OPEN ACCESS

KING'S College LONDON

King's Research Portal

DOI:
10.1002/wrna.1430

Document Version
Peer reviewed version

Link to publication record in King's Research Portal

Citation for published version (APA):
Maraia, R. J., Mattijssen, S., Cruz-Gallardo, I., & Conte, M. R. (2017). The La and related RNA-binding proteins (LARPs): Structures, functions, and evolving perspectives. *Wiley Interdisciplinary Reviews: RNA*, [e1430]. https://doi.org/10.1002/wrna.1430

Citing this paper
Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

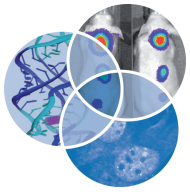
General rights
Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

•Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
•You may not further distribute the material or use it for any profit-making activity or commercial gain
•You may freely distribute the URL identifying the publication in the Research Portal

Take down policy
If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 26. Oct. 2023

Advanced Review



The La and related RNA-binding proteins (LARPs): structures, functions, and evolving perspectives

Richard J. Maraia,[1,2*] Sandy Mattijssen,[1] Isabel Cruz-Gallardo[3] and Maria R. Conte[3*]

La was first identified as a polypeptide component of ribonucleic protein complexes targeted by antibodies in autoimmune patients and is now known to be a eukaryote cell-ubiquitous protein. Structure and function studies have shown that La binds to a common terminal motif, UUU-3′-OH, of nascent RNA polymerase III (RNAP III) transcripts and protects them from exonucleolytic decay. For precursor-tRNAs, the most diverse and abundant of these transcripts, La also functions as an RNA chaperone that helps to prevent their misfolding. Related to this, we review evidence that suggests that La and its link to RNAP III were significant in the great expansions of the tRNAomes that occurred in eukaryotes. Four families of La-related proteins (LARPs) emerged during eukaryotic evolution with specialized functions. We provide an overview of the high-resolution structural biology of La and LARPs. LARP7 family members most closely resemble La but function with a single RNAP III nuclear transcript, 7SK, or telomerase RNA. A cytoplasmic isoform of La protein as well as LARPs 6, 4, and 1 function in mRNA metabolism and translation in distinct but similar ways, sometimes with the poly(A)-binding protein, and in some cases by direct binding to poly(A)-RNA. New structures of LARP domains, some complexed with RNA, provide novel insights into the functional versatility of these proteins. We also consider LARPs in relation to ancestral La protein and potential retention of links to specific RNA-related pathways. One such link may be tRNA surveillance and codon usage by LARP-associated mRNAs. © 2017 Wiley Periodicals, Inc.

How to cite this article:
WIREs RNA 2017, e1430. doi: 10.1002/wrna.1430

INTRODUCTION AND BACKGROUND

Brief Overview Including General Properties of La and Related Proteins

The La motif (LaM) was established as a protein fold coincident with the emergence of the Eucarya and became readily associated with a downstream RNA recognition motif (RRM), together comprising the La module[1] (Figure 1). The La module in turn became allied to additional motifs as members of the major families of La-related proteins (LARPs) 7, 6, 4, and 1 emerged and diversified with specialized functions in the eukaryotes.[1,17,18] Developments in

*Correspondence to: maraiar@mail.nih.gov; sasi.conte@kcl.ac.uk

[1]Intramural Research Program, *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD, USA

[2]Commissioned Corps, US Public Health Service, Rockville, MD, USA

[3]Randall Division of Cell and Molecular Biophysics, King's College London, London, UK

Conflict of interest: The authors have declared no conflicts of interest for this article.

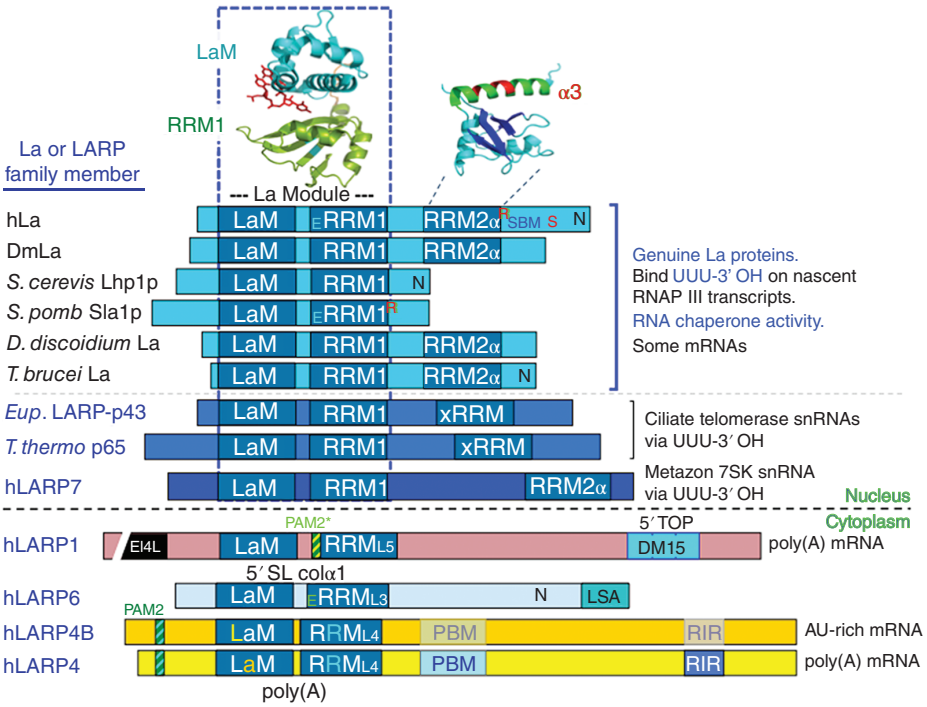


FIGURE 1 | Schematic representation of architectures of La proteins and selected La-related proteins (LARPs). The La motif (LaM) and RNA recognition motif (RRM) comprise the La module. The PDBs used to depict the three-dimensional structures for the La module and hLa RRM2α are 2VOD and 1OWX, respectively.[2,3] Other symbols refer to the following: N, nuclear localization (import) sequence[4–8]; E, nuclear export sequence[5,9,10]; R, nuclear retention element[9,10]; S, Serine-366[11,12]; SBM, short basic motif important for recognition of 5′ pppG of nascent RNA and nucleolar localization[11,13,14]; DM15, important for direct binding to 7mGpppC-cap-5′TOP motif.[15] The RRM2α, xRRM, and DM15 RNA-interaction motifs/domains are reviewed in a separate section. PAM2, poly(A)-binding protein interaction motif-2; PAM2*, LARP1-associated PAM2 candidate with atypical features (see text); LSA, LaM and S1-associated motif; PBM, poly(A)-binding protein interaction protein motif; RIR, RACK1-interaction region. EI4L in LARP1 refers to eIF4-like region.[16] The different species' La proteins depicted are referred to in the text; for simplicity, only the human versions of the LARPs (hLARP) are shown.

basic and in some cases clinical-related research on La and LARPs over the past few years have advanced our understanding to a point where we can begin to appreciate how commonality and individuality contribute to their unique functions.

LARPs encompass a range of functions. Some have adopted functions specific to a limited set of RNA target ligands, for example, LARP7 members for specific small nuclear RNAs (snRNAs). Accumulating evidence indicates that others can coregulate subsets of mRNAs that produce functionally related proteins, e.g., LARP1 for the 5′TOP mRNAs that encode ribosomal proteins as well as other mRNAs,[16,19,20] according to the posttranscriptional RNA regulon model.[21,22] Human LARP6 (hLARP6) targets a conserved stem-loop (SL) motif found in mRNAs that encode α-collagens I and III to coordinate their translation.[23,24] LARPs 4 and 4B appear to coordinate somewhat larger sets to promote stability and translation of their mRNA targets.[25–27]

Due to space limitations, we can neither review all LARP progress nor offer highly detailed critiques. Our goal is an overview preceded by adequate background to provide what we believe are emerging conceptual advances. Central to the LARPs is the La module RNA-binding unit, to which other protein motifs have been added in an LARP-specific manner. These include other RRMs and/or motifs that interact with other proteins, some of which are common to multiple LARPs. This has led to a model in which the La module has adapted, structurally and functionally, to befit the LARP-specific function at hand. This overview contains sections focused on each of the LARP families. However, several themes are generally applicable and relevant information in other sections may not be referred to by LARP-specific nomenclature.

Genuine La proteins (Figure 1) exhibit sequence and length specificity for UUU-3′-OH, the terminal motif common to transcripts synthesized by RNA polymerase (RNAP III).[28] Studies on La proteins from protists, yeasts, frogs, mouse, and human cells indicate that its function is to protect nascent RNAP III transcripts from untimely 3′ exonucleolytic

© 2017 Wiley Periodicals, Inc.

digestion and for some of these it also helps prevent their misfolding by virtue of a separate chaperone activity reviewed in Refs 17 and 29. Although general RNA chaperone functions have been attributed to the La modules of La and the LARPs,[30] independent lines of evidence suggest that motifs in the C-terminal regions of La proteins and some LARPs can also contribute to these activities[31] including in mRNA translation[32] (below).

Apparently early in eukaryal evolution, a second RRM (termed RRM2) emerged downstream of the La module and this may have been the gene arrangement that gave rise to La and LARP7 members (see figure 5 in Ref 1). LARP7 family members are the closest relatives to genuine La in sequence, architecture, and function. The RRM2 of human La (hLa) has an extra structural element that is appended onto the canonical RRM fold, namely the α3 helix that lies over its β-sheet surface.[2] Although α-helices located on top of the central β-sheet comprise one of the known variations of the RRM protein fold,[33] the distinctive elements and functional significance attributed to the RRM2 α3 helix make it a distinguishing feature, as described in a later section; we shall hereafter refer to the RRMs2 of La and metazoan LARP7 as RRM2α. The RRM2α would appear to have been an ancient part of La, as it is found in the proteins from extant representatives of phylogenetically deep-rooted eukaryotes, *Trypanosoma*, *Leishmania*, *Giardia*, and *Dictyostelium* species[34] (Figure 1 and below).

A later section will review structural details and functional attributes of p65, the LARP7 member in the protist ciliate, *Tetrahymena thermophila*, in which this type of RRM2α, designated as xRRM, contributes to telomerase (TER) RNA folding and hierarchical assembly of the ribonucleic protein (RNP) complex.[35,36] However, La proteins which lack an RRM2α such as yeast Lhp1 contain a C-terminal extension that can assist pre-tRNAs in correct folding[31] (below).

Thus, with two versions of both La and LARP7, each with and without an RRM2α as residents in early eukaryotes, see figure 5 in Ref 1, the evolutionary stage would have been set for emergence of additional LARP architectures. The genuine La proteins of human and yeast contain multiple intracellular trafficking elements that control nuclear, cytoplasmic, and nucleolar distribution[4,9,10,13,37] (Figure 1). Unlike La and LARP7, which are predominantly nuclear and associated with noncoding snRNAs, the LARPs 6, 4, and 1 appear mostly cytoplasmic at steady state as mRNA-associated proteins. It is noteworthy here that majority distribution at steady state does not preclude dynamic nuclear–cytoplasmic shuttling. Indeed, *Schizosaccharomyces pombe* (*S. pombe*) and hLa proteins are mostly (~80%) nuclear but shuttle by virtue of conserved nuclear import and export elements, the latter of which map to the α-helical backside surfaces of their RRMs1, and are controlled by a conserved nuclear retention element in the α3 helix of human RRM2α and of a comparable region of RRM1 of *S. pombe* La[9,10] (Figure 1). This reflects on the versatility of the RRM as a multifunctional folding platform.[38,39] These elements in La constitute a conserved circuit for presumably coordinating RNA binding and subcellular trafficking.[6,14]

Nucleo-cytoplasmic shuttling elements were also identified in hLARP6/Acheron.[40] A classic Crm1-consensus nuclear export sequence that was functionally mapped[40,41] to residues that align to the β1 strand of the RRM of the La module[1] was indeed found as part of β1 in the solution structure of hLARP6[23] ('E' in Figure 1). Other LARPs also use RNA-binding modes that are juxtaposed to subcellular trafficking elements, other interaction motifs, and/or regulatory/signaling elements (below).

Another commonality among several LARPs is the interaction with poly(A)-binding protein (PABP), which is clear for hLARPs 1, 4A, and 4B, and *Arabidopsis thaliana* (At)LARPs, 6b and 6c.[18,25,26,42,43] As will be detailed below, existing and emerging data indicate that the La modules of hLARPs 1 and 4 and plant 6c exhibit preferential binding to poly(A) RNA itself[18,26] and for hLARP4B, A-rich RNA.[27] Therefore, a theme for the cytoplasmic LARPs is the association with the 3′ end poly(A) region of their substrate mRNAs, via PABP and the mRNA 3′ UTR or the poly(A) tail, in any case a polar orientation reminiscent of La and LARP7 on the UUU-3′-OH ends of their snRNA ligands. From this another common theme follows: mRNA circularization, first proposed for LARP4B,[25] mediated by simultaneous interactions with the 5′ and 3′ regions. We will review emerging data on how the different LARPs manage this. mRNA circularization promotes ribosome cycling and ensures that only intact mRNAs, i.e., with a 5′ m7GpppN-cap and 3′ poly(A) tail are efficiently translated, and this is commonly referred to as the closed-loop model of translation.[44–46]

The La Protein La Module and RRM2α Existed in Ancient Eukaryotes and Persists Widely in Extant Species

Recent advances relevant to the phylogenetics of La came from two sources, a study that sought to

© 2017 Wiley Periodicals, Inc.

uncover the ancestral gene repertoire of the most ancient animal stem cells,[47] and secondary structure examination of La protein from the amoebozoa *Dictyostelium discoideum*,[34,48] which represents one of the earliest known branches from the last common ancestor of all free-living eukaryotes.[49] *D. discoideum* La protein contains a typical La module comprised a LaM and RRM1 as well as an RRM2α with a short β4 strand followed by an α3 helix.[34,48] The La protein of the intracellular parasite *Trypanosoma brucei* was also known to contain a typical La module.[7,50] The report of RRM2α in *D. discoideum* La prompted further examination of more phylogenetically distant La proteins of some of the simplest eukaryotes of the *Trypanasoma*, *Leishmania*, and *Giardia* species. We began with a JPred-4 sequence structure prediction analysis of the C-terminal region of *T. brucei* La. This predicted a characteristic RRM2α with an additional β strand following β4 (β4′) and an extended α3 (Figure 2(a)), in agreement with the RRM2α of hLa whose high-resolution structure has been determined[54] (Figure 2 (b) and see below). The protein sequences of other intracellular parasites, from two *Leishmania* species as well as two parasites, *Giardia* species were aligned with other La proteins and compared to the RRM2α primary and tertiary structures from hLa (Figure 2 (c)). The sequence mapping to hLa RRM2 helix α3 is enclosed in the orange-green box and its salient conserved residues are annotated by asterisks in Figure 2 (c). Although not shown in Figure 2(c), the *bona fide* La protein in At, AtLa1 also has an RRM2 (as well as La homologs in other plants)[55] bearing the key conserved residues in helix α3 (not shown).

The alignment also revealed that in several of the La proteins of the deep-rooted eukaryotes, the RRM2α is followed by a conserved GK/R-rich region (Figure 2 (c), blue rectangle), also conserved in AtLa1 (not shown). In hLa, this region is termed the short basic motif (SBM, Figure 1) which was proposed to share sequence features with a Walker A-like nucleotide-binding motif[56] and was later found to recognize the 5′-ppp triphosphate terminus of nascent pre-tRNAs and whose binding is attenuated by phosphoSer-366 to control 5′ processing by RNAse P.[11,57]

It is noteworthy that several of the La proteins, including AtLa1 (not shown), contain multiple copies of the potential arginine methylation motif, GRR/GR[58] in this region. We also note that this part of hLa overlaps with the sequence recently found to be required for cyclin D1-internal ribosome entry site (CCND1-IRES) mRNA chaperone activity.[32]

As *Trypanosoma*, *Leishmania*, *Giardia*, and *Dictyostelium* are more deep rooted on the phylogenetic tree than are the yeasts, see figure 5 in Ref 49, which lack an RRM2, these observations suggest that RRM2-GR/K existed widely in eukaryotes and that the yeast/fungal lineage lost their RRM2 (Figure 1). However, the yeasts *Saccharomyces cerevisiae* Lhp1 and *S. pombe* Sla1, as well as other La proteins have intrinsically unstructured regions in their C-terminal regions.[31] For Lhp1, the disordered sequence bears some resemblance to GR/GK region and was shown to function in post-transcriptional biogenesis of tRNA and other cellular RNAs.[31] For hLa, this disordered region predicted by the ISOPRED2 server[31] and confirmed by nuclear magnetic resonance (NMR) and circular dichroism experiments[2] includes the SBM GR/GK region and mRNA chaperone domain.[32]

The second phylogenetic observational source referred to above is also related to *Dictyostelium*. *D. discoideum* can undergo regulated differentiation of homogeneous cells into distinct types, and many of the genes involved in the process were inherited by the metazoa.[49] Toward defining a relevant ancestral gene repertoire of animal stem cells, Alie et al. found that of the 44 conserved RNA-binding proteins (RBPs) involved, La was among the 25 that were previously shown to be expressed in, and in the case of La, essential for the survival of,[59] mouse ES cells.[47] In the ES cell, essential group of 25 were a number of RBPs involved siRNA/piRNA and Piwi-related pathways.[47] In this regard, it is notable that La in animal cell nuclei is important for preventing nascent pre-tRNAs with potential for folding into alternate structures from entering the miRNA processing pathway.[60] These observations add support to the idea that ancient La was critical during early eukaryotic evolution and was fixed as a fundamental cellular component thereafter. As an ancient RBP with a complexity of RNA binding and subcellular trafficking motifs, it should not be surprising that La function would extend to multiple aspects of RNA metabolism and that the La module (and RRM2α) would be appended to other motifs in the LARPs.

LA IS A EUKARYOTIC PRE-tRNA CHAPERONE THAT SUPPORTS tRNAome DIVERSITY

La Acts on RNAP III Transcripts

La is a eukaryote-specific factor whose phylogenetic emergence is consistent with co-evolutionary appearance with RNAP III in an ancient common ancestor prior to eukaryotic radiation.[1,61,62] The term tRNAome used here refers to the tRNA gene content

© 2017 Wiley Periodicals, Inc.

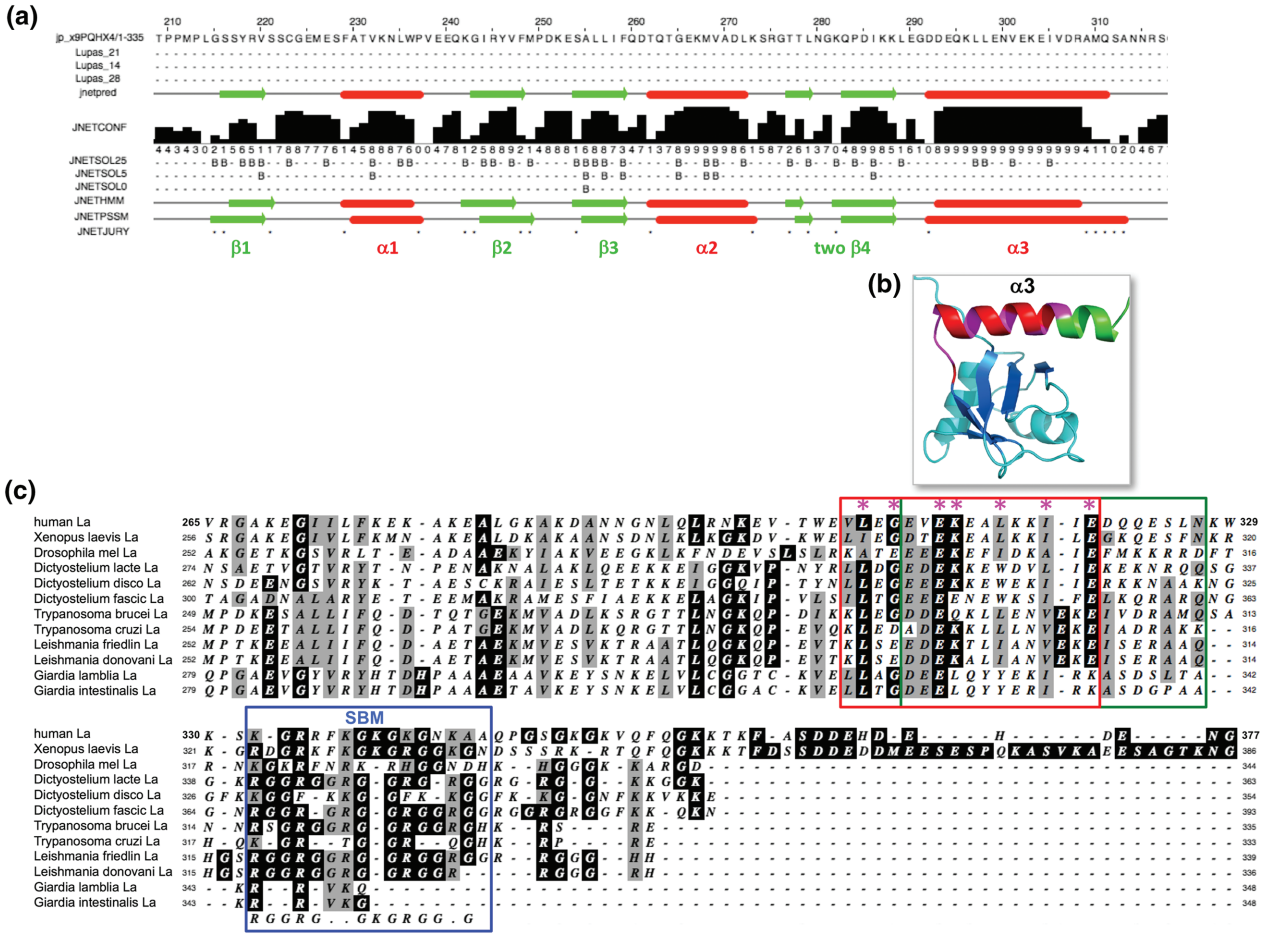


FIGURE 2 | RRM2α and short basic motif (SBM) regions of La proteins representing early eukaryotes. (a) Results of secondary structure prediction for sequence of putative RRM2 region of T. brucei La protein (GenBank: AAF34598.1) performed by Jpred-4.[51] The secondary structure elements are indicated below in the sequence: α-helices depicted by red cylinders and β-strands by green arrows. (b) Solution structure of hLa RRM2α shown for reference (PDB 1OWX),[2] with α3 colored to match underlying (c): red to coincide with boundaries of red rectangle, magenta to match conserved positions demarcated by asterisks, and green to coincide with rectangle downstream sequence boundary in C. (c) Alignment of La protein sequences beginning from the approximate middle of RRM2 extending toward their C-termini. The overlapping green and orange rectangles represent the sequences aligning with hLa α3 region of RRM2α and a conserved sequence block, respectively. Asterisks above the hLa sequence demarcate conserved residues, the first three of which reside in the tract connecting β4′ to the α3 helix, colored in magenta in the structure in (b). The blue rectangle encompasses the SBM region described for hLa (see text).

of a genome. Most tRNA genes in bacteria and archaea are included in operons with other RNA types, all of which are under the control of a single RNAP.[63,64] In contrast, all eukaryotes divide transcription among three RNAPs, I, II, and III, with the tRNAs further distinguished as individual transcription units under the dedicated control of RNAP III. The large ribosomal RNA is synthesized by RNAP I, while RNAP II synthesizes the mRNAs, as well as long and short noncoding regulatory RNAs. RNAP III synthesizes only short RNAs and does so with high efficiency for its most abundant products, the tRNAs.[65,66] For example, the *S. cerevisiae* RNAP III transcriptome is limited to just seven or so types of genes for conserved small RNAs: U6 spliceosomal snRNA, scR1 small cytoplasmic RNA of the signal recognition particle, the RPR1 RNase P RNA, NME1 RNase MRP RNA, snr52 a small nucleolar RNA, the 5S rRNA, and the tRNAs (and a few tRNA-like transcripts, *TLT*s1-6, of unknown function).[67] Except for the tRNAs and 5S rRNA, the others are single-copy genes that encode RNAs that undergo minimal if any processing as compared to pre-tRNAs and become stably associated with one or more polypeptides in their corresponding RNPs. In contrast, *S. cerevisiae* tRNAs are synthesized from 275 individual genes that are widely dispersed to all chromosomes and occupy the great majority of RNAP III.[67–69]

© 2017 Wiley Periodicals, Inc.

Nascent pre-tRNAs undergo the most complex processing and maturation process of the RNAP III transcripts, requiring multiple cleavage and modification activities. tRNA biogenesis is further elaborated by their collective sequence diversity and potential to form alternate structures yet they must compete for several shared processing activities and other limiting factors which include the first protein they interact with, La, and certain tRNA modification enzymes.[66,70–73]

With 17 integral subunits, RNAP III is specialized for efficient termination–reinitiation recycling and high production of short RNAs.[74,75] In all eukaryotes examined, the critical control element for transcription termination of RNAP III-transcribed genes is the oligo(dT) tract following the coding sequence, which in the nascent RNA becomes the high-affinity UUU-3′-OH motif to which La binds. Three RNAP-III-specific subunits, C53, C37, and C11 contribute to termination and efficient reinitiation.[72,75–77] The C11 subunit is a RNAP-III-intrinsic RNA 3′ cleavage factor that can impose 3′-oligo(U) length differences on terminated transcripts.[72] Thus, termination is a critical feature of RNAP III, and La is mechanistically directed to survey the termini of its newly synthesized transcripts.[29] La binding is oligo(U) length dependent[28,72] and on this basis, it sorts or channels nascent RNAP III transcripts to different processing and maturation pathways reviewed in Ref 29. In fission yeast, La is limiting relative to the pool of nascent RNAP III transcripts so the efficacy with which a pre-tRNA competes for it is dependent on the length of its oligo(U) tract and abundance.[72] This was demonstrated *in vivo* by genetic mutations in C11 that cause increase in the 3′ oligo(U) length of the nascent transcript for a suppressor-tRNA.[72] By increasing 3′ oligo(U) length, a pre-tRNA can be more efficiently bound by La, more efficiently processed to mature tRNA, and lead to higher levels of tRNA-mediated suppression.[72] Thus although La is very abundant, it can be functionally limiting, because the rate of nascent transcript synthesis by RNAP III is exceedingly high.[78]

While La interacts with UUU-3′-OH on the spectrum of RNAP III transcripts, its most functional impact is on the pre-tRNAs.[70,79] This distinction reflects that La has two activities that confer biological function, 3′ end protection from exonucleases[80–82] and RNA chaperone[17] the latter of which differentially benefits pre-tRNAs. Genome-wide screens in yeast uncovered multiple tRNA genes in which single-nucleotide substitutions cause cells to require La for growth.[31,70,79,80,83] To our knowledge, no gene for a non-tRNA RNAP III transcript has been uncovered this way although mutations in genes encoding proteins involved in U6 snRNP assembly have been identified.[84,85] That La can shield scR1 and other non-tRNA RNAP III transcripts from untimely decay reflects its 3′ end protection activity.[82,86] Also, while pre-5S rRNA is the next most abundant RNAP III transcript after tRNAs, it undergoes little end processing and is largely independent of La, e.g., see Ref 60. Thus, pre-tRNAs are the RNAP III genes whose transcripts would appear to be the critical beneficiaries of La chaperone activity.

La function extends beyond 3′ end protection for pre-tRNAs by preventing misfolding of those with propensity to form alternate structures.[31,70,79,83,87] Structurally challenged pre-tRNAs succumb to nuclear surveillance-mediated decay in the absence of La.[80,81,87] The RNA chaperone activity is important for yeast growing at low temperature, due to insufficient amino acid charging of one or more susceptible tRNAs in the absence of La.[79] As reviewed above, some tRNA gene alleles have been found to be conditionally dependent on La for efficient maturation of their transcripts.

Molecular Chaperones Can Mask Mutations in Their Substrates

Chaperones can buffer mutations in their substrates and act as capacitors by allowing mutation-bearing gene alleles to exist as a means to evolutionary change.[88,89] The highly abundant molecular chaperone, Hsp90, assists nascent polypeptides in avoiding 'off-pathway' folding events and in achieving their correctly folded form. Beyond this important function, it has been established that Hsp90 additionally serves as a capacitor of genetic variation that can affect evolutionary change by buffering polymorphisms.[90] This results from the abundance and power of Hsp90 chaperone activity such that it can mask mutations in its substrate polypeptides that would otherwise disrupt their correct folding. Any number of these polymorphisms can accumulate throughout the genome in a variety of independent genes, because they are silent under the protective activity of the chaperone. However, when future conditions change including when the amount of Hsp90 is decreased relative to its substrate load, some of these mutations will be unmasked and if beneficial (or detrimental) may become deterministic toward phenotype.[90] Thus, the chaperone acts as a 'capacitor,' serving to store the genetic variations and release them.[90] This has indeed been observed in wild

© 2017 Wiley Periodicals, Inc.

populations of genetically distinct *Drosophila*; as Hsp90 is debilitated, different phenotypes emerge in different populations due to unmasking of a number of polymorphisms in otherwise unrelated genes,[91] linked by the fact that they share the same chaperone.

La is a chaperone that has been reproducibly shown to mask mutations in tRNAs that would otherwise disrupt their structure and lead to their decay.[50,72,79–81,87] La can also mask mutations in tRNA modification enzymes, without which their tRNA substrates fail to accumulate.[92–94] It is therefore reasonable to consider that La may have a similar but unique relationship with tRNA genes as Hsp90 has with its substrates, unique because of the natural high copy number, inherent variability, and propensity for amplification of tRNA genes. As tRNA gene number appears to be dynamic in eukaryotes (below), the potential storage of variation could be large.

tRNAome Gene Expansions in Eukaryotes

Comprehensive analysis found that the only eukaryotic group lacking a genuine La homolog comprised the intracellular parasites, *Plasmodia*.[1] Most *Plasmodium* species have only 40–60 tRNA genes, and the rest a maximum of 78,[95] similar to many bacteria and archaea which typically contain 45–85.[96] By comparison, free-living eukaryotes such as yeast contain 200–300 tRNA genes while plants and animals typically contain 400–700, although quite a few contain several thousand.[96] Thus, tRNA gene numbers can vary very widely among related species[97] and even among members of a species.[98–101] tRNA genes number from 171 to 322 in four fission yeast *Schizosaccharomyces* species despite very high similarity in all other gene types, overall architecture, and other components of their genomes.[99] tRNA diversity accompanies differential tRNA gene amplification including appearance of new isoacceptors.[99,102] Although the mechanisms of tRNA gene amplification are unknown, encounters between DNA replication forks and stable RNAP III transcription complexes, and known links between the molecular machineries involved[103–105] likely underlie this propensity.

Nascent transcripts from individual tRNA genes contain sequence complexity in the 5′ leaders, 3′ trailers, and for some, the introns.[102] It is reasonable to expect that this diversity would render some pre-tRNAs more prone than others to misfolding. Some eukaryotic tRNAomes include diversity within the mature tRNAs of the same anticodon families, referred to as isodecoders,[102] a source of additional potential for misfolding of the pre-tRNA. It is proposed that under the protective and chaperone activities of La, the transcripts of some tRNA genes might survive to maturity that otherwise might not. Hence, La might support a capacitor function of tRNA gene variation, consistent with a role for chaperones as capacitors for evolution,[88,89] as discussed above. It is therefore tempting to speculate that the functional organization of eukaryotic tRNA genes, which includes RNAP-III-mediated expression that is mechanistically linked to the UUU-3′-OH-binding activity of La, contributed to the expansion and variation of eukaryotic tRNAomes by also linking tRNA fate to the chaperone activity of La. Other factors would have contributed, for example, the tRNA gene amplification mechanism(s) itself, as well as certain nuclear tRNA modification enzymes that are known to act redundantly with La to assist structurally challenged pre-tRNAs avoid decay and promote maturation[94,106–109] reviewed in Ref 110.

The eukaryotic arrangement of tRNA genes as separate transcription units also empowers them as individual genetic units that may be acted on as selectable. Diversity and flexibility in the tRNAome might lead to or coevolve with enhanced potential for biased codon use by the transcriptome and adaptation to translational demands,[66,99,111] a perspective consistent with the possibility that the La module was keyed to and rooted in tRNA surveillance.

CONSERVED AND DIVERGED FEATURES OF THE LA MODULE RNA-BINDING UNIT

The La Module: A Novel RNA-Binding Platform with Unique Features and LARP Adaptability

'La module' is the term used to refer to the tandem arrangement of the two RNA-binding motifs in which a LaM (a unique fold, below) is followed by an RRM (termed RRM1 in La and LARPs)[18,112] (Figure 1). The LaM and RRM1 are composed of approximately 70 and 80 amino acids, respectively, and each is an independent structural 'domain' and will be referred to as such hereafter as appropriate. The La module is an exceptional RNA-binding unit discovered from structural studies of hLa.[3,113,114] While several hundred RRMs are found encoded in metazoan genomes,[38,39] their pairing as part of a La module is limited to six to nine unique LARPs per genome. Furthermore, phylogenetic analysis revealed the LaM and RRM1 coevolved in LARPs,[1] with only

© 2017 Wiley Periodicals, Inc.

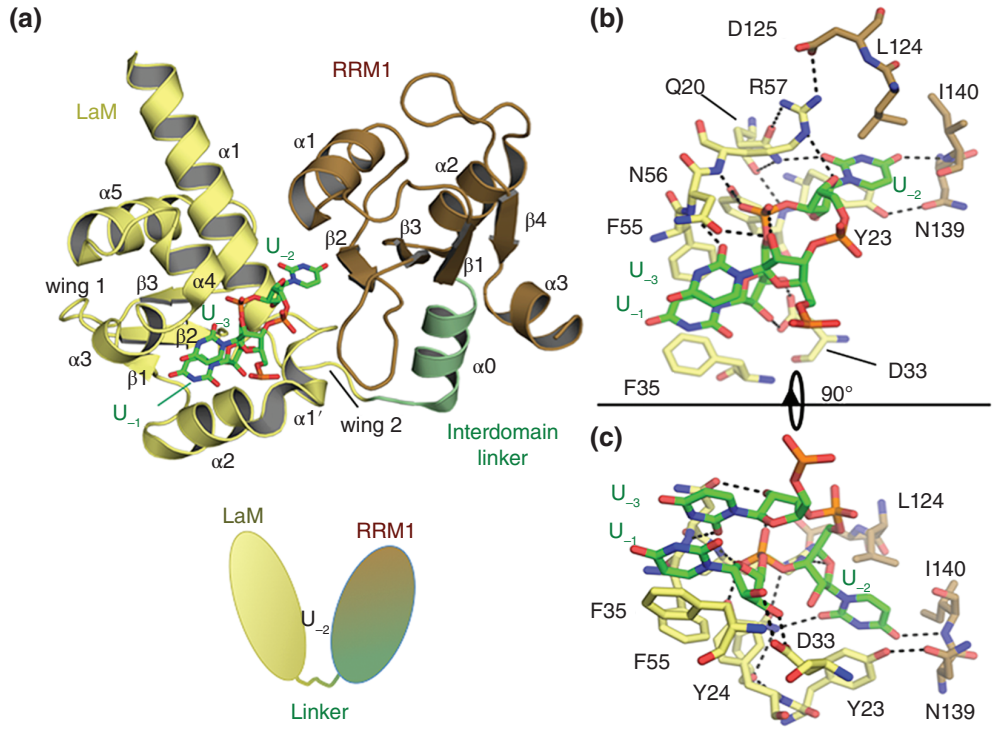


FIGURE 3 | The La module of human La protein in complex with UUU 3′OH RNA. (a) The La module of La (PDB 2VOS) is composed of the LaM (yellow), an interdomain linker (pale green), and the RRM1 (brown). The UUU 3′OH RNA is shown as sticks, color coded by atom type. For clarity, the bases beyond U-3 have been omitted. Cartoon under the structure shows schematic V-shaped model with U-2 in the cleft between the LaM and RRM1 (see text). (b) Close-up view of La–RNA interaction showing the three 3′ terminal bases in the same orientation as (a). Selected side chains are shown as sticks; dashed lines indicate hydrogen bonds. (c) Close-up view of the interaction, rotated by 90°.

5 proteins out of the 134 analyzed suggested to contain the LaM alone, underscoring the importance of the La module in its entirety to RNA recognition. Notably, LARP-specific features of the La module have been emerging from phylogenetic and structural studies as discussed below.

In La, the LaM and RRM work synergistically to interact with their RNA targets[3,113] and, albeit not demonstrated yet, this is likely true for the LARP superfamily. In hLa protein, the LaM and RRM1 are connected by a largely flexible linker and move about each other independently in the absence of RNA,[3] but adopt a fixed 'V-shaped' conformational clamp onto the UUU-3′-OH target, forming a sequence-specific binding cleft at the interdomain interface that appears to be perfectly tailored for this target (Figure 3(a)). While all but one of the specific interactions to the RNA are by side chains of the LaM, recognition of the most sequence-specific determinant of the target, the penultimate uridylate, U-2, is achieved by a cooperative set of protein–RNA and protein–protein interactions involving residues from both domains[3,114] (Figure 3(b) and (c)) reviewed in Ref 115. This induced fit of the binding pocket around U-2 explains the high synergic nature of this interaction (Figure 3). This mode of binding was also replicated in LARP7.[116]

An intriguing aspect of the LARPs is that despite the multiple conserved features detailed below, their different La modules recognize rather distinct RNA targets. Thus, a challenge is to reconcile differential RNA recognition among LARPs given a high degree of primary structure conservation, especially in their LaMs. In this review, we will summarize key findings of how structural studies have started to unveil divergent features that may account for or at least explain in part the different RNA-binding properties of LARPs.

The La Motif

The LaM was discovered in La as an elaboration of a previously known fold, the 'winged-helix motif,' which in LaM contains three additional α-helices[50,113] appended onto the canonical winged-helix scaffold. These extra elements (α1′, α2, and α4, Figure 3(a)) allow the formation of a hydrophobic cavity with RNA-binding capabilities that enables La

© 2017 Wiley Periodicals, Inc.

to adopt a novel mode of RNA recognition that uses entirely different binding surfaces as compared to canonical winged-helix domains. Within this hydrophobic pocket, specific and nonspecific contacts are made with the UUU-3′-OH ligand, in particular mediated by side chains of six key residues that are conserved in LARPs: Q20, Y23, Y24, D33, F35, and F55 (Figures 3(b) and (c) and 4) (numbering is for hLa, which will also be used to refer to corresponding positions in LARPs throughout this review).[3,114] Thus, the LaM-specific extra helices attached to the winged-helix scaffold are integral to RNA recognition and delineate unique function in the context of the La modules of La and LARPs.

Structural and mutagenic analyses have revealed that the hydrophobic RNA-binding pocket identified in the LaM of La appears functional in LARP7 for its interaction with 7SK RNA UUU-3′-OH,[116] and surprisingly also for hLARP6 for its internal recognition with collagen 5′UTR SL RNA.[23] Whereas this was anticipated for LARP7 based on high degree of amino acid homology with La as well as similarities in RNA termini recognition, for hLARP6 this finding is particularly significant given the difference in size and overall SL structure of its ligand, the collagen SL RNA. More intriguingly, the six key residues that in La mediate UUU-3′-OH-specific recognition appear to retain a central RNA-binding role in hLARP6, albeit the 3′-OH does not appear to be an element of recognition for LARP6 proteins (see below). Consistent with this is the high conservation of the six residues across all LARPs: considering the non-PAM2 (PABP-interacting motif 2) proteins (below), position 20 (hLa numbering) is occupied by Q in 96.6%; positions 23 and 24 by F/Y/W in 97.9 and 99.9%, respectively; position 33 by D in 99.7%; position 35 by F/Y in 96.2%; and position 55 by F in 97.6% of the sequences examined.[18] This is astounding conservation over the presumed 2 billion year phylogenesis of all of the LARPs. Figure 4 shows an alignment of the non-PAM2- and PAM2-containing LARPs.

Invariant Asp-33 in the Hydrophobic Pocket and Potential for 3′-OH End Binding

A particularly intriguing issue that is potentially relevant to the La modules of all LARPs is the role of invariant D33 which in La proteins resides deep within the LaM binding pocket and is critical for recognition of the 2′-OH and 3′-OH moieties of the terminal ribose, via formation of a bifurcated hydrogen bond with its side chain carboxylate[114] (Figure 3 (b) and (c)). There is phylogenetic divergence of two to three key hydrophobic pocket residues among LARPs that acquired a PAM2 sequence[18]; of these, D33 was maintained as universally conserved in LARPs4 and 4B. However, for plant LARPs 6B and 6C, which also acquired a PAM2 sequence and diverged in key hydrophobic pocket residues,[18] this included divergence of D33 but acquisition of D and E, respectively at their adjacent 34 positions (Figure 4). In the LARPs for which the structures and 3′-end recognition have been examined,,[3,23,50,114,116] i.e., La and LARP7, bearing D at position corresponding to hLa 33, simply altering the RNA terminus from 3′-OH to 3′-PO4, 3′-*O*-CH3, or 2′-*O*-CH3, significantly impairs binding to the wild-type (WT) proteins.. For La, depending on how the experiments are done, the source of the protein, the specific RNA target, and which 3′ or 2′ terminal ribose modification is examined, the negative effects range from small to more than 50-fold.[28,50,114] For La–UUU-3′-OH interaction, in agreement with the structural data, mutagenesis experiments have shown that D33 accounts for La's ability to distinguish RNAs with 3′-OH from 3′-*O*-phosphate and 2′-*O*-CH3 termini.[28,50,114] The snugness of fit of the 3′-terminal uridylate (U-1) in the LaM binding pocket in close proximity of D33 (Figure 3(a) and (b)) explains the increasingly deleterious effects caused by D33 substitutions with progressively bulkier amino acid side chains and modifications of the 2′-OH or 3′-OH of the RNA ligands, as larger chemical groups can be less well accommodated in the tight binding slot.[114] Still, the nonbulky D33A substitution decreases affinity for the 3′-OH ligands, two to threefold,[50,114] and the D-to-A substitution in *T. brucei* La protein led to a nearly sixfold decrease in discriminatory affinity for 3′-PO4 versus 3′-OH.[50] It should be noted that such discriminatory activity would help target La to newly transcribed transcripts, which end with 3′-OH, and disfavor binding to products of RNA cleavage and/or decay enzymes that produce 3′-PO4 and related termini.

Burying the RNA terminus in a deep binding pocket is likely critical to a principal biological activity of genuine La protein, protection of its ligands from 3′ exonucleases.[80–82] This 3′-end protection activity serves three roles in tRNA biogenesis: (1) orders the early phase of pre-tRNA maturation by directing 5′ processing to precede 3′ processing,[80] (2) directs 3′-end maturation by the endonuclease RNase Z rather than by the 3′ exonuclease, Rex1,[82,117] and (3) protects structurally challenged pre-tRNAs from 3′ polyadenylation by the TRAMP complex and subsequent degradation by the exosome-Rrp6, nuclear surveillance system.[81,92,108] Relevant *in vivo*, a La D33R point mutant is

© 2017 Wiley Periodicals, Inc.

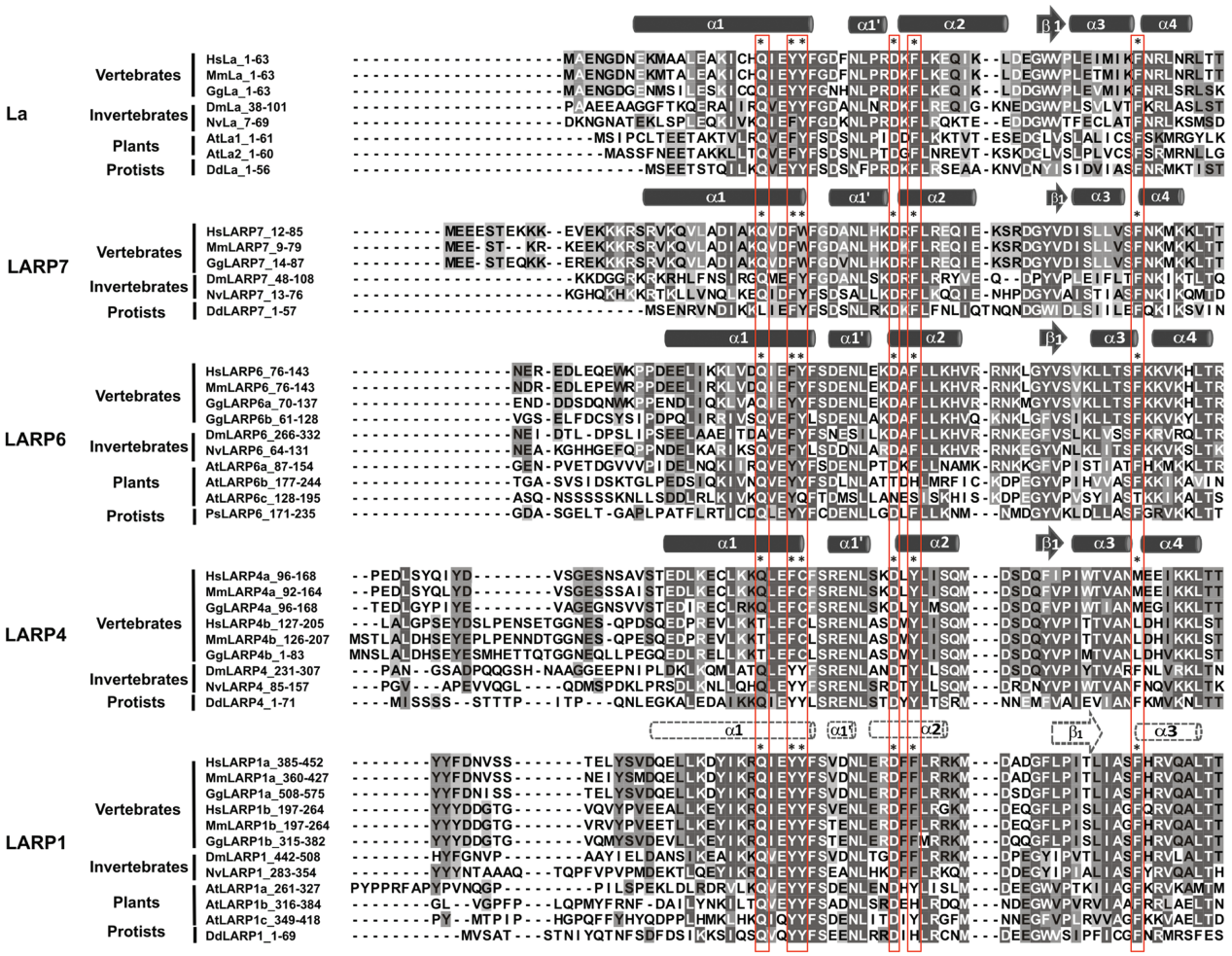


FIGURE 4 | Sequence alignment of the LaM and their N-terminal regions in La and La-related proteins (LARPs). For each family, the sequence of the human LaM was aligned with proteins of different species using Clustal Omega in Uniprot portal (http://www.uniprot.org/align/).[52] The alignments were edited and analyzed with Jalview software.[53] Residues were colored in gray scale according to the extent of conservation. Species selected include vertebrates-eutherians (Hs, Homo sapiens, Mm, Mus musculus), vertebrates (Gg, Gallus gallus), invertebrates (Dm, Drosophila melanogaster; Nv, Nematostella vectensis), plants (At, Arabidopsis thaliana), and protists (Dd, Dictyostelium discoideum; Ps, Phytophora sojae). The secondary structure elements for human LaMs appear at the top of the sequence for each family, where α-helices are depicted by cylinders and β-strands by arrows. For La, and LARPs 7, 6, and 4, these structured elements have been determined experimentally for the human proteins (I. Cruz-Gallardo and M.R. Conte, unpublished).[23,113,116] For HsLARP1a, a prediction of secondary structure was performed with Jpred-4 server (http://www.compbio.dundee.ac.uk/jpred/).[51] The conserved residues of the hydrophobic pocket of the LaM are labeled with asterisks and boxed in orange.

defective in stabilization of pre-tRNALys (ref 81). By this La module-mediated RNA-binding mode, La acts as a 3′ cap on its UUU-3′-OH RNA ligands, protects them from a variety of exonucleases, and can also chaperone them through different processing pathways.[29,118] Is a related activity relevant for LARPs?

An Alternative Role for Invariant Asp-33 in the Hydrophobic Pocket?

As noted, D33 has been universally conserved by LARPs with exception of plant 6b and 6c, which interestingly possess a D and E, respectively at the +1 adjacent positions (Figure 4). In considering the possible significance of this conservation, there have been some curious observations on LARPs that do not entirely fit with 3′-end protection observed in La and worthy of review. hLARP4 was shown not to significantly distinguish SL hairpin RNAs that differed only in their 3′-OH or -PO4 termini.[26] However, more recent experiments using single-stranded oligo(A) suggest that hLARP4 is somewhat more discriminatory against 3′-PO4 and 2′-O-CH3 termini than observed for the SL RNA (S. Gaidamakov and R.J. Maraia, unpublished). For hLARP6, while modification of the 3′-OH of collagen 5′ SL RNA to 2′–3′ cyclic phosphate had no effect on high-affinity

© 2017 Wiley Periodicals, Inc.

binding, interestingly, D33A substitution in the LaM binding pocket was associated with a significant 3.5-fold reduction in binding affinity.[23] This echoes what was observed with hLa D33A and the D33A equivalent of *T. brucei* La,[50,114] suggesting a comparable contribution of the hLARP6 D33 equivalent to RNA-binding affinity as La, despite the expected differences in its contacts to the ribose, the triphosphate, and/or base moieties of the respective ligand.[23] We note that recognition by hLARP6 of the SL in the 5′ UTR of collagen mRNA is not expected to involve 3′-OH terminal binding *per se*, and this will be considered in a model below.

Finally, recent data indicate that the LaM of hLa can exhibit context-dependent sensitivity to the nature of ligand RNA 3′ termini.[112] Binding of hLa to SL IV of the IRES element of the hepatitis C virus (HCV) RNA involves an alternative mode of molecular recognition compared to UUU-3′-OH, requiring cooperative interplay of the La module with the C-terminal RRM2α, which is largely insensitive to distinguishing 3′-OH versus 3′-*O*-CH3.[112] This was surprising because analogous 2′-*O*-CH3 alteration of a U(4)-terminated RNA decreased its affinity for hLa by nearly 40-fold.[114] Here too, high-affinity binding to this SL in its native setting would be connected to the rest of the HCV IRES and downstream coding sequences and therefore is not expected to include recognition of the 3′-OH terminus of the RNA.

The findings that both hLa and hLARP6 when binding their respective HCV and α-collagen SL RNAs respectively use their LaM binding pocket and invariant D33, but in a way not involving 3′-OH recognition, raises an important question: can an RNA that engages D33, continue on a course out of the hydrophobic pocket? Examination of the structure of hLa bound to its RNA ligand suggests such an exit path might exist, probably through a space between helices α2 and α3 or α2 and α4 of the LaM, although it is noteworthy that the overall conformation of the non-3′-OH-bound RNA is likely to differ somewhat from the one that is 3′-OH recognized (Figure 3(a), U-1 is the 3′ end of the RNA). For hLa–HCV SL RNA interaction studies, NMR chemical shift perturbation (CSP) analysis may provide some support for this as a larger number of residues in and around the hydrophobic binding pocket experience CSP with the SL RNA compared with the UUU-3′-OH (figure 5 of Ref 112). An exit path for the RNA could also be speculated for collagen SL binding to hLARP6, whereby the internal bulge of the RNA has been identified as the putative recognition element.[23,24] Beyond this initial speculative model, a deeper understanding of the binding of hLa or hLARP6 to structured SLs (HCV or collagen SL, respectively) requires further structural data of the protein–RNA complexes.

The collective results suggest that while D33 (and equivalent thereof) appears to contribute to RNA binding across the LARPs, its ability to distinguish and/or discriminate 3′ termini is dependent on overall architecture and mechanism of the particular molecular interaction. An outstanding issue is whether a La module binding pocket–D33 interaction is used to protect ligand RNAs from 3′ exonucleases in the LARPs, two of which have been shown or suspected to have 3′-end binding.[119,120] In any case, the observations suggest the La module as an extremely versatile platform capable of more than one mode of RNA recognition.

LaM Domain Boundaries

Current understanding of the structural basis of differential RNA target specificity by the La modules of LARPs 1, 4, and 6 is limited. Nonetheless, interesting features of the LaM boundaries that emerged from structural work on hLARP6 may provide valuable insights. At the N-terminal region of LARP6 LaM, the structure revealed an α1 helix that is shorter than in La, with the stretch preceding this helix (residues 81–89) forming an integral part of the LaM domain, establishing nonpolar interactions with helix α5.[23] Mutations of W85 and K86 here did not alter RNA-binding activity of hLARP6,[23] although a deletion mutant of this region failed to bind RNA, suggesting unfolding of the truncated domain.[24] Interestingly, this N-terminal stretch is highly conserved in eutherian and other vertebrate LARP6 proteins[23] but not in invertebrates and plants. Notably, LARP6 orthologs from different species have different RNA-binding properties[18] but whether the N-terminal region contributes to this remains to be established. Interestingly, inspection of sequence preceding the LaM in other LARP families highlights similar patterns of species-dependent conservation for this region (Figure 4).

Another divergent feature of the LaM that was not anticipated from sequence alignment, but became evident from recent structure determination, applies to wing 2 at its C-terminal boundary (Figure 5(a)–(c)). In La, wing 2 begins with Arginines 90 and 91 and is characterized by a nearby PLP motif (96–98) (Figure 5(a)), which represents the C-terminal end of the LaM, with P96 and L97 establishing hydrophobic interactions with α1′[113] (Figure 5(b) and (c)). Residues beyond P98 were shown by NMR backbone dynamics analysis to be flexible, thereby demarcating the end of the LaM and

© 2017 Wiley Periodicals, Inc.

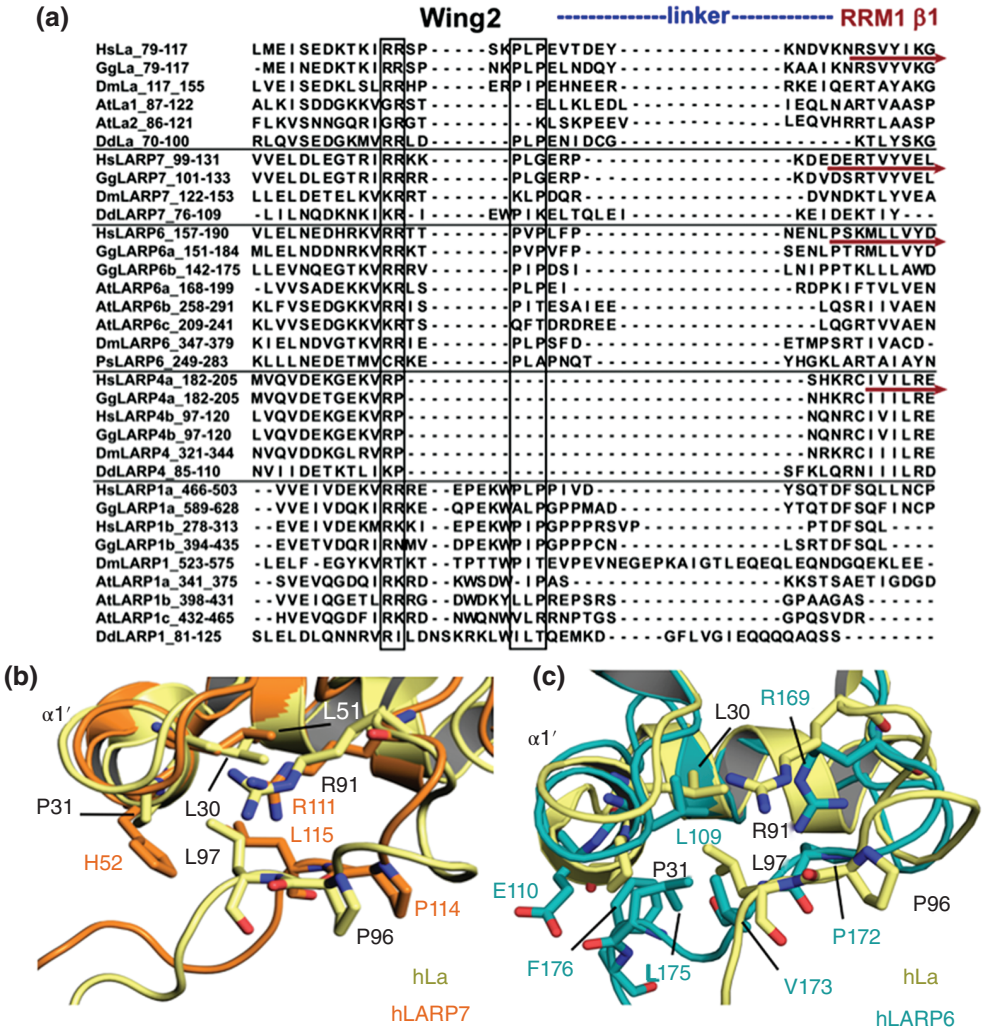


FIGURE 5 | Comparison of the wing 2 motifs and interdomain linkers in La-related proteins (LARPs). (a) Sequence alignment focusing on LaM wing 2, interdomain linker, and the beginning of RRM1 of LARPs. Human (Hs) La sequence was aligned with 32 LARPs as for Figure 4. Black vertical rectangles indicate conserved signatures characterizing wing 2; the dark red horizontal arrows denote the beginning of RRM1 strand β1 determined experimentally for hLa, LARP7, LARP6, and LARP4 (I. Cruz-Gallardo and M.R. Conte, unpublished).[23,113,116] The potential extent of the variable linker regions is indicated with a dotted line above the sequences. (b, c) Close-up views of the wing 2 region of hLa superimposed with hLARP7 (b) and hLARP6 (c). Selected residues involved in the interaction of wing 2 with the rest of the domain are highlighted as sticks. hLa (PDB 2VOS) is depicted in yellow, hLARP7 in light orange (PDB 4WKR), and hLARP6 in cyan (PDB 2MTF).

the beginning of the interdomain linker.[3] In the hLARP6 LaM, despite the two Arginines and a PVP signature (residues 172–174) (Figure 5(a)), P172 and V173 were surprisingly not involved in interactions with α1′ instead, a different configuration of wing 2 positions these residues somewhat away from the rest of the molecule, leaving the downstream residues, L175 and F176, to stabilize it via contacts with α1′ (Figure 5(c)). The divergence between La and LARP6 here is twofold: (1) different structure at the C-terminal of LaM and (2) shorter interdomain linker for hLARP6, and both parameters are very likely to play a dominant role in determining the relative domain orientation of the tethered LaM and RRM1 within the La module.[23]

Sequence alignment of LARPs (Figure 5(a)) indicates that as compared to hLa, hLARP7 contains a shorter wing 2, and the signature PLP is replaced by PLG (residues 114–116). Nonetheless, L115 makes stabilizing contacts with α1′ and the overall configuration of wing 2 is not very dissimilar to that of La[116] (Figure 5(b)). The wing 2 sequence is somewhat longer in LARP1 members and is interestingly absent from LARP4 members (Figure 5(a)).

© 2017 Wiley Periodicals, Inc.

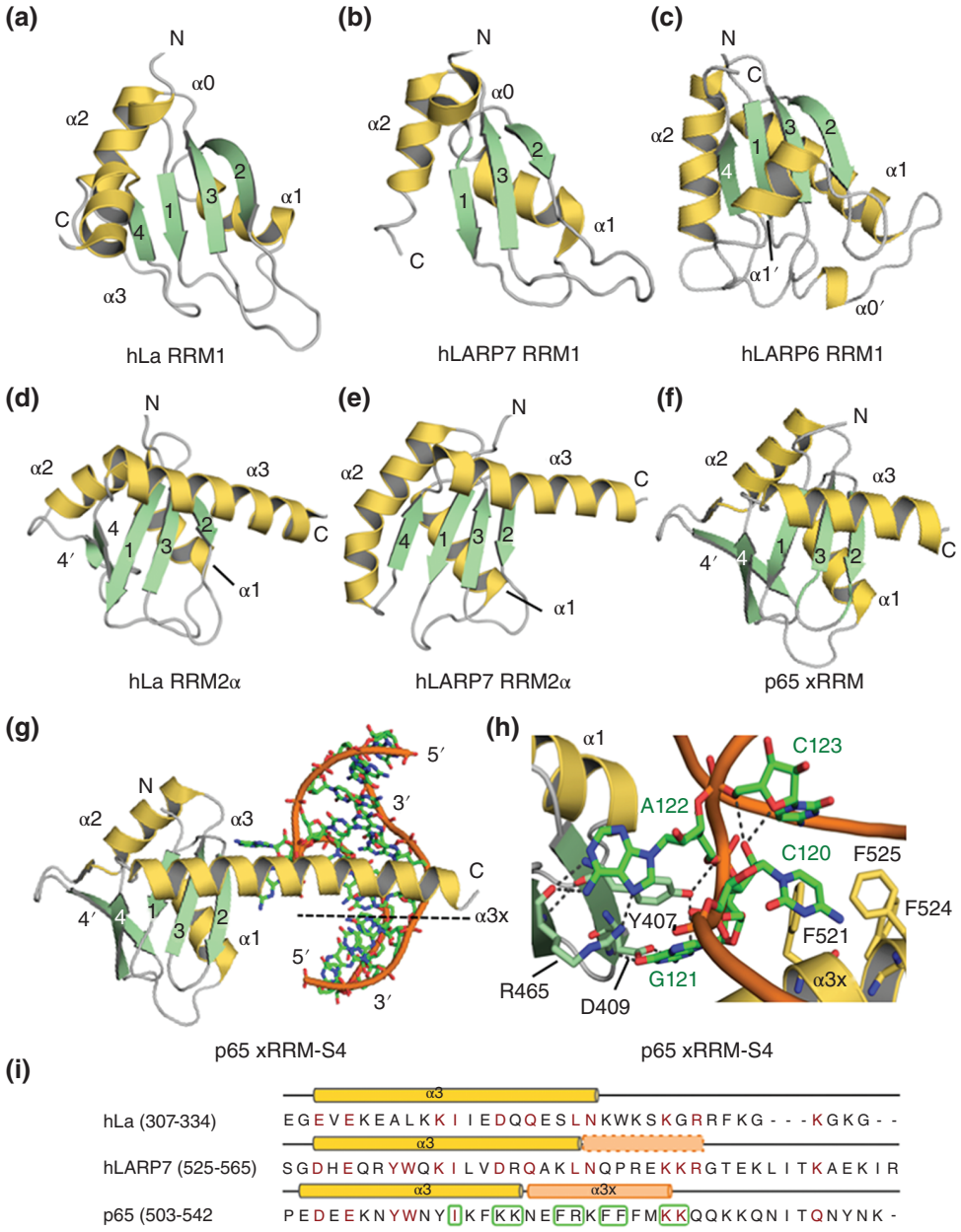


FIGURE 6 | Comparison of RRM1 and RRM2 domains in La-related proteins (LARPs). The RRM1 domains of LARPs are structurally diverse: (a) hLa RRM1 (PDB 1S79); (b) hLARP7 RRM1 (PDB 4WKR); (c) hLARP6 RRM1 (PDB 2MTG). hLARP7 RRM1 lacks strand β4; hLARP6 RRM1 contains additional helices α0′ and α1′. The RRMs2 of (d) hLa (PDB 1OWX), (e) hLARP7 (PDB 5KNW), and (f) p65 (PDB 4EYT) all contain a long C-terminal helix (α3) that obscures the β-sheet platform. (g) In p65, the unstructured C-terminal of the helix (α3x) refolds upon RNA binding. (h) Close-up view of the interaction between p65 and the S4 RNA. Selected residues are highlighted in stick representation. (i) Sequence alignment of the α3 region for hLa, hLARP7, and p65 performed with Clustal Omega in Uniprot portal (http://www.uniprot.org/align/)[52] and edited and analyzed with Jalview.[53] Residues colored in dark red indicate conservation. Boxed residues denote amino acids that interact with RNA in p65.

Overall, there are significant differences in the wing 2 sequences across LARPs that could potentially have considerable repercussions in the mode of RNA binding, influencing the relative positioning of the LaM and RRM domains of the respective La modules.

The RRM1

Although the structure of RRM1 in hLa closely resembles the canonical RRM fold (Figure 6(a)), it was surprising to find that it uses a noncanonical RNA-binding surface to bind short oligoU sequences.[3,114,121,122] Phylogenetic classification anticipated LARP family-specific

characteristics in the predicted RRMs,[1] and the newly solved structures of hLARP7 and LARP6 uncovered significant RRM1 variability: only three β strands instead of the canonical four were observed in LARP7 RRM1[116] (Figure 6(b)), whereas LARP6 RRM1 contains additional helices[23] (Figure 6(c)). In addition, work on hLARP7 and hLARP6 suggested that the canonical RNA-binding surface of their RRMs1 is also not a positive determinant of sequence-specific RNA interaction, similar to La protein.[23,116]

The distinctive LARP family-associated traits of RRM1 suggest that this domain plays a key role in their RNA substrate selection specificity. This was demonstrated by loss of binding specificity or affinity when RRMs1 of hLARP7 and hLARP6, respectively, were replaced by RRM1 of hLa.[23,123] Although the features in hLARP6 RRM1 that mediate specific interaction with the SL in collagen mRNA are yet to be defined, preliminary data indicate more extensive contacts as compared to hLa and hLARP7, and this may also be the case for hLARP4 (I. Cruz-Gallardo and M.R. Conte, unpublished). Detailed structural information on the La modules of additional LARPs in complex with their targets will be necessary to further address this issue.

Interdomain Linker and Domain–Domain Orientation in the La Module

Robust synergism between LaM and RRM1 for RNA binding that was seen in La is recapitulated in both LARP7 and LARP6 La modules.[23,116] In hLa, the linker is flexible in the apo protein but adopts helical conformation in the RNA-bound form, assisting the correct positioning of both domains in an orientation optimized for RNA recognition.[3]

Demonstration of the important topological role of the linker in RNA recognition was provided by the observation that replacing the short linker of hLARP6 with the longer one from hLa, which would perturb interdomain distance and degree of conformational sampling, decreased binding affinity for collagen SL RNA by approximately 10-fold. Importantly, the cumulative analyses from the LARP6 study were not readily compatible with a V-shaped architecture for LaM and RRM1 as occurs in La and LARP7 bound to UUU-3′-OH.[23] Instead, they envisioned a more elongated tandem domain orientation in the RNA-bound form, in which the interdomain linker played an integral role in domain architecture and RNA recognition[23] (Figure 7).

Perhaps, of more general significance, the biophysical studies of LARP6 provided initial evidence to suggest that the two domains of the La module may be capable of synergistic RNA recognition via different topological arrangements in different LARPs. An elongated rather than a deep V-shaped (Figure 7) configuration may also be expected for the hLARP4 La module from preliminary structural data (I. Cruz-Gallardo and M.R. Conte, unpublished), and this would be consistent with a binding requirement for a relatively long tract, 15 nucleotide of single-stranded RNA.[26] Whether the La module tandem domain arrangement in hLARP6 and hLARP4 differs from hLa and hLARP7 awaits confirmation from structural data of their complexes with target RNAs.

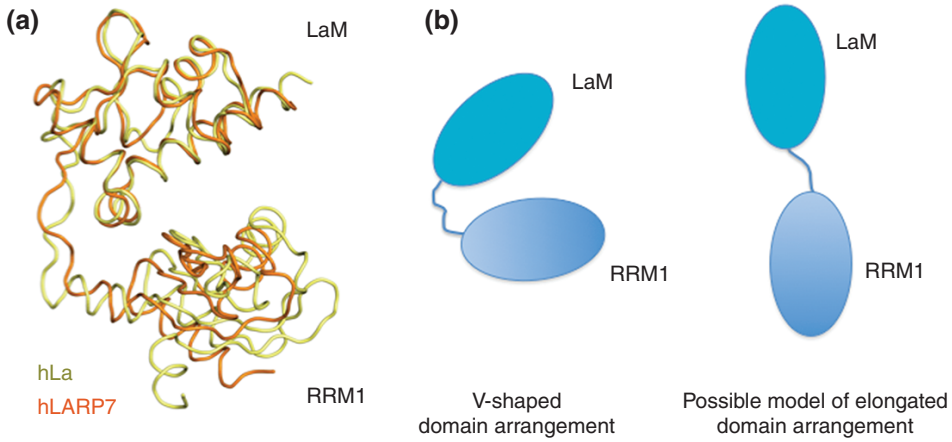


FIGURE 7 | Proposed arrangements of the La module for RNA recognition by La-related proteins (LARPs). (a) Domain arrangement of hLa and hLARP7 in complex with RNA (RNA removed for clarity). The LaM and RRM1 adopt a V-shaped conformation to create the binding pocket to accommodate the RNA ligand. (b) Cartoon representation of possible domain–domain orientation of the La modules of hLARP6 and hLARP4. Current knowledge on wing 2 conformation and interdomain linker for these proteins suggests that LaM and RRM1 will adopt a more elongated arrangement to interact with RNA (see text).

© 2017 Wiley Periodicals, Inc.

In summary, the length and amino acid composition of the LaM-RRM1 interdomain linker should be expected to be key in determining the functional structure of La modules across the LARP superfamily, with consequent effects on RNA binding. Interestingly, sequence and length of the linker indeed vary among LARP families, similar to that noted for wing 2 (Figure 5(a)).

Use of Canonical and Noncanonical RNA-Binding Surfaces

Formation of the unique binding pocket for UUU-3′OH recognition in La protein by the LaM and RRM1 does not involve the RNA-binding surfaces expected for canonical winged helix and RRM domains, respectively, and therefore leaves open the possibility for additional, perhaps simultaneous, modes of RNA binding.[114,121,122] In support of this possibility, it should be underscored that La RRM1 bears a number of hallmark features of the canonical β-sheet RNA-binding surface, including RNP-1 and RNP-2 motifs on β-strands 3 and 1, respectively, whose key aromatic residues have been conserved, e.g., see figure 3 in Ref 81. Functional involvement of the canonical RRM surface in pre-tRNA binding was demonstrated for hLa.[81,87] Moreover, the UUU-3′-OH and the additional RRM1 canonical surface binding mode could be functionally linked to the two different activities of La, 3′ end protection, and pre-tRNA chaperone activity, respectively.[81] Mutation of hydrophobic pocket residues in the LaM impaired UUU-3′-OH binding, while substitution of RRM1 loop-3 residues weakened binding to other parts of the pre-tRNA.[87] Mutagenesis coupled with *in vivo* functional assays indicate that the conserved aromatic residues on RNPs 1 and 2 and key basic residues in loop 3 that connect β2 and β3 of RRM1 are required for pre-tRNA chaperone activity.[87] Analogously, the β-sheet of LARP7 RRM1 has been suggested to contact 7SK RNA beyond the terminal U triplet.[116]

It is noteworthy that a relatively short stretch of RNA, (U)UUU-3′OH is specifically recognized by the La module[3,114] of La and LARP7 to comprise most of its high-affinity binding despite the potential extra RNA-binding surface. In contrast, the La modules of hLARPs 4 and 6 do not recognize equally short RNAs: e.g., hLARP4 requires at least 15 nucleotides of its highest known affinity ligand, oligo(A).[26]

In summary, evidence exists for canonical and noncanonical-type RNA binding by the RRM1 in La protein, the latter of which provides the clearest and best established structural models by which it recognizes its most prevalent and biologically relevant ligands, the UUU-3′OH termini of nascent RNAP III transcripts. However, as reviewed above, biochemical and functional biological data also indicate use of the canonical RRM surface of hLa in tRNA biogenesis.[81,87]

While our collective understanding of RNA recognition by La modules has advanced enormously due to detailed structural studies coupled to mutagenesis and affinity measurements, many interesting and important questions remain unanswered. Genuine La protein bears multiple RNA-binding surfaces and can exhibit more than one mode of RNA recognition. Therefore, models based on UUU-3′OH binding by La should not necessarily be expected to explain RNA binding by a LARP. Evidence of noncanonical arrangement and combinatorial potential of the La module should also caution predictions based on sequence and structure alone. On the other hand, the possibility that the La module can adopt alternate architectures (Figure 7) enlightens another perspective. The extent to which a La module may be able to switch between topological architectures upon interactions with different RNAs is an intriguing possibility.

Proteins with a LaM and No Apparent RRM

It is also noteworthy that a few proteins have been identified that contain a LaM without an adjacent RRM, including *S. pombe* Slr1p, *Aspergillus niger* AnLARP, *Phytophora sojae* PsLARP, as well as the characterized *S. cerevisiae* Sro9p and Slf1p.[1] Sro9p and Slf1p appear to function in mRNA translation[124] and were shown to selectively bind homopolymer RNA.[124] Slf1 was known to be involved in copper metabolism in yeast before it was known to harbor a LaM.[125] RNA co-IP experiments revealed association of Slf1 with mRNAs related to copper resistance and oxidative stress, apparently dependent on the LaM, as the triple mutant Y24A–F35A–F55A (hLa protein numbering) resulted in decreased association.[126] An oxidative link to mitochondria through Slf1 was uncovered.[127,128] Slf1 and Sro1 appear functionally related to LARPs 1 and 4 families and similarly to these, associate with translating ribosomes and PABP and interact with overlapping sets of abundant mRNAs including those that encode ribosomal proteins.[128] Indeed, Slf1p was found as one of a few nonribosomal proteins in the yeast 77S complex that contains closed-loop-specific translation components,[129] reminiscent of other cytoplasmic LARPs (below). Although much remains to be uncovered about the mode of RNA binding of these unusual

© 2017 Wiley Periodicals, Inc.

LARPs, it may be possible to hypothesize the existence of an additional noncanonical RNA recognition mechanism by LARPs that involves the LaM but does not require a RRM1 at all. Alternatively, a cryptic or nonrecognizable RNA-binding motif may reside C-terminal to these LaMs. Further analysis of their RNA binding requirements will be required to test this.

LARP REGIONS BEYOND THE LA MODULE INVOLVED IN RNA RECOGNITION

Although the La module is a unique RNA-binding unit and a main locus of RNA recognition, it is allied to other motifs and domains that are distinctive to each LARP family, and these can contribute to discrete RNA binding by these proteins. Here, we will use three recent examples of RNA binding by downstream motifs of different LARPs, the xRRM, RRM2α, and DM15.

The RRM2α and xRRM

La and LARP7 families contain an additional RRM in their C-terminal region, of a subtype with an α3 helix that lies across the β-sheet,[33] referred to here as RRM2α (Figures 1 and 6(d)–(f)). In p65, the *Tetrahymena* homolog of LARP7, the RRM2α was dubbed xRRM based on specific novel features not shared by La RRM2α. The p65 xRRM interacts specifically with the terminal stem-loop 4 (S4) of TER RNA independently of the La module using an extension of helix α3 (termed α3x) that transitions from an unstructured state in the apo protein to helical conformation upon binding to S4 RNA[36] (Figure 6(g)–(i)). Structural studies revealed that α3x binds S4 across the major groove adjacent to a GA bulge with aromatic residues F521, F524, and F525 engaging in hydrophobic interactions (Figure 6(g) and (h)). Residues Y407 and D409 on β2 and R465 on β3 provide further hydrophobic contacts and hydrogen bonds between the edge of the β-sheet and the RNA[36] (Figure 6(g) and (h)). Additional interactions, largely with the RNA backbone, involve basic residues on the α3x helix (K517, K518, R522, K528, and K529).

The unusual features and novel mechanism of RNA recognition that led to the definition of xRRM are based on: (1) absence of RNP2 and RNP1 consensus sequences in β1 and β3, respectively; (2) conservation of arginine (R465 in p65) within the nonconsensus RNP1 which recognized both nucleotides in the GA bulge; (3) presence of a new aromatic-X-D/Q/E/N motif on β2, involved in nucleotide recognition, that was termed RNP3; (4) an α3 helix that lies across the β-sheet surface with a C-terminal tail required for RNA binding that is disordered in the free protein but forms the α3x extension upon RNA binding; and (5) high-affinity binding to RNA that depends on contacts with α3x and is independent of the La module[36] (Figure 6(g)–(i)).

Recent studies on hLARP7 RRM2α reveal similarities and differences with p65 xRRM. First, contrary to p65, LARP7 RRM2α does not bind a two-nucleotide bulge, but interacts with both unpaired and base-paired nucleotides in the stem and apical loop of hairpin 4 (HP4) of 7SK RNA[116,130] (also see section on LARP7-RNA interactions below). LARP7 RRM2α shares with p65 xRRM the hallmark RNP3 sequence in β2 (YVD) and the conserved Arg on β3, which are involved in the 7SK HP4 RNA binding according to NMR CSP analysis.[130] Despite the presence of helix α3 in LARP7 RRM2α that appears to participate in RNA binding, the extent of involvement of a C-terminal tail that would form an α3x helical extension when bound to RNA is still somewhat unclear. The very weak CSP detected upon RNA binding beyond residue 546[130] are inconsistent with a newly folded α-helix which would typically be accompanied by extremely large CSP.[3,131] Furthermore, C-terminal truncation of five residues (551–546) decreases LARP7 binding affinity for 7SK HP4 by approximately 10-fold; however, the equivalent truncation in p65 xRRM (to residue 519) completely abolishes binding to TER S4,[36] indicating dissimilarities in RNA recognition mode between LARP7 and p65 RRM2. Supporting this, α3 sequence alignment shows that hLARP7 lacks the three Phe aromatic residues and some of the basic residues that bind to TER S4 in p65 (Figure 6(i)).

The structure of the RRM2α of hLa was the first that showed presence of a long C-terminal helix α3[2] adopting a configuration that was largely recapitulated in the RRMs2 of LARP7 and p65. RRM2α of La behaves quite differently from LARP7 and p65, in that it has very little, if any, RNA-binding capability in isolation, and appears to contribute to RNA recognition only when working in synergism with the La module to recognize internal SL structures, such as in the HCV IRES and pre-miRNAs.[112,132] RRM2α also appears important for binding to HBV SL and other mRNAs.[133,134]

Closer inspection of sequence alignment shows that in La the RNP-3 sequence is somewhat conserved on β2 (WID),[130] whereas the key Arg on β3 and the basic and aromatic residues on α3 and α3x are not (Figure 6(i)), perhaps providing an explanation for the difference in RNA-binding capability of the RRM2α in La as compared to the LARP7 family.

While the isolated RRM2α of hLa (225–334) exhibits very little if any RNA binding on its own,

© 2017 Wiley Periodicals, Inc.

addition of sequences about 20 residues downstream of α3 which includes a tract of highly basic residues (a.k.a., SBM, in Figure 1) increases RNA binding[2,11,135] and probably contributes to yet other modes of RNA recognition by La, including to the 5′ triphosphate of nascent RNAP III transcripts.[11,136] A new study employing time-resolved electrospray ionization hydrogen–deuterium exchange identified potential changes that accompany RNA binding in the entire hLa C-terminal domain, encompassing RRM2α and the SBM.[135] Interestingly, a region just following helix α3 (residues 321–326) was proposed to become less structured in the presence of ssRNA and more structured when mixed with SL IV of HCV RNA.[135] However, exact molecular details await further studies, including clarification as to whether any cooperation exists between the basic region and the RRM2α for RNA recognition.

The DM15 Domain

LARPs 1A and 1B possess a highly conserved C-terminal region, namely DM15, composed of one to four conserved DM15 repeats/boxes.[1] The DM15 domain of hLARP1A was found to adopt an atypical HEAT-like fold containing three helix–turn–helix repeats that map to the three conserved DM15 boxes (A, B, and C).[137] Importantly, this domain is capable of mediating RNA binding, independent of the La module.[137] Very recently, the DM15 domain was unveiled as a new cap-binding protein, recognizing specifically the 5′ cap motif m^{7}GpppC—with a C as the first base—which is characteristic of 5′TOP mRNAs[15] (Figure 8). In agreement with previous observations, in a classical case of convergent evolution the specific mode of recognition of the m^{7}G cap by DM15 follows a common theme shared with other cap-binding proteins, whilst the overall domain fold diverges.[115] Specifically in DM15, the charged methylated m^{7}G base is stacked between two tyrosines, Y883 and Y992, in a classical 'cation-π-sandwich,' whilst its nitrogen atoms N1 and N2 establish hydrogen bonds with an acidic residue (E886) in the binding pocket, all of which are common features of cap recognition. Additional H bonds are formed between N1/N2 and Ser882 in the cap-binding crevice (Figure 8). Interestingly, the first C base engages in a specific network of H bonds with arginines 847 and 879. The functional repercussions of this specific recognition are considerable, as most cellular mRNAs have a Guanine in the +1 position, contrary to 5′TOP mRNAs. The net preference of DM15 for a C in the +1 position was unambiguously demonstrated in electrophoretic mobility shift assay (EMSA) experiments comparing different capped oligo-

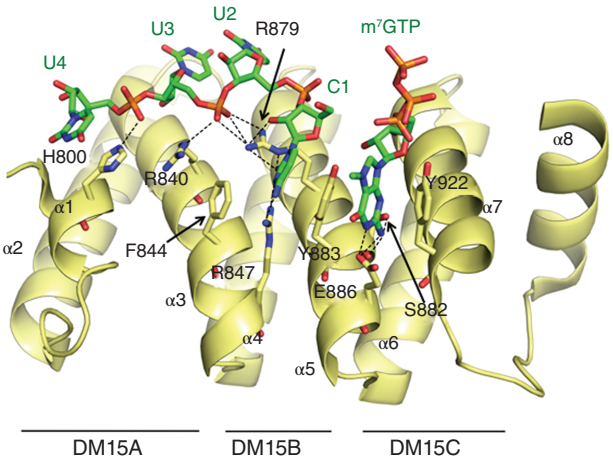


FIGURE 8 | Structure of the DM15 domain of hLARP1 in complex with m^{7}GTP and 5′TOP RNA. The DM15 domain is composed of three helix–turn–helix repeats denoted by DM15 boxes A, B, and C. The figure shows a superimposition of two crystallographic structures: (1) DM15 bound to the RNA sequence 5′-CUUUUCCG-3′ (PDB 5V7C) and (2) the complex DM15-m^{7}GTP (PDB 5V4R). For clarity, the bases beyond U4 in the 5′TOP mRNA have been omitted. Selected side chains involved in protein–RNA contacts are shown as sticks and dashed lines indicate hydrogen bonds.

RNAs.[15] Strikingly, the LARP1 DM15 domain both outcompetes cap-binding protein eIF4E for binding to m^{7}GpppC-TOP RNA and can displace the cap-C TOP RNA from eIF4E.[15]

Beyond the cap and the first base, only the first four nucleotides of the bound RPS6 5′TOP mRNA sequence contact the positively charged concave surface of the DM15 domain, with interactions between the phosphate backbone of the RNA and side chains of R879, R840, and H800[15] (Figure 8). No further contacts were seen with ribose or base moieties beyond C + 1, suggesting that the 5′TOP mRNA sequence would not be a determinant for substrate discrimination by DM15. Thus, the basis of the preference reported for the DM15 domain for some 5′TOP sequences over others appears to remain unexplained by the recent structure. Nonetheless, direct interaction of DM15 with the 5′TOP motif[15] is particularly interesting given that a putative mTORC1-recognition sequence is located in a flexible loop C-terminal to DM15 repeats, thereby ascribing a role for LARP1 in directly linking TOR signaling to ribosome biogenesis (see below).

THE NUCLEAR LA PROTEIN

La protein has two recognized functions, nuclear, to assist nascent RNAP III transcripts in their processing and maturation, and cytoplasmic, in mRNA

© 2017 Wiley Periodicals, Inc.

translation. Furthermore, studies from multiple laboratories have shown that a fraction of La concentrates in nucleoli in various cell types although its function remains unknown. In addition to binding cellular transcripts, the intimate association of La with specific transcripts synthesized by viral RNAPs, first elucidated by work on RNA viruses, continues to deepen our understanding of its impact on an expanding field of virology. Thus, it is appreciated that La is a multifunctional protein. Below we will review recent findings on nuclear functions of La and in a later section its cytoplasmic functions.

Nuclear La Function in Nascent RNA Sorting

As noted above, a recent advance in La function was in the channeling of nascent RNAP III transcripts.[60] Related to this, it had been reported that the activity of Ago2, a Piwi-homologous, active component of the siRNA-mediated silencing complex (RISC) that uses siRNA to guide cleavage of cytoplasmic mRNA, could be enhanced by La.[138] Evidence was later presented that La was additionally involved in another part of the miRNA-mediated silencing pathway. It was shown to associate with, stabilize and regulate nuclear pre-miRNAs by SL binding in a UUU-independent manner requiring its La module and RRM2α,[132] similar to the domain requirement used for binding to the HCV mRNA.[112] The recent results indicate that La prevents nascent pre-tRNAs with propensity for misfolding from routing to the pre-miRNA pathway.[60] In the absence of La some pre-tRNAs undergo alternate folding to structures recognized by the nuclear export factor exportin-5 and the processing factor Dicer and get loaded into Ago-containing RISC; however, when La is present its early access to nascent pre-tRNAs prevents their mischanneling.[60]

La Binding to Spliced Pre-tRNA Intermediates in Animal Cell Nuclei

As noted in preceding sections, the La proteins in all species examined representing ancient eukaryotes that predate the evolutionary emergence of yeast, as well as other species examined including plants, contain an RRM2α, whereas the yeast lineage appears to have lost its RRM2α.[1] A feature here that is of potential relevance to the nuclear function of La is that the miRNA pathway was lost from some yeast.[139] However, a more striking difference is that pre-tRNA splicing occurs in the nuclei of animal cells but in the cytoplasm of yeast.[140] This is relevant because La is found associated with pre-tRNA splicing intermediates in animal cells, whereas it is physically separated from the splicing pathway in yeast. Indeed, mouse brain cells genetically deleted of La protein become deficient in a 3′ trailer-containing spliced intermediate pre-tRNA that is normally bound to La in WT brain cells.[118]

LARP7 FAMILY MEMBERS

These are the LARPs most similar to La, not only by homology and phylogeny,[1] but their La modules exhibit the same UUU-3′OH sequence-specific RNA recognition mode and they also share similar RRM2α architecture (Figures 1 and 6). However, their distinctiveness is that while La binds UUU-3′OH terminal motifs on all RNAP III transcripts, LARP7 members recognize this motif on a specific subset of RNAP III transcripts, the identity and function of which can differ in different species. Another functional distinction is that while La binding is almost always transient for the spectrum of nascent RNAP III transcripts, LARP7 members remain associated with their RNA as an integral part of a stable RNP, e.g., metazoan 7SK snRNP or ciliate TER RNP.

LARPs7 provide a clear important example of how sequence homologous members of a single LARP family can adopt different ligands and functions in different species. Vertebrate and other metazoan LARP7 members stably bind to the ~331 nt 7SK RNA as a component of an abundant snRNP. Physiologic roles of the 7SK snRNP include balancing the growth and other regulatory functions of P-TEFb (positive transcription elongation factor-b) activity, which releases RNAP II from its promoter proximal paused state to a productive elongation state.[141–143] LARP7 association with 7SK RNA is highly stable, making it an integral component of the snRNP.[144–146] The LARP7 members of the two ciliate species that have been studied, *Tetrahymena* p65 and *Euplotes* p43 (Figure 1), are stably associated with TER RNA, which in ciliates is a ~160 nt product of RNAP III (whereas RNAP II synthesizes TER RNA in metazoa). As polyploid ciliate macronuclei contain a large number (200–300) of linear chromosomes, they need very high levels of TER RNA, a function well suited for RNAP III.[65,66] p65 and p43 are stable components of their respective multisubunit RNP TER enzyme complex and integral to function.[147,148] Both of these types of RNPs are responsible for critical functions that are central to proper growth and development in their respective biological systems, and much work has gone into their structure–function analyses.

© 2017 Wiley Periodicals, Inc.

Data suggest that hLARP7 and p65 exhibit RNA chaperone-like activity. This involves RNA folding and/or other facets of dynamic RNP organization upon binding to 7SK snRNA or TER snRNAs.[36,144,149]

Transfer of Pre-7SK RNA from La to LARP7 Preceding Assembly of 7SK snRNP

It is known that as a nascent transcript, pre-7SK RNA is associated with La via its UUU-3′OH before transfer to LARP7 for assembly of the 7SK snRNP[144] reviewed in Ref 17, but an open question remains on how transfer from La to LARP7 occurs.[17] A first observation is that LARP7 appears tailored for 7SK because in addition to La module recognition of UUU-3′OH, its RRM2α makes tight interactions with the apical loop of HP4. Both docking sites, UUU-3′OH and the HP4 apical loop are required for LARP7 interaction: the U tail alone is unable to support *in vivo* binding[144] and LARP7 mutants lacking the C-terminal region that contains RRM2α are unable to bind 7SK *in vivo*.[123] In contrast, the La RRM2α is unable to interact with HP4 and appears to have different ability for RNA binding[112] (and see above). It is not clear however that increased avidity of LARP7 for 7SK RNA is sufficient to displace La.

A nucleotide triphosphate is a universal motif incorporated at the 5′ end (5′-pppG/A) of all newly synthesized RNAP III transcripts, which has been shown to be a target of the short basic region (SBM) in the C-terminal region of hLa.[11,136] The 7SK RNA sequence contains a 5′ proximal motif that promotes the modification of its 5′-pppG to a monomethyl γ-phosphate cap (5′-me-ppG) by the methyltransferase, MeCPE,[150] which decreases the affinity for La protein.[136] Moreover, MeCPE remains stably associated with 7SK RNA in the complex, prompting the suggestion that it may participate in the transfer process. Intriguingly, immunoprecipitation of La from HeLa cells detected 7SK RNA and MePCE but not other components of the regulatory 7SK snRNP (LARP7, HEXIM, CycT1, and hnRNP A1),[144] suggesting that MeCPE may associate with La-bound 7SK and methylate the 5′-pppG-RNA, which in turn reduces the affinity for La.[17,136]

LARP7–RNA Interactions Are Necessary for LARP7 Functions

In the first characterization of the biochemical and biological activity of hLARP7, it was named PIP7S.[123] LARP7 specifically recognizes the 3′ hairpin of 7SK (termed HP4), in that the UUU-3′OH element is recognized by the La module and the apical loop is specifically contacted by the RRM2α in the C-terminal part of the protein[116] (Figure 1). Both elements have been shown *in vivo* to be essential docking sites of LARP7.[144] The atomic details of these interactions have been discussed previously. Intriguingly, recent EMSA experiments revealed that LARP7 can also bind to HP1, the conserved N-terminal hairpin of 7SK that is also the locus for HEXIM interactions.[151] It was proposed that the long, mostly unfolded linker between RRM1 and RRM2α (encompassing residues 210–450 in hLARP7), which contains several stretches of basic residues, may mediate this binding.[151] These findings also suggest intriguing functional interplay between HEXIM and LARP7 that merits further investigations.

Interactions between LARP7 and 7SK RNA is prerequisite to recruitment of P-TEFb to the 7SK RNP,[144] with inhibition of P-TEFb kinase activity controlled by the interaction of CycT1 with the C-terminal domain of HEXIM1/2.[141,152] The molecular association of HEXIM1/2 and P-TEFb with 7SK RNA/RNP is a dynamic process involving conformational rearrangements and intricate temporal and spatial multipartner interplay that is not yet fully understood.[144,151,153] Similar dynamic assembly of TER RNP has been attributed to LARP7 p43.[149]

A recent biochemical study that reconstituted the 7SK RNP from multiple purified components showed that it was functional for release of active P-TEFb and that its MePCE-mediated 5′-me-ppG capping activity was inhibited by LARP7.[154] This further showed that the region C-terminal to LARP7 xRRM is required for interaction with MeCPE in the context of 7SK RNA, for inhibiting its capping activity. This also provided evidence that xRRM in full-length LARP7 recognizes the 7SK RNA bulge at nucleotides CU 320–321 as important for the inhibition,[154] and that this recognition was proposed to occur in a manner similar to how p65 xRRM interacts with TER RNA.[35,36] The new data also supported two conformational states of the 7SK RNA,[154] including a ‘closed state’ consistent with juxtaposition of the 3′ and 5′ ends of 7SK RNA.[116]

Involvement of LARP7 in Other Activities

In agreement with a proposed role of LARP7 in p-TEFb recruitment to 7SK RNP, LARP7 mutations or its down regulation are associated with gastric, breast, and cervical tumorigenesis,[155–157] presumably by increasing the growth-promoting activity of P-TEFb.[17,123,158] However, data have shown that in ES cells knockdown of LARP7 is paradoxically not

© 2017 Wiley Periodicals, Inc.

associated with increased cell growth and proliferation as expected via activation of P-TEFb, but instead leads to growth failure.[159] This apparent embryonic-specific effect may help explain cases of loss-of-function mutations of LARP7 associated with primordial dwarfism (PD), a condition of severe growth restriction and associated symptoms.[160–163] The exact mechanism of the unexpected activity of LARP7 in early development is yet unclear although it was proposed to be mediated by interaction with Lin28 mRNA and the poly(A) polymerase Star-PAP in a pTEFb-independent manner.[159]

Another study proposed that some symptoms of Alazami syndrome (specific PD with LARP7 mutations) may result from translational and/or nucleolar stress in neural cells.[164] Although LARP7 is generally nucleoplasmic, it was convincingly found among proteins specifically enriched in nucleoli of rat neurons.[164] Interestingly, in hippocampal neurons, LARP7 knockdown reduced perikaryal ribosome content and protein synthesis.[164] How LARP7 executes this function had been unclear but is explained by recent findings that show that the 7SK snRNP plays a positive role in promoting the transcription of snRNAs and the small nucleolar (sno)RNAs, U3, U8, and U13 which are involved in processing the precursors of large ribosomal RNA during ribosome biogenesis.[165] This function of 7SK snRNP is mediated by interacting with an RNAP II elongation complex that is dedicated to the transcription of snRNAs, known as the little elongation complex, and requires LARP7.[165]

RECENT ADVANCES ON THE CYTOPLASMIC LARPS

Ample evidence indicates that members of LARPs 6, 4, and 1 families are mostly cytoplasmic (although nuclear–cytoplasmic shuttling suggests nuclear functions for some) and share a common function in mRNA binding, translation, and/or stability. Therefore, we will first provide a brief overview of relevant regulatory factors and other issues related to mRNA translation and metabolism, after which we review advances on cytoplasmic La followed by LARPs 6, 1, and 4. Several excellent reviews on mechanisms of translation initiation are available.[166–168]

PABP: A Central Factor in mRNA Metabolism

For the purposes of this review, we will highlight a few relevant points that are schematized in Figure 9 (a). mRNAs contain poly(A) 3′ tails of variable length that are bound by one or more molecules of PABP, a multifunctional protein that interacts with a variety of other proteins and heteromeric complexes that carry various enzymatic activities, and integrates translation initiation, termination, and ribosome reinitiation or cycling, as well as mRNA decay/stability functions. Eukaryotic mRNAs contain a 7-methyl-G(5′)ppp(5′)N 'cap' on their 5′ ends, that is recognized by the cytoplasmic cap-binding protein known as eukaryotic initiation factor-4E (eIF4E), which associates with eIF4G, a large protein that serves as a central hub for several other translation factors.[168] PABP contains four RRMs and a C-terminal region that includes an mademoiselle (MLLE) domain. The eIF4G interacts with the backside of RRM2 of PABP when bound to poly(A).[169] Interaction of PABP with eIF4G somehow increases the functional affinity of eIF4E for the 7mG-cap.[172,173,175] Bimodal interaction of PABP with poly(A) at the 3′ end of the mRNA and with eIF4G at the 5′ end accounts for the 'closed-loop' model of translation.[174,176–178] Accordingly, mRNAs with 5′ capped ends and 3′ poly(A) tails are efficiently engaged in translation (Figure 9(a)). In addition, by binding to the RNA termini, proteins can protect against exonucleases, a major source of mRNA decay.[179] In this model, which can be conditional on cell type,[167] PABP activities promote initiation, termination and ribosome recycling, as well as stabilization of the mRNA[44,180] (Figure 9).

Members of the cytoplasmic LARPs 1, 4, and 6 have been shown to interact with PABP, by use of a short peptide sequence that conforms to various degrees to a consensus motif termed PAM2,[16,18,26] and for LARPs 1, 4, and 4B evidence of a second interacting region also exists.[19,25,26,43,181] Poly(A) is bound by PABP via its four tandem RRMs with collective binding affinity of ~20 nM.[182] Poly(A) tail length can vary from ~25 to 250 nucleotides,[170,180] and multiple PABP molecules can bind with a periodicity of ~30 nucleotides.[183]

In addition to interacting with poly(A), PABP is a hub for a multitude of proteins that interact with one or another of its RRMs and/or its C-terminal MLLE domain.[170] Paip1 is a translation initiation regulatory factor that interacts with two separate regions of PABP via two Paip1 motifs, one of which is a 12 amino acid PAM2 sequence.[184] Several other proteins involved in translation and/or mRNA poly(A) metabolism also contain PAM2 sequences that bind the MLLE of PABP.[171,185] These 'PAM2 proteins' include stimulators and inhibitors of translational initiation (Paips 1 and 2, respectively), the translation

© 2017 Wiley Periodicals, Inc.

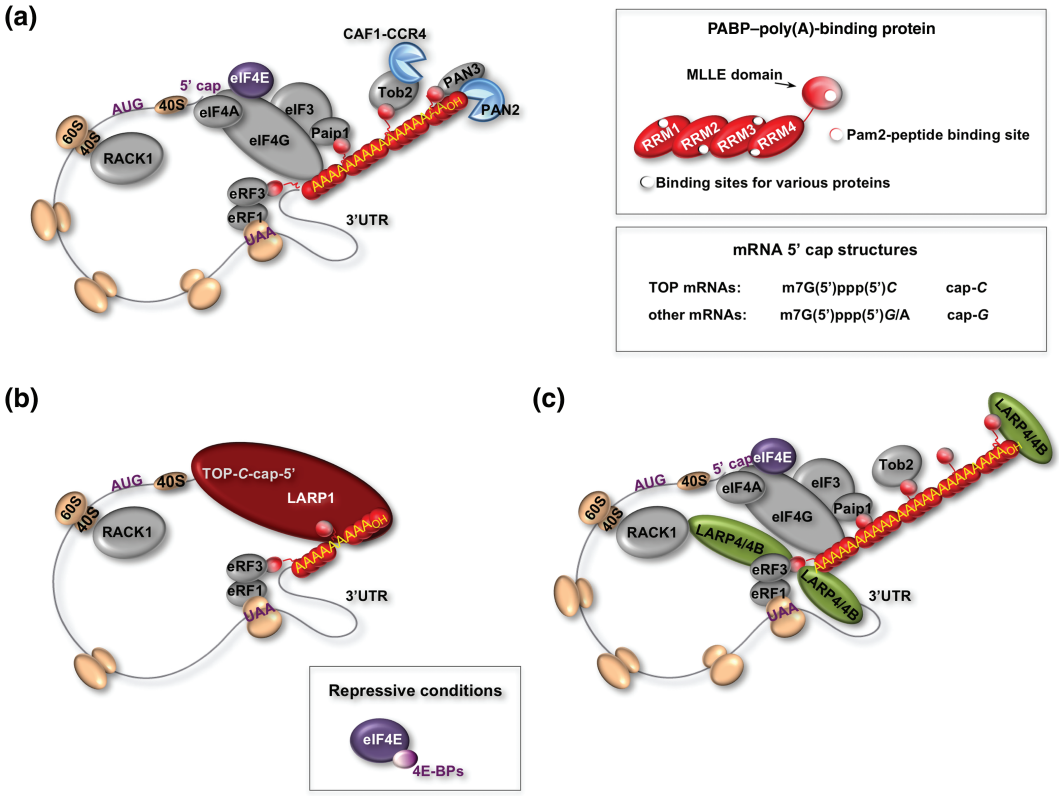


FIGURE 9 | Working models of translation and La-related protein-poly(A)-binding protein (LARP-PABP) factors. (a) Schematic showing factors involved in an example network of protein–protein and RNA–protein interactions that serve to bridge the 5′ and 3′ end regions of the mRNA according to a closed-loop model of translation. A translation initiation (start) codon AUG, and a termination (stop) codon, UAA are to orient polarity. Individual factors are referred to in the text. Multiple copies of PABP can bind to the poly(A) tail via its four RNA recognition motif (RRM) domains (depicted in highly schematized form), some of which also serve as separate docking sites for other factors including eIF4G and Paip1.[169,170] The MLLE domain of PABP, depicted as red circles, is used to interact with different PAM2-consensus sequence-containing proteins involved in translation initiation (Paip1), termination (eRF3), or recruitment of deadenylase-containing complexes (CAF1-CCR4 and PAN3-PAN2).[170,171] The eIF4E cap-binding positive initiation factor and its association with eIF4G are depicted (for more comprehensive reviews see Refs 172–174). An inset reveals the eukaryotic mRNA two different cap structures; 5′-TOP mRNAs contain a cap-**C** (m7GpppC), and other mRNAs contain a cap-G/A (m7Gppp-G/A). (b) Under repressive conditions such as nutritional or other stress, 4E-BPs (eIF4E-binding proteins), sequester the cap-binding protein eIF4E. A potential working model for LARP1 involvement in mRNA metabolism in relation to translation initiation and stabilization. It can bind via the DM15 domain to the 7mGppp**C**-TOP motif of TOP mRNAs and to PABP, and possibly to poly(A) (or other regions) of the mRNA via its La module (see text), although whether it would do so simultaneously as depicted here is unknown. LARP1 may either stimulate or inhibit translation and stabilize poly(A) under different conditions (see text). (c) Proposed working model for LARP4/4B involvement in translation and mRNA stability. LARP4/4B interacts with RACK1 and with PABP, although whether either of them would do so simultaneously is unknown, three of the combinatorial possibilities are depicted. LARP4B is proposed to bind to mRNA 3′ UTR sequences (see text).[27] LARP4 has been shown to bind PABP via the PAM2 motif, and through a second PABP-interaction motif (PBM). It is unknown to which part of PABP the PBM binds. LARP4 can also bind poly(A) RNA. Through competition with the PAM2 motifs of the deadenylases for the MLLE domain of PABP, LARP4/4B may protect mRNA 3′ ends from deadenylation, leading to poly(A) length modulation.[120] 4E-BPs, 4E-binding proteins, repressors of translation; Paip1, poly(A)-binding protein interacting protein 1, a stimulator of translation; eIF4E, eukaryotic initiation factor 4E, a.k.a., cytoplasmic cap-binding protein; eIF4G, eukaryotic initiation factor 4G; eIF3, eukaryotic initiation factor 3; eIF4A, eukaryotic initiation factor 4A; RACK1, receptor for activated kinase C, a 40S ribosome subunit; eRF1 and eRF3 are eukaryotic translation termination/release factors. Tob2, transducer of ERBB2; CAF1-CCR4, chromatin assembly factor 1 and CCR4-NOT transcription complex subunit 6, two proteins with poly(A) deadenylase activity; PAN2, PAB-dependent poly(A)-specific ribonuclease subunit PAN2, catalytic subunit of the poly(A)-nuclease (PAN) deadenylation complex; PAN3, PAB-dependent poly(A)-specific ribonuclease subunit PAN3, regulatory subunit of the poly(A)-nuclease (PAN) deadenylation complex.

termination/release factor eRF3, the poly(A) deadenylases (PAN3 and CCR4 complexes), and several others, and their regulated competition for PABP likely underlies important physiologic homeostasis mechanisms.[171] The PAM2 consensus sequence contains conserved aromatic and hydrophobic residues at key positions that when mutated diminish interaction with PABP MLLE.[171] LARPs 4 and 4B have a PAM2

© 2017 Wiley Periodicals, Inc.

just upstream of their La modules that was shown for LARP4 to be functional.[17,26]

Similar to Paip1, other PAM2 proteins contain a second region that interacts with one of the four RRM domains of PABP.[170] LARP4 contains a second region, downstream of its La module (Figure 1), that interacts with PABP although the interacting region of PABP is unknown.[26] Mapping of interaction between LARP4B and PABP is consistent with this.[25] Thus, the La module of LARP4 binds poly(A) RNA and is flanked by two PABP-interacting motifs, PAM2 and PBM[26] (Figure 1).

Phylogenetic analyses suggest that acquisition of a PAM2 sequence proximal to the La module occurred recurrently during evolution of LARPs.[18] In addition to LARP4 and 4B, a PAM2 was also found just upstream of the La modules of plant LARP6b and 6c and shown to be important for PABP interaction.[18] We are unaware of any reports of hLARP6/Acheron interactions with PABP, consistent with lack of the same for plant LARP6a.[18]

LARP1 was shown to bind PABP in an RNA-independent manner.[43,181] Co-IP experiments showed that partial or full truncation of the DM15 region of LARP1 substantially reduced but did not ablate interaction with PABP.[19] A PAM2 sequence candidate that resides in the predicted α1 helix of hLARP1 RRM[1] was identified.[16] Partial reduction of PABP co-IP with DM15-truncated LARP1 is consistent with two interacting regions.[19] It should be noted that this candidate lacks sequence spacing features important in other PAM2 motifs[171] and it maps to a predicted α-helix in the RRM whereas other PAM2 peptides are unstructured upon interacting with PABP MLLE.[171] These features together with significantly decreased binding to PABP after mutation of a key aromatic residue in this putative PAM2[16] suggest the possibility that this is an atypical interaction. Curiously, acquisition of a PAM2 or variant thereof within the La module of LARP1 would strengthen the prior noted association.[18] It is tempting to speculate that because its N-terminal region contains an eIF4G-like domain,[16] localization of the PAM2 within RRM would better ensure tight linkage of PABP and La module activity. In any case, we note that this PABP-interaction region maps to the α1 backside of the canonical RNA-binding surface, a region of human and *S. pombe* La proteins to which residues required for nuclear export reside.[9] Nonetheless, it will be important to address if this LARP1 candidate PAM2 sequence motif is mediating poly(A) RNA-independent interaction with PABP and if so via which domain.

The m^{7}GpppN cap, Translation Initiation Factors, and 5′TOP mRNAs

Genetic and biochemical studies of yeast have been invaluable toward understanding basic mechanisms of translation and regulation thereof, including central importance of the m^{7}G cap-binding factor, eIF4E and associated proteins, PABP and mRNA poly(A) metabolism.[186] However, certain differences between yeast and mammalian cells in their mechanistic and regulatory control of translation are pertinent here (e.g., see Ref 168). The mRNAs that encode the ~90 ribosome protein subunits and other translation factors, including PABP are each present at high copy number and collectively comprise ~20% of total cellular poly(A) mRNA.[187] In mammals these share a consensus known as the 5′ terminal oligopyrimidine (5′TOP) motif[187] (not found on the analogous yeast mRNAs). In addition to a 5′ tract of 5–15 pyrimidines followed by G + C-rich sequence, the 5′TOP motif is distinguished by a **C** at the 5′ position, whereas the ~80% of other mRNAs have A or G. Thus, 5′TOP mRNAs bear an m^{7}GpppC-cap (5′cap-C) while the others have a m^{7}GpppG/A-cap (5′cap-G)[187] (Figure 9(a), right lower).

The mammalian target of rapamycin (mTOR) is a protein kinase with multiple substrates that promotes cell growth and proliferation in response to nutrition-associated factors such as insulin and other cues.[188] Effects of mTOR are to increase the protein synthetic capacity of cells in part by regulating the 5′ cap-binding protein, eIF4E, and promoting translation of 5′TOP mRNAs that encode ribosome subunits and translation factors. Positive regulation of eIF4E activity, as with stimulation of mTOR activity after treatment of cells with insulin involves its release from the eIF4E-binding proteins (4E-BPs), which are repressors of translation initiation.[189,190]

Poly(A) Binding by Some La Modules

As alluded to above, *C. elegans* LARP1 preferentially binds poly(G) over the other homopolymers whereas hLARP1 was found to associate with poly(A) but not G, C, or U in an extract-based system.[119] The basis of this species-specific difference is unknown and emphasizes the need for more well-defined RNA-binding analyses of the La module. Comparison of the four homopolymers as well as RNAs with various combinations of mixed nucleotides, identified poly(A) as a preferred ligand of the hLARP4 La module[26] (S. Gaidamakov and R.J. Maraia, unpublished). Photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation (PAR-CLIP)[191] was used to evaluate hLARP4B for *in vivo* RNA-binding sites.[27]

© 2017 Wiley Periodicals, Inc.

Because this crosslinking method requires a UV-activatable Uridine analog, 4-thioU, it introduces strong bias for U in the target sequence. Nonetheless, the top eight LARP4B binding sites isolated, comprising 76% of all of the target sites quantified, are composed of 66% A residues and 27% U residues.[27] Thus, despite an intrinsic bias for U imposed by PAR-CLIP, LARP4B revealed a preference for A-rich RNA.[27] To the best of our knowledge LARP4B has not been examined for relative binding to the homopolymers, nor has LARP4 been subjected to PAR-CLIP.

The well-recognized binding specificity of hLARP6 is a conserved SL in the 5′ UTR of the collagen mRNA[23,24] and although it is capable of interacting with oligo(U), oligo(A), and oligo(G) these interactions are between 30 and 100-fold weaker than with collagen SL[23] (U. Abongwa and M.R. Conte, unpublished). Of the plant LARP6 members tested, AtLarp6a, exhibited preference for poly(A), which is curious because this protein has conserved all of the six key LaM residues (see Figure 4 and below), while AtLarp6c preferred poly(U).[18]

The La module of hLa exhibits binding to poly(A) with apparent affinity similar to LARP4 (S. Gaidamakov and R.J. Maraia, unpublished). Also, hLa exhibits poly(A) binding in the cytoplasm (M.A. Bayfield, manuscript in preparation). This suggests that the La module may have some propensity for poly(A) recognition that further evolved for some of the cytoplasmic LARPs.

CYTOPLASMIC LA PROTEIN

The great majority of La protein is obviously nuclear by cell staining. La is soluble and unlike many nuclear proteins it readily leaks out of nuclei upon cell lysis. This feature as well as the transient nature of its binding (i.e., dissociation) to nascent transcripts noted above, led some researchers to question if La had any real function in the cytoplasm of living cells. However, several studies helped characterize cytoplasmic La as a distinct entity of functional significance. As noted above La contains multiple subcellular trafficking elements and was demonstrated to shuttle from nuclei to cytoplasm.[4–6,9,10] It was known early that hLa was targeted to small nuclear (sn) transcript precursors of cytoplasmic tRNA and 5S rRNA and U6 snRNA synthesized by RNAP III,[192] as well as the small VAI/II and EBER RNAs encoded by the adenovirus and Epstein-Barr DNA viruses respectively, also transcribed by RNAP III.[193,194] However, work on La and RNA viruses demonstrated that only the cytoplasmic fraction of La associated with short leader transcripts synthesized by the viral RNA-dependent RNAP.[195] Nonetheless La exhibited preference for the leader transcript with the longer 3′ U tail whereas the viral N protein showed no preference.[195] Another development that distinguished cytoplasmic La was antibodies for phosphoSer-366-specific and nonphosphoSer-366-specific isoforms of hLa which demonstrated that these discrete isoforms were separately localized in the nucleoplasm and cytoplasm, respectively.[37] Parallel biochemical analyses suggested that ~80% of hLa is phosphorylated on Ser 366.[37] Since ours and other data indicate about 2×10^7 molecules of La per HeLa cell,[196] the amount of cytoplasmic nonphosphoSer-366 was estimated at ~10^6 molecules/cell.

Multiple phosphorylation sites were mapped to hLa, Thr-302, Ser-325, Thr-362, Ser-366, and Thr-389[12,197,198] and some were shown to control an activity *in vitro* and/or *in vivo* and/or to be of regulatory significance.[13,32,37,57,199–202] Of the multiple phosphorylated sites characterized by Pruijn et al., it was noted that Ser-366 was by far the most abundant[197] consistent with this isoform associated with pre-tRNAs.[37] It is important to note that phosphoSer-366 does not direct the nuclear or subcellular localization of La protein.[57,197] Rather, Ser-366 phosphorylation status more likely reflects colocalization and/or accessibility of the responsible kinase, CKII[201] and relative paucity of phosphatase. This further suggests that upon relocalization from nucleus to cytoplasm, La phosphoSer-366 may be dephosphorylated although this has not been examined. In any case, it has been shown that the phosphorylation status of Ser-366 does not influence the nuclear import or nuclear retention of La, and nor does a nonphosphorylatable hLa-A366 mutant distribute differently from WT hLa.[197] In addition, although the other phosphorylation sites may appear minor relative to total La levels they could be a substantial fraction of cytoplasmic La. hLa also undergoes sumoylation of the LaM on K41 which promotes its retrograde transport from neuronal axon processes to the nucleus,[203] and on two sites between RRMs 1 and 2 which decrease its binding to 5′TOP mRNAs and internal RNA elements.[204]

Phosphorylation Status of hLa Ser-366 Controls SBM Activity

Contrary to the lack of effect of Ser-366 phosphorylation on subcellular distribution is its striking effect *in vitro* and *in vivo* on RNA binding and related activities mediated by the adjacent SBM region

© 2017 Wiley Periodicals, Inc.

C-terminal of RRM2α (Figures 1, 2). As noted above, nucleoplasmic La is phosphorylated on Ser-366, but when it is nonphosphorylated, biochemical and genetic evidence indicate that the SBM confers high-affinity recognition of the 5′ end of pre-tRNAs, and leads to inhibition of their processing by RNase P (*in vitro* and *in vivo*) that is alleviated by Ser-366 phosphorylation.[11,57] Analysis of tRNA alleles epistatic for RNase P processing support these findings[57] as does a La-A366 mutant that cannot be phosphorylated *in vivo* which is dominant negative for pre-tRNA maturation.[57,205] Direct binding showed that hLa exhibited higher affinity for RNA with 5′ triphosphate versus 5′ dephosphorylated RNA, that the 5′ triphosphate moiety increased nonphospho-La's inhibitory effect on processing by RNase P, and that inhibition required an intact SBM region in hLa.[11]

This 5′ triphosphate binding that was localized to the SBM of hLa is sensitive to methylation of the γ-phosphate of the 5′ triphosphate as is found on the 7SK, U6 and a few other RNAP III transcripts, and is also weakened by addition of the m^7G cap found on mRNAs.[136] The SBM and sequences up to position 366 are part of an unstructured region that is basic whereas amino acid side chains following Ser-366 are acidic. It is plausible that phosphorylation of Ser-366 would alter the local conformation of this region and inhibit the RNA-binding activity of the SBM (see figure 3 in Ref 13 and figure 5 in Ref 32). Again, nucleoplasmic La is phosphorylated on Ser-366 and does not interfere with pre-tRNA processing whereas the cytoplasmic form is nonphosphorylated on Ser-366.[37]

Cytoplasmic La and the Short Noncoding Leader Transcripts of RNA Viruses

The first identified specific association of La with an RNA other than from RNAP III was the short leader noncoding transcript of the negative strand RNA virus, vesicular stomatitis virus (VSV).[195] Similar results were soon confirmed for La specific association with short leader transcripts of additional RNA viruses including the deadly rabies virus[206,207] and extended by others to the short noncoding leaders of Parainfluenza and Rinderpest viruses.[208,209] These short 5′ triphosphate-containing leader RNAs of negative strand RNA viruses are the first transcripts synthesized by the RNA-dependent RNAPs and are not modified with a 7-methyl-guanosine-cap (m^7G-cap), although the subsequent longer protein-coding transcripts are.[210] General commonalities of these leader RNAs are short length, ~50 nt, tendency toward A + U-rich 3′ tails, and 5′ ppp termini. Also notable is that the leader RNAs originate from the 3′ end of the viral template RNA, that VSV and others cause redistribution of La to the cytoplasm, and that the VSV leader was shown to be bound by cytoplasmic La rather than nuclear La.[195]

The negative strand RNA viruses cause diverse diseases that range in severity from mild to lethal. They typically produce 10^5 copies of RNA-dependent RNAP in the cytoplasm of an infected host cell within hours of infection.[210] As with other infectious bodies, foreign material, especially cytoplasmic nonself RNAs, can be strong activators of the host cell antiviral innate immune response. Among other host cell RNA sensors, RIG-I contributes to the early induction of antiviral type I interferons in negative strand viruses while another RNA sensor, MDA5 is important for detecting a distinct overlapping set of viral RNAs.[210,211] RIG-I requires a 5′ triphosphate on the RNA for activation of its interferon response activity[212] (reviewed in Ref 210). The 55 nt noncoding leader of the measles virus ends with a 3′ CUU-OH terminus and contains a 5′ triphosphate that is required for RIG-I activation.[212] Viruses have been in an arms race with hosts and have collectively developed numerous ways to avoid or delay detection by host cell surveillance, in some cases by hijacking host factors.[213]

The negative strand RNA respiratory syncytial virus (RSV) is a major cause of mortality in infants on all continents, second only to malaria as a lethal pathogen worldwide.[214] The RSV 44 nt noncoding leader RNA contains a 5′ triphosphate, terminates with U-OH 3′ and is efficiently bound by La early in infection.[215] RSV infection causes redistribution of La to the cytoplasm; and binding of leader RNA by La facilitates RSV proliferation.[215] Decreasing La levels by knockdown increases leader RNA association with RIG-I and interferon induction.[215] Thus, it would appear as if the leader RNAs of negative strand viruses engage cytoplasmic La to mask their 5′ triphosphate danger signal which would otherwise activate the surveillance factor, interferon activator, RIG-I. These viruses hijack La to shield their leader transcripts from detection and attenuate early activation of innate immune signaling.[215]

Cytoplasmic La and Virus mRNA IRES-Dependent Translation

That La associates with viral leader transcripts prompted studies of additional RNA viruses that were found to use La for translation of their mRNAs, including poliovirus, HIV, and HCV.[216–222] Some of

© 2017 Wiley Periodicals, Inc.

these, including poliovirus do not m^7G-cap their RNAs for cap-dependent translation initiation and instead use an IRES structural element that encompasses the start site AUG to direct initiation.[223] Indeed, viruses that use IRES-mediated initiation employ mechanisms to deactivate cap-dependent initiation and other means to direct the translation machinery to viral protein production.[224] Poliovirus, HCV, and other viruses rely on a number of host factors to promote IRES-mediated translation of their mRNAs.[213] La was the first such host factor found that promoted IRES-mediated translation initiation.[216,217] La function in IRES-mediated mRNA translation was demonstrated *in vivo* for HCV mRNA[225] and was shown to promote IRES-mediated translation of several viral and cellular mRNAs.[213,226,227] Of the ~10 proteins produced by poliovirus, two proteases target host cell cap-dependent initiation factors for cleavage. Protease 3C substrates include PABP and La, the latter of which is cleaved at position 358, which removes Ser-366 and the downstream nuclear localization signal (NLS) resulting in cytoplasmic accumulation of a high-affinity RNA-binding form of La presumably to facilitate virus mRNA translation.[228,229]

Biochemical experiments using mammalian translation systems reconstituted with purified factors indicate that La can enhance eIF4E-, PABP-, and m^7G cap-dependent translation by inhibiting spurious initiation at aberrant internal start sites.[230] Moreover, inhibition of cap-dependent translation is exacerbated by increasing La as PABP becomes limiting.[231] These studies provide insight into La function as a positive factor in directing IRES-mediated translation initiation when cap-dependent mechanisms are inhibited.

Of note are recent advances toward understanding the roles of La in the translation and replication of HCV RNA virus.[232–237] These include mechanisms of HCV RNA recognition by La, granzyme H-mediated redistribution of La to the cytoplasm, and effects on telomerase activity in HCV infected cells.[112,238,239] Some studies have led to development of compounds to target La by pharmaceutical inhibition of HCV.[233] Another line of work has established that La binds and stabilizes Hepatitis B virus RNA.[133,240] Phosphorylation of hLa Ser-366 by protein kinase CKII is required for the positive effect on HBV expression.[200]

La and TOP mRNAs

As an oligo(U)-binding protein it was reasonable that La may function in translation of 5′TOP mRNAs.[241,242] Subsequent investigations showed that cytoplasmic nonphosphoSer-366 La is specifically associated with 5′TOP mRNAs[37] and that increasing the fraction of nonphoshoSer-366 by inhibiting protein kinase CKII increases the amount of 5′TOP association.[201] As noted above, while 5′TOP mRNAs are efficiently translated in proliferating cells their physiologic regulation involves translational repression in times of metabolic and/or growth quiescence. Consistent with this, in a Xenopus cell line, several 5′TOP mRNAs were 80 and 40% loaded on translating polysomes in growing and resting cells, respectively.[243] In the same study, overexpression of WT La and a dominant-negative (dn)La had no effect on 5′TOP mRNA translation efficiency in growing cells whereas in resting cells wtLa increased translation efficiency from 40 to 50% and dnLa decreased it from 40 to ~30%.[243] Thus, these data indicate that La promoted 5′TOP mRNA translation but only in the resting cells, in what would appear to be in repressing conditions.

Other studies have indicated a negative impact of La on 5′TOP mRNA translation.[201,244] Later experiments revealed no specific effects of La on the translation efficiency of 5′TOP mRNAs.[245]

It is conceivable that with overexpression of La or under conditions in which La redistributes to the cytoplasm, and when m^7G cap-dependent translation is limited or weak, the 5′-ppp cap like binding activity of the SBM of nonphosphoSer-366 La could negatively modulate m^7G cap-dependent translation initiation[201,230,231,244] and promote IRES translation.

La Chaperone Activity in mRNA Translation

Some cellular mRNAs use IRES elements of one type or another to direct cap-independent translation initiation[226] and several of these have been shown to engage La for optimal function.[32,134,246–249] La can therefore function as an mRNA translation factor, possibly employing both its N-terminal and C-terminal fragments both of which harbored comparable RNA chaperone activities.[250]

It was known from early studies that the ability of La to promote IRES-mediated translation was dependent on amino acids 293–380.[217,219,251] La was shown to promote IRES-mediated translation of CCND1 mRNA in tumors.[134] A significant advance toward linking the RNA chaperone and translation activities of La came from studies of the CCND1-IRES mRNA.[32] This IRES mRNA chaperone activity was mapped to amino acids 333–390 of hLa and regulated by AKT-mediated phosphorylation of

© 2017 Wiley Periodicals, Inc.

Thr-389.[32] It was noted that the disordered C-terminal region of yeast La that is important for correct folding of tRNA may bear some similarity to this region.[31,32] As noted in an earlier section, this is in a very intricate region of hLa, comprised in part of the SBM (Figures 1 and 2) proposed to share features with a Walker A-like nucleotide-binding motif[56] and found to interact with the 5′ pppG end of nascent pre-tRNAs as part of a phosphoSer-366-mediated control of their 5′ processing by RNase P,[11,57] and to overlap with the nucleolar localization signal.[13,14]

RECENT ADVANCES ON LARP6 FAMILY MEMBERS

Brief Overview of LARP6 Phylogeny

LARP6 (a.k.a. Acheron in metazoa) members are characterized by a La module followed by a LARP6-specific domain, denoted LaM and S1-associated (LSA) due to homology to part of a repeat found in S1 ribosomal protein and structurally related nucleic acid binding folds.[1] LARP6 is encoded by one gene in protists, invertebrates, and eutherians, two in other vertebrates, and 3–6 in vascular plants.[18] The phylogeny of LARP6 appears more complex than for other LARPs. In higher plants there are three LARP6 subclusters, a, b, and c, of which b and c contain N-terminal PAM2 peptides for PABP binding, and these have been characterized for the *Arabidopsis* At LARP6 proteins.[18] As noted, phylogenetic emergence of the PAM2 was associated with divergence of some of the key hydrophobic pocket residues in AtLARPs 6b, 6c and LARPs 4, 4B.[18] Thus, plant LARP6a and hLARP6 are most comparable, whereas plant LARPs 6b and 6c are more distinct from these.[18]

We should clarify that LARPs 6b and 6c exist in plants but not in vertebrates, LARPs 4 and 4B exist in vertebrates but do not exist in plants, whereas LARP6/a is common to vertebrates and plants. Thus, vertebrates and plants each contain two LARPs that share N-terminal PAM2 sequences and whose LaM hydrophobic RNA-binding pockets underwent partial remodeling but are of different LARP subfamily memberships.[18] It seems plausible that the two PAM2-containing LARPs in vertebrates (LARPs4/4B) and plants (LARPs6b/c) participate in similar mRNA-associated functions.

Arabidopsis thaliana LARPs6

The three *At*LARPs6 likely fulfill distinct functions as mRNA-associated proteins with different subcellular localizations and RNA- and PABP-interaction networks.[18] AtLARP6a mainly localizes to nucleolus and cytoplasm, recognizes poly(A), neither bind or colocalize with PABP nor relocate to stress granules upon hypoxia.[18] AtLARP6b has similar localization as 6a but associates with PABP and colocalizes to stress granules.[18] AtLARP6c binds PABP and goes to stress granules, but was also observed in nucleolus, nucleoplasm, and cytoplasm.[18] RNA interaction profiles of two AtLARPs6 show clear differences: 6a prefers oligo(A) while 6c prefers oligo(U) (any preference by 6b is unknown due to limitations of protein stability).[18] Also as noted, while AtLARPs 6b and 6c bind PABP, AtLARP6a does not.[18]

Functional studies on metazoan LARPs6

Studies on the tobacco hawkmoth, *Manduca sexta*, first uncovered LARP6, referred to as Acheron, as critical for myogenesis in a manner that involves programmed apoptosis; a role in myogenesis for LARP6 was also found in mice and fish.[41,252,253] In vertebrates, LARP6 has been reported to control muscle differentiation and development[41] by acting upstream of the key myogenic transcription factor MyoD.[252] LARP6 may regulate myogenesis via interaction with the CaMKII-like domain of the developmental transcription factor CASK-C and/or association with the Id (inhibitor of differentiation) transcription repressors.[253] Since Ids modulate regulatory factors such as MyoD,[254] LARP6 may contribute to myogenic differentiation through interactions with Id proteins. Indeed, ectopic MyoD can rescue differentiation in C2C12 myoblasts expressing a dominant-negative form of LARP6.[253] The complex LARP6/CASC-K/Id may therefore regulate development at the transcriptional level, thereby assigning this function to the nuclear fraction of LARP6. While it is unclear if the above function of LARP6 entails interaction with yet unknown RNA targets, specific RNA binding is unequivocally required for its function in regulated synthesis of collagen α1(I), α2(I), and α1(III) chains, which as reviewed in an earlier section, is mediated by interaction with the SL in the 5′-UTRs of these mRNAs.[24] These 5′ SL sequences encompass the translation initiation codon and are highly conserved in all vertebrates.[255]

The role of LARP6 in orchestrating collagen production was recently reviewed.[256] While synthesis of type I collagen is critical for normal growth and regulated response to wound healing, excessive production can cause fibrosis and be pathologic. Disruption of the LARP6–SL colα1–mRNA interaction decreases collagen production, making LARP6 appealing as a target for treatment of

© 2017 Wiley Periodicals, Inc.

fibroproliferative disorders.[257] On the other hand, the LARP6–SL interaction is a critical parameter in mediating protective collagen synthesis in atherosclerosis by reducing vulnerability of plaque rupture. In this case, insulin-like growth factor-1 (IGF-1) regulates LARP6 and resultant collagen expression in vascular smooth muscle cells.[258] It is notable that this study showed that a Colα1 5′ SL RNA decoy could block collagen synthesis in response to IGF-1.[258]

At a molecular level, the La module of LARP6 is proposed to facilitate coordinated translation of collagen α1(I) and α2(I) mRNAs, which occurs with 2:1 stoichiometric ratio, by binding to a bulge in the SL and modulating accessibility to the translation start codon located in the adjacent double helix stem.[24] LARP6 also participates in trafficking of these mRNAs in the nucleus and cytoplasm.[24] It was postulated that LARP6 interacts with a number of molecular chaperones to choreograph collagen production, including nonmuscle myosin[259] SEC61 translocon,[260] vimentin,[261] and RNA helicase A (RHA)[262] (reviewed in Ref 256). LARP6 has also been reported to recruit the serine-threonine kinase receptor-associated protein (STRAP) to collagen mRNA, thereby preventing uncontrolled translation mainly of collagen α2(I) mRNA by a mechanism thought to involve eIF4A.[263] STRAP (a.k.a. unrip) belongs to a family of WD-repeat proteins and was found to interact with the last 27 amino acids of hLARP6,[263] which map to the LSA domains[1] (Figure 1). Importantly, in addition to its significance with regard to STRAP, this work also attributes a specific protein–protein interaction function to the LSA motif. On the other hand, vertebrate LARP6 members share with numerous other proteins a slightly larger region encompassing the LSA, termed SUZ-C (supplementary figure 2 in Ref 264), that was proposed to participate in RNA binding by the SYZ-20 protein. In hLARP6, SUZ-C maps to residues 453–488. Clarification on the functional role of the LSA in the LARP6 family of proteins awaits further investigations.

LARP6 appears to play a role in cancer as it enhances proliferation, lamellipodia formation, and invasion activity of breast cancer cells. These functions appear linked to nuclear activity of LARP6, as loss of its NLS abrogates these effects.[40] Interestingly, the effects on cell morphology and motility appear to be the reverse of what is observed with LARP4, where cell motility was increased by depleting LARP4[265] (below).

LARP6 function in collagen synthesis can be modulated by phosphorylation.[266,267] Of note, hLARP6 appears to be a downstream target of Akt and phosphorylated on S451, although a known consensus sequence for Akt is not recognizable surrounding S451,[266] suggesting that it may be phosphorylated by a kinase downstream of Akt. In any case, S451A mutation or inhibition of Akt reduces the type I collagen secretion.[266] Furthermore, hLARP6 was reported to be phosphorylated on S348 and S409 by mTORC1, and in cells where the phosphorylation is prevented by mutating these serines to alanines, or by mTORC1 inhibition, collagen secretion decreases and hyper-modified collagen polypeptides are produced.[267] Also, lack of S348/S409 phosphorylation appears to weaken the interaction of hLARP6 with STRAP, although these serines are far from the putative STRAP interaction site on hLARP6 (464–491). Interestingly, LARP1 and LARP4 as well as many other proteins are also phosphorylated by mTOR.[268]

RECENT ADVANCES ON LARPs1

LARP1 was discovered prior to other LARPs and has been studied in a variety of organisms including *C. elegans*, drosophila, plants, humans, and mice. Several early studies, although not limited to those,[269] uncovered LARP1 through genetic screens to better understand developmental programs or other processes.[43,270–272] A recent review covers phylogeny and function of LARP1 in reproduction, proliferation, and oncogenic potential, as well as emerging links to translation, mRNA metabolism, and stress response pathways.[273] LARP1 has also received directed focus as a tumor progressor in human cancers, consistent with its mechanistic links to the mTOR pathway that controls cellular growth and proliferation in response to nutritional status, hypoxia, and other states relevant to cancer,[188] and has also been reviewed along with the other LARPs from this perspective.[42] In keeping with the aims of this review, we will focus on recent advances on the structure and function aspects of LARP1 and its mechanism of action as an RNA-binding protein.

Advances came from phylogenetic comparisons and recognition of the conserved region initially referred to as the LARP1 domain (Ref 270 and references therein) and more recently, DM15 as will be used hereafter.[1] A conserved region corresponding to DM15 was previously noted in La-related homologs of human, *C. elegans*, and two yeast proteins, Sro9p and Slf1p although neither of the latter contain DM15.[124] The DM15 box is a ~40 amino acid motif found in three consecutive repeats in the LARPs1 of

© 2017 Wiley Periodicals, Inc.

protists, fungi, plants, and animals that was shown to bind 5′TOP RNA.[137]

C. elegans LARP1 fragments containing its La module or DM15 region were examined for RNA binding[270]; the latter exhibited higher avidity for poly(U) consistent with recent 5′TOP binding[137] whereas the La module-containing fragment showed preference for poly(G).[270] An approach to identify new RNA-binding proteins, employing bait constructs carrying 3′ poly(A) either with or without a 5′-m^7GpppG cap was used on native human cell extracts.[119] Mass spectrometry analysis showed that LARP1 was higher ranking over three other proteins, PARN, PATL1, and eIF4A1, that showed preference for poly(A) specifically at the 3′ end of the bait.[119] Using noncapped RNA, LARP1 was bound preferentially by A(**10**) with specificity for A at the 3′ end but not A(**9**) followed by a single U, C, or G.[119] Although the assay as used cannot ensure direct binding to the RNA bait by LARP1, this specificity is reminiscent of RNA 3′ binding during protection from 3′ exonucleases by La protein although the mechanisms remain to be distinguished.[119] Accordingly, because of the noted relation of LARP1 to mRNA metabolism,[16,20,119,270,274–276] one may suspect that its La module might protect against deadenylation (Figure 9(b)). Cytoplasmic mRNA decay can be initiated by endonucleolytic cleavage or 3′ exonucleolytic poly(A) shortening,[277] and followed by 5′ and/or 3′ exonucleolytic digestion.[179,278,279] In plants, LARP1 associates with the 5′→3′ exonuclease XRN4 and is required for a heat-activated mRNA decay via 5′-decapping although if this pathway involves prior 3′ deadenylation is yet unknown.[275,276]

Human LARP1, 5′TOP mRNA Translation, and mTOR

Affinity matrix of m^7Gppp-sepharose was used to identify cellular proteins that associate with the m^7G-cap in an mTOR-dependent manner.[19] This revealed PABP as the isolated protein most increased by insulin, followed by other translation factors as expected based on earlier studies. It also identified hLARP1, which was increased by insulin and decreased by mTOR inhibitors, in both cases concomitant with reciprocal changes in the 4E-BPs.[19] In this study, hLARP1 was complexed with PABP, and was associated in polysome profiles made from serum-stimulated cells with 40S-80S ribosomes as well as in the same fractions as PABP in polysomes. Also, in hLARP1 knock down cells, 5′TOP mRNA levels were substantially reduced.[19] The results provide strong evidence that LARP1 associates with and stabilizes 5′TOP transcripts, which comprise a substantial fraction of the total pool of actively translating mRNAs in stimulated cells. The data were also consistent with a positive role for LARP1 in translation of 5′TOP mRNAs.

Fonseca et al. also examined hLARP1 interactions in response to mTOR and its inhibitors.[16] In agreement with Tcherkezian et al.,[19] they found direct interactions with PABP and Raptor[16] and also showed interactions with 5′TOP mRNAs. TOR inhibition decreased LARP1-Raptor and increased LARP1-5′TOP mRNA association, consistent with the identification of a TOR signaling (TOS) motif in the DM15 (Ref 16 and references therein).[137] Both studies showed most LARP1 in the RNP region as well as 40S–80S region, with less on polysomes with translating mRNAs and PABP.[16,19] In contrast to Tcherkezian et al. who propose a stimulatory role for hLARP1 in 5′TOP mRNA translation, data in Fonseca et al. led to a model in which LARP1 represses 5′TOP mRNA translation in response to mTOR activity. It has been proposed that LARP1 may function to either stimulate or inhibit translation, dependent on phosphorylation status or competitive binding to other components of the translational or mTOR machinery.[42]

As noted previously for LARP1, distinct RNA recognition modes by DM15 and the La module could juxtapose the 5′TOP and 3′ poly(A) and contribute to a closed loop configuration of mRNA, presumably facilitated by direct LARP1–PABP interactions.[42] New evidence shows that in addition to binding to the 5′TOP sequence, the DM15 of hLARP1 also binds to the m^7GpppC (cap-C) and sequesters the 5′TOP motif away from eIF4E, thereby preventing the formation of eIF4F complexes on 5′TOP mRNAs.[15] In this capacity LARP1 would inhibit translation initiation by the eIF4F pathway[15] (Figure 9(b)). In this model, the La module of LARP1 would protect the poly(A) 3′ tail from exonucleolytic decay and DM15 would protect the 5′ cap from exonucleolytic decay while the 5′TOP mRNA is in a state of repression.[187]

Mitochondria-Targeted mRNA Translation and *Drosophila* LARP1

Recent work identified LARP1 as a translation stimulator for a set of embryo mRNAs whose protein products must be targeted to the inner mitochondria for *Drosophila melanogaster* viability.[269] Earlier work identified *d*LARP as colocalized with mitochondria.[271] The mitochondrial outer membrane protein MDI was found to bind LARP1 in association with

© 2017 Wiley Periodicals, Inc.

PABP and eIF4G, and is required for insertion of cytoplasmic translated proteins that are essential for mtDNA replication into mitochondria.[269] Remarkably, targeting of LARP1 to the outer mitochondrial membrane in the absence of MDI restored protein synthesis to mitochondria and rescued *mdi*-mutant phenotypes.[269]

The 5′TOP motif is conserved among vertebrates and extends to Drosophila.[187] A cursory examination revealed that the *D. melanogaster* cytoplasmic mRNAs that are targeted by LARP1 for mitochondrial outer membrane translation do not show evidence of enrichment of the 5′TOP motif (R.J. Maraia, unpublished). The work on mitochondria[269] provides biological evidence that *D. melanogaster* LARP1 functions to promote translation in a 5′TOP motif-independent manner.

LARP1 DM15 Mediates m^{7}GpppC–TOP mRNA Interactions

As described above, biochemical and biophysical analysis of DM15-RNA cocrystals are uncovering structure-function aspects of hLARP1 and insight into its translational control of 5′TOP mRNAs.[15] Structural and biochemical analyses revealed that DM15 exhibits almost 100-fold greater binding specificity for TOP RNA bearing a 7mGpppC than a -G.[15] Moreover, the DM15 domain could outcompete the cap-G binding factor, eIF4E, when the RNA bears a 7mGpppC and can displace eIF4E from RNA with 7mGpppC but not with a 7mGpppG.[15] These results provide evidence that LARP1 is indeed a candidate for a most proximal translational regulator of 5′TOP mRNAs[187] and raise unanticipated questions. Do LARPs1 of other species exhibit similar activity in mTOR and nutrition control of 5′TOP mRNA translation according[187]? It will also be interesting to see how this DM15 may be used to identify other cap-C-containing RNAs. Other questions arise. To what extent can LARP1 bind cap-C RNA that is less pyrimidine-rich than typical 5′TOP mRNAs? What is the nature of the cap-C transcriptome, is it tissue- and species-specific, and to what extent is it under the control of LARP1? What is its stoichiometry relative to the 5′TOP mRNAs; in practical terms, to what extent must mRNAs compete for LARP1? Does LARP1 protect mRNA poly(A) from destabilization during TOR inhibition?

It is interesting to note from phylogenetic perspective that the DM15 region as a 5′-pppCap-C-RNA-binding motif is functionally reminiscent of the C-terminal region SBM in La that binds to the 5′-pppG-end of nascent pre-tRNAs, is sensitive to m^{7}G-G-cap, and can inhibit translation of 5'TOP mRNAs (ref 201).[11,136]

LARP1 and Cancer

There is strong association of LARP1 with cancer (reviewed in Ref 42).[274] In this capacity LARP1 appears to bind to a multitude of mRNAs including that encoding mTOR, the latter of which was shown to be elevated and stabilized by LARP1.[20] Although direct binding between LARP1 and PABP may account for much of the LARP1-interactome, it is also possible that LARP1 exhibits sequence-specific direct recognition of some of the mRNAs although this remains undetermined.

The TOR pathway regulates ribosome biogenesis to meet the needs of nutrition and factor-dependent cell growth and proliferation, and includes transient repression of 5′TOP mRNA translation when appropriate.[187,188] Ribosome biogenesis is a hallmark of cancer.[188,280] It will be interesting and challenging to see how cancers not only overcome the otherwise inhibitory role of LARP1 in translation initiation of the 5′TOP RNAs,[15,16] but also use LARP1 to their benefit.[42]

RECENT ADVANCES ON LARP4 FAMILY MEMBERS

Protists and invertebrates have one LARP4 gene, whereas gene duplication appears to have established independent lineages for LARP4 and LARP4B (previously known as LARP5) paralogs in an ancestral vertebrate.[1,18] Zebrafish is the only known organism with more than two LARP4-homologous genes,[18] designated LARPs 4aa, 4ab, and 4b. No LARP4 homologs appear in plants or yeasts.[1,18]

hLARPs 4 and 4B are cytoplasmic proteins encoded on chromosomes 12 and 10, respectively, and although they share general overall architecture (Figure 1), amino acid identity is 40% throughout including 74% in their La modules. They vary in degree to which their other motifs have been mapped.[25,26] Two PABP-interaction motifs flank their La modules and they also interact with the receptor for activated kinase C (RACK1), an integral component of the 40S subunit of the ribosome.[281–283] Thus, LARPs 4 and 4B interact with PABP and RACK1, associate with polysomes, promote mRNA translation, and localize to stress granules.[25,26] LARP4 expression stabilized reporter mRNA and the cellular mRNAs for which it was examined and found to be associated, e.g., FAIM.[26]

© 2017 Wiley Periodicals, Inc.

LARPs4 with Distinctive LaM Binding Pocket Sequence and RNA Preference

The hLARP4 La module was examined for binding to different RNA homopolymers using gel-shift assays and isothermal titration calorimetry which revealed specificity for poly(A) of ≥15 nt.[26] As detailed in an earlier section, PAR-CLIP revealed that hLARP4B was bound to A-rich sequences; the same study identified numerous cellular mRNA targets that were stabilized by hLARP4B[27] some of which were confirmed in other cells.[284] As noted previously and shown in Figure 3, the LaMs of vertebrate LARPs4 differ from each other, La, and other LARPs in some of the six otherwise highly conserved key amino acids involved in UUU-3′OH recognition[17,18] (Figure 4). hLARPs 4 and 4B differ from La and the other LARPs by having a cysteine at position 24 (hLa numbering) whereas all others have the aromatic Tyr (or Trp, LARP7) (Figure 4) contributing to the hydrophobicity of the binding pocket and making hydrogen bonds to the phosphate of U-1[3,114] (Figure 3(c)). For hLa and TbLa, mutation of Y24 was severely detrimental to RNA binding.[50,114] Moreover, hLARPs 4 and 4B differ from each other at two other of the six LaM binding pocket residues, Q20 and F55 (Figure 4). LARP4 has conserved Q20 whereas LARP4B has T at this position (Figure 4) and notably Q20 in hLa and LARP7 make base-specific contacts to U-2 which resides in the cleft between the LaM and the RRM1[3,114] (Figure 3(b)). Regarding the F55 position, hLARPs 4 and 4B contain M and L respectively in place of this otherwise nearly invariant aromatic residue (Figure 4). LARPs 4 and 4B also differ in 10 amino acids in their RRMs. It would seem that these differences may contribute to differential RNA preferences of LARPs 4 and 4B.

LARP4 Activity in mRNA and Poly(A) Metabolism

Unexpected insight into LARP4 function came from examination of its negative regulation which identified a central region in its mRNA as a codon-biased coding region determinant of instability that when substituted with synonymous codons, increased its mRNA and protein levels.[120] This revealed activity of cellular LARP4 to affect mRNA poly(A) tail (PAT) length[120] concomitant with mRNA stabilization[26] (Figure 9(c)).

RACK1 is a WD40 family member that is stably associated with the mRNA exit channel of ribosomes from protists, yeasts, and humans, via contacts to three 40S proteins as well as 18S rRNA.[281–283,285,286] RACK1 is a hub for interactions with several proteins in addition to protein kinase C, mediated through its WD repeats.[287] The yeast RACK1 homolog was shown to enhance translation of short ORF mRNAs according to the closed loop structure model.[288] Using yeast-2-hybrid assays, LARP4 was shown to interact with RACK1 propellers 6 and 7.[26] Reciprocal yeast-2-hybrid and biophysical methods have recently mapped the RACK1-interaction region (RIR) to a C-terminal sequence of LARP4 (Figure 1) (unpublished).

LARPs 4 and 4B Links to Cancer

LARP4 would appear to regulate cancer cell migration and invasion. Upon its knockdown, prostate cancer cells exhibited very elongated phenotype consistent with increased migration.[289] LARP4 knockdown increased cell migration and invasion of other cancer cells, whereas overexpression decreased elongation.[265] Moreover, more than 140 mutations in LARP4 have been identified in a multiple cancers, and six of these, chosen because they reside in the C-terminal regions of interaction with PABP and RACK1, were subjected to examination in the experimental cell migration system. Overexpression of some of the mutants were more effective than WT LARP4 toward reversing the antiinvasion phenotypes.[265] Of these, the S388* truncation mutant exhibited elevated association with PABP, which as noted had also been observed for other LARP4 C-terminal truncation constructs,[26] suggesting that a downstream region is inhibitory to PABP binding.[265]

LARP4B has been found to act as a tumor suppressor and oncogene, possibly dependent on the type of cancer, e.g., solid or blood. LARP4B was identified as a tumor suppressor by a genetic screen in mice[290] and more recently characterized in human glioma cells[284] (for more comprehensive review see Ref 291). Overexpression of LARP4B induced apoptosis and mitotic arrest, partially dependent on the La module suggesting that interaction with RNA and its other interacting proteins is important.[284] Knockdown of LARP4B in primary mouse astrocytes lacking p53 and Nf1 promoted cell proliferation, tumor size, and invasiveness.[284] In a study of an acute myeloid leukemia mouse model, LARP4B was found to act as an oncogene.[292]

LARP4 Regulation

The hLARP4 mRNA 3′UTR contains conserved AU-rich element (ARE) binding sites for Tristetraprolin (TTP) protein which negatively regulates mRNA

© 2017 Wiley Periodicals, Inc.

stability and translation.[293,294] LARP4 mRNA and protein but not LARP4B transiently decrease then increase reciprocally with TTP levels in response to tumor necrosis factor (TNFα),[293] an apical cytokine in initiation of inflammation that up-regulates other cytokines.[295] Unrestrained TNFα activity becomes pathogenic unless controlled in part by TTP.[296,297] Regulation by TNFα-TTP and evidence of function in mRNA stability[26,293] suggest that LARP4 may contribute to mRNA metabolism during fine tuning of the inflammatory response.[297]

hLARP4 mRNA contains a translation-dependent coding region determinant of instability that may be sensitive to tRNA levels.[120] Overexpression of limiting tRNA cognate to codons in this region increase LARP4 protein levels and its activity to mediate poly(A) length maintenance of heterologous mRNAs including those that encode ribosomal proteins[120] (Figure 9(c)). As increased tRNA production occurs in cancer as well as in response to growth promoting signals including insulin through mTOR,[298] it is possible that this may contribute to the regulation of LARP4 and its activity. This curious focus on codons by LARP4 may reflect an ancestral connection to La protein and tRNA surveillance that may extend to other LARPs.

CONCLUSIONS

When the LARP family was discovered less than a decade ago, most of its members were largely unknown. Conserved features were eye-catching but it soon became evident that each family had a distinct profile and in some cases species-specific members of the same family clearly had evolved to perform specialized tasks. The LARPs would appear to be phylogenetically rooted in an ancestral eukaryotic La protein which functions via 3′ end binding and chaperone activities directed toward the intranuclear handling of precursor tRNAs and other nascent transcripts of RNAP III. As investigations proceeded, commonalities among La and the LARPs have been emerging as discussed in this review, the most clear to date being links to the 3′ end regions of their RNA targets, which for the cytoplasmic LARPs include 3′ poly(A) binding and stability, mRNA translation, and interactions with PABP.

A good deal of high quality focused research has revealed LARPs as a key superfamily that share a unique and versatile RNA-binding module, that surely play important functional roles in eukaryotic biology that operate at fundamental points in RNA control. Collectively the LARPs tell a tale of 3′ and 5′ recognition of RNAs, details of which have revealed novel ways of interacting with nucleic acids. We believe that a lot more will be uncovered, from details of molecular interaction with RNA ligands as well as protein partners, to additional roles in translation and possibly extending back to tRNA. New challenges await that we and others will relish to take on.

ACKNOWLEDGMENTS

The authors thank Cécile Bousquet-Antonelli, Anne Catherine Dock-Bregeon for comments and discussion; Andrea Berman, Bruno Fonseca, and Mark Bayfield for generously and graciously reporting their data for mention prior to publication; Jack Keene for inspiration; and B. Fonseca for his rendition of Figure 9. This work was supported by the Intramural Research Program of the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development, NIH, the EU Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement 655341, and Royal Society Newton International fellowship NF140482.

REFERENCES

1. Bousquet-Antonelli C, Deragon JM. A comprehensive analysis of the La-motif protein superfamily. *RNA* 2009, 15:750–764.
2. Jacks A, Babon J, Kelly G, Manolaridis I, Cary PD, Curry S, Conte MR. Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif coupled to a helical nuclear retention element. *Structure* 2003, 11:833–843.
3. Kotik-Kogan O, Valentine ER, Sanfelice D, Conte MR, Curry S. Structural analysis reveals conformational plasticity in the recognition of RNA 3′ ends by the human La protein. *Structure* 2008, 16:852–862.
4. Fok V, Friend K, Steitz JA. Epstein-Barr virus noncoding RNAs are confined to the nucleus, whereas their partner, the human La protein, undergoes

© 2017 Wiley Periodicals, Inc.

nucleocytoplasmic shuttling. *J Cell Biol* 2006, 173:319–325.

5. Simons FH, Broers FJ, Van Venrooij WJ, Pruijn GJ. Characterization of cis-acting signals for nuclear import and retention of the La (SS-B) autoantigen. *Exp Cell Res* 1996, 224:224–236.
6. Rosenblum JS, Pemberton LF, Bonifaci N, Blobel G. Nuclear import and the evolution of a multifunctional RNA-binding protein. *J Cell Biol* 1998, 143:887–899.
7. Marchetti MA, Tschudi C, Kwon H, Wolin SL, Ullu E. Import of proteins into the trypanosome nucleus and their distribution at karyokinesis. *J Cell Sci* 2000, 113:899–906.
8. Long KS, Cedervall T, Walch-Solimena C, Noe DA, Huddleston MJ, Annan RS, Wolin SL. Phosphorylation of the *Saccharomyces cerevisiae* La protein does not appear to be required for its functions in tRNA maturation and nascent RNA stabilization. *RNA* 2001, 7:1589–1602.
9. Bayfield MA, Kaiser TE, Intine RV, Maraia RJ. Conservation of a masked nuclear export activity of La proteins and its effects on tRNA maturation. *Mol Cell Biol* 2007, 27:3303–3312.
10. Intine RV, Dundr M, Misteli T, Maraia RJ. Aberrant nuclear trafficking of La protein leads to disordered processing of associated precursor tRNAs. *Mol Cell* 2002, 9:1113–1123.
11. Fan H, Goodier JL, Chamberlain J, Engelke DR, Maraia RJ. 5′ Processing of tRNA precursors can be modulated by the human La antigen phosphoprotein. *Mol Cell Biol* 1998, 18:3201–3211.
12. Fan H, Sakulich AL, Goodier JL, Zhang X, Qin J, Maraia RJ. Phosphorylation of the human La antigen on serine 366 can regulate recycling of RNA polymerase III transcription complexes. *Cell* 1997, 88:707–715.
13. Intine RV, Dundr M, Vassilev A, Schwartz E, Zhao Y, Depamphilis ML, Maraia RJ. Nonphosphorylated human La antigen interacts with nucleolin at nucleolar sites involved in rRNA biogenesis. *Mol Cell Biol* 2004, 24:10894–10904.
14. Horke S, Reumann K, Schweizer M, Will H, Heise T. Nuclear trafficking of La protein depends on a newly identified NoLS and the ability to bind RNA. *J Biol Chem* 2004, 279:26563–26570.
15. Lahr RM, Fonseca BD, Ciotti GE, Al-Ashtal HA, Jia JJ, Niklaus MR, Blagden SP, Alain T, Berman AJ. La-related protein 1 (LARP1) binds the mRNA cap, blocking eIF4F assembly on TOP mRNAs. *Elife* 2017, 6:e24146.
16. Fonseca BD, Zakaria C, Jia JJ, Graber TE, Svitkin Y, Tahmasebi S, Healy D, Hoang HD, Jensen JM, Diao IT, et al. La-related Protein 1 (LARP1) represses terminal oligopyrimidine (TOP) mRNA translation downstream of mTOR Complex 1 (mTORC1). *J Biol Chem* 2015, 290:15996–16020.
17. Bayfield MA, Yang R, Maraia RJ. Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim Biophys Acta* 1799, 2010:365–378.
18. Merret R, Martino L, Bousquet-Antonelli C, Fneich S, Descombin J, Billey E, Conte MR, Deragon JM. The association of a La module with the PABP-interacting motif PAM2 is a recurrent evolutionary process that led to the neofunctionalization of La-related proteins. *RNA* 2013, 19:36–50.
19. Tcherkezian J, Cargnello M, Romeo Y, Huttlin EL, Lavoie G, Gygi SP, Roux PP. Proteomic analysis of cap-dependent translation identifies LARP1 as a key regulator of 5′TOP mRNA translation. *Genes Dev* 2014, 28:357–371.
20. Mura M, Hopkins TG, Michael T, Abd-Latip N, Weir J, Aboagye E, Mauri F, Jameson C, Sturge J, Gabra H, et al. LARP1 post-transcriptionally regulates mTOR and contributes to cancer progression. *Oncogene* 2015, 34:5025–5036.
21. Keene JD. RNA regulons: coordination of post-transcriptional events. *Nat Rev Genet* 2007, 8:533–543.
22. Blackinton JG, Keene JD. Post-transcriptional RNA regulons affecting cell cycle and proliferation. *Semin Cell Dev Biol* 2014, 34:44–54.
23. Martino L, Pennell S, Kelly G, Busi B, Brown P, Atkinson RA, Salisbury NJ, Ooi ZH, See KW, Smerdon SJ, et al. Synergic interplay of the La motif, RRM1 and the interdomain linker of LARP6 in the recognition of collagen mRNA expands the RNA binding repertoire of the La module. *Nucleic Acids Res* 2015, 43:645–660.
24. Cai L, Fritz D, Stefanovic L, Stefanovic B. Binding of LARP6 to the conserved 5′ stem-loop regulates translation of mRNAs encoding type I collagen. *J Mol Biol* 2009, 395:309–326.
25. Schaffler K, Schulz K, Hirmer A, Wiesner J, Grimm M, Sickmann A, Fischer U. A stimulatory role for the La-related protein 4B in translation. *RNA* 2010, 16:1488–1499.
26. Yang R, Gaidamakov SA, Xie J, Lee J, Martino L, Kozlov G, Crawford AK, Russo AN, Conte MR, Gehring K, et al. LARP4 binds poly(A), interacts with poly(A)-binding protein MLLE domain via a variant PAM2w motif and can promote mRNA stability. *Mol Cell Biol* 2011, 31:542–556.
27. Kuspert M, Murakawa Y, Schaffler K, Vanselow JT, Wolf E, Juranek S, Schlosser A, Landthaler M, Fischer U. LARP4B is an AU-rich sequence associated factor that promotes mRNA accumulation and translation. *RNA* 2015, 21:1294–1305.
28. Stefano JE. Purified lupus antigen La recognizes an oligouridylate stretch common to the 3′ termini of

 © 2017 Wiley Periodicals, Inc.

RNA polymerase III transcripts. *Cell* 1984, 36:145–154.

29. Maraia RJ, Lamichhane TN. 3′ Processing of eukaryotic precursor tRNAs. *WIREs RNA* 2011, 2:362–375.
30. Hussain RH, Zawawi M, Bayfield MA. Conservation of RNA chaperone activity of the human La-related proteins 4, 6 and 7. *Nucleic Acids Res* 2013, 4I:8715–8725.
31. Kucera NJ, Hodsdon ME, Wolin SL. An intrinsically disordered C terminus allows the La protein to assist the biogenesis of diverse noncoding RNA precursors. *Proc Natl Acad Sci USA* 2011, 108:1308–1313.
32. Kuehnert J, Sommer G, Zierk AW, Fedarovich A, Brock A, Fedarovich D, Heise T. Novel RNA chaperone domain of RNA-binding protein La is regulated by AKT phosphorylation. *Nucleic Acids Res* 2015, 43:581–594.
33. Afroz T, Cienikova Z, Clery A, Allain FH. One, two, three, four! How multiple RRMs read the genome sequence. *Methods Enzymol* 2015, 558:235–278.
34. Argyriou AI, Chasapis CT, Apostolidi M, Konstantinidou P, Stathopoulos C, Bentrop D, Spyroulias GA. Backbone and side chain NMR assignment, along with the secondary structure prediction of RRM2 domain of La protein from a lower eukaryote exhibiting identical structural organization with its human homolog. *Biomol NMR Assign* 2015, 9:219–222.
35. Singh M, Choi CP, Feigon J. xRRM: a new class of RRM found in the telomerase La family protein p65. *RNA Biol* 2013, 10:353–359.
36. Singh M, Wang Z, Koo BK, Patel A, Cascio D, Collins K, Feigon J. Structural basis for telomerase RNA recognition and RNP assembly by the holoenzyme La family protein p65. *Mol Cell* 2012, 47:16–26.
37. Intine RV, Tenenbaum SA, Sakulich AS, Keene JD, Maraia RJ. Differential phosphorylation and subcellular localization of La RNPs associated with precursor tRNAs and translation-related mRNAs. *Mol Cell* 2003, 12:1301–1307.
38. Clery A, Blatter M, Allain FH. RNA recognition motifs: boring? Not quite. *Curr Opin Struct Biol* 2008, 18:290–298.
39. Maris C, Dominguez C, Allain FH. The RNA recognition motif, a plastic RNA-binding platform to regulate post-transcriptional gene expression. *FEBS J* 2005, 272:2118–2131.
40. Shao R, Scully SJ Jr, Yan W, Bentley B, Mueller J, Brown C, Bigelow C, Schwartz LM. The novel lupus antigen related protein acheron enhances the development of human breast cancer. *Int J Cancer* 2012, 130:544–554.
41. Valavanis C, Wang Z, Sun D, Vaine M, Schwartz LM. Acheron, a novel member of the Lupus Antigen family, is induced during the programmed cell death of skeletal muscles in the moth *Manduca sexta*. *Gene* 2007, 393:101–109.
42. Stavraka C, Blagden S. The La-related proteins, a family with connections to cancer. *Biomolecules* 2015, 5:2701–2722.
43. Blagden SP, Gatt MK, Archambault V, Lada K, Ichihara K, Lilley KS, Inoue YH, Glover DM. Drosophila Larp associates with poly(A)-binding protein and is required for male fertility and syncytial embryo development. *Dev Biol* 2009, 334:186–197.
44. Kahvejian A, Svitkin YV, Sukarieh R, M'Boutchou MN, Sonenberg N. Mammalian poly(A)-binding protein is a eukaryotic translation initiation factor, which acts via multiple mechanisms. *Genes Dev* 2005, 19:104–113.
45. Tomek W, Wollenhaupt K. The "closed loop model" in controlling mRNA translation during development. *Anim Reprod Sci* 2012, 134:2–8.
46. Thompson MK, Gilbert WV. mRNA length-sensing in eukaryotic translation: reconsidering the "closed loop" and its implications for translational control. *Curr Genet* 2016. https://doi.org/10.1007/s00294-016-0674-3.
47. Alie A, Hayashi T, Sugimura I, Manuel M, Sugano W, Mano A, Satoh N, Agata K, Funayama N. The ancestral gene repertoire of animal stem cells. *Proc Natl Acad Sci USA* 2015, 112: E7093–E7100.
48. Apostolidi M, Vourtsis DJ, Chasapis CT, Stathopoulos C, Bentrop D, Spyroulias GA. (1)H, (1)(5)N, (1)(3)C assignment and secondary structure determination of two domains of La protein from *D. discoideum*. *Biomol NMR Assign* 2014, 8:47–51.
49. Eichinger L, Pachebat JA, Glockner G, Rajandream MA, Sucgang R, Berriman M, Song J, Olsen R, Szafranski K, Xu Q, et al. The genome of the social amoeba *Dictyostelium discoideum*. *Nature* 2005, 435:43–57.
50. Dong G, Chakshusmathi G, Wolin SL, Reinisch KM. Structure of the La motif: a winged helix domain mediates RNA binding via a conserved aromatic patch. *EMBO J* 2004, 23:1000–1007.
51. Drozdetskiy A, Cole C, Procter J, Barton GJ. JPred4: a protein secondary structure prediction server. *Nucleic Acids Res* 2015, 43:W389–W394.
52. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 2011, 7:539.
53. Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2—a multiple

© 2017 Wiley Periodicals, Inc.

sequence alignment editor and analysis workbench. *Bioinformatics* 2009, 25:1189–1191.

54. Jacks A, Kelly G, Curry S, Conte MR. Resonance assignment and secondary structure determination of a C-terminal fragment of the lupus autoantigen (La) protein containing a putative RNA recognition motif (RRM). *J Biomol NMR* 2002, 22:387–388.
55. Fleurdepine S, Deragon JM, Devic M, Guilleminot J, Bousquet-Antonelli C. A bona fide La protein is required for embryogenesis in *Arabidopsis thaliana*. *Nucleic Acids Res* 2007, 35:3306–3321.
56. Topfer F, Gordon T, McCluskey J. Characterization of the mouse autoantigen La (SS-B). *J Immunol* 1993, 150:3091–3100.
57. Intine RVA, Sakulich AL, Koduru SB, Huang Y, Pierstorrf E, Goodier JL, Phan L, Maraia RJ. Control of transfer RNA maturation by phosphorylation of the human La antigen on serine 366. *Mol Cell* 2000, 6:339–348.
58. Thandapani P, O'Connor TR, Bailey TL, Richard S. Defining the RGG/RG motif. *Mol Cell* 2013, 50:613–623.
59. Park JM, Kohn MJ, Bruinsma MW, Vech C, Intine RV, Fuhrmann S, Grinberg A, Mukherjee I, Love PE, Ko MS, et al. The multifunctional RNA-binding protein La is required for mouse development and for the establishment of embryonic stem cells. *Mol Cell Biol* 2006, 26:1445–1451.
60. Hasler D, Lehmann G, Murakawa Y, Klironomos F, Jakob L, Grasser FA, Rajewsky N, Landthaler M, Meister G. The lupus autoantigen La prevents mis-channeling of tRNA fragments into the human microRNA pathway. *Mol Cell* 2016, 63:110–124.
61. Carter R, Drouin G. The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. *Mol Biol Evol* 2010, 27:1035–1043.
62. Drouin G, Carter R. Evolution of eukaryotic RNA polymerases. *eLS* 2010, https://doi.org/10.1002/9780470015902.a0022872.
63. Inokuchi HaY F. Structure and expression of prokaryotic tRNA genes. In: Söll D, RajBhandary U, eds. *tRNA: Structure, Biosynthesis, and Function*. Washington, DC: ASM Press; 1995.
64. Thomm M, Hausner W. Genes for stable RNAs and their expression in archaea. In: Sebald M, ed. *Genetics and Molecular Biology of Anaerobic Bacteria*. New York: Springer Science & Business Media; 2012.
65. Maraia RJ, Rijal K. Structural biology: a transcriptional specialist resolved. *Nature* 2015, 528:204–205.
66. Arimbasseri AG, Maraia RJ. RNA polymerase III advances: structural and tRNA functional views [review]. *Trends Biochem Sci* 2016, 41:546–559.
67. Turowski TW, Lesniewska E, Delan-Forino C, Sayou C, Boguta M, Tollervey D. Global analysis of transcriptionally engaged yeast RNA polymerase III reveals extended tRNA transcripts. *Genome Res* 2016, 26:933–944.
68. Roberts DN, Stewart AJ, Huff JT, Cairns BR. The RNA polymerase III transcriptome revealed by genome-wide localization and activity-occupancy relationships. *Proc Natl Acad Sci USA* 2003, 100:14695–14700.
69. Harismendy O, Gendrel CG, Soularue P, Gidrol X, Sentenac A, Werner M, Lefebvre O. Genome-wide location of yeast RNA polymerase III transcription machinery. *EMBO J* 2003, 22:4738–4747.
70. Wolin SL, Cedervall T. The La protein. *Annu Rev Biochem* 2002, 71:375–403.
71. Arimbasseri AG, Blewett NH, Iben JR, Lamichhane TN., Cherkasova V, Hafner M, Maraia RJ. RNA polymerase III output is functionally linked to tRNA dimethyl-G26 modification. *PLoS Genet* 2015. 11(12):e1005671. https://doi.org/10.1371/journal.pgen.1005671.
72. Huang Y, Intine RV, Mozlin A, Hasson S, Maraia RJ. Mutations in the RNA polymerase III subunit Rpc11p that decrease RNA 3′ cleavage activity increase 3′-terminal oligo(U) length and La-dependent tRNA processing. *Mol Cell Biol* 2005, 25:621–636.
73. Karkusiewicz I, Turowski TW, Graczyk D, Towpik J, Dhungel N, Hopper AK, Boguta M. Maf1, repressor of RNA polymerase III, indirectly affects tRNA processing. *J Biol Chem* 2011, 286:39478–39488.
74. Hoffmann N, Jakobi A, Moreno-Morcillo M, Glatt S, Kosinski J, Hagen W, Sachse C, Muller C. Molecular structures of unbound and transcribing RNA polymerase III. *Nature* 2015, 528:231–236.
75. Arimbasseri AG, Rijal K, Maraia RJ. Comparative overview of RNA polymerase II and III transcription cycles, with focus on RNA polymerase III termination and reinitiation. *Transcription* 2014, 5:e27639. https://doi.org/10.4161/trns.27369.
76. Landrieux E, Alic N, Ducrot C, Acker J, Riva M, Carles C. A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. *EMBO J* 2006, 25:118–128.
77. Arimbasseri AG, Rijal K, Maraia RJ. Transcription termination by the eukaryotic RNA polymerase III. *Biochim Biophys Acta* 2013, 1829:318–330.
78. Moir RD, Willis IM. Regulation of pol III transcription by nutrient and stress signaling pathways. *Biochim Biophys Acta* 1829, 2013:361–375.
79. Chakshusmathi G, Kim SD, Rubinson DA, Wolin SL. A La protein requirement for efficient pre-tRNA folding. *EMBO J* 2003, 22:6562–6572.

© 2017 Wiley Periodicals, Inc.

80. Yoo CJ, Wolin SL. The yeast La protein is required for the 3′ endonucleolytic cleavage that matures tRNA precursors. *Cell* 1997, 89:393–402.
81. Huang Y, Bayfield MA, Intine RV, Maraia RJ. Separate RNA-binding surfaces on the multifunctional La protein mediate distinguishable activities in tRNA maturation. *Nat Struct Mol Biol* 2006, 13:611–618.
82. Copela LA, Fernandez CF, Sherrer RL, Wolin SL. Competition between the Rex1 exonuclease and the La protein affects both Trf4p-mediated RNA quality control and pre-tRNA maturation. *RNA* 2008, 14:1214–1227.
83. Wolin SL, Wurtmann EJ. Molecular chaperones and quality control in noncoding RNA biogenesis. *Cold Spring Harb Symp Quant Biol* 2006, 71:505–511.
84. Xue D, Rubinson DA, Pannone BK, Yoo CJ, Wolin SL. U snRNP assembly in yeast involves the La protein. *EMBO J* 2000, 19:1650–1660.
85. Pannone B, Xue D, Wolin SL. 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* 1998, 17:7442–7453.
86. Leung E, Schneider C, Yan F, Mohi-El-Din H, Kudla G, Tuck A, Wlotzka W, Doronina VA, Bartley R, Watkins NJ, et al. Integrity of SRP RNA is ensured by La and the nuclear RNA quality control machinery. *Nucleic Acids Res* 2014, 42:10698–10710.
87. Bayfield MA, Maraia RJ. Precursor-product discrimination by La protein during tRNA metabolism. *Nat Struct Mol Biol* 2009, 16:430–437.
88. Jarosz DF, Taipale M, Lindquist S. Protein homeostasis and the phenotypic manifestation of genetic diversity: principles and mechanisms. *Annu Rev Genet* 2010, 44:189–216.
89. Jarosz DF, Lindquist S. Hsp90 and environmental stress transform the adaptive value of natural genetic variation. *Science* 2010, 330:1820–1824.
90. Lindquist S. Protein folding sculpting evolutionary change. *Cold Spring Harb Symp Quant Biol* 2009, 74:103–108.
91. Rutherford SL, Lindquist S. Hsp90 as a capacitor for morphological evolution. *Nature* 1998, 396:336–342.
92. Anderson J, Phan L, Cuesta R, Carlson BA, Pak M, Asano K, Bjork GR, Tamame M, Hinnebusch AG. The essential Gcd10p-Gcd14p nuclear complex is required for 1-methyladenosine modification and maturation of initiator methionyl-tRNA. *Genes Dev* 1998, 12:3650–3662.
93. Calvo O, Cuesta R, Anderson J, Gutierrez N, Garcia-Barrio MT, Hinnebusch AG, Tamame M. GCD14p, a repressor of GCN4 translation, cooperates with Gcd10p and Lhp1p in the maturation of initiator methionyl-tRNA in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1999, 19:4167–4181.
94. Johansson MJ, Bystrom AS. Dual function of the tRNA(m(5)U54)methyltransferase in tRNA maturation. *RNA* 2002, 8:324–335.
95. Plasmodb. 2017. Available at: http://plasmodb.org/plasmo/showApplication.do. (Accessed Oct. 19, 2016).
96. Chan PP, Lowe TM. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res* 2015, 44(D1): D184–D189.
97. Glanzmann B, Moller M, le Roex N, Tromp G, Hoal EG, van Helden PD. The complete genome sequence of the African buffalo (*Syncerus caffer*) *BMC Genomics* 2016, 17:1001.
98. Iben JR, Maraia RJ. tRNA gene copy number variation in humans. *Gene* 2014, 536:376–384.
99. Iben JR, Maraia RJ. Yeast tRNAomics: tRNA gene copy number variation and codon use provide bioinformatics evidence of a new wobble pair in a eukaryote. *RNA* 2012, 18:1358–1372.
100. Iben JR, Epstein JA, Bayfield MA, Bruinsma MW, Hasson S, Bacikova D, Ahmad D, Rockwell D, Kittler EL, Zapp ML, et al. Comparative whole genome sequencing reveals phenotypic tRNA gene duplication in spontaneous *Schizosaccharomyces pombe* La mutants. *Nucleic Acids Res* 2011, 39:4728–4742.
101. Parisien M, Wang X, Pan T. Diversity of human tRNA genes from the 1000-genomes project. *RNA Biol* 2013, 10:1853–1867.
102. Goodenbour JM, Pan T. Diversity of tRNA genes in eukaryotes. *Nucleic Acids Res* 2006, 34:6137–6146.
103. Nguyen VC, Clelland BW, Hockman DJ, Kujat-Choy SL, Mewhort HE, Schultz MC. Replication stress checkpoint signaling controls tRNA gene transcription. *Nat Struct Mol Biol* 2010, 17:976–981.
104. Clelland BW, Schultz MC. Genome stability control by checkpoint regulation of tRNA gene transcription. *Transcription* 2010, 1:115–125.
105. Tran PLT, Pohl TJ, Chen CF, Chan A, Pott S, Zakian VA. PIF1 family DNA helicases suppress R-loop mediated genome instability at tRNA genes. *Nat Commun* 2017, 8:15025.
106. Copela LA, Chakshusmathi G, Sherrer RL, Wolin SL. The La protein functions redundantly with tRNA modification enzymes to ensure tRNA structural stability. *RNA* 2006, 12:644–654.
107. Kadaba S, Wang X, Anderson JT. Nuclear RNA surveillance in *Saccharomyces cerevisiae*: Trf4p-dependent polyadenylation of nascent hypomethylated tRNA and an aberrant form of 5S rRNA. *RNA* 2006, 12:508–521.
108. Kadaba S, Krueger A, Trice T, Krecic AM, Hinnebusch AG, Anderson J. Nuclear surveillance and degradation of hypomodified initiator tRNAMet in *S. cerevisiae*. *Genes Dev* 2004, 18:1227–1240.

© 2017 Wiley Periodicals, Inc.

109. Qiu H, Hu C, Anderson J, Björk G, Sarkar S, Hopper A, Hinnebusch AG. Defects in tRNA processing and nuclear export induce GCN4 translation independently of phosphorylation of the alpha subunit of eIF2. *Mol Cell Biol* 2000, 20:2505–2516.
110. Maraia RJ, Arimbasseri AG. Factors that shape eukaryotic tRNAomes: processing, modification and anticodon–codon use. *Biomolecules* 2017, 7(1), 26; https://doi.org/10.3390/biom7010026.
111. Maraia RJ, Iben JR. Different types of secondary information in the genetic code. *RNA* 2014, 20:977–984.
112. Martino L, Pennell S, Kelly G, Bui TTT, Kotik-Kogan O, Smerdon SJ, Drake AF, Curry S, Conte MR. Analysis of the interaction with the hepatitis C virus mRNA reveals an alternative mode of RNA recognition by the human La protein. *Nucleic Acids Res* 2012, 40:1381–1394.
113. Alfano C, Sanfelice D, Babon J, Kelly G, Jacks A, Curry S, Conte MR. Structural analysis of cooperative RNA binding by the La motif and central RRM domain of human La protein. *Nat Struct Mol Biol* 2004, 11:323–329.
114. Teplova M, Yuan Y-R, Ilin S, Malinina L, Phan AT, Teplov A, Patel DJ. Structural basis for recognition and sequestration of UUU-OH 3′-termini of nascent RNA pol III transcripts by La, a rheumatic disease autoantigen. *Mol Cell* 2006, 21:75–85.
115. Curry S, Kotik-Kogan O, Conte MR, Brick P. Getting to the end of RNA: structural analysis of protein recognition of 5′ and 3′ termini. *Biochim Biophys Acta* 1789, 2009:653–666.
116. Uchikawa E, Natchiar KS, Han X, Proux F, Roblin P, Zhang E, Durand A, Klaholz BP, Dock-Bregeon AC. Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. *Nucleic Acids Res* 2015, 43:3373–3388.
117. Ozanick SG, Wang X, Costanzo M, Brost RL, Boone C, Anderson JT. Rex1p deficiency leads to accumulation of precursor initiator tRNAMet and polyadenylation of substrate RNAs in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 2009, 37:298–308.
118. Blewett NH, Iben JR, Gaidamakov S, Maraia RJ. La deletion from mouse brain alters pre-tRNA metabolism and accumulation of pre-5.8S rRNA, with neuron death and reactive astrocytosis. *Mol Cell Biol* 2017. 37(10). pii: e00588-16. https://doi.org/10.1128/MCB.00588-16.
119. Aoki K, Adachi S, Homoto M, Kusano H, Koike K, Natsume T. LARP1 specifically recognizes the 3′ terminus of poly(A) mRNA. *FEBS Lett* 2013, 587:2173–2178.
120. Mattijssen S, Iben JR, Arimbasseri AG, G S, Lee J, Hafner M, Maraia RJ. Codon-tRNA fit can impact LARP4 activity for mRNA poly(A) tail lengthening. Submitted.
121. Maraia RJ, Bayfield MA. The La protein-RNA complex surfaces [review]. *Mol Cell* 2006, 21:149–152.
122. Curry S, Conte MR. A terminal affair: 3′-end recognition by the human La protein. *Trends Biochem Sci* 2006, 31:303–305.
123. He N, Jahchan NS, Hong E, Li Q, Bayfield MA, Maraia RJ, Luo K, Zhou Q. A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol Cell* 2008, 29:588–599.
124. Sobel SG, Wolin SL. Two yeast La motif-containing proteins are RNA-binding proteins that associate with polyribosomes. *Mol Biol Cell* 1999, 10:3849–3862.
125. Yu W, Farrell RA, Stillman DJ, Winge DR. Identification of SLF1 as a new copper homeostasis gene involved in copper sulfide mineralization in *Saccharomyces cerevisiae*. *Mol Cell Biol* 1996, 16:2464–2472.
126. Schenk L, Meinel DM, Strasser K, Gerber AP. La-motif-dependent mRNA association with Slf1 promotes copper detoxification in yeast. *RNA* 2012, 18:449–461.
127. Chatenay-Lapointe M, Shadel GS. Repression of mitochondrial translation, respiration and a metabolic cycle-regulated gene, SLF1, by the yeast Pumilio-family protein Puf3p. *PLoS One* 2011, 6: e20441.
128. Kershaw CJ, Costello JL, Castelli LM, Talavera D, Rowe W, Sims PF, Ashe MP, Hubbard SJ, Pavitt GD, Grant CM. The yeast La related protein Slf1p is a key activator of translation during the oxidative stress response. *PLoS Genet* 2015, 11:e1004903.
129. Zhang C, Wang X, Park S, Chiang YC, Xi W, Laue TM, Denis CL. Only a subset of the PAB1-mRNP proteome is present in mRNA translation complexes. *Protein Sci* 2014, 23:1036–1049.
130. Eichhorn CD, Chug R, Feigon J. hLARP7 C-terminal domain contains an xRRM that binds the 3′ hairpin of 7SK RNA. *Nucleic Acids Res* 2016, 44:9977–9989.
131. Sanfelice D, Kelly G, Curry S, Conte MR. NMR assignment of the N-terminal region of human La free and in complex with RNA. *Biomol NMR Assign* 2008, 2:107–109.
132. Liang C, Xiong K, Szulwach KE, Zhang Y, Wang Z, Peng J, Fu M, Jin P, Suzuki HI, Liu Q. Sjogren syndrome antigen B (SSB)/La promotes global microRNA expression by binding microRNA precursors through stem-loop recognition. *J Biol Chem* 2013, 288:723–736.
133. Horke S, Reumann K, Rang A, Heise T. Molecular characterization of the human La protein-hepatitis B virus RNA.B interaction *in vitro*. *J Biol Chem* 2002, 277:34949–34958.

© 2017 Wiley Periodicals, Inc.

134. Sommer G, Dittmann J, Kuehnert J, Reumann K, Schwartz PE, Will H, Coulter BL, Smith MT, Heise T. The RNA-binding protein La contributes to cell proliferation and CCND1 expression. *Oncogene* 2011, 30:434–444.

135. Brown KA, Sharifi S, Hussain R, Donaldson L, Bayfield MA, Wilson DJ. Distinct dynamic modes enable the engagement of dissimilar ligands in a promiscuous atypical RNA recognition motif. *Biochemistry* 2016, 55:7141–7150.

136. Bhattacharya R, Perumal K, Sinha K, Maraia R, Reddy R. Methylphosphate cap structure in small RNAs reduces the affinity of RNAs to La protein. *Gene Expr* 2002, 10:243–253.

137. Lahr RM, Mack SM, Heroux A, Blagden SP, Bousquet-Antonelli C, Deragon JM, Berman AJ. The La-related protein 1-specific domain repurposes HEAT-like repeats to directly bind a 5′TOP sequence. *Nucleic Acids Res* 2015, 43:8077–8088.

138. Liu Y, Tan H, Tian H, Liang C, Chen S, Liu Q. Autoantigen La promotes efficient RNAi, antiviral response, and transposon silencing by facilitating multiple-turnover RISC catalysis. *Mol Cell* 2011, 44:502–508.

139. Drinnenberg IA, Weinberg DE, Xie KT, Mower JP, Wolfe KH, Fink GR, Bartel DP. RNAi in budding yeast. *Science* 2009, 326:544–550.

140. Phizicky EM, Hopper AK. tRNA biology charges to the front. *Genes Dev* 2010, 24:1832–1860.

141. Peterlin BM, Price DH. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* 2006, 23:297–305.

142. Guo J, Li T, Price DH. Runaway transcription. *Genome Biol* 2013, 14:133.

143. Guo J, Price DH. RNA polymerase II transcription elongation control. *Chem Rev* 2013, 113:8583–8603.

144. Muniz L, Egloff S, Kiss T. RNA elements directing *in vivo* assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. *Nucleic Acids Res* 2013, 41:4686–4698.

145. Krueger BJ, Jeronimo C, Roy BB, Bouchard A, Barrandon C, Byers SA, Searcey CE, Cooper JJ, Bensaude O, Cohen EA, et al. LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res* 2008, 36:2219–2229.

146. Jeronimo C, Forget D, Bouchard A, Li Q, Chua G, Poitras C, Therien C, Bergeron D, Bourassa S, Greenblatt J, et al. Systematic analysis of the protein interaction network for the human transcription machinery reveals the identity of the 7SK capping enzyme. *Mol Cell* 2007, 27:262–274.

147. Jiang J, Miracco EJ, Hong K, Eckert B, Chan H, Cash DD, Min B, Zhou ZH, Collins K, Feigon J. The architecture of tetrahymena telomerase holoenzyme. *Nature* 2013, 496:187–192.

148. Aigner S, Cech TR. The Euplotes telomerase subunit p43 stimulates enzymatic activity and processivity *in vitro*. *RNA* 2004, 10:1108–1118.

149. Stone MD, Mihalusova M, O'Connor CM, Prathapam R, Collins K, Zhuang X. Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature* 2007, 446:458–461.

150. Shumyatsky G, Shimba S, Reddy R. Capping signals correspond to the 5′ end in four eukaryotic small RNAs containing gamma-monomethylphosphate cap structure. *Gene Expr* 1994, 4:29–41.

151. Martinez-Zapien D, Saliou JM, Han X, Atmanene C, Proux F, Cianferani S, Dock-Bregeon AC. Intermolecular recognition of the non-coding RNA 7SK and HEXIM protein in perspective. *Biochimie* 2015, 117:63–71.

152. Egloff S, Van Herreweghe E, Kiss T. Regulation of polymerase II transcription by 7SK snRNA: two distinct RNA elements direct P-TEFb and HEXIM1 binding. *Mol Cell Biol* 2006, 26:630–642.

153. Fujinaga K, Luo Z, Peterlin BM. Genetic analysis of the structure and function of 7SK small nuclear ribonucleoprotein (snRNP) in cells. *J Biol Chem* 2014, 289:21181–21190.

154. Brogie JE, Price DH. Reconstitution of a functional 7SK snRNP. *Nucleic Acids Res* 2017, https://doi.org/10.1093/nar/gkx262. [Epub ahead of print].

155. Biewenga P, Buist MR, Moerland PD, Ver Loren van Themaat E, van Kampen AH, ten Kate FJ, Baas F. Gene expression in early stage cervical cancer. *Gynecol Oncol* 2008, 108:520–526.

156. Cheng Y, Jin Z, Agarwal R, Ma K, Yang J, Ibrahim S, Olaru AV, David S, Ashktorab H, Smoot DT, et al. LARP7 is a potential tumor suppressor gene in gastric cancer. *Lab Invest* 2012, 92:1013–1019.

157. Ji X, Lu H, Zhou Q, Luo K. LARP7 suppresses P-TEFb activity to inhibit breast cancer progression and metastasis. *Elife* 2014, 3:e02907.

158. Markert A, Grimm M, Martinez J, Wiesner J, Meyerhans A, Meyuhas O, Sickmann A, Fischer U. The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep* 2008, 9:569–575.

159. Dai Q, Luan G, Deng L, Lei T, Kang H, Song X, Zhang Y, Xiao ZX, Li Q. Primordial dwarfism gene maintains Lin28 expression to safeguard embryonic stem cells from premature differentiation. *Cell Rep* 2014, 7:735–746.

160. Najmabadi H, Hu H, Garshasbi M, Zemojtel T, Abedini SS, Chen W, Hosseini M, Behjati F, Haas S,

© 2017 Wiley Periodicals, Inc.

Jamali P, et al. Deep sequencing reveals 50 novel genes for recessive cognitive disorders. *Nature* 2011, 478:57–63.

161. Alazami AM, Al-Owain M, Alzahrani F, Shuaib T, Al-Shamrani H, Al-Falki YH, Al-Qahtani SM, Alsheddi T, Colak D, Alkuraya FS. Loss of function mutation in LARP7, chaperone of 7SK ncRNA, causes a syndrome of facial dysmorphism, intellectual disability, and primordial dwarfism. *Hum Mutat* 2012, 33:1429–1434.
162. Hollink IH, Alfadhel M, Al-Wakeel AS, Ababneh F, Pfundt R, de Man SA, Jamra RA, Rolfs A, Bertoli-Avella AM, van de Laar IM. Broadening the phenotypic spectrum of pathogenic LARP7 variants: two cases with intellectual disability, variable growth retardation and distinct facial features. *J Hum Genet* 2016, 61:229–233.
163. Ling TT, Sorrentino S. Compound heterozygous variants in the LARP7 gene as a cause of Alazami syndrome in a Caucasian female with significant failure to thrive, short stature, and developmental disability. *Am J Med Genet A* 2016, 170A:217–219.
164. Slomnicki LP, Malinowska A, Kistowski M, Palusinski A, Zheng JJ, Sepp M, Timmusk T, Dadlez M, Hetman M. Nucleolar enrichment of brain proteins with critical roles in human neurodevelopment. *Mol Cell Proteomics* 2016, 15:2055–2075.
165. Egloff S, Vitali P, Tellier M, Raffel R, Murphy S, Kiss T. The 7SK snRNP associates with the little elongation complex to promote snRNA gene expression. *EMBO J* 2017, 36:934–948.
166. Sonenberg N, Hinnebusch AG. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell* 2009, 136:731–745.
167. Hinnebusch AG, Lorsch JR. The mechanism of eukaryotic translation initiation: new insights and challenges. *Cold Spring Harb Perspect Biol* 2012, 4: pii: a011544. https://doi.org/10.1101/cshperspect.a011544.
168. Merrick WC. eIF4F: a retrospective. *J Biol Chem* 2015, 290:24091–24099.
169. Safaee N, Kozlov G, Noronha AM, Xie J, Wilds CJ, Gehring K. Interdomain allostery promotes assembly of the poly(A) mRNA complex with PABP and eIF4G. *Mol Cell* 2012, 48:375–386.
170. Eliseeva IA, Lyabin DN, Ovchinnikov LP. Poly(A)-binding proteins: structure, domain organization, and activity regulation. *Biochemistry (Mosc)* 2013, 78:1377–1391.
171. Xie J, Kozlov G, Gehring K. The "tale" of poly(A) binding protein: the MLLE domain and PAM2-containing proteins. *Biochim Biophys Acta* 1839, 2014:1062–1068.
172. Borman AM, Michel YM, Kean KM. Biochemical characterisation of cap-poly(A) synergy in rabbit reticulocyte lysates: the eIF4G-PABP interaction increases the functional affinity of eIF4E for the capped mRNA 5′-end. *Nucleic Acids Res* 2000, 28:4068–4075.
173. von der Haar T, Gross JD, Wagner G, McCarthy JE. The mRNA cap-binding protein eIF4E in post-transcriptional gene expression. *Nat Struct Mol Biol* 2004, 11:503–511.
174. Tarun SZ Jr, Sachs AB. Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G. *EMBO J* 1996, 15:7168–7177.
175. Wei CC, Balasta ML, Ren J, Goss DJ. Wheat germ poly(A) binding protein enhances the binding affinity of eukaryotic initiation factor 4F and (iso)4F for cap analogues. *Biochemistry* 1998, 37:1910–1916.
176. Tarun SZ Jr, Wells SE, Deardorff JA, Sachs AB. Translation initiation factor eIF4G mediates *in vitro* poly(A) tail-dependent translation. *Proc Natl Acad Sci USA* 1997, 94:9046–9051.
177. Wells SE, Hillner PE, Vale RD, Sachs AB. Circularization of mRNA by eukaryotic translation initiation factors. *Mol Cell* 1998, 2:135–140.
178. Amrani N, Ghosh S, Mangus DA, Jacobson A. Translation factors promote the formation of two states of the closed-loop mRNP. *Nature* 2008, 453:1276–1280.
179. Chen CY, Shyu AB. Mechanisms of deadenylation-dependent decay. *WIREs RNA* 2011, 2:167–183.
180. Mangus DA, Evans MC, Jacobson A. Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol* 2003, 4:223.
181. Burrows C, Latip NA, Lam SJ, Carpenter L, Sawicka K, Tzolovsky G, Gabra H, Bushell M, Glover DM, Willis AE, et al. The RNA binding protein Larp1 regulates cell division, apoptosis and cell migration. *Nucleic Acids Res* 2010, 38:5542–5553.
182. Deo RC, Bonanno JB, Sonenberg N, Burley SK. Recognition of polyadenylate RNA by the poly(A)-binding protein. *Cell* 1999, 98:835–845.
183. Baer BW, Kornberg RD. Repeating structure of cytoplasmic poly(A)-ribonucleoprotein. *Proc Natl Acad Sci USA* 1980, 77:1890–1892.
184. Roy G, De Crescenzo G, Khaleghpour K, Kahvejian A, O'Connor-McCourt M, Sonenberg N. Paip1 interacts with poly(A) binding protein through two independent binding motifs. *Mol Cell Biol* 2002, 22:3769–3782.
185. Albrecht M, Lengauer T. Survey on the PABC recognition motif PAM2. *Biochem Biophys Res Commun* 2004, 316:129–138.
186. Hinnebusch AG. Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol Mol Biol Rev* 2011, 75:434–467, first page of table of contents..

 © 2017 Wiley Periodicals, Inc.

187. Meyuhas O, Kahan T. The race to decipher the top secrets of TOP mRNAs. *Biochim Biophys Acta* 1849, 2015:801–811.

188. Laplante M, Sabatini DM. mTOR signaling in growth control and disease. *Cell* 2012, 149:274–293.

189. Pause A, Belsham GJ, Gingras AC, Donze O, Lin TA, Lawrence JC Jr, Sonenberg N. Insulin-dependent stimulation of protein synthesis by phosphorylation of a regulator of 5′-cap function. *Nature* 1994, 371:762–767.

190. Mamane Y, Petroulakis E, LeBacquer O, Sonenberg N. mTOR, translation initiation and cancer. *Oncogene* 2006, 25:6416–6422.

191. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. *Cell* 2010, 141:129–141.

192. Rinke J, Steitz JA. Precursor molecules of both human 5S ribosomal RNA and transfer RNAs are bound by a cellular protein reactive with anti-La lupus antibodies. *Cell* 1982, 29:149–159.

193. Lerner MR, Boyle JA, Hardin JA, Steitz JA. Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus. *Science* 1981, 211:400–402.

194. Lerner MR, Andrews NC, Miller G, Steitz JA. Two small RNAs encoded by Epstein-Barr virus and complexed with protein are precipitated by antibodies from patients with systemic lupus erythematosus. *Proc Natl Acad Sci USA* 1981, 78:805–809.

195. Kurilla MG, Keene JD. The leader RNA of vesicular stomatitis virus is bound by a cellular protein reactive with anti-La lupus antibodies. *Cell* 1983, 34:837–845.

196. Gottlieb E, Steitz JA. The RNA binding protein La influences both the accuracy and the efficiency of RNA polymerase III transcription *in vitro*. *EMBO J* 1989, 8:841–850.

197. Broekhuis CH, Neubauer G, van der Heijden A, Mann M, Proud CG, van Venrooij WJ, Pruijn GJ. Detailed analysis of the phosphorylation of human La (SS-B) autoantigen. (De)phosphorylation does not affect subcellular distribution. *Biochemistry* 2000, 39:3023–3033.

198. Moritz A, Li Y, Guo A, Villen J, Wang Y, MacNeill J, Kornhauser J, Sprott K, Zhou J, Possemato A, et al. Akt-RSK-S6 kinase signaling networks activated by oncogenic receptor tyrosine kinases. *Sci Signal* 2010, 3:ra64.

199. Brenet F, Socci N, Sonenberg N, Holland E. Akt phosphorylation of La regulates specific mRNA translation in glial progenitors. *Oncogene* 2009, 28:128–139.

200. Tang J, Zhang ZH, Huang M, Heise T, Zhang J, Liu GL. Phosphorylation of human La protein at Ser 366 by casein kinase II contributes to hepatitis B virus replication and expression *in vitro*. *J Viral Hepat* 2013, 20:24–33.

201. Schwartz E, Intine RV, Maraia RJ. CK2 is responsible for phosphorylation of human La protein serine-366 and can modulate 5′TOP mRNA metabolism. *Mol Cell Biol* 2004, 24:9580–9591.

202. Fairley JA, Kantidakis T, Kenneth NS, Intine RV, Maraia RJ, White RJ. Human La is found at RNA polymerase III-transcribed genes *in vivo*. *Proc Natl Acad Sci USA* 2005, 102:18350–18355.

203. Kota V, Sommer G, Durette C, Thibault P, van Niekerk EA, Twiss JL, Heise T. SUMO—modification of the La protein facilitates binding to mRNA *in vitro* and in cells. *PLoS One* 2016, 11: e0156365.

204. van Niekerk EA, Willis DE, Chang JH, Reumann K, Heise T, Twiss JL. Sumoylation in axons triggers retrograde transport of the RNA-binding protein La. *Proc Natl Acad Sci USA* 2007, 104:12913–12918.

205. Park JM, Intine RV, Maraia RJ. Mouse and human La proteins differ in kinase substrate activity and activation mechanism for tRNA processing. *Gene Expr* 2007, 14:71–81.

206. Wilusz J, Kurilla MG, Keene JD. A host protein (La) binds to a unique species of minus-sense leader RNA during replication of vesicular stomatitis virus. *Proc Natl Acad Sci USA* 1983, 80:5827–5831.

207. Kurilla MG, Cabradilla CD, Holloway BP, Keene JD. Nucleotide sequence and host La protein interactions of rabies virus leader RNA. *J Virol* 1984, 50:773–778.

208. De BP, Gupta S, Zhao H, Drazba JA, Banerjee AK. Specific interaction *in vitro* and *in vivo* of glyceraldehyde-3-phosphate dehydrogenase and LA protein with cis-acting RNAs of human parainfluenza virus type 3. *J Biol Chem* 1996, 271:24728–24735.

209. Raha T, Pudi R, Das S, Shaila MS. Leader RNA of Rinderpest virus binds specifically with cellular La protein: a possible role in virus replication. *Virus Res* 2004, 104:101–109.

210. Weber M, Weber F. RIG-I-like receptors and negative-strand RNA viruses: RLRly bird catches some worms. *Cytokine Growth Factor Rev* 2014, 25:621–628.

211. Schlee M. Master sensors of pathogenic RNA–RIG-I like receptors. *Immunobiology* 2013, 218:1322–1335.

212. Plumet S, Herschke F, Bourhis JM, Valentin H, Longhi S, Gerlier D. Cytosolic 5′-triphosphate ended viral leader transcript of measles virus as activator of the RIG I-mediated interferon response. *PLoS One* 2007, 2:e279.

© 2017 Wiley Periodicals, Inc.

213. Lloyd RE. Nuclear proteins hijacked by mammalian cytoplasmic plus strand RNA viruses. *Virology* 2015, 479-480:457–474.

214. McLellan JS, Chen M, Leung S, Graepel KW, Du X, Yang Y, Zhou T, Baxa U, Yasuda E, Beaumont T, et al. Structure of RSV fusion glycoprotein trimer bound to a prefusion-specific neutralizing antibody. *Science* 2013, 340:1113–1117.

215. Bitko V, Musiyenko A, Bayfield MA, Maraia RJ, Barik S. Cellular La protein shields nonsegmented negative-strand RNA viral leader RNA from RIG-I and enhances virus growth by diverse mechanisms. *J Virol* 2008, 82:7977–7987.

216. Meerovitch K, Pelletier J, Sonenberg N. A cellular protein that binds to the 5′-noncoding region of poliovirus RNA: implications for internal translation initiation. *Genes Dev* 1989, 67:3798–3807.

217. Meerovitch K, Svitkin YV, Lee HS, Lejbkowicz F, Kenan DJ, Chan EK, Agol VI, Keene JD, Sonenberg N. La autoantigen enhances and corrects aberrant translation of poliovirus RNA in reticulocyte lysate. *J Virol* 1993, 67:3798–3807.

218. Chang Y-N, Kenan DJ, Keene JD, Gatignol A, Jeang K-T. Direct interactions between autoantigen La and human immunodeficiency virus leader RNA. *J Virol* 1994, 68:7008–7020.

219. Svitkin YV, Meerovitch K, Lee HS, Dholakia JN, Kenan DJ, Agol VI, Sonenberg N. Internal translation initiation on poliovirus RNA: further characterization of La function in poliovirus translation *in vitro*. *J Virol* 1994, 68:1544–1550.

220. Svitkin YV, Pause A, Sonenberg N. La autoantigen alleviates translational repression by the 5′ leader sequence of the human immunodeficiency virus type 1 mRNA. *J Virol* 1994, 68:7001–7007.

221. Ali N, Siddiqui A. The La antigen binds 5′ noncoding region of the hepatitis C virus RNA in the context of the initiator AUG codon and stimulates internal ribosome entry site-mediated translation. *Proc Natl Acad Sci USA* 1997, 94:2249–2254.

222. Ali N, Pruijn GJ, Kenan DJ, Keene JD, Siddiqui A. Human La antigen is required for the hepatitis C virus internal ribosome entry site (IRES)-mediated translation. *J Biol Chem* 2000, 275:27531–27540.

223. Pelletier J, Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 1988, 334:320–325.

224. Walsh D, Mohr I. Viral subversion of the host protein synthesis machinery. *Nat Rev Microbiol* 2011, 9:860–875.

225. Costa-Mattioli M, Svitkin Y, Sonenberg N. 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* 2004, 24:6861–6870.

226. Macejak DG, Sarnow P. Internal initiation of translation mediated by the 5′ leader of a cellular mRNA [see comments]. *Nature* 1991, 353:90–94.

227. Heise T, Kota V, Brock A, Morris AB, Rodriguez RM, Zierk AW, Howe PH, Sommer G. The La protein counteracts cisplatin-induced cell death by stimulating protein synthesis of anti-apoptotic factor Bcl2. *Oncotarget* 2016, 7:29664–29676.

228. Shiroki K, Isoyama T, Kuge S, Ishii T, Ohmi S, Hata S, Suzuki K, Takasaki Y, Nomoto A. Intracellular redistribution of truncated La protein produced by poliovirus 3Cpro-mediated cleavage. *J Virol* 1999, 73:2193–2200.

229. Kuyumcu-Martinez NM, Van Eden ME, Younan P, Lloyd RE. Cleavage of poly(A)-binding protein by poliovirus 3C protease inhibits host cell translation: a novel mechanism for host translation shutoff. *Mol Cell Biol* 2004, 24:1779–1790.

230. Svitkin YV, Ovchinnikov LP, Dreyfuss G, Sonenberg N. General RNA binding proteins render translation cap dependent. *EMBO J* 1996, 15:7147–7155.

231. Svitkin YV, Evdokimova VM, Brasey A, Pestova TV, Fantus D, Yanagiya A, Imataka H, Skabkin MA, Ovchinnikov LP, Merrick WC, et al. General RNA-binding proteins have a function in poly(A)-binding protein-dependent translation. *EMBO J* 2009, 28:58–68.

232. Shwetha S, Kumar A, Mullick R, Vasudevan D, Mukherjee N, Das S. HuR displaces polypyrimidine tract binding protein to facilitate La binding to the 3′ untranslated region and enhances hepatitis C virus replication. *J Virol* 2015, 89:11356–11371.

233. Manna AK, Kumar A, Ray U, Das S, Basu G, Roy S. A cyclic peptide mimic of an RNA recognition motif of human La protein is a potent inhibitor of hepatitis C virus. *Antiviral Res* 2013, 97:223–226.

234. Kumar A, Ray U, Das S. Human La protein interaction with GCAC near the initiator AUG enhances hepatitis C virus RNA replication by promoting linkage between 5′ and 3′ untranslated regions. *J Virol* 2013, 87:6713–6726.

235. Pudi R, Srinivasan P, Das S. La protein binding at the GCAC site near the initiator AUG facilitates the ribosomal assembly on the hepatitis C virus RNA to influence internal ribosome entry site-mediated translation. *J Biol Chem* 2004, 279:29879–29888.

236. Pudi R, Abhiman S, Srinivasan N, Das S. Hepatitis C virus internal ribosome entry site-mediated translation is stimulated by specific interaction of independent regions of human La autoantigen. *J Biol Chem* 2003, 278:12231–12240.

237. Das S, Ott M, Yamane A, Venkatesan A, Gupta S, Dasgupta A. Inhibition of internal entry site (IRES)-mediated translation by a small yeast RNA: a novel

 © 2017 Wiley Periodicals, Inc.

strategy to block hepatitis C virus protein synthesis. *Front Biosci* 1998, 3:D1241–D1252.

238. Romero V, Fellows E, Jenne DE, Andrade F. Cleavage of La protein by granzyme H induces cytoplasmic translocation and interferes with La-mediated HCV-IRES translational activity. *Cell Death Differ* 2009, 16:340–348.
239. Shirasaki T, Honda M, Mizuno H, Shimakami T, Okada H, Sakai Y, Murakami S, Wakita T, Kaneko S. La protein required for internal ribosome entry site-directed translation is a potential therapeutic target for hepatitis C virus replication. *J Infect Dis* 2010, 202:75–85.
240. Heise T, Guidotti LG, Chisari FV. La autoantigen specifically recognizes a predicted stem-loop in hepatitis B virus RNA. *J Virol* 1999, 73:5767–5776.
241. Pellizzoni L, Cardinali B, Lin-Marq N, Mercanti D, Pierandrei-Amaldi P. A Xenopus laevis homologue of the La autoantigen binds the pyrimidine tract of the 5′ UTR of ribosomal protein mRNAs *in vitro*: implication of a protein factor in complex formation. *J Mol Biol* 1996, 259:904–915.
242. Meyuhas O, Hornstein E. Translational control of TOP mRNAs. In: Sonenberg N, Hershey JWB, Mathews M, eds. *Translational Control of Gene Expression*: Cold Spring Harbor, NY. Cold Spring Harbor Press; 2000, 671–693.
243. Crosio C, Boyl PP, Loreni F, Pierandrei-Amaldi P, Amaldi F. La protein has a positive effect on the translation of TOP mRNAs *in vivo*. *Nucleic Acids Res* 2000, 28:2927–2934.
244. Zhu J, Hayakawa A, Kakegawa T, Kaspar RL. Binding of the La autoantigen to the 5′ untranslated region of a chimeric human translation elongation factor 1A reporter mRNA inhibits translation *in vitro*. *Biochim Biophys Acta* 2001, 1521:19–29.
245. Caldarola S, De Stefano MC, Amaldi F, Loreni F. Synthesis and function of ribosomal proteins—fading models and new perspectives. *FEBS J* 2009, 276:3199–3210.
246. Kim YK, Back SH, Rho J, Lee SH, Jang SK. La autoantigen enhances translation of BiP mRNA. *Nucleic Acids Res* 2001, 29:5009–5016.
247. Trotta R, Vignudelli T, Pecorari L, Intine RV, Guerzoni C, Santilli G, Candini O, Byrom MW, Goldoni S, Ford LP, et al. BCR/ABL activates mdm2 mRNA translation via the La antigen. *Cancer Cell* 2003, 13:145–160.
248. Petz M, Them N, Huber H, Beug H, Mikulits W. La enhances IRES-mediated translation of laminin B1 during malignant epithelial to mesenchymal transition. *Nucleic Acids Res* 2012, 40:290–302.
249. Petz M, Them NC, Huber H, Mikulits W. PDGF enhances IRES-mediated translation of Laminin B1 by cytoplasmic accumulation of La during epithelial to mesenchymal transition. *Nucleic Acids Res* 2012, 40:9738–9749.
250. Naeeni AR, Conte MR, Bayfield MA. RNA chaperone activity of the human La protein is mediated by a variant RNA recognition motif. *J Biol Chem* 2012, 287:5472–5482.
251. Craig AW, Svitkin YV, Lee HS, Belsham GJ, Sonenberg N. The La autoantigen contains a dimerization domain that is essential for enhancing translation. *Mol Cell Biol* 1997, 17:163–169.
252. Wang Z, Glenn H, Brown C, Valavanis C, Liu JX, Seth A, Thomas JE, Karlstrom RO, Schwartz LM. Regulation of muscle differentiation and survival by Acheron. *Mech Dev* 2009, 126:700–709.
253. Weng H, Kim C, Valavanis C, Wang Z, Schwartz LM. Acheron: an novel LA antigen family member, binds to CASK and forms a complex with id transcription factors. *Cell Mol Biol Lett* 2009, 14:273–287.
254. Tapscott SJ. The circuitry of a master switch: myod and the regulation of skeletal muscle gene transcription. *Development* 2005, 132:2685–2695.
255. Stefanovic B, Hellerbrand C, Brenner DA. Regulatory role of the conserved stem-loop structure at the 5′ end of collagen alpha 1(I) mRNA. *Mol Cell Biol* 1999, 19:4334–4342.
256. Zhang YJ, Stefanovic B. LARP6 meets collagen mRNA: specific regulation of type I collagen expression. *Int J Mol Sci* 2016, 17:419. https://doi.org/10.3390/ijms17030419.
257. Fritz D, Cai L, Stefanovic L, Stefanovic B. Progress towards discovery of antifibrotic drugs targeting synthesis of type I collagen. *Curr Med Chem* 2011, 18:3410–3416.
258. Blackstock CD, Higashi Y, Sukhanov S, Shai SY, Stefanovic B, Tabony AM, Yoshida T, Delafontaine P. Insulin-like growth factor-1 increases synthesis of collagen type I via induction of the mRNA-binding protein LARP6 expression and binding to the 5′ stem-loop of COL1a1 and COL1a2 mRNA. *J Biol Chem* 2014, 289:7264–7274.
259. Cai L, Fritz D, Stefanovic L, Stefanovic B. Nonmuscle myosin-dependent synthesis of type I collagen. *J Mol Biol* 2010, 401:564–578.
260. Stefanovic L, Longo L, Zhang YJ, Stefanovic B. Characterization of binding of LARP6 to the 5′ stem-loop of collagen mRNAs: implications for synthesis of type I collagen. *RNA Biol* 2014, 11:1386–1401.
261. Challa AA, Stefanovic B. A novel role of vimentin filaments: binding and stabilization of collagen mRNAs. *Mol Cell Biol* 2011, 31:3773–3789.
262. Manojlovic Z, Stefanovic B. A novel role of RNA helicase A in regulation of translation of type I collagen mRNAs. *RNA* 2012, 18:321–334.

© 2017 Wiley Periodicals, Inc.

263. Vukmirovic M, Manojlovic Z, Stefanovic B. Serine-threonine kinase receptor-associated protein (STRAP) regulates translation of type I collagen mRNAs. *Mol Cell Biol* 2013, 33:3893–3906.

264. Song MH, Aravind L, Muller-Reichert T, O'Connell KF. The conserved protein SZY-20 opposes the Plk4-related kinase ZYG-1 to limit centrosome size. *Dev Cell* 2008, 15:901–912.

265. Seetharaman S, Flemyng E, Shen J, Conte MR, Ridley AJ. The RNA-binding protein LARP4 regulates cancer cell migration and invasion. *Cytoskeleton (Hoboken)* 2016, 73:680–690.

266. Zhang YJ, Stefanovic B. Akt mediated phosphorylation of LARP6: critical step in biosynthesis of type I collagen. *Sci Rep* 2016, 6:22597. https://doi.org/10.1038/srep22597.

267. Zhang Y, Stefanovic B. mTORC1 phosphorylates LARP6 to stimulate type I collagen expression. *Sci Rep* 2017, 7:41173.

268. Hsu PP, Kang SA, Rameseder J, Zhang Y, Ottina KA, Lim D, Peterson TR, Choi Y, Gray NS, Yaffe MB, et al. The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science* 2011, 332:1317–1322.

269. Zhang Y, Chen Y, Gucek M, Xu H. The mitochondrial outer membrane protein MDI promotes local protein synthesis and mtDNA replication. *EMBO J* 2016, 35:1045–1057.

270. Nykamp K, Lee MH, Kimble J. *C. elegans* La-related protein, LARP-1, localizes to germline P bodies and attenuates Ras-MAPK signaling during oogenesis. *RNA* 2008, 14:1378–1389.

271. Ichihara K, Shimizu H, Taguchi O, Yamaguchi M, Inoue YH. A Drosophila orthologue of larp protein family is required for multiple processes in male meiosis. *Cell Struct Funct* 2007, 32:89–100.

272. Chauvet S, Maurel-Zaffran C, Miassod R, Jullien N, Pradel J, Aragnol D. dlarp, a new candidate Hox target in Drosophila whose orthologue in mouse is expressed at sites of epithelium/mesenchymal interactions. *Dev Dyn* 2000, 218:401–413.

273. Deragon JM, Bousquet-Antonelli C. The role of LARP1 in translation and beyond. *WIREs RNA* 2015, 6:399–417.

274. Hopkins TG, Mura M, Al-Ashtal HA, Lahr RM, Abd-Latip N, Sweeney K, Lu H, Weir J, El-Bahrawy M, Steel JH, et al. The RNA-binding protein LARP1 is a post-transcriptional regulator of survival and tumorigenesis in ovarian cancer. *Nucleic Acids Res* 2016, 44:1227–1246.

275. Merret R, Descombin J, Juan YT, Favory JJ, Carpentier MC, Chaparro C, Charng YY, Deragon JM, Bousquet-Antonelli C. XRN4 and LARP1 are required for a heat-triggered mRNA decay pathway involved in plant acclimation and survival during thermal stress. *Cell Rep* 2013, 5:1279–1293.

276. Merret R, Nagarajan VK, Carpentier MC, Park S, Favory JJ, Descombin J, Picart C, Charng YY, Green PJ, Deragon JM, et al. Heat-induced ribosome pausing triggers mRNA co-translational decay in *Arabidopsis thaliana*. *Nucleic Acids Res* 2015, 43:4121–4132.

277. Schoenberg DR. Mechanisms of endonuclease-mediated mRNA decay. *WIREs RNA* 2011, 2:582–600.

278. Roy B, Jacobson A. The intimate relationships of mRNA decay and translation. *Trends Genet* 2013, 29:691–699.

279. Parker R. RNA degradation in *Saccharomyces cerevisiae*. *Genetics* 2012, 191:671–702.

280. Quin JE, Devlin JR, Cameron D, Hannan KM, Pearson RB, Hannan RD. Targeting the nucleolus for cancer intervention. *Biochim Biophys Acta* 1842, 2014:802–816.

281. Nilsson J, Sengupta J, Frank J, Nissen P. Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome. *EMBO Rep* 2004, 5:1137–1141.

282. Rabl J, Leibundgut M, Ataide SF, Haag A, Ban N. Crystal structure of the eukaryotic 40S ribosomal subunit in complex with initiation factor 1. *Science* 2011, 331:730–736.

283. Chandramouli P, Topf M, Menetret JF, Eswar N, Cannone JJ, Gutell RR, Sali A, Akey CW. Structure of the mammalian 80S ribosome at 8.7 Å resolution. *Structure* 2008, 16:535–548.

284. Koso H, Yi H, Sheridan P, Miyano S, Ino Y, Todo T, Watanabe S. Identification of RNA-binding protein LARP4B as a tumor suppressor in glioma. *Cancer Res* 2016, 76:2254–2264.

285. Sengupta J, Nilsson J, Gursky R, Spahn CM, Nissen P, Frank J. Identification of the versatile scaffold protein RACK1 on the eukaryotic ribosome by cryo-EM. *Nat Struct Mol Biol* 2004, 11:957–962.

286. Coyle SM, Gilbert WV, Doudna JA. Direct link between RACK1 function and localization at the ribosome *in vivo*. *Mol Cell Biol* 2009, 29:1626–1634.

287. McCahill A, Warwicker J, Bolger GB, Houslay MD, Yarwood SJ. The RACK1 scaffold protein: a dynamic cog in cell response mechanisms. *Mol Pharmacol* 2002, 62:1261–1273.

288. Thompson MK, Rojas-Duran MF, Gangaramani P, Gilbert WV. The ribosomal protein Asc1/RACK1 is required for efficient translation of short mRNAs. *Elife* 2016, 5. pii: e11154. https://doi.org/10.7554/eLife.11154.

289. Bai SW, Herrera-Abreu MT, Rohn JL, Racine V, Tajadura V, Suryavanshi N, Bechtel S, Wiemann S, Baum B, Ridley AJ. Identification and

 © 2017 Wiley Periodicals, Inc.

characterization of a set of conserved and new regulators of cytoskeletal organization, cell morphology and migration. *BMC Biol* 2011, 9:54.

290. Koso H, Takeda H, Yew CCK, Ward JM, Nariai N, Ueno K, Nagasaki M, Watanabe S, Rust AG, Adams DJ, et al. Transposon mutagenesis identifies genes that transform neural stem cells into glioma-initiating cells. *Proc Natl Acad Sci USA* 2012, 109:E2998–E3007.
291. Blagden S, Schneider C, Fischer U. Loss of LARP4B, an early event in the tumorigenesis of brain cancer? *Transl Cancer Res* 2016, 5:S1196–S1199.
292. Zhang Y, Peng L, Hu T, Wan Y, Ren Y, Zhang J, Wang X, Zhou Y, Yuan W, Wang Q, et al. La-related protein 4B maintains murine MLL-AF9 leukemia stem cell self-renewal by regulating cell cycle progression. *Exp Hematol* 2015, 43:309.e2–318.e2.
293. Mattijssen S, Maraia RJ. LARP4 is regulated by tumor necrosis factor alpha in a tristetraprolin-dependent manner. *Mol Cell Biol* 2015, 36:574–584.
294. Brooks SA, Blackshear PJ. Tristetraprolin (TTP): interactions with mRNA and proteins, and current thoughts on mechanisms of action. *Biochim Biophys Acta* 1829, 2013:666–679.
295. Walczak H. TNF and ubiquitin at the crossroads of gene activation, cell death, inflammation, and cancer. *Immunol Rev* 2011, 244:9–28.
296. Taylor GA, Carballo E, Lee DM, Lai WS, Thompson MJ, Patel DD, Schenkman DI, Gilkeson GS, Broxmeyer HE, Haynes BF, et al. A pathogenetic role for TNF alpha in the syndrome of cachexia, arthritis, and autoimmunity resulting from tristetraprolin (TTP) deficiency. *Immunity* 1996, 4:445–454.
297. Tiedje C, Diaz-Munoz MD, Trulley P, Ahlfors H, Laass K, Blackshear PJ, Turner M, Gaestel M. The RNA-binding protein TTP as a global post-transcriptional regulator of feedback control in inflammation. *Nucleic Acids Res* 2016, 44:7418–7440.
298. Orioli A, Praz V, Lhote P, Hernandez N. Human MAF1 targets and represses active RNA polymerase III genes by preventing recruitment rather than inducing long-term transcriptional arrest. *Genome Res* 2016, 26:624–635.

© 2017 Wiley Periodicals, Inc.