SUPPLEMENTARY DATA 1

A retroelement-derived mammalian Arc protein exhibits selective RNA recognition and nucleic acid chaperone functions

Julita Gumna-Mikina, Angelika Andrzejewska-Romanowska, Maciej Antczak, Ewa Tykwińska, Marta Szachniuk, Katarzyna Pachulska-Wieczorek

Figure S1. Raw MST traces for recombinant GST protein interactions with RNA. No binding was detected.

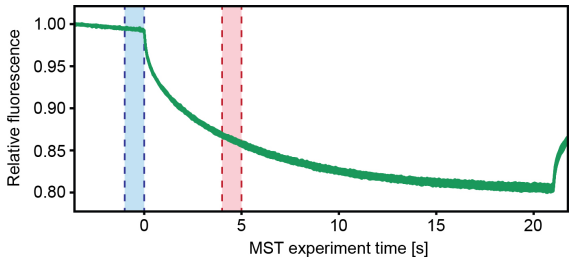

Figure S2. Dose-response binding curves of Arc to F1, F1ΔhUTR, F1ΔUTR, F2, F3, F4 Arc RNAs, and 18S rRNA without competing RNA (circle) and in the presence of 150- (square), 250- (triangle), and 350-fold (inverted triangle) molar excess of total yeast RNA. Lines represent fits of the data points using the Hill equation (EC50). Dashed vertical lines represent EC50 values for Arc complexes with RNA in the two-component system (black) and in the presence of a 350-fold molar excess of total yeast RNA (light orange).

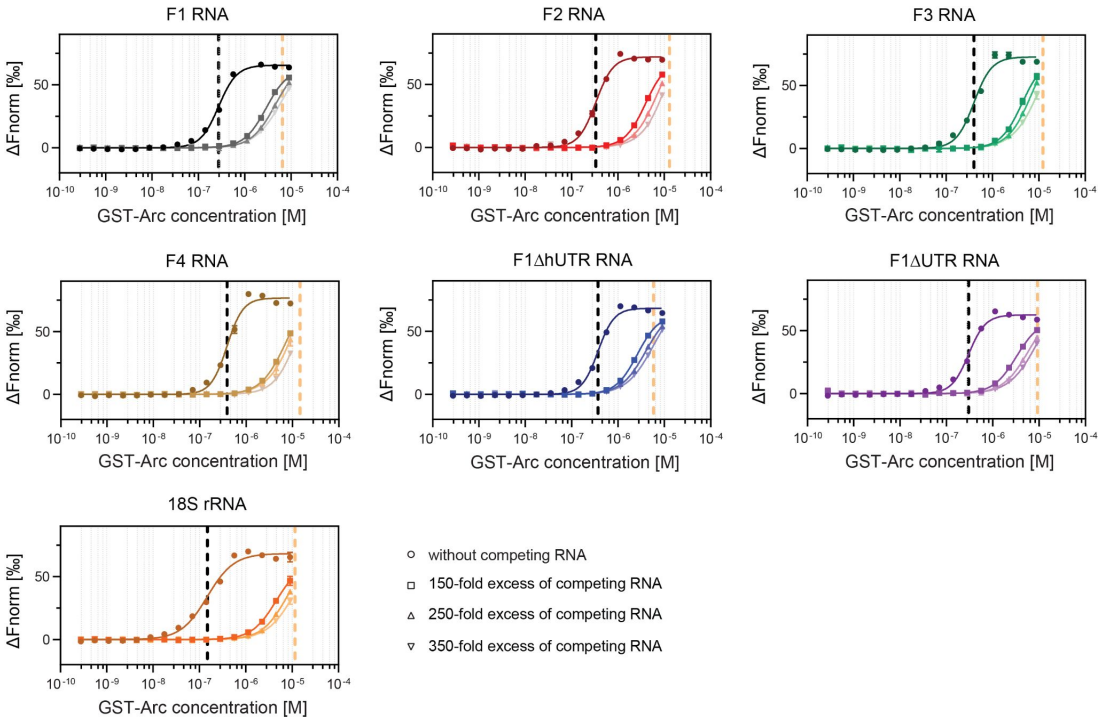

Figure S3. Dose-response binding curves of ArcΔMA to F1 RNA and 18S rRNA without competing RNA (circle) and in the presence of 25- (square), 50- (triangle), 100- (inverted triangle), and 150-fold (diamond) molar excess of total yeast RNA. Lines represent fits of the data points using the Hill equation (EC50). Dashed vertical lines represent EC50 values for ArcΔMA complexes with RNA in the two-component system (black) and in the presence of a 150-fold molar excess of total yeast RNA (light orange).

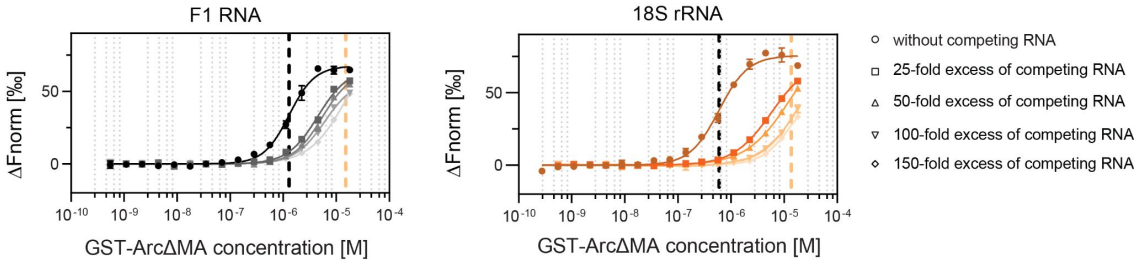

Figure S4. Pearson correlation plot of normalized SHAPE reactivities from two biological repetitions for Arc RNA in the unbound **(A)** and bound **(B)** states.

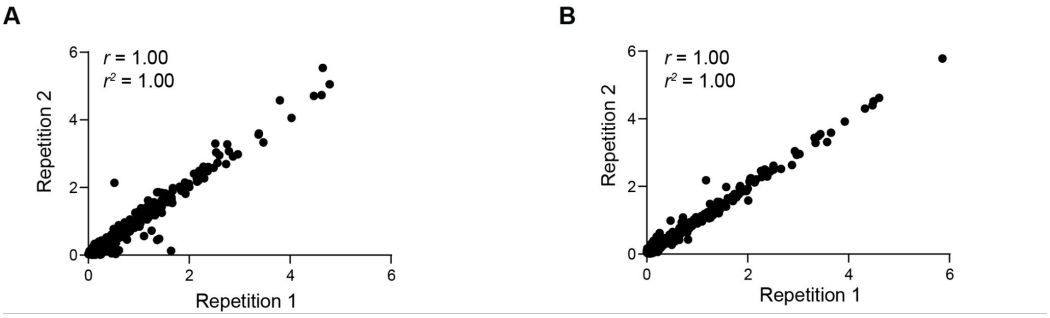

Figure S5. SHAPE-based MFE structures for +1-2926 Arc mRNA in the unbound and bound states.

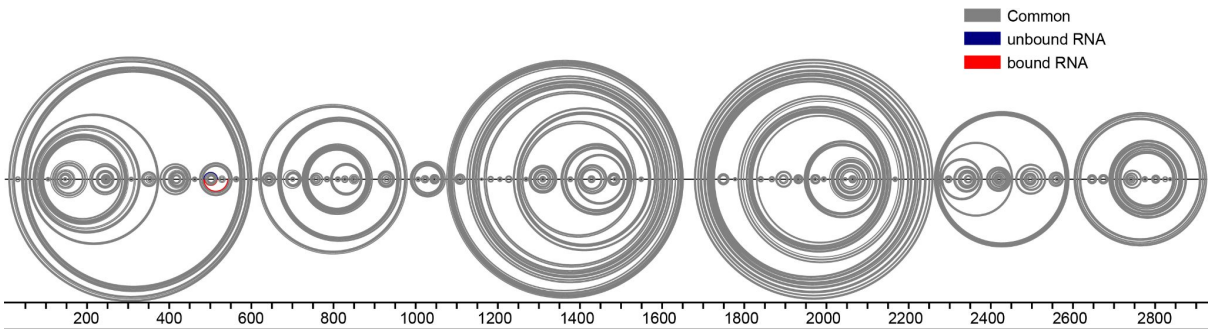

Figure S6. Pearson correlation plot of normalized SHAPE reactivities for F1ΔhUTR and F1ΔUTR and their corresponding Arc RNA fragments (+1-2926) in the unbound **(A)** and bound **(B)** states.

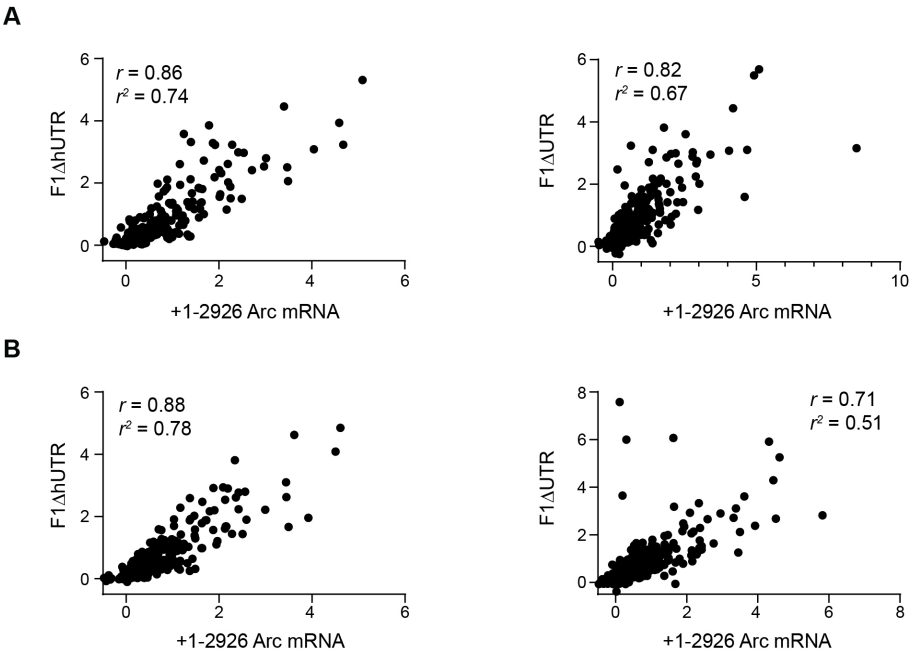

Figure S7. SHAPE-based comparison between structures of F1ΔhUTR **(A)** and F1ΔUTR **(B)** RNA in the unbound and bound states. SHAPE reactivity and Gini index distributions with medians. Significance was computed by the Wilcoxon rank-sum test; *ns* no significance; ***p*-value < 0.01; ****p*-value < 0.001. Profiles of the median SHAPE (upper plot), Gini index distributions (middle plot), and Pearson correlation (lower plot), smoothed with a 55-nt sliding window. The nucleotide numbering in the figure follows the full-length Arc mRNA sequence.

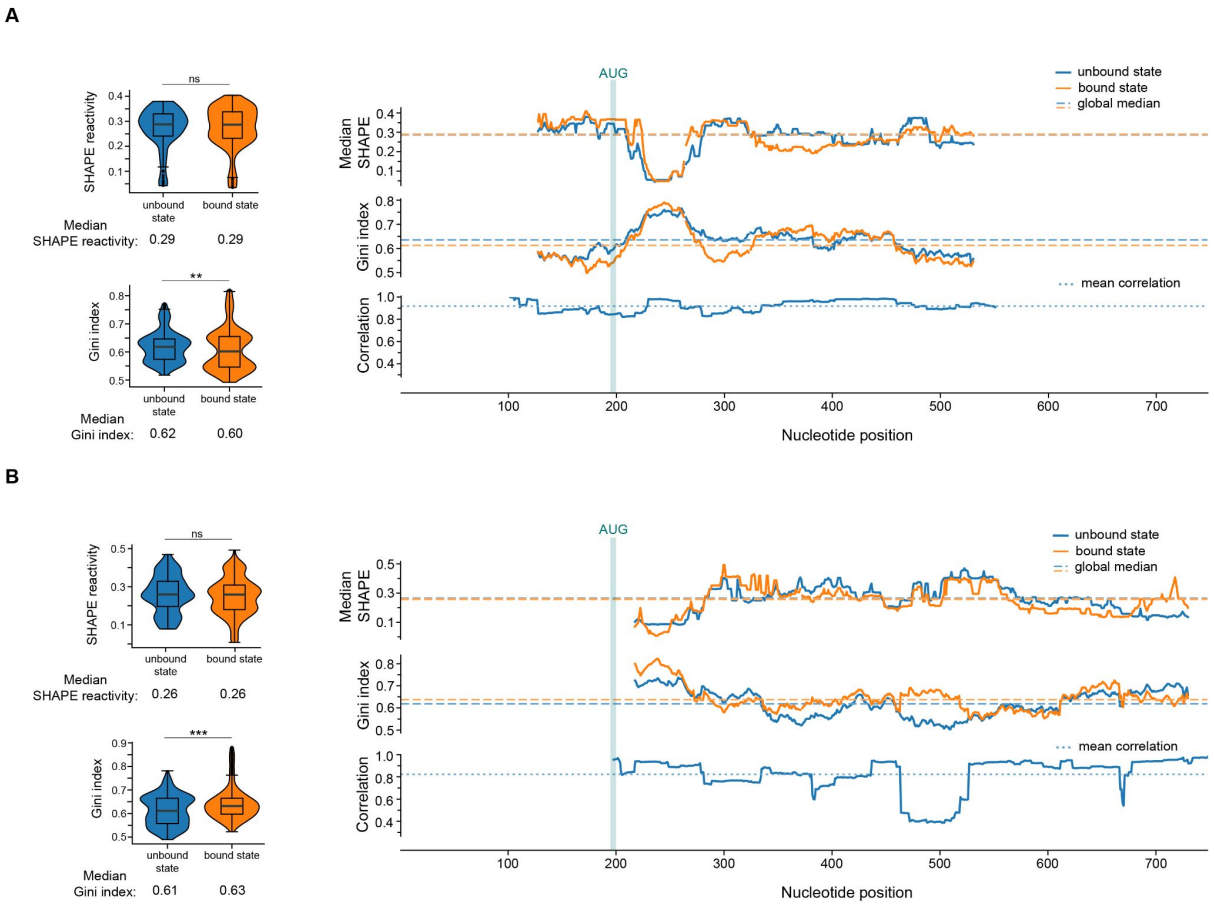

Figure S8. SHAPE-based MFE structures for F1ΔhUTR **(A)** and F1ΔUTR **(B)** in the unbound and bound states.

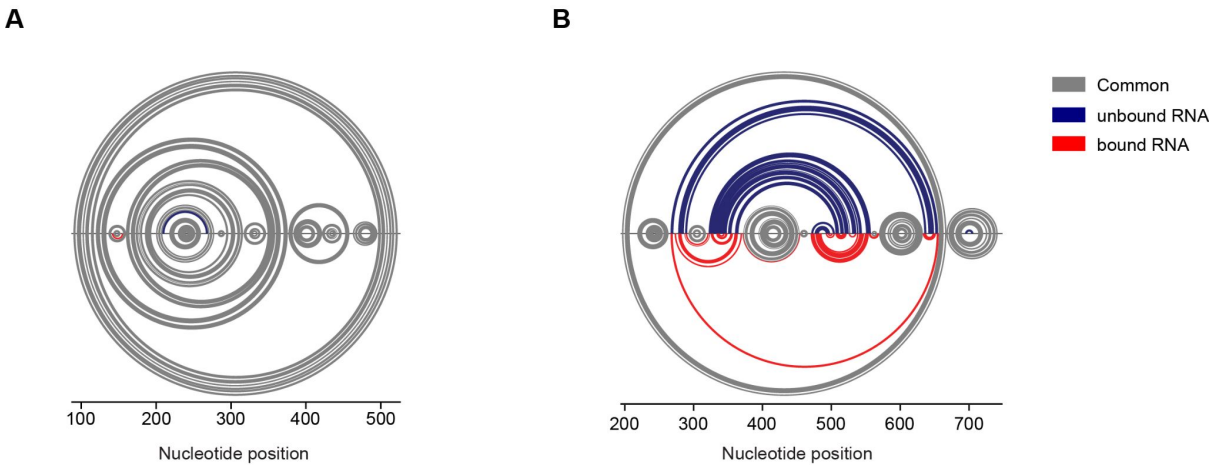

Figure S9. Arc mRNA fragments employed in the docking procedure.

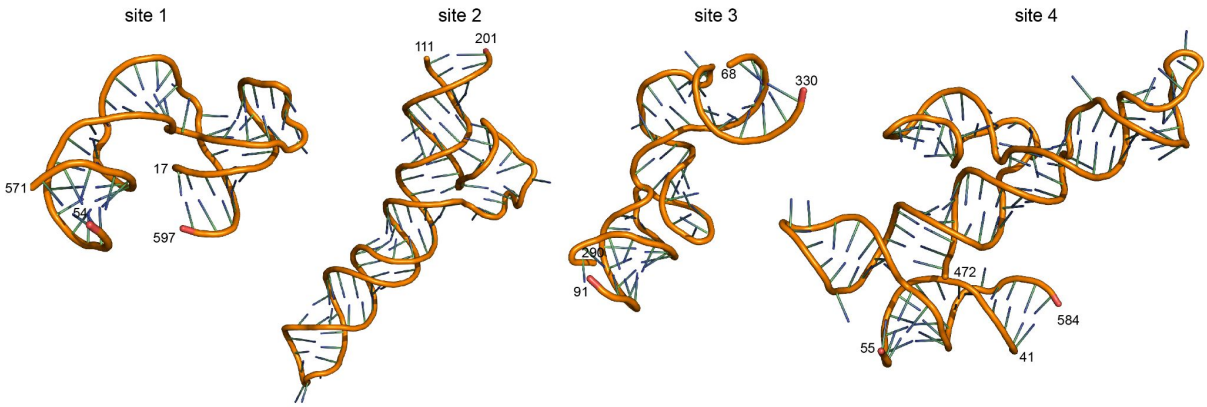

Table S1. Statistics for the first-best and second-best clusters according to HADDOCK 2.4 web server.

	Arc-RNA_site1	**Arc-RNA_site2**	**Arc-RNA_site3**	**Arc-RNA_site4**
The best cluster				
HADDOCK score	−83.8 ± 4.1	−96.2 ± 4.7	−69.1 ± 6.4	−83.2 ± 7.3
Cluster size	40	79	73	60
RMSD from the overall lowest-energy structure	12.4 ±0.5	0.9 ± 0.6	2.5 ± 1.9	2.6 ± 2.5
Van der Waals energy	−79.5 ± 8.2	−69.1 ± 5.3	−91.3 ± 5.3	−75.2 ± 14.6
Electrostatic energy	−179.8 ± 32.0	−211.8 ± 20.7	−119.0 ± 13.2	−129.7 ± 49.9
Desolvation energy	12.4 ± 5.1	9.4 ± 2.6	18.9 ± 3.6	10.2 ± 1.8
Restraints violation energy	192.6 ± 83.4	58.4 ± 28.9	271.5 ± 20.9	77.2 ± 20.0
Buried Surface Area	1970.4 ± 238.8	1714.8 ± 65.5	2121.0 ± 107.0	1852.2 ± 260.2
Z-Score	−1.4	−2.2	−1.7	−1.5
The second cluster				
HADDOCK score	−81.4 ± 12.1	−75.0 ± 5.1	−61.3 ± 3.9	−76.8 ± 5.5
Cluster size	14	28	74	18
RMSD from the overall lowest-energy structure	3.1 ± 2.8	32.2 ± 2.5	21.7 ± 0.8	36.1 ± 1.6
Van der Waals energy	−65.5 ± 9.9	−57.8 ± 2.0	−59.3 ± 8.2	−47.5 ± 6.9
Electrostatic energy	−221.5 ± 65.0	−157.0 ± 17.5	−162.1 ± 42.2	−203.4 ± 17.8
Desolvation energy	17.1 ± 8.0	5.8 ± 1.6	12.7 ± 5.3	4.6 ± 1.5
Restraints violation energy	113.6 ± 46.7	83.9 ± 25.4	177.3 ± 65.9	67.6 ± 25.0
Buried Surface Area	1705.5 ± 140.5	1383.8 ± 119.2	1486.6 ± 246.0	964.0 ± 101.0
Z-Score	−1.2	−0.8	−0.6	−1.1

Table S2. Primers used to prepare templates for *in vitro* transcription.

Primer Name	Sequence (5'-3')
PF_F1	GATTTAGGTGACACTATAGGAGCTCCCTCAGCCGCAGTCTCTGGGCCTCTTCAGCTTGA
PR_F1	CTCCCTCCACACGTGCATCTCACGCTTGACC
PF_F1ΔhUTR	GATTTAGGTGACACTATAGCACTTGCCACTGCCACTG
PF_F1ΔUTR	GATTTAGGTGACACTATAGGATGGAGCTGGACCATATGACGACCG
PR_F1ΔUTR	GGCAGCTCTCCTGCGGCAGGTGGCGGGGT
PF_F2	GATTTAGGTGACACTATGAGCAGGGTGAGCCACTT
PR_F2	AGAGGGTGAGGCCACTGGAGA
PF_F3	GATTTAGGTGACACTATGTTTGGCTTCCTCATTCTGCATCAGTGTCCAG
PR_F3	CTTTGTAATCCTATTTTCTCTGCCTTGAAAGTGTC
PF_F4	GATTTAGGTGACACTATGGAGGTGCCTGGCAGAGG
PR_F4	GGGCAGATGGCCTGGCTG
PF_5'_Ty3	TAATACGACTCACTATAGTAAGTAACATTCCGT
PR_5'_Ty3	GCCCACTGAGCAGCGGGGTT
PF_3'_Ty3	TAATACGACTCACTATAGAGGGGCCAGC
PR_3'_Ty3	GTCAAAACAGTTTATCAGATTAATTCACGGAATG
PF_18S	GATTTAGGTGACACTATAGTATCTGGTTGATCCT
PR_18S	CGCGGCTGCTGGCACCAGAC

Table S3. Primers used for amplicon-specific reverse transcription and PCR.

Primer Name	Sequence (5'-3')
Arc_long_amp1_PF	CTGATCATCTCCCTCAGCCGCAGTCTCT
Arc_long_amp1_PR	CTGATCATGGCAGGGTAGGCGTGGAG
Arc_long_amp2_PF	GACATCGCGCCTGTCCCGAACCGTAACC
Arc_long_amp2_PR	GACATCGCGTTCTCCAGCTTGCCCACCG
Arc_long_amp3_PF	CAAGTCTACACCGGCATCTGTTGACCGA
Arc_long_amp3_PR	CAAGTCTACGGTCGGCCCACCTCTC
Arc_long_amp4_PF	AGCACTGTCAACCTGGAGCGCTGGGT
Arc_long_amp4_PR	AGCACTGTGGCAGCTCTCCTGCGGCA
Arc_MaP_PR	GGCAGCTCTCCTGCGGCAGG
F1ΔhUTR_MaP_PF	GCACTTGCCACTGCCACTGC
F1ΔhUTR_MaP_PR	CTCCCTCCACACGTGCATCT
F1ΔUTR_MaP_PF	ATGGAGCTGGACCATATGAC