

SUPPLEMENTAL DATA for

Maria D. Mamońska, Maciej M. Basczok, Ewa M. Stein, Julia Kurzawska, and Mikołaj Olejniczak

Different RNA recognition by ProQ and FinO depends on the sequence surrounding intrinsic terminator hairpins

SUPPLEMENTAL METHODS

Prediction of secondary structures involving nucleotides adjacent to the terminator hairpins

The secondary structures were analyzed for RNA ligands of ProQ and FinO identified in previous studies using CLIP-seq, RIL-seq and RIP-seq methods (Holmqvist et al. 2018; Melamed et al. 2020; El Mouali et al. 2021). For the analysis of RNA ligands of ProQ 20 RNA sequences with the highest averaged read counts were selected from CLIP-seq data when the peak covered the terminator loop or when it was located no further than ten nucleotides upstream of it (Suppl. Fig. S1) (Holmqvist et al. 2018). Additionally, 20 RNA sequences with the highest averaged read counts, which were annotated as 3'-UTRs or sRNAs, were selected from RIL-seq data obtained in LB medium (Suppl. Fig. S2) (Melamed et al. 2020). The two main RNA ligands of F-like plasmid FinO protein were analyzed, which were previously identified using RIP-seq method (Suppl. Fig. S3) (El Mouali et al. 2021). For each RNA the secondary structure was predicted for a fragment consisting of the transcription terminator hairpin together with 10 nucleotides upstream of the terminator hairpin and 3'-poly(U) tail. The secondary structures were predicted using *RNAstructure* 6.4 software and visualized using its Structure Editor (Reuter and Mathews 2010).

Competition Assay

RNAs were denatured for 2 min at 90°C followed by refolding on ice for 5 min. To initiate the binding reaction a ³²P-labeled RNA (1 nM final concentration) was mixed with either ProQ or FinO or an equimolar mixture of both ProQ and FinO (100 nM final concentration of each protein) in the binding buffer (150 mM NaCl, 25 mM Tris-HCl pH 7.5, 5% glycerol, 1 mM MgCl₂) and incubated for 30 min at room temperature. Incubation was performed in low-

protein binding Eppendorf tubes. After 30 min of incubation, 10 µl reaction aliquots were loaded onto a native 6% polyacrylamide gel (19:1), and run in 0.5% TBE at 4°C. Gels were dried using a vacuum dryer, and exposed to phosphor screens, followed by data quantification using a phosphorimager and MultiGauge software (Fuji FLA-5000). The fraction bound was calculated by dividing the signal value corresponding to the gel band of the specific ³²P-RNA-protein complex (either ³²P-RNA-ProQ or ³²P-RNA-FinO complex) by the sum of the signal values of all gel bands of ³²P-RNA either free or in ³²P-RNA-protein complexes, using the equation:

$$F_b = \frac{S_{p-RNA}}{S_{p-RNA} + S_{RNA} + S_{p*-RNA}}$$

where F_b is fraction bound, S_{p-RNA} is signal corresponding to the specific protein-RNA complex (with either ProQ or FinO), S_{RNA} is signal corresponding to free RNA and S_{p*-RNA} is signal corresponding to protein-RNA complexes with the other protein. The average fraction bound was calculated from at least three independent experiments and the data were presented using GraphPad Prism software.

Gelshift assays of RNA binding by ProQ, ProQ^{NTD}, and FinO

Gelshift assays monitoring RNA binding to ProQ, ProQ^{NTD}, and FinO, which are presented in Supplemental Figures S8, S9 and S11 have been performed as described in the Materials and Methods section of the main text.

SUPPLEMENTAL TABLES

Supplemental Table S1. DNA oligonucleotides used to prepare templates for *in vitro* transcription.

Name	Sequence (5' -> 3')
cspE81_wt_F	TAATACGACTCACTATAGGCCCTTCTGCTGCAAACGTAATCGC TCTGTAAGATACGTCAGCAAG
cspE81_wt_R	AAAAAAAAACCCGCTGATTAAGCGGGTTTTGAATTCTTGCTGAC GTATCTTACAGAG
cspE81-FinP-R	AAAATCCCCGCTGATTAAGCGGGTGTGAATTCTTGCTGACGT ATCTTACAGAG
cspE81-FinP-stem-R	AAAATCGCCGCTGATTAAGCGACTGTTGAATTCTTGCTGACGT ATCTTACAGAG
FinP_wt_F	TAATACGACTCACTATAGGATACATAGGAACCTCCTCACAAAG GATTCTATGGACAGTCGATGCAGGG
FinP_wt_R	AAAATCGCCGATGCAGGGAGACGTGAACTCCCTGCATCGACTG TCCATAGAATCCT
FinP_malM_F	TAATACGACTCACTATAGGATACATAGGAACCTCCTCACAAAG GATTCTATGGACTGTTCGATGCAGGG
FinP_malM_R	AAAAAAGCCGATGCAGGGAGACGTGAACTCCCTGCATCGAC AGTCCATAGAATCCT
FinP-U-UAU ₄ -R	AAAATAGCCGATGCAGGGAGACGTGAACTCCCTGCATCGACA GTCCATAGAATCCT
FinP-U-U ₆ -R	AAAAAAGCCGATGCAGGGAGACGTGAACTCCCTGCATCGACA GTCCATAGAATCCT
FinP-U-U ₄ - R	AAAAGCCGATGCAGGGAGACGTGAACTCCCTGCATCGACAGT CCATAGAATCCT
malM_wt_F	TAATACGACTCACTATAGCTTTATCAGCAGTGTAAGGCAAG GGGTAATTACGCCCCACAGTGCTGATT
malM_wt_R	AAAAAAAGGTGCGCCAGGAGACGCACCAGTTGTTGCAAAATC AGCACTGTGGGGCGTA
malM_FinP_R	AAAATCGGTGCGCCAGGAGACGCACCTGTTGTTGCAAAATCAG CACTGTGGGGCGTA

malM_RepX_R	AGAGCGGTGCGCCAGGAGACGCACCTAATGTTGCAAAATCAG CACTGTGGGGCGTA
malM-A-GU ₆ -R	AAAAAACGGTGCGCCAGGAGACGCACCTGTTGTTGCAAAATC AGCACTGTGGGGCGTA
malM-A-GAU ₅ - R	AAAAATCGGTGCGCCAGGAGACGCACCTGTTGTTGCAAAATCA GCACTGTGGGGCGTA
malM-A-GAU ₇ - R	AAAAAAATCGGTGCGCCAGGAGACGCACCTGTTGTTGCAAAAT CAGCACTGTGGGGCGTA
RepX_F	TAATACGACTCACTATAGCATTGATTGCCTCCTTTGCAGGCAGT TGGTGGTTAGGCGCTGGCGGGG
RepX_R	AGAGCGGCGCAGGGCGGGGGTAGTGACCCCGCCAGCGCCTAA CCACCAACTGCC
RepX-malM-F	TAATACGACTCACTATAGCATTGATTGCCTCCTTTGCAGGCAGT TGGTGGACTGGCGCTGGCGGGG
RepX-malM-R	AAAAAAAGGCGCAGGGCGGGGGTAGTGACCCCGCCAGCGCCA GTCCACCAAC

Supplemental Table S2. DNA oligonucleotides used to prepare the insert to clone *E.coli* *finO* gene into the pET-15b expression plasmid.

Name	Sequence (5' → 3')	Additional information
FinO-1_F	CTTTATTTCCAATCCATGACAGAGCAG AAGC	Insertion of a part of TEV Protease recognition sequence at the 5'-end of the coding sequence
FinO-2_F	GCCATATGCTCGAGGATCCgGAAAAT CTTTATTTCCAA	Insertion of a part of TEV Protease recognition sequence and BamHI restriction site at the 5'-end of the coding sequence
FinO-1_R	GCATTGGTTCAATGGATCCTTATTGCT CATCAAGCA	Insertion of the BamHI restriction site at the 3'-end of the coding sequence

GAAAATCTTTATTTCCAATCC – TEV Protease recognition sequence

g – nucleotide introduced to maintain the open reading frame

GGATCC – BamHI restriction site

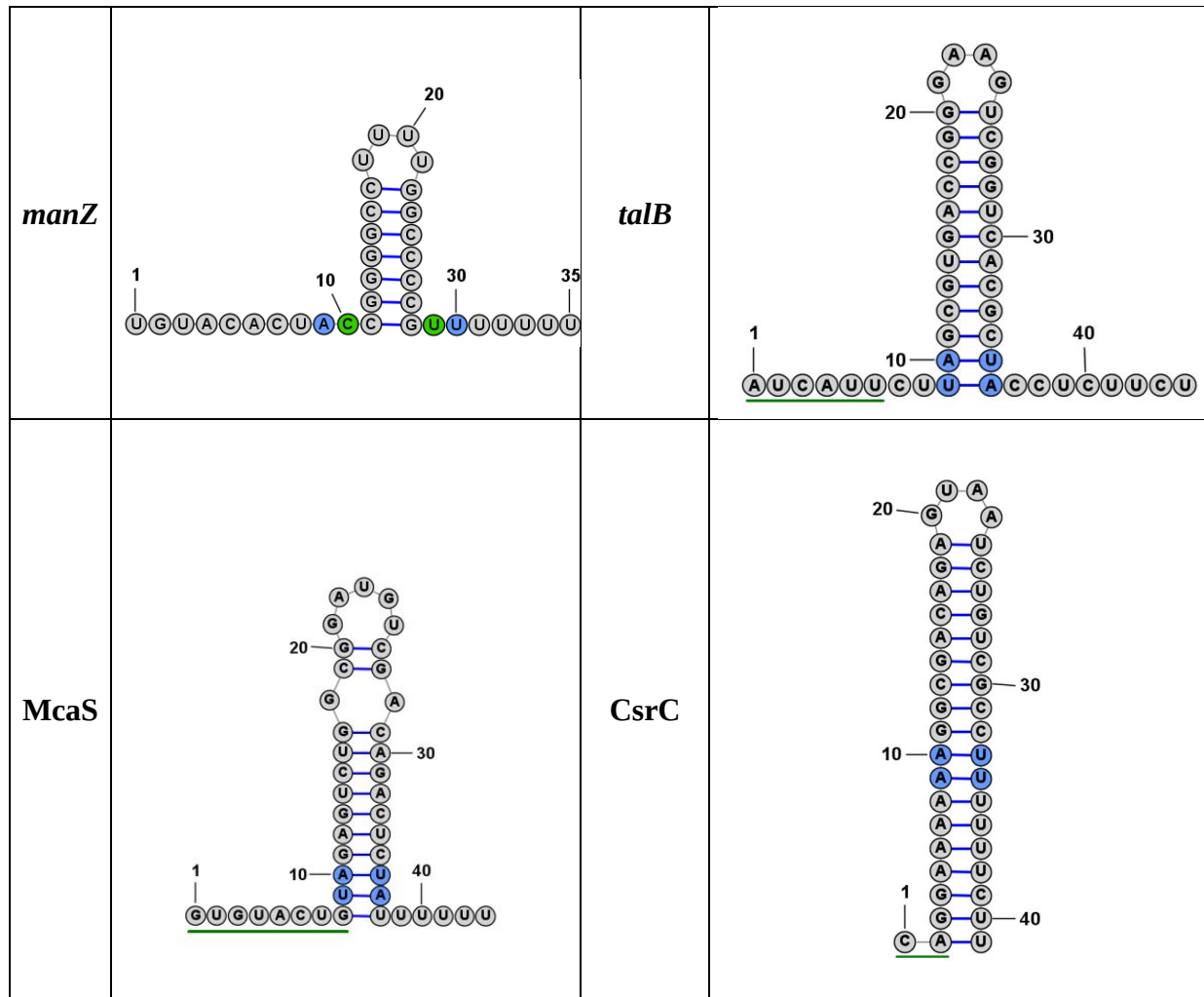
XXX – FinO coding sequence

XXX – additional sequences at 5'- and 3'-ends of the construct with FinO coding sequence inserted to improve the efficiency of BamHI cleavage

SUPPLEMENTAL FIGURES

Gene	Terminator hairpin	Gene	Terminator hairpin
<i>cspC</i>		<i>cspE</i>	
<i>uspA</i>		<i>rmf</i>	
<i>CsrB</i>		<i>ompC</i>	
<i>gadC</i>		<i>dps</i>	

<i>tnaA</i>		<i>malM</i>	
<i>cspD</i>		<i>upd</i>	
<i>SibC</i>		<i>rpsQ</i>	
<i>lpp</i>		<i>hupB</i>	



Supplemental Figure S1. Secondary structures of the terminator hairpins of the top 20 ProQ-specific RNA ligands identified in the previous CLIP-seq study in *Escherichia coli* (Holmqvist et al. 2018), whose peaks map to the regions containing intrinsic transcription terminator. RNAs were ranked according to averaged read counts (Holmqvist et al. 2018). The secondary structures were predicted using the *RNAstructure* program (Reuter and Mathews 2010) for the 3'-terminal sequences, which included the terminator hairpin, the 3' terminal polyU tail and a 10-nucleotide sequence upstream of the closing G-C or C-G base pair of the terminator hairpin. The two nucleotide pairs immediately below the closing G-C or C-G base pair of the terminator hairpin are marked in color, where blue denotes canonical base pairs (A-U, U-A, or G-U), and green denotes pyrimidine-pyrimidine mismatches (U-U, C-U, or C-C). 5' adjacent sequences, which are involved in secondary structure in the context of longer RNA are underlined green.

Gene	Terminator hairpin	Gene	Terminator hairpin
<i>malM</i>		<i>cspD</i>	
<i>acpP</i>		<i>garD</i>	
<i>RyfA</i>		<i>lpp</i>	
<i>cspE</i>		<i>manZ</i>	

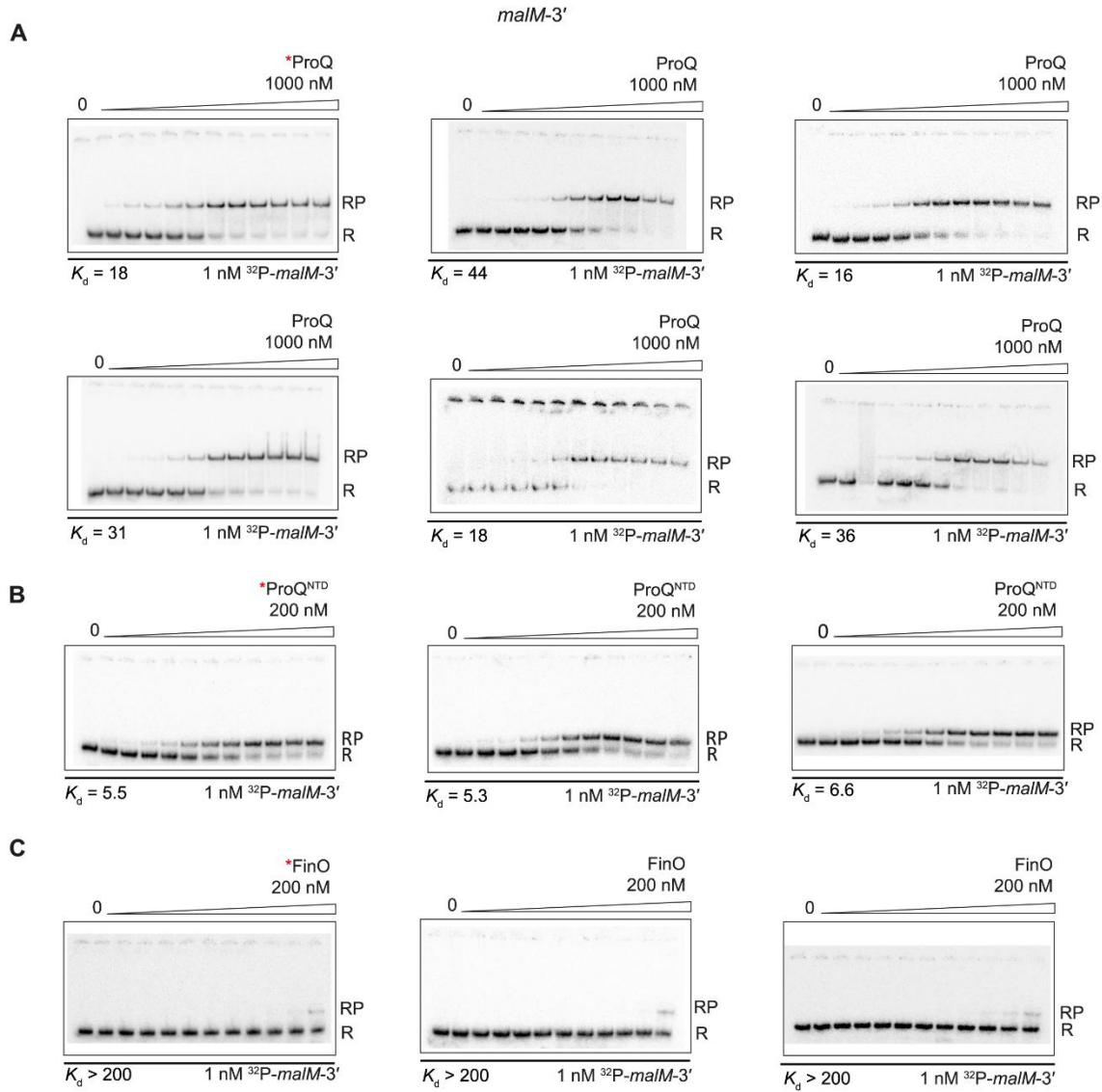
<i>cspA</i>		<i>ompF</i>	
<i>RybB</i>		<i>fbaA</i>	
<i>RaiZ</i>		<i>rraB</i>	
<i>hupB</i>		<i>McaS</i>	

<i>infA</i>		<i>adhE</i>	
<i>ihfA</i>		<i>sodB</i>	

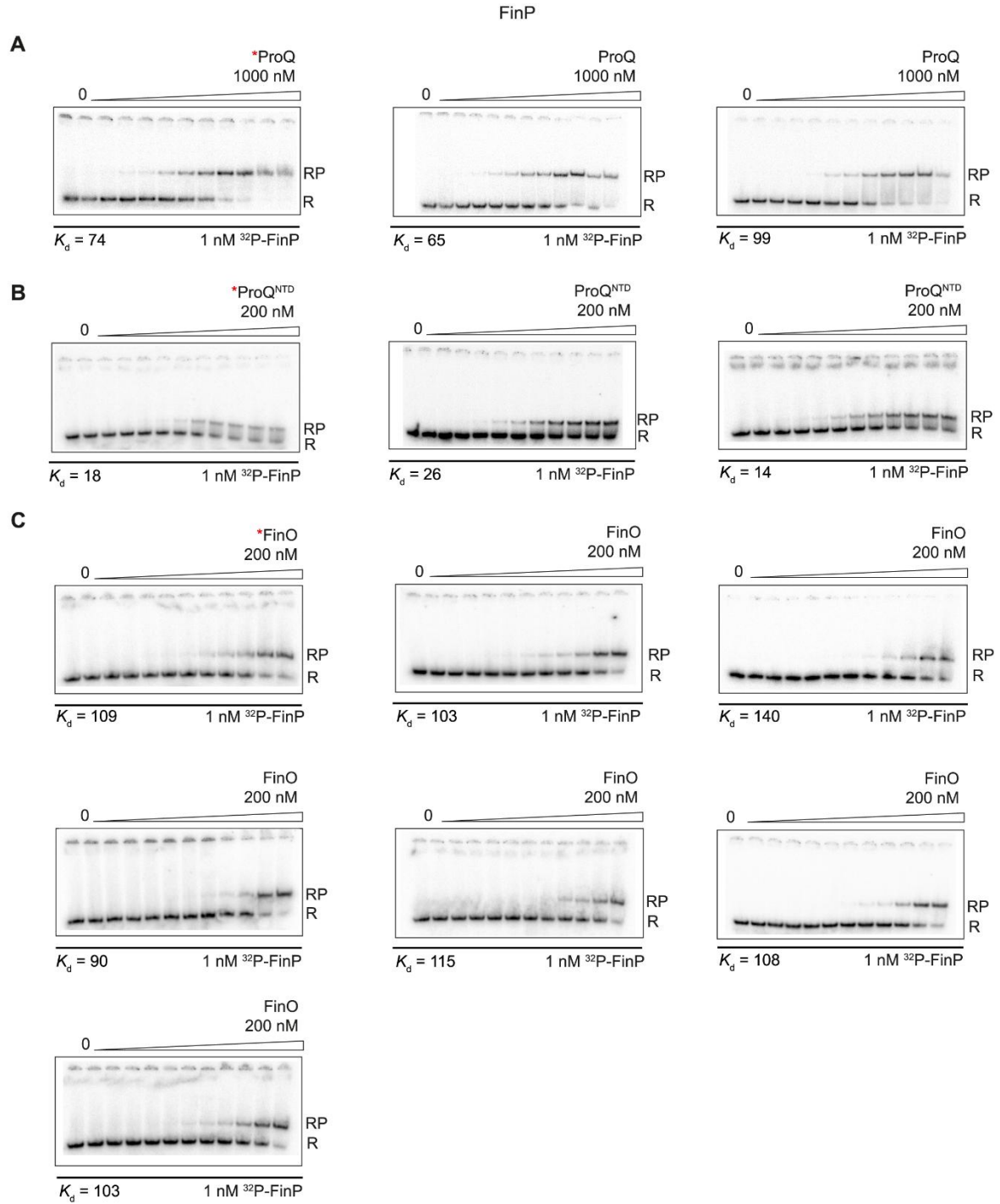
Supplemental Figure S2. Secondary structures of the terminator hairpins of the top 20 ProQ-specific RNA ligands identified in the previous RIL-seq study in *Escherichia coli* (Melamed et al. 2020), which were annotated either as 3'-UTRs or sRNAs. RNAs were ranked according to averaged read counts (Melamed et al. 2020). The secondary structures were predicted using the *RNAstructure* program (Reuter and Mathews 2010) for the 3'-terminal sequences, which included the terminator hairpin, the 3' terminal polyU tail and a 10-nucleotide sequence upstream of the closing G-C or C-G base pair of the terminator hairpin. The two nucleotide pairs immediately below the closing G-C or C-G base pair of the terminator hairpin are marked in color, where blue denotes canonical base pairs (A-U, U-A, or G-U), and green denotes pyrimidine-pyrimidine mismatches (U-U, or C-U). 5' adjacent sequences, which are involved in secondary structure in the context of longer RNA are underlined green.

sRNA	Terminator hairpin
FinP	
RepX	

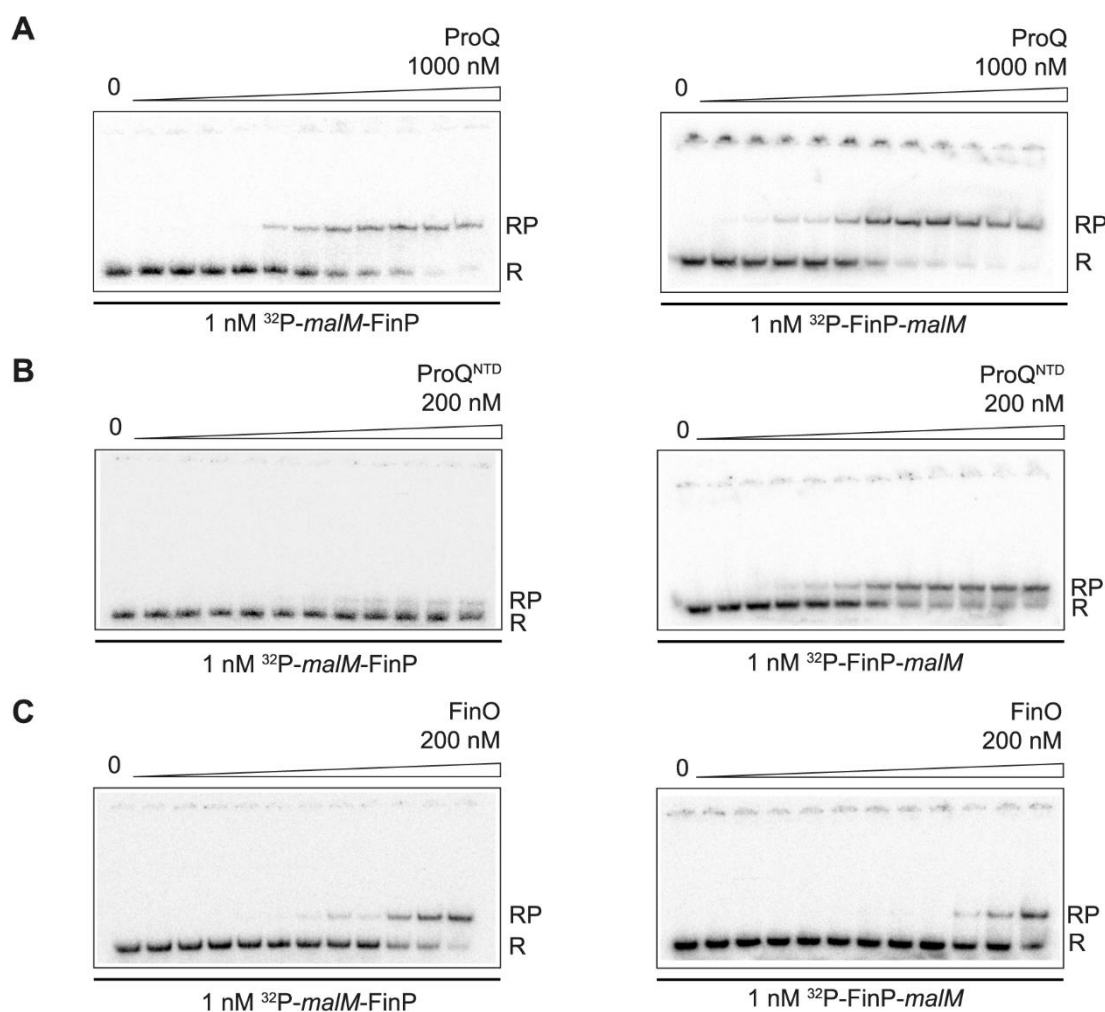
Supplemental Figure S3. Secondary structures of the two main *Salmonella enterica* FinO-specific RNA ligands, FinP and RepX, which were detected using RIP-seq (El Mouali et al. 2021). The secondary structures were predicted using the *RNAstructure* program (Reuter and Mathews 2010) for the 3'-terminal sequences, which included the terminator hairpin, the 3' terminal tail and a 10-nucleotide sequence upstream of the closing G-C or C-G base pair of the terminator hairpin. The two nucleotide pairs immediately below the closing G-C or C-G base pair of the terminator hairpin are marked in color, where red denotes purine-purine mismatches (A-G), green denotes a pyrimidine-pyrimidine mismatch (U-C), and light blue denotes a C-A mismatch. 5' adjacent sequences, which are involved in secondary structure in the context of longer RNA are underlined green.



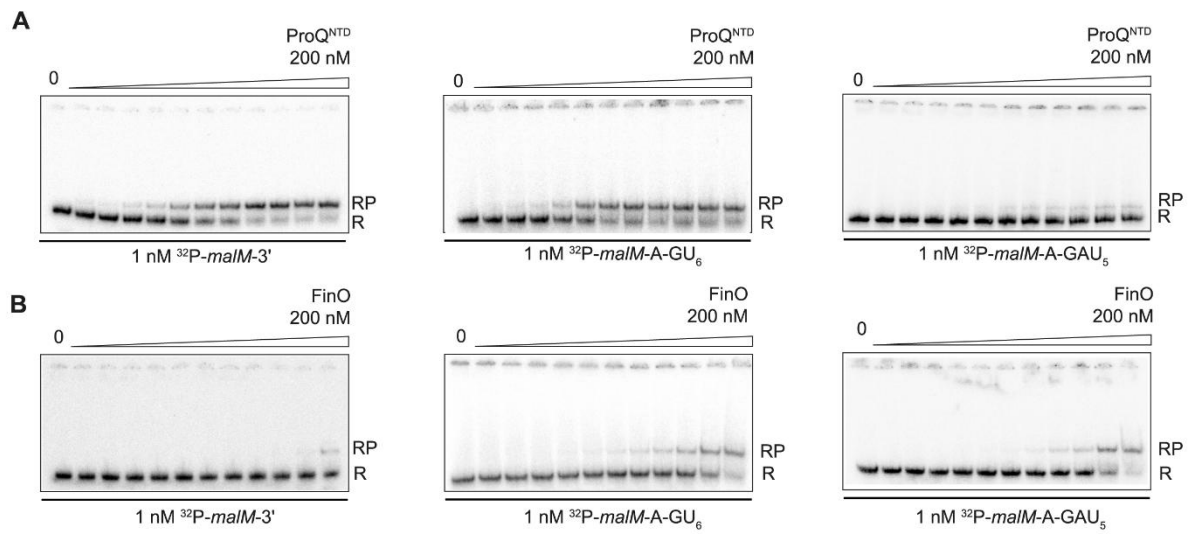
Supplemental Figure S4. The gelshift analysis of the binding of *malM-3'* RNA to ProQ (A), ProQ^{NTD} (B) and FinO (C) proteins. All replicates are shown. The gels presented in Figure 1 in main text are marked with red asterisks. Free ^{32}P -RNA is marked as R and RNA-protein complex as RP. The fitting of the quadratic equation to the binding data provided the K_d values presented below each gel, while the average K_d values are presented in Table 1 in the main text. The range of protein concentrations is indicated above the gels. The concentration series of each protein was obtained by 2-fold sequential dilutions.



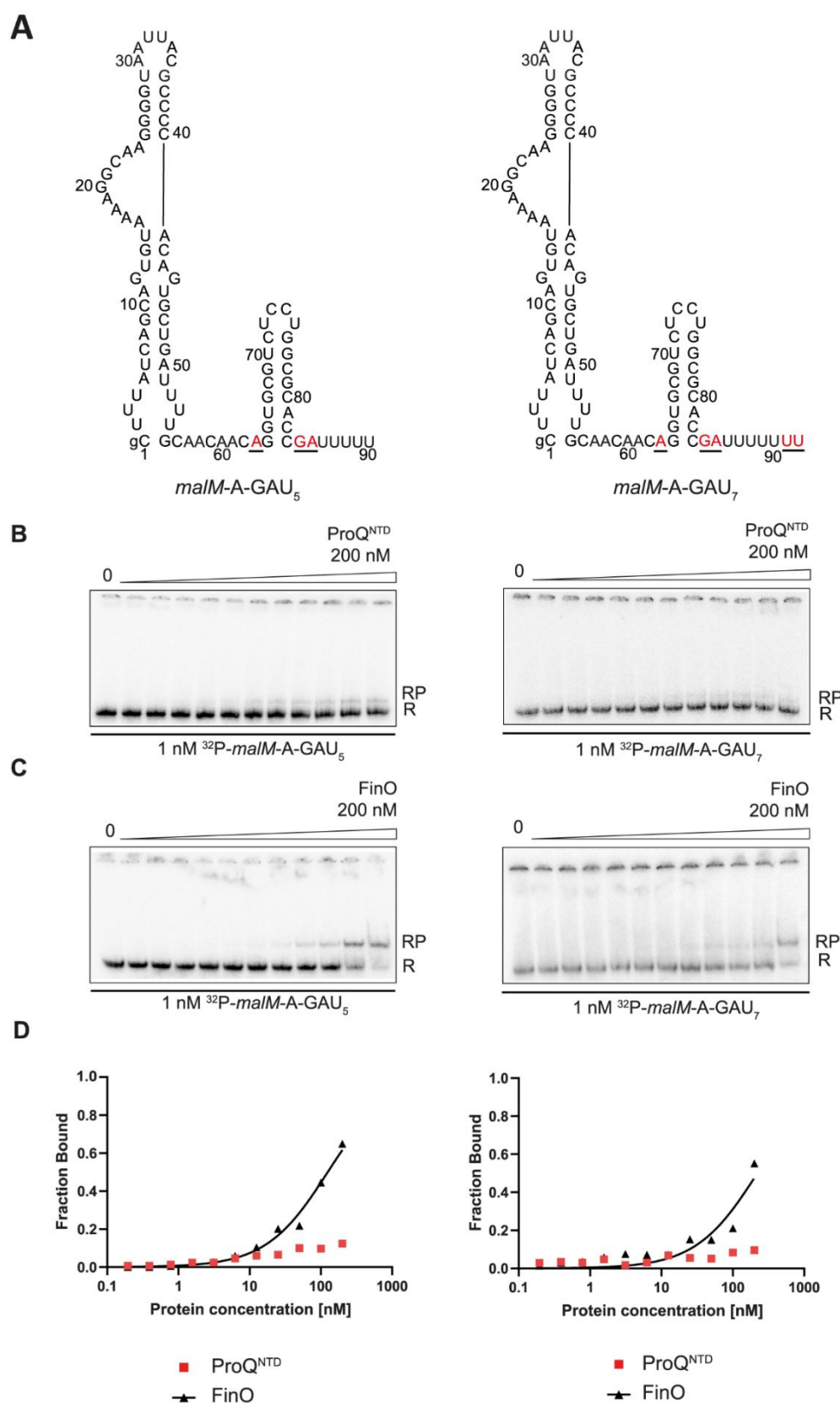
Supplemental Figure S5. The gelshift analysis of the binding of FinP RNA to ProQ (A), ProQ^{NTD} (B) and FinO (C) proteins. All replicates are shown. The gels presented in Figure 1 in main text are marked with red asterisks. Free ^{32}P -RNA is marked as R and RNA-protein complex as RP. The fitting of the quadratic equation to the binding data provided the K_d values presented below each gel, while the average K_d values are presented in Table 1 in the main text. The range of protein concentrations is indicated above the gels. The concentration series of each protein was obtained by 2-fold sequential dilutions.



Supplemental Figure S6. Gelshift analysis of the binding of *malM*-FinP and FinP-*malM* to full-length ProQ (A), ProQ^{NTD} (B), and FinO (C). The raw data in the gels correspond to the data presented in plots in Figure 2 in the main text. Free ^{32}P -labeled RNA is marked as R and RNA-protein complexes as RP.

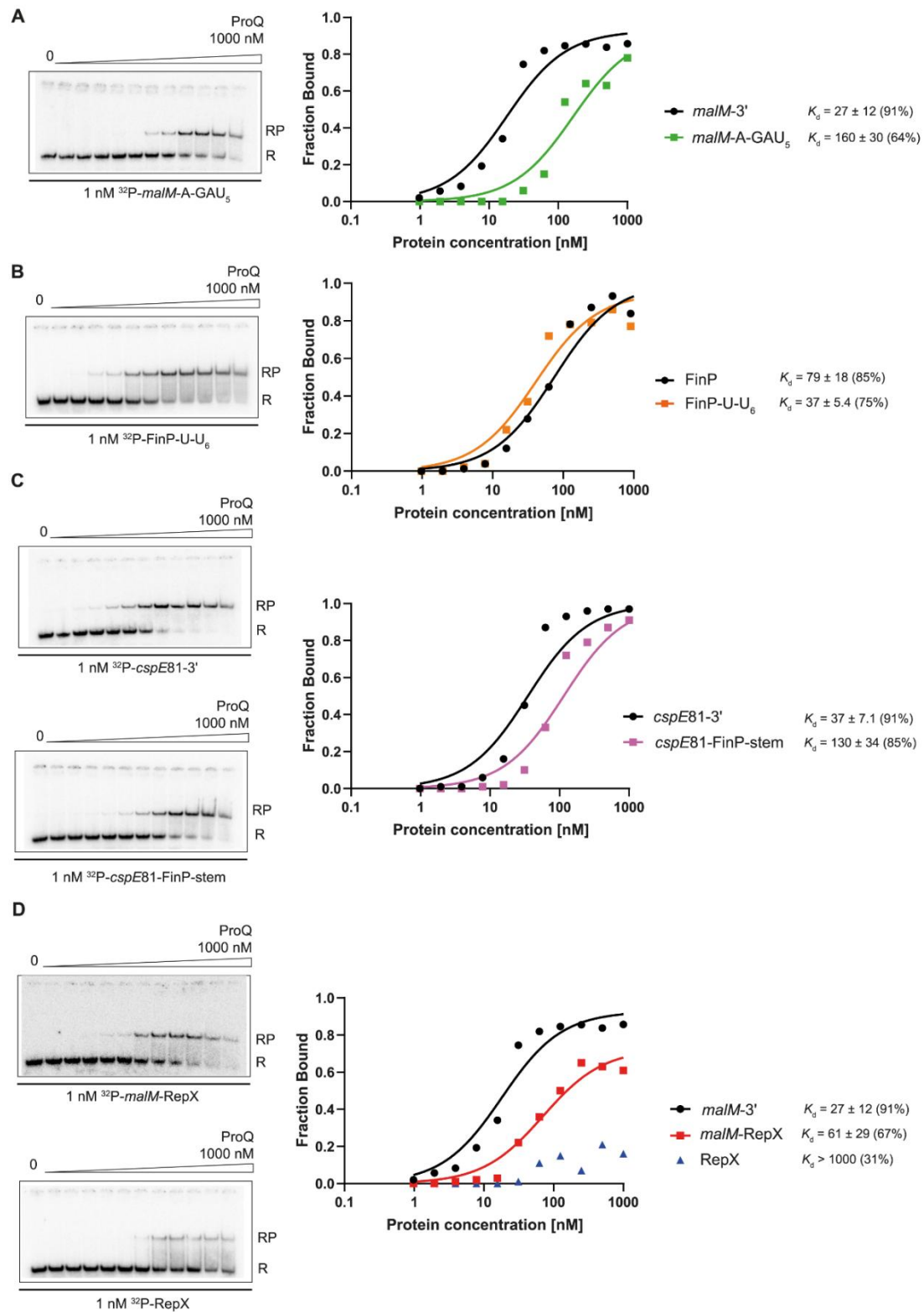


Supplemental Figure S7. Gelshift analysis of the binding of *malM*-3', *malM*-A-GU₆ and *malM*-A-GAU₅ to ProQ^{NTD} (A), and FinO (B). The raw data in the gels correspond to the data presented in plots in Figure 3 in the main text. Free ³²P-labeled RNA is marked as R and RNA-protein complexes as RP.

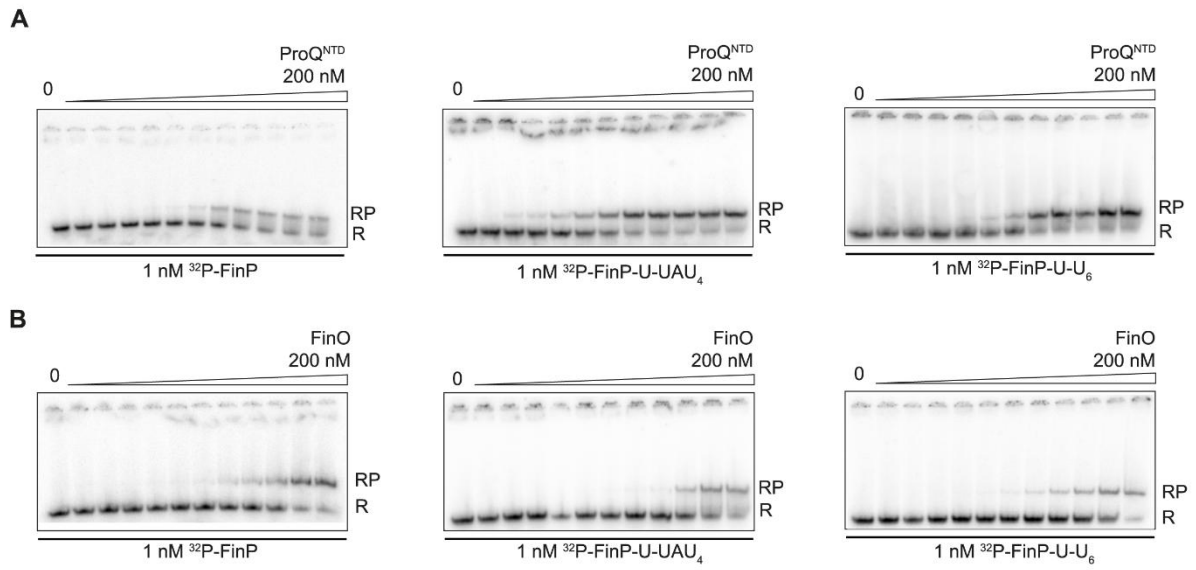


Supplemental Figure S8. Comparison of *malM-A-GAU₅* and *malM-A-GAU₇* binding to ProQ^{NTD}, and to FinO. (A) Secondary structures of *malM-A-GAU₅* and *malM-A-GAU₇*, which were predicted using *RNAstructure* software (Reuter and Mathews 2010). The substitutions introduced into *malM-3'* are shown in red, underlined font. The lower case g

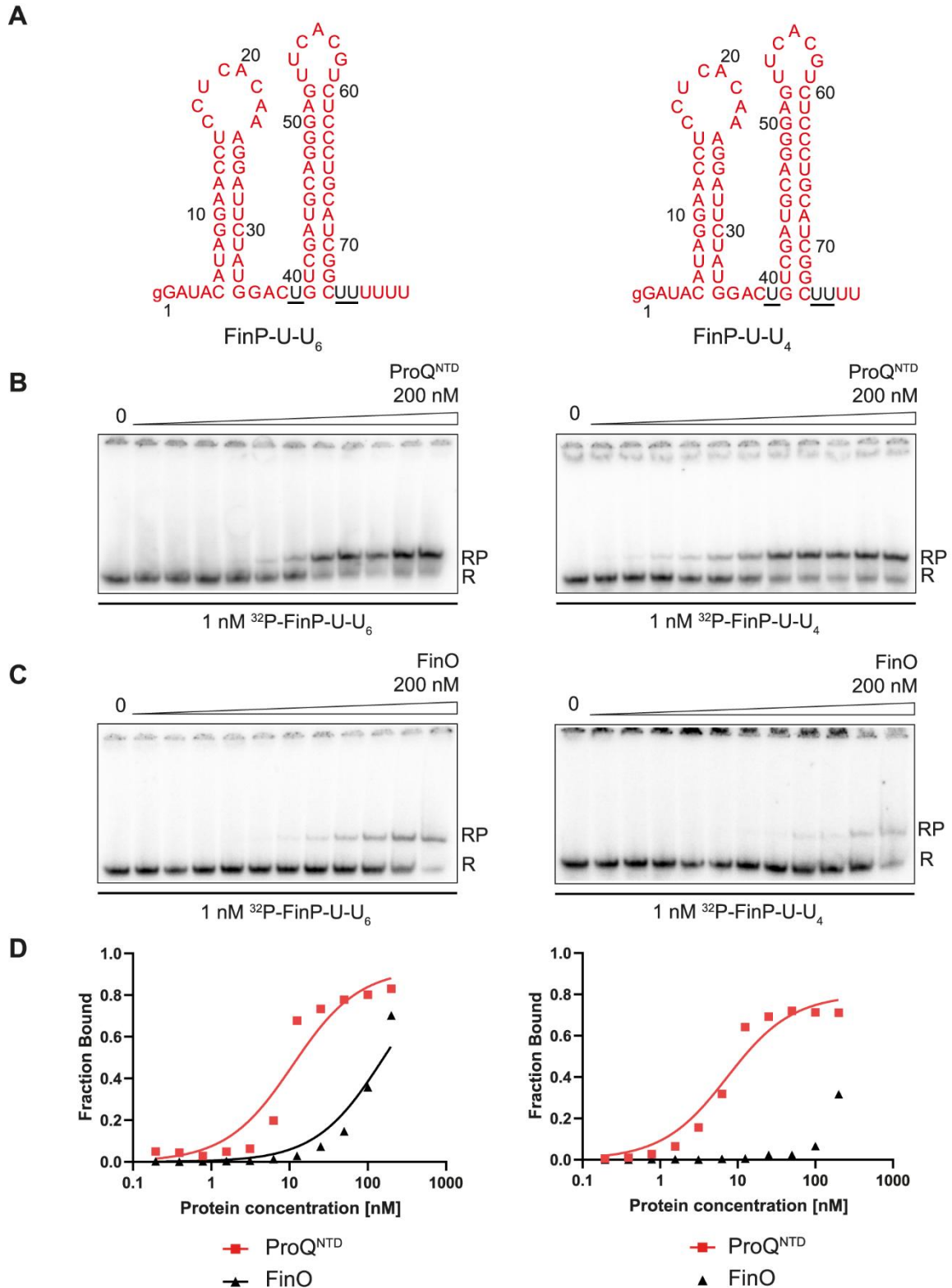
denotes guanosine residue added on 5' ends of RNA molecules to enable T7 RNA polymerase transcription. (B), (C) The gelshift analysis of *malM*-A-GAU₅ and *malM*-A-GAU₇ binding to ProQ^{NTD} (B), and FinO (C). (D) The plots of fraction bound data from (B) and (C) versus protein concentration are presented. The data shown for *malM*-A-GAU₅ are the same as in Figure 3 (main text), and the corresponding average K_d values are shown in Table 1 (main text). Because the maximum fraction bound of *malM*-GAU₇ binding to ProQ^{NTD} was 8%, the K_d value was assumed as bigger than 200 nM. The fitting of the quadratic equation into data for *malM*-A-GAU₇ binding to FinO provided the K_d value of 197 ± 86 with maximum fraction bound of 56%. The values were calculated from at least three independent experiments.



Supplementary Figure S9. The binding of *malM*-RepX, RepX (A), *malM*-A-GAU₅ (B), *cspE81*, *cspE81*-FinP-stem (C), and FinP-U-U₆ (D) RNAs to the full-length ProQ protein. Free ^{32}P -RNA is marked as R, and RNA-protein complex as RP. The fitting of the quadratic equation into the data provided the average K_d values presented at RNA names in the plot legends, and in Table 1 in the main text. The range of the protein concentrations is indicated above the gels. The concentration series were made by 2-fold sequential dilutions.

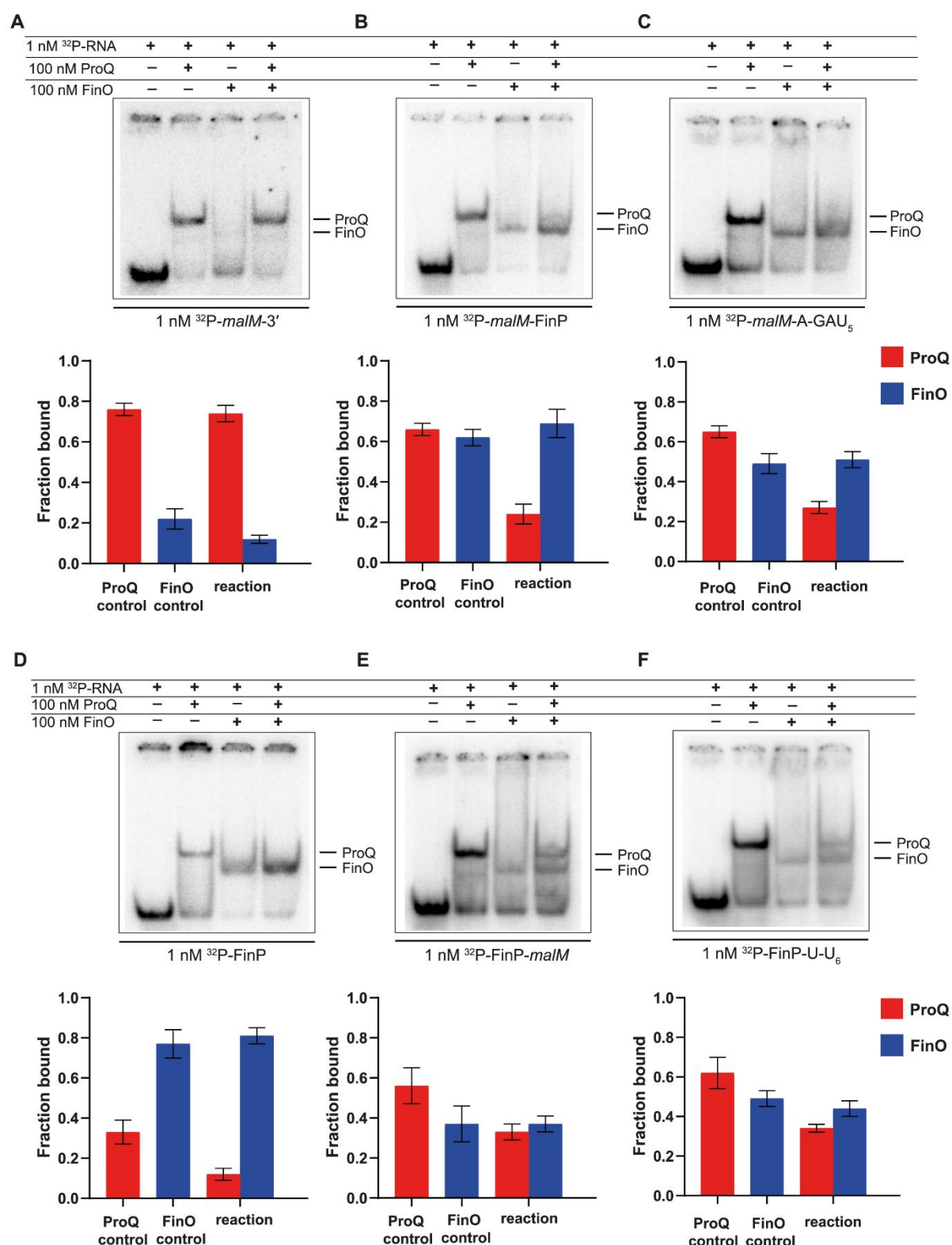


Supplemental Figure S10. Gelshift analysis of the binding of FinP, FinP-U-UAU₄ and FinP-U-U₆ to ProQ^{NTD} (A), and FinO (B). The raw data in the gels correspond to the data presented in plots in Figure 4 in the main text. Free ³²P-labeled RNA is marked as R and RNA-protein complexes as RP.

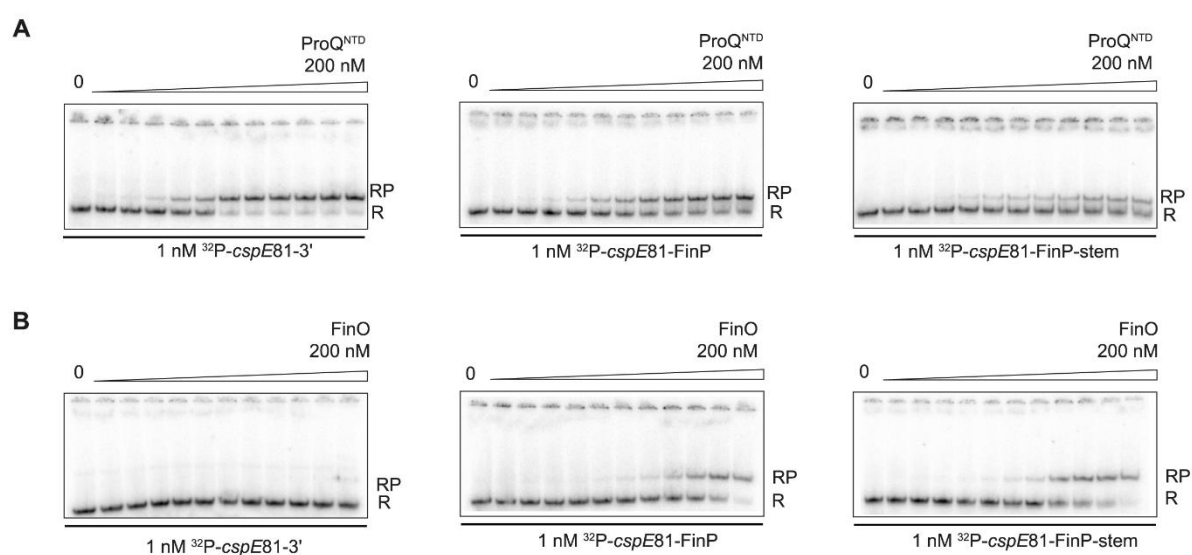


Supplemental Figure S11. Comparison of FinP-U-U₆ and FinP-U-U₄ binding to ProQ^{NTD} and to FinO. (A) Secondary structures of FinP-U-U₆ and FinP-U-U₄, which were predicted using *RNAstructure* software (Reuter and Mathews 2010). The substitutions introduced into FinP are shown in black, underlined font. The lower case g denotes guanosine residue added on 5' ends of RNA molecules to enable T7 RNA polymerase transcription. (B), (C) The gelshift analysis of FinP-U-U₆ and FinP-U-U₄ binding to ProQ^{NTD} (B), and FinO (C). (D) The

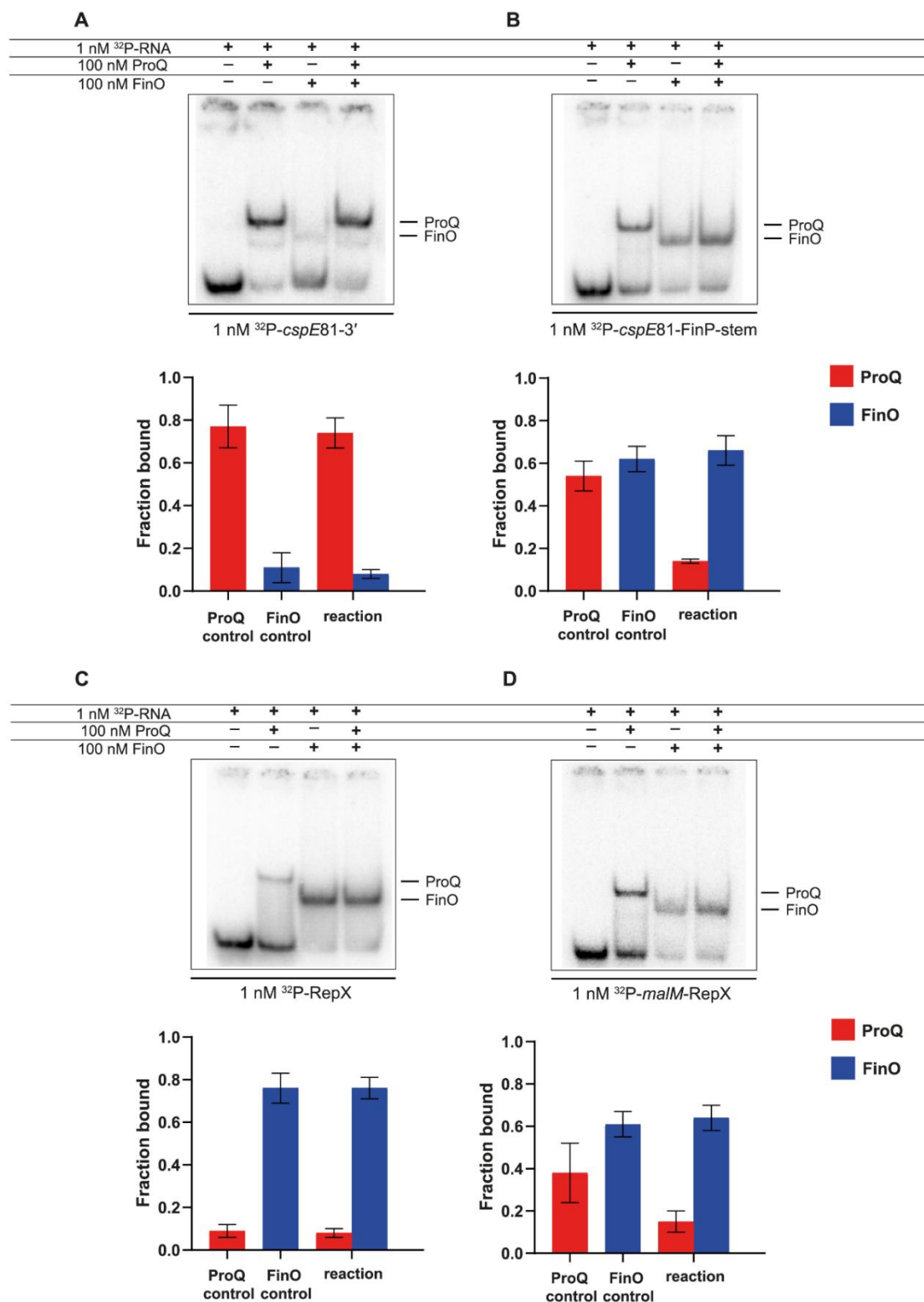
plots of fraction bound data from (B) and (C) versus protein concentration are shown for FinP-U-U₆ and FinP-U-U₄. The data shown for FinP-U-U₆ are the same as in Figure 4 (main text), and the corresponding average K_d values are shown in Table 1 (main text). The fitting of the quadratic equation into data for FinP-U-U₄ binding to ProQ^{NTD} provided the K_d value of 7.6 ± 3.0 nM with a maximum fraction bound of 71%. Because the maximum fraction bound of FinP-U-U₄ binding to FinO was 30%, the K_d value was assumed as bigger than 200 nM. The K_d values were calculated from at least three independent experiments.



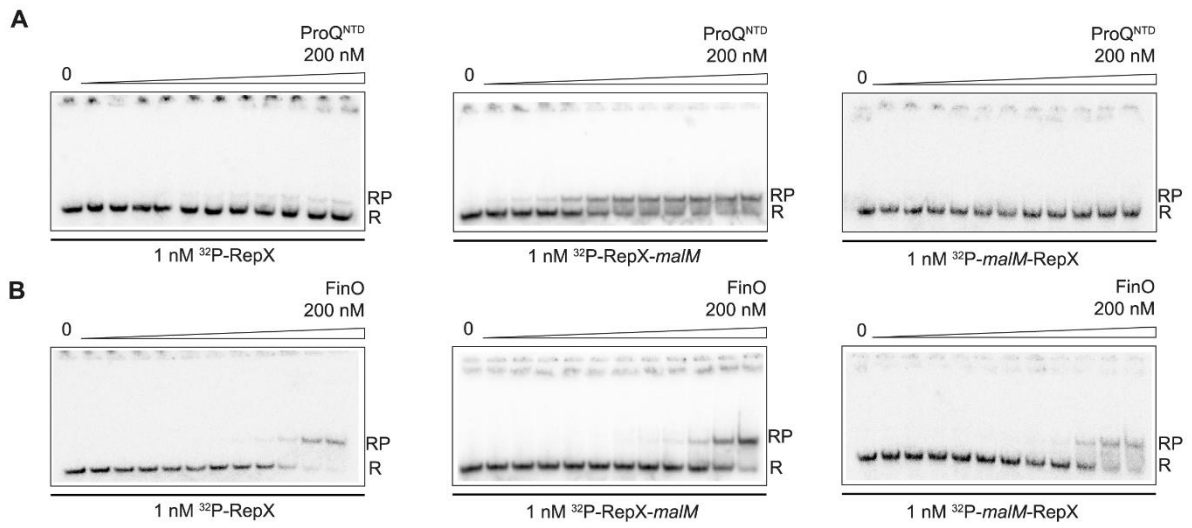
Supplemental Figure S12. Competition of ProQ and FinO proteins for binding to *malM*-3', *malM*-FinP, *malM*-A-GAU₅, FinP, FinP-*malM*, and FinP-U-U₆. The gelshift analysis and the corresponding plots of fraction bound data for the binding of ProQ and FinO to (A) *malM*-3', (B) *malM*-FinP, (C) *malM*-A-GAU₅, (D) FinP, (E) FinP-*malM*, and (F) FinP-U-U₆. The average fractions bound were calculated from at least three independent experiments.



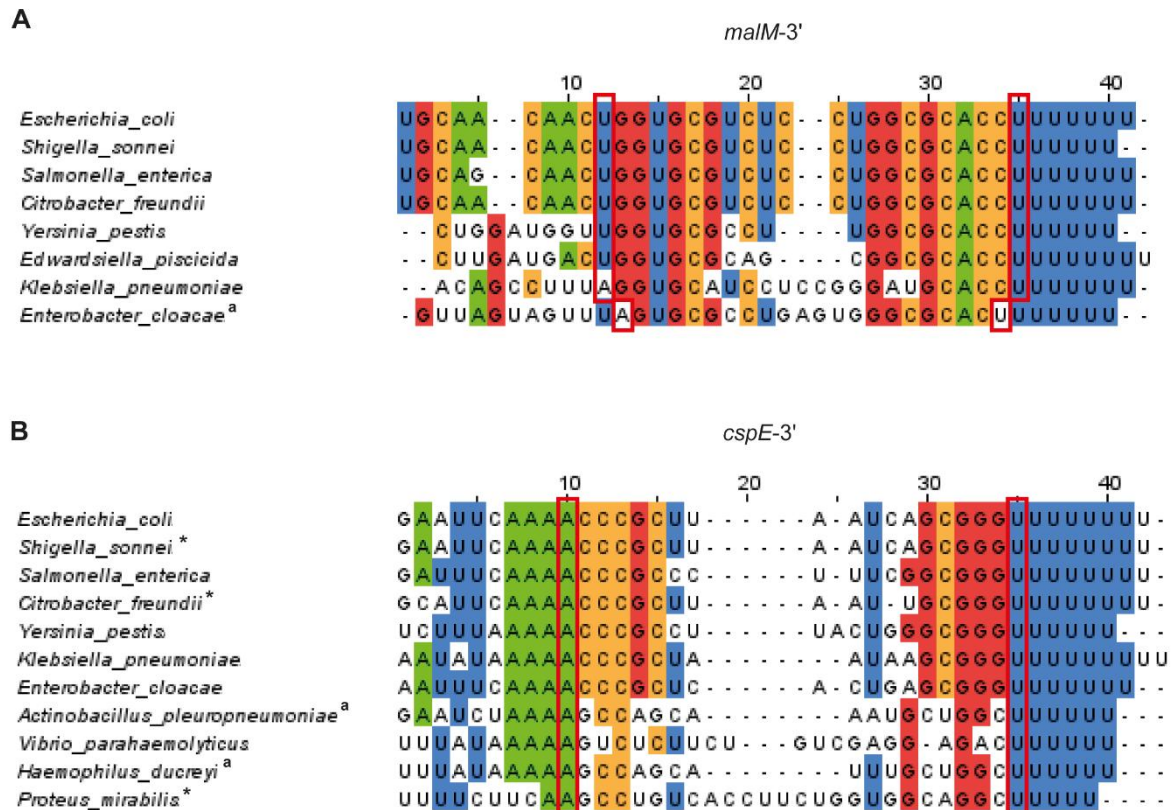
Supplemental Figure S13. Gelshift analysis of the binding of *cspE81-3'*, *cspE81-FinP* and *cspE81-FinP-stem* binding to ProQ^{NTD} (A), and FinO (B). The raw data in the gels correspond to the data presented in plots in Figure 5 in the main text. Free ³²P-labeled RNA is marked as R and RNA-protein complexes as RP.



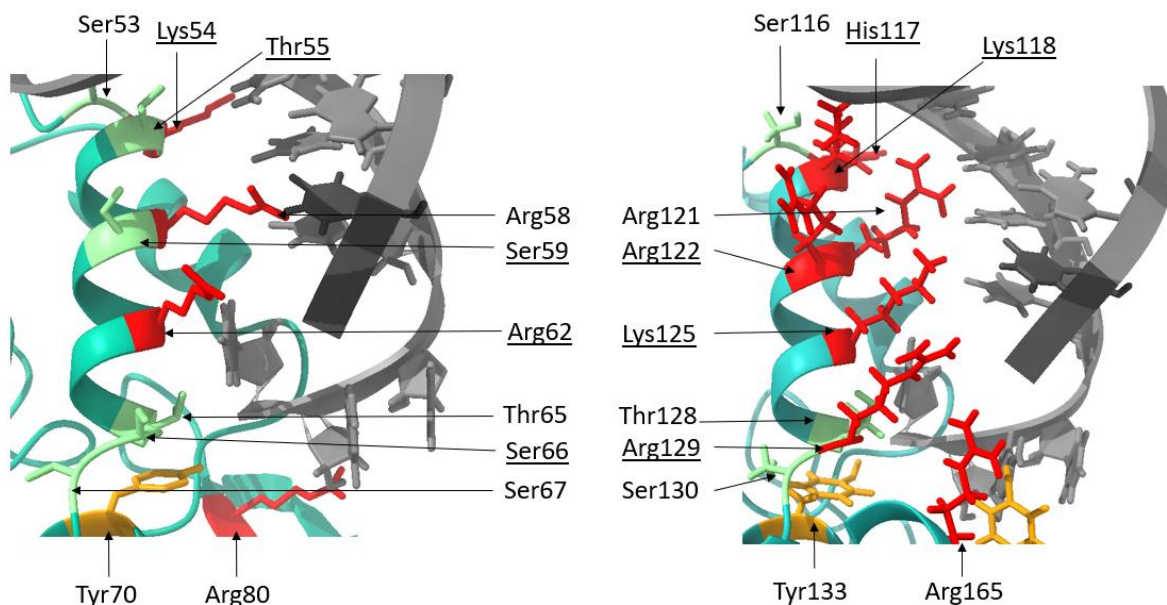
Supplemental Figure S14. Competition of ProQ and FinO proteins for binding to *cspE81-3'*, *cspE81-FinP-stem*, RepX, and *malM-RepX*. The gelshift analysis and the corresponding plots of fraction bound data for the binding of ProQ and FinO to (A) *cspE81-3'*, (B) *cspE81-FinP-stem*, (C) RepX, and (D) *malM-RepX*. The average fractions bound were calculated from at least three independent experiments.



Supplemental Figure S15. Gelshift analysis of the binding of RepX, RepX-*malM* and *malM*-RepX to ProQ^{NTD} (A), and FinO (B). The raw data in the gels correspond to the data presented in plots in Figure 6 in the main text. Free ³²P-labeled RNA is marked as R and RNA-protein complexes as RP.



Supplemental Figure S16. Alignment of the sequences surrounding intrinsic transcription terminators of *malM* and *cspE* mRNAs. The sequences used in the analysis were obtained from the NCBI database. For comparison, we selected regions including 10 nucleotides upstream of the terminator hairpin on the 5' side and the entire polyU tail on the 3' side. The opposing nucleotides located immediately below the closing G-C or C-G base pair of the terminator hairpin are highlighted with the red boxes. Nucleotides with $\geq 50\%$ identity are highlighted with colors. Sequences were aligned by Clustal Omega and visualized in Jalview. (A) Alignment of the sequences surrounding intrinsic transcription terminators in *malM* mRNA in different bacterial species. A gene homologous to *malM*, which was identified in *Enterobacter cloacae* using BLAST software is marked with "a" in the figure. This gene is listed in the NCBI database under the name A3UG_01455 (GenBank: CP003678.1). (B) Alignment of the sequences surrounding intrinsic transcription terminators in *cspE* mRNA in different bacterial species. Two genes homologous to *cspE*, which were identified using BLAST software are marked with "a" in the figure. The gene present in *Actinobacillus pleuropneumoniae* is listed in the NCBI database under the name D1099_00625 (GenBank: CP069796.1), and the gene present in *Haemophilus ducreyi* is listed under RZ68_03410 (GenBank: CP011230.1). In cases where two genes with the same name are present in the same species in the NCBI database, this has been indicated with the "*" sign. The gene selected for analysis starts at position 3,272,654 in *Shigella sonnei* (GenBank: CP176614.1), at position 3,657,114 in *Citrobacter freundii* (GenBank: CP174402.1), and at position 2,475,002 in *Proteus mirabilis* (GenBank: CP140127.1).



Supplemental Figure S17. The modeling of RNA binding surfaces in ColabFold-predicted structures of *E. coli* ProQ and F-like plasmid FinO. The figure shows the α -helix 3 from ProQ (left) and the corresponding α -helix 4 from FinO (right) with marked amino acid residues pointing towards modeled location of RNA helix. Both structures were predicted using ColabFold software (Mirdita et al. 2022). The modelling of interactions was done using Chimera X (Pettersen et al. 2021) by aligning the ColabFold-predicted structure of the FinO domain of *E. coli* ProQ (A) and the ColabFold-predicted structure of the F-like plasmid FinO protein (B) with the X-ray structure of the FinO domain of *L. pneumophila* RocC in complex with the terminator hairpin of RocR RNA (Kim et al. 2022). The side chains of amino acid residues located in the corresponding positions of both proteins are marked in color, with arginine, lysine and histidine residues marked in red, serine and threonine in green, and tyrosine in orange. The descriptions of corresponding amino acids are located in corresponding places on the figure. The descriptions of those amino acids, which are located in corresponding positions, but are different, are underlined. The structure of the *L. pneumophila* RocR hairpin is shown in grey.

SUPPLEMENTAL LITERATURE

- El Mouali Y, Gerovac M, Mineikaite R, Vogel J. 2021. In vivo targets of Salmonella FinO include a FinP-like small RNA controlling copy number of a cohabitating plasmid. *Nucleic Acids Res* **49**: 5319-5335.
- Holmqvist E, Li L, Bischler T, Barquist L, Vogel J. 2018. Global Maps of ProQ Binding In Vivo Reveal Target Recognition via RNA Structure and Stability Control at mRNA 3' Ends. *Molecular Cell* **70**: 971-982.e976.
- Kim HJ, Black M, Edwards RA, Peillard-Fiorente F, Panigrahi R, Klingler D, Eidelpes R, Zeindl R, Peng S, Su J et al. 2022. Structural basis for recognition of transcriptional terminator structures by ProQ/FinO domain RNA chaperones. *Nature Communications* **13**.
- Melamed S, Adams PP, Zhang A, Zhang H, Storz G. 2020. RNA-RNA Interactomes of ProQ and Hfq Reveal Overlapping and Competing Roles. *Molecular Cell* **77**: 411-425.e417.
- Mirdita M, Schutze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. *Nat Methods* **19**: 679-682.
- Pettersen EF, Goddard TD, Huang CC, Meng EC, Couch GS, Croll TI, Morris JH, Ferrin TE. 2021. UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci* **30**: 70-82.
- Reuter JS, Mathews DH. 2010. RNAstructure: software for RNA secondary structure prediction and analysis. *BMC Bioinformatics* **11**: 129.