Small RNA biology is systems biology

Daniel Jost, Andrzej Nowojewski & Erel Levine*
Department of Physics, FAS Center for Systems Biology, Harvard University, Cambridge MA 02138, USA

During the last decade small regulatory RNA (srRNA) emerged as central players in the regulation of gene expression in all kingdoms of life. Multiple pathways for srRNA biogenesis and diverse mechanisms of gene regulation may indicate that srRNA regulation evolved independently multiple times. However, small RNA pathways share numerous properties, including the ability of a single srRNA to regulate multiple targets. Some of the mechanisms of gene regulation by srRNAs have significant effect on the abundance of free srRNAs that are ready to interact with new targets. This results in indirect interactions among seemingly unrelated genes, as well as in a crosstalk between different srRNA pathways. Here we briefly review and compare the major srRNA pathways, and argue that the impact of srRNA is always at the system level. We demonstrate how a simple mathematical model can ease the discussion of governing principles. To demonstrate these points we review a few examples from bacteria and animals. (BMB reports 2011; 44(1): 11-21)

INTRODUCTION

The role of RNA molecules in the regulation of gene expression was first suggested by Jacob and Monod 50 years ago (1). Still, for over 4 decades evidence for RNA-based regulation was limited to the control of mobile elements (plasmids and transposons) and isolated examples of endogenous bacterial small RNAs in *Escherichia coli*. But at the turn of the millennium evidence for the real impact of small RNA were accumulating: the first animal microRNA, the heterochronic gene *lin-4* of the nematode *Caenorhabditis elegans*, was discovered (2, 3); microarray experiments, comparative genomics and bioinformatics (4-6) revealed dozens of small RNAs in *E. coli*; and the ability of double-stranded RNA to completely silence endogenous gene indicated the existence of the RNAi pathway (7). But full appreciation for the role of RNA in controlling gene regulation came only when another *C. elegans* microRNA, *let-7*, was shown to have homologs in species across the animal kingdom, including mouse and human (8, 9). Since then, the accumulation of genome sequences that allow bioinformatic identification of small RNA, and in particular the emergence of Next-Generation Sequencing methods enabling deep sequencing of the transcriptome, led to identification of countless microRNAs and multiple small RNA pathways (For recent reviews see 10, 11).

Small RNAs in Animals

Multiple srRNA pathways exist in animals, including the well studied microRNAs (miRNAs), endogenous small-interfering RNAs (siRNAs) in the RNAi pathway, and Piwi-interacting RNAs (piRNAs) (Fig. 1). MiRNAs have been shown to play significant roles in all fields of biology, including development (reviewed in 12, 13), stem cell differentiation (14), germline maintenance (15, 16), cancer biology (17) and stress response (18). piRNAs have been implicated in maintenance of genomic integrity and suppression of transposon activity (19). The role endogenous siRNAs is even more diverse and less

![Fig. 1. Outline of three srRNA pathways in animals (miRNA and siRNA) and bacteria (Hfq-dependent srRNAs).](image-url)
Small RNA biology is systems biology
Daniel Jost, et al.

understood.

Small RNA pathways differ in the biogenesis of the small RNAs. miRNAs are derived from distinctive hairpin precursors that are genomically encoded, transcribed by RNA polymerase II and undergo several processing steps (20, 21). Endogenous siRNAs and piRNAs in animals come from multiple sources, including transposons, long RNA transcripts, and mRNAs (22-31); siRNAs are processed from double stranded RNAs, while piRNAs are not. Different sets of proteins are involved in biogenesis of siRNAs in different pathways, although some of these enzymes can work in multiple pathways (see below).

The mechanism of gene silencing is also divergent: siRNAs lead to direct cleavage of the bound mRNA (32, 33), while the mechanism of miRNA silencing is less direct, and may involve both translation inhibition and mRNA destabilization (reviewed in 11). siRNAs require perfect base-pairing with their target, whereas miRNA hybridization to their target is more complex, and involves a perfectly matched “seed” 6-8 nt long and some mismatches outside of it. In some animals (and in plants), siRNA or piRNA binding to its target can trigger the synthesis of secondary siRNAs (34-37), thus amplifying the silencing signal. In nematodes (and in plants) this allows for systemic silencing, where the amplified silencing signal can travel across cells and tissues (38).

At the same time, much is common to the diverse RNA pathways. Importantly, specificity is attributed to Watson-Crick basepairing between the srRNA and its targets. Dicer is involved in srRNA processing in multiple pathways. All known srRNAs bind to a member of the Argonaute protein family, which confers stability to the small RNA molecule, facilitates its binding with its mRNA target, and is involved in recruiting additional factors required for silencing.

Small RNAs in Bacteria

Bacterial small regulatory RNAs (recently reviewed in 39, 40) are longer than their eukaryotic counterparts, and are typically between 60-150 nt long. This can be attributed to the rarity of RNA processing in bacteria, as animal miRNAs are also synthesized as longer molecules that are processed to their final small size post-transcriptionally. Bacterial srRNAs typically bind the 5’UTR of their target in the vicinity of the ribosomal binding site (RBS), resulting in modulation of translation efficiency and stability of the mRNA molecules. While RNA regulation in eukaryotes typically results in gene silencing, bacterial small RNAs can both repress or activate their targets.

The physiological role of bacterial small RNAs is diverse. Notably, bacterial small RNAs are abundant in stress response pathways (41), and have been implicated in the regulation of virulence genes (42-47). While the molecular interaction between a bacterial small RNA and its target may be more complex than that of the short srRNA in eukaryotes, recognition is still typically limited to a short (7-10 nt) sequence of almost perfect complementarity (40).

Small RNA Recycling Determines its Functional Properties

Post-transcriptional silencing of a gene requires targeting the multiple mRNA molecules that reside in the cytoplasm. These molecules need to be efficiently found by the srRNA, bound and silenced. In the RNAi pathway, a siRNA - bound to an Argonaute protein in an RNA-induced silencing complex (RISC) (48) - can be recycled and undergo multiple rounds of silencing (49-51). In some cases (as mentioned above) the interaction of the siRNA with its target can even lead to the synthesis of secondary siRNAs. With these two properties, a relatively small dsRNA trigger can result in efficient - and sometimes systemic - gene silencing.

Gene silencing by miRNAs, however, is believed to be more complex than simple direct cleavage of the target, leading to some titration of the miRNA to its target. miRNAs therefore act in a more stoichiometric way, and their levels should be better matched with that of their targets (52). Indeed, in cases where a miRNA is known to strongly suppress its target, the miRNA is highly expressed.

An abundant class of bacterial small RNAs marks the other extreme. These srRNAs require the bacterial RNA chaperone Hfq for efficient gene silencing (40). Hfq binds to these srRNAs and stabilizes them, thus forming the bacterial equivalent of RISC. Small RNAs of this class may be directly affected by the interaction with their target. In such cases, a small RNA directs cleavage of the srRNA:mRNA complex, thus reducing the stability of the mRNA target while promoting its own degradation (53, 54).

Recent quantitative studies of small RNA regulation focused on the implications of incomplete (or absent) srRNA recycling on the properties of target repression (55-59). Using simple kinetic models (Box 1) it has been demonstrated that srRNAs can impose a threshold for the activation of gene expression: a significant gene silencing is achieved when the rate of mature mRNA production is smaller than the threshold, and is effectively relieved above it (Fig. 2). This mechanism for silencing has been shown to exhibit reduced level of fluctuations, as compared with a more catalytic reaction (such as transcriptional repression). Dynamically, stoichiometric srRNA-target interactions result in a delayed response to abrupt changes in the activity of either gene, allowing filtering of environmental noise. These models suggest that the efficiency of srRNA regulation (i.e. the sharpness of the transition, noise reduction etc) is controlled by a single effective parameter which lumps together the biochemical properties of the srRNA and its target (Box 1 and Fig. 2).

Coordinated Response in Bacteria (1): srRNA Regulation of Master Regulators

Two of the better characterized srRNAs in bacteria are regulating multiple mRNA species including global transcriptional
Box 1. Minimal Model for Small RNA Regulation

The motivation behind composing a generic mathematical model for small RNA based regulation comes from two "big" questions. First, on the face of it small RNA regulation and transcriptional regulation seem redundant. Nevertheless, the modes of regulation exist almost in all organisms, suggesting that the two complement each other in some non-trivial way. It was therefore suggested that small RNA regulation might exhibit quantitative features that make it different from protein-based transcriptional regulation. Second, given that small RNAs have been missed over decades of molecular biology, it has been suggested that their role in regulation is minor, idea that was supported by the fact that deletion of almost all miRNAs in C. elegans, one by one, has no effect on development or viability. It was therefore suggested that small RNAs have a more quantitative role in tuning gene expression.

In order to answer such high-level questions, one needs to construct a model that is stripped down to the essential ingredients, without losing the essence of the modeled system. When done right, details can either be lumped together into effective parameters, or can be accounted for by functions whose properties - albeit not detailed form - can be postulated.

Generic models for small RNA regulation typically focus on a small RNA and one (or a small number of) targets, and account for synthesis of the RNA molecules (with rates denoted below by \( \alpha_1 \)) or degradation (by \( \beta_1 \)) of the targets, and the two-species interaction (denoted by \( k \)). Recycling of the small RNA can be parameterized via a probability parameter \( q \). In its simplest form, the model is represented by a set of mass-action equations.

\[
\frac{dn_i}{dt} = \alpha_1 - \beta_1 n_i - (k_i m_i - k_i c_i)
\]

\[
\frac{dc_i}{dt} = \alpha_1 - \beta_1 s - \sum_i [k_i m_i - (k_i - q_1 c_i)]
\]

\[
\frac{ds}{dt} = k_i m_i - k_i c_i - \gamma s
\]

Here \( n_i \) is the concentration of miRNAs of species \( i \), \( s \) is the concentration of the srRNA, and \( c_i \) the concentration of srRNA:mRNA complex. To complete the model one needs to specify the process of translation. For example, in cases where translation is inhibited directly by binding of RISC, the concentration of protein \( i \) will be proportional to that of its free mRNA, \( m_i \). However if RISC binding only facilitates degradation of the molecules (e.g. in the sRNA pathway) then protein concentration is proportional to \( (m_i + c_i) \).

More generally, other linear combinations of the two are possible. Even in this simplified scheme, these are highly coupled non-linear differential equations, and one cannot solve them exactly. However, if one focuses on steady state properties, all time derivatives (the left-hand side of these equations) are set to zero, and one remains with a simpler set of algebraic equations. These are readily solved for the case of one or two targets. In this case the kinetics of the interaction between the srRNA and a target is given by the binding rate \( k = qk'/(k_i' + q's')\). The last approximation holds in cases where the complex is efficiently degraded (as in the bacterial case and in the RNAi pathway).

A key feature of this model is a non-linear relation between the transcription rate of a gene and its steady-state mRNA concentration (Fig. 2). In this "threshold-linear" response, the gene is efficiently silenced when its transcription rate is below a prescribed threshold. Above the threshold, the mRNA level increases linearly with the transcription rate. The behavior near the threshold is determined by a combination of all other model parameters, that measures the efficiency of srRNA binding and regulation. Efficient regulation (e.g. large \( k \) and/or small \( \beta \)') makes the transition sharp and well defined, less efficient regulation makes it broader and leaky.

The "threshold-linear" picture already provides one way to see the cross-talk between different targets of an srRNA. From the perspective of one target (say, \( i = 1 \)), the threshold in \( \alpha_1 \) is given by \( \alpha_1 = \alpha_2 \). Thus, at one level of \( \alpha_1 \), the first gene may be either in the silenced or active state, depending on the rate of transcription of the second target. Alternatively, one can either use numerical methods, or group all "other" targets into a single effective species.

Increasing the complexity of the mathematical description can reveal more quantitative features. To study the dynamics of regulatory response, one retains the temporal derivatives. While limited analytical progress can be made, integrating these equations numerically is straightforward, and one can identify scenarios where the temporal response may be faster or slower than the expected response in transcriptional regulatory circuits. The model can be developed further to account of the discreteness in molecules number and the burstiness in transcription and translation processes, showing that small RNA acts efficiently to suppress intrinsic noise in the silenced state.

Without major modifications, this model can integrate more details of the pathway and account for more players, assuming that these act to modify the kinetic parameters. This approach is taken in the text to account for enzymes of the srRNA pathways (such as the bacterial RNA chaperone Hfq, and the eukaryotic Dicer and Argonaute proteins).
Mediates in the quorum sensing signal transduction pathway (71–73). The Quorum sensing mechanism integrates multiple chemical signals to sense the density of like cells in the environment (74). Based on these signals, a master regulator is activated and goes on to regulate multiple response pathways. Interestingly, the (redundant multi-copied) Qrr sRNAs are positioned downstream of signal integration (by transcriptional response regulators) and up-stream of the master regulator.

The function of this pathway can be described in terms of a computational device called a classifier. A classifier takes multiple inputs, and needs to classify them into one of several possible classes. In the case of the quorum sensing pathways, a linear "weighted" sum of the input signals are passed through the thresholding filter provided by the sRNAs to perform a classification between "high" and "low" cell density (71). This classification scheme is known as a linear classifier, and is one of the basic ingredients in the theory of machine learning.

These two examples show how small RNAs can play a role in global regulation by targeting just a few strategic targets. The role of the sRNA in this case is to link between the input signal (or signals) and the response by setting a threshold for the response. It is easy to see how global regulation would benefit from the quantitative features of sRNA regulation, in particular the tight repression and noise suppression in the silenced state.

The Many Targets of Small RNAs

The sequence that is responsible for specificity of sRNA-target interaction is limited to around 6–8 nt. It is therefore no surprise that naive attempts at predicting targets of a small RNA yield astounding number of potential targets per sRNA. As target prediction algorithms improve to incorporate target structure, thermodynamic predictions, and - most importantly - evolutionary conservation (75), the idea that most miRNA have multiple potential targets becomes substantiated. Indeed, some miRNAs are predicted to have even hundreds of targets (76, 77).

MiRNAs certainly affect the evolution of mRNA 3'UTR (78, 79), such that binding targets for a particular miRNA would be specifically absent from genes that are co-expressed with the miRNA. Thus, in many cases putative miRNA binding sites have no real biological implication, as they are present in miRNAs that are rarely co-expressed with the targeting miRNA.

Still, the association of multiple targets with one miRNA raises the possibility that a small RNA rewires the genetic network at the system level. Incomplete recycling of the small RNA results in indirect interactions among targets of a small RNA, as the expression of each target affects the shared sRNA. From a quantitative standpoint, these interactions can be accounted for without significantly complicating the mathematical model (Box 1). This model suggests that this crosstalk may in fact be highly efficient and sensitive (55, 56).

In the reminder of this review we discuss a few examples for the system-level behavior of small RNA, starting from a small pathway and working our way up towards the cell-level system.

Coordinated Response (2): Recycling-dependent Cross Talk between Targets of a Small RNA

Targets of a common sRNA can interact indirectly by modulating sRNA abundance. This is exemplified clearly by a conserved bacterial small RNA, ChiX, which acts in the chitin metabolism network (54, 80). Chitin is an abundant organic polymer, whose main degradation product is chitobiose (N-acetylglucosamine dimer).

Two conserved targets of ChiX are chb and chiP. The chb operon in E. coli and Salmonella encodes genes for the transport and degradation of chitobiose. This operon also encodes its own regulator, chbR, which represses transcription in the absence of chitobiose and activates it in its presence. The chitotriphin ChiP was recently found to be absolutely required for Salmonella to grow on chitotriose, but is not essential for uptake of chitobiose. While the chb operon is transcriptionally active only in the presence of the inducer, chiP is transcribed constitutively.

The small RNA ChiX is also constitutively expressed. In the absence of the inducer, it strongly suppresses the accumulation of ChiP mRNA, but is not co-degraded. Conversely, in the presence of chitotriose the Chb mRNA binds the sRNA and degrades it (presumably with little or no effect on the expression of genes in the operon). Thus, transcription of chb allows ChiP mRNA (and proteins) to accumulate.

Interestingly, ChbR, that up-regulates the transcription of chb (81), also activates (moderately) the transcription of chiP. What is the need, then, for the indirect activation through ChiX? A possible answer comes from the quantitative model
shown to be most strongly affected by a served seed sequence complementarity among all putative targets of RyhB. First, the SodB mRNA has the longest coding sequence, and therefore the longest binding sequence for RyhB (90), however it failed to explain why a small RNA is required in this position. Possibly, this hierarchy reflects the relative abundance of the different degradation products of chitin.

This simple model allows querying other aspects of the relationship between the targets (82). For example, one may expect that efficient depletion of ChiX would require a preference for binding to chb rather than to ChiP. Model results, however, indicates that once the binding rates to both targets are at the same range, no significant improvement is achieved from preferential attachment to chb. This is due to the cooperative non-linearity of this mode of action.

**Coordinated Response (3): The Oxidative Stress Hypothesis**

The fact that a small RNA is not only affecting its targets but is also affected by them offers an attractive interpretation for the role of a small RNA in coordinating stress response. For clarity, we focus on RyhB, a small RNA that is involved in iron metabolism in enteric bacteria (83-87). The ferric uptake repressor Fur, the key regulator of genes involved in iron uptake and metabolism, transcriptionally controls RyhB. Transcription of the RyhB gene is activated under iron limitation. Multiple genes were found to be under the control of RyhB, including genes encoding for non-essential iron-using proteins, ferritins and importantly the superoxide dismutase (SOM) SodB (83, 85). RyhB is also a positive regulator of ShiA, a permease of shikimate, a compound believed to function in the biosynthesis of siderophores that acquire extracellular iron (88). Fur itself is also repressed by RyhB (89).

Given that most targets of RyhB utilize iron, it has been suggested that RyhB assists in controlling the flux of iron in the cell, along with other regulators (85). Quantitative analysis of iron metabolism in *E. coli* suggested a fairly robust control system (90), however it failed to explain why a small RNA is required in this position.

We propose here an alternative, system-level view of the role of RyhB. We note that one target, SodB, stands out as a major target of RyhB. First, the SodB mRNA has the longest conserved seed sequence complementarity among all putative targets of RyhB. Second, the mRNA level of SodB has been shown to be most strongly affected by a RyhB deletion (83, 85). Finally, some enteric bacteria carry multiple copies of RyhB, some of which conserve only the base-pairing with SodB (A Nowojewski, unpublished data). SodB, which encodes for the Fe-SOM, also stands out for its function. While it indeed obeys the definition of an iron-carrying gene that is not essential under iron starvation, SodB is the only RyhB target that functions as a stress response gene.

Superoxide is only mildly reactive physiologically. However, iron interacts with this species to generate a highly reactive and extremely damaging hydroxyl radical. *In vivo* superoxide concentrations are considered too low to cause iron reduction, but can be sufficiently high to damage the exposed (4Fe-4S) clusters of iron-sulfur proteins, some of which are targets of RyhB (91). Iron-mediated sensitivity to superoxide stress is also overcome by increasing the levels of iron chelators and the storage capacity through ferritins.

Taken together, we suggest that the targets of RyhB can be classified in two classes: those targets that are regulated by RyhB, and those that modulate RyhB. SodB is our major candidate for the first class. The levels of all other genes, together with transcriptional regulation by Fur, make the effective accumulation rate of RyhB a faithful gauge for the iron-related severity of superoxide-stress, that goes beyond the instantaneous iron concentration (82). Targets of the second class also play a role in further reducing fluctuations in the main target (82). The ability of utilizing RyhB to estimate future levels of free iron may also be useful during pathogenesis, as the bacterial colony engages in a battle for iron with the host.

Might similar target classification be relevant to animal microRNA? One evidence in this direction comes from the two founding fathers of the miRNAs family, the miRNAs *let-7* and *lin-4* of *C. elegans*. Each of these miRNA is predicted to have tens of targets. However, a compensating mutation in only one gene (*lin-41* and *lin-14*, respectively) is enough to suppress a mutation in the *miRNA* sequence (92, 93). It is possible, then, that these are the “class-1” targets of these miRNAs, while other targets serve to regulate (or buffer) the miRNA.

**Competition for the Enzymes of the srRNA Pathways**

Many of the enzymes that make the srRNA pathways are shared among different pathways. This feature, however, is not conserved. Mammals and *C. elegans*, for example, have only one Dicer that is shared between the miRNA and srRNA pathways (94, 95), while in *D. melanogaster* two enzymes, Dicer-1 and Dicer-2, are used exclusively in the miRNA and srRNA pathways, respectively (96). *C. elegans*, on the other hand, has at least 27 Argonaute proteins, each with a unique function, as opposed to humans where one of 4 Argonaute proteins, Ago2, functions in the miRNA and multiple srRNA pathways (97). Therefore, under conditions where one or more of these - and other enzymes in the srRNA pathways - become rate limiting, shared enzymes may introduce cross-talk among different pathways.

Due to the high efficiency of RNAi regulation, synthetic siRNAs or miRNAs have been widely developed for genomic studies or potential therapeutic treatments. To ensure complete silencing of the target gene one is typically motivated to use high levels of siRNA. However, many side effects have been noticed when introducing highly concentrated synthetic siRNAs necessary to silence the desired target gene (98). In particular, the down-regulation of many off-targets of the introduced siRNA is observed (98, 99) and more surprisingly, the

http://bmbreports.org
global up-regulation of transcripts which do not share complementary site with the sRNA (100). These changes in the regulatory pathways could potentially lead to toxic or lethal modifications (101). While the down-regulation effect is straightforward and could simply be explained by the presence of extra-regulatory molecules in the cell, it has been advanced that the up-regulation effect could be due to the saturation of the endogenous RNAi machinery by a high number of exogenous sRNAs that compete for the enzymes (102).

In studying the function of endogenous miRNAs (or siRNAs) it is a common practice to study both loss-of-function and gain-of-function mutants. One way of achieving the latter is creating a transgenic line that over expressed the endogenous miRNA. Once again, strong over expression (e.g. through the use of a high-copy plasmid) may oucome other sRNAs, expressed at endogenous levels, for RNAi enzymes.

To account for competition over enzymes, the simplified model introduced above needs to be augmented with a quantitative relation between the kinetic parameters of the model (e.g. synthesis and binding rates) and the concentrations of the enzymes. For example, if the rate limiting enzyme are components of the RISC (e.g. Ago2, as believed to be the case in RNAi experiments (102,103)) the siRNA:mRNA complex formation rate k would need to be specified as $k = k_0 \frac{[\text{RISC}]}{([K]+[\text{srRNA}]_{\text{total}})}$. Here $[\text{srRNA}]_{\text{total}}$ accounts for the total concentration of all free sRNAs that compete for the enzyme, and may include sRNAs from different pathways. It is therefore organism dependent. The parameter K sets the scale for the total sRNA concentration at which the [RISC] component becomes saturated. Straightforward algebraic manipulation yields a characteristic dependence of the expression level of a silenced target on $[\text{srRNA}]_{\text{total}}$ as $(A+[\text{srRNA}]_{\text{total}})/(A+B+[\text{srRNA}]_{\text{total}})$. The positive amplitudes A and B depend on the model parameters, in a way that becomes messier as one introduces more features into the model. In any case, the underlying parameters are not expected to be known for a particular case. Still, this functional form can easily be identified (or ruled out) if one plots the level of the miRNA target as a function of the concentration of exogenous siRNAs on a double-log scale (Fig. 3).

Alternatively, the rate-limiting enzyme may be Dicer. In this case, one modified the expression for the biogenesis rate of mature miRNA in the model as $\alpha_{\text{miRNA}} = \alpha_{\text{Dicer}}/[K+[\text{srRNA}]_{\text{total}}]$. This time, the functional form for the dependence of the miRNA target concentration on $[\text{srRNA}]_{\text{total}}$ is more complex and non-linear than in the RISC case, a reminiscent of the threshold-linear behavior described above. Plotting this dependence on a double-log scale is quantitatively different, in a way that allows one to discriminate between the two even without knowing any of the underlying parameters (Fig. 3, and unpublished data from our lab).

De-repression by Exogenous Targets - miRNA “Sponges”

Not only saturating levels of sRNA can affect the activity of RNAi enzymes, but also high levels of targets, as in cases where the sRNA:target complex is long-lived. This is particularly true for bacterial sRNAs that rely on the RNA chaperone Hfq for regulation (40, 104-106). Hfq binds not only the small RNA (making the bacterial equivalent of RISC) but also to the miRNA. Binding of Hfq to the target accelerates the kinetics of sRNA:mRNA binding, and in some cases may play a role in stabilizing the miRNA molecule in vivo, perhaps by protecting it from RNaseE-dependent degradation by sequestering its initiation site. Hfq has been reported to exist in the E.coli cell at extremely high numbers (on the order of 10,000 copies of hexameric Hfq, similar to the abundance of ribosomes in the fast growing cell) (107). However recent single-molecule data suggest numbers that are order of magnitude smaller (108). Contradicting evidence about the level of Hfq at stationary state and under stress conditions (107), when many sRNAs are activated, make it even harder to determine the relevant level of Hfq. Indeed over-expression an Hfq-binding target can alleviate sRNA repression globally (109). It is straightforward to account for Hfq in the mathematical model (Box 1), in a way similar to our treatment of RISC above. One readily concludes that sRNA targets that experience “efficient” sRNA regulation (due for example to fast miRNA:sRNA binding) would be more sensitive to the level of Hfq. If the abundance of Hfq changes significantly during the life cycle of a bacterial population (e.g. between vegetative growth and steady-state), a regulating small RNA may couple the expression of such genes to the global state of the cell.

In animals, the centrality of miRNAs in many biological pathways motivated the need for miRNA loss-of-function approaches. One such approach is the introduction of a “spoon” miRNA, which contains multiple target sites complementary to the miRNA of interest (110). When the spoon is
expressed at high levels, it specifically inhibits the activity of a whole family of miRNAs that share a common seed. This approach has major advantages over genetic knockout; among which is the fact that srRNAs are frequently redundant, and family members may appear in multiple genetic loci.

The efficiency of miRNA sponges has been shown to depend not only on the affinity and avidity of binding sites, but also on the concentration of sponge RNAs relative to the concentration of the miRNA (52). To maximize sponge expression, sponges are cloned behind strong promoters, and high-copy number plasmids are used for transfections.

Introduction of highly abundant miRNA target affects not only the specific miRNA. It may also titrate limiting components of the srRNA pathways, such as Argonaute proteins, resulting in global deregulation of multiple endogenous miRNA and siRNA targets (e.g. in humans, where Ago2 is used in both pathways), as discussed above. It is therefore of merit to limit the expression of the sponge RNA.

The same prototypical model discussed above (Box 1) can be used to consider the effect of a miRNA sponge. First it is noted that the miRNA sponge should compete for all miRNAs, not only the free ones. This distinction may sometimes be elusive both in modeling and in interpreting experimental results, but is critical in cases where a large fraction of the miRNA population is titrated to targets. Effective de-repression requires a match between the synthesis rates of the miRNA sponge and the mature miRNA, not their abundance.

The quantitative model allows one to investigate ways to trade expression rates with other properties. For example, within this simple model, one can readily show that the effect of a miRNA sponge relies on the product \( \frac{k_{\text{endogenous target}}}{k_{\text{target}} \times \alpha_{\text{sponge}} / \beta_{\text{sponge}}} \), where \( k \) denotes the binding rate to the miRNA and \( \alpha \) the biogenesis rate. Thus one can trade affinity with expression in a linear fashion.

Conclusions: Unifying View of srRNA Systems Biology

In this mini review we toured multiple small RNA pathways in bacteria and animals to point out the system level aspects of small RNA regulation. While these pathways are very different-in-principles and in details - they also share some common aspects, especially at the coarse-grained and system level.

We suggest that one should always take a systems approach when studying small RNA regulation. This is due to the fact that post-transcriptional regulation has to act on multiple molecules and species at the same time, thus introducing indirect coupling among targets and between pathways. These indirect interactions may be essential part of the story (as in the examples of chitin and iron metabolism in bacteria), or may just complicate the interpretation of experimental data (as in the case of RNAi and miRNA sponges). Moreover, the non-linear response of an mRNA to its small RNA may complicate the results of large scale experiments meant to identify the targets of a particular srRNA, as srRNAs only have significant effect when a target is expressed at low levels; highly abundant target is only mildly affected by an srRNA (but may have a large effect on the srRNA itself). Consequently, we suggest that some srRNA:target interactions are there to affect the miRNA, while some are placed to modulate the level of the srRNA.

Attempts to answer fundamental questions, that are true across organisms and pathways, require high-level models. For a given question, it is not clear that such models even exist, as it may be possible that the details that have to be ignored to construct it cannot be omitted. Here, however, it seems that under some simplifying assumptions, a fairly simple mathematical model can surface interesting possibilities and generate testable predictions across a wide range of biological realms.

REFERENCES


Small RNA biology is systems biology

Daniel Jost, et al.

J. 28, 429-439.


