

# The L7Ae proteins mediate a widespread and highly functional protein–RNA interaction

David M.J. Lilley

(The University of Dundee,  
UK)

The k-turn is a ubiquitous structural motif in RNA forming a very tight kink in the axis of helical RNA that plays an important role in many aspects of RNA function. L7Ae is a member of a superfamily of proteins that bind k-turns in RNA, stabilizing the tightly kinked conformation. They are extremely widespread and are important in the assembly of RNA–protein complexes central to translation, splicing and site-specific RNA modification. The interaction is exploited in order to regulate the synthesis of L7Ae proteins and is itself subject to regulation in box C/D snoRNP assembly by N6 methylation of a key adenine in the k-turn. Lastly, we can exploit the L7Ae–k-turn interaction in the construction of nanoscale assemblies.

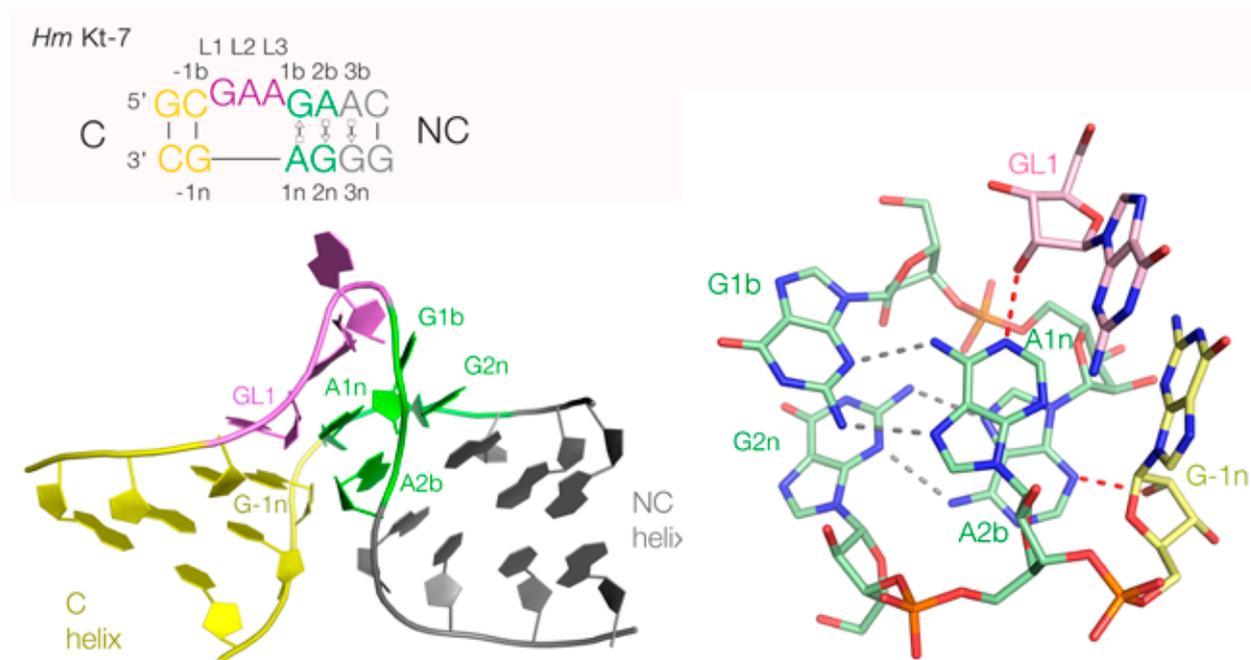
RNA is a biopolymer that is essential for life, with various roles in genetic decoding and the regulation and expression of genes. RNA can form complexes by binding to proteins; these complexes play an integral part in a number of important biological functions that includes the regulation of gene expression. Unlike DNA, which is normally double-stranded and forms the familiar double helix, RNA comes in a variety of different conformations. RNA is almost always single-stranded, although in general it folds back on itself to generate locally double-stranded regions that adopt the A-form helix (that is different from the B-form helix of DNA) connected by junctions. Moreover, the addition of the 2'-hydroxyl group significantly alters its chemical and structural properties.

The L7Ae proteins are a superfamily that bind to a widespread structural motif of RNA called the kink-turn (k-turn) in archaea and eukaryotes. The L7Ae family includes L7Ae and L30e, human 15.5k protein and yeast snu31p. These interactions are very important in the assembly of the ribosome, the snoRNP complexes that direct the site-specific modification of RNA, in the formation of the spliceosome and in additional species such as telomerase and the RNaseP ribozyme required for the maturation of transfer RNA (tRNA). These interactions are ubiquitous and are involved in nearly all the key transactions of RNA including translation, RNA covalent modification and splicing.

## The L7Ae–k-turn interaction

A typical k-turn comprises a three-nucleotide bulge followed by tandem *trans*G•A and A•G base pairs (Figure 1). The RNA folds into a tightly kinked conformation whereby the helical axes include an angle of around 50°. The conserved adenine nucleobases of the G•A base pairs are directed into the minor groove of the facing helix, and accept cross-strand hydrogen bonds from specific 2-hydroxyl groups. In the absence of protein or metal ions, the k-turns are unfolded in an extended conformation. However, upon binding an L7Ae-class protein the RNA adopts the kinked conformation. This can be shown by studying fluorescence resonance energy transfer (FRET) between fluorophores attached to the ends of the helical arms flanking the k-turn. As the population of kinked RNA rises, the average end-to-end distance becomes shorter and FRET efficiency increases. In principle, this might be used to determine the apparent affinity of binding. Yet it turns out this can be so high that at the RNA concentration required to ensure equilibrium, the fluorescence is too low to measure using fluorimetry. The apparent  $K_d$  (dissociation constant) was therefore calculated indirectly by measuring the rates of association (nearly diffusion limited, measured by stopped-flow) and dissociation, from which picomolar affinity was determined. This is an extremely high affinity.

We tend to speak loosely of the L7Ae 'inducing' the kinked conformation in RNA. Yet in principle this might occur in one of two ways. One is indeed induced fit, where binding the protein coerces the RNA to change conformation. The alternative is conformational selection



**Figure 1.** The structure of k-turns in RNA. Top left shows the sequence of a standard k-turn (*H. marismortui* Kt-7), with the nucleotide nomenclature shown. The RNA comprises a three-nucleotide bulge (purple) followed by tandem G•A and A•G base pairs (green).

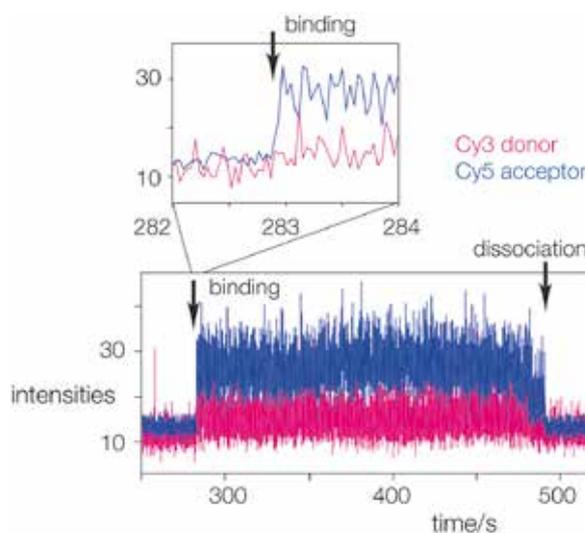
The left-hand side details the structure of the k-turn, showing the juxtaposition of the minor grooves of the two helices. The structure was determined at a resolution of 2.2 Å and is deposited with PDB ID 4C40.

The core of the structure is shown on the right. The sheared G•A and A•G base pairs are both *trans* sugar (G)-Hoogsteen (A) pairs, and the cross-strand O2' to adenine ring N atom hydrogen bonds are highlighted red. In the *trans* sugar Hoogsteen base pair the N2 of G donates a H bond to AN7, and AN6 to GN3. So that the sugar edge of the G faces the Hoogsteen edge of the A, and the two ribose rings are on opposite sides.

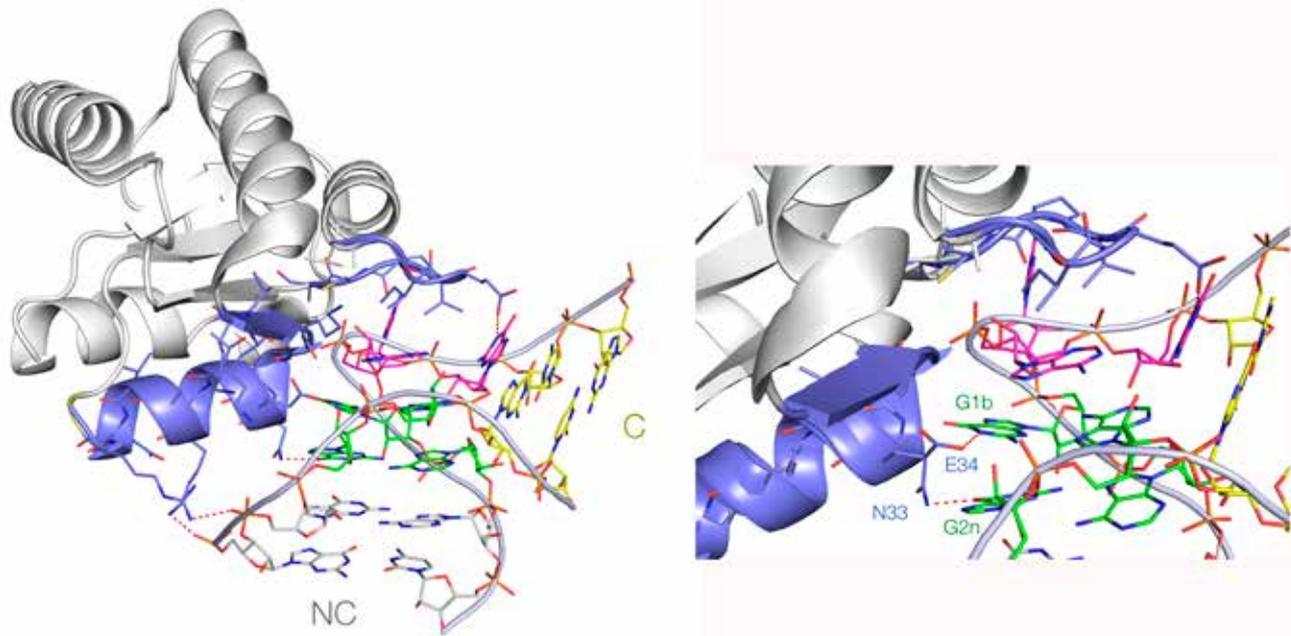
whereby the protein selectively binds to a transiently formed kinked subpopulation of the RNA and thus drives the equilibrium to the new conformation. Observation of real-time binding of L7Ae to single k-turn molecules using FRET failed to show any unfolded-but-bound RNA even at the shortest times (Figure 2). This is consistent with conformational selection, although transitions on a faster timescale cannot be excluded.

### The recognition of k-turns by L7Ae

How is the structure of the folded k-turn recognized by an L7Ae-family protein? We determined the structure of an archaeal L7Ae protein bound to a very well-characterized ribosomal k-turn called Kt-7 from the archaeon *Haloarcula marismortui* at atomic resolution (Figure 3). In forming the kinked geometry, the major groove of the RNA is opened and splayed around the outer side of the structure. An  $\alpha$ -helix from the protein enters the groove to make both non-specific backbone contacts and specific hydrogen-bonding contacts with the guanine nucleobases of the G•A base pairs. This is very reminiscent of the recognition helix of a bacterial repressor protein located



**Figure 2.** Real-time single-molecule FRET observation of RNA folding on binding the L7Ae protein. Terminally donor-acceptor fluorescently labelled RNA is observed binding to the immobilized L7Ae protein. The RNA folds on binding to the L7Ae, leading to an increase in FRET efficiency, such that the intensity of the Cy5 acceptor (blue trace) increases. The inset shows an expansion of the region where the RNA binds to the protein, exhibiting an immediate increase in FRET efficiency consistent with a conformational capture mechanism.



**Figure 3.** The molecular recognition of a k-turn in RNA by the L7Ae protein. A crystal structure was determined at 2.3 Å resolution of *A. fulgidus* L7Ae bound to *H. marismortui* Kt-7. The structure is deposited with PDB ID 4BW0. The complete structure is shown on the left, and the region of the protein–RNA interaction is shown on the right. The regions of the protein that interact with the RNA are highlighted in blue, comprising the recognition helix in the major groove, and the hydrophobic loop that caps the loop region of the k-turn. The residues E34 and N33 make specific hydrogen bonds with the guanine nucleobases of the G•A and A•G base pairs. The O6 of guanine G1b sits on the positive pole of the recognition helix dipole.

in the major groove of DNA. In double-stranded helical nucleic acid structures, the major groove may be defined by the presence of the N7 atom of the purine nucleobases and normally the major groove of RNA is deep and narrow because of its A-form conformation. While double-stranded DNA (dsDNA) generally adopts the B structure—with a central axis and major and minor grooves of equal depth and typically a C2'-endo sugar conformation—by contrast dsRNA always adopts an A-conformation structure, in which the helical axis moves about 4Å into the major groove and has a C3'-endo sugar pucker. This causes the dsRNA major groove to become very deep and narrow, and hence normally inaccessible. It is the kinked conformation of the RNA that opens the groove and allows this manner of interaction. A second feature of the interaction is a hydrophobic loop of protein that sits over the bases of the k-turn loop nucleotides. Structures of archaeal L7Ae and human 15.5k proteins bound to different k-turns have been determined, and all conform to this general style of interaction.

### L7Ae autoregulates its own synthesis

For such a versatile and widely used protein in the cell, unsurprisingly it has been found that the synthesis of L7Ae

is regulated in many archaea. Using RIP-Seq experiments (a combination of RNA immunoprecipitation followed by next-generation sequencing), Randau and colleagues found that L7Ae binds to the 5'-untranslated regions (UTR) of its own structural gene *l7ae* to suppress translation. These elements can form potential stem-loop structures that contain putative k-turn structures, and this overlaps the ribosome binding site. However, in general, these k-turns have sequences that according to our previously determined rules cannot fold unless L7Ae is bound. We have determined the crystal structure of *Archaeoglobus fulgidus* L7Ae bound to its cognate 5'-UTR sequence, showing that it forms a standard k-turn in the complex. In addition, using in-line probing experiments (where self-cleavage reactions on end-labelled RNA are used to probe local flexibility), we demonstrated that the hairpin loop only forms on binding the L7Ae protein. In this conformation it is likely that the RNA cannot bind the ribosome to initiate translation and this is how L7Ae down-regulates its own synthesis.

### Possible epigenetic regulation of snoRNP assembly

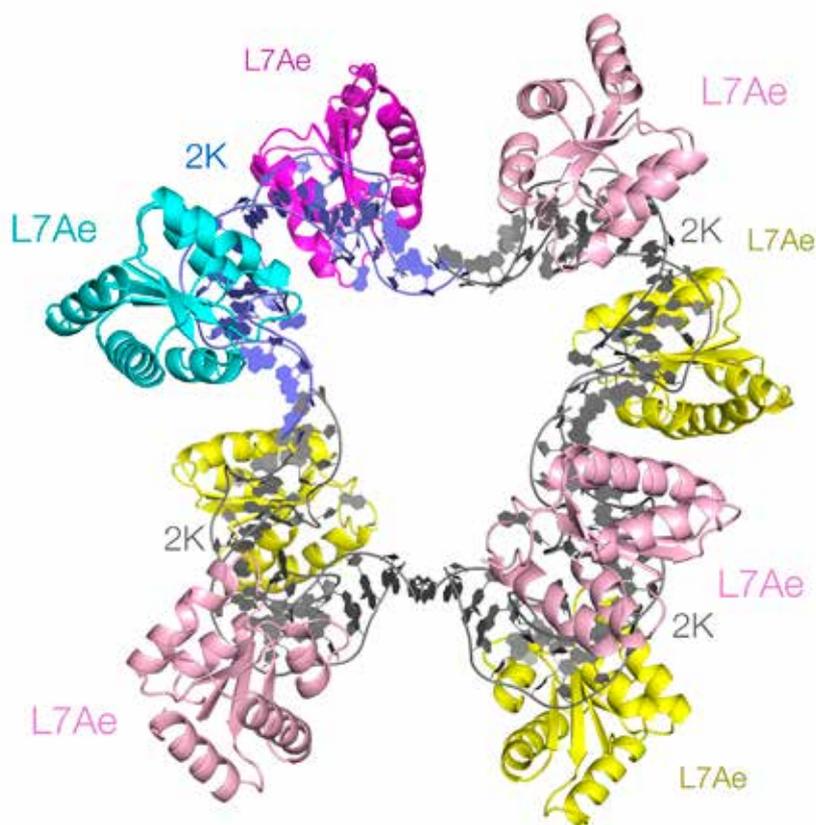
We have uncovered another form of potential regulation involving k-turn interactions in box C/D snoRNP

assembly that relates to epigenetic RNA modification. These assemblies carry out guided 2'-*O*-methylation of RNA in archaea and eukaryotes. Each snoRNP comprises a kind of open loop that is complementary to two target RNAs, flanked at either end by k-turns. In the initial step of assembly, L7Ae or 15.5k bind to the k-turns and then recruit other proteins, the last of which is the methyl transferase enzyme. However, if L7Ae/15.5k binding is prevented, the assembly process goes no further. So how does the epigenetic aspect fit into this? The most common covalent modification of RNA is the addition of a methyl group at the N6 position of adenine (N<sup>6</sup>mA). By determining a series of crystal structures of modified RNA we have shown that *cis* Watson–Crick base pairs (A–U and A–G pairings) tolerate N<sup>6</sup>mA perfectly well, but by contrast *trans*-sugar–Hoogsteen G•A base pairs (sometimes called sheared base pairs; the N2 of G donates a H bond to AN7, and AN6 to GN3 and the two ribose rings are on opposite sides) are totally disrupted by this modification. Why is this relevant? Well, the conserved G•A base pairs in the core of the k-turn form exactly this kind of sheared base pair, and so should be disrupted by N<sup>6</sup>mA inclusion. However, for the critical A to be methylated this requires the k-turn to have a particular sequence so that it is recognized by the METTL3–METTL14 methyl transferase. Bioinformatic analysis of human box C/D sequences revealed that about 10% conform to this sequence, and so are potentially methylated, and that in fact about half of these are indeed methylated *in vivo*. We then showed *in vitro* that inclusion of N<sup>6</sup>mA into human box C/D snoRNA prevented proper binding of the 15.5k protein, and the correct folding of the RNA. Very recent experiments to show that this prevents the function of the box C/D snoRNP in the O<sup>2</sup>'-methylation of the target RNA. This is not quite yet a proven regulatory mechanism based upon epigenetic modification of RNA structure, but it's close.

It is generally thought that the inclusion of N<sup>6</sup>mA into RNA is recognized by methyl-specific RNA binding proteins (e.g. the YTHDF2 protein), called 'reader' proteins in this field. Yet this provides a clear example of where the modification exerts its potential biological effect directly on the local RNA structure.

### L7Ae–k-turn complexes in RNA-based nanotechnology

The L7Ae–k-turn interaction is widely exploited in the cell, but it is also a potentially valuable building block in RNA-based nanotechnology. We have found that a short dsRNA comprising two k-turns related by two-fold symmetry (the 2K unit) can assemble in a variety of crystal lattices as two, three or four



**Figure 4.** A nanoscale assembly of 4 two-k-turn RNA units and eight molecules of L7Ae protein in a square arrangement. This structure was determined at 2.87 Å resolution and deposited with PDB ID 5G4V. The crystal lattice is monoclinic C 1 2 1. Four modified two-k-turn units associate by end-on-end stacking with four-fold rotational symmetry. The proteins bind sequentially on alternative sides of the plane of the square. In 1 two-k-turn unit (top left) the L7Ae molecules are coloured cyan and magenta, with the RNA coloured blue. The remaining L7Ae molecules are coloured pink (bound on the front face) and yellow (back face), with the RNA coloured grey.

2K units, forming dumb-bells, triangles and squares by helical end-on-end stacking. Some of these will form with RNA alone, but others require the k-turn to be bound by an L7Ae protein. For example, four variant 2K units each bound by two *A. fulgidus* L7Ae molecules assembled in a monoclinic lattice as a square structure; altogether there are eight k-turns and eight L7Ae molecules in the assembly (Figure 4). Such species hold great promise for the construction of functional nanotechnological tools. Perhaps this is no great surprise, as nature already uses them in just this manner to construct its nanomachines such as ribosomes, spliceosomes and snoRNP assemblies. The L7Ae–k-turn complex is an important component of nature's RNA nanotech toolkit!

Research in David Lilley's laboratory is funded by Cancer Research UK. ■

## Further reading

- Klein, D.J., Schmeing, T.M., Moore, P.B. and Steitz, T.A. (2001) The kink-turn: a new RNA secondary structure motif. *EMBO J.* **20**, 4214–4221
- Goody, T.A., Melcher, S.E., Norman, D.G. and Lilley, D.M.J. (2004) The kink-turn motif in RNA is dimorphic, and metal ion-dependent. *RNA* **10**, 254–264
- Moore, T., Zhang, Y., Fenley, M.O. and Li, H. (2004) Molecular basis of box C/D RNA-protein interactions; cocrystal structure of archaeal L7Ae and a box C/D RNA. *Structure* **12**, 807–818
- Vidovic, I., Nottrott, S., Hartmuth, K., Luhrmann, R. and Ficner, R. (2000) Crystal structure of the spliceosomal 15.5kD protein bound to a U4 snRNA fragment. *Molec. Cell* **6**, 1331–1342
- Huang, L. and Lilley, D.M.J. (2018) The kink-turn in the structural biology of RNA. *Quart. Rev. Biophys.* **51**, 1–32
- Liu, J. and Lilley, D.M.J. (2007) The role of specific 2'-hydroxyl groups in the stabilization of the folded conformation of kink-turn RNA. *RNA* **13**, 200–210
- Turner, B., Melcher, S.E., Wilson, T.J., Norman, D.G. and Lilley, D.M.J. (2005) Induced fit of RNA on binding the L7Ae protein to the kink-turn motif. *RNA* **11**, 1192–1200
- Turner, B. and Lilley, D.M.J. (2008) The importance of G-A hydrogen bonding in the metal ion- and protein-induced folding of a kink turn RNA. *J. Molec. Biol.* **381**, 431–442
- Wang, J., Fessl, T., Schroeder, K.T., Ouellet, J., Liu, Y., Freeman, A.D. and Lilley, D.M.J. (2012) Single-molecule observation of the induction of k-turn RNA structure on binding L7Ae protein. *Biophys. J.* **103**, 2541–2548
- Huang, L. and Lilley, D.M.J. (2013) The molecular recognition of kink-turn structure by the L7Ae class of proteins. *RNA* **19**, 1703–1710
- Huang, L., Ashraf, S. and Lilley, D.M.J. (2019) The role of RNA structure in translational regulation by L7Ae protein in archaea. *RNA* **25**, 60–69
- Huang, L., Ashraf, S., Wang, J. and Lilley, D.M. (2017) Control of box C/D snoRNP assembly by N6-methylation of adenine. *EMBO Rep.* **18**, 1631–1645
- Huang, L. and Lilley, D.M.J. (2016) A quasi-cyclic RNA nano-scale molecular object constructed using kink turns. *Nanoscale* **8**, 15189–15195



*David Lilley FRS is Professor and Director of the Cancer Research UK Nucleic Acids Research Group at the University of Dundee, and a visiting Professor at Xiamen University and Nankai Universities. He combines structural, biophysical and mechanistic studies to explore RNA structure and function, and the origins of chemical catalysis in the ribozymes. His laboratory also studies the structure and dynamics of DNA Holliday junctions and their interactions with proteins. Email: d.m.j.lilley@dundee.ac.uk*



A BIOCHEMICAL SOCIETY SCIENTIFIC MEETING

## The PI3K/PTEN pathway: from basic science to clinical translation

18–20 July 2019  
Buxton, UK

Find out more: <http://bit.ly/PI3K-PTEN>