The Sm-like protein Hfq (host factor Q-beta phage) facilitates regulation by bacterial small noncoding RNAs (sRNAs) in response to stress and other environmental signals. Here, we present a low-resolution model of *Escherichia coli* Hfq bound to the *rpoS* mRNA, a bacterial stress response gene that is targeted by three different sRNAs. Selective 2'-hydroxyl acylation and primer extension, small-angle X-ray scattering, and Monte Carlo molecular dynamics simulations show that the distal face and lateral rim of Hfq interact with three sites in the *rpoS* leader, folding the RNA into a compact tertiary structure. These interactions are needed for sRNA regulation of *rpoS* translation and position the sRNA target adjacent to an sRNA binding region on the proximal face of Hfq. Our results show how Hfq specifically distorts the structure of the *rpoS* mRNA to enable sRNA base pairing and translational control.

**Results**

**Hfq Binds A-Rich and U-Rich Motifs in *rpoS* mRNA.** We used SHAPE footprinting to identify Hfq interaction sites in the *rpoS* leader RNA. Previous experiments showed that the (AAN)_4 motif upstream of the sRNA target site binds the distal face of the Hfq and recruits Hfq to the *rpoS* mRNA (20, 22, 23). Hfq has the potential to also interact with a "U" loop motif (5′ UUAUUU) downstream of the sRNA target site (21, 24).

For footprinting experiments, we used *rpoS*301, a 284-nt variant of the 576-nt *rpoS* leader that lacks a nonessential upstream domain of the *rpoS* mRNA, how it remodels the *rpoS* mRNA to seed base pairing by a complementary sRNA (22), and why sRNA binding displaces Hfq from the inhibitory stem loop.

Here, we show that Hfq enables sRNA regulation by folding the *rpoS* mRNA leader into a specific tertiary structure that partially unwinds the inhibitory stem and poises Hfq to bring both RNAs together. Small-angle X-ray scattering (SAXS), functional assays, and SHAPE (selective 2'-hydroxyl acylation and primer extension) footprinting revealed that Hfq contacts three distinct sites in the *rpoS* mRNA, folding the 5′ leader of the *rpoS* mRNA into a compact structure. Three-dimensional models of the *rpoS*Hfq complex refined against the SAXS data show that the two domains of the *rpoS* mRNA wrap around the Hfq hexamer, placing the inhibitory stem over the arginine patch and adjacent to the sRNA binding sites on the rim and proximal face. These results demonstrate that multiple RNA binding surfaces on Hfq enable the protein to distort the structure of the *rpoS* mRNA, poising the complex for sRNA entry and translation.

**Significance**

Small noncoding RNAs optimize bacterial gene expression under stress and increase the virulence of many bacterial pathogens. The RNA-binding protein Hfq (host factor Q-beta phage) promotes base pairing between small RNAs and target mRNAs, but it is not known how Hfq brings the two RNAs together in the proper orientation. We used chemical footprinting, small-angle X-ray scattering, and molecular dynamics simulations to model the structure of Hfq bound to an mRNA in solution. The surprising result is that the mRNA wraps entirely around the Hfq protein, specifically contacting both surfaces. This destabilizes the mRNA structure around the small RNA target site, poising it to base pair with a complementary small RNA also bound to Hfq.
domain but retains the Hfq binding domain and inhibitory stem needed for translational control and Hfq and sRNA binding (24). The rpoS301 RNA folds homogeneously in vitro and retains the native secondary structure (Fig. S1 A and B) based on its similar SHAPE modification as the full-length rpoS leader (24). We compared the SHAPE modification levels of free rpoS301 RNA with rpoS301 RNA bound to DsrA sRNA or to Hfq (Fig. 1A and Fig. S1C). We then categorized the decrease or increase in relative SHAPE reactivity based on a histogram of the entire dataset (Fig. S1D), which reflects a change in the accessibility of the ribose 2′OH or the flexibility of the RNA backbone (25).

As expected, base pairing between rpoS mRNA and DsrA sRNA protected the DsrA binding site in the inhibitory stem from modification, reducing the SHAPE reactivity by ∼30–40% (Fig. 1B). The SHAPE reactivity of the upstream and downstream domains did not change appreciably, however, suggesting they are unaffected by DsrA (green trace in Fig. 1A and B).

By contrast, Hfq remodeled the rpoS mRNA structure extensively (magenta trace in Fig. 1A and C). First, the reactivity of the inhibitory stem and the helix connecting the (AAN)₄ and A₆ motifs increased two- to threefold over that of the free RNA. These residues were uniformly and moderately modified in the rpoS·Hfq complex, suggesting that Hfq partially opens the mRNA secondary structure. An Hfq-induced structural change in the inhibitory stem was also reported based on RNase footprinting experiments (26).

Second, Hfq binding resulted in unusually strong modification of three regions that we deduced make specific contacts with Hfq: the (AAN)₄ motif previously known to bind Hfq, the U₅ motif in the downstream domain, and A157 in the inhibitory stem near the 5′ end of the sRNA target site. The first A of every AAN triplet was four to nine times more modified in the Hfq complex than in the RNA control (A80, A83, and A85 in Fig. 1A) (also Fig. S1C). This hyperreactivity was explained by a structure showing that the A-specific pocket on the distal face
of the DsrA target site was determined by SHAPE experiments showing that Hfq cycles off the DsrA–rpoS mRNA duplex after the RNAs have base paired (21). The DsrA target site was ~50–80% less modified in the ternary complex than in the DsrA–rpoS complex (Fig. 1D), consistent with tighter DsrA–rpoS binding in the presence of Hfq (19). Meanwhile, nucleotides upstream of the Shine–Dalgarno sequence became two- to threefold more accessible in the ternary complex.

A U5 Motif Binds the Lateral Rim of Hfq. To test which surfaces of Hfq contact rpoS mRNA, we repeated the SHAPE experiments with Hfq mutants Y25D, R16A, and K56A that disrupt RNA binding in the presence of Hfq (19). When the (AAN)5 motif in the DsrA–rpoS complex was mutated, expression of the (AAN)5 ΔrpoS::lacZ fusions was reduced a further 50% (compare Δ2 and Δ3 in Fig. S2 A and B), showing that the (AAN)5, A50, and U50 motifs not only interact with different surfaces of Hfq, but also make distinct contributions to the regulation of rpoS translation by sRNAs and Hfq.

To investigate whether the U5 motif is important for DsrA annealing in vitro, we next measured the stability of the DsrA–rpoS complex, using native gel mobility shift assays (Fig. 3). Strikingly, the R16A rim mutation abolished interactions with the inhibitory stem and the U5 motif while leaving intact interactions with the (AAN)5 motif (Fig. 2B, orange trace). The lost hypermodification of the U5 motif (A194 and U195) suggested that this loop directly contacts the lateral rim of Hfq. Modification of A157 returned to the average level, and modification of C137, C140, and C165 increased approximately threefold (Fig. 2B, orange and gray traces), indicating that the perturbed interaction with the rim also changed the conformation of the inhibitory stem. Finally, the K56A mutant did not appreciably change the modification pattern (Fig. 2C), confirming that the proximal face of Hfq does not bind rpoS mRNA directly.

U5 Motif Binding at Hfq Rim Facilitates DsrA Annealing. The SHAPE results showed that the lateral rim of Hfq contacts the downstream U5 motif in the rpoS mRNA leader, whereas the distal face remains bound to the upstream (AAN)5 motif. To investigate whether the U5 motif is required for regulation of rpoS translation by Hfq and sRNAs, we replaced the UUAUUU loop with UCCGC (Fig. 3A, ΔU5), shortened the stem by 3 bp (Fig. 3A, ΔU3S), or enlarged the loop by 9 nt (Fig. 3A, ΔU3UL).

All three mutations in the U5 stem loop diminished the ability of DsrA and RprA sRNAs to up-regulate expression of full-length rpoS::lacZ fusions in the E. coli chromosome by 20–40% (Fig. 3B, green and gold bars). The magnitude of this effect was similar to that of mutating the upstream (AAN)5 and ΔrpoS motifs (Fig. S2A, Δ2), although the U5 mutations had a smaller effect on up-regulation by ArcZ sRNA (Fig. 3B, red bars). When the (AAN)5, A50, and U50 motifs were all mutated, expression of rpoS::lacZ fusions was reduced a further 50% (compare Δ2 and Δ3 in Fig. S2 A and B), showing that the (AAN)5, A50, and U50 motifs not only interact with different surfaces of Hfq, but also make distinct contributions to the regulation of rpoS translation by sRNAs and Hfq.

To investigate whether the U5 motif is important for DsrA annealing in vitro, we next measured the stability of the DsrA–rpoS complex, using native gel mobility shift assays (Fig. 3).
We titrated $^{32}$P-labeled rpoS mRNA with DsrA sRNA (0–2 μM) and quantified the total fraction of rpoS•DsrA and rpoS•DsrA•Hfq complexes as a function of DsrA concentration (Fig. 3C). Without Hfq, the ΔU$_3$ mutation did not change the strength of the DsrA–rpoS RNA interaction, suggesting that this mutation does not alter the structure of free rpoS mRNA (Table S1). With Hfq present, however, DsrA bound the ΔU$_3$ complex about twofold better than the WT rpoS•DsrA•Hfq complex, perhaps owing to better release of the downstream domain (Table S1).

We next measured the ability of Hfq to increase the rate of DsrA annealing with rpoS mRNA (Fig. 3D). Without Hfq, both WT and ΔU$_3$ rpoS mRNA base paired with DsrA at the same rate (0.03 min$^{-1}$). In the presence of Hfq, however, a lower proportion of ΔU$_3$ than WT rpoS mRNA annealed with DsrA during the first 30 s (Table S1). Thus, these results suggest that interactions between Hfq and the U$_3$ motif distort the rpoS mRNA conformation for efficient DsrA entry.

Further SHAPE footprinting results on the ΔU$_3$ mRNA confirmed that this defect in DsrA annealing was due to impaired Hfq binding at the U$_3$ motif, based on the loss of hyperreactivity at this position (Fig. S3A). The ΔU$_3$ mutation also lowered modification of the upstream (AAN)$_4$ motif by ~80%, consistent with an overall reduction in Hfq affinity (24). Surprisingly, we still observed strong modification of A157 in the rpoS•DsrA•Hfq complex, perhaps owing to be...

**SHAPE footprinting results on the ΔU$_3$ mRNA confirmed that this defect in DsrA annealing was due to impaired Hfq binding at the U$_3$ motif, based on the loss of hyperreactivity at this position (Fig. S3A). The ΔU$_3$ mutation also lowered modification of the upstream (AAN)$_4$ motif by ~80%, consistent with an overall reduction in Hfq affinity (24). Surprisingly, we still observed strong modification of A157 in the rpoS•DsrA•Hfq complex, perhaps owing to be...**
data and that suggested how Hfq enables sRNA regulation of *rpoS* translation.

We first built an all-atom model of the full-length *E. coli* Hfq hexamer by appending disordered N- and C-terminal residues to a crystallographic model of the stable Sm core [Protein Data Bank (PDB) ID: 4HT8] (32). We used the program SASSIE (33) to simulate conformations of the N and C termini that fitted the experimental SAXS data (Fig. S6A). In the best structures, the N termini (amino acids 1–5) projected from the center of the proximal face (Fig. 5A, purple), whereas the C termini were mostly oriented toward the distal face (Fig. 5A, pink). This distal orientation differs from the radial projection of the C termini in previous ab initio models (30) (Fig. S5C).

To model the tertiary structure of the free *rpoS* mRNA, we divided the *rpoS*301i sequence into six fragments, using our SHAPE-determined secondary structure as a guide (Fig. S7A). We generated structures for each fragment with MC-Sym (34) and arranged the fragments in space by rigid-body modeling (SASREF) (35) against the experimental SAXS data (Fig. S7B and SI Materials and Methods). In the resulting model, the upstream and inhibitory domains again form an L connected by a flexible hinge at nucleotides 128–129 (Fig. 5C). Because these domains likely sample different orientations in solution, we used this hinge as a pivot point in a SASSIE Monte Carlo simulation, which generated an ensemble of 27,427 structures spanning the experimental *R*ₚ (Fig. S6B). The best-fit structures from this ensemble resembled the initial L-shaped model.

**Structural Models of the *rpoS*-Hfq Complex.** We repeated this modeling procedure to visualize the structure of the *rpoS*-Hfq complex, using the scattering data from the 1:1 *rpoS*-Hfq, sample as an experimental constraint. We used a crystallographic structure of the Hfq core bound to *rA* as an experimental constraint. We used a crystallographic structure of the stable Sm core [Protein Data Bank (PDB) ID: 4HT8] (32). We used the program SASSIE (33) to simulate conformations of the N and C termini that fitted the experimental SAXS data (Fig. S6A). In the best structures, the N termini (amino acids 1–5) projected from the center of the proximal face (Fig. 5A, purple), whereas the C termini were mostly oriented toward the distal face (Fig. 5A, pink). This distal orientation differs from the radial projection of the C termini in previous ab initio models (30) (Fig. S5C).

To model the tertiary structure of the free *rpoS* mRNA, we divided the *rpoS*301i sequence into six fragments, using our SHAPE-determined secondary structure as a guide (Fig. S7A). We generated structures for each fragment with MC-Sym (34) and arranged the fragments in space by rigid-body modeling (SASREF) (35) against the experimental SAXS data (Fig. S7B and SI Materials and Methods). In the resulting model, the upstream and inhibitory domains again form an L connected by a flexible hinge at nucleotides 128–129 (Fig. 5C). Because these domains likely sample different orientations in solution, we used this hinge as a pivot point in a SASSIE Monte Carlo simulation, which generated an ensemble of 27,427 structures spanning the experimental *R*ₚ (Fig. S6B). The best-fit structures from this ensemble resembled the initial L-shaped model.

The resulting model (Fig. 5 E and F) showed the *rpoS* mRNA wrapped around the Hfq hexamer, with the U₅ motif on the proximal side of the rim opposite the second AAN triplet and A157 at the rim on the other side of the ring. Strikingly, this orientation projected the inhibitory stem across the proximal face of Hfq, with the sRNA complementary strand toward Hfq and the ribosome binding site away from Hfq. This wrapped structure necessitates a slight unwinding of the inhibitory stem, consistent with the moderate increase in SHAPE modification of this region when Hfq binds. Hfq may induce additional RNA conformational changes that are not captured by our rigid-body modeling procedure. Overall, the model explained how Hfq folds the *rpoS* mRNA into a more compact structure and why interactions with both the AAN motif and the U₅ motif are needed for efficient sRNA entry.

To determine whether other conformations also fit the SAXS data, we used SASSIE to vary the orientation of the downstream RNA domain about the flexible hinge (nucleotides 128–129). Structures of the *rpoS*-Hfq complex that best represent the data (χ² < 1.5, 917 structures) were symmetrically distributed about *R*ₚ = 55 Å (Fig. 5H) and collectively sampled a restricted wedge of space that could reflect an oscillatory path of the inhibitory stem in which the U₅ motif detaches and rebinds the Hfq lateral rim (Fig. 5G and Fig. S6D). This ensemble of “open” structures described the scattering data nearly as well as the initial “closed” structure (Fig. 5F).

In all of these structures, nucleotide A157 remained close to the Hfq rim, consistent with our SHAPE data showing that hypermodification of this residue in the inhibitory stem depends on the (AAN)₄ motif binding rather than the U₅ motif. By contrast, the sRNA annealing site, the ribosome binding site, and the U₅ motif moved away from Hfq in the more open structures.

**Discussion**  
Our SHAPE footprinting results, SAXS data, and all-atom models collectively show that Hfq folds the *rpoS* mRNA leader into a compact tertiary structure. This folded structure positions the inhibitory stem of the *rpoS* leader over the proximal face of Hfq where sRNAs are known to bind. This unexpected result explains many features of *rpoS* regulation by sRNAs and Hfq, such as how Hfq brings together the complementary regions of the mRNA and sRNA near the arginine patches along the rim and why sequences upstream and downstream of the sRNA target site are important. Moreover, our SHAPE results show that Hfq partially opens the secondary structure of the inhibitory stem to enhance sRNA annealing and ribosome binding (22, 26).

Remodeling of the *rpoS* mRNA requires interactions with both (AAN)₄ and U₅ motifs.

As the SAXS data do not provide information about local structure, our model cannot capture the details of the RNA–Hfq interactions. Moreover, the model does not account for local perturbations to the RNA structure. Nevertheless, the overall arrangement of the *rpoS* mRNA leader with respect to Hfq in our model is well supported by experimental data. First, the dramatic change in the scattering function provides direct physical evidence for compaction of the RNA by Hfq. Second, the marked change in RNA backbone modification (SHAPE) in response to the Hfq and AU₅ mutations is consistent with specific Hfq interactions, rather than nonspecific effects of the protein on the RNA structure. Unusually strong ribose modification may serve as a diagnostic for direct Hfq–RNA interactions. Third, mutational studies showed that the position and orientation of the (AAN)₄ and U₅ sequences are important for Hfq-mediated sRNA regulation, suggesting they bind Hfq simultaneously (Fig. S3) (32). Finally, an unbiased search of structural models indicated that only a subset of RNA conformations recapitulates the SAXS data (Fig. 5 G and H and Fig. S6D).

Our data show that Hfq folds the *rpoS* leader into a compact, closed conformation by simultaneously recognizing an upstream (AAN)₄ motif and downstream U₅ motif flanking the sRNA target site. In this closed mode, the inhibitory stem is partially melted, and the 5′ end of the target site interacts with the Hfq rim where we propose the arginine patch promotes base pairing with a complementary sRNA. The SHAPE data show that Hfq disengages from the downstream U₅ motif after a sRNA base pairs with the inhibitory stem, while remaining bound to the (AAN)₄ motif. The potential to form more open structures explains how the *rpoS* leader can flex to allow Hfq to cycle off the DsrA–*rpoS* duplex, exposing the ribosome binding site.

The potential for opening and closing the *rpoS*-Hfq complex is clearly captured in our structural models. The closed *rpoS*-Hfq model obtained by constraining the U₅ motif to interact with the Hfq rim was reasonably consistent with the SAXS data. However, the Monte Carlo simulations showed that more open structures fitted the scattering data equally well, even assuming a small fraction of free RNA. The *rpoS*-Hfq complex may fluctuate between open and closed conformations in solution. As the scattering curves for 2:1 Hfq-*rpoS* also indicate a folded structure, our data do not exclude models in which the open *rpoS* leader binds a second Hfq hexamer.

Although AAN sequences are known to recruit Hfq via its distal face (7, 8, 20), here we find that the U₅ motif in *rpoS* also contributes to sRNA annealing by interacting with the Hfq rim. This distorts the mRNA structure, making it more accessible to sRNAs (22). Multilateral Hfq interactions may be widespread among bacterial sRNA–mRNA pairs and important for regulation. The *flhA* mRNA leader was proposed to contact both distal and proximal faces of Hfq based on competitive binding experiments (36). Hfq inhibits translation of *cirA* by binding to an upstream (AAN) motif and two U-rich patches close to the Shine–Dalgarno sequence (37), raising the possibility that Hfq
also folds the cirA mRNA for translational control. Our results show that Hfq forms a specific, folded ppsO mRNA that springs loads the regulatory helix for sRNA entry.

Materials and Methods

SHAPE Footprinting. Complexes of 50 mM ppsO301 RNA, 333 mM E. coli Hfq hexamer, and 200 mM DsrA sRNA were prepared as previously described (24) in 10 μL annealing buffer [50 mM Tris HCl, pH 7.5, 50 mM NaCl, 50 mM KCl, 50 mM NaH2PO4, 2% (vol/vol) glycerol] at 25 °C for 2 h. Complexes were modified with N-methylsulfato anhydride (Molecular Probes) and analyzed by reverse transcription as described in SI Materials and Methods. Reported values of relative SHAPE reactivities are the average of at least three independent experiments.

Hfq Binding and Translational Activation. E. coli strains and β-galactosidase assays of ppsO-lacZ expression were performed as previously described (19, 24). Gel mobility shift binding assays with ~70 nM 32P-labeled ppsO301 RNA and DsrA or Hfq were performed in annealing buffer for 2 h at 25 °C as previously described (20, 24) before native 6% polyacrylamide gel electrophoresis in 66 mM Heps, 34 mM Tris, 0.1 mM EDTA, and 2 mM MgCl2.

SAXS, ppsO301 RNA and Hfq protein were prepared under native conditions as described in SI Materials and Methods. Small-angle X-ray scattering data were collected at room temperature at the Advanced Photon Source 12-ID-B, over the range 0.005 < q < 1.007 Å−1 after background subtraction. Data collected at three different sample concentrations showed the expected increase in I(0) and constant Rg and ratios of scattering intensity, indicating a lack of interparticle interactions (Fig. S4). Parameters of the fits and estimates of the particle mass are listed in Table S2.

Structural Models. Three-dimensional models of ppsO mRNA secondary structure fragments (Fig. S7) were generated using MC-Sym web server (34) and oriented in three dimensions with SASREF (35). The RNA chain connectivity and the SAXS experimental data as constraints. CORAL was used to model the full ppsO-Hfq complex against the SAXS data for the 1:1 RNA-Hfq complex (35). In 125 complexes, ppsO P 195 (U3 motif) and P 157 (inhibitory stem) were constrained to c12 or 15 Å, respectively, from R16 C6 in any Hfq monomer. Monte Carlo simulations were performed using the program SASSIE (33) to identify conformations of free Hfq, free ppsO mRNA, and the ppsO-Hfq complex consistent with the scattering data for each sample. The coordinates of the Hfq core were fixed during the simulations, whereas the N and C termini (amino acids 1–5 and amino acids 66–102) were allowed to move. The RNA was allowed to pivot between nucleotides 128 and 129. Whereas the residues between the best 91 models and the experimental data for the 1:1 ppsO-Hfq complex showed some positive serial correlation (Durbin–Watson <2), the magnitudes of the residues were on the order of the statistical error of the data (Fig. S52). See SI Materials and Methods for details of the modeling.

Certain commercial equipment, instruments, materials, suppliers, or software are identified in this paper to foster understanding. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the materials or equipment identified are necessarily the best available for the purpose.

ACKNOWLEDGMENTS. The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.

The authors thank X. Zuo and the APS 12-ID-B staff, S. Krueger, S. Gottesman, G. Storz, and Y.-X. Wang for helpful discussion and D. Kilburn, N. Majdalani, N. Kim, and T. Schlick for their assistance.
Supporting Information

Peng et al. 10.1073/pnas.1410114111

SI Materials and Methods

SHAPE Footprinting. Complexes of *E. coli* Hfq protein, *rpoS*301 RNA, and DsrA sRNA were assembled in 10 μL annealing buffer (50 mM Tris-HCl, pH 7.5, 50 mM NaCl, 50 mM KCl, 50 mM NH₄Cl, 2% glycerol) at 25 °C for 2 h. Reactions contained 50 nM *rpoS*301 plus 200 nM DsrA and 333 nM Hfq hexamer as stated in the text. Complexes were modified for 2 h at 37 °C with 1 μL 20 mM N-methylisatoic anhydride (Molecular Probes) dissolved in anhydrous DMSO and then analyzed by reverse transcription as previously described (1). cDNA products were quantified using SAFA (2) and normalized to reference bands that showed constant intensity in different lanes of the gel (3). The SHAPE reactivity of *rpoS* RNA complexes relative to *rpoS* RNA alone was calculated from the ratios of adjusted band intensities and ranged from 0.15 to 8.88 (Fig. S1D). Outliers with values 30–100 times above the background usually occurred next to very strong RT pauses and were manually excluded from the data. Each SHAPE reactivity profile is the average of at least three independent experiments. Error bars represent the SD from the mean. A histogram for the entire dataset was constructed with bin sizes equal to the average SD of the SHAPE reactivity, and the nucleotides were grouped and colored as illustrated in Fig. S1D. Secondary structure schemes were prepared with xrna (rna.ucsc.edu/rnacent/xrna/xrna.html).

SAXS. Hfq was purified as previously described (4). Concentrated protein was dialyzed twice against SAXS buffer (annealing buffer plus 2 mM MgCl₂) at a final concentration of 1.90 mg/mL (170 μM monomer). *rpoS*301 RNA was purified by a native 6% polyacrylamide gel in 1x THEM2, eluted from the gel overnight at 4 °C in SAXS buffer, concentrated by ultrafiltration (Amicon Ultra-15 centrifugal filter unit, 50 kDa), and washed five times with fresh SAXS buffer (5). The final concentration was 0.43 mg/mL (4.6 μM). Samples were wrapped on ice and stored at 4 °C before use. Remaining SAXS buffer from the sample preparation was used for diluting samples and measuring background scattering. Small-angle X-ray scattering data were collected at the Advanced Photon Source 12-ID-B, over the range 0.005 < q < 1.007 Å⁻¹ as described previously (5). Guinier fits and real space inversions were done using Primus and GNOM from the ATSAS software package (6). Other plots were generated using the ScAtter software package (bl1231.als.lbl.gov/scatter). SAXS data collected on fresh samples with different concentrations showed constant Rg and ratios of scattering intensity, indicating a lack of interparticle interactions (Fig. S4), although dynamic light scattering of frozen samples showed ~11% scattering from an RNA dimer. Estimates of the molecular mass by the method of Rambo and Tainer (7) were within 10% of the nominal value (Table S2). The unordered C terminus of Hfq may result in higher than expected estimates by this method (8).

MC-Sym. Three-dimensional models of *rpoS* mRNA secondary structure fragments (Fig. S7) were generated using MC-Sym web server (9). The RNA sequence and secondary structure based on SHAPE experiments were used as the input with default settings; two-stranded fragments were first connected with a GAAA tetraloop. The output structures were ranked using the MC-Sym web server tools. The top five predicted structures for each fragment did not differ substantially. The highest-scoring structure for each fragment was selected for rigid-body modeling, after removal of GAAA tetraloops and energy minimization with UCSF Chimera (10). The predicted structures were consistent with the experimental SHAPE data, with highly modified residues occurring in hairpin loops or kinks in the RNA backbone (Fig. S7). The “hinge” region of the structure (light blue in Fig. S7) is the least well determined, but even the model for this region is in reasonable agreement with the data. The two three-helix junctions (type C) (11) were also predicted using a knowledge-based method (12–14). This method returned the same stacking geometry as in our model for the inhibitory stem-loop domain (pink in Fig. S7), but a slightly different stacking geometry for the upstream three-helix junction (red in Fig. S7).

Rigid-Body Modeling of *rpoS* and *rpoS*Hfq Complexes. The tertiary structure of *rpoS* RNA was modeled by orienting the RNA by rigid-body modeling (SASREF) (15), using the SAXS experimental data as a constraint (χ² = 0.59). The RNA connectivity was enforced by setting the distance between adjacent phosphorous atoms to ≤ 7 Å (res 40–41, 57–58, and 127–128). To allow more flexibility in the fitting procedure, the distance between connecting phosphorus atoms was constrained to ≤ 8 Å in only one of the two strands in the inhibitory stem (res 162–163 or res 272–273 for connecting fragments 4 and 5, and res 184–185 or res 248–249 for connecting fragments 5 and 6). Because we could not model single-stranded regions of the upstream domain (res 73–86 and res 102–116) with MC-Sym, the missing sequences were initially built assuming an A-form conformation, and all of the fragments were manually connected into a continuous strand in UCSF Chimera (10). Structures resulting from repeated calculations were similar, with two structural domains oriented at ~90°. The predicted structures were not sensitive to changes in distance constraints.

CORAL was used to model the full *rpoS*•Hfq complex against the SAXS data for the 1:1 RNA:Hfq sample (16). An initial model of Hfq core bound to the AAN₄ motif was obtained by importing the coordinates of six protomers of *E. coli* Hfq in complex with *A*₇ RNA (PDB ID: 4HT8) (17). To this model, we added the missing C termini of Hfq (resi 66–102) and the *rpoS* fragments used for SASREF modeling of free *rpoS* RNA with the same contact restraints. The A7 RNA bound to Hfq was joined to the rest of the RNA by setting the distance between *rpoS*301 P88 and the 3’ end of A7 ≤ 7 Å. Two additional constraints were introduced based on the SHAPE footprinting data: *rpoS*301 P195 (U₃ motif) ≤ 12 Å from the α-carbon of R16 in any Hfq monomer and *rpoS*301 P157 (inhibitory stem) ≤ 15 Å from R16 α-carbon in any Hfq monomer. Finally, the missing nucleotides (nucleotides 73–80 and 102–116) were built and connected manually in UCSF Chimera.

Monte Carlo Simulations of Hfq, *rpoS*, and Hfq•*rpoS* Complexes by SASSIE. Monte Carlo simulations (SASSIE) (18) were used to identify conformations of free Hfq, free *rpoS* mRNA, and the Hfq•*rpoS* complex consistent with the scattering data for each sample. The coordinates of the Hfq core (alone or with tRNA bound to the distal face) (17) were fixed during the simulations, whereas the N- and C-terminal residues (res 1–5 and res 66–102) were allowed to vary. The initial full-length Hfq structure was energy minimized using CHARMM (19), before it was used as the input for the Monte Carlo simulation. During the Monte Carlo simulation, Crysol 2.7 (20) was used to calculate scattering profiles of simulated structures after renaming atoms to C, H, N, O, P, S to avoid reading errors in Crysol. The averaged SAXS profile of Hfq was interpolated (43 points, Δq = 0.005 from 0.005 to 0.017 Å⁻¹).
to 0.21) and was used to evaluate the theoretical scattering profiles. Models generated by SASSIE that best fitted the experimental SAXS data were minimized using CHARMM.

The free \( rpoS \) structure generated by rigid-body modeling was modeled using SASSIE as described above. However, the RNA was allowed to pivot around the flexible hinge connecting the upstream and downstream domains (res 128–129). The Monte Carlo dihedral sampling of RNA backbone configurations was carried out using CHARMm 36 force-field parameters (21) for \( \alpha, \beta, \gamma, \delta, \epsilon, \eta \) angles, using the same energetic sampling as described previously (18).

The RNA coordinates of the \( rpoS\cdot Hfq \) complex generated from rigid-body modeling and the full-length Hfq structure generated from SASSIE were merged into a starting structure for Monte Carlo simulations of the complex. CHARMm was used to minimize the energy in three steps by first restraining all RNA atoms, then restraining the RNA carbon and phosphate atoms, and then allowing all atoms to move. A SASSIE Monte Carlo simulation was carried out as above, using \( rpoS \) res 128–129 as a pivot point and allowing Hfq N and C termini to vary. The SAXS profile of the 1:1 \( rpoS\cdot Hfq \) sample was used to evaluate the theoretical scattering profiles, using \( \chi^2 \) (Eq. S1) as a measure of statistical goodness-of-fit,

\[
\chi^2 = \frac{1}{(N-1)} \sum_i \frac{(Q_{\text{exp}}(Q_i) - Q_{\text{calc}}(Q_i))^2}{\sigma_{\text{exp}}(Q_i)^2},
\]

in which \( Q_{\text{exp}}(Q_i) \) is the experimentally determined SAXS scattering curve, \( Q_{\text{calc}}(Q_i) \) is the value obtained using SASSIE, \( \sigma_{\text{exp}}(Q_i) \) is the experimentally determined \( Q \)-dependent variance, and the sum was taken over \( i = 1 \) to \( N \), with \( n = 43 \) grid points of momentum transfer \( Q \). Some values of \( \chi^2 < 1 \), indicating that either our estimate of \( \sigma_{\text{exp}}(Q_i) \) is too large or more likely that the there are fewer than \( N - 1 \) df in our dataset. To account for variance in the value of \( \chi^2 \) (3σ ~ ± 0.66) given the limited number of data points and the uncertainty in the true number of degrees of freedom in the data, we arbitrarily considered all models with \( \chi^2 < 1.5 \), which span the experimentally determined \( R_q \) value. An alternative measure of error \( V(r) \) (7) reached a minimum around a similar range of \( R_q \) values (Fig. S6). The best-fit structures were energy minimized using CHARMm. Final models were compared with the scattering data, using FoXS (22).

Fig. S1. SHAPE footprinting of rpoS mRNA with sRNA and Hfq. (A) Modification of ribose 2′-OH was quantified by primer extension and analyzed by sequencing gel electrophoresis. To determine the overall secondary structure, traces from primer extensions covering different regions of the rpoS301 RNA were overlapped. (B) rpoS301 secondary structure consistent with experimental SHAPE modification data (colors). The linker between the upstream and downstream domains is predicted by MC-Sym to fold into a short stem loop, in agreement with its low reactivity. Gray, no data. (C) Selected sequencing gel images showing hyperreactive residues induced by Hfq binding at the (AAN)₄ motif (Top Left) and the Us motif (Bottom Left). Hfq also opens the secondary structure at the A₆ motif (Top Right) but does not affect other single-stranded regions (Bottom Right). (D) Histogram of SHAPE reactivity relative to rpoS RNA for the entire dataset. Values below 0.6 (~3% of nucleotides) were considered protected; values above 1.75 (~10% of nucleotides). Nucleotides were clustered based on their relative SHAPE reactivity: dark to light blue, protected, 0.140–0.287 (~0.2%); 0.287–0.434 (~0.3%); 0.434–0.581 (~0.6%); white, unchanged, 0.728–1.757 (~87.1%); and light to dark red, enhanced, 1.757–2.051 (~4.5%); 2.051–2.639 (~3.4%); 2.639–3.521 (~1.5%); and 3.521–8.813 (~0.8%). RNA secondary structure schematics were drawn with XRNA (rna.ucsc.edu/rnaserver/xrna.html). (E) Structure of AAN triple bound to the distal face of Hfq (PDB ID: 3GIB) showing the short 3.8-Å distance between the first 2′-OH and the phosphate of the second A. This conformation is associated with strong reactivity toward electrophiles used for SHAPE chemistry (1).

The U5 motif functions independently of the AAN motif. (A and B) Full-length rpoS leader (576 nt) fused to lacZ was used to measure in vivo expression of β-galactosidase as in Fig. 3B. Δ2, upstream (AAN) and A6 motifs were replaced with GC-rich sequence (1); Δ3, ΔU5 plus Δ2 mutation; 366A18 and 441A18, an A18 insertion at positions 366 and 441 to rescue Δ2 rpoS function (2). (C and D) Native polyacrylamide gel mobility shift assay for DsrA binding to rpoS mRNA. (C) Equilibrium binding at different DsrA concentrations. Fraction bound was calculated from counts in each lane in rpoS•DsrA (RD) or rpoS•DsrA•Hfq (RDH) bands. (D) Binding kinetics, as in C. Samples were loaded at various times during electrophoresis.


Hfq specifically recognizes the U5 motif. Shown is comparison of relative SHAPE reactivity of WT rpoS•Hfq complex (gray trace) and U5 mutant rpoS•Hfq complexes (red trace). Mutations (highlighted in red lines) are as shown in Fig. 3A. These mutations were predicted by MFOLD (1) not to alter the rest of the RNA secondary structure.

Fig. S4. Solution scattering of Hfq and rpoS mRNA. Shown are SAXS scattering profiles of (A and B) full-length Hfq and (C and D) rpoS301 RNA at three concentrations (colored green, orange, and red) in solution. (B and D, Top) Guinier plot of averaged data for free Hfq and free rpoS301 RNA. (Bottom) The ratio of scattering intensity from two Hfq or RNA concentrations remained constant over the Guinier region, confirming the absence of interparticle interactions (shown here for the lowest two concentrations that differ by a factor of 2). (E) The averaged scattering curves for Hfq (magenta), rpoS RNA (purple), and 1:1 rpoS•Hfq complex (green) were used to calculate $R_g$ from the Guinier region (33.6 ± 0.5 Å, 68.1 ± 1.6 Å, and 58.0 ± 1.0 Å, respectively). (F) Guinier plot of 1:1 rpoS•Hfq complex.
Ab initio models of rpoS RNA, Hfq protein, and the rpoS•Hfq complex. (A and B) Five ab initio structures predicted by DAMMIF for (A) free rpoS RNA and (B) the rpoS•Hfq complex (1:1 molar ratio). The rpoS envelopes adopt an extended L-shaped conformation, consistent with the all-atom model (Fig. 5). Hfq binding folds rpoS RNA into a more compact conformation. DAMMIF assumes the electron density is evenly distributed in the complex and does not account for differential SAXS scattering intensity from Hfq protein and rpoS mRNA. As a result, the space occupied by Hfq protein appears empty in the molecular envelope. Averaging these ab initio models resulted in a significant loss of structural detail. (C) The averaged ab initio model of Hfq predicted by DAMAVER (1) is very similar to previous models based on SAXS data (2, 3), which assumed oblate P6 symmetry. Twenty DAMMIF bead models were averaged and DAMMIN was restarted to fit the experimental data. (D) P(R) distribution of free rpoS RNA (purple) and the rpoS•Hfq complex (green), showing the change in the average conformation of the particle. The mass of the RNA and Hfq particles was calculated from the molecular volume, Vc, using the method of Rambo and Tainer (4). Real space parameters for the scattering data are given in Table S2.

Fig. S6. Monte Carlo and molecular dynamics simulations of free Hfq and rpoS RNA. (A–C) Comparison of experimental SAXS data with Monte Carlo simulations of (A) Hfq (24,991 structures), (B) free rpoS mRNA (27,427 structures), and (C) rpoS•Hfq complex (19,132 structures), performed with the program SASSIE (18). In A, the conformations of the intrinsically disordered N- and C-terminal residues were varied until the space around the hexamer core was fully sampled. (B and C) Residues 128–129 were used as a pivot point to produce an ensemble of structures spanning the experimental $R_g$ for the free RNA (68.1 Å) and the complex (58 Å). The metric $V_r$ (7) (blue in C) reaches a minimum at similar $R_g$ values to $\chi^2$, but is less sensitive to variations in the structures. The “best-fit” structure in this ensemble is depicted as an example of the “open” conformation of the rpoS•Hfq complex (Fig. 5E). (D) Spatial distribution of downstream rpoS mRNA domain in simulated Hfq•rpoS RNA structures (gray surface). Models were aligned to Hfq and superimposed. One conformation in the trajectory is shown as a ribbon; RNA, violet; Hfq, yellow. (Top) In the full trajectory, the downstream rpoS mRNA domain sampled the entire space around the Hfq Sm core. (Bottom) Structures that best match the experimental SAXS data ($\chi^2 \leq 1.5$) are confined to a wedge of space around the proximal face of Hfq, as also shown in Fig. 5G. (E) A plot of the residual (calculated profile minus experimental profile) for each of the 917 structures of the complex with $\chi^2 \leq 1.5$. Although some positive serial correlation was observed (Durbin–Watson statistical test <2), the magnitude of the residuals was on the order of the statistical error of the data.
Fig. S7. Tertiary structures of rpoS301 fragments predicted by MC-Sym. (A) Using the SHAPE-determined secondary structure as a guide (Fig. S1), rpoS301 was divided into six fragments. The tertiary structure of each fragment was predicted by MC-Sym (9): an upstream four-way junction (red), a long helix connecting the (AAN)$_4$ motif and A$_6$ loops (orange), a short hairpin next to the (AAN)$_4$ motif (green), a linker region (blue), the inhibitory stem (purple), and the downstream four-way junction containing the U$_5$ motif (magenta). The single-stranded (AAN)$_4$ and A$_6$ motifs were assumed to be unstructured linkers, based on moderate SHAPE reactivity (gray). The predicted secondary structures were consistent with SHAPE modification (Fig. S1). Highly modified nucleotides are shown as spheres on the ribbons. (B) SASREF rigid-body model of free rpoS RNA. (Left) Initial output of SASREF with disconnected fragments arranged to satisfy proximity constraints and the experimental SAXS scattering. (Right) Final models were built after filling in missing fragments and connecting adjacent RNA modules. (C) CORAL rigid-body model of the rpoS•Hfq complex, as in B.

Table S1. Summary of rpoS•DsrA binding constants and annealing rate constants

<table>
<thead>
<tr>
<th>rpoS</th>
<th>DsrA $K_d$, nM</th>
<th>DsrA annealing rate, min$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-Hfq</td>
<td>+Hfq</td>
</tr>
<tr>
<td>WT</td>
<td>26 ± 2</td>
<td>0.7 ± 0.1</td>
</tr>
<tr>
<td>ΔU$_5$</td>
<td>23 ± 2</td>
<td>0.32 ± 0.04</td>
</tr>
</tbody>
</table>

Shown are binding equilibria and rate constants measured by native polyacrylamide gel electrophoresis, based on data in Fig. 3 C and D. Experiments were without Hfq or with 0.6 μM Hfq monomer. The binding constants were calculated from DsrA titrations; rate constants were obtained from the formation of DsrA•rpoS RNA binary complex plus DsrA•rpoS•Hfq ternary complex over time. Values ± SD are based on three or more independent trials. Amplitudes of kinetic phases are given in parentheses.
<table>
<thead>
<tr>
<th>Sample</th>
<th>μM</th>
<th>$R_g$ Å ±</th>
<th>I(0) ±</th>
<th>$R_g$ Å ±</th>
<th>I(0) ±</th>
<th>$R_{max}$</th>
<th>SAXS</th>
<th>Nominal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hfq</td>
<td>42.4</td>
<td>32.3</td>
<td>0.7</td>
<td>1.86E-02</td>
<td>4.6E-04</td>
<td>32.8</td>
<td>0.8</td>
<td>1.90E-02</td>
</tr>
<tr>
<td>84.8</td>
<td>33.7</td>
<td>0.9</td>
<td>3.95E-02</td>
<td>5.4E-04</td>
<td>34.0</td>
<td>0.4</td>
<td>3.95E-02</td>
<td>4.0E-04</td>
</tr>
<tr>
<td>170</td>
<td>34.3</td>
<td>0.6</td>
<td>8.05E-02</td>
<td>6.1E-04</td>
<td>35.2</td>
<td>0.3</td>
<td>8.10E-02</td>
<td>5.0E-04</td>
</tr>
<tr>
<td>Hfq avg</td>
<td>33.6</td>
<td>0.5</td>
<td>4.45E-02</td>
<td>34.2</td>
<td>0.4</td>
<td>4.40E-02</td>
<td>4.0E-04</td>
<td>120</td>
</tr>
<tr>
<td>rpoS RNA</td>
<td>1.2</td>
<td>63.8</td>
<td>2.0</td>
<td>3.95E-02</td>
<td>1.3E-03</td>
<td>72.0</td>
<td>2.0</td>
<td>4.20E-02</td>
</tr>
<tr>
<td>2.3</td>
<td>63.8</td>
<td>1.2</td>
<td>7.80E-02</td>
<td>1.0E-03</td>
<td>71.5</td>
<td>1.4</td>
<td>8.20E-02</td>
<td>1.4E-03</td>
</tr>
<tr>
<td>4.6</td>
<td>64.2</td>
<td>0.9</td>
<td>1.55E-01</td>
<td>2.0E-03</td>
<td>71.8</td>
<td>0.9</td>
<td>1.64E-01</td>
<td>1.8E-03</td>
</tr>
<tr>
<td>rpoS avg</td>
<td>66.0</td>
<td>1.6</td>
<td>8.97E-02</td>
<td>2.0E-03</td>
<td>76.7</td>
<td>2.7</td>
<td>9.60E-02</td>
<td>2.0E-03</td>
</tr>
<tr>
<td>1:1 Hfq:rpoS</td>
<td>2.3</td>
<td>58.0</td>
<td>1.0</td>
<td>1.38E-01</td>
<td>2.1E-03</td>
<td>61.1</td>
<td>0.7</td>
<td>1.40E-01</td>
</tr>
</tbody>
</table>

SAXS data for Hfq protein, rpoS301 RNA, and 1:1 [Hfq]:rpoS mixtures were acquired at room temperature over a momentum transfer range of $0.005 < q < 1.007 \, \text{Å}^{-1}$ at APS 12-ID-B as described in Materials and Methods. For the 1:1 complex, the sample contained 2.3 μM RNA and 13.8 μM Hfq monomer. The $P(r)$ real space parameters were calculated for $q = 0–0.5$, using GNOM. The particle mass was estimated from the SAXS scattering curves by the method of Rambo and Tainer (21) for $q = 0–0.3$ to reduce contributions of noise at high $q$. 

Peng et al. www.pnas.org/cgi/content/short/1410114111