An intrinsically disordered C terminus allows the La protein to assist the biogenesis of diverse noncoding RNA precursors

Nathan J. Kucera^a, Michael E. Hodsdon^b, and Sandra L. Wolin^{a,c,1}

^aDepartment of Cell Biology, Yale University School of Medicine, New Haven, CT 06510; ^bDepartment of Laboratory Medicine, Yale University School of Medicine, New Haven, CT 06510; and ^cDepartment of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, CT 06510

Edited by Jennifer A. Doudna, University of California, Berkeley, CA, and approved December 6, 2010 (received for review November 15, 2010)

The La protein binds the 3' ends of many newly synthesized noncoding RNAs, protecting these RNAs from nucleases and influencing folding, maturation, and ribonucleoprotein assembly. Although 3' end binding by La involves the N-terminal La domain and adjacent RNA recognition motif (RRM), the mechanisms by which La stabilizes diverse RNAs from nucleases and assists subsequent events in their biogenesis are unknown. Here we report that a conserved feature of La proteins, an intrinsically disordered C terminus, is required for the accumulation of certain noncoding RNA precursors and for the role of the *Saccharomyces cerevisiae* La protein Lhp1p in assisting formation of correctly folded pre-tRNA anticodon stems in vivo. Footprinting experiments using purified Lhp1p reveal that the C terminus is required to protect a pre-tRNA anticodon stem from chemical modification. Although the C terminus of Lhp1p is hypersensitive to proteases in vitro, it becomes protease-resistant upon binding pre-tRNAs, U6 RNA, or pre-5S rRNA. Thus, while high affinity binding to 3' ends requires the La domain and RRM, a conformationally flexible C terminus allows La to interact productively with a diversity of noncoding RNA precursors. We propose that intrinsically disordered domains adjacent to well characterized RNA-binding motifs in other promiscuous RNA-binding proteins may similarly contribute to the ability of these proteins to influence the cellular fates of multiple distinct RNA targets.

RNA-binding protein | unstructured domain | noncoding RNA biogenesis

Many RNA-binding proteins interact with multiple ligands, influencing the fate of the RNAs and in some cases altering structure. For example, the pre-mRNA binding protein hnRNP A1 and the mRNA decay mediator KSRP also bind miRNA precursors to promote their maturation (1, 2). Similarly, the Y-box protein YB-1, a component of cytoplasmic mRNPs, functions in both pre-mRNA splicing and translational repression of some mRNAs (3). Although most of these proteins contain well characterized RNA-binding motifs, the mechanisms by which they interact with diverse targets and influence their conformations and outcomes are not understood.

One such protein, called La, is a nuclear phosphoprotein whose best characterized role is to assist noncoding RNA biogenesis. La binds the UUU_{OH} that is the initial end of all RNA polymerase III transcripts and protects the RNAs from exonucleases (4, 5). In budding yeast, La also binds processing intermediates of RNA polymerase II-transcribed small nuclear and nucleolar RNAs that end in UUU_{OH} (6–8). Because most La-bound RNAs undergo 3' maturation, La is usually not bound to the mature RNAs. Binding by La results in diverse consequences. For pre-tRNAs, La binding favors endonucleolytic removal of the trailer (9) and is required for wild-type levels of some tRNAs (10). For the yeast U4 snRNA, the La-bound precursor preferentially associates with the core Sm proteins in vitro, suggesting that La assists snRNP formation by protecting the favored substrate for Sm protein binding from nucleases (7).

In addition to its role in 3' end protection, La can influence RNA structure. Although La is dispensable for growth in budding

and fission yeast, La becomes essential in the presence of mutations that compromise the structure of essential pre-tRNAs or impair snRNP assembly (4, 5). While many of the mutations cause the affected RNA to be degraded without La, *Saccharomyces cerevisiae* La is required for correct folding of pre-tRNA^{Arg}_{CCG} when the tRNA carries a mutation that weakens the anticodon stem, allowing formation of an incorrect helix. In the absence of La, the mature tRNA accumulates but is not aminoacylated (11). The role of La in assisting pre-tRNA folding may be general, as La is required for efficient charging of two wild-type tRNAs when yeast are grown at low temperature (11). Also, because human and yeast La influence the translation of certain mRNAs, it is proposed that La may stabilize mRNA structural elements, such as those found in some internal ribosome entry sites (8, 12).

Although structural studies have revealed how La recognizes 3' ends, the mechanisms by which La contacts other regions of these RNAs to influence folding and contribute to their subsequent biogenesis are poorly understood. La contains an N-terminal La domain that adopts a winged helix-turn-helix fold, one or two RNA recognition motifs (RRMs), and a disordered C terminus (13–15). As the UUU_{OH} binds in a cleft between the La and RRM domains (16, 17), but does not involve the canonical nucleic acid-binding surfaces of either the winged helix or the RRM, it is proposed that these surfaces assist folding (5, 17, 18). However, this model is difficult to reconcile with the finding that binding of the *S. cerevisiae* La to a pre-tRNA reduces the accessibility of the anticodon stemloop to chemical and enzymatic probes (11), because neither the La domain nor the RRM is sufficient to span the ~70 Å between the 3' end and the anticodon loop.

To elucidate how La assists the biogenesis of multiple distinct RNAs, we carried out genetic and biochemical analyses of the *S. cerevisiae* La protein Lhp1p. Here we report that a previously unexplored feature of La proteins, the intrinsically disordered C terminus, is required for several Lhp1p functions. A truncated Lhp1p lacking the C terminus does not support efficient growth in several yeast strains that require Lhp1p. We show that the C terminus is required for the accumulation of precursors to the U4 spliceosomal snRNA and the U3 small nucleolar RNA and for the role of Lhp1p in allowing correct folding of a pre-tRNA with a weakly base paired anticodon stem. Although the C terminus is hypersensitive to proteases, consistent with data showing that this portion of Lhp1p is disordered, it becomes protease-resistant upon binding multiple noncoding RNA precursors. Our results indicate that while high affinity binding of

Author contributions: N.J.K., M.E.H., and S.L.W. designed research; N.J.K. and M.E.H. performed research; N.J.K., M.E.H., and S.L.W. analyzed data; and N.J.K., M.E.H., and S.L.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: sandra.wolin@yale.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1017085108/-DCSupplemental.

UUU_{OH}-containing RNAs requires the La motif and the RRM, a flexible C terminus allows La proteins to interact productively with other RNA elements.

Results

The C terminus is important for some Lhp1p functions. Lhp1p consists of an N-terminal La domain, a RRM, and a C terminus that is predicted to be disordered in Lhp1p and other La proteins (Fig. 1A and Fig. S1). In addition, a nonclassical nuclear import signal has been mapped to the RRM (19). To determine if the La domain and the RRM are sufficient for function, we tested whether a truncated Lhp1p containing only these domains [*lhp1*(1-227)] could support growth in strains that carry mutations that cause them to require *LHP1*. Plasmids containing the truncated *lhp1*(1-227) and the *TRP1* gene were introduced into *lhp1Δ* strains that also carried a mutation in tRNA^{Ser}_{CGA} (*sup61-10*), tRNA^{Arg}_{CCG} (*trr4-1*), or the U6 snRNP protein Lsm8p (*lsm8-1*) (9, 11, 20). In the presence of the mutations, Lhp1p is required for accumulation of mature tRNA^{Ser}_{CGA}, aminoacylation of tRNA^{Arg}_{CCG}, and accumulation of U6 snRNA, respectively (9, 11, 20). As these strains also contained full-length *LHP1* on an *URA3* gene-containing plasmid, we assayed whether the cells could lose the *URA3* plasmid and survive on media containing 5-fluoroorotic acid (5-FOA), which selects against *URA3*. Although all strains showed some growth on 5-FOA media in the presence of *lhp1*(1-227), growth with the truncated protein was always less robust than with full-length *LHP1* (Fig. 1B–D).

Because sequences outside the RRM may contribute to nuclear localization (19), we tested if failure of the truncated Lhp1p to accumulate in nuclei accounted for the growth defects. To this end, we performed immunofluorescence on *sup61-10* cells expressing the full-length or truncated protein as the only source of Lhp1p. We chose the *sup61-10* strain because it showed the

strongest growth in the presence of *lhp1*(1-227) (Fig. 1B–D). Western blotting revealed that *sup61-10* cells expressed both proteins at levels comparable to wild-type cells (Fig. S24). Colocalization experiments using DAPI to visualize DNA and Nop1p as a nucleolar marker revealed that, as described for Lhp1p (21), the full-length and truncated proteins localized to nuclei and nucleoli (Fig. 1E and Fig. S2B). Because it was reported that the RRM does not contain as strong an import signal as full-length Lhp1p (19), we tested if appending an SV40 nuclear localization signal (NLS) to the proteins affected function. Although adding the NLS to the N terminus of Lhp1p resulted in a nonfunctional protein in the growth assays, adding it to a variable loop within the RRM (*LHP1*-NLS; Fig. S2C) did not affect function in the tested strains (Fig. 1A–D). Importantly, although the truncated Lhp1p with the SV40 NLS [*lhp1*(1-227)-NLS] was strongly nuclear (Fig. 1E and Fig. S2B), it did not confer wild-type growth in the mutant strains. We conclude that the C terminus is required for optimal Lhp1p function.

Accumulation of Certain Noncoding RNA Precursors Requires the Lhp1p C terminus. To determine if the C terminus was important for the accumulation of newly synthesized RNAs, we examined RNA from *sup61-10* cells expressing the full-length and truncated proteins. Probing to detect pre-tRNA^{Ser}_{CGA} confirmed that, as described (9), wild-type cells contain three intron-containing forms: a primary transcript with 5' and 3' extensions, an RNase P cleavage product containing a mature 5' end and a 3' trailer, and the end-matured pre-tRNA generated by trailer cleavage (Fig. 24). Because Lhp1p protects the two trailer-containing pre-tRNAs from exonucleases (9), the primary transcript is shorter and less abundant in *lhp1Δ* cells, and the intermediate containing a mature 5' end and a 3' trailer is not detected (Fig. 24). In the presence of the *sup61-10* mutation, which disrupts the

Fig. 1. The C terminus is required for some Lhp1p functions. (A) Constructs encoding Lhp1 proteins. Lhp1p contains a Kap108-dependent NLS that maps to the RRM (19). In some constructs, an SV40 NLS was added to a loop in the RRM (Fig. S2C). (B–D) Plasmids containing the indicated *lhp1* alleles, or the empty vectors, were introduced into *sup61-10* *lhp1Δ* (B), *lsm8-1* *lhp1Δ* (C), or *trr4-1* *lhp1Δ* (D) strains that also carried *LHP1* on a *URA3*-containing vector (pSL28). *LHP1* was in the plasmid pAD12, which contains *TRP1* and *ADE2*, while *LHP1*-NLS, *lhp1*(1-227), and *lhp1*(1-227)-NLS were in the *TRP1* plasmid pRS314. After growing cells in media lacking tryptophan, fivefold serial dilutions were spotted on agar lacking tryptophan (-Trp) or containing 5-fluoroorotic acid (FOA) and grown at 25 °C (B, C, and D, left) or 16 °C (D, right). (E) Wild-type, *lhp1Δ*, and *sup61-10* *lhp1Δ* cells carrying the indicated plasmids were subjected to immunofluorescence with anti-Lhp1p (green) and anti-Nop1p (red). DNA was stained with DAPI (blue). As described (21), staining of *lhp1Δ* cells with anti-Lhp1p reveals a background cytoplasmic fluorescence. (Scale bar, 10 μm).

Fig. 2. Accumulation of some nascent RNAs requires the Lhp1p C terminus. (A–D) Total RNA from the indicated strains was fractionated in denaturing gels and subjected to Northern blotting to detect precursor and mature forms of tRNA^{Ser}_{CGA} (A), tRNA^{Leu}_{UAG} (B), U4 snRNA (C), and U3 snRNA (D). As loading controls, blots were reprobed to detect 5S rRNA (A–C) or scR1 RNA (D).

tRNA^{Ser}_{CGA} anticodon stem, all three precursors are more abundant, consistent with a processing defect, and *LHP1* is required for accumulation of the mature tRNA (9). Consistent with data that the La motif and RRM are sufficient for UUU_{OH} binding, both trailer-containing forms of pre-tRNA^{Ser}_{CGA} were detected in strains expressing *lhp1(1-227)* (Fig. 2A). Probing for wild-type tRNA^{Leu}_{UAG} revealed that Lhp1p(1-227) also allowed accumulation of these pre-tRNAs (Fig. 2B). However, while the levels of the pre-tRNA^{Leu}_{UAG} species were unchanged in strains expressing *lhp1(1-227)* (Fig. 2B), quantitation and normalization to the 5S rRNA control revealed that all three mutant pre-tRNA^{Ser}_{CGA} species decreased in the presence of the truncated alleles (Fig. 2A). Specifically, the primary transcript decreased 2.1-fold, the 3' trailer-containing intermediate decreased 5.0-fold, and the end-matured pre-tRNA decreased 1.8-fold compared to their levels with full-length Lhp1p, indicating the C terminus is important for accumulation of the mutant pre-tRNAs.

Surprisingly, probing to detect other precursors revealed that accumulation of some wild-type RNAs requires the Lhp1p C terminus. The pre-U4 spliceosomal snRNA was only detected in the presence of full-length Lhp1p (Fig. 2C), and the two major 3' extended forms of the U3 snoRNA (U3 + 18 and U3 + 12) were replaced in the presence of the truncated protein by slightly shorter, more heterogeneous pre-U3 snoRNAs (Fig. 2D), as previously described for *lhp1Δ* cells (6). Thus, the C terminus is required for the accumulation of some noncoding RNA precursors.

The C terminus is important for aminoacylation of a mutant tRNA^{Arg}.

For many of the mutations that cause yeast to require *LHP1*, the affected RNAs are degraded, making it difficult to determine if Lhp1p influences folding. However, for pre-tRNA^{Arg}_{CCG}, which is the only *S. cerevisiae* tRNA with a mismatch in the anticodon stem, Lhp1p is required for efficient folding at low temperature (16°C) and when the pre-tRNA contains a cold-sensitive mutation (*trr4-1*) that further weakens the anticodon stem, allowing formation of an alternative helix (11). In the absence of Lhp1p, the mature tRNA is mostly uncharged, indicating it is in a non-functional conformation (11). Consistent with a requirement for the C terminus, the growth of *trr4-1* strains is reduced in the pre-

sence of the truncated Lhp1p at 25°C and is nearly undetectable at 16°C (Fig. 1D).

To examine the role of the C terminus, we generated *trr4-1* strains that contained *LHP1* under control of the methionine-repressible *MET3* promoter and also carried plasmids with *LHP1-NLS*, *lhp1(1-227)-NLS*, or the empty vector. After growing the strains in the absence of methionine to allow *LHP1* expression, transcription was repressed by adding 2 mM methionine to the media. Western blotting of cells carrying only the empty vector revealed that Lhp1p became undetectable ~12 hr after methionine addition (Fig. S3). Northern analyses, followed by quantitation and normalization to the tRNA^{Arg}_{UCU} control, confirmed that although the *trr4-1* mutation results in a 7.7-fold decrease in the levels of tRNA^{Arg}_{CCG} (Fig. 3A, lanes 1 and 3), depletion of Lhp1p in the empty vector-containing cells results in only a 2.0-fold further decrease in the mutant tRNA levels (Fig. 3A, lane 11; also ref. 11). Fractionation of the tRNA in acidic acrylamide gels (Fig. 3B), together with quantitation of three independent experiments (Fig. 3C), revealed that ~58% of the mutant tRNA^{Arg}_{CCG} was aminoacylated in all strains when *LHP1* was expressed from the *MET3* promoter. Although the fraction of the charged mutant tRNA in the strain carrying *LHP1-NLS* declined only slightly after methionine addition, reaching 43 ± 7%

Fig. 3. The C terminus is required for efficient aminoacylation of a mutant tRNA^{Arg}_{CCG}. (A) *lhp1Δ* cells and *trr4-1 lhp1Δ* cells containing the indicated plasmids and pMETLHP1 were grown without methionine and switched to 2 mM methionine media at time 0. At intervals, RNA was extracted and subjected to Northern blotting to detect tRNA^{Arg}_{CCG}. As a loading control, the blot was reprobed to detect tRNA^{Arg}_{UCU}. (B) To separate aminoacylated and deacylated tRNAs, RNA was extracted under acidic conditions, fractionated in acidic polyacrylamide gels, and subjected to Northern blotting to detect tRNA^{Arg}_{CCG} and tRNA^{Arg}_{UCU}. Lane 1, deacylated tRNA from *lhp1Δ* cells. As described (11), both the charged and uncharged forms of *trr4-1* tRNA^{Arg}_{CCG} migrate differently in these gels from the wild-type tRNA, consistent with adoption of an altered conformation by the mutant tRNA. (C) Quantification of the fraction of charged tRNA^{Arg}_{CCG} at intervals after methionine addition. Data from three independent experiments are graphed. Error bars are not visible for points where the S.E.M. was ≤1%.

at 24 h, the levels of the charged tRNA in the strain carrying *lhp1* (1-227)-NLS were similar to the strain carrying the empty vector ($16 \pm 0\%$ vs. $14 \pm 1\%$ at 24 h, respectively). We conclude that the Lhp1p C terminus is important for aminoacylation of the mutant tRNA.

The C terminus is Required to Protect a Pre-tRNA Anticodon Stem from Chemical Modification. We tested whether the C terminus contributes to binding of target RNAs in vitro. Binding of Lhp1p and Lhp1p(1-227) proteins to wild-type pre-tRNA^{Arg}_{CCG} revealed that Lhp1p(1-227) bound the pre-tRNA with an affinity ($K_d = 8.8 \pm 0.3$ nM) that was indistinguishable from full-length Lhp1p (Fig. 4A), consistent with results using similar truncated human and *Trypanosoma brucei* La proteins (5, 15). Experiments using pre-U4 RNA revealed that the affinities of Lhp1p and Lhp1p(1-227) for this RNA were also similar [$K_d = 7.7 \pm 2.0$ nM for Lhp1p and 9.8 ± 0.4 nM for Lhp1p(1-227)]. Finally, binding studies using only the C terminus (amino acids 228–275) revealed that even at high protein concentrations (2.4 μ M), the C-terminal fragment did not form a detectable complex with pre-tRNA^{Arg}_{CCG} (Fig. S4). We conclude that, compared to the high affinity 3' end binding site formed by the La motif and RRM (16, 17), the C terminus does not contribute significantly to binding affinity.

Previously, it was found that purified Lhp1p is not sufficient to fully convert the misfolded mutant pre-tRNA^{Arg}_{CCG} structure to that of the wild-type tRNA (11), suggesting that other proteins, such as helicases, contribute in vivo. However, as Lhp1p decreases the accessibility of residues in both the wild-type and mutant pre-tRNA^{Arg}_{CCG} anticodon stemloops to chemical probes (11), we determined if the C terminus was required. As described (11), G36 and G37 in the anticodon loop are accessible to kethoxal, which reacts with single-stranded Gs (Fig. 4B). In addition, the accessibility of G39 and G40, which are components of base pairs in the stem, is enhanced at 37 °C (Fig. 4B and C), possibly due to increased breathing of the fragile stem (11). In the presence of Lhp1p, but not Lhp1p(1-227), G36, G37, G9, and G40 are not

modified (Fig. 4B), indicating that the C terminus is required to prevent the opening of the anticodon stem and/or to protect it from kethoxal (11). Quantitation confirmed that these guanines were less accessible in the presence of Lhp1p (Fig. 4D–F). In addition, Lhp1p, but not Lhp1p(1-227), conferred protection from kethoxal at G17 and G18 in the D-loop (Fig. 4B). As the UUU_{OH} binds in a cleft between the La motif and the RRM (16, 17), our results are consistent with a model in which the C terminus contacts other features of pre-tRNAs, including the anticodon stemloop.

Lhp1p N and C termini Become Protease-Resistant upon RNA Binding.

Previous NMR studies revealed that the C terminus of human La is unstructured (13) and secondary structure prediction programs suggest that the Lhp1p terminus is similarly disordered (22; see Fig. S1). Because unstructured regions are often hypersensitive to proteolysis (23), we examined the protease-sensitivity of Lhp1p. Although Lhp1p has a predicted molecular mass of 32 kDa, it migrates in gels at ~ 38 kDa (24; also Fig. 5A–D). Limited proteolysis with trypsin (Fig. 5A,C, and D) or chymotrypsin (Fig. 5B) results in the production of a 24–25 kDa fragment that was identified by mass spectrometry and N-terminal sequencing to consist solely of the La motif and adjacent RRM, lacking 42–46 C-terminal amino acids and 16–20 amino acids N-terminal to the RRM (Fig. 5 and Table S1). Consistent with the finding that N-terminal amino acids are susceptible to proteolysis, this region is predicted by DISOPRED (22) to be unstructured in Lhp1p and several other La proteins, although not human La (Fig. S1). NMR studies of full length and truncated forms of Lhp1p also supported the lack of well ordered secondary structure for the N and C termini (Fig. S5).

As many unstructured domains exhibit increased resistance to proteolysis on binding of a partner (23, 25), we determined if the C terminus becomes protease-resistant upon RNA binding. On binding pre-tRNA^{Arg}_{CCG} or pre-tRNA^{Phe}_{GAA}, the 24 kDa fragment was mostly replaced by two larger bands (Fig. 5A–D,

Fig. 4. The C terminus is required to protect the pre-tRNA^{Arg}_{CCG} anticodon stem from kethoxal. (A) ³²P-labeled wild-type pre-tRNA^{Arg}_{CCG} was incubated with the indicated concentrations of Lhp1p (left) or Lhp1p(1-227) (right) and fractionated in native gels to separate RNAs from RNPs. Dissociation constants ($\pm$ s.d.) were derived from three independent experiments. As described for Lhp1p (21), upon addition of increasing protein, both proteins form complexes that migrate with slower mobility. As these complexes are preferentially dissociated by competitor RNAs, they may represent binding by a second protein molecule to a less specific site on the pre-tRNA (21). (B) Unlabeled pre-tRNA^{Arg}_{CCG} was incubated in the absence (lanes 1–3) or presence of Lhp1p(1-227) (lanes 4–6) or Lhp1p (lanes 7–9). Kethoxal was added and the reaction incubated an additional 15 min at the indicated temperature. Sites of modification were detected by primer extension; thus, no information was obtained about the 3' end of the RNA. Additional bands which do not correspond to Gs may be due to degradation or secondary structure. (C) Sites of kethoxal modification on pre-tRNA^{Arg}_{CCG} are indicated by arrowheads. (D–F) Phosphorimager quantitation of anticodon stem modifications in the presence of no protein (D), Lhp1p(1-227) (E), or full-length Lhp1p (F).

Fig. 5. The Lhp1p N- and C-termini increase in protease-resistance upon binding pre-tRNAs or U6 snRNA. (A–D) Lhp1p was incubated with either no RNA or the indicated RNAs, followed by digestion with trypsin (A, C, and D) or chymotrypsin (B). In (C), Lhp1p was incubated with pre-tRNA^{Arg} or pre-tRNA^{Arg} lacking the anticodon stemloop. After SDS-PAGE, protein fragments were visualized by Coomassie blue staining. (E and F) Major products from the trypsin (E) and chymotrypsin (F) digests were identified using mass spectrometry and N-terminal sequencing. Squares, fragments consisting of the La motif and RRM. Open circles, bands detected in the presence of pre-tRNAs or U6 snRNA. The largest band is full-length Lhp1p, while the shorter fragment contains 15–19 amino acids N-terminal to the La motif but lacks the C terminus (Table S1).

dots) that were identified as full-length Lhp1p and a fragment lacking the C terminus but containing the N-terminal amino acids (Fig. 5, dots and Table S1). Increased protease resistance did not occur in the presence of *Escherichia coli* tRNA, which does not bind Lhp1p, or a polyuridine-containing oligonucleotide (GCUUUUUUU; Fig. 5A and B) indicating that binding to a larger RNA is required. Consistent with our result that the C terminus is required to protect the pre-tRNA^{Arg} anticodon stem from kethoxal, increased protease resistance did not occur on binding a pre-tRNA lacking the anticodon stem (Fig. 5C). However, Lhp1p also became protease-resistant when bound to U6 snRNA (Fig. 5D) or pre-5S rRNA (Fig. S6.4), indicating the C terminus may similarly contact portions of other RNA targets.

Discussion

Although crystallographic studies have elucidated how La recognizes RNA 3' ends (16, 17), it has been unclear how La can influence RNA structure and biological outcomes at other sites. We discovered that a previously unexplored feature of La proteins, an intrinsically disordered C terminus, is required for the stable accumulation of several noncoding RNA precursors and for the role of the La protein Lhp1p in assisting pre-tRNA folding. Based on NMR data of La from humans (13) and our proteolysis and NMR studies, this region is unstructured in the absence of RNA. As the C terminus becomes protease-resistant upon binding two different pre-tRNAs, U6 snRNA and pre-5S rRNA, but not a pre-tRNA lacking the anticodon stemloop, our experiments demonstrate that this conformationally heterogeneous domain is critical for the ability of La proteins to interact with diverse structural elements at sites distant from the 3' end.

Our results, together with structural studies of human La (16, 17) indicate that while high affinity binding by La to 3' ends

takes place in the cleft formed by the La and RRM, the C terminus contacts additional RNA elements (Fig. S7). Because preliminary experiments in which NMR was performed on a Lhp1p/pre-tRNA complex gave ambiguous results (in part due to the increased size of the complex), additional structural studies of full length La bound to multiple RNA targets will be required to determine if the domain becomes more ordered upon RNA binding and whether it interacts in distinct ways with diverse substrates. Moreover, because structures of human La bound to five different oligonucleotides have revealed plasticity in the conformation of nucleotides near the terminal uridylates (16, 17), both the 3' end binding site and the unstructured C terminus contribute to the ability of La to stabilize diverse RNAs. It is also likely that other portions of La, such as the N terminus (which in Lhp1p also becomes protease-resistant upon RNA binding), the La motif, and the RRM add to La's plasticity. In addition, many metazoan La proteins contain a second RRM, which could contribute to binding some RNAs, such as those mRNAs whose translation may be influenced by human La (5).

How might the C terminus contribute to the multiple roles of Lhp1p in RNA biogenesis? For the pre-U4 snRNA and pre-U3 snoRNAs, we speculate that in addition to the role of Lhp1p in protecting newly synthesized 3' ends from exonucleases (9, 26), interactions with the C terminus could protect other vulnerable sites from endonucleases. For the mutant pre-tRNA^{Arg}, where binding by Lhp1p to the pre-tRNA is required for production of aminoacylated mature tRNA (11), we favor a model in which contacts with the C terminus stabilize the correctly folded anticodon stem. In support of this model, experiments in which an Lhp1p/pre-tRNA complex was digested with high amounts of protease before kethoxal addition revealed that protection of the anticodon stem required bound Lhp1p, indicating Lhp1p functions by stoichiometric binding (11). Moreover, our result that Lhp1p becomes protease-resistant on binding U6 snRNA and pre-tRNA^{Phe}_{GAA}, which are not known to misfold, argues against an alternative model in which the C terminus selectively binds misfolded elements. Our data do not rule out the possibility that the C terminus contributes to correct pre-tRNA folding by loosening structures of kinetically trapped folding intermediates, as proposed in some models for how proteins can act as RNA chaperones (27, 28).

As Lhp1p is removed upon 3' end maturation, other mechanisms must contribute to protecting mature RNAs from nucleases and maintaining misfolding-prone tRNAs in the correct conformations. For some RNAs, specific RNA-binding proteins function redundantly with La to protect the RNAs from nucleases. For example, Lhp1p becomes required for accumulation of the U6 spliceosomal snRNA when yeast contain mutations in the Lsm2-Lsm8 proteins that are components of the mature RNP (4, 5). For pre-tRNAs, both modifications and binding to subsequent proteins such as synthetases likely contribute to maintaining correct folds in vivo, because Lhp1p is important for growth when yeast contain mutations in either the arginyl-tRNA synthetase or the tRNA modifying enzyme *TRM1* (29). Consistent with a redundant function in stabilizing tRNA structure, Trm1p catalyzes dimethylguanosine formation at position 26 in many tRNAs, a modification that prevents formation of alternate anticodon stems (30).

We note that several other proteins that bind multiple RNAs resemble La in that known RNA-binding motifs, such as RRMs, KH domains and/or cold shock domains, are adjacent to regions predicted to be disordered (27, 28) (also Fig. S1). It is possible that some of these proteins, which include the RRM proteins hnRNP A1 and hnRNP C1, the KH domain protein KSRRP, and the Y-box protein YB-1, similarly use their RNA-binding motifs to bind specific ligands with high affinity and their disordered regions to interact with other structural elements in ways that profoundly influence functional outcomes.

Materials and Methods

Yeast Strains, Media, and Plasmids. Yeast media was as described (31). Strains and plasmids are listed in Table S2. Construction of plasmids is described in *SI Text*.

Immunoblotting, Immunofluorescence, and Northern Analyses. For immunoblotting, cells were lysed in 40 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.05% NP-40, 0.5 mM phenylmethylsulfonyl chloride (PMSF) by vortexing with glass beads. After sedimenting at $1,500 \times g$, extracts were subjected to Western blotting with anti-Lhp1p (24) or anti-Sro9p (32) antibodies as described (24). Immunofluorescence was as described (21). Monoclonal anti-Nop1p was a gift of J. Aris (University of Florida, Gainesville). Northern blotting was as described (29). To examine aminoacylation, RNA was extracted at low pH and fractionated in acidic polyacrylamide gels (33). Oligonucleotide probes are listed in Table S2.

Lhp1p Depletion. Strain SK130.1 carrying pMETLHP1 (11) was transformed with plasmids pLHP1-NLS, plhp1(1-227)-NLS and pRS314 to create strains NK5, NK6, and NK7, respectively. Strains were grown in synthetic complete medium lacking histidine, tryptophan (SC-his-trp), and methionine at 25 °C to $OD_{600} = 0.3$. Cells were then diluted into SC-his-trp containing 2 mM methionine and grown for 36 h. Cultures were diluted to keep the OD_{600} below 0.3.

Purification of Lhp1 Proteins. BL21 cells (Novagen) containing plasmids (*SI Text* and Table S2) were grown to $OD_{600} = 0.5$ and induced with 1 mM IPTG for 4 h at 37 °C. After lysing in buffer A (50 mM Tris HCl pH 8.0, 3 mM $MgCl_2$, 0.1 mM EDTA, 1 mM DTT, 10% glycerol), extracts were sedimented at $20,000 \times g$ for 30 min and passed through a 0.2 μm filter. Both Lhp1p and Lhp1p(1-227) were purified as described for Lhp1p (9). For Lhp1p(228-275), which carried a Strep-Tag, extracts were applied to Strep-Tactin Superflow (Qiagen) and eluted with 5 mM desthiobiotin in buffer A. NMR analyses are described in *SI Text*.

Electrophoretic Mobility Shift Assays. Plasmids containing pre-tRNA^{Arg}_{CCG} and pre-U4 RNA behind T7 promoters were linearized with DraI and transcribed (15). After gel purification, RNAs were resuspended in 50 mM Tris pH 8.0, 10 mM $MgCl_2$, and refolded by heating to 70 °C for 1 min and slow cooling

to 23 °C. Binding assays were as described (15). The fraction of bound RNA(θ) was quantitated using a PhosphorImager (Molecular Dynamics). ImageQuant was used to fit the data by nonlinear least-squares analysis as a function of protein concentration to the equation $\theta = [\text{protein}]/([\text{protein}] + K_d)$, where K_d is the dissociation constant.

Protease Protection. 62.5 pmol of Lhp1p was incubated with 300 pmol of pre-tRNA^{Arg}_{CCG}, pre-tRNA^{Phe}_{GAA}, U6 snRNA, the oligonucleotide 5'-GCUUUU UUU-3' (Dharmacon), or 10 μg E. coli tRNA (~400 pmol) in a volume of 10 μL for 10 min. Trypsin or chymotrypsin was added to 10 $\mu g/mL$ and incubated for 1 h at 23 °C. Proteolysis was stopped by adding SDS-PAGE loading buffer and boiling for 10 min. Although experiments using lower RNA amounts (120 and 180 pmol) resulted in identical protease protection (Fig. S6B), reactions were performed at the higher amount of RNA to ensure that all Lhp1p was RNA-bound. For mass spectrometry, reactions were scaled up 10-fold and stopped by adding formic acid to 1%. MALDI-TOF and ESI-TOF were performed by the Scripps Center for Mass Spectrometry. Fragment identities were assigned with PAWS (Genomic Solutions). For N-terminal sequencing, gels were transferred to PVDF membranes, stained with Coomassie blue and subjected to N-terminal sequencing at the Tufts Proteomics Facility.

Kethoxal Modification. To maximize correct folding of pre-tRNA^{Arg}_{CCG}, the in vitro-transcribed pre-tRNA was isolated on DEAE-Sephacose without denaturants (11). Kethoxal modification was as described (11). Briefly, 0.68 pmol of pre-tRNA^{Arg}_{CCG} was incubated with 40 $\mu g/mL$ E. coli tRNA and 6.8 pmol of Lhp1p, Lhp1p(1-227), or no protein in modification buffer for 15 min. After adding 3 μL of kethoxal stock (37 $\mu g/mL$), the reaction was incubated at 22 or 37 °C for 15 min. After phenol extraction and ethanol precipitation, modifications were detected by extending a primer complementary to nt 55–89 of the pre-tRNA.

ACKNOWLEDGMENTS. We thank Gang Dong and Karin Reinisch for advice, Yousif Shamoo, David Brow, and John Aris for reagents, and Susan Baserga, Karin Reinisch, and Elisabetta Ullu for comments on the manuscript. This work was supported by National Institutes of Health Grants GM048410 to S.L.W. and CA108992 to M.E.H.

- Michlewski G, Guil S, Semple CA, Caceres JF (2008) Posttranscriptional regulation of miRNAs harboring conserved terminal loops. *Mol Cell* 32:383–393.
- Trabucchi M, et al. (2009) The RNA-binding protein KSRP promotes the biogenesis of a subset of microRNAs. *Nature* 459:1010–1014.
- Kohn K, Izumi H, Uchiumi T, Ashizuka M, Kuwano M (2003) The pleiotropic functions of the Y-box-binding protein, YB-1. *Bioessays* 25:691–698.
- Wolin SL, Cedervall T (2002) The La protein. *Annu Rev Biochem* 71:375–403.
- Bayfield MA, Yang R, Marais RJ (2010) Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* 1799:365–378.
- Kufel J, et al. (2000) Precursors to the U3 small nucleolar RNA lack small nucleolar RNP proteins but are stabilized by La binding. *Mol Cell Biol* 20:5415–5424.
- Xue D, Robinson DA, Pannone BK, Yoo CJ, Wolin SL (2000) U snRNP assembly in yeast involves the La protein. *EMBO J* 19:1650–1660.
- Inada M, Guthrie C (2004) Identification of Lhp1p-associated RNAs by microarray analysis in *Saccharomyces cerevisiae* reveals association with coding and noncoding RNAs. *Proc Natl Acad Sci USA* 101:434–439.
- Yoo CJ, Wolin SL (1997) The yeast La protein is required for the 3' endonucleolytic cleavage that matures tRNA precursors. *Cell* 89:393–402.
- Arhin GK, Shen S, Perez IF, Tschudi C, Ullu E (2005) Downregulation of the essential *Trypanosoma brucei* La protein affects accumulation of elongator methionyl-tRNA. *Mol Biochem Parasitol* 144:104–108.
- Chakshumathi G, Kim SD, Robinson DA, Wolin SL (2003) A La protein requirement for efficient pre-tRNA folding. *EMBO J* 22:6562–6572.
- Costa-Mattioli M, Vitkin Y, Sonenberg N (2004) La autoantigen is necessary for optimal function of the poliovirus and hepatitis C virus internal ribosome entry site in vivo and in vitro. *Mol Cell Biol* 24:6861–6870.
- Jacks A, et al. (2003) Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif coupled to a helical nuclear retention element. *Structure* 11:833–843.
- Alfano C, et al. (2004) Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nat Struct Mol Biol* 11:323–329.
- Dong G, Chakshumathi G, Wolin SL, Reinisch KM (2004) Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. *EMBO J* 23:1000–1007.
- Teplova M, et al. (2006) Structural basis for recognition and sequestration of UUU(OH) 3' termini of nascent RNA polymerase III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* 21:75–85.
- Kotik-Kogan O, Valentine ER, Sanfelice D, Conte MR, Curry S (2008) Structural analysis reveals conformational plasticity in the recognition of RNA 3' ends by the human La protein. *Structure* 16:852–862.
- Bayfield MA, Marais RJ (2009) Precursor-product discrimination by La protein during tRNA metabolism. *Nat Struct Mol Biol* 16:430–437.
- Rosenblum JS, Pemberton LF, Blobel G (1997) A nuclear import pathway for a protein involved in tRNA maturation. *J Cell Biol* 139:1655–1661.
- Pannone BK, Xue D, Wolin SL (1998) A role for the yeast La protein in U6 snRNP assembly: evidence that the La protein is a molecular chaperone for RNA polymerase III transcripts. *EMBO J* 17:7442–7453.
- Long KS, et al. (2001) Phosphorylation of the *Saccharomyces cerevisiae* La protein does not appear to be required for its functions in tRNA maturation and nascent RNA stabilization. *RNA* 7:1589–1602.
- Ward JJ, McGuffin LJ, Bryson K, Buxton BF, Jones DT (2004) The DISOPRED server for the prediction of protein disorder. *Bioinformatics* 20:2138–2139.
- Receveur-Brechot V, Bourhis JM, Uversky VN, Canard B, Longhi S (2006) Assessing protein disorder and induced folding. *Proteins* 62:24–45.
- Yoo CJ, Wolin SL (1994) La proteins from *Drosophila melanogaster* and *Saccharomyces cerevisiae*: a yeast homolog of the La autoantigen is dispensable for growth. *Mol Cell Biol* 14:5412–5424.
- Dyson HJ, Wright PE (2005) Intrinsically unstructured proteins and their functions. *Nat Rev Mol Cell Biol* 6:197–208.
- Copela LA, Fernandez CF, Sherrer RL, Wolin SL (2008) Competition between the Rex1 exonuclease and the La protein affects both Trf4p-mediated RNA quality control and pre-tRNA maturation. *RNA* 14:1214–1227.
- Tomba P, Csermely P (2004) The role of structural disorder in the function of RNA and protein chaperones. *Faseb J* 18:1169–1175.
- Rajkowitz L, et al. (2007) RNA chaperones, RNA annealers and RNA helicases. *RNA Biol* 4:118–130.
- Copela LA, Chakshumathi G, Sherrer RL, Wolin SL (2006) The La protein functions redundantly with tRNA modification enzymes to ensure tRNA structural stability. *RNA* 12:644–654.
- Steinberg S, Cedergren R (1995) A correlation between N2-dimethylguanosine presence and alternate tRNA conformers. *RNA* 1:886–891.
- Sherman F (1991) Getting started with yeast. *Methods Enzymol* 194:3–21.
- Sobel SG, Wolin SL (1999) Two yeast La motif-containing proteins are RNA-binding proteins that associate with polyribosomes. *Mol Biol Cell* 10:3849–3862.
- Varshney U, Lee CP, RajBhandary UL (1991) Direct analysis of aminoacylation levels of tRNAs in vivo. Application to studying recognition of *Escherichia coli* initiator tRNA mutants by glutamyl-tRNA synthetase. *J Biol Chem* 266:24712–24718.