Escherichia coli Hfq Binds A_{18} and DsrA Domain II with Similar 2:1 Hfq_6/RNA Stoichiometry Using Different Surface Sites†

Xueguang Sun and Roger M. Wartell*

School of Biology, Georgia Institute of Technology, Atlanta, Georgia 30332

Received November 18, 2005; Revised Manuscript Received February 17, 2006

ABSTRACT: Hfq is a RNA-binding protein in *Escherichia coli* that plays an essential role in post-transcriptional regulation of mRNAs by facilitating pairing of noncoding RNAs (ncRNAs) to mRNA target sites. Recent work has provided evidence that *E. coli* Hfq has two distinct RNA-binding surfaces. In this study, a comparative sequence−structure analysis of *hfq* genes in bacterial genomes was employed to identify conserved residues that may be involved in binding RNA. A covariance of residue properties at neighboring positions 12 and 39 and conserved surface residues with high propensities at binding sites of RNA-binding proteins suggested several sites for Hfq−RNA interactions. On the basis of these predictions, eight mutant Hfq proteins were produced and their interactions were examined with the 38 nucleotide (nt) domain II of DsrA ncRNA ($DsrA_{DII}$) and A_{18} by a gel-mobility shift assay, fluorescence anisotropy, and fluorescence quenching. Mutations on the proximal surface of Hfq had a small affect on Hfq binding to A_{18} ($\leq$2-fold), while the mutations Y25A and K31A on the distal surface decreased affinity to A_{18} by 100-fold in solution. Mutations F39A and R16A on the proximal surface reduced affinity to $DsrA_{DII}$ by 6−8-fold, while other mutations on the distal or proximal surfaces affected affinity to $DsrA_{DII}$ by $\leq$2-fold using the gel-mobility shift assay. The F39A/L12F double mutation partially regained the affinity for $DsrA_{DII}$ lost by the F39A mutation. The latter observation is consistent with the implied importance of an aromatic residue at position 12 or 39 suggested by the sequence covariance. Titration experiments indicate a 2:1 Hfq_6/RNA stoichiometry for the strong binding complexes of Hfq with either A_{18} or $DsrA_{DII}$ and suggests that RNA-induced dimer formation of Hfq_6 is a common feature of Hfq−RNA interactions.

Hfq is a highly conserved and abundant RNA-binding protein in bacteria. It was originally discovered as an host factor in *Escherichia coli* (also called HF-I) required for bacteriophage Qβ RNA replication (*1*, *2*). Later, it was shown to be a global regulator of bacterial metabolism because disruption of the *hfq* gene causes a pleiotropic phenotype (*3*). The broad impact of the protein appears to stem from its role in regulating the stability and/or translation of mRNAs from a number of genes that respond to environmental stress (*4*, *5*). Many regulatory noncoding RNAs, such as DsrA, OxyS, RprA, and Spot 42, require Hfq as an essential component in their regulation of mRNA translation (*6*−*9*). The functional interactions between these small RNAs and their target mRNA sites are greatly decreased in the absence of Hfq, suggesting a cooperative interaction between Hfq, the mRNA site, and the ncRNA (*10*, *11*). One mRNA that is regulated by several ncRNAs in conjunction with Hfq is the *rpoS* mRNA (*6*, *7*, *9*, *12*−*15*). This mRNA encodes the stationary phase σ factor σ^s of RNA polymerase. Hfq has also been shown to affect the in vivo stability and post-transcriptional expression of mRNAs from the *ompA*, *mutS*, *miaA*, and *hfq* genes (*4*, *16*). Hfq is also involved in regulating the addition of poly(A) tails onto mRNAs (*17*−*19*).

Sequence and structure analysis has shown that Hfq is a member of Sm protein family (*5*, *6*, *8*, *20*, *21*). This family of proteins includes the Sm and Sm-like (LSm) proteins found in eukaryotes and archea. Eukaryotic Sm proteins are involved in a variety of RNA-processing events including pre-mRNA splicing (*22*), telomere replication (*23*), and mRNA degradation (*24*). Members of this protein family share a highly conserved Sm motif that dictates a common folding domain, enabling Sm proteins to assemble into a donut-like heptamer (*25*, *26*). Structural studies show that Hfq also forms a donut-like oligomer but with six rather than seven subunits (*6*, *27*−*29*).

Crystal structures of complexes formed by the *Staphylococcus aureus* Hfq and *Archaeoglobus fulgidus* LSm proteins with 5−7 nucleotide (nt)[1] U-rich oligonucleotides reveal that these two proteins share similar features in binding to this sequence (*27*, *30*). In both cases, the RNA molecules are bound to one side of the protein (proximal surface, Figure 2) in a circular path near the central cavity. Most of the oligonucleotide bases stack adjacent to the periodically repeated side chain of a residue located in loop 3 of neighboring subunits (Tyr at position 42 in Hfq and His at position 37 in LSm). A short U-rich segment ($PuAU_{4-6}GPu$) is a common sequence within RNAs bound by Sm proteins (*31*). A U-rich segment is also observed and/or protected in

† This work was supported by funding from the College of Sciences, Georgia Institute of Technology.

* To whom correspondence should be addressed. Telephone: 404-894-8421. Fax: 404-894-2291: E-mail: roger.wartell@biology.gatech.edu.

[1] Abbreviations: nt, nucleotide; SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; UV, ultraviolet.

10.1021/bi0523613 CCC: $33.50
Published on Web 03/24/2006

a number of RNAs bound by *E. coli* Hfq such as DsrA, Spot42, RydC, and the RpoS mRNA target region (*5*, *8*, *10*, *32*, *33*). These observations suggest that an unpaired U-rich stretch, 4−6 nt long, is part of a RNA motif bound by Hfq.

While some structural details of Hfq−RNA binding have been revealed and multiple functional roles for Hfq in regulating gene expression have been demonstrated, the molecular mechanisms governing its functions remain unclear. Evidence indicates that Hfq can act as a RNA chaperone and modulate the RNA structure (*34*). Studies also show that Hfq may act as a "matchmaker" to facilitate RNA−RNA complex formation (*10*, *11*, *35*) or polyadenylation of mRNAs (*17*). How Hfq recognizes its RNA targets, the surface patches involved in binding RNA, and the stoichiometry of Hfq−RNA complexes is not well-understood.

A recent study by Mikulecky et al. suggested that *E. coli* Hfq has distinct interaction surfaces for DsrA and poly(A) sequences (*36*). Mutant Hfq proteins were generated with single-site changes on the proximal surface near the central cavity, on the proximal surface at a larger radius from the central cavity, and on the distal surface. Mutations changing Tyr at position 25 and Ile at position 30 on the distal surface decreased Hfq binding to A_{27} by about 10-fold using a gel-shift assay, but they had little affect on binding to DsrA. Two cavity mutants (Y55A and K56A) showed decreased binding to DsrA. Several other mutations on the proximal surface (e.g., F42A, Q8A, and Q41A) did not have a significant effect on Hfq binding to DsrA using a gel-shift assay. Because the U-rich stretch of DsrA is protected from nucleases by *E. coli* Hfq binding and studies with *S. aureus* Hfq and *E. coli* Hfq imply that a U-rich segment binds to the proximal surface, other contacts between DsrA and the *E. coli* Hfq proximal surface may exist.

To identify potential RNA-binding sites on Hfq, a comparative sequence and structure analysis of over 300 bacterial *hfq* genes was carried out. Residues with a high propensity at protein−RNA binding sites that were highly conserved and mapped onto the Hfq surface were determined. The influence of several of these residues on RNA binding was examined by generating eight mutant Hfq proteins and characterizing their interactions with two target RNAs: the 38 nt domain II of DsrA ($DsrA_{DII}$) shown to bind Hfq (*33*) and A_{18}. The gel-mobility shift assay, fluorescence anisotropy, and fluorescence quenching were employed.

Under conditions in which Hfq is a hexamer in solution, titration experiments indicated that Hfq binds to $DsrA_{DII}$ or A_{18} with a 2:1 Hfq_6/RNA stoichiometry. This stoichiometry is the same as that previously noted for the strong binding complex with the complete 87 nt DsrA and a 140 nt segment of the RpoS mRNA (*10*). These results suggest a common feature of Hfq binding to RNA. Mutations on the proximal surface of Hfq had a small influence on Hfq binding to A_{18} ($\leq$2-fold), while two mutations on the distal surface Y25A and K31A decreased affinity to A_{18} by approximately 100-fold in solution. This result confirms that the distal surface is specific for polyA binding (*36*) and shows that Lys at position 31 is an important element in this binding specificity. Mutations on this patch of the distal surface may influence the function of Hfq in regulating polyA addition to mRNAs (*37*).

Mutations F39A and R16A on the proximal surface reduced Hfq affinity to $DsrA_{DII}$ by 6−8-fold using the gel-mobility shift assay, while other mutations on the proximal and distal surfaces affected the apparent K_d by $\leq$2-fold. Hfq with a F39A/L12F double mutation partially regained the affinity for $DsrA_{DII}$ lost by the F39A mutation. The latter observation is consistent with the implied importance of an aromatic group at this site based on the covariance observed at these two positions in *hfq* genes. Differences in affinity between wild-type (wt) and mutant Hfq for $DsrA_{DII}$ were not, however, detected using fluorescence anisotropy.

MATERIALS AND METHODS

Sequence Analysis and Structural Modeling. BLAST (*38*) searches were carried out against the Microbial Genome database at NCBI (http://www.ncbi.nlm.nih.gov/cig-bin/Entrez/genome_table_cgi), using *E. coli* Hfq or *A. fulgidus* Sm-1 sequences as the query. The multiple alignments were constructed by CLUSTAL W (*39*) using the output of BLAST searches. Mapping the location of residues onto the *E. coli* Hfq surface employed the crystal structure determined by Sauter et al. (*28*).

Purification and Characterization of wt and Mutant Hfq. The Impact-CN intein system (New England Biolabs, Beverly, MA) was used to purify Hfq proteins. The *E. coli hfq* gene was amplified by PCR using oligonucleotides EC-N (5′-GGTGGTTGCTTCCAACATGGCTAAGGGGCAATCTTTACAAGATC-3′) and EC-C (5′-TTATTCGGTTTCTTCGCTGTCCTGTT-3′) as primers and *E. coli* chromosomal DNA as the template. PCR products were digested with *Sap*I and cloned into a *Sap*I−*Sma*I digested pTYB11 plasmid. Protein purification was carried out according to the recommendation of the manufacturer using strain ER2566. The lysis/wash buffer that was used contained 20 mM Tris (pH 8.3) and 1 M NaCl. Triton X100 was added at 0.1%. The chitin column was extensively washed with this buffer prior to incubation of the column with this buffer plus 40 mM dithiothreotol. The eluted protein was concentrated and buffer-exchanged to 0.5 M NaCl and 20 mM Tris at pH 8.3 using centrifugation filtration units.

Plasmids containing mutant *hfq* genes were generated using the QuikChange Mutagenesis Kit from Stratagene according to the instructions of the manufacturer. The mutations F39A, F42A, F39A/F42A, F39A/L12F, Q8A, R16A, Y25A, and K31A were produced in the *hfq* gene using the following oligonucleotides: F39A, 5′-GGGCAAATCGAGTCTGCTGATCAGTTCGTGATCCTGTTG-3′ and 5′-CAACAGGATCACGAACTGATCAGCAGACTCGATTTGCCC-3′; F42A, 5′-CGAGTCTTTTGATCAGGCCGTGATCCTGTTGAAAAACACGG-3′ and 5′-CCGTGTTTTTCAACAGGATCACGGCCTGATCAAAAGACTCG-3′; F39A/F42A, 5′-CGAGTCTGCTGATCAGGCCGTGATCCTGTTGAAAAACAC-3′ and 5′-GTGTTTTTCAACAGGATCACGGCCTGATCAGCAGACTCG-3′; L12F/F39A, 5′-CAATCTTTACAAGATCCGTTCTTCAACGCACTGCGTCGGG-3′ and 5′-CCCGACGCAGTGCGTTGAAGAACGGATCTTGTAAAGATTG-3′; Q8A, 5′-GGCTAAGGGGCAATC TTTAGCAGATCCGTTCCTGAACGC-3′ and 5′-GCGTTCAGGAACGGATCTGCTAAAGATTGCCCCTTAGCC-3′; R16A, 5′-5CCTGAACGCACTGGCTCGGGAACGTGTTCC-3′ and 5′-GGAACACGTTC-

CCGAGCCAGTGCGTTCAGG-3′; Y25A, 5′-GAACGTGTTCCAGTTTCTATTGCTTTGGTGAATGGTATTAAGCTG-3′ and 5′-CAGCTTAATACCATTCACCAAAGCAATAGAAACTGGAACACGTTC-3′; and K31A, 5′-TTTATTTGGTGAATGGTATTGCGCTGCAAGGGCAAATCGAGTC-3′ and 5′-TTTATTTGGTGAATGGTATTGCGCTGCAAGGGCAAATCGAGTC-3′. Plasmid constructs were verified by DNA sequencing. The same procedure used to purify wt Hfq was used to purify the mutant proteins.

All proteins showed the 11 kD monomer band by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE) with purity estimated to be ~95% from Commassie blue staining. A faint band at 66 kD was occasionally observed. It has been attributed to incomplete denaturation of the hexamer prior to loading the gel lane (*40*). Silver staining showed no other bands. Analytical sedimentation velocity centrifugation showed that Hfq exists as a hexamer species (~97%) in 0.5 M NaCl and 20 mM Tris (pH 8.3) (Correia, J., Sun, X., and Wartell, R. M., unpublished data). This solvent was used in the RNA-binding studies described below. Ultraviolet (UV) absorbance at 276 nm of Hfq in 5.5 M guanidinium hydrochloride and aqueous solution was used to determine molar concentrations (*41*, *42*). Extinction coefficients of 3850 M^{-1} cm^{-1} at 280 nm or 4250 M^{-1} cm^{-1} at 276 nm were employed for all Hfq proteins except Hfq-Y25A. For Hfq-Y25A, the extinction coefficient employed was 2833 M^{-1} cm^{-1} at 276 nm. For some Hfq preparations, a DEAE column was employed to remove contaminating 260 nm absorbing material (*43*).

UV spectra were also used to assess RNA contamination. Pure Hfq was assumed to have a spectrum predicted empirically from the composition of its amino acids and studies with previous proteins (*42*). The RNA contaminant was assumed to be A_{18}. Using experimental Hfq absorbance values at pairs of wavelengths from 255 to 280 nm, a two-component analysis gave 1−6% RNA. Because a similar analysis of the UV spectrum of ribonuclease A gave similar results, RNA contamination appeared minimal. CD spectra from 200 to 300 nm were similar for the Hfq proteins and consistent with a previous spectrum (*33*).

RNA Oligonucleotides and Gel-Mobility Shift Assay. RNAs were purchased commercially and purified by HPLC. The gel-mobility shift experiment and fluorescence anisotropy experiments utilized A_{18} with 6-carboxyfluorescein (FAM) linked to the 5′ end and the 38 nt $DsrA_{DII}$ (AACGAAUUUUUUAAGUGCUUCUUGCUUAAGCAAGUUUC) with Oregon Green-514 (OG-514) linked to its 5′ end. RNA concentrations were determined by absorbance readings at 260 and 494 nm. For some gel-mobility shift assays, 10 μL of FAM−A_{18} or OG-514−$DsrA_{DII}$ was mixed with the indicated amounts of wt or mutant Hfq proteins to give a total volume of 20 μL in a binding buffer of 20 mM Tris (pH 8.3) and 0.5 M NaCl. The $DsrA_{DII}$ was heated for 3 min at 85 °C and quick-cooled on ice for 10 min prior to mixing with the proteins. Reactions were incubated at 25 °C for 10 min, whereupon 2 μL of loading buffer (0.25% bromophenol blue, 0.25% xylene cyanol FF, and 40% glycerol) was added. Samples were run on a 6% PAG in 0.5× TBE with 4% glycerol at room temperature at 120 V. After electrophoresis, the gel was scanned using a Fujifilm Image Reader FLA-3000 using the excitation and emission modes of 473 and 520 nm, respectively. Band intensities were evaluated using AlphaImager 950 software. Gel-mobility shift experiments were also carried out with ^{32}P 5′-end-labeled $DsrA_{DII}$. The $DsrA_{DII}$ molecule was 5′-end-labeled using ^{32}P-ATP and T4 polynucleotide kinase and purified from the ATP using a Bio-Spin P-30 column (Bio-Rad, Inc.). The concentration of the purified RNA was determined by comparing its band intensity with the labeled unpurified RNA of a known concentration.

Fluorescence Measurements. Fluorescence measurements were performed on an ISS spectrofluorimeter. The solution that was employed contained 20 mM Tris (pH 8.3) and 0.5 M NaCl. Emission spectra of the intrinsic fluorescence of Hfq proteins because of its tyrosines were scanned from 290 to 390 nm with excitation at 277 nm. Fluorescence quenching experiments were carried out with 50 μL of wt or mutant Hfq at 5 μM (in hexamer) and serially adding 2 μL aliquots of 10 μM (unlabeled) A_{18} until saturation of quenching was reached. The observed intensities at 305 nm were corrected for the loss of signal because of dilution. If F^0 is the corrected fluorescence of the free protein and F is the signal when an amount of A_{18} is added, the percent quenching, Q, is expressed as $Q = (F^0 - F)/F^0$. The possibility of the inner-filter effect disturbing the measurement was examined by adding aliquots of A_{18} to a tyrosine solution with similar fluorescence intensity as the Hfq solution. No change in the tyrosine emission intensity was observed. The titration of Hfq with a DNA oligomer also gave no significant quenching (<5%).

Fluorescence anisotropy measurements were carried out at room temperature (25 °C) in the 20 mM Tris (pH 8.3) and 0.5 M NaCl solvent. The L format was employed with the excitation monochromator at 490 nm and emission monochromator at 517 nm. Anisotropy values were obtained from the average of 10 iterations, using an integration time of 10 s for each measurement. Slits (2 mm) were employed (16 mm bandwidth). The wt and mutant Hfq proteins were serially titrated into a 3 mL cuvette that contained 2 mL of 2 nM FAM−A_{18} or 5 nM of 5′-OG-514-labeled $DsrA_{DII}$. The total fluorescent intensity and emission spectra of A_{18} showed no significant change with Hfq binding after accounting for the dilution. The $DsrA_{DII}$ molecule showed a 3−5% decrease in intensity in addition to the dilution effect after adding 600 nM Hfq.

Analysis of Fluorescence Anisotropy Data. Two models were employed in the analysis of the equilibrium binding of Hfq with the RNA molecules. One model assumed a 1:1 complex forms between a Hfq hexamer and RNA molecule. An equation describing the fluorescence anisotropy in terms of the dissociation constant K_d and other parameters of the experiment can be derived (*44*) and is given by eq 1

$$A = A_f + (A_b - A_f)\{\beta - [\beta^2 - 4R_tP_t]^{1/2}\}/2R_t \quad (1)$$

where $\beta = R_t + P_t + K_d$. A is the measured anisotropy of the fluorescent RNA during the titration; A_f and A_b are the anisotropy of the free and bound RNA, respectively; and R_t and P_t are the total concentrations of RNA and the Hfq hexamer, respectively. A nonlinear least-squares fit of the equation to the data was made fitting the parameters A_b and K_d.

The second model assumed that the Hfq hexamer binds RNA in a two-step reaction. The binding reaction is described by a dissociation constant K_1 for binding one Hfq hexamer to RNA and a dissociation constant K_2 for binding a subsequent Hfq hexamer

$$R + P \rightleftarrows RP \text{ with } K_1 = (R)(P)/(RP) \quad (2a)$$

$$RP + P \rightleftarrows RP_2 \text{ with } K_2 = (RP)(P)/(RP_2) \quad (2b)$$

Data were fit to the second model using BIOEQS (*45*, *46*). This numerical algorithm performs a least-squares fit to the anisotropy data with the parameters corresponding to the standard state free energies related to K_1 and K_2 and the anisotropies of the free RNA, the 1:1 complex, and the 2:1 complex. In general, the anisotropy of the free RNA was fixed to the experimental value, and the remaining four parameters fit to the data.

RESULTS

Sequence−Structure Analysis: *Identifying Potential RNA-Binding Sites in Hfq Proteins*. Previous work has shown that RNA-binding proteins divide into two main classes based on their mode of RNA recognition (*47*). Groove binding proteins were designated as class I and position a secondary-structure element such as an α helix or a loop into a groove of an RNA helix. β-Sheet-binding proteins, designated as class II, utilize β-sheet surfaces to create binding pockets that bind unpaired RNA bases. Hfq is a β-sheet-rich protein (a α helix followed by five β strands), suggesting that its RNA-binding sites have characteristics of class II RNA-binding proteins. This form of binding was observed in the crystal structure of the AU_5G oligomer with the *S. aureus* Hfq (*27*). Statistical analysis of 32 protein−RNA complexes indicates that RNA-binding proteins have a preference for contacting guanine and uracil and that the residues lysine, tyrosine, phenylalanine, isoleucine, and arginine have the highest propensities at RNA-binding sites (*48*).

Using the above observations as guidelines, we sought to identify potential RNA-binding sites of *E. coli* Hfq by determining surface locations containing residues with high propensities at RNA-binding sites (i.e., Lys, Phe, Tyr, and Arg) and that are highly conserved among bacterial Hfq proteins. Inherent in this analysis are the assumptions that Hfq binds RNA using its hexamer surfaces, the structures of different Hfq are very similar, and the RNA-binding sites of different Hfq proteins are at the same locations. A multiple alignment of Hfq sequences was built based on a search of over 300 completed or partially completed eubacterial genomes available at the NCBI database using BLAST. The core of the *E. coli* Hfq sequence (from residues 1−72) was used as the query sequence. More than half of the bacterial genomes contain at least one gene that is readily identified as Hfq. Figure 1A shows the alignment of Hfq sequences from amino acid positions 1−67 for a limited number of microbial genomes separated into bacterial groups by phylogenetic analysis (*49*). In a few bacterial species, two or three distinct copies of Hfq were found. They are designated in Figure 1 by letters A, B, and C.

As has been previously noted (*5*, *49*), a number of positions have strikingly conserved residues or residues with similar chemical properties. Focusing on amino acids with high propensities at RNA-binding sites, we note that Phe or Tyr at positions 11, 25, 39, 42, and 55 and Lys or Arg at 3, 16, 17, 31, 47, and 56 are highly conserved. Figure 2 shows the exposure of these residues on the proximal and distal surfaces of *E. coli* Hfq as well as the edge-on or side view of the structure. Residues in only one of the six subunits are labeled for clarity in parts A and B of Figure 2, and residues in only one of the six subunits are shown in Figure 2C.

Phe 11 is located near the outer rim of the structure. It is not visible when viewing the proximal or distal surfaces but is exposed in the side view. Lys 47 is also at the outer rim of the structure, and its side-chain atoms in subunit B are ∼10 Å from the side-chain atoms of Phe11 in subunit A (Figure 2C). This Phe11-Lys47 pair is reminiscent of the Phe and Lys pairs on the outer rim of TRAP protein subunits that interact with RNA bases (*50*). Viewed from the proximal surface, Phe42, Tyr55, and Lys56 cluster near the cavity and Phe39, Lys16, and Lys17 form a radial-oriented patch near the outer part of this surface (Figure 2A). Lys3, which was not visualized in the crystal structure of *E. coli* Hfq (*28*), extends from the proximal surface. Tyr25 and Lys31 are adjacent to each other on the distal surface.

Evidence that several of these residues are involved in Hfq binding to RNA is indicated from previous work. The crystal structure of *S. aureus* Hfq and AU_5G shows residues at or adjacent to positions corresponding to 42, 55, and 56 interacting with this RNA. Mikulecky et al. (*36*) showed that although *E. coli* Hfq binds much more weakly to AU_5G (K_d ∼ 2 μM) than *S. aureus* Hfq [K_d ∼ 50 nM (*27*)], the Y55A and K56A mutant forms of *E. coli* Hfq each reduce affinity to this RNA by about 10-fold. We note however that the mutation F42A in *E. coli* Hfq did not significantly affect Hfq binding to AU_5G or DsrA (*36*). Substitution of Lys3 with several residues disrupted the Hfq function in Qβ RNA virus replication (*51*). On the distal surface of *E. coli* Hfq, mutations at Tyr25 and nearby Ile30 reduced Hfq affinity to A_{27}, implicating this surface in polyA binding (*36*).

Covariance of Residue Properties at Spatially Adjacent Positions. It was noted during the alignment of Hfq sequences that approximately 4% of the Hfq sequences did not have a Phe or Tyr at position 39. Figure 1B lists four of the bacterial sequences with this property. In each case, a smaller side chain replaced the aromatic residue expected at position 39 and, simultaneously, Tyr replaced the highly conserved Leu at position 12. These two positions, 12 and 39, are adjacent to each other in the Hfq structure at the same radial distance from the center of the hexamer (Figure 2). A covariance of side-chain properties appears to be needed at these highly conserved positions. This observation supports the notion that an aromatic residue at or near position 39 is important for the Hfq function.

One Hfq sequence listed in Figure 1B is from the archeal species *Methanococcus jannaschii*. This is the only archeal species found to date to contain an *hfq* gene while missing Sm genes (*28*). A phylogenetic analysis shows that this *M. jannaschii* gene belongs to a single phylic clade in the evolutionary tree of Hfq, separated from other bacterial species (data not shown). Thus, there is no evidence for lateral gene transfer from bacteria. Results described in the Appendix indicate that a single aromatic group at the surface patch of residues 39 and 12 in Hfq finds a close counterpart

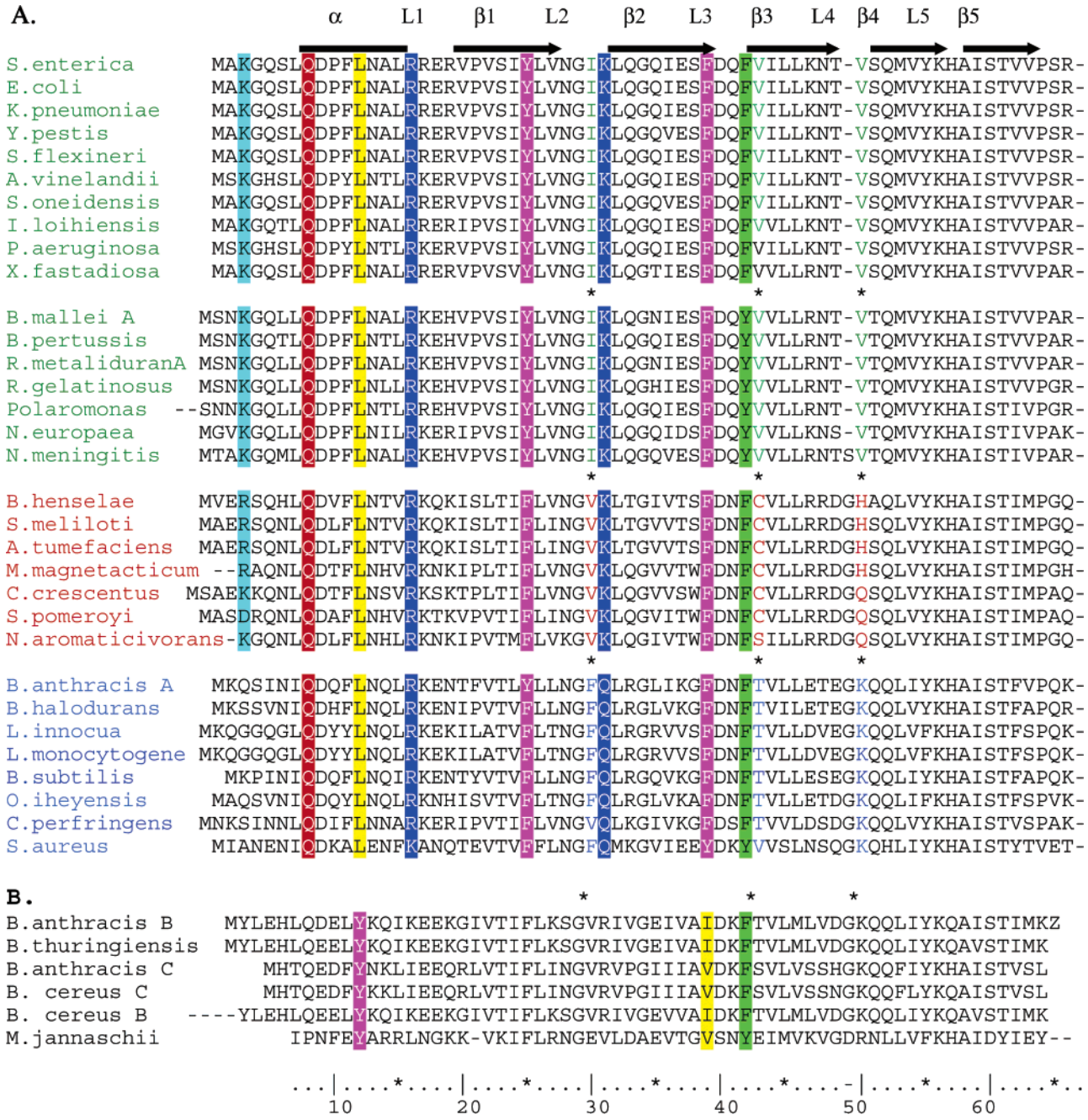

FIGURE 1: Multiple alignment of 38 representative bacterial Hfq coding regions. Residues from 1 to 67 are shown for most proteins. The location of secondary-structure elements of Hfq proteins are shown as lines above the sequences: α, α helix; L, loop; and *β*, *β* strand. (A) Hfq sequences are grouped according to bacterial clades. If a genome of a species has several distinct copies of the *hfq* gene, each is designated with letters A, B, etc. Residues shown as black or white letters with a colored background are conserved surface-exposed residues identified as potential RNA-binding sites. (B) Hfq sequences with a covariant switch of aromatic/aliphatic residues at positions 12 and 39.

in the structurally similar Sm proteins and covariance at the adjacent positions is also observed.

Effect of Hfq Mutations on RNA Binding. Guided by the above analysis, we examined the effect of several Hfq surface residues on RNA binding. Eight mutant *E. coli* Hfq proteins were constructed and expressed as described in the Materials and Methods. Six of the mutant proteins, designated Hfq-Q8A, Hfq-R16A, Hfq-F39A, Hfq-F42A, Hfq-F39A/F42A, and Hfq-L12F/F39A, are on the proximal face of the hexamer. Two mutant proteins, Hfq-Y25A and Hfq-K31A, are located in the distal face. The two RNA targets employed were $DsrA_{DII}$ (domain II of DsrA, nucleotides 23−60, and A_{18}. Domain II of DsrA competes with the intact DsrA and contains the U-rich region implicated in Hfq−DsrA binding (*33*). It is missing RNA regions that increase affinity for wt Hfq. This smaller RNA may be more sensitive to specific mutations at the sites where it binds. A_{18} represents the 3′-terminal segment of several mRNAs known to be regulated by Hfq (*17*, *18*).

Gel-Mobility Shift Assay of Hfq−RNA Binding. Figure 3A shows gel-shift experiments of wt Hfq and several mutant Hfq proteins with A_{18}. The wt Hfq, Hfq-F39A, and Hfq-F42A show similar behavior. Increasing amounts of protein initially shifts A_{18} to a low mobility complex (C_1) and then

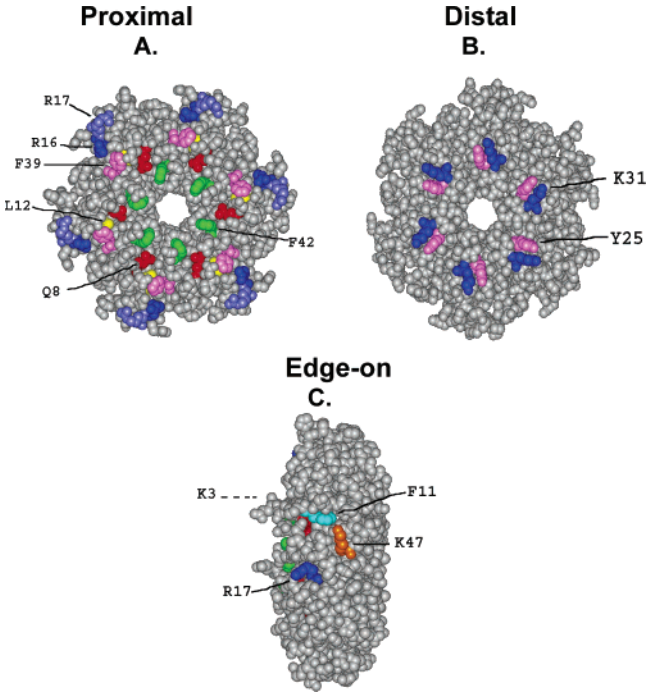

FIGURE 2: Proximal, distal, and edge-on surface views of the *E. coli* Hfq structure (*28*). Color-coded and labeled residues are potential RNA-binding sites from sequence−structure analysis. K3 is not observed in the crystal structure, but its approximate position in one subunit is indicated in the edge-on view.

to the complex labeled C_2. Gel-shift experiments of the other mutant Hfq proteins were similar to wt Hfq with two exceptions. Hfq-K31A (bottom panel of Figure 3A) showed only a small decrease in intensity of the free A_{18} band with increasing Hfq and no well-defined complex bands. A similar result was obtained with Hfq-Y25A (not shown). Both Tyr25 and Lys31 are situated on the distal face of the hexamer structure (Figure 2B). Tyr25 and the adjacent Ile30 have been previously identified to be important for binding A_{27} (*36*). The results confirm the importance of Tyr25 in binding oligoA and demonstrate that Lys31 is also needed to interact with polyA sequences. Figure 3B summarizes the change in the fraction of A_{18} in the free A_{18} band as a function of the Hfq concentration for the wt and eight mutant Hfq proteins examined.

The experiments conducted in Figure 3 employed the fluorescent FAM−A_{18} and used a total RNA concentration of 250 nM. Because this is well above the estimated K_d of the wt Hfq−A_{18} complex measured from a gel-shift assay (*36*) and fluorescence anisotropy (results below), it should reflect stoichiometric binding. The titration of A_{18} with wt Hfq saturated the C_1 complex at a ratio of Hfq_6/A_{18} of approximately 2:1 (Figure 3A). Although an accurate evaluation of this ratio was not established from the gel data, two independent approaches described below are consistent with a 2:1 ratio for Hfq hexamer/A_{18} in the strong binding complex.

Figure 4A shows gel-mobility shift experiments of wt and several mutant Hfq proteins to $DsrA_{DII}$. Radioactively labeled $DsrA_{DII}$ was used in these experiments at a 2 nM concentration. The wt Hfq initially shifts unbound $DsrA_{DII}$ to a low mobility complex (C_1) and then shifts the RNA to lower mobility complexes with increasing amounts of Hfq. The observation of several gel-mobility complexes was previously observed between wt Hfq and the complete DsrA (*10*). The second panel of Figure 4A indicates that Hfq-F39A has a lower affinity for $DsrA_{DII}$. Intensity of the unbound $DsrA_{DII}$ band was observed at protein concentrations where no free $DsrA_{DII}$ bands were detected for wt Hfq. Additionally, the bands at the position of the complexes were more diffuse, indicating that complexes of Hfq-F39A and $DsrA_{DII}$ dissociate more readily in the gel. The third panel shows results with Hfq-R16A, which were similar to Hfq-F39A. The bottom panel of Figure 4A shows the gel-shift experiment of Hfq-L12F/F39A binding to $DsrA_{DII}$. The relative amount of the free $DsrA_{DII}$ band decreased more rapidly with additions of Hfq-L12F/F39A than with equal amounts of Hfq-F39A. This double mutation partially restored the binding defect caused by F39A.

Figure 4B displays the fraction of bound $DsrA_{DII}$ as a function of the Hfq concentration assessed from the change in intensity of free $DsrA_{DII}$ bands. The concentration at which $F_B = 0.5$ ($F_{0.5}$) for wt Hfq was 470 nM. Most mutant Hfq proteins had a $F_{0.5}$ value between 550 nM and 1 μM; however, for Hfq-R16A, the value was 2.6 μM, while the midpoint for Hfq-F39A and Hfq-F39A/F42A was 3.8 μM. The 6−8-fold increase in the apparent K_d for Hfq-R16A and Hfq-F39A compared to wt Hfq suggests that these sites interact with $DsrA_{DII}$. The observation that the double-mutant Hfq-L12F/F39A increased the affinity for $DsrA_{DII}$ relative to Hfq-F39A supports the notion suggested from the bioinformatics analysis that an aromatic group at this location is functionally important.

Gel-shift experiments using fluorescent-labeled $DsrA_{DII}$ at 100 nM produced better defined C_1 bands, but the results were otherwise qualitatively similar to Figure 4 (data not shown). Figure 5 shows results from a gel-shift assay carried out with 1 μM $DsrA_{DII}$ and wt Hfq. The C_1 complex reached a maximum intensity at a molar ratio of 2:1 $Hfq_6/DsrA_{DII}$ and was constant until 5 μM Hfq_6 before its intensity decreased in conjunction with the increased intensity of slower mobility complexes. A similar saturating ratio of 2:1 was observed by Lease and Woodson (*10*) for the strong binding complex formed by wt Hfq_6 and the complete DsrA at high concentrations.

Fluorescence Anisotropy Measurements of Hfq−RNA Binding. Fluorescence anisotropy was employed to analyze the binding affinity of the wt and mutant Hfq proteins with the RNAs. Figure 6A shows the anisotropy change for 2 nM A_{18} with an increasing concentration of wt Hfq. The dashed line shows the nonlinear least-squares fit of a 1:1 model of Hfq binding to A_{18} to the data. The inability of this model to accurately fit the data indicates a more complex binding interaction.

The numerically based equilibrium binding algorithm, BIOEQS, was employed to analyze the anisotropy data using a model in which the RNA can bind one Hfq_6 and subsequently a second Hfq_6. The solid line in Figure 6A shows the least-squares fit of this model to the data. The quality of the fit is considerably improved with a χ^2 of 1.3. The free energies evaluated correspond to dissociation constants for binding the first and second Hfq_6 of 10.1 and 4.7 nM, respectively. Confidence limit testing of the recovered free energies gave deep quadratic well-like plots of χ^2 versus free energy, with minimums at both evaluated free energies.

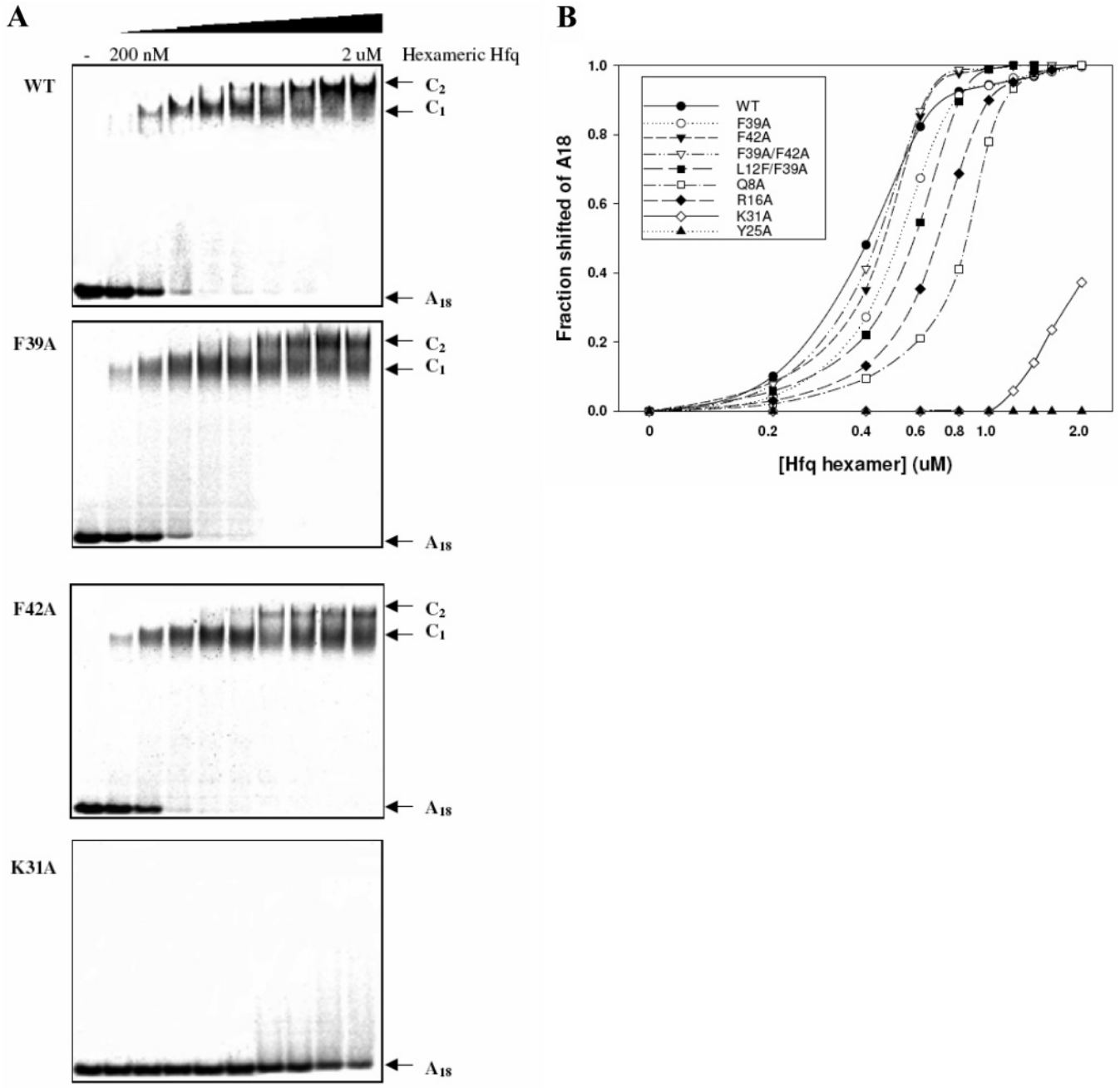

FIGURE 3: (A) Binding of wt and three mutant Hfq to A_{18} by the gel-shift assay. Concentration of Hfq from left to right in nanomolars hexamer: 0, 200, 400, 600, 800, 1000, 1200, 1400, 1600, and 2000. (B) Fraction of A_{18} bound as a function of the Hfq concentration for wt and eight mutant Hfq.

Figure 6B shows the results of anisotropy experiments for two of the mutant Hfq proteins, Hfq-F39A and Hfq-K31A. As expected from the gel-shift experiments, Hfq-F39A binds A_{18} similarly to wt Hfq, while Hfq-K31A shows very weak binding. Table 1 summarizes equilibrium dissociation constants evaluated from the analyses of the anisotropy measurements. In all cases, except for Hfq-K31A and Hfq-Y25A, the two-step binding model was needed to produce a good fit to the anisotropy curves and binding was in the nanomolar range. For wt Hfq and four of the mutant Hfq, the analysis yielded binding constants for the second bound Hfq_6 that were 2−10-fold lower than K_1. Analyses of data for Hfq-L12F/F39A, Hfq-Q8A, and Hfq-F39A/F42A produced similar K_2 and K_1 values. For the two proteins, K31A and Y25A, the anisotropy change was barely higher than the background. A rough estimate of K_d based on a 1:1 binding model gave $K_d \geq 1.3\ \mu M$ for these two single-site mutations, indicating a $\sim 100\times$ decrease in affinity.

Figure 7 shows the results of fluorescence anisotropy experiments of wt Hfq binding to $DsrA_{DII}$. The fluorescence anisotropy reached a plateau between ~600 nM and 1.1 μM Hfq_6 and then increased slightly to another plateau with additions of Hfq_6 above 1.1 μM. The overall anisotropy change was considerably less than observed for Hfq binding to A_{18}. The second plateau, although quite small, was reproduced in several experiments. Given the signal-to-noise of the data, we could not discriminate between 1:1 and 2:1 binding models for Hfq_6 binding $DsrA_{DII}$ based on the quality of fit. If one assumes a 1:1 binding model, the least-squares fit to the first plateau region using eq 1 yielded $K_d \approx 470$

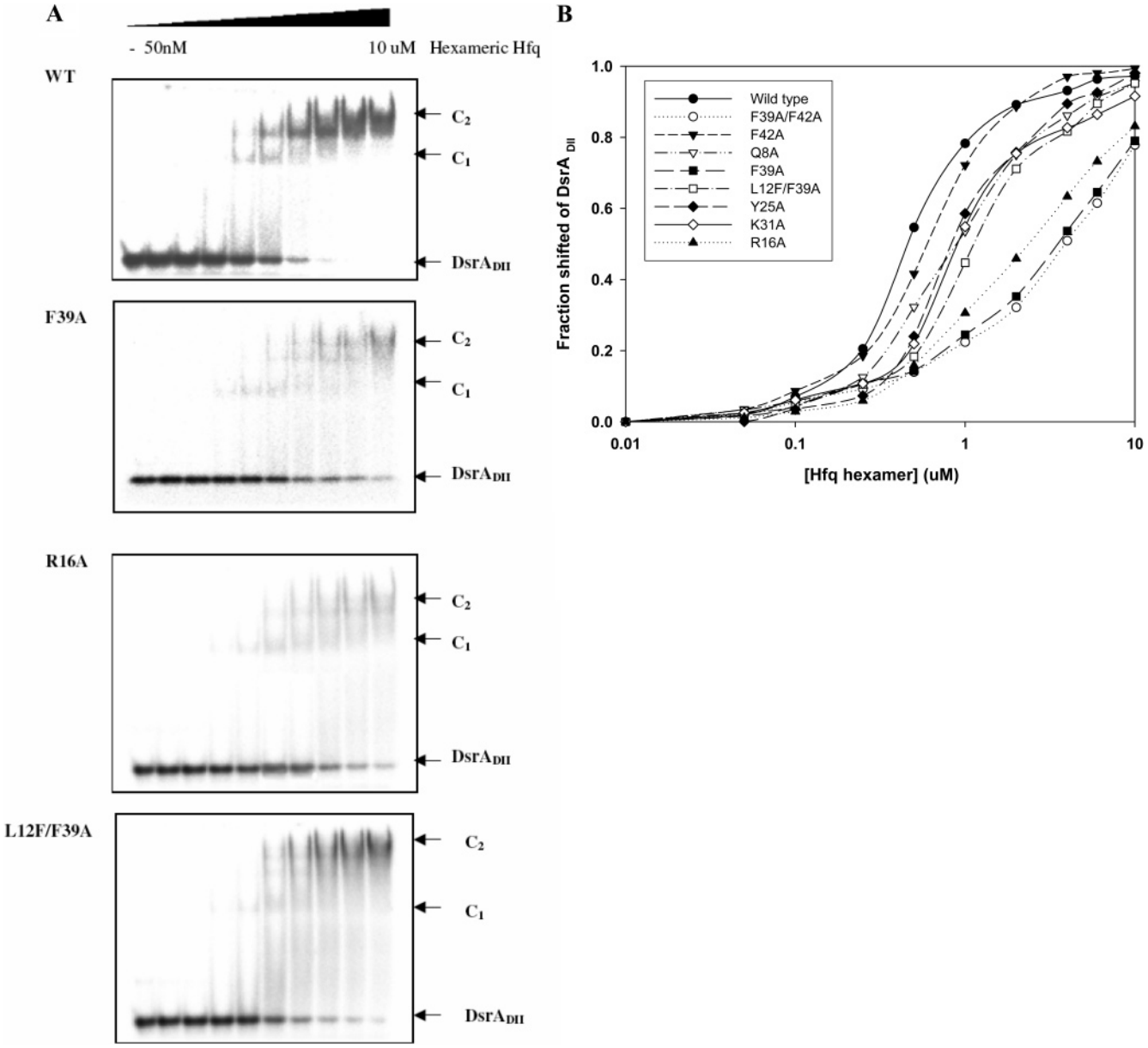

FIGURE 4: (A) Binding of wt and three mutant Hfq to $DsrA_{DII}$ by the gel-shift assay. Hfq concentrations from left to right in micromolars hexamer: 0, 0.05, 0.1, 0.2, 0.5, 1.0, 2.0, 4.0, 6.0, and 10.0. (B) Fraction of $DsrA_{DII}$ bound as a function of the Hfq concentration for wt and eight mutant Hfq.

nM (— in Figure 7). If one uses all of the data (- - - in Figure 7), the best fit K_d is approximately 1 μM. Fitting either all of the data or just the first plateau region to the two-step binding model using BIOEQS yielded a dissociation constant of $K_1 \approx 440$ nM for binding the first Hfq_6 and $K_2 \approx 1.9$ μM for binding a second Hfq_6. However, confidence limit testing indicated that K_2 could not be accurately resolved.

Surprisingly, the titration of $DsrA_{DII}$ with the mutant Hfq proteins produced anisotropy curves similar to that for wt Hfq. The anisotropy change as a function of the Hfq-F39A concentration is shown in Figure 7 as △. Data for the other mutant Hfq were similar. The data were not consistent with an 8-fold reduction in affinity indicated by the gel-shift experiments. The discrepancy between the anisotropy measurements and gel-shift experiments with regard to binding wt and mutant Hfq to $DsrA_{DII}$ is discussed below.

Fluorescence Quenching Measurements. A different approach that can be employed to investigate RNA binding to Hfq is to monitor the fluorescence quenching of Hfq as a function of the RNA concentration. DeHaseth and Uhlenbeck showed that polyA binding quenches Hfq fluorescence (*52*, *53*). *E. coli* Hfq has a moderate intrinsic fluorescence because of its three tyrosines Tyr25, Tyr55, and Tyr83. Tyr25 is located on the distal surface of the hexamer; Tyr55 is located near the central cavity close to the proximal surface; and Tyr83 is in the unstructured C-terminal end.

Figure 8A shows the fluorescence emission spectrum of wt Hfq at 5 μM (hexamer). The titration of A_{18} decreased the fluorescence of Hfq until a saturation level (~73% at 305 nm) was reached. Quenching may result from the stacking of an adenine base with exposed Tyr and/or the transfer of the Tyr hydroxyl group proton to an acceptor on A_{18} and formation of the nonfluorescent tyrosinate anion. Figure 8B plots the percent quenching as a function of the ratio of A_{18}/wt Hfq. The value of $[A_{18}]/[Hfq_6]$ at the break point in the curve is 0.5, consistent with a 2:1 stoichoimetry

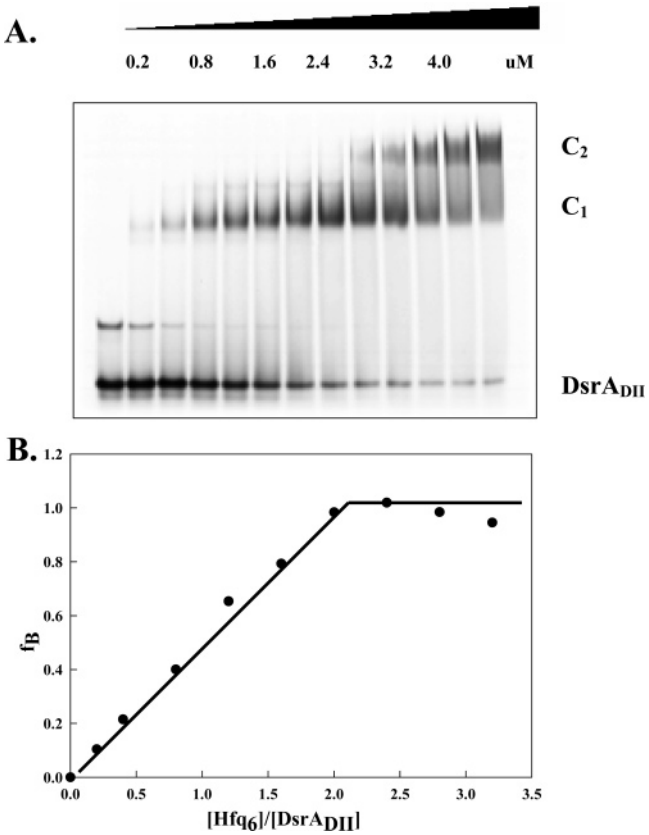

FIGURE 5: (A) Gel-shift assay of wt Hfq binding to 1 μM $DsrA_{DII}$. Hfq_6 (in micromolars) from left to right are 0, 0.2, 0.4, 0.8, 1.2, 1.6, 2.0, 2.4, 2.8, 3.2, 3.6, and 4.0. Weak band above the free $DsrA_{DII}$ band is the dimer of this RNA. (B) Fraction of the C_1 band intensity divided by its maximum intensity as a function of the molar ratio of $Hfq_6/DsrA_{DII}$.

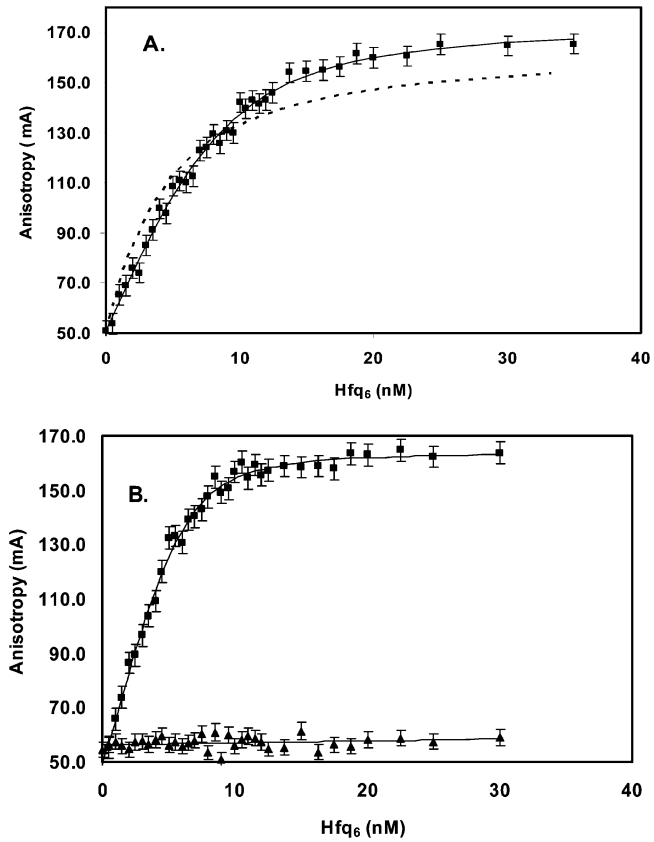

FIGURE 6: (A) Fluorescence anisotropy data of A_{18} as a function of the wt Hfq concentration. The solid line is a least-squares fit of a two-step binding model using BIOEQS. The dashed line is a least-squares fit of a 1:1 binding model described by eq 1. (B) Anisotropy data of A_{18} as a function of F39A-Hfq (■) and K31A-Hfq (▲). Solid lines are a least-squares fit of a two-step binding model to F39A-Hfq data and a 1:1 binding model to K31A-Hfq data.

of Hfq_6/A_{18}. Similar results were obtained for the other mutant Hfq proteins with the exceptions of Hfq-Y25A, Hfq-K31A, and Hfq-Q8A.

The fluorescence of Hfq-Y25A and Hfq-K31A was quenched by only 12 and 15%, respectively, at saturating amounts of A_{18} and showed no clear break points in the titrations. These results indicate weaker binding of A_{18} to these mutant proteins, consistent with the data from the mobility shift assay and anisotropy measurements. The fluorescence intensity of free Hfq-Y25A was about one-third that of wt Hfq at the same concentration. Reduced intensity is expected for this Hfq missing the surface-exposed Tyr25. Free Hfq-Q8A had about 140% higher fluorescence intensity compared to wt Hfq at the same concentration.

The latter observation can be understood in terms of two crystal structures of Hfq and the effect that a Q8A mutation may have on Tyr55. The crystal structure of *E. coli* Hfq reveals that Gln8 in one subunit is adjacent to Tyr55 in a neighboring subunit (*28*). Additionally, a recent high-resolution structure of Hfq from *Pseudomonas aeruginosa* shows a conserved hydrogen bond between the OH group of Tyr55 and OE1 of Gln8 that stabilizes the α helix in Hfq (*29*). In wt *E. coli* Hfq, it is likely that the fluorescence of Tyr55 is quenched because of the hydrogen bond between its hydroxyl group and Gln8. The Gln to Ala substitution appears to release the hydroxyl group of Tyr55, increasing the fluorescence originating from this residue and the overall fluorescence of the protein. The percent quenching of Hfq-Q8A fluorescence at saturating amounts of A_{18} was ~60% compared to 73−75% for wt Hfq. This is qualitatively consistent with the notion that the quenching effect of A_{18} is due to its interaction with Tyr25. Because this residue contributes less to the overall fluorescence of Hfq-Q8A than to wt Hfq, the saturation quenching percentage is expected to be reduced. Additions of $DsrA_{DII}$, rGU_5A, and rU_{20} had a negligible effect on Hfq fluorescence.

Table 1: Equilibrium Dissociation Constants Evaluated for wt and Mutant Hfq Binding to A_{18} Using Fluorescence Anisotropy[a]

Hfq	K_1[b] (nM)	K_2[b] (nM)	Hfq	K[c] (μM)
wt	10.1 ± 1.2	4.73 ± 0.8	K31A	>1.6
F39A	8.7	0.75	Y25A	>1.3
L12F/F39A	4.4	3.61		
F42A	3.1	0.62		
F39A/F42A	8.3	8.27		
Q8A	12.0	14.7		
R16A	14.9	3.0		

[a] The solvent was 0.5 M NaCl and 20 mM Tris (pH 8.3). [b] Binding constants were evaluated by a nonlinear least-squares fit of data to a two-step binding model with BIOEQS. [c] Binding constants were estimated by a nonlinear least-squares fit of data to a 1:1 binding model using eq 1.

DISCUSSION

The results on the interaction of A_{18} with wt and mutant Hfq proteins indicate that this RNA sequence interacts

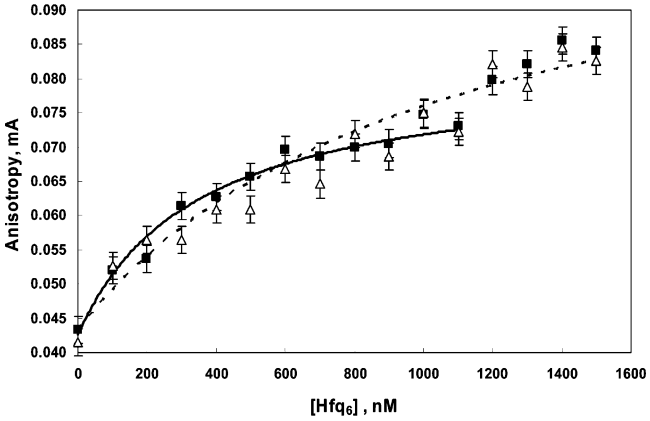

FIGURE 7: Fluorescence anisotropy of $DsrA_{DII}$ as a function of wt Hfq (■) and F39A-Hfq (△). The solid line is best fit of a 1:1 model to wt Hfq data up to 1100 nM Hfq_6, and the dashed line is a best fit to 1500 nM.

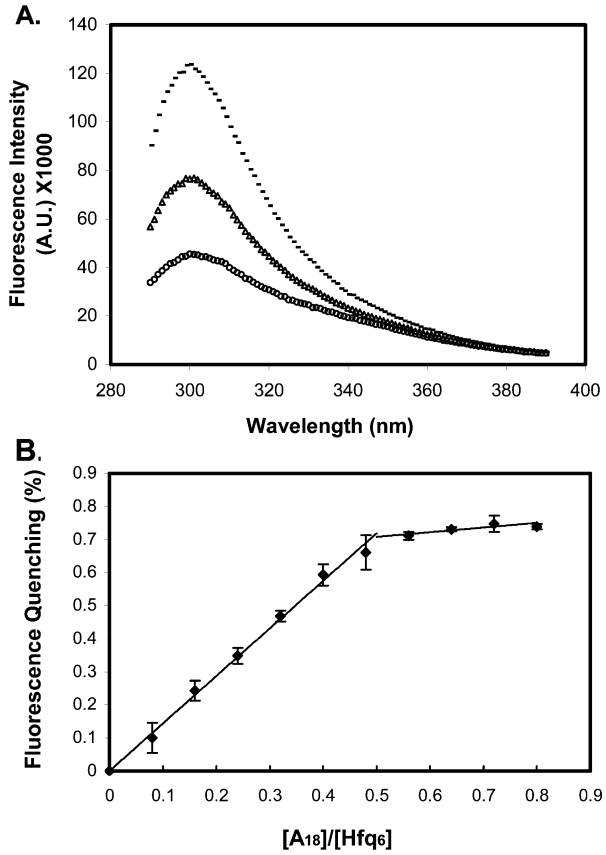

FIGURE 8: (A) Fluorescence emission spectra of 50 μL of wt Hfq at 5 μM with 0, 2, and 4 μL of A_{18} at 50 μM added. (B) Percentage of quenching measured at 305 nm as a function of $[A_{18}]/[Hfq_6]$.

specifically with residues on the distal surface of the *E. coli* Hfq. The gel-shift assay, fluorescence anisotropy, and fluorescence quenching experiments all showed that mutations on the proximal surface of Hfq (F39A, F42A, Q8A, R16A, F39A/F42A, and L12F/F39A) had a relatively minimal effect on binding A_{18}, while single-site mutations Y25A and K31A on the distal surface reduced affinity by 100-fold or more. The results are consistent with the conclusion of Mikulecky et al. (*36*) that polyA sequences bind specifically to the distal surface.

In addition to confirming the importance of Tyr25, our results demonstrate that a mutation to Lys31, which is adjacent to Tyr25, also produces a major reduction in Hfq binding to a polyA sequence. Several studies indicate that Hfq is required for the regulation of mRNA polyadenylation (*17*, *24*, *37*). This activity requires Hfq interacting with polyA polymerase and its ability to preferentially bind to the 3′ end of Rho-independent transcripts prior to polyadenylation (*17*). The role of the high affinity of Hfq proteins to polyA sequences in regulating polyadenylation of mRNAs is uncertain. Studies suggest that Hfq binding to polyA tails protects the mRNA from ribonuclease degradation (*37*). If this is the case, the Y25A and K31A mutants would be expected to influence polyadenylation-dependent mRNA decay in vivo.

The gel-shift experiments of A_{18} in the presence of an increasing Hfq concentration showed the existence of two protein−RNA bands designated C_1 and C_2 (Figure 3A). Using concentrations greater than the value of the equilibrium dissociation constant, the C_1 band saturated at a stoichiometry of approximately 2:1 for wt Hfq_6/A_{18} and then decreased as the C_2 band increased. The fluorescence anisotropy data on Hfq binding to A_{18} required a model in which two Hfq_6 molecules bind one A_{18} to obtain a good fit of theory to the experiment. In addition, fluorescence quenching of wt Hfq by A_{18} at micromolar concentrations also indicated a 2:1 Hfq_6/A_{18} stoichiometry. This stoichiometry differs from the value obtained previously (*36*) from isothermal titration calorimetric measurements that indicated a 1:2 Hfq_6/A_{18} complex. Although we do not know the cause for this difference, we note that a different solvent was employed in this study.

A very faint band that was faster in mobility than the C_1 band was detected at Hfq_6 concentrations lower than those employed in Figure 3A; however, it never exceeded 5% of the total RNA and disappeared above 200 nM Hfq_6 (data not shown). This faint band may correspond to a 1:1 Hfq_6/A_{18} complex. Its absence as the C_1 band became evident implies that Hfq binds A_{18} with positive cooperativity. The anisotropy data (Table 1) also suggests a positive cooperativity for binding Hfq_6 to A_{18}. K_1 is always higher than or similar to K_2. If A_{18} has two identical binding sites for Hfq_6, a microscopic cooperativity constant k_{12} as defined by Ackers et al. (*54*) is equal to $4K_1/K_2$ and values of $k_{12} > 1$ represent positive cooperativity. A calculation indicates that the positive cooperativity inferred from the anisotropy measurements is less than that indicated by the gel-shift data.

If we employ the K_1 and K_2 values from Table 1 for wt Hfq and calculate the relative amount of A_{18} expected in a 1:1 Hfq_6/A_{18} complex for the concentrations employed in the gel-shift assay, one predicts a much higher maximum (∼20%) than is observed. Although gel conditions may enhance RNA-induced dimerization of Hfq_6 (as suggested above), the solution studies on A_{18} argue against this being solely due to the gel conditions. The relative mobility of the C_2 band in Figure 3A and its appearance at the expense of the C_1 band indicates complexes of Hfq_6 and A_{18} with a higher Hfq_6/RNA stoichiometry than C_1.

The 2:1 stoichiometry determined for the strong binding complex of Hfq_6 and A_{18} was also determined for the C_1 complex of Hfq_6 and $DsrA_{DII}$ from Figure 5. This same value was previously observed for the strong binding complex of Hfq_6 with the 87 nt DsrA and 140 nt RpoS RNAs (*10*, *36*). This common stoichiometry for RNAs of different lengths and sequence suggests that it is a common feature of Hfq_6

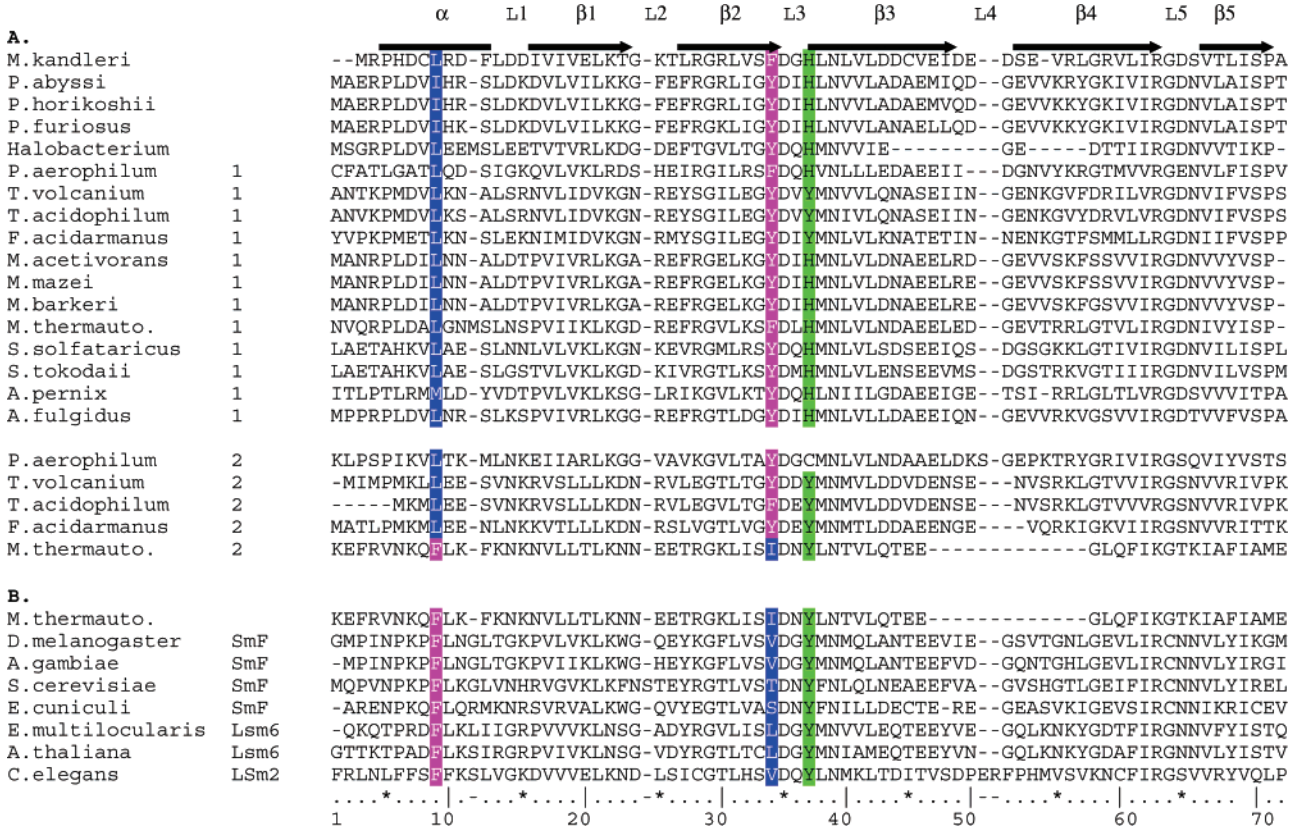

FIGURE 9: (A) Multiple alignment of archeal Sm proteins built up using a genome BLAST search. Each archeal genome contains either one or two copies of a Sm protein, termed as Smα and Smβ, and numbered 1 and 2, respectively. (B) Covariance between Leu9 and Phe/Tyr34 observed in Sm proteins of eukaryotes as well as archaea. A BLAST search was made against the nonredundant archaea and eukaryotic database using the Smβ protein of *M. thermoautotrophicum* as a query sequence. The covariance found in *M. thermoautotrophicum* Smβ was observed in *D. melanogaster* SmF, *A. thaliana* Lsm6, and *C. elegans* Lsm2.

binding to RNA. Dimerization of Hfq_6 appears to be induced for RNA binding to a patch on the distal surface, A_{18}, as well as for RNA binding that is influenced by residues on the proximal surface and not the distal surface [DsrA (*36*) and $DsrA_{DII}$ (this study)]. The above observations may indicate that RNA-induced dimerization of Hfq_6 may utilize more than one type of dimerization interface. Recent work has shown that Hfq_6 can form fibers in which each hexamer can partially stack on another hexamer using interactions on both surfaces (*55*). A crystal structure of the *S. aureus* Hfq_6 exhibits distal−distal surface interactions between two Hfq_6 with seven nucleotide oligomers bound to their proximal surfaces. Sm-like proteins form crystals in which small RNAs are sandwiched between two Sm-like proteins (*30*).

Gel-shift studies of Hfq binding to $DsrA_{DII}$ indicated that the F39A and R16A mutations affect the affinity of Hfq to $DsrA_{DII}$ and imply that these residues interact with this RNA. These two mutations reduced the apparent K_d by 6−8-fold (assuming a 1:1 model), while other single-site mutations altered the apparent K_d by 2-fold or less. The ability of the L12F/F39A-Hfq to partially restore the affinity of Hfq to $DsrA_{DII}$ supports the notion that Phe39 is involved in $DsrA_{DII}$ binding. It is consistent with the bioinformatics analysis, indicating that an aromatic residue at this surface location is important for the Hfq function.

Contrary to the gel-shift studies, the fluorescence anisotropy experiments with $DsrA_{DII}$ did not show a significant difference between the binding of the wt and mutant Hfq proteins. This discrepancy is not currently understood; however, several factors may contribute to it. The anisotropy change observed for the fluorescently labeled $DsrA_{DII}$ upon Hfq binding is relatively small. Calculations using the model described by eq 1 indicate that a 2−3-fold difference in the apparent K_d for Hfq variants would be difficult to detect. Another factor worth noting is that the gel-shift experiment is a pseudo-equilibrium method. Although this method is more sensitive than the anisotropy measurements for quantifying Hfq−RNA complexes, it may be discriminating between binding intermediates or complexes existing in the gel assay that are not prevalent under equilibrium conditions in solution.

During the time RNA−protein complexes enter the gel and separate from free RNA and protein, the low ionic strength running buffer is exchanging with the 0.5 M Na^+-loading buffer lowering the ionic strength. The electrophoresis process in the gel also results in dissociation of Hfq−$DsrA_{DII}$ complexes at the 2 nM RNA concentration used to evaluate binding constants. This was evident from the intensity between the free RNA band and the C_1 band and the broad bands observed for the complexes with the mutant Hfq proteins (Figure 5a). An investigation of Hfq−$DsrA_{DII}$ interactions in low ionic strength solutions and an evaluation of dissociation and association rates of wt and mutant Hfq for $DsrA_{DII}$ may illuminate the discrepancy.

APPENDIX

The covariance described in the Results of an aromatic residue and small aliphatic residue at positions 39 and 12 in

bacterial genomes was also observed in *M. jannaschii*. This is the first and thus far only archeal species found to contain a *hfq* gene and no Sm gene. All other archeal species examined to date have Sm-type proteins that form stable heptamer oligomers. It is of interest to examine if the covariance exhibited by the archeal *M. jannaschii* Hfq occurs in Sm proteins and whether other surface residues conserved in Hfq are preserved in Sm proteins.

The exact evolutionary relationship between Hfq and Sm proteins in bacteria, archaea, and eukaryotes is still not clear, but the most generally accepted view is that the Sm proteins in archaea and eukaryotes evolved from a common ancestor that diverged from bacterial Hfq (*49*). It has been found that Hfq and Sm proteins have many similar functions, such as serving as a host factor for RNA virus replication, participating in RNA degradation, and facilitating RNA duplex formation (*5*, *56*). The function(s) of this protein family appears to have been conserved during evolution, even though its oligomeric form changed from hexamer to heptamer. One may thus anticipate that some residues that are conserved and functionally important in Hfq may have counterparts in Sm proteins.

To examine this hypothesis, a multiple alignment of archeal Sm proteins was built up using a genome BLAST search (Figure 9A). Each archeal genome contains either one or two copies of a Sm protein, termed as Smα and Smβ, and numbered 1 and 2, respectively. In all of the Smα proteins, the three residues that correspond to Leu12, Phe39, and Phe42 on the proximal surface of Hfq are highly conserved but with His, the dominating residue at position 37 (corresponding to Phe at position 42 in Hfq). The covariant switch of residue properties at positions equivalent to 12 and 39 was observed in a Smβ protein of *Methanobacteriaum thermoautotrophicum*. In contrast, the highly conserved positively charged Arg residues at positions 16 and 17 in Hfq were not observed in these Sm proteins. A previous study on the *M. thermoautotrophicum* Smα protein also noticed that the spatially adjacent Leu9 and Phe34 (corresponding to Leu12 and Phe39 in Hfq) are a highly conserved hydrophobic pair in archaea and suggested that this region may be involved in a protein−protein interaction (*57*). Recently, however, Thore et al. showed that Tyr34 in a crystal structure of *Pyrococcus abyssi* Smα protein can form contacts with RNA (*58*). Mutating this residue to Val reduced the binding of this Smα protein to the Sm consensus RNA oligomer.

Figure 9B shows that the covariance between Leu9 and Phe/Tyr34 is observed in Sm proteins of eukaryotes as well as archaea. A BLAST search was made against the nonredundant archaea and eukaryotic database using the Smβ protein of *M. thermoautotrophicum* as a query sequence. The covariance found in *M. thermoautotrophicum* Smβ was observed in *Drosophila melanogaster* SmF, *Arabidopsis thaliana* Lsm6, and *Caenorhabditis elegans* Lsm2. If one examines the residues on the distal surface of the Sm proteins listed in Figure 9, the counterpart of the three polyA binding residues Tyr25, Ile30, and Lys31 in Hfq are not found. This is consistent with the observation that almost all RNAs known to interact with Sm proteins are U-rich RNA. No A-rich RNA has been reported.

ACKNOWLEDGMENT

We are grateful to Catherine Royer for providing BIOEQS and for her help in providing information on its application. We also thank Scott Bradley for assisting in the purification of some Hfq proteins and Cameron Thompson and Andrew Lyon for allowing the use of their fluorescence spectrometers during the course of this work.

REFERENCES

1. Franze de Fernandez, M. T., Eoyang, L., and August, J. T. (1968) Factor fraction required for the synthesis of bacteriophage Qβ-RNA, *Nature 219*, 588−590.
2. Franze de Fernandez, M. T., Hayward, W. S., and August, J. T. (1972) Bacterial proteins required for replication of phage Q ribonucleic acid. Pruification and properties of host factor I, a ribonucleic acid-binding protein, *J. Biol. Chem. 247*, 824−831.
3. Tsui, H. C., Leung, H. C., and Winkler, M. E. (1994) Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in *Escherichia coli* K-12, *Mol. Microbiol. 13*, 35−49.
4. Storz, G., Opdyke, J. A., and Zhang, A. (2004) Controlling mRNA stability and translation with small, noncoding RNAs, *Curr. Opin. Microbiol. 7*, 140−144.
5. Valentin-Hansen, P., Eriksen, M., and Udesen, C. (2004) The bacterial Sm-like protein Hfq: A key player in RNA transactions, *Mol. Microbiol. 51*, 1525−1533.
6. Zhang, A., Wassarman, K. M., Ortega, J., Steven, A. C., and Storz, G. (2002) The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs, *Mol. Cell 9*, 11−22.
7. Sledjeski, D. D., Whitman, C., and Zhang, A. (2001) Hfq is necessary for regulation by the untranslated RNA DsrA, *J. Bacteriol. 183*, 1997−2005.
8. Moller, T., Franch, T., Hojrup, P., Keene, D. R., Bachinger, H. P., Brennan, R. G., and Valentin-Hansen, P. (2002) Hfq: A bacterial Sm-like protein that mediates RNA−RNA interaction, *Mol. Cell 9*, 23−30.
9. Majdalani, N., Chen, S., Murrow, J., St John, K., and Gottesman, S. (2001) Regulation of RpoS by a novel small RNA: The characterization of RprA, *Mol. Microbiol. 39*, 1382−1394.
10. Lease, R. A., and Woodson, S. A. (2004) Cycling of the Sm-like protein Hfq on the DsrA small regulatory RNA, *J. Mol. Biol. 344*, 1211−1223.
11. Geissmann, T. A., and Touati, D. (2004) Hfq, a new chaperoning role: Binding to messenger RNA determines access for small RNA regulator, *EMBO J. 23*, 396−405.
12. Brown, L., and Elliott, T. (1996) Efficient translation of the RpoS σ factor in *Salmonella typhimurium* requires host factor I, an RNA-binding protein encoded by the *hfq* gene, *J. Bacteriol. 178*, 3763−3770.
13. Cunning, C., Brown, L., and Elliott, T. (1998) Promoter substitution and deletion analysis of upstream region required for rpoS translational regulation, *J. Bacteriol. 180*, 4564−4570.
14. Zhang, A., Altuvia, S., Tiwari, A., Argaman, L., Hengge-Aronis, R., and Storz, G. (1998) The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein, *EMBO J. 17*, 6061−6068.
15. Muffler, A., Fischer, D., and Hengge-Aronis, R. (1996) The RNA-binding protein HF-I, known as a host factor for phage Qβ RNA replication, is essential for rpoS translation in *Escherichia coli*, *Genes Dev. 10*, 1143−1151.
16. Majdalani, N., Vanderpool, C. K., and Gottesman, S. (2005) Bacterial small RNA regulators, *Crit. Rev. Biochem. Mol. Biol. 40*, 93−113.
17. Mohanty, B. K., Maples, V. F., and Kushner, S. R. (2004) The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in *Escherichia coli*, *Mol. Microbiol. 54*, 905−920.
18. Hajnsdorf, E., and Regnier, P. (2000) Host factor Hfq of *Escherichia coli* stimulates elongation of poly(A) tails by poly-(A) polymerase I, *Proc. Natl. Acad. Sci. U.S.A. 97*, 1501−1505.
19. Le Derout, J., Folichon, M., Briani, F., Deho, G., Regnier, P., and Hajnsdorf, E. (2003) Hfq affects the length and the frequency of short oligo(A) tails at the 3′ end of *Escherichia coli* rpsO mRNAs, *Nucleic Acids Res. 31*, 4017−4023.
20. Fromont-Racine, M., Mayes, A. E., Brunet-Simon, A., Rain, J. C., Colley, A., Dix, I., Decourty, L., Joly, N., Ricard, F., Beggs,

J. D., and Legrain, P. (2000) Genome-wide protein interaction screens reveal functional networks involving Sm-like proteins, *Yeast 17*, 95−110.
21. Achsel, T., Stark, H., and Luhrmann, R. (2001) The Sm domain is an ancient RNA-binding motif with oligo(U) specificity, *Proc. Natl. Acad. Sci. U.S.A. 98*, 3685−3689.
22. He, W., and Parker, R. (2000) Functions of Lsm proteins in mRNA degradation and splicing, *Curr. Opin. Cell Biol. 12*, 346−350.
23. Seto, A. G., Zaug, A. J., Sobel, S. G., Wolin, S. L., and Cech, T. R. (1999) *Saccharomyces cerevisiae* telomerase is an Sm small nuclear ribonucleoprotein particle, *Nature 401*, 177−180.
24. Bouveret, E., Rigaut, G., Shevchenko, A., Wilm, M., and Seraphin, B. (2000) A Sm-like protein complex that participates in mRNA degradation, *EMBO J. 19*, 1661−1671.
25. Kambach, C., Walke, S., and Nagai, K. (1999) Structure and assembly of the spliceosomal small nuclear ribonucleoprotein particles, *Curr. Opin. Struct. Biol. 9*, 222−230.
26. Mura, C., Cascio, D., Sawaya, M. R., and Eisenberg, D. S. (2001) The crystal structure of a heptameric archaeal Sm protein: Implications for the eukaryotic snRNP core, *Proc. Natl. Acad. Sci. U.S.A. 98*, 5532−5537.
27. Schumacher, M. A., Pearson, R. F., Moller, T., Valentin-Hansen, P., and Brennan, R. G. (2002) Structures of the pleiotropic translational regulator Hfq and an Hfq−RNA complex: A bacterial Sm-like protein, *EMBO J. 21*, 3546−3556.
28. Sauter, C., Basquin, J., and Suck, D. (2003) Sm-like proteins in Eubacteria: The crystal structure of the Hfq protein from *Escherichia coli*, *Nucleic Acids Res. 31*, 4091−4098.
29. Nikulin, A., Stolboushkina, E., Perederina, A., Vassilieva, I., Blaesi, U., Moll, I., Kachalova, G., Yokoyama, S., Vassylyev, D., Garber, M., and Nikonov, S. (2005) Structure of *Pseudomonas aeruginosa* Hfq protein, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 61*, 141−146.
30. Toro, I., Thore, S., Mayer, C., Basquin, J., Seraphin, B., and Suck, D. (2001) RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex, *EMBO J. 20*, 2293−2303.
31. Jarmolowski, A., and Mattaj, I. W. (1993) The determinants for Sm protein binding to *Xenopus* U1 and U5 snRNAs are complex and non-identical, *EMBO J. 12*, 223−232.
32. Antal, M., Bordeau, V., Douchin, V., and Felden, B. (2005) A small bacterial RNA regulates a putative ABC transporter, *J. Biol. Chem. 280*, 7901−7908.
33. Brescia, C. C., Mikulecky, P. J., Feig, A. L., and Sledjeski, D. D. (2003) Identification of the Hfq-binding site on DsrA RNA: Hfq binds without altering DsrA secondary structure, *RNA 9*, 33−43.
34. Moll, I., Leitsch, D., Steinhauser, T., and Blasi, U. (2003) RNA chaperone activity of the Sm-like Hfq protein, *EMBO Rep. 4*, 284−289.
35. Barrera, I., Schuppli, D., Sogo, J. M., and Weber, H. (1993) Different mechanisms of recognition of bacteriophage Qβ plus and minus strand RNAs by Qβ replicase, *J. Mol. Biol. 232*, 512−521.
36. Mikulecky, P. J., Kaw, M. K., Brescia, C. C., Takach, J. C., Sledjeski, D. D., and Feig, A. L. (2004) *Escherichia coli* Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs, *Nat. Struct. Mol. Biol. 11*, 1206−1214.
37. Folichon, M., Allemand, F., Regnier, P., and Hajnsdorf, E. (2005) Stimulation of poly(A) synthesis by *Escherichia coli* poly(A)-polymerase I is correlated with Hfq binding to poly(A) tails, *FEBS J. 272*, 454−463.
38. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, *Nucleic Acids Res. 25*, 3389−3402.
39. Higgins, D. G. (1994) CLUSTAL V: Multiple alignment of DNA and protein sequences, *Methods Mol. Biol. 25*, 307−318.
40. Sukhodolets, M. V., and Garges, S. (2003) Interaction of *Escherichia coli* RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq, *Biochemistry 42*, 8022−8034.
41. Gill, S. C., and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem. 182*, 319−326.
42. Edelhoch, H. (1967) Spectroscopic determination of tryptophan and tyrosine in proteins, *Biochemistry 6*, 1948−1954.
43. Carmichael, G. G., Weber, K., Niveleau, A., and Wahba, A. J. (1975) The host factor required for RNA phage Qβ RNA replication in vitro. Intracellular location, quantitation, and purification by polyadenylate−cellulose chromatography, *J. Biol. Chem. 250*, 3607−3612.
44. Lundblad, J. R., Laurance, M., and Goodman, R. H. (1996) Fluorescence polarization analysis of protein−DNA and protein−protein interactions, *Mol. Endocrinol. 10*, 607−612.
45. Royer, C. A., and Beechem, J. M. (1992) Numerical analysis of binding data: Advantages, practical aspects, and implications, *Methods Enzymol. 210*, 481−505.
46. Royer, C. A. (1993) Improvements in the numerical analysis of thermodynamic data from biomolecular complexes, *Anal. Biochem. 210*, 91−97.
47. Draper, D. E. (1999) Themes in RNA−protein recognition, *J. Mol. Biol. 293*, 255−270.
48. Jones, S., Daley, D. T., Luscombe, N. M., Berman, H. M., and Thornton, J. M. (2001) Protein−RNA interactions: A structural analysis, *Nucleic Acids Res. 29*, 943−954.
49. Sun, X., Zhulin, I., and Wartell, R. M. (2002) Predicted structure and phyletic distribution of the RNA-binding protein Hfq, *Nucleic Acids Res. 30*, 3662−3671.
50. Hopcroft, N. H., Wendt, A. L., Gollnick, P., and Antson, A. A. (2002) Specificity of TRAP−RNA interactions: Crystal structures of two complexes with different RNA sequences, *Acta Crystallogr., Sect. D: Biol. Crystallogr. 58*, 615−621.
51. Sonnleitner, E., Napetschnig, J., Afonyushkin, T., Ecker, K., Vecerek, B., Moll, I., Kaberdin, V. R., and Blasi, U. (2004) Functional effects of variants of the RNA chaperone Hfq, *Biochem. Biophys. Res. Commun. 323*, 1017−1023.
52. de Haseth, P. L., and Uhlenbeck, O. C. (1980) Interaction of *Escherichia coli* host factor protein with oligoriboadenylates, *Biochemistry 19*, 6138−6146.
53. de Haseth, P. L., and Uhlenbeck, O. C. (1980) Interaction of *Escherichia coli* host factor protein with Qβ ribonucleic acid, *Biochemistry 19*, 6146−6151.
54. Ackers, G. K., Shea, M. A., and Smith, F. R. (1983) Free energy coupling within macromolecules. The chemical work of ligand binding at the individual sites in co-operative systems, *J. Mol. Biol. 170*, 223−242.
55. Arluison, V., Mura, C., Guzman, M. R., Liquier, J., Pellegrini, O., Gingery, M., Regnier, P., and Marco, S. (2006) Three-dimensional structures of fibrillar Sm proteins: Hfq and other Sm-like proteins, *J. Mol. Biol. 356*, 86−96.
56. Diez, J., Ishikawa, M., Kaido, M., and Ahlquist, P. (2000) Identification and characterization of a host protein required for efficient template selection in viral RNA replication, *Proc. Natl. Acad. Sci. U.S.A. 97*, 3913−3918.
57. Collins, B. M., Harrop, S. J., Kornfeld, G. D., Dawes, I. W., Curmi, P. M., and Mabbutt, B. C. (2001) Crystal structure of a heptameric Sm-like protein complex from archaea: Implications for the structure and evolution of snRNPs, *J. Mol. Biol. 309*, 915−923.
58. Thore, S., Mayer, C., Sauter, C., Weeks, S., and Suck, D. (2003) Crystal structures of the *Pyrococcus abyssi* Sm core and its complex with RNA. Common features of RNA binding in archaea and eukarya, *J. Biol. Chem. 278*, 1239−1247.

BI0523613