Published online 26 September 2016

Nucleic Acids Research, 2016, Vol. 44, No. 20 ***9977–9989***
doi: 10.1093/nar/gkw833

hLARP7 C-terminal domain contains an xRRM that binds the 3′ hairpin of 7SK RNA

Catherine D. Eichhorn, Rahul Chug and Juli Feigon*

Department of Chemistry and Biochemistry, P.O. Box 951569, University of California, Los Angeles, CA 90095-1569, USA

Received June 28, 2016; Revised September 02, 2016; Accepted September 10, 2016

ABSTRACT

The 7SK small nuclear ribonucleoprotein (snRNP) sequesters and inactivates the positive transcription elongation factor b (P-TEFb), an essential eukaryotic mRNA transcription factor. The human La-related protein group 7 (hLARP7) is a constitutive component of the 7SK snRNP and localizes to the 3′ terminus of the 7SK long noncoding RNA. hLARP7, and in particular its C-terminal domain (CTD), is essential for 7SK RNA stability and assembly with P-TEFb. The hLARP7 N-terminal La module binds and protects the 3′ end from degradation, but the structural and functional role of its CTD is unclear. We report the solution NMR structure of the hLARP7 CTD and show that this domain contains an xRRM, a class of atypical RRM first identified in the *Tetrahymena thermophila* telomerase LARP7 protein p65. The xRRM binds the 3′ end of 7SK RNA at the top of stem-loop 4 (SL4) and interacts with both unpaired and base-paired nucleotides. This study confirms that the xRRM is general to the LARP7 family of proteins and defines the binding site for hLARP7 on the 7SK RNA, providing insight into function.

INTRODUCTION

The RNA 7SK is one of the first discovered and among the most abundant long noncoding RNAs (lncRNAs), with over 100 000 copies per cell in HeLa cells (1). 7SK, an RNA Polymerase III (RNAPIII) transcript, functions as a scaffold that dynamically binds several proteins, forming a small nuclear ribonucleoprotein (snRNP) complex whose primary known function is to sequester and inactivate the positive transcription elongation factor b (P-TEFb), an essential eukaryotic transcription activator (2,3). P-TEFb, a heterodimer of the kinase cyclin dependent kinase 9 (CDK9) and cofactor Cyclin T1, phosphorylates negative elongation factors (NEFs) and Ser2 in YSPTSPS repeats in the RNAPII C-terminal tail, which transitions RNAPII initiation complexes from stalled to productive elongation complexes (4). In this way, the 7SK snRNP is an essential regulator of messenger RNA transcription. Improper regulation of P-TEFb can give rise to many human diseases, e.g. cardiac hypertrophy (5), leukemia (6–8), lymphoma (9), cervical cancer (10), and gastric cancer (11,12). Furthermore, in the HIV replication cycle, P-TEFb is hijacked from the 7SK snRNP by the transactivator of transcription (TAT) protein and recruited to the HIV 5′ UTR transactivation response (TAR) element (13–17) to facilitate HIV replication, making P-TEFb an essential host cofactor for replication of the HIV genome (18,19).

The 7SK secondary structure determined by early chemical footprinting studies showed that 7SK consists of four stem-loops (Figure 1A) (20), although more recent sequence alignment of 79 7SK sequences including both vertebrates and invertebrate species suggests that up to eight stem-loops can form (21). In both secondary structure models, stem loop 4 (SL4) has the same overall secondary structure, consisting of a hairpin containing a two-pyrimidine nucleotide bulge, a five-nucleotide apical loop, and a single-stranded UUU-3′OH. SL4 is highly conserved in both sequence and secondary structure, with loop residue G312 100% conserved (22). The solution NMR structure of a human 7SK SL4 construct bound to arginine showed that bulge residue C320 forms an unusual base triple with the G303-C323 base pair (23), similar to the base triple observed in HIV TAR (24–26).

Two proteins, methylphosphate capping enzyme (MePCE) and the La-related group 7 protein (LARP7), are constitutively bound to 7SK *in vivo* (27,28). In addition to this 'core' ternary complex, the kinase inhibitor hexamethylene bis-acetamide inducible protein 1, HEXIM1 (and less frequently HEXIM2) (29–31), and P-TEFb bind to form the minimal 'active' 7SK snRNP capable of sequestering and inactivating the kinase activity of P-TEFb (2,3). Several other proteins dynamically associate with 7SK to form different 7SK snRNP complexes, including CTIP2 (32); hnRNPs A, Q, and R (33); HMGA1 (34–38); KAP1 (39); PPM1G (40) and DDX21 helicase (41). While HEXIM1 is required to sequester P-TEFb in the 7SK snRNP complex, the remainder of these proteins appears to be involved in P-TEFb recruitment to RNAPII and/or

*To whom correspondence should be addressed. Tel: +1 310 206 6922 Fax: +1 310 825 0982; Email: feigon@mbi.ucla.edu

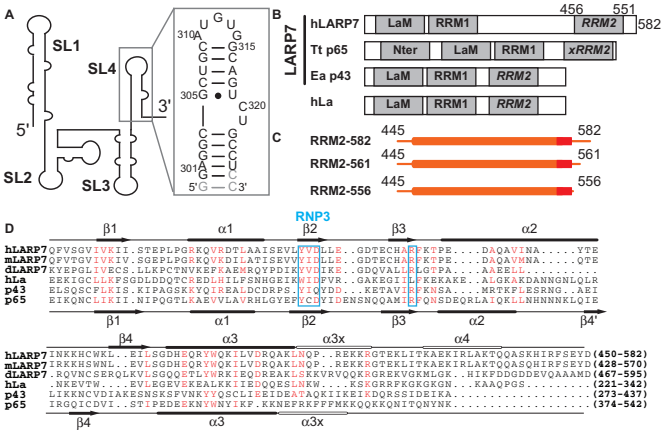

Figure 1. LARP7 family protein domain structure and homology. (**A**) Cartoon schematic of 7SK RNA with SL4 construct used for studies shown inset, with residues that deviate from wild-type shown in gray. (**B**) Domain structure of LARP7 family proteins human LARP7 (a component of 7SK RNP), *Tetrahymena thermophila* p65 and *Euplotes* p43 (telomerase holoenzyme components), and human genuine La protein. All contain a La module (La+RRM1) followed by a second RRM2 domain. (**C**) Schematic of the main RRM constructs used in this study. (**D**) Sequence alignment of human (hLARP7), mouse (mLARP7), and *Drosophila melanogaster* (dLARP7) LARP7 putative RRM2s, human genuine La (hLa), *Euplotes* p43, and *Tetrahymena* p65. Residues in red denote highly conserved equivalent residues. The RNP3 sequence and conserved Arg in β3, first identified in p65, are boxed in blue. Secondary structure elements (arrow, β-strand; solid cylinder, α-helix; open cylinder, predicted α-helix) at top and bottom denote those experimentally determined for hLARP7 and p65 xRRM2, respectively.

release from the 7SK snRNP. *In vitro*, LARP7 remains bound even under high salt (800 mM NaCl) conditions where HEXIM1, Cyclin T1 and CDK9 dissociate (27). LARP7's known function is to protect the 7SK RNA from 3′ end degradation, and it also appears to be required for P-TEFb association to the 7SK snRNP (27,42).

Human LARP7 (hLARP7), also known as PIP7S (43) and HDCMA18p (44), is a member of the genuine La and La-related protein (LARP) superfamily (44,45) and contains two structured domains: an N-terminal La module, comprised of a Lupus antigen motif (LaM) and a RNA recognition motif (RRM) that together recognize the UUU-3′OH of RNAPIII transcripts, and a C-terminal domain (CTD) that is predicted to contain a second RRM (Figure 1B) (27,42). While the N-terminal La module binds the 7SK UUU-3′OH terminus and is necessary in preventing 7SK RNA degradation (27,43,46), the role of the CTD remains speculative. Deletion of the hLARP7 CTD resulted in loss of the ability of hLARP7 to discriminate 7SK against other RNAPIII transcripts with UUU-3′OH ends (43). hLARP7 has been identified as a tumor suppressor and is downregulated in gastric cancer (11) and metastasized cervical cancer tumors (10). Frameshift mutations resulting in premature truncation of hLARP7 are correlated with human gastric cancer (10–12) and can cause a form of primordial dwarfism (47–49). An hLARP7 construct lacking the CTD failed to sequester P-TEFb in the 7SK snRNP complex (43,50), suggesting that this domain plays an important role in the stable association of P-TEFb to the 7SK snRNP.

The predicted LARP7 RRM2 sequence is highly conserved from *Drosophila melanogaster* to humans and shares sequence homology with the RRM2 of the other known LARP7 family proteins p65, found in *Tetrahymena thermophila*, and its homolog p43 in *Euplotes*, as well as with genuine La protein (Figure 1D). p65 is an essential protein component of the *Tetrahymena* telomerase holoenzyme and is required for telomerase assembly, activity, and processivity *in vivo* (51). Structural studies revealed the p65 CTD contained a novel class of atypical RRM, named xRRM (52,53). Unusual features of the xRRM include the absence of conserved RNP1 and RNP2 aromatic sequences on the β3 and β1 strands, respectively, typically involved in nucleotide recognition; the presence of an additional helix α3 that lies across the β-sheet surface, where single-stranded nucleotides usually bind; and a C-terminal tail required for RNA binding that is disordered in the free xRRM but forms an α3 extension (α3x) on binding RNA. On binding p65, *Tetrahymena* telomerase RNA stem 4 (S4) undergoes a large conformational change including inducing a ∼105° bend. Based on sequence and predicted structural homology it was proposed that the xRRM might be common to other LARP7 and La family proteins (52,53). To investigate this possibility and determine the structural basis for the role of the hLARP7 CTD in 7SK snRNP assembly, we solved the solution NMR structure of the hLARP7 RRM2, identified the domain boundaries for 7SK RNA binding, and mapped the protein and RNA binding interface. We find that the hLARP7 RRM2 is an xRRM that binds to the 7SK SL4 upper stem and apical loop residues, thereby recognizing both unpaired and base paired nucleotides in a manner similar to p65 xRRM but at a hairpin rather than an internal bulge. These findings show that the xRRM is not unique to *Tetrahymena* p65 but rather appears to be common to LARP7 family proteins and establish the binding site of hLARP7 xRRM on 7SK SL4 RNA, providing insight into function.

MATERIALS AND METHODS

Sample preparation

hLARP7 RRM2 constructs (Figure 1C) were cloned into the pet30 Xa/LIC vector with a His_6 affinity tag, modified using site-directed mutagenesis to remove the Xa sequence between the His tag and RRM2, and expressed in *Escherichia coli* BL21 (DE3) cells with a final concentration of 1 mM IPTG, grown at 18°C for 18–20 h. Cells were pelleted, resuspended in buffer R (20 mM Tris pH 8.0, 500 mM NaCl, 15 mM imidazole, 1 mM TCEP, 0.002% NaN_3), and lysed via sonication. Cells were centrifuged to separate the soluble fraction that was subsequently run on a 5 mL HisTrap HP nickel affinity column (GE Healthsciences). The column was washed with buffer R and eluted with buffer E (20 mM Tris pH 8.0, 500 mM NaCl, 200 mM imidazole, 1 mM TCEP). The elution fraction was concentrated and further purified on an S75 column attached to AKTA FPLC (GE Healthsciences). Fractions containing pure protein were pooled, buffer exchanged into protein NMR buffer (20 mM $NaPO_4$ pH 6.1, 100 mM NaCl, 1 mM TCEP, 0.002% NaN_3) using a 3 kDa Amicon (EMD Millipore), and concentrated to 0.8–1 mM for NMR studies. Un-

labeled protein was purified from cells grown in LB media (Fisher) and uniformly ^{15}N- and/or ^{13}C- labeled proteins were purified from cells grown in M9 minimal media with ^{15}N ammonium chloride and/or ^{13}C D-glucose (Cambridge Isotope Labs) as the sole nitrogen and/or carbon source, respectively.

The 7SK SL4 RNA hairpin (called SL4, residues 300–326), modified by substituting the terminal G–U base pair with a G–C base pair and adding a terminal G–C base pair (Figure 1A) was *in vitro* transcribed using T7 RNAP (P266L mutant) (54) and chemically synthesized DNA templates (Integrated DNA Technologies) following established protocols (55). Briefly, the transcription reaction (40 mM Tris pH 8, 25 mM $MgCl_2$, 1 mM spermidine, 0.01% Triton X, 2.5 mM DTT, 2 mM each rATP, rCTP, rUTP, rGTP, 0.5 μM DNA template) was incubated at 37°C for 4–6 h. Transcribed RNA was purified by 15–20% denaturing polyacrylamide gel electrophoresis (PAGE), the band containing RNA was visualized by UV shadowing and excised from the gel, and RNA was electroeluted from the gel pieces using an Elutrap device (GE Waters). RNA was further purified by ion-exchange chromatography using a DEAE column (GE Healthcare) and eluted into 10 mM sodium phosphate pH 7.6, 1 mM EDTA, 1.5 M KCl. RNA was diluted to <100 μM in ddH_2O and annealed by heating to 95°C for 3 min, followed by incubation on ice for 1 h. RNA was then buffer exchanged into appropriate buffer using a 3 kDa Amicon and concentrated to 0.4–1 mM. For NMR studies the RNA was exchanged into protein-RNA complex NMR buffer (20 mM sodium phosphate pH 6.05, 50 mM KCl, 1 mM TCEP).

The xRRM-RNA complex was prepared using a minimal xRRM construct called RRM2-556 (residues 445–556), in which a Tobacco etch virus (TEV) protease site was placed (ENLYQ) between the His_6 tag and residue S445. The protein was expressed and purified as described above. After affinity purification on the NTA column, RRM2-556 was dialyzed (20 mM Tris HCl pH 7.5, 50 mM NaCl, 1 mM TCEP) for 1–2 h at ambient temperature in the presence of TEV protease (recombinantly expressed and purified in house) to cleave the His_6 tag. RRM2-556 was purified on the HisTrap HP column a second time to separate cleaved from uncleaved protein, and the cleaved protein was further purified on an S75 column. Purified RRM2-556 protein was added to a SL4 RNA hairpin at a 1:1.2 ratio, respectively, under dilute conditions (∼25 μM) in protein-RNA complex NMR buffer and concentrated to ∼500 μl using a 3 kDa Amicon to a final concentration of ∼0.8 mM.

NMR spectroscopy

NMR protein samples were concentrated to 0.5–1.0 mM in protein NMR buffer in either 90% H_2O/10% D_2O or lyophilized and resuspended in 99.996% D_2O. NMR experiments were performed at 298 K on AVANCE 800 MHz Bruker spectrometer equipped with HCN cryoprobe. Backbone (N, H, C, Cα, Cβ, Hα, Hβ) and side-chain (Cδ, Cγ, Cϵ, Hδ, Hγ, Hϵ) assignments were obtained using standard triple resonance assignment experiments (56,57). Briefly, 3D CBCACONH, HNCACB, HNCA, HBHACONH, HNCO and HNCACO experiments from the Bruker experimental suite were acquired and analyzed to assign backbone resonances. The CCCONH experiment was acquired to assign all carbon atoms in a given residue. Once backbone assignments were completed the sample was lyophilized and re-dissolved into 99.996% D_2O for side-chain assignments. 3D TOCSY and HCCH-COSY experiments were acquired to assign side-chain resonances. ^{15}N- and ^{13}C-edited 3D NOESY experiments were acquired on samples in 90% H_2O/10% D_2O or 99.996% D_2O respectively with a mixing time of 120 ms to obtain nuclear Overhauser effect crosspeaks (NOEs) for structure determination. NMR data was collected using XwinNMR 3.5 (Bruker), processed using NMRPipe (58), and analyzed using Sparky 3.110 (59). Hydrogen–deuterium exchange experiments were acquired to determine residues protected from solvent through hydrogen bonding as follows. A 500 μl sample of 0.5 mM protein in protein NMR buffer was lyophilized for 18 h. The sample was resuspended in the equivalent volume 99.996% D_2O and quickly transferred to an NMR tube. 1.5 min SOFAST HMQC ^{1}H–^{15}N experiments (60) were run successively for 24 h to observe hydrogen-to-deuterium exchange of backbone amide protons. One-bond ^{1}H–^{15}N residual dipolar couplings (RDCs) were measured from analysis of 2D ^{1}H–^{15}N HSQC-IPAP experiments (61) in the absence and presence of 10 mg/ml Pf1 phage (ASLA Biotech, Ltd), with an observed deuterium splitting of ∼3 Hz. A total of 51 RDCs were used to refine the hLARP7 RRM2 structure (Table 1). The weighted average chemical shifts were calculated using the following equation $\sqrt{\Delta i^2 + x\Delta j^2}$, where Δi is the change in chemical shift in the ^{1}H dimension, Δj is the change in chemical shift in the second dimension (^{1}H, ^{15}N or ^{13}C), and x is a coefficient that scales the chemical shift changes of the second dimension according to the ratio between the gyromagnetic ratios γ_j/γ_i, where x = 1 for ^{1}H, 0.1 for ^{15}N and 0.25 for ^{13}C (57).

NMR structure calculation

The hLARP7 RRM2-561 structure was calculated in a semi-automated iterative manner using CYANA 2.1 (62). Intra- and inter-residue NOEs were manually picked from the 3D NOESY experiments. Dihedral angle restraints were generated using the software Talos+ (63). Hydrogen bond restraints were determined by hydrogen-deuterium exchange experiments and observation of NOE crosspeaks characteristic to α-helices and β-sheets. NOE peaklists, dihedral angle restraints, hydrogen bond restraints, and chemical shift assignments were used as input for CYANA 2.1 to generate 100 starting conformers. After each round of calculations NOESY spectra were re-analyzed to update the NOE peak list and add new NOEs consistent with the preliminary model. After multiple iterations, the preliminary structure was refined in XPLOR with additional RDC restraints. The 20 lowest energy structures out of 200 models were selected for the final ensemble.

Electrophoretic mobility shift assay (EMSA)

SL4 RNA and hLARP7 RRM2 samples were prepared separately in protein-RNA complex NMR buffer at 50 μM concentrations. RNA and protein were added together at

Table 1. Restraints and structure statistics for solution NMR structure of hLARP7 xRRM (residues 445 to 561)

Distance and dihedral angle restraints	
Total NOE restraints	1306
Intraresidue	327
Sequential	407
Medium (i+2 to i+4)	189
Long range (>i+4)	383
Hydrogen bond restraints	88
RDC restraints	51
Dihedral angle restraints	135
Structure Statistics (20 lowest energy structures)	
No. NOE violations >0.2 Å	0
No. Dihedral violations >5°	0
No. RDC violations >2 Hz	0
RMSD of RDC (Hz)	2
RMSD from ideal covalent geometry	
Bond lengths (Å)	0.00108 ± 0.00005
Bond angles (°)	0.313 ± 0.003
Impropers (°)	0.207 ± 0.005
RMSD from mean structure	
Residues 20–24, 32–42, 46–50, 56–61, 64–74, 84–88, 91–104	
Backbone (Å)	0.37
Heavy atoms (Å)	0.80
Ramachandran statistics	
Most favored regions	88.2%
Additional allowed regions	10.2%
Generously allowed regions	1.6%
Disallowed regions	0.0%

various ratios in a total volume of 8 μl with 2 μl 30% glycerol added as loading dye. Samples were loaded on a native 10% polyacrylamide gel (29:1 crosslinking ratio) and run at 100 V for ∼45 min at ambient temperature in running buffer (20 mM $NaPO_4$, pH 7, 50 mM NaCl). Gels were stained with toluidine blue.

Isothermal titration calorimetry (ITC)

The binding dissociation coefficient (K_D) for binding of hLARP7 RRM2 constructs to 7SK SL4 constructs was determined using a MicroCal 200 ITC instrument (GE). RNA and protein were individually exchanged into ITC buffer (20 mM Tris pH 7.0, 50 mM NaCl). Protein at concentrations of 100–250 μM was titrated into 5–10 μM RNA at 295 K. Calorimetric data was fit using ORIGIN 7 (MicroCal). The binding parameters stoichiometry (N), entropy (ΔS), enthalpy (ΔH) and association constant (K_a) were kept as floating variables during each fit. Experiments were performed in duplicate or triplicate with each experiment fit individually and binding parameters averaged.

RESULTS

hLARP7 C-terminal boundary requirements for RNA binding

The NCBI conserved domain database predicts an RRM at the hLARP7 CTD with a start site at residue 454, although the predicted C-terminal end residue is ambiguous, ending at residue 535 or 552 (accession Q4G0J3.1) (64,65). Based on our sequence alignment to p65, p43 and hLa, we predicted that the hLARP7 RRM2 encompassed residues 455–547 (52) (Figure 1D). To more precisely define the C-terminal residues required for high affinity binding to 7SK SL4, we made a series of protein constructs with serial C-terminal truncations (Figure 1C and Table 2). Previous chemical footprinting studies on hLARP7 La module and CTD domains indicated that the La module but not the CTD binds to the UUU-3′OH (66). For binding studies, we used an RNA construct of SL4 (residues 300–326) modified by substituting the terminal G–U base pair with a G–C base pair and adding an additional G–C base pair at the end to facilitate transcription by T7 RNAP (Figure 1A). To verify specific binding of this SL4 construct to the hLARP7 CTD, we measured ^1H–^{15}N HSQC NMR spectra of uniformly ^{15}N-labeled RRM2-582, consisting of a N-terminal His_6 affinity tag and residues 445–582, in the absence and presence of SL4 (Figure 2A). On titration of SL4, many peaks in the RRM2-582 construct initially broadened and/or disappeared and subsequently new peaks appeared as the concentration of SL4 increased, characteristic of intermediate-to-slow chemical exchange (Supplemen-

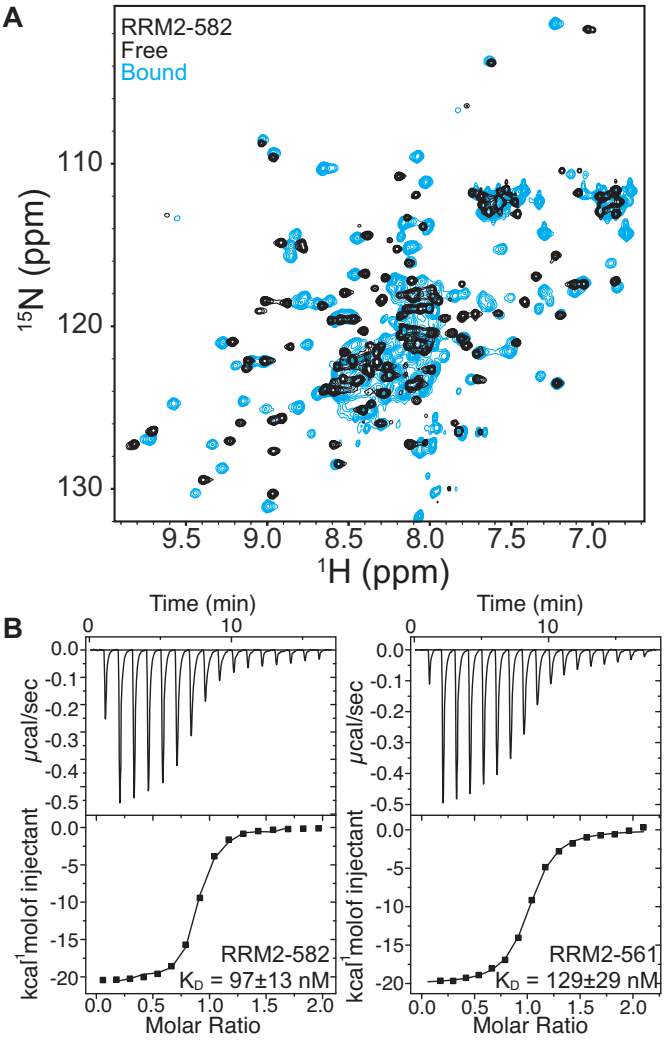

Figure 2. The hLARP7 CTD binds 7SK SL4. (**A**) ^1H-^{15}N HSQC of free (black) and bound at a 1:1.2 protein:RNA ratio (blue) RRM2-582 show large chemical shift perturbations on binding 7SK SL4 RNA. (**B**) Representative ITC plots of RRM2-582 and RRM2-561. Additional ITC plots can be found in Supplementary Data.

tary Figure S1). No further chemical shift changes in the ^{1}H–^{15}N HSQC spectra were observed above a ratio of 1:1.2 of RRM2-582:SL4. Using isothermal titration calorimetry (ITC) measurements we found that RRM2-582 bound SL4 RNA with high affinity (97 ± 13 nM) with a one-to-one stoichiometry (Figure 2B and Table 2) The one-to-one stoichiometry was also confirmed by electrophoretic mobility shift assay (EMSA) (Supplementary Figure S1). Deletion of up to 31 C-terminal amino acids (aa, RRM2-561, RRM2-556 and RRM2-551), had a minimal effect on binding with a 1.3, 1.5-fold and 2-fold reduction, respectively (Figure 2, Supplementary Figure S2, and Table 2). Truncation of another 5 aa (RRM2-546) reduced binding 18-fold, to 1824 ± 45 nM. Finally, deletion of another 9 aa (RRM2-537) reduced binding 40-fold compared to RRM2-582 (Table 2 and Supplementary Figure S2). These experiments define the minimal hLARP7 RRM2 C-terminal boundary required for high-affinity RNA binding as residue 551.

Solution NMR structure of hLARP7 RRM2

Initial NMR studies were performed on RRM2-582, which has well dispersed peaks in the backbone amide ^{1}H–^{15}N HSQC spectrum indicative of a folded protein (Supplementary Figure S3). Backbone resonance assignments (H, N, C, Cα, Cβ, Hα, Hβ) were completed for 88% of resonances using standard NMR experiments on a uniformly ^{15}N,^{13}C-labeled sample (Supplementary Table S1). The first eight residues, including the His_6 tag and first two residues of the construct, were not observed in the 2D ^{1}H–^{15}N HSQC spectrum. Chemical shift indexing (CSI) using Talos+ (63) indicates that RRM2-582 residues 455–570 adopt a βαββαβαα topology, consistent with an RRM having additional α3 and α4 helices at the C-terminus (Supplementary Figure S4). In agreement with CSI, residues 455–537 had ^{1}H–^{15}N heteronuclear NOE (hetNOE) values of ∼0.9 (Supplementary Figure S5), consistent with the RRM2 boundaries predicted by the NCBI conserved domain database. Predicted loop regions β1-α1, β2-β3, and α2-β4 have lower ^{1}H–^{15}N hetNOE values. The α2-β4 loop, particularly residues 517–520, appears to be the most disordered, with a minimum of ∼0.7 for residue 518. These amide resonances also show line broadening indicative of chemical exchange. ^{1}H–^{15}N hetNOE values decrease gradually from residues 538 (0.8), the fourth residue from the C-terminal end of α3 predicted by CSI, to terminal residue 582 (-0.1) (Supplementary Figure S5). Interestingly, residues 559–571, predicted to form an α4 helix by CSI, have elevated ^{1}H–^{15}N hetNOE s of ∼0.6 compared to the C-terminal end (residues 552–582) (Supplementary Figure S5).

Due to the high degree of disorder and resonance overlap in the C-terminal residues of RRM2-582, RRM2-561 was used for structure determination. As shown above, deletion of residues beyond 551 does not significantly affect RNA binding. 2D ^{1}H–^{15}N HSQC spectra are nearly identical between the two constructs, confirming that they have the same global fold (Supplementary Figure S3). Backbone and sidechain resonances were assigned by standard 3D NMR protein assignment experiments. Small chemical shift differences between RRM2-582 and RRM2-561 are only observed for C-terminal residues 556–561. In RRM2-582 these residues are at the N-terminus of the putative helix α4 whereas in RRM2-561 these residues are at the C-terminus, explaining the difference in chemical shift. The secondary structure determined from CSI for residues up to 559 are identical (Supplementary Figure S4). ^{1}H–^{15}N het-NOE values (Supplementary Figure S3) for RRM2-561 are also nearly identical to RRM2-582, with the exception of C-terminal residues 558–561, which all have values below zero, indicating a high degree of flexibility.

A solution NMR structure of RRM2-561 was solved to a backbone RMSD of 0.38 Å (heavy atom RMSD 0.80 Å) for the 20 lowest energy structures (Figure 3A and Table 1). The hLARP7 RRM2 structure is an atypical RRM consisting of four β-sheets and three α-helices. The front face of the RRM consists of an antiparallel β-sheet comprising β4 (residues 520–523), β1 (residues 456–460), β3 (residues 493–496), and β2 (residues 482–486) strands with helix α3 (residues 537–544) lying across the β-sheet perpendicular to the β-strand axis (Figure 3A). The back side of the RRM consists of two α-helices: helix α1 (residues 468–478) lies under β2 and β3 while helix α2 (residues 500–516) lies under β4. There are 5 structured loops: β1-α1 (residues 461–467), α1-β2 (residues 479–481), β2-β3 (residues 487–492), β3-α2 (residues 497–499), α2-β4 (residues 517–519), and β4-α3 (residues 524–536). The interface between helix α3 and the β-sheet consists of mainly aromatic and hydrophobic residues (H494, L524, H528, E529, Y532, W533, I536) with a π-π stacking interaction between H494 and Y532 (Figure 3B). At the C-terminal end of helix α3 and at the interface with the β2 strand, residues are predominantly charged or aromatic (Y483, D485, R496, K535, D539, R540, K543) (Figure 3B). The C-terminal seven residues of helix α3 (residues 538–544) extend past the β-sheet surface. Helix α3x, predicted to encompass residues 545–552, are disor-

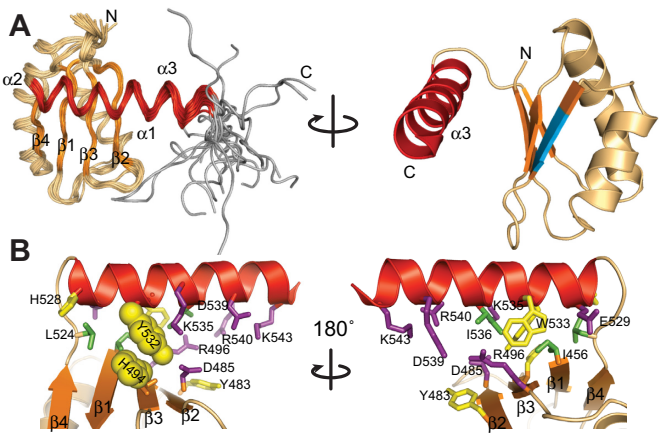

Figure 3. Solution NMR structure of hLARP7 RRM2-561. (**A**) (left) Ensemble of the 20 lowest energy NMR structures. Disordered residues 445–449 and 553–561 that are outside of the xRRM boundary are not shown in the figure. β strands shown in orange, with α3 shown in red and putative α3x residues shown in gray. (right) View of lowest energy conformation rotated to show RNA binding surface and RNP3 (blue). Residues 545–561 are not shown. (**B**) Helix α3 is stabilized on the β-sheet by hydrophobic (green) and aromatic (yellow) stacking interactions. Left: β3 residue H494 and α3 residue Y532 stack. Right: 180° rotation to show β-sheet and α3 interface on other side of xRRM. Charged residues (in purple) are located at the putative RNA binding site in helix α3 past the β-sheet.

Table 2. hLARP7 xRRM Boundary Determinants

Protein	K_D (nM)	N	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔG (kcal/mol)
RRM2-582	97 ± 13	0.82 ± 0.02	−21 ± 0.2	11 ± 0.2	−9.5 ± 0.3
RRM2-561	129 ± 29	0.98 ± 0.01	−21 ± 1.4	12 ± 1.3	−9.3 ± 1.9
RRM2-556	155 ± 6	1.0 ± 0.03	−18 ± 0.9	8.5 ± 0.9	−9.2 ± 1
RRM2-551	193 ± 27	1.0 ± 0.02	−19 ± 0.7	10 ± 0.7	−9.1 ± 1
RRM2-546	1824 ± 45	1.1 ± 0.01	−16 ± 1	8.1 ± 1	−7.7 ± 2
RRM2-537	4056 ± 721	1.0 ± 0.2	−20 ± 9	13 ± 9	−7.3 ± 13

dered as are the nine C-terminal residues (residues 553–561).

The hLARP7 RRM2 structure gives insight into the ITC data above. Deletion of the predicted helix α4 and disordered C-terminal aa up to residue 551 reduced the binding affinity only 1.3 to 2-fold (Figure 2, Supplementary Figure S2, and Table 2). However, deleting six of eight putative α3x residues in addition to helix α4 (RRM2-546) reduced binding 18-fold indicating that the putative α3x is important for RNA binding. Further truncation of the last 6 residues of helix α3 (RRM2-537), reduced binding 40-fold compared to RRM2-582 (Table 2 and Supplementary Figure S2). These data implicate helix α3 and the putative α3x in RNA binding, as expected for an xRRM.

Canonical RRMs characteristically bind RNA on the β-sheet surface with conserved RNA-recognition sequences RNP1 [(R/K)-G-(F/Y)-(G/A)-(F/Y)-(I/L/V)-X-(F/Y)] and RNP2 [(I/L/V)-(F/Y)-(I/L/V)-X-N-L] on β3 and β1, respectively (53,67,68). hLARP7 RRM2 lacks both RNP1 and RNP2: the β3 sequence is CHAR and the β1 sequence is IVKII (Figure 1D). p65 also lacks canonical RNP1 and RNP2 sequences; rather, a conserved arginine in β3 along with a proposed xRRM-specific RNP3 sequence in β2 [(Y/W)-X-D] interacts with single-stranded nucleotides of its target telomerase RNA. The hLARP7 RRM2 also contains the xRRM-specific RNP3 sequence (YVD) and conserved arginine in β3 (Figures 1D and 3A), suggesting it is an xRRM.

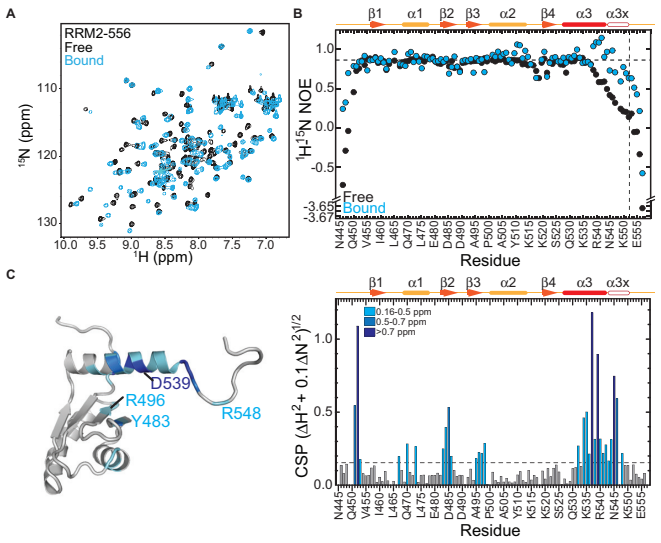

Figure 4. Mapping the hLARP7 xRRM2 RNA binding surface. (**A**) ^{1}H–^{15}N HSQC of free (black) and bound at 1:1.2 protein:RNA ratio (blue) RRM2-556 show large CSPs on binding 7SK SL4 RNA. (**B**) ^{1}H–^{15}N het-NOE of free (black) and bound (blue) to SL4 RNA show similar values for folded regions of the protein, with higher values for the putative α3x helix indicating an increase in order. (**C**) Left: CSPs mapped onto free RRM2-561 solution NMR structure, right: plot of backbone amide CSP of RRM2-556 on binding 7SK SL4. Dashed line indicates the average CSP (0.16 ppm). CSPs were binned as follows: 0.16–0.5 ppm (cyan), 0.5–0.7 ppm (blue), 0.7 ppm (dark blue).

Interaction of hLARP7 xRRM2 with 7SK SL4

To investigate the structural and dynamic properties of the hLARP7 RRM2 in the presence of 7SK SL4 RNA and determine if residues 545–552 form α3x on binding RNA, we prepared a complex of uniformly ^{13}C,^{15}N-labeled RRM2-556 bound to unlabeled SL4 RNA (Figure 4A). As observed for RRM2-582, RRM2-556 binding to SL4 occurred on the intermediate-to-slow exchange regime, and the RRM2-556:SL4 complex saturated at a 1:1.2 ratio. RRM2-556 bound to SL4 had the best quality NMR spectra compared with longer RRM2 constructs, and complete backbone assignments were obtained for RRM2-556:SL4. ^{1}H–^{15}N het-NOEs were measured on free and SL4-bound RRM2-556 to determine if the residues of the putative helix α3x become more ordered in the presence of SL4 (Figure 4B). While ^{1}H–^{15}N hetNOE values of structured regions (residues 550-539) do not change significantly when SL4 is added, helix α3 residues 538–544 and putative helix α3x residues 545–553 have significantly increased ^{1}H–^{15}N hetNOE values, indicating an increase in order at the C-terminus (Figure 4B). The requirement for these C-terminal residues for high-affinity binding to SL4 RNA, coupled with the increase in order on binding RNA, indicates that helix α3x forms in the presence of SL4.

Three of the defining features of the xRRM are that (i) it binds RNA on the 'side' of its β-sheet rather than on the surface of the β-sheet, as is characteristic of canonical RRMs, (ii) it features a disordered C-terminal tail that becomes ordered on binding RNA, forming a helix α3x extension and (iii) it binds both unpaired and base paired RNA nucleotides. To investigate the RNA binding surface of the hLARP7 RRM2 we performed chemical shift mapping experiments using the weighted average chemical shift perturbation (CSP) of unbound and bound backbone amide resonances in ^{1}H–^{15}N HSQC spectra. Significant CSPs (defined as greater than the average change of 0.16 ppm) occur primarily at β2, β3, α3, and the four disordered C-terminal residues after helix α3 (Figure 4C). Notably, conserved residues in β2, β3, and α3 (Y483, D485, R496, I536, R540), which in p65 were involved in direct contacts to *Tetrahymena* telomerase S4 RNA, have CSPs, suggesting they are involved in RNA recognition. Some of the largest

CSPs were observed for residues at the C-terminal end of helix α3 (L537-D539) and the beginning of the putative α3x (N545-Q546), which may also reflect in part formation of helix α3x helix on binding SL4, in agreement with hetNOE data. HetNOE and CSP data combined with the requirement for the C-terminal residues confirm that hLARP7 RRM2 is an xRRM.

SL4 requirements for recognition by hLARP7 xRRM

Muniz *et al* showed that many elements of the 7SK 3′ stem-loop SL4, including both upper and lower stems, an internal bulge, and nucleotides G312 and U313 in the apical loop, were essential for co-immunoprecipitation of full-length hLARP7 with a 7SK RNA construct comprising SL1 and SL4 (nucleotides 1–115 linked to 296–331) *in vivo* (69). Recent chemical footprinting studies of individual hLARP7 domains showed that the La module bound the lower half of SL4 including the UUU-3′OH and the CTD bound the upper half of SL4, consistent with the small-angle X-ray scattering (SAXS) envelope of the hLARP7 CTD bound to a 7SK SL4 construct (66). A single G312C substitution resulted in a loss of affinity for the C-terminal RRM2 observed by EMSA (66). To identify the important SL4 nucleotides for high-affinity binding to the hLARP7 xRRM, the binding affinity of different SL4 RNA constructs to RRM2-561 was measured by ITC (Figure 5A and Table 3). Substitution of the 5′-AUGUG-3′ apical loop with the canonical UUCG tetraloop completely abolished binding to hLARP7 xRRM2 (Figure 5C). Substitution of apical loop residue G312C, which has been shown *in vivo* and *in vitro* to abrogate hLARP7 binding, reduced the binding affinity 5-fold (K_D = 607 ± 130 nM) (Figure 5C and Table 3). These results indicate that the apical loop in general and G312 specifically are required for specific binding to the hLARP7 xRRM2. A construct of SL4 with only the upper stem and apical loop (SL4(306-318)) binds RRM2-561 with a K_D of 35 ± 13 nM (Figure 5B and Table 3), a ∼4-fold increase in binding affinity compared to 129 ± 29 nM for SL4. Deletion of the C320, U321 bulge from SL4 (ΔCU) also resulted in an increase in binding affinity (∼6-fold, 19 ± 2 nM) (Figure 5B). These results provide further evidence that the CU bulge is not part of the xRRM binding site, and in fact decreases binding affinity. C320 in the CU bulge was proposed to form a base triple with the G303-C323 base pair in the lower stem in the free SL4 (23); whether this triple forms in the free RNA or when hLARP7 binds under the ITC buffer conditions, and if so, interferes with binding of helix α3x remains to be determined. Together, these data define the binding site for hLARP7 xRRM as the upper stem-loop of SL4 and indicate that the CU bulge does not contribute to specific binding.

To further investigate the binding site of hLARP7 xRRM2 on the 7SK SL4 RNA, we mapped chemical shift changes in 1D and 2D spectra of a SL4(302-324), consisting of residues 302–324 with a terminal G-C base pair, on addition of RRM2-556 to form a RRM2-556:SL4 complex that saturated at a 1:1.2 ratio. CSPs were observed for imino resonances for the top four base pairs (C306-G318, U307-A317, G308-C316, C309-G315) (Figure 6A). Comparison of 2D H5-H6 TOCSY spectra shows CSPs to pyrimidines in the upper helix (residues C306, U307, C309, C316) and apical loop (residues U311 and U313) in agreement with ITC data above (Figure 6B and Table 3). Also in agreement with this data, when unlabeled RRM2-582 was added to uniformly ^{13}C,^{15}N-labeled SL4 large CSPs (>0.1 ppm) were observed for all apical loop residues as well as base paired nucleotides U307, G308, C309 and G318, and modest CSPs (0.07–0.1 ppm) were observed for C306, G315, C316 and A317 in the ^{1}H–^{13}C C6H6/C8H8/C2H2 HSQC spectrum (Supplementary Figure S6). Interestingly, small CSPs (0.03-0.07 ppm) were also observed for the CU bulge and flanking base pairs (C304, G305, U319, C320, U321 and G322). RRM2-582 contains an additional α4 helix (residues 559–571) that contains several basic residues and may contribute to the small observed difference in binding affinity (Table 2) between RRM2-582 and RRM2-561. Together, the NMR and ITC data show that a combination of base paired and single-stranded residues are involved in SL4 binding to the hLARP7 xRRM2 (Figure 6C).

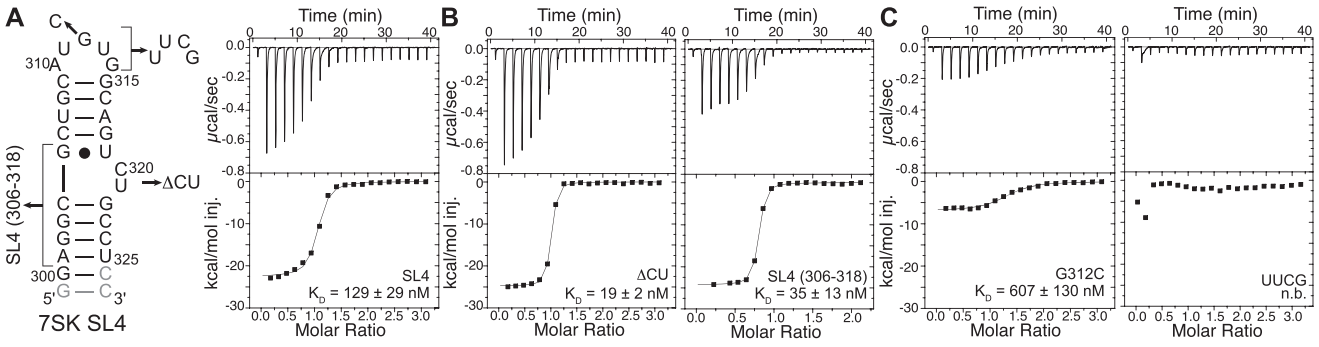

Figure 5. The upper stem and apical loop of SL4 RNA are required for high-affinity binding to the hLARP7 xRRM2. (**A**) (left) SL4 construct with substitutions marked with arrows, (right) ITC plot of SL4 binding to RRM2-561. (**B**) ITC plots of (left) ΔCU and (right) ΔStem1 constructs binding to RRM2-561 show that the lower 7SK stem is not required for binding to hLARP7 xRRM. (**C**) ITC plots of (left) G312C and (right) UUCG substitutions to the apical loop show that the apical loop, and G312 in particular, are required for high affinity binding to RRM2-561.

Table 3. 7SK SL4 determinants for binding to hLARP7 xRRM2

RNA	K_D (nM)	N	ΔH (kcal/mol)	$-T\Delta S$ (kcal/mol)	ΔG (kcal/mol)
SL4	129 ± 29	0.98 ± 0.01	−21 ± 1.4	12 ± 1.3	−9.3 ± 1.9
ΔCU	19 ± 2	0.92 ± 0.03	−24 ± 1.0	13 ± 1.1	−10.4 ± 1.5
SL4(306-318)	35 ± 13	0.91 ± 0.1	−25 ± 1.0	15 ± 1.1	−10.1 ± 1.5
G312C	607 ± 130	1.5 ± 0.3	−7.0 ± 0.1	−1.4 ± 0.2	−8.4 ± 0.3
UUCG	n.b.	n.b.	n.b.	n.b.	n.b.

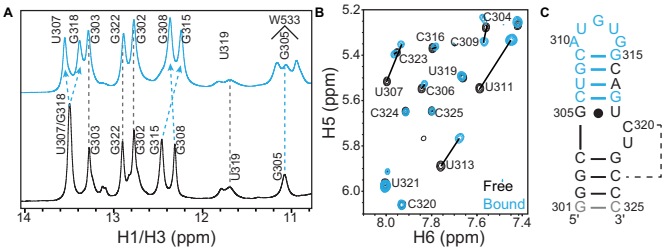

Figure 6. hLARP7 xRRM2 interacts with both the upper stem and loop residues of SL4. (**A**) 1D imino proton and (**B**) 2D H5-H6 TOCSY spectra of SL4 RNA (residues 302–325) free (black) and bound to hLARP7 RRM2-556 (blue). (**C**) CSPs from (A) and (B) with CSPs from addition of RRM2-582 to SL4 (Supplementary Data) over 0.07 ppm are combined to map the xRRM binding surface.

DISCUSSION

Comparison of RNA recognition by hLARP7 and p65 xRRM

Based on sequence homology, we previously proposed that the xRRM identified in the telomerase LARP7 protein p65 was common to genuine La and LARP7 family proteins (52,53). Here we provide a second example in hLARP7, confirming our hypothesis at least for LARP7 family proteins. Canonical RRMs typically bind single-stranded nucleotides and there are few examples of RRMs that bind double stranded RNA (68). The xRRM is a rare example of an RRM that recognizes both single and double stranded RNA residues. Although the hLARP7 CTD was shown *in vivo* to be an important specificity factor for 7SK recognition, it was unclear how it achieved this function (43). Here we show that binding of the xRRM to the upper stem and loop of 7SK SL4 confers the specificity of hLARP7 for the 7SK transcript vs. other RNAPIII transcripts that have a UUU-3′OH.

Although p65 and hLARP7 xRRM recognize both unpaired and base paired nucleotides, they interact with different RNA motifs. The p65 xRRM binds an internal two-nucleotide bulge and G-C base pairs on either side of the bulge, whereas the hLARP7 xRRM binds an apical loop and adjacent base pairs in the stem (Figure 7A and B). Co-immunoprecipitation studies showed that apical loop residues G312 and U313 are essential for hLARP7 recognition, while other apical loop residues could be substituted with little effect on hLARP7 binding, and that the upper stem structure but not sequence is required (69). The chemical shift mapping experiments reported here show that both the apical loop and base paired residues at the top of SL4 are involved in xRRM recognition (Figure 6). Consistent with these data, our ITC studies also show that the apical loop is required for xRRM binding, and a single G312C substitution significantly reduces binding affinity (Table 3). In contrast, the C320, U321 bulge does not contribute to specific binding of the hLARP7 xRRM.

The structure of the p65 xRRM2:S4 complex revealed that a proposed RNP3 sequence [(F/Y/W)-X-(D/Q/E/N)] on the β2 strand and R465 in the β3 strand are important for recognition of unpaired nucleotides (52). hLARP7 xRRM2 has an RNP3 sequence nearly identical to p65 xRRM (Y483-V484-D485 and Y407-C408-D409, respectively) and shares a conserved arginine residue on β3 (R496) (Figures 1D and 7C and D). In the X-ray crystal structures of free and RNA-bound p65 xRRM, RNP3 residue Y407 rotates 93° about the C-Cα-Cβ-Cγ dihedral bond on binding RNA (Figure 7E and F). This rotation stacks Y407 onto bulged nucleotide G121 and was proposed to act as a locking mechanism to stabilize the xRRM-RNA complex (52). Interestingly, in the solution NMR structure of the free hLARP7 xRRM the analogous residue, Y483, is already in the 'bound' position (Figure 7D) for all 20 lowest energy structures, suggesting it may be poised for recognition of an unpaired nucleotide, which is likely G312, the conserved apical loop residue required for high affinity binding of the hLARP7 xRRM to SL4. The other p65 RNP3 residue that interacts with RNA, D409, hydrogen bonds to the Watson-Crick face of G121 while the conserved arginine in β3 hydrogen bonds to the Hoogsteen face of G121. In hLARP7, the equivalent residues D485 and R496 have weighted average CSPs of 0.53 ppm and 0.23 ppm, respectively, suggesting these residues are involved in RNA recognition.

The helix α3–α3x is critical for high affinity binding to target RNA for both p65 and hLARP7. For p65, no binding was observed when α3x was deleted (52). For hLARP7, deletion of helix α3x residues reduced the binding affinity by an order of magnitude (Table 2). p65 contains several positively charged amino acids in helix α3-α3x that make contacts with the RNA phosphate backbone. For example, helix α3 residues K517 and K518 stack on bulged residue G121 and hydrogen bond to its phosphate moiety, respectively. In hLARP7, equivalent residues are R540 and Q541, both of which have sidechain amide groups capable of hydrogen bonding to the RNA phosphate backbone, and have modest CSPs of 0.32 ppm and 0.22 ppm indicating they interact with SL4. In p65, helix α3x residues R522, K528, and K529 interact with the phosphate moieties of bulge nucleotide G121, base paired nucleotide G146, and base paired nucleotide G147, respectively. In hLARP7 the equivalent residue to R522 is N545, which still has a sidechain amide group that can hydrogen bond to the RNA backbone and has a large CSP, which may also reflect conformational changes occurring at the α3-α3x boundary. In hLARP7 the equivalent residues to p65 K528 and K529 are E549 and K550, respectively. Both E549 and K550 have small back-

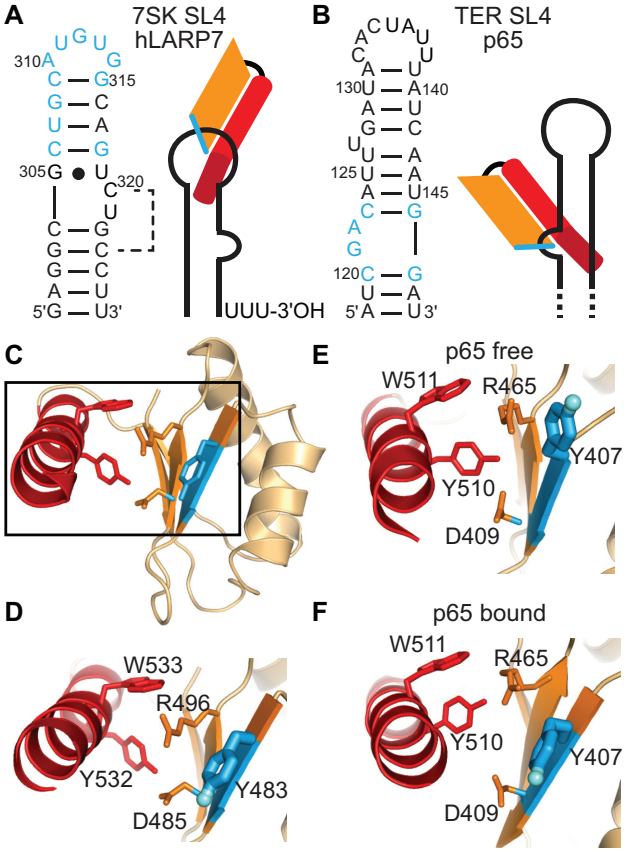

Figure 7. Comparison of RNA binding mode of hLARP7 and p65 xRRM. (**A**) 7SK SL4 sequence with nucleotides undergoing CSPs shown in blue (left) and schematic of hLARP7 xRRM2 binding to SL4 (right). (**B**) *Tetrahymena* telomerase RNA (TER) partial SL4 sequence with nucleotides interacting with p65 xRRM2 shown in blue (left) and schematic of p65 xRRM2 recognition of TER S4 (right). (**C**) hLARP7 xRRM2 solution NMR structure with black box denoting RNA binding site. (**D**) hLARP7 xRRM2 solution NMR structure with residues D485, R496, Y532, and W533 shown in stick format, with Y483 shown as a thicker stick in blue with hydroxyl moiety shown in cyan. (**E** and **F**) X-ray crystal structure of p65 xRRM2 (4EYT) (E) and bound to TER S4 (4ERD) (F). Residues D409, R465, Y510 and W511 are shown in stick format and Y407 as thicker stick in blue with hydroxyl radical shown in light blue.

bone amide CSPs. In p65, K528 and K529 interact with the base paired residues on either side of the GA bulge; since hLARP7 xRRM binds an apical loop, the mechanism for base paired RNA recognition may differ between p65 and hLARP7. In p65, aromatic residues F521, F524 and F525 in α3x insert between the base pairs on either side of the GA bulge to induce a ∼105° bend in the RNA. Strikingly, in hLARP7 equivalent residues are hydrophobic or charged (L544, P547, R548). These differences may reflect the fact that hLARP7 α3x binds at the top of a hairpin rather than between base pairs on either side of a bulge. The different modes of RNA recognition between p65 and hLARP7 xRRMs likely gave rise to differences in residues at the α3x-RNA interface.

Comparison of p65, hLARP7 and hLa RRM2

Hundreds of RRM structures have been solved, many of which have atypical features (67,68). Some classes of atypical RRMs lack RNP1 and/or RNP2, but still bind RNA via loops, e.g. quasiRRMs (qRRMs) (70–72), or helices, e.g. pseudoRRMs (73) and occluded RRMs (oRRMs) (74). Some atypical RRMs have an α-helix on the β-sheet, but this helix is not used for RNA binding: it either binds protein (75,76) and the RRM is not involved in RNA recognition (77–80), or moves out of the way for RNA binding (81,82). Although other atypical RRMs contain some unusual features in common with the xRRM, the combination of helix α3-α3x and conserved RNP3 sequence that together form a unique RNA–protein binding interface sets the xRRM apart.

The hLARP7 xRRM2 (RRM2-561) structure is highly similar to the solution NMR structures of the p65 xRRM2 (2LSL) (52) and atypical human genuine La protein hLa RRM2 (1OWX) (83) (Figure 8). p65 contains an unusually long β2-β3 loop (45 aa) while hLARP7 RRM2 contains a 3 aa β2-β3 loop. The long β2-β3 loop in p65 was shown to be dispensable for RNA binding and telomerase RNP assembly. Likewise, the hLARP7 xRRM2 β2-β3 loop is not involved in RNA recognition. p65 also contains an additional β4′ strand, which has been observed in other RRMs, including hLa (67). Instead of a β4′ strand, the hLARP7 xRRM2 has an extended α2 helix. Helix α3 in both p65 and hLARP7 xRRM2 are similarly positioned on the β-sheet, although hLARP7 helix α3 is one turn longer. The composition of the α3:β-sheet interface is highly conserved, particularly at the β4-α3 turn and beginning of helix α3 (Figure 1D). For example, hLARP7 hydrophobic residues I456 and L524 are conserved in p65 (L380 and I502, respectively) with helix α3 beginning with aspartic acid for both proteins. Several residues are identical (E529, Y532, W533, I536 in hLARP7); these residues all face the β-sheet and appear to stabilize helix α3 on the β-sheet (Figure 8). In hLARP7, Y532 stacks upon H494 in β3; in p65, analogous residue Y510 stacks on M462 in an aryl-sulfur interaction (Figure 8). Overall, the xRRM appears to have a conserved hydrophobic interface between α3 and the β-sheet.

Genuine La protein performs a broad range of essential cellular functions both in the nucleus and cytoplasm (44,84). In higher eukaryotes, La contains a C-terminal RRM2 that recognizes and is proposed to act as a chaperone for several RNA targets, e.g. pre-tRNA (85,86), the IRES domain IV of the HCV RNA virus (87), and stem-loops of precursor microRNA (88). However, in contrast to hLARP7 and p65, in which the xRRM binds with nanomolar binding affinity even in the absence of the La module, the isolated hLa RRM2 has not been shown to stably bind RNA but rather appears to enhance the binding affinity when in tandem with the La module, hinting at weak and/or allosteric binding (87,88). While the hLa RRM2 has xRRM features it has not yet been confirmed to be an xRRM. Overall hLa RRM2 has a very similar structure to p65 and hLARP7 xRRM2. There are small differences in the positions and length of helix α1 and α2. The β4–α3 turn and RNP3 sequence are conserved between hLARP7 and p65 xRRM2 and hLa RRM2 (Figure 8). Although helix α3 lies

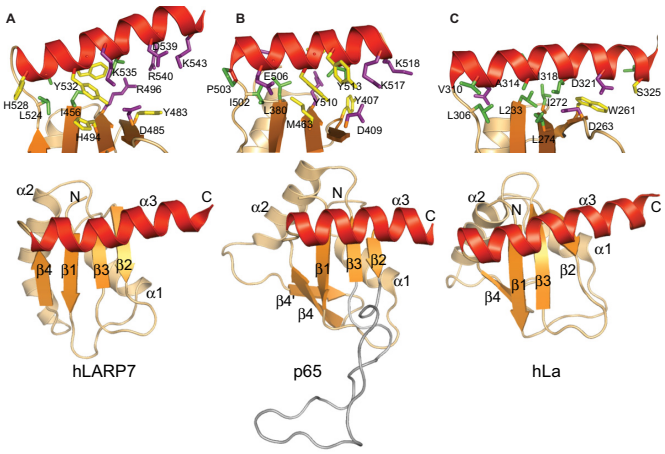

Figure 8. Comparison of structures and interaction of α3 with the β-sheet of RRM2 of hLARP7, p65, and hLa. Solution NMR structures of (top) side and (bottom) front views of (**A**) hLARP7 xRRM2, (**B**) p65 xRRM2 (PDB: 2LSL) and (**C**) hLa RRM2 (PDB: 1OWX). Side chains at the α3-β-sheet interface are shown colored by type (aromatic, yellow; hydrophobic, green; charged, purple).

across the β-sheet in all the structures, in hLARP7 and to a lesser extent p65 xRRM2 the C-terminal end of helix α3 tilts away from the β-sheet surface, whereas in the hLa RRM2 helix α3 lies at a uniform distance across the β-sheet and is closer to the β-sheet. In hLa, helix α3 is stabilized on the β-sheet predominantly by hydrophobic interactions between leucines and isoleucines, compared to p65 and hLARP7 which have similar interfaces to hLa RRM2 at the β4-α3 turn but have more aromatic and charged residues at the interface between β3-β2 and the C-terminal end of helix α3.

The putative RNA binding surface of hLa RRM2 has some differences compared to known xRRMs p65 and hLARP7. Although it has the RNP3 sequence on β2, it has a leucine on β3 instead of an arginine. In p65, R465 (R496 in LARP7) hydrogen bonds to unpaired nucleotides and is important for high affinity binding to RNA (52). In addition, hLa RRM2 has relatively few charged or aromatic residues in helix α3 that face the putative RNA binding site compared to hLARP7 and p65 (Figure 8). Residues 316–332 of hLa RRM2, comprising helix α3 and the putative helix α3x, function as a nuclear retention element that is required for binding and processing pre-tRNA (83,89). The putative hLa helix α3x (residues 327–332) is disordered in the solution NMR structure (83) and is extremely basic (Figure 1D), although it remains to be shown whether it forms helix α3x on binding specific RNA targets. In contrast to p65 and hLARP7, hLa RRM2 has multiple binding partners and may have evolved an xRRM2 to have weak affinity to RNA so that it can be released.

hLARP7 xRRM2 is required for specific recognition of 7SK SL4 RNA

A small fraction of 7SK RNA is found bound to hLa, and it is thought that hLa binds first to 7SK RNA to protect the UUU-3′OH end and is replaced with hLARP7 (44,90). Previously it was shown that deletion of the hLARP7 CTD (aa 408–582) caused it to lose specificity to 7SK RNA *in vivo* (43), indicating that the CTD contained a required specificity factor for 7SK RNA recognition. Here we show that the CTD contains an xRRM that binds to the SL4 apical loop and upper stem with high affinity. We propose the following model for hLARP7 recognition of 7SK RNA. Genuine hLa binds the nascent 7SK RNA at the UUU-3′OH end as it exits RNAPIII. The hLARP7 xRRM2 recognizes the 7SK SL4 apical loop and binds with high affinity. hLa is either actively or passively displaced at the UUU-3′OH end and replaced with the hLARP7 La module, resulting in the stable association of hLARP7 with 7SK RNA.

hLARP7 appears to enhance the association of MePCE, HEXIM, and P-TEFb to assemble the 'active' 7SK snRNP *in vitro* and *in vivo* (28,43). *In vitro* co-immunoprecipitation experiments showed that an hLARP7 CTD construct (starting at residue 375) pulled down P-TEFb subunit CDK9 as efficiently as full-length hLARP7, suggesting the hLARP7 CTD directly interacts with CDK9 (42). An hLARP7 construct (1–561) lacking the last 21 aa expressed *in vivo* was unable to co-immunoprecipitate P-TEFb and pulled down only a fraction of 7SK RNA in separate co-immunoprecipitation experiments (43). However, our studies show that *in vitro* the C-terminal 21 residues are beyond the boundary of the xRRM and are not required for 7SK SL4 recognition. Therefore, the hLARP7 C-terminus may be important for interaction with P-TEFb subunit CDK9 or for protein stability *in vivo*.

Although 7SK SL4 has been proposed to interact with the HIV TAT protein and the P-TEFb subunit Cyclin T1 at the SL4 apical loop (23,91), it has been conclusively shown that TAT binds to SL1 (92,93). Nevertheless, 7SK SL4 has similar structural features as the HIV transactivation response (TAR) element and so it is useful to discuss here. The X-ray crystal structure of the EIAV TAR:Cyclin T1:TAT ternary complex shows that Cyclin T1 interacts with two extruded EIAV apical loop nucleotides and a TAT helix interacts with the TAR major groove (94). Our studies show that the hLARP7 xRRM2 binds at the SL4 apical loop and upper stem, and we propose that the helix α3x inserts into the major groove (Figure 7A), occluding TAT binding and potentially Cyclin T1 binding to 7SK SL4.

hLa and LARPs frequently act as chaperones, facilitating RNA re-folding or protein recruitment onto RNA, in addition to protecting RNAP transcripts from degradation via binding of the La module to UUU-3′OH (44,45). For example, the p65 xRRM2 induces a large conformational change in *Tetrahymena* telomerase RNA that enhances binding of telomerase reverse transcriptase (TERT) by positioning *Tetrahymena* telomerase S4 RNA for interaction with TERT (52,95). It is possible that the LARP7 xRRM:SL4 complex positions the SL4 apical loop for favorable contacts with P-TEFb (Cyclin T1), similarly to how TAT enhances Cyclin T1 binding to TAR and p65 xRRM enhances telomerase assembly. Future studies on the structure of the LARP7 xRRM2 bound to 7SK SL4 RNA will further elucidate the protein-RNA interactions that enable specificity and provide further insights into the functional role of LARP7 xRRM2 in P-TEFb assembly onto the 7SK snRNP. The studies presented here show that the xRRM is not unique to p65 but rather is conserved in the LARP7 pro-

tein family and provide insight into how the hLARP7 CTD confers specificity toward 7SK RNA over other RNAPIII transcripts.

ACCESSION NUMBER

Coordinates for the 20 lowest energy structures of hLARP7 xRRM (residues 445–561) have been deposited in the Protein Data Base under accession code 5KNW and chemical shifts and NMR restraints used for structure calculations have been deposited in the BioMagResBank under accession code 30126.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

The authors thank Drs Mahavir Singh and Zhonghua Wang for help with NMR data analysis.

FUNDING

National Institues for Health (NIH) [GM107567 to J.F.]; NIH Tumor Cell Biology Training Grant (USHHS Ruth L. Kirschstein Institutional National Research Service Award T32 CA009056); American Heart Association Postdoctoral Fellowship [14POST20380737]; American Cancer Society Postdoctoral Fellowship [126777-PF-14-179-01-DMC to C.D.E.]; The UCLA-DOE NMR facility is supported in part by Department of Energy (DOE) [DE-FC03-02ER63421]. Funding for open access charge: NIH [GM107567].
Conflict of interest statement. None declared.

REFERENCES

1. Gurney,T. Jr and Eliceiri,G.L. (1980) Intracellular distribution of low molecular weight RNA species in HeLa cells. *J. Cell Biol.*, **87**, 398–403.
2. Yang,Z., Zhu,Q., Luo,K. and Zhou,Q. (2001) The 7SK small nuclear RNA inhibits the CDK9/cyclin T1 kinase to control transcription. *Nature*, **414**, 317–322.
3. Nguyen,V.T., Kiss,T., Michels,A.A. and Bensaude,O. (2001) 7SK small nuclear RNA binds to and inhibits the activity of CDK9/cyclin T complexes. *Nature*, **414**, 322–325.
4. Kohoutek,J. (2009) P-TEFb- the final frontier. *Cell Div.*, **4**, 19.
5. Sano,M., Abdellatif,M., Oh,H., Xie,M., Bagella,L., Giordano,A., Michael,L.H., DeMayo,F.J. and Schneider,M.D. (2002) Activation and function of cyclin T-Cdk9 (positive transcription elongation factor-b) in cardiac muscle-cell hypertrophy. *Nat. Med.*, **8**, 1310–1317.
6. Shilatifard,A., Lane,W.S., Jackson,K.W., Conaway,R.C. and Conaway,J.W. (1996) An RNA polymerase II elongation factor encoded by the human ELL gene. *Science*, **271**, 1873–1876.
7. Lin,C., Smith,E.R., Takahashi,H., Lai,K.C., Martin-Brown,S., Florens,L., Washburn,M.P., Conaway,J.W., Conaway,R.C. and Shilatifard,A. (2010) AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia. *Mol. Cell*, **37**, 429–437.
8. Chen,R., Keating,M.J., Gandhi,V. and Plunkett,W. (2005) Transcription inhibition by flavopiridol: mechanism of chronic lymphocytic leukemia cell death. *Blood*, **106**, 2513–2519.
9. Bellan,C., De Falco,G., Lazzi,S., Micheli,P., Vicidomini,S., Schurfeld,K., Amato,T., Palumbo,A., Bagella,L., Sabattini,E. *et al.* (2004) CDK9/CYCLIN T1 expression during normal lymphoid differentiation and malignant transformation. *J. Pathol.*, **203**, 946–952.
10. Biewenga,P., Buist,M.R., Moerland,P.D., Ver Loren van Themaat,E., van Kampen,A.H., ten Kate,F.J. and Baas,F. (2008) Gene expression in early stage cervical cancer. *Gynecol. Oncol.*, **108**, 520–526.
11. Cheng,Y., Jin,Z., Agarwal,R., Ma,K., Yang,J., Ibrahim,S., Olaru,A.V., David,S., Ashktorab,H., Smoot,D.T. *et al.* (2012) LARP7 is a potential tumor suppressor gene in gastric cancer. *Lab Invest.*, **92**, 1013–1019.
12. Mori,Y., Sato,F., Selaru,F.M., Olaru,A., Perry,K., Kimos,M.C., Tamura,G., Matsubara,N., Wang,S., Xu,Y. *et al.* (2002) Instabilotyping reveals unique mutational spectra in microsatellite-unstable gastric cancers. *Cancer Res.*, **62**, 3641–3645.
13. Barboric,M. and Lenasi,T. (2010) Kick-sTARting HIV-1 transcription elongation by 7SK snRNP deporTATion. *Nat. Struct. Mol. Biol.*, **17**, 928–930.
14. D'Orso,I., Jang,G.M., Pastuszak,A.W., Faust,T.B., Quezada,E., Booth,D.S. and Frankel,A.D. (2012) Transition step during assembly of HIV Tat:P-TEFb transcription complexes and transfer to TAR RNA. *Mol. Cell. Biol.*, **32**, 4780–4793.
15. He,N. and Zhou,Q. (2011) New insights into the control of HIV-1 transcription: when Tat meets the 7SK snRNP and super elongation complex (SEC). *J. Neuroimmune Pharmacol.*, **6**, 260–268.
16. Lu,H., Li,Z., Xue,Y., Schulze-Gahmen,U., Johnson,J.R., Krogan,N.J., Alber,T. and Zhou,Q. (2014) AFF1 is a ubiquitous P-TEFb partner to enable Tat extraction of P-TEFb from 7SK snRNP and formation of SECs for HIV transactivation. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E15–24.
17. Sobhian,B., Laguette,N., Yatim,A., Nakamura,M., Levy,Y., Kiernan,R. and Benkirane,M. (2010) HIV-1 Tat assembles a multifunctional transcription elongation complex and stably associates with the 7SK snRNP. *Mol. Cell*, **38**, 439–451.
18. Bieniasz,P.D., Grdina,T.A., Bogerd,H.P. and Cullen,B.R. (1999) Recruitment of cyclin T1/P-TEFb to an HIV type 1 long terminal repeat promoter proximal RNA target is both necessary and sufficient for full activation of transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **96**, 7791–7796.
19. Peterlin,B.M. and Price,D.H. (2006) Controlling the elongation phase of transcription with P-TEFb. *Mol. Cell*, **23**, 297–305.
20. Wassarman,D.A. and Steitz,J.A. (1991) Structural analyses of the 7SK ribonucleoprotein (RNP), the most abundant human small RNP of unknown function. *Mol. Cell. Biol.*, **11**, 3432–3445.
21. Marz,M., Donath,A., Verstraete,N., Nguyen,V.T., Stadler,P.F. and Bensaude,O. (2009) Evolution of 7SK RNA and its protein partners in metazoa. *Mol. Biol. Evol.*, **26**, 2821–2830.
22. Gruber,A.R., Koper-Emde,D., Marz,M., Tafer,H., Bernhart,S., Obernosterer,G., Mosig,A., Hofacker,I.L., Stadler,P.F. and Benecke,B.J. (2008) Invertebrate 7SK snRNAs. *J. Mol. Evol.*, **66**, 107–115.
23. Durney,M.A. and D'Souza,V.M. (2010) Preformed protein-binding motifs in 7SK snRNA: structural and thermodynamic comparisons with retroviral TAR. *J. Mol. Biol.*, **404**, 555–567.
24. Huthoff,H., Girard,F., Wijmenga,S.S. and Berkhout,B. (2004) Evidence for a base triple in the free HIV-1 TAR RNA. *RNA*, **10**, 412–423.
25. Brodsky,A.S., Erlacher,H.A. and Williamson,J.R. (1998) NMR evidence for a base triple in the HIV-2 TAR C-G.C+ mutant-argininamide complex. *Nucleic Acids Res.*, **26**, 1991–1995.
26. Tao,J., Chen,L. and Frankel,A.D. (1997) Dissection of the proposed base triple in human immunodeficiency virus TAR RNA indicates the importance of the Hoogsteen interaction. *Biochemistry*, **36**, 3491–3495.
27. Krueger,B.J., Jeronimo,C., Roy,B.B., Bouchard,A., Barrandon,C., Byers,S.A., Searcey,C.E., Cooper,J.J., Bensaude,O., Cohen,E.A. *et al.* (2008) LARP7 is a stable component of the 7SK snRNP while P-TEFb, HEXIM1 and hnRNP A1 are reversibly associated. *Nucleic Acids Res.*, **36**, 2219–2229.
28. Xue,Y., Yang,Z., Chen,R. and Zhou,Q. (2010) A capping-independent function of MePCE in stabilizing 7SK snRNA and facilitating the assembly of 7SK snRNP. *Nucleic Acids Res.*, **38**, 360–369.
29. Michels,A.A., Fraldi,A., Li,Q., Adamson,T.E., Bonnet,F., Nguyen,V.T., Sedore,S.C., Price,J.P., Price,D.H., Lania,L. *et al.* (2004) Binding of the 7SK snRNA turns the HEXIM1 protein into a P-TEFb (CDK9/cyclin T) inhibitor. *EMBO J.*, **23**, 2608–2619.

30. Kohoutek,J., Blazek,D. and Peterlin,B.M. (2006) Hexim1 sequesters positive transcription elongation factor b from the class II transactivator on MHC class II promoters. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 17349–17354.
31. Czudnochowski,N., Vollmuth,F., Baumann,S., Vogel-Bachmayr,K. and Geyer,M. (2010) Specificity of Hexim1 and Hexim2 complex formation with cyclin T1/T2, importin alpha and 7SK snRNA. *J. Mol. Biol.*, **395**, 28–41.
32. Cherrier,T., Le Douce,V., Eilebrecht,S., Riclet,R., Marban,C., Dequiedt,F., Goumon,Y., Paillart,J.C., Mericskay,M., Parlakian,A. *et al.* (2013) CTIP2 is a negative regulator of P-TEFb. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 12655–12660.
33. Hogg,J.R. and Collins,K. (2007) RNA-based affinity purification reveals 7SK RNPs with distinct composition and regulation. *RNA*, **13**, 868–880.
34. Eilebrecht,S., Le Douce,V., Riclet,R., Targat,B., Hallay,H., Van Driessche,B., Schwartz,C., Robette,G., Van Lint,C., Rohr,O. *et al.* (2014) HMGA1 recruits CTIP2-repressed P-TEFb to the HIV-1 and cellular target promoters. *Nucleic Acids Res.*, **42**, 4962–4971.
35. Eilebrecht,S., Wilhelm,E., Benecke,B.J., Bell,B. and Benecke,A.G. (2013) HMGA1 directly interacts with TAR to modulate basal and Tat-dependent HIV transcription. *RNA Biol.*, **10**, 436–444.
36. Eilebrecht,S., Benecke,B.J. and Benecke,A. (2011) 7SK snRNA-mediated, gene-specific cooperativity of HMGA1 and P-TEFb. *RNA Biol.*, **8**, 1084–1093.
37. Eilebrecht,S., Becavin,C., Leger,H., Benecke,B.J. and Benecke,A. (2011) HMGA1-dependent and independent 7SK RNA gene regulatory activity. *RNA Biol.*, **8**, 143–157.
38. Eilebrecht,S., Brysbaert,G., Wegert,T., Urlaub,H., Benecke,B.J. and Benecke,A. (2011) 7SK small nuclear RNA directly affects HMGA1 function in transcription regulation. *Nucleic Acids Res.*, **39**, 2057–2072.
39. McNamara,R.P., Reeder,J.E., McMillan,E.A., Bacon,C.W., McCann,J.L. and D'Orso,I. (2016) KAP1 recruitment of the 7SK snRNP complex to promoters enables transcription elongation by RNA polymerase II. *Mol. Cell*, **61**, 39–53.
40. Gudipaty,S.A., McNamara,R.P., Morton,E.L. and D'Orso,I. (2015) PPM1G binds 7SK RNA and Hexim1 To block P-TEFb assembly into the 7SK snRNP and sustain transcription elongation. *Mol. Cell. Biol.*, **35**, 3810–3828.
41. Calo,E., Flynn,R.A., Martin,L., Spitale,R.C., Chang,H.Y. and Wysocka,J. (2015) RNA helicase DDX21 coordinates transcription and ribosomal RNA processing. *Nature*, **518**, 249–253.
42. Markert,A., Grimm,M., Martinez,J., Wiesner,J., Meyerhans,A., Meyuhas,O., Sickmann,A. and Fischer,U. (2008) The La-related protein LARP7 is a component of the 7SK ribonucleoprotein and affects transcription of cellular and viral polymerase II genes. *EMBO Rep.*, **9**, 569–575.
43. He,N., Jahchan,N.S., Hong,E., Li,Q., Bayfield,M.A., Maraia,R.J., Luo,K. and Zhou,Q. (2008) A La-related protein modulates 7SK snRNP integrity to suppress P-TEFb-dependent transcriptional elongation and tumorigenesis. *Mol. Cell*, **29**, 588–599.
44. Bayfield,M.A., Yang,R. and Maraia,R.J. (2010) Conserved and divergent features of the structure and function of La and La-related proteins (LARPs). *Biochim. Biophys. Acta*, **1799**, 365–378.
45. Bousquet-Antonelli,C. and Deragon,J.M. (2009) A comprehensive analysis of the La-motif protein superfamily. *RNA*, **15**, 750–764.
46. Barboric,M., Lenasi,T., Chen,H., Johansen,E.B., Guo,S. and Peterlin,B.M. (2009) 7SK snRNP/P-TEFb couples transcription elongation with alternative splicing and is essential for vertebrate development. *Proc. Natl. Acad. Sci. U.S.A.*, **106**, 7798–7803.
47. Alazami,A.M., Al-Owain,M., Alzahrani,F., Shuaib,T., Al-Shamrani,H., Al-Falki,Y.H., Alsheddi,T., Colak,D. and Alkuraya,F.S. (2012) Loss of function mutation in LARP7, chaperone of 7SK ncRNA, causes a syndrome of facial dysmorphism, intellectual disability and primordial dwarfism. *Hum. Mutat.*, **33**, 1429–1434.
48. Hollink,I.H., Alfadhel,M., Al-Wakeel,A.S., Ababneh,F., Pfundt,R., de Man,S.A., Jamra,R.A., Rolfs,A., Bertoli-Avella,A.M. and van de Laar,I.M. (2016) Broadening the phenotypic spectrum of pathogenic LARP7 variants: two cases with intellectual disability, variable growth retardation and distinct facial features. *J. Hum. Genet.*, **61**, 229–233.
49. Ling,T.T. and Sorrentino,S. (2016) 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*, **170**, 217–219.
50. Egloff,S., Van Herreweghe,E. and Kiss,T. (2006) Regulation of polymerase II transcription by 7SK snRNA: two distinct RNA elements direct P-TEFb and HEXIM1 binding. *Mol. Cell. Biol.*, **26**, 630–642.
51. O'Connor,C.M. and Collins,K. (2006) A novel RNA binding domain in tetrahymena telomerase p65 initiates hierarchical assembly of telomerase holoenzyme. *Mol. Cell. Biol.*, **26**, 2029–2036.
52. Singh,M., Wang,Z., Koo,B.K., Patel,A., Cascio,D., Collins,K. and Feigon,J. (2012) Structural Basis for Telomerase RNA Recognition and RNP Assembly by the Holoenzyme La Family Protein p65. *Mol. Cell*, **47**, 16–26.
53. Singh,M., Choi,C.P. and Feigon,J. (2013) xRRM: a new class of RRM found in the telomerase La family protein p65. *RNA Biol.*, **10**, 353–359.
54. Guillerez,J., Lopez,P.J., Proux,F., Launay,H. and Dreyfus,M. (2005) A mutation in T7 RNA polymerase that facilitates promoter clearance. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 5958–5963.
55. Gonzalez,C., Stec,W., Kobylanska,A., Hogrefe,R.I., Reynolds,M. and James,T.L. (1994) Structural study of a DNA.RNA hybrid duplex with a chiral phosphorothioate moiety by NMR: extraction of distance and torsion angle constraints and imino proton exchange rates. *Biochemistry*, **33**, 11062–11072.
56. Reid,D.G., MacLachlan,L.K., Edwards,A.J., Hubbard,J.A. and Sweeney,P.J. (1997) Introduction to the NMR of proteins. *Methods Mol. Biol.*, **60**, 1–28.
57. Cavanagh,J. (2007) *Protein NMR Spectroscopy : Principles and Practice*. 2nd edn. Academic Press, Amsterdam; Boston.
58. Delaglio,F., Grzesiek,S., Vuister,G.W., Zhu,G., Pfeifer,J. and Bax,A. (1995) Nmrpipe - a Multidimensional Spectral Processing System Based On Unix Pipes. *J. Biomol. NMR*, **6**, 277–293.
59. Goddard,T.D. and Kneller,D.G. (2004) *SPARKY 3*. University of California, San Franscisco.
60. Schanda,P., Kupce,E. and Brutscher,B. (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR*, **33**, 199–211.
61. Ottiger,M and D.F.,Bax A. (1998) Measurement of J and dipolar couplings from simplified two-dimensional NMR spectra. *J. Magn. Reson.*, **131**, 373–378.
62. Guntert,P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.*, **278**, 353–378.
63. Shen,Y., Delaglio,F., Cornilescu,G. and Bax,A. (2009) TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J. Biomol. NMR*, **44**, 213–223.
64. Altschul,S.F., Gish,W., Miller,W., Myers,E.W. and Lipman,D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.*, **215**, 403–410.
65. Gish,W. and States,D.J. (1993) Identification of protein coding regions by database similarity search. *Nat. Genet.*, **3**, 266–272.
66. Uchikawa,E., Natchiar,K.S., Han,X., Proux,F., Roblin,P., Zhang,E., Durand,A., Klaholz,B.P. and Dock-Bregeon,A.C. (2015) Structural insight into the mechanism of stabilization of the 7SK small nuclear RNA by LARP7. *Nucleic Acids Res.*, **43**, 3373–3388.
67. Clery,A., Blatter,M. and Allain,F.H. (2008) RNA recognition motifs: boring? Not quite. *Curr. Opin. Struct. Biol.*, **18**, 290–298.
68. Muto,Y. and Yokoyama,S. (2012) Structural insight into RNA recognition motifs: versatile molecular Lego building blocks for biological systems. *Wiley Interdiscip. Rev. RNA*, **3**, 229–246.
69. Muniz,L., Egloff,S. and Kiss,T. (2013) RNA elements directing in vivo assembly of the 7SK/MePCE/Larp7 transcriptional regulatory snRNP. *Nucleic Acids Res.*, **41**, 4686–4698.
70. Dominguez,C., Fisette,J.F., Chabot,B. and Allain,F.H. (2010) Structural basis of G-tract recognition and encaging by hnRNP F quasi-RRMs. *Nat. Struct. Mol. Biol.*, **17**, 853–861.
71. Dominguez,C. and Allain,F.H. (2006) NMR structure of the three quasi RNA recognition motifs (qRRMs) of human hnRNP F and interaction studies with Bcl-x G-tract RNA: a novel mode of RNA recognition. *Nucleic Acids Res.*, **34**, 3634–3645.
72. Stamm,S., Smith,C., Lührmann,R. and ebrary Inc. (2012) Wiley-Blackwell, Weinheim, pp. xxxvii, 622 p.
73. Clery,A., Sinha,R., Anczukow,O., Corrionero,A., Moursy,A., Daubner,G.M., Valcarcel,J., Krainer,A.R. and Allain,F.H. (2013) Isolated pseudo-RNA-recognition motifs of SR proteins can regulate

splicing using a noncanonical mode of RNA recognition. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, E2802–E2811.
74. Martin-Tumasz,S., Richie,A.C., Clos,L.J., Brow,D.A. and Butcher,S.E. (2011) A novel occluded RNA recognition motif in Prp24 unwinds the U6 RNA internal stem loop. *Nucleic Acids Res.*, **39**, 7837–7847.
75. Law,M.J., Lee,D.S., Lee,C.S., Anglim,P.P., Haworth,I.S. and Laird-Offringa,I.A. (2013) The role of the C-terminal helix of U1A protein in the interaction with U1hpII RNA. *Nucleic Acids Res.*, **41**, 7092–7100.
76. Schellenberg,M.J., Edwards,R.A., Ritchie,D.B., Kent,O.A., Golas,M.M., Stark,H., Luhrmann,R., Glover,J.N. and MacMillan,A.M. (2006) Crystal structure of a core spliceosomal protein interface. *Proc. Natl. Acad. Sci. U.S.A.*, **103**, 1266–1271.
77. Kielkopf,C.L., Lucke,S. and Green,M.R. (2004) U2AF homology motifs: protein recognition in the RRM world. *Genes Dev.*, **18**, 1513–1526.
78. Eulalio,A., Tritschler,F., Buttner,R., Weichenrieder,O., Izaurralde,E. and Truffault,V. (2009) The RRM domain in GW182 proteins contributes to miRNA-mediated gene silencing. *Nucleic Acids Res.*, **37**, 2974–2983.
79. Arieti,F., Gabus,C., Tambalo,M., Huet,T., Round,A. and Thore,S. (2014) The crystal structure of the Split End protein SHARP adds a new layer of complexity to proteins containing RNA recognition motifs. *Nucleic Acids Res.*, **42**, 6742–6752.
80. Martinez-Lumbreras,S., Taverniti,V., Zorrilla,S., Seraphin,B. and Perez-Canadillas,J.M. (2016) Gbp2 interacts with THO/TREX through a novel type of RRM domain. *Nucleic Acids Res.*, **44**, 437–448.
81. Perez Canadillas,J.M. and Varani,G. (2003) Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *EMBO J.*, **22**, 2821–2830.
82. Deka,P., Rajan,P.K., Perez-Canadillas,J.M. and Varani,G. (2005) Protein and RNA dynamics play key roles in determining the specific recognition of GU-rich polyadenylation regulatory elements by human Cstf-64 protein. *J. Mol. Biol.*, **347**, 719–733.
83. Jacks,A., Babon,J., Kelly,G., Manolaridis,I., Cary,P.D., Curry,S. and Conte,M.R. (2003) Structure of the C-terminal domain of human La protein reveals a novel RNA recognition motif coupled to a helical nuclear retention element. *Structure*, **11**, 833–843.
84. Wolin,S.L. and Cedervall,T. (2002) The La protein. *Annu. Rev. Biochem.*, **71**, 375–403.
85. Fan,H., Goodier,J.L., Chamberlain,J.R., Engelke,D.R. and Maraia,R.J. (1998) 5′ processing of tRNA precursors can Be modulated by the human La antigen phosphoprotein. *Mol. Cell. Biol.*, **18**, 3201–3211.
86. Chakshusmathi,G., Kim,S.D., Rubinson,D.A. and Wolin,S.L. (2003) A La protein requirement for efficient pre-tRNA folding. *EMBO J.*, **22**, 6562–6572.
87. Martino,L., Pennell,S., Kelly,G., Bui,T.T., Kotik-Kogan,O., Smerdon,S.J., Drake,A.F., Curry,S. and Conte,M.R. (2012) 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.*, **40**, 1381–1394.
88. Liang,C., Xiong,K., Szulwach,K.E., Zhang,Y., Wang,Z., Peng,J., Fu,M., Jin,P., Suzuki,H.I. and Liu,Q. (2013) Sjogren syndrome antigen B (SSB)/La promotes global microRNA expression by binding microRNA precursors through stem-loop recognition. *J. Biol. Chem.*, **288**, 723–736.
89. Intine,R.V., Dundr,M., Misteli,T. and Maraia,R.J. (2002) Aberrant nuclear trafficking of La protein leads to disordered processing of associated precursor tRNAs. *Mol. Cell*, **9**, 1113–1123.
90. Chambers,J.C., Kurilla,M.G. and Keene,J.D. (1983) Association between the 7 S RNA and the lupus La protein varies among cell types. *J. Biol. Chem.*, **258**, 11438–11441.
91. Lu,J., Wong,V., Zhang,Y., Tran,T., Zhao,L., Xia,A., Xia,T. and Qi,X. (2014) Distinct conformational transition patterns of noncoding 7SK snRNA and HIV TAR RNAs upon Tat binding. *Biochemistry*, **53**, 675–681.
92. Lebars,I., Martinez-Zapien,D., Durand,A., Coutant,J., Kieffer,B. and Dock-Bregeon,A.C. (2010) HEXIM1 targets a repeated GAUC motif in the riboregulator of transcription 7SK and promotes base pair rearrangements. *Nucleic Acids Res.*, **38**, 7749–7763.
93. Muniz,L., Egloff,S., Ughy,B., Jady,B.E. and Kiss,T. (2010) Controlling cellular P-TEFb activity by the HIV-1 transcriptional transactivator Tat. *PLoS Pathog.*, **6**, e1001152.
94. Anand,K., Schulte,A., Vogel-Bachmayr,K., Scheffzek,K. and Geyer,M. (2008) Structural insights into the cyclin T1-Tat-TAR RNA transcription activation complex from EIAV. *Nat. Struct. Mol. Biol.*, **15**, 1287–1292.
95. Stone,M.D., Mihalusova,M., O'Connor C,M, Prathapam,R., Collins,K. and Zhuang,X. (2007) Stepwise protein-mediated RNA folding directs assembly of telomerase ribonucleoprotein. *Nature*, **446**, 458–461.