-like protein involved in the endocytic pathway that has no obvious direct interactions with RNA or RNA-binding proteins. The Sup35 prion domain is not essential for its role in translation but in a sense, it becomes 'essential' when added to Nab3 in the aforementioned grafting experiment.

A simple model suggests that any amyloid-forming LCD can substitute for any other such domain, if the only critical function it provides is that of self-assembly. That clearly is not the case here. Alternatively, the domains might also encode specific protein-protein docking sites for which their amino acid primary structure would be important. That does not seem to be the case here either, based on the surprising finding that distinct LCDs could cover for one another, albeit not universally. Proteins with similar biological functions might be expected to have cross-compatible LCD's. Yet neither termination-related protein Rat1 nor Pcf11 had amyloid-forming LCDs that could supply function to the Nab3 stem.

What properties create the specificity of function within the family of LCDs? Table 3 shows the composition of the domains studied here with respect to those amino acids over-represented in amyloid-forming prion-like domains. The six amino acids glutamine, asparagine, proline, serine, tyrosine, glycine, and methionine compose 65-90% of the yeast domains we examined (Table 3). In general, successful LCDs were no closer to the amino acid composition of the piece of the LCD removed from Nab3, than were those LCDs that failed the test (Table 3). For example, Pcf11's LCD (39% Q) could not be swapped for the region removed from Nab3 (27%Q) in our experiments, whereas the 15% Q domain from Hrp1 could. There has been a lot of attention paid to Q-rich domains because of their ability to assemble into amyloid filaments and their involvement in human disease. The abundance and length of Q-tracts in proteins is important for their aggregation properties. This is true for Nab3 where the glutamines in Nab3's LCD are important since a reduction in Q-content by mutation destroys the LCD's ability to form an amyloid polymer and leads to cellular lethality [36]. A similar discordance was seen for proline content where Nab3's

18%-proline domain could be replaced by Hrp1's which was only 4% proline. Likewise, the ability of an LCD to rescue is not correlated with the LCD's: 1) asparagine content, 2) proline plus glutamine content, or 3) glutamine plus asparagine content, as can be seen in Table 3. Our data are consistent with a model in which a diffuse and redundant amino acid composition that leads to the host protein's assembly, is the important property provided to the Nab3 stem. However, different LCDs must be able to assemble into distinct conformers, assemble to different extents, or otherwise generate different structures to which other proteins might join. This model is attractive because it obviates the need to postulate that a specific primary sequence is needed, while proposing that not all assembly/polymerization functions of LCDs are equivalent. However, without additional experimentation, we cannot resolve the specific and complex primary and secondary structural characteristics that are responsible for LCD function in the context of Nab3. In any case, there appear to be subclasses of LCDs, the effectiveness of which cannot be predicted by the function of their host protein. One possibility is that each category of LCD may be a determinant of its host proteins' ability to localize to a specific subcellular location, *e.g.* during recruitment of RNA-binding proteins to stress granules.

The viable, chimeric Nab3 proteins were capable of operating the *IMD2* terminator that regulates Nab3. Flow cytometric analysis found near wild type levels of GFP expression, implying that the heterologous LCDs efficiently complement Nab3's function in our termination assay. While the exact *in vivo* function of Nab3's LCD is unclear, the findings here support prior suggestions [36, 38] that it mediates either homotypic or heterotypic protein-protein interactions. It is interesting that the LCDs of proteins that localize to different compartments of the cell and possess disparate functions, can mediate

these same interactions. The spontaneous emergence of cross generationally stable, termination defective yeast containing the Nab3Sup35 chimera was an unexpected outcome. The finding suggests that Nab3Sup35 is undergoing some prion-like event resulting in stably heritable, termination defective yeast. The phenomenon is analogous to the molecular genetic protein nullification tool called 'anchor-away' which is a method of re-localizing to the cytoplasm, and thereby inactivating, a nuclear protein by genetically fusing it to a protein that binds a cytoplasmic counterpart [141]. It is also reminiscent of the aberrant aggregation-dependent cytosolic migration seen for nuclear RNA-binding proteins with LCDs such as TDP-43, FUS, or hnRNP-A1 which aberrantly accumulate in the cytosol in diseases such as amyotrophic lateral sclerosis [105].

Since the heterologous LCD's shared little primary sequence similarity, we performed an unbiased screen using calf thymus DNA as a source of coding sequence to search for a fusion protein capable of sustaining viability. The bovine sequence we found was intriguing. The evidence that it was functional was strongly supported by its ability to rescue the lethality of the loss of either Nab3 or Hrp1's LCD. It did not score well as a prion, due in part to the relative paucity of Q and N residues (4% and 0%, respectively) and relatively high content of charged residues (29%), but it does have a biased sequence makeup, significant hydrophobicity, and low complexity with a repeated set of cysteines, valines, and leucines composing 43% of the sequence. It seems plausible that the sequence rescues Nab3 function by facilitating Nab3 polymerization, although it may very well do so in a manner distinct from that of prion-like LCDs.

We have previously characterized an hnRNP-C like α-helical domain at the C-terminus of Nab3 adjacent to the LCD [38]. Loss of this helix, or even subtle alterations

to its sequence such as the L800A substitution, had profound impacts on growth and termination, although it was not lethal. Similarly, removal of the Q/P-rich region alters growth and termination but alone is not lethal (Fig. 3-5). While either the Q/P-rich domain or the hnRNP-C structural homology segment are sufficient for viability, loss of both is lethal, implying possible functional redundancy between the two domains. The fact that the C-terminal α-helical domain's ability to ameliorate the loss of a large stretch of Nab3's LCD is reduced when it contains the L800A mutation that affects termination, suggests that the contribution of this region of the protein to viability is related to its role in transcription termination.

This report expands our understanding of a well characterized termination factor, Nab3, and the role of its LCD in termination function. It also suggests that while a simple primary sequence does not determine an LCD's function, there does appear to be specialization of function within the family of LCDs. This adds to the growing literature on the important roles of these domains in biology and shows that the LCDs are more complex than previously appreciated.

Acknowledgements

The authors thank Anna Kania and Juan Rodriguez for their contributions to this work during their rotation research. We also thank Dr. Natalya Degtyareva for helpful discussions and tetrad analysis and Dr. Graeme Conn for a critical reading of the manuscript. This research project was supported in part by the Emory University School of Medicine Flow Cytometry Core. Contributions of strains or antibodies from the labs of Drs. Pamela Silver, Maurice Swanson, and Michael Henry are acknowledged. Research reported in this publication was supported the National Institute of General Medical Sciences of the National Institutes of Health under awards R01GM120271. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Figure 3-1

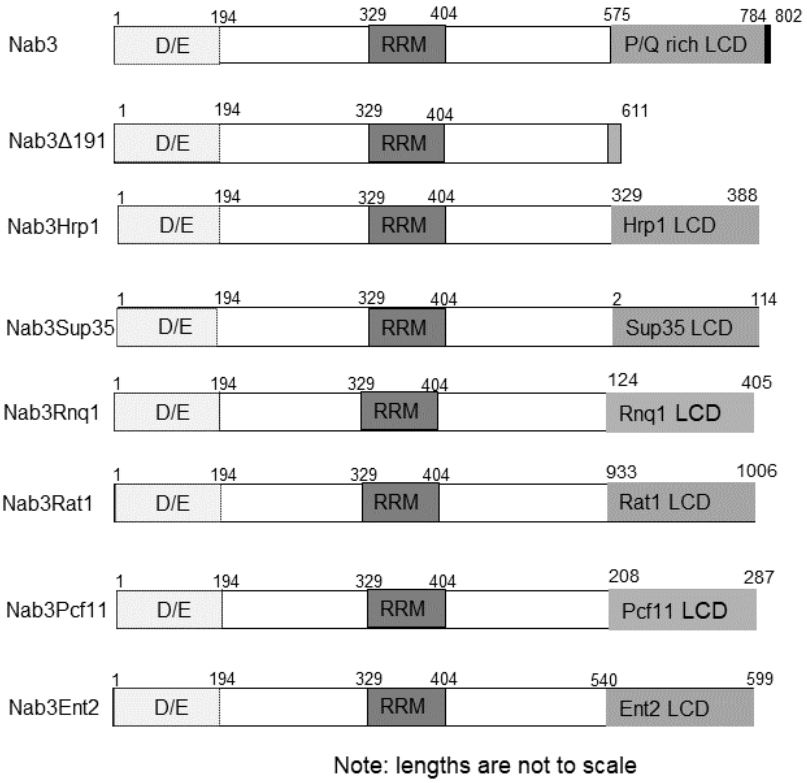


Fig. 3-1- Schematic of Nab3 chimeras. The D/E-rich and RRM domains of Nab3 are indicated. The positions of the amino acids derived from the respective heterologous proteins are indicated. The hnRNP-C homology domain (784-802) is indicated in black at the C-terminus.

Figure 3-2

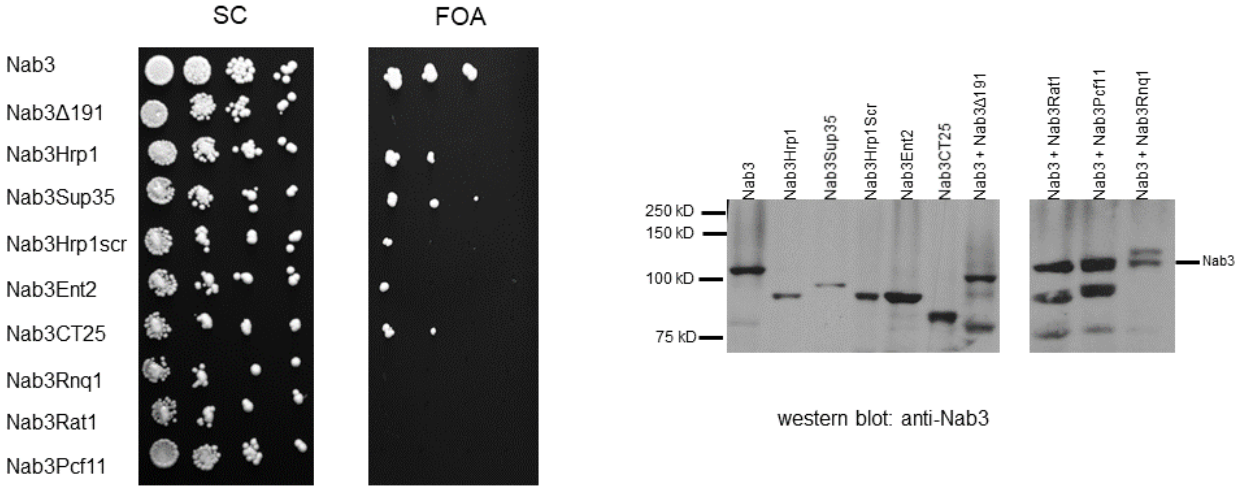


Fig. 3-2- Growth and protein expression by yeast strains with Nab3-chimeras. Yeast cultures were diluted to 0.05 OD_{600} and serially ten-fold thereafter, 10 µl were spotted solid media, and incubated at 30°C. From the top down, the strains were expressing wildtype Nab3 and a second Nab3 protein, either wildtype Nab3 (DY351), the Nab3Δ191 'stem' (DY3183), the Nab3Hrp1 chimera (DY3193), the Nab3Sup35 (DY4002), the Nab3Rnq1 chimera (DY3186), the Nab3Rat1 chimera (DY3184), the Nab3Pcf11 chimera (DY3185), the Nab3Hrp1scrambled chimera (DY3213), the Nab3Ent2 chimera (DY3244), or the Nab3CT25 chimera (DY4001). The top seven strains were grown on a single plate of SC or FOA, as indicated. The bottom 3 were grown on separate plates of SC and FOA. (The blemishes visible in the Nab3-Rat1 sample spotted onto FOA

were divots in the agar surface.) The right side of the figure is a western blot of whole cell lysates from logarithmically grown strains containing the indicated Nab3 protein (left to right: DY3033, DY3196, DY3182, DY4006, DY3245, DY4004, DY3183) or wildtype Nab3 protein with an additional Nab3 derivative that could not support growth on its own (DY3184, DY3185, DY3186) lysed in sample buffer and subjected to SDS-PAGE and western blotting with an antibody against Nab3, as described in Methods. Each of the two western blot panels were run on separate gels and processed separately.

Figure 3-3

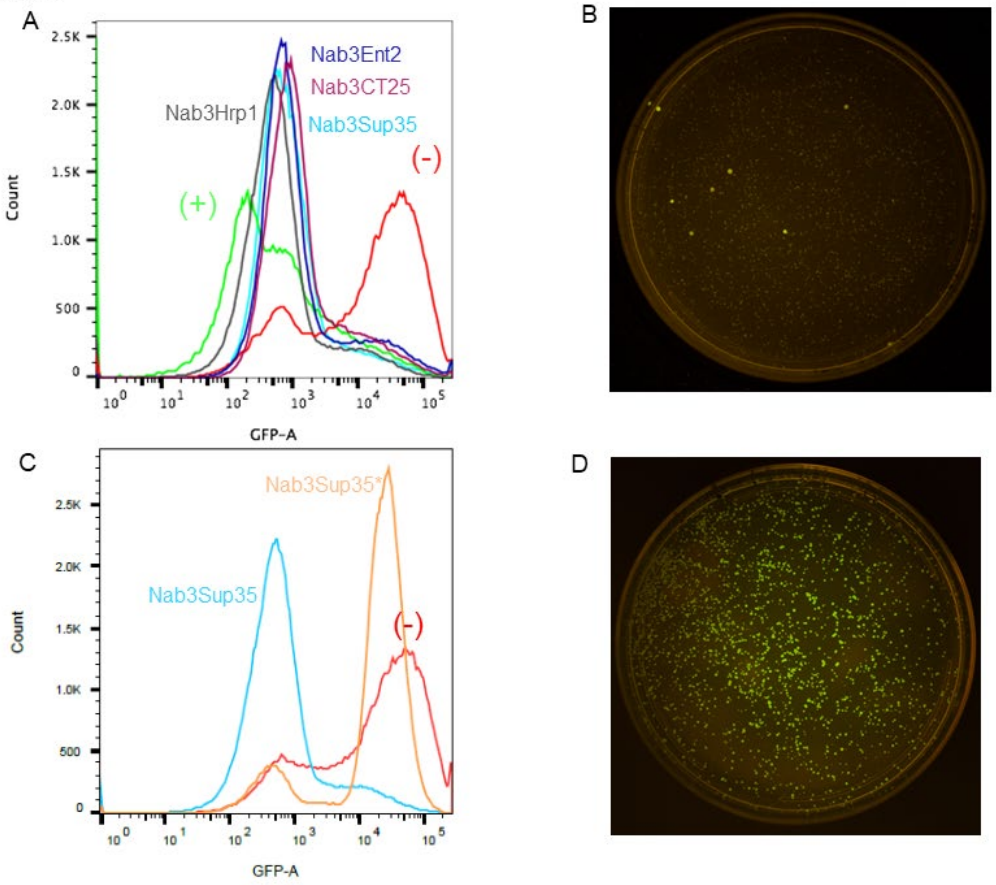


Fig. 3-3- GFP expression in strains with chimeric Nab3 proteins. (A) Flow cytometry of control strains showing populations with fully functional termination [DY3217 (+)] or no termination [DY30239 (-)]. Results for experimental strains expressing Nab3Hrp1 (DY3197), Nab3Sup35 (DY3187), or Nab3Ent2 (DY3246) expressed from a plasmid as the sole Nab3 protein, are shown. (B) GFP fluorescence for DY3187 was monitored for colonies grown on SC leu$^-$ solid medium supplemented with galactose and illuminated with the Dark Reader optical system (Clare Chemical Research, Inc.). (C) Flow cytometry of a Nab3-Sup35-expressing strain before (DY3187) and after (DY3204) conversion to the reporter-positive state (see text). A control strain was analyzed that lacks terminator function [DY30239 (-)]. (D) GFP fluorescence for DY3204 was monitored for

colonies grown on SC leu$^-$ containing galactose after illumination with the Dark Reader optical system (Clare Chemical Research, Inc.).

Figure 3-4

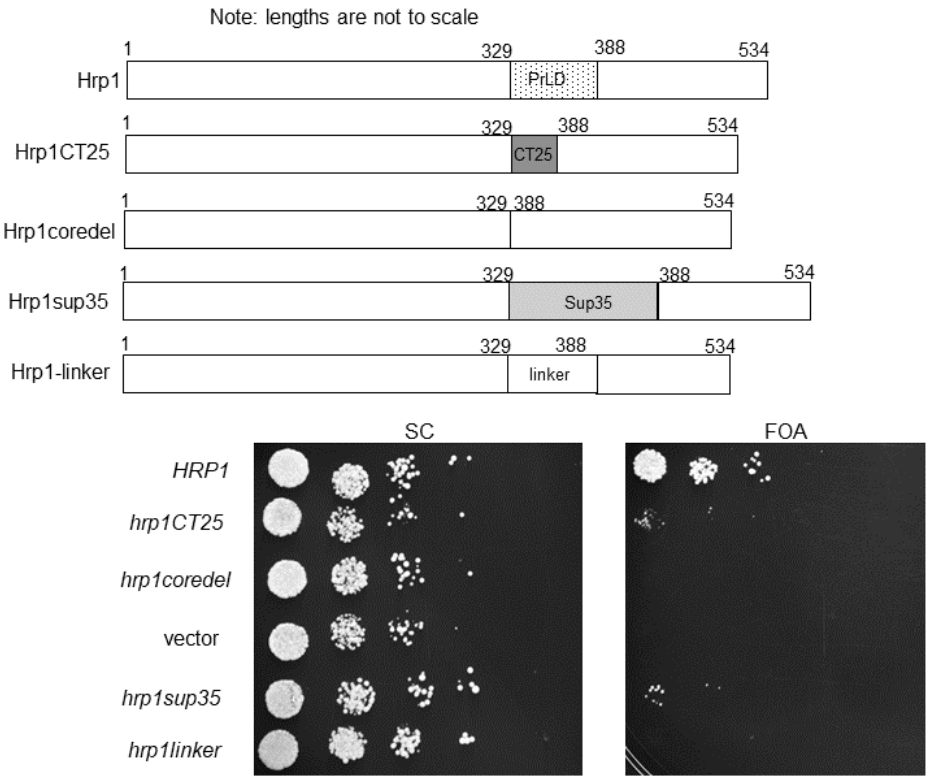


Fig. 3-4- Cell viability and expression of Hrp1-chimeric proteins. A schematic depiction of the wildtype Hrp1 and the indicated LCD derivatives is shown at the lower right. Strains (DY388, DY387, DY389, DY1638, and DY3242; from top to bottom) carrying the respective plasmid-encoded proteins as well as a *URA3*-marked plasmid with *HRP1*, were grown in SC leu$^-$, diluted to OD_{600} of 0.05 and serially diluted ten-fold four times. Ten microliters of each dilution were spotted onto the indicated solid media and cells were grown at 30°C. For western blotting, whole cell lysates from strains DY388, DY387, DY387F, DY389, DY3242, and DY3243 were boiled in loading buffer, resolved by SDS-PAGE, transferred to nitrocellulose, and probed with anti-Hrp1 antibody, as described in Methods. Pre-stained marker proteins were run in an adjacent lane and their sizes are indicated at left.

Figure 3-5

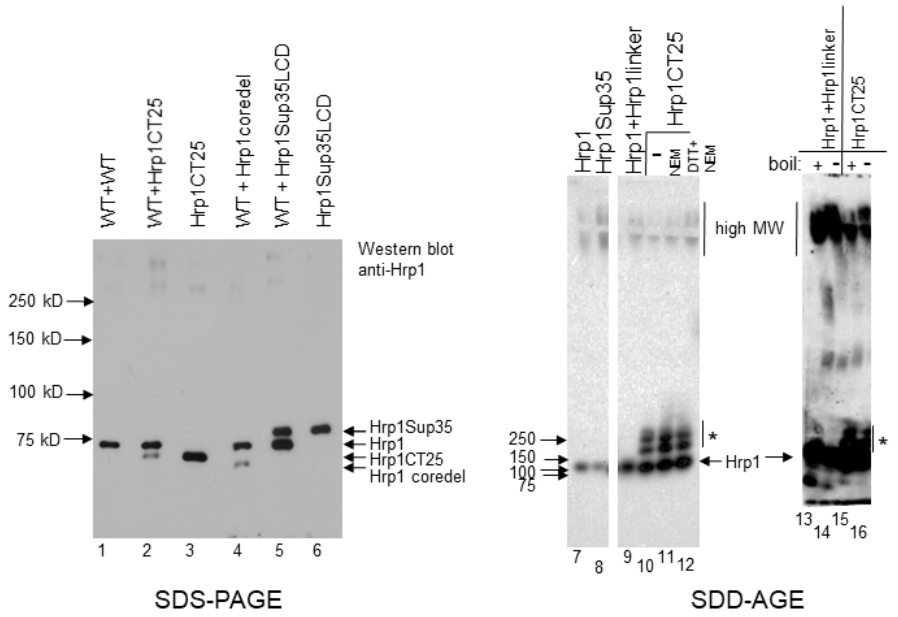


Fig. 3-5- Western blotting and SDD-AGE analysis of Hrp1 and related chimera.

Whole cell lysates from strains DY388, DY387, DY387F, DY389, DY3242, and DY3243 were boiled in loading buffer, resolved by SDS-PAGE, transferred to nitrocellulose, and probed with anti-Hrp1 antibody, as described in Methods. Pre-stained marker proteins were run in an adjacent lane and their sizes are indicated at left. To the right are SDD-AGE assays analyzing lysates from strains bearing the indicated Hrp1-derivatives (lane 7 = DY308, lane 8 = DY3243, lane 9 = DY4500, lanes 10–12 = DY387F, lanes 13 and 14 = DY388, lane 15 and 16 = DY387F), and developed with anti-Hrp1 antibody, as described in Methods. Lanes 7–12 (all lanes are derived from the same gel and filter) show that the oligomeric forms of Hrp1CT25 ("*") are specific for that Hrp1 derivative, and that the proteins are relatively resistant to reducing agent. Lanes 13–16

show that the oligomers ("*") and high molecular weight (MW) forms of Hrp1 and Hrp1CT25 are relatively resistant to boiling in SDS for 5 min before electrophoresis.

Figure 3-6

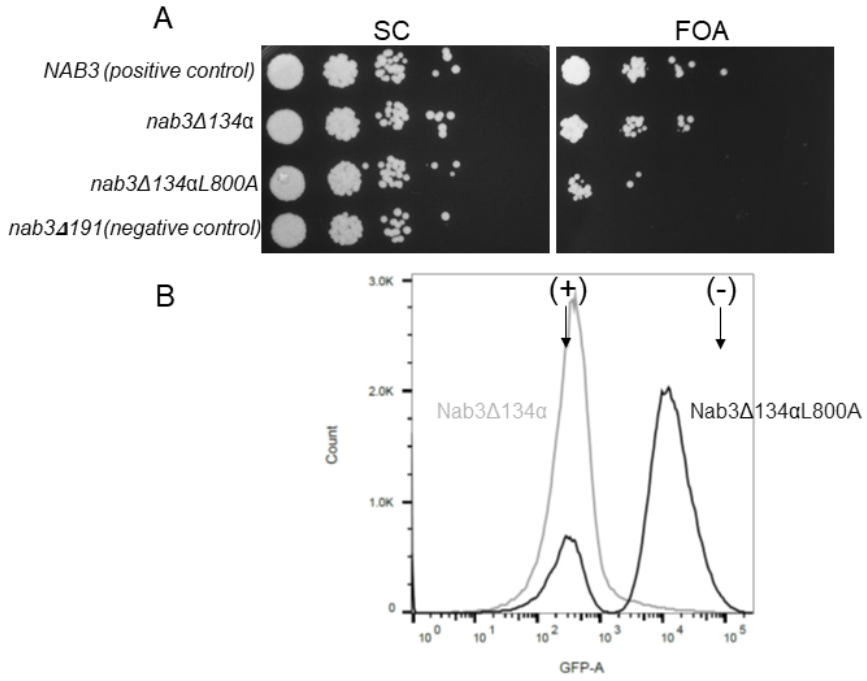


Fig 3-6. Bipartite nature of the Nab3 C-terminal essential region.

A) Yeast strains containing a plasmid bearing wildtype *NAB3* on a URA3-marked plasmid and either the *nab3Δ134α* allele (strain DY353) or *nab3Δ134αL800A* allele (strain DY377) on *LEU2*-marked plasmids, were grown on complete medium (SC) or medium with FOA to test for viability after the loss of the former plasmid. Cells were diluted, spotted, and grown on a single plate of each indicated solid media at 30°C, along with a positive control (DY351; with wildtype *NAB3* on a *LEU2*-marked plasmid) and negative control strain with a non-viable allele (*nab3Δ191*) of *NAB3* (DY3183). B) The resulting shuffled FOA-resistant strains (DY359 [gray line] and DY379 [black line]) were tested for their termination competence as described in

Materials and Methods. In lieu of reproducing the control strain peaks seen in Fig 3, only the position of the fluorescence maxima (arrows) for the termination-competent (+) (DY3217) and termination-defective (-) (DY3218) control strains are shown for reference.

Supplemental Figures

Figure 3-S1

Fig. 3-S1- Sequences of proteins studied here. The entire Nab3 sequence is shown with position 191 identified. Nab3$^{1\text{-}191}$ was used as the 'stem' for the addition of LCDs from the remaining group of proteins. Italicized residues are the prion-like domain based on Alberti *et al.*'s algorithm [34]. The bold sequence is the region of Nab3 with structural homology to hnRNP-C's α-helix [38]. LCDs from yeast are listed with their native coordinates. The scrambled Hrp1 sequence and sequences from the bovine genome described here are also listed.

Figure 3-S2

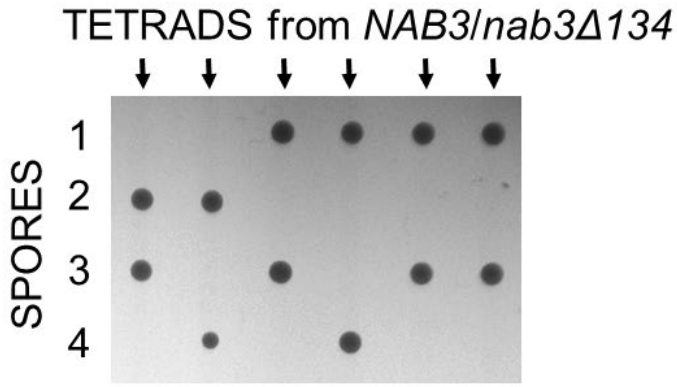


Fig. 3-S2- Tetrad dissection of a diploid with wildtype *NAB3* and *nab3Δ134* alleles. The arrows indicate six columns each from different tetrads. The rows are the four positions in which spores were deposited.

Table 3-1- Oligonucleotides used.

RNQ1 F	5’ ATATATCATATGGTAGTGGTTCTGGCGGC 3’
RNQ 1 R	5’ ATATATCTCGAGTCAGTAGCGGTTCTGGTTGC 3’
PCF11 F	5’ ATATATCATATGGTCAAGTTCAAATGCAACTA AGGC 3’
PCF11 R	5’ ATATATCTCGAGCTAAGAATTCTGTTGGTTCGTTGT ATC 3’
RAT1 F	5’ ATATATCATATGGTAATAATGTCCAACCCGCC 3’
RAT1 R	5’ ATATATCTCGAGCTAACGCCTATTTGCTCTTGAA 3’
HRP1 F	5’ ATATATCATATGGTAACAATGGTGGTAACAATGG 3’
HRP1 R	5’ ATATATCTCGAGCTATTGCATTTGTTGGTAATATTCTTG 3’
SUP35 F	5’ ATATCATATGGTTCGGATTCAAACCAAGGC 3’
SUP 35 R	5’ ATATCTCGAGCTATTGATATCCTTGCAAATTGTTATTG 3’

Table 3-2- Yeast Strains studied here.

DY307	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS315-*HRP1* (*LEU2*)] [pRS316-HRP1 *(URA3)*]	this study
DY308	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS316-*HRP1 (URA3)*]	this study
DY318	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS316-*HRP1 (URA3)*] [pRS315-HRP1coredel (*LEU2*)]	this study
DY351	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanM*X [pRS316-*NAB3* (*URA3*)] [pRS315-*NAB3* (*LEU2*)]	this study
DY353	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3* (*URA3*)] [pRS315-Nab3Δ134α (*LEU2*)]	this study
DY359	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanM*X [pRS315-Nab3Δ134α (*LEU2*)]	this study
DY377	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-NAB3 (*URA3*)] [pRS315-Nab3Δ134α-L800A(*LEU2*)]	this study
DY379	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3Δ134α-L800A(*LEU2*)]	this study
DY387	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS316-*HRP1 (URA3)*] [*pRS315Hrp1coredelCT25 (LEU2)*]	this study
DY387F	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS315-Hrp1coredelCT25 *(LEU2)*]	this study
DY388	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS316-*HRP1 (URA3)*] *[pRS315-HRP1(LEU2)*]	this study
DY389	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS316-*HRP1 (URA3)*] *[pRS315-Hrp1coredel (LEU2)*]	this study
DY1638	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS316-*HRP1 (URA3)*] [pRS315 *(LEU2)*]	this study
DY3033	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3FL (*LEU2*)]	this study
DY3111	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3* (*URA3*)]	this study
DY3134	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3 (URA3)*] [pRS315-Nab3Δ134 (*LEU2*)]	this study
DY3182	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3Sup35 (*LEU2*)]	this study
DY3183	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3* (*URA3*)] [pRS315-Nab3Δ191STOP (*LEU2*)]	this study
DY3184	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3 (URA3)*] [pRS315-Nab3Rat1 (*LEU2*)]	this study
DY3185	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3 (URA3)*] [pRS315-Nab3Pcf11 (*LEU2*)]	this study
DY3186	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3 (URA3)*] [pRS315Nab3Rnq1 (*LEU2*)]	this study
DY3187	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3Sup35 (*LEU2*)] [pREFGFPKan (*URA3*)]	this study
DY3193	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3 (URA3)*] [pRS315-Nab3Hrp1 (*LEU2*)]	this study
DY3196	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3Hrp1 (*LEU2*)]	this study

DY3197	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3Hrp1 (*LEU2*)] [pREFGFPKan *(URA3)*]	this study
DY3204	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3Sup35]* (*LEU2*) [pREFGFPKan *(URA3)*] *epigenetic change to GFP expression (see text)	this study
DY3210	*MATa/α ade5-1/ade5-1 his7-2/his7-2 leu2-3/leu2-3 trp1-289/trp1-289 ura3-52/ura3-52 NAB3/nab3□134*	this study
DY3212	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3CT25 (*LEU2*)]	this study
DY3213	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3 (URA3)*] [pRS315-Nab3hrp1scr (*LEU2*)]	this study
DY3217	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3 (*LEU2*)] [pREFGFPKan (*URA3*)]	this study
DY3242	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS316-HRP1 *(URA3)*] [pRS315 Hrp1Sup35LCD *(LEU2)*]	this study
DY3243	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS315 Hrp1Sup35LCD *(LEU2)*]	this study
DY3244	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS315 Nab3ENT2LCD *(LEU2)*] [pRS316-*NAB3 (URA3)*]	this study
DY3245	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS315 Nab3ENT2LCD *(LEU2)*]	this study
DY3246	*MATa hrp1::HIS3 ura3 ade2 ade8 his3 leu2 lys1* Trp$^-$ [pRS315 Nab3ENT2LCD *(LEU2)*] [pREF-GFPKan(URA3)]	this study
DY3912	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3□134α (*LEU2*)] [pGAL-REF-GFP *(URA3)*]	this study
DY3913	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3Δ134αL800A (*LEU2*)][pGAL-REF-GFP *(URA3)*]	this study
DY4001	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3 (URA3)*] [pRS315-Nab3CT25 (*LEU2*)]	this study
DY4002	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3 (URA3)*] [pRS315-Nab3Sup35 (*LEU2*)]	this study
DY4004	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3CT25 (*LEU2*)]	this study
DY4006	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315-Nab3hrp1scr (*LEU2*)]	this study
DY30239	*MATα ura3Δ0 his3Δ1 leu2Δ0 lys2Δ0* [pGalGFP-Kan (*URA3*)] [pRS315 (*LEU2*)]	this study
PSY818	*MATa hrp1::HIS3 ura3 ade2-1 ade8 his3 leu2-3 lys1* Trp$^-$ [pRS315-HRP1 (*LEU2*)]	Henry *et al*. 2003
YH990	*MATa/α ade5-1/ade5-1 his7-2/his7-2 leu2-3/leu2-3 trp1-289/trp1-289 ura3-52/ura3-52*	Morrison *et al*. 1991

Table 3-3–Frequency in proteins studied here of the residues commonly over-represented in amyloid-forming LCDs.

	Length	%Q	%N	%S	%P	%Y	%G	%M	%PQ	%QN	%QNP	%QNSPYGM
Nab3$^{609-803}$	194	27	6	10	18	5	7	4	45	6	51	77
Hrp1$^{329-388}$	**60**	**15**	**22**	**0**	**5**	**10**	**17**	**15**	**20**	**22**	**42**	**83**
Sup35$^{2-114}$	**113**	**28**	**18**	**4**	**4**	**18**	**17**	**0**	**33**	**18**	**50**	**88**
Ent2$^{540-599}$	**60**	**65**	**5**	**0**	**10**	**8**	**0**	**2**	**75**	**5**	**80**	**90**
Pcf11$^{208-287}$	80	39	8	8	1	4	3	4	40	8	48	65
Rat1$^{933-1006}$	74	7	30	15	3	12	10	0	10	37	40	77
Rnq1$^{124-405}$	282	25	15	17	1	4	17	2	26	15	41	81
CT25	28	4	0	4	0	0	7	4	4	0	4	18

The regions of the indicated yeast proteins (and the bovine sequence CT25 described herein) were analyzed for their content of the amino acids commonly enriched in amyloid-forming, low complexity domains. The fraction of residues that are glutamine (Q), asparagine (N), serine (S), proline (P), tyrosine (Y), glycine (G), methionine (M), proline+glutamine (PQ), glutamine+asparagine (QN), glutamine+asparagine+proline (QNP), or glutamine+asparagine+serine+proline+tyrosine+glycine+methionine (QNSPYGM) was calculated and rounded to the nearest percent. The bold sections are those LCDs that could substitute for Nab3's, those in plain text did not. The bovine sequence is shown at the bottom.

Chapter 4

Nab3's localization to a nuclear granule in response to nutrient deprivation is determined by its essential prion-like domain

Dr. Joshua Kelley and William Simke are responsible for Figs 4-5 and 4-6. Thomas W. O'Rourke is responsible for Figs 4-S4 and 4-S5A and the western blot on image 4-S2

This chapter is part of a manuscript currently under review for publication

4.1 ABSTRACT

Ribonucleoprotein (RNP) granules are higher order assemblies of RNA, RNA-binding proteins, and other proteins, that regulate the transcriptome and protect RNAs from environmental challenge. There is a diverse range of RNP granules, many cytoplasmic, which provide various levels of regulation of RNA metabolism. Here we present evidence that the yeast transcription termination factor, Nab3, is targeted to intranuclear granules in response to glucose starvation by Nab3's proline/glutamine-rich, PrLD, which can assemble into amyloid *in vitro*. Localization to the granule is reversible and sensitive to the chemical probe 1,6-hexanediol suggesting condensation is driven by phase separation. Nab3's RNA recognition motif is also required for localization as seen for other PrLD-containing RNA-binding proteins that phase separate. Although the PrLD is necessary, it is not sufficient to localize to the granule. A heterologous PrLD that functionally replaces Nab3's essential PrLD, directed localization to the nuclear granule, however a chimeric Nab3 molecule with a heterologous PrLD that cannot restore termination function or viability, does not form granules. The Nab3 nuclear granule shows properties similar to well characterized cytoplasmic compartments formed by phase separation, suggesting that, as seen for other elements of the transcription machinery, termination factor condensation is functionally important.

4.2 INTRODUCTION

In yeast, transcription by RNA polymerase II can be terminated in two major ways: The NNS pathway which primarily generates short noncoding transcripts, or the polyadenylation-coupled termination pathway, where termination is associated with nascent transcript cleavage and polyadenylation of the RNA [142]. The NNS termination pathway contains Nrd1 and Nab3, essential RNA-binding proteins with canonical RRMs [11, 14, 25, 26, 71, 143]. Nab3 and Nrd1 heterodimerize and bind specific sequences in target RNAs, as well as interact with the C-terminal domain (CTD) of the large subunit of RNA polymerase II *via* Nrd1's CTD-interacting domain [123]. Upon binding a target RNA, Nrd1-Nab3 recruits Sen1, a helicase responsible for termination of elongation [41, 56, 69]. The terminated transcript can be further processed by the TRAMP complex and the nuclear exosome.

Prior work in our lab has established that a C-terminal segment of Nab3 is essential for viability and plays a role in termination in both a reporter assay and at endogenous targets of NNS regulation [27, 35]. This Nab3 domain is scored as prion-like by a computer algorithm trained on prion sequences from yeast [34]. The PrLD of Nab3 is of low complexity with a skewed overrepresentation of glutamine (Q; 27%) and proline (P; 17%). It is relatively unstructured but can assemble into authentic amyloid polymers with a characteristic cross-β architecture that further organize into hydrogels. A purified portion of the PrLD (C-terminal 134 residues), as well as full length Nab3, form amyloid filaments *in vitro* [27, 36]. Removal of even part of the PrLD from full length Nab3 abrogates this property of the protein [36]. A previously described derivative of the Nab3 PrLD in which glutamates are substituted for each of 24 glutamines, fails to polymerize *in vitro* and, when

included in full-length Nab3, fails to support cell viability [36]. As well, exchanging Nab3's PrLD for some, but not all, heterologous yeast PrLDs with amyloid polymerizing properties, can rescue cell viability and termination activity [37]. Thus, although a PrLD can polymerize in vitro, that property is not sufficient to rescue Nab3 function, leading to the idea that there are subcategories of PrLDs amongst RNA-binding proteins that govern protein activity. PrLDs are over-represented in RNA-binding proteins where they are important for supporting phase separation by their client protein as a means of assembling a subcellular compartment in which metabolic processes take place [77, 78]. Some of these subcellular compartments include the nucleolus and Cajal bodies in the nucleus, and P-bodies and stress granules in the cytoplasm, all of which are enriched in RNA-binding proteins and RNAs [103, 144]. Recently, RNA polymerase II and mediator have been shown to cluster in transcription-dependent phase-separated nuclear condensates in mammalian cells [145-149]. This suggests that many aspects of the biogenesis of mature RNAs from initiation to termination, may be governed by phase separation of the cognate machinery.

In addition to possessing a PrLD, other features of Nab3 are suggestive that it may operate *via* phase separation. Genetic and biochemical data support a model in which Nab3, and its dimerization partner Nrd1, function as multimers while binding with sequence-specificity to nascent RNA. The complex can bind RNA polymerase II through the CTD-interaction domain of Nrd1. These protein-protein and protein-nucleic acid interactions are important for the termination process [26, 27]. Thus, the function of the PrLD could serve as a device used to assemble a functional termination complex. The ability of some heterologous PrLDs to substitute for Nab3's PrLD in supporting the

protein's termination function, as well as cell viability [37], suggests that the essential role played by Nab3's PrLD is to facilitate the assembly of the termination factor into a specific polymeric complex. Indeed, Nab3 interacts with itself in living yeast when tested in a protein-protein interaction assay and in a yeast strain in which two different mutant *nab3* alleles can cross-complement [27, 36].

Changes in the transcriptome following glucose-deprivation are mediated in part by the NNS pathway. Here we exploit a recently discovered, glucose deprivation-induced, re-localization of Nab3 to a unique granule within the nucleus to learn if Nab3's PrLD is involved in this process [96]. We show that truncation or a reduction in the glutamine-richness of the Nab3 PrLD prevented the recruitment of Nab3 to the granule when subjected to the glucose starvation paradigm. Though necessary, the PrLD was not sufficient to be recruited to the granule. Further characterization of granule formation revealed that it is dynamic; time lapse microscopy using microfluidics showed that minutes after the re-addition of glucose to starved cells, Nab3 rapidly delocalizes and becomes distributed throughout the nucleus. A functional Nab3 chimera employing a heterologous PrLD in place of its own, also reorganizes into the granule following a glucose-deprivation challenge while another heterologous PrLD that does not rescue Nab3 function, fails to localize to granules when substituted for Nab3's own PrLD. These findings suggest that there is an important inducible rearrangement of Nab3 that resembles a phase separation event in which a new compartment can be reversibly assembled in yeast nuclei. This property relies upon on its essential PrLD.

4.3 MATERIALS AND METHODS

Plasmid construction

The pRS315GFPNab3 plasmid was generated (Genscript Inc.) by inserting eGFP after the start codon of the Nab3 open reading frame in pRS315Nab3. The pRS315GFPNab3Δ191 plasmid was generated from pRS315GFPNab3 by PCR deletion using oligonucleotides 5'–ACCATATGGCTGTTGCTG–3' and 5'–TAGACTCCCTTTTTTCAATCTTTTCCATTTCTTG-3'. The NdeI site in GFP was deleted from plasmid pRS315GFPNab3 by PCR mutagenesis using the oligonucleotide 5'–CCCGGATCACATGAAACGGC–3' and 5'-TATCTTGAAAAGCACTGAACACCATAA–3' to generate pRS315GFPNab3NdeI. The plasmids pRS315GFPNab3Sup35 and pRS315GFPNab3Rat1 were constructed by releasing part of the PrLD coding sequence from pRS315GFPNab3NdeI with *Nde*I and *Xho*I followed by insertion of the heterologous PrLDs released from *Nde*I and *Xho*I cut pRS315Nab3Sup35 and PRS315Nab3Rat1, respectively. The plasmid pRS315-GFP-PrLD was created by PCR mutagenesis with the oligonucleotides 5'-AATTATGGGATGGGATGCCACCACC-3' and 5'-GCAGCCGGATCCTTTGT-3'. pRS315yomKate2-Nab3 incorporated yeast-optimized mKate2 sequences [150] in frame at the N-terminus of Nab3 (Genscript, Inc.). The plasmid pRS315-GFPNab3ΔRRM was generated by PCR-based deletion using the oligonucleotides 5'-GACAACCCTCAAAGCGTTAG-3' and 5'-ATTGTGCATCTCGGTGATTTT-3' and the template pRS315-GFPNab3NdeI.

Yeast strains

Yeast strains used in this study are presented in Table 4-1. Cells were grown in rich medium (YPD), synthetic medium (SC), or standard selective drop-out medium [SC ura$^-$, SC leu$^-$, or SC ura$^-$ leu$^-$; [132]] at 30°C unless otherwise indicated. Plasmids were transformed into yeast strains using the lithium acetate method [133]. DY3111 (Loya et al. 2017) was transformed with pRS315GFPNab3 and pRS315-Nab3Δ191 to yield yeast strains DY3205 and DY3183, respectively. DY3205 was cured of pRS316Nab3 by selection on 5-fluroorotic acid to create DY3206. HTB2 tagged with mCherry and marked with *HIS4* was amplified from genomic DNA of yeast strain YOL890 [gift from Dr. S. Wente] using 5'-GCAGCTGAACCAGCTTTACC-3' and 5'-GCTTTCAGTCGAAAACAGC–3'. The PCR product was transformed into DY3206 using high efficiency lithium acetate transformation [151], transformants were isolated by growth on SC leu$^-$ his$^-$ media and verified by PCR using 5'-GGGAATGTTAAACCAGCTTTAGC-3' and 5'-GCTTTCAGTCGAAAACAGC–3' creating strain DY3228.

DY3228 was transformed with pRS316Nab3 to create DY3238 and grown in SC ura$^-$ to encourage loss of pRS315GFPNab3. A derivative lacking this *LEU2*-marked plasmid was isolated yielding strain DY32359. DY32359 was transformed with pRS315GFPNab3, pRS315GFPNab3Δ191, pRS315GFPNab3Sup35, or pRS315Nab3Rat1 to create DY32309, DY3233, DY4524, and DY4525, respectively.

DY4540 was generated by transforming a strain containing chromosomal GFP-tagged *NIC96* (ThermoFisher) with pRS315-yomkate2-Nab3. DY4538 was generated by transforming a strain deleted for *BTN2* (courtesy of A. Corbett, Emory U.) with pRS315-

GFP-Nab3. DY4546 was generated by transforming DY32359 with pRS315 Nab3Q-->E [36]. DY4549 was generated by transforming DY32359 with pRS315-GFP-PrLD. DY4551 was generated by transforming DY32359 with pRS315-GFPNab3ΔRRM.

For treatment with hexanediol, glucose starved cells were adjusted to 5% 1,6 hexanediol and 10μg/ml digitonin as described [152]. Cells were incubated at 30°C for 30 min before confocal imaging.

Western blotting

Cells were grown in SC ura$^-$ leu$^-$ with or without glucose, as indicated. Cycloheximide (0.1%, Sigma-Aldrich was added where indicated to a final concentration of 125μM for the indicated times. Cells were collected, washed with water, and boiled for 5 min in sample buffer [136], before being resolved on a 6% SDS-polyacrylamide gel. Proteins were electrophoretically transferred to nitrocellulose blocked with 5% nonfat dry milk [137], and probed with an antibody against Nab3 [2F12-2; Dr. M. Swanson [11]]. Signal was detected using horse radish peroxidase-conjugated anti-mouse IgG (Sigma Chemical Co.) and enhanced chemiluminescence was performed in 100mM Tris pH 8.5, 1.25 mM luminol, 200nM p-coumaric acid, and 0.01% H_2O_2.

Confocal Microscopy

Cells were grown overnight to mid-log phase in SC ura$^-$leu$^-$glucose media, washed three times in appropriate media for the experiment, and resuspended in the wash media.

After 120 min of incubation, cells were deposited on 1.5% agarose pads. Confocal images were taken at room temperature using a confocal microscope (SP8; Leica) using a HC PL APO 63X 1.40 NA oil immersion lens, WD 0.14mm (Leica). Images were acquired using Leica Application Suite X 3.0.2.16120 (Leica). Representative cells are shown as maximum intensity projections or as single Z-planes as indicated from experiments repeated using three biological replicates. Images were processed using FIJI [153]. For counts of granules/cell, images of at least 100 cells expressing GFP-Nab3 or its variants, were coded and counted by a blinded scorer. The number of cells with granules divided by the total number of cells with detectable pan-nuclear staining of that Nab3 derivative were multiplied by 100 to give the % with granules. Statistical significance was assessed using GraphPad Prism (v7) software and Fisher's exact test.

Microfluidics

Cultures were grown in SC ura$^-$ leu$^-$ to an OD_{600} between 0.2-0.6 at 30°C. Live-cell microfluidics experiments were performed using an IX83 (Olympus, Waltham MA) microscope with a Prime 95B CMOS Camera (Photometrics). Fluorescence and Differential Interference Contrast (DIC) images were acquired using an Olympus-APON-60X-TIRF objective. Five 290nm Z-stacks of GFP and RFP images were acquired using 3% intensity light from an Xcite 120 LEDBoost (Excelitas) with 35 ms and 50 ms exposure, respectively. Cells were imaged in a microfluidic device based on the Dial-a-wave design that allows for the rapid switching of media while holding the yeast in place [154, 155]. Glucose addition and removal was verified using AlexaFluor 647 dye (Life Technologies) present only in glucose-free media, imaged at 3% intensity for 100ms and

1 Z-slice. Cells were exposed to glucose for 1 hour, starved of glucose for 2 hours, and reintroduced to glucose in order to observe formation and dissolution of granules. Images were deconvolved using Huygens Software (Scientific Volume Imaging, Hilversum, Netherlands) Classic Maximum Likelihood Estimation (CMLE) Deconvolution Algorithm with a signal to noise ratio of 5. Masks of nuclei were made using ImageJ (National Institute of Health, Bethesda, MD) and data analysis was performed using MATLAB (MathWorks, Natick, MA).

Quantification of Granule Formation

To quantify the fraction of granule formation, fluorescent intensities of the nucleus were determined for single cells through all time points. Fluorescent values were normalized to 8-bit scaling (0-255). The maximum and minimum fluorescent intensities for each cell were found through the time course for histogram normalization. Histograms were created from 0-255 with a bin width of five and normalized to sum to one. An average histogram was calculated for each cell during the first hour to serve as a baseline reading. Using MATLAB's' 'fit' function, a Gaussian curve was fit to each average histogram to acquire the mean of the histogram and the standard deviation. The MATLAB 'fit' function was used with 'gauss1' fit type and a 'NonlinearLeastSquares' method. Lower limits for the mean of the Gaussian were set to zero, as the fluorescence values cannot be negative. Granule formation involves an increase in fluorescence intensity from the baseline reading, as there is a higher concentration of GFP-Nab3 in each pixel. To measure the increased fluorescence derived from granule formation, we quantified the fraction of the fluorescence that was present in bins three standard deviations above the mean, as seen in Fig. 4-6. We

call this value the “spot fraction” and it represents the amount of fluorescence which registers as being in a granule, or other spot, by this method. We scored cells as positive for granule formation if more than two thirds of the total granule fraction measured occurred during glucose starvation.

4.4 RESULTS

Nab3 localization to a granule is PrLD-dependent

Glucose restriction is commonly used to provoke the assembly of inducible compartments such as stress granules. In response to glucose starvation, Nrd1 and Nab3 shift from a diffuse pan-nuclear distribution to a concentrated perinuclear granule of unknown function [96]. Numerous studies have shown that low complexity PrLDs are often critical for proper localization of RNA-binding proteins to granules involved in the stress response, protein homeostasis, RNA metabolism, and development [95, 105, 156]. Thus, we asked: Is Nab3's well-studied PrLD responsible for its ability to re-localize in response to carbon source restriction? To test this model, a series of plasmid-borne, N-terminally GFP-tagged Nab3 variants were introduced into cells expressing untagged Nab3. Nab3 variants were supplied on a *CEN* plasmid and expressed from their endogenous promoter in a strain deleted for chromosomal *NAB3*. To mark the nucleus, strains also included a chromosomal copy of mCherry-tagged histone H2B [157]. Cells were grown overnight to mid-log phase and washed free of glucose. As a control, an aliquot of cells was resuspended in fresh glucose-containing medium. After two hours, cells were imaged *via* confocal microscopy to assess the distribution of fluorescent protein. GFP-tagged wild type Nab3 exhibited a pan-nuclear distribution in glucose rich conditions. In glucose starved cells, Nab3 localized to a granule in roughly 20% of the population, consistent with a previous report. [Fig. 4-1, Fig. 4-2, and Fig. 4-S1]. This is similar to the fraction of cells that acquire cytoplasmic stress granules following glucose-restriction [158, 159]. If starvation-induced recruitment of Nab3 to granules is PrLD-dependent, there should be a loss of recruitment of GFP-Nab3 to the granule when the PrLD is deleted.

Indeed, this is what was observed when the C-terminal 191 amino acids were removed from GFP-Nab3 ("nab3Δ191", Figs. 4-2 and 4-S1).

We previously characterized an allele of Nab3 in which 24 Q residues distributed across its C-terminal PrLD, were substituted with glutamate [36], reducing the domain's Q content from 27% to 17%. Collectively these substitutions abrogate *in vitro* amyloid formation by a purified recombinant portion of the domain and cells with this allele as the sole source of Nab3 are not viable [36]. When expressed with an N-terminal GFP-tag, this PrLD-mutant Nab3 showed a strongly reduced ability to form perinuclear granules following glucose restriction (Fig. 4-2). This more targeted alteration of the PrLD (in contrast to deletion of 191 amino acids) confirms a role for the domain in nuclear granule formation.

Nab3 contains a canonical RRM through which it binds nascent RNA during termination [11, 25, 28]. Mutations in the domain impair termination and cell growth. The yeast protein Whi3, which binds RNA through its RRM and contains a Q-rich PrLD, its interaction with RNA is important for liquid-like condensate formation in cells[160, 161]. To test if Nab3's RRM is needed for formation of nuclear granules, a deletion derivative (amino acids 326-371) was generated that lacks its two consensus RNP-motifs. The GFP-variant of this Nab3 derivative was severely reduced in its ability to form nuclear granules following glucose-starvation and was lethal when the only Nab3 present. (Figs. 4-2 and 4-S2). This suggests that, as observed for Whi3, RNA-binding is important for Nab3's recruitment into granules.

The sequence of Nab3's PrLD, its propensity to homopolymerize *in vitro*, and the dependence of the protein upon its PrLD to localize to nuclear granules, suggests the

granule is a dynamic, liquid-like assembly similar to P-bodies. The aliphatic alcohol 1,6-hexanediol has been used as a probe for such assemblies due to its ability to dissolve them; in contrast, solid structures such as protein aggregates and the cytoskeleton resist solubilization [162, 163]. Cells were glucose starved to form nuclear granules and treated for 30 min with 5% 1,6 hexanediol. This caused a virtually complete loss of granules from cells (Fig. 4-2), similar to that seen for yeast P-bodies [163].

Since the PrLD sequence isolated from Nab3 can assemble into polymers *in vitro* [36], we asked if a GFP-tagged version of the PrLD itself (GFP-PrLD) could be recruited to nuclear granules. Under glucose-rich conditions, GFP-PrLD was distributed broadly in the cytoplasm and nuclei of cells (Fig. 4-S3). When glucose was withdrawn, the protein formed cytoplasmic aggregates, perhaps at cytoplasmic bodies such as stress granules (Fig. 4-S3). These results indicate that the Nab3 PrLD is not sufficient for recruitment to nuclear granules although surprisingly, the PrLD alone was responsive to glucose starvation in forming a cytoplasmic granule.

Characteristics of the glucose starvation-induced Nab3 granule

Corden and co-workers determined that the Nab3-containing granule is distinct from three well characterized ribonucleoprotein structures in yeast: stress granules, P-bodies, and the nucleolus [96]. Other cellular granules are thought to be locations in which substrates are collected for degradation, including insoluble protein deposit (IPOD) and juxtanuclear quality control/intranuclear quality control (JUNQ/INQ) deposits [80, 164]. However, IPOD is a cytoplasmic and perivacuolar depot for terminally segregating

proteins, and therefore cannot be the Nab3 granule. JUNQ/INQ has been characterized as a perinuclear granule and in some studies as a nuclear granule in which proteins are prepared for degradation [80]. To test if glucose starvation prompted degradation of Nab3, its levels were measured by western blots of cells treated with cycloheximide to prevent new Nab3 synthesis during glucose deprivation. Nab3 levels were not diminished when glucose was removed, and cells were treated with cycloheximide (Fig. 4-S4, lane 3 vs lane 4), implying that Nab3's recruitment to a granule is not a prelude to Nab3 degradation and is not a site of Nab3 turnover. To further explore if the Nab3 nuclear granule was the same as the INQ body, we exploited the finding that INQ formation is dependent upon *BTN2* [80]. Yeast deleted for *BTN2* were transformed with a *LEU2*-marked plasmid expressing GFP-Nab3 and the resulting strain was glucose starved and scored for Nab3 granule formation. *BTN2* null cells could efficiently form Nab3-positive granules (Fig. 4-2). This suggests that the conditional Nab3-containing nuclear granule is not INQ. [81].

Previous studies described Nab3 granules as perinuclear but did not formally test which side of the nuclear membrane the granule was on [96]. To resolve this, we exploited a strain containing a chromosomal copy of GFP-tagged *NIC96*, a linker nucleoporin known to associate with the nuclear pore complex basket and other nucleoporins [165]. Since Nic96 is a component of the nuclear pore, it provided a definitive demarcation of the nuclear periphery [80, 166], allowing assessment of where Nab3 granules reside. This strain was transformed with a plasmid encoding yomKate2-tagged Nab3 and cells were glucose-starved. Images of single Z-axis focal planes across multiple nuclei clearly showed that the Nab3 granule resided within the nucleus as defined by the Nic96 ring, identifying the structure as unequivocally intranuclear (Fig. 4-3).

Resuspending cells in glucose-free medium could represent an osmotic shock [158]. To control for the possibility that such a hypoosmotic challenge is responsible for inducing nuclear granule formation, we substituted equimolar sorbitol for glucose following washout of the latter. Sorbitol serves as an osmotically active solute in growth media, it serves to mimic the osmotic environment of glucose-containing media without serving as a metabolizable carbon source for some yeast [158, 167]. In the strain studied here, sorbitol did not support growth while galactose, sucrose, and fructose could (Fig. 4-S5A). After 2 hours in sorbitol, nuclear granules were observed as seen above for cells washed out of glucose without a sugar replacement (Figs. 4-4A). The fraction of cells displaying granules was reduced by half relative to the no glucose condition (Fig. 4-2), suggesting there could be a partial osmotic component to glucose withdrawal. A similar result of a combined osmotic and metabolic basis for cytoplasmic stress granule formation was observed previously when glucose was replaced with sorbitol [158].

Quantification and dynamics of the Nab3 nuclear granule

Examination of the dynamics of granule formation and relaxation requires the ability to follow a single cell over time while altering the extracellular glucose concentration. To this end, we employed a microfluidics chamber that holds yeast cells stationary while allowing rapid switching of extracellular media [154]. In this way, single live cells can be followed by differential interference contrast bright-field and fluorescence microscopy to monitor Nab3 dynamics before, during, and after glucose starvation. Consistent with the confocal microscopy finding (Figs. 4-1 and 4-2), a subset of cells in the chamber formed readily observable granules in strains with GFP-tagged wild type

NAB3 following glucose-restriction (Fig. 4-5A, white arrowheads and Supplemental Movie SM1). In comparison, nab3Δ191 did not form a similar granule in glucose-free media (Fig. 4-5B). To perform an unbiased quantification of granule formation and dynamics, we developed an algorithm based on histogram analysis to quantify the fraction of nuclear fluorescence present in spots. In short, due to its accumulation in a granule, the Nab3 signal is more focused than soluble, nucleoplasmic Nab3. As a baseline, the mean fluorescence intensity peak was calculated during the one hour of growth in glucose prior to starvation for each cell (blue lines, Fig. 4-6A). Pixels with higher Nab3 accumulation show up as a rightward shift in the scaled fluorescent intensity value. We found that the granules contained pixels with fluorescence three standard deviations above the mean. We define the fraction of pixels at or above this threshold as the "spot fraction" (Fig. 4-6A). Using the spot fraction as a metric, cells with GFP-Nab3 (n=114) were compared to those with GFP-nab3Δ191 (n=109) as a function of glucose starvation (Fig. 4-6B, shaded area). The average of all cells positive for a spot fraction for both strains are plotted versus time (Fig. 4-6B). This stringent criterion showed that distinct granule-containing and granule-free cells could be resolved (Figs. 4-6A and 4-6B). Since individual cells could be identified and followed over time, we found that 18% of cells with wild type Nab3 cells formed a stable granule during starvation, consistent with the results seen above using confocal microscopy. Cells that displayed granules kept their granules for the duration of glucose deprivation and those that did not display granules failed to do so for the duration. Strikingly, following the re-addition of glucose, Nab3 granule intensity shifted back to a pan-nuclear baseline distribution within minutes, showing the rapidity of the response to the re-introduction of glucose (Fig. 4-6B and Supplemental Movie SM1). This is the first

demonstration of the reversible nature of the Nab3-containing granule. Importantly, Nab3 lacking its PrLD did not form stable granules (Fig. 4-6B and Supplemental Movie SM2), recapitulating the observations from confocal microscopy. Once formed in glucose-free conditions, Nab3 granules appeared relatively fixed in location, typically at the nuclear periphery, and persisted until cells were re-fed with glucose (Fig. 4-6B). In contrast, nab3Δ191 formed smaller, lower intensity puncta in 10% of cells. (Fig. 4-6B and 6C). These spots were distinct from granules formed by wild type Nab3 in that they were dim, short lived, and distributed broadly throughout the nucleus, as opposed to the typical, strongly fluorescent, more localized granules seen with GFP-Nab3 (Fig. 4-6C). These granules were not dependent upon exposure to glucose-free media. This again indicates that the PrLD is important for the localization behavior of Nab3 in response to glucose restriction.

Since granule formation correlated with metabolizable or non-metabolizable in the media, we examined if granule dissolution was also related to sugar 'usability'. Cells were grown in glucose, washed free of sugar, and incubated at 30°C for 2 hrs to allow granule formation. The culture was then split and cells were re-fed with different sugars. Cells that were re-fed with the metabolizable sugars glucose, fructose, or sucrose, lost their granules after 10 min (Fig. 4-4B). In contrast, the population re-fed with non-metabolizable sorbitol, which did not support growth (Fig. 4-S5B), retained about half their granules (Fig. 4-4B). Interestingly, cells washed out of glucose and re-fed with galactose showed an unusually long lag period (> 10 hrs) before resuming cell division (Fig. 4-S5B). A more complete loss of granules required a long incubation with galactose (Fig. 4-5B) consistent with the extended time it took to resume cell proliferation in the presence of this

sugar (Fig. 4-S5). This correlation between the resumption of growth and magnitude of loss of granules, implies that consumption of the carbon source is important for reversing the effect of glucose-starvation on Nab3 condensation.

We next asked if overall cell viability is compromised following glucose restriction by assaying for colony forming units after 2 hrs without glucose. Cells were washed into glucose-free medium and plated on standard solid medium immediately or after 2 hrs of incubation at 30°C in the continued absence of glucose. Plates were incubated overnight and colony forming units were counted. A similar level of viability was observed for cells before and after glucose starvation indicating there is no deleterious effect of this restrictive regimen (Fig. 4-S6).

Heterologous PrLDs facilitate Nab3 puncta formation

We previously showed that some, but not all, heterologous yeast PrLDs can substitute for Nab3's PrLD thereby conferring viability and termination function onto this transcription factor [37]. We tested whether chimeric Nab3 proteins would re-localize to granules as wild type Nab3 does under glucose starvation. Strains containing untagged, wild type Nab3 were transformed with plasmids expressing GFP-Nab3Rat1, a chimera that does not support cell viability, or GFP-Nab3Sup35, a viable chimera. The chimeric proteins were well-expressed in these strains, though at somewhat reduced levels *versus* wildtype untagged Nab3 (Fig. 4-S2). GFP-Nab3Rat1-containing cells exhibited no GFP-positive granules following glucose restriction (Fig. 7, right), while GFP-Nab3Sup35 strains were positive for granules similar to cells with wild type GFP-Nab3 (Fig. 4-7, left).

The granules formed from GFP-Nab3Sup35 were somewhat different from those assembled from GFP-Nab3 with a higher frequency of multiple, lobulated, or fragmented granules apparent when compared to GFP-Nab3.

GFP-Nab3Sup35 exhibited a qualitatively different puncta compared to wild type Nab3. Is this the result of Nab3Sup35 localizing to a different granule than Nab3 or does it simply et localized to other granules as well? To test this a strain containing GFP-Nab3Sup35 as well as yomKate2-Nab3 was imaged after starvation to assess localization of both proteins (4-S7).

4.5 DISCUSSION

Evidence is accumulating in favor of a model in which RNA metabolism takes place in compartments assembled through liquid-liquid phase separation by RNA-binding proteins that contain low complexity, prion-like domains [77, 78, 105]. Many cytoplasmic granules are inducible such as P-bodies and stress granules. Here we provide evidence that a compartment that harbors the hnRNP-like Nab3 termination factor, is likely to employ such a mechanism in the nucleus. The functional consequence of this subnuclear granule is not clear, although its inducible assembly serves as a valuable paradigm to study the rearrangement of these RNA binding proteins.

The Nab3-Nrd1-containing granule was first identified [96] as a novel structure in or near the nucleus that resembled an RNA quality control site seen in cells with mutated RNA processing factors [168]. Here we observe the reversible induction of Nab3 granules in a PrLD-dependent manner. The well-established homopolymerization properties of this essential domain suggest that the Nrd1-Nab3 termination machinery forms a functionally important compartment for RNA metabolism. This process may be akin to 'transcription factories' at which one or more transcribed genes congregate and where the many steps of pre-mRNA synthesis may take place in a concerted fashion [169]. In metazoan cells, other nuclear compartments that perform a more limited task are so-called nuclear speckles which are enriched for splicing factors [85] or the speckles formed by the elongation factor P-TEFb, a kinase that associates with RNA polymerase II [146]. The cyclin T1 subunit of P-TEFb employs a low complexity domain to associate with RNA polymerase II through a phase separation mechanism [146]. Along this line of thinking, the Nab3-Nrd1 granule may form through phase separation at sites in need of its termination activity.

An alternative possibility is that the Nab3 granule may be a location for the storage of inactive proteins that can then reorganize when a physiological stimulus is detected, in this case, the return of sugar to the growth medium. One possibility is that the granule forms at a specific stage of the cell cycle. In any event, we postulate that detection of the granule after glucose starvation is a special case of Nab3 assembly, and that during normal vegetative growth, Nab3 and Nrd1 regularly carry out routine transcription termination as a higher order assembly in an RNA polymerase II-containing termination complex. The form they take during vegetative growth would be submicroscopic. This idea is supported by evidence of the self-assembly property of Nab3 *in vivo* and *in vitro* [27, 36, 38] and is derived from an early model that multiple Nab3-Nrd1 complexes are loaded onto nascent RNA [26]. It is possible that Nrd1 also forms oligomers with itself during termination, as it bears a low complexity domain [34], a recombinant version of which forms amyloid filaments [36]. These two proteins also dimerize efficiently through domains outside of their PrLDs [26], leading to the possibility of large Nab3-Nrd1 co-polymers.

The Nab3 rearrangement noted here was detected by confocal microscopy as well as differential interference contrast microscopy using flow chambers in which single cells can be observed over time. The dynamic Nab3-Nrd1 granules are similar to the glucose deprivation-induced distribution of cytoplasmic stress granules and processing-bodies seen in the yeast cytoplasm [95, 158, 159]. Interestingly, and in contrast to Nab3, yeast Hrp1 migrates from the nucleus to cytoplasmic stress granules following glucose starvation [95]. Hrp1 is a polyadenylation-coupled transcription termination factor that bears a PrLD and assembles into amyloid polymers *in vitro* [34, 37, 170]. Here we show the Nab3 granule is clearly intranuclear and that its formation is reversible. It is also notable that formation

of the granule takes hours whereas its dissolution takes minutes. The biochemical changes that regulate this process are an obvious next important question and are the focus of many studies into ribonucleoprotein granule assembly through liquid-liquid phase separation. Our results with 1,6 hexanediol are consistent with this mechanism. In this case, the stimulus for granule formation derives from the cell's access to a usable carbon source. Nutrient depletion and the Ras/protein kinase A pathway have been genetically linked to Nab3 and Nrd1 [96]. The biochemical details of how directly this signaling pathway is connected to granule formation remains to be defined.

Another important question is what other macromolecules may be harbored in this nuclear granule. It does not appear to be a location at which the Nab3 protein is degraded, such as that seen for proteins that aggregate in the JNQ and INQ compartments [80, 81]. If it is a site of active termination, then we predict RNA polymerase II and nascent RNA will co-localize with Nab3. Resolving the role of this cellular compartment and how its assembly is regulated will require further examination.

FUNDING

Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under award number [R01GM120271] to D.R. and R15 [GM128026] to J.K. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Funding for open access charge: National Institutes of Health.

CONFLICT OF INTEREST

The authors declare there are no conflicts of interest.

ACKNOWLEDGEMENT

The authors thank Drs. Susan Wente (Vanderbilt University) and Anita Corbett (Emory University) for yeast strains and Dr. Graeme Conn for critical reading of the manuscript.

This research project was supported in part by the Emory University Integrated Cellular Imaging Core. We thank Laura Fox-Goharioon for assistance with microscopy.

Figure 4-1

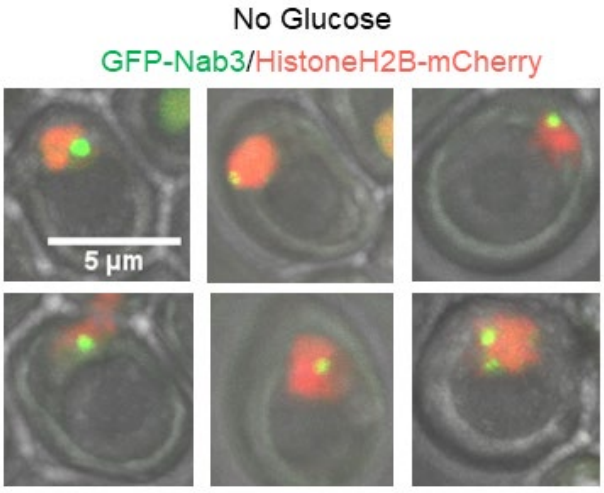


Fig. 4-1. Re-localization of Nab3 to a granule during glucose starvation. *S. cerevisiae* containing an integrated C-terminal mCherry-tagged histone H2B were transformed with an expression vector plasmid containing N-terminal GFP-tagged Nab3 (DY32309). Cells were grown to log phase and washed into SC ura$^-$ leu$^-$ glucose-free medium for two hours. Cells were placed on 1.5% agarose pads dissolved in SC ura$^-$ leu$^-$ glucose-free media and Z-stacks were imaged using a confocal microscope. Maximum intensity projections of representative granule-containing cells are shown.

Figure 4-2

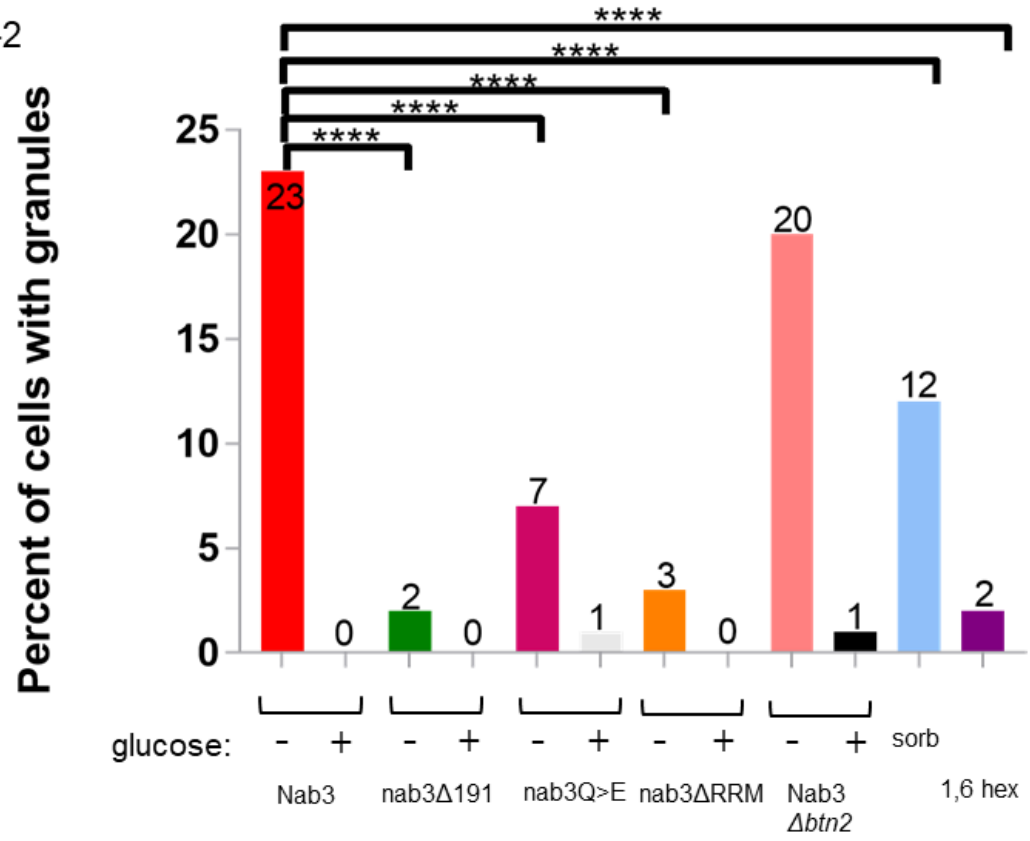


Fig. 4-2. Nab3's PrLD is a key driver of granule recruitment. *S. cerevisiae* strains were transformed with GFP tagged versions of Nab3 (DY32309), Nab3 lacking its C-terminal 191 residues, (Nab3Δ191; DY3233), Nab3 with a Q to E substituted PrLD (DY4546), or Nab3 deleted for its RRM motif (DY4551; ΔRRM). DY4538 had plasmid-borne GFP-Nab3 transformed into a *BTN2* deletion strain. Cells were washed into either glucose-free or glucose-containing SC leu- media, incubated at 30°C for 2 hrs, and Z-stacks were imaged using confocal microscopy. For hexanediol treatment, DY32309 was induced to form granules and incubated with the alcohol as described in Materials and Methods. Maximum intensity projections were created in FIJI and cells were binned into granule-containing or granule-free bins. The percent of cells containing granules

were compared across groups under similar growth conditions using Fisher's exact test. P-values of <0.0001 are indicated by ****.

Figure 4-3

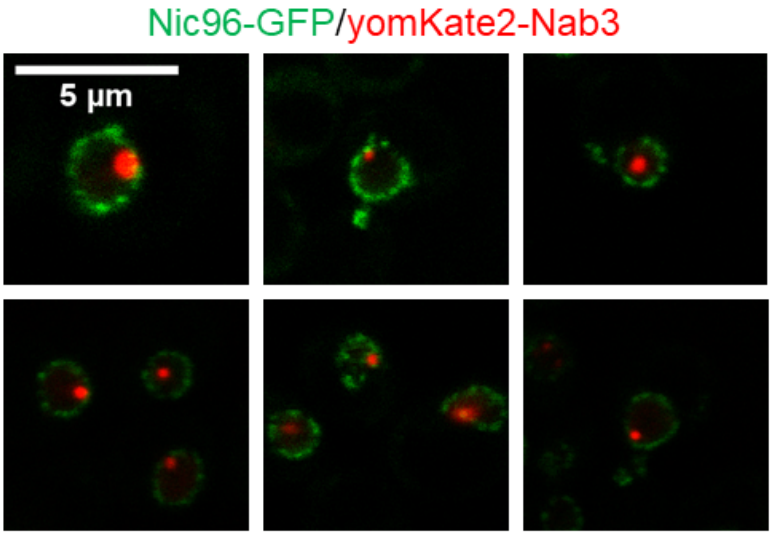


Fig. 4-3. Establishing the cellular location of the Nab3 granule. A yeast strain containing an integrated, GFP-tagged *NIC96* was transformed with an expression vector containing yomKate2-taggedNab3 (DY4538). Cells were washed into SC ura$^-$ leu$^-$ glucose-free medium incubated at 30 degrees C for 2hr and imaged using a confocal microscope. A single Z-plane of representative granule-containing cells is shown

Figure 4-4

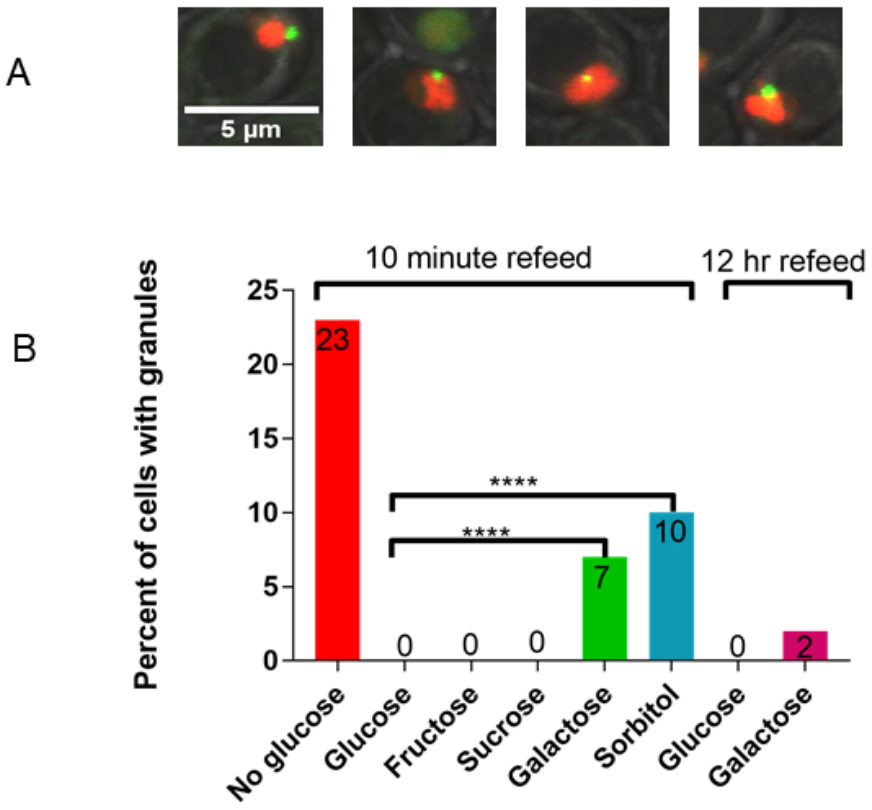


Fig. 4-4. Sorbitol substitution for glucose yields Nab3-granules and does not efficiently reverse granules obtained after glucose starvation. A) DY32309 (GFP-Nab3) cells were grown in glucose, washed, incubated for 2 hrs at 30ºC with sorbitol, and imaged by confocal microscopy. B) A culture of DY32309 was grown in glucose-containing medium, washed, and incubated for 18 hrs at 30°C in SC ura-leu- lacking glucose. Cells were washed, equivalent numbers were inoculated into media with the indicated sugars (2%), incubated at 30°C for 10 min or 12 hrs, imaged by confocal microscopy, and the percent of granule-containing cells were quantified. The P-values for a Fisher's exact test of <0.0001 are indicated by ****.

Figure 4-5

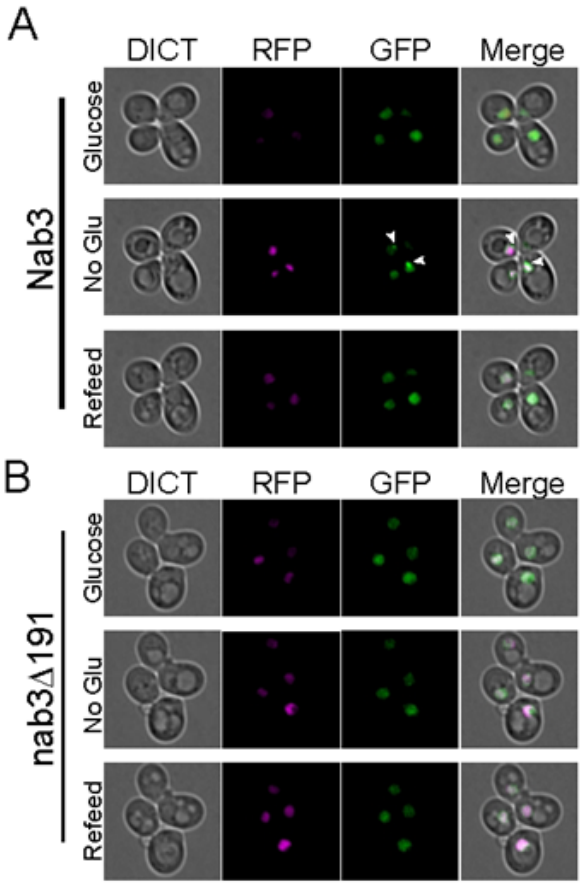


Fig. 4-5. Dynamic recruitment and dissociation of the Nab3 granule. Strains (A) DY32309 (GFP-Nab3) and (B) DY3233 (GFP-nab3Δ191; B) were examined in a microfluidics chamber using bright field, time-lapse microscopy prior to, during, and after glucose starvation. The nucleus is marked with mCherry-tagged histone H2B (RFP), Nab3 and nab3Δ191 are GFP-tagged (GFP). Images are single frames of maximum intensity projections from Z-stacks. Cells in glucose-containing medium show a pan-nuclear distribution of Nab3 (first row of A and B). After 2 hrs of glucose depletion the Nab3 signal condenses (white arrowheads) in wild type but not in nab3Δ191 (second row of A and B). Ten min following refeeding with glucose-containing medium, the Nab3 granule expands back into a pan-nuclear signal (third row) A and B. Time lapse video is available as Supplemental Material for both GFP-Nab3 and GFP-Nab3Δ191 expressing strains.

Figure 4-6

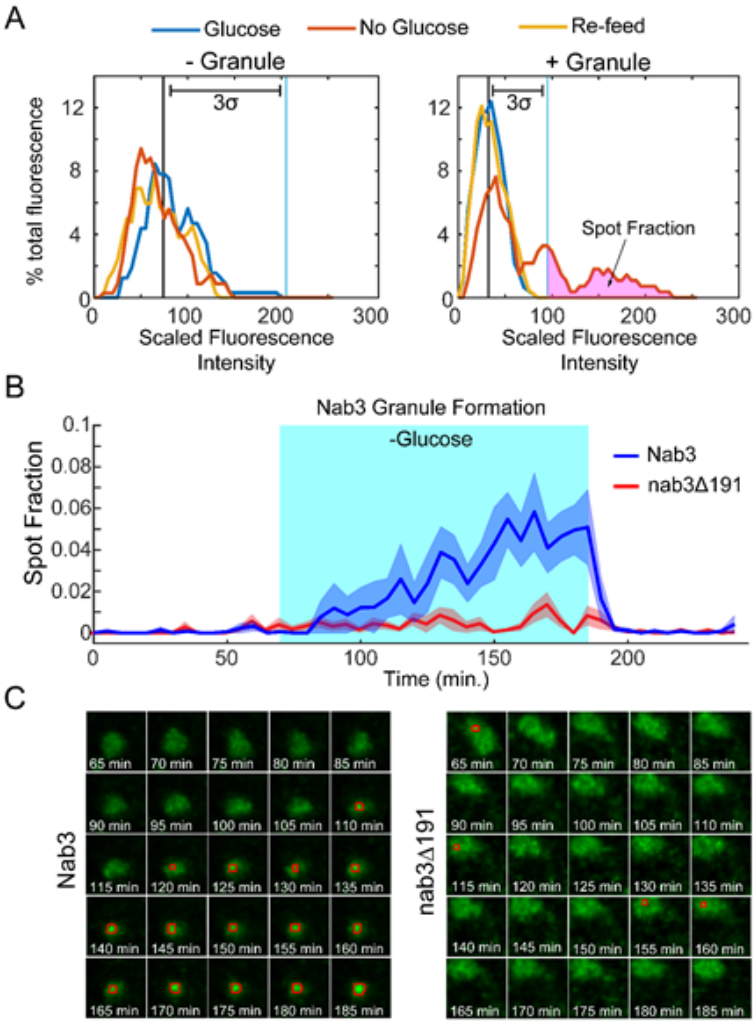


Fig 4-6. The Nab3 granule is reversible and dependent upon Nab3's PrLD. A) Illustrative histograms for a single cell scored as lacking (left) and a single cell scored as containing (right), a GFP-Nab3 granule. For each cell, fluorescence over 1 hr growth in glucose (blue peak) was used to calculate a baseline mean and standard deviation. Fluorescent intensity (scale of 0-255, bins of 5) is plotted versus % of fluorescence intensity each bin is of the total. The amount of fluorescence present more than 3σ above the mean was used to calculate the spot. B) Time course of accumulation of fluorescence in the spot fraction for GFP-Nab3 (blue) and GFP-nab3Δ191 (red). Cells were analyzed by real time fluorescence microscopy in a microfluidic device that allowed the

exchange of glucose-containing and glucose-free media. Cells were imaged for 60 min prior in glucose, then switched to no glucose for 2 hours (turquoise area), followed by refeeding with glucose for an hour. The average spot fractions (lines) and standard deviations (shaded areas) for GFP-Nab3 and GFP-nab3Δ191cells were calculated over the time course. C) Time course microscopy of a single nucleus from GFP-Nab3 (left) and GFP-nab3Δ191 (right) starting 5 min after glucose-free conditions (t=65'). Pixels with fluorescence >3σ for each cell are circled in red. Cells with Nab3 lacking the PrLD generally did not form granules, but, as seen in this example, those that scored positively show sporadic, small spots, in contrast to the robust and persistent spots formed in wild type Nab3.

Figure 4-7

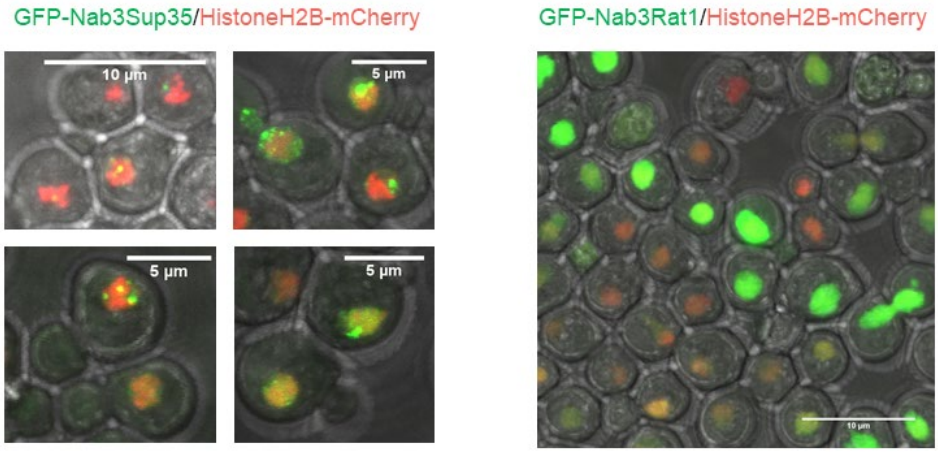


Fig. 4-7. GFP-Nab3Sup35 chimera forms granules following glucose-depletion. *S. cerevisiae* containing an integrated mCherry-tagged histone H2B gene were transformed with a plasmid containing N-terminal GFP-tagged Nab3Sup35 (left; DY4524) and Nab3Rat1 (right; DY4525). Cells were grown to log phase, washed into SC ura$^-$ leu$^-$ glucose-free media for 2 hrs, and Z-stacks were imaged using a confocal microscope. Maximum intensity projections of representative granule-containing cells are shown.

Figure 4-S1

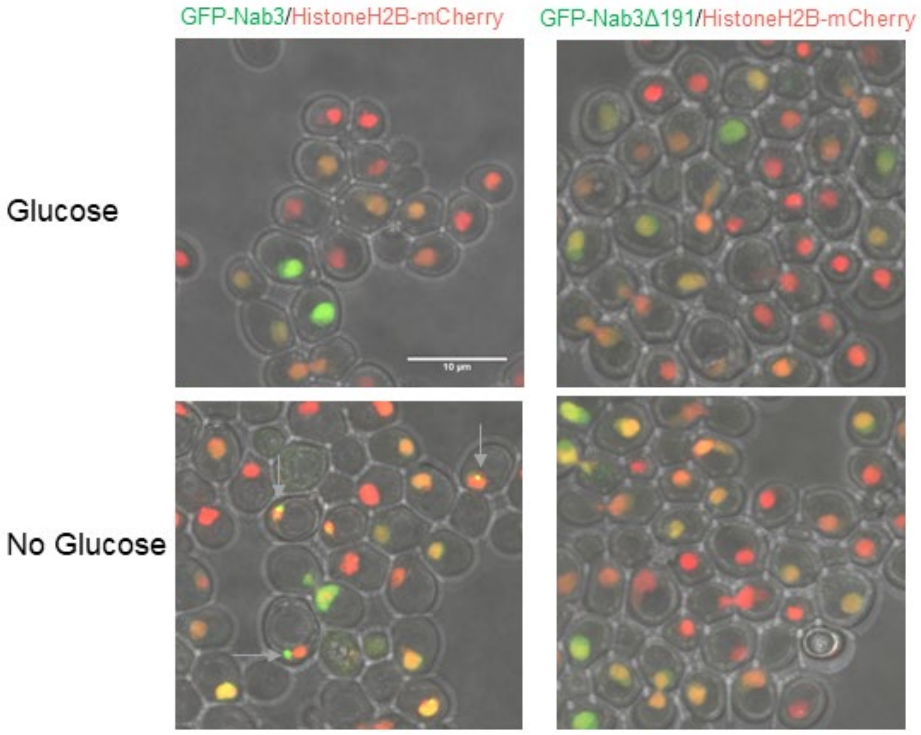


Fig. 4-S1. Re-localization of Nab3 to a granule during glucose starvation. *S. cerevisiae* containing an integrated mCherry-tagged histone H2B were transformed with an expression vector plasmid containing GFP-tagged Nab3 (DY32309), or GFP-tagged Nab3ΔLCD (DY3233). Cells were washed into SC ura$^-$ leu$^-$ glucose-free or glucose-containing medium, incubated at 30 degrees C for 2 hrs, and Z-stacks were imaged using confocal microscopy. Maximum intensity projections of representative fields of cells are shown for both strains under both conditions. GFP-Nab3 localized to a granule within the mCherry-tagged histone H2B field (bottom left, white arrows).

Figure 4-S2

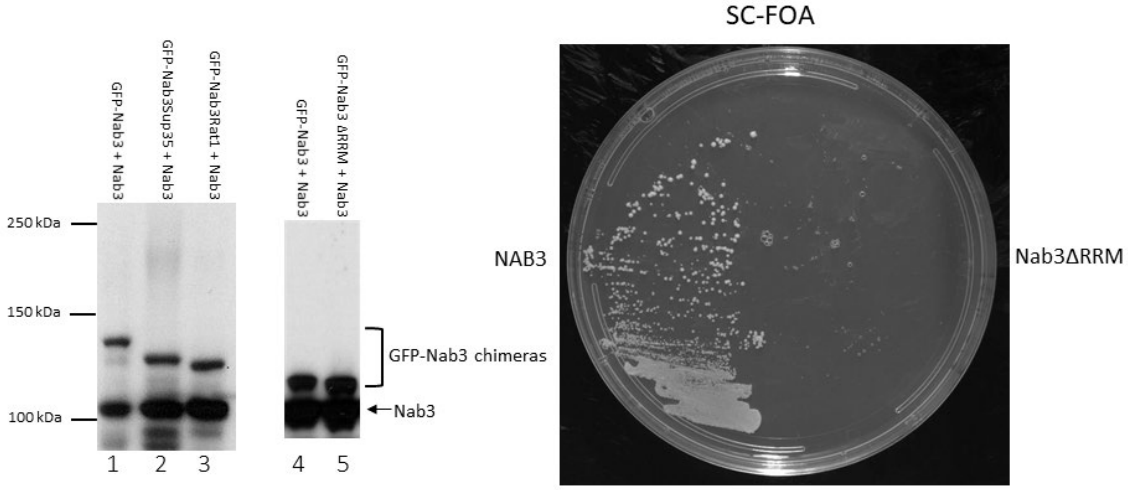


Fig. 4-S2. Expression of GFP-Nab3 and GFP-Nab3 chimeric proteins. Log phase cultures of DY32309, DY4524 and DY4525 were collected, lysed in boiling electrophoresis sample buffer and subjected to SDS-PAGE and western blotting using an anti-Nab3 antibody as described in Materials and Methods. Nab3ΔRRM is non-viable, right panel right side of plate.

Figure 4-S3

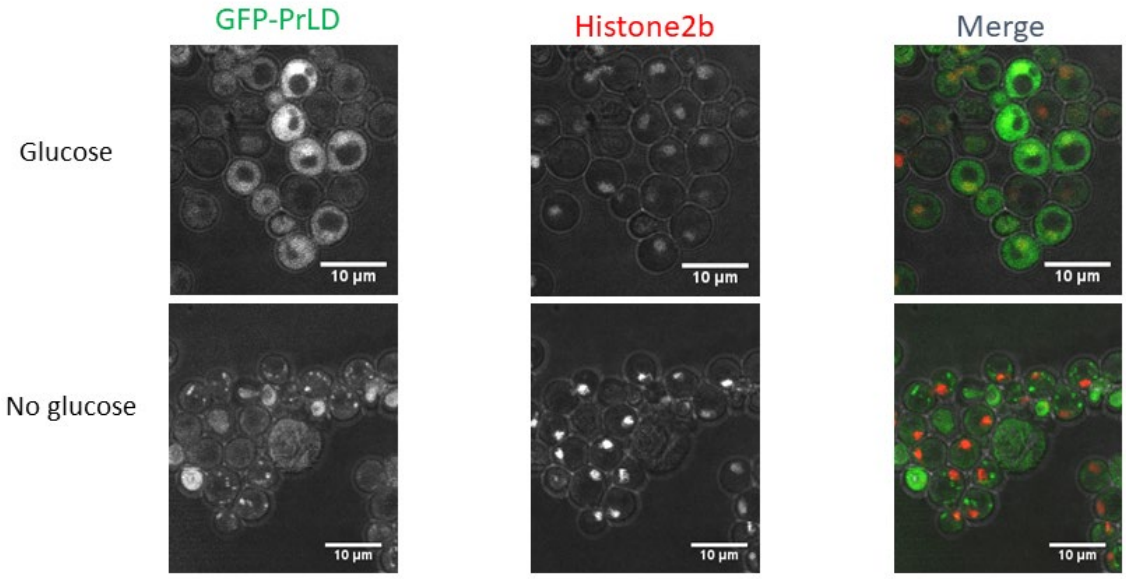


Fig. 4-S3. Nab3's 191-amino acid PrLD forms cytoplasmic granules after glucose restriction. *S. cerevisiae* containing an integrated mCherry-tagged histone H2B was transformed with a plasmid containing GFP fused to the 191-amino acid Nab3 PrLD (DY4549). Cells were washed into SC ura$^-$ leu$^-$ glucose-free medium, incubated at 30°C for 2hr and a single Z plane was imaged using confocal microscopy.

Figure 4-S4

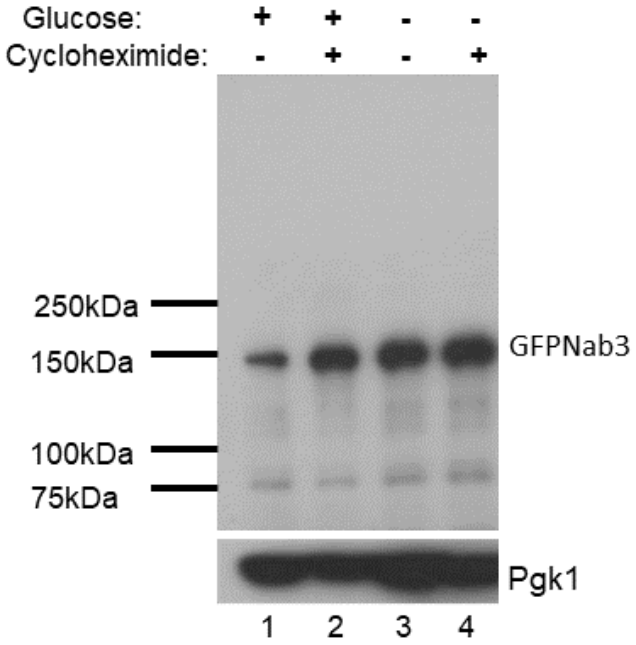


Fig. 4-S4. Nab3 levels persist during starvation. To assess the stability of Nab3 under the varying growth conditions, cycloheximide (125 μM,) was added as indicated to logarithmically growing DY32309 cells washed into or glucose-free media. Samples were taken immediately ("glucose") or after the medium exchange and 2.5 hours at 30°. Cells were lysed by boiling in electrophoresis sample buffer and subjected to SDS-PAGE. And western blotting using an anti-Nab3 monoclonal antibody as described in Materials and Methods.

Figure 4-S5

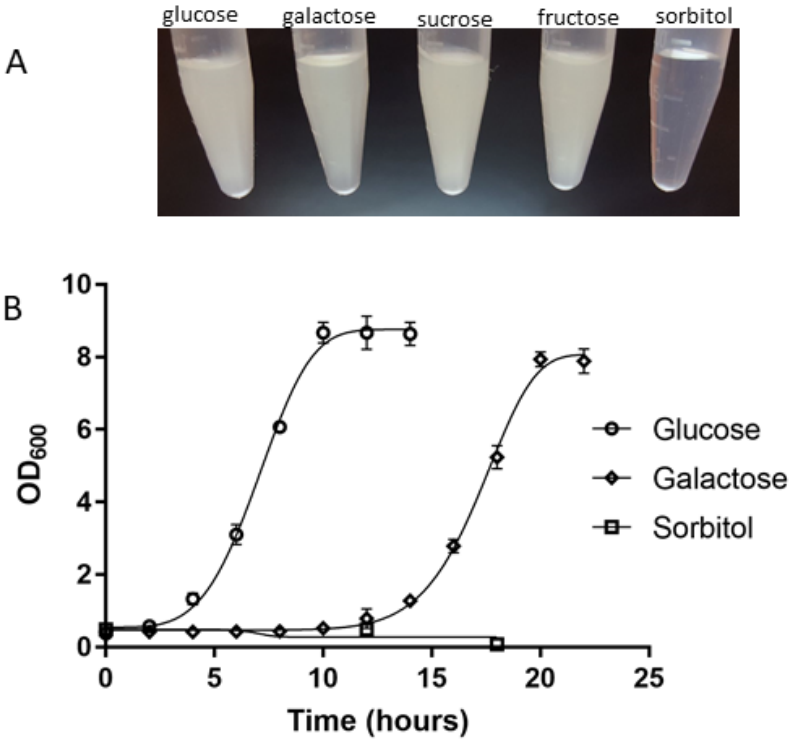


Fig. 4-S5. Growth of yeast on various sugars. A) DY32309 cultures were grown overnight to OD_{600} of approximately 0.4, washed free of glucose, and inoculated into media with 2% of the indicated sugar. Cultures were grown at 30ºC for 18 hrs before photographing. B) Cultures of DY32309 were grown in glucose-containing medium to OD_{600} of approximately 0.4, washed, and grown in glucose-free medium for 2 hrs at 30ºC. Equivalent numbers of cells were inoculated into glucose-, galactose-, or sorbitol (2%)-containing medium at returned to 30ºC. OD600 was scored in triplicate biological repeats. The mean and standard deviations were plotted.

Figure 4-S6

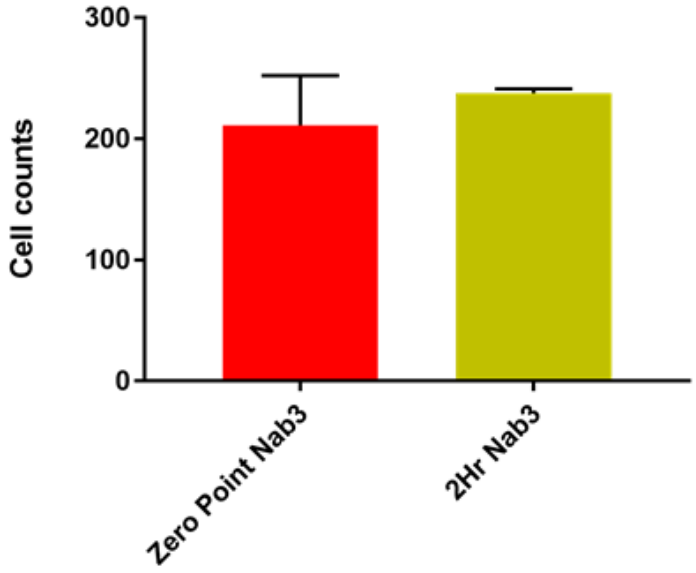


Fig. 4-S6. Survivability post glucose-depletion. Three separate cultures (biological replicates) were grown to log phase in SC ura$^-$ leu$^-$ with glucose and washed into glucose-free medium. Equivalent cell numbers were plated onto glucose-containing medium immediately or after 2 hrs in glucose-free medium at 30 degrees C. Colonies (viable cells) were counted and plotted. Error bars are the standard error of the mean (SEM). For "zero point Nab3" mean ± SEM = 215 ± 17.2, n=3, for "2 hour Nab3" mean ± SEM = 231.7 ± 6.0, n=3

Fig 4-S7

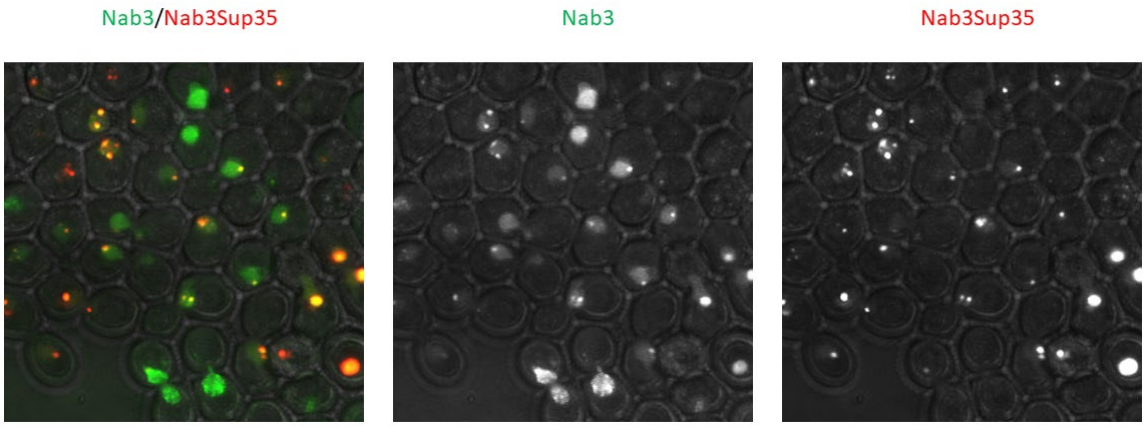


Fig. 4-S7 Nab3up35 is co-localized with Nab3 in a granule. *S. cerevisiae* were transformed with a *URA3* marked expression vector containing GFP -tagged Nab3Sup35 and a *LEU2* marked expression vector containing Kate2-tagged Nab3 (DY4553. Cells were grown to log phase overnight, starved of glucose for 2 hours, then z-stacks were captured. Nab3Sup35 co-localizes with Nab3 in granules.

Table 4-1: Strain Table

Strain	Genotype	Reference
DY3111	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3* (*URA3*)]	O'Rourke & Reines, 2016
DY3121	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 NRD1-TAP* [pREFGFPKan (*URA3*)]	This Study
DY3123	*MATa his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 NAB3-TAP* [pREFGFPKan (*URA3*)]	This Study
DY3183	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-NAB3 (*URA3*)]	Loya & Reines, 2017
DY3205	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-NAB3 (URA3)] [pRS315GFPNab3 (*LEU2*)]	This Study
DY3206	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS315GFPNab3 (*LEU2*)]	This Study
DY3228	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS315GFPNab3 (*LEU2*)]	This Study
DY3233	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS316Nab3 (*URA3*) [pRS315GFPNab3Δ191 (*LEU2*)]	This Study
DY32309	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS315GFPNab3 (*LEU2*)][pRS316Nab3 (*URA3*)]	This Study
DY32359	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS316Nab3 (*URA3*)]	This Study
DY4524	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS316Nab3 (*URA3*)] [pRS315GFPNab3Sup35 (*LEU2*)]	This Study
DY4525	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS316Nab3 (*URA3*)] [pRS315GFPNab3Rat1 (*LEU2*)]	This Study
DY4538	*MATa ura3Δ0 his3Δ1 leu2Δ0 btn2Δ0::kanMX* [pRS315GFPNab (*LEU2*)]	This Study
DY4540	*MATa ura3Δ0 his3Δ1 leu2Δ0 nic96-GFP* [pRS315yomkate2Nab3 (*LEU2*)]	This Study
DY4546	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS316Nab3 (*URA3*)][pRS315GFPNab3Q-->E (*LEU2*)]	This Study
DY4549	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS316Nab3 (*URA3*)] [pRS315-GFP-PrLD (*LEU2*)]	This Study
DY4551	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX htb2-mCherry:HIS3* [pRS316Nab3 (*URA3*)][pRS315GFPNab3ΔRRM (*LEU2*)]	This Study
DY4553	*MATα ura3Δ0 his3Δ1 leu2Δ0 nab3Δ0::kanMX* [pRS316-*NAB3Sup35* (*URA3*)][pRS315GFP-Nab3(*LEU2*)]	This Study
YOL890	*ura3-1 his3-11,15 leu2-3,112 lys2 htb2-mCherry:HIS3*	Wente Lab

Chapter 5

Discussion and future directions

Chapter 5: Discussion and future directions

At the outset of the research project for this dissertation, the goal was expanding our understanding of the role of the constituents of the NNS complex in regulating transcription termination. A genetic screen using a reporter whose promoter region contained a consensus Nab3-binding site identified a series of mutants with termination defects. These mutants were characterized and the affected genes identified as encoding proteins involved in termination, or post-transcriptional processing of transcripts [35]. Attention was focused on the C-terminus of Nab3 since multiple nonsense mutations contained in it were isolated in the screen. Prior to this work, the only well characterized domain of Nab3 was its RRM and the function of the essential C-terminus of Nab3 was poorly understood. The mutant with the smallest change on Nab3's primary sequence was a 19 amino acid C-terminal truncation. This region contained 18 amino acids with sequence homology to an α-helix in human hnRNP-C, whose α-helix is known to be needed for formation of tetramers of the protein during RNA processing [93]. A series of short truncations of the α-helix of Nab3, as well as point mutants of residues within the α-helix, all produced termination defects. Interestingly, a genetic interaction of two Nab3 mutants, one in the RRM, the other in the C-terminal domain, showed complementation in *trans.* These results, and others, support a model of Nab3 function that involves homotypic interactions *in vivo* dependent upon the C-terminal domain. Further analysis showed loss of a larger region of the C-terminus of Nab3 was lethal, confirming earlier reports that an essential function resides within the C-terminal domain. These data support a model where

Nab3 assembly, mediated by the C-terminus, is necessary for proper termination and establish the C-terminus as a new functional domain important in NNS complex function.

5.1 Discovery of an essential assembly domain in the C-terminus of Nab3

The above findings set the stage for the findings in chapter 2. A growing body of research showed that RNA-binding proteins form higher order complexes, and formation of these complexes were often mediated by a PrLD [77, 78]. These data suggested PrLD mediated assembly is involved in RNA-binding protein activity *in vivo*. Nab3 and its binding partner, Nrd1, were identified as PrLD containing proteins in an *in silico* screen of the yeast proteome [34]. Our lab explored whether these two essential PrLD containing proteins form amyloid-based assemblies and if this amyloid-like assembly was needed for cellular function.

Nrd1 is an essential RRM-containing RNA-binding protein with a PrLD at its C-terminus. An open question was whether this PrLD had amyloid forming properties like other well characterized PrLDs. The 60 amino acid PrLD, as defined in Alberti *et al* [34], was purified and incubated at 4℃ and 20℃ to allow for formation of amyloid-like fibers. The resulting protein was examined by electron microscopy for formation of amyloid filaments (Fig 2-4).

Nab3 contains as essential PrLD. It scored as prion-like in one of the four assays used by Alberti *et al* [34]. Based on these observations we embarked on a more thorough examination of the biochemical properties of Nab3's PrLD. The purified 134 C-terminal amino acids of Nab3 showed a spontaneous ability to form hydrogels *in vitro*, consistent with observations of other PrLD-containing proteins. A series of standard assays to test

amyloid-like properties all showed Nab3's PrLD as an amyloid-like domain, *e.g.*, X-ray fiber diffraction, electron microscopy, Thioflavin T binding, circular dichroism, and SDD-AGE. The PrLD, in the context of full length Nab3, drove amyloid-like assembly of the full-length protein. What is the essential role of this amyloid forming property in Nab3? A set of PrLD mutants that abrogate *in vitro* amyloid formation are lethal *in vivo*, suggesting that Nab3's essential function is dependent upon PrLD-mediated assembly. This assembly is most likely related to termination as this is the best-established cellular function of Nab3.

The findings in chapter 2 establish the presence of an amyloid forming PrLD in Nab3 and Nrd1. This fits a model of RNA-binding protein function being modulated by amyloid mediated higher order assemblies. Coupling this with previous data about the Nab3 C-terminal α-helix suggested both domains play a role in Nab3 assembly *in vivo*. Chapter 3 further assessed the separability of the PrLD and the C-terminal α-helix for viability and probed the question of how functionally versatile PrLDs are across the proteome, *i.e.*, are all PrLDs equivalent in their ability to replace a protein's native PrLD?

5.2 Exploring the interoperability of PrLDs in RNA-binding proteins

A simple model of PrLDs is that the assembly function of the domain is its functionally important aspect. If so, any domain capable of amyloid formation will suffice for the proper function of the larger protein. To examine this model, we exploited the fact that Nab3 contains an essential PrLD, permitting us to test heterologous PrLDs for their ability to complement the essential function of Nab3's PrLD. A series of PrLDs, as defined by Alberti et al.[34] were selected as candidates for a heterologous PrLD swap onto a truncated version of Nab3 where 191 amino acids of the C-terminus were removed,

(Nab3Δ191). The PrLDs of three proteins involved in termination (Pcf11, Hrp1, and Rat), two established prions (Sup35 and Rnq1), and one epsin-like protein with no known association to RNA metabolism (Ent2), were grafted onto Nab3Δ191 and assessed first for viability and, if viable, assessed for termination competence. Three of the six candidate PrLDs were viable and termination, as assayed by a GFP reporter was only slightly altered. What was most surprising was the three viable candidates were disparate in their cellular function. Hrp1 is a component of cleavage factor 1 in yeast and like Nab3 is an RNA-binding protein with evidence suggesting it is involved in termination of short non-coding RNAs. This is the least surprising result as one would expect other RNA-binding proteins involved in termination would have PrLDs modulating similar assembly states. The second viable PrLD was from Sup35, the most studied yeast prion and a translation termination factor whose PrLD is not essential for its cellular function. In this swap the Sup35 PrLD does provide the essential function for Nab3. Finally, the PrLD of Ent2, an epsin-like cell membrane associated protein involved in the endocytic pathway with no links to RNA metabolism also restored viability to Nab3 lacking most of its PrLD.

The simple model we began with clearly was not predictive of the results obtained. There is no general cross compatibility of PrLDs. Going further, one may hypothesize that proteins of similar function would share compatible PrLDs. In one case, Hrp1, that was true but Pcf11 and Rat1 are termination proteins whose PrLDs failed to complement function. From this it seems clear there are classes or types of PrLD as dissected in our assays. An interesting possibility is that each PrLD serves an addressing function for its native host protein perhaps with regard to the set of proteins with which it assembles. Table 3-3 contains a breakdown of amino acid compositions of the family of PrLDs. No obvious

pattern at the primary sequence level emerges between the viable and non-viable groups. A more exhaustive analysis of grafting more PrLDs onto Nab3Δ191 may provide further insight. Given the data, it is reasonable to postulate that the primary sequence composition is not key to determining function, and possibly something about the conformation assumed by the PrLD drives function. An avenue of further inquiry would be a structural analysis of wild-type Nab3 and Nab3 with various heterologous PrLDs in solution via cryo-electron microscopy or other solution-based structural determination techniques to learn if common structural patterns emerge in the group of viable heterologous PrLDs.

Two other interesting results came from this study. First, Nab3Sup35 showed an unusual phenotype where a subset of the yeast population containing the chimera showed growth alterations and a strong termination readthrough defect. The conversion of this subset of GFP negative (termination competent) yeast colonies into a GFP positive state (termination defective) became stably heritable. It is possible that the subset of cells with Nab3Sup35 that spontaneously switch from termination competent to incompetent have experienced a conversion of Nab3Sup35 into a heritable prion-like form. To further assess this phenomenon, it would be interesting to attempt to cure the prion-like state using established techniques such as growth on guanidinium containing media, which blocks prion formation through inhibition of Hsp104, or placing Nab3Sup35 in an Hsp104 mutant, a chaperone essential for fragmenting prion fibers allowing them to propagate to daughter cells. Loss of Hsp104 is known to result in loss of [PSI+], the prion form of Sup35 [171, 172].

In probing the importance of primary sequence composition on PrLD function we devised an unbiased screen in which segments of calf thymus genomic DNA were

appended to Nab3Δ191 to search for fusions capable of sustaining viability. One viable fusion protein that could substitute for the loss of either Nab3's or Hrp1's PrLD was examined in detail. Surprisingly, the amino acid sequence scored poorly in algorithms to detect prions and shared little similarity to the endogenous PrLDs of Nab3, Hrp1, and the aforementioned yeast PrLD donors in the domain swapping experiments. The segment identified is a simple 28 aa sequence, biased for cysteine, valine, and leucine which sum to 43% of the sequence. SDD-AGE suggests that it does support Nab3 polymerization, but possibly by a distinct manner from PrLDs [37]. Due to the prevalence of cysteines, a model where cysteine bridges between Nab3Ct25 mediate formation was tested. The cysteines were individually and pairwise substituted, but these changes have no effect on viability or assembly. A second, untested model is one whereby the hydrophobic residues form an α-helix that drives assembly similar to the α-helix of Nab3 and hnRNP-C.

Those questions outstanding, it is clear that Nab3 has an essential function contained within its C-terminal domain, that the function is mediated by an amyloid-forming PrLD, and that loss of the ability to form amyloid is lethal. The question remains why is assembly necessary for cellular survival? What essential processes are impacted by loss of assembly? A candidate process was revealed by Darby et al [96] who found that Nab3 and Nrd1 are localized to a granule in response to glucose starvation. Chapter 4 covered our investigation of Nab3's PrLD and its role in modulating granule recruitment.

5.3 Nab3's RRM and PrLD are both required for optimal granule recruitment

There is a growing body of evidence for a model of RNA metabolism involving the assembly of membrane-free compartments through liquid-liquid phase separation. This process results in locally higher concentrations of compartment constituents compared to

the surrounding milieu, with compartment constituents often able to dynamically enter and exit the granule. Such compartments include the nucleolus and Cajal bodies in the nucleus, and P-bodies and stress granules in the cytoplasm, all of which are enriched in RNA-binding proteins and RNAs [103]. While not all mechanisms regulating this de-mixing process are not known, it is known to often be mediated by PrLDs within RNA-binding proteins [77, 78, 105]. Nab3 was found in a previously uncharacterized granule induced by glucose starvation[96]. Fitting with the model, it is plausible that Nab3's PrLD regulates its assembly into the glucose starvation-induced granule in a reversible, dynamic manner (Fig. 5-1). The first question addressed was if Nab3's PrLD-mediated granule recruitment. Loss of the PrLD as well as a set of substitution mutations that reduce the glutamine content of the PrLD and that were known to destroy amyloid formation in Nab3, prevented the protein from localizing to the granule, adding support to the model of PrLD mediated granule formation. Not only is the PrLD necessary for recruitment to the granule, but the RRM is also needed for granule recruitment, suggesting a cooperativity between RNA binding and protein-protein interactions in Nab3 granule recruitment, which has been observed elsewhere [77, 84, 90, 173-177]. An extension of the question about the mechanism of granule formation is the role of LLPS. 1,6-hexanediol is an aliphatic alcohol commonly used to probe the liquid-like nature of other granules. Upon exposure to 1,6-hexanediol, LLPS-mediated granules dissolve, except for those containing irreversibly aggregated amyloid fibers. Indeed, Nab3-containing granules were dissolved upon exposure, fitting with the model of a dynamic, LLPS granule.

Second, upon refeeding granule containing cells with a metabolizable sugar, they begin to divide and Nab3 and Nrd1 dissociate from the granule. The kinetics of formation,

over an hour, appear much slower than dissolution, on the order of minutes, suggesting the mechanism behind dissolution responds rapidly upon reintroduction of metabolizable sugars. It is noteworthy that galactose had an extremely long lag period relative to the other metabolizable sugars before granule dissolution, suggesting some difficulty in recovering from glucose deprivation and moving into galactose metabolism. The dissolution coincides with cells re-entering the cell cycle and dividing. Perhaps the transcriptional changes necessary for galactose metabolism are difficult to orchestrate in a cell that has been starved for extended periods of time explaining the extended lag phase.

Currently, none of the markers used to identify known granules by co-localization are present in the Nab3-containing granule[96]. This leaves open the question of the identity of the Nab3 granule, what other proteins or molecules are found there, and its specific function.

5.4 Future directions

This work is at an exciting point where numerous open questions remain offering many paths to explore. While not an exhaustive list I will cover several next steps I would like to see explored in future studied.

Is the granule a site of active or inactive NNS termination? At this point, we know Nab3 and its binding partner Nrd1 both localize to this granule, the interaction is reversible and translation independent. The active or inactive state of those components would require a readout of active transcription within the granule as opposed to measuring steady state RNA levels. Starving cells to induce granule formation, then pulsing with uridine analogue 5-bromouridine (BrU) followed by fixing and staining with anti-BrUTP antibody is a

potential means of probing this question or newer "click" chemistry-based methods using 5-ethynyluridine [178, 179].

What other modifications may play a role in Nab3 assembly in the cell? It is known that Nrd1 undergoes a de-phosphorylation during starvation, suggesting PTMs as modulators of assembly [96]. Nab3 also has putative phosphorylation as well as acetylation sites, which make them potential effectors of assembly state, and acetylation of Nab3 by the lysine acetyltransferase Esa1 seems to alter its nuclear distribution [180]. A first step to addressing this question would be a simple *in vitro* assay whereby a kinase or acetylase is added to Nab3. If successfully modified, a band shift should be detectable by SDS-PAGE and reversed by phosphatase or deacetylase treatment. Positive results from this first pass experiment would warrant in vivo analysis if mutants to assess viability, response to starvation, and granule formation. This would then open up exploration of which cellular kinases/acetylases regulate modification, potentially linking back to established signaling pathways in the cell, *e.g.*, Darby et al.'s data suggesting a genetic interaction between Ras signaling and NNS[96]. This avenue of experimentation would help expand our understanding of NNS regulation and give insights into potential regulation of other RNA-binding proteins and their recruitment to various granules.

Why do a subset of cells exhibit granules? We observe granules in 20% of cells during starvation. While this is consistent with other reports, this fact makes it difficult to imagine an essential role for a granule not formed in all cells. Even though not all cells form granules in our experiments, this does not mean they cannot form them at all. One possibility is that the stage of the cell cycle in which a cell resides, is a determinant of granule formation. Nab3 is known to effect cyclin 3 expression, a regulator of G1 exit

[181]. If granule containing cells are all in the same part of the cell cycle it may provide an underlying cause for only a subset containing granules. An established method for assessing where yeast are in the cell cycle is bud size analysis. Immediately upon separation of the daughter cell, a new bud begins to form, therefore they are always present on the mother and they always start forming as M transitions to G1. If analysis of the phase of the cell cycle correlates to granule formation in the cell, the time-lapse live cell technique used in chapter 4 could be applied to monitoring one cell across multiple rounds of the cell cycle. In each division it could be stressed at one phase of the cell cycle to confirm that a given cell forms granules in a cell cycle-dependent manner. Showing this is reproducible across a large number of cells would help confirm this model. An alternative experiment would be to synchronize a population of yeast in specific phases of the cell cycle [182], then starve the cells post release to see if granule formation correlates more strongly to certain cell cycle phases.

What, if any, benefit arises from forming Nab3-containing granules? A test of this would be to assess survival of cells with granule forming versions of Nab3 versus those with granule defective Nab3 for differential survival in response to starvation. We are currently testing this with wild-type Nab3 and Nab3Δ191α, a viable form of Nab3 missing a large portion of the PrLD but containing the α-helix, that shows diminished ability to form granules via microscopy. The data suggest a reduction in viability in Nab3Δ191α containing cells post starvation. It is possible that the essential function is not the ability to form larger, microscopic granules, this could be examined using a recently reported technique to assess smaller scale molecular assemblies at CTD phosphorylated pol II [146]. Perhaps Nab3 assembly at genome wide transcriptional termination events is the essential

function for the PrLD and this can now be probed using the aforementioned technique. To formally establish which PrLD function is essential one would need cells containing only Nab3Δ191. As Nab3Δ191 alone would result in a lethal phenotype, a system in which an inducible wild-type Nab3 can be used to leave only a PrLD-defective form that would allow probing the cellular roles altered by loss of the PrLD with regard to viability.

What else is in the granule? At this point only two proteins are known to be in the Nab3-containing granule, Nab3 and Nrd1. Beyond that there is data on which established granule markers are not present. A candidate-based approach for other translational machinery within the granule can be tested using C-terminal GFP fusions from the yeast GFP collection [183]. Ideally, purifying granules then subjecting them to downstream mass spectroscopy and nucleic acid analysis would provide a clear answer on what constitutes the granule. Technically this has not been achieved yet in our lab, but others have shown the ability to recover stress granule-containing cells through a flow cytometry-based method then performing downstream analysis [184]. Successfully replicating this method would allow us to look at changes at the whole cell level with regards to transcriptome and proteomic changes in granule-containing versus granule-free cells. If the Nab3-containing granule could be purified cells that would provide the cleanest answer, but liquid-like granules have not been shown to be separable at this point.

In this work we have presented a careful characterization of a previously unstudied domain of the essential RNA-binding protein Nab3. We established that the C-terminus of Nab3 contains a PrLD capable of forming amyloid-like fibers and that this amyloid mediated assembly is essential for viability. PrLDs seem to fall into complementation groups whereby PrLDs from proteins with disparate cellular functions are interoperable

while those from proteins in similar processes are not, illustrating that a protein's identity or gene ontology is not a good predictor of the ability to rescue Nab3 function. Finally, we showed that a glucose starvation induced Nab3-containing granule is dependent upon Nab3's PrLD for Nab3 recruitment and have preliminary data suggesting PrLD mutants may have diminished survival post starvation relative to wild type Nab3 containing cells. From this work, a large number of new avenues are open to explore and continued experimentation in Nab3, that will deepen our understanding of the regulation and function of PrLDs in RNA metabolism across eukarya.

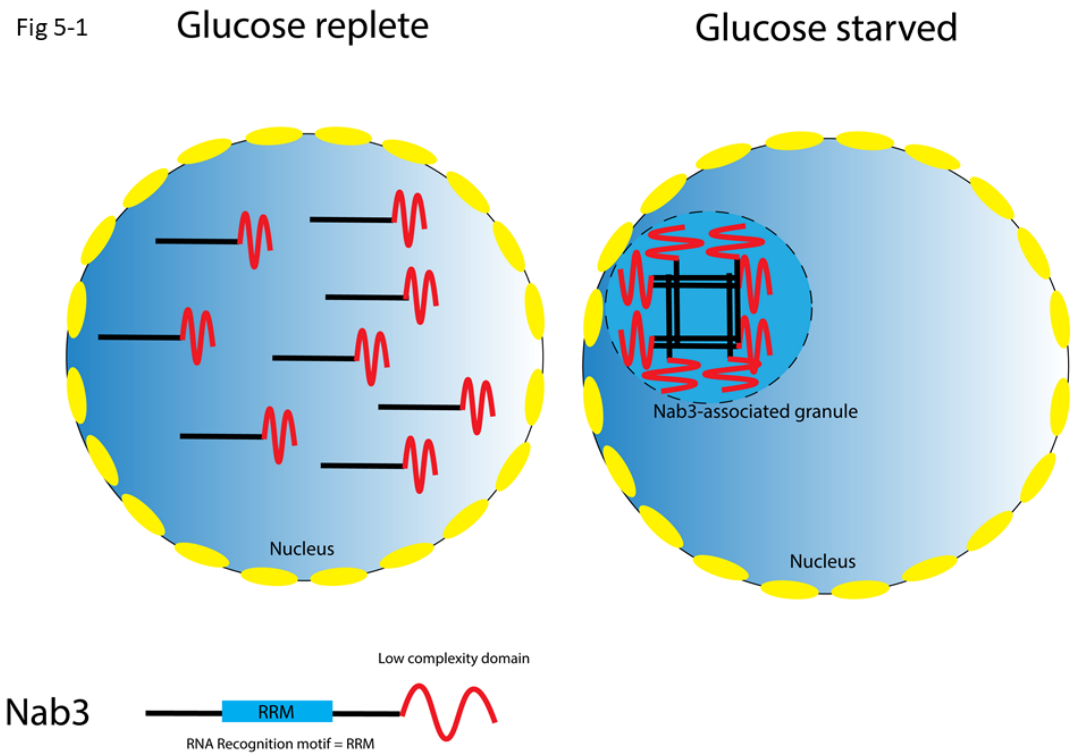


Fig 5-1 Model of Nab3's PrLD regulated granule recruitment. Under glucose rich conditions Nab3 is pan-nuclear. Upon glucose starvation Nab3 localizes to a granule due to PrLD mediated assembly.

Works Cited

1. Jensen, T.H., A. Jacquier, and D. Libri, *Dealing with pervasive transcription.* Mol Cell, 2013. **52**(4): p. 473-84.
2. Meinhart, A. and P. Cramer, *Recognition of RNA polymerase II carboxy-terminal domain by 3'-RNA-processing factors.* Nature, 2004. **430**(6996): p. 223-6.
3. Krishnamurthy, S., et al., *Ssu72 Is an RNA polymerase II CTD phosphatase.* Mol Cell, 2004. **14**(3): p. 387-94.
4. Murray, S., et al., *Phosphorylation of the RNA polymerase II carboxy-terminal domain by the Bur1 cyclin-dependent kinase.* Mol Cell Biol, 2001. **21**(13): p. 4089-96.
5. Mosley, A.L., et al., *Rtr1 is a CTD phosphatase that regulates RNA polymerase II during the transition from serine 5 to serine 2 phosphorylation.* Mol Cell, 2009. **34**(2): p. 168-78.
6. Ahn, S.H., M.-C. Keogh, and S. Buratowski, *Ctk1 promotes dissociation of basal transcription factors from elongating RNA polymerase II*, in *Embo Journal*. 2009. p. 205-212.
7. Nikolov, D.B. and S.K. Burley, *RNA polymerase II transcription initiation: a structural view.* Proc Natl Acad Sci U S A, 1997. **94**(1): p. 15-22.
8. Sims, R.J., 3rd, R. Belotserkovskaya, and D. Reinberg, *Elongation by RNA polymerase II: the short and long of it.* Genes Dev, 2004. **18**(20): p. 2437-68.
9. Bowman, E.A. and W.G. Kelly, *RNA Polymerase II transcription elongation and Pol II CTD Ser2 phosphorylation: A tail of two kinases.* Nucleus, 2014. **5**(3): p. 224-236.
10. Mischo, H.E. and N.J. Proudfoot, *Disengaging polymerase: terminating RNA polymerase II transcription in budding yeast.* Biochim Biophys Acta, 2013. **1829**(1): p. 174-85.
11. Wilson, S.M., et al., *Characterization of nuclear polyadenylated RNA-binding proteins in Saccharomyces cerevisiae.* The Journal of cell biology, 1994. **127**(5): p. 1173-84.
12. Winey, M. and M.R. Culbertson, *Mutations affecting the tRNA-splicing endonuclease activity of Saccharomyces cerevisiae.* Genetics, 1988. **118**(4): p. 609-17.
13. Steinmetz, E.J. and D.A. Brow, *Repression of gene expression by an exogenous sequence element acting in concert with a heterogeneous nuclear ribonucleoprotein-like protein, Nrd1, and the putative helicase Sen1.* Molecular and cellular biology, 1996. **16**(12): p. 6993-7003.
14. Steinmetz, E.J., et al., *RNA-binding protein Nrd1 directs poly(A)-independent 3'-end formation of RNA polymerase II transcripts.* Nature, 2001. **413**(6853): p. 327-31.
15. Grzechnik, P. and J. Kufel, *Polyadenylation linked to transcription termination directs the processing of snoRNA precursors in yeast.* Molecular cell, 2008. **32**(2): p. 247-58.
16. Xu, Z., et al., *Bidirectional promoters generate pervasive transcription in yeast.* Nature, 2009. **457**(7232): p. 1033-7.
17. Wyers, F., et al., *Cryptic pol II transcripts are degraded by a nuclear quality control pathway involving a new poly(A) polymerase.* Cell, 2005. **121**(5): p. 725-37.
18. Davis, C.A. and M. Ares, Jr., *Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in Saccharomyces cerevisiae.* Proc Natl Acad Sci U S A, 2006. **103**(9): p. 3262-7.
19. Neil, H., et al., *Widespread bidirectional promoters are the major source of cryptic transcripts in yeast.* Nature, 2009. **457**(7232): p. 1038-42.

20. David, L., et al., *A high-resolution map of transcription in the yeast genome.* Proc Natl Acad Sci U S A, 2006. **103**(14): p. 5320-5.
21. Xu, Z., et al., *Antisense expression increases gene expression variability and locus interdependency.* Mol Syst Biol, 2011. **7**: p. 468.
22. Aravind, L., et al., *Lineage-specific loss and divergence of functionally linked genes in eukaryotes.* Proc Natl Acad Sci U S A, 2000. **97**(21): p. 11319-24.
23. Berretta, J., M. Pinskaya, and A. Morillon, *A cryptic unstable transcript mediates transcriptional trans-silencing of the Ty1 retrotransposon in S. cerevisiae.* Genes Dev, 2008. **22**(5): p. 615-26.
24. Martens, J.A., L. Laprade, and F. Winston, *Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene.* Nature, 2004. **429**(6991): p. 571-4.
25. Conrad, N.K., et al., *A yeast heterogeneous nuclear ribonucleoprotein complex associated with RNA polymerase II.* Genetics, 2000. **154**(2): p. 557-71.
26. Carroll, K.L., et al., *Interaction of yeast RNA-binding proteins Nrd1 and Nab3 with RNA polymerase II terminator elements.* RNA, 2007. **13**(3): p. 361-73.
27. Loya, T.J., et al., *A network of interdependent molecular interactions describes a higher order Nrd1-Nab3 complex involved in yeast transcription termination.* J Biol Chem, 2013. **288**(47): p. 34158-67.
28. Carroll, K.L., et al., *Identification of cis elements directing termination of yeast nonpolyadenylated snoRNA transcripts.* Molecular and cellular biology, 2004. **24**(14): p. 6241-52.
29. Creamer, T.J., et al., *Transcriptome-wide binding sites for components of the Saccharomyces cerevisiae non-poly(A) termination pathway: Nrd1, Nab3, and Sen1.* PLoS genetics, 2011. **7**(10): p. e1002329.
30. Porrua, O., et al., *In vivo SELEX reveals novel sequence and structural determinants of Nrd1-Nab3-Sen1-dependent transcription termination.* The EMBO journal, 2012. **31**(19): p. 3935-48.
31. Webb, S., et al., *PAR-CLIP data indicate that Nrd1-Nab3-dependent transcription termination regulates expression of hundreds of protein coding genes in yeast.* Genome Biol, 2014. **15**(1): p. R8.
32. Ursic, D., et al., *Multiple protein/protein and protein/RNA interactions suggest roles for yeast DNA/RNA helicase Sen1p in transcription, transcription-coupled DNA repair and RNA processing.* Nucleic Acids Res, 2004. **32**(8): p. 2441-52.
33. Fasken, M.B., R.N. Laribee, and A.H. Corbett, *Nab3 Facilitates the Function of the TRAMP Complex in RNA Processing via Recruitment of Rrp6 Independent of Nrd1.* PLoS Genetics, 2015. **11**(3): p. e1005044.
34. Alberti, S., et al., *A systematic survey identifies prions and illuminates sequence features of prionogenic proteins.* Cell, 2009. **137**(1): p. 146-58.
35. Loya, T.J., T.W. O'Rourke, and D. Reines, *A genetic screen for terminator function in yeast identifies a role for a new functional domain in termination factor Nab3.* Nucleic acids research, 2012. **40**(15): p. 7476-91.
36. O'Rourke, T.W., et al., *Amyloid-like assembly of the low complexity domain of yeast Nab3.* Prion, 2015. **9**(1): p. 34-47.
37. Loya, T.J., T.W. O'Rourke, and D. Reines, *The hnRNP-like Nab3 termination factor can employ heterologous prion-like domains in place of its own essential low complexity domain.* PLoS One, 2017. **12**(10): p. e0186187.
38. Loya, T.J., T.W. O'Rourke, and D. Reines, *Yeast Nab3 protein contains a self-assembly domain found in human heterogeneous nuclear ribonucleoprotein-C (hnRNP-C) that is*

necessary for transcription termination. The Journal of biological chemistry, 2013. **288**(4): p. 2111-7.

39. Yuryev, A., et al., *The C-terminal domain of the largest subunit of RNA polymerase II interacts with a novel set of serine/arginine-rich proteins.* Proceedings of the National Academy of Sciences of the United States of America, 1996. **93**(14): p. 6975-80.
40. Vasiljeva, L., et al., *The Nrd1-Nab3-Sen1 termination complex interacts with the Ser5-phosphorylated RNA polymerase II C-terminal domain.* Nature structural & molecular biology, 2008. **15**(8): p. 795-804.
41. Jenks, M.H., T.W. O'Rourke, and D. Reines, *Properties of an intergenic terminator and start site switch that regulate IMD2 transcription in yeast.* Mol Cell Biol, 2008. **28**(12): p. 3883-93.
42. Gudipati, R.K., et al., *Phosphorylation of the RNA polymerase II C-terminal domain dictates transcription termination choice.* Nature structural & molecular biology, 2008. **15**(8): p. 786-94.
43. Steinmetz, E.J. and D.A. Brow, *Control of pre-mRNA accumulation by the essential yeast protein Nrd1 requires high-affinity transcript binding and a domain implicated in RNA polymerase II association.* Proceedings of the National Academy of Sciences of the United States of America, 1998. **95**(12): p. 6699-704.
44. Tudek, A., et al., *Molecular basis for coordinating transcription termination with noncoding RNA degradation.* Mol Cell, 2014. **55**(3): p. 467-81.
45. Kim, M., et al., *Distinct pathways for snoRNA and mRNA termination.* Molecular cell, 2006. **24**(5): p. 723-34.
46. Thiebaut, M., et al., *Transcription termination and nuclear degradation of cryptic unstable transcripts: a role for the nrd1-nab3 pathway in genome surveillance.* Molecular cell, 2006. **23**(6): p. 853-64.
47. Schulz, D., et al., *Transcriptome surveillance by selective termination of noncoding RNA synthesis.* Cell, 2013. **155**(5): p. 1075-87.
48. Jamonnak, N., et al., *Yeast Nrd1, Nab3, and Sen1 transcriptome-wide binding maps suggest multiple roles in post-transcriptional RNA processing.* RNA, 2011. **17**(11): p. 2011-25.
49. Nedea, E., et al., *Organization and function of APT, a subcomplex of the yeast cleavage and polyadenylation factor involved in the formation of mRNA and small nucleolar RNA 3'-ends.* The Journal of biological chemistry, 2003. **278**(35): p. 33000-10.
50. Riordan, D.P., D. Herschlag, and P.O. Brown, *Identification of RNA recognition elements in the Saccharomyces cerevisiae transcriptome.* Nucleic Acids Research, 2011. **39**(4): p. 1501-9.
51. DeMarini, D.J., et al., *SEN1, a positive effector of tRNA-splicing endonuclease in Saccharomyces cerevisiae.* Mol Cell Biol, 1992. **12**(5): p. 2154-64.
52. Ghaemmaghami, S., et al., *Global analysis of protein expression in yeast.* Nature, 2003. **425**(6959): p. 737-41.
53. Suraweera, A., et al., *Senataxin, defective in ataxia oculomotor apraxia type 2, is involved in the defense against oxidative DNA damage.* J Cell Biol, 2007. **177**(6): p. 969-79.
54. De Amicis, A., et al., *Role of senataxin in DNA damage and telomeric stability.* DNA Repair (Amst), 2011. **10**(2): p. 199-209.
55. Becherel, O.J., et al., *Senataxin plays an essential role with DNA damage response proteins in meiotic recombination and gene silencing.* PLoS Genet, 2013. **9**(4): p. e1003435.
56. Porrua, O. and D. Libri, *A bacterial-like mechanism for transcription termination by the Sen1p helicase in budding yeast.* Nat Struct Mol Biol, 2013. **20**(7): p. 884-91.

57. Kim, H.D., J. Choe, and Y.S. Seo, *The sen1(+) gene of Schizosaccharomyces pombe, a homologue of budding yeast SEN1, encodes an RNA and DNA helicase.* Biochemistry, 1999. **38**(44): p. 14697-710.
58. Han, Z., D. Libri, and O. Porrua, *Biochemical characterization of the helicase Sen1 provides new insights into the mechanisms of non-coding transcription termination.* Nucleic Acids Research, 2017. **45**(3): p. 1355-1370.
59. LaCava, J., et al., *RNA degradation by the exosome is promoted by a nuclear polyadenylation complex.* Cell, 2005. **121**(5): p. 713-24.
60. Vanacova, S., et al., *A new yeast poly(A) polymerase complex involved in RNA quality control.* PLoS biology, 2005. **3**(6): p. e189.
61. Mitchell, P., *Exosome substrate targeting: the long and short of it.* Biochem Soc Trans, 2014. **42**(4): p. 1129-34.
62. van Hoof, A., P. Lennertz, and R. Parker, *Yeast exosome mutants accumulate 3'-extended polyadenylated forms of U4 small nuclear RNA and small nucleolar RNAs.* Mol Cell Biol, 2000. **20**(2): p. 441-52.
63. Jia, H., et al., *RNA unwinding by the Trf4/Air2/Mtr4 polyadenylation (TRAMP) complex.* Proc Natl Acad Sci U S A, 2012. **109**(19): p. 7292-7.
64. Jia, H., et al., *The RNA helicase Mtr4p modulates polyadenylation in the TRAMP complex.* Cell, 2011. **145**(6): p. 890-901.
65. Kadaba, S., et al., *Nuclear surveillance and degradation of hypomodified initiator tRNAMet in S. cerevisiae.* Genes Dev, 2004. **18**(11): p. 1227-40.
66. Schneider, C. and D. Tollervey, *Threading the barrel of the RNA exosome.* Trends Biochem Sci, 2013. **38**(10): p. 485-93.
67. Kuehner, J.N. and D.A. Brow, *Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation.* Mol Cell, 2008. **31**(2): p. 201-11.
68. Rondon, A.G., et al., *Fail-safe transcriptional termination for protein-coding genes in S. cerevisiae.* Mol Cell, 2009. **36**(1): p. 88-98.
69. Steinmetz, E.J., et al., *Genome-wide distribution of yeast RNA polymerase II and its control by Sen1 helicase.* Mol Cell, 2006. **24**(5): p. 735-46.
70. Arigo, J.T., et al., *Regulation of yeast NRD1 expression by premature transcription termination.* Molecular cell, 2006. **21**(5): p. 641-51.
71. Steinmetz, E.J., et al., *cis- and trans-acting determinants of transcription termination by yeast RNA polymerase II.* Molecular and Cellular Biology, 2006. **26**(7): p. 2688-2696.
72. Jenks, M.H. and D. Reines, *Dissection of the molecular basis of mycophenolate resistance in Saccharomyces cerevisiae.* Yeast, 2005. **22**(15): p. 1181-90.
73. Glover, J.R., et al., *Self-seeded fibers formed by Sup35, the protein determinant of [PSI+], a heritable prion-like factor of S. cerevisiae.* Cell, 1997. **89**(5): p. 811-9.
74. Pan, K.M., et al., *Conversion of alpha-helices into beta-sheets features in the formation of the scrapie prion proteins.* Proc Natl Acad Sci U S A, 1993. **90**(23): p. 10962-6.
75. Liebman, S.W. and Y.O. Chernoff, *Prions in yeast.* Genetics, 2012. **191**(4): p. 1041-72.
76. King, O.D., A.D. Gitler, and J. Shorter, *The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease.* Brain Res, 2012. **1462**: p. 61-80.
77. Kato, M., et al., *Cell-free formation of RNA granules: low complexity sequence domains form dynamic fibers within hydrogels.* Cell, 2012. **149**(4): p. 753-67.
78. Han, T.W., et al., *Cell-free formation of RNA granules: bound RNAs identify features and components of cellular assemblies.* Cell, 2012. **149**(4): p. 768-79.
79. Mao, Y.S., B. Zhang, and D.L. Spector, *Biogenesis and function of nuclear bodies.* Trends Genet, 2011. **27**(8): p. 295-306.

80. Miller, S.B., et al., *Compartment-specific aggregases direct distinct nuclear and cytoplasmic aggregate deposition.* Embo j, 2015. **34**(6): p. 778-97.
81. Gallina, I., et al., *Cmr1/WDR76 defines a nuclear genotoxic stress body linking genome integrity and protein quality control.* Nat Commun, 2015. **6**: p. 6533.
82. Kurischko, C. and J.R. Broach, *Phosphorylation and nuclear transit modulate the balance between normal function and terminal aggregation of the yeast RNA-binding protein Ssd1.* Mol Biol Cell, 2017. **28**(22): p. 3057-3069.
83. McCann, C., et al., *The Ataxin-2 protein is required for microRNA function and synapse-specific long-term olfactory habituation.* Proc Natl Acad Sci U S A, 2011. **108**(36): p. E655-62.
84. Van Treeck, B., et al., *RNA self-assembly contributes to stress granule formation and defining the stress granule transcriptome.* Proc Natl Acad Sci U S A, 2018. **115**(11): p. 2734-2739.
85. Mao, Y.S., et al., *Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs.* Nat Cell Biol, 2011. **13**(1): p. 95-101.
86. Gilks, N., et al., *Stress granule assembly is mediated by prion-like aggregation of TIA-1.* Mol Biol Cell, 2004. **15**(12): p. 5383-98.
87. Mittag, T. and R. Parker, *Multiple Modes of Protein-Protein Interactions Promote RNP Granule Assembly.* J Mol Biol, 2018.
88. Boulay, G., et al., *Cancer-Specific Retargeting of BAF Complexes by a Prion-like Domain.* Cell, 2017.
89. Molliex, A., et al., *Phase Separation by Low Complexity Domains Promotes Stress Granule Assembly and Drives Pathological Fibrillization.* Cell, 2015. **163**(1): p. 123-33.
90. Lin, Y., et al., *Formation and Maturation of Phase-Separated Liquid Droplets by RNA-Binding Proteins.* Mol Cell, 2015. **60**(2): p. 208-19.
91. Conicella, A.E., et al., *ALS Mutations Disrupt Phase Separation Mediated by alpha-Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain.* Structure, 2016. **24**(9): p. 1537-49.
92. Patel, A., et al., *A Liquid-to-Solid Phase Transition of the ALS Protein FUS Accelerated by Disease Mutation.* Cell, 2015. **162**(5): p. 1066-77.
93. McCloskey, A., et al., *hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export.* Science, 2012. **335**(6076): p. 1643-6.
94. Wang, J.T., et al., *Regulation of RNA granule dynamics by phosphorylation of serine-rich, intrinsically disordered proteins in C. elegans.* Elife, 2014. **3**: p. e04591.
95. Buchan, J.R., D. Muhlrad, and R. Parker, *P bodies promote stress granule assembly in Saccharomyces cerevisiae.* The Journal of Cell Biology, 2008. **183**(3): p. 441-455.
96. Darby, M.M., et al., *The Saccharomyces cerevisiae Nrd1-Nab3 transcription termination pathway acts in opposition to Ras signaling and mediates response to nutrient depletion.* Mol Cell Biol, 2012. **32**(10): p. 1762-75.
97. Kato, M., Y. Lin, and S.L. McKnight, *Cross-beta polymerization and hydrogel formation by low-complexity sequence proteins.* Methods, 2017. **126**: p. 3-11.
98. Kuehner, J.N., E.L. Pearson, and C. Moore, *Unravelling the means to an end: RNA polymerase II transcription termination.* Nature reviews. Molecular cell biology, 2011. **12**(5): p. 283-94.
99. Hyle, J.W., R.J. Shaw, and D. Reines, *Functional distinctions between IMP dehydrogenase genes in providing mycophenolate resistance and guanine prototrophy to yeast.* J Biol Chem, 2003. **278**(31): p. 28470-8.

100. Romero, P., et al., *Sequence complexity of disordered protein.* Proteins, 2001. **42**(1): p. 38-48.
101. Radivojac, P., et al., *Intrinsic disorder and functional proteomics.* Biophys J, 2007. **92**(5): p. 1439-56.
102. Frieden, C., *Protein aggregation processes: In search of the mechanism.* Protein Sci, 2007. **16**(11): p. 2334-44.
103. Weber, S.C. and C.P. Brangwynne, *Getting RNA and protein in phase.* Cell, 2012. **149**(6): p. 1188-91.
104. Kwon, I., et al., *Phosphorylation-regulated binding of RNA polymerase II to fibrous polymers of low-complexity domains.* Cell, 2013. **155**(5): p. 1049-60.
105. Kim, H.J., et al., *Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS.* Nature, 2013. **495**(7442): p. 467-73.
106. Remy, I. and S.W. Michnick, *Clonal selection and in vivo quantitation of protein interactions with protein-fragment complementation assays.* Proc Natl Acad Sci U S A, 1999. **96**(10): p. 5394-9.
107. Mehta, A.K., et al., *Context dependence of protein misfolding and structural strains in neurodegenerative diseases.* Biopolymers, 2013. **100**(6): p. 722-30.
108. LeVine, H., 3rd, *Quantification of beta-sheet amyloid fibril structures with thioflavin T.* Methods Enzymol, 1999. **309**: p. 274-84.
109. Halfmann, R., et al., *Opposing effects of glutamine and asparagine govern prion formation by intrinsically disordered proteins.* Mol Cell, 2011. **43**(1): p. 72-84.
110. Shen, C.L. and R.M. Murphy, *Solvent effects on self-assembly of beta-amyloid peptide.* Biophys J, 1995. **69**(2): p. 640-51.
111. Biancalana, M. and S. Koide, *Molecular mechanism of Thioflavin-T binding to amyloid fibrils.* Biochim Biophys Acta, 2010. **1804**(7): p. 1405-12.
112. Walsh, D.M., et al., *Amyloid beta-protein fibrillogenesis. Structure and biological activity of protofibrillar intermediates.* J Biol Chem, 1999. **274**(36): p. 25945-52.
113. Tarassov, K., et al., *An in vivo map of the yeast protein interactome.* Science, 2008. **320**(5882): p. 1465-70.
114. Holmes, D.L., et al., *Heritable remodeling of yeast multicellularity by an environmentally responsive prion.* Cell, 2013. **153**(1): p. 153-65.
115. Heinrich, S.U. and S. Lindquist, *Protein-only mechanism induces self-perpetuating changes in the activity of neuronal Aplysia cytoplasmic polyadenylation element binding protein (CPEB).* Proc Natl Acad Sci U S A, 2011. **108**(7): p. 2999-3004.
116. Suzuki, G., N. Shimazu, and M. Tanaka, *A yeast prion, Mod5, promotes acquired drug resistance and cell survival under environmental stress.* Science, 2012. **336**(6079): p. 355-9.
117. Chi, A., et al., *Analysis of phosphorylation sites on proteins from Saccharomyces cerevisiae by electron transfer dissociation (ETD) mass spectrometry.* Proc Natl Acad Sci U S A, 2007. **104**(7): p. 2193-8.
118. Ficarro, S.B., et al., *Phosphoproteome analysis by mass spectrometry and its application to Saccharomyces cerevisiae.* Nat Biotechnol, 2002. **20**(3): p. 301-5.
119. Ling, S.C., M. Polymenidou, and D.W. Cleveland, *Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis.* Neuron, 2013. **79**(3): p. 416-38.
120. Ramaswami, M., J.P. Taylor, and R. Parker, *Altered ribostasis: RNA-protein granules in degenerative disorders.* Cell, 2013. **154**(4): p. 727-36.
121. Pezza, J.A., et al., *Amyloid-associated activity contributes to the severity and toxicity of a prion phenotype.* Nat Commun, 2014. **5**: p. 4384.

122. Raveendra, B.L., et al., *Characterization of prion-like conformational changes of the neuronal isoform of Aplysia CPEB.* Nat Struct Mol Biol, 2013. **20**(4): p. 495-501.
123. Arndt, K.M. and D. Reines, *Termination of Transcription of Short Noncoding RNAs by RNA Polymerase II.* Annu Rev Biochem, 2015.
124. Bresson, S., et al., *Nuclear RNA Decay Pathways Aid Rapid Remodeling of Gene Expression in Yeast.* Mol Cell, 2017. **65**(5): p. 787-800.e5.
125. van Nues, R., et al., *Kinetic CRAC uncovers a role for Nab3 in determining gene expression profiles during stress.* Nat Commun, 2017. **8**(1): p. 12.
126. Scherzinger, E., et al., *Huntingtin-encoded polyglutamine expansions form amyloid-like protein aggregates in vitro and in vivo.* Cell, 1997. **90**(3): p. 549-58.
127. Boke, E., et al., *Amyloid-like Self-Assembly of a Cellular Compartment.* Cell, 2016. **166**(3): p. 637-50.
128. Courchaine, E. and K.M. Neugebauer, *Paraspeckles: Paragons of functional aggregation.* J Cell Biol, 2015. **210**(4): p. 527-8.
129. Walker, L.C., *Proteopathic Strains and the Heterogeneity of Neurodegenerative Diseases.* Annu Rev Genet, 2016. **50**: p. 329-346.
130. Altschul, S.F., et al., *Gapped BLAST and PSI-BLAST: a new generation of protein database search programs.* Nucleic Acids Res, 1997. **25**(17): p. 3389-402.
131. Sikorski, R.S. and P. Hieter, *A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in Saccharomyces cerevisiae.* Genetics, 1989. **122**(1): p. 19-27.
132. Sherman, F., *Getting started with yeast.* Methods Enzymol, 2002. **350**: p. 3-41.
133. Chen, D.C., B.C. Yang, and T.T. Kuo, *One-step transformation of yeast in stationary phase.* Curr Genet, 1992. **21**(1): p. 83-4.
134. Morrison, A., et al., *Eukaryotic DNA polymerase amino acid sequence required for 3'----5' exonuclease activity.* Proc Natl Acad Sci U S A, 1991. **88**(21): p. 9473-7.
135. Henry, M.F., et al., *The yeast hnRNP-like protein Hrp1/Nab4 sccumulates in the cytoplasm after hyperosmotic stress: a novel Fps1-dependent response.* Mol Biol Cell, 2003. **14**(9): p. 3929-41.
136. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
137. Reines, D. and M. Clarke, *Immunochemical analysis of the supramolecular structure of myosin in contractile cytoskeletons of Dictyostelium amoebae.* The Journal of biological chemistry, 1985. **260**(26): p. 14248-54.
138. Ross, E.D., U. Baxa, and R.B. Wickner, *Scrambled prion domains form prions and amyloid.* Mol Cell Biol, 2004. **24**(16): p. 7206-13.
139. Ross, E.D., et al., *Primary sequence independence for prion formation.* Proc Natl Acad Sci U S A, 2005. **102**(36): p. 12825-30.
140. Chen, X., et al., *Transcriptomes of six mutants in the Sen1 pathway reveal combinatorial control of transcription termination across the Saccharomyces cerevisiae genome.* PLoS Genet, 2017. **13**(6): p. e1006863.
141. Haruki, H., J. Nishikawa, and U.K. Laemmli, *The anchor-away technique: rapid, conditional establishment of yeast mutant phenotypes.* Mol Cell, 2008. **31**(6): p. 925-32.
142. !!! INVALID CITATION !!! {}.
143. Kopcewicz, K.A., T.W. O'Rourke, and D. Reines, *Metabolic regulation of IMD2 transcription and an unusual DNA element that generates short transcripts.* Mol Cell Biol, 2007. **27**(8): p. 2821-9.
144. Anderson, P. and N. Kedersha, *RNA granules.* J Cell Biol, 2006. **172**(6): p. 803-8.

145. Boehning, M., et al., *RNA polymerase II clustering through carboxy-terminal domain phase separation.* Nat Struct Mol Biol, 2018. **25**(9): p. 833-840.
146. Lu, H., et al., *Phase-separation mechanism for C-terminal hyperphosphorylation of RNA polymerase II.* Nature, 2018. **558**(7709): p. 318-323.
147. Chong, S., et al., *Imaging dynamic and selective low-complexity domain interactions that control gene transcription.* Science, 2018. **361**(6400).
148. Sabari, B.R., et al., *Coactivator condensation at super-enhancers links phase separation and gene control.* Science, 2018. **361**(6400).
149. Cho, W.K., et al., *Mediator and RNA polymerase II clusters associate in transcription-dependent condensates.* Science, 2018. **361**(6400): p. 412-415.
150. Lee, C., et al., *Protein aggregation behavior regulates cyclin transcript localization and cell-cycle control.* Dev Cell, 2013. **25**(6): p. 572-84.
151. Gietz, D., et al., *Improved method for high efficiency transformation of intact yeast cells.* Nucleic acids research, 1992. **20**(6): p. 1425.
152. Wheeler, J.R., et al., *Distinct stages in stress granule assembly and disassembly.* Elife, 2016. **5**.
153. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.
154. Bennett, M.R., et al., *Metabolic gene regulation in a dynamically changing environment.* Nature, 2008. **454**(7208): p. 1119-22.
155. Kelley, J.B., et al., *RGS proteins and septins cooperate to promote chemotropism by regulating polar cap mobility.* Curr Biol, 2015. **25**(3): p. 275-285.
156. Harrison, A.F. and J. Shorter, *RNA-binding proteins with prion-like domains in health and disease.* Biochem J, 2017. **474**(8): p. 1417-1438.
157. Lord, C.L., et al., *Altering nuclear pore complex function impacts longevity and mitochondrial function in S. cerevisiae.* The Journal of Cell Biology, 2015. **208**(6): p. 729-744.
158. Shah, K.H., et al., *A Hybrid-Body Containing Constituents of Both P-Bodies and Stress Granules Forms in Response to Hypoosmotic Stress in Saccharomyces cerevisiae.* PLoS One, 2016. **11**(6): p. e0158776.
159. LaFleur, D.W., et al., *Monoclonal antibody therapeutics with up to five specificities: functional enhancement through fusion of target-specific peptides.* MAbs, 2013. **5**(2): p. 208-18.
160. Zhang, H., et al., *RNA Controls PolyQ Protein Phase Transitions.* Mol Cell, 2015. **60**(2): p. 220-30.
161. Langdon, E.M., et al., *mRNA structure determines specificity of a polyQ-driven phase separation.* Science, 2018. **360**(6391): p. 922-927.
162. Patel, S.S., et al., *Natively unfolded nucleoporins gate protein diffusion across the nuclear pore complex.* Cell, 2007. **129**(1): p. 83-96.
163. Kroschwald, S., et al., *Promiscuous interactions and protein disaggregases determine the material state of stress-inducible RNP granules.* Elife, 2015. **4**: p. e06807.
164. Kaganovich, D., R. Kopito, and J. Frydman, *Misfolded proteins partition between two distinct quality control compartments.* Nature, 2008. **454**(7208): p. 1088-95.
165. Grandi, P., V. Doye, and E.C. Hurt, *Purification of NSP1 reveals complex formation with 'GLFG' nucleoporins and a novel nuclear pore protein NIC96.* Embo j, 1993. **12**(8): p. 3061-71.
166. Guet, D., et al., *Combining Spinach-tagged RNA and gene localization to image gene expression in live yeast.* Nat Commun, 2015. **6**: p. 8882.

167. Quain, D.E. and C.A. Boulton, *Growth and metabolism of mannitol by strains of Saccharomyces cerevisiae.* J Gen Microbiol, 1987. **133**(7): p. 1675-84.
168. Carneiro, T., et al., *Inactivation of cleavage factor I components Rna14p and Rna15p induces sequestration of small nucleolar ribonucleoproteins at discrete sites in the nucleus.* Mol Biol Cell, 2008. **19**(4): p. 1499-508.
169. Zorio, D.A. and D.L. Bentley, *The link between mRNA processing and transcription: communication works both ways.* Exp Cell Res, 2004. **296**(1): p. 91-7.
170. Chen, S. and L.E. Hyman, *A specific RNA-protein interaction at yeast polyadenylation efficiency elements.* Nucleic Acids Res, 1998. **26**(21): p. 4965-74.
171. Kryndushkin, D.S., et al., *Yeast [PSI+] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104.* J Biol Chem, 2003. **278**(49): p. 49636-43.
172. Chernoff, Y.O., et al., *Role of the chaperone protein Hsp104 in propagation of the yeast prion-like factor [psi+].* Science, 1995. **268**(5212): p. 880-4.
173. Patel, A., et al., *ATP as a biological hydrotrope.* Science, 2017. **356**(6339): p. 753-756.
174. Decker, C.J., D. Teixeira, and R. Parker, *Edc3p and a glutamine/asparagine-rich domain of Lsm4p function in processing body assembly in Saccharomyces cerevisiae.* J Cell Biol, 2007. **179**(3): p. 437-49.
175. Nonhoff, U., et al., *Ataxin-2 interacts with the DEAD/H-box RNA helicase DDX6 and interferes with P-bodies and stress granules.* 2007. **18**(4): p. 1385-1396.
176. Hennig, S., et al., *Prion-like domains in RNA binding proteins are essential for building subnuclear paraspeckles.* J Cell Biol, 2015. **210**(4): p. 529-39.
177. Nott, T.J., et al., *Phase transition of a disordered nuage protein generates environmentally responsive membraneless organelles.* Mol Cell, 2015. **57**(5): p. 936-947.
178. Jao, C.Y. and A. Salic, *Exploring RNA transcription and turnover in vivo by using click chemistry.* Proc Natl Acad Sci U S A, 2008. **105**(41): p. 15779-84.
179. Wansink, D.G., et al., *Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus.* J Cell Biol, 1993. **122**(2): p. 283-93.
180. Chang, C.S., A. Clarke, and L. Pillus, *Suppression analysis of esa1 mutants in Saccharomyces cerevisiae links NAB3 to transcriptional silencing and nucleolar functions.* G3 (Bethesda), 2012. **2**(10): p. 1223-32.
181. Sugimoto, K., et al., *Dosage suppressors of the dominant G1 cyclin mutant CLN3-2: identification of a yeast gene encoding a putative RNA/ssDNA binding protein.* Mol Gen Genet, 1995. **248**(6): p. 712-8.
182. Rosebrock, A.P., *Synchronization and Arrest of the Budding Yeast Cell Cycle Using Chemical and Genetic Methods.* Cold Spring Harb Protoc, 2017. **2017**(1).
183. Huh, W.K., et al., *Global analysis of protein localization in budding yeast.* Nature, 2003. **425**(6959): p. 686-91.
184. Hidalgo, I.H., et al., *Characterization of aggregate load and pattern in living yeast cells by flow cytometry.* Biotechniques, 2016. **61**(3): p. 137-48.