

Structural basis for RNA recognition in roquin-mediated post-transcriptional gene regulation

Andreas Schlundt^{1,2,6}, Gitta A Heinz^{3,4,6}, Robert Janowski^{5,6}, Arie Geerlof¹, Ralf Stehle², Vigo Heissmeyer^{3,4}, Dierk Niessing⁵ & Michael Sattler^{1,2}

Roquin function in T cells is essential for the prevention of autoimmune disease. Roquin interacts with the 3' untranslated regions (UTRs) of co-stimulatory receptors and controls T-cell activation and differentiation. Here we show that the N-terminal ROQ domain from mouse roquin adopts an extended winged-helix (WH) fold, which is sufficient for binding to the constitutive decay element (CDE) in the *Tnf* 3' UTR. The crystal structure of the ROQ domain in complex with a prototypical CDE RNA stem-loop reveals tight recognition of the RNA stem and its trilob. Surprisingly, roquin uses mainly non-sequence-specific contacts to the RNA, thus suggesting a relaxed CDE consensus and implicating a broader spectrum of target mRNAs than previously anticipated. Consistently with this, NMR and binding experiments with CDE-like stem-loops together with cell-based assays confirm roquin-dependent regulation of relaxed CDE consensus motifs in natural 3' UTRs.

The post-transcriptional control of gene regulation is essential for a range of cellular functions and is linked to human disease^{1,2}. Regulation of mRNA stability and translation mediates fast changes of protein expression and involves *trans*-acting protein factors, which recognize specific *cis* elements in the UTRs of target mRNAs. The molecular requirements for targeting elements and binding factors are often poorly understood.

Roquin is an essential factor in the prevention of autoimmune disease³. Both mouse and human roquin are found in two isoforms, roquin-1 and roquin-2, of high sequence identity. The *Rc3h1* gene encoding the roquin-1 protein has been identified in an *N*-ethyl-*N*-nitrosourea (ENU) mouse mutagenesis screen, in which one point mutation changed Met199 to arginine. Homozygous mice of this so-called *sanroque* strain (*Rc3h1*^{san/san}) develop high-affinity autoantibodies and a pathology that resembles the human autoimmune disease systemic lupus erythematosus³. The phenotype is related to T cell-intrinsic causes because the mice spontaneously develop activated T cells and accumulate follicular helper T (T_{FH}) cells. T_{FH} cells in *sanroque* mice provide inappropriate B-cell help so that high-affinity autoantibodies against self-antigens are produced^{3,4}. In addition, *sanroque* mice show increased susceptibility to autoimmune diabetes^{3,5} and autoantibody-induced arthritis⁶ as well as angioimmunoblastic T-cell lymphomas⁷. Surprisingly, deletion of roquin-1 did not manifest autoimmune phenotypes⁸. This apparent paradox was explained by the ability of roquin-2 to exert redundant functions⁹. Roquin-2 may therefore compensate when roquin-1 is missing but not when it is present as the *sanroque* M199R mutant^{6,9}.

Regulation of T-cell activity by roquin has been correlated to its mRNA binding and downregulation of the expression of the inducible co-stimulator (ICOS) and another co-stimulatory receptor, Ox40, encoded by the *Tnfrsf4* gene (here denoted *Ox40*)^{9–12}. The minimal RNA-binding domain of roquin is currently unknown. More than half of the 125-kDa roquin protein is predicted to be intrinsically disordered. The N-terminal region of roquin is highly conserved and comprises a RING finger with a potential E3 ubiquitin-ligase function¹³ and a conserved but not well-defined ROQ domain (residues 64–400) as well as a CCCH-type zinc finger that is potentially involved in RNA recognition. It has been suggested that the ROQ domain mediates binding to the ICOS mRNA^{10,11}. However, the ROQ domain has been predicted solely on the basis of sequence conservation, and RNA binding analysis did not exclude the involvement of additional regions^{3,10,11}. Recently, it was reported that an N-terminal fragment of roquin-1 (residues 131 to 360) mediates binding to a region in the 3' UTR of the *Tnf* mRNA comprising a 23-mer CDE¹⁴. However, structural details of the roquin-CDE interaction are not known, and the sequence requirements for functional decay elements are poorly understood.

Here, we present structural, biochemical and functional analyses of the roquin-1 RNA-binding domain and its interaction with CDEs. Our results unravel structural requirements for recognition of CDE RNA by the ROQ domain. The data indicate that CDE motifs with confined variations of the consensus sequence are functional and occur in natural targets of roquin-mediated post-transcriptional gene regulation.

¹Biomolecular NMR Spectroscopy, Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg, Germany. ²Center for Integrated Protein Science Munich at Biomolecular NMR Spectroscopy, Department Chemie, Technische Universität München, Garching, Germany. ³Institute of Molecular Immunology, Research Unit Molecular Immune Regulation, Helmholtz Zentrum München, Germany. ⁴Institute for Immunology, Ludwig-Maximilians-Universität München, Munich, Germany. ⁵Group Intracellular Transport and RNA Biology, Institute of Structural Biology, Helmholtz Zentrum München, Neuherberg, Germany. ⁶These authors contributed equally to this work. Correspondence should be addressed to V.H. (vigo.heissmeyer@helmholtz-muenchen.de), D.N. (niessing@helmholtz-muenchen.de) or M.S. (sattler@helmholtz-muenchen.de).

Received 14 April; accepted 3 June; published online 13 July 2014; doi:10.1038/nsmb.2855

RESULTS

The roquin-1 RNA-binding domain and its target RNA

To define the boundaries of the roquin-1 RNA-binding domain, we performed limited proteolysis of the longest predicted ROQ domain of roquin-1 (amino acids 64 to 411) (Fig. 1a). NMR spectra of a stable protein fragment identified the presence of a globular domain (Supplementary Fig. 1a,b) and superimposed well with a highly conserved roquin-1 fragment spanning amino acids 147–326 (Supplementary Fig. 1c,d).

Roquin binds with nanomolar affinity to a 23-mer oligonucleotide corresponding to a CDE RNA element of the *Tnf* 3' UTR (*Tnf*CDE)¹⁴. We confirmed the secondary structure of the *Tnf* CDE, predicted to form a hairpin with a triloop, by a two-dimensional imino NOESY NMR spectrum. We observed base pairs with sequential imino-imino NOE connectivities for only the six bases preceding the triloop. The lack of observable imino NOEs for U4 and G20 suggests that this previously predicted U-G base pair¹⁴ is not formed or dynamic (Fig. 1b,c).

We quantitatively assessed binding of full-length roquin-1 (roquin-1 fl), the N-terminal region of roquin-1 (N-term) or the identified ROQ fragment alone to the *Tnf* CDE by using radioactive electrophoretic mobility shift assays (EMSAs) (Fig. 1d). Roquin-1 fl and the N-terminal region bound to this RNA with equilibrium dissociation constants (K_d) of 145 nM and 240 nM, respectively (Fig. 1d,e). The ROQ fragment alone yielded a K_d of 119 nM. We thus conclude that this fragment comprises the functional ROQ domain and is necessary and sufficient for binding to CDE RNA. NMR spectra showed that the ROQ domain adopts a globular fold, although the N-terminal residues (147 to 170) are unstructured and highly flexible in solution (Supplementary Fig. 2a–e).

Structure of the roquin-1 ROQ domain

To understand the molecular details of the CDE RNA recognition by roquin-1, we solved the crystal structure of the ROQ domain (residues 147–326) to a resolution of 1.9 Å (Fig. 1f, Supplementary Fig. 3a,b and Table 1). Gel filtration combined with

static light scattering demonstrated that the ROQ domain and also the N-terminal region of roquin-1 are monomeric in solution (Supplementary Fig. 3c), even though we observed a dimeric arrangement in the asymmetric unit (Supplementary Fig. 3a). Lack of observable electron density for the N-terminal 18 residues suggests conformational flexibility, in agreement with the NMR data (Fig. 1f and Supplementary Fig. 2a,b). The globular fold comprises residues 165–326 and consists of seven α -helices and five β -strands (Fig. 1f,g). Following helix α 1–strand β 1, the region comprising α 2– β 2– α 3– α 4– β 3– β 4 adopts a canonical WH fold^{15,16} (Supplementary Fig. 3d). Line-broadening of NMR signals for residues in the W1 wing (Fig. 1f,g and Supplementary Fig. 2a–e) suggests conformational dynamics on a microsecond-to-millisecond time scale. This is consistent with structural variations seen for the β 3–W1– β 4 region in the two molecules of the asymmetric unit in the crystal structure (Supplementary Fig. 3e).

The region C terminal to the WH domain contains a three-helix bundle (α 5– α 7) (Fig. 1f,g), which contacts the N-terminal helix α 1 of the WH fold so that the N- and C-terminal regions of the ROQ domain are in spatial proximity. A structural-homology search failed to identify similar folds, thus demonstrating that the ROQ domain adopts a new RNA-binding fold (Supplementary Fig. 3f).

Structure of the ROQ–*Tnf* CDE RNA complex

We next crystallized the complex of the ROQ domain with the *Tnf* CDE RNA. Crystals obtained belonged to space group *P*1 and diffracted to 3.0-Å resolution (Table 1 and Supplementary Fig. 4a,b). Overall, the structure of the ROQ domain in the RNA complex is

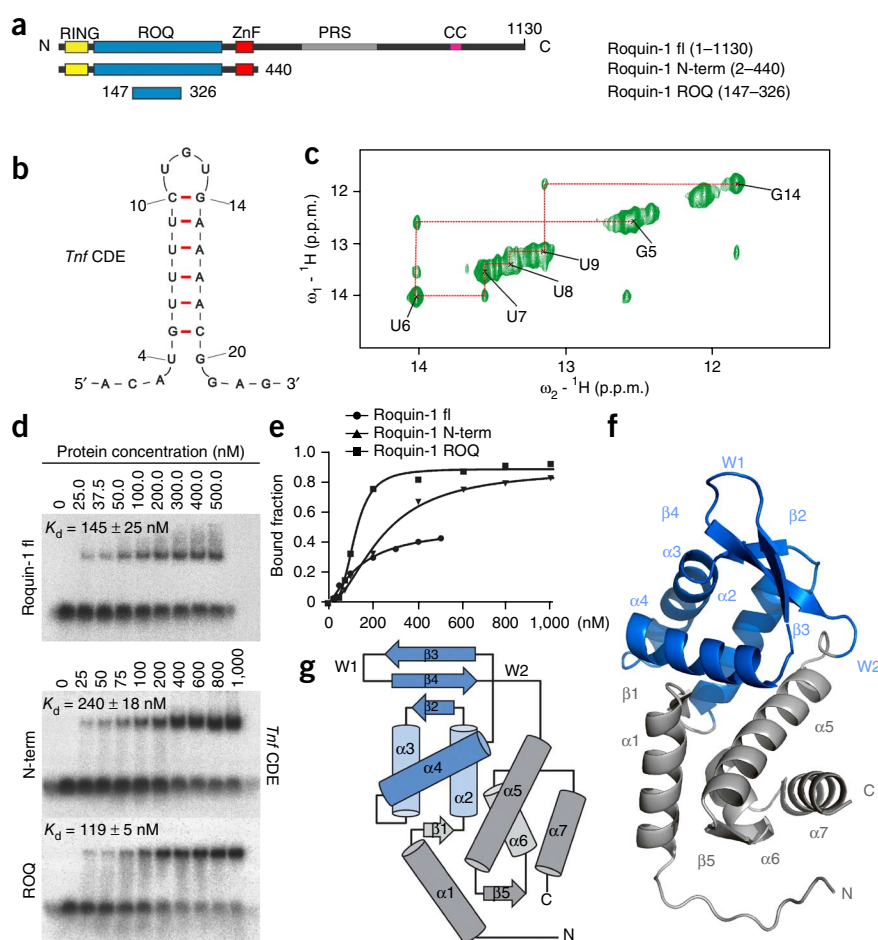


Figure 1 The ROQ domain is sufficient for CDE RNA binding. (a) Domain organization of mouse roquin-1 and the three length variations of roquin-1 used in the study: full-length (fl), N terminus (N-term) and ROQ domain, as defined from limited proteolysis. ZnF, zinc finger; PRS, proline-rich sequence; CC, coiled-coil region. (b) CDE stem-loop of mouse *Tnf* mRNA (5'-ACA UGU UUU CUG UGA AAA CGG AG-3'), as suggested by base-pairing observed in the spectrum in c. (c) 2D imino proton NOESY of the *Tnf* CDE element, with assigned G- and U-imino protons annotated. (d) EMSA experiments with the three length versions of roquin-1 shown in a and *Tnf* CDE RNA. (e) Quantification of the EMSA experiments. (f) Crystal structure of the roquin-1 ROQ domain (residues 147–326). α -helices, β -sheets and loops (W) of the WH motifs are numbered according to sequential order. The WH domain with the recognition helix α 4 is highlighted in blue. The two wings correspond to W1 (Arg260–Asp263), connecting the β 3 and β 4 strand, and W2 (Glu271–Thr275) linking β 4 and α 5. (g) Domain topology scheme of roquin-1 ROQ with color code and numbering as in f.

very similar to that of the free protein (r.m.s. deviation of 0.77 Å for 151 C α atoms of residues 175 to 325), with the exception of the W1 wing region, which adopts a more closed conformation upon RNA binding (Supplementary Note and Supplementary Fig. 3e). The *Tnf* CDE RNA folds into a hairpin structure with a unique U11-G12-U13 triloop and a helical stem comprising six Watson-Crick base pairs (Fig. 2a and Supplementary Fig. 4c,d). The first and last nucleotides with visible electron densities were A3, U4, G20 and G21. We found U4-G20 and A3-G21 base-pairing in only a subset of the four molecules of the asymmetric unit, and NMR experiments did not indicate formation of base pairs between U4 and G20 or between A3 and G21 in the unbound RNA (Fig. 1b,c). Together these data suggest that A3 and G21 as well as U4 and G20 do not form stable base pairs.

The ROQ domain recognizes the *Tnf* CDE RNA mainly through non-sequence-specific contacts with the phosphodiester backbone or RNA bases. The W1 region of the ROQ domain interacts with the UGU triloop, and the WH fold provides numerous electrostatic contacts with the *Tnf* CDE RNA stem (Fig. 2b and Supplementary Fig. 4d). The UGU triloop is clamped by contacts involving residues in the recognition helix α 4 (Gln247, Tyr250, Arg251 and Ser253), helix α 3 (Arg219) and the β 3-W1- β 4 wing (Asp263, Ser264 and Ser265) characteristic of the WH fold (Fig. 2a-e). The first and the third residue of the triloop (U11 and U13) are flipped out and show numerous interactions with the ROQ domain (Fig. 2a-e and

Supplementary Fig. 4b). Recognition of U11 is non-sequence specific, but for steric reasons only a pyrimidine base is compatible at this position (Fig. 2c). The purine base of G12 stacks with the G14 purine base of the closing C10-G14 base pair in the CDE stem (Fig. 2a,d and Supplementary Fig. 4c,d) and is locked in this position by stacking with Arg219 from the opposite side (Fig. 2d). The G12 phosphate contacts the Tyr250 side chain. This provides a rigid and tight coordination that seems optimal for a purine base at position 12. With the exception of Ser253, the exposed U13 base does not exhibit sequence-specific interactions with the CDE RNA (Fig. 2e). The space available at the position of U13 suggests that any nucleotide could be allowed at this position.

Almost all phosphate groups in the 5' half of the CDE RNA stem are recognized by residues in helices α 1 (Arg188), α 3 (Lys220) and α 4 (Ser238, Lys239 and Thr240) of the ROQ domain (Fig. 2b,f). In contrast, we observed no contacts with the 3' stem region (Fig. 2b). In two of the four molecules in the asymmetric unit of the crystal, Trp184, Arg188 and Phe194 interact with a flipped-out U4 base (Fig. 2g and Supplementary Fig. 4a), for which we also observed no base-pairing by NMR in the free RNA (Fig. 1b,c). This indicates that single-stranded sequences flanking the CDE stem-loop may contribute to roquin binding.

The recognition of the *Tnf* CDE RNA by the ROQ domain is distinct from any other WH-RNA interaction reported previously (Supplementary Fig. 4e-i). We therefore conclude that the overall arrangement and specific recognition of the CDE hairpin is unique and enables the specific recognition of CDE RNAs by roquin-1. To characterize the ROQ-RNA interaction in solution, we performed NMR titrations with the 23-mer *Tnf* CDE RNA. Large chemical-shift changes upon RNA binding mapped to the ROQ region comprising the recognition helix, the β 3-W1- β 4 wing and additional structural elements (Fig. 3a-c). Interestingly, in the ROQ-RNA complex we did not observe the line-broadening of NMR signals present in the free protein (Fig. 3a,b); this suggests that the W1 wing, which is dynamic in the free protein, becomes ordered by RNA contacts. NMR relaxation data (Supplementary Fig. 2c-e) and small-angle X-ray scattering (SAXS) measurements (Supplementary Note and Supplementary Fig. 4j-l) indicated nonspecific dimerization of the ROQ domain at higher concentrations. In contrast to those for the free ROQ domain, SAXS data unequivocally demonstrated that ROQ-RNA forms a defined 1:1 complex. The shape of the ROQ-RNA complex in solution is fully consistent with the crystal structure, as indicated by the excellent agreement of the experimental SAXS data with the theoretical scattering curve (Supplementary Fig. 4j-l and Supplementary Table 1). Chemical-shift perturbations (CSPs) observed upon RNA binding (Fig. 3a-c) are consistent with the ROQ-RNA binding interface seen in the crystal structure (Fig. 2). Notably, CSPs for Phe194 and Leu195 (Fig. 3a,b) support the interaction of U4 with the ROQ domain by aromatic stacking with the Trp184 side chain (Fig. 2g).

Comparison of imino NOESY NMR spectra of the free and ROQ-bound *Tnf* CDE RNA (Supplementary Fig. 5a,b) suggests that the stem-loop undergoes an induced fit upon protein binding. For example, the large up-field shift for the NMR signal of the G14 imino proton in the bound RNA is consistent with ring-current shift induced by stacking with the closing C-G base pair. Also, we observed more than the expected six imino signals for the free *Tnf* CDE RNA, thus suggesting conformational heterogeneity. A single set of imino signals was present for the six expected base pairs when they were bound to ROQ. This suggests that the ROQ domain may induce or select the specific conformation of the RNA stem-loop observed in the ROQ-RNA complex.

Table 1 Data collection and refinement statistics

	Roquin-1 ROQ [Ta ₆ Br ₁₂] ²⁺ × 2 Br ⁻ derivative	Roquin-1 ROQ	Roquin-1 ROQ- <i>Tnf</i> CDE RNA
Data collection			
Space group	C222 ₁	C222 ₁	P1
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	55.34, 78.90, 184.28	55.22, 79.51, 184.91	56.56, 60.41, 84.37
α , β , γ (°)	90, 90, 90	90, 90, 90	105.68, 101.36, 95.72
Resolution (Å)	100–2.5 (2.57–2.50)	50–1.94 (1.99–1.94)	50–3.0 (3.08–3.00)
<i>R</i> _{merge}	9.3 (31.0)	13.7 (62.1)	7.0 (51.3)
<i>I</i> / σ <i>I</i>	14.9 (4.1)	9.7 (2.1)	13.9 (3.0)
Completeness (%)	99.7 (96.7)	99.8 (99.9)	95.7 (96.9)
Redundancy	6.6 (3.9)	6.4 (5.6)	3.6 (3.6)
Refinement			
Resolution (Å)		1.94	3.00
No. reflections		30,555	19,951
<i>R</i> _{work} / <i>R</i> _{free}		16.3 / 20.6	20.3 / 24.7
No. atoms			
Protein		2,642	4,895
Ligand/ion		8	1,619
Water		323	32
<i>B</i> factors			
Protein		15.7	78.5
Ligand/ion		24.9	84.4
Water		31.9	51.0
r.m.s. deviations			
Bond lengths (Å)		0.019	0.011
Bond angles (°)		1.88	1.50

For each data set, only one crystal was used. Values in parentheses are for highest-resolution shell.

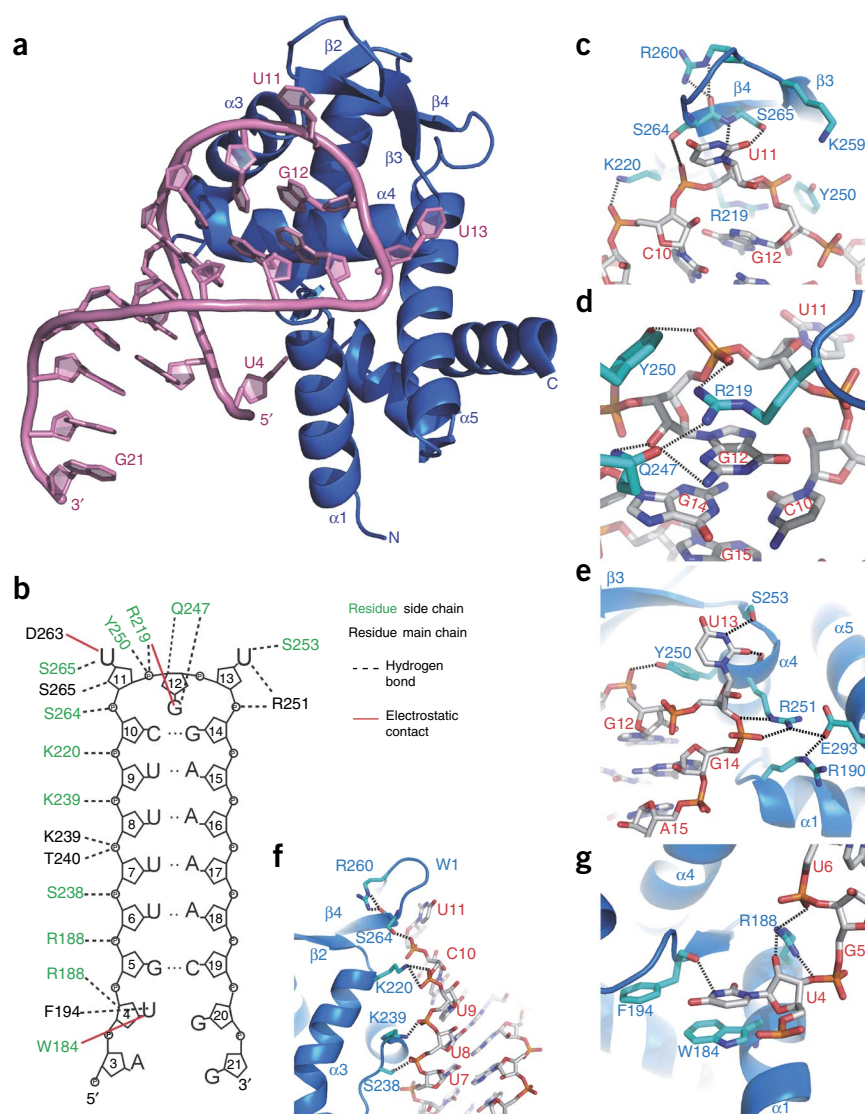
Figure 2 Structure of the roquin-1 ROQ–*Tnf* CDE RNA complex. (a) Cartoon presentation of the crystal structure of the ROQ domain (residues 174–325) (blue) and the *Tnf* CDE stem-loop RNA (magenta). Selected RNA bases are labeled. (b) Schematic drawing of the protein–RNA interactions and RNA–RNA base-pairing observed in the ROQ–RNA complex structure. Different types of contacts are indicated by color and line style as shown. (c) Close-up view of the contacts between the ROQ domain and nucleotides C10 and U11 in the CDE RNA triloop. U11 interacts with the W1 region of ROQ (Supplementary Fig. 4b). Protein backbone and side chains are shown in blue and turquoise, respectively, and the RNA is shown in gray. Atoms are color-coded according to charge. (d) Close-up view of the contacts between the ROQ domain and G12 in the triloop. G12 and U13 contact the C-terminal end of helix $\alpha 4$. (e) Close-up view of the contacts between the ROQ domain and U13. (f) Close-up view of the contacts between the ROQ domain and the phosphates in the CDE RNA stem from U8 to U11. (g) Close-up view of the contacts between the ROQ domain and U4. In c–e, amino acids are labeled in blue and nucleotides in red.

Mutational analysis of the ROQ–*Tnf* CDE RNA interaction

To examine the importance of individual ROQ–CDE RNA contacts for binding, we probed the protein–RNA interactions by mutational analysis and monitored complex formation in EMSA experiments. We designed amino acid changes for residues that mediate crucial contacts seen in the crystal structure (Fig. 2b) and exhibit large NMR CSPs upon RNA binding (Fig. 3a,b and Supplementary Fig. 1d). We observed considerably weaker binding of the ROQ domain upon mutation of conserved lysine residues (K220A, K239A and K259A) (Fig. 3d–f and Supplementary Fig. 5c). Attenuated complex formation was also present for the S238A, Y250A, S264A and S265Y mutants. We also probed the effect of an R260A mutation, on the basis of the large CSP seen in the NMR titration. This mutant ROQ domain showed almost complete loss of RNA binding in EMSAs (Fig. 3d–f). Inspection of the structure revealed that the side chain of Arg260 contacts the main chain of Ser264, the side chain of which forms a hydrogen bond with U11 (Fig. 2c,f). This indicates that RNA recognition by the ROQ W1 wing is associated with rigidification and compacting of the protein–RNA complex, assisted by Arg260-mediated indirect interactions. These observations are in line with the NMR relaxation data of the complex (Supplementary Fig. 2f,g).

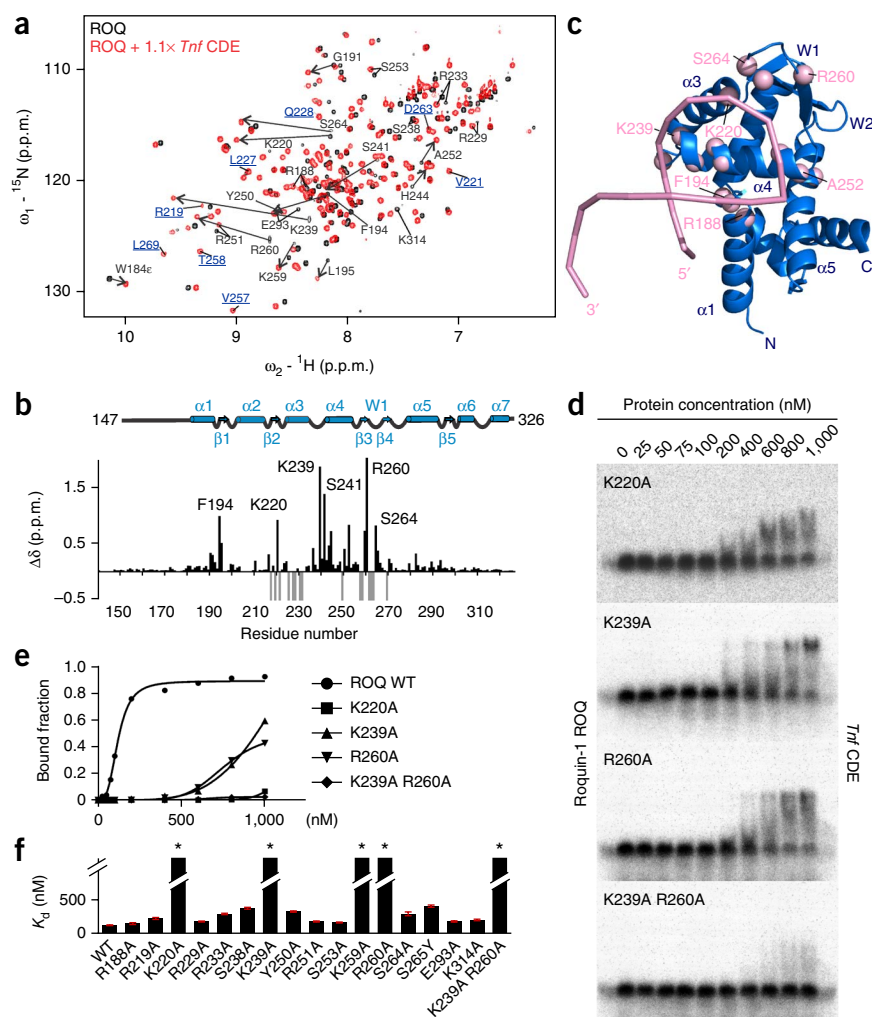
Target-mRNA regulation in cells requires roquin-1 binding

Having identified residues critical for *in vitro* RNA binding, we analyzed the importance of protein–RNA contacts for roquin-1 function in cells. We established comparable expression of the mutant and wild-type proteins after retroviral transduction of mouse embryonic fibroblasts (Fig. 4a). To analyze regulation of the physiological mRNA targets *Ox40* and *ICOS*^{3,9,12}, we first infected roquin-deficient mouse embryonic fibroblasts with retroviruses encoding full-length



ICOS or *Ox40*. We then superinfected the cells with retroviruses that bicistronically expressed roquin-1 and the marker Thy1.1 (Fig. 4b–d and Supplementary Fig. 6a–c). Using flow cytometry, we found that wild-type roquin-1 effectively downregulated *ICOS* or *Ox40* surface expression in infected cells, which we identified by Thy1.1 surface expression. Individual roquin-1 K220A, K239A or R260A mutations revealed a partial impairment and combined K239A R260A mutations an almost complete impairment of *Ox40* or *ICOS* repression in Thy1.1-positive cells (Fig. 4b,c and Supplementary Fig. 6a,b). We also analyzed the *Tnf* CDE *cis* element by fusing the coding sequence of *ICOS* as a reporter gene to an artificial 3' UTR of 260 nt (CDE₂₆₀) from the mouse *Tnf* 3' UTR. This region contained the roquin-1-recognized CDE but lacked the AU-rich elements of the *Tnf* 3' UTR¹⁴. Again, the *ICOS* reporter expression under the control of the CDE₂₆₀ was effectively repressed by wild-type roquin-1. The single mutants (roquin-1 K220A, K239A and R260A) with strongly reduced affinity for the *Tnf* CDE in EMSAs (Fig. 3d–f) were still able to repress the *ICOS* CDE₂₆₀ reporter (Fig. 4d and Supplementary Fig. 6c). Moreover, the roquin-1 K239A R260A double mutant was functionally impaired but nevertheless reduced *ICOS* surface expression by 40%. Together these findings demonstrate that the extent to which roquin-1 is able to interact with its target mRNAs is directly

Figure 3 Mutational analysis of the ROQ–*Tnf* CDE RNA interface. **(a)** Overlay of ^1H – ^{15}N HSQC spectra of the roquin-1 ROQ domain, both free (black) and bound to *Tnf* CDE RNA (red). Selected residues are labeled, and their chemical-shift changes upon RNA addition are indicated by arrows. Blue underlined residues are amide signals that are line-broadened in the absence of RNA. **(b)** Amide chemical-shift perturbations (CSP, $\Delta\delta$) upon CDE RNA binding. The secondary structure of the ROQ domain is shown at top. Gaps in the plot result from incomplete assignments or prolines. Residues that could be assigned unambiguously only in one of the spectra because of overlap or line-broadening are indicated with a gray bar with CSP values set to -0.5 . The most strongly affected residues are indicated. **(c)** Mapping of the CSPs onto the crystal structure of the ROQ–RNA complex. Amides with the largest CSPs are shown as pink spheres, and residues are annotated according to **a** and **b**. The RNA backbone is shown as semitransparent ribbon in pink. **(d)** EMSA experiments with the ROQ mutants K220A, K239A, R260A and K239A R260A. **(e)** Binding curves of the ROQ-domain mutants in **d** and wild-type ROQ as derived from EMSA quantification. **(f)** Bar plot comparing $K_d \pm$ s.e. for ROQ wild type and mutants binding to *Tnf* CDE RNA. Asterisks indicate extrapolated K_d values $>1,000$ nM when exact determination was not applicable from EMSAs. All values and details are given in **Supplementary Table 2**. S.e. values represent nonlinear least-square-regression deviations from fitting of binding curves to ten individual data points.



correlated to its activity in post-transcriptional target regulation. Consistently with this, the roquin-1 mutant with the strongest reduction in RNA binding was least repressive in cells.

Assessing the CDE consensus for ROQ-domain recognition

We next studied the contribution of individual nucleotides in the *Tnf* CDE RNA element to complex formation. Analysis of the ROQ–CDE RNA structure indicated that CDE recognition mainly involves non-sequence-specific interactions with the phosphate backbone and steric restrictions for interactions with the RNA bases in the triloop (Fig. 2). We used EMSA experiments to examine the ROQ interaction for a number of *Tnf* CDE variants, with alteration of nucleotides in the triloop (loop mutation, LM) and base pairs in the stem region (stem mutation, SM) or with extended base-pairing at the base of the stem (Ext) (Fig. 5). We confirmed the secondary structure of the variant stem-loops by imino NOESY NMR spectra (Supplementary Fig. 7).

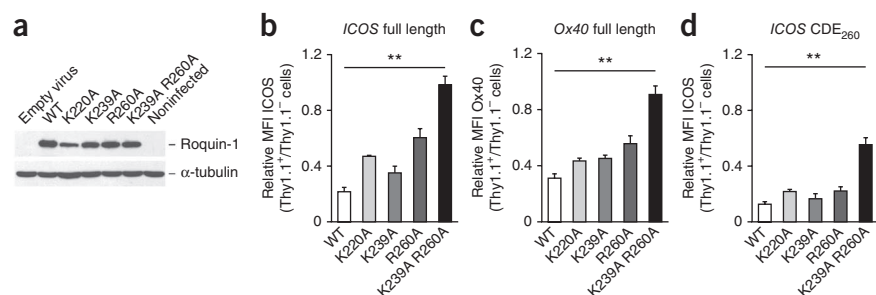
Complex formation compared to the wild-type *Tnf* CDE (Fig. 5a) was fully abolished upon inversion of the purine (Pu)–pyrimidine (Py) composition of the triloop from Pu–Pu–Pu to Pu–Py–Pu (LM3; Fig. 5b). This demonstrates that the configuration of the triloop is essential for roquin-1 binding. Exchange of the central-loop purine G12 to a pyrimidine (G12C; LM1) caused no impairment in ROQ binding, and replacement of the pyrimidine in position 13 by a purine (U13G; LM2) resulted in an only moderately reduced interaction (Fig. 5b). Thus, a cytosine at position 12 can mediate comparable base-stacking interactions as seen for G12, whereas pyrimidine or purine nucleotides are tolerated in position 13. Variations of the base pairs in the

Tnf CDE stem (Fig. 5c) confirm the importance of a closing C10–G14 base pair for efficient stacking with G12 in the triloop (mutants SM1 and SM2). Replacing the central two U–A pairs by C–G (SM3) moderately affected the ROQ interaction, consistently with the primarily observed contacts with the RNA backbone in the structure (Fig. 5c). Because a tract comprising a stretch of successive U–A base pairs (A tract) exhibits reduced thermodynamic stability and conformational plasticity^{17,18}, this feature may be important for high-affinity ROQ binding. Extending the *Tnf* CDE stem by two additional G–C pairs (Ext, Fig. 5d) reduced ROQ binding by four-fold, consistently with contributions involving the U4 base in a flipped-out, non-stacking conformation. This is further supported by comparison of NMR CSP of the ROQ domain for *Tnf* RNA and the Ext mutant (Supplementary Fig. 8), in which we observed different CSPs for amides (Phe194, Leu195 and Trp184) located near U4 (Fig. 2b,g). Thus, interactions of the unpaired U4 base with residues Trp184, Arg188 and Phe194 provide important contributions to the ROQ–CDE interface.

Natural CDE-like motifs are accepted by the ROQ domain

We also tested variants of the consensus CDE that are naturally occurring in the 3' UTRs of human *ICOS* and mouse *Ox40* (Fig. 6). We determined secondary structures and base-pairing of these stem-loops from imino NOESY NMR spectra (Supplementary Fig. 7c,d). Using EMSA experiments, we observed binding of the ROQ domain to both CDE-like stem-loop structures, although with reduced affinity

Figure 4 Functional analysis of roquin-1 ROQ-domain mutations. (a) Immunoblot analysis of the expression levels of roquin-1 and different roquin-1 mutants upon retroviral expression in roquin-deficient MEFs (uncropped images in **Supplementary Fig. 9**). α -tubulin is a loading control. (b) Coexpression of ICOS and roquin-1 (assessed by ICOS and Thy1.1 surface expression) in roquin-1- and roquin-2-deficient MEFs. Samples are cells transduced with a retrovirus encoding *ICOS* full-length mRNA and superinfected with the indicated retroviruses containing roquin-1 wild type (WT) or mutants linked to an internal ribosomal entry site and the Thy1.1-encoding gene. In flow-cytometric analyses, cells with or without ectopic roquin-1 expression were defined, and quantifications of ICOS surface expression via mean fluorescence intensity of the ICOS surface staining are shown as relative values after normalizing Thy1.1⁺ to Thy1.1⁻ populations (also **Supplementary Fig. 6a**). Bars indicate either the range of measured values from two independent experiments (K220A) or the s.d. from four independent experiments (WT, K239A, R260A and K239A R260A). Statistical significance was calculated by one-way ANOVA Kruskal-Wallis test followed by Dunn's multiple comparison test (** $P < 0.01$). (c) As in b, but with *Ox40* full-length mRNA as reporter and *Ox40* surface expression as readout (**Supplementary Fig. 6b**). (d) As in b, but with 260 nt containing the prototypic CDE from the *Tnf* 3' UTR as an artificial 3' UTR fused to the *ICOS* coding region (**Supplementary Fig. 6c**).



(**Fig. 6a,b**). The reduced affinity reflects that both CDE-like elements lack an A-tract stem, and human *ICOS* additionally has an inverted closing base pair (GC versus CG in *Tnf* CDE). This is also consistent with a comparison of NMR CSPs of the ROQ domain for *Tnf* RNA and the *Ox40* CDE-like RNA (**Supplementary Fig. 8**). Although the overall CSPs confirm that the binding interface of *Tnf* and *Ox40* with the ROQ domain is similar, distinct CSPs seen, for example for Lys220, reflect the different base pairs in the stem of the *Tnf* and *Ox40* RNAs.

We then determined whether 3' UTRs that contain non-consensus CDE variants are post-transcriptionally repressed by roquin-1 and

roquin-2 in cells. These experiments required functional analyses at endogenous roquin expression levels. We therefore generated a mouse embryonic fibroblast cell line that allowed tamoxifen-inducible deletion of roquin-1- and roquin-2-encoding genes. We monitored the conditional deletion of alleles by measuring induction of a reporter gene (*CARD1*, a cytoplasmic deletion mutant of the *CXADR* gene) and reduction of roquin expression (**Fig. 6c**). *Ox40* or *ICOS* surface levels expressed from full-length mRNAs increased after roquin deletion in tamoxifen-treated compared to untreated cells (**Fig. 6d,e**). However, this derepression did not occur when the same proteins were expressed from constructs containing only the coding

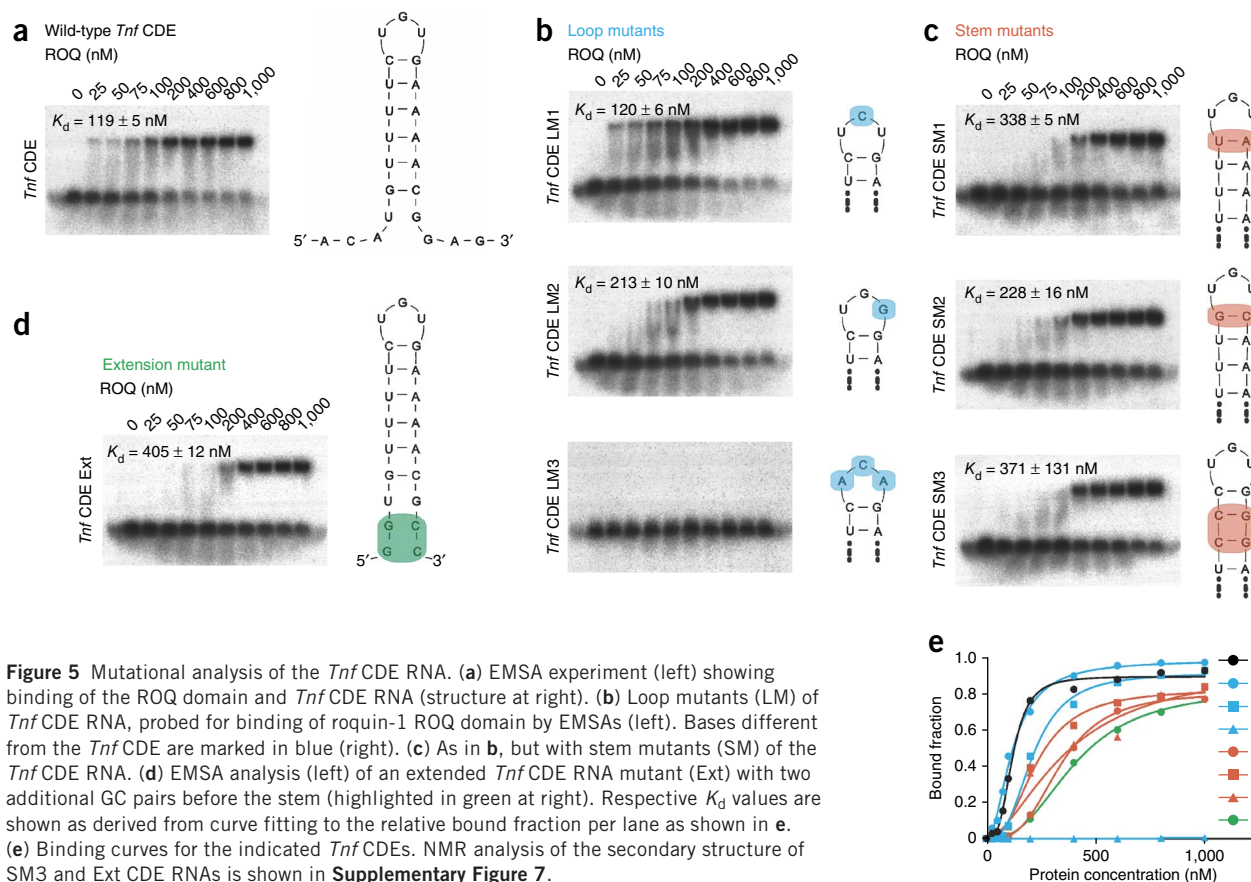


Figure 5 Mutational analysis of the *Tnf* CDE RNA. (a) EMSA experiment (left) showing binding of the ROQ domain and *Tnf* CDE RNA (structure at right). (b) Loop mutants (LM) of *Tnf* CDE RNA, probed for binding of roquin-1 ROQ domain by EMSAs (left). Bases different from the *Tnf* CDE are marked in blue (right). (c) As in b, but with stem mutants (SM) of the *Tnf* CDE RNA. (d) EMSA analysis (left) of an extended *Tnf* CDE RNA mutant (Ext) with two additional GC pairs before the stem (highlighted in green at right). Respective K_d values are shown as derived from curve fitting to the relative bound fraction per lane as shown in e. (e) Binding curves for the indicated *Tnf* CDEs. NMR analysis of the secondary structure of SM3 and Ext CDE RNAs is shown in **Supplementary Figure 7**.

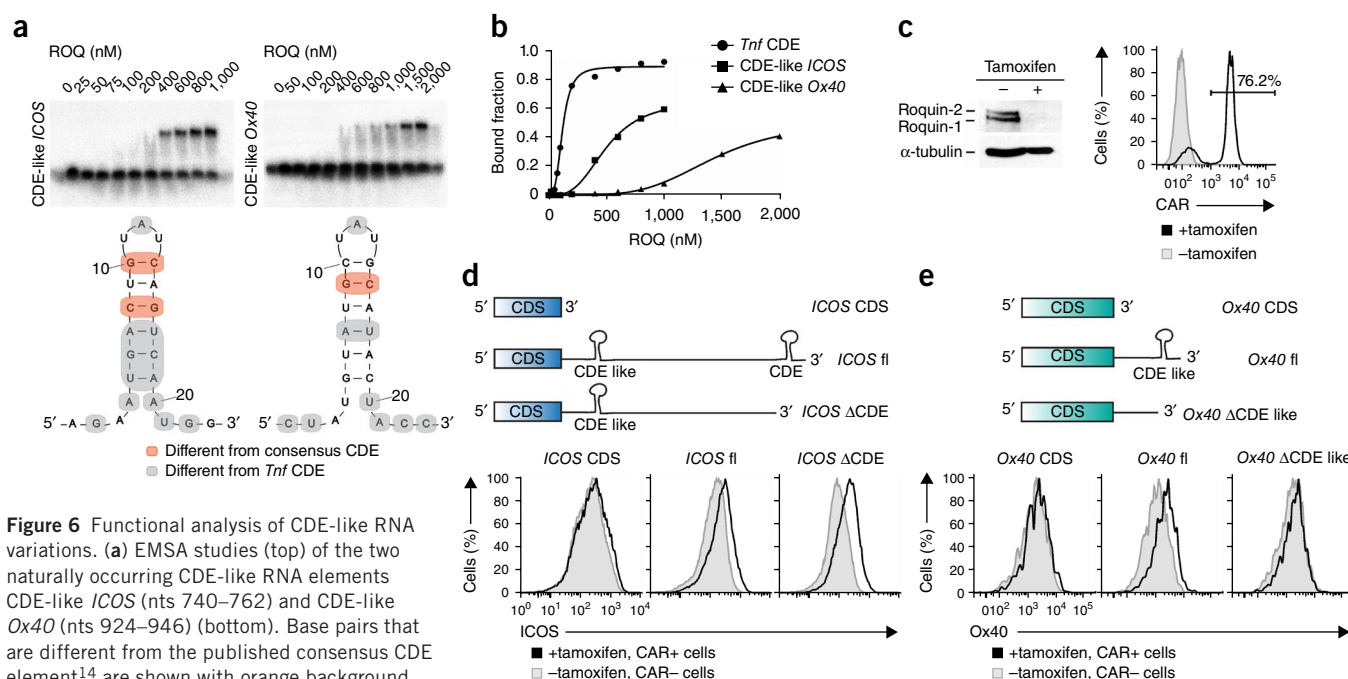


Figure 6 Functional analysis of CDE-like RNA variations. **(a)** EMSA studies (top) of the two naturally occurring CDE-like RNA elements CDE-like ICOS (nts 740–762) and CDE-like Ox40 (nts 924–946) (bottom). Base pairs that are different from the published consensus CDE element¹⁴ are shown with orange background, and nucleotide differences from the *Tnf* CDE are highlighted with gray background. **(b)** Quantification of EMSA gels shown in **a**. The wild-type *Tnf* CDE is included for comparison. **(c)** Loss of roquin-1 and roquin-2 and surface expression of CAR after tamoxifen treatment of mouse embryonic fibroblasts expressing a Cre recombinase–Ert2 fusion protein, a CAG promoter–driven *CAR* reporter containing a loxP-flanked stop cassette, and *Rc3h1* and *Rc3h2* alleles with loxP-flanked exons (details in Online Methods). Left, immunoblotting (uncropped images in **Supplementary Fig. 9**). Right, upregulation of CAR on the cell surface as assessed by flow cytometry. **(d,e)** Schematic representation (top) and ICOS and Ox40 expression measured by flow cytometry (bottom) for three different reporter constructs of ICOS (**d**) or of Ox40 (**e**) that were retrovirally expressed in MEFs and followed by tamoxifen-induced ablation of roquin-1– and roquin-2–encoding genes. Plots are representative of two independent experiments.

sequences (CDS) (**Fig. 6d,e**). The *ICOS* 3' UTR contains a consensus CDE at its 3' end¹⁴ and a CDE-like sequence upstream, and the *Ox40* 3' UTR contains only one CDE-like element in its 3' UTR (**Fig. 6d,e**). Importantly, the expression of an *ICOS* construct with a truncated 3' UTR lacking the consensus CDE¹⁴ was nevertheless derepressed upon roquin deletion (**Fig. 6d**). Moreover, deleting the CDE-like element in the *Ox40* 3' UTR by removing the 3'-end sequences strongly reduced *Ox40* regulation by endogenous roquin expression (**Fig. 6e**).

Together, these findings demonstrate that the physiologic targets *Ox40* and *ICOS* do not require high-affinity CDE sequences for roquin-mediated post-transcriptional regulation. Instead, these mRNAs contain lower-affinity CDE-like sequences that are recognized by roquin and can be critical for post-transcriptional repression by roquin in cells.

DISCUSSION

Here, we have identified the ROQ domain as a minimal roquin-1 RNA-binding region. Our crystal structures of ROQ with and without RNA reveal unprecedented features in the recognition of the *Tnf* CDE RNA by a new extended WH fold of the ROQ domain. Unique features of the RNA, i.e., a tri-loop and an AU-rich RNA stem, are recognized by the ROQ domain. Notably, most contacts are non-sequence specific, and the CDE recognition depends on the shape of this RNA element. For the tri-loop motif, a relaxed sequence consensus is tolerated, and, as compared to the U11-G12-U13 tri-loop in the *Tnf* CDE, other nucleotides in positions 12 and 13 still support roquin-1 binding (**Fig. 5b**). The shape-specific rather than sequence-specific recognition of the CDE RNA hairpin explains the specificity of roquin-1 in the regulation of mRNAs containing CDE-like RNA elements. Interestingly, the presence of an A tract in the RNA stem

seems to be necessary for high-affinity binding. A tract has been shown to reduce the thermodynamic stability of double-stranded RNAs^{17,18}. Therefore, an A tract in CDE-like RNAs might be required for an optimal alignment of the stem-loop to the ROQ domain and the simultaneous recognition of the tri-loop at the top and the unpaired U4 at the base of the stem (**Fig. 2a,b,g**).

Our structural and mutational analyses of the RNA ligand suggest that a relaxed CDE consensus can mediate roquin-dependent regulation. A comparison of roquin-1 binding to the wild-type *Tnf* CDE with binding to either mutated versions of the CDE or to other CDE-like stem-loops suggests that RNA-protein interactions tolerate deviations from the canonical CDE. We confirmed the interaction and structural similarities by EMSA (**Figs. 5 and 6**) and NMR experiments (**Supplementary Figs. 7 and 8**), respectively. Although some of these deviations resulted in a reduction of *in vitro* binding affinity, they were still functionally active in our cell-based assays (**Fig. 6e**). From these observations and our structural analysis, we conclude that a variety of CDE-like RNA structures can function as *cis* elements in roquin-mediated post-transcriptional regulation.

Our mutational analysis of the protein-RNA interface on the protein side has demonstrated a critical role of residues Lys220, Lys239 and Arg260 for RNA binding. Most importantly, we found comparable dependencies on these residues in post-transcriptional repression for the physiologic roquin targets *ICOS* and *Ox40* as well as for a CDE-containing fragment from the *Tnf* 3' UTR. Our findings thus formally establish that binding of roquin-1 to RNA is required for post-transcriptional gene regulation.

We also compared the extent to which roquin-1 mutants with compromised binding to the *Tnf* CDE showed impaired regulation

of the different 3' UTRs in cells (Fig. 4). The experiments revealed weaker repression for ICOS and Ox40. Consistently with this, the ROQ domain interacted with decreased affinity with CDE-like structures present in the 3' UTRs of ICOS and Ox40 mRNAs (Fig. 6). Furthermore, the cellular regulation of ICOS was not impaired when the consensus CDE sequence in the 3' terminal region was deleted. The inability of this CDE to contribute to regulation in our cellular system establishes that the predicted consensus CDEs¹⁴ may be sufficient to confer functional repression only in the appropriate context. Such context may enable folding or accessibility or allow the interaction with additional components of post-transcriptional gene regulation.

Together, these data suggest that physiologic targets of roquin contain functional *cis* elements with a range of roquin binding affinities in their 3' UTRs. These *cis* elements may be CDEs as well as a broad range of CDE-like sequences and may additionally include yet-unrecognized roquin-binding sites. The different high- and low-affinity binding sites as well as combinations of these may, on the one hand, allow for a variable degree of post-transcriptional gene regulation at a given cellular level of roquin proteins. On the other hand, changes in roquin expression will establish differential gene expression because targets with low-affinity binding sites but not those with high-affinity binding sites become post-transcriptionally derepressed upon reduction of roquin levels.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes [4QI0](#) and [4QI2](#) for the ROQ domain and the ROQ–RNA complex, respectively. Chemical shifts of the ROQ domain alone and in complex with *Tnf* CDE RNA have been deposited in the Biological Magnetic Resonance Data Bank under accession codes [25021](#) and [19996](#), respectively.

Note: Any Supplementary Information and Source Data files are available in the [online version of the paper](#).

ACKNOWLEDGMENTS

We acknowledge the use of the X-ray crystallography platform and the protein expression and purification facility of the Helmholtz Zentrum München, the Bavarian NMR Center, and SAXS measurements at the facility of the SFB1035, Technische Universität München. We thank S. Hauck (Helmholtz Zentrum München) and J. Winter (Technische Universität München) for MS analysis and A. Berns (The Netherlands Cancer Institute) for providing the Cre-ERT2 mouse line. G.A.H. acknowledges the Boehringer Ingelheim Fonds for a PhD fellowship. This work was supported by the Deutsche Forschungsgemeinschaft through grants

SCHL2062/1-1 (to A.S.), SFB1035 and GRK1721 (to M.S.), NI-1110/4-1 (to D.N.), SFB646 and FOR855 (to D.N.) and SFB1054 TP-A03 (to V.H.) and by the European Commission through a European Research Council grant to V.H.

AUTHOR CONTRIBUTIONS

A.S. and A.G. carried out cloning, protein expression and static light scattering. A.S. performed and analyzed NMR experiments. R.J. performed crystallization and structure determination. G.A.H. carried out EMSA assays and functional experiments. A.S. and R.S. recorded and analyzed SAXS data. A.S., G.A.H., R.J., V.H., D.N. and M.S. designed the project and wrote the paper. All authors discussed the results and commented on the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Garneau, N.L., Wilusz, J. & Wilusz, C.J. The highways and byways of mRNA decay. *Nat. Rev. Mol. Cell Biol.* **8**, 113–126 (2007).
- Houseley, J. & Tollervey, D. The many pathways of RNA degradation. *Cell* **136**, 763–776 (2009).
- Vinuesa, C.G. *et al.* A RING-type ubiquitin ligase family member required to repress follicular helper T cells and autoimmunity. *Nature* **435**, 452–458 (2005).
- Linterman, M.A. *et al.* Follicular helper T cells are required for systemic autoimmunity. *J. Exp. Med.* **206**, 561–576 (2009).
- Silva, D.G. *et al.* Anti-islet autoantibodies trigger autoimmune diabetes in the presence of an increased frequency of islet-reactive CD4 T cells. *Diabetes* **60**, 2102–2111 (2011).
- Pratama, A. *et al.* Roquin-2 shares functions with its paralog Roquin-1 in the repression of mRNAs controlling T follicular helper cells and systemic inflammation. *Immunity* **38**, 669–680 (2013).
- Ellyard, J.I. *et al.* Heterozygosity for *Roquin*^{san} leads to angioimmunoblastic T-cell lymphoma-like tumors in mice. *Blood* **120**, 812–821 (2012).
- Bertossi, A. *et al.* Loss of Roquin induces early death and immune deregulation but not autoimmunity. *J. Exp. Med.* **208**, 1749–1756 (2011).
- Vogel, K.U. *et al.* Roquin paralogs 1 and 2 redundantly repress the Icos and Ox40 costimulator mRNAs and control follicular helper T cell differentiation. *Immunity* **38**, 655–668 (2013).
- Athanasopoulos, V. *et al.* The ROQUIN family of proteins localizes to stress granules via the ROQ domain and binds target mRNAs. *FEBS J.* **277**, 2109–2127 (2010).
- Glasmacher, E. *et al.* Roquin binds inducible costimulator mRNA and effectors of mRNA decay to induce microRNA-independent post-transcriptional repression. *Nat. Immunol.* **11**, 725–733 (2010).
- Yu, D. *et al.* Roquin represses autoimmunity by limiting inducible T-cell co-stimulator messenger RNA. *Nature* **450**, 299–303 (2007).
- Maruyama, T. *et al.* Roquin-2 promotes ubiquitin-mediated degradation of ASK1 to regulate stress responses. *Sci. Signal.* **7**, ra8 (2014).
- Leppek, K. *et al.* Roquin promotes constitutive mRNA decay via a conserved class of stem-loop recognition motifs. *Cell* **153**, 869–881 (2013).
- Clark, K.L., Halay, E.D., Lai, E. & Burley, S.K. Co-crystal structure of the HNF-3/fork head DNA-recognition motif resembles histone H5. *Nature* **364**, 412–420 (1993).
- Ramakrishnan, V., Finch, J.T., Graziano, V., Lee, P.L. & Sweet, R.M. Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* **362**, 219–223 (1993).
- Conte, M.R., Conn, G.L., Brown, T. & Lane, A.N. Conformational properties and thermodynamics of the RNA duplex r(CGCAAUUUGCG)₂: comparison with the DNA analogue d(CGCAAATTTGCG)₂. *Nucleic Acids Res.* **25**, 2627–2634 (1997).
- Huang, Y., Weng, X. & Russu, I.M. Enhanced base-pair opening in the adenine tract of a RNA double helix. *Biochemistry* **50**, 1857–1863 (2011).

ONLINE METHODS

Cloning, expression and purification of recombinant proteins. Expression of full-length roquin-1 has been described previously¹¹. Cloning of expression vectors for roquin-1 ROQ(147–326), ROQ(171–326), N-term(2–440) and various ROQ domain-length versions spanning amino acids between 64 and 411 was carried out by standard procedures. We used the pETM11 and pETTrx1a vectors based on pET24d, provided by the Protein Expression and Purification Facility (PEPF) at Helmholtz Zentrum München. All vectors contained TEV protease-recognition sites for subsequent proteolytic removal of the tags. All variable-length roquin-1 expression constructs were designed and cloned via restriction sites NcoI (5') and XhoI (3'). ROQ-domain RNA-binding mutants were cloned by QuikChange PCR with high-fidelity Phusion DNA polymerase and subsequent treatment with DpnI. Alternatively, we used conventional cloning with a two-step PCR protocol and enzymatic restriction.

The roquin-1 fragments (147–326) and (171–326) were expressed as N-terminal His₆-thioredoxin fusion proteins. A single fresh clone of BL21 DE3 cells was inoculated and an overnight culture grown in LB with 35 mg/l kanamycin. 1 ml was used to inoculate an expression culture. Cells were grown to an OD₆₀₀ of 0.9 at 37 °C, induced with 0.5 mM IPTG and grown overnight at 20 °C before harvesting. Pellets were resuspended in lysis buffer (50 mM Tris, 300 mM NaCl, 4 mM TCEP, 15 mM imidazole, 1 mg/ml lysozyme, 10 µg/ml DNase I, and protease inhibitors, pH 8.0), incubated on ice for 30 min and sonicated. Cleared lysates were subjected to Ni²⁺-agarose beads. After intensive washing, beads were incubated with 500 µg/l culture of TEV protease in lysis buffer for 3 h with gentle shaking at RT. Subsequently, the bead supernatant was collected, concentrated and gel-filtrated in 20 mM Tris, 500 mM NaCl, and 2 mM TCEP, pH 7.0. The respective protein-monomer peak was pooled and salt concentration adjusted to 150 mM. Isotope-labeled protein for NMR studies was expressed in M9 minimal medium supplemented with 0.5 g/l [¹⁵N]ammonium chloride and 2 g/l unlabeled or [U-¹³C]glucose. For the preparation of deuterated proteins, cells were adapted and grown as described previously¹⁹.

The roquin-1 N-terminal domain (N-term, residues 2–440) and smaller fragments were expressed and purified essentially as described above for the ROQ domain, but no thioredoxin tag was used. For roquin-1 N-term, all expression media and the final buffer contained 100 µM or 25 µM of zinc chloride, respectively.

Full-length roquin-1 was expressed from a construct containing a TEV-cleavable N-terminal His₆-tag and a C-terminal strep tag. The construct was transformed into *Escherichia coli* Rosetta2 (DE3) and cultured at 20 °C in ZYM 5052 autoinduction medium²⁰ containing kanamycin (30 µg/ml) and chloramphenicol (33 µg/ml). Cells from 2 l of culture were harvested after reaching saturation (at OD₆₀₀ of approximately 10), resuspended in lysis buffer (50 mM Tris, 300 mM NaCl, 20 mM imidazole, 10 mM MgSO₄, 10 µg/ml DNase I, 1 mM AEBSE-HCl, 1 mg/ml lysozyme, 0.2% (v/v) NP-40, and 0.02% (v/v) 1-thioglycerol, pH 8.0) and lysed by sonication. The cleared supernatant was applied to a 5-ml HiTrap chelating column (GE Healthcare) equilibrated with 50 mM Tris, 300 mM NaCl, 20 mM imidazole, and 0.01% (v/v) 1-thioglycerol, pH 8.0, with an Äkta Purifier (GE Healthcare). The column was washed and eluted with 50 mM Tris, 300 mM NaCl, 300 mM imidazole, and 0.01% (v/v) 1-thioglycerol, pH 8.0. Fractions containing roquin-1 fl were collected and dialyzed overnight at 4 °C against 1 l of buffer C (50 mM Tris, 300 mM NaCl, and 0.01% (v/v) 1-thioglycerol, pH 8.0). The solution was subsequently applied to a 5-ml StrepTrap column (GE Healthcare). The column was washed and eluted with Strep-Tactin elution buffer (Novagen). Fractions containing roquin-1 fl were pooled and dialyzed against buffer C again. A typical yield of 1–2 mg purified roquin-1 fl per liter of culture was obtained.

Limited proteolysis and mass spectrometry. Limited proteolysis of roquin-1 was carried out starting from a roquin-1(64–411) construct. The protein was expressed and purified as described and digested in a 1:200 (w/w) ratio of proteinase K or trypsin to ROQ for 2 h in 20 mM Tris, 150 mM NaCl, 4 mM CaCl₂, and 2 mM TCEP, pH 8.0, at room temperature, and the reaction was followed by SDS gel analysis (Supplementary Fig. 1a). Products were concentrated and gel-filtrated in NMR buffer for further use, yielding monomeric ROQ domain as the main species. For MS analysis, the corresponding fragment was excised from the gel and further used in in-gel tryptic digestion to identify the protein sequence region by identifiable peptides. Tryptic digestion of postproteolytic ROQ fragments was carried out with the help of the MS facility of Helmholtz Zentrum München. Tryptic in-gel digestion was performed essentially as described previously²¹ with

ESI-Q-TOF devices with a preceding micro-HPLC step. Data were analyzed with Scaffold. Complete mass determination of the small proteolytic fragment of ROQ was carried out by the MS facility of the Technische Universität München. Freshly produced protein fragment was diluted to almost-salt-free conditions and subjected to MALDI-TOF. Data were visualized and analyzed with mmass²². The most likely region of sequence covered by the isolated ROQ fragment was chosen by comparison of digestion patterns of proteinase K and trypsin. We also took into account NMR-fingerprint spectra, suggesting an estimable number of glycines and tryptophans in the isotope-labeled fragment spectra.

RNAs. RNAs were synthesized and purified via PAGE followed by two steps of desalting and were purchased from IBA GmbH (Göttingen, Germany). No major impurities were seen in NMR spectra. Complex formation for crystallography and NMR experiments was achieved by dissolving the lyophilized RNA in water or NMR buffer. This stock solution was snap-cooled by boiling at 95 °C for 5 minutes and transfer to an ice-cooled bath for 10 min before being divided into aliquots. All RNAs were stored at –80 °C to avoid degradation and thermodynamically favored duplex formation.

Crystallization, diffraction data collection and processing. The crystallization experiments were performed at the X-ray Crystallography Platform at HMGU. The initial crystallization screening of roquin-1 ROQ(147–326) was set up at 292 K with 11 mg/ml of protein with a Mosquito nanodrop dispenser (TTP Labtech) in sitting-drop 96-well plates and commercial screens. Optimization was performed with the sitting-drop vapor-diffusion method at 292 K in 24-well plates. The best-diffracting native crystals were obtained from 100 mM bis-Tris-propane, pH 6.45, 340 mM sodium/potassium tartrate and 22% (v/w) PEG 3350. For the anomalous data collection, the native crystals were soaked with [Ta₆Br₁₂]²⁺ × 2 Br[–] (Jena Bioscience). ROQ–RNA complex was prepared by mixing roquin-1(147–326) with RNA (23-mer, *Thf*CDE, 5'-ACAUGUUUUGUGAAAA CGG AG-3') at a ratio of 1:1.25. Crystallization screening was performed as described above. The best crystals were obtained in the condition containing 100 mM bis-Tris, pH 6.0, 200 mM NaCl and 25% (v/w) PEG 3350.

Native diffraction data were collected on the ID23-1 beamline (ESRF, Grenoble, France) with a PILATUS detector at a wavelength of 1.90745 Å. A crystal derivatized with [Ta₆Br₁₂]²⁺ × 2 Br[–] was used for the anomalous data collection at the tantalum absorption edge (1.25427 Å). The data were collected at the same beamline. The diffraction data for the RNA-complex crystal were collected at 1.03321 Å wavelength at the P11 beamline at PETRA III (DESY, Hamburg, Germany). All data sets were collected at 100 K. All data sets were indexed and integrated with XDS²³ and scaled with SCALA^{24,25}. Intensities were converted to structure-factor amplitudes with TRUNCATE^{25,26}. Table 1 summarizes data collection and processing statistics for all data sets.

Structure determination and refinement. The structure of roquin-1 ROQ(147–326) was solved with the SAD protocol of Auto-Rickshaw, the EMBL-Hamburg automated crystal structure-determination platform^{27,28}. The input diffraction data were prepared and converted for use in Auto-Rickshaw with programs of the CCP4 suite²⁵. F_o values were calculated with SHELXC²⁹. On the basis of an initial analysis of the data, the maximum resolution for substructure determination and initial phase calculation was set to 2.5 Å. 34 heavy atom positions were located with SHELXD²⁹. The correct hand for the substructure was determined with ABS³⁰ and SHELXE²⁹. The occupancy of all substructure atoms was refined with MLPHARE²⁵ and phases improved by density modification with DM^{25,31}. The initial model was partially built with ARP/wARP^{32,33}. Further model building and refinement with the native data set (1.94 Å) were performed with COOT³⁴ and REFMAC5 (ref. 35), respectively, with the maximum-likelihood target function including TLS parameters³⁶. The final model is characterized by R and R_{free} factors of 16.3% and 20.6% (Table 1). The co-structure of roquin-1 with 23-mer *Thf*CDE RNA was solved by molecular replacement with the native roquin-1 ROQ(147–326) structure as a search model. Model building and refinement was performed in REFMAC5 to 3.0-Å resolution including NCS averaging. RNA molecules were modeled manually. The final model is characterized by R and R_{free} factors of 20.3% and 24.7% (Table 1). Stereochemical analysis of both final models with PROCHECK³⁷ indicates no residues with generously allowed or unfavorable backbone dihedral angles, whereas 99.4% (ROQ domain) and 92.3% (ROQ–RNA complex) of all residues are in the core region of the Ramachandran plot.

NMR spectroscopy. NMR measurements of roquin-1 ROQ(147–326) and ROQ(171–326) were performed in buffers as described mixed with 10% D₂O. Backbone chemical-shift assignments of ROQ(147–326) alone and with 1.1-fold excess of *Thf*CDE RNA were recorded with protein concentrations of 300 μ M and 220 μ M, respectively. HNCA, HNCACB, HNCO, HNCA and 3D ¹⁵N-edited NOESY spectra³⁸ were acquired at 298 K on Bruker Avance III spectrometers equipped with TCI cryogenic probe heads, at field strengths corresponding to 600, 800 and 900 MHz proton Larmor frequency. Amide ¹⁵N R₁, R₁ ρ and R₂ relaxation data and steady-state heteronuclear {¹H}¹⁵N NOE experiments were performed as described³⁹. For the estimation of τ_c values, only residues with heteronuclear NOE >0.5 were considered. Spectra were processed with Topspin3.2 and analyzed with CCPNMR Analysis⁴⁰ and Sparky (<http://www.cgl.ucsf.edu/home/sparky/>). 1D and 2D imino NOESY spectra with water-flip-back WATERGATE^{41,42} were recorded at 600–900 MHz at 278 and 298 K at 85–250 μ M RNA concentrations. Sequential assignments were guided by secondary-structure predictions with Mfold⁴³.

Static light scattering. Static light-scattering (SLS) experiments of roquin-1 fragments were performed by connecting a Viscotek TDA 305 triple detector array to an Äkta Purifier equipped with an analytical size-exclusion column at 4 °C. Roquin-1 samples were run in gel-filtration buffer on a GE Superdex200 10/300 column with concentrations as described in **Supplementary Figure 3c** at a flow of 0.5 ml/min. The molecular masses of the samples were calculated from the experimentally determined refractive index and right-angle light-scattering signals with Omniscan (Malvern Instruments). The SLS detector was calibrated with a 4 mg/ml BSA solution with 66.4 kDa for the BSA monomer and a dn/dc value of 0.185 ml/g for all protein samples.

SAXS measurements. SAXS measurements were performed with a RIGAKU Biosaxs1000 system, a microfocus rotating anode (Cu-K α , 0.154 nm) at 40 kV and 30 mA, and a Pilatus 100K Detector. Transmissions were measured by an integrated photodiode, and q values were calibrated via silver behenate. Samples were measured in a flow cell of 0.77-mm diameter. All measurements were carried out in four time frames of 900 s, and they did not show radiation damage. Concentrations of 10, 5, 2.5, and 1.25 mg/ml ROQ(147–326) protein were measured at 20 °C. The protein–RNA complex was measured at concentrations of 8 and 2 mg/ml (protein concentration) with a 1.1-fold excess of RNA. Buffers were identical to those in NMR conditions. Radial averaging, q -calibration and solvent subtraction were performed with the Rigaku SaxsLab software 3.0.1r1. $P(r)$ curves were calculated with the ATSAS package⁴⁴ version 2.5.0-2, separately normalized to the concentration and scaled to a maximum of the highest peak of the protein to 1 and the highest peak of the protein–RNA complex to 2. The ratios of the volume fractions of the protein monomer and dimer equilibria were calculated with Oligomer. The form factors of both species required were calculated from the corresponding crystal structures with *ffmaker*. χ^2 values of a fit of the theoretical scattering curve derived from the monomer to the different concentrations of the protein were calculated with crysol. Molecular weights were calculated from the Porod volumes as described previously⁴⁴ and have an error of 20%.

Electromobility shift assay (EMSAs). The EMSAs were performed according to Leppek *et al.*¹⁴ with slight modifications. For the binding reaction, a master mix containing tRNA, ³²P-labeled CDE RNA and reaction buffer was prepared and then mixed with dilutions of the recombinant proteins to achieve the indicated protein concentrations. The binding was performed for 10 min at RT or 20 min on ice in 20- μ l reaction volume in the presence of 2.5 μ g/ μ l tRNA from baker's yeast (Sigma), 500 pM ³²P-labeled CDE RNA, 20 mM HEPES, pH 7.4, 50 mM NaCl, 1 mM MgCl₂, 1 mM DTT and 1 μ g/ μ l BSA. After addition of 4 μ l 30% (v/v) glycerol as loading buffer, the RNP complexes were resolved by PAGE (6% polyacrylamide, 5% glycerol, and 0.5 $\times$ TBE) at 120 V for 40 min at RT. Gels were then fixed, dried and exposed to a phosphorimager screen. Quantification of the bound and unbound RNA fraction to calculate K_d values was performed with the AIDA Image Analyzer. The bound fraction was normalized to the total signal for each lane and plotted against the protein concentration. With the GraphPad Prism software, a nonlinear regression curve was fitted to the data points with the equation $y = B_{max}x^h/(K_d^h + x^h)$, in which B_{max} is maximum specific binding, and h is the Hill coefficient.

Mice and cell culture. The mouse line *Gt(ROSA)26Sor^{tm2(cre/ERT2)Brn}* was provided by A. Berns⁴⁵. Compound mutant mice with the *Rc3h1^{fl/fl}* (ref. 8) and *Rc3h2^{fl/fl}* (ref. 9) (denoted *Rc3h1-2^{fl/fl}*) and *CAG-CAR^{stop-fl}* (ref. 9) mutations and transgene were maintained on a C57BL/6 genetic background. Mouse embryonic fibroblasts were grown in DMEM with 10% (v/v) FCS, penicillin-streptomycin (1,000 U/ml), and 10 mM HEPES, pH 7.4, at 37 °C, 10% CO₂.

Functional assays. Expression of roquin-1 in fibroblasts used the MSCV-IRES-Thy1.1 retroviral expression vector. Roquin-1 mutations were introduced by QuikChange mutagenesis (QuikChange II XL site-directed mutagenesis Kit, Agilent). The reporter constructs *ICOS fl* (nts 1–2540; [NM_012092.3](#)), *ICOS CDS* (nts 1–600), *ICOS Δ CDE* (nts 1–2438), *Ox40 fl* (nts 1–976; [NM_011659](#)), *Ox40 CDS* (nts 1–819), *Ox40 Δ CDE-like* (nts 1–939) and *ICOS CDE₂₆₀* (*ICOS* nts 1–600 fused to *Thf* nts 1198–1456; [NM_013693](#)) were cloned into the KMV-IRES-GFP vector. To produce amphotropic or ecotropic retroviruses, the expression plasmid as well as the corresponding packaging plasmids were introduced into HEK293T cells by calcium-phosphate transfection. 72 h after transfection, cell-culture supernatants containing the retrovirus particles were harvested, filtered through 0.45- μ m filter units and stored at –80 °C.

For retroviral infection of the *Rc3h1* and *Rc3h2* double-deleted mouse embryonic fibroblast (MEFs) cell line⁹, the virus supernatants were thawed, supplemented with 5 μ g/ml polybrene and added to the MEFs for spin infection (1 h, 32 °C, 300g). After primary infection with the reporter constructs, the cells were split and spin-infected again with different roquin-1-expressing viruses or an empty virus as control. For the secondary infection, the viruses were titrated to achieve infection levels of 30–70%. 60 h after the secondary infection, the expression of the infection markers (GFP and Thy1.1) as well as the expression level of the reporter genes (*ICOS* or *Ox40*) was analyzed by cell-surface staining (1:200 anti-mouse Thy1.1 OX-7, BD Biosciences; 1:1,000 anti-human ICOS-biotin ISA-3, eBioscience; 1:200 anti-mouse Ox40 OX86, eBioscience; antibodies have been tested for application in flow cytometry by the manufacturers) and subsequent flow cytometry (LSR Fortessa, Beckton Dickinson). Statistical analysis was performed with GraphPad Prism 5.0d, and P values were calculated with the Kruskal-Wallis test followed by Dunn's multiple comparison test.

Expression levels of the different roquin-1 mutants were determined by immunoblotting with the 3F12 monoclonal antibody against roquin-1 and roquin-2 (ref. 9) and anti- α -tubulin (1:2,000, B-5-1-2, Santa Cruz; tested for immunoblotting by the manufacturer) as loading control. The rat anti-roquin antibody clone 3F12 was used in a 1:10 dilution of in house-produced hybridoma supernatants. Details have been described previously⁹.

For measuring roquin regulation of reporter genes, acute deletion of roquin-1- and roquin-2-encoding genes was performed in *Cre Ert2*; *CAG-CAR^{stop-fl}*; *Rc3h1-2^{fl/fl}* MEFs, which were retrovirally transduced with the respective reporter and subsequently treated with 0.3 μ M tamoxifen to translocate the Cre recombinase–Ert2 fusion protein to the nucleus and thereby to induce deletion of loxP-flanked exons in the *Rc3h1* and *Rc3h2* alleles as well as of the stop cassette that prevents CAR expression from the CAG promoter. After 2–4 d of tamoxifen treatment, the reporter expression in untreated cells was compared to the reporter expression of treated (CAR⁺, 1:25 anti-CAR E1-1, Santa Cruz; tested for flow cytometry by the manufacturer) cells by flow cytometry.

19. Günther, S. *et al.* Bidirectional binding of invariant chain peptides to an MHC class II molecule. *Proc. Natl. Acad. Sci. USA* **107**, 22219–22224 (2010).
20. Studier, F.W. Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* **41**, 207–234 (2005).
21. Schlundt, A. *et al.* Proline-rich sequence recognition: II. Proteomics analysis of Tsg101 ubiquitin-E2-like variant (UEV) interactions. *Mol. Cell. Proteomics* **8**, 2474–2486 (2009).
22. Strohal, M., Hassman, M., Kosata, B. & Kodicek, M. mMass data miner: an open source alternative for mass spectrometric data analysis. *Rapid Commun. Mass Spectrom.* **22**, 905–908 (2008).
23. Kabsch, W. Xds. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 125–132 (2010).
24. Evans, P. Scaling and assessment of data quality. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 72–82 (2006).
25. Winn, M.D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. D Biol. Crystallogr.* **67**, 235–242 (2011).



26. French, S. & Wilson, K. On the treatment of negative intensity observations. *Acta Crystallogr. A* **34**, 517–525 (1978).
27. Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S. & Tucker, P.A. Auto-rickshaw: an automated crystal structure determination platform as an efficient tool for the validation of an X-ray diffraction experiment. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 449–457 (2005).
28. Panjikar, S., Parthasarathy, V., Lamzin, V.S., Weiss, M.S. & Tucker, P.A. On the combination of molecular replacement and single-wavelength anomalous diffraction phasing for automated structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **65**, 1089–1097 (2009).
29. Sheldrick, G.M. Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 479–485 (2010).
30. Hao, Q. ABS: a program to determine absolute configuration and evaluate anomalous scatterer substructure. *J. Appl. Crystallogr.* **37**, 498–499 (2004).
31. Cowtan, K.D. & Zhang, K.Y. Density modification for macromolecular phase improvement. *Prog. Biophys. Mol. Biol.* **72**, 245–270 (1999).
32. Morris, R.J. *et al.* Breaking good resolutions with ARP/wARP. *J. Synchrotron Radiat.* **11**, 56–59 (2004).
33. Perrakis, A., Harkiolaki, M., Wilson, K.S. & Lamzin, V.S. ARP/wARP and molecular replacement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 1445–1450 (2001).
34. Emsley, P., Lohkamp, B., Scott, W.G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
35. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D Biol. Crystallogr.* **53**, 240–255 (1997).
36. Winn, M.D., Isupov, M.N. & Murshudov, G.N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* **57**, 122–133 (2001).
37. Laskowski, R., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
38. Sattler, M., Schleucher, J. & Griesinger, C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. *Prog. Nucl. Magn. Reson. Spectrosc.* **34**, 93–158 (1999).
39. Farrow, N.A. *et al.* Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by ¹⁵N NMR relaxation. *Biochemistry* **33**, 5984–6003 (1994).
40. Vranken, W.F. *et al.* The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
41. Grzesiek, S. & Bax, A. The importance of not saturating H₂O in protein NMR: application to sensitivity enhancement and NOE measurements. *J. Am. Chem. Soc.* **115**, 12593–12594 (1993).
42. Piotto, M., Saudek, V. & Sklenář, V. Gradient-tailored excitation for single-quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR* **2**, 661–665 (1992).
43. Zuker, M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31**, 3406–3415 (2003).
44. Petoukhov, M.V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* **45**, 342–350 (2012).
45. Hameyer, D. *et al.* Toxicity of ligand-dependent Cre recombinases and generation of a conditional Cre deleter mouse allowing mosaic recombination in peripheral tissues. *Physiol. Genomics* **31**, 32–41 (2007).