Small RNAs in *Escherichia coli*

Karen Montzka Wassarman, Aixia Zhang and Gisela Storz

Bacterial cells contain several small RNAs (sRNAs) that are not translated. These stable, abundant RNAs act by multiple mechanisms, such as RNA–RNA basepairing, RNA–protein interactions and intrinsic RNA activity, and regulate diverse cellular functions, including RNA processing, mRNA stability, translation, protein stability and secretion.

K. Montzka Wassarman, A. Zhang and G. Storz are in the Cell Biology and Metabolism Branch, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, MD 20892-5430, USA.*
**tel: +1 301 402 0968,*
fax: +1 301 402 0078,
e-mail: storz@helix.nih.gov

Ten small RNAs (sRNAs), in addition to 5S rRNA and tRNAs, are currently known to be encoded by the *Escherichia coli* genome. The majority of these sRNAs can be defined as regulatory RNAs, as all of them modulate an aspect of bacterial metabolism. Some tRNAs and sRNAs encoded by plasmids, phage and transposable elements have also been shown to have regulatory functions. These RNAs, in addition to *Staphylococcus aureus* RNAIII and the multitude of sRNAs found in eukaryotes, are the subject of other reviews[1–5,48]. Here, we focus on the ten *E. coli* sRNAs; we summarize their general properties, discuss their mechanisms of action, compare them with eukaryotic sRNAs and outline topics that need to be addressed by future studies.

Discovery of sRNAs

Evidence of sRNAs, other than 5S rRNA and tRNAs, came with the development of polyacrylamide gel electrophoresis to analyze [^{32}P]orthophosphate-labeled total RNA. The 4.5S and 6S RNAs were the first to be identified[6,7]. Improved techniques, such as two-dimensional gel electrophoresis, led to the discoveries of Spot 42 RNA (Band IV) and 10S RNA (Band IX)[8]. Characterization of the 10S RNA band subsequently revealed that it corresponded to two distinct RNAs of similar size, 10Sa and 10Sb (Ref. 9; also called M2 and M1 and now denoted tmRNA and RNase P RNA, respectively). Analyses of these five sRNAs, together with some serendipity, led to the identification of functions for the 4.5S, tmRNA and RNase P RNAs, but the roles of 6S and Spot 42 RNAs are still unknown. A second group of sRNAs, including CsrB and OxyS, was discovered under conditions that suggested possible functions, as described below. A third group of sRNAs, including MicF, DicF and DsrA, was identified as cloned genomic fragments that modulated certain activities. In the following section, we describe what is known about these ten sRNAs in *E. coli*, although homologs of all of the sRNAs have been detected in other bacterial species. Many of their properties and functions are summarized in Tables 1,2, and proposed secondary structures are shown in Fig. 1.

4.5S RNA

The 4.5S RNA has been characterized extensively and found to function in protein secretion as an integral component of the signal recognition particle (SRP) (reviewed in Refs 10,11). The 114-nt mature 4.5S RNA is processed from a precursor molecule at the 5′ end by RNase P. The gene encoding 4.5S [*ffs* (four-point five S)] is essential. The earliest defect in cells conditionally depleted of 4.5S is an inhibition of protein synthesis, and 4.5S RNA is normally associated with ribosomes. These results suggest a role for 4.5S RNA in translation. However, proposed secondary structures have revealed similarities between 4.5S RNA and 7SL, the RNA component of eukaryotic SRP. Support for the involvement of 4.5S RNA in protein secretion came from the finding that p48 (a bacterial protein with similarity to the eukaryotic SRP54 protein) binds 4.5S RNA. In addition, the p48/4.5S ribonucleoprotein (RNP) particle can bind signal sequences, and genetic studies have revealed that these highly conserved mammalian and bacterial counterparts can partially substitute for one another. The findings that 4.5S RNA affects both protein synthesis and protein secretion have not been resolved, but 4.5S might function in two different pathways. The exact role of 4.5S in the SRP is also unclear, although recent results suggest that 4.5S RNA might stabilize an active conformation of the p48 protein[12].

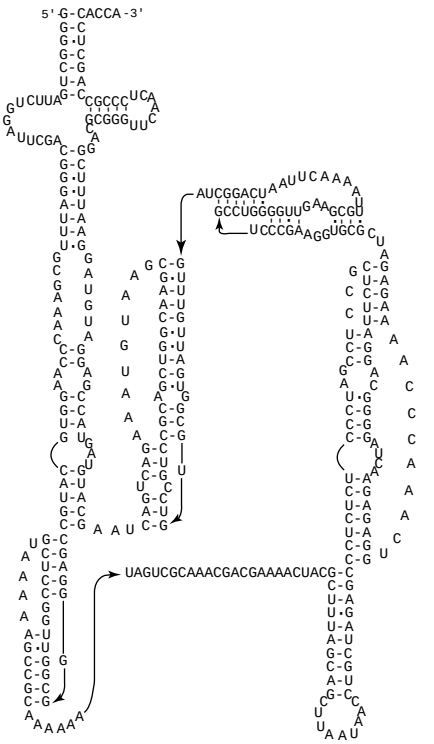
tmRNA

The 363-nt tmRNA is processed from a primary transcript at the 5′ end by RNase P and at the 3′ end by RNase III (reviewed in Refs 13,14). A proposed secondary structure, supported by chemical and enzymatic structure probing and phylogenetic analysis, revealed that tmRNA can form a tRNA half molecule. The presence of a tRNA-like structure is further supported by RNase P cleavage of pre-tmRNA (similar to pre-tRNA) and aminoacylation of tmRNA by alanyl-tRNA synthetase *in vitro*. The tmRNA contains a short open reading frame (ORF), the significance of which was not originally apparent. Strains carrying mutations in the gene encoding tmRNA [*ssrA* (small stable RNA)] show a confusing array of subtle phenotypes, including slightly reduced growth, increased Alp protease activity, and failure to support

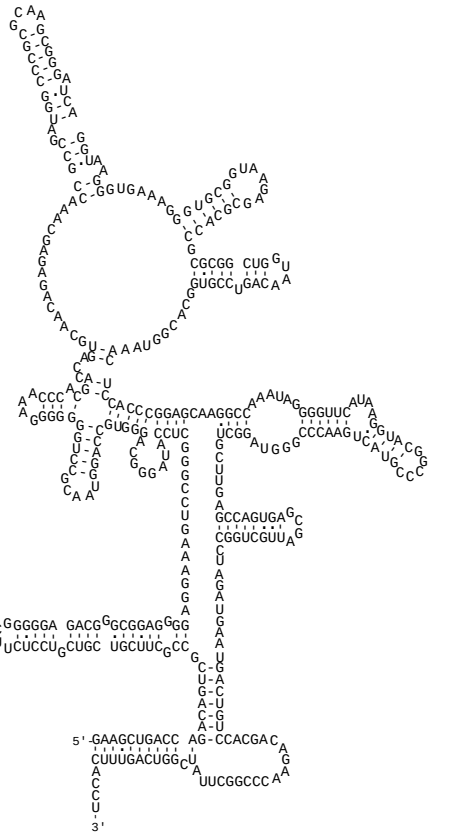
Published by Elsevier Science Ltd. PII: S0966-842X(98)01379-1

Trends in Microbiology

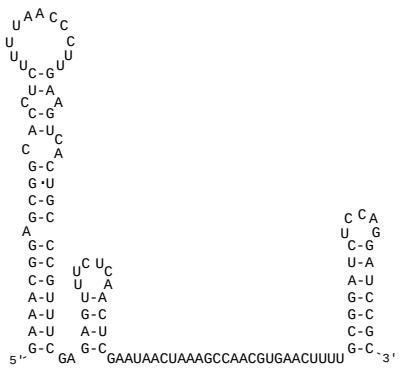
(a) tmRNA



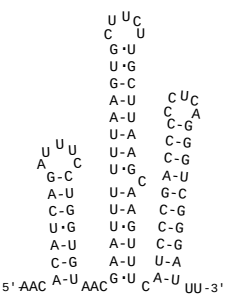
(b) RNase P RNA



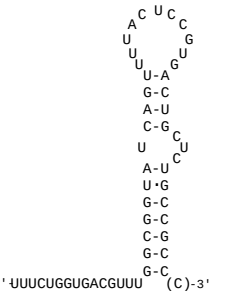
(d) OxyS RNA



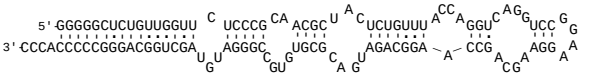
(e) DsrA RNA



(f) DicF RNA



(i) 4.5S RNA



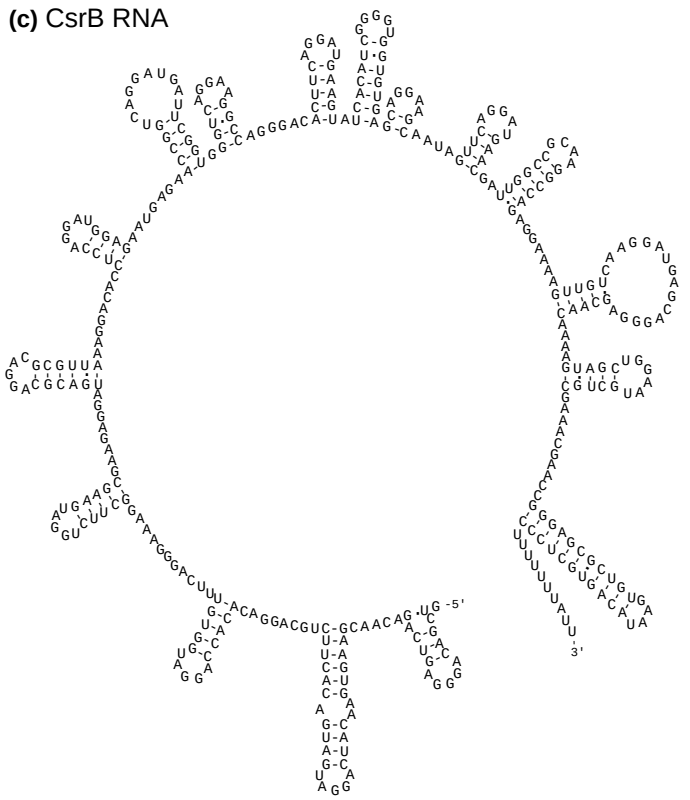
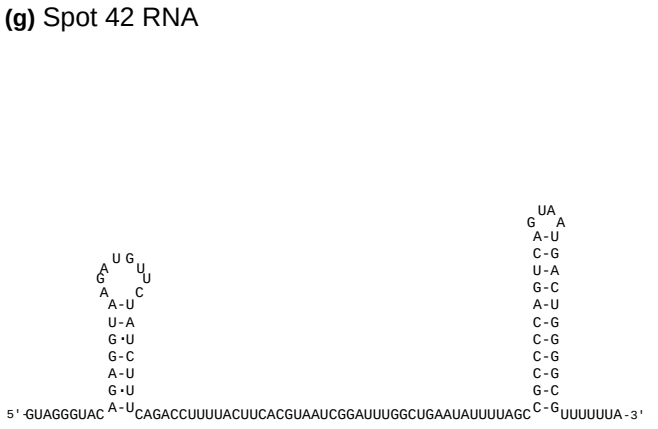
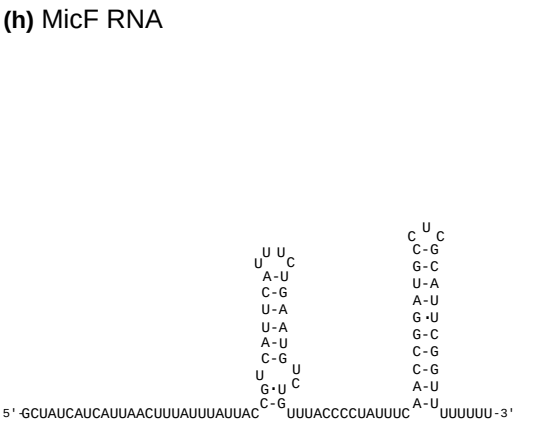


Fig. 1. Proposed structures of the ten *Escherichia coli* sRNAs described in this review. Structures of **(a)** tmRNA and **(b)** RNase P RNA are based on biochemical evidence and extensive phylogenetic studies[14,16]. **(c)** The CsrB RNA structure is based on computer prediction and is supported by the sequence of a homolog from *Erwinia carotovora* in which the repeats are conserved in number and structure[27]. **(d)** The OxyS RNA structure is supported by direct structural probing of the 5′ end[28] and by comparison with the *Salmonella typhimurium oxyS* sequence (G. Storz, unpublished). The predicted structure of **(e)** DsrA RNA is supported by comparisons with *Klebsiella pneumoniae*[37] and *S. typhimurium* homologs[40]. **(f)** The stem–loop structure of DicF RNA is supported by the sequence of a homolog in *Shigella*[42]. **(g)** The Spot 42 RNA structure is based on computer predictions (J. Dahlberg, unpublished), and the RNA is detected on northern blots of related species (K.M. Wassarman and G. Storz, unpublished). **(h)** The MicF RNA structure is based on biochemical probing and phylogenetic comparisons[46]. **(i)** The 4.5S RNA structure is similar to the eukaryotic 7SL RNA (Ref. 10), and **(j)** the 6S RNA structure is based on sequence comparison with a *Pseudomonas aeruginosa* homolog[47].





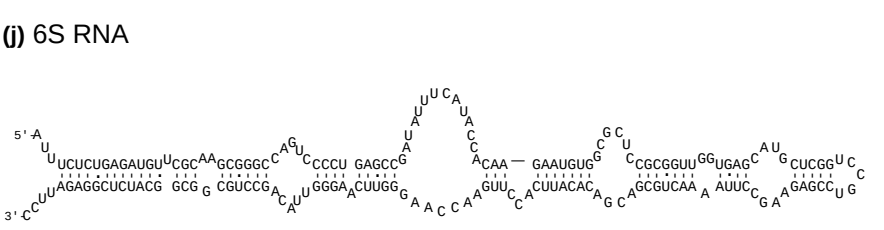


Table 1. Properties of sRNAs[a]

RNA[b]	Other names	Gene	Map location[c]	Length[d] (nt)	Abundance[e]	RNA biogenesis[f]	Interactions	Refs
4.5S		*ffs*	10.2′	114	0.13–0.28	Processed from primary transcript RNase P (5′ end)	Associates with 70S ribosomes	10,11, 25
tmRNA	10Sa M2	*ssrA*	59.3′	363	0.10–0.26	Processed from primary transcript RNase P (5′ end) RNase III (3′ end)	Associates with 70S ribosomes	13,14, 28
RNase P	10Sb M1	*rnpB*	70.4′	377	ND[f]	Processed from primary transcript RNase E (3′ end)	RNP with RNase P protein	15–17
Spot 42		*spf*	87.2′	109	0.01	Primary transcript	Associates with nucleoid and ribosomes	18,19
6S		*ssrS*	65.8′	184	0.2–0.3	Processed from dicistronic transcript by an unknown mechanism	11S RNP	23–25
CsrB		*csrB*	63′	369	ND	Unknown	RNP with CsrA protein	27
OxyS		*oxyS*	89.6′	109	0.09	Primary transcript	Binds Hfq protein	28,30
MicF		*micF*	49.8′	93	ND	Primary transcript	Crosslinks to 80-kDa protein	3,32
DicF		*dicF*	35.5′	53	ND	Processed from polycistronic transcript RNase E (5′ end) RNase III (3′ end)		32,35
DsrA		*dsrA*	43.6′	87	ND	Primary transcript		37

[a]Abbreviations: RNP, ribonucleoprotein; ND, not determined.
[b]See also: www.wi.mit.edu/bartel/tmRNA/home; psyche.uthct.edu/dbs/tmRDB/tmRDB.html; www.mbio.ncsu.edu/RNaseP/home.html and psyche.uthct.edu/dbs/SRPDB/SRPDB.html.
[c]As reported in cgsc.biology.yale.edu; susi.bio.uni-giessen.de.
[d]Reported sizes of RNAs varied by 1–5 nt in some references. Lengths shown here are based on the RNA structures shown in Fig. 1.
[e]RNA levels determined as a fraction of 5S rRNA. Jain *et al.*[9] report 300 molecules of RNase P RNA per cell.
[f]Most sRNAs are subject to further trimming of 1–5 nucleotides by various exonucleases[43].

the growth of some hybrid phages. The function of tmRNA was unclear until the fortuitous discovery that a novel carboxy-terminal tag, encoded by tmRNA rather than the mRNA, could be added to recombinant proteins expressed in *E. coli*. The tag is added to polypeptides generated from mRNAs lacking stop codons in a *trans*-translation reaction in which the tmRNA molecule acts as both a tRNA and an mRNA. Proteins containing the tmRNA-encoded tag are targeted for degradation. The *trans*-translation reaction also serves to release ribosomes stalled on incomplete mRNAs.

RNase P RNA

Characterization of 10Sb RNA and biochemical and enzymatic studies on RNase P soon revealed that 10Sb is the RNA component of RNase P (a key enzyme in RNA processing; reviewed in Refs 15,16). The 377-nt RNase P RNA is processed from a primary transcript at the 3′ end by RNase E (Ref. 17), and the gene encoding the RNA (*rnpB*) is essential for viability. The RNase P RNA alone can catalyze appropriate processing of pre-tRNAs *in vitro*, indicating that RNase P RNA is sufficient for substrate recognition and cleavage activity. However, the RNase P protein, also called C5, is required for activity *in vivo* and increases the efficiency of processing *in vitro*. Analyses of RNase P RNA and substrate mutants have revealed that substrate recognition is via structural features and not by basepairing interactions. The RNase P RNA is conserved across all organisms, and, thus, phylogenetic analysis has been an exceptionally powerful tool for determination of the secondary and tertiary structure of the RNA, as well as for the

Table 2. Functions of sRNAs[a]

RNA	Function	Null phenotype	Overexpression phenotype[b]	Mode(s) of action	Physiological system regulated	Refs
4.5S	Component of signal recognition particle Potential role in translation	Death	None	RNA–protein interaction (p48 and EF-G) Unique activity as RNP	Protein export Potential role in protein synthesis	10,11, 44
tmRNA	Mediator of *trans*-translation reaction	Slow growth	None	Intrinsic activity of RNA (tRNA structure and encoded peptide tag)	Stalled ribosome release Protein degradation	13,14
RNase P	Component of RNase P	Death	None	Catalytic RNA	RNA processing	9,15, 16,45
Spot 42	Unknown	Increased sensitivity to MMS Slight growth impairment	Decreased sensitivity to MMS Growth defect	Unknown	Unknown	20–22
6S	Unknown	None	None	Unknown	Unknown	24,26
CsrB	Inhibitor of CsrA	ND	Increased glycogen accumulation Altered cell surface properties	RNA–protein interaction (inhibits CsrA activity)	mRNA decay	27
OxyS	Regulator of ~40 genes	Increased mutagenesis	Altered expression of multiple genes Decreased mutagenesis	Short RNA–RNA pairing (blocks ribosome binding to *fhlA* mRNA) RNA–protein interaction (inhibits Hfq activity)	Hydrogen peroxide stress response	28–30
MicF	Inhibitor of OmpF	Increased OmpF	Decreased OmpF	Long RNA–RNA pairing (blocks ribosome binding to *ompF* mRNA)	Outer membrane composition in response to environmental stimuli	3, 31–33
DicF	Inhibitor of FtsZ	ND	Blocked cell division	Long RNA–RNA pairing (blocks ribosome binding to *ftsZ* mRNA)	Cell division	3,32, 36
DsrA	Activator of RpoS H-NS antagonist	Decreased RpoS	Derepression of H-NS silenced genes	Short RNA–RNA pairing (prevents inhibitory *rpos* mRNA secondary structure and inhibits *hns* mRNA translation) Potential RNA–protein interaction (modifies H-NS and Hfq activities)	Response to low temperature	37–40[c]

[a]Abbreviations: MMS, methyl methanesulfonate; ND, not determined; RNP, ribonucleoprotein.
[b]As sRNAs were expressed from a variety of plasmids, the relative levels of overexpression vary.
[c]S. Gottesman, unpublished.

identification of domains important for substrate recognition and catalysis.

Spot 42 RNA
The role of Spot 42 RNA is still unclear, even though the 109-nt RNA has been studied extensively[18–22]. The gene encoding Spot 42 RNA [*spf* (spot forty-two)] is negatively regulated by cAMP (Ref. 19). Spot 42 RNA appears to be present in many forms in the cell, as shown by its representation in cytosolic, ribosomal and nucleoid fractions[19]. Strains carrying deletions of *spf* are viable but show a slight growth impairment under some conditions[20]. Tenfold overproduction of Spot 42 RNA results in a pronounced growth defect under a variety of conditions[21]. In addition, strains carrying deletions of *spf* show decreased DNA polymerase I (Pol I) activity, whereas strains that overproduce Spot 42 have increased Pol I activity[22]. Although the exact function of Spot 42 RNA has not been discovered, it is intriguing that *spf* is directly downstream of the gene encoding Pol I (*polA*).

6S RNA
6S RNA was the first non-rRNA, non-tRNA to be sequenced from *E. coli*[23] and yet its function remains elusive. This sRNA is transcribed as part of a dicistronic message from which the 184-nt mature 6S RNA is processed by an unknown mechanism[24]. The ORF downstream of 6S RNA encodes a 170 amino acid protein with homology to 5,10-methenyltetrahydrofolate synthetase, but a function for this putative protein has not been demonstrated. Intriguingly, 6S RNA is present in a stable RNP complex in *E. coli* extracts[25]. Cells carrying insertions in the gene encoding 6S [*ssrS* (small stable RNA)] are viable in the absence of 6S RNA, with no detectable growth defects on various carbon sources at temperatures ranging from 23°C to 42°C (Ref. 26). Likewise, eightfold overexpression of 6S RNA has no detectable phenotype[24].

CsrB RNA
CsrB RNA was identified by co-purification with overproduced CsrA protein, which negatively regulates glycogen biosynthesis by binding mRNAs and facilitating their decay (reviewed in Ref. 27). The ~360-nt CsrB RNA contains numerous imperfect repeats of a seven-nucleotide consensus sequence, the majority of which are in the loops of a predicted secondary structure. As the purified CsrB–CsrA RNP contains ~18 molecules of CsrA protein per CsrB RNA, one CsrA protein could bind to each repeat sequence in CsrB. Overexpression of CsrB RNA appears to antagonize all regulatory functions of CsrA *in vivo*, and studies *in vitro* suggest that CsrB RNA directly inhibits CsrA activity.

OxyS RNA
The 109-nt OxyS RNA is highly induced by hydrogen peroxide treatment[28]. Cells carrying deletions of the *oxyS* gene or overexpressing the RNA are viable. However, changes in the protein expression pattern observed with constitutive *oxyS* expression suggest that the RNA might be a regulator. In fact, biochemical and genetic approaches have confirmed that many genes are regulated by OxyS RNA. Surprisingly, OxyS might function differently at two of its target genes[29]: (1) OxyS RNA represses *fhlA* translation by directly basepairing to the *fhlA* mRNA near the Shine–Dalgarno sequence, thus blocking ribosome binding[29] and (2) OxyS RNA represses *rpoS* translation by binding to and altering the activity of a protein factor (Hfq) normally required for *rpoS* translation[30]. Although the physiological role of the hydrogen-peroxide-induced RNA is not fully understood, the observation that OxyS represses the expression of two transcriptional regulators (*fhlA* and *rpoS*) suggests that OxyS RNA integrates adaptation to hydrogen peroxide with other regulatory networks. In addition, OxyS RNA acts as an antimutator; cells deficient in *oxyS* show increased mutagenesis by hydrogen peroxide, and cells overexpressing OxyS RNA show decreased mutagenesis[28].

MicF RNA
The *micF* gene was first identified[31] as a cloned fragment that inhibits expression of OmpF, an outer membrane porin (reviewed in Refs 3,32). The 93-nt MicF RNA is induced by a variety of environmental conditions, including temperature shifts from 24°C to 37°C, high osmolarity and superoxide stress (reviewed in Ref. 33). The RNA is highly complementary to the region surrounding the *ompF* Shine–Dalgarno sequence and, therefore, it was proposed that MicF RNA represses *ompF* via an antisense mechanism[31]. The effects of deleting *micF* and overexpressing MicF RNA on OmpF expression support this hypothesis. Subsequent biochemical studies have shown that a stable duplex between MicF RNA and the *5′* untranslated region of the *ompF* message is formed, and that MicF inhibits *ompF* translation and destabilizes the *ompF* message. The exact mechanism of MicF RNA antisense repression is still being studied, but UV crosslinking experiments suggest that an unidentified 80-kDa *E. coli* protein binds the MicF RNA. The *micF* gene is distant from *ompF* but is encoded directly upstream of *ompC*, which encodes a second outer membrane protein. These outer membrane proteins are regulated such that the total amount of OmpF and OmpC protein remains constant, but the level of each protein varies with environmental conditions. As the expression of *micF* and *ompC* is regulated similarly under some conditions, it appears that the MicF RNA represents one highly sensitive mechanism to maintain the coordinate, opposing regulation of *ompF* and *ompC*.

DicF RNA
The *dicF* gene was discovered[34] as a small fragment of the *dicB* operon that could inhibit cell division when present on a multi-copy plasmid (reviewed in Refs 3,32). DicF RNA is processed from a

polycistronic transcript by RNase III and RNase E (Ref. 35). The minimal functional RNA is 53 nt, although a 190-nt RNA is also present and is presumably active. A cell-division gene, *ftsZ*, has been identified as the target of DicF RNA by a genetic screen for suppressors of *dicF*-dependent inhibition of cell division[36]. DicF RNA has significant complementarity to the *ftsZ* mRNA in the region surrounding the Shine–Dalgarno sequence, and *dicF* has been shown to regulate *ftsZ* expression post-transcriptionally, which is consistent with RNA-hybrid formation interfering with ribosome binding. Although the initial effects of DicF were seen by overexpression, the chromosomally encoded DicF RNA can also prevent cell division (J-P. Bouché, unpublished).

DsrA RNA

The *dsrA* gene was first identified as a derepresser of *rcsA* (a regulator of capsular polysaccharide synthesis) when present in multiple copies[37]. The expression from *dsrA–lacZ* fusions is increased at 20°C, in contrast with expression at 42°C, which is consistent with a role for DsrA at low temperature[38]. Overexpression of the 87-nt DsrA RNA leads to the induction of multiple genes repressed by the histone-like protein H-NS, and strains devoid of *dsrA* have substantially lower levels of RpoS, indicating that DsrA is acting as a regulator. DsrA RNA acts as an antisilencer of H-NS-regulated genes, by reducing H-NS expression by pairing with the *hns* message[39] and, possibly, by modulating H-NS oligomerization (S. Gottesman, unpublished). The DsrA RNA appears to regulate *rpoS* translation by basepairing to the *rpoS* mRNA, thereby altering its secondary structure and allowing enhanced translation[40]. Mutational analyses suggest that the effects of DsrA on *rpoS* translation are independent of the effects on H-NS-regulated genes; however, both effects require the Hfq protein[40]. Although the mechanisms of action need further investigation, there are several intriguing parallels between the OxyS and DsrA RNAs: both appear to function by more than one mechanism and both affect *rpoS* expression together with the Hfq protein, albeit under different conditions and with opposing effects.

Mechanisms of action

Although *E. coli* sRNAs have varied functions, their mechanisms of action can be grouped into three broad categories: RNAs that act via direct RNA–RNA basepairing (OxyS, MicF, DicF and DsrA); RNAs that act via RNA–protein interactions (4.5S, CsrB, OxyS and DsrA); and RNAs that have intrinsic activities (tmRNA and RNase P RNA). Some of these RNAs (OxyS and DsrA) appear to act by more than one mechanism on different targets. In addition, these categories represent our current understanding of a direct interaction required for activity, but are not exclusive. For instance, RNase P RNA is clearly the catalytic component of RNase P and can act independently *in vitro*. However, *in vivo*, RNase P RNA is present in a stable RNP complex with the RNase P protein and, *in vitro*, RNase P protein binding to the RNA increases its catalytic activity[15]. Thus, RNA–protein interactions are also critical for RNase P activity. sRNAs that act via RNA–RNA interactions can also bind proteins that could act in a variety of ways, including stabilization of the sRNA–target RNA interaction, stabilization of a specific sRNA structure facilitating activity, or could even act directly, once brought to the target by the RNA–RNA interaction. Further analyses will be necessary to identify and characterize all the interactions made by sRNAs and to understand fully the mechanisms of action.

The sRNAs that act by basepairing can be subdivided further. MicF and DicF RNAs have the potential for long basepaired regions, analogous to the antisense mechanisms described for many plasmid- and phage-encoded sRNAs (Ref. 3). OxyS and DsrA RNAs interact with their target RNAs through shorter basepaired regions. These short basepairing interactions are more analogous to many snRNA–substrate interactions in eukaryotes. Although the significance of long versus short basepairing interactions is currently unknown, it is interesting to speculate about the reasons for the difference. The strength, and therefore the duration, of sRNA–target RNA binding might depend directly on the number of basepairs formed. Are there mechanisms of action or environmental conditions that require stronger or weaker, long-term or transient, interactions? Does the length of interaction suggest different requirements for accessory factors for efficient activity? The extent of basepairing will also determine the likelihood of duplicating the target sequence throughout the genome. Do shorter interactions represent sRNAs that might act via the same RNA–RNA interaction on multiple targets? It is also interesting to note that all the *E. coli* sRNAs known to act via RNA–RNA interactions function at the level of translation. The sRNAs that act via protein interactions can also be divided into two groups: CsrB, OxyS and DsrA RNAs appear to bind target proteins and inhibit their normal function; by contrast, the 4.5S RNA binding to p48 is required for SRP activity.

Comparison with eukaryotic RNAs

There are literally hundreds of sRNAs present in eukaryotic cells, many of which are well characterized (reviewed in Ref. 5). As with *E. coli* sRNAs, the eukaryotic sRNAs use RNA–RNA interactions, RNA–protein interactions and intrinsic RNA function to mediate their respective activities. However, the functions of most of the characterized eukaryotic sRNAs are distinct from those known for the bacterial sRNAs, and it is worthwhile contemplating the differences. For instance, the majority of known eukaryotic sRNAs are snoRNAs (small nucleolar RNAs), which act in rRNA biogenesis, either in pre-rRNA processing or in the modification of rRNAs. Although bacterial rRNAs are also processed from polycistronic precursors, and the mature rRNAs

contain modifications similar to their eukaryotic counterparts, currently there are no known prokaryotic sRNAs involved in rRNA metabolism. By contrast, many of the *E. coli* sRNAs are involved in the regulation of translation of target mRNAs. The only characterized eukaryotic RNA that appears to modulate translation is the *Caenorhabditis elegans* lin4 RNA, which represses translation of its target genes during development[32].

It is intriguing to speculate whether the differences in apparent modes of action between eukaryotic sRNAs and bacterial sRNAs represent divergence between these types of organisms, or whether the differences are a function of how sRNAs have been discovered and characterized, suggesting, therefore, the presence of additional sRNAs. For instance, many of the eukaryotic sRNAs were first identified by immunoprecipitation with patient's autoimmune sera that recognize Sm and fibrillarin proteins, which are present primarily in the nucleus and nucleolus. Therefore, the dearth of eukaryotic sRNAs known to be involved in translational regulation might reflect the fact that few cytoplasmic sRNAs have been identified. Some of the prokaryotic sRNAs are expressed only under very defined conditions, and eukaryotic sRNAs similarly expressed under limited conditions might not have been discovered by approaches used to date. Many of the prokaryotic sRNAs were found by chance and, thus, the number of identified sRNAs and the range of known functions are probably limited. Given the known roles of eukaryotic sRNAs, for example in rRNA modification, it might be possible to carry out directed approaches to identify prokaryotic sRNAs of similar function.

Unresolved issues

There are several important directions for future studies of sRNAs. First is the further characterization of the sRNAs with known functions. Although the roles of the 4.5S, tmRNA, RNase P, OxyS, CsrB, MicF, DicF and DsrA RNAs are at least partially understood, many general properties of these RNAs still need to be evaluated. What are the relative amounts of the sRNAs? When and how are they expressed? How stable are the sRNAs? Do the sRNAs contain modified bases? For example, tmRNA contains two types of modified nucleosides[41], and the 4.5S, Spot 42 and 6S RNAs have been reported to be unmodified[7,8,24], but the status of the other sRNAs is not known. What are the secondary and tertiary structures of the sRNAs? The secondary structures of 4.5S, tmRNA and RNase P RNAs have been elucidated by extensive phylogenetic comparisons and biochemical analysis. Similar studies of the other sRNAs would be extremely informative for the identification of important structural and functional domains. Do the sRNAs clearly act as RNAs? Mutational or biochemical studies indicate that 4.5S, tmRNA, RNase P, Spot 42, OxyS, CsrB, MicF, DicF and DsrA all function as RNAs, but in some cases further proof is warranted. In addition, what is the subcellular localization of the sRNAs, and which proteins and RNAs interact with these regulators? Information about all of these characteristics is important for the evaluation of models of sRNA action.

A second important direction for future studies is the identification of functions for sRNAs such as Spot 42 RNA and 6S RNA. The abundance and conservation of both of these RNAs suggest that they have cellular functions. Identification of interacting factors and structural analysis might elucidate their functions in a manner similar to the identification of the role of 4.5S RNA in protein export. Further evaluation of overexpression or deletion phenotypes might also provide clues, as was the case for the OxyS, MicF, DicF and DsrA RNAs. Alternatively, the functions of these RNAs might be discovered fortuitously, as was the case for tmRNA.

Third, it is worth discussing how novel sRNAs might be discovered. Computer searching of complete genomes, based on parameters common to several sRNAs, could identify candidates. For example, searches for a promoter sequence within a short distance of a terminator could be carried out. Of course, not all sRNA genes will contain recognizable promoters or terminators. Multi-array chips or filters can be probed with total RNA isolated under a variety of conditions. This approach might allow the discovery of sRNAs that are expressed only under very defined conditions. As sRNAs can be encoded separately or within other messages, the arrays would need whole genome representation and not just ORFs or operons. In addition to these global approaches, further characterization of known sRNAs could reveal features that might be used to identify other RNAs. For example, searches could be expanded to include parameters such as common sequence or structural motifs. Other features of the known sRNAs might also provide clues. It is intriguing that DicF RNA is encoded as part of a cryptic prophage[42]. Are other sRNAs associated with cryptic prophages? OxyS and DsrA RNAs both require the Hfq protein for function. Does Hfq interact with any other sRNAs? If so, can it be used as an antigenic marker to find additional sRNAs? A combination of systematic approaches, further characterization of the known sRNAs, and chance might identify novel sRNAs in *E. coli* and other organisms.

Further studies of sRNAs are important for several reasons. Information about the mechanisms of action

Questions for future research

- Why do the characterized sRNAs from prokaryotes and eukaryotes target different cellular functions (i.e. translational control and mRNA stability versus RNA processing and modification)?
- What are general properties of the sRNAs (i.e. relative abundances, stabilities, modifications, secondary and tertiary structures, subcellular locations and patterns of expression)?
- How can functions for sRNAs be identified?
- How might additional sRNAs be discovered? How can genome sequences be probed for information about sRNAs?

would be extremely useful for biotechnological purposes. For example, many companies have tried to exploit antisense RNAs as regulators, but with only limited success. Further understanding of cellular antisense mechanisms might elucidate alternative, more favorable approaches. In addition, given the flexibility of RNA, the yet-to-be-characterized sRNAs are likely to have functions we have not yet imagined. An understanding of these functions should provide important general insights into cellular regulation.

Acknowledgements

We appreciate the comments, suggestions and information about unpublished data from M. Belfort, H. Bernstein, J-P. Bouché, S. Brown, N. Delihas, S. Gottesman, N. Pace, T. Romeo and D. Sledjeski. We sincerely apologize for being unable to cite all relevant original work owing to space constraints.

References

1 Inouye, M. and Delihas, N. (1988) *Cell* 53, 5–7
2 Eguchi, Y., Itoh, T. and Tomizawa, J. (1991) *Annu. Rev. Biochem.* 60, 631–652
3 Wagner, E.G.H. and Simons, R.W. (1994) *Annu. Rev. Microbiol.* 48, 713–742
4 Novick, R. in *Cell–Cell Signalling in Bacteria* (Dunny, G. and Winans, S., eds), ASM Press (in press)
5 Grosjean, H. and Benne, R. (1998) *Modification and Editing of RNA*, ASM Press
6 Hindley, J. (1967) *J. Mol. Biol.* 30, 125–136
7 Griffin, B.E. (1971) *FEBS Lett.* 15, 165–168
8 Ikemura, T. and Dahlberg, J.E. (1973) *J. Biol. Chem.* 248, 5024–5032
9 Jain, S.K., Gurevitz, M. and Apirion, D. (1982) *J. Mol. Biol.* 162, 515–533
10 Brown, S. (1991) *New Biol.* 3, 430–438
11 Luirink, J. and Dobberstein, B. (1994) *Mol. Microbiol.* 11, 9–13
12 Zheng, N. and Gierasch, L.M. (1997) *Mol. Cell* 1, 79–87
13 Muto, A., Ushida, C. and Himeno, H. (1998) *Trends Biochem. Sci.* 23, 25–29
14 Williams, K.P. and Bartel, D.P. (1998) *Nucleic Acids Res.* 26, 163–165
15 Altman, S. (1989) *Adv. Enzymol. Rel. Areas Mol. Biol.* 62, 1–36
16 Frank, D.N. and Pace, N.R. (1998) *Annu. Rev. Biochem.* 67, 153–180
17 Kim, S. *et al.* (1996) *J. Biol. Chem.* 271, 19330–19337
18 Sahagan, B.G. and Dahlberg, J.E. (1979) *J. Mol. Biol.* 131, 573–592
19 Sahagan, B.G. and Dahlberg, J.E. (1979) *J. Mol. Biol.* 131, 593–605
20 Hatfull, G.F. and Joyce, C.M. (1986) *J. Bacteriol.* 166, 746–750
21 Rice, P.W. and Dahlberg, J.E. (1982) *J. Bacteriol.* 152, 1196–1210
22 Polayes, D.A., Rice, P.W. and Dahlberg, J.E. (1988) *J. Bacteriol.* 170, 2083–2088
23 Brownlee, G.G. (1971) *Nat. New Biol.* 229, 147–149
24 Hsu, L.M. *et al.* (1985) *J. Bacteriol.* 161, 1162–1170
25 Lee, S.Y., Bailey, S.C. and Apirion, D. (1978) *J. Bacteriol.* 133, 1015–1023
26 Lee, C.A., Fournier, M.J. and Beckwith, J. (1985) *J. Bacteriol.* 161, 1156–1161
27 Romeo, T. (1998) *Mol. Microbiol.* 29, 1321–1330
28 Altuvia, S. *et al.* (1997) *Cell* 90, 43–53
29 Altuvia, S. *et al.* (1998) *EMBO J.* 17, 6069–6075
30 Zhang, A. *et al.* (1998) *EMBO J.* 17, 6061–6068
31 Mizuno, T., Chou, M.Y. and Inouye, M. (1984) *Proc. Natl. Acad. Sci. U. S. A.* 81, 1966–1970
32 Delihas, N. (1995) *Mol. Microbiol.* 15, 411–414
33 Pratt, L.A. *et al.* (1996) *Mol. Microbiol.* 20, 911–917
34 Bouché, F. and Bouché, J-P. (1989) *Mol. Microbiol.* 3, 991–994
35 Faubladier, M., Cam, K. and Bouché, J-P. (1990) *J. Mol. Biol.* 212, 461–471
36 Tétart, F. and Bouché, J-P. (1992) *Mol. Microbiol.* 6, 615–620
37 Sledjeski, D. and Gottesman, S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* 92, 2003–2007
38 Sledjeski, D.D., Gupta, A. and Gottesman, S. (1996) *EMBO J.* 15, 3993–4000
39 Lease, R.A., Cusick, M. and Belfort, M. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 12456–12461
40 Majdalani, D. *et al.* (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 12462–12467
41 Felden, B. *et al.* (1998) *EMBO J.* 17, 3188–3196
42 Faubladier, M. and Bouché, J-P. (1994) *J. Bacteriol.* 176, 1150–1156
43 Li, Z., Pandit, S. and Deutscher, M.P. (1998) *Proc. Natl. Acad. Sci. U. S. A.* 95, 2856–2861
44 Brown, S. and Fournier, M.J. (1984) *J. Mol. Biol.* 178, 533–550
45 Motamedi, H. *et al.* (1982) *J. Mol. Biol.* 162, 535–550
46 Schmidt, M., Zheng, P. and Delihas, N. (1995) *Biochemistry* 34, 3621–3631
47 Vogel, D.W. *et al.* (1987) *Nucleic Acids Res.* 15, 4583–4593

Reference added in proof

48 Amundsen, S.K., Taylor, A. and Smith, G.R. (1998) *Nucleic Acids Res.* 26, 2125–2131