Supporting Information

Sauer and Weichenrieder 10.1073/pnas.1103420108

SI Materials and Methods

Competition Experiments and Analysis of Chromatography Fractions. For competition experiments with the 5' Cy3-labeled 16-mer RNA (R16, 5' GCCACUGCUUUUCUUU 3'), Hfq (5 μM final concentration) was preincubated with annealed RybB RNA [7 μM RybB-OH or RybB-cP (cP, 2'-3' cyclic phosphate)] prior to the addition of Cy3-R16 RNA (10 μM final concentration). This mixture (150 μL) was incubated for 10 min at room temperature and subsequently analyzed by analytical size exclusion chromatography. The elution of Cy3-R16 was monitored selectively at 550 nm.

To monitor contaminating 5′ truncation products of RybB that were present in the chemically synthesized sample of RybB-OH, chromatography fractions were analyzed on gel. To this end, we used unlabeled R16 and increased sample concentrations (Hfq, 10 μM; RybB-OH or RybB-cP, 17.5 μM; R16, 25 μM), ensuring sufficient amounts for final analysis. The sequences of RNA addition and incubation times were identical to the experiments with Cy3-R16. Fractions (300 μL) were split for a separate analysis of RNA (200 μL) and protein (100 μL) content. RNA from each fraction was phenol-extracted and ethanol-precipitated (80% final) before analysis on a denaturing 15% polyacrylamide gel, containing 8 M urea. Protein was precipitated with trichloroacetic acid (10% final) before analysis on a 15% denaturing SDS-polyacrylamide gel. For visualization, RNA was stained with methylene blue and protein with Coomassie brilliant blue according to standard protocols.

Data Collection and Refinement Procedure. Diffraction data were collected at beamline PXII (X10SA) of the Swiss Light Source (Villigen, Switzerland) at a temperature of 90 K (see Table S1 for data collection and refinement statistics). Data of *Salmonella typhimurium* (*St*) Hfq72 were recorded at 0.827 Å on a MAR225 CCD detector, whereas data of the *St* Hfq72/U_6 complex were collected at 0.979 Å on a PILATUS 6 M detector (1). Images were processed with XDS (2).

The structure of *St* Hfq72 was solved by molecular replacement (MR). We searched for six copies of an *Escherichia coli* Hfq monomer (derived from Protein Data Bank ID 1HK9) using PHASER MR (3) from within the CCP4 package (4). The top solution was fed into ARP/wARP (5) for automated model (re)building. The model was finished manually in COOT (6), alternating with rounds of refinement using PHENIX (7).

Although the data for *St* Hfq72/U_6 merge perfectly well in space group P6, they were initially processed in space group P1 to allow for an entire Hfq hexamer per asymmetric unit and for asymmetry within. Molecular replacement (PHASER MR) was done with an entire *St* Hfq72 ring, followed by automated model (re)building (ARP/wARP) to remove any potential model bias. The protein model was finished manually and water molecules were added in COOT, alternating with rounds of refinement using PHENIX. Difference density was calculated from this model, clearly defining the RNA ligand. It reveals central "holes" for the rings of the base and of the ribose, and clearly defines the sugar pucker as C3′ endo. However, the 5′- and 3′-terminal riboses cannot be distinguished from internal ones in this density, indicating that the termini of the hexauridine substrate fail to pack the *St* Hfq72/U_6 complexes in a register that would be reflected by the crystal symmetry. Consequently, we merged the data in space group P6, corresponding to one Hfq monomer per asymmetric unit and refined the uridine ligand with covalent bonds to its nucleotide neighbors and with a phosphate occupancy of 5/6. This procedure leads to a circular, sixfold symmetric model for the entire hexauridine ligand. The absence of significant residual density indicates that the 5′- and 3′-terminal nucleotides are indeed oriented very similarly to their internal counterparts, and the refinement statistics show an excellent quality of the model. Thiocyanate ions were assigned based on the cone-shaped difference density, on suitable coordination distances, and on crystallographic evidence (i.e., refinement leads to temperature factors that are similar to the surrounding residues and to a good agreement with the crystallographic data).

Final refinement rounds for both structures were done in PHENIX. The refinement protocol included hydrogen atoms, anisotropic B factors for all non-hydrogen atoms, occupancy refinement of atoms in alternative conformations, and target weight optimization. Stereochemical properties were analyzed with MOLPROBITY (8) and WHATCHECK (9), and structure figures were generated in PYMOL (http://pymol.org/).

1. Broennimann C, et al. (2006) The PILATUS 1M detector. *J Synchrotron Radiat* 13:120–130.
2. Kabsch W (2010) XDS. *Acta Crystallogr D Biol Crystallogr* 66:125–132.
3. McCoy AJ, et al. (2007) Phaser crystallographic software. *J Appl Crystallogr* 40:658–674.
4. CCP4 (1994) The CCP4 Suite: Programs for protein crystallography. *Acta Crystallogr D Biol Crystallogr* 50:760–763.
5. Cohen SX, et al. (2008) ARP/wARP and molecular replacement: The next generation. *Acta Crystallogr D Biol Crystallogr* 64:49–60.
6. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486–501.
7. Adams PD, et al. (2010) PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213–221.
8. Davis IW, et al. (2007) MolProbity: All-atom contacts and structure validation for proteins and nucleic acids. *Nucleic Acids Res* 35:W375–383.
9. Hooft RW, Vriend G, Sander C, Abola EE (1996) Errors in protein structures. *Nature* 381:272.

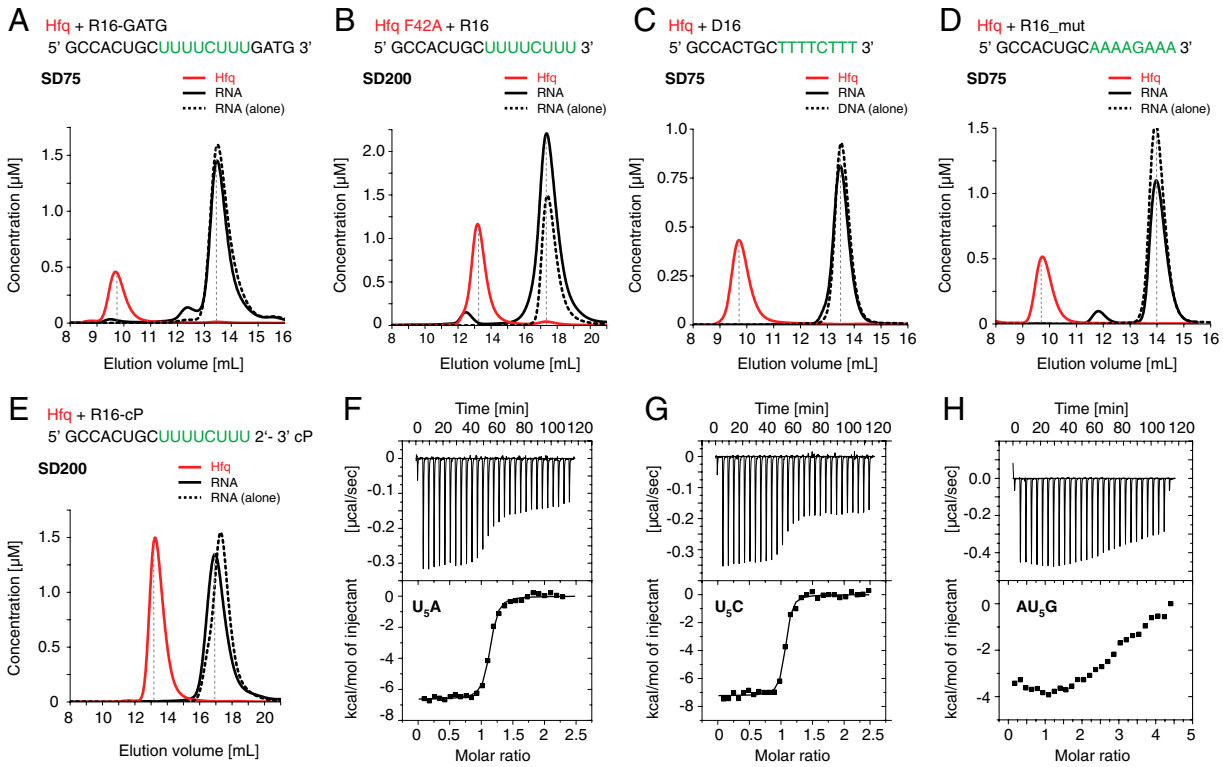

Fig. S1. Specificity of the proximal RNA binding site of *St* Hfq (*A–E*) Analytical size exclusion chromatography of Hfq (red, solid lines, 10 μM at start) in the presence of nucleic acid substrate (black, solid lines, 15 μM at start). Elution profiles show apparent concentrations for hexameric Hfq_6 rings and nucleic acid, calculated from the relative absorption properties of the components. Elution profiles for nucleic acid substrate alone (black, dashed lines, 15 μM at start) are superimposed. R16 RNA forms a stable equimolar complex with *St* Hfq_6 rings (see Fig. 1*B*), whereas the 3′ extension by a GATG sequence (*A*, R16_GATG) abolishes Hfq-binding as well as a mutation of the proximal RNA binding site (*B*, Hfq F42A). Furthermore, a DNA derivative of R16 fails to interact (*C*, D16), and complex formation is prevented also by other modifications of the RNA 3′ end. These are the mutation of the 3′-terminal uridines (*D*, R16_mut) and the terminal addition of a 2′–3′ cyclic phosphate (*E*, R16-cP). SD200, Superdex200 resin; SD75, Superdex75 resin. (*F–H*) Isothermal titration calorimetry with hexanucleotide substrates. Individual titration experiments with U_5A (*F*) and U_5C (*G*) substrates fit a model with a single high-affinity binding site (for analysis, see Fig. 1*D*), whereas a previously published (1) AU_5G substrate (*H*) does not. Upper panels show thermal power removed to maintain a constant temperature in the sample cell, recorded over the time of multiple RNA ligand injections until saturation. Lower panels show integrated binding enthalpy associated with every injection, plotted against the molar ratio of RNA over Hfq.

1 Schumacher MA, Pearson RF, Møller T, Valentin-Hansen P, Brennan RG (2002) Structures of the pleiotropic translational regulator Hfq and an Hfq–RNA complex: A bacterial Sm-like protein. *EMBO J* 21:3546–3556.

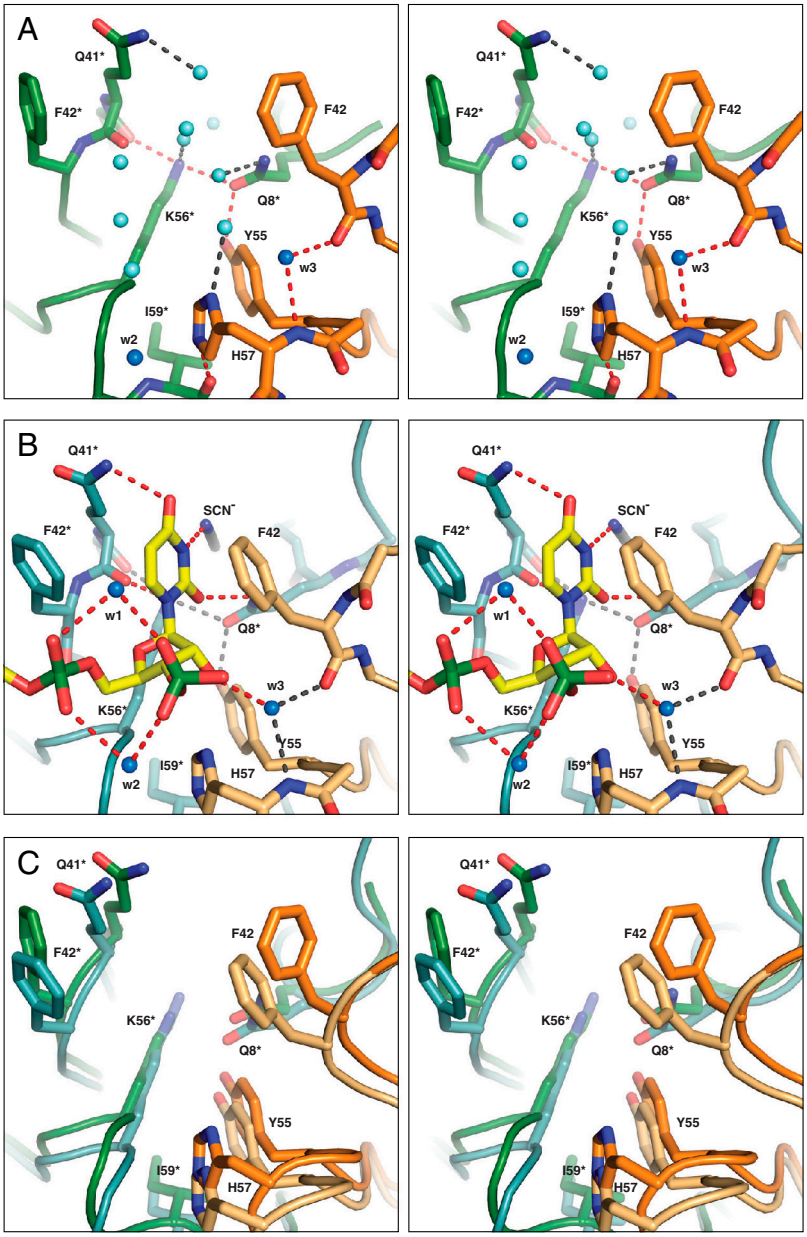

Fig. S2. Structural details of the RNA binding pocket in *St* Hfq (stereo). (*A*) Structure of the binding pocket in the absence of ligand (*St* Hfq72, apo-structure). Hydrogen bonds remaining upon ligand binding are red; displaced water molecules are cyan. (*B*) Structure of the binding pocket in the presence of ligand (*St* Hfq72/U_6). Newly established hydrogen bonds are red. (*C*) Superposition of the apo- and ligand-bound structures (ligands and water molecules removed), illustrating the adjustments of individual residues. RNA and important amino acids are shown as sticks, water molecules as spheres, and hydrogen bonds as dashed lines. Nitrogen, blue; oxygen, red; sulfur, gray; phosphorus, green.

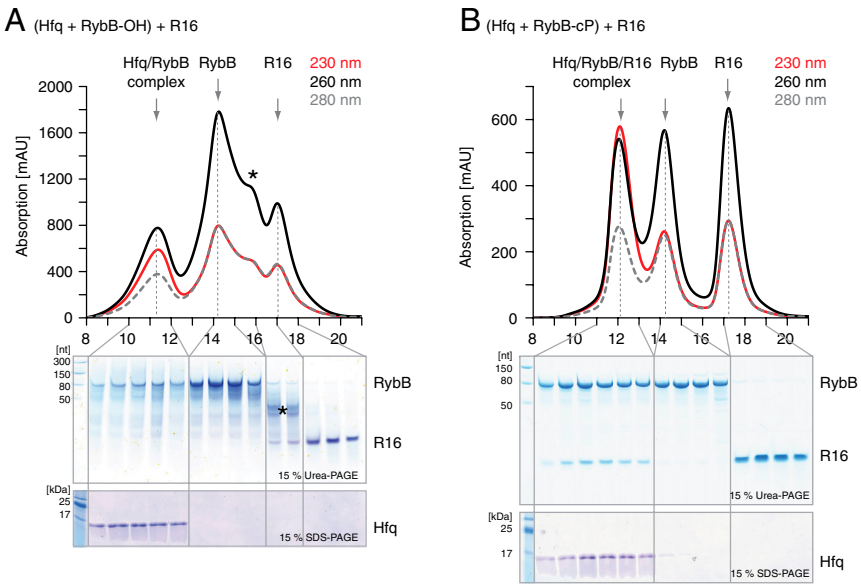

Fig. S3. Competition between an sRNA (RybB) and the R16 RNA oligonucleotide. Analytical size exclusion chromatography of Hfq in the presence of RybB RNA and of unlabeled R16 RNA. Elution volumes are indicated by gray vertical lines, together with the molecular species. Eluted fractions were split and analyzed separately for their RNA (top gel) and protein content (bottom gel). (*A*) Hfq_6 rings (10 μM) in the presence of RybB-OH (17.5 μM) and R16 (25 μM). Chemically synthesized RybB-OH has a free 3′ hydroxyl group and elutes with trailing peaks (5′ truncation products resulting from chemical synthesis, marked by an asterisk). RybB-OH coelutes with Hfq_6 and prevents R16 from binding Hfq_6. R16 does not interact with RybB-OH. (*B*) Hfq_6 rings (10 μM) in the presence of RybB-cP (17.5 μM) and R16 (25 μM). RybB-cP terminates in a 2′–3′ cyclic phosphate, but still coelutes with Hfq_6. However, it allows the coassociation of R16 to form a ternary complex (see Fig. 3).

Table S1. Data collection and refinement statistics

Dataset	*St* Hfq72	*St* Hfq72/U_6
Space group	$P6_1$	P6
Unit cell		
dimensions (*a*/*b*/*c*), Å	61.4/61.4/167.0	61.2/61.2/28.4
angles ($\alpha/\beta/\gamma$), °	90/90/120	90/90/120
Data collection*		
Wavelength, Å	0.827	0.979
Resolution range, Å	20–1.15 Å (1.18–1.15 Å)	30.6–1.3 Å (1.4–1.3 Å)
R_{sym}, %	6.7 (63.6)	7.7 (69.0)
Completeness, %	98.0 (96.9)	99.9 (99.9)
Mean $I/\sigma(I)$	12.81 (2.52)	16.6 (3.6)
No. of unique reflections	123,284	15,112
Multiplicity	4.9 (3.9)	9.1 (7.4)
Refinement		
Data range, Å	18.84–1.15	30.60–1.30
R_{cryst}, %	16.8	13.2
R_{free}, %	20.8	18.2
No. of atoms per asymmetric unit		
all atoms	7,229	1,179
Protein	6,853	1,093
Ligand/ion	0	30/6
Water	376	50
Average *B* factor, $Å^2$		
All atoms	15.6	19.4
Protein	13.9	18.1
Ligand/ion		20.8/24.9
Water	15.6	32.1
Ramachandran plot		
Favored regions, %	98.4	98.4
Disallowed regions, %	0.0	0.0
Rmsd from ideal geometry		
Bond lengths, Å	0.015	0.013
Bond angles, °	1.614	1.500

*Values in parentheses are for highest-resolution shell.

Table S2. sRNA 3′ ends and Hfq binding in *Salmonella typhimurium*

sRNA	Alternative IDs	5′ end	3′ end	Hfq*	Control*	Enrichment*	3′-terminal sequence†
Hfq-binding sRNAs (containing terminator sequences), *‡§							
micA	sraD	2,966,853	2,966,926	128	1	128.0	aucccuguuuucagcgaugaaauuuuGGCCACUCCGUGAGUGGCCUUUUUU
sroB	rybC	556,005	556,085	1,530	27	56.7	acgaugaaauuccucuuugacgGGCCAAUAGCGAUAUUGGCCAUUUUUUu
omrB	t59, rygA, sraE	3,170,408	3,170,322	52	1	52.0	gaucacacgaauuacaccAACCUGCGUAGAGAUGCGCAGGUUUUUUUUug
omrA	rygB	3,170,208	3,170,122	51	0	≥51.0	cuucguaccugucucuugcaccAACCUGCGCGGAUGCGCAGGUUUUUUUU
rydC	IS067	1,729,673	1,729,738	245	5	49.0	guccuccuuCGCCUGCGUCACGGGUCCUGGUUAGACGCAGGCGUUUUUCUg
sraH	ryhA	3,490,383	3,490,500	2,292	55	41.7	ugguguuggcgcaguauucgcgcACCCCGGUCAAACCGGGGUCAUUUUUUU
sroC	—	728,913	728,761	898	26	34.5	aaaucgccgggcaacaaucuucgAGGGUAGCGCUGCUACCCUUUUUUUuc
rybB	p25	942,632	942,554	103	3	34.3	ugugguAGCCCAUCAACCCCGCCAUUUCGGUUCAAGGUUGGUGGGUUUUUU
gcvB	IS145	3,135,317	3,135,522	402	12	33.5	ucuguccauagugauuaauguAGCACCGCCAUAUUGCGGUGCUUUUUUUu
invR	STnc270	3,044,924	3,045,014	3,236	113	28.6	aucaaucugcuuuuugauaCAGCAGCACCUCGCUGCUGCUUUUUUUuauuu
ryeB	tpke79	1,968,155	1,968,053	653	24	27.2	aagagccauuucccugGACCGAAUACAGGAAUCGUAUUCGGUCUCUUUUU
sraF	tpk1, IS160	3,392,069	3,392,261	25	0	≥25.0	gcauAAAAGCAACGGCUAAUGCCAUUCGGCGUUAGCCGUUUUUUUUUUUU
dsrA	—	2,068,736	2,068,649	149	6	24.8	cuucuugcauaagcaaguuugAUCCCGACCCGUAGGGCCGGGAUUUUUUU
cyaR	ryeE	2,231,130	2,231,216	659	31	21.3	ggaaccaccuccuugGCCUGCGUAAUCUCCCUUACGCAGGCUUAUUUUUU
sgrS	ryaA	128,574	128,812	61	3	20.3	gcgauagccuaaaaucacccGCCAGCAGAUAAUAUCUGCUGGCUUUUUUU
spf	—	4,209,066	4,209,175	33	2	16.5	ggcugaauauuuuagcCGCCCCAGUCAGUUUAUGACUGGGGCGUUUUUUa
micC	IS063, tke8	1,745,786	1,745,678	15	0	≥15.0	cuuuccagcguauaaauugacAAGCCCGAACGGAUGUUCGGGCUUUUUUU
micF	—	2,366,913	2,367,005	11	0	≥11.0	augucuguuuaccccuauuucAACCGGAUGCUUCGCAUUCGGUUUUUUUU
oxyS	—	4,342,986	4,342,866	10	0	10.0	uaacuaaagccaacgugaacuuuuGCGGACCUCUGGUCCGCUUUUUUUuuu
glmZ	k19, ryiA, sraJ	4,141,650	4,141,854	196	20	9.8	cauacagaccuguuuugACGCCUGCCCCUUAACCGGGCAGGCGUUUUUUU
rprA	IS083	1,444,938	1,444,832	286	37	7.7	gaauugcuguguguagucuuuGCCCGUCUCCUACGAUGGGCUUUUUUUuuu
ryhB-1	sraI, IS176	3,715,495	3,715,401	2	0	≥2.0	ucacauugcuuccaguauuauuuugGCCAGCUUUUGCUGGCUUUUUUUuuu
Hfq-independent sRNAs (containing terminator sequences)*‡							
csrB	—	3,117,059	3,116,697	67	69	1.0	ugcugcgaAACGAACCGGGAGCACUGUUUAUACAGUGCUCCCUUUUUUUU
csrC	sraK, ryiB, tpk2	4,210,157	4,210,400	64	63	1.0	gucgugaguca AGAAAAAAGGCGGCAGAUUACUCUGUCGCCUUUUUUCUU
glmY	tke1, sroF	2,707,847	2,707,664	92	20	4.6	uuaggcaccgccuuauuccauAACAAAGCCGGGUUUAGCCCGGCUUUGUU
istR-1	—	3,998,147	3,998,018	0	0		ggcuaccacccgcauuacgcuguAAGAACCUCGCUUCGGCGAGGUUUUUU
rydB	tpe7, IS082	1,450,415	1,450,519	10	4	2.5	cggaugggccggAACCCUUAAGCCUGUAUUGAACAGACUUAAGGGUUUU
sraL	ryjA	4,505,010	4,504,870	0	0		agacaaagaccggaaacAAACUAAAGCGCCCUUAGCGGCGCUUUAGUUUU
Hfq-independent sRNAs (processed 3′ends, missing terminator sequences)*‡¶							
ffs	SRP_RNA	524,423	524,536				ccaaggcagaugacgcgugugccggguguagcuggcagggccccccacc
rnpB	RNaseP_RNA	3,415,177	3,414,803				agaugaaugacugucc acgacagaacccggcuuaucggucaguuucacuu
ssrA	tm_RNA	2,843,947	2,844,309				cgaggauguaggaauuucggacgcggguucaacucccgccagcucccacca
ssrS	6S_RNA	3,222,098	3,222,280	451	836	0.5	uugaaccaaggguucaagggguuacagccugcggcggcaucucggagauuc

*Deep sequencing analysis of sRNAs associated with Hfq in *Salmonella typhimurium* (LT2). For data and additional references see Sittka et al. (2008) (1), table 2. Hfq: Number of blastable cDNA reads obtained from Hfq coimmunoprecipitation (Hfq coIP). Control: Number of blastable cDNA reads obtained from control coimmunoprecipitation (control coIP). Enrichment: Enrichment factor calculated by the number of blastable reads from Hfq coIP over control coIP.

†Terminal sequences are in uppercase letters. Terminal uridines are highlighted in bold when mostly single-stranded (six positions, tolerating C, A). Guanosines potentially interfering with Hfq binding are italicized. Nucleotides that are base-paired as part of the terminator hairpin are underlined. For sRNAs with processed 3′ ends, all base-paired nucleotides are underlined.

‡Compare also Zhang et al. (2003) (2) for a microarray-based analysis in *Escherichia coli*.

§sRNAs are sorted according to the enrichment on Hfq; only previously named sRNAs with an enrichment of ≥7.5 (plus ryhB-1) are listed.

¶5′ and 3′ ends are according to Li et al. (1998) (3).

1 Sittka A, et al. (2008) Deep sequencing analysis of small noncoding RNA and mRNA targets of the global posttranscriptional regulator, Hfq. *PLoS Genet* 4:e1000163.

2 Zhang A, et al. (2003) Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* 50:1111–1124.

3 Li Z, Pandit S, Deutscher MP (1998) 3′ exoribonucleolytic trimming is a common feature of the maturation of small, stable RNAs in Escherichia coli. *Proc Natl Acad Sci USA* 95:2856–2861.