Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

🔓 OPEN ACCESS | Check for updates

Evolution of the RNA alternative decay *cis* element into a high-affinity target for the immunomodulatory protein Roquin

Jan-Niklas Tants [a], Katharina Friedrich[a], Jasmina Neumann[a], and Andreas Schlundt [a,b]

aInstitute for Molecular Biosciences and Biomolecular Resonance Center (BMRZ), Goethe University Frankfurt, Frankfurt, Germany; bInstitute of Biochemistry, University of Greifswald, Greifswald, Germany

ABSTRACT

RNA *cis* elements play pivotal roles in regulatory processes, e.g. in transcriptional and translational regulation. Two stem-looped *cis* elements, the constitutive and alternative decay elements (CDE and ADE, respectively) are shape-specifically recognized in mRNA 3' untranslated regions (UTRs) by the immune-regulatory protein Roquin. Roquin initiates mRNA decay and contributes to balanced transcript levels required for immune homoeostasis. While the interaction of Roquin with several CDEs is described, our knowledge about ADE complex formation is limited to the mRNA of *Ox40*, a gene encoding a T-cell costimulatory receptor. The *Ox40* 3'UTR comprises both a CDE and ADE, each sufficient for Roquin-mediated control. Opposed to highly conserved and abundant CDE structures, ADEs are rarer, but predicted to exhibit a greater structural heterogeneity. This raises the question of how and when two structurally distinct *cis* elements evolved as equal target motifs for Roquin. Using an inter-disciplinary approach, we here monitor the evolution of sequence and structure features of the *Ox40* ADE across species. We designed RNA variants to probe en-detail determinants steering Roquin-RNA complex formation. Specifically, those reveal the contribution of a second RNA-binding interface of Roquin for recognition of the ADE basal stem region. In sum, our study sheds light on how the conserved Roquin protein selected ADE-specific structural features to evolve a second high-affinity mRNA target *cis* element relevant for adaptive immune regulation. As our findings also allow expanding the RNA target spectrum of Roquin, the approach can serve a paradigm for understanding RNA-protein specificity through back-tracing the evolution of the RNA element.

Introduction

RNA elements regulate numerous cellular processes, such as gene silencing, translation, mRNA decay and localization or integrate external stimuli. While the first discovered so-called *cis*-regulatory elements were single-stranded, e.g. mRNA start codons [1] and AU-rich elements (ARE) [2,3], we continue identifying folded elements comprising a great structural variety [4–6]. Hairpin elements account for a large number of these structures, and they occur in different flavours with varying loop sizes or internal bulges and branches. Tertiary and long-range contacts in RNAs are rare [7], but crucial as they mediate the formation of highly complex structures, e.g. in viral replication elements [8]. Evolutionarily, these critical contacts are often highly conserved [9]. RNA *cis* elements serve as a binding platform for *trans*-acting factors, e.g. proteins that confer functionality to the protein RNA complex (RNP). RNA target recognition can be either sequence- [10] or shape-specific [11]. Several proteins exploit their modular design to integrate both sequence and shape specificity [12,13] or to increase affinity towards their targets [14] via multiple RNA-binding domains. The limited number of tertiary RNA

contacts favours the formation of isolated structured building blocks that often form *cis* element hubs and can exert cooperative or antagonistic effects [15,16]. Changes in *cis* element sequence or structure have been associated with diseases, as subtle alterations can lead, e.g. to the disruption of functional RNPs [17].

RNA shape recognition requires a precise/defined geometry, as sequence-specific contacts between protein and RNA are limited because RNA structure is accompanied by shielding nucleobases [18]. Complex formation is then dependent on recognition of exposed single-stranded nucleotides as found in loops or kinks [18], or the overall RNA shape. By now, there are only a few RNPs that provide specific recognition of *cis* element shape. Most prominent are U1A [19], FUS [12], double-stranded RNA-binding proteins (dsRBPs) [20,21] and Roquin [22]. Roquin controls transcript levels through initiation of mRNA decay [23]. It engages with hairpin structures that comprise tri- and hexaloops, which were termed constitutive and alternative decay elements (CDE and ADE, respectively) [22,24]. The Roquin coreROQ domain harbours the stem-loop binding interface, called A-site, which is highly

conserved across vertebrates, while a second interface (B-site) composed of flanking helical extensions of the ROQ domain (extROQ) was suggested to engage with double-stranded RNA [25]. CDEs and ADEs occur solitary or in clusters in mRNA 3' untranslated regions (UTR) [16,26]. At least 13 functional CDEs have been confirmed in nine 3'UTRs [16,23,24,27,28], while only four ADEs [16,24,27] were discovered so far. Predictions of further ADEs, however, indicated a great structural variety in contrast to the seemingly more conserved CDE fold [28]. Currently, we do not understand how Roquin integrates this structural variety of targets into high-affinity RNPs.

Both decay elements have proven sufficient for Roquin-mediated posttranscriptional gene regulation, but they can also act cooperatively, e.g. in *NFKBID* [16]. For the *Ox40* (*TNFRSF4*) tandem ADE-CDE cassette, it was shown that Roquin engages with both elements independently to form a 2:1 complex with low nanomolar affinities [26]. Evolution of two functionally redundant elements might serve as a backup strategy for tight regulation of critical genes. It could also offer yet undiscovered hidden contributions to the fine-tuning of gene regulation, which might be coupled to the structural differences between ADE and CDE. Evolution of a tri- and hexaloop RNA hairpin element as equal high-affinity targets for a single protein highlights the great plasticity of Roquin RNPs. Protein-RNA pairs of high functional relevance often co-evolved, meaning that sequence and structure alterations in one binding partner entailed adaptations in the other interacting molecule [29,30]. One driving force for such co-evolution can be the affinity between the involved interaction partners [31]. For RNA, its poor structural predictability hampers a reliable judgement of preservation or modulation of RNA element structure upon changes in sequence throughout evolution [32,33], and it requires sophisticated bioinformatic approaches to also take co-variation into account [34]. Co-evolution of *cis-trans* pairs has led to specialized binding interfaces [35], implying beneficial functions for the organism, and was found to occur even over comparably short time periods [36]. In fact, host pathogen interactions have been identified as a hotspot for rapid co-evolution of protein RNA pairs [37]. Consequently, the immune system is expected to reflect evolutionary adaptations on the protein and RNA levels

Roquin controls many targets with immune-regulatory functions. Up to date we lack information on when and how the two types of decay elements evolved. The CDE was described first, however, no study so far analysed whether ADE and CDE evolved simultaneously or sequentially. In the latter case, this could point at additional functions of the second, structurally different element required in immune system control. Since comparative analyses of *cis* elements often focus on mammals or even apes, we do have no examples of progenitor ADEs. Early ancestors of known ADEs and CDEs could deepen our knowledge of Roquin–RNA interaction, contribute to Roquin target definition by revealing RNA features key for recognition and shed light on the benefits of the evolution of two RNA elements similar in function, yet different in structure. Monitoring the evolution of an RNA *cis* element should reflect key steps in structure optimization

towards a high-affinity target for a cognate RBP. It is thus tempting to track back the occurrence of decay elements. So far, high-resolution structural studies of Roquin RNPs focused on small stem-loop elements bound by the coreROQ domain [22,24], omitting the full network of interactions of the extROQ interface with extended RNA elements. The precise interplay of Roquin A and B-sites is thus unknown as well as their contribution to structure selection during the evolution of target *cis* elements. Of note, there is evidence that both sites contribute to the recognition of distinct RNA elements [25,38]. For a full understanding of Roquin target specificity, both RNA binding interfaces and an extended *cis* element contexts need to be considered that exceeds the apical stem-loops used so far. In fact, in this context, the larger structural plasticity of ADEs compared to CDEs can be exploited to reveal structural features of Roquin RNP formation that may have been overseen so far.

We here assessed the degree of secondary structure in *Ox40* ADEs from six mammals as well as two putative ancestor ADEs from reptiles and birds. Mammalian ADEs showed a high conservation of stability and shape, which significantly enhanced Roquin binding compared to possible ancestral ADEs. Technical mutants of human and murine *Ox40* ADEs revealed that the Roquin B-site is particularly sensitive to sequence changes leading to a lowered stability of the central stem. Further, geometric hallmarks and asymmetric base distribution within the RNA stem fine-tuned affinity during evolution. We observed that the orientation and size of a central bulge affects RNP formation. Together, we give evidence for the divergent roles of and target requirements in the Roquin A and B-sites. Our evolutionary study highlights the formation of stable hairpin structures as a key event in the evolution of ADEs as Roquin targets, while technical mutants reveal geometric parameters for the fine-tuning of complex formation that may have developed during evolution. We suggest that the mammalian ADE has evolved as a second and equal Roquin target *cis* element for critical genes under the pressure of the immune system.

Material and methods

RNA sequence alignment and ADE prediction

Ox40 3'UTR sequences were obtained from the UCSC Genome Browser [39]. For opossum, chicken and lizard predicted 3'UTR sequences were used. As no clear alignment could be obtained in the UCSC Genome Browser for these species, the three putative UTRs were checked manually in frames of 100 nucleotides for potential ADE-like folds that harbour a U-rich loop with at least six nucleotides. Tandem ADE-CDE alignment was performed with Clustal Omega [40]. Phylogenetic trees were obtained from alignments. For the visualization of exon and 3'UTR conservation of genes the Multiz Alignment of 100 vertebrates from the UCSC Genome Browser was used. RNA secondary structure prediction was done with the RNAfold tool from the Vienna RNA Websuite [41,42]. ΔG values were obtained from the same webserver. RNA constructs for *in vitro* transcription were designed to be similar in size and to contain the predicted central bulge as

well as the lower stem adjacent to the bulge, whenever possible.

Protein expression and purification

Murine protein constructs of coreROQ (aa 171–326) and extROQ (aa 89–404) were expressed and purified as described before [26]. Briefly, an LB pre-culture of *E. coli* BL21 (DE3) with 50 μg/mL Kanamycin was inoculated from a single colony and incubated at 37°C overnight shaking. From this, a preparative culture in M9 minimal medium containing 1 g/L ^{15}N NH_4Cl was started. Protein expression was induced with IPTG overnight, cells were harvested and lysed in 50 mm Tris pH 8.0, 500 mm NaCl, 4 mm β–mercaptoethanol. Proteins were purified via IMAC (Ni-NTA), subsequent TEV-cleavage and a reverse IMAC. After SEC (HiLoad S75 16/600 from GE) proteins were flash frozen, stored at −80°C and buffer exchanged to 150 mm NaCl, 20 mm Tris pH 7.0, 2 mm TCEP before usage. Purity was confirmed by SDS-PAGE and NMR spectroscopy.

RNA in vitro transcription and purification

ADE RNAs (Table 1) were *in vitro* transcribed from a plasmid containing the HDV ribozyme for 3'-end homogeneity. After transcription by T7 polymerase at 37°C RNAs were purified via denaturing 12% polyacrylamide gels and eluted from gels by 'crush-and-soak'. Purified RNAs were stored at −20°C in water. Quality was assessed by denaturing PAGE and NMR spectroscopy. For buffer exchange and concentrating Amicon® filter units were used. All RNAs were snap-cooled for folding prior to use, i.e. RNAs were heated to 95°C for 5 min and rapidly cooled down in an ice bath.

NMR spectroscopy

All measurements were performed at Bruker AV spectrometers (600, 700, 900 MHz proton Larmor frequencies) equipped with triple-resonance cryoprobes. Spectra were processed with Topspin 4.0.6 and analysed with NMRFAM-Sparky 1.414 [43]. RNA-only samples were buffer exchanged to 25 mm potassium phosphate pH 7.0, 50 mm potassium chloride, while protein and RNP samples were dialysed to 20 mm Tris pH 7.0, 150 mm NaCl, 2 mm TCEP prior measurement. For RNA imino proton assignments, $^{1}H,^{1}H$-NOESY experiments were recorded at 283 K with mixing times of 150 or 200 msec. 1D imino proton spectra were recorded at 283, 298 and 310 K. Additionally, spectra in the presence of 1 mm $MgCl_2$ were recorded at 298 K. To monitor protein RNA interaction, $^{1}H,^{15}N$-HSQC spectra were recorded of apo protein and a 1.5-fold excess of RNA at 298 K.

Small-angle X-ray scattering (SAXS)

RNAs were dialysed to 25 mm potassium phosphate pH 7.0, 50 mm potassium chloride with optional addition of 1 mm $MgCl_2$, refolded by snap-cooling and flash frozen for transportation. Measurements were carried out at beamline P12 (PETRA III) at DESY, Hamburg, at 293 K. The ATSAS [44] package and RAW [45 2.3.0] were used for data processing and analysis. R_gs were obtained from RAW/ATSAS. Dimensionless Kratky plots were generated with RAW. RNA tertiary structure models were generated with RNAMasonry [46] in 50 simulation steps based on our SAXS curves, and the fit quality to the experimental SAXS data was estimated with the implemented CRYSOL tool as X^2.

Circular dichroism (CD) spectroscopy

CD spectra were recorded in a JASCO J-810 spectrometer using 20 μM samples in 25 mm potassium phosphate pH 7.0, 50 mm potassium chloride in High Precision Cell Quartz cuvettes (Hellma Analytics) with a 1 mm path. Spectra were recorded in triplicates at 20°C across wavelengths from 180 to 320 nm with a data interval of 0.5 nm. Melting curves were recorded from 5°C to 95°C at 260 nm in an interval of 0.2°C with a temperature ramp of 1°C/min. Melting curves were normalized to 5°C. Data were plotted and melting points were calculated with OriginPro 2024 using a bi-dose response curve fit covering the full temperature range.

Table 1. ADE RNAs used in this study.

Species	Variant	Abbreviation	Sequence (5'→3')	MW (kDa)
Homo sapiens	wt	HS	GGCCAACUCUGCACCGUUCUAGGUGCCGAUGGCU	11.06
Homo sapiens	mirrored	HS_{mirr}	GGCCUAGCCGUGGGUUCUACCACGUCUCAAGGCU	11.06
Homo sapiens	swapped bulge	HS 3'bulge	GGCCAUCUGCACCGUUCUAGGUGCCGAACUGGCU	11.06
Homo sapiens	low purine in 3' stem	HS low pur.	GGCGAACAGUGCAGCGUUCUAGCUGCCCUUCGCU	11.06
Mus musculus	wt	MM	GCCUGCCAGUACCCUCCACACCGUUCUAGGUGCUGGGCUGGCUCUGGGC	15.84
Mus musculus	destabilized in central stem	d_{stem1}	GCCUGUAAGUACCCUCCACACCGUUCUAGGUGCUGGGCUUACUCUGGGC	15.81
Mus musculus	destabilized in central stem	d_{stem2}	GCCUGUAAGUACCCUCUACACCGUUCUAGGUGCUAGGCUUACUCUGGGC	15.80
Mus musculus	destabilized in apical stem	d_{apical}	GCCUGCCAGUACCCUCCAUAAAGUUCUAUUUACUGGGCUGGCUCUGGGC	15.80
Mus musculus	deleted bulge	Δbulge	GCCUGCCAGUCCACACCGUUCUAGGUGCUGGGCUGGCUCUGGGC	14.29
Mus musculus	swapped bulge	MM 3'bulge	GCCUGCCAGUCCACACCGUUCUAGGUGCUGGACCCUGCUGGCUCUGGGC	15.84
Tursiops truncates (Dolphin)/ *Orcinus orca* (Killer whale)	wt	DK	GCCCGUGCCGUUUUAGGCGCUCCUGGGC	9.44
Loxodonta africana (Elephant)	wt	EL	GUCCCCUCUGCGCCGUUUUAGGCGUUUCUGGGCUGGC	11.97
Erinaceus europaeus (Hedgehog)	wt	HE	GCACCCUCCCACACCGUUUUAGGUGUUCUCGGGUG	11.00
Monodelphis domestica (Opossum)	wt	OP	GGCCAGGACGCAGUAUUACUCUCGGGUC	9.19
Gallus gallus (Chicken)	putative	CH	GCCGAGCCUGAGUUUAUACAGUCCUUUGUGC	10.07
Anolis carolinensis (Lizard)	putative	LI	GGUGUGUUUAUUUCAGGAGACA	7.28

Corresponding abbreviations used are given. wt = wildtype; MW = molecular weight.

Electrophoretic mobility shift assays (EMSA)

For gel shift assays RNAs were 5'-fluorescently labelled [26]. First, RNAs were dephosphorylated using Quick CIP (NEB) at 37°C for 1.5 h. After phenol/chloroform extraction and ethanol precipitation, RNAs were phosphorylated using ATPγS (Cayman Chemical) and T4 PNK (NEB) at 37°C overnight. Again, RNAs were phenol/chloroform extracted, precipitated and resuspended in 25 mm HEPES pH 7.4. RNAs were labelled at room temperature for 4 h in the dark by adding 5'-(iodoacetamido)fluorescein (IAF) (ThermoFisher) dissolved in DMSO to a final concentration of 7.5 mm. Labelled RNAs were purified via denaturing PAGE as described above and refolded in 20 mm Tris pH 7.0, 150 mm NaCl, 2 mm TCEP as before. For EMSAs, 2 μl RNA were incubated for 10 min at room temperature with 0.6 μg baker's yeast $tRNA_{Phe}$ (Roche), 1 mm $MgCl_2$ and increasing amounts of protein in a volume of 20 μl. Protein concentrations of 0, 100, 200, 400, 700, 1,000, 2,000, 3,000, 5,000, 20000 nM were used. Samples were mixed with loading buffer (1× TB, 60% glycerol, 0.02% Bromphenolblue) and run on 6% native polyacrylamide gels for 1 h as described before [26]. Gels were imaged on a ChemiDoc Imager (Bio-Rad). For analysis, ImageJ was used to quantify the free RNA. Fraction bound was calculated as before [26]. K_D values are given as average plus standard deviation from individually fit triplicates using Hill1 fit in Origin.

Results

### Prediction of an Ox40 ADE in multiple species

The *Ox40* 3'UTR provides posttranscriptional control over a critical T-cell costimulatory receptor and is thus a key regulator within the mammalian immune system. Regulation occurs through Roquin binding to either or both ADE and CDE, and we have previously shown that Roquin recognizes the two *cis* elements independently [26]. As typical for CDEs, the *Ox40* CDE shows a high degree of conservation among mammals (Figure 1A) with a near 100% sequence identity among apes and monkeys as assessed by a sequence alignment of the tandem ADE-CDE cassette. For all other species, only positions 1 and 2 within the loop showed variation as well as the 5' nucleotide of the closing base pair. However, this position is always a pyrimidine and allows the formation of a C-G or U-G base pair, conserving the overall hairpin fold with a triloop. All CDEs are predicted to adopt a six-base pair stem. In contrast, the ADE is prone to larger sequence variations, which is reflected in a larger score within the phylogenetic tree for the individual ADE compared to the CDE (Supplementary Figure S1B and C). As for the CDE, conservation within the ADE stem is high among apes and monkeys, but low for other species. Based on our sequence alignment, ape ADEs have an insertion in 5' of the loop and show small deletions and sequence variations in 3' of the loop compared to other mammals (Figure 1A). Interestingly, the ADE loop is more conserved than the CDE loop with only position four switching between C and U. Oft note, the high degree of conservation of other known ADEs (*NFKBID* [16], *]ICOS* [27] suggests that these genes might have been targets for

Roquin control earlier in evolution (Supplementary Figure S2).

The high degree of CDE conservation suggests that it developed earlier and that ADE evolution was coupled to advancements of the immune system. In line with this all species from the alignment showed a high sequence conservation for Roquin exons (Supplementary Figure S3), pointing at possible Roquin–CDE interactions throughout all mammals. This is supported by high sequence conservation of known Roquin target 3'UTRs, i.e. *NFKBIZ* and *NFKBID* (Supplementary Figure S3). Of note, chicken and lizard sequences aligned well with mammalian Roquin and target gene exons, while the corresponding 3'UTRs showed poor conservation. Assuming a similar immune-regulating function of Roquin in birds and reptiles, this could imply regulation of target genes by other *cis* element structures than the known CDE or ADE, like linear binding elements (LBEs) [16] or AREs [47]. Roquin binding could also occur through conservation of a CDE/ADE fold despite high sequence variation. This could reveal possible progenitor ADEs and highlight structural features that were selected by Roquin during evolution. Predictions of mammalian ADEs (Supplementary Figure S1D) suggested the formation of stable hairpin structures.

Ox40 ADEs adopt a stable fold

We designed ADE constructs of four mammals (Table 1, Figure 1B) based on the sequence alignment to assess critical structural ADE features and included the murine ADE from an earlier study [26] to cover ADEs from multiple animal orders. The high degree of conservation suggests (Figure 1A and Supplementary Figure S1B) that these are indeed functional ADEs in their respective species. We added an ADE-like sequence from an early mammal, i.e. the opossum, which showed poor conservation in the *Ox40* 3'UTR (Supplementary Figure S3). Additionally, we included analogous RNA regions as potential ADE ancestors in birds and reptiles to analyse their fold and stability. Therefore, we designed two putative ADEs from the aligned *Ox40* 3'UTR of chicken and lizard. Defined sequences were guided by the mammalian ADEs, i.e. expected to form hairpin structures and contain a U-rich loop of a size of 6–10 nucleotides. All ADE constructs were designed to be similar in size to avoid charge-driven bias in RNP interaction studies and to start with a G to facilitate *in vitro* transcription (Table 1). To assess the degree of secondary structure, we recorded ^1H NMR imino spectra of the RNAs at three temperatures (Figure 2A and Supplementary Figure S4A). The putative chicken and lizard ADEs showed only a few and broad imino proton signals, suggesting no stable and/or single conformation. Mammalian ADEs on the contrary, showed dispersed peaks, indicative of a stable fold. Interestingly, peaks are more dispersed and uniform in linewidth for dolphin, mouse and human than for the hedgehog or elephant. This points at potential structural changes within the mammalian development that led to a further stabilization of the ADE fold. We confirmed the secondary structure of the human ADE by NMR spectroscopy (Supplementary Figure S4B). For all ADEs, only minor changes in the imino proton pattern were
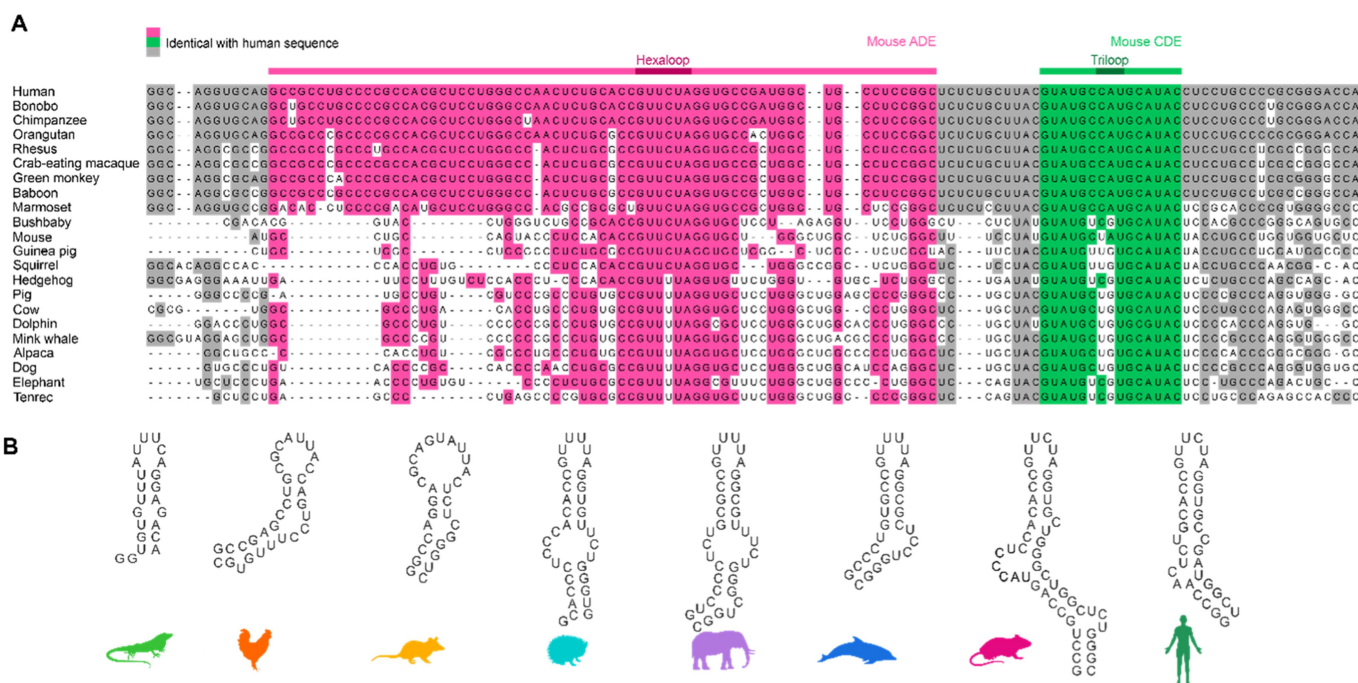
Figure 1. Evolutionary conservation of *Ox40* tandem ADE-CDE cassette. A) Sequence alignment of *Ox40* tandem ADE-CDE from multiple mammals. Bases identical with the human sequence are coloured. Residues corresponding to the murine ADE and CDE are coloured in magenta and green, respectively. Dark magenta and green highlight the hexa- and triloop, respectively. B) Secondary structure predictions of ADEs used in this study. These ADE constructs were designed to be comparable and small in size, to contain the central bulge and the basal stem and start with a stable GC or GU basepair. Note that these constructs were derived from Supplementary Figure S1D, where the extended ADE context is shown.
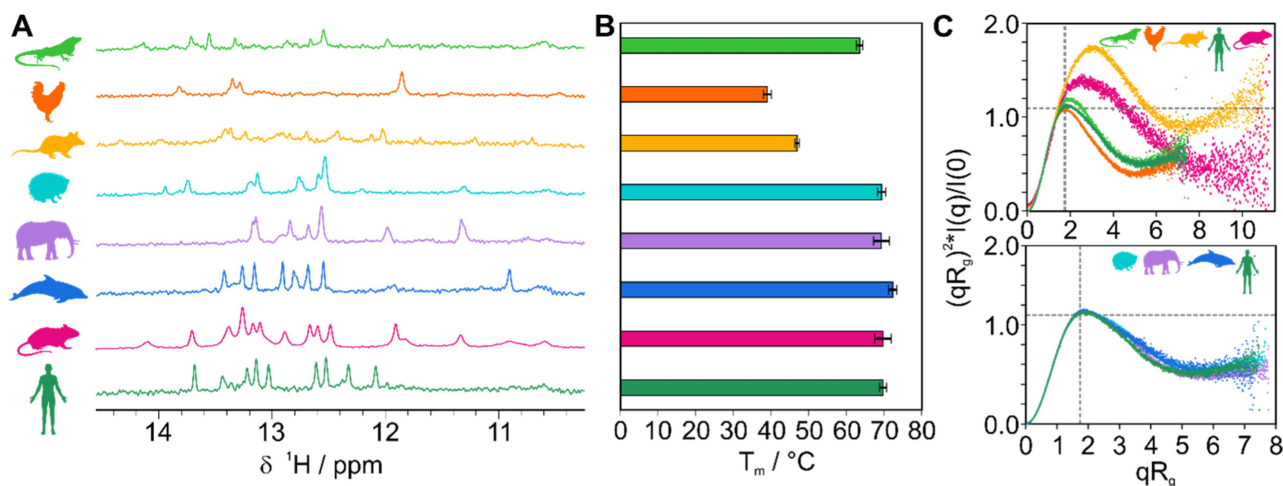
Figure 2. *Ox40* ADEs from multiple species form stable structures. A) Imino ^{1}H spectra of *Ox40* species ADEs (from top to bottom: lizard, chicken, opossum, hedgehog, elephant, dolphin, mouse, and human). B) CD-derived melting points of ADEs aligned and colour coded as in panel A). Melting points are given as average from triplicates with standard deviations (error bars). C) Dimensionless Kratky plots derived from SAXS measurements of ADEs. A curve maximum at the intersection of the two grey dotted lines corresponds to a globular (compact) fold. Curves with a shift to the upper right indicate partially unfolded structures (e.g. as observed for opossum) or elongated shapes (e.g. as for the murine ADE due to its increased construct and stem size).

observed across the tested temperatures (Supplementary Figure S4A), suggesting formation of stable structures. To corroborate our observations, we measured CD melting curves (Supplementary Figure S5). Mammalian ADEs yielded melting temperatures of 69.3–72.3°C (Figure 2B and Supplementary Table S1), in line with the predicted low ΔG values (Supplementary Figure S1D). For chicken and lizard, we observed melting temperatures of 39.0°C and 63.5°C, respectively. The lowered melting temperatures can be accounted for by an overall increased AU-content of the RNAs compared to the mammalian ADEs or increased loop-size (chicken) (Supplementary Figure S1D). The increased thermal stability of mammalian ADEs agrees well with our observation of secondary structure formation in NMR experiments. Despite sequence and structure variations, the mammalian *Ox40* ADEs are comparably stable, suggesting the melting temperature to be optimal for gene regulation of *Ox40*, e.g. through an RNA fold-recognizing protein.

We next used small-angle X-ray scattering (SAXS) to probe the three-dimensional shape of the ADEs (Figure 2C and Supplementary Table S2) [48]. Dimensionless Kratky plots allow to assess the fold of an RNA independent of its molecular weight [49,50]. A bell-shaped curve maximum at the so-called Guinier-Kratky-point (i.e. the intersection of the two dashed lines in Figure 2C) indicates a globular fold, while shifts in the maximum to the upper right point at elongated or partially unstructured molecules. For hedgehog, dolphin, elephant and human we observe the curve maximum close to the intersection, pointing at a globular and similar fold between these four ADEs (Figure 2C bottom panel). This agrees well with SAXS-based tertiary structure models (Supplementary Figure S6C) that show a high degree of similarity. In contrast, the putative chicken and lizard ADE curves show a visible deviation from the mammalian ones (Figure 2C top panel). This suggests a more opened conformation for the lizard, in line with weak and broad imino proton signals in our NMR spectra. The chicken ADE shows a more bell-shaped curve, which can be explained by the formation of a partial duplex stem and the large loop, which probably samples a large conformational space. Our NMR spectra indeed confirm the existence of only four stable base pairs. This renders the overall shape more globular than the slightly extended form of the mammalian ADEs (Supplementary Figure S1D and Supplementary Figure S6). Our results show

that the *Ox40* ADE forms a stable moiety in mammals and that this fold is robust across a broad range of temperature, which allows formation of hairpin structures at physiological conditions. Analyses of reptile and bird RNAs suggested that ADEs could have evolved from less stable and only transiently structured hairpins.

Roquin developed a high affinity for mammalian ADEs

We wondered whether Roquin (Supplementary Figure S7A) binds all species ADEs and how subtle differences in RNA sequence and shape would affect binding. Therefore, we performed gel shift assays (EMSA) of ADEs with murine/human coreROQ and extROQ (Figure 3 and Supplementary Figure S7B). For mammalian ADEs both protein constructs yielded discrete complex bands, confirming RNP formation (Figure 3A). For lizard and chicken ADEs, micromolar protein concentrations caused a smear in the gel, which suggests formation of a less stable complex that partially dissociates within the gel matrix. However, for extROQ, binding occurred at lower protein concentrations (lizard) or led to formation of complex bands (chicken), albeit less sharp than for mammalian ADEs. Quantification of EMSAs (Figure 3B,C and Table 2) confirmed that Roquin binds mammalian ADEs with nanomolar affinity, while putative avian and reptilian ADEs are bound with ~10× lower affinity (micromolar).
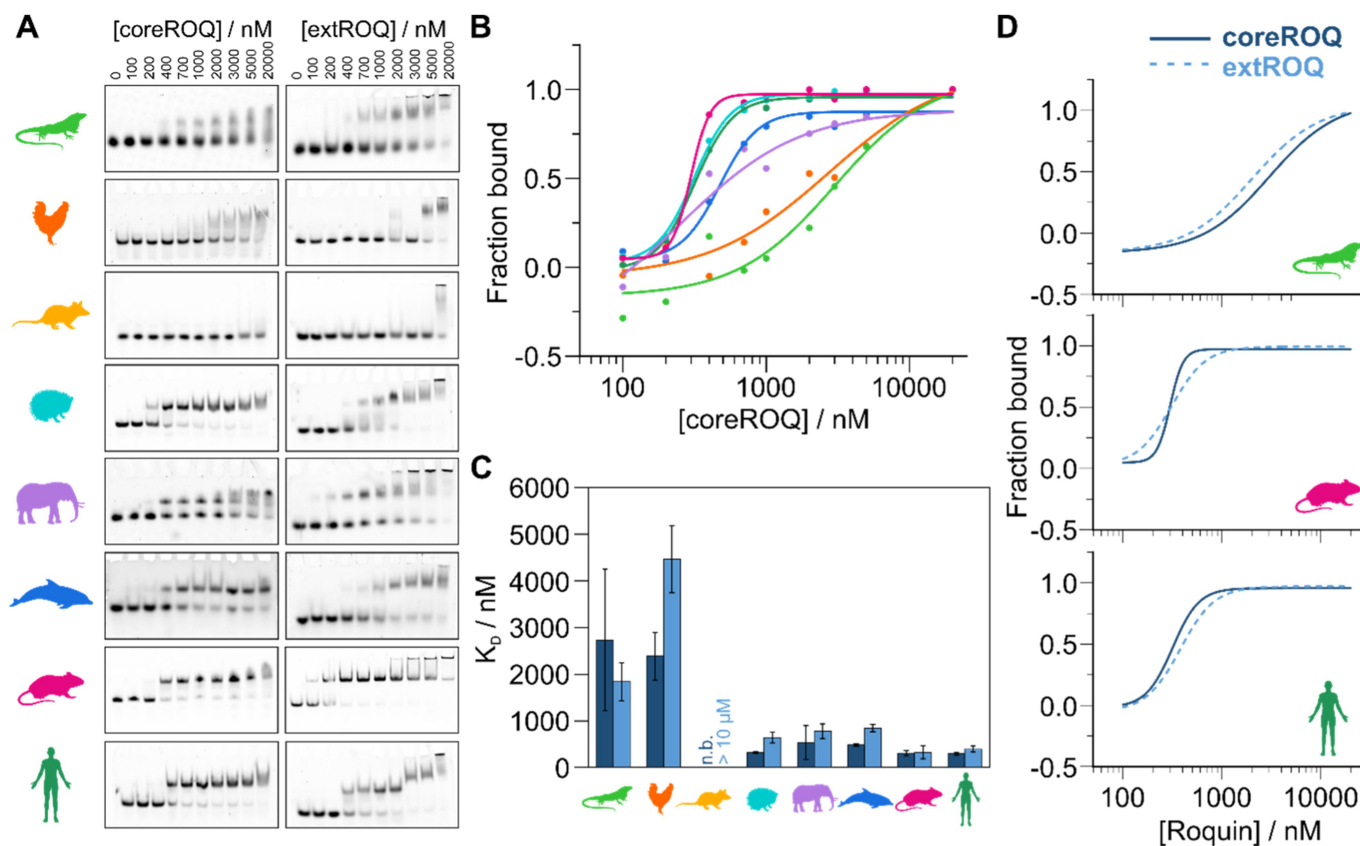

Figure 3. Affinity of Roquin for ADEs increased during evolution. A) Representative EMSAs of coreROQ and extROQ with species ADEs. Protein concentrations are given on top. Triplicates are shown in supplementary figure S6. B) CoreROQ binding curves derived from EMSAs. Plotted are averaged curves from triplicates. C) K_D values of coreROQ (dark blue) and extROQ (light blue) for species ADEs as obtained from EMSAs shown as bar plot. K_D values are averages from triplicates and errors are standard deviations (see Table 2). n.b. = no binding D) Comparison of coreROQ (dark blue, solid line) and extROQ (light blue, dashed line) binding curves with lizard, mouse and human ADEs.

Table 2. Affinities of coreROQ and extROQ for ADE RNAs.

ADE	K_D (coreROQ)/nM	K_D (extROQ)/nM
HS	295.3 ± 29.6	385.5 ± 63.7
HS$_{mirr}$	546.6 ± 15.5	769.5 ± 158.1
HS 3'bulge	292.8 ± 148.9	474.7 ± 62.6
HS low pur.	432.2 ± 62.6	672.6 ± 65.2
MM	307.8 ± 57.2	322.8 ± 138.7
d$_{stem1}$	381.2 ± 94.8	1041.4 ± 138.8
d$_{stem2}$	450.9 ± 46.9	977.7 ± 76.9
d$_{apical}$	n.b.	487.6 ± 373.2
Δbulge	299.1 ± 41.6	479.4 ± 188.3
MM 3'bulge	240.6 ± 15.6	525.4 ± 60.1
DK	479.9 ± 24.7	847.2 ± 77.4
EL	533.8 ± 365.5	778.4 ± 157.1
HE	318.4 ± 17.2	641.6 ± 118.8
OP	n.b.	>10 μM
CH	2387.9 ± 510.9	4463.5 ± 716.9
LI	2732.1 ± 1515.9	1841.0 ± 403.1

K_D values are averages of $n = 3$; errors are standard deviations.
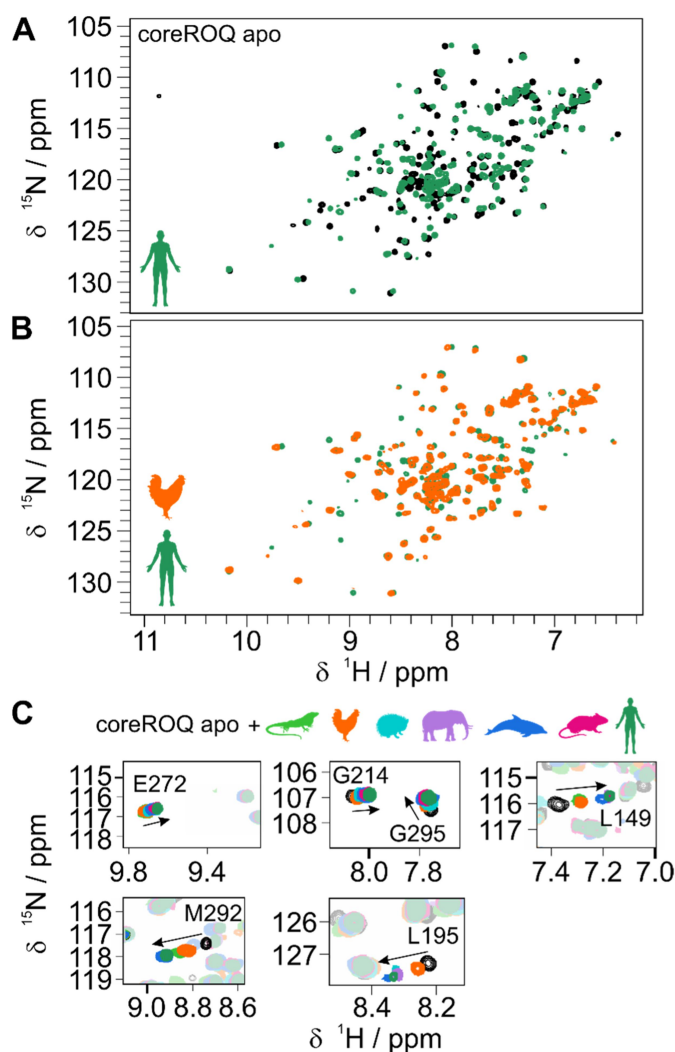
Figure 4. Evolution of the conserved Roquin ADE binding mode. A) ^{1}H,^{15}N-HSQC spectra of apo coreROQ (black) and in complex with the human *Ox40* ADE (green). B) ^{1}H,^{15}N-HSQC spectra of coreROQ in complex with the putative chicken (orange) or human (green) ADE. For all species comparisons refer to supplementary figure S7. C) Zoom-ins of ^{1}H,^{15}N-HSQC overlays of apo coreROQ (black) and in presence of species ADEs (colour).

Interestingly, elephant and dolphin ADEs showed an intermediate high-nanomolar affinity. These findings suggest that ADE binding at relevant protein concentrations only occurred

in mammals, but binding was further enhanced throughout development of individual mammalian species. Cognate *cis-trans* pairs often co-evolved to derive RNPs with high affinity or specificity. While the coreROQ domain is conserved among all species tested, the flanking B-site contains sequence variations (Supplementary Table S3). Of note, no species extROQ showed a significantly increased or decreased affinity for its cognate ADE compared to the human extROQ (Supplementary Figure S8A and B). We thus conclude that indeed changes in RNA sequence or structure were selected by Roquin to obtain a high affinity RNP.

A comparison of K_D values from both protein constructs (Figure 3C,D) revealed a higher affinity of coreROQ compared to extROQ for mammalian ADEs, except for the murine one, which bound both proteins equally well. The putative lizard ADE had a higher affinity for extROQ. This can point at an accessory function of the Roquin B-site in early ADE development, where the additional RNA binding interface could facilitate RNP formation through charges in lack of a stable hairpin.

The ADE binding mode changed during evolution

We next set out to determine when in evolution ADEs started to be bound in the known 'ADE-like' fashion [24]. To monitor changes in binding on the atomic level, we recorded ^{1}H,^{15}N-HSQC fingerprint spectra of the murine/human coreROQ protein alone (apo) and in complex with species ADEs (Figure 4 and Supplementary Figure S9). Including earlier work [24,26], the human and murine ADEs were found to show the same pattern of chemical shift perturbations (CSPs) (Figure 4A). Spectra of the mammalian ROQ-ADE complexes (Supplementary Figure S9A) show near-complete overlap with each other with only minor deviations in CSPs, suggesting a conserved binding mode. However, avian and reptilian ADEs show large differences in CSPs between each other and compared to mammals (Figure 4B and Supplementary Figure S9B). This correlates well with the observed differences between these two clades and mammals with respect to predicted structure, stability and affinity.

Tracking of individual amino acids in overlays of all species complexes allowed to monitor changes in CSPs along evolution: Several residues showed increased CSPs for mammalian ADEs compared to chicken and lizard ADEs (e.g. L149, M292; Figure 4C). Others, e.g. L195, exhibited stronger line broadening for chicken and lizard, which is often indicative of changes in affinity. Together, our NMR data capture and support the gradual changes in affinity throughout ADE evolution and recapitulate the formation of a conserved ADE binding mode.

Stem stability affects Roquin binding

From our CD and EMSA experiments, we conclude that the formation of a stable stem-loop structure was a key step in evolution to generate high-affinity interactions between RNA and Roquin. We hence tested murine ADE variants with destabilizing stem mutants for protein binding (Figure 5). Mutations of GC base pairs to AU base pairs in different

stem regions (Figure 5A) showed a shift in CD melting points to lower temperatures (Figure 5B), confirming the destabilizing effect of the mutations. Additionally, we deleted the central bulge (Figure 5A), which stabilized the ADE as observed in an increased melting point (Figure 5B). Gel-shift assays of the murine ADE variants with core and extROQ (Figure 5C and Supplementary Figure S10) showed formation of complex bands. However, no band shift was observed for coreROQ incubated with the ADE comprising destabilizing mutations in the apical stem (d_{apical}). This confirms that the apical stem-loop is the only binding site for coreROQ and that a strong reduction in stem stability severely perturbs RNP formation through the Roquin A-site. In extROQ the B-site partially compensates this effect as assessed by complex band formation. Quantification of EMSAs (Figure 5D) revealed that destabilization of the central stem regions (d_{stem1} and d_{stem2}) had only a minor effect on coreROQ binding. On the contrary, extROQ engagement with these ADE variants was significantly perturbed and led to a decrease in affinity of ~3× (Table 2). Interestingly, K_D values of both coreROQ and extROQ for the Δbulge ADE variant were almost not altered from the wildtype (Figure 5D). A slight difference for extROQ, which is within the error, could point at a structure-mediated effect recognized through the B-site. Despite this small effect, further stabilization of the wildtype ADE stability does neither enhance nor reduce the affinity of Roquin.

## RNA determinants required for high-affinity Roquin binding

Changes in stem stability do not explain the increase in affinity throughout mammalian development (Figure 3C), where ADEs are similar in both shape and stability

(Figure 2B,C). We hence examined features of sequence and structure that differ among mammals. Interestingly, an internal bulge with an asymmetric number of unpaired nucleotides evolved directly below the upper four-to-five base pair hairpin stem (Supplementary Figure S1D), which harbours the known coreROQ binding site. While elephant and dolphin possess a 3' bulge (Supplementary Figure S1D and Supplementary Table S4), hedgehog, human and mouse ADEs form a 5' bulge. We observed a positive correlation of affinity with the presence of a 5' bulge (Figure 6A) for extROQ. Binding affinity also correlated well with the size of the bulge. Interestingly, binding affinity of coreROQ to ADEs did not correlate with a 5' or 3' bulge, confirming no interaction of coreROQ with this region of the ADE hairpin (Figure 6A). We thus hypothesized that bulge inclusion in either stem half and the bulge size make up a structural feature that is read out by Roquin, in particular by the B-site. Potentially, a bulge could facilitate conformational rearrangements of the RNA to allow proper B-site engagement by serving as an anchoring point. We tested our hypothesis with two technical ADE mutants: For the human and murine ADE, we swapped the bulge to the 3' stem (Supplementary Figure S11A) and analysed Roquin binding by EMSAs (Figure 6B and Supplementary Figure S11B). For the human ADE, we observed no change in affinity for coreROQ (Figure 6C), while the murine ADE 3' bulge variant even showed a small increase in affinity. Affinity of extROQ showed a small decrease for the human 3' bulge ADE compared to the wild-type ADE, while the reduction in affinity by a factor of ~1.6× was larger for the murine version (Figure 6C). We suggest that the difference in responsiveness between the two species is caused by the increased size of the bulge for the mouse (Figure 6A).
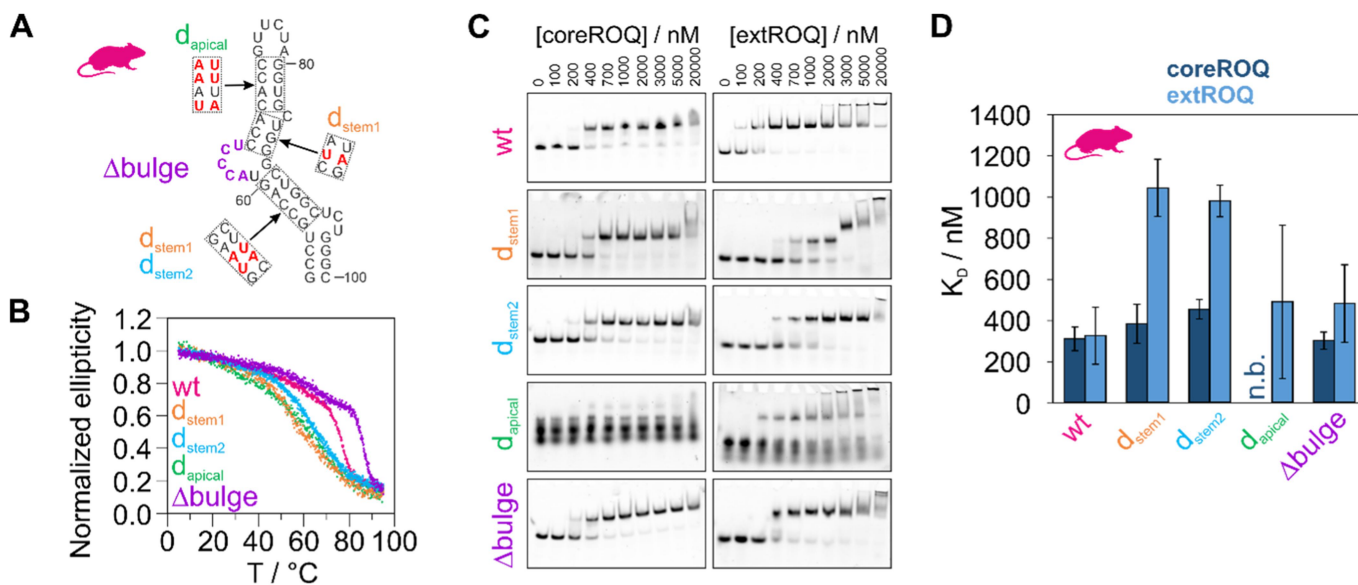
Figure 5. ADE stability mutants modulate Roquin binding affinity. A) Secondary structure of murine *Ox40* ADE wt. Red nucleotides in dashed boxes indicate sites of mutations. The corresponding constructs, in which these mutations occur are denoted besides the boxes. Nucleotides deleted in a Δbulge version are coloured in purple. B) Normalized CD melting curves of ADE variants shown in A). C) Representative EMSAs of coreROQ and extROQ with stability variants of the murine *Ox40* ADE from A). Protein concentrations are given on top. Triplicates are shown in supplementary figure S8. D) K_D values of coreROQ (dark blue) and extROQ (light blue) for ADE stability variants as obtained from EMSAs shown as bar plot. K_D values are averages from triplicates and errors are standard deviations (see Table 2). n.b. = no binding.
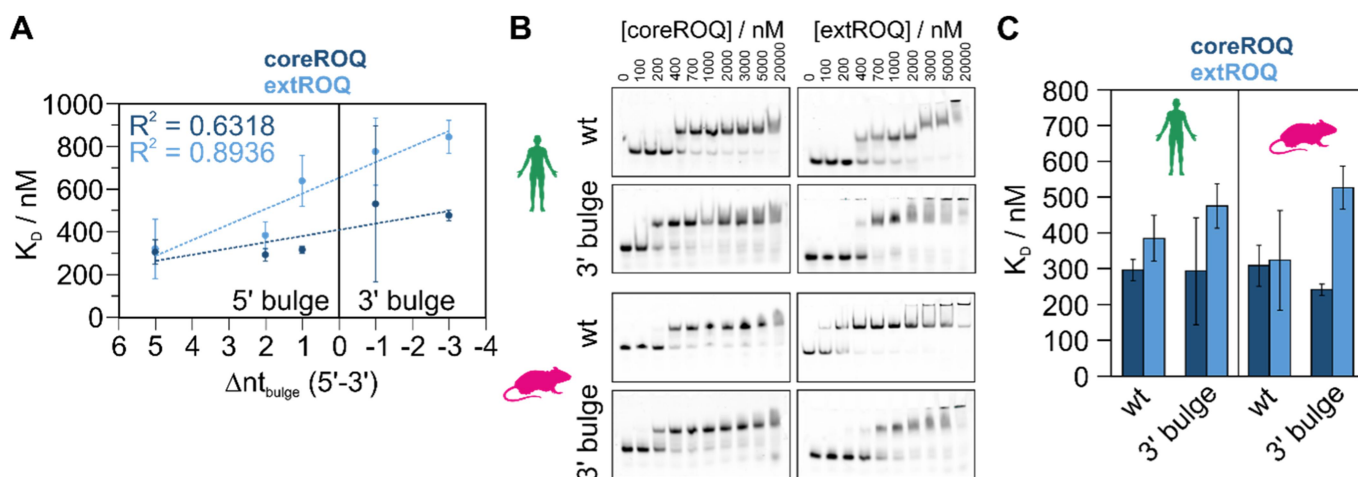
Figure 6. The central ADE bulge affects Roquin binding affinity. A) Correlation plot of coreROQ (dark blue) and extROQ (light blue) affinities for mammalian ADEs (taken from Figure 3C) with the number of unpaired nucleotides in the central bulge of the respective ADE (see supplementary table 3). Note that the difference between 5' and 3' unpaired nucleotides is plotted. Positive values thus indicate a bulge within the 5' stem and negative values within the 3' stem. B) Representative EMSAs of coreROQ and extROQ with human and murine wt ADEs and a version with a 3' bulge (see Supplementary Figure S9). Protein concentrations are given on top. Triplicates are shown in Supplementary Figure S9. Wt EMSAs are the same as in Figure 3A. C) K_D values of coreROQ (dark blue) and extROQ (light blue) for wt and 3' bulge ADEs as obtained from EMSAs shown as bar plot. K_D values are averages from triplicates and errors are standard deviations (see Table 2).

Roquin reads out stem geometry

A close inspection of the species ADE sequences revealed a purine content in the 3' stem of 46.2% to 57.1% (Supplementary Table S5). Binding affinity for both coreROQ and extROQ showed a good correlation with the 3' purine content, albeit a stronger dependency was observed for extROQ binding (Figure 7A). To probe for purine sensitivity, we lowered the 3' purine content of the human ADE by swapping base pairs (Figure 7B) from 53.8% to 23.1% (Supplementary Table S5). In EMSAs (Figure 7C) a low 3' purine content led to a reduction in affinity by 1.5-fold for coreROQ and 1.7-fold for extROQ (Figure 7D and Table 2). Our data thus highlight a potential role of the central stem bulge and the purine content of the 3' stem for ADE complex formation with Roquin. As our alterations in stability and RNA geometry are built on sequence changes, we assessed the impact of sequence variations on binding affinity in EMSAs with a human and murine ADE, where bp were shifted within duplex stretches and loop/bulge bases were altered (Supplementary Figure S12A): Roquin bound both sequence-varied ADEs with the same affinity as the respective wildtype ADE (Supplementary Figure S12B and C), confirming that RNA geometry is the dominant factor for complex formation. Based on our determined rules for optimal ADE geometry, we designed an artificial ADE (Supplementary Figure S12B, C), which showed a slight increase in affinity for extROQ compared to the human ADE, thereby serving as a proof-of-principle for our geometry centred approach.

At this stage, our experiments with swapped bulges, lower purine content and changes in stem stability showed detectable yet moderate effects on binding affinity. We next speculated that the precise interplay of loop sequence, geometry and stability renders a high-affinity Roquin target. We thus created a mirrored version of the human ADE (Figure 7B) where only the loop and basal stem orientation was preserved

and confirmed its structural integrity and the mirrored shape by NMR spectroscopy (compare Supplementary Figure S13A and Supplementary Figure S4B). EMSAs confirmed formation of stable complexes (Figure 7C and Supplementary Figure S13B), but with K_D values of 546 and 769 nM for core and extROQ, respectively (compared to 295 and 385 nM for the human wildtype ADE; Table 2). The affinity reduction for the mirrored ADE strongly indicates that Roquin recognizes a defined geometry and requires a precise interplay of loop sequence orientation and stem geometry (Figure 7D). However, NMR spectra confirmed a similar binding mode for both wildtype and mirrored human ADEs (Supplementary Figure S13C). This suggests that Roquin forces the suboptimal geometry of the mirrored ADE into a preferred fold, resulting in the lowered affinity.

Discussion

CDEs are abundant RNA *cis* elements specifically recognized by Roquin to control transcript levels of target genes. Shape-recognition assures highly specific complex formation and enables effective posttranscriptional control. ADEs are much less abundant in 3'UTRs than CDEs and structural information about ADEs and their engagement with Roquin is limited. High-resolution structures of a SELEX-derived hexaloop-stem and the *Ox40* ADE [24,26] confirmed a preference for a GUUUUA/GUUCUA loop sequence by Roquin, leading to comparable affinities as for CDEs (171 nM for the *Ox40* ADE, 149 nM for the *Ox40* CDE) [26]. For *NFKBID* a functionally relevant ADE has been confirmed *in vivo* [16]. Braun et al. predicted further 19 ADEs in human genes, of which three were found in proximity to a CDE [28]. Folding probabilities of these ADEs varied significantly questioning an ADE-like fold. However, these predicted ADEs [28] as well as available experimental data on the *Ox40* ADE [26] suggested a large structural variety within this class of RNA *cis* elements.
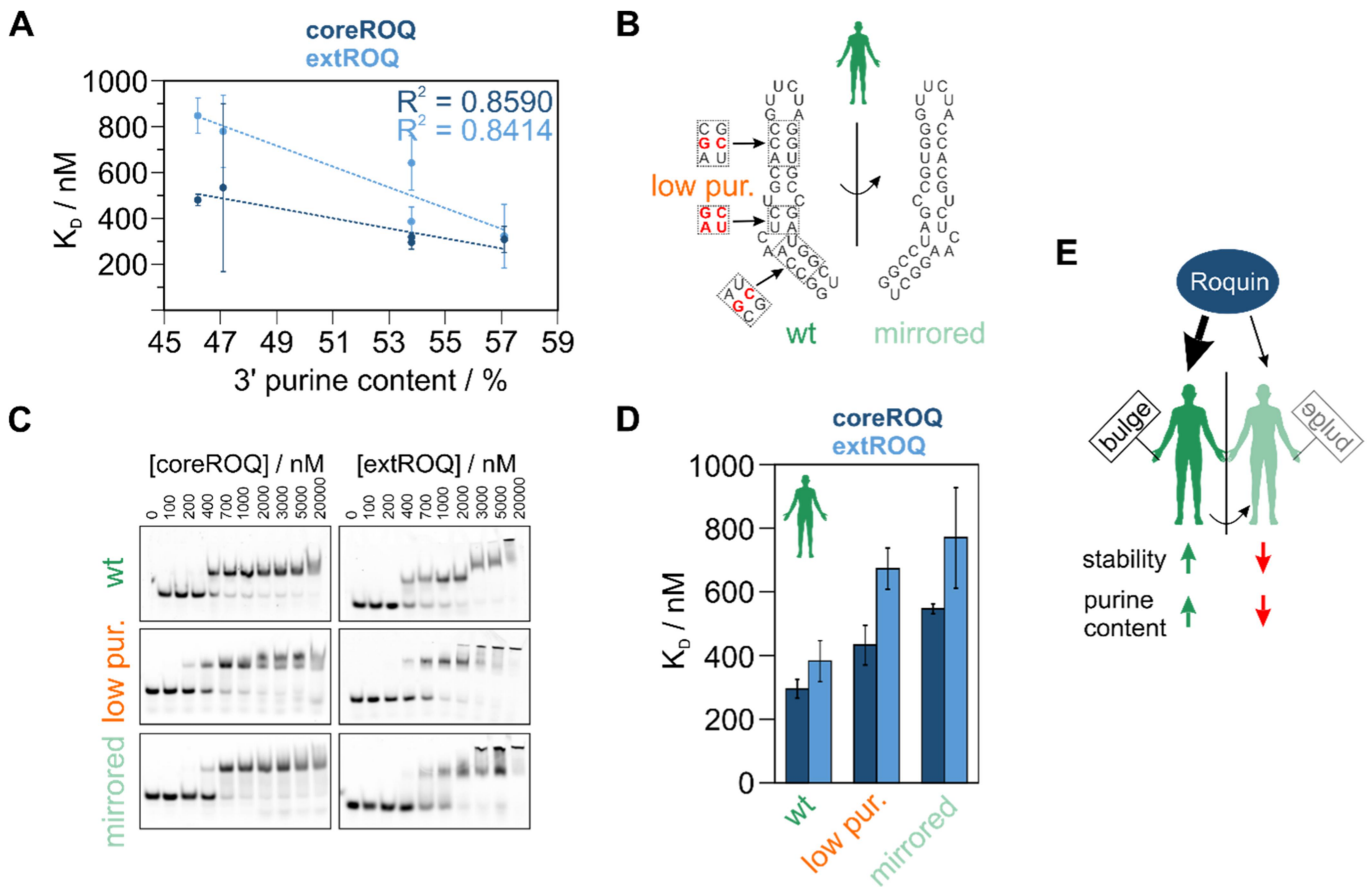
Figure 7. The impact of ADE geometry and purine content on Roquin binding affinity. A) Correlation of coreROQ and extROQ affinity with the 3' purine content of ADEs. B) Secondary structure of human *Ox40* ADE wt and a mirrored version. Note that the central stem is mirrored, while the basal stem and the loop orientation are maintained with respect to the wildtype. Red nucleotides in dashed boxes indicate sites of mutations within an ADE mutant with low purine content in the 3' stem. C) Representative EMSAs of coreROQ and extROQ with human ADE variants from A). Protein concentrations are given on top. Triplicates are shown in supplementary figure 10. Wt EMSAs are the same as in Figure 3A. D) K_D values of coreROQ (dark blue) and extROQ (light blue) for human ADE variants as obtained from EMSAs shown as bar plot. K_D values are averages from triplicates and errors are standard deviations (see Table 2). E) Roquin shows a binding preference for ADEs with a bulge in the 5' stem. Increased stability and a high purine content in the 3' stem increase affinity. Overall, the interplay of geometric factors, stability and sequence fine tunes Roquin-ade complex formation.

Our comparative analysis of *Ox40* ADEs from different species contributes to empirically filling the gap in sparse ADE data. We showed that mammalian ADEs are similar in size, shape, and stability and are bound by Roquin with nanomolar affinity, suggesting positive selection of RNA parameters during evolution by the protein. A comparison to bird and reptile analoga suggested that an increase in thermal stability and formation of a stable hairpin structure led to increased binding affinity and was a critical event in the evolution of ADEs. Despite the missing annotation of the herein presented chicken and lizard ADEs, our findings illustrate how ADEs (and other structured *cis* elements) could have evolved from AU-rich, i.e. rather unstructured RNAs. Roquin binding to these putative ADEs in the low micromolar range is comparable to Roquin ZnF domain binding to AREs [13]. As Roquin is highly conserved across vertebrates, this might point at an increasing relevance of the ROQ domain in more highly developed immune systems, while at early time points in evolution gene regulation could have relied on (low-affinity) AU-rich elements. This agrees well with the observed induction of RNA structure by Roquin in AU-rich, unstructured RNAs in *UCP* that can fold into CDEs [47].

Interestingly, in line with this, we observed that the Roquin B-site had a greater contribution to binding of these progenitor ADEs. We conclude that the B-site, which was shown to interact with dsRNA [25], i.e. rather charge driven and sequence unspecific, could have selected for RNAs that form (transient) hairpin structures until stable RNA folds evolved for specific engagement with the coreROQ domain.

Our data expands the target spectrum of Roquin. Different from the known consensus loop motif, we show that substitutions at loop positions 1 and 4 are tolerated, and Roquin can even cope with increased loop sizes: The octaloop from chicken could be recognized in a hexaloop like fashion with a rather weak A–A closing base pair. This agrees with Roquin binding to an octaloop in the *ICOS* 3'UTR with low micro-molar affinity [27]. However, a poly(U) loop background needs to be embedded in a hairpin context, as we observed no binding to the large and flexible opossum ADE loop. ADEs could thus have evolved from AU-rich, unstable elements and initially might have included larger loops.

While RNA elements are often compared across different genes, our comparison of ADEs from the same gene but of different species allows to evaluate the contribution of subtle

changes in sequence and structure to complex formation. In line with previous observations [51], we showed that high-affinity Roquin binding requires formation of a stable stem. We gave evidence that a central bulge 4–5 bp basal to the loop affected Roquin binding, especially via the B-site. Bulge integration into the 5' ADE stem combined with an increase purine content in the 3' stem enhanced interaction of the B-site with the central stem. While an increased 3' purine content had been suggested before, not all predicted ADEs possess such a purine stack [28]. 3' purine stacks were shown for CDEs to widen the RNA major groove and thereby facilitate Roquin engagement with the 5' stem side [51]. This causes optimal prearrangement of the RNA structure and facilitates complex formation compared to hairpins without purine stacks. Hence, purine rich 3' stems provide yet another layer of fine-tuning of Roquin RNPs required for a subset of targets only and are especially recognized by the Roquin B-site. In line with this, our data on *Ox40* species ADEs and predicted ADEs showed that a central bulge is no absolute requirement for Roquin binding, despite its positive effect on affinity. Together, we confirmed that the Roquin B-site prefers duplex RNA regions [25], but it is sensitive to geometric deviations from a perfect duplex and can readout geometric parameters, e.g. mirrored shapes. While the coreROQ domain has a well-defined target consensus for a stable apical stem-loop, which our data corroborate, the B-site is more promiscuous with respect to sequence and structure. We conclude that the precise interplay of Roquin A and B-site integrates high specificity for stem-loops with a platform to readout RNA stem stability and geometry (Figure 7E). For a high-affinity interaction, the concerted recognition of these defined apical and central stem elements is crucial.

Altogether, our study provides insight into the evolution of ADE stem-loop features from potential progenitor ADEs, which differ significantly from the known ADE consensus forming stable hairpin structures. This development of RNA *cis* element structure is accompanied by an increase in Roquin binding affinity. The interaction of Roquin with ADEs might not have existed early in evolution but could have rather evolved as a second RNP besides the Roquin-CDE complex. The evolutionary snapshots in our study provide us detailed information on key steps of the evolution of RNA determinants for an optimized, highly adapted RNP in coincidence with the development of the adaptive immune system.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Data availability statement

All constructs, NMR spectra and SAXS scattering curves will be provided upon reasonable request.

ORCID

Jan-Niklas Tants http://orcid.org/0000-0002-2455-8135
Andreas Schlundt http://orcid.org/0000-0003-2254-7560

References

[1] Shine J, Dalgarno L. The 3′-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA. 1974;71 (4):1342–1346. doi: 10.1073/pnas.71.4.1342

[2] Shaw G, Kamen R. A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell. 1986;46(5):659–667. doi: 10.1016/0092-8674(86)90341-7

[3] Caput D, Beutler B, Hartog K, et al. Identification of a common nucleotide sequence in the 3′-untranslated region of mRNA molecules specifying inflammatory mediators. Proc Natl Acad Sci USA. 1986;83(6):1670–1674. doi: 10.1073/pnas.83.6.1670

[4] Bao C, Zhu M, Nykonchuk I, et al. Specific length and structure rather than high thermodynamic stability enable regulatory mRNA stem-loops to pause translation. Nat Commun. 2022;13 (1):988. doi: 10.1038/s41467-022-28600-5

[5] Costantino DA, Pfingsten JS, Rambo RP, et al. tRNA–mRNA mimicry drives translation initiation from a viral IRES. Nat Struct Mol Biol. 2008;15(1):57–64. doi: 10.1038/nsmb1351

[6] Keane SC, Heng X, Lu K, et al. RNA structure. Structure of the HIV-1 RNA packaging signal. Science. 2015;348(6237):917–921. doi: 10.1126/science.aaa9266

[7] Klovins J, Berzins V, van Duin J. A long-range interaction in Qβ RNA that bridges the thousand nucleotides between the M-site and the 3′ end is required for replication. RNA. 1998;4 (8):948–957. doi: 10.1017/S1355838298980177

[8] Li D, Lu HT, Ding YZ, et al. Specialized cis-acting RNA elements balance genome cyclization to ensure efficient replication of yellow fever virus. J Virol. 2023;97(4):e0194922. doi: 10.1128/jvi.01949-22

[9] Weinreb C, Riesselman AJ, Ingraham JB, et al. 3D RNA and Functional interactions from evolutionary couplings. Cell. 2016;165(4):963–975. doi: 10.1016/j.cell.2016.03.030

[10] Ye X, Yang W, Yi S, et al. Two distinct binding modes provide the RNA-binding protein RbFox with extraordinary sequence specificity. Nat Commun. 2023;14(1):701. doi: 10.1038/s41467-023-36394-3

[11] Sanchez de Groot N, Armaos A, Grana-Montes R, et al. RNA structure drives interaction with proteins. Nat Commun. 2019;10 (1):3246. doi: 10.1038/s41467-019-10923-5

[12] Loughlin FE, Lukavsky PJ, Kazeeva T, et al. The solution structure of FUS bound to RNA reveals a bipartite mode of RNA recognition with both sequence and shape specificity. Mol Cell. 2019;73 (3):490–504.e6. doi: 10.1016/j.molcel.2018.11.012

[13] Tants JN, Oberstrass L, Weigand JE, et al. Structure and RNA-binding of the helically extended Roquin ccch-type zinc finger. Nucleic Acids Res. 2024;52(16):9838–9853. doi: 10.1093/nar/gkae555

[14] Schneider T, Hung LH, Aziz M, et al. Combinatorial recognition of clustered RNA elements by the multidomain RNA-binding protein IMP3. Nat Commun. 2019;10(1):2266. doi: 10.1038/s41467-019-09769-8

[15] Kristjansdottir K, Fogarty EA, Grimson A. Systematic analysis of the Hmga2 3′ UTR identifies many independent regulatory sequences and a novel interaction between distal sites. RNA. 2015;21(7):1346–1360. doi: 10.1261/rna.051177.115

[16] Essig K, Kronbeck N, Guimaraes JC, et al. Roquin targets mRNAs in a 3′-utr-specific manner by different modes of regulation. Nat Commun. 2018;9(1):3810. doi: 10.1038/s41467-018-06184-3

[17] Halvorsen M, Martin JS, Broadaway S, et al. Disease-associated mutations that alter the RNA structural ensemble. PLOS Genet. 2010;6(8):e1001074. doi: 10.1371/journal.pgen.1001074

[18] Gupta A, Gribskov M. The role of RNA sequence and structure in RNA–protein interactions. J Mol Biol. 2011;409(4):574–587. doi: 10.1016/j.jmb.2011.04.007

[19] Allain FH, Howe PW, Neuhaus D, et al. Structural basis of the RNA-binding specificity of human U1A protein. Embo J. 1997;16 (18):5764–5772. doi: 10.1093/emboj/16.18.5764

[20] Tants JN, Fesser S, Kern T, et al. Molecular basis for asymmetry sensing of siRNAs by the drosophila loqs-PD/Dcr-2 complex in RNA interference. Nucleic Acids Res. 2017;45(21):12536–12550. doi: 10.1093/nar/gkx886

[21] Masliah G, Maris C, Konig SL, et al. Structural basis of siRNA recognition by TRBP double-stranded RNA binding domains. Embo J. 2018;37(6). doi: 10.15252/embj.201797089

[22] Schlundt A, Heinz GA, Janowski R, et al. Structural basis for RNA recognition in roquin-mediated post-transcriptional gene regulation. Nat Struct Mol Biol. 2014;21(8):671–678. doi: 10.1038/nsmb.2855

[23] Leppek K, Schott J, Reitter S, et al. Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. Cell. 2013;153(4):869–881. doi: 10.1016/j.cell.2013.04.016

[24] Janowski R, Heinz GA, Schlundt A, et al. Roquin recognizes a non-canonical hexaloop structure in the 3′-UTR of Ox40. Nat Commun. 2016;7(1):11032. doi: 10.1038/ncomms11032

[25] Tan D, Zhou M, Kiledjian M, et al. The ROQ domain of roquin recognizes mRNA constitutive-decay element and double-stranded RNA. Nat Struct Mol Biol. 2014;21(8):679–685. doi: 10.1038/nsmb.2857

[26] Tants JN, Becker LM, McNicoll F, et al. Nmr-derived secondary structure of the full-length Ox40 mRNA 3′UTR and its multi-valent binding to the immunoregulatory RBP Roquin. Nucleic Acids Res. 2022;50(7):4083–4099. doi: 10.1093/nar/gkac212

[27] Rehage N, Davydova E, Conrad C, et al. Binding of NUFIP2 to Roquin promotes recognition and regulation of ICOS mRNA. Nat Commun. 2018;9(1):299. doi: 10.1038/s41467-017-02582-1

[28] Braun J, Fischer S, Xu ZZ, et al. Identification of new high affinity targets for Roquin based on structural conservation. Nucleic Acids Res. 2018;46(22):12109–12125. doi: 10.1093/nar/gky908

[29] Francisco-Velilla R, Embarc-Buh A, Rangel-Guerrero S, et al. RNA-protein coevolution study of Gemin5 uncovers the role of the PXSS motif of RBS1 domain for RNA binding. RNA Biol. 2020;17(9):1331–1341. doi: 10.1080/15476286.2020.1762054

[30] Brandman R, Brandman Y, Pande VS, et al. Sequence coevolution between RNA and protein characterized by mutual information between residue triplets. PLOS ONE. 2012;7(1):e30022. doi: 10.1371/journal.pone.0030022

[31] Dunstan MS, Guhathakurta D, Draper DE, et al. Coevolution of protein and RNA structures within a highly conserved ribosomal domain. Chem Biol. 2005;12(2):201–206. doi: 10.1016/j.chembiol.2004.11.019

[32] Gruber AR, Bernhart SH, Hofacker IL, et al. Strategies for measuring evolutionary conservation of RNA secondary structures.

BMC Bioinformatics. 2008;9(1):122. doi: 10.1186/1471-2105-9-122

[33] Ariza-Mateos A, Briones C, Perales C, et al. Archaeological approaches to RNA virus evolution. J Physiol. 2023;602 (11):2469–2478. doi: 10.1113/JP284416

[34] Rivas E, Schlick T. RNA structure prediction using positive and negative evolutionary information. PLOS Comput Biol. 2020;16 (10):e1008387. doi: 10.1371/journal.pcbi.1008387

[35] Chao JA, Patskovsky Y, Almo SC, et al. Structural basis for the coevolution of a viral RNA–protein complex. Nat Struct Mol Biol. 2008;15(1):103–105. doi: 10.1038/nsmb1327

[36] Rigby CV, Sabsay KR, Bisht K, et al. Evolution of transient RNA structure–RNA polymerase interactions in respiratory RNA virus genomes. Virus Evol. 2023;9(2):vead056. doi: 10.1093/ve/vead056

[37] Smyth RP, Negroni M, Lever AM, et al. RNA structure—A neglected puppet master for the evolution of virus and host immunity. Front Immunol. 2018;9:2097. doi: 10.3389/fimmu.2018.02097

[38] Schlundt A, Niessing D, Heissmeyer V, et al. RNA recognition by Roquin in posttranscriptional gene regulation. Wiley Interdiscip Rev RNA. 2016;7(4):455–469. doi: 10.1002/wrna.1333

[39] Nassar LR, Barber GP, Benet-Pages A, et al. The UCSC genome browser database: 2023 update. Nucleic Acids Res. 2023;51(D1): D1188–D1195. doi: 10.1093/nar/gkac1072

[40] Madeira F, Madhusoodanan N, Lee J, et al. The EMBL-EBI job dispatcher sequence analysis tools framework in 2024. Nucleic Acids Res. 2024;52(W1):W521–W5. doi: 10.1093/nar/gkae241

[41] Gruber AR, Lorenz R, Bernhart SH, et al. The Vienna RNA websuite. Nucleic Acids Res. 2008;36(Web Server):W70–4. doi: 10.1093/nar/gkn188

[42] Lorenz R, Bernhart SH, Honer Zu Siederdissen C, et al. ViennaRNA package 2.0. Algorithms Mol Biol. 2011;6(1):26. doi: 10.1186/1748-7188-6-26

[43] Lee W, Tonelli M, Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics. 2015;31(8):1325–1327. doi: 10.1093/bioinformatics/btu830

[44] Manalastas-Cantos K, Konarev PV, Hajizadeh NR, et al. ATSAS 3.0: expanded functionality and new tools for small-angle scattering data analysis. J Appl Crystallogr. 2021;54(1):343–355. doi: 10.1107/S1600576720013412

[45] Hopkins JB. BioXTAS RAW 2: new developments for a free open-source program for small-angle scattering data reduction and analysis. J Appl Crystallogr. 2024;57(1):194–208. doi: 10.1107/S1600576723011019

[46] Chojnowski G, Zaborowski R, Magnus M, et al. RNA 3D structure modeling by fragment assembly with small-angle X-ray scattering restraints. Bioinformatics. 2023;39(9):39. doi: 10.1093/bioinformatics/btad527

[47] Binas O, Tants JN, Peter SA, et al. Structural basis for the recognition of transiently structured au-rich elements by Roquin. Nucleic Acids Res. 2020;48:7385–7403. doi: 10.1093/nar/gkaa465

[48] Tants JN, Schlundt A. Advances, applications, and perspectives in small-angle X-ray scattering of RNA. Chembiochem. 2023;24(17): e202300110. doi: 10.1002/cbic.202300110

[49] Burger VM, Arenas DJ, Stultz CM. A structure-free method for quantifying conformational flexibility in proteins. Sci Rep. 2016;6 (1):29040. doi: 10.1038/srep29040

[50] Kikhney AG, Svergun DI. A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins. FEBS Lett. 2015;589(19PartA):2570–2577. doi: 10.1016/j.febslet.2015.08.027

[51] Codutti L, Leppek K, Zalesak J, et al. A distinct, sequence-induced conformation is required for recognition of the constitutive decay element RNA by Roquin. Structure. 2015;23(8):1437–1447. doi: 10.1016/j.str.2015.06.001