RNA binding by Hfq and ring-forming (L)Sm proteins

A trade-off between optimal sequence readout and RNA backbone conformation

Oliver Weichenrieder

Department of Biochemistry; Max Planck Institute for Developmental Biology; Tübingen, Germany

The eukaryotic Sm and the Sm-like (LSm) proteins form a large family that includes LSm proteins in archaea and the Hfq proteins in bacteria. Commonly referred to as the (L)Sm protein family, the various members play important roles in RNA processing, decay, and riboregulation. Particularly interesting from a structural point of view is their ability to assemble into doughnut-shaped rings, which allows them to bind preferentially the uridine-rich 3'-end of RNA oligonucleotides. With an emphasis on Hfq, this review compares the RNA-binding properties of the various (L)Sm rings that were recently co-crystallized with RNA substrates, and it discusses how these properties relate to physiological function.

Heptamer-forming (L)Sm Proteins Have a Conserved Nucleotide Binding Pocket

Sm proteins were originally identified as the targets of auto-antibodies from the Sm (Stephanie Smith) serotype of patients with systemic lupus erythematosus.[1] Soon thereafter they were found to be part of the spliceosomal U snRNPs[2,3] and to bind to internal, uridine-rich RNA sequences of the U snRNAs[4] around which they form heteroheptameric Sm rings.[5] A series of other Sm-like proteins were identified in eukaryotic species and were suggested to assemble into at least two additional heteroheptameric rings, the cytoplasmic LSm1–7 ring involved in mRNA decay, and the nuclear LSm2–8 ring that binds to the 3'-end of the U6 snRNA (**Fig. 1**).[6-8] However, the first fully assembled LSm ring to be crystallized was the homoheptameric ring formed by the LSm protein *Af*Sm1 from the archaeon *Archaeoglobus fulgidus*.[9] *Af*Sm1 can be regarded as a simplified prototype for the Sm fold (**Fig. 1A and B**), which consists of an N-terminal α-helix that packs against a five-stranded half-open β-barrel.[5] The structure of the *Af*Sm1 ring confirmed

that the suggested assembly of (L)Sm proteins into heptamers[5] via the outer β-strands (β4 and β5) was correct (**Fig. 1A and B**). Furthermore, the structure demonstrates how one uridine per LSm monomer is recognized in a specific nucleotide binding pocket (**Fig. 1A and C**). The uridine base is sandwiched between H37 and R63 as the upper and lower π-stacking partners. Its Watson-Crick edge is read out by hydrogen bonds to N39 on β-strand β3, explaining why uridines are a preferred substrate. The highly conserved R63 can also reach and recognize the ribose on the RNA backbone, whereas D65 on the L(β4-β5) loop seems to limit the space of the backbone phosphates inside the pore of the ring. This nucleotide binding pocket is not present in all LSm proteins,[6,7,10] but it is conserved in ring-forming Sm and LSm proteins from eukaryotes, and therefore, amino acids will frequently be referred to in the following by their position in the *Af*Sm1 protein.

The Structure of Hexameric Hfq Reveals Modified Nucleotide Binding Pockets at the Subunit Interfaces

The Hfq protein was originally identified as a bacterial host factor for the Qβ phage replication.[11] Since then, it has been implicated in numerous processes in bacterial RNA metabolism and by now it is probably one of the best studied RNA binding protein in bacteria, which has recently been reviewed extensively.[12-21] Its most prominent role is in riboregulation by bacterial small RNAs (sRNA). The realization that Hfq shares the Sm-fold came relatively late,[22,23] because Hfq sequences lack an important sequence signature that commonly identifies eukaryotic and archaeal (L)Sm proteins[6-8,24] and that results from their conserved nucleotide binding pocket (**Fig. 1A, C, and E**). Moreover, the Hfq protein assembles into homohexameric rings[22,23] rather than into heptamers, as confirmed by the first crystal structures.[25,26] These structures also revealed a modified nucleotide binding pocket as compared with the heptameric (L)Sm proteins (**Fig. 1A, C, and E**).[25,27] Whereas F42 from the prototypic *Escherichia coli* (*Ec*) Hfq (the structural equivalent of H37 in *Af*Sm1) still provides a stacking
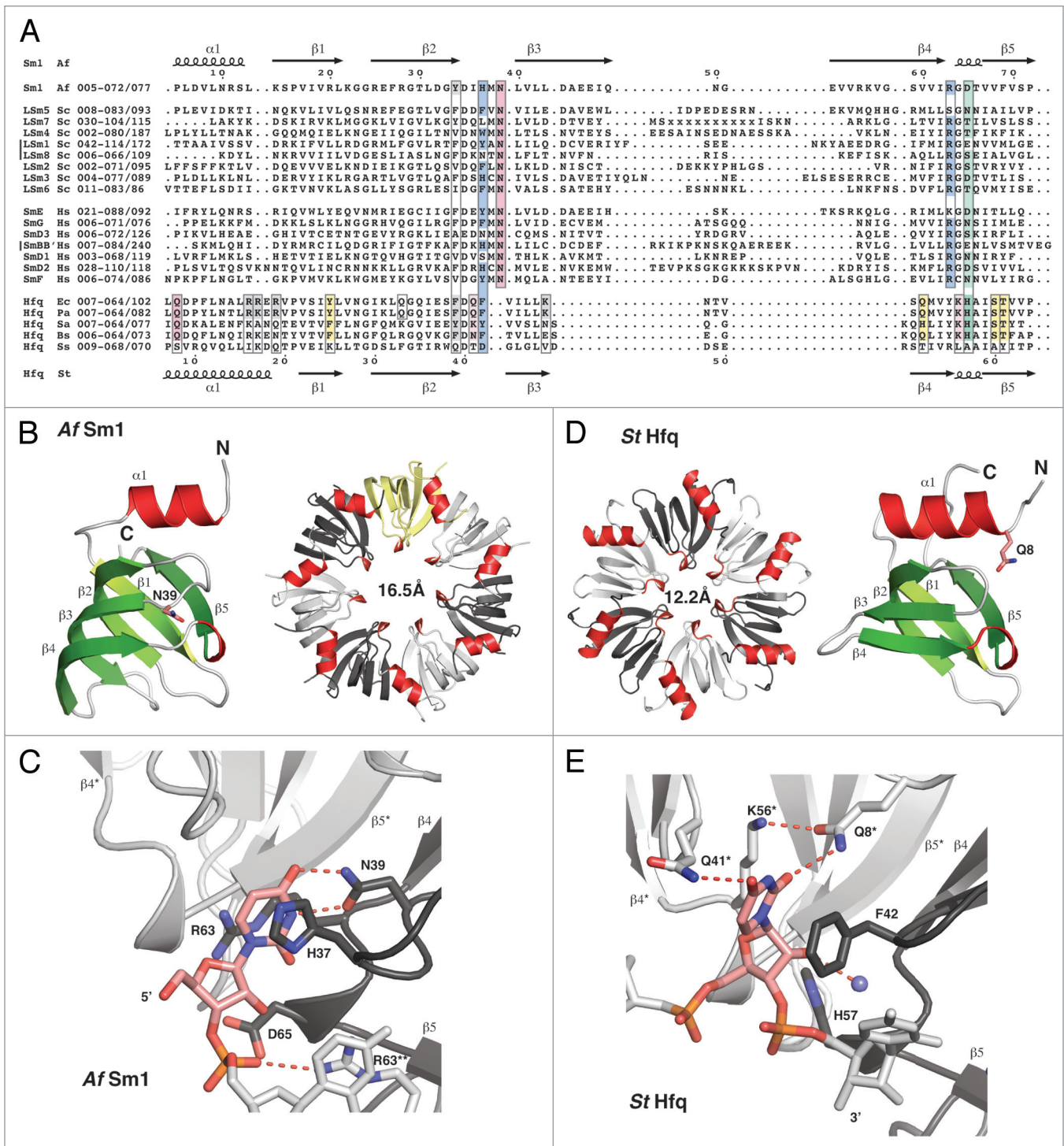
Figure 1. For figure legend, see page 539.

platform for the base, and whereas H57 (the equivalent of D65 in *Af*Sm1) still is in a position to contact the RNA backbone, the equivalent of N39 is missing. Instead, the Watson-Crick edge of a bound uridine is read out by Q8* on helix α1 from the neighboring Hfq monomer (indicated by the asterisk). Q8* is highly conserved among Hfq proteins and assisted in base recognition by K56* and Q41*. The pocket shows a preference for uridine, but can also accommodate cytidine and adenine.[26,28-30] It is rather clear by now however that it strongly discriminates against guanine.[26,28,29] In the hexamer, the six-nucleotide binding pockets are located in close proximity on the inner rim of the so-called proximal surface of the ring. Collectively, they are also referred to as the proximal RNA binding site (**Fig. 2**).

Figure 1 (see opposite page). Alignment and structure of crystallized ring-forming (L)Sm and Hfq proteins. (**A**) Structure-based sequence alignment. Secondary structure elements and numbering above and below the alignment are from the *Af*Sm1 and *St*Hfq proteins, respectively. Only the crystallized (L)Sm-core of each protein is aligned. The numbers preceding each sequence refer to the aligned residues and to the total length of each protein. They also indicate the presence and size of C-terminal extensions. The crystallized sequences of *St*Hfq and *Ec*Hfq are identical. Positions involved in RNA or nucleotide binding are boxed. Blue: Stacking partners for the nucleobase in the proximal site. Hydrogen bonds to the RNA backbone in the case of R63 and its (L)Sm homologs. Red: Base recognition by hydrogen bonds in the proximal site. Green: Constriction of the RNA backbone in the proximal site and RNA 3'-end recognition in the case of Hfq. Yellow: Conserved "R"-site on the distal surface of Hfq; Q33 in *Ec*Hfq: "A"-site on the distal surface of *Ec*Hfq, not present in many gram-positive bacteria. Gray: Conserved uridine-site on the outer rim of the proximal surface of Hfq and of archaeal LSm proteins, plus members of the basic patch on the lateral surface of *St*Hfq and *Ec*Hfq (R16, R17, R19, K47) that is less well conserved in gram-positive bacteria. Sequence information is from the respective PDB files listed by PDB-ID, chain-ID, resolution, and species: Sm1, 1i4k_A,[9] 2.50 Å, *Archaeoglobus fulgidus*, *Af*; LSm1–7, 4c92_ABCDEFG,[84] 2.30 Å, and 4m75_A,[85] 2.95 Å, *Saccharomyces cerevisiae*, *Sc*; LSm8, 4m77_A,[86] 3.11 Å, *Saccharomyces cerevisiae*, *Sc*; Sm proteins, 2y9a_ABCDEFG,[87] 3.60 Å, *Homo sapiens*, *Hs*; EcHfq, 2y1c_A,[29] 1.30 Å, *Salmonella typhimurium, St* and *Escherichia coli*, *Ec*; PaHfq, 4j6y_A,[28] 2.14 Å, *Pseudomonas aeruginosa*, *Pa*; SaHfq, 1kq2_A,[26] 2.71 Å, *Staphylococcus aureus*, *Sa*; BsHfq, 3ahu_A,[39] 2.20 Å, *Bacillus subtilis*, *Bs*; SsHfq, 3hfo_A,[79] 1.30 Å, *Synechocystis* ssp., *Ss*. (**B**) LSm monomer and homoheptamer. The monomer of the archaeal *Af*Sm1 protein[9] is shown with α-helices in red and β-strands in green. N39 (sticks) marks the nucleotide binding pocket shown in (**C**). The loop L(β4-β5) forms a short 3–10 helix that is highly conserved. The protomers in the heptameric ring are colored with in white, gray and yellow, and helices in red for orientation. The diameter of the pore is 16.5 Å (proximal view), using the Cβ atom of D65 as a reference. (**C**) Nucleotide binding in the proximal site of an LSm protein. Uridine recognition is shown for *Af*Sm1,[9] with optimal geometry for hydrogen bonds (dotted red). Important side chains are shown as sticks with nitrogens in blue and oxygens in red. The asterisk and double asterisk mark the preceding and following protomer in the 5'-3' direction of the bound RNA, respectively. (**D**) Hfq monomer and homohexamer. Compared with the *Af*Sm1 protein, the monomer of the bacterial *St*Hfq protein[29] reveals an extended α-helix α1, a loop L(β2-β3) that has a distinct structure, and a variable loop L(β3-β4) that is very short. Q8 (sticks) marks the modified nucleotide binding pocket shown in (**E**). The diameter of the pore is 12.2 Å (proximal view), using the Cβ atom of H57 as a reference. (**E**) Nucleotide binding in the proximal site of Hfq. Uridine recognition by hexameric *St*Hfq[29] is structurally distinct from (L)Sm proteins and involves residues from two neighboring protomers. Base recognition geometry is rather poor, a potential consequence of the constricted RNA backbone conformation. A conserved water molecule is shown as a blue sphere (see also **Fig. 2**). H57 is positioned close to the 3'-oxygen of the ribose. The alignment was done with the help of ESPRIPT,[110] structural analysis and figures were done with the help of COOT[111] and PYMOL (http://www.pymol.org).

The Proximal RNA Binding Site on Hfq Hexamers Preferably Binds to the 3'-Ends of RNA

An initial crystal structure of Hfq from the gram-positive bacterium *Staphylococcus aureus* (*Sa*) revealed how the proximal RNA binding site can bind to an oligonucleotide (5′-AUUUUUG-3′) with internal uridines that was derived from the RNA binding site[4] of the eukaryotic Sm ring on U snRNAs. In this structure, the six-nucleotide binding pockets of the *Sa*Hfq ring are occupied by the first six residues of the oligonucleotide, whereas the 3′-terminal guanine is expelled (**Fig. 2A**).[26]

Subsequent binding studies with *Salmonella typhimurium* (*St*) Hfq demonstrated however that the proximal site binds uridine-rich oligonucleotides much better if the 3′-terminal nucleotide is a uridine.[29] In this case, the 3′-terminal nucleotide occupies one of the binding pockets and the 3′-hydroxyl group of the last ribose in the chain is directly recognized by H57, as indicated by a high-resolution crystal structure of *St*Hfq in complex with hexauridine (5′-UUUUUU-3′, **Fig. 2B**). The same recognition is possible with an adenine or cytidine in the terminal position, whereas phosphate modification of the 3′ hydroxyl group leads to a strong reduction of affinity.[29]

The crystal structure also reveals a much more constricted backbone conformation than observed with the *Sa*Hfq complex and a different binding register for the phosphate and ribose moieties that is likely triggered by the positioning of the 3′-terminal hydroxyl group (**Fig. 2A and B**).[29] These differences indicate a general difficulty to accommodate the RNA backbone in the constricted space of the pore and a trade-off between optimal nucleotide binding and optimal backbone geometry. Indeed, such a trade-off might also explain the poor base-recognition geometry observed in the crystal structures[26,29,30] and it might

be another reason for Hfq hexamers to preferably bind RNA 3′-ends, where the strain on the RNA backbone is reduced.

Regarding Hfq function, the realization that Hfq binds RNA 3′-ends[29,31] was particularly important, because it suggested and explained how Hfq protects and stabilizes RNA transcripts with uridine-rich 3′-ends resulting from Rho-independent transcription termination.[32] These transcripts include the small regulatory sRNAs for the function of which Hfq plays a central role.[12-14,16,18-20]

The trade-off between nucleotide binding and an optimal backbone geometry can also lead to more complex modes of oligonucleotide binding that might quite frequently occur in nature, but are hard to crystallize and observe. A good example is the remarkable crystal structure of a 5′-AUUUUUA-3′ oligonucleotide bound in the proximal site of *Ec*Hfq (**Fig. 2C**, the crystallized sequence of *Ec*Hfq is 100% identical to *St*Hfq).[30] The structure demonstrates once more the preference of the proximal site for RNA 3′-ends and it shows how an adenine is recognized and tolerated at the 3′-end. Furthermore, the 3′-hydroxyl group is indeed recognized by H57 in the way as suggested before,[29] and a magnesium ion plays a crucial role in the stabilization of the irregular conformation of the RNA backbone (**Fig. 2C**). Intriguingly, only four of the six nucleotide binding pockets are occupied, and the corresponding uridines are either looped out (**Fig. 2C**, protomer 4) or participate in the stabilization of the RNA backbone (**Fig. 2C**, protomer 1) by a hydrogen bond from the N3 position to the last phosphate five nucleotides downstream.

In summary, the preference of the proximal RNA binding site for RNA 3′-ends over internal RNA sequences likely results from the direct recognition of the 3′-hydroxyl group and from the possibility to engage the nucleotide binding pockets with less strain on the phosphoribose backbone and with more contacts to the protein surface.
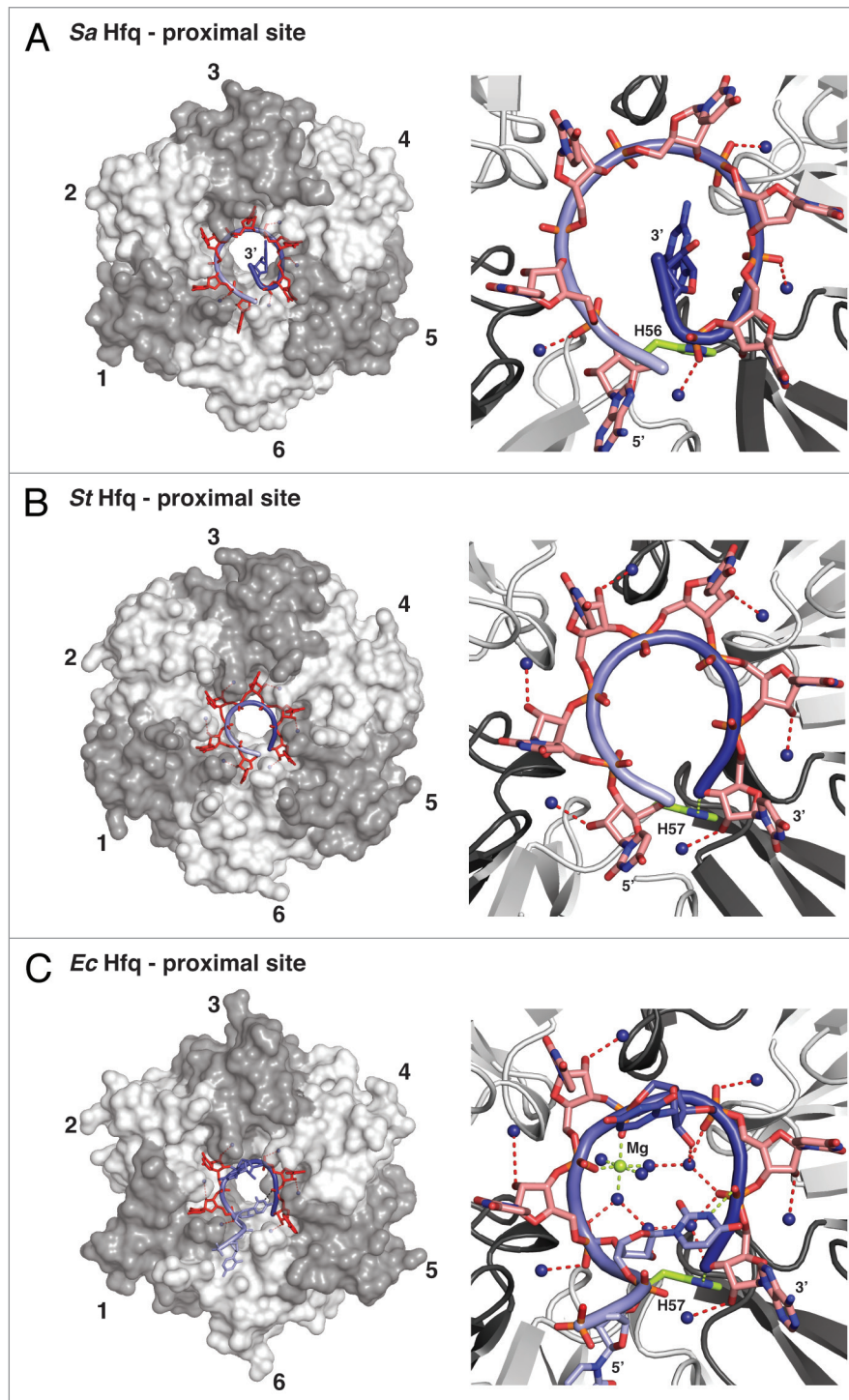
Figure 2. RNA binding and 3'-end recognition in the proximal site of Hfq. Left panels: Overview, with the protein as transparent surfaces (ring orientation as in **Fig. 1D**). Right panels: Zoom, with the protein as cartoon. (**A**) Dilated RNA conformation and an expelled 3'-end on *Sa*Hfq.[26] Hfq protomers are colored in white and gray and numbered clockwise for orientation. RNA is shown as sticks with a backbone cartoon that is colored with a gradient from light to dark blue to indicate the 5'-3' direction. Nucleotides that reside in specific binding pockets are in red, the expelled terminal guanine is in dark blue. Dark blue spheres indicate conserved water molecules that, in the dilated conformation, contact the phosphates of the RNA backbone (hydrogen bonds dotted red). (**B**) Constricted RNA conformation on *St*Hfq.[29] The constricted conformation allows the recognition of the 3'-end by H57 (residue and dotted hydrogen bond in lime). The conserved water molecules contact the ribose of the RNA backbone. (**C**) Mixed RNA conformation on *Ec*Hfq.[30] The irregular backbone alternates between the dilated and constricted conformations as illustrated by the contacts to the conserved waters. One of the uridines (dark blue) is expelled from its pocket, and another uridine (light blue) stabilizes the backbone conformation by a hydrogen bond (lime) to the last phosphate in the RNA chain. The 3'-hydroxyl group of the terminal adenine is recognized via a hydrogen bond to H57 (both in lime). The stabilizing Mg^{2+}-ion and its coordination are also shown in lime, the remaining hydrogen bond network as dotted red lines.

The Distal Surface of Hfq Hexamers Binds Adenine-Rich RNAs

*Ec*Hfq also binds oligo-(A) tracts with high affinity.[16,33-35] This has been puzzling until it was realized that the distal surface of the Hfq hexamer provides a second RNA binding surface,[36] and crystal structures revealed how adenine-rich oligonucleotides are recognized (**Fig. 3**).[37-39] Each subunit contains a conserved nucleotide binding pocket (originally termed the "R" site)[38] that allows the base to stack on the aromatic Y25 and that is lined by Q52, S60, and T61 to recognize the Watson-Crick edge (**Fig. 1A**, sequence and numbers according to *Ec*Hfq). Experimentally, the "R"-site has been shown to be highly selective for adenine, although other nucleotides might occasionally be accommodated as well.[28,37-41]

The "R"-site pockets allow *Ec*Hfq to recognize every third adenine of an oligo-(A) tract (**Fig. 3A**),[38] whereas the "R"-site pockets on Hfq hexamers from *Bacillus subtilis* (*Bs*), *Staphylococcus aureus*, and probably many other gram-positive bacteria recognize every second adenine (**Fig. 3B**).[37,39] In the latter case, the nucleotide that links two recognized adenines (linker nucleotide) frequently seems to adopt the rare syn-conformation, but linker nucleotide identity does not seem to be conserved between species. In the case of *Ec*Hfq, there are two linker residues between the adenines in the "R"-site pockets. The first one of these protrudes into the bulk solvent and is not recognized specifically, whereas the second one fits into an additional shallow and adenine-specific binding pocket that is not present in *Bs*Hfq and *Sa*Hfq and that reads out the Hoogsteen-edge of the base via contacts to the peptide-backbone of Q33 on β-strand β2 (**Figs. 1A and 3**).

As a consequence, the Hfq hexamers from *Escherichia coli* and related gram-negative bacteria preferably bind RNA sequences with an (5′-ANA-3′)$_n$ signature on their distal surface, whereas Hfq hexamers from many gram-positive bacteria seem to prefer a (5′-AN-3′)$_n$ signature.[37-39,41] Furthermore, at least four Hfq subunits need to be engaged for an optimal binding, and longer oligo-(A) tracts can be bound cooperatively by Hfq hexamers.[16,42]

Additional RNA Binding Surfaces on Hfq Hexamers are Responsible for the Specific Recognition of Regulatory sRNA

Small regulatory RNAs in bacteria (sRNAs) are characterized by (1) a usually 5′-terminal "seed" sequence that is complementary to the target mRNA, (2) a central sRNA "body," and (3) a 3′-terminal stem-loop structure that ends in an oligo-(U) tail as a consequence of Rho-independent transcription termination.[19,32,43,44] The size and shape of the RNA body varies considerably among sRNAs and they frequently contain additional uridine-rich stretches that alternate with small, often only transiently stable base-pairs and stem-loop structures.[42-46] Nevertheless, Hfq recognizes these sRNAs specifically and distinguishes them from other RNAs in the cell.[22,23,47-49] With the described RNA binding properties of the proximal and distal surfaces of the Hfq ring, the specificity of the recognition and the protection of the

Figure 3. RNA binding on the distal surface of Hfq. (**A**) Recognition of an RNA with a (5'-ANA-3')$_6$ signature on *Ec*Hfq.[38] Transparent surfaces of Hfq protomers are colored in white and gray and numbered counterclockwise to relate the arrangement to **Figure 2**. RNA is shown as sticks with a backbone cartoon that is colored with a gradient from light to dark blue to indicate the 5'-3' direction. Nucleotides that reside in specific binding pockets are in dark red (conserved "R"-site) and light red ('A'-site for the second linker nucleotide), the unbound linker nucleotide is in blue. (**B**) Recognition of an RNA with a (5'-AN-3')$_6$ signature on *Bs*Hfq,[39] a representative from a gram-positive bacterium. Nucleotides that reside in specific binding pockets are in dark red (conserved "R"-site), the unbound linker nucleotide is in blue.

entire sRNA body from nuclease attack[12,22,23,42,43,50-52] were hard to explain.

This situation changed with the realization that certain sRNAs still bind Hfq even when their 3′-end is blocked by a cyclic phosphate and even when the proximal and distal RNA binding surfaces of the Hfq hexamer are mutated or occupied by other small oligonucleotides.[42] Obviously, there had to be additional RNA-binding surfaces on Hfq (**Fig. 4**) that were subsequently identified to include a basic patch on the lateral surface of each Hfq protomer. On *St*Hfq and *Ec*Hfq, the basic patch consists of R16, R17, R19, and K47, and it is conserved primarily in gram-negative bacteria (**Figs. 1A and 4A**).[42] Furthermore, an additional uridine binding site was identified by crystallography on
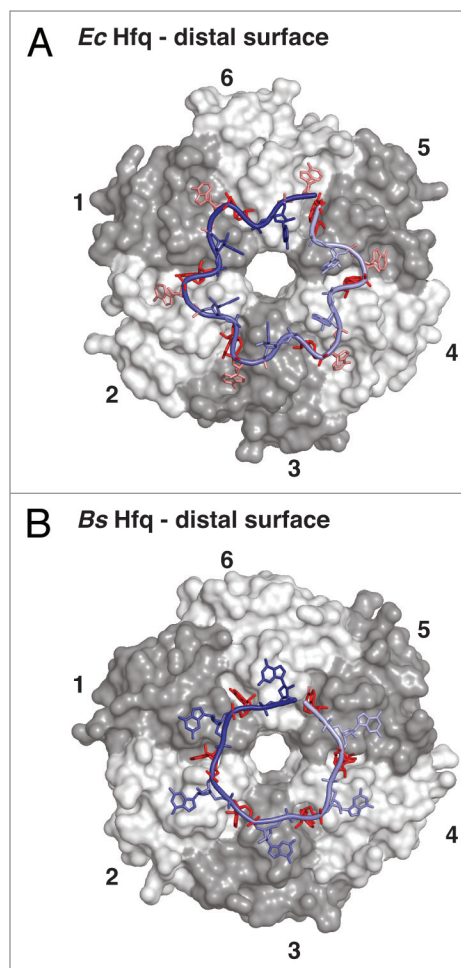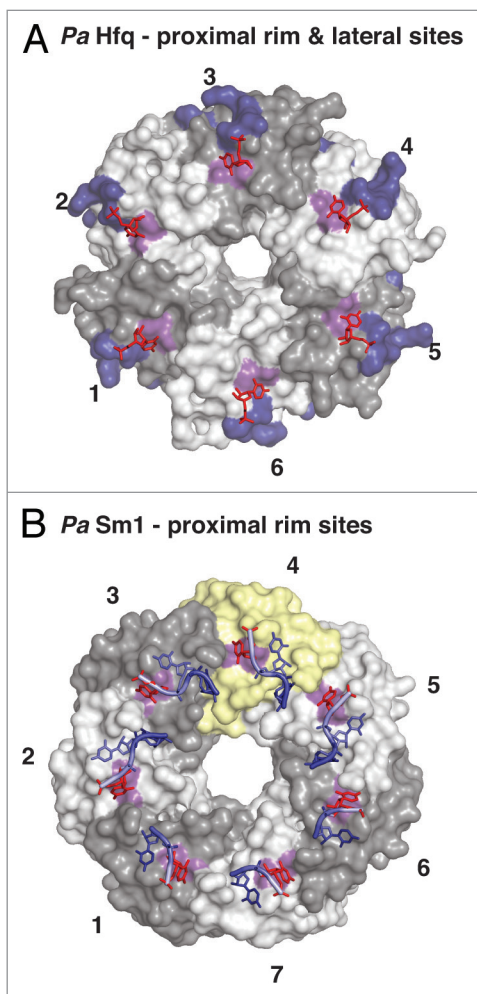
Figure 4. Additional RNA binding surfaces on Hfq and (L)Sm proteins. (**A**) RNA binding on the outer rim of the proximal surface and on the lateral surface of Hfq. Transparent surfaces of Hfq protomers are colored in white and gray and numbered for orientation. A conserved uridine binding site on the proximal surface of bacterial *Pa*Hfq (*Pseudomonas aeruginosa*) is colored in purple (corresponding to F39 in *Ec*Hfq).[28] The areas corresponding to the basic patch on the lateral surface of *St*Hfq are colored in dark blue (corresponding to R16, R17, R19, K47 in *St*Hfq and *Ec*Hfq).[29] The basic patch allows the specific recognition of regulatory sRNAs and catalyzes base-pair formation and exchange. The uridines co-crystallized with *Pa*Hfq are shown as red sticks. (**B**) Conservation of the uridine binding site on an archaeal LSm heptamer.[53] The surface of the additional protomer in the homoheptameric ring of the archaeal *Pa*Sm1 (*Pyrococcus abyssii*) protein is shown in yellow. The conserved uridine binding site is colored in purple (corresponding to Y34 in *Af*Sm1) and the bound uridines are shown as red sticks. Other nucleotides and the RNA backbone of the co-crystallized oligomers are drawn in blue as described before.

the outer rim of the proximal surface in *Pseudomonas aeruginosa* (*Pa*) Hfq, where the base stacks on the aromatic F39 in β-strand β2 (F39 in *Ec*Hfq) and where specificity is achieved by hydrogen bonds from the Watson-Crick edge of the uridine to the peptide backbone of the amino acid (**Figs. 1A and 4A**).[28] This uridine binding site is close to the basic patch and may indeed be quite conserved, because it is also found in the crystal structure of the heptameric LSm ring of the *Pa*Sm1 protein from the archaeon

Pyrococcus abyssii (Y34 in *Af*Sm1) that was co-crystallized with a heptauridine (5′-UUUUUUU-3′) oligonucleotide (**Fig. 1A and 4B**).[28,53]

Importantly, the basic patch is accessible by oligonucleotides both from the proximal and from the distal surface of hexameric Hfq,[42,54] and the respective arginines with their intrinsic ability to stack and pair with nucleotide bases were shown by mutational analysis to participate actively in the nucleation of RNA helices.[54] Consequently, the basic patch (possibly extended by F39) may therefore not only bind and expose single-stranded uridine-rich sequences from the sRNA body, but it may also recognize and stabilize transient base-pairs and stem-loop structures as they occur within the sRNA body or between the seed and target sequences as they begin to anneal.[42,54]

Multiple Roles of Hfq in Riboregulation

Given the abundance of Hfq in the bacterial cell and its rather general and structure-specific RNA binding properties, it is not surprising that Hfq has been identified in a very large number of regulatory processes where RNA is involved.[12-14,16,19,20] The most important functions and selected examples are listed below with increasing complexity (**Fig. 5**).

Protection against and recruitment of nucleases

One of the easiest ways for Hfq to affect RNA stability and gene expression is the attachment to an RNA molecule and its protection from nuclease attack. For example, it has long been known that Hfq stabilizes A-tracts and oligo-(A) tails, which likely happens via the distal surface and arguably in a cooperative manner.[16,33-35,42,55-57] Furthermore, the proximal binding site can protect the 3′-oligo-(U) ends of Rho-independent transcripts from the attack of PNPase,[50] and uridine-rich sequences that are present in the interior of both mRNAs and sRNAs and are protected from RNaseE probably because they interact with the additional RNA-binding surfaces of Hfq.[16,34,51] In particular, many sRNAs are strongly destabilized in the absence of Hfq.[22,23,52]

On the other hand, Hfq also interacts with PNPase and RNaseE,[12,13,19,58,59] potentially even directly.[60] In the case of the RyhB sRNA, this property of Hfq is exploited specifically for the targeted degradation of sodB mRNA by RNaseE. This works even in the absence of sodB mRNA translation, outlining a pathway of active, sRNA-induced mRNA degradation.[61]

Rapid recycling of sRNAs on Hfq

In the bacterial cell, the amount of Hfq is considered to be limiting such that numerous sRNAs compete for Hfq binding.[20,46,62-64] Consequently, the ability of Hfq to rapidly exchange bound sRNAs while maintaining a high binding affinity and specificity becomes an important property.[20,62] This property allows Hfq to act like a hub, integrating different sRNA responses and it allows the cell to switch rapidly between the regulation by different sRNAs, simply because freshly induced sRNA molecules will displace previous sRNAs from Hfq, leading to their rapid degradation.[20,42,46,62-64] Mechanistically, the dynamic interaction with sRNAs and

their rapid competitive displacement can be explained by a "peeling" mechanism, where the sRNA disengages the lateral surfaces consecutively one by one rather than breaking all of the contacts with Hfq in a single step.[42,62]

Induction of conformational changes

Another important function of Hfq is its action as an RNA chaperone,[65-68] inducing conformational changes in the bound RNA. A recent example of a positional effect is from the 5′ leader sequence of rpoS mRNA, where the distal surface of Hfq binds to an upstream $(5'-ANA-3')_4$ motif and stabilizes an RNA conformation that allows sRNAs (DsrA, RprA, ArcZ) to act on a downstream target site.[69] Alternatively, Hfq may also interfere directly with the formation of an inhibitory secondary structure and expose the bound nucleotides in a conformation that is optimal for base-pairing with a target RNA as suggested for the lateral binding surface of Hfq.[42]

Assistance of sRNA targeting

A single Hfq ring can stably bind two RNA molecules at the same time, such as an oligo-$(A)_{20}$ RNA on the distal surface and a RybB sRNA that is 3′-anchored in the proximal site and covers several of the basic patches.[42] Consequently, Hfq can stimulate sRNA targeting not only by conformational activation, but also by bringing the binding partners closely together in a ternary sRNA-Hfq-mRNA complex that should exist at least transiently before further conformational changes and downstream events take place.[19,36,42,70] Indeed, many mRNAs that are known sRNA targets contain adenine-rich $(5'-ANA-3')_n$ motifs in their 5′-UTRs close to the translation start site,[19,38,47,71] supporting the notion that Hfq actively assists bound sRNAs to associate with their mRNA targets.

Direct catalysis of base-pair formation

Finally, the basic patch not only recognizes transient base-pairs and stem-loops as part of the sRNA body,[42,45] but the respective arginine residues also promote RNA-helix formation between distinct RNA oligonucleotides that are anchored on the proximal and distal surfaces of the Hfq hexamer, respectively. In this sense, the basic patch has enzyme-like active site properties, stabilizing initial base-pairs as a transition state.[54]

Future Challenges Regarding Hfq

It becomes clear from this list that the contribution of Hfq to riboregulation is highly complex and variable for each RNA partner, and this will continue to present a challenge for future investigations. Mutational approaches in vivo reveal
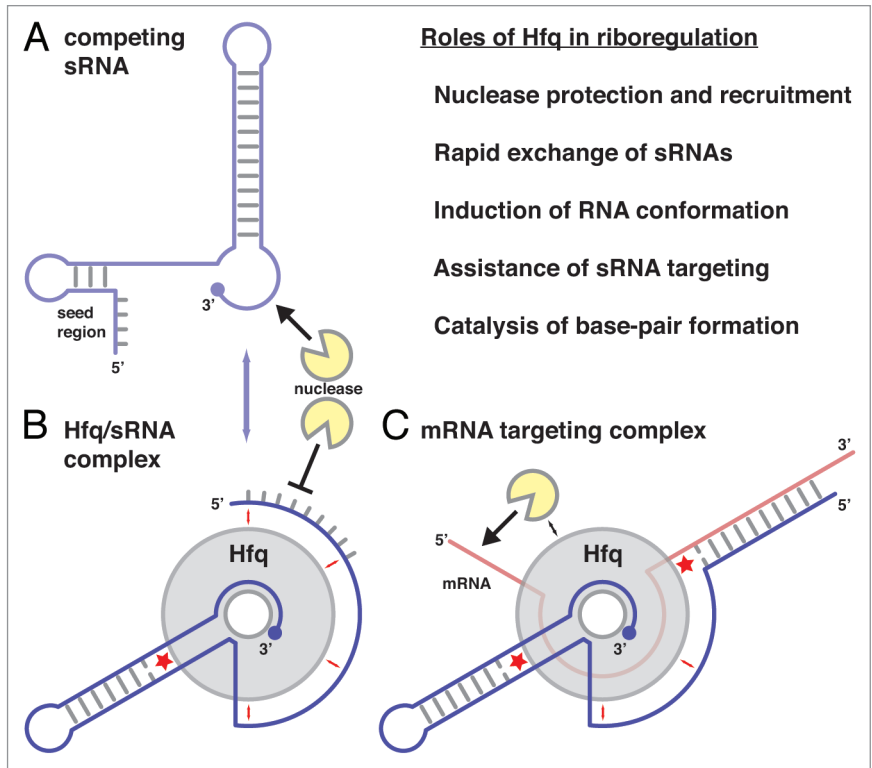

Figure 5. Roles of Hfq in riboregulation. (**A**) Schematic structure of an sRNA with a 5'-terminal seed region and a 3'-terminal terminator stem-loop structure. sRNAs compete rapidly for Hfq binding (blue arrow). (**B**) Hfq/sRNA complex. The sRNA is shown with its 3'-end anchored in the proximal site of Hfq and its body adapted to the lateral sites (small red arrows). (**C**) mRNA targeting complex with an mRNA bound to the distal surface of Hfq. The lateral sites assist in the (de-) formation of base-pairs (red asterisks). The topology of the complexes may vary from case to case;[72] the directionality of the RNA on the proximal and distal surfaces of Hfq corresponds to the crystal structures. Figure modified from Sauer, et al.[42]

considerable differences on how individual sRNAs interact with Hfq,[72] an observation which is supported by structural studies in solution.[73,74] However, to capture the dynamics of sRNA Hfq interactions requires novel approaches such as single molecule techniques. These would also allow it to address the roles of the presumably disordered C-terminal extensions that are present in many gram-negative bacteria (**Fig. 1A**). In the simplest view, these C-terminal tails might act as a "molecular bumper" or as a protective cage around a bound sRNA, but more specific roles in riboregulation have also been suggested.[75,76]

Furthermore, the roles of Hfq in many bacterial species may differ substantially from what has been learned from the gram-negative species that have been studied most intensively. For example, in the Hfq hexamers from the gram-positive *Staphylococcus aureus*, the C-terminal tails are short, the basic patch on the lateral surface is only poorly conserved (**Fig. 1A**),[42] and the distal surface reveals an altered mode of RNA recognition.[37] Functional assays suggest a much less prominent role of Hfq-bound sRNAs in *Staphylococcus aureus* riboregulation.[77,78] Another example is cyanobacterial Hfq (*Synechocystis*, *Ss*Hfq) that appears not to be involved in riboregulation at all[79] and many bacterial pathogens such as *Helicobacter pylori* do not even encode a protein with an Sm fold.[80,81] It shall be highly
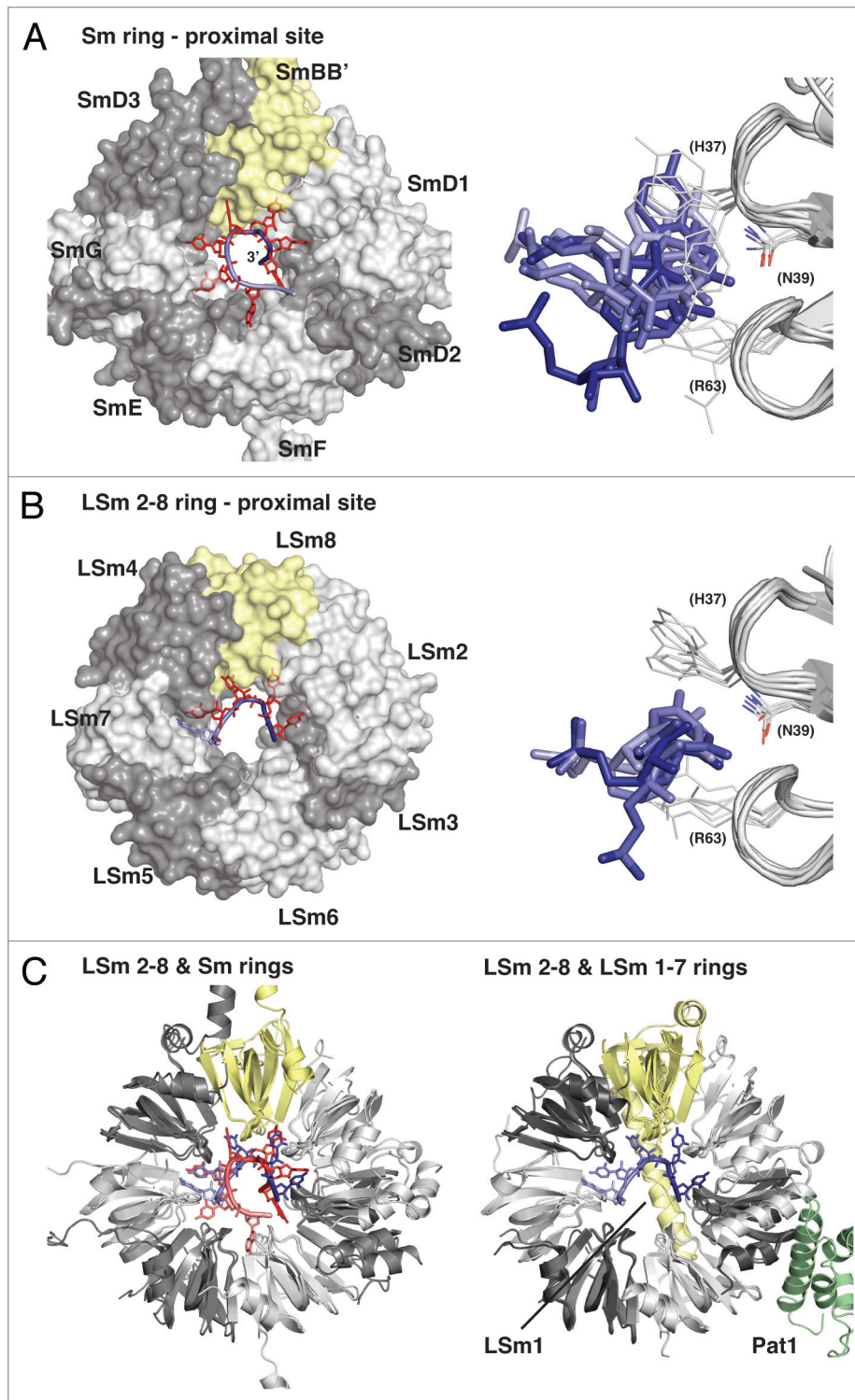
Figure 6. For figure legend, see page 545.

interesting to learn how riboregulation is achieved in these species.

Finally, in species where Hfq does exist, its impact on the molecular biology of the entire system deserves particular attention.[82] Indeed, the entire transcriptome is affected by the presence of Hfq, because RNA molecules have not only evolved to positively engage the protein and exploit it for riboregulation, but apparently also to prevent the inadvertent association with Hfq by chance. Examples can be seen by the underrepresentation of $(5'-ANA-3')_n$ stretches in protein coding regions of mRNAs[47,71] or in the sequestration of the uridine-rich 3′-tail into a stable stem-loop structure in GlmY sRNA to prevent

Hfq association in the context of amino sugar pathway regulation.[83]

The Recognition of RNA 3'-Ends by Eukaryotic (L)Sm Proteins

Recently, crystal structures became available of the LSm1–7[84,85] and the LSm2–8 ring,[86] the latter also in complex with a short RNA oligonucleotide (5′-GCUUUU-3′). Together with the previously published structures of the U1 and U4 snRNPs that contain the Sm ring,[87-89] these data allow a first comparison of the hetero-heptameric rings from eukaryotic species (**Fig. 6**), especially with respect to the mechanism of RNA 3′-end recognition that has been demonstrated for Hfq.[29,31]

The Sm Ring in U snRNPs

The Sm ring (**Fig. 6A**) is specialized for the recognition of the Sm site on U snRNAs,[4] which is a single-stranded and defined RNA sequence (5′-AUUUUUG-3′ in the crystallized structures[87-89]), located internally, in between base-paired parts of U4 RNA that are fixed above and below the Sm ring. As a consequence, the Sm ring needs to assemble around the Sm site on U snRNAs, a process for which several assembly factors have been identified in mammalian cells.[90,91] The Sm ring consists of the protomers SmE, SmG, SmD3, SmBB', SmD1, SmD2, and SmF, listed in the order of engagement by the RNA in the 5′ to 3′ direction. The RNA enters via the protomer SmE from the proximal surface and exits via the protomer SmF before threading through the pore to the distal side (**Fig. 6A**). Each protomer binds a nucleotide in its binding pocket. The pockets are in the same structural positions as in the archaeal *Af*Sm1 prototype (**Fig. 1C**) and the amino acid corresponding to asparagine N39 that contacts the nucleotide at the Watson-Crick edge is strictly conserved. Similarly conserved is R63, the arginine that serves as the lower π-stacking partner and that contacts the phosphoribose backbone. An exception is SmE that binds the adenine at the RNA entrance and which contains a lysine instead. The upper π-stacking partner corresponding to H37 in *Af*Sm1 is less well conserved, but it provides π-electrons for base-stacking in all cases except SmD1, which contains a serine (**Fig. 1A**). Importantly, the pockets reveal considerable variation in the orientation of their binding partner, even for the five central uridines where not every Watson-Crick edge is read out with optimal geometry (**Fig. 6A**).[87] Similar to the situation in bacterial Hfq (**Fig. 1E**), there appears to be a trade-off between optimal base recognition by the conserved asparagine and an ideal backbone conformation in the constricted space of the pore. For the Sm ring, the placement of the phosphoribose backbone inside the narrow pore might pose a particular challenge because it assembles around an internal RNA binding site with the need to accommodate the continuing 3′-sequence. As a consequence, the more limited conformational freedom of the RNA molecule in the absence of a 3′-end might be the reason why none of the bases is in an optimal position with respect to the interacting asparagine (N39 in *Af*Sm1).

The LSm2–8 Ring

In contrast to the Sm ring, the LSm2–8 ring (**Fig. 6B**) assembles in the absence of RNA[92] and it specifically recognizes the processed 3′-end of the spliceosomal U6 snRNA, which terminates with five uridines and a 2'-3'-cyclic phosphate instead of the usual 3′-hydroxyl group.[93] This substrate specificity is reminiscent of the bacterial Hfq rings, only that a terminal phosphate decreases the affinity in the case of Hfq, whereas it increases the affinity in the case of the LSm2–8 ring.[29,93] Based on sequence similarity, the LSm protomers (LSm5, LSm7, LSm4, LSm8, LSm2, LSm3, and LSm6) can be aligned to the corresponding Sm protomers (SmE, SmG, SmD3, SmBB', SmD1, SmD2, and SmF), and they assemble in the same order (**Fig. 1A**).[86] The preference for uridines is reflected by conserved nucleotide binding pockets, with the exceptions of LSm7 and LSm8, which lack the aromatic π-stacking partner in the upper position, and especially LSm5, where the lower π-stacking partner, the highly conserved arginine (R63 in *Af*Sm1), is replaced by a serine (**Fig. 1A and C**). The LSm2–8 ring was co-crystallized with 5′-GCUUUU-3′ and 5′-GCUUU-3′ RNA oligonucleotides that terminate with a 3′-hydroxyl group.[86] The structures reveal that the 3′-terminal uridine of these substrates is preferentially bound by the LSm3 protomer (**Fig. 6B**), which shows the best base-recognition geometry and permits a contact between the 2' hydroxyl group of the terminal ribose and the conserved arginine of LSm6. LSm3 is the last protomer in the 5′-3′ direction of the RNA where this contact is possible. If LSm6 accommodated the terminal uridine, LSm5 could not provide the corresponding contact because it

lacks the arginine.[86] Uracil base recognition geometry deteriorates with increasing distance from the 3′-end,[86] consistent again with the notion that increasing conformational constraints on the phosphoribose backbone lead to suboptimal use of the uracil binding pockets (**Fig. 6B and C**). As a consequence, substrates with terminal uridines are preferred over substrates with internal uridines, and uridine substitutions in the terminal position have the strongest effect on affinity,[86] as previously observed for Hfq.[29]

However, the present structures do not show how the physiological RNA substrate with its five terminal uridines and the 2′-3′-cyclic phosphate[93] would fit into the ring. Zhou, et al.[86] argue that LSm3 could also accommodate the cyclic phosphate if the conserved arginine from LSm6 adjusts its position. However, there is an attractive alternative where the additional uridine would enter the empty binding pocket on LSm6, placing the terminal phosphate for specific recognition into the space that is left by the missing arginine from LSm5. In this case, all five uridines of the substrate would be recognized without the need to change the path for RNA entry into the LSm ring between the LSm7 and LSm4 protomers. It should be easy to experimentally distinguish between these two alternatives.

The LSm1–7 Ring

Whereas the Sm ring and the LSm2–8 ring form specific and stable complexes with spliceosomal U snRNAs in the nucleus, the LSm1–7 ring is not involved in mRNA splicing.[6-8] Instead, it plays an important role in cytoplasmic mRNA decay, linking mRNA deadenylation to decapping.[94,95] LSm1–7 binds and protects the 3′-ends of deadenylated mRNA substrates and it initiates mRNA decapping most likely through its interaction with the structured, C-terminal domain of the decapping activator Pat1,[96,97] with which it forms a stable complex that has recently been crystallized (**Fig. 6C**).[84]

LSm1–7 appears to recognize a variety of RNA 3′-ends.[21] It participates in the 3′-oligouridinylation-dependent decay of mRNAs[98,99] that has been described for the special case of mammalian histone mRNAs,[100] for the mRNAs of certain proteins in the yeast *Saccharomyces pombe*,[101] and for many oligoadenylated mRNAs in the plant *Arabidopsis thaliana*.[102] Most importantly, however, LSm1–7 is thought to trigger the decapping of 3′-oligoadenylated mRNA decay intermediates, which result from mRNA deadenylation by the PAN2-PAN3 and CCR4-NOT complexes in the mRNA 5′-3′ decay pathway.[103-105] Such decay intermediates can be recognized directly by the LSm1–7/Pat1 complex, as suggested by in vitro analysis.[106] Furthermore, very recent data suggest that oligoadenylated mRNA decay intermediates may get oligouridinylated quite frequently also in mammalian cells,[107] providing an additional explanation for the association of the LSm1–7 ring with these RNA substrates.

Structurally, the LSm1–7 ring is highly similar to the LSm2–8 ring, with an rmsd of 0.89 Å for the entire ring (461 Cα atoms) and an rmsd of 0.97 Å for the alternative protomers LSm1 and LSm8 (54 Cα atoms).[85] Similar to the C-terminal extension of LSm8, the LSm1 C-terminal extension is positioned to obstruct the pore on the distal side of the ring, probably preventing an RNA 3′-end from exiting on this side.[84] In the case of LSm1, the C-terminal extension forms a long helix that lies across the pore, reaching out for the distal surface of the LSm6 protomer (**Fig. 6C**).[84,85] The nucleotide binding pocket of LSm1 is intact and, in contrast to LSm8, even contains an aromatic π-stacking partner in the upper position again (**Fig. 1A**). Consequently, both the LSm1–7 and the LSm2–8 ring present an arc of six putative nucleotide binding pockets that starts with LSm7 and ends with LSm6, with the notable exception of LSm5 that lacks the otherwise highly conserved arginine in the lower stacking position (R63 in *Af*Sm1, see above and **Fig. 1A**). It is therefore surprising that LSm1–7 binds octauridine (5′-UUUUUUUU-3′) with an affinity of only ~2 μM, i.e., ~100-fold less well than the LSm2–8 ring.[85]

Although at present one can only speculate about the cause of the reduced affinity, an increased difficulty to accommodate the phosphoribose backbone in the constricted space of the pore might be a possible explanation. It is possible that even minor differences in the relative orientation of the nucleotide binding pockets have important effects. More likely, however, residues inside the pore, like the bulky E107 of LSm1 (S59 in LSm8, D65 in *Af*Sm1, **Fig. 1A**) might obstruct the path of the RNA backbone, promoting it to loop out of the LSm ring and preventing the consecutive interaction of more than two to three nucleotides with their pockets. Such a scenario is feasible as illustrated by the crystal structure of the *Ec*Hfq ring with 5′-GUUUUA-3′ RNA (**Fig. 2C**).[30] This structure also provides an example of how an adenine can be accommodated in the 3′-terminal position, and it demonstrates how an upstream uridine can help to stabilize the RNA backbone of the terminal nucleotides by a hydrogen bond from its N3 position.

The described scenario therefore would not only account for the lower affinity of the LSm1–7 ring for uridine-rich RNA 3′-ends,[85] but it also suggests how the LSm1–7 ring could recognize 3′-ends on oligoadenylated mRNA.[106] In this case, the ring would simultaneously bind two or three upstream uridines, as are frequently found near the end of mRNA 3′-UTRs,[108] and one to three 3′-terminal adenines that could fit the nucleotide binding pockets reasonably well. The intermediate nucleotides would be looped out like the single uridine in the *Ec*Hfq structure (**Fig. 2C**), increasing the difficulty to bind both elements simultaneously with an increasing number of adenines in the loop. Like this, the LSm1–7 ring could sense and protect the 3′-end of the RNA molecule and distinguish oligoadenylated from polyadenylated substrates, which in vitro are known to bind less well.[106] Additional contributions that increase substrate specificity and compensate for the generally lower RNA affinity of the LSm1–7 ring probably come from Pat1 (**Fig. 6C**). Pat1 associates with the protomers LSm2 and LSm3[84,109] and also binds RNA independently.[96] Again, it should be fairly easy to test whether the reconstituted LSm1–7 ring still recognizes the 3′-terminal nucleotide on its suggested RNA substrates and what are the contributions of the Pat1 protein.

Final Remarks

The present overview demonstrates that Hfq and the ring-forming (L)Sm proteins share many common principles. These include the common protein fold and the assembly into hexameric and heptameric rings with similar RNA recognition principles on the proximal surface. Hfq and LSm rings are suited particularly well to recognize and bind RNA 3′-ends, combining direct sequence readout with the recognition of the 3′-terminal ribose and RNA backbone distorsion.

However, many rings also acquired highly specialized functions. In eukaryotes, this includes the sequence-specific Sm ring, where the RNA continues through the central pore, but also rare LSm proteins that assemble with the described ones into modified rings of novel function.[8] In archaea, even the most basic RNA binding functions of the LSm rings remain to be explored, whereas bacterial Hfq rings are probably the functionally most diverse members of the family.[12-21] In many species, Hfq has become a central player in riboregulation and has evolved additional RNA binding surfaces on the distal and lateral surfaces of the ring. It remains an important challenge to determine precisely how Hfq recognizes its sRNA partners in these species. Furthermore, the interactions of Hfq and of the other (L)Sm rings with additional, modulating protein partners like RNaseE and Pat1 is just at the beginning to be understood in molecular terms.

Clearly, there still is much to be learned on how these tiny rings express their power to rule the RNA world.

References

1. Notman DD, Kurata N, Tan EM. Profiles of antinuclear antibodies in systemic rheumatic diseases. Ann Intern Med 1975; 83:464-9; PMID:1080976; http://dx.doi.org/10.7326/0003-4819-83-4-464

2. Lerner MR, Steitz JA. Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. Proc Natl Acad Sci U S A 1979; 76:5495-9; PMID:316537; http://dx.doi.org/10.1073/pnas.76.11.5495

3. Lührmann R, Kastner B, Bach M. Structure of spliceosomal snRNPs and their role in pre-mRNA splicing. Biochim Biophys Acta 1990; 1087:265-92; PMID:2147394; http://dx.doi.org/10.1016/0167-4781(90)90001-I

4. Liautard JP, Sri-Widada J, Brunel C, Jeanteur P. Structural organization of ribonucleoproteins containing small nuclear RNAs from HeLa cells. Proteins interact closely with a similar structural domain of U1, U2, U4 and U5 small nuclear RNAs. J Mol Biol 1982; 162:623-43; PMID:6187926; http://dx.doi.org/10.1016/0022-2836(82)90392-8

5. Kambach C, Walke S, Young R, Avis JM, de la Fortelle E, Raker VA, Lührmann R, Li J, Nagai K. Crystal structures of two Sm protein complexes and their implications for the assembly of the spliceosomal snRNPs. Cell 1999; 96:375-87; PMID:10025403; http://dx.doi.org/10.1016/S0092-8674(00)80550-4

6. Albrecht M, Lengauer T. Novel Sm-like proteins with long C-terminal tails and associated methyltransferases. FEBS Lett 2004; 569:18-26; PMID:15225602; http://dx.doi.org/10.1016/j.febslet.2004.03.126

7. Anantharaman V, Aravind L. Novel conserved domains in proteins with predicted roles in eukaryotic cell-cycle regulation, decapping and RNA stability. BMC Genomics 2004; 5:45; PMID:15257761; http://dx.doi.org/10.1186/1471-2164-5-45

8. Wilusz CJ, Wilusz J. Eukaryotic Lsm proteins: lessons from bacteria. Nat Struct Mol Biol 2005; 12:1031-6; PMID:16327775; http://dx.doi.org/10.1038/nsmb1037

9. Törö I, Thore S, Mayer C, Basquin J, Séraphin B, Suck D. RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. EMBO J 2001; 20:2293-303; PMID:11331594; http://dx.doi.org/10.1093/emboj/20.9.2293

10. Tritschler F, Eulalio A, Truffault V, Hartmann MD, Helms S, Schmidt S, Coles M, Izaurralde E, Weichenrieder O. A divergent Sm fold in EDC3 proteins mediates DCP1 binding and P-body targeting. Mol Cell Biol 2007; 27:8600-11; PMID:17923697; http://dx.doi.org/10.1128/MCB.01506-07

11. Franze de Fernandez MT, Eoyang L, August JT. Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. Nature 1968; 219:588-90; PMID:4874917; http://dx.doi.org/10.1038/219588a0

12. Bandyra KJ, Luisi BF. Licensing and due process in the turnover of bacterial RNA. RNA Biol 2013; 10:627-35; PMID:23580162; http://dx.doi.org/10.4161/rna.24393

13. De Lay N, Schu DJ, Gottesman S. Bacterial small RNA-based negative regulation: Hfq and its accomplices. J Biol Chem 2013; 288:7996-8003; PMID:23362267; http://dx.doi.org/10.1074/jbc.R112.441386

14. Gottesman S, Storz G. Bacterial small RNA regulators: versatile roles and rapidly evolving variations. Cold Spring Harb Perspect Biol 2011; 3:3; PMID:20980440; http://dx.doi.org/10.1101/cshperspect.a003798

15. Mura C, Randolph PS, Patterson J, Cozen AE. Archaeal and eukaryotic homologs of Hfq: A structural and evolutionary perspective on Sm function. RNA Biol 2013; 10:636-51; PMID:23579284; http://dx.doi.org/10.4161/rna.24538

16. Régnier P, Hajnsdorf E. The interplay of Hfq, poly(A) polymerase I and exoribonucleases at the 3′ ends of RNAs resulting from Rho-independent termination: A tentative model. RNA Biol 2013; 10:602-9; PMID:23392248; http://dx.doi.org/10.4161/rna.23664

17. Sauer E. Special focus Hfq. RNA Biol 2013; 10:590-1; PMID:23774896; http://dx.doi.org/10.4161/rna.24617

18. Sauer E. Structure and RNA-binding properties of the bacterial LSm protein Hfq. RNA Biol 2013; 10:610-8; PMID:23535768; http://dx.doi.org/10.4161/rna.24201

19. Vogel J, Luisi BF. Hfq and its constellation of RNA. Nat Rev Microbiol 2011; 9:578-89; PMID:21760622; http://dx.doi.org/10.1038/nrmicro2615

20. Wagner EG. Cycling of RNAs on Hfq. RNA Biol 2013; 10:619-26; PMID:23466677; http://dx.doi.org/10.4161/rna.24044

21. Wilusz CJ, Wilusz J. Lsm proteins and Hfq: Life at the 3′ end. RNA Biol 2013; 10:592-601; PMID:23392247; http://dx.doi.org/10.4161/rna.23695

22. Møller T, Franch T, Højrup P, Keene DR, Bächinger HP, Brennan RG, Valentin-Hansen P. Hfq: a bacterial Sm-like protein that mediates RNA-RNA interaction. Mol Cell 2002; 9:23-30; PMID:11804583; http://dx.doi.org/10.1016/S1097-2765(01)00436-1

23. Zhang A, Wassarman KM, Ortega J, Steven AC, Storz G. The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. Mol Cell 2002; 9:11-22; PMID:11804582; http://dx.doi.org/10.1016/S1097-2765(01)00437-3

24. Hermann H, Fabrizio P, Raker VA, Foulaki K, Hornig H, Brahms H, Lührmann R. snRNP Sm proteins share two evolutionarily conserved sequence motifs which are involved in Sm protein-protein interactions. EMBO J 1995; 14:2076-88; PMID:7744013

25. Sauter C, Basquin J, Suck D. Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from Escherichia coli. Nucleic Acids Res 2003; 31:4091-8; PMID:12853626; http://dx.doi.org/10.1093/nar/gkg480

26. Schumacher MA, Pearson RF, Møller T, Valentin-Hansen P, Brennan RG. Structures of the pleiotropic translational regulator Hfq and an Hfq-RNA complex: a bacterial Sm-like protein. EMBO J 2002; 21:3546-56; PMID:12093755; http://dx.doi.org/10.1093/emboj/cdf322

27. Murina VN, Nikulin AD. RNA-binding Sm-like proteins of bacteria and archaea. similarity and difference in structure and function. Biochemistry (Mosc) 2011; 76:1434-49; PMID:22339597; http://dx.doi.org/10.1134/S0006297911130050

28. Murina V, Lekontseva N, Nikulin A. Hfq binds ribonucleotides in three different RNA-binding sites. Acta Crystallogr D Biol Crystallogr 2013; 69:1504-13; PMID:23897473; http://dx.doi.org/10.1107/S090744491301010X

29. Sauer E, Weichenrieder O. Structural basis for RNA 3′-end recognition by Hfq. Proc Natl Acad Sci U S A 2011; 108:13065-70; PMID:21737752; http://dx.doi.org/10.1073/pnas.1103420108

30. Wang W, Wang L, Zou Y, Zhang J, Gong Q, Wu J, Shi Y. Cooperation of Escherichia coli Hfq hexamers in DsrA binding. Genes Dev 2011; 25:2106-17; PMID:21979921; http://dx.doi.org/10.1101/gad.16746011

31. Otaka H, Ishikawa H, Morita T, Aiba H. PolyU tail of rho-independent terminator of bacterial small RNAs is essential for Hfq action. Proc Natl Acad Sci U S A 2011; 108:13059-64; PMID:21788484; http://dx.doi.org/10.1073/pnas.1107050108

32. Wilson KS, von Hippel PH. Transcription termination at intrinsic terminators: the role of the RNA hairpin. Proc Natl Acad Sci U S A 1995; 92:8793-7; PMID:7568019; http://dx.doi.org/10.1073/pnas.92.19.8793

33. de Haseth PL, Uhlenbeck OC. Interaction of Escherichia coli host factor protein with oligoriboadenylates. Biochemistry 1980; 19:6138-46; PMID:6162476; http://dx.doi.org/10.1021/bi00567a029

34. Folichon M, Arluison V, Pellegrini O, Huntzinger E, Régnier P, Hajnsdorf E. The poly(A) binding protein Hfq protects RNA from RNase E and exoribonucleolytic degradation. Nucleic Acids Res 2003; 31:7302-10; PMID:14654705; http://dx.doi.org/10.1093/nar/gkg915

35. Le Derout J, Folichon M, Briani F, Dehò G, Régnier P, Hajnsdorf E. Hfq affects the length and the frequency of short oligo(A) tails at the 3′ end of Escherichia coli rpsO mRNAs. Nucleic Acids Res 2003; 31:4017-23; PMID:12853618; http://dx.doi.org/10.1093/nar/gkg456

36. Mikulecky PJ, Kaw MK, Brescia CC, Takach JC, Sledjeski DD, Feig AL. Escherichia coli Hfq has distinct interaction surfaces for DsrA, rpoS and poly(A) RNAs. Nat Struct Mol Biol 2004; 11:1206-14; PMID:15531892; http://dx.doi.org/10.1038/nsmb858

37. Horstmann N, Orans J, Valentin-Hansen P, Shelburne SA 3rd, Brennan RG. Structural mechanism of Staphylococcus aureus Hfq binding to an RNA A-tract. Nucleic Acids Res 2012; 40:11023-35; PMID:22965117; http://dx.doi.org/10.1093/nar/gks809

38. Link TM, Valentin-Hansen P, Brennan RG. Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. Proc Natl Acad Sci U S A 2009; 106:19292-7; PMID:19889981; http://dx.doi.org/10.1073/pnas.0908744106

39. Someya T, Baba S, Fujimoto M, Kawai G, Kumasaka T, Nakamura K. Crystal structure of Hfq from Bacillus subtilis in complex with SELEX-derived RNA aptamer: insight into RNA-binding properties of bacterial Hfq. Nucleic Acids Res 2012; 40:1856-67; PMID:22053080; http://dx.doi.org/10.1093/nar/gkr892

40. Hämmerle H, Beich-Frandsen M, Večerek B, Rajkowitsch L, Carugo O, Djinović-Carugo K, Bläsi U. Structural and biochemical studies on ATP binding and hydrolysis by the Escherichia coli RNA chaperone Hfq. PLoS One 2012; 7:e50892; PMID:23226421; http://dx.doi.org/10.1371/journal.pone.0050892

41. Robinson KE, Orans J, Kovach AR, Link TM, Brennan RG. Mapping Hfq-RNA interaction surfaces using tryptophan fluorescence quenching. Nucleic Acids Res 2014; 42:2736-49; PMID:24288369

42. Sauer E, Schmidt S, Weichenrieder O. Small RNA binding to the lateral surface of Hfq hexamers and structural rearrangements upon mRNA target recognition. Proc Natl Acad Sci U S A 2012; 109:9396-401; PMID:22645344; http://dx.doi.org/10.1073/pnas.1202521109

43. Balbontín R, Fiorini F, Figueroa-Bossi N, Casadesús J, Bossi L. Recognition of heptameric seed sequence underlies multi-target regulation by RybB small RNA in Salmonella enterica. Mol Microbiol 2010; 78:380-94; PMID:20979336; http://dx.doi.org/10.1111/j.1365-2958.2010.07342.x

44. Papenfort K, Bouvier M, Mika F, Sharma CM, Vogel J. Evidence for an autonomous 5′ target recognition domain in an Hfq-associated small RNA. Proc Natl Acad Sci U S A 2010; 107:20435-40; PMID:21059903; http://dx.doi.org/10.1073/pnas.1009784107

45. Ishikawa H, Otaka H, Maki K, Morita T, Aiba H. The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3′ poly(U) tail. RNA 2012; 18:1062-74; PMID:22454537; http://dx.doi.org/10.1261/rna.031575.111

46. Olejniczak M. Despite similar binding to the Hfq protein regulatory RNAs widely differ in their competition performance. Biochemistry 2011; 50:4427-40; PMID:21510661; http://dx.doi.org/10.1021/bi102043f

47. Sittka A, Lucchini S, Papenfort K, Sharma CM, Rolle K, Binnewies TT, Hinton JC, Vogel J. Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. PLoS Genet 2008; 4:e1000163; PMID:18725932; http://dx.doi.org/10.1371/journal.pgen.1000163

48. Sittka A, Sharma CM, Rolle K, Vogel J. Deep sequencing of Salmonella RNA associated with heterologous Hfq proteins in vivo reveals small RNAs as a major target class and identifies RNA processing phenotypes. RNA Biol 2009; 6:266-75; PMID:19333007; http://dx.doi.org/10.4161/rna.6.3.8332

49. Zhang A, Wassarman KM, Rosenow C, Tjaden BC, Storz G, Gottesman S. Global analysis of small RNA and mRNA targets of Hfq. Mol Microbiol 2003; 50:1111-24; PMID:14622403; http://dx.doi.org/10.1046/j.1365-2958.2003.03734.x

50. Andrade JM, Pobre V, Matos AM, Arraiano CM. The crucial role of PNPase in the degradation of small RNAs that are not associated with Hfq. RNA 2012; 18:844-55; PMID:22355164; http://dx.doi.org/10.1261/rna.029413.111

51. Moll I, Afonyushkin T, Vytvytska O, Kaberdin VR, Bläsi U. Coincident Hfq binding and RNase E cleavage sites on mRNA and small regulatory RNAs. RNA 2003; 9:1308-14; PMID:14561880; http://dx.doi.org/10.1261/rna.5850703

52. Urban JH, Vogel J. Translational control and target recognition by Escherichia coli small RNAs in vivo. Nucleic Acids Res 2007; 35:1018-37; PMID:17264113; http://dx.doi.org/10.1093/nar/gkl1040

53. Thore S, Mayer C, Sauter C, Weeks S, Suck D. 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 2003; 278:1239-47; PMID:12409299; http://dx.doi.org/10.1074/jbc.M207685200

54. Panja S, Schu DJ, Woodson SA. Conserved arginines on the rim of Hfq catalyze base pair formation and exchange. Nucleic Acids Res 2013; 41:7536-46; PMID:23771143; http://dx.doi.org/10.1093/nar/gkt521

55. de Haseth PL, Uhlenbeck OC. Interaction of Escherichia coli host factor protein with Q β ribonucleic acid. Biochemistry 1980; 19:6146-51; PMID:6162477; http://dx.doi.org/10.1021/bi00567a030

56. Folichon M, Allemand F, Régnier P, Hajnsdorf E. Stimulation of poly(A) synthesis by Escherichia coli poly(A)polymerase I is correlated with Hfq binding to poly(A) tails. FEBS J 2005; 272:454-63; PMID:15654883; http://dx.doi.org/10.1111/j.1742-4658.2004.04485.x

57. Mohanty BK, Maples VF, Kushner SR. The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. Mol Microbiol 2004; 54:905-20; PMID:15522076; http://dx.doi.org/10.1111/j.1365-2958.2004.04337.x

58. De Lay N, Gottesman S. Role of polynucleotide phosphorylase in sRNA function in Escherichia coli. RNA 2011; 17:1172-89; PMID:21527671; http://dx.doi.org/10.1261/rna.2531211

59. Ikeda Y, Yagi M, Morita T, Aiba H. Hfq binding at RhlB-recognition region of RNase E is crucial for the rapid degradation of target mRNAs mediated by sRNAs in Escherichia coli. Mol Microbiol 2011; 79:419-32; PMID:21219461; http://dx.doi.org/10.1111/j.1365-2958.2010.07454.x

60. Morita T, Maki K, Aiba H. RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. Genes Dev 2005; 19:2176-86; PMID:16166379; http://dx.doi.org/10.1101/gad.1330405

61. Prévost K, Desnoyers G, Jacques JF, Lavoie F, Massé E. Small RNA-induced mRNA degradation achieved through both translation block and activated cleavage. Genes Dev 2011; 25:385-96; PMID:21289064; http://dx.doi.org/10.1101/gad.2001711

62. Fender A, Elf J, Hampel K, Zimmermann B, Wagner EG. RNAs actively cycle on the Sm-like protein Hfq. Genes Dev 2010; 24:2621-6; PMID:21123649; http://dx.doi.org/10.1101/gad.591310

63. Hussein R, Lim HN. Disruption of small RNA signaling caused by competition for Hfq. Proc Natl Acad Sci U S A 2011; 108:1110-5; PMID:21189298; http://dx.doi.org/10.1073/pnas.1010082108

64. Moon K, Gottesman S. Competition among Hfq-binding small RNAs in Escherichia coli. Mol Microbiol 2011; 82:1545-62; PMID:22040174; http://dx.doi.org/10.1111/j.1365-2958.2011.07907.x

65. Geissmann TA, Touati D. Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. EMBO J 2004; 23:396-405; PMID:14739933; http://dx.doi.org/10.1038/sj.emboj.7600058

66. Moll I, Leitsch D, Steinhauser T, Bläsi U. RNA chaperone activity of the Sm-like Hfq protein. EMBO Rep 2003; 4:284-9; PMID:12634847; http://dx.doi.org/10.1038/sj.embor.embor772

67. Rajkowitsch L, Chen D, Stampfl S, Semrad K, Waldsich C, Mayer O, Jantsch MF, Konrat R, Bläsi U, Schroeder R. RNA chaperones, RNA annealers and RNA helicases. RNA Biol 2007; 4:118-30; PMID:18347437; http://dx.doi.org/10.4161/rna.4.3.5445

68. Rajkowitsch L, Schroeder R. Dissecting RNA chaperone activity. RNA 2007; 13:2053-60; PMID:17901153; http://dx.doi.org/10.1261/rna.671807

69. Peng Y, Soper TJ, Woodson SA. Positional effects of AAN motifs in rpoS regulation by sRNAs and Hfq. J Mol Biol 2014; 426:275-85; PMID:24051417; http://dx.doi.org/10.1016/j.jmb.2013.08.026

70. Bandyra KJ, Said N, Pfeiffer V, Górna MW, Vogel J, Luisi BF. The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. Mol Cell 2012; 47:943-53; PMID:22902561; http://dx.doi.org/10.1016/j.molcel.2012.07.015

71. Lorenz C, Gesell T, Zimmermann B, Schoeberl U, Bilusic I, Rajkowitsch L, Waldsich C, von Haeseler A, Schroeder R. Genomic SELEX for Hfq-binding RNAs identifies genomic aptamers predominantly in antisense transcripts. Nucleic Acids Res 2010; 38:3794-808; PMID:20348540; http://dx.doi.org/10.1093/nar/gkq032

72. Zhang A, Schu DJ, Tjaden BC, Storz G, Gottesman S. Mutations in interaction surfaces differentially impact E. coli Hfq association with small RNAs and their mRNA targets. J Mol Biol 2013; 425:3678-97; PMID:23318956; http://dx.doi.org/10.1016/j.jmb.2013.01.006

73. Henderson CA, Vincent HA, Casamento A, Stone CM, Phillips JO, Cary PD, Sobott F, Gowers DM, Taylor JE, Callaghan AJ. Hfq binding changes the structure of Escherichia coli small noncoding RNAs OxyS and RprA, which are involved in the riboregulation of rpoS. RNA 2013; 19:1089-104; PMID:23804244; http://dx.doi.org/10.1261/rna.034595.112

74. Vincent HA, Henderson CA, Stone CM, Cary PD, Gowers DM, Sobott F, Taylor JE, Callaghan AJ. The low-resolution solution structure of Vibrio cholerae Hfq in complex with Qrr1 sRNA. Nucleic Acids Res 2012; 40:8698-710; PMID:22730296; http://dx.doi.org/10.1093/nar/gks582

75. Beich-Frandsen M, Večerek B, Konarev PV, Sjöblom B, Kloiber K, Hämmerle H, Rajkowitsch L, Miles AJ, Kontaxis G, Wallace BA, et al. Structural insights into the dynamics and function of the C-terminus of the E. coli RNA chaperone Hfq. Nucleic Acids Res 2011; 39:4900-15; PMID:21330354; http://dx.doi.org/10.1093/nar/gkq1346

76. Olsen AS, Møller-Jensen J, Brennan RG, Valentin-Hansen P. C-terminally truncated derivatives of Escherichia coli Hfq are proficient in riboregulation. J Mol Biol 2010; 404:173-82; PMID:20888338; http://dx.doi.org/10.1016/j.jmb.2010.09.038

77. Jousselin A, Metzinger L, Felden B. On the facultative requirement of the bacterial RNA chaperone, Hfq. Trends Microbiol 2009; 17:399-405; PMID:19733080; http://dx.doi.org/10.1016/j.tim.2009.06.003

78. Tomasini A, François P, Howden BP, Fechter P, Romby P, Caldelari I. The importance of regulatory RNAs in Staphylococcus aureus. Infect Genet Evol 2014; 21:616-26; PMID:24291227; http://dx.doi.org/10.1016/j.meegid.2013.11.016

79. Bøggild A, Overgaard M, Valentin-Hansen P, Brodersen DE. Cyanobacteria contain a structural homologue of the Hfq protein with altered RNA-binding properties. FEBS J 2009; 276:3904-15; PMID:19777643; http://dx.doi.org/10.1111/j.1742-4658.2009.07104.x

80. Chao Y, Vogel J. The role of Hfq in bacterial pathogens. Curr Opin Microbiol 2010; 13:24-33; PMID:20080057; http://dx.doi.org/10.1016/j.mib.2010.01.001

81. Sharma CM, Hoffmann S, Darfeuille F, Reignier J, Findeiss S, Sittka A, Chabas S, Reiche K, Hackermüller J, Reinhardt R, et al. The primary transcriptome of the major human pathogen Helicobacter pylori. Nature 2010; 464:250-5; PMID:20164839; http://dx.doi.org/10.1038/nature08756

82. Tsui HC, Leung HC, Winkler ME. Characterization of broadly pleiotropic phenotypes caused by an hfq insertion mutation in Escherichia coli K-12. Mol Microbiol 1994; 13:35-49; PMID:7984093; http://dx.doi.org/10.1111/j.1365-2958.1994.tb00400.x

83. Göpel Y, Papenfort K, Reichenbach B, Vogel J, Görke B. Targeted decay of a regulatory small RNA by an adaptor protein for RNase E and counteraction by an anti-adaptor RNA. Genes Dev 2013; 27:552-64; PMID:23475961; http://dx.doi.org/10.1101/gad.210112.112

84. Sharif H, Conti E. Architecture of the Lsm1-7-Pat1 complex: a conserved assembly in eukaryotic mRNA turnover. Cell Rep 2013; 5:283-91; PMID:24139796; http://dx.doi.org/10.1016/j.celrep.2013.10.004

85. Zhou L, Zhou Y, Hang J, Wan R, Lu G, Yan C, Shi Y. Crystal structure and biochemical analysis of the heptameric Lsm1-7 complex. Cell Res 2014; 24:497-500; PMID:24513854; http://dx.doi.org/10.1038/cr.2014.18

86. Zhou L, Hang J, Zhou Y, Wan R, Lu G, Yin P, Yan C, Shi Y. Crystal structures of the Lsm complex bound to the 3′ end sequence of U6 small nuclear RNA. Nature 2014; 506:116-20; PMID:24240276; http://dx.doi.org/10.1038/nature12803

87. Leung AK, Nagai K, Li J. Structure of the spliceosomal U4 snRNP core domain and its implication for snRNP biogenesis. Nature 2011; 473:536-9; PMID:21516107; http://dx.doi.org/10.1038/nature09956

88. Pomeranz Krummel DA, Oubridge C, Leung AK, Li J, Nagai K. Crystal structure of human spliceosomal U1 snRNP at 5.5 A resolution. Nature 2009; 458:475-80; PMID:19325628; http://dx.doi.org/10.1038/nature07851

89. Weber G, Trowitzsch S, Kastner B, Lührmann R, Wahl MC. Functional organization of the Sm core in the crystal structure of human U1 snRNP. EMBO J 2010; 29:4172-84; PMID:21113136; http://dx.doi.org/10.1038/emboj.2010.295

90. Meister G, Eggert C, Fischer U. SMN-mediated assembly of RNPs: a complex story. Trends Cell Biol 2002; 12:472-8; PMID:12441251; http://dx.doi.org/10.1016/S0962-8924(02)02371-1

91. Pellizzoni L, Yong J, Dreyfuss G. Essential role for the SMN complex in the specificity of snRNP assembly. Science 2002; 298:1775-9; PMID:12459587; http://dx.doi.org/10.1126/science.1074962

92. Zaric B, Chami M, Rémigy H, Engel A, Ballmer-Hofer K, Winkler FK, Kambach C. Reconstitution of two recombinant LSm protein complexes reveals aspects of their architecture, assembly, and function. J Biol Chem 2005; 280:16066-75; PMID:15711010; http://dx.doi.org/10.1074/jbc.M414481200

93. Licht K, Medenbach J, Lührmann R, Kambach C, Bindereif A. 3′-cyclic phosphorylation of U6 snRNA leads to recruitment of recycling factor p110 through LSm proteins. RNA 2008; 14:1532-8; PMID:18567812; http://dx.doi.org/10.1261/rna.1129608

94. Tharun S, He W, Mayes AE, Lennertz P, Beggs JD, Parker R. Yeast Sm-like proteins function in mRNA decapping and decay. Nature 2000; 404:515-8; PMID:10761922; http://dx.doi.org/10.1038/35006676

95. Tharun S, Parker R. Targeting an mRNA for decapping: displacement of translation factors and association of the Lsm1p-7p complex on deadenylated yeast mRNAs. Mol Cell 2001; 8:1075-83; PMID:11741542; http://dx.doi.org/10.1016/S1097-2765(01)00395-1

96. Braun JE, Tritschler F, Haas G, Igreja C, Truffault V, Weichenrieder O, Izaurralde E. The C-terminal α-α superhelix of Pat is required for mRNA decapping in metazoa. EMBO J 2010; 29:2368-80; PMID:20543818; http://dx.doi.org/10.1038/emboj.2010.124

97. He W, Parker R. The yeast cytoplasmic LsmI/Pat1p complex protects mRNA 3′ termini from partial degradation. Genetics 2001; 158:1445-55; PMID:11514438

98. Song MG, Kiledjian M. 3′ Terminal oligo U-tract-mediated stimulation of decapping. RNA 2007; 13:2356-65; PMID:17942740; http://dx.doi.org/10.1261/rna.765807

99. Wilusz CJ, Wilusz J. New ways to meet your (3′) end oligouridylation as a step on the path to destruction. Genes Dev 2008; 22:1-7; PMID:18172159; http://dx.doi.org/10.1101/gad.1634508

100. Mullen TE, Marzluff WF. Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5′ to 3′ and 3′ to 5′. Genes Dev 2008; 22:50-65; PMID:18172165; http://dx.doi.org/10.1101/gad.1622708

101. Rissland OS, Norbury CJ. Decapping is preceded by 3′ uridylation in a novel pathway of bulk mRNA turnover. Nat Struct Mol Biol 2009; 16:616-23; PMID:19430462; http://dx.doi.org/10.1038/nsmb.1601

102. Sement FM, Ferrier E, Zuber H, Merret R, Alioua M, Deragon JM, Bousquet-Antonelli C, Lange H, Gagliardi D. Uridylation prevents 3′ trimming of oligoadenylated mRNAs. Nucleic Acids Res 2013; 41:7115-27; PMID:23748567; http://dx.doi.org/10.1093/nar/gkt465

103. Jonas S, Izaurralde E. The role of disordered protein regions in the assembly of decapping complexes and RNP granules. Genes Dev 2013; 27:2628-41; PMID:24352420; http://dx.doi.org/10.1101/gad.227843.113

104. Tharun S. Lsm1-7-Pat1 complex: a link between 3′ and 5′-ends in mRNA decay? RNA Biol 2009; 6:228-32; PMID:19279404; http://dx.doi.org/10.4161/rna.6.3.8282

105. Wahle E, Winkler GS. RNA decay machines: deadenylation by the Ccr4-not and Pan2-Pan3 complexes. Biochim Biophys Acta 2013; 1829:561-70; PMID:23337855; http://dx.doi.org/10.1016/j.bbagrm.2013.01.003

106. Chowdhury A, Mukhopadhyay J, Tharun S. The decapping activator Lsm1p-7p-Pat1p complex has the intrinsic ability to distinguish between oligoadenylated and polyadenylated RNAs. RNA 2007; 13:998-1016; PMID:17513695; http://dx.doi.org/10.1261/rna.502507

107. Chang H, Lim J, Ha M, Kim VN. TAIL-seq: genome-wide determination of poly(A) tail length and 3′ end modifications. Mol Cell 2014; 53:1044-52; PMID:24582499; http://dx.doi.org/10.1016/j.molcel.2014.02.007

108. Tian B, Graber JH. Signals for pre-mRNA cleavage and polyadenylation. Wiley Interdiscip Rev RNA 2012; 3:385-96; PMID:22012871; http://dx.doi.org/10.1002/wrna.116

109. Wu D, Muhlrad D, Bowler MW, Jiang S, Liu Z, Parker R, Song H. Lsm2 and Lsm3 bridge the interaction of the Lsm1-7 complex with Pat1 for decapping activation. Cell Res 2014; 24:233-46; PMID:24247251; http://dx.doi.org/10.1038/cr.2013.152

110. Gouet P, Robert X, Courcelle E. ESPript/ENDscript: Extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Res 2003; 31:3320-3; PMID:12824317; http://dx.doi.org/10.1093/nar/gkg556

111. Emsley P, Lohkamp B, Scott WG, Cowtan K. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 2010; 66:486-501; PMID:20383002; http://dx.doi.org/10.1107/S0907444910007493