Yeast Nab3 Protein Contains a Self-assembly Domain Found in Human Heterogeneous Nuclear Ribonucleoprotein-C (hnRNP-C) That Is Necessary for Transcription Termination*

Travis J. Loya, Thomas W. O’Rourke, and Daniel Reines

From the Department of Biochemistry, Emory University School of Medicine, Atlanta, Georgia 30322

Background: Termination factor Nab3 binds RNA, but its function and structure are poorly understood.

Results: The carboxyl terminus of Nab3 functions as a self-assembly region important for termination.

Conclusion: Nab3 resembles hnRNP-C in this new function, indicating that it is related to mammalian RNA-packaging proteins.

Significance: Nab3 coats RNA to form a higher order assembly due to interactions among RNA, RNA polymerase, and hnRNPs.

Nab3 is an RNA-binding protein whose function is important for terminating transcription by RNA polymerase II. It co-assembles with Nrd1, and the resulting heterodimer of these heterogeneous nuclear ribonucleoprotein-C (hnRNP)-like proteins interacts with the nascent transcript and RNA polymerase II. Previous genetic analysis showed that a short carboxyl-terminal region of Nab3 is functionally important for termination and is located far from the Nab3 RNA recognition domain in the primary sequence. The domain is structurally homologous to hnRNP-C from higher organisms. Here we provide biochemical evidence that this short region is sufficient to enable self-assembly of Nab3 into a tetrameric form in a manner similar to the cognate region of human hnRNP-C. Within this region, there is a stretch of low complexity protein sequence (16 glutamines) adjacent to a putative α-helix that potentiates the ability of the conserved region to self-assemble. The glutamine stretch and the final 18 amino acids of Nab3 are both important for termination in living yeast cells. The findings herein describe an additional avenue by which these hnRNP-like proteins can polymerize on target transcripts. This process is independent of, but acts in concert with, the interactions of the proteins with RNA and RNA polymerase and extends the relationship of Nab3 as a functional orthologue of a higher eukaryotic hnRNP.

Saccharomyces cerevisiae Nab3 is an hnRNP2-like, RNA-binding protein that is involved in the termination of transcription of small, noncoding RNAs such as snRNAs and cryptic unstable (CUT) RNAs such as that involved in regulation of the IMD2 gene (1, 2). IMD2 encodes IMP dehydrogenase, an enzyme important for GTP synthesis. Cellular GTP levels control its expression. When GTP is abundant, RNA polymerase II initiates transcription at an upstream location where the incorporation of tandem guanines is called for by the template. Just downstream, a terminator is encountered that results in a short CUT RNA. When GTP concentrations are too low to support this start site choice, RNA polymerase II initiates with an adenyate at +1 that is downstream of the terminator. This terminator bypass results in synthesis of IMD2 mRNA in GTP-depleted states.

During termination, Nab3 associates with Nrd1, another hnRNP-like protein that also contains a conserved RNA recognition motif (RRM) (3). One model of transcription termination suggests that this complex associates with the repeated carboxyl-terminal domain of the large subunit of RNA polymerase II through the Nrd1 subunit (4). As well, Nab3 and Nrd1 recognize specific RNA sequences in target RNA enabling the proteins to associate directly with RNA (4). The number and arrangement of such sites vary from substrate to substrate. Thus, nascent RNA destined for termination by this mechanism accumulates a coating of Nrd1 and Nab3 proteins during RNA chain elongation, marking it for the specific mode of termination that characterizes this class of RNA. The process also involves the putative RNA helicase Sen1 and the TRAMP complex, both of which interact with the Nrd1/Nab3 proteins to activate termination and RNA processing (2).

Whereas Nrd1 appears to be the yeast orthologue of human hnRNPs SCAF8 and SCAF4 (3–5) and it contains a CTD-interacting domain and an RRM, Nab3 is less obviously related to a particular mammalian hnRNP protein based upon sequence homology. Like a number of other hnRNPs, it contains amino acid repeats and tracts of homopolymeric residues including aspartic acid/glutamic acid and proline/glutamine rich domains that have not yet been assigned function. Polyglutamine stretches are of biomedical interest because a variety of proteins suffer mutations that lengthen the homopolymeric tract resulting in multiple human diseases (6). Whereas long tracts are associated with protein aggregation, wild type runs of polyglutamine in proteins are important for protein function where they can augment protein-protein interactions depending upon their length and the sequences adjacent to them (6, 7). In particular, polyglutamine regions can stabilize the co-assembly of adjacent coiled-coil structures (8, 9).
Self-association of Yeast Termination Factor Nab3

Little is known of the function of most of the essential Nab3 polypeptide outside of the RRM and Nrd1-interaction domains, including the numerous runs of consecutive glutamines and other amino acids. Furthermore, the protein is difficult to express in recombinant form to study its function as a specialized ribonucleoprotein complex (10).

Recently, we revealed a role for a small segment of the carboxyl-terminal “tail” of Nab3 in transcription termination at the cut associated with IMD2 regulation (1, 11). From spontaneous mutations in yeast, we learned that the final 19 amino acids at the Nab3 carboxy terminus are important for function. Interestingly, deletion of the last three amino acids recapitulated the phenotype observed when the entire tail was lost, and even deletion of only the last amino acid of Nab3 resulted in significant loss of function (11). Furthermore, this region shares the phenotype observed when the entire tail was lost, and Interestingly, deletion of the last three amino acids recapitulated the phenotype observed when the entire tail was lost, and even deletion of only the last amino acid of Nab3 resulted in significant loss of function (11). Furthermore, this region shares the phenotype observed when the entire tail was lost, and even deletion of only the final 40 amino acids was generated from the plasmid pRS315Nab3FL (11) using oligonucleotides 5′-aaatgcgctgctatttttgagcttttgctaaactatctaatagactttgaacat-3′ and 5′-tcgacctatttttagcttttgctaaactatctaatagactttgaacat-3′ (for L800A) or 5′-aaatgcgctgctatttttgagcttttgctaaactatc-3′ and 5′-tcgacctatttttagcttttgctaaactatc-3′ (for Q801A) were annealed, phosphorylated, and ligated into it. The resulting Nab3 region was excised from the respective intermediates with HindIII and Xhol and inserted into similarly cut pRS315Nab3 to generate pRS315-Nab3-L800A and pRS315-Nab3-Q801A.

pRS315-Nab3 was the starting plasmid from which the Gln769 to Gln783 polyglutamine stretch was deleted by in vitro replication with Phusion DNA polymerase (New England Biolabs) and mutagenic oligonucleotides 5′-ctctgctgctatttttgagcttttgctaaactatctaatagactttgaacat-3′ and 5′-aaatgcgctgctatttttgagcttttgctaaactatctaatagactttgaacat-3′. Replicated DNA was used to transform E. coli, and the loss of the region encoding the glutamine stretch in the new plasmid (pRS315-Nab3-NoQ) was confirmed by DNA sequencing.

The strains used in this study are presented in Table 1. The LEU2-marked plasmids pRS315-Nab3-L800A, pRS315-Nab3-Q801A, and pRS315-Nab3-NoQ were introduced independently into a nab3 strain that also contained the plasmid pRS316-nab3-11 (DY30229 (11)). Following growth on SC ura− leu− plates, colonies were grown at 30°C on SC leu− and tested for loss of the URA3-bearing plasmid by growth on 5-fluoroorotic acid. Following this plasmid shuffle, yeast were then transformed with the GAL-terminator-GFP reporter plasmid described previously (11) and were named DY3126, DY3127, and DY3129. The resulting yeast strains, and a control strain expressing wild type NAB3 (DY3036 (11)), were assayed at an excitation wavelength of 488 nm on a LSRII (BD Bioscience) flow cytometer as described (11). The lithium acetate method of transformation was used throughout (13).

Protein Cross-linking—Cross-linking reactions were prepared as described by Whitson et al. (12). Briefly, purified recombinant proteins (0.5–1 μg/μl) were incubated for 30 min at 22°C with freshly dissolved 5 mM (unless otherwise indicated) bis(sulfosuccinimidyl)suberate (BS3; Thermo Scientific) in 20 mM Tris, pH 7.5, 50 mM NaCl, 1 mM EDTA. Reactions were terminated by adjustment to 50 mM glycine. Proteins were precipitated with 10% (v/v) trichloroacetic acid and resolved on SDS gels composed of a resolving slab of 9.5% polyacrylamide layered upon an 11% gel to better separate monomeric from polymeric species. Where indicated, 16% polyacrylamide-Tris-
Tricine gels were used to resolve small proteins as described by Schagger (14). Gels were stained with fresh Coomassie Brilliant Blue (0.2% w/v).

**RESULTS**

Using chemical cross-linking, a 27-residue α-helix from hnRNP-C was shown to self-assemble into a tetrameric bundle when appended to a heterologous protein (*E. coli* thioredoxin (12)). Because the carboxyl-terminal tail of Nab3 is structurally related to the hnRNP-C self-assembling helix (11), we employed this assay to test whether the cognate region of the yeast protein could similarly polymerize. A family of recombinant fusion proteins in which various portions of the Nab3 carboxyl terminus were fused to thioredoxin was expressed in *E. coli* strains and purified. The homobifunctional reagent BS3 was used to cross-link the purified proteins, and the resulting multimers were separated by SDS-PAGE. The 18-amino acid terminus of Nab3 (residues 785–802; Fig. 1, Nab3–18aa) was able to facilitate intermolecular cross-linking when attached to thioredoxin (Fig. 2A, Nab3–18). More effective cross-linking was observed when the carboxyl-terminal 40 amino acids of Nab3 (Fig. 1, Nab3–40aa) were added to the carboxyl terminus of thioredoxin (Fig. 2A, Nab3–40). This construct included a stretch of 16 glutamines immediately adjacent to the 18-amino acid tail that was identified as important by mutation (11). The pattern of higher order multimers resembled what was seen for the cross-linking of a thioredoxin fusion protein to which the human hnRNP-C assembly helix was attached (Fig. 2A, hnRNP-C-27), as described by Whitson et al. (12). These cross-linking data are evidence that the Nab3 tail is functionally similar to the hnRNP-C assembly domain and that it is likely to possess α-helical character. The similarity in disposition of hydrophobic and hydrophilic residues seen when the primary sequence of yeast Nab3 and human hnRNP-C were displayed in a helical wheel representation is also consistent with this interpretation (Fig. 1).

The structure of hnRNP-C self-assembly domain and the helical wheel modeling of the related Nab3 region suggest that this protein-protein interaction results from a packing of hydrophobic faces of the helices into a bundle. This implies the interaction would be relatively insensitive to high ionic strength. To test this, we repeated the cross-linking assay in 0.5 M and 1 M NaCl. The thioredoxin-Nab3 40 amino acid fusion protein could be readily cross-linked at these high ionic strengths (data not shown).

A spontaneous mutation in Nab3 (tov1) changed the final glutamine in the 16-residue repeat to a stop codon, resulting in a Nab3 protein that lacks this residue and the carboxyl-terminal 18 amino acid tail (Fig. 1, Nab3-tov1) (11). To test whether this derivative of Nab3 could self-associate, we generated a thioredoxin fusion derivative containing only the 21 amino acids (763–783) of Nab3 on the amino-terminal side of the stop
Codon. In other words, we tested a fusion protein with the last 40 amino acids of Nab3 from which the 18-amino acid tail was removed, leaving 15 glutamines at the fusion protein carboxyl terminus. This protein also generated multimers (Fig. 2B, tov1).

Similarly, we recapitulated a second tov mutant, tov4, which had a nonsense mutation within the glutamine repeat tract. The construct had a thioredoxin-Nab3 fusion polypeptide that extends only 7 residues into the glutamine repeat (residues 763–775; Fig. 1, Nab3-tov4). This derivative also assembled into higher order species as determined by chemical cross-linking (Fig. 2B, tov4). For the two tov mutant derivatives of Nab3, as well as the complete 40-amino acid tail, a small amount of the dimeric form was detectable in the absence of chemical cross-linking reagent (Fig. 2B, lanes). More of this denaturation-resistant form was observed for the 40-amino acid piece of Nab3 versus the 13- or 21-amino acid segments representing tov mutants. We conclude that both the terminal 18 amino acids and the adjacent 16 glutamines of Nab3 contribute independently to the ability to self-associate. Together they provide a multimerization interface such as that seen for hnRNP-C.

To confirm that multimerization was due to the Nab3 portion of the fusion proteins, the 18- and 40-amino acid portions of Nab3 were removed from thioredoxin by digestion with thrombin at a site engineered into the expression plasmid. Free thioredoxin was separated from the cleaved Nab3 tails by metal affinity chromatography, and the small isolated tails were cross-linked with BS3 and analyzed by Tris-Tricine PAGE. Both the 18- and 40-amino acid pieces of Nab3 formed higher order polypeptides following cross-linking, with the larger 40-amino acid piece doing so more robustly (Fig. 3). Interestingly, multimers of the 40-amino acid piece were apparent in the absence of cross-linking reagent, similar to what has been seen for the 27-amino acid α-helix from human hnRNP-C and other polyglutamine-containing proteins (12, 15). This independently confirmed that the carboxyl-terminal region of Nab3 can self-associate and that the glutamine tract augments this property.

As an additional test of the ability of the Nab3 tail to self-assemble, we subjected the thioredoxin-Nab3 40-amino acid fusion protein to gel filtration on Superdex 200. Two protein peaks were resolved in migration positions consistent with the sizes of monomer and tetramer based on the positions of reference size markers (data not shown). When the monomeric peak was pooled and gel-filtered a second time, it again was separable as two peaks, consistent with an equilibrium between monomeric and assembled versions of the fusion protein (Fig. 4).

We next tested whether yeast could grow following removal of the carboxyl-terminal 40 amino acids of Nab3 that contains the glutamine stretch and the hnRNP-C homology region (Fig. 1). Through plasmid shuffling, we obtained a strain of yeast containing such an allele of Nab3 (nab3Δ40) on a plasmid covering a deletion of chromosomal NAB3. This strain was viable but grew slowly, as observed previously for the spontaneous mutant of Nab3 (tov1) that lacked its carboxyl-terminal 19 res-
Self-association of Yeast Termination Factor Nab3

When the penultimate residue (Gln⁸₀₁; Fig. 1) was mutated to alanine (blue), the Nab3 termination function was compromised to an extent similar to the loss of the glutamine tract. When the residue before it (Leu⁸₀⁰) was mutated to alanine, termination function was severely compromised resulting in GFP intensities as strong as that observed from loss of the carboxyl-terminal three amino acids ((11) Fig. 5B, orange). These data underscore the importance of the very carboxyl-terminal sequence of Nab3 and the polyglutamine stretch. The difference in severity between the alanine substitutions in residues 800 versus 801 may be related to structural modeling, which suggested that Leu⁸₀⁰ ends an α-helix whereas Gln⁸₀₁ is not in that helix (11).

DISCUSSION

Mutations in Nab3 that impaired its termination function have focused attention on the extreme carboxyl terminus of the protein (11). This domain has structural homology with an α-helical region of hnRNP-C that was shown to enable this mammalian protein to tetramerize in its role as a RNA-packaging protein (12). Adjacent to that is an unusually long stretch of glutamines. Here we show that this region of Nab3 is functionally similar to the cognate region of human hnRNP-C in that both self-assemble into tetramers. This strengthens the case that Nab3 is a yeast orthologue of mammalian hnRNP-C, a suggestion that initially arose due to homology between the RRM domains of the two proteins (16). A Nrd1 binding region near the Nab3 RRM has also been described (10), but other functions of this large yeast protein have remained mysterious. Study of full-length recombinant protein has proven difficult due to its unusual sequence composition which seems responsible for its extreme lability (data not shown (10)).

Although hnRNP-C1 and -C2 (293 and 306 residues, respectively) and Nab3 (802 residues) share similar RNA binding and tetramORIZATION domains, the positioning of these motifs in the respective primary sequences differs between the proteins, and the surrounding sequences are not closely related. Thus, although the RRM and assembly modules are conserved between the yeast and human proteins, the overall protein structure and function are more divergent. Whereas both polypeptides serve to size-select newly made RNAs, they do so in opposite ways. hnRNP-C binds and measures relatively long RNAs for processing and export (17) whereas Nab3 preferentially facilitates termination of short RNAs in yeast (18). hnRNP-C is highly abundant (~10⁶/cell (19)), whereas the level of Nab3 is more modest (~<6,000 copies/cell (20)). Genetic probing revealed that the self-assembly region of Nab3 is important for the biological function of the protein (11). The significance of the cognate region in hnRNP-C was shown biochemically, due in part to the ability to reconstitute RNA-protein particles with recombinant full-length hnRNP-C (21). hnRNP-C does not have a polyglutamine stretch and is toxic when expressed in S. cerevisiae (22).

Low complexity protein sequences, including homopolymeric runs of certain amino acids, are distributed nonrandomly in proteins. RNA-binding proteins are overrepresented in this group (23). The sequences tend to be involved in protein-protein interactions resulting in aggregation or assembly of
**Self-association of Yeast Termination Factor Nab3**

![Image of a graph and figure](image-url)

**FIGURE 5. Growth and in vivo termination activity of Nab3 tail mutants.** A, growth of tail deletion mutants. Yeast lacking their chromosomal copy of NAB3 and containing a plasmid with either wild type NAB3 (DY3034), NAB3 lacking the 16-glutamine stretch (nab3ΔQ16; DY3128), the nab3Δ19 deletion allele (DY2233), or the nab3Δ40 deletion allele (DY3132), were diluted to an A at 600 nm of 0.05, serially 10-fold diluted, and grown on SC leu−ura−medium at 30 °C. B, flow cytometry on yeast expressing GFP from a reporter plasmid. Yeast strains DY3129 (green), DY3127 (orange), DY3126 (blue), and DY3036 (red) were grown in SC leu−ura−medium and induced with galactose. 30,000 cells of each strain were subjected to flow cytometry as described under “Materials and Methods,” and the cell count was plotted versus the GFP intensity. The mean intensity is indicated next to the NAB3 status of each strain.

Dynamics polymers including hydrogels. This property has been suggested to be functionally important for the reversible formation of cytoplasmic granules that harbor RNA-binding proteins (23).

Nab3 contains numerous regions of low complexity sequence across its length, including 12 consecutive aspartates and 10 consecutive glutamates that are embedded in a stretch of 30 residues that are one or the other of these residues. Nab3 also has numerous uninterrupted glutamine runs including three stretches of 6 and one of 16. Our results suggest a role for the latter, which is the longest of the Nab3 homopolymeric tracts and which is immediately adjacent to the hnRNP-C self-assembly region important for transcription termination. This glutamine stretch augments the Nab3 protein–protein interaction, as has been seen elsewhere for polyglutamine tracts (9), and is consistent with suggestions that polyglutamate tracts form an extended helix when placed next to sequences with an α-helical character (9). Whether this is the case here will require more detailed structural analysis. Polyglutamine stretches are seen in human diseases in which the repeat expands by mutation. In general, the repeat expansions in humans are not sufficient for pathology. Instead, it seems that the gain of function in such mutant proteins results from enhancement of its native function rather than acquisition of a new role (6). This report provides some insight into a role for the natural function of a 16-glutamine stretch in a yeast protein whose structure is not well understood. We also learned there is a region neighboring the low complexity hnRNP-C-like multimerization tail that is essential for cell viability. Thus, there is another important domain in this unexplored region of Nab3. Whether its role is related to that of the extreme carboxyl-terminal tail remains to be elucidated.

These findings suggest an extension of current models positing that the Nab3/Nrd1 termination machinery forms a network with nascent RNA and the RNA polymerase II CTD (10, 24). Based upon the roles of the Nab3 glutamine tracts in both termination and protein–protein interaction, this nuclear complex could resemble the higher order assemblage seen in cytoplasmic RNA-binding proteins that contain low complexity primary amino acid sequences (23, 25). Self-association of Nab3 subunits provides an additional tether between the Nab3-Nrd1 heterodimers as they form an array on RNA substrates in which they recognize multiple consensus sequence motifs (10). The number and positioning of Nab3-Nrd1 heterodimers attracted to a single RNA could be governed, not only by the quantity of consensus sites on the RNA, but also through stabilization by the additional association of multiple Nrd1 subunits with the repeats in the CTD of RNA polymerase II (10). Reversible assembly may regulate the ability of the Nrd1-Nab3 complex to terminate transcription and guide substrate RNAs through their respective processing or degradative pathways. Determination of the structure and stoichiometry of such a hypothetical (Nab3-Nrd1)ₙ RNA complex likely depends upon the specific RNA substrate. Future analysis of these interactions will be aided by in vitro reconstitution from purified RNA and protein components.

Acknowledgments—We thank Dr. A. Krezel for samples of thioredoxin-hnRNP-C fusion protein, Drs. X. Cheng and G. Conn for helpful discussions, and Dr. G. Conn and his laboratory for access to their AKTA.

REFERENCES

1. Jenks, M. H., O’Rourke, T. W., and Reines, D. (2008) Properties of an intergenic terminator and start site switch that regulate IMD2 transcription in yeast. Mol. Cell. Biol. 28, 3883–3893
2. Kuehner, J. N., Pearson, E. L., and Moore, C. (2011) Unravelling the means to an end: RNA polymerase II transcription termination. Nat. Rev. Mol.
Self-association of Yeast Termination Factor Nab3

...