Title
Reverse Engineering of RNA Regulatory Networks

Author
Adamson, David Nellinger

Publication Date
2013

Peer reviewed|Thesis/dissertation

Reverse Engineering of RNA Regulatory Networks

by

David Nellinger Adamson

A dissertation in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Biophysics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:
Professor Han N. Lim, Chair
Professor Adam P. Arkin
Professor Jan T. Liphardt
Professor Jasper Rine

Fall 2013

Reverse Engineering of RNA Regulatory Networks

Abstract

Reverse Engineering of RNA Regulatory Networks

by

David Nellinger Adamson

Doctor of Philosophy in Biophysics

University of California, Berkeley

Professor Han N. Lim, Chair

Understanding gene regulation is of central importance to biology because controlling when, where, and how genes are expressed accounts for a large proportion of the complexity and organization of living systems. While for many years RNA was largely thought of as the middleman for converting genetic information into biochemically active proteins, it has become increasingly clear that RNAs are also major players gene regulation. This thesis examines the functional properties of regulatory RNAs in the context of bacterial RNA networks. More specifically, the three studies presented here examine signaling dynamics and crosstalk within the CsrA and Hfq-dependent small RNA regulatory systems; each of these systems (i) contain non-coding RNAs which regulate the translation of multiple target messenger RNAs, (ii) have central roles in coordinating their cells' adaptation to environmental change, and (iii) are evolutionarily conserved across a wide range of bacterial species. In this work, a combination of synthetic biology and mathematical modeling are brought to bear to uncover the operational behavior and potential of the these two model RNA networks.

Table of Contents

Manuscripts and Publications

Material from the following manuscripts and publications has been incorporated into this work:

Adamson DN, Lim HN (2011) "Essential requirements for robust signaling in Hfq dependent small RNA networks" *PLoS Comput Biol* 7(8): e1002138

Adamson DN, Lim HN (2013) "Rapid and robust signaling in the CsrA cascade via RNA-protein interactions and feedback regulation" *Proc Natl Acad Sci* 110(32): 13120-13125

Adamson DN, Lim HN (2014) "Non-coding RNAs buffer crosstalk in the CsrA regulatory network" *(manuscript in preparation)*

Acknowledgements

I have many to thank for their support over the past six years; without their encouragement, strategic guidance, technical assistance, companionship and camaraderie this work would not have been possible.

To my family: thank you for supporting me unwaveringly through the trials and uncertainty of graduate school. I thank my parents in particular for over and over again helping me maintain perspective when I needed to the most. I also thank my sister: you are a patient listener and your sixth sibling-sense helps me feel understood even when I am at my most ridiculous.

To Kristen Stone and to Doran & Nina Bennett: you have been my family away from home; you turned California from a strange and surreal place (in the eyes of a midwesterner) into the host of a thousand happy memories. Kristen: thank you for being my partner; every day with you is a joy and I look forward to many more to come.

To my colleagues in the lab: thank you for your friendship. Razika Hussein: your pranks, banter and unyielding need to celebrate each lab birthday more elaborately than the last will be forever memorable. Dena Block: your contagious laughter can be heard from corridors away and I hope its volume never diminishes. Emily Chang: you have remained ever skeptical of my antics, and for that I applaud you. Katya Frazier: your baked goods are the stuff of legends. Michelle Marcus: your incessantly upbeat lab singing has been sorely missed. Tiffany Lee: your fighting to be the first to perform lab chores has set a strange but admirable precedent; the battles have continued to this day.

To my advisor Han Lim: thank you for your mentorship and generosity; your scientific and career guidance has been invaluable and I expect will continue to be so for years to come. With our daily conversations spanning all topics practical, philosophical, and bizarre, you should know that working with you has been a privilege and a pleasure.

To Adam Arkin, Jan Liphardt and Jasper Rine: thank you for serving on my thesis committee for your valuable advice and support. To Kate Chase: thank you for being a resolute advocate for your students and tirelessly working to ensure their wellbeing.

Last but not least, to the millions of bacteria who made this work possible: thank you for making the ultimate sacrifice on the front lines of scientific progress.

1. Introduction

1.1 Functional properties of gene circuits

While the content of individual genes plays a crucial role in an organism's operation and survival, much of the complexity and structure of living systems emerges from how the expression of those genes is controlled. Gene regulation is a broad topic because individual genes can be controlled by staggering numbers of distinct mechanisms even within a single cell. Among gene regulatory mechanisms, we observe many distinct functional properties; as an example, some regulatory mechanisms modulate gene expression in a graded fashion in response to an input stimuli (much like a rheostat) while others will keep gene expression off until the input exceeds a threshold after which expression is fully on (much like a transistor or switch) [1-5]. Some more sophisticated regulatory mechanisms can either filter or amplify 'noise' arising from the randomness inherent in biochemical reactions (*e.g.* to ensure signal fidelity and developmental reproducibility or to increase phenotypic variation in a population) [6-8] while others can even cause gene expression to oscillate as a function of time (*e.g.* for generating circadian rhythms) [9,10].

Because many properties of regulatory mechanisms can be described using analogies to electrical engineering and because gene regulation is often accomplished through the interaction of multiple gene product 'components', collections of these interacting gene products are often referred to either as "gene circuits" or "gene networks". Determining the functional properties of a regulatory gene circuit is important for understanding the physiological role of the gene(s) that the circuit controls. Additionally, identifying general principles for how a gene circuit may acquire these properties can allow us both (i) to more readily discern the properties of uncharacterized gene circuits and (ii) to rationally design synthetic gene circuits for industrial, therapeutic, diagnostic and/or bioremediation purposes. In the studies presented in this thesis, naturally occurring gene circuits within *Escherichia coli* are reverse engineered in an effort (i) to identify new general principles that dictate gene circuit properties and (ii) to determine what known principles are at work within native gene circuits.

1.2 The role of mathematical modeling

In the study of genetic circuits, it is useful to create mathematical models for two significant reasons. First, mathematical modeling helps manage complexity. The study of gene circuits is the study of how interactions among gene products produce emergent behavior and biological function. Because the non-linear interaction of even a very small number of components can produce highly complex behavior (*e.g.* three components is sufficient for generating deterministic chaos) [11,12] modeling is an important tool for testing the conceptual soundness and internal consistency of a hypothesis before it is applied against an actual biological system [13].

Second, modeling can also help maintain generality. The ultimate goal in studying gene circuits is not to characterize circuits on a case-by-case basis with a high degree of specificity; instead, the goal is to extrapolate general mechanisms that are applicable to large classes of circuits from the detailed observation of a select few. With this philosophy in mind, the most valuable models are those that represent a system in the simplest way possible while still reproducing that systems' fundamental behavior. The use of minimal, granular models allows for the omission of extraneous detail that would unnecessarily restrict comparisons between the specific circuit studied and other related circuits. The models presented in this thesis were constructed with this mindset; as an example: modeling of the binding between the CsrA protein and the CsrB non-coding RNA largely omitted cooperativity/anti-cooperativity because that level of detail was shown to be unnecessary for the models to reproduce the fundamental dynamic and buffering properties of the system (see below) [14].

1.3 The role of synthetic gene circuits

The ability to produce sophisticated synthetic gene circuits is of great value when reverse engineering natural circuits because it allows the researcher to reconstruct the natural system incrementally from the ground up. By adding one component at a time to the synthetic version of the native circuit, the researcher can (as with modeling) manage the complexity of the system, this time by comparing the *in vivo* behavior of circuits that differ from one another by only a single genetic change. Additionally, synthetic circuits allow both the magnitude and timing of the expression of genes to be precisely controlled by the researcher (*e.g.* via inducible promoters); this control enables

the researcher to measure the dynamic behavior of and dose/response relationships between the involved genetic components [14,15].

1.4 Why RNA regulatory networks?

In recent years, it has become increasingly clear that *trans*-acting RNAs play a far bigger role in regulating gene expression than was originally thought; small regulatory RNAs (srRNAs) have now been identified in organisms spanning all kingdoms of life [16]. Many of these srRNAs act by either interfering with or promoting the translation of targeted messenger RNAs (mRNAs). Additionally, a large proportion of srRNAs have multiple mRNA targets and/or make use of a shared protein complex in order to execute their function (*e.g.* Hfq-dependent small RNAs (sRNAs) which share the Hfq chaperone, the CsrB and CsrC non-coding RNAs which compete with mRNAs for CsrA binding or microRNAs (miRNAs) which incorporate into "RNA-induced silencing complexes" (RISCs)) [16,17]. Because of the widespread sharing of substrates and mediating complexes, srRNA regulatory circuits become well-connected 'networks' and thus gene circuits composed of srRNAs, their target mRNAs and their mediating proteins are referred to here as "RNA regulatory networks". While many *trans*-acting srRNAs act upon their targets via complementary RNA-RNA base pairing (*e.g.* Hfq-dependent sRNAs, miRNAs) [18], some act via sequestration of a shared regulatory protein (*e.g.* CsrB, CsrC) [19] among other mechanisms. Because the general importance of RNA regulatory networks has only become clear recently, many of the functional properties and design principles of these networks remain to be characterized; this lack of characterization, in conjunction with the significant role these systems play in gene regulation throughout the biosphere makes RNA regulatory networks attractive systems for reverse engineering.

1.5 The specific RNA networks and signaling properties investigated

The studies presented in this thesis investigate the signaling properties of two specific RNA regulatory networks in *Escherichia coli*: the Hfq-dependent small RNA network and the carbon storage regulatory (Csr) network (also known as the CsrA network). The specific physiology of these networks is described below. This thesis discusses the dynamics of signaling in these networks as well as the and the susceptibility of these networks to crosstalk among parallel srRNA pathways [14,15,20-22]. Specifically, this

thesis includes studies of signaling dynamics in the CsrA regulatory network [14] as well as crosstalk within both the Hfq-dependent small RNA network [20] and the CsrA regulatory network [15]. The dynamics of signaling within the Hfq-dependent small RNA network has been characterized previously [22] so those results are summarized.

1.5.1 THE HFQ-DEPENDENT SMALL RNA NETWORK

Hfq-dependent sRNAs play an important role in regulating bacterial virulence and stress responses [23]. This importance is demonstrated in part by how significantly the gene encoding the Hfq protein is conserved among bacterial species: putative *hfq* homologs have been identified in approximately 50% of all sequenced bacteria [17,24]. Hfq-dependent sRNAs regulate their mRNA targets by base paring with segments of complementary sequence. Most often, the binding of a sRNA silences translation of one or more target mRNAs (*e.g.* MicC silences *ompC* [25]; RyhB silences *sodB, ftnA, bfR, acnA, fumA and sdhCDAB* [26-28]), however, several sRNAs are known to activate translation of their mRNA targets (*e.g.* DsrA activates *rpoS* [29]; GlmZ activates *glmS* [28]). The Hfq protein complex (which is a hexamer of individual Hfq proteins) serves as a chaperone to assist in the annealing between sRNAs and their target mRNA [17]. As several experimental studies have shown, the various sRNA and mRNA molecules that depend upon Hfq for annealing can compete with one another for access to Hfq, disrupting one another's activity [21,30,31].

1.5.2 THE CsrA REGULATORY NETWORK

Like Hfq-dependent sRNAs, the CsrA regulatory network plays an important role in regulating bacterial virulence and stress; additionally, the CsrA system controls genes integral to cell metabolism, motility, and biofilm formation [19]. As with *hfq*, *csrA* homologs (*e.g. rsmA*) exist in a large subset of bacteria, including gram negative and gram positive species [19,32,33]. The CsrA protein regulates its mRNA targets by binding to specific sequences within the target mRNAs that form small hairpin loops. Typically, these loops are found within the 5' untranslated region (UTR) of the mRNA, allowing the CsrA protein to block the ribosomal binding site of the downstream gene; this occlusion silences translation of the mRNA (*e.g.* CsrA silences *glgCAP*, *pgaABCD* [19]). More rarely, CsrA has been observed to increase translation of its target mRNAs upon binding by blocking RNAse E access to the mRNA (*e.g.* CsrA activates *flhDC* [34]).

1.5.3 SIGNALING DYNAMICS

Cells can exhibit a broad range of dynamic responses to changes in their environment. Often, a cell's survival will depend upon rapid adaptation to an environmental change (*e.g.* when a cell encounters a source of physiological stress). Alternatively, some environmental variations are met with a more delayed response so that the cell does not overreact to what may only be a transient change [35,36]. Additionally, some cell pathways respond only to quick changes in the environment while averaging out slower changes; this form of adaptive (or acclimating) response is useful in bacterial chemotaxis where cells need to detect any nutrient gradient as they move [37-40] and when differentiating between signals that occur on different time scales but that are transmitted using sharing signaling components [41].

Because both the CsrA network and the Hfq-dependent sRNA network are intimately associated with stress response pathways, this thesis focuses on how genetic circuits can provide rapid responses to changes in a cell's environment (thus allowing for quick adaptation) [14]. It should be noted, however, that the CsrA system exhibits the potential for both rapid and delayed signaling depending upon how a particular signal is processed; this affords the CsrA system flexibility in terms of how quickly its targets respond to environmental cues [14]. How the dynamic response of the CsrA system varies across different environmental contingencies remains to be fully characterized.

1.5.4 SIGNALING CROSSTALK

As mentioned above, RNA regulatory networks contain many shared components either in the form of shared regulators (*e.g.* CsrA, RyhB) or chaperones (*e.g.* Hfq). Because these shared regulators control a pool of target mRNAs that is variable in both size and composition, the transcription of individual target mRNAs has the potential to influence the translation of other target mRNAs through competition for the shared molecule. This example of "substrate availability" crosstalk [42], while not unique to RNA and translational regulatory networks, is a fundamental concern for translational regulation since even the translational machinery itself (*e.g.* ribosomes) can be competed for by mRNAs [43]. Transcriptional regulation (and thus transcriptional regulatory networks) does not face this problem since the copy number of individual genes in the cell (*i.e.* the DNA substrate to which transcription factors bind) does not vary to the

degree that the copy number of mRNAs does. In this thesis, we examine crosstalk in both the Hfq-dependent sRNA network [20] and the CsrA regulatory network [15].

2. Signaling dynamics in the CsrA system

Biological mechanisms of signal transduction can operate on a variety of timescales. Some mechanisms are fast-acting; for example, signaling via small molecule binding to cytoplasmic proteins (*e.g.* transcription factors such as LacI and TetR which change their activity in the presence of sugars or antibiotics) or via transmembrane receptors that trigger the covalent modification of downstream effectors (*e.g.* two component systems and phosphorelays such as BarA/UvrY and the MAP kinase cascade) can occur in a few seconds. In contrast, other mechanisms act much more slowly; for example, endocrine signaling in mammals. Because hormones must be produced in one part of the body, be transported by the circulatory system, and then accumulate in another part of the body signaling can take hours or days. While typically slower than small molecule and phosphorelay signaling mechanisms, genetic circuits that operate within single cells tend to lie on the faster end of the signaling spectrum. Their speed allows them to be used in a reactive way: to allow an organism to adapt quickly to a changing environment. Signaling via genetic circuits is not always fast, however, and the factors that control the speed of signaling within gene circuits are the subject of active study.

A previous study by Razika Hussein and Han N. Lim examined the dynamics of signaling via Hfq-dependent sRNAs [22]. As a control, the dynamics of sRNA signaling was compared with and contrasted against the dynamics of signaling via protein transcription factors. It had been postulated that inducing transcription of a sRNA would illicit a more rapid response from a downstream target than would inducing a transcription factor because (i) the sRNA need only be transcribed before it is able to regulate its target gene (while transcription factors must also be translated and folded) and (ii) because the sRNA can block translation from and promote the clearance of existing target gene mRNA (while the transcription factor can only prevent the production more mRNA from the target gene) [22]. As was shown, these described differences are in fact negligible because mRNA half-lives are already quite small and the time delay between starting transcription and starting translation is small when compared to the time required for a molecule to reach a steady-state concentration. Instead, the biggest difference in signaling dynamics between sRNAs and transcription factors arises when their production is halted. Because transcription factors are far more stable than sRNAs, it can take much longer for them to be cleared and for their influence to diminish once their expression has been turned off. This effect is an

important factor governing the dynamics of the CsrA regulatory system which we discuss in more detail in the study below.

2.1 ABSTRACT

Bacterial survival requires the rapid propagation of signals through gene networks during stress but how this is achieved is not well understood. This study systematically characterizes the signaling dynamics of a cascade of RNA-protein interactions in the CsrA system, which regulates stress responses and biofilm formation in *Escherichia coli*. Non-coding RNAs are at the center of the CsrA system; target *mRNAs* are bound by CsrA *proteins* that inhibit their translation, CsrA *proteins* are sequestered by CsrB non-coding *RNAs*, and the degradation of CsrB *RNAs* is increased by CsrD *proteins*. Here we show using *in vivo* experiments and quantitative modeling that the CsrA system integrates three strategies to achieve rapid and robust signaling. These strategies include: (i) the sequestration of stable proteins by non-coding RNAs, which rapidly inactivates protein activity; (ii) the degradation of stable non-coding RNAs, which enables their rapid removal; and (iii) a negative feedback loop created by CsrA repression of CsrD production, which reduces the time for the system to achieve steady state. We also demonstrate that sequestration in the CsrA system results in signaling that is robust to growth rates because it does not rely on the slow dilution of molecules via cell division; therefore signaling can occur even during growth arrest induced by starvation or antibiotic treatment.

2.2 INTRODUCTION

A quantitative understanding of signaling dynamics is critical to determining how bacteria adapt to sudden environmental changes, combating pathogenesis, and designing synthetic circuits with specific dynamic properties. Recent studies have characterized signaling by transcription factor proteins [22,44-46] and by small RNAs that bind to target mRNAs to modulate their translation and/or degradation [22,46-48]. However, the signaling properties of non-coding RNAs that sequester proteins have not been defined and are of particular interest because of their theoretical potential for very rapid signaling (see **Results**).

In this study we chose the CsrA system in *Escherichia coli* as a model to investigate signaling by protein-sequestering non-coding RNAs (**Fig. 2.1**). CsrA regulation is

important for carbon storage, motility, biofilm formation and pathogenesis, and is evolutionarily conserved across distant groups of bacteria [19,32,33]. The system consists of CsrA protein, CsrB and CsrC non-coding RNAs, and CsrD protein. CsrA binds as a dimer to the 5′ untranslated region of target mRNAs to silence their translation [19,32,33]. CsrB and CsrC, which have binding sites for 9 and 3-4 CsrA dimers respectively [49], sequester CsrA to prevent it from silencing target mRNA translation [49,50]. CsrB and CsrC concentrations are in turn regulated by CsrD which acts as a specificity factor to increase their degradation by RNase E [51].

The study has three major parts. The first part describes the dynamics of signaling in a synthetic CsrA cascade without native control mechanisms in response to turning on and off transcription at each level in the cascade (*target gene*, *csrA*, *csrB* and *csrD*). Our model and experiments show that sequestration of CsrA and degradation of CsrB enable rapid signaling by eliminating the need for multiple generations of cell division to dilute out these stable molecules. The second part describes how the sequestration of CsrA by CsrB enables signaling to occur in cells with growth arrest caused by starvation and antibiotic induced stress. The third part describes the systematic reintroduction of wild-type components with their native transcriptional and translational regulatory sequences into the CsrA cascade. These experiments show that negative feedback control enables signaling in the native system to be even faster than in the synthetic system. Together these findings highlight general strategies for rapid intracellular signaling that are important for reprogramming gene expression during stress.

2.3 RESULTS

2.3.1 Modeling summary

We constructed a simple and general mathematical model to qualitatively predict and interpret how turning on and off the production of CsrA, CsrB and CsrD affect the dynamics of target gene expression. Briefly, the model uses ordinary differential equations (details below) to describe the production, clearance, association, dissociation and/or catalytic activity of the target mRNA, target protein, CsrA, CsrB, CsrD and their complexes (**Fig. 2.1**). For simplicity, all binding reactions are independent and the production of CsrA dimers and CsrD occur in a single step. CsrA primarily exists as a dimer in solution [52] therefore dimerization is presumably rapid compared to the CsrA turnover rate. Inclusion of separate transcription, translation and dimerization steps is

unnecessary since these reactions are assumed to be fast relative to the overall dynamics. The model is later extended to include feedback regulation.

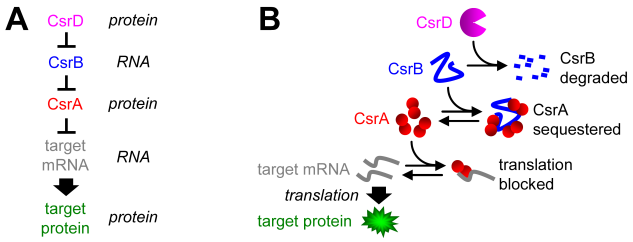

Fig. 2.1 | CsrA system. (***A***) Simplified schematic of the synthetic CsrA system. (***B***) Mechanistic description of the synthetic CsrA system (see text).

Parameter values were obtained from our data or the literature (see below). In the synthetic CsrA systems (**Figs. 2.2**, **2.4 & 2.5**), the only parameters fitted to our dynamics data were the production rates. Each production rate was fitted in a circuit where it was the only free parameter and once fitted, the parameter was constant across all simulations. In addition, we validated the model and parameter values by predicting the effect of steady state CsrA, CsrB and CsrD concentrations ([CsrA], [CsrB], [CsrD] respectively) on target protein (GlgC-GFP) expression and then confirmed the predictions qualitatively with *in vivo* measurements (**Figs. 2.2*D*, 2.4*D* & 2.5*D***). In the native CsrA system, parameter values for feedback and saturation of CsrD activity were also obtained from the dynamics data.

2.3.2 Experimental system and signaling metrics

To test the model's predictions, signaling was initially measured in a completely synthetic CsrA system. In this system, CsrA binding sites and flanking sequences from *glgC* (-61 to +8 nucleotides relative to the start codon) [53] were fused to *gfp*, thereby enabling GlgC-GFP expression ("target expression") to be quantified by fluorescence. The transcription of each component (*glgC-gfp*, *csrA*, *csrB* and/or *csrD*) was controlled by an inducible promoter (PLlacO-1 or PLtetO-1) [54] or a constitutive promoter (variants of Pcon/O3 [55]). CsrB and CsrC are believed to behave similarly therefore only the more potent (CsrB) was used. These synthetic circuits were constructed on plasmids and transformed into strains with chromosomal *csrA*, *csrB*, *csrC*, *csrD*, *glgCAP* and/or *pgaABCD* deleted. Deletion of *glgCAP* was necessary for *csrA* knockouts to survive [56] and deletion of *pgaABCD* was required to prevent the overproduction of biofilm adhesins [57] and to enable efficient transformation.

Table 2.1 | Synthetic circuit signaling times

figure	*experiment*	τ_{20} *(min)*	τ_{50} *(min)*	τ_{80} *(min)*
Fig. 2.2*B*	*glgC-gfp* turned on	39 ± 1	97 ± 1	>240
(turning target on)	*csrA* turned off	187 ± 1	>240	>240
Fig. 2.4*B*	*glgC-gfp* turned on	27 ± 1	74 ± 1	196 ± 3
(turning target on)	*csrB* turned on	26 ± 1	81 ± 2	>240
Fig. 2.5*B*	*glgC-gfp* turned on	28 ± 1	72 ± 2	159 ± 6
(turning target on)	*csrD* turned off	77 ± 18	158 ± 5	>240
Fig. 2.8*G*	*csrA* turned off	190 ± 9	>240	>240
(turning target on)	*csrB* turned on	44 ± 9	119 ± 18	>240
Fig. 2.2*C*	*glgC-gfp* turned off	90 ± 1	47 ± 1	9 ± 1
(turning target off)	*csrA* turned on	103 ± 3	49 ± 1	8 ± 3
Fig. 2.4*C*	*glgC-gfp* turned off	89 ± 2	46 ± 2	13 ± 1
(turning target off)	*csrB* turned off	147 ± 4	97 ± 9	67 ± 3
Fig. 2.5*C*	*glgC-gfp* turned off	78 ± 1	41 ± 1	14 ± 1
(turning target off)	*csrD* turned on	88 ± 7	52 ± 4	17 ± 2
Fig. 2.8*H*	*csrB* turned off	103 ± 1	67 ± 2	42 ± 4
(turning target off)	*csrD* turned on	83 ± 3	42 ± 12	16 ± 19
Fig. 2.9*C*	*glgC-gfp* turned off	85 ± 7	57 ± 6	35 ± 5
(turning target off)	*csrA* turned on	86 ± 1	54 ± 2	33 ± 2
Fig. 2.9*D*	*glgC-gfp* turned off	44 ± 3	27 ± 2	14 ± 3
(turning target off)	*csrB* turned off	>90	70 ± 7	58 ± 6
Fig. 2.9*E*	*glgC-gfp* turned off	55 ± 5	35 ± 6	18 ± 8
(turning target off)	*csrD* turned on	65 ± 4	43 ± 6	25 ± 8

Table 2.1 | Synthetic circuit signaling times. The τ_{20}, τ_{50} and τ_{80} values for turning on and off target gene expression (calculated from normalized dynamics data shown in **Figs. 2.8 & 2.9**). Uncertainty represents the s.e.m. of duplicate measurements.

Dynamics experiments were performed by turning transcription on or off for each component in the synthetic cascade and measuring GlgC-GFP expression at regular intervals. The rate of change in target expression reflects the convolved effects of target protein clearance, the difference between the initial and final steady states for the target protein, and the time required for the CsrA cascade upstream of the target protein to reach equilibrium [22]. Since the target protein degradation rate is constant and we rescaled the initial and final steady state of the target expression so the dynamic range is the same for all experiments, any observed differences in the dynamics are due to the delay in the CsrA cascade reaching equilibrium. We quantified the delay by measuring the time to reach 20%, 50% and 80% of the maximum GFP level (τ_{20}, τ_{50} and τ_{80}) (**Table 2.1 & Fig. 2.8**). Because target expression was often slow to turn off or on (and therefore

did not fall or rise to 50% of maximum expression in the experimental timeframe), τ_{80} or τ_{20} respectively were used to quantify increases in target expression.

We compared time delays using stable GFP (with constant and predictable clearance by dilution) to destabilized GFP (with a tail specific degradation tag [58] that decreases its half-life by up to 3.8 ± 0.5 fold) (**Fig. 2.9**). We chose not to use the latter because the time delays not only measure the delay in the CsrA system but also include saturation effects for active degradation (see below). We also found that destabilized GFP did not appreciably improve the resolution of time delays, and measurements were less reproducible than with stable GFP.

2.3.3 CsrA signaling: Stable signaling molecules can cause delays

In our first set of experiments, we turned on target expression *directly* by turning on *glgC-gfp* transcription (with *csrA* transcription kept off) or *indirectly* by turning off *csrA* transcription (with *glgC-gfp* transcription kept on) (**Fig. 2.2*A* & *B***). The model predicts that turning on transcription of *glgC-gfp* mRNA will increase target expression after a very short period whereas a long delay will occur between turning off *csrA* transcription and a significant increase in target expression. The delay will occur for two major reasons: (i) clearance of a stable protein such as CsrA occurs slowly via dilution due to cell growth [59] (**Fig. 2.10*C***); and (ii) the transfer function reveals a high initial CsrA concentration at 100% transcription that must first be cleared before a significant increase in target expression can occur. That is, when moving from high to low CsrA concentration (open to solid circle) there is minimal effect on target expression until the CsrA level is quite low (shaded region) (**Fig. 2.2*D***). We tested the predictions in our experimental system (**Fig. 2.2*B***) and confirmed that target expression took longer to rise to 20% of maximum when *csrA* transcription was turned off compared to when *glgC-gfp* transcription was turned on (τ_{20} = 187 ± 1 and 39 ± 1 min respectively; **Table 2.1 & Fig. 2.8**).

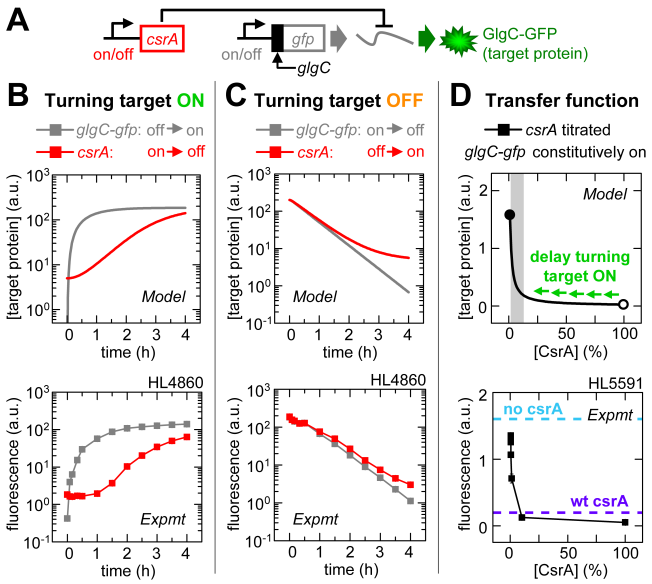

Fig. 2.2 | CsrA signaling: Stable signaling molecules can cause delays. Error bars are s.e.m. of duplicate measurements. (***A***) Experimental schematic. (***B***) Target expression turned on *directly* by turning on *glgC-gfp* transcription (IPTG added) or *indirectly* by turning off *csrA* transcription (aTc removed). (***C***) Target expression turned off *directly* by turning off *glgC-gfp* transcription (IPTG removed) or *indirectly* by turning on *csrA* transcription (aTc added). (***D***) GlgC-GFP expression as a function of % maximum *csrA* transcription (calibrated using PLlacO-1:st7:*gfp*; **Fig. 2.8*Q***). Target expression was also measured in strains without *csrA* (HL5594; cyan dash line) or native (wt) *csrA* (HL5562 and HL5596; purple dash line indicates both as the data overlay). Gray shading indicates the range over which the CsrA concentration has a significant effect on target expression ("regulatory range"). Open and closed circles are 100% and 0% of maximum [CsrA] respectively.

We next turned off target expression *directly* by turning off *glgC-gfp* transcription (with *csrA* transcription kept off) or *indirectly* by turning on *csrA* transcription (with *glgC-gfp* transcription kept on) (**Fig. 2.2*A* & *C***). The model predicts similar signaling delays for the two mechanisms (note: the curves eventually diverge because they have different steady-states). Turning off *glgC-gfp* transcription causes target expression to fall almost immediately because pre-existing target mRNAs are quickly degraded (**Fig. 2.10*B***). Turning on *csrA* transcription quickly decreases target expression despite the signal having to propagate through an extra regulatory step. This is because a small amount of CsrA is sufficient to silence most of the *glgC-gfp* mRNA (due to high affinity binding [53] and the stability of CsrA) as shown by the transfer function (**Fig. 2.2*D***). We confirmed the predictions experimentally; turning off *glgC-gfp* transcription and turning on *csrA* transcription caused target expression to fall to 80% of maximum on similar timescales (τ_{80} = 9 ± 1 and 8 ± 3 min respectively; **Table 2.1 & Fig. 2.8**).

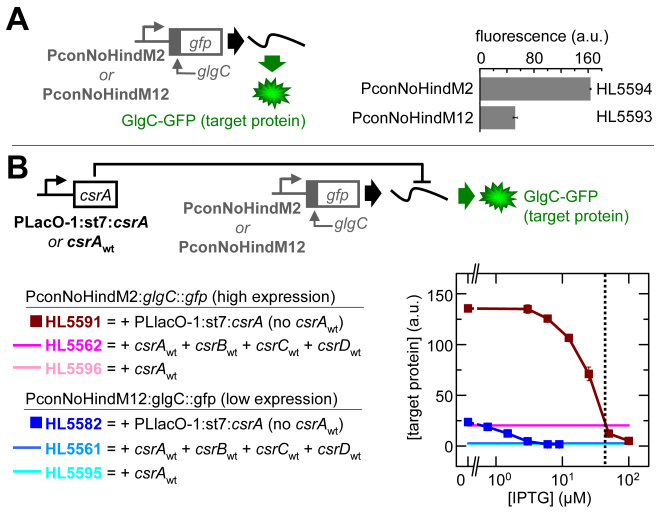

Fig. 2.3 | Effect of target mRNA concentrations on CsrA silencing. Error bars are s.e.m. of duplicate measurements. (***A***) A comparison of the expression of the target mRNA reporter under two different promoters; one has a constitutively high rate of target mRNA transcription (PconNoHindM2) and the other has a lower rate of target mRNA transcription (PconNoHindM12). (***B***) A comparison of the effect of the two different levels of target mRNA expression on the measurement of the CsrA transfer function and native CsrA concentrations. To obtain the transfer functions we varied the level of induction of the PLlacO-1:st7:*csrA* gene and measured its effect on target protein expression with the PconNoHindM2:*glgC-gfp* reporter or the PconNoHindM12:*glgC-gfp* reporter. To assess the effect of the native *csrA* ("$csrA_{wt}$") we measured the expression of both reporters in strains with the complete native *csrA* system (+ $csrA_{wt}$ + $csrB_{wt}$ + $csrC_{wt}$ + $csrD_{wt}$) or with only native *csrA* (+ $csrA_{wt}$). The synthetic system with PLlacO-1:st7:*csrA* and high levels of target mRNA transcription (PconNoHindM2:*glgC-gfp*) achieved an amount of target protein expression that was similar to the native *csrA* at 25 - 50 μM IPTG (black dot line), which is ~10% of the maximum CsrA that we induced (**Fig. 2.2*D***). In contrast, with the synthetic system with PLlacO-1:st7:*csrA* and lower levels of target mRNA transcription (PconNoHindM12:*glgC-gfp*), we could not accurately determine native CsrA levels because the lower quantity of target mRNA meant that it was completely silenced by CsrA at concentrations less than the native concentration (*i.e.* it was fully silenced at <10 μM IPTG).

We measured the level of target gene expression in a strain with native *csrA* instead of synthetic *csrA* (purple dash line, **Fig. 2.2*D***) and this indicated that physiological levels of CsrA are close to but not quite at a saturating concentration. That is, for target mRNAs at high concentrations, the native CsrA concentration minimizes delays in signaling without a significant trade-off in effectiveness (*i.e.* if CsrA levels were reduced too much then the dynamic range of CsrA activity would be severely diminished). These findings may not only apply to highly transcribed individual target mRNAs, but also to target mRNAs that are transcribed concurrently with many others (*i.e.* when the total target mRNA pool is large due to a global transcriptional response to stress). However in the case of target mRNAs that are at low concentrations, native CsrA concentrations are saturating (**Fig. 2.3**) and this is expected to cause long signaling delays when CsrA is removed solely by dilution (hence the importance of CsrB; see below).

In summary, our model and experiments demonstrate that stable proteins such as CsrA can introduce long delays when signaling depends on their removal; this is consistent with predictions [48,60] and experiments in other systems [22,61].

2.3.4 CsrB signaling: Sequestration can bypass downstream delays

In our second set of experiments, we turned on target expression *directly* by turning on *glgC-gfp* transcription (with *csrA* and *csrB* transcription kept on) or *indirectly* by turning on *csrB* transcription (with *glgC-gfp* and *csrA* transcription kept on) (**Fig. 2.4*A* & *B***). The latter decreases free CsrA levels to increase *glgC-gfp* mRNA translation. Our model predicts similar signaling times for turning on target expression via these two mechanisms. The transfer function shows that CsrB can quickly reach a level that is sufficient to sequester CsrA away from its target mRNAs (**Fig. 2.4*D***) due to its rapid production, long half-life [51] and multiple CsrA binding sites (**Fig. 2.10*E***). We experimentally confirmed that turning on *glgC-gfp* transcription and turning on *csrB* transcription caused target expression to rise to 20% of maximum in comparable periods (τ_{20} = 27 ± 1 and 26 ± 1 min respectively; **Table 2.1 & Fig. 2.8**).

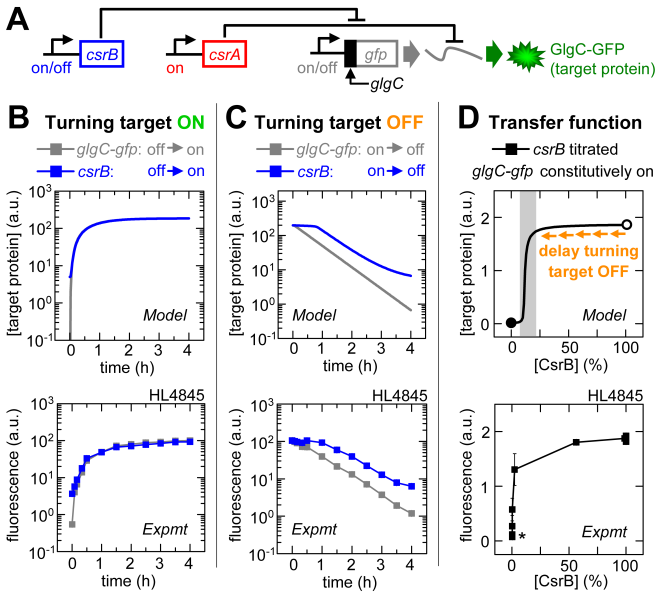

Fig. 2.4 | CsrB signaling: Sequestration can bypass downstream delays. Error bars are s.e.m. of duplicate measurements. (***A***) Experimental schematic. (***B***) Target expression turned on *directly* by turning on *glgC-gfp* transcription (IPTG added) or *indirectly* by turning on *csrB* transcription (aTc added). (***C***) Target expression turned off *directly* by turning off *glgC-gfp* transcription (IPTG removed) or *indirectly* by turning off *csrB* transcription (aTc removed). (***D***) GlgC-GFP expression as a function of % maximum *csrB* transcription (calibrated using PLtetO-1:st7:*gfp*; **Fig. 2.8*R***). Gray shading indicates the regulatory range for CsrB. Open and closed circles are 100% and 0% of maximum [CsrB] respectively. *Incomplete silencing of GlgC-GFP expression occurs if the total [CsrA] is less than the total [target mRNA] or if there is "leaky" CsrB expression.

We next turned off target expression either *directly* by turning off *glgC-gfp* transcription (with *csrA* and *csrB* transcription kept on) or *indirectly* by turning off *csrB* transcription (with *glgC-gfp* and *csrA* transcription kept on) (**Fig. 2.4*A* & *C***). Our model predicts a significant delay in signaling for the latter. Turning off *csrB* transcription decreases sequestration of CsrA and consequently decreases *glgC-gfp* mRNA translation. The delay will occur for the same basic reasons that CsrA activity is slow to turn off: (i) CsrB is slowly cleared from the cell in the absence of CsrD [51]; and (ii) the CsrB concentration at 100% transcription corresponds to the saturating part of the transfer function (open circle, **Fig. 2.4*D***) and therefore most of this excess CsrB must be cleared before it reaches a concentration that significantly decreases target expression (shaded region, **Fig. 2.4*D***). Our *in vivo* experiments confirmed this prediction; turning off *glgC-gfp* transcription caused target expression to fall to 80% of maximum in less time than turning off *csrB* transcription (τ_{80} = 13 ± 1 and 67 ± 3 min respectively; **Table 2.1 & Fig. 2.8**).

Our model and experiments demonstrate that CsrB can rapidly sequester and turn off the activity of the downstream CsrA thereby bypassing delays due to slow CsrA clearance. However, clearance of CsrB is itself slow, therefore it can delay signal propagation when turned off.

2.3.5 CsrD signaling: Degradation can prevent downstream delays

In our third set of experiments we investigated signaling using CsrD, a specificity factor that decreases the CsrB half-life from >30 min to <2 min at wild-type CsrD levels [51]. We turned on target expression *directly* by turning on *glgC-gfp* transcription (with *csrA* and *csrB* transcription kept on) or *indirectly* by turning off *csrD* transcription (with *glgC-gfp*, *csrA* and *csrB* transcription kept on) (**Fig. 2.5*A* & *B***). The model predicts it will take longer to turn on target expression by turning off *csrD* transcription because: (i) CsrD is cleared slowly by dilution; and (ii) the initial CsrD concentration corresponds to the saturating part of the transfer function (open circle, **Fig. 2.5*D***) therefore most of it must be cleared before there will be a significant increase in target expression (shaded region, **Fig. 2.5*D***). Our experimental results agree with the model. Target expression took longer to increase to 20% of maximum when *csrD* transcription was turned off than when *glgC-gfp* transcription was turned on (τ_{20} = 77 ± 18 and 28 ± 1 min respectively; **Table 2.1 & Fig. 2.8**).

Part 2: Signaling dynamics in the CsrA system

We next turned off target expression *directly* by turning off *glgC-gfp* transcription (with *csrA* and *csrB* transcription kept on) or *indirectly* by turning on *csrD* transcription (with *glgC-gfp*, *csrA* and *csrB* transcription kept on) (**Fig. 2.5*A* & *C***). The model predicts that turning on *csrD* transcription will have an almost immediate effect on target expression. This is because a very small amount of CsrD is sufficient to dramatically increase CsrB degradation and consequently decrease CsrA sequestration and target mRNA translation (**Fig. 2.5*D***). Our *in vivo* experiments confirmed this prediction. Turning on *csrD* transcription and turning off *glgC-gfp* transcription caused target expression to fall to 80% of maximum after a similar delay (τ_{80} = 17 ± 2 and 14 ± 1 min respectively; **Table 2.1 & Fig. 2.8**).

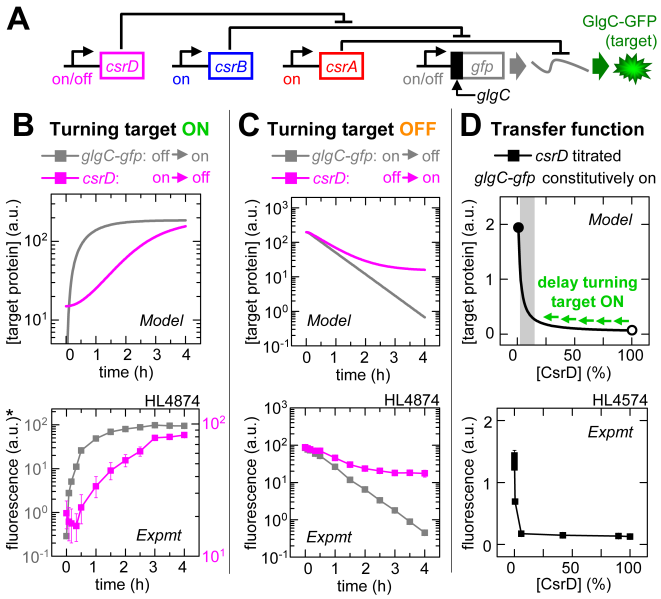

Fig. 2.5 | CsrD signaling: Degradation can prevent downstream delays. Error bars are s.e.m. of duplicate measurements. (***A***) Experimental schematic. (***B***) Target expression turned on *directly* by turning on *glgC-gfp* transcription (IPTG added) or *indirectly* by turning off *csrD* transcription (aTc removed). *Left and right y-axis correspond to turning on *glgC-gfp* (gray) and turning off *csrD* (magenta) respectively. (***C***) Target expression turned off *directly* by turning off *glgC-gfp* transcription (IPTG removed) or *indirectly* by turning on *csrD* transcription (aTc added). (***D***) GlgC-GFP expression as a function of % maximum *csrD* transcription (calibrated using PLlacO-1:st7:*gfp*; **Fig. 2.8*Q***). Gray shading indicates the regulatory range for CsrD. Open and closed circles are 100% and 0% of the maximum [CsrD] respectively.

2.3.6 Faster signaling is possible in longer cascades

We have shown that turning on *csrB* transcription caused target expression to turn on faster than turning off *csrA* transcription (**Fig. 2.2*B*** and **Fig. 2.4*B***). Similarly, turning on *csrD* transcription caused target expression to turn off faster than turning off *csrB* transcription (**Fig. 2.4*C*** and **Fig. 2.5*C***). These results show, somewhat counterintuitively,

that regulating an upstream molecule can alter target expression more rapidly than regulating a downstream molecule under some conditions (*i.e.* when the shorter cascade requires the slow clearance of a stable molecule and the longer cascade does not). Because these measurements were in different strains, we constructed circuits to directly compare signaling between CsrA and CsrB, or between CsrB and CsrD in the same strain (**Fig. 2.8*G* & *H***); these measurements confirmed our earlier findings. Therefore the CsrA cascade differs from purely transcriptional cascades in that the number of regulatory connections in the cascade does not predict signaling delay [45,60].

2.3.7 Robust signaling during stress

Rapid cell signaling is most crucial when cells are exposed to potentially lethal environmental stress. However, stress often causes growth arrest which prevents the clearance of stable molecules by dilution causing signaling lock-up. In theory, signaling by sequestration bypasses this problem. To test this proposal we attempted to turn on target expression by turning off *csrA* transcription and turning on *csrB* transcription during starvation (M9 media without a carbon source) or in the presence of antibiotics for which the cells were not resistant.

We tested several classes of antibiotics including: sulfamethoxazole (100-500 μg/mL) and trimethoprim (5 μg/mL) which disrupt folate synthesis; novobiocin (200-3000 μg/mL) and norfloxacin (12.5-250 ng/mL) which inhibit DNA gyrase; and polymyxin B (0.25-50 μg/mL) which destabilizes the outer membrane. Trimethoprim (5ug/mL) and novobiocin (200mg/mL) in LB media caused significant growth arrest without rapid lysis. Doubling times in M9, novobiocin and trimethoprim were 700 ± 300, 400 ± 100 and 180 ± 30 min respectively (mean ± s.e.m.). In all three stress conditions, target expression could be turned on by inducing *csrB* transcription to sequester CsrA but not by turning off *csrA* transcription (**Fig. 2.6**).

These experiments with three independent sources of stress confirm the generality of our prediction that sequestration can be *essential* for signaling in pathways with stable signaling molecules during stress.

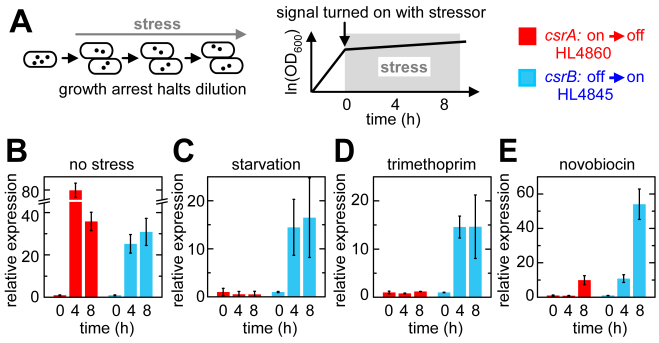

Fig. 2.6 | Robust signaling during stress. Error bars are s.e.m. of two or more measurements. (***A***) Experimental schematic. (***B-E***) Experiments performed using circuits shown in **Fig. 2.2*A*** and **Fig. 2.4*A*** without stress treatment (***B***), in M9 with no carbohydrate (***C***), in LB with 5μg/mL trimethoprim (***D***), or in LB with 200μg/mL novobiocin (***E***). In HL4860, *csrA* transcription was kept off (control) or turned off at t = 0. In HL4845, *csrB* transcription was kept on (control) or turned on at t = 0. Fluorescence levels were normalized to their respective control at each time point to correct for general effects of stress. The normalized values were rescaled to the initial measurement to determine the fold-change in expression ("relative expression").

2.3.8 Feedback in the native CsrA system

We next probed the architecture of the native CsrA cascade using the synthetic CsrA system as a benchmark. Specifically, we sought to determine whether putative transcriptional and translational feedback loops described in the native system affect signaling dynamics under our experimental conditions [32,62]. Feedback is known to influence signaling dynamics (1) and CsrA has been reported to: (i) positively and negatively regulate its own expression via mechanisms that have not been fully elucidated [62]; (ii) inhibit its own activity by increasing transcription of CsrB and CsrC [49,63]; and (iii) inhibit its own activity by decreasing CsrD production which increases CsrB [64].

In these experiments we systematically replaced components of the synthetic cascade with native genes in the chromosome that have intact regulatory sequences. We turned on and off synthetic *csrB* transcription (input) and measured the effect of the native gene(s) on the dynamics of target expression (output) (**Fig. 2.7*A***). Initially we started with all four native genes (*csrA*, *csrB*, *csrC*, *csrD*) present (**Fig. 2.7*B***), and remarkably we found that turning on *csrB* transcription with the entire native CsrA cascade caused target expression to turn on more rapidly than the synthetic cascade (**Fig. 2.7*B***). To isolate the regulatory interaction responsible for this "enhanced signaling" speed in the native CsrA cascade we incrementally replaced or removed native genes (**Fig. 2.7*C-F***).

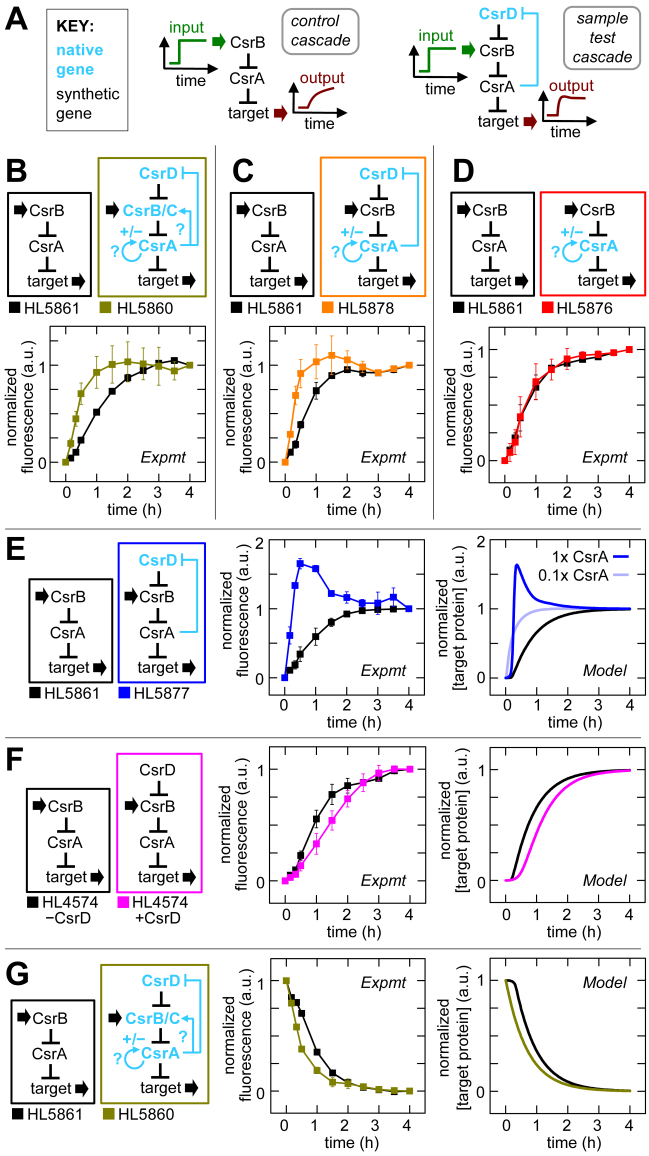

Fig. 2.7 | Feedback in the native CsrA system. Error bars are s.e.m. of duplicate measurements. (***A***) Experimental schematic with synthetic (black) and native genes with reported feedback loops (blue). The "synthetic cascade", which was a benchmark for comparison, was composed of synthetic *csrA* and *csrB*. Normalized fluorescence was determined by dividing each value by the fluorescence value in a control with *csrB* transcribed constitutively; the resulting ratio was rescaled so the start and end points were 0 and 1 respectively (**Fig. 2.8**). (***B-F***) Comparison of systems with synthetic and native genes where synthetic *csrB* was induced at t = 0. (***B***) Cascade with native *csrA*, *csrB*, *csrC* and *csrD*, and synthetic *csrB* (gold) *versus* synthetic cascade (black). (***C***) Cascade with native *csrA* and *csrD* (orange) *versus* synthetic cascade (black). (***D***) Cascade with native *csrA* (red) *versus* synthetic cascade (black). (***E***) Cascade with native *csrD* (blue) *versus* synthetic cascade (black). Native *csrD* is modeled at low (light blue) and high (dark blue) CsrA levels. (***F***) Cascade with and without synthetic *csrD* expression (magenta and black respectively). (***G***) Comparison of cascade with native *csrA*, *csrB*, *csrC* and *csrD*, and synthetic *csrB* (gold) *versus* synthetic cascade (black) where synthetic *csrB* was turned off at t = 0.

Part 2: Signaling dynamics in the CsrA system

We started by removing native *csrB* and *csrC* and found their removal did not eliminate enhanced signaling (**Fig. 2.7*C***). However, when we also removed native *csrD* enhanced signaling no longer occurred (**Fig. 2.7*D***). Therefore native *csrD* was necessary for enhanced signaling. Enhanced signaling was restored when native *csrD* was reinstated with synthetic *csrA* (**Fig. 2.7*E***). This latter strain also showed a transient overshoot of the steady state target expression. Given that enhanced signaling and overshoot are known to be generated by negative feedback under certain conditions [44] and that native *csrD* was necessary for these behaviors, our findings are consistent with the reported negative feedback regulation in which CsrA represses the production of CsrD. In further support of this, we demonstrated that enhanced signaling does not occur with the synthetic *csrD* gene which lacks the flanking sequences (including the native *csrD* promoter and 5′ UTR sequences) that are necessary for CsrA repression of CsrD production [51,64]. Instead we found that the induction of a small, constant amount of CsrD from synthetic *csrD* resulted in *slower* signaling than in the control cascade without CsrD induction (**Fig. 2.7*F***). Additionally, we showed by quantitative RT-PCR that the CsrD mRNA concentration decreases with increased CsrA production (*see below*), which is consistent with negative feedback and previous reports [51,64].

We used the model to determine whether negative feedback explains the enhanced signaling and overshoot observed. The model was modified to include repression of CsrD production by CsrA and the capacity for CsrD binding to be saturated by CsrB (see below). We found that an increase in *csrB* transcription is countered by negative feedback which decreases CsrB levels after a delay (**Fig. 2.10*J* & *K***). The net effect is a pulse of CsrB resulting in a brief surge in target mRNA translation that causes target expression to reach its final steady state faster (right panel, **Fig. 2.7*E***). If the delay in the negative feedback is sufficiently large, which depends on the total CsrA concentration, then target expression can briefly overshoot the new steady state (**Fig. 2.10*J* & *K***) [44]. This explains the overshoot with synthetic *csrA* and why it does not occur with native *csrA* that produces approximately one tenth the CsrA concentration (**Fig. 2.2*D*** and **Fig. 2.7*B***, ***C*** and ***E***). Negative feedback can also allow target expression to reach a new steady state sooner when *csrB* transcription is decreased. In this case, the response is faster because increased CsrD concentrations minimize the accumulation of saturating amounts of CsrB in the first place and increase the CsrB clearance rate (**Fig. 2.7*G***). The model also explains why slower signaling is observed with synthetic *csrD* (**Fig. 2.7*F***). Initially the CsrB concentration increases slowly because it is cleared by CsrD therefore target expression increases more slowly than in the control cascade (**Fig. 2.10*I* & *L***).

However as CsrB accumulates, CsrD becomes saturated causing the overall half-life of CsrB and the CsrB concentration to increase at a faster rate resulting in a rapid increase in target expression. That is, there is a slow initial increase in target expression followed by a fast increase with synthetic *csrD*, which is the opposite of native *csrD* with negative feedback.

In summary, characterization of the native CsrA system reveals that negative feedback (*i.e.* inhibition of CsrD production by CsrA) enhances the speed of turning on and off target protein expression.

2.3.9 RT-PCR Corroborates Inhibition of CsrD Production by CsrA

Above, we examined feedback by measuring the dynamic response of the CsrA system (**Fig. 2.7**) and determined that the only component that had a significant effect on the dynamics was CsrD. In particular the data indicated that the inhibition of CsrD expression by CsrA increased the speed of the system's response to changes in CsrB transcription ("enhanced signaling"). We also evaluated this feedback regulation of CsrD by CsrA under steady-state conditions.

We measured CsrD mRNA concentrations in parallel for two strains (HL 5944 and HL5947) by quantitative RT-PCR. One strain carried a gene that constitutively expressed synthetic *csrA* (HL5947) while a second strain contained a control plasmid without *csrA* (HL5944). In the strain expressing CsrA, CsrD mRNA levels were very low ($-9.28 \times 10^{-8} \pm 3.90 \times 10^{-7}$ a.u., 8 samples, uncertainty is the s.e.m.) and indistinguishable from background (*i.e.* plus RT measurements were the same as the minus RT control, which contains the harvested RNA but without reverse transcriptase). In contrast, the strain lacking CsrA had significantly elevated CsrD mRNA levels ($1.82 \times 10^{-6} \pm 8.13 \times 10^{-7}$ a.u., 8 samples, uncertainty is the s.e.m.). The difference between these two measurements is statistically significant (two-tailed t-test: $p = 0.04$, t-value = 2.26). We note that the variances of the values for the two strains were not found to be different (Levene's test: $p = 0.26$), so homoscedasticity was assumed for the t-test calculation. Our observed decrease in the *csrD* mRNA concentration in the presence of CsrA suggests that CsrD mRNA levels are directly or indirectly repressed by CsrA; this result is consistent with CsrD mRNA measurements from previous reports [51,64].

In summary, in this study we have demonstrated via two independent types of experiments (dynamics measurements and steady state RT-PCR) that CsrA represses

CsrD production, and these findings are in agreement with prior measurements by others [64].

2.4 DISCUSSION

In this study we reverse engineered and quantitatively modeled the CsrA regulatory network to obtain a detailed and coherent picture of its signaling dynamics. We found that it is capable of rapid signal propagation due to three general principles: (i) sequestration of stable signaling molecules; (ii) degradation of stable signaling molecules; and (iii) negative feedback regulation. Sequestration and degradation of CsrA and CsrB respectively avoids long delays due to the slow removal of these stable signaling molecules by dilution through cell division. Negative feedback regulation (due to CsrA inhibition of CsrD production) enables target expression to reach its new steady state faster in response to changes in *csrB* transcription (as explained above).

How might rapid signaling in the CsrA system aid adaptation to stress? Unfortunately, the pathways which sense environmental stresses and transmit this information to the CsrA system have not been fully elucidated. However, one regulator that has been identified to activate the CsrA cascade is the BarA/UvrY two-component system. Two component systems also control expression of CsrB and CsrC homologs in other bacteria (*e.g.* GacS/GacA in *Pseudomonas aeruginosa*, LetS/LetA in *Legionella pneumophila*, and VarS/VarA in *Vibrio cholerae*) [33]. Extracellular weak acids are thought to activate BarA which alters the phosphorylation state of UvrY; UvrY then interacts with the CsrA system by increasing *csrB* and *csrC* transcription [19,32,33]. As with other two component systems, the BarA/UvrY response to a signal is believed to be extremely rapid therefore propagation of the signal through the CsrA cascade is likely to be rate-limiting. As a consequence, increasing the speed of signaling through the CsrA system should directly shorten delays in adapting gene expression patterns and phenotypes to environmental perturbations. Our findings with stable and destabilized GFP show that shortening the time delay in the CsrA system benefits the dynamics of target proteins with short lifetimes as well as those with long lifetimes (**Fig. 2.9**).

There are many parallels between the CsrA system and the architecture of other bacterial stress response pathways. In the *Salmonella* ChiP and *E. coli* YbfM pathways, a small RNA silences target mRNA translation (ChiX and MicM respectively) in a manner analogous to CsrA silencing of mRNAs [65,66]. These small RNAs are themselves

sequestered by a sRNA or mRNA to turn off their activity (in a manner analogous to CsrB). In *Bacillus subtillus*, a CsrA homolog and AarB (another global regulator of biofilm formation) are also regulated by sequestration but by proteins (FliW and AbbA respectively) instead of non-coding RNAs [67,68]. In *E. coli*, the regulation of the extracytoplasmic stress sigma factor (σ^E) that controls the transcription of genes in stress response pathways also has features in common with the CsrA system [69]. σ^E is inactivated by sequestration to the membrane by the RseA protein (analogous to CsrA and CsrB respectively) and reactivated when RseA is cleaved by DegS and YeaL (analogous to CsrD) [69]. These examples highlight the general importance of sequestration and degradation in regulating the dynamics of stress response pathways.

The regulation of signaling molecules by sequestration and degradation not only increases signaling speed but it also makes signaling robust to growth rates. As we showed, this is important for preventing signaling lock-up when stresses such as starvation and antibiotics lead to growth arrest.

In conclusion, this study shows that CsrB, which is at the center of the CsrA network, synergistically integrates multiple mechanisms to achieve rapid and robust signaling. These findings provide further evidence that non-coding RNAs, which include small RNAs, have evolved a prominent role in connecting genetic pathways due to their general ability to rapidly propagate signals needed for prompt adaptation to stress.

2.5 METHODS

2.5.1 Gene Expression Measurement and Analysis

2.5.1.1 Cell growth protocols

Steady state experiments (*e.g.* measuring transfer functions) were performed by inoculating 5-50 mL of overnight culture in 5 mL of LB media with 100 μg/mL ampicillin and/or 50 μg/mL kanamycin. Cultures were grown for 3.5 h at 37°C and 200 rpm then 5 mL of culture was inoculated into 5 mL of fresh LB with antibiotics and IPTG (0.01 to 1 mM), aTc (0.01 to 1 μM), both or neither. Harvested cells were placed on ice and GFP expression was measured using a Coulter EPICS-XL flow cytometer (488nm/15mW argon ion laser) and analyzed as described [21]. GFP distributions were unimodal for all measurements included in this study. However, bimodality was

observed at very high free CsrA concentrations that exceeded levels presented in this study and caused aberrant morphologies consistent with severe cellular stress.

Dynamics experiments were performed as above except as follows. There was a pre-induction step (to achieve steady state on or off transcription) and a post-induction step (to reverse or maintain the transcription states of the pre-induction step). In the pre-induction step, cells were grown with IPTG (0.1 to 1 mM), aTc (1 μM), both or neither in the initial 3.5 h growth phase. After the initial growth phase, cells were inoculated into 2.5 mL of fresh media (with the same antibiotic and inducer concentrations) to produce an OD_{600} = 0.01 to 0.05. The culture was grown for 30 min, diluted 1 in 2 in the same fresh media, and grown for an additional 30 min. Post-induction cultures were created by taking cells from pre-induction cultures, centrifuging them (16, 100 × *g* for 1 min), discarding the supernatant, and resuspending the pellets in LB with the same antibiotics and IPTG (0.1 to 1 mM), aTc (1 μM), both or neither. The cultures were grown for 4 h and diluted ~ 1 in 2 every 30 min to maintain OD_{600} within a range. HL5877 was inoculated from glycerol stock instead of overnight cultures to prevent selection of a "locked-on" phenotype with constitutively high GlgC-GFP expression.

Growth arrest experiments were performed as the dynamics experiments except as follows. Pre-induction cultures were 10 mL and grown for 4 h with IPTG (1 mM), aTc (1 μM), both or neither. Post-induction cultures were created by resuspending pre-induction cells in M9 media with no carbohydrate or in LB with trimethoprim or novobiocin. IPTG, aTc, both or neither were added to the media. Cells were harvested after 4 and 8 h of stress.

Dynamics experiments using destabilized GFP (**Fig. 2.9*B-E***) were performed and measured as described for stable GFP except that 5-50 μM of overnight culture was inoculated into 10 mL of LB with 100 μg/mL ampicillin and 50 μg/mL kanamycin with IPTG (0.01 to 1 mM), aTc (1 μM), both or neither. Cells were grown for 3.5 hours, centrifuged (16,000 x g for 1 min), the supernatant discarded, and the pellets resuspended in 15 mL of fresh media to produce a final OD_{600} = 0.01 to 0.05 and a volume of ~15 mL. The culture was then grown for 90 minutes with cells harvested every 10 minutes. No serial dilution was necessary due to the shortened experimental timeframe.

2.5.1.2 Fluorescence data analysis

In the dynamics experiments we measure the change in target expression as a function of time after an input (*e.g. csrB* transcription) has been altered at t = 0. The change in gene expression reflects the convolved effects of: (i) the target protein degradation rate, (ii) the ratio of the initial and final steady states for the target protein (*i.e.* dynamic range), and (iii) the time required for the CsrA cascade upstream of the target protein to reach equilibrium [22]. We eliminate the first factor as a cause for any differences in the dynamics of different CsrA systems by only directly comparing systems that the same target protein. We eliminate the second factor by rescaling the difference between the initial and final steady states so that it is the same for all experiments; specifically we set the minimum and maximum relative steady state levels of target expression to zero and one respectively ("zero-to-one rescaling"). By eliminating the first and second factors as possible causes for differences in the time delay for different cascades we are left with a metric that primarily measures the time required for the CsrA cascade upstream of the target protein to reach equilibrium (**Figs. 2.2, 2.4 & 2.5** and **Fig. 2.8*A*-*H***).

In all of our dynamics experiments we had duplicate "physiological" control samples where target protein was constitutively expressed (*i.e.* GlgC-GFP was on before t = 0 and it was maintained in the on state for the duration of the dynamics experiment) (**Fig. 2.8**). These control samples are included to demonstrate that conditions are constant (*e.g.* cell density and inducer activity) throughout the experiment and that there are no changes in non-specific physiological factors that generally affect gene expression (*e.g.* ribosome concentration). Moreover, if there is a change in these general factors that affect gene expression we can correct for them using these control samples which are measured in parallel.

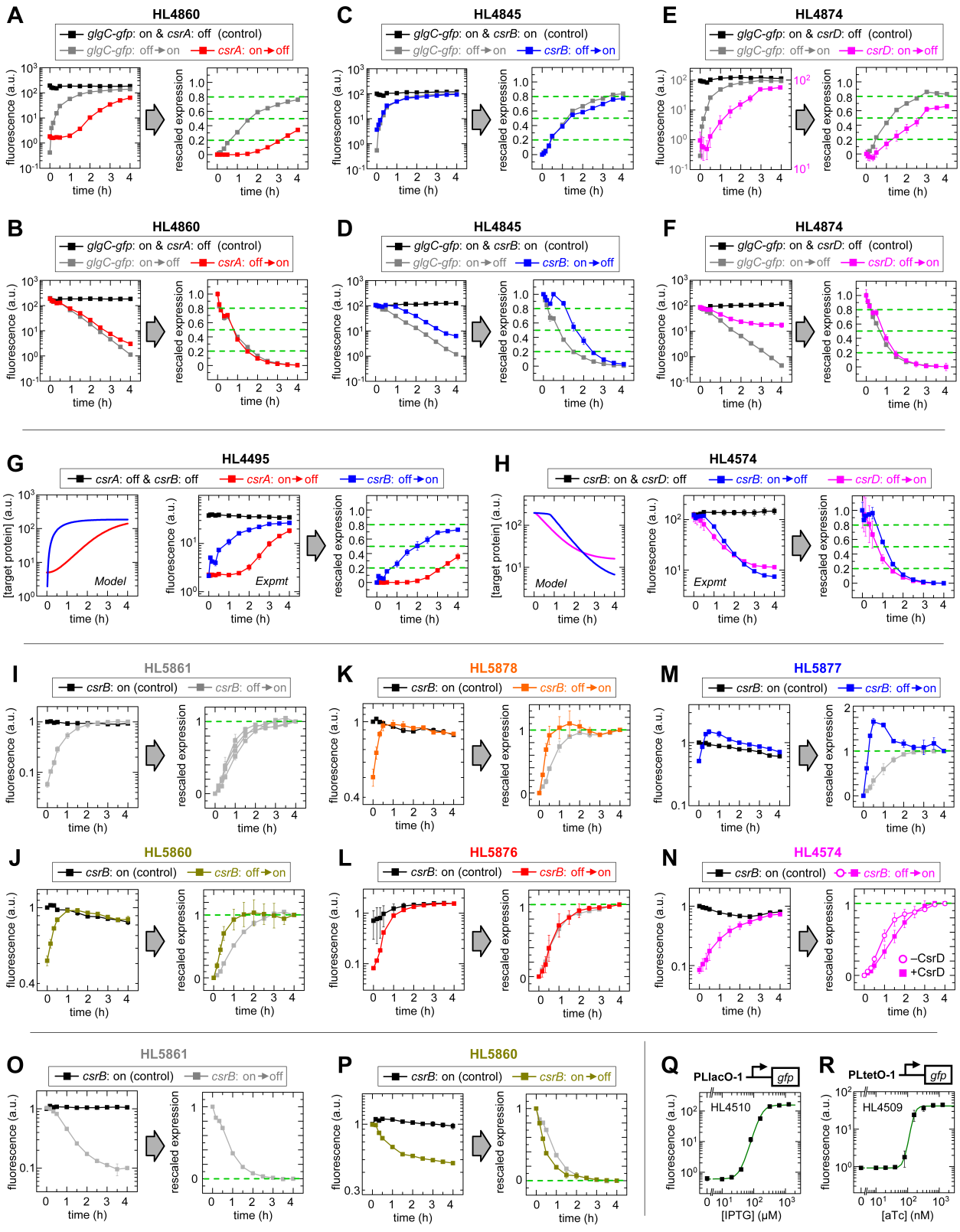
A
HL4860
glgC-gfp: on & csrA: off (control)
glgC-gfp: off→on
csrA: on→off
B
HL4860
glgC-gfp: on & csrA: off (control)
glgC-gfp: on→off
csrA: off→on
C
HL4845
glgC-gfp: on & csrB: on (control)
glgC-gfp: off→on
csrB: off→on
D
HL4845
glgC-gfp: on & csrB: on (control)
glgC-gfp: on→off
csrB: on→off
E
HL4874
glgC-gfp: on & csrD: off (control)
glgC-gfp: off→on
csrD: on→off
F
HL4874
glgC-gfp: on & csrD: off (control)
glgC-gfp: on→off
csrD: off→on
G
HL4495
csrA: off & csrB: off
csrA: on→off
csrB: off→on
[target protein] (a.u.)
Model
fluorescence (a.u.)
Expmt
rescaled expression
time (h)
H
HL4574
csrB: on & csrD: off
csrB: on→off
csrD: off→on
[target protein] (a.u.)
Model
fluorescence (a.u.)
Expmt
rescaled expression
time (h)
I
HL5861
csrB: on (control)
csrB: off→on
J
HL5860
csrB: on (control)
csrB: off→on
K
HL5878
csrB: on (control)
csrB: off→on
L
HL5876
csrB: on (control)
csrB: off→on
M
HL5877
csrB: on (control)
csrB: off→on
N
HL4574
csrB: on (control)
csrB: off→on
–CsrD
+CsrD
O
HL5861
csrB: on (control)
csrB: on→off
P
HL5860
csrB: on (control)
csrB: on→off
Q
PLlacO-1
gfp
HL4510
[IPTG] (μM)
R
PLtetO-1
gfp
HL4509
[aTc] (nM)
fluorescence (a.u.)
rescaled expression
time (h)

Fig. 2.8 | Normalization and analysis of *in vivo* dynamics and transfer function data. Error bars are s.e.m. of two or more measurements. (***A-F***) Normalization and analysis of data presented in **Figs. 2.2, 2.4 & 2.5**. Each left plot is raw data with the physiological control (black curve); each right plot is data from the left plot rescaled to lie between zero and one. τ_{20}, τ_{50} and τ_{80} values (green dash lines) were calculated as described in the text using linear interpolation from adjacent data points (**Table 2.1**). (***A***) **Fig. 2.2*B*** data. (***B***) **Fig. 2.2*C*** data. (***C***) **Fig. 2.4*B*** data. (***D***) **Fig. 2.4*C*** data. (***E***) **Fig. 2.5*B*** data. (***F***) **Fig. 2.5*C*** data. (***G, H***) Longer cascades can signal faster than shorter cascades (under some conditions) within the same strain. Model predictions, raw data and rescaled data are left, center and right respectively. (***G***) CsrA versus CsrB. (***H***) CsrB versus CsrD. (***I-P***) Normalization and analysis of data presented in **Fig. 2.7**. Each left plot is raw data from two replicates that were rescaled to eliminate constant offsets between them. The black curve is the physiological control. Each right plot is data from the left plot normalized 'time point-by-time point' by the physiological control and then rescaled so the initial and final concentrations are 0 and 1 or *vice versa*. (***I***) The right plot shows all the dynamics data for the synthetic cascade that was used as a benchmark for comparison in the right plot of panels ***J***, ***K***, ***L*** and ***M*** (grey curve) (note: the synthetic cascade is the same as the black curve shown in **Fig. 2.7*B***, ***C***, ***D*** & ***E*** respectively). (***J***) **Fig. 2.7*B*** data. (***K***) **Fig. 2.7*C*** data. (***L***) **Fig. 2.7*D*** data. (***M***) **Fig. 2.7*E*** data. (***N***) **Fig. 2.7*F*** data. (***O***) The right plot shows the dynamics data for the synthetic cascade that was used as a benchmark for comparison in the right plot of panel ***P*** (grey curve) (note: the synthetic cascade is the same as the black curve shown in **Fig. 2.7*G***). (***P***) **Fig. 2.7*G*** data. (***Q***, ***R***) PLlacO-1 (panel ***Q***) and PLtetO-1 (panel ***R***) induction with IPTG and aTc respectively. Green curves are Hill function fits used to calibrate the x-axis of the transfer functions in **Figs. 2.2, 2.4 & 2.5**.

In purely synthetic systems using stable GFP the physiological control samples showed only a small amount of "drift" over the course of the dynamics experiments and therefore there was no need to correct for it (**Fig. 2.8*A-H***). In systems with native genes (**Fig. 2.8*I-P***) or with destabilized GFP (**Fig. 2.9*C-E***), the drift was large compared to the dynamic range (*i.e.* ratio of maximum to minimum steady state expression) of the target protein. To correct for this drift we performed a 'point-to-point' normalization (prior to the zero-to-one rescaling) where the fluorescence value at each time point for each sample was normalized by the fluorescence value in the corresponding physiological control samples at the same time point. This point-to-point normalization was important for HL5860, HL5877 and HL5878 (**Fig. 2.8*J***, ***K*** & ***M***) to prevent drift from exaggerating the observed overshoot (**Fig. 2.7*B*, *C* & *E***).

2.5.1.3 Transfer function calibration

We collected our *in vivo* transfer function data by varying the amount of IPTG and aTc added to the media to induce different levels of expression of CsrA, CsrB and CsrD from the PLlacO-1 and PLtetO-1 promoters. We then measured the effect of this induction on target protein expression via GFP fluorescence (**Fig. 2.2*D***, **2.4*D*** & **2.5*D***). To estimate the relative rate of production of CsrA, CsrB and CsrD from the PLlacO-1 and PLtetO-1 promoters at each concentration of inducer, we expressed GFP under these promoters to obtain their induction curves (**Fig. 2.8*Q*** & ***R***). We modeled these induction curves using the Hill function in **Eq. 2.i**.

$$\text{Gene Expression} = a \cdot \left(\frac{[\text{inducer}]^n}{k^n + [\text{inducer}]^n} \right) + c \qquad [2.\text{i}]$$

In **Eq. 2.i**, a is the maximum expression in the presence of inducer, n is the Hill coefficient, k is the concentration at which expression is half the maximum, and c is the minimum expression in the absence of inducer. We determined the parameter values by performing a non-linear least squares fit of the logarithm of the Hill function to the logarithm of the data (green curves, **Fig. 2.8*Q* & *R***) and then subtracted *E. coli* autofluorescence (determined by the fluorescence of HL716 which does not have a fluorescent reporter gene) from c to more accurately quantify the amount of "leaky" expression. The resulting values for the parameters for IPTG were: c = 0.100 fluorescence a.u., a = 155 fluorescence a.u., k = 176 μM IPTG, and n = 3.53 (unitless). The parameter values for aTc were: c = 0.419 fluorescence a.u., a = 40.7 fluorescence a.u., k = 153 nM aTc, and n = 5.87 (unitless).

The above parameter values and Hill functions were used with **Eq. 2.ii** to determine the relative amount of CsrA, CsrB or CsrD produced at any given amount of IPTG or aTc (*i.e.* to calibrate the x-axis of the transfer functions presented in **Fig. 2.2*D***, **2.4*D* & 2.5*D***).

$$\text{Relative Expression } (\%) = \left(\frac{f([\text{IPTG}])}{a}\right) \cdot 100 \qquad \text{[2.ii]}$$

In **Eq. 2.ii**, f is the Hill function for IPTG induction and a is the maximum expression obtained for the above IPTG induction curve. The relative expression for aTc was calculated in the same manner except that the Hill function and maximum expression associated with aTc were used. The concentration of each of these species is assumed to be proportional to their production rates because the rate constants for the clearance of CsrA, CsrB, and CsrD do not vary under the conditions in which the transfer functions were measured (**Fig. 2.2*D***, **2.4*D* & 2.5*D*** respectively).

2.5.2 Stable versus Destabilized GFP

To determine whether an actively degraded reporter would improve our ability to resolve differences in signaling delays, we repeated some of the dynamics experiments (**Fig. 2.2*C***, **Fig. 2.4*C*** and **Fig. 2.5*C***) using a destabilized form of GFP as a reporter. These additional experiments with destabilized GFP (**Fig. 2.9*B-E***) were performed identically to those with stable GFP (*i.e.* **Fig. 2.2*C***, **Fig. 2.4*C*** and **Fig. 2.5*C***) except the measurements were performed more frequently and for a shorter length of time (protocol below). Specifically, we used GlgC-GFP-LVA which has the LVA degradation tag fused to the

C-terminal end of GlgC-GFP [58] (**Fig. 2.9*B-E***). The LVA degradation tag decreased the half-life of the GlgC-GFP target protein by as much as 3.8 ± 0.5 fold relative to the untagged target (**Fig. 2.9*A* & *B***).

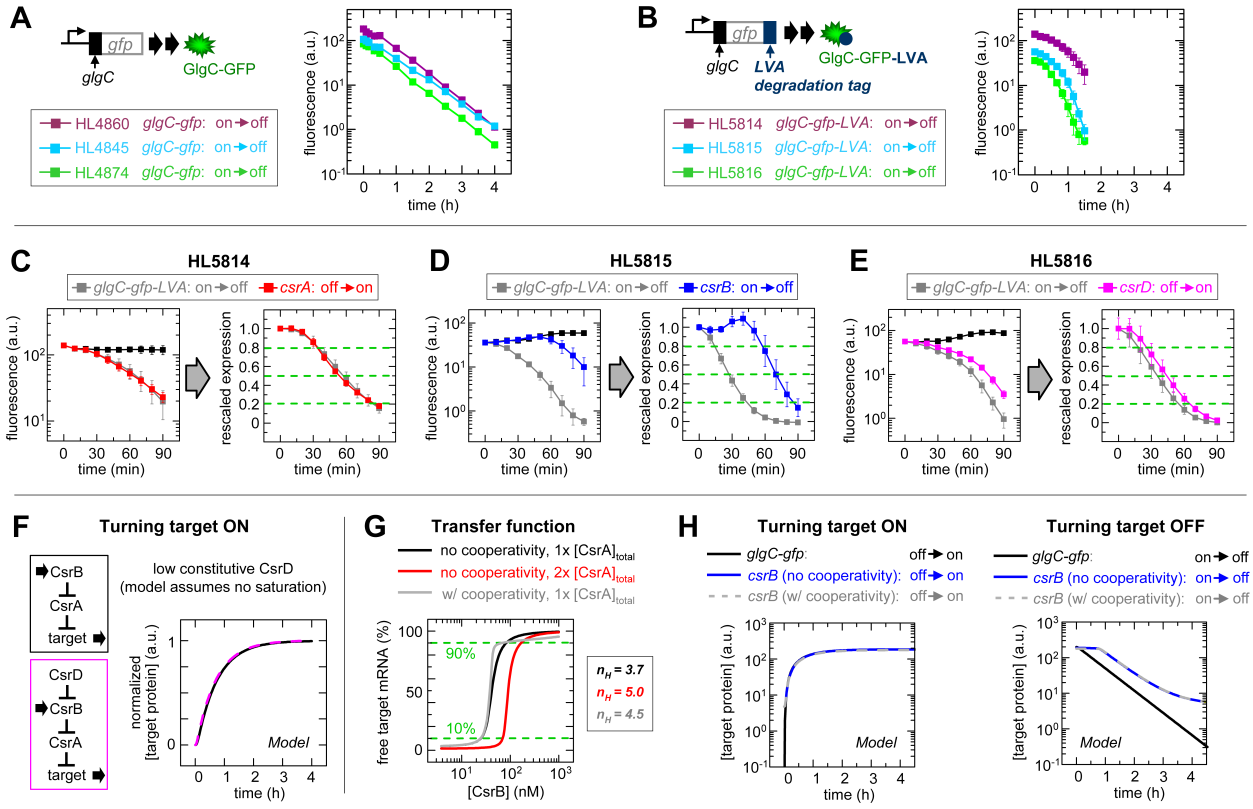

Fig. 2.9 | Alternate experimental and model designs. (***A-E***) Dynamics experiments using target protein without (passive clearance of the target protein) and with the LVA degradation tag (active and passive clearance of the target protein). Transcription of the target gene is turned off at t = 0. (***A***) Clearance of GlgC-GFP without the degradation tag. Dynamics data from **Fig. 2.2*C***, **2.4*C* & 2.5*C*** are replotted for comparison. (***B***) Clearance of GlgC-GFP with the LVA degradation tag. Data are replotted in panels ***B-D***. Note: the curves no longer have a predictable linear decay on a logarithmic plot and instead degradation is initially slow and then becomes faster as the concentration of target protein decreases. (***C-E***) Experiments with a target protein without the LVA tag in **Fig. 2.2*C***, **Fig. 2.4*C* & Fig. 2.5*C*** repeated with a target protein that has the LVA tag in panels ***C***, ***D*** and ***E*** respectively. Green dash lines indicate 80%, 50% and 20% of maximum expression. (***F***) Simulated induction of *csrB* in the presence or absence of a low CsrD concentration without negative feedback (pink dash and black curves respectively). Simulations were performed using the "three step cascade" model that does not include the CsrB-CsrD complex. (***G, H***) Simulations demonstrating that cooperativity has negligible effect on the steady state and dynamic behavior of the CsrA system. (***G***) Free target mRNA levels as a function of CsrB concentration (*i.e.* CsrB transfer functions). CsrB has 9 sites for CsrA dimers. There are three comparisons: (i) CsrA dimers bind independently at 1x concentration (black); (ii) CsrA dimers bind independently but are present at twice the concentration (2x) (red); and (iii) CsrA dimers bind cooperatively at 1x concentration (grey) to CsrB. n_H is the Hill coefficient that would produce a Hill function with the same "steepness" as the function shown; steepness was measured using the levels of CsrB required to reach 10% and 90% occupancy of the target mRNA by CsrA (green dash lines). (***H***) Dynamics of target protein expression following the turning on and off *glgC-gfp* transcription or the transcription of CsrB (thereby turning target expression on and off respectively), with and without cooperative binding between CsrA dimers and CsrB. Direct induction of target protein expression is provided for comparison.

Part 2: Signaling dynamics in the CsrA system

To determine the effect of destabilizing the target protein on our ability to resolve differences between the time delays caused by direct and indirect signaling we make use of the metric δ:

$$\delta = \tau_{80(\text{indirect})} - \tau_{80(\text{direct})} \qquad [2.iii]$$

In **Eq. 2.iii**, $\tau_{80(\text{indirect})}$ is the time delay to reach 80% of the maximum expression with (i) CsrA turned off, (ii) CsrB turned on, or (iii) CsrD turned off, and $\tau_{80(\text{direct})}$ is the time delay to reach 80% of maximum expression after the *glgC-gfp* control is turned off. We found the δ values were similar with stable and destabilized GFP when CsrA is turned off (-1 ± 3 and -2 ± 5 min respectively), CsrB is turned on (54 ± 3 and 44 ± 7 min respectively) and CsrD is turned off (3 ± 2 and 7 ± 11 min respectively) (**Table 2.1** and **Fig. 2.9*C-E***). In short, the addition of a degradation tag to the GFP protein did not improve the measurement of signaling delays.

The dynamics measurements with the LVA-tagged reporter also revealed that its clearance rate is more complicated than that of stable GFP. The clearance of stable GFP by dilution has a clear first-order dependence on reporter concentration (as shown by the linear decrease in fluorescence on the log-linear time course plot in **Fig. 2.9*A***). In contrast, the active clearance of LVA-tagged GFP is slow initially and depends on the starting concentrations of the target protein (**Fig. 2.9*B***). This finding is not unexpected given that active degradation requires specific enzymes and cellular machinery (such as tail-specific proteases) that can become saturated if there are high concentrations of substrate such as GlgC-GFP-LVA. Therefore the addition of the LVA tag introduces extra complexity into the dynamic behavior of the reporter and as a consequence the τ values may convolve delays caused by the CsrA cascade with those caused by the target protein degradation machinery. In summary, stable GFP provides a more suitable target protein to isolate the dynamic properties of the CsrA cascade because its clearance is highly predictable and the measured time delays reflect only the delays incurred by the CsrA cascade. In addition, using destabilized GFP to measure time delays was found to increase the uncertainty in our experiments (**Table 2.1**).

These experiments demonstrated that LVA-tagged GFP was not the best available choice for measuring dynamics within the CsrA cascade. However, the fact that we have observed similar time delays using experiments with the stable and destabilized GFP demonstrates that the time delays that we have measured in the CsrA system are

general and applicable to target proteins that have short lifetimes as well as those that are stable.

2.5.3 Quantitative RT-PCR

Total RNA was extracted from eight exponentially growing cell cultures for each strain (HL5944 and HL5947) using the RNeasy RNA extraction kit (Qiagen). The RNA was treated with DNase I and cDNA was synthesized using the iScript select cDNA synthesis kit and random primers (Bio-Rad). Quantitative RT-PCR was performed to determine the concentration of cDNA using iQ SYBR Green Supermix and the iQ5 Real-Time PCR detection system (Bio-Rad). Samples without reverse transcriptase ("minus RT") were created and measured in parallel to determine the concentration of any contaminating DNA and non-specific amplification. The reported amount of CsrD mRNA for each sample represents the difference between the amount of cDNA in the RNA sample prepared with reverse transcriptase and the same RNA sample prepared identically except without reverse transcriptase. CsrD mRNA was amplified using the oligonucleotides CsrDinF and CsrDinR (**Table 2.4**).

2.5.4 Detailed Model Description

Differential equations, parameter values and initial conditions are discussed below. Simulations were performed using MATLAB (Mathworks).

2.5.4.1 Model overview

We constructed a model with ordinary differential equations for each of the following species: GlgC-GFP protein ("GFP"), *glgC-gfp* mRNA ("m"), CsrA dimer ("A"), CsrB non-coding RNA ("B"), CsrD ("D"), CsrA-target mRNA complex ("Am") and the CsrA-CsrB complex ("AB"). The CsrB-CsrD complex ("BD") is also included in later versions of the model. It should be noted that in our model, "B" represents a pair of CsrA binding sites within the CsrB molecule that can be bound by a single CsrA dimer, and thus the CsrA-CsrB complex "AB" represents a CsrA dimer occupying one pair of CsrA binding sites.

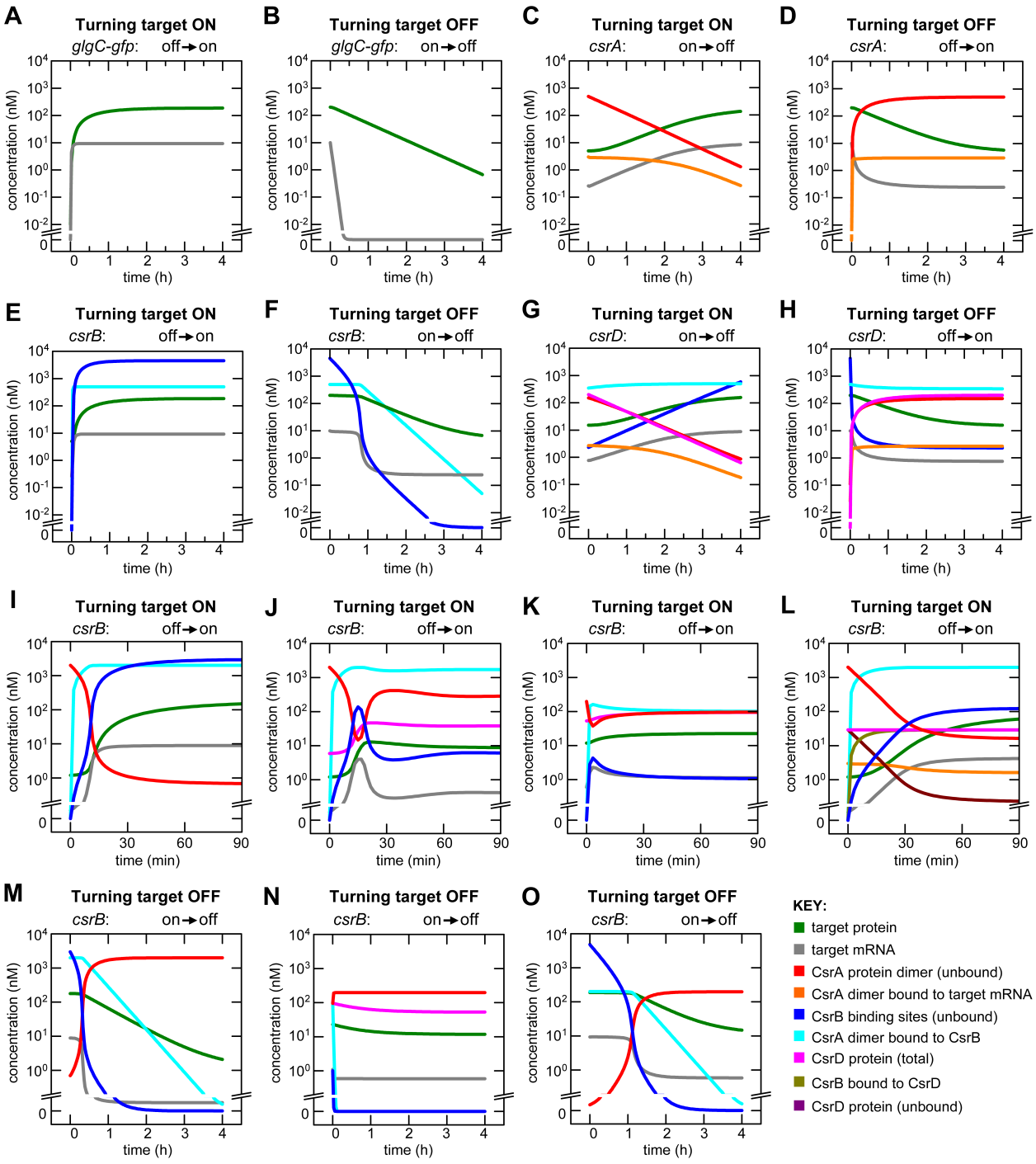

Fig. 2.10 | Simulated concentrations CsrA cascade components. Components in the CsrA cascade are color coded (key is in the bottom right of the figure) and not all are shown in each panel. Colors used here do not correspond to those in other figures. (***A***, ***B***) Turning on and off *glgC-gfp* mRNA transcription (CsrA, CsrB and CsrD are absent). These simulations correspond to the grey curves in **Fig. 2.2*B, C***, **Fig. 2.4*B, C*** and **Fig. 2.5*B, C***. (***C***, ***D***) Turning off and on CsrA expression with *glgC-gfp* mRNA at constant maximal transcription (CsrB and CsrD are absent). These simulations correspond to the red curves in **Fig. 2.2*B, C***. (***E***, ***F***) Turning

on and off CsrB expression with both *glgC-gfp* mRNA and CsrA at constant maximal expression (CsrD is absent). These simulations correspond to the blue curves in **Fig. 2.4*B, C***. (***G***, ***H***) Turning off and on CsrD expression with *glgC-gfp* mRNA, CsrA and CsrB all at constant maximal expression. These simulations correspond to the magenta curves in **Fig. 2.5*B***, ***C***. (***I***) Turning on CsrB expression in the synthetic CsrA cascade without feedback and both *glgC-gfp* mRNA and CsrA at constant maximal expression (*csrD* is absent). $[CsrA]_{total}$ = 1x. This simulation corresponds to the control curve (black) in **Fig. 2.7*E, F***. This is a reference panel for comparison with panels ***J***, ***K*** and ***L***. (***J***, ***K***) Turning on CsrB expression in the three step cascade with negative feedback and saturation of CsrD activity by CsrB. These simulations correspond to the dark blue and light blue curves in **Fig. 2.7*E*** which have $[CsrA]_{total}$ = 1x (panel ***J***) and $[CsrA]_{total}$ = 0.1x (panel ***K***) respectively. (***L***) Turning on CsrB expression in a system where CsrD is constitutively produced at a low rate (no feedback is present), and *glgC-gfp* mRNA and CsrA are at constant maximal expression. $[CsrA]_{total}$ = 1x. This simulation corresponds to the magenta curve in **Fig. 2.7*F***. (***M***) Turning off CsrB expression in the synthetic CsrA cascade as described in panel ***I***. This simulation corresponds to the control curve (black) in **Fig. 2.7*G***. This is a reference panel for comparison with panels ***N*** and ***O***. (***N***) Turning off CsrB expression in the cascade as described in panel ***K***. This simulation corresponds to the gold curve in **Fig. 2.7*G***. (***O***) Same as ***M*** except $[CsrA]_{total}$ = 0.1x.

To our knowledge, our model is the first to quantitatively and specifically describe the CsrA regulatory cascade. A previously reported model described the general features of repression of target mRNA translation by a protein and CsrA was cited as example [47]; however the reported model did not incorporate any upstream regulators of CsrA (*i.e.* CsrB or CsrD) (**Fig. 2.4-2.7**) or negative feedback resulting from the repression of CsrD production by CsrA (**Fig. 2.7**). The basic assumptions of our model are described below.

Initially we modeled the synthetic CsrA system in which each gene is under the control of a synthetic promoter without any flanking native regulatory sequences (and therefore without negative feedback). This model was constructed in three sequential steps. In the first step we modeled only the action of CsrA on the target protein concentration ("one step cascade"). In the second step we added CsrB to the one step cascade model thereby creating a "two step cascade" model. In the third step we added CsrD to the two step cascade model to create a "three step cascade" model. Parameter values were obtained from the literature, reasonable *a priori* estimates, and by fitting the simulated dynamics to the experimental data. The latter was used to determine the production rates of CsrA, CsrB and CsrD in the one, two and three step cascade models respectively. The one, two and three step cascade models were each validated by comparing the transfer functions they predicted to those measured experimentally. Finally, we expanded the three step cascade to include negative feedback and the saturation of CsrD by CsrB to model the behavior of the native CsrA system.

In each dynamics simulation, we turned on or off the production of one component in the cascade ("modulated component") and measured the effect on the concentrations of all the other components of the system as a function of time. When we turned on the production of the modulated component (*e.g.* CsrA), its initial concentration was zero

and the initial concentrations of the other species (*e.g. glgC-gfp* mRNA and target protein) were at their steady state levels in the absence of the modulated component. These steady levels were obtained by simulating the system with the production of the modulated component turned off (*e.g.* CsrA) and allowing the system to converge to steady state. When we turned off the production of a modulated component (*e.g.* CsrA), its initial concentration and the initial concentrations of the other species (*e.g. glgC-gfp* mRNA and target protein) were at their steady state with maximal production of the modulated component (*i.e.* CsrA). These initial steady state levels (*e.g.* CsrA, *glgC-gfp* mRNA and target protein) were obtained by simulating the system with the production of the modulated component turned on (*e.g.* CsrA) and allowing it to converge to steady state. The simulations were performed by integrating the model's equations using the ode15s solver in MATLAB. All the initial conditions are provided in **Table 2.2**.

Table 2.2 | Initial conditions for dynamic simulations

simulation	*[GFP] (nM)*	*[m] (nM)*	*[A] (nM)*	*[Am] (nM)*	*[B] (nM)*	*[AB] (nM)*	*[D] (nM)*	*[BD] (nM)*
Fig. 2.2*B*, *glgC-gfp* turned on	0	0	0	0	n/a	n/a	n/a	n/a
Fig. 2.2*B*, *csrA* turned off	5	0.26	497	3	n/a	n/a	n/a	n/a
Fig. 2.2*C*, *glgC-gfp* turned off	200	10	0	0	n/a	n/a	n/a	n/a
Fig. 2.2*C*, *csrA* turned on	200	10	0	0	n/a	n/a	n/a	n/a
Fig. 2.4*B*, *glgC-gfp* turned on	0	0	0.11	0	4500	500	n/a	n/a
Fig. 2.4*B*, *csrB* turned on	5	0.26	497	3	0	0	n/a	n/a
Fig. 2.4*C*, *glgC-gfp* turned off	198	9.9	0.11	0.026	4500	500	n/a	n/a
Fig. 2.4*C*, *csrB* turned off	198	9.9	0.11	0.026	4500	500	n/a	n/a
Fig. 2.5*B*, *glgC-gfp* turned on	0	0	0.11	0	4500	500	0	n/a
Fig. 2.5*B*, *csrD* turned off	15	0.76	150	2.7	2.3	347	200	n/a
Fig. 2.5*C*, *glgC-gfp* turned off	198	9.9	0.11	0.026	4500	500	0	n/a
Fig. 2.5*C*, *csrD* turned on	198	9.9	0.11	0.026	4500	500	0	n/a
Fig. 2.7*E* and ***F*** control (black)	1.2	0.06	1997	3	0	0	0	0
Fig. 2.7*E*, 1.0x CsrA (drk blue)	1.2	0.06	1997	3	0	0	5.9	0
Fig. 2.7*E*, 0.1x CsrA (lght blue)	11.8	0.6	197	3	0	0	53	0
Fig. 2.7*F*, fixed CsrD (pink)	1.2	0.06	1997	3	0	0	29	0
Fig. 2.7*G*, control (black)	178	8.9	0.7	0.1	3000	1999	0	0
Fig. 2.7*G*, native system (gold)	22.8	1.1	95	2.6	1	102	47	49
Fig. 2.10*O*, native w/o *csrD*	187	9.3	0.042	0.009	4800	200	0	0

Table 2.2 | Initial conditions for dynamic simulations. The initial conditions for dynamics simulations are included here. [GFP], [m], [A], [Am], [B], [AB], [D] and [BD] represent the concentrations of the target protein, the target mRNA, CsrA dimers, the CsrA-target mRNA complex, CsrB, the CsrA-CsrB complex, CsrD and the CsrB-CsrD complex respectively as described.

Part 2: Signaling dynamics in the CsrA system

The transfer functions were simulated as described above for the dynamics with the production rate of the modulated component (CsrA, CsrB or CsrD) assigned a range of values between zero (turned off) and its maximum (turned on). For each value of the production rate we ran the dynamics simulations to convergence to determine the steady-state target protein concentration. We emphasize that the model was constructed from first principles, parameter values obtained from the literature and production rates obtained from the dynamics experiments. No information from our steady state experiments was incorporated into the models used to predict the transfer functions. Furthermore, predicting the transfer functions is non-trivial because they convolve multiple non-linear functions. For example, predicting the effect of varying the steady state CsrD concentration on the steady state target protein concentration requires the model to accurately predict how the CsrD concentration alters the CsrB concentration, how the CsrB concentration alters the free CsrA concentration, and how the free CsrA concentration alters the target mRNA's concentration and translation. We performed *in vivo* experiments to determine the transfer functions and found them to be in good agreement with the predictions. That is, the steady state experiments confirmed the predictive capacity of our model and thus the appropriateness of its assumptions and parameters.

2.5.4.2 One-step cascade (CsrA and target)

The simplest model is the "one step" cascade that contains only target proteins, target mRNAs, CsrA dimers, and CsrA-target mRNA complexes ("CsrA-*glgC* mRNA") (**Fig. 2.2**). Our model describes the production and removal (by dilution, active degradation and/or sequestration) of each component.

The GlgC-GFP concentration ([GFP]) is determined by (i) its production rate, which is proportional to the free *glgC-gfp* mRNA concentration ([m]) and the rate constant α_G, and (ii) its degradation rate, which is proportional to [GFP] and the sum of the rate constants for passive dilution (β_{dil}) and active degradation of the protein (β_G).

$$\frac{d[\mathrm{GFP}]}{dt} = \alpha_{\mathrm{G}}[\mathrm{m}] - (\beta_{\mathrm{G}} + \beta_{\mathrm{dil}})\,[\mathrm{GFP}] \qquad [2.1]$$

The concentration of free *glgC-gfp* mRNA ([m]) is determined by (i) its production rate α_m which depends on promoter activity, and (ii) its removal rate which depends on passive dilution, active degradation, and sequestration by CsrA. The rates of active

degradation and passive dilution for the mRNA are proportional to [m] and the rate constants β_m and β_{dil} respectively. The rate of removal of free *glgC-gfp* mRNA via sequestration into the CsrA-*glgC* mRNA complex (Am) is proportional to the *glgC-gfp* mRNA and CsrA concentrations and the rate constant for association (k_1). Release of free *glgC-gfp* mRNA from the CsrA-*glgC* mRNA complex is proportional to the concentration of the complex ([Am]) and the dissociation rate constant k_{-1}. For simplicity, we assume the *glgC* leader sequence in the *glgC-gfp* mRNA is capable of binding to one CsrA dimer and that the bound CsrA completely inhibits translation. Since the predictions of our model are in good agreement with our *in vivo* dynamics data (see above), this simplifying assumption indicates that more complex modes of interaction between CsrA and RNA molecules (*e.g.* cooperativity) are not required to explain the observed dynamics (note: this is modeled and discussed in detail below). In addition, the details of the reaction steps involved in CsrA binding to its RNA partners have not been fully elucidated and are still being investigated [70]. Theoretically, free *glgC-gfp* mRNA could also be generated by active degradation of CsrA in the CsrA-*glgC* mRNA complex; while we include this possibility to be systematic, active degradation of CsrA is negligible and therefore we set the rate constant for this reaction (β_{AAm}) to zero.

$$\frac{d[\mathrm{m}]}{dt} = \alpha_{\mathrm{m}} - (\beta_{\mathrm{m}} + \beta_{\mathrm{dil}})[\mathrm{m}] - k_1[\mathrm{A}][\mathrm{m}] + (k_{-1} + \beta_{\mathrm{AAm}})[\mathrm{Am}] \qquad [2.2]$$

The concentration of CsrA-*glgC* mRNA complex ([Am]) is determined by it creation rate (which depends on the association rate constant and the concentrations of *glgC-gfp* mRNA and CsrA) and the clearance rate (which is proportional to the concentration of CsrA-*glgC* mRNA complex and the rate constants for dissociation (k_{-1}), active degradation of the *glgC-gfp* mRNA within the complex (β_{mAm}), active degradation of CsrA within the complex (β_{AAm}), and passive dilution (β_{dil})).

$$\frac{d[\mathrm{Am}]}{dt} = k_1[\mathrm{A}][\mathrm{m}] - (k_{-1} + \beta_{\mathrm{mAm}} + \beta_{\mathrm{AAm}} + \beta_{\mathrm{dil}})[\mathrm{Am}] \qquad [2.3]$$

The concentration of free CsrA dimer ([A]) is determined by (i) its production rate α_A and (ii) its removal rate which depends on passive dilution, active degradation, and sequestration into the CsrA-*glgC* complex. The production of CsrA dimers was treated as a single step to maintain the simplicity of the model. It was unnecessary to explicitly

model the production and degradation of *csrA* mRNA because its half-life is short compared to the CsrA protein's half-life (therefore the dynamics of the *csrA* mRNA has negligible effect on the CsrA protein dynamics). Similarly, we do not separately model the monomeric and dimeric forms of CsrA because the vast majority of intracellular CsrA exists as a dimer [52]. Furthermore, because very little CsrA is found in the monomeric form this implies that it associates rapidly compared to the turnover rate of CsrA. In the unlikely event that CsrA dimerization is slow; its contribution to the dynamics would be accounted for because it would be incorporated into the CsrA production rate α_A which is obtained from a fit to the experimental data (see below). The rate of removal of free CsrA dimers via active degradation and passive dilution is proportional to [A] and the rate constants β_A and β_{dil} respectively. We included the possibility of active degradation of CsrA in the model for completeness; however, because CsrA is a stable protein we assumed β_A to be zero. Free CsrA dimers are removed by sequestration into the CsrA-*glgC* mRNA complex as described above. Free CsrA dimers are generated both by degradation of the *glgC-gfp* mRNA within the CsrA-*glgC* mRNA complex (rate constant β_{mAm}) and by dissociation of the CsrA-*glgC* mRNA complex (rate constant k_{-1}).

$$\frac{d[\mathrm{A}]}{dt} = \alpha_{\mathrm{A}} - (\beta_{\mathrm{A}} + \beta_{\mathrm{dil}})[\mathrm{A}] - k_1[\mathrm{A}][\mathrm{m}] - (k_{-1} + \beta_{\mathrm{mAm}})[\mathrm{Am}] \qquad [2.4a]$$

We now establish numerical values for the above kinetic parameters. The rate constant for passive dilution (β_{dil}) was determined to be 4×10^{-4} s^{-1} by dividing ln(2) by the average doubling time of four strains (HL4495, HL4574, HL4845 and HL4860) grown in LB in the absence of inducer molecules. We used stable GFP in our simulated experiments (*i.e.* GFP is not actively degraded) therefore $\beta_G = 0$. We estimated that the steady state target protein concentration was likely to be on the high end of the physiological range (2×10^2 nM) [71] because the *glgC-gfp* mRNA is transcribed from a strong promoter, the *glgC-gfp* gene is located on a medium copy plasmid, and the *glgC-gfp* mRNA has an efficient RBS. Multiplying this steady state concentration of target protein by β_{dil} yields the GFP production rate (8×10^{-2} nM·s^{-1}), which is the product of the rate constant for GFP production (α_G) and the steady state *glgC-gfp* mRNA concentration. The steady state *glgC-gfp* mRNA concentration was also estimated to be on the high side of the physiological range (1×10^1 nM) due to its strong promoter and because the gene is on a medium copy plasmid [71]. Dividing the total GFP production rate by the steady state mRNA concentration provides $\alpha_G = 8 \times 10^{-3}$ s^{-1}.

Part 2: Signaling dynamics in the CsrA system

We next specify the rate constants related to the *glgC-gfp* mRNA. We multiplied the steady state concentration of target mRNA (1 x 10^1 nM) by the sum of the rate constants for passive dilution (β_{dil}) and active degradation (β_m) to calculate the production rate for the *glgC-gfp* mRNA (α_m). β_m was estimated to be 6 x 10^{-3} s^{-1} by dividing ln(2) by the half-life (~120 s) determined from reported plots of native *glgC* mRNA degradation [72]. Therefore $\beta_m + \beta_{dil}$ = 6.0 x 10^{-3} s^{-1} + 0.4 x 10^{-3} s^{-1} = 6.4 x 10^{-3} s^{-1} and α_m = 6.4 x 10^{-2} nM·s^{-1}. We assumed that the active degradation of CsrA, bound or unbound, is negligible; this is supported both by our experiments (**Fig. 2.2*B* & Fig. 2.6**) and studies that have shown that the vast majority of intracellular proteins in *E. coli* are stable [73]. We therefore assume that the rate constant for active degradation of CsrA from within the CsrA-*glgC* mRNA complex (β_{AAm}) is zero. The rate constants k_1 and k_{-1} are unknown but their ratio ($K_{d,Am} = k_{-1}/k_1$) has been determined for native *glgC* mRNA (~40 nM) [53]. We used the association rate constant for another RNA-protein complex (a 16 nucleotide RNA substrate binding to Hfq) [74], which has been determined (~ 10^{-1} nM^{-1}·s^{-1}), as a starting point for estimating k_1. Thinking ahead, the association rate constant of the CsrA-CsrB complex is likely to be larger than that of the CsrA-*glgC* mRNA complex due to the difference in their reported K_d values (see below). We therefore selected a conservative value for the association rate constant of the CsrA-*glgC* mRNA complex (k_1 = 10^{-2} nM^{-1}·s^{-1}), leaving the higher value for the association rate constant of the CsrA-CsrB complex (10^{-1} nM^{-1}·s^{-1}). Having estimated k_1, we calculated k_{-1} = 4 x 10^{-1} s^{-1} from the above ratio (k_{-1}/k_1 ~ 40 nM). The rate constant for the active degradation of *glgC-gfp* mRNA from within the CsrA-glgC mRNA complex (β_{mAm} = 2 x 10^{-2} s^{-1}) was calculated by dividing ln(2) by the half-life (~35 s) estimated from plots of the degradation of native *glgC* mRNA bound to CsrA [72].

We now establish the remaining rate constants for CsrA that have not been described above. As we discussed, CsrA is not believed to be actively degraded therefore β_A = 0 s^{-1} and therefore CsrA is cleared at a rate determined by β_{dil}. The value for the production rate of CsrA dimer, α_A (2 x 10^{-2} nM·s^{-1}), was determined by fitting the model to our dynamics data. Specifically, α_A was selected so that the simulated delay in turning on target expression caused by turning off *csrA* transcription (which depends on the CsrA concentration) was comparable in magnitude to that measured experimentally (**Fig. 2.2*B***).

Simulation of the one step cascade with the above parameters and only a single fit parameter (α_A) was able to describe the dynamics behavior of the system with CsrA production turned off (α_A = 0, **Fig. 2.2*B***) or on (α_A = 2 x 10^{-2} nM·s^{-1}, **Fig. 2.2*C***). Furthermore, simulation of the system with a range of different CsrA production rates ($0 \leq \alpha_A \leq 2 \times 10^{-2}$ nM·s^{-1}) qualitatively predicted the steady state CsrA transfer function (**Fig. 2.2*D***).

2.5.4.3 Two-step cascade (CsrB, CsrA and target)

The two step cascade model extends the one step cascade to include CsrB and the CsrA-CsrB complex (**Fig. 2.4**). This model is described by **Eq. 2.1-2.3**, **2.4b, 2.5a & 2.6**. The concentration of free CsrB ([B]) is determined by its (i) production rate α_B and (ii) removal rate due to passive dilution, active degradation (via non-CsrD pathways), and sequestration into the CsrA-CsrB complex. For simplicity, [B] represents the concentration of CsrB binding sites rather than the concentration of full length CsrB RNA molecules. In other words, we measure the CsrB concentration in terms of CsrA dimer equivalents where one unit of CsrB can bind one CsrA dimer. This assumes that the binding of each CsrA dimer to CsrB is independent, which is the simplest model. More complex binding interactions between CsrA and CsrB are unnecessary for qualitative agreement between our model and experimental measurements (**Fig. 2.2, 2.4 & 2.5**). To be clear, we refer to the independent binding of CsrA *dimers* to CsrB not to the binding of the *individual* CsrA molecules within a dimer to CsrB. In the latter case, the binding of a CsrA molecule within a dimer to CsrB is known to increase the probability that the other CsrA molecule will also bind to CsrB [70]; this behavior is included in our model with the assumption that CsrA dimers occupy two binding sites within CsrB (*i.e.* CsrB has 18 sites for CsrA, but only 9 sites for CsrA dimers).

As with all other species in our model, the rate of passive dilution of CsrB is proportional to its concentration and the rate constant β_{dil}. The rate of active degradation of CsrB via non-CsrD mediated pathways is proportional to [B] and the rate constant β_B. The rate of sequestration of CsrB is proportional to [B], [A], and the association rate constant for the CsrA-CsrB complex (k_2). CsrB can be released from the CsrA-CsrB complex by active degradation of CsrA or by dissociation; the rates of these reactions are proportional to the concentration of the complex ([AB]) and the constants β_{AAB} or k_{-2} respectively. Active degradation of CsrB in the CsrA-CsrB complex may also generate free CsrA at a rate that is proportional to the concentration of the complex ([AB]) and the rate constant β_{BAB}.

Part 2: Signaling dynamics in the CsrA system

A model of CsrB molecules with 9 sites for CsrA dimers versus a model of single binding sites for CsrA dimer will produce equivalent results in a deterministic model. However the latter is much simpler because there is no need to model all the different types of CsrA-CsrB complexes. The reason the two types of models are equivalent is that differences in the production rate are taken into account when we fit this parameter. Furthermore, the degradation rate constants for CsrB molecules and individual CsrA dimer binding sites, which depend on their lifetimes, are the same under conditions where CsrB clearance is not saturated. That is, under non-saturating conditions the time taken to cleave an individual molecule is independent of the number of other molecules present. Under these circumstances, the rate of clearance of single CsrA dimer sites will be nine times greater than for CsrB whole molecules but the *rate constant* and overall flux of dimer sites is identical. Under conditions when CsrD is saturated with CsrB we fit our model to the experimental data and therefore the parameter values obtained account for differences in the clearance of CsrB molecules and single binding sites for CsrA dimer.

The equations for the two step cascade include the above process as well as those described in the one step cascade; that is, **Eq. 2.1-2.3** from the one step cascade and a modified equation for the free CsrA concentration (**Eq. 2.4b**) that includes the association, dissociation and active degradation of the CsrA-CsrB complex:

$$\frac{d[\mathrm{A}]}{dt} = \alpha_{\mathrm{A}} - (\beta_{\mathrm{A}} + \beta_{\mathrm{dil}})[\mathrm{A}] - k_1[\mathrm{A}][\mathrm{m}] + (k_{-1} + \beta_{\mathrm{mAm}})[\mathrm{Am}] - k_2[\mathrm{A}][\mathrm{B}] + (k_{-2} + \beta_{\mathrm{BAB}})[\mathrm{AB}] \qquad [2.4b]$$

In addition, the model for the two step cascade has equations for free CsrB and the CsrA-CsrB complex:

$$\frac{d[\mathrm{B}]}{dt} = \alpha_{\mathrm{B}} - (\beta_{\mathrm{B}} + \beta_{\mathrm{dil}})[\mathrm{B}] - k_2[\mathrm{A}][\mathrm{B}] + (k_{-2} + \beta_{\mathrm{AAB}})[\mathrm{AB}] \qquad [2.5a]$$

$$\frac{d[\mathrm{AB}]}{dt} = k_2[\mathrm{A}][\mathrm{B}] - (k_{-2} + \beta_{\mathrm{BAB}} + \beta_{\mathrm{AAB}} + \beta_{\mathrm{dil}})[\mathrm{AB}] \qquad [2.6]$$

The values of the additional parameters were obtained as follows. CsrB stability is not believed to be affected by binding to CsrA [63] (unlike target mRNAs that bind to CsrA)

and therefore the active degradation rates for bound and unbound CsrB are the same ($\beta_{BAB} \equiv \beta_B$). The rate constant for the active degradation of CsrB in the absence of CsrD ($\beta_B \leq 4 \times 10^{-4}$ s^{-1}) was determined by dividing ln(2) by the reported half-life of CsrB (≥ 30 min) in a *csrD* deletion strain [51]. It is unclear how much the half-life was influenced by cell growth therefore the above value should be considered as an upper bound. In other words, active degradation contributes the same or less than passive dilution to the clearance of CsrB. We include the active degradation of CsrA from the CsrA-CsrB complex (β_{AAB}) for completeness. However, since CsrA is stable, $\beta_{AAB} = 0$. The association rate constant for CsrA dimer binding to CsrB (k_2) was estimated to be similar in magnitude to RNA binding to Hfq (see above; $k_2 = 10^{-1}$ $nM^{-1} \cdot s^{-1}$). We calculated the dissociation rate constant for the CsrA-CsrB complex ($k_{-2} = 10^{-1}$ s^{-1}) from the above value for k_2 and the reported equilibrium dissociation constant for this complex ($K_{d,AB} = k_{-2}/k_2 \sim 1$ nM) as measured by gel mobility shift assays [49]. The production rate of pairs of CsrB binding sites ($\alpha_B = 4$ $nM \cdot s^{-1}$) was determined by fitting the model to our dynamics data so the simulated signaling delay was comparable in magnitude to that measured experimentally (**Fig. 2.4*C***). This production rate corresponds to a steady state concentration of CsrB binding sites [B] of 5×10^3 nM. Since each CsrB molecule has approximately 9 binding sites for CsrA dimers (assuming both faces of the CsrA dimer bind simultaneously) [50] the concentration of CsrB molecules will be ~9-fold lower than [B] (and therefore the production rate for CsrB molecules would also be 9-fold lower).

Simulation of the two step cascade with the above parameters and only a single fit parameter (α_B) was able to describe the dynamics behavior of the system with CsrB production turned on ($\alpha_B = 4$ $nM \cdot s^{-1}$, **Fig. 2.4*B***) or off ($\alpha_B = 0$, **Fig. 2.4*C***). Furthermore, simulation of the system with a range of different CsrB production rates ($0 \leq \alpha_B \leq 4$ $nM \cdot s^{-1}$) qualitatively predicted the steady state CsrB transfer function (**Fig. 2.4*D***).

2.5.4.4 Three-step cascade (CsrD, CsrB, CsrA and target)

The three step cascade model extends the two step cascade to include CsrD (**Fig. 2.5**). This model is described by **Eq. 2.1-2.3**, **Eq. 2.4b**, **Eq. 2.5b**, **Eq. 2.6 & Eq. 2.7a**. The concentration of CsrD ([D]) depends on (i) its production rate (α_D) and (ii) its removal rate by passive dilution and active degradation. The production of CsrD is modeled as a single reaction step as with CsrA dimers. The clearance of CsrD by passive dilution and active degradation is proportional to [D] and the rate constants β_D and β_{dil} respectively.

Part 2: Signaling dynamics in the CsrA system

$$\frac{d[\mathrm{D}]}{dt} = \alpha_{\mathrm{D}} - (\beta_{\mathrm{D}} + \beta_{\mathrm{dil}})[\mathrm{D}] \qquad [2.7a]$$

CsrD is known to bind to and facilitate the degradation of free CsrB by RNase E [51]. The kinetics of this process have not been completely characterized and therefore we begin by modeling the action of CsrD as a simple first order process. Active degradation of free CsrB by CsrD is therefore proportional to [B], [D] and a single kinetic parameter that defines the catalytic efficiency of CsrD-mediated degradation (ω). Incorporating active degradation of CsrB by the CsrD mediated pathway into the model converts **Eq. 2.5a** into **Eq. 2.5b**:

$$\frac{d[\mathrm{B}]}{dt} = \alpha_{\mathrm{B}} - (\omega[\mathrm{D}] + \beta_{\mathrm{B}} + \beta_{\mathrm{dil}})[\mathrm{B}] - k_2[\mathrm{A}][\mathrm{B}] + (k_{-2} + \beta_{\mathrm{AAB}})[\mathrm{AB}] \qquad [2.5b]$$

This equation assumes the free CsrB concentration is not greatly in excess of the CsrD concentration, which our experiments show is appropriate for dynamics experiments using the synthetic CsrA cascade (**Fig. 2.5**). However, at the lower concentrations of CsrD that can occur with native *csrD* (**Fig. 2.7*B*, *C*, *E*, *G***) or at very low induction levels of synthetic *csrD* (**Fig. 2.7*F***), CsrD can become saturated by high levels of free CsrB. Under these conditions where CsrD becomes saturated it is important to consider the CsrB-CsrD complex (see next section). As stated above, the three step cascade model combines all the equations for the two step cascade except **Eq. 2.5a** (*i.e.* **Eq. 2.1, 2.2, 2.3, 2.4b, 2.6**) plus **Eq. 2.5b & Eq. 2.7a**.

The additional parameter values for **Eq. 2.5b & 2.7** were determined as follows. It is believed that CsrD is primarily cleared by dilution for the same reasons as CsrA therefore the rate constant for active degradation of CsrD is assumed to be near zero (β_D = 0). Since CsrD partners with RNase E to degrade CsrB, we selected a catalytic efficiency value ($\omega = 8 \times 10^{-3}$ nM^{-1}·s^{-1}) within the range of reported catalytic efficiencies for RNase E [75,76]. The production rate for CsrD ($\alpha_D = 8 \times 10^{-2}$ nM·s^{-1}) was determined by fitting the model to our dynamics data so that the simulated signaling delay was comparable in magnitude to that measured experimentally (**Fig. 2.5*B***). It should be noted that all other parameter values (including α_A and α_B) were held fixed during this fit process. We found the production rate for CsrD is less than that for CsrA which is consistent with *csrD* having a weaker ribosomal binding sequence (st3 instead of the stronger st7 ribosomal binding sequence of synthetic *csrA*) [77].

Simulation of the three step cascade with the above parameters and only a single fit parameter (α_D) was able to describe the dynamic behavior of the system with CsrD production turned off ($\alpha_D = 0$, **Fig. 2.5*B***) or on ($\alpha_D = 8 \times 10^{-2}$ nM·s^{-1}), **Fig. 2.5*C***). Furthermore, simulation of the system with a range of different CsrD production rates ($0 \leq \alpha_D \leq 8 \times 10^{-2}$ nM·s^{-1}) qualitatively predicted the CsrD transfer function at steady state (**Fig. 2.5*D***).

2.5.4.5 Three-step cascade with feedback

To interpret and explain our experimental observations for the CsrA cascade with native *csrD* we incorporated into our model the repression of CsrD production by CsrA (**Fig. 2.7**). This repression is necessary for negative feedback regulation. Because native *csrD* is not as highly expressed as synthetic *csrD*, CsrD may become saturated by CsrB therefore we needed to explicitly include CsrB-CsrD complexes (BD) in the model to account for this possibility. The model verified that negative feedback was necessary for the "enhanced signaling" observed in **Fig. 2.7*E***. In addition, the model demonstrated that saturation of CsrD activity by CsrB is required for the delay (**Fig. 2.7*F***) and the delay does not occur if CsrD is absent (**Fig. 2.7*F***) or with low constant expression of CsrD without saturation (**Fig. 2.9*F***).

The repression of CsrD expression by CsrA was incorporated into the model by having the production rate of CsrD from native *csrD* depend on the CsrA concentration according to a simple Hill-type function. The Hill-type function [78] is $k_f/(k_f + [A])$, where the constant k_f determines the CsrA concentration [A] at which CsrD production is half its maximum value. Use of the Hill-type function requires fewer parameters and assumptions than additional equations that explicitly consider the *csrD* mRNA and its association, dissociation, degradation and translation in the CsrA-*csrD* mRNA complex (note: to our knowledge, this complex has not been identified *in vivo*).

The association of CsrB and CsrD results in the CsrB-CsrD complex and this reaction occurs at a rate proportional to [B], [D] and the rate constant k_{ES}. The CsrB-CsrD complex dissociates to free CsrB and CsrD at a rate that depends on its concentration and the rate constant k_{-ES}. Clearance of the CsrB-CsrD complex can occur by dilution, active degradation of CsrB in the complex by RNase E (which releases free CsrD), and active degradation of CsrD in the complex (which releases free CsrB); the rate constants for these respective processes are β_{dil}, k_P and β_D.

Part 2: Signaling dynamics in the CsrA system

Our model of the three step cascade with feedback regulation extends the previous model by the modification of **Eq. 2.5b & 2.7a** to include (i) the regulation of CsrD production by CsrA and (ii) the addition of an equation for the CsrB-CsrD complex. The final set of equations for the three step model with feedback is therefore:

$$\frac{d[\mathrm{GFP}]}{dt} = \alpha_{\mathrm{G}}[\mathrm{m}] - (\beta_{\mathrm{G}} + \beta_{\mathrm{dil}})\,[\mathrm{GFP}] \qquad [2.1]$$

$$\frac{d[\mathrm{m}]}{dt} = \alpha_{\mathrm{m}} - (\beta_{\mathrm{m}} + \beta_{\mathrm{dil}})[\mathrm{m}] - k_1[\mathrm{A}][\mathrm{m}] + (k_{-1} + \beta_{\mathrm{AAm}})[\mathrm{Am}] \qquad [2.2]$$

$$\frac{d[\mathrm{Am}]}{dt} = k_1[\mathrm{A}][\mathrm{m}] - (k_{-1} + \beta_{\mathrm{mAm}} + \beta_{\mathrm{AAm}} + \beta_{\mathrm{dil}})[\mathrm{Am}] \qquad [2.3]$$

$$\frac{d[\mathrm{A}]}{dt} = \alpha_{\mathrm{A}} - (\beta_{\mathrm{A}} + \beta_{\mathrm{dil}})[\mathrm{A}] - k_1[\mathrm{A}][\mathrm{m}] + (k_{-1} + \beta_{\mathrm{mAm}})[\mathrm{Am}] - k_2[\mathrm{A}][\mathrm{B}] + (k_{-2} + \beta_{\mathrm{BAB}})[\mathrm{AB}] \qquad [2.4b]$$

$$\frac{d[\mathrm{B}]}{dt} = \alpha_{\mathrm{B}} - (\beta_{\mathrm{B}} + \beta_{\mathrm{dil}})[\mathrm{B}] - k_2[\mathrm{A}][\mathrm{B}] + (k_{-2} + \beta_{\mathrm{AAB}})[\mathrm{AB}] - k_{\mathrm{ES}}[\mathrm{B}][\mathrm{D}] + (k_{\text{-ES}} + \beta_{\mathrm{D}})[\mathrm{BD}] \qquad [2.5c]$$

$$\frac{d[\mathrm{AB}]}{dt} = k_2[\mathrm{A}][\mathrm{B}] - (k_{-2} + \beta_{\mathrm{BAB}} + \beta_{\mathrm{AAB}} + \beta_{\mathrm{dil}})[\mathrm{AB}] \qquad [2.6]$$

$$\frac{d[\mathrm{D}]}{dt} = \alpha_{\mathrm{D}}\left(\frac{k_{\mathrm{f}}}{k_{\mathrm{f}} + [\mathrm{A}]}\right) - (\beta_{\mathrm{D}} + \beta_{\mathrm{dil}})[\mathrm{D}] - k_{\mathrm{ES}}[\mathrm{B}][\mathrm{D}] + (k_{\text{-ES}} + k_{\mathrm{P}})[\mathrm{BD}] \qquad [2.7b]$$

$$\frac{d[\mathrm{BD}]}{dt} = k_{\mathrm{ES}}[\mathrm{B}][\mathrm{D}] - (\beta_{\mathrm{D}} + k_{\text{-ES}} + k_{\mathrm{P}} + \beta_{\mathrm{dil}})[\mathrm{BD}] \qquad [2.8]$$

The parameter values used to simulate the dynamics of the three step cascade *with* feedback were identical to those used in the three step cascade *without* feedback unless otherwise stated. The additional parameters values that were not in the original three step cascade without feedback model (*i.e.* k_f, k_{ES}, k_{-ES} and k_P) were determined as follows. We estimated k_{ES} (the association rate constant for the binding of CsrB to CsrD)

to be the same as the association rate constant between CsrB and CsrA ($k_{ES} = k_2 = 10^{-1}$ $nM^{-1} \cdot s^{-1}$). k_P, k_{-ES}, k_f, and α_A were obtained by simultaneously fitting the model to the dynamic behavior of the experimental results in **Fig. 2.7*B***, ***C***, ***E***, ***F*** **&** ***G***. Identical values for k_P and k_{-ES} were used for all simulations in **Fig. 2.7**. The value for the rate constant for the conversion of CsrB-CsrD complex into free CsrD and degraded CsrB ($k_P = 8 \times 10^{-2}$ s^{-1}) falls within the experimentally measured range reported for RNase E (1.4 to 3 x 10^{-2} s^{-1}) [75,76]. Our value for the rate constant for CsrB dissociation from CsrD (k_{-ES}) was 2 x 10^{-2} s^{-1}, which is comparable to reported RNA-protein dissociation rate constants [79-81]. k_f pertains only to the simulations demonstrating the effects of feedback (dark and light blue lines in **Fig. 2.7*E*** and gold line in **Fig. 2.7*G***); its value (k_f = 30 nM) was identical in all three simulations and was comparable to the levels of free CsrA observed in the model. We estimate that the level of CsrA produced by the native *csrA* gene to be approximately one tenth the concentration that we expressed in our synthetic system (**Fig. 2.2*D***). Therefore $\alpha_A = 8 \times 10^{-2}$ $nM \cdot s^{-1}$ for the native *csrA* gene (0.1x CsrA, **Fig. 2.7*E*, Fig. 2.8**) and $\alpha_A = 8 \times 10^{-1}$ $nM \cdot s^{-1}$ for the synthetic *csrA* gene (1x CsrA, **Fig. 2.7*E*, Fig. 2.8**).

Simulations of the three step cascade with feedback using the above parameters were able to describe the dynamic behavior of the system. In the models we turned *csrB* transcription either on (**Fig. 2.7*B-F***) or off (**Fig. 2.7*G***) under different combinations of conditions. These conditions include two different CsrA concentrations: (i) native CsrA levels ($\alpha_A = 8 \times 10^{-2}$ $nM \cdot s^{-1}$; light blue line in **Fig. 2.7*E*** and gold line in **Fig. 2.7*G***) or (ii) synthetic CsrA levels ($\alpha_A = 8 \times 10^{-1}$ $nM \cdot s^{-1}$; dark blue line in **Fig. 2.7*E*** and magenta line in **Fig. 2.7*F***, and black line in **Fig. 2.7*E-G***). We also modeled three possible regulatory patterns for CsrD: (i) no expression ($\alpha_D = 0$; black control curve in **Fig. 2.7*E-G***), (ii) low induced expression without feedback ($\alpha_D = 1.16 \times 10^{-2}$ $nM \cdot s^{-1}$; magenta curve in **Fig. 2.7*F***) and (iii) native expression with feedback ($\alpha_D = 1.6 \times 10^{-1}$ $nM \cdot s^{-1}$; gold line in **Fig. 2.7*B*** **&** ***G***, orange line in **Fig. 2.7*C***, and blue line in **Fig. 2.7*E***). In the case of the low induced expression of CsrD without feedback (**Fig. 2.7*F***), the term $\alpha_D \cdot (k_f/(k_f + [A]))$ in the model was replaced with just α_D. The low expression of CsrD was achieved experimentally (**Fig. 2.7*F***) from PLlacO-1 by adding 20 μM IPTG to the media rather than the 0.5-1 mM IPTG required for full transcription (**Fig. 2.10*Q***). The maximum production rates for *glgC-gfp* mRNA (α_m), GlgC-GFP protein (α_{gfp}), and CsrB (α_B) are the same in **Fig. 2.7*E-G*** as described in prior sections.

2.5.4.5 The Minimal Effects of Cooperative CsrA Binding

In our model, CsrA binds as a dimer to its target mRNAs and non-coding RNAs. This feature of the model is based on experimental evidence showing that CsrA primarily exists in the cell as a dimer [52]. It has also been shown experimentally that an individual CsrA subunit can facilitate the binding of its dimer partner to CsrB [70]; this phenomenon is included in our model in that we assume that both CsrA molecules in the dimer bind together to CsrB. One CsrA molecule facilitating the binding of the other CsrA molecule within the *same* dimer is distinct from cooperative binding *between* CsrA dimers in which the binding of one CsrA dimer promotes the binding of another CsrA dimer to CsrB or an mRNA. There is one report that may indicate cooperative binding of CsrA dimers to CsrB [49], however many details remain unclear. The binding of CsrA dimers to the CsrB non-coding RNA is likely to be very different to that of a small molecule binding a protein. When a small molecule binds to a protein (*e.g.* dioxygen binding to hemoglobin) it can induce a conformational change in the protein that alters small molecule binding at other sites on the protein. For the CsrB non-coding RNA, there is no described mechanism by which the binding of a CsrA dimer at one location on CsrB enhances the binding of other CsrA dimers at other locations on CsrB; although it is of course possible that this does occur.

In this section, we explore how cooperative binding between CsrA and CsrB would affect the transfer function (**Fig. 2.9*G***) and dynamic behavior (**Fig. 2.9*H***) of our system. To evaluate the effects of cooperativity in CsrB binding on the CsrB transfer function we compared a model where each site for a CsrA dimer was bound independently (*i.e.* "non-cooperative") to a model where the binding of each dimer increased the equilibrium association constant for the binding of the next CsrA dimer by 2-fold (*i.e.* "cooperative"). In the non-cooperative model, there are 9 sites for CsrA dimers on each CsrB molecule and the binding of each CsrA dimer to each site on CsrB is independent. In the cooperative model, there are also 9 sites for CsrA dimers on each CsrB molecule; however in this case the binding of the first CsrA dimer increases the association rate constant of the second CsrA dimer by two-fold which increases the association rate constant of the third CsrA dimer by a further two-fold, and so on until the ninth and last CsrA dimer has an association rate constant that is 2^8-fold greater than that of the first CsrA dimer. In the non-cooperative and cooperative models, we varied the total CsrB concentration and measured the amount of free target mRNA in the presence of fixed total CsrA (**Fig. 2.9*G***). Cooperativity is often considered from the perspective of its effect on the steepness of an input-output relationship. In this case, we will quantify the

effect of cooperative binding on the change in the free target mRNA concentration (output) as the CsrB concentration (input) increases. We measured the steepness of each function using the "response coefficient" (*i.e.* the input value that yields 90% of maximum output divided by the input value that yields 10% of maximum output) and then converted these values to an equivalent Hill coefficient as described by Goldbeter and Koshland [1].

We found that the extreme example of cooperative binding described above only increased the Hill coefficient for the CsrB transfer function from 3.7 (for non-cooperative binding) to 4.5 (cooperative binding) (**Fig. 2.9*G***). By comparison, a two-fold increase in the total CsrA concentration in the system with non-cooperative binding increased the Hill coefficient from 3.7 to 5.0. Cooperativity has such a comparatively small effect on the transfer function because the stoichiometric point (where the number of CsrA dimers equals the number of available sites for those dimers on CsrB) creates a sharp transition even in the absence of cooperativity. That is, when the concentration of CsrB sites [B] is less than the CsrA dimer concentration [A] there is substantial silencing of the target mRNA; on the other hand, when the concentration of CsrB sites [B] is greater than the CsrA dimer concentration [A] there is minimal silencing of the target mRNA. The sharpness of this transition across the stoichiometric point (from silencing to non-silencing) depends on the equilibrium dissociation constants for CsrA dimers binding to CsrB sites, and on the absolute concentrations of CsrA dimers and CsrB sites. The transfer function for CsrB is very different from that of a protein acting "catalytically" (*e.g.* an enzyme or transcription factor) where there is no stoichiometric point and cooperative binding can have a much greater impact on the steepness of the transfer function.

We next compared signaling dynamics using a version of our two level cascade model that was modified to incorporate cooperative and non-cooperative binding as described above. Using this model, we turn on and off *csrB* transcription and monitor target protein levels (**Fig. 2.9*H***) as we did previously (**Fig. 2.4*B*, *C***). The dynamic behavior of the system with cooperative binding was found to be very similar to the system with non-cooperative binding (**Fig. 2.9*H***). Note: cooperative binding implemented by changing dissociation rate constants (rather than association rate constants) produced dynamics traces that are visually indistinguishable from those produced by changing the association rate constants (**Fig. 2.9*H***).

Part 2: Signaling dynamics in the CsrA system

In summary, our model demonstrates that the qualitative dynamic behavior of our system can be explained without incorporating cooperative binding. Furthermore, our simulations show that if the binding of CsrA dimers is cooperative it is unlikely to have a substantial qualitative effect on the transfer functions or dynamic behavior of the CsrA system.

Table 2.3 | Plasmids and strains

*ID**	*Description*
pHL177	contains chloramphenicol resistance cassette flanked by FRT sites
pHL600	PLtetO-1:*csrB*, PLlacO-1:RBS(st7):*csrA*
pHL661	PLtetO-1:RBS(st7):*gfp*, PLlacO-1:RBS(st7):*mCherry*
pHL662	PLtetO-1:RBS(st7):*mCherry*, PLlacO-1:RBS(st7):*gfp*
pHL1318	PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st2):*csrA*
pHL1335	PconNoHindM12:*glgC*::*gfp*†, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st2):*csrA*
pHL1355	PconNoHindM12:*glgC*::*gfp*†, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st3):*csrA*
pHL1490	PLtetO-1:*csrB*, PLlacO-1:RBS(st3):*csrD*
pHL1506	PLlacO-1:RBS(st7):*csrA*
pHL1529	PLlacO-1:*glgC*::*gfp*†, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st2):*csrA*
pHL1530	PLlacO-1:*glgC*::*gfp*†, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st3):*csrA*
pHL1559	PLtetO-1:*csrB*
pHL1561	PLtetO-1:RBS(st7):*csrA*
pHL1575	PconNoHind:*csrB*, PLtetO-1:RBS(st3):*csrD*
pHL1756	PconNoHindM12:*glgC*::*gfp*†, PconNoHind:RBS(st3):*tetR*
pHL1757	PconNoHindM2:*glgC*::*gfp*†, PconNoHind:RBS(st3):*tetR*
pHL1801	PLlacO-1:RBS(st7):*gfp*
pHL1853	PLlacO-1:*glgC*::*gfp*::*LVA*†, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st2):*csrA*
pHL1855	PLlacO-1:*glgC*::*gfp*::*LVA*†, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st3):*csrA*
HL3721	HL716‡ + *ΔcsrB* + *ΔcsrC* + *ΔglgCAP* + *ΔpgaABCD*
HL3796	HL716‡ + *ΔcsrA* + *ΔcsrB* + *ΔcsrC* + *ΔglgCAP* + *ΔpgaABCD*
HL4018	HL716‡ + *ΔcsrB* + *ΔcsrC* + *ΔcsrD* + *ΔglgCAP* + *ΔpgaABCD*
HL4142	HL716‡ + *ΔcsrA* + *ΔcsrB* + *ΔcsrC* + *ΔcsrD* + *ΔglgCAP* + *ΔpgaABCD*
HL4495	HL4142 + pHL1335 + pHL600
HL4574	HL4142 + pHL1355 + pHL1490
HL4845	HL4142 + pHL1530 + pHL1559
HL4860	HL4142 + pHL1529 + pHL1561
HL4874	HL4142 + pHL1530 + pHL1575
HL4509	HL4142 + pHL1318 + pHL661
HL4510	HL4142 + pHL1318 + pHL662
HL5561	MG1655 + pHL1756
HL5562	MG1655 + pHL1757
HL5582	HL4142 + pHL1506 + pHL1756
HL5591	HL4142 + pHL1506 + pHL1757
HL5593	HL4142 + pHL1756
HL5594	HL4142 + pHL1757
HL5595	HL4018 + pHL1756
HL5596	HL4018 + pHL1757

HL5814	HL4142 + pHL1561 + pHL1853
HL5815	HL4142 + pHL1559 + pHL1855
HL5816	HL4142 + pHL1575 + pHL1855
HL5840	HL716‡ + *ΔpgaABCD*
HL5860	HL5840 + pHL1559 + pHL1756
HL5861	HL4142 + pHL1559 + pHL1355
HL5876	HL4018 + pHL1559 + pHL1756
HL5877	HL3796 + pHL1559 + pHL1355
HL5878	HL3721 + pHL1559 + pHL1756
HL5944	HL3796 + pHL1801
HL5947	HL4142 + pHL1561

Table 2.3 | Plasmids and strains. Selected plasmids and strains were submitted to the Addgene repository (www.addgene.org). *ID numbers beginning with "pHL" correspond to plasmids; ID numbers beginning with "HL" correspond to strains. †To construct the *glgC*::*gfp* reporter gene, the DNA sequence of the *glgCAP* mRNA between -61 and +8 nucleotides (relative to the *glgC* start codon) was translationally fused to the *gfp* coding region. ‡This strain is MG1655 with *lacI*$_q$ added to the chromosome [68].

2.5.5 Bacterial Strains and Plasmids

Bacterial plasmids and strains are listed in **Table 2.3**. Oligonucleotide sequences for knockouts, plasmid construction and RT-PCR are listed in **Table 2.4**. Plasmid maps are in **Fig. 2.11**. The chromosomal genes and operons *csrA*, *csrB*, *csrC*, *csrD*, *glgCAP* and/or *pgaABCD* were deleted from the chromosome using the λ-red method [82] or removed from their host strain via phage transduction. All deletions were confirmed by PCR. The *csrA* deletion was performed after *glgCAP* was deleted because CsrA is essential for cell survival when *glgCAP* is present [32]. CsrA was also deleted only after *pgaABCD* was deleted because in the absence of CsrA, cells with the *pgaABCD* operon overproduced biofilm adhesins [57] which prevented their resuspension and further genetic manipulation.

Table 2.4 | Oligonucleotides

name	*description*	*sequence*
csrAKOpkD1F	For deletion of *csrA* using pHL177 as template	TGCCGGGATACAGAGAGACCCGACTCTTTTAATCTTTCAAGGAGCAAAGAGTGT AGGCTGGAGCTGCTTC
csrAKOpkD4R	For deletion of *csrA* using pHL177 as template	GGAGAAATTTTGAGGGTGCGTCTCACCGATAAAGATGAGACGCGGAAAGAATTC CGGGGATCCGTCGACC
csrBKOpkD1F	For deletion of *csrB* using pKD13 as template	AGCGCCTTGTAAGACTTCGCGAAAAAGACGATTCTATCTTCGTCGACAGGGTGT AGGCTGGAGCTGCTTC
csrBKOpkD4R	For deletion of *csrB* using pKD13 as template	GTGGTCATAAAGCAACCTCAATAAGAAAAACTGCCGCGAAGGATAGCAGGATTC CGGGGATCCGTCGACC
csrCKOpkD1F	For deletion of *csrC* using pKD13 as template	ACTGATGGCGGTTGATTGTTTGTTTAAAGCAAAGGCGTAAAGTAGCACCCGTGT AGGCTGGAGCTGCTTC
csrCKOpkD4R	For deletion of *csrC* using pKD13 as template	GCCGTTTTATTCAGTATAGATTTGCGGCGGAATCTAACAGAAAGCAAGCAATTC CGGGGATCCGTCGACC
csrDpkD1F	For deletion of *csrD* using pKD13 as template	ATCTGATTTGCTAGTATGCCCGCTTCCTCACTATCGGAGTTAACACAAGGGTGT AGGCTGGAGCTGCTTC
csrDpkD4R	For deletion of *csrD* using pKD13 as template	CATGAGACGCAGCGCGCATTATTCTACGTGAAAACGGATTAAACGGCAGGATTC CGGGGATCCGTCGACC
glgCpkD1F	For deletion of *glgCAP* using pKD13 as template	CCTGCACACGGATTGTGTGTGTTCCAGAGATGATAAAAAAGGAGTTAGTCGTGT AGGCTGGAGCTGCTTC
glgPpkD4R	For deletion of *glgCAP* using pKD13 as template	TTACAATCTCACCGGATCGATATGCCAGATATGATCGGCGTACTCTTTGAATTC CGGGGATCCGTCGACC
pgaApkD4R	For deletion of *pgaABCD* using pKD13 as template	CTGTAATTAGATACAGAGAGAGATTTTGGCAATACATGGAGTAATACAGGATTC CGGGGATCCGTCGACC
pgaDpkD1F	For deletion of *pgaABCD* using pKD13 as template	AGTGTGTTATCGGTGCAGAGCCCGGGCGAACCGGGCTTTGTTTTGGGTGTGTGT AGGCTGGAGCTGCTTC
csrARBSXmaF	PCR amplifies *csrA* with a synthetic RBS (st7)	CCTCCCGGGTAAGGAGGAAAAAAAATGCTGATTCTGACTCGTCGAGTTG
csrAKpnHindR	PCR amplifies *csrA*	GGCCAAGCTTCTTTCAGGTACCTTAGTAACTGGACTGCTGGGATTTTTCAG
csrBSalF	PCR amplifies *csrB*	CAAGTCGACGAGTCAGACAACGAAGTGAACATC
csrBApaR	PCR amplifies *csrB*	CATGGGCCCAATAAAAAAAGGGAGCACTGTATTCACAGC
csrARBS2NotIF	PCR amplifies *csrA* with a synthetic RBS (st2)	TCCTGCGGCCGCTAAGGAGGAAATGCTGATTCTGACTCGTCGAGTTG
csrARBS3NotIF	PCR amplifies *csrA* with a synthetic RBS (st3)	TCCTGCGGCCGCTAAGGAGGAAAATGCTGATTCTGACTCGTCGAGTTG
csrAApaR	PCR amplifies *csrA*	TAAGGGCCCTTAGTAACTGGACTGCTGGGATTTTTCAG
glgCleadSalF	PCR amplifies 5'UTR of *glgC* leader for fusion to	CCTGTCGACTCTGGCAGGGACCTGCACACGGATTG
glgCleadSphR	PCR amplifies 5'UTR of *glgC* for fusion to *gfp*	TACGCATGCTAACCATGACTAACTCCTTTTTTATCATCTCTGG
GFPRBSSalSphF	PCR amplified *gfp* with *Sal*I & *Sph*I sites	TTAGTCGACTAAGGAGGAAAAAGCATGCGTAAAGGAGAAGAACTTTTC
PconNoHindBamHF	PCR synthesis of Pcon promoter with no *Hin*dIII	CGCGGATCCTCGAGCACCGTCGTTGTTGACATTTTTATGCTTGGCGGTTATAAT

PconNoHindXmaR	PCR synthesis of Pcon promoter with no *Hin*dIII	CCTCCCGGGTGTGTGGAATCCATTATAACCGCCAAGCATAAAAATGTCAACAAC
PconNoHindEcoRF	PCR synthesis of Pcon promoter with no *Hin*dIII	CCGGAATTCTCGAGCACCGTCGTTGTTGACATTTTTATGCTTGGCGGTTATAAT
PconNoHindNotlR	PCR synthesis of Pcon promoter with no *Hin*dIII	TCCTGCGGCCGCCTGTGTGGAATCCATTATAACCGCCAAGCATAAAAATGTCAACAAC
PconM2NoHindAatF	PCR synthesis of PconM2 promoter with no *Hin*dIII	CGCGACGTCTCGAGCACCGTCGTTGTTTACATTTTTATGCTTGGCGGTTATGAT
PconM2NoHindSalR	PCR synthesis of PconM2 promoter with no *Hin*dIII	TTAGTCGACCTGTGTGGAATCCATCATAACCGCCAAGCATAAAAATGTAAACAAC
PconM8NoHindAatF	PCR synthesis of PconM12 promoter with no *Hin*dIII	CGCGACGTCTCGAGCACCGTCGTTGTTTACATTTTTATGCTTGGCGGTTATGGT
PconM12NoHindSalR	PCR synthesis of PconM12 promoter with no *Hin*dIII	TTAGTCGACCTGTGTGGAATCCACCATAACCGCCAAGCATAAAAATGTAAACAAC
PconNoHindBamHF	PCR synthesis of Pcon promoter with no *Hin*dIII	CGCGGATCCTCGAGCACCGTCGTTGTTGACATTTTTATGCTTGGCGGTTATAAT
csrDRBS3NotIF	PCR amplifies *csrD* with a synthetic RBS (st3)	TCCTGCGGCCGCTAAGGAGGAAAATGAGATTAACGACGAAATTTTCG
csrDHindR	PCR amplifies *csrD*	GCCAAGCTTTTAAACCGAGTATCTTTGTGAATA
csrDinF	For RT-PCR measurement of *csrD* mRNA	CTGGCGCGTTACCACCGCAGTGAT
csrDinR	For RT-PCR measurement of *csrD* mRNA	CCAATGTGGATCATATCGTCGCGA

Table 2.4 | Oligonucleotides. Oligonucleotides were used (i) to construct the plasmids, (ii) to create PCR products for the deletion of CsrA system components from the chromosome or (iii) to perform quantitative RT-PCR measurements of intracellular mRNA concentrations.

PconNoHind, PconNoHindM2 and PconNoHindM12 are variants of Pcon/O3 [55] without a *Hin*dIII site. The PconNoHindM2 promoter is the same as PconNoHind except for a substitution mutation at the -10 site that modestly increases transcription. The PconNoHindM12 promoter is the same as PconNoHind except for two substitution mutations at the -10 and -35 sites that moderately decrease transcription. The PLlacO-1 and PLtetO-1 promoters and T1 terminators were obtained from the pZ system [54]. The st2, st3 and st7 RBS sequences [83] were synthesized. The *gfp* gene and the T1T2 terminator sequence were obtained from pTAK102 [4]. The Asp terminator sequence was PCR-amplified from pLex (Invitrogen). The sources and sequences for *tetR* and *lacIq* have been reported [21,22]. *csrA*, *csrB*, *csrC* and *csrD* genes were PCR amplified from *E. coli* MG1655 chromosomal DNA (Yale E. coli Stock Center, CGSC #7740). The *mCherry* gene was amplified from a plasmid provided by R. Tsien (University of California, San Diego, CA) [84].

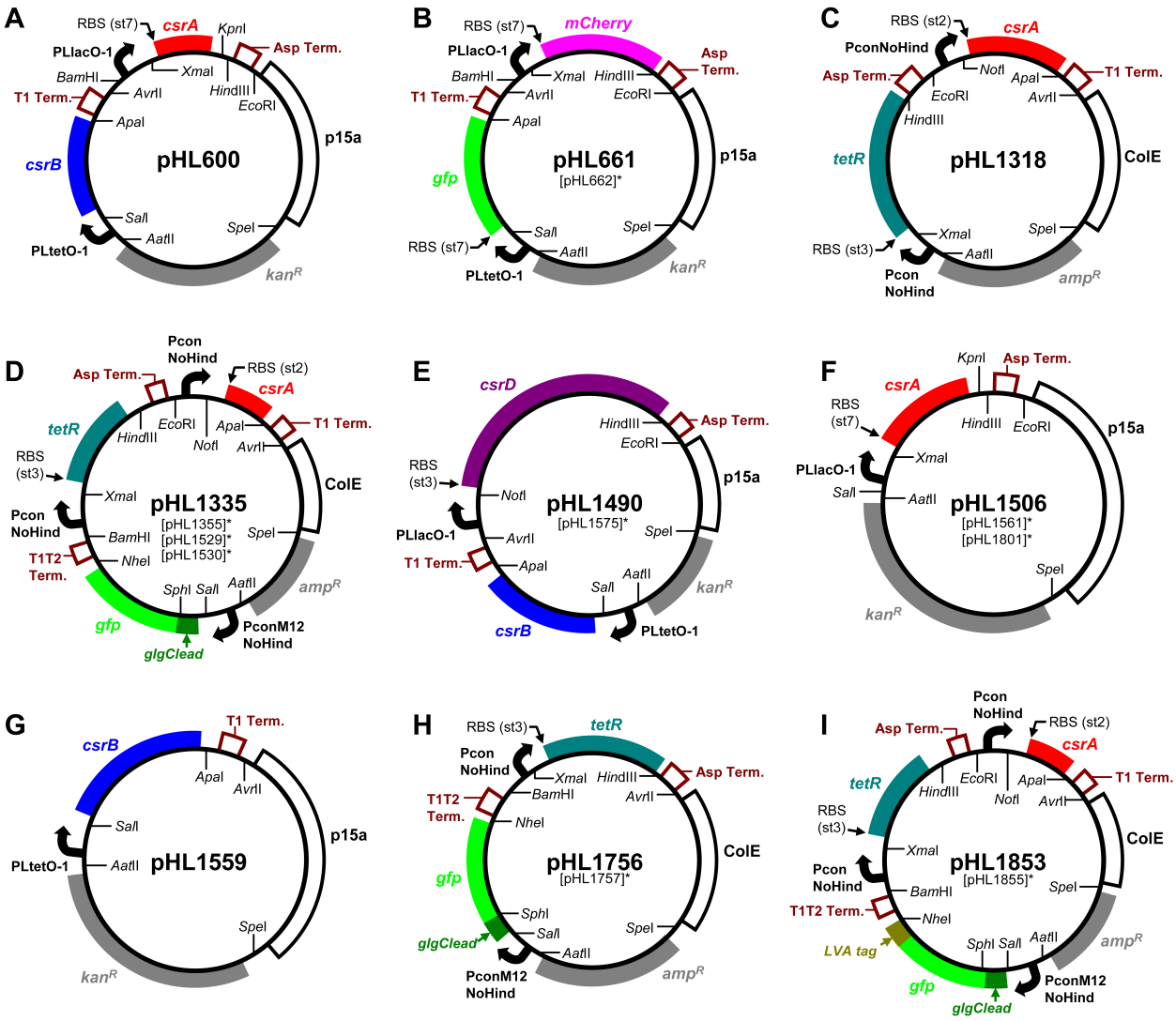

Fig. 2.11 | Plasmid maps. *Plasmids listed in brackets are similar to the plasmid shown except for the differences described here. Relative expression levels from the different ribosome binding sequence (RBS) used: st7 > st3 > st2. p15a and ColE are origins of replication. T1 term, T1T2 Term and Asp Term are terminator sequences. (***A***) pHL600. (***B***) pHL661. In pHL662 the st7-*gfp* replaces st7-*mCherry* and st7-*mCherry* replaces st7-*gfp*. (***C***) pHL1318. (***D***) pHL1335. In pHL1355, the st3 RBS replaces st2 for *csrA*. In pHL1529, PLlacO-1 replaces PconNoHindM12. In pHL1530, st3 replaces st2 and PLlacO-1 replaces PconNoHindM12. (***E***) pHL1490. In pHL1575, PconNoHind replaces PLtetO-1 and PLtetO-1 replaces PLacO-1. (***F***) pHL1506. In pHL1561, PLtetO-1 replaces PLlacO-1 (*SalI* is removed). In pHL1801, *gfp* replaces *csrA*. (***G***) pHL1559. (***H***) pHL1756. In pHL1757, a stronger PconM2NoHind promoter replaces PconM12NoHind. (***I***) pHL1853. In pHL1855, st3 replaces st2 for *csrA*.

3. Crosstalk among Hfq-dependent small RNAs

Having examined in detail how RNA regulators and their associated gene networks can accelerate and otherwise control the speed of signal propagation, we now examine how these same RNA regulators and networks control crosstalk and manage interference among signaling pathways that share common molecules.

As discussed in the introduction, shared components can become a source of competition for signaling molecules that require them. Through this competition, otherwise distinct signaling pathways can influence one another's activity. While some crosstalk in biological networks can be useful for signal integration and coordination, indiscriminate crosstalk degrades signal specificity and effectiveness. In the following study, crosstalk and signaling robustness are examined in the context of the Hfq-dependent small RNA network and a framework is constructed to characterize the kinetic constraints on the operation of Hfq.

3.1 ABSTRACT

Bacteria possess networks of small RNAs (sRNAs) that are important for modulating gene expression. At the center of many of these sRNA networks is the Hfq protein. Hfq's role is to quickly match cognate sRNAs and target mRNAs from among a large number of possible combinations and anneal them to form duplexes. Here we show using a kinetic model that Hfq can efficiently and robustly achieve this difficult task by minimizing the sequestration of sRNAs and target mRNAs in Hfq complexes. This sequestration can be reduced by two non-mutually exclusive kinetic mechanisms. The first mechanism involves heterotropic cooperativity (where sRNA and target mRNA binding to Hfq is influenced by other RNAs bound to Hfq); this cooperativity can selectively decrease singly-bound Hfq complexes and ternary complexes with non-cognate sRNA-target mRNA pairs while increasing cognate ternary complexes. The second mechanism relies on frequent RNA dissociation enabling the rapid cycling of sRNAs and target mRNAs among different Hfq complexes; this increases the probability the cognate ternary complex forms before the sRNAs and target mRNAs degrade. We further demonstrate that the performance of sRNAs in isolation is not predictive of their performance within a network. These findings highlight the importance of experimentally characterizing duplex formation in physiologically

relevant contexts with multiple RNAs competing for Hfq. The model will provide a valuable framework for guiding and interpreting these experiments.

3.2 AUTHOR SUMMARY

Bacteria have small RNAs (sRNAs) which are important modulators of gene expression. Many of these sRNAs require the Hfq protein to mediate their binding to specific target mRNAs which alters the translation and/or degradation of the mRNAs. The Hfq protein has a difficult task; it has to correctly pair cognate sRNAs and target mRNAs from among a large number of possible combinations and anneal them before the RNAs degrade. Furthermore, the process must be robust to changes in the number and types of sRNAs and target mRNAs that are transcribed and changes in the Hfq concentration. Here we show that Hfq can most successfully achieve its task when sRNAs and target mRNAs are not unnecessarily sequestered in Hfq complexes. The cell can accomplish this via cooperative binding of sRNAs and target mRNAs to Hfq and/or by rapid RNA dissociation from Hfq complexes. These findings reveal the requirements for efficient and robust sRNA signaling which are important for understanding the regulation of gene expression in diverse cell processes, for devising strategies that inhibit Hfq activity during pathogenesis and for the rational construction of synthetic circuits.

3.3 INTRODUCTION

Small RNAs (sRNAs) regulate a wide variety of pathways in prokaryotes [85]. An important subset of these small RNAs act in *trans* with the aid of the Hfq protein to decrease ("silencing") or increase ("activation") the expression of specific target mRNAs. These *trans*-acting, Hfq-dependent sRNAs, which have important roles in the cellular response to stress and the virulence of major pathogens [86-90], are the focus of this study.

sRNAs typically function by binding to target mRNAs at or near the site of the ribosome binding sequence (RBS) [85]. This results in sRNA-target mRNA duplexes which decrease or less commonly increase the translation of mRNAs. The decreased mRNA translation can be accompanied by an increase in mRNA degradation. The binding between a sRNA and its cognate target mRNA is sequence specific. However, this does not mean that each sRNA can only bind to one target mRNA; a sRNA can act

on multiple target mRNAs and a target mRNA can have binding sites for more than one sRNA.

The Hfq protein, which was originally identified as an essential host factor for the replication of the bacteriophage Qβ [91], plays an important role in bringing many sRNAs and target mRNAs together and assisting their annealing. Hfq is a small 11 kDa protein which forms stable cyclic homo-hexamers [92] that have a "Proximal face" and a "Distal face". The Proximal face binds uridine rich sequences and the distal face binds poly(A) tracts and poly(A-R-N) repeats, where R is a purine nucleotide and N is any nucleotide [93]. Competition studies indicate substantial overlap in the binding sites for sRNAs and target mRNAs on Hfq and they indicate interactions between the RNAs bound to these sites [94,95]. Hfq also binds to proteins including RNase E [96], polynucleotide phosphorylase (PNPase) [97] and ribosomal subunit S1 [98]. In addition to its role in mediating sRNA activity, Hfq also binds DNA [99,100] and regulates the degradation of polyadenylated mRNAs [97,101].

In many sRNA-target mRNA pairs, both members can bind free Hfq. This has been demonstrated in co-immunoprecipitation studies [102,103], *in vitro* Hfq binding assays [104,105] and with *in vivo* competition studies [21]. The affinity of the sRNA and the target mRNA for free Hfq in many pairs appears to be comparable [80,101,104] (although the RyhB-sodB pair appears to be an exception [105]). Therefore most sRNAs have two potential paths to duplex formation; one where the sRNA binds to free Hfq followed by target mRNA binding ("sRNA-Hfq branch") and another where the target mRNA binds to free Hfq followed by sRNA binding ("target mRNA-Hfq branch") (**Fig. 3.1**).

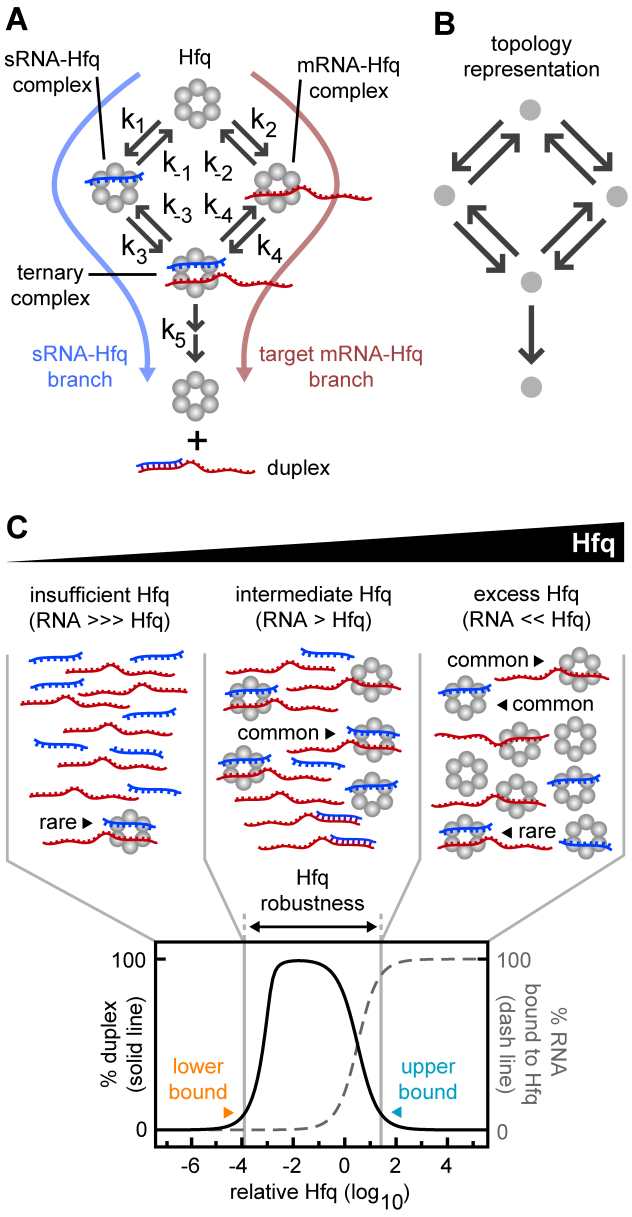

Fig 3.1 | General reaction scheme for Hfq dependent duplex formation. (A) A kinetic model of Hfq-dependent duplex formation showing two paths to duplex formation. The annealing of the sRNA and target mRNA and the release of the duplex from Hfq are treated as a single step. (B) The kinetic model as a simplified topology representation. (C) Percentage duplex formation and percentage of RNA bound to Hfq at different concentrations of Hfq. The illustration (top) shows mechanistically why insufficient and excess Hfq result in decreased duplex formation. At low concentrations of Hfq, the formation of cognate ternary complexes is limited by the number of Hfq hexamers. At high concentrations of Hfq, the formation of cognate ternary complexes is limited because the probability of a sRNA and its cognate target mRNA binding to the same Hfq hexamer is low. That is, sRNA and target mRNA molecules are sequestered from one another on separate Hfq complexes. The lower and upper bounds indicate the minimum and maximum Hfq concentrations respectively that result in at least 10% duplex formation.

Part 3: Crosstalk among Hfq-dependent small RNAs

Once the sRNA and its cognate target mRNA are bound to Hfq (forming a cognate sRNA-Hfq-target mRNA ternary complex), Hfq can promote duplex formation by providing a structure for strand exchange to take place [106,107] or by acting as a chaperone that alters the sRNA and target mRNA structures to expose sites necessary for annealing [108,109]. Early *in vitro* evidence suggested that the conformational change in the bound RNAs and annealing was slow and required ten minutes or more to occur [108]. However, more recently it was observed that a partial duplex can form and be released from Hfq with the whole process taking seconds rather than minutes [74]. The basis for the discrepancy between the two studies is unclear but the latter is consistent with *in vivo* experiments which have shown target mRNA silencing occurring within three minutes of sRNA induction [110]. Once the duplex is formed it can be released or remain bound to Hfq while it is degraded or translated (the latter will decrease the availability of free Hfq).

Experimental and theoretical studies have typically focused on duplex formation for individual sRNA-target mRNA pairs in isolation [47,48]. However, most sRNAs act within a network with dozens of different sRNAs and target mRNAs competing for Hfq [111]. Furthermore, the network is not static but changes its composition of sRNAs, target mRNAs and the amount of Hfq in response to environmental conditions [112-114]. Because the actions of sRNAs are so interdependent due to their shared need for Hfq, these changes in the network's composition can dramatically alter sRNA activity [21]. In this study we sought to address the fundamental question of how a large network with many ligands (sRNAs and target mRNAs) competing for a single protein (Hfq) can function efficiently and robustly.

In the first part of the study we modeled the kinetics of duplex formation for a single cognate sRNA-target mRNA pair. We identified two non-mutually exclusive mechanisms that can increase the efficiency of sRNA signaling: 1. heterotropic cooperativity for the binding of sRNAs and/or target mRNAs to Hfq (we simply refer to this as "cooperativity"); and 2. frequent RNA dissociation. These mechanisms increase duplex formation by reducing the sequestration of sRNAs and target mRNAs in singly-bound Hfq complexes. In the second part of the study we show the same two mechanisms also promote signaling in sRNA networks with many sRNAs and target mRNAs competing for Hfq. In this case, cooperativity and/or frequent RNA dissociation also decrease the sequestration of sRNAs and target mRNAs in non-cognate ternary complexes (*i.e.* where Hfq is bound by a sRNA and a target mRNA that

do not form a cognate duplex). These mechanisms make duplex formation more efficient as well as more robust to changes in the Hfq concentration and the composition of the network.

3.4 RESULTS

3.4.1 A General Model of Hfq Kinetics

Current evidence indicates that sRNAs and target mRNAs have separate binding sites on Hfq as well as shared sites. We simply assumed that sRNAs and target mRNAs have separate sites, which provides a conservative estimate of the difficulties faced by Hfq. However, the qualitative findings of this study are still applicable when sRNAs and target mRNAs compete for shared sites on Hfq (see Discussion). There are two possible paths to duplex formation; one where the sRNA binds first to Hfq followed by the target mRNA ("sRNA-Hfq branch") and another where the target mRNA binds first to Hfq followed by the sRNA ("target mRNA-Hfq branch") (**Fig. 3.1**). The reaction scheme, which may be a random order *bi uni* or compulsory order *bi uni* enzymatic reaction [115], can be topologically represented as a graph with weighted, directed arrows indicating the relative magnitude of the rate constant for each reaction (**Fig. 3.1*B***).

The complete reaction scheme for a single sRNA-target mRNA pair has three categories of rate constants: 1. association rate constants which describe the binding of sRNAs and target mRNAs to free Hfq or an Hfq complex (k_1, k_2, k_3 and k_4 with units of concentration^{-1}·time^{-1}); 2. dissociation rate constants for the unbinding of sRNAs and target mRNAs from Hfq complexes (k_{-1}, k_{-2}, k_{-3} and k_{-4} with units of time^{-1}); and 3. a "duplex" rate constant which is an overarching constant for the steps involved in sRNA-target mRNA annealing and the release of the duplex from Hfq (k_5 with units of time^{-1}). While duplexes can rebind Hfq *in vitro* [116], *in vivo* many duplexes are rapidly degraded [85] and therefore we do not include duplex rebinding. The fraction of total target mRNA converted to a specific cognate duplex at steady state is the measured output of the pathway; this output does not depend on whether the sRNA is silencing or activating gene expression.

We varied the rate constants to generate reaction schemes with different topologies. Unless otherwise stated, we kept the production and degradation of the sRNAs and target mRNAs constant and equal (unless otherwise stated) and varied the total Hfq concentration by altering its production (a typical plot is shown in **Fig. 3.1*C***). That is,

the concentration of Hfq varies relative to the total target mRNA. Therefore when the "relative Hfq" = 1 (*i.e.* 10^0), it indicates the concentration of Hfq in all forms is equal to the total concentration of target mRNA for the cognate pair being measured (which includes free target mRNA, target mRNA bound to Hfq and target mRNA in duplexes). The minimum and maximum relative Hfq concentrations that permit at least 10% duplex formation, were termed the "lower bound" and the "upper bound" respectively. The logarithmic range of Hfq concentrations over which at least 10% of target mRNA is converted to duplex (*i.e.* the fold-difference between the upper and lower bounds) is termed "Hfq robustness". Hfq robustness is a useful overall measure of how efficiently the system copes with changes in Hfq, sRNA and target mRNA concentrations and with competition for Hfq.

At the lower bound, the Hfq concentration is insufficient to process the quantity of sRNAs and target mRNAs present (**Fig. 3.1*C***). Typically the lower bound is determined by the Hfq recycling rate; that is, the rate at which free Hfq is converted to ternary Hfq complex and then to free duplex and free Hfq. At the upper bound, the sRNA and target mRNA concentrations are low compared to Hfq. Therefore there is a low probability that the sRNA and its cognate target mRNA will bind to the same Hfq hexamer to form the ternary complex, and a high probability they will bind to separate Hfq hexamers to form singly-bound complexes (sRNA-Hfq and target mRNA-Hfq). The upper bound is consequently a measure of the susceptibility of a pathway to sequester sRNAs and target mRNAs in singly-bound Hfq complexes. It has been shown *in vitro* that relatively high Hfq concentrations do indeed increase singly-bound Hfq complexes and reduce duplex formation [80,104].

To keep the model as simple as possible, the rate constant β (equal to 1 unit of $time^{-1}$) for degradation and dilution is the same for all species. While this simplification does not reflect the relative degradation rates in biological systems (*e.g.* sRNAs often have longer half-lives when bound to Hfq), it does not alter the basic qualitative results. This was demonstrated by showing that a 10-fold greater degradation rate constant for the free sRNA compared to the sRNA bound to Hfq had minimal effect on the behavior of the system with several different kinetic schemes (**Fig. 3.14**). The reason the free sRNA degradation rate has a minor effect at high Hfq concentrations is that most of the sRNA is bound to Hfq. At low Hfq concentrations, the free sRNA degradation has minimal effect because there is insufficient Hfq to bind all the sRNA.

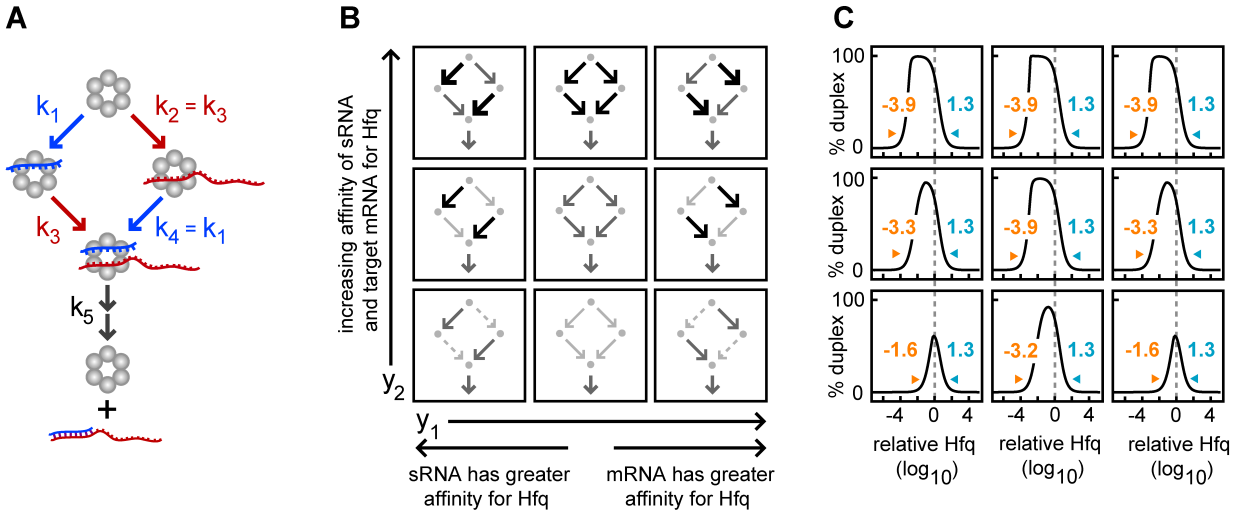

Fig. 3.2 | Independent binding of sRNAs and target mRNAs to Hfq. (A) Reaction scheme with independent binding of sRNAs and target mRNAs to free Hfq hexamers and Hfq complexes ($k_1 = k_4$ and $k_2 = k_3$). In this scheme there is no sRNA or target mRNA dissociation. (B) Topological representation of the reaction schemes with low, medium and high values for the y_1 and y_2 parameters. y_1 determines the relative affinity of target mRNAs and sRNAs for Hfq [$y_1 \equiv ((k_2 \cdot k_3) / (k_1 \cdot k_4))^{1/2}$] and it has values of 10^{-4}, 10^{0} and 10^{4} (unitless) in the simulations. y_2 determines the overall magnitude of the association rate constants for the sRNAs and target mRNAs [$y_2 \equiv (k_1 \cdot k_2 \cdot k_3 \cdot k_4)^{1/4}$] and it has values of $10^{0.5}$, $10^{2.5}$ and $10^{4.5}$ concentration^{-1}·time^{-1} in the simulations. The relative magnitude of the kinetic parameters is represented graphically by the weight of the arrows. (C) Percentage duplex formation at different concentrations of Hfq. Each panel corresponds to the reaction scheme shown in the previous panel at the same position. The grey dash line indicates a 1:1 ratio of [total target mRNA] to [Hfq], where the [total target mRNA] ≡ [T] + [HT] + [HST] + [D] and [T], [HT], [HST] and [D] are the concentrations of free target mRNA, target mRNA-Hfq complex, cognate ternary complex and duplex respectively. Yellow values indicate the "lower bound" while blue values indicate the "upper bound" as defined in the text.

3.4.2 Part 1: Duplex Formation for a Single sRNA-Target mRNA Pair

3.4.2.1 Independent Binding of sRNAs and Target mRNAs to Hfq

To understand the basic behavior of Hfq mediated duplex formation, we began with a highly simplified model with few kinetic degrees of freedom and gradually introduced additional parameters. We first examined duplex formation in the absence of RNA dissociation ($k_{-1} = k_{-2} = k_{-3} = k_{-4} = 0$) and with independent binding (**Fig. 3.2*A***). Independent binding means the probability that a sRNA or target mRNA binds to Hfq does not depend on whether the Hfq is already bound (*i.e.* $k_1 = k_4$ and $k_2 = k_3$). With independent binding, a system with a fixed value for duplex annealing and release (k_5) has only two free parameters. We reparameterized them to yield two new parameters, y_1 and y_2 (**Fig. 3.2*B, C***). y_1 determines the relative affinity of sRNAs and target mRNAs for Hfq [$y_1 \equiv ((k_2 \cdot k_3)/(k_1 \cdot k_4))^{1/2}$; unitless]. y_2 specifies the overall magnitude of the sRNA and target mRNA binding [$y_2 \equiv (k_1 \cdot k_2 \cdot k_3 \cdot k_4)^{1/4}$; units of concentration^{-1}·time^{-1}].

Part 3: Crosstalk among Hfq-dependent small RNAs

We selected low, intermediate and high values for y_1 and for y_2 resulting in nine representative reaction schemes [y_1 = 10^{-4}, 10^{0} and 10^{4}; y_2 = $10^{0.5}$, $10^{2.5}$ and $10^{4.5}$ concentration^{-1}·time^{-1}]. For each reaction scheme, the percentage of the target mRNA converted to duplex was measured at varying Hfq concentrations (**Fig. 3.2*C***). Duplex formation was shown to require less Hfq when sRNAs and target mRNAs bind more rapidly to Hfq (*i.e.* increasing y_2 decreases the lower bound). The lower bound decreases, because with all other factors being equal, increasing RNA binding to Hfq increases Hfq recycling. The maximum concentration of Hfq at which duplex formation occurred (*i.e.* the upper bound) was invariant to y_1 and y_2. That is, sRNA and target mRNA sequestration in singly-bound Hfq complexes is unaffected by the kinetic parameters in a system with independent binding and without RNA dissociation.

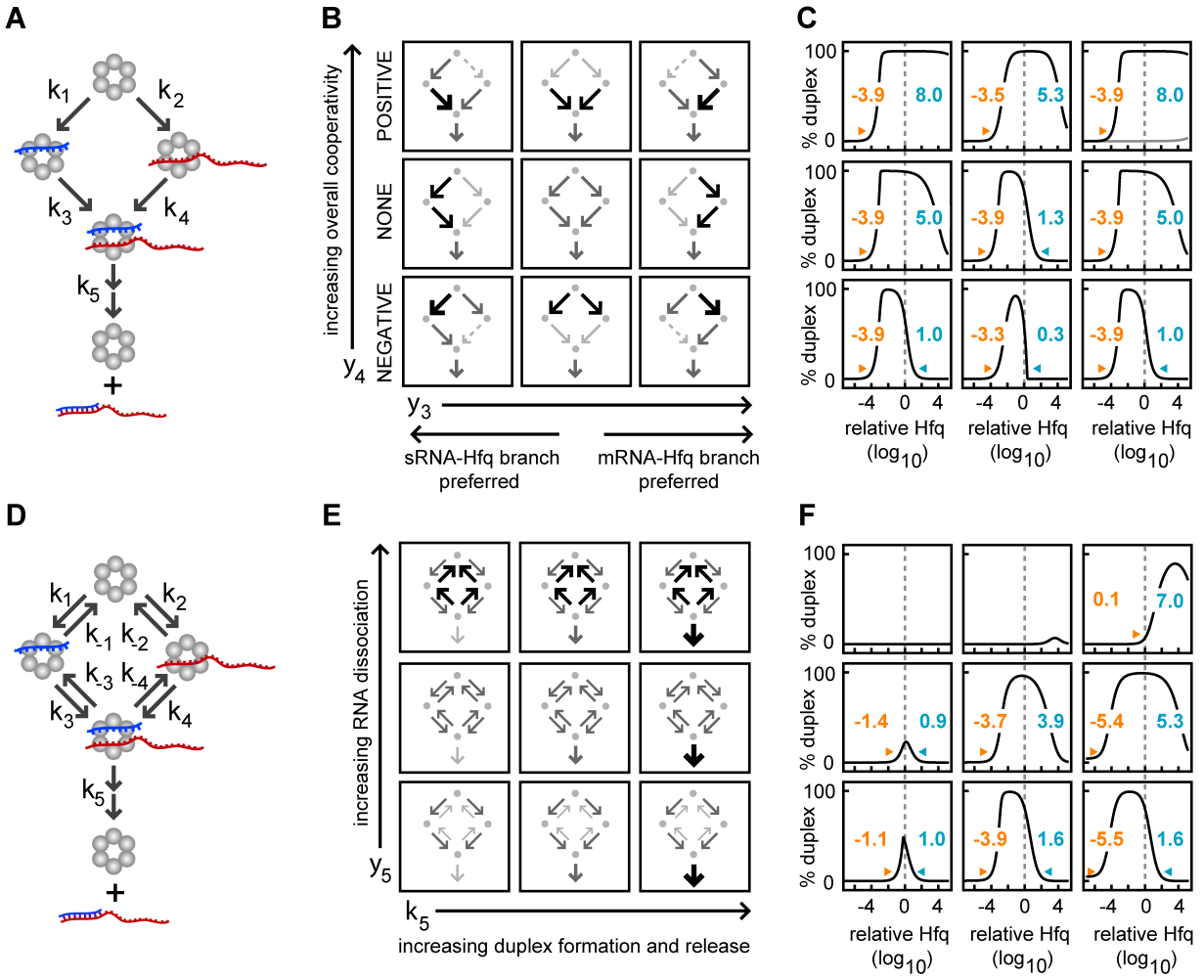

Fig. 3.3 | Cooperative binding and dissociation reactions can increase the efficiency and robustness of duplex formation. (A) Reaction scheme without sRNA and target mRNA dissociation. (B) Topological representation of the cooperative reaction schemes with low, medium and high values for y_3 and y_4. y_3 determines whether the association rate constants favor the sRNA-Hfq branch or the target mRNA-Hfq branch [$y_3 \equiv ((k_2 \cdot k_4)/(k_1 \cdot k_3))^{1/2}$] and it has values of 10^{-4}, 10^0 and 10^4 (unitless) in the simulations. y_4 biases the system from negative to positive cooperative association [$y_4 \equiv ((k_3 \cdot k_4)/(k_1 \cdot k_2))^{1/2}$] and it has values of 10^{-4}, 10^0 and 10^4 (unitless) in the simulations. The relative magnitude of the kinetic parameters is represented graphically by the weight of the arrows. (C) As described for **Fig. 3.2*C***. (D) Reaction scheme for duplex formation with association and dissociation reactions that are independent. (E) Topological representation of the reaction schemes with low, medium and high values for the k_5 and y_5 parameters. k_5 is the overall rate of duplex formation and release with values of 10^0, 10^3 and 10^6 time^{-1} in the simulations. y_5 determines the overall magnitude of the dissociation rate constants for the sRNA and target mRNA [$y_5 \equiv (k_{-1} \cdot k_{-2} \cdot k_{-3} \cdot k_{-4})^{1/4}$] and it has values of 10^0, 10^4 and 10^8 time^{-1} in the simulations. (F) As described for **Fig. 3.2*C***.

3.4.2.2 Cooperative Binding and Dissociation Reactions can Increase the Efficiency and Robustness of Duplex Formation

Heterotropic cooperativity, which we simplify to "cooperativity", exists when the binding and unbinding of a sRNA or target mRNA to a given Hfq hexamer is not independent but depends upon whether that Hfq hexamer is already bound to an RNA. Heterotropic cooperativity could be due to an allosteric change in the Hfq hexamer or a direct or indirect interaction between the bound sRNA and target mRNA. "Positive cooperativity" occurs when the affinity of a sRNA or target mRNA is greater for the singly-bound Hfq complex than for the free Hfq and "negative cooperativity" occurs when the reverse is true. When positive and negative cooperativity arise because the *binding* of a sRNA and target mRNA to the Hfq is altered by the presence of a bound RNA then we use the more specific terms "positive cooperative association" and "negative cooperative association" respectively. Whereas when positive and negative cooperativity occur because the *unbinding* of a sRNA and a target mRNA from Hfq is altered by the presence of a bound RNA then we use the terms "positive cooperative dissociation" and "negative cooperative dissociation" respectively.

We relaxed our assumption of independent binding and created two parameters, y_3 and y_4, to comprehensively explore the effect of positive and negative cooperative association while keeping fixed the relative affinity of sRNAs and target mRNAs for Hfq (y_1) and the total magnitude of the association rate constants (y_2) (**Fig. 3.3*A-C***). y_3 tunes the relative cooperative association of the sRNA compared to that of the target mRNA which alters the bias for the sRNA-Hfq branch and the target mRNA-Hfq branch [$y_3 \equiv (k_2 \cdot k_4)/(k_1 \cdot k_3))^{1/2}$; unitless] (**Fig. 3.3*B***). When $y_3 < 1$, cooperative association is greater for the target mRNA than the sRNA (*i.e.* $k_3/k_2 > k_4/k_1$) resulting in a bias for the sRNA-Hfq branch. Alternatively when $y_3 > 1$, cooperative association is greater for the sRNA than for the target mRNA (*i.e.* $k_4/k_1 > k_3/k_2$) resulting in a bias for the target mRNA-Hfq branch. When $y_3 = 1$, sRNA and target mRNA cooperative association are

equal and therefore duplex formation occurs equally via both branches. y_4 determines whether the RNA is more or less likely to bind to Hfq after its partner has bound ($y_4 > 1$ and $y_4 < 1$ respectively) [$y_4 \equiv ((k_3 \cdot k_4)/(k_1 \cdot k_2))^{1/2}$; unitless]. When an RNA is more likely to bind to Hfq after its partner has bound we term this positive cooperative association ($k_1 < k_4$ and/or $k_2 < k_3$). When an RNA is less likely to bind after its partner has bound we term this negative cooperative association ($k_1 > k_4$ and/or $k_2 > k_3$).

We selected low, intermediate and high values for y_3 and y_4, resulting in nine representative reaction schemes [$y_3 = 10^{-4}$, 10^0 and 10^4; $y_4 = 10^{-4}$, 10^0 and 10^4] (**Fig. 3.3*B*, *C***). Our analysis shows that a bias for one RNA binding order (*i.e.* sRNA-Hfq branch or target mRNA-Hfq branch) diminishes sRNA and target mRNA sequestration in singly-bound Hfq complexes which increases the upper bound (compare left or right columns with center column in **Fig. 3.3*C***). Positive cooperative association ($y_4 > 1$) also alleviates the sequestration of sRNAs and target mRNAs in singly-bound complexes at high Hfq concentrations, while negative cooperative association ($y_4 < 1$) exacerbates it (compare upper bound in the top and bottom rows in **Fig. 3.3*C***). In summary, rate constants that result in a compulsory order of RNA binding (*i.e.* a strong bias for the sRNA-Hfq branch or the target mRNA-Hfq branch) and/or positive cooperative association increase the maximum Hfq concentration at which duplex formation can occur thereby increasing Hfq robustness.

3.4.2.3 Frequent RNA Dissociation Increases the Efficiency and Robustness of Duplex Formation

We next incorporated RNA dissociation, governed by the rate constants k_{-1}, k_{-2}, k_{-3}, k_{-4}, into a model with independent RNA binding and balanced affinity of sRNAs and target mRNAs for Hfq (**Fig. 3.3*D***). In the context of the reverse reactions, these criteria imply that $k_1 = k_2 = k_3 = k_4$ and $k_{-1} = k_{-2} = k_{-3} = k_{-4}$. The parameter y_5 determines the overall probability of non-duplex RNA dissociating from Hfq [$y_5 \equiv (k_{-1} \cdot k_{-2} \cdot k_{-3} \cdot k_{-4})^{1/4}$, units of $time^{-1}$]. Because the dissociation of non-duplex RNA from Hfq constitutes backtracking along the paths to duplex formation, we vary y_5 in conjunction with k_5 (the rate constant for duplex annealing and release) which opposes its action. To systematically explore the impact of these competing effects, we selected low, intermediate and high values for y_5 and for k_5, resulting in nine reaction schemes [$y_5 = 10^0$, 10^4 and 10^8 $time^{-1}$; $k_5 = 10^0$, 10^3 and 10^6 $time^{-1}$] (**Fig. 3.3*E*, *F***).

Increasing duplex annealing and release (k_5) increases Hfq recycling and reduces RNA sequestration in singly-bound complexes. This decreases the minimum Hfq concentration and increases the maximum Hfq concentration for duplex formation leading to increased Hfq robustness (**Fig. 3.3*E, F***). Moderately increasing RNA dissociation such that y_5 is greater than the RNA degradation rate can diminish the sequestration of sRNA and target mRNA molecules in singly-bound Hfq complexes resulting in an increased upper bound (**Fig. 3.3*F***). The increased dissociation of sRNAs and target mRNAs from Hfq complexes enables them to bind multiple Hfq hexamers before they degrade which increases the likelihood that they will encounter their partner on the same Hfq complex (*i.e.* forming a cognate ternary complex). However, the increased RNA dissociation also diminishes the overall RNA affinity for Hfq and therefore higher Hfq concentrations are needed for duplex formation (increased lower bound in **Fig. 3.3*F***). If RNA dissociation is too great then duplex formation is prevented (upper left panel, **Fig. 3.3*F***).

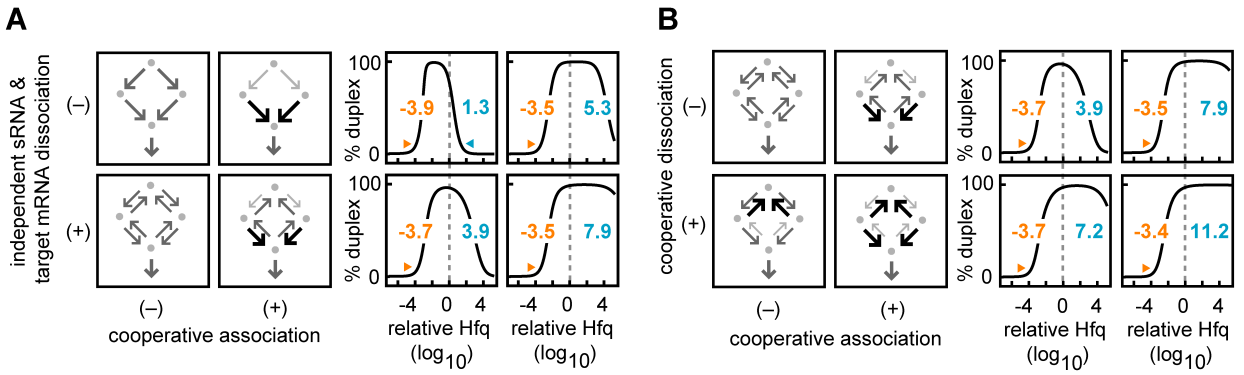

Fig. 3.4 | Cooperative binding and dissociation synergistically promote duplex formation. The combination of cooperative association plus independent dissociation reactions (A) or cooperative association plus cooperative dissociation (B). (A) Duplex formation was simulated with or without independent, RNA dissociation ($k_{-1} = k_{-2} = k_{-3} = k_{-4} = y_5 = 10^4$ or 0 time^{-1}; lower and upper panels respectively) and with or without cooperative association ($y_4 = 10^4$ or 10^0 unitless, right and left panels respectively). The four plots on the right of this panel are as described in **Fig. 3.2*C***. (B) Duplex formation was simulated with or without cooperative dissociation [$y_6 = ((k_{-1} \cdot k_{-2})/(k_{-3} \cdot k_{-4}))^{1/2}$ and it has values of 10^4 or 10^0; lower and upper panels respectively] and with or without cooperative association [$y_4 = 10^4$ or 10^0 unitless; right and left panels respectively]. The four plots on the right of this panel are as described in **Fig. 3.2*C***.

3.4.2.4 Cooperativity Combined with Frequent RNA Dissociation can Result in Greater Robustness than Either Mechanism Alone (Synergy)

We next demonstrated that cooperativity and RNA dissociation, which promote duplex formation via two different mechanisms, can have synergistic effects. We first showed that positive cooperative association plus *independent* unbinding of sRNAs and target

mRNAs from Hfq complexes reduces singly-bound Hfq complexes more than either mechanism alone (upper bound increases; **Fig. 3.4*A***). We then examined the effects of positive cooperativity in the dissociation reactions (*i.e.* positive cooperative dissociation as defined earlier) which acts to relatively increase dissociation from singly-bound Hfq complexes compared with dissociation from Hfq ternary complexes ($k_{-4} < k_{-1}$ and/or $k_{-3} < k_{-2}$) also promotes Hfq robustness (upper bound increases, **Fig. 3.4*B***). Cooperativity in the dissociation reactions is quantified by the unitless metric $y_6 \equiv ((k_{-1} \cdot k_{-2})/(k_{-3} \cdot k_{-4}))^{1/2}$; $y_6 > 1$ indicates positive cooperativity while $y_6 < 1$ indicates negative cooperativity. The combination of positive cooperative association plus positive cooperative dissociation further reduces sequestration in singly-bound Hfq complexes (**Fig. 3.4*B***).

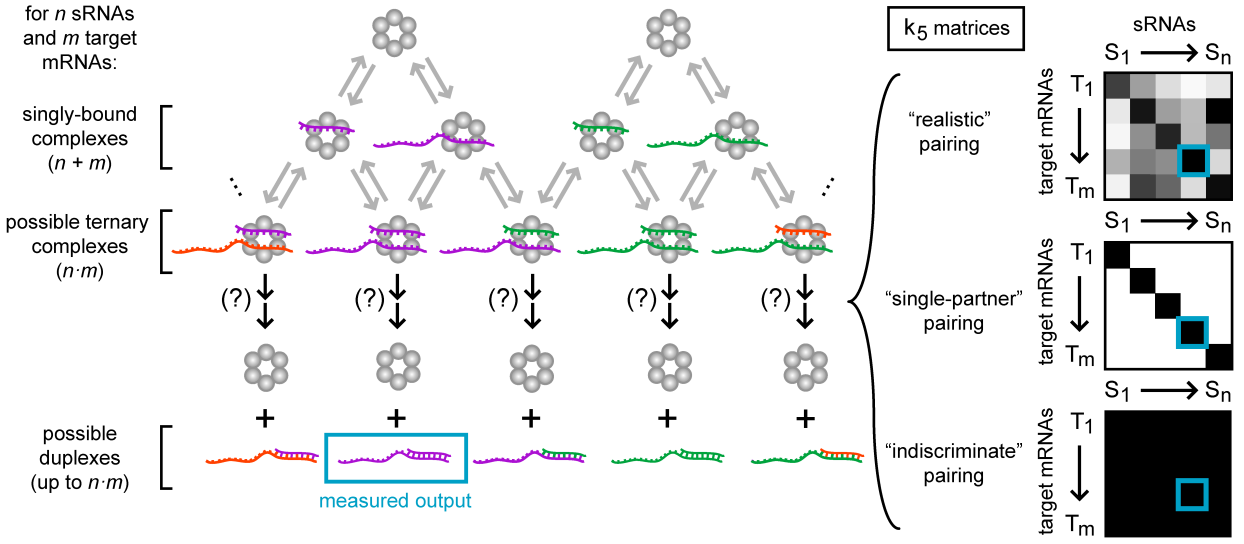

Fig. 3.5 | General reaction scheme for duplex formation in a network with multiple sRNAs and target mRNAs. In a network with multiple sRNAs and target mRNAs many types of Hfq ternary complexes are possible. If these ternary complexes cannot form duplexes they are termed non-cognate ternary complexes. In other words, $k_5 \sim 0$ for a non-cognate ternary complex and $k_5 >> 0$ for a cognate ternary complex. The k_5 values that specify whether a ternary complex is cognate or non-cognate can be represented in a $n \times m$ matrix for n sRNAs and m target mRNAs with the darkness of the shading representing the magnitude of k_5. Three examples of this k_5 matrix are shown: 1. a "mock" native sRNA network (upper right panel); 2. a simplified system where each sRNA and target mRNA has only one specific partner (middle right panel); and 3. a simplified system where each sRNA and target mRNA can indiscriminately form duplexes with any other target mRNA or sRNA respectively (lower, right panel). For subsequent plots, we stress that "relative Hfq" = 1 (*i.e.* 10^0) indicates the concentration of Hfq in all forms is equal to the total concentration of target mRNA for the cognate pair being measured (as opposed to the total concentration of all the target mRNAs in the network).

3.4.3 Part 2: Duplex formation in the sRNA network

3.4.3.1 General Reaction Scheme for Duplex Formation in a Network with Multiple sRNAs and Target mRNAs

In a network, we need to consider that sRNAs and target mRNAs can bind in multiple combinations to Hfq resulting in many types of ternary complexes [85] (**Fig. 3.5**). Therefore we categorized ternary complexes as cognate or non-cognate according to whether the bound sRNAs and target mRNAs can form a duplex or not. It is possible for any given sRNA or target mRNA to have several cognate RNA partners. By definition, if a ternary complex is non-cognate then $k_5 \sim 0$ and if a ternary complex is cognate then $k_5 >> 0$.

In any real biological system k_5 will have a range of values for every possible ternary complex that can form and therefore the distinction between cognate and non-cognate may not be clear. That is, in a real network there will be many sRNAs and target mRNAs that can form duplexes with more that one type of target RNA or sRNA respectively (as stated above). Furthermore, each sRNA or target mRNA may form these different types of duplexes at different rates. The k_5 value, which determines the rate of duplex annealing and release, can be represented in a "k_5 matrix" for each possible sRNA-target mRNA pairing on Hfq (a "mock" example of a k_5 matrix for a "real" system is shown in the upper right panel, **Fig. 3.5**; darker shading indicates higher magnitudes of k_5 with white representing zero).

We reduced the complexity of the system by adding the constraint that every sRNA and target mRNA behaves identically and the number of sRNAs (n) is equal to the number of target mRNAs (m) (*i.e.* $n = m$) unless otherwise stated. We examined two extreme cases of this simplified system. In the first case, each sRNA and target mRNA has only one specific partner. That is, $sRNA_i$ forms a duplex with $mRNA_j$ if and only if $i = j$ (by definition this means that $k_{5(i,j)} >> 0$ if and only if $i = j$ in the k_5 matrix) (middle right panel, **Fig. 3.5**). In the second case, each sRNA and target mRNA can indiscriminately form duplexes with any other target mRNA or sRNA respectively. That is, $sRNA_i$ forms duplex with $mRNA_j$ for all possible pairs i and j (by definition this means $k_{5(i,j)} >> 0$, for all values of i and j in the k_5 matrix) (lower right panel, **Fig. 3.5**). We examined the first case where each sRNA and target mRNA has exactly one cognate partner in **Figs. 3.6-3.10** and **Fig. 3.12** and the second case where sRNAs and target mRNAs can form duplexes in all possible combinations in **Fig. 3.10**.

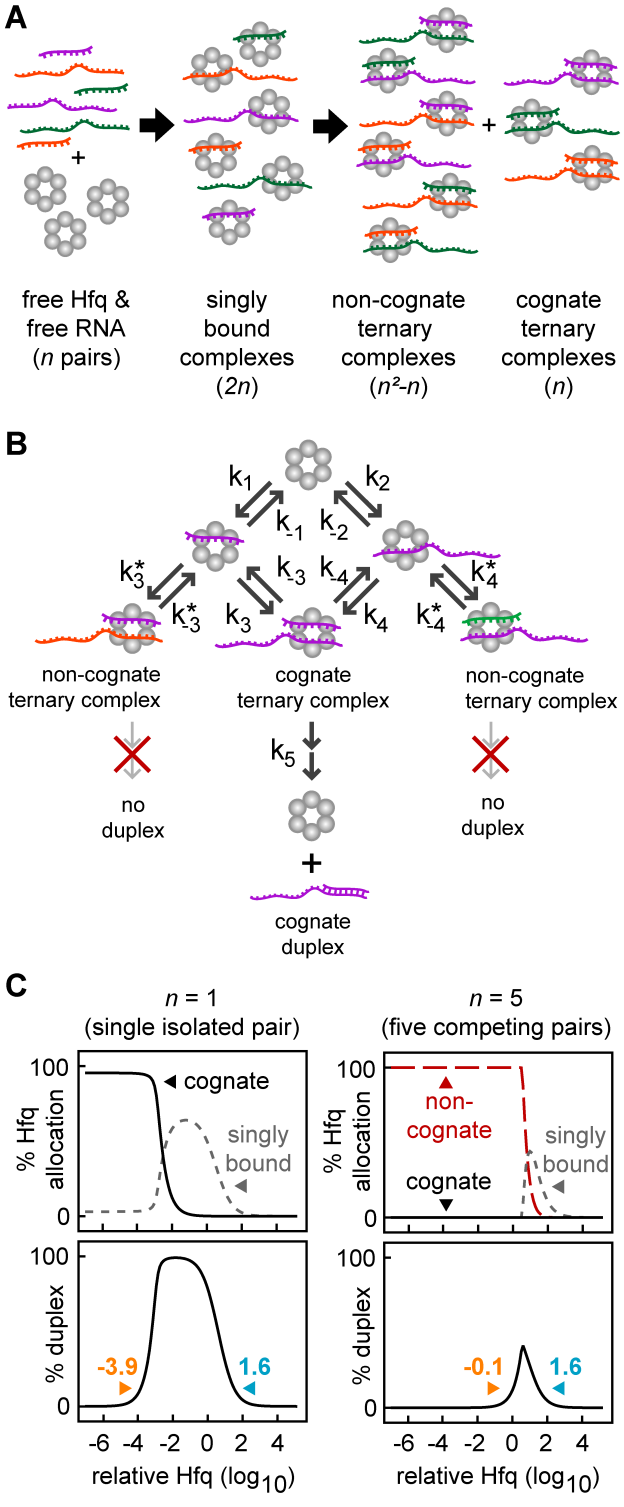
A
free Hfq &
free RNA
(n pairs)
singly
bound
complexes
(2n)
non-cognate
ternary
complexes
(n²-n)
cognate
ternary
complexes
(n)
B
k1
k2
k-1
k-2
k*3
k*-3
k-3
k-4
k3
k4
k*4
k*-4
non-cognate
ternary complex
cognate
ternary complex
non-cognate
ternary complex
k5
no
duplex
no
duplex
cognate
duplex
C
n = 1
(single isolated pair)
n = 5
(five competing pairs)
100
% Hfq
allocation
0
cognate
singly
bound
non-
cognate
cognate
singly
bound
100
% duplex
0
-3.9
1.6
-0.1
1.6
-6 -4 -2 0 2 4
relative Hfq (log10)
-6 -4 -2 0 2 4
relative Hfq (log10)

Fig. 3.6 | Non-cognate ternary complexes decrease the efficiency and robustness of duplex formation. (A) Schematic showing the possible Hfq complexes formed by three sRNAs and three target mRNAs with single partner pairing. The number of possible types of Hfq complexes that can form with *n* sRNA-target mRNA pairs if both the sRNA and the target mRNA in each pair can bind to Hfq hexamers and there is only one cognate partner for each sRNA and target mRNA. (B) Schematic showing the kinetics of duplex formation for a single cognate sRNA-target mRNA pair in the presence of multiple sRNAs and target mRNAs competing for Hfq. Non-cognate ternary complexes form but by definition do not result in duplexes. The rate constants k^*_3 and k^*_4 specify the association of the target mRNA and sRNA respectively to a sRNA-Hfq and target mRNA-Hfq complex resulting in the formation of a non-cognate ternary complex. The rate constants k^*_{-3} and k^*_{-4} specify the dissociation of the target mRNA and sRNA respectively from the non-cognate ternary complex. (C) The percentage allocation of Hfq and percentage duplex formed for a sRNA-target mRNA pair in isolation (left panels) and in a network with five sRNA-target mRNA pairs competing equally for Hfq (right panels). In these simulations, all sRNA-target mRNA pairs have independent RNA binding and unbinding. Yellow and blue values indicate the lower and upper bounds respectively.

3.4.3.2 Non-cognate Ternary Complexes Decrease the Efficiency and Robustness of Duplex Formation

The effect of non-cognate ternary complexes on duplex formation was simulated in a system with *n* sRNA-target mRNA pairs, where each sRNA and target mRNA has only one partner (*i.e.* middle right panel of **Fig. 3.5**). In this system there are potentially n^2-n different types of non-cognate ternary complexes that can form (**Fig. 3.6*A***). Therefore with five or more sRNA-target mRNA pairs there are more possible types of non-cognate ternary complexes than singly–bound Hfq complexes ($2n$) and cognate ternary complexes (n) combined. We demonstrated that when sRNAs and target mRNAs bind independently and with equal affinity for Hfq (**Fig. 3.6*B***), duplex formation can be so impaired in this system by the formation of non-cognate ternary complexes that there is no Hfq concentration at which duplexes form efficiently (lower right, **Fig. 3.6*C***). Therefore in a network where non-cognate ternary complexes can form, there need to be mechanisms that allow duplexes to form efficiently.

3.4.3.3 Decreasing Non-cognate Ternary Complexes by Cooperativity Increases the Efficiency and Robustness of Duplex Formation

We demonstrate that cooperativity is a mechanism that can increase the proportion of cognate ternary complexes and decrease the proportion of non-cognate ternary complexes. Cooperativity can achieve this by "cognate selection" and "non-cognate exclusion" (**Fig. 3.7*A***)**.** Cognate selection requires a stabilizing interaction between cognate sRNA-target mRNA pairs in the ternary complex while non-cognate exclusion requires a destabilizing interaction between non-cognate sRNAs and target mRNAs in the ternary complex.

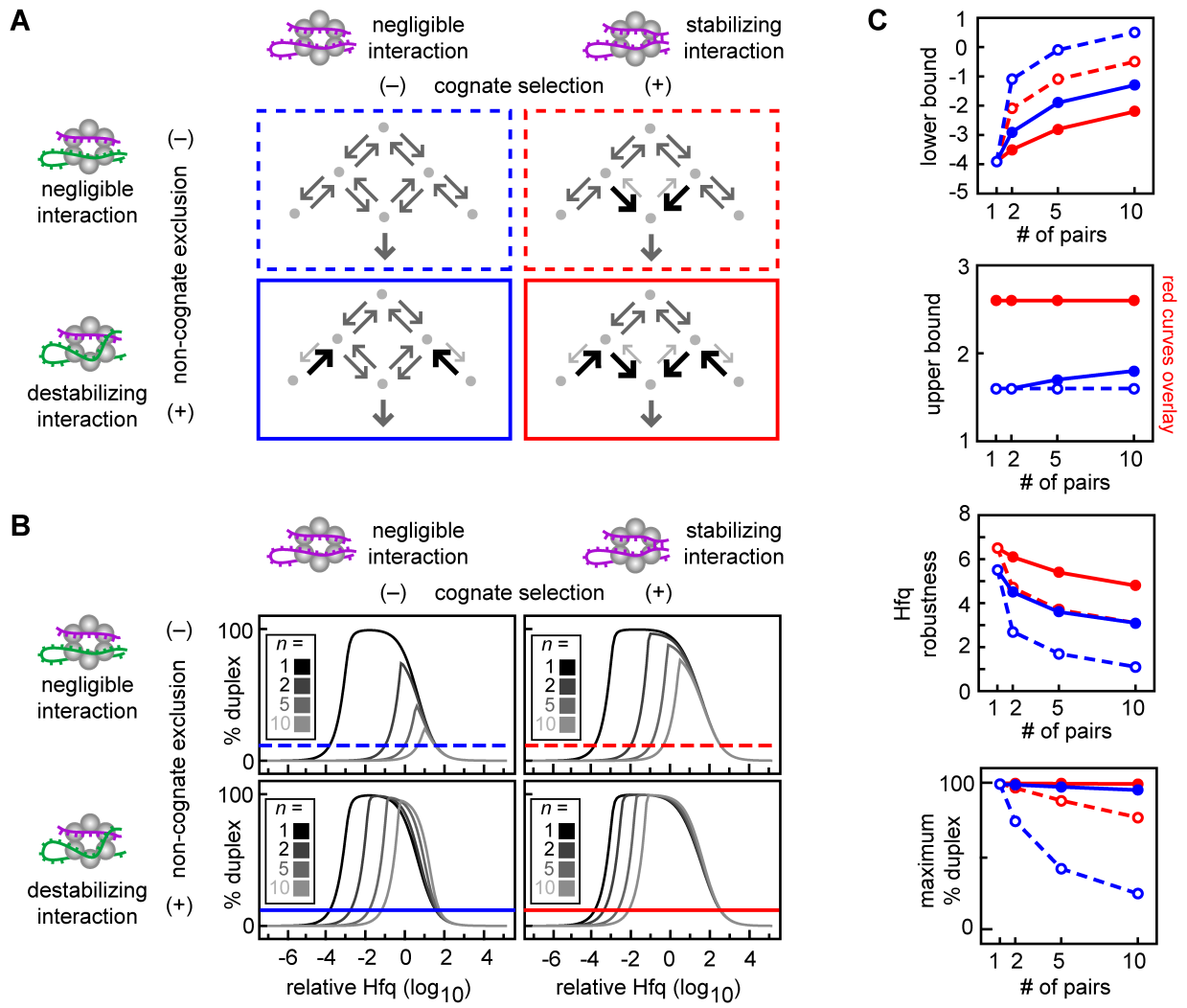

Fig. 3.7 | Decreasing non-cognate ternary complexes by cooperativity increases the efficiency and robustness of duplex formation. (A) Simplified topology representations showing the reaction schemes with or without cognate selection and with or without non-cognate exclusion. Each RNA can form duplexes with only one specific partner. Cognate selection occurs when the rate constants increase the formation and/or the stability of the cognate ternary complex such that the unitless ratio $y_7 < 1$ (right panels). Non-cognate exclusion when rate constants decrease the formation and/or the stability of the non-cognate ternary complex such that the unitless ratio $y_8 > 1$ (lower panels). The color and style of the border surrounding each topology indicates the properties of the reaction scheme. Dash blue is independent binding ($y_7 = 1$; $y_8 = 1$); solid blue is non-cognate exclusion ($y_7 = 1$; $y_8 = 10^1$); dash red is cognate selection ($y_7 = 10^{-1}$; $y_8 = 1$); solid red is cognate selection plus non-cognate exclusion ($y_7 = 10^{-1}$; $y_8 = 10^1$). (B) Percentage duplex formed in sRNA networks with *n* identical sRNA-target mRNA pairs having the reaction scheme shown in the corresponding panel in (A). The horizontal line indicates 10% duplex formation which defines the upper and lower bounds as previously described. (C) The lower bound, upper bound, Hfq robustness and maximum percentage duplex for each kinetic scenario described in (A). The color and style of each curve indicates the corresponding topology shown in (A).

Cognate selection can occur when Hfq-bound sRNAs or target mRNAs assist the binding of their cognate partner (*i.e.* increasing the RNA binding rate constants k_3 and

k_4) or when sRNA-target mRNA pairs in cognate ternary complexes stabilize their partners binding (*i.e.* reducing the RNA dissociation rate constants k_{-3} and k_{-4}) (**Figs. 3.6*B*, 3.7*A***). We created a parameter, y_7, which measures the affinity of sRNAs and target mRNAs for cognate ternary complexes relative to singly-bound complexes [$y_7 \equiv ((k_1 \cdot k_2 \cdot k_{-3} \cdot k_{-4})/(k_{-1} \cdot k_{-2} \cdot k_3 \cdot k_4))^{1/4}$ which is unitless; $y_7 < 1$ for cognate selection].

Non-cognate exclusion can occur when Hfq-bound sRNAs or target mRNAs occlude the binding sites for non-cognate partners (*i.e.* decreasing the binding rate constants k^*_3 and k^*_4) or when sRNAs and target mRNAs in non-cognate ternary complexes destabilize one another's binding (*i.e.* increasing the dissociation rate constants k^*_{-3} and k^*_{-4}) (**Figs. 3.6*B*, 3.7*A***). We created a parameter, y_8, to specify the affinity of sRNAs and target mRNAs for the non-cognate ternary complexes relative to singly-bound complexes [$y_8 \equiv ((k_1 \cdot k_2 \cdot k^*_{-3} \cdot k^*_{-4})/(k_{-1} \cdot k_{-2} \cdot k^*_3 \cdot k^*_4))^{1/4}$ which is unitless; $y_8 > 1$ for non-cognate exclusion].

In the absence of any cooperativity (*i.e.* with independent RNA binding and unbinding), increasing the number of sRNA-target mRNA pairs in the network reduces the maximum percentage duplex achievable and increases the minimum Hfq concentration required for 10% duplex formation (upper left panel in **Fig. 3.7*B*, *C***). The same network with cognate selection or non-cognate exclusion has a greater maximum percentage duplex formation and a lower minimum Hfq concentration for 10% duplex formation (**Fig. 3.7*B*, *C***). Cognate selection combined with non-cognate exclusion reduces Hfq sequestration and improves duplex formation more than either mechanism separately (**Fig. 3.7*B*, *C***). These results show that reducing the sequestration of sRNAs, target mRNAs and Hfq in non-cognate ternary complexes by these forms of cooperativity enables more pairs to signal in parallel over a wider range of Hfq concentrations.

3.4.3.4 Increasing RNA Dissociation Increases the Efficiency and Robustness of Duplex Formation in Networks

In this section, rapid RNA dissociation is shown to be an important mechanism for decreasing sequestration in non-cognate ternary complexes (**Fig. 3.8**). The dissociation rate constants determine how stably sRNAs and target mRNAs are bound in non-cognate ternary complexes (and also in cognate ternary complexes). High dissociation kinetics prevent sRNAs and target mRNAs being sequestered in singly-bound Hfq complexes as well as ternary complexes, which allows them to bind Hfq multiple times before they degrade. As a result, there is an increased probability that the sRNAs and

target mRNAs will form cognate ternary complexes and this will lead to increased cognate duplex production [if the total magnitude of the association rate constants (y_2) and duplex annealing and release (k_5) are sufficiently high].

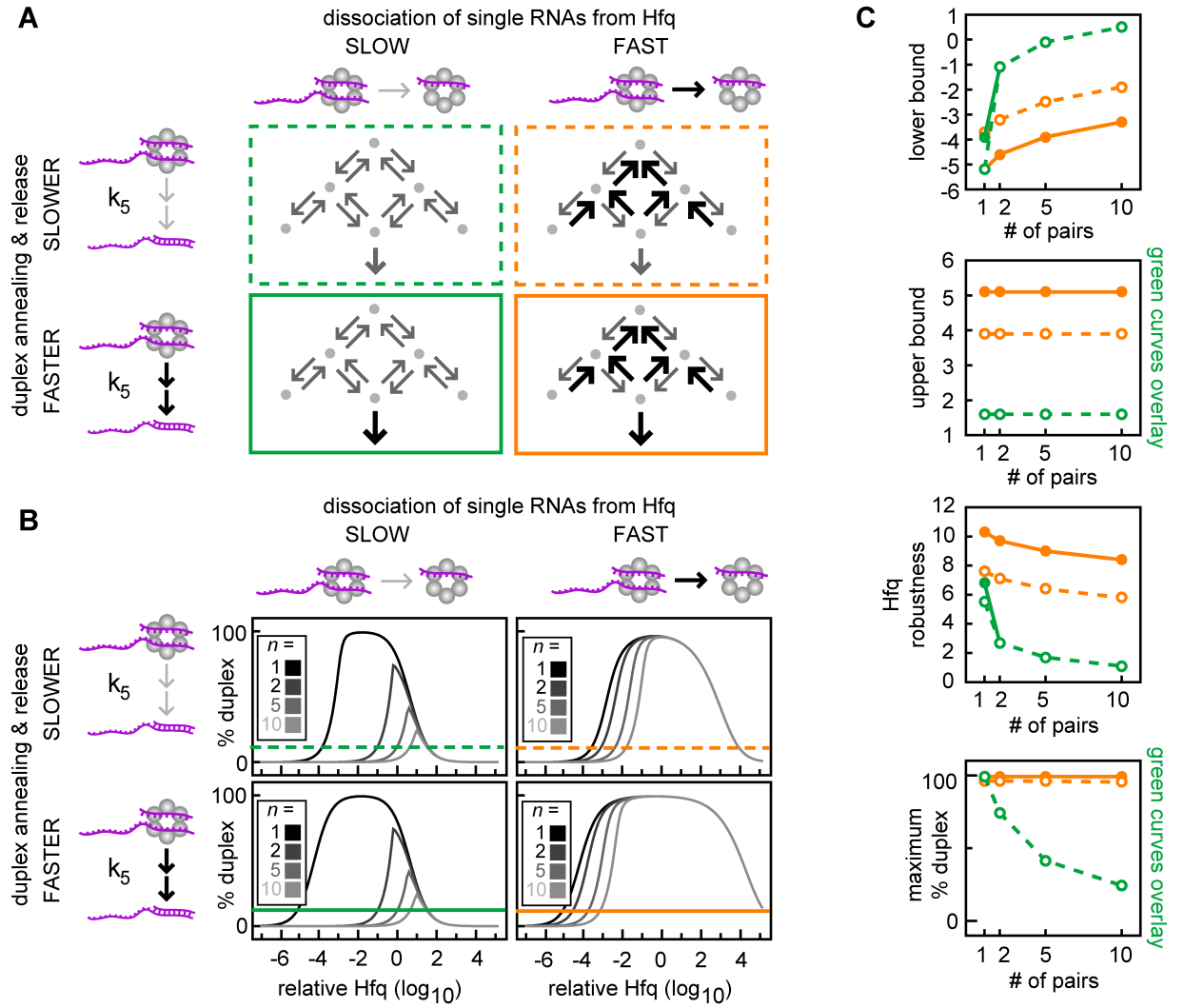

Fig. 3.8 | Increasing RNA dissociation increases the efficiency and robustness of duplex formation in networks. (A) Simplified topology representations showing the reaction schemes with and without increased RNA dissociation and with and without increased duplex annealing. Each RNA can form duplexes with only one specific partner. The parameter for duplex annealing and release (k_5) has values of 10^3 and 10^6 time^{-1} for these simulations (upper and lower panels respectively). RNA dissociation from singly-bound Hfq complexes (quantified by y_9) is non-cooperative and is equal for the sRNA and mRNA within these topologies (*i.e.* $k_{-1} = k_{-2} = k_{-3} = k_{-4} = k^*_{-3} = k^*_{-4} = y_9 = 10^0$ and 10^4 time^{-1}; left and right panels respectively). The color and style of the border surrounding each topology indicates the properties of the reaction scheme. Dash green serves as a basis for comparison and is identical to dash blue from **Fig. 3.7** ($k_5 = 10^3$; $y_9 = 10^0$); solid green indicates increased duplex annealing and release ($k_5 = 10^6$; $y_9 = 10^0$); dash yellow indicates increased RNA dissociation ($k_5 = 10^3$; $y_9 = 10^4$); and solid yellow indicates both increased RNA dissociation and increased duplex annealing and release ($k_5 = 10^6$; $y_9 = 10^4$). (B) Percentage duplex formed in sRNA networks with *n* identical sRNA-target mRNA pairs with the reaction scheme shown in the corresponding panel in (A). The horizontal

line indicates 10% duplex formation which defines the upper and lower bounds as previously described. (C) The lower bound, upper bound, Hfq robustness and maximum percentage duplex for each of the four kinetic scenarios described in (A). The color and style of each curve indicates the corresponding topology shown in (A).

We created a parameter, y_9, which measures the total magnitude of RNA dissociation from Hfq complexes [$y_9 \equiv (k_{-1} \cdot k_{-2} \cdot k_{-3} \cdot k^*_{-3} \cdot k_{-4} \cdot k^*_{-4})^{1/6}$, units of time^{-1}] (Note: k^*_{-3} and k^*_{-4} are defined above). We found that increasing RNA dissociation (y_9) did indeed decrease the minimum amount of Hfq required for 10% duplex formation (decreased lower bound in **Fig. 3.8*B***, ***C***). As before, the impact of duplex annealing and release (k_5) was examined in combination with RNA dissociation (y_9) as these reactions compete for cognate ternary complexes. We found that increasing duplex annealing and release together with increased RNA dissociation further decreased the amount of Hfq required for duplex formation. However, increasing duplex annealing and release (k_5) in the absence of sufficient RNA dissociation had limited effect (**Fig. 3.8*B***, ***C***) because the prevalence of the cognate ternary complex is small.

3.4.3.5 Cognate sRNA-Target mRNA Pairs with Distinct Dissociation Kinetics Perform Differently in Isolation than in Networks

We next examined networks with a mix of "stable" cognate pairs that stably bind Hfq due to slow RNA dissociation or "unstable" cognate pairs that unstably bind Hfq due to rapid RNA dissociation (**Fig. 3.9*A***). These simulations showed that duplex formation for stable pairs is sensitive to the number of other stable pairs in the network but relatively insensitive to the number of unstable pairs in the network (left panels, **Fig. 3.9*B***, ***C***). In contrast, duplex formation for unstable pairs is relatively insensitive to both the number of stable and unstable pairs in the network (right panels, **Fig. 3.9*B***, ***C***).

Together the results reveal that in terms of duplex formation, a network is more scalable if it is primarily composed of unstable cognate pairs. The sRNAs and target mRNAs in unstable pairs permit scalability because they do not get sequestered in non-cognate ternary complexes due to rapid RNA dissociation. Whereas cognate pairs that form stable Hfq complexes and perform well in isolation, may perform relatively poorly in a network. This result shows that an assessment of the properties of cognate pairs needs to take into account the other sRNAs and target mRNAs that are acting at the same time; this has important implications for future studies as we discuss below.

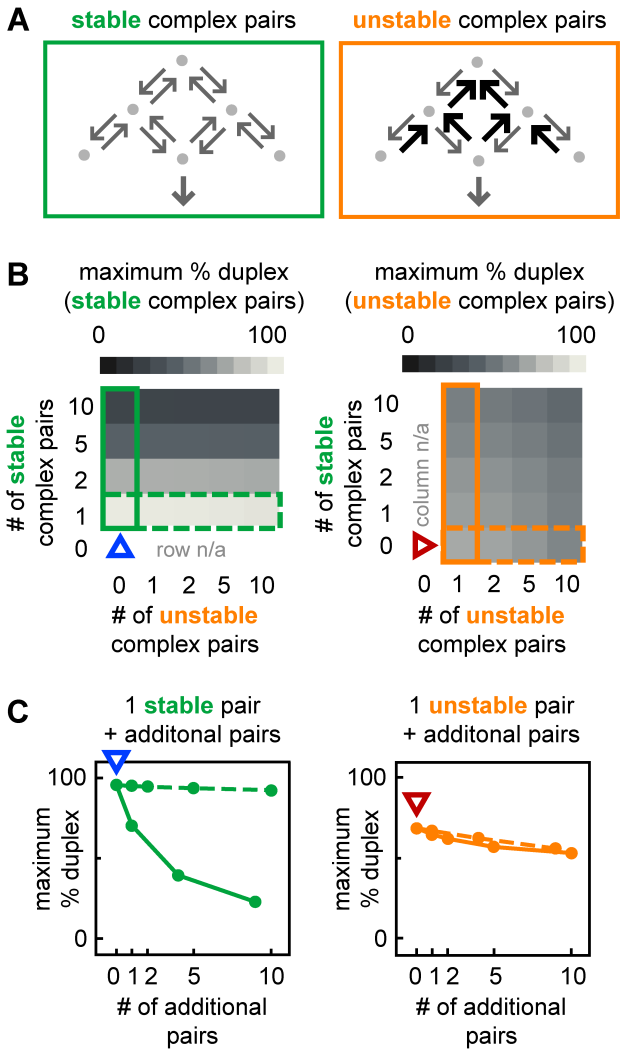

Fig. 3.9 | Cognate sRNA-target mRNA pairs with distinct dissociation kinetics perform differently in isolation and in networks. (A) Topologies for cognate pairs which form stable Hfq complexes due to low RNA dissociation rate constants or unstable Hfq complexes due to high RNA dissociation rate constants. (B) Maximum percentage duplex production in networks with different numbers of cognate sRNA-target mRNA pairs that form unstable (horizontal axis) and stable Hfq (vertical axis) complexes. The left and right panels show the maximum duplex formation in cognate pairs that form stable and unstable complexes respectively. The vertical boxed regions (solid green and orange) and the horizontal boxed regions (dash green and orange) indicate the data shown in (C). Blue and red arrowheads indicate stable and unstable cognate pairs acting in isolation. (C) Duplex formation for a single stable pair (left) and a single unstable pair (right) in a network with an increasing number of unstable pairs (dot line) or stable pairs (solid line).

3.4.3.6 Indiscriminate Duplex Formation can Increase Hfq Robustness at the Cost of Decreased Maximum Duplex Yield

Until this point, we have only simulated networks where sRNAs and target mRNAs have one cognate partner. Here, we directly compare a network where sRNAs and

target mRNAs have one cognate partner (upper panels, **Fig. 3.10**) to a network where all ternary complexes can form duplexes (lower panels, **Fig. 3.10**). In other words, in the latter network all sRNAs and target mRNAs can form a duplex with any other RNA (*i.e.* $k_{5(i,j)} > 0$, for all values of *i* and *j*); this represents ubiquitous cross-talk or non-specific interactions. Of course in a wild-type network not all sRNAs and target mRNAs will indiscriminately form duplexes with each other. However, this scenario will most clearly shed light on how the formation of multiple types of duplexes by each type of sRNA and target mRNA, which does occur to some extent in wild-type sRNA networks, affects the efficiency and robustness of duplex formation.

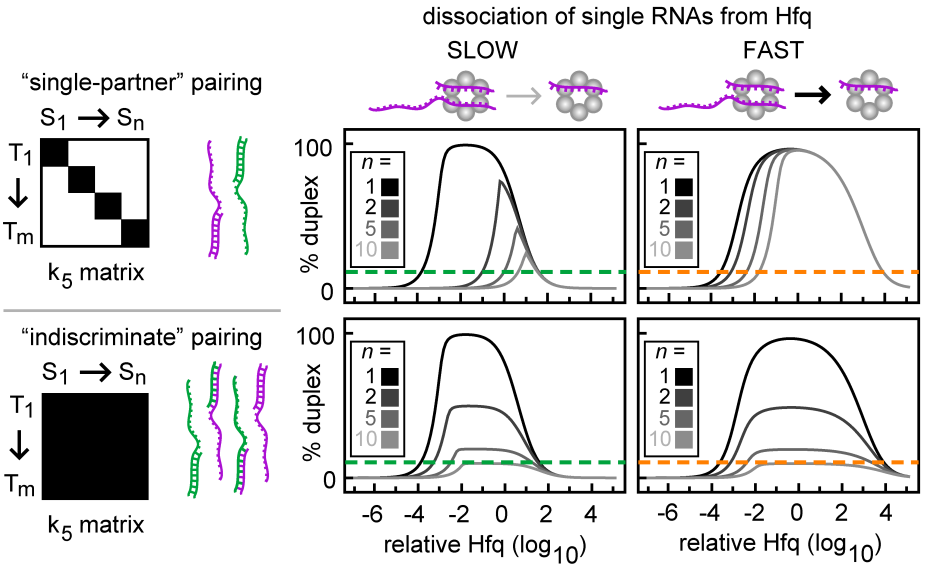

Fig. 3.10 | Indiscriminate duplex formation can increase Hfq robustness at the cost of maximum duplex yield. Percentage duplex formed with specific or indiscriminant duplex formation (upper and lower panels) and slow or fast RNA dissociation (left and right panels). For specific duplex formation, each sRNA and target mRNA has only one cognate partner (*i.e.* $k_{5(i,j)} = 0$ time^{-1} for $i \neq j$ and $k_{5(i,j)} = 10^3$ time^{-1} for $i = j$) (illustrated in the upper left panel). We define indiscriminate duplex formation as when all sRNAs and target mRNAs can form duplexes with any target mRNA and sRNA respectively (*i.e.* $k_{5(i,j)} = 10^3$ time^{-1} for all values of *i* and *j*) (illustrated in the lower left panel). The rate constants for RNA dissociation ($k_{-1} = k_{-2} = k_{-3} = k_{-4} = k^*_{-3} = k^*_{-4} = y_9$) were equal to 10^0 or 10^4 time^{-1} for the slow and fast categories respectively. *n* is the number of types of pairs sRNAs and target mRNAs in the network.

We found that at low rates of RNA dissociation, the indiscriminate formation of duplexes prevents Hfq sequestration because there are no non-cognate ternary complexes and therefore duplexes are released from all ternary complexes (upper and lower left panels, **Fig. 3.10**). This increases Hfq recycling which decreases the lower bound resulting in increased Hfq robustness. However, the increased Hfq robustness comes at the cost of a large reduction in the maximum yield for any given duplex. The amount of duplex decreases because sRNAs and target mRNAs are being incorporated

into a wider variety of duplexes and therefore there is less cognate sRNA and target mRNA for any particular duplex. It should be noted that indiscriminate sRNA-target mRNA pairing has little impact on Hfq robustness when RNA dissociation from Hfq complexes is high. This is because under these conditions Hfq recycling is already high (upper and lower right panels, **Fig. 3.10**) and therefore the sequestration of Hfq and RNAs in non-cognate ternary complexes is low.

3.4.3.7 Imbalances in sRNA and Target mRNA Production can Globally Alter Duplex Formation

In the previous simulations, the production and degradation rates were equal for both members of each cognate sRNA-target mRNA pair. We now investigate how unequal production rates for sRNAs and target mRNAs impacts duplex formation in different scenarios (**Figs. 3.11-3.13**).

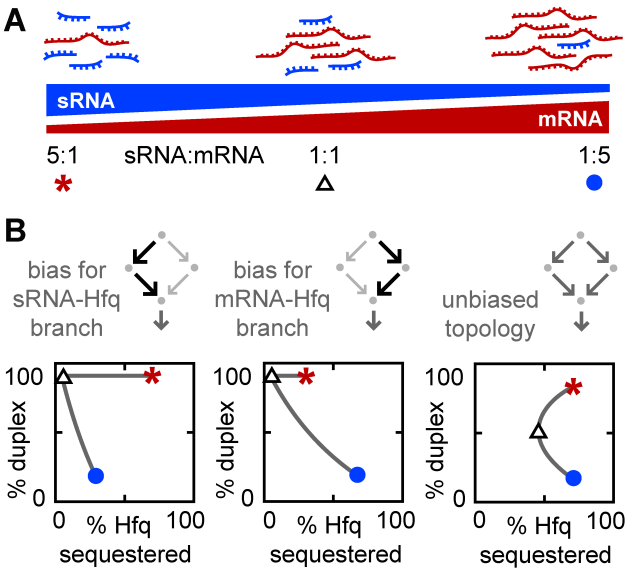

Fig. 3.11 | Sequestration of Hfq by imbalanced RNA expression is minimized by a compulsory RNA binding order. (A) The relative concentration of sRNA to target mRNA within a cognate sRNA-target mRNA pair is varied from 5:1 to 1:5 while the combined concentration of sRNA and target mRNA is kept constant. sRNA/target mRNA ratios of 5:1, 1:1 and 1:5 are indicated by a red asterisk, black triangle and blue circle respectively. (B) Percentage duplex versus percentage Hfq sequestered for cognate sRNA-target mRNA pairs with imbalanced ratios. The shape of the curve depends on whether the kinetics of the sRNA-target mRNA pair is biased towards the sRNA-Hfq branch, the target mRNA-Hfq branch or neither. The bias was altered by changing the values for y_3 while keeping y_1, y_2 and y_4 constant ($y_3 = 10^{-8}$, 10^{8} and 10^{0} in the left, middle and right panels respectively).

We first show that for a single cognate sRNA-target mRNA pair there will be less Hfq sequestration if the member of the pair that is in excess does not efficiently bind free

Hfq (**Fig. 3.11**). If the concentration of the target mRNA exceeds its sRNA partner (blue circles in **Fig. 3.11**) then sequestration will be lower if only the sRNA binds free Hfq (*i.e.* a sRNA-Hfq branch bias). Conversely, if the concentration of sRNA exceeds its target mRNA partner (red asterisks in **Fig. 3.11**), sequestration will be less if only the target mRNA binds free Hfq (*i.e.* a target mRNA-Hfq branch bias).

We next examined a simple network composed of two cognate sRNA-target mRNA pairs ($sRNA_1$-target $mRNA_1$ and $sRNA_2$-target $mRNA_2$) where each RNA can only form duplexes with one partner. The production of one sRNA ($sRNA_1$) was varied and the production of the other RNAs was kept constant and equal (*i.e.* the production of $sRNA_1 \geq$ production of $sRNA_2$, target $mRNA_1$ and target $mRNA_2$) (**Fig. 3.12*A***). All sRNAs and target mRNAs had the same degradation rate, independent binding to Hfq and identical RNA dissociation kinetics. In this scenario, increasing $sRNA_1$ production increased Hfq robustness for the $sRNA_1$-target $mRNA_1$ pair because the additional sRNA drives the concentration of the $sRNA_1$-Hfq complex higher which increases the formation of the $sRNA_1$-Hfq-target $mRNA_1$ ternary complex and $duplex_1$ (**Fig. 3.12*B***). However, the excess $sRNA_1$-Hfq also leads to increased formation of non-cognate ternary complexes which diminishes duplex formation and robustness for the other sRNA-target mRNA pair ($duplex_2$ in **Fig. 3.12*B***). That is, imbalances in sRNA-target mRNA concentrations can increase Hfq robustness for one pair at the cost of reducing duplex formation and Hfq robustness for other pairs.

We then investigated a common scenario where most target mRNAs are transcribed without their partners and a single sRNA is activated in response to a specific stressor [117,118]. In this case, a single cognate sRNA-target mRNA pair with balanced production needs to signal in a pool of unpartnered target mRNAs (**Fig. 3.13**). The large pool of unpartnered target mRNAs (100-fold the concentration of the cognate target mRNA in this simulation) increases the formation of non-cognate ternary complexes and therefore decreases Hfq recycling. The decreased Hfq recycling results in an increased upper bound and decreased duplex formation (compare right and left panels, **Fig. 3.13**).

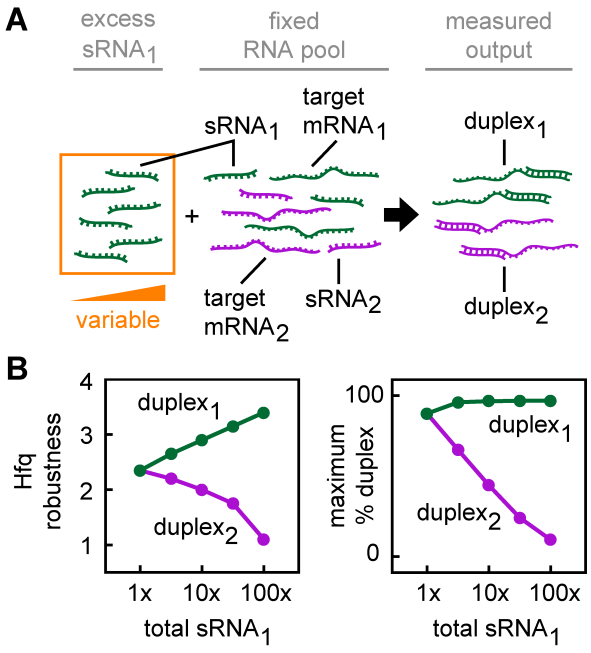

Fig. 3.12 | Imbalances in sRNA and target mRNA production can globally alter duplex formation. (A) A network with two cognate sRNA-target mRNA pairs where each sRNA and target mRNA can only bind one partner. Initially the concentrations of $sRNA_1$, $sRNA_2$, target $mRNA_1$ and target $mRNA_2$ are equal. Additional $sRNA_1$ is then added to the system by increasing the production of $sRNA_1$. (B) Hfq robustness and the maximum percentage duplex for the $sRNA_1$-target $mRNA_1$ pair (green) and the $sRNA_2$-target $mRNA_2$ pair (purple) as a function of the total concentration of $sRNA_1$.The total $sRNA_1$ concentration is measured relative to target $mRNA_1$.

As with the other scenarios described above, rapid RNA dissociation and cooperativity (only non-cognate exclusion is shown) minimize non-cognate ternary complexes resulting in increased Hfq robustness and duplex formation (upper panels, **Fig. 3.13**). Hfq robustness is also improved by increasing the affinity of the sRNA for Hfq (middle panels, **Fig. 3.13**) because this increases the probability that the limited numbers of sRNA that are transcribed in response to a specific stressor are bound to Hfq. Therefore when the cognate target mRNA binds to Hfq there is a higher probability that the cognate ternary complex will form. Because there is a large pool of target mRNAs, increasing the affinity of the target mRNAs (cognate and unpartnered) for Hfq reduces the fraction of target mRNAs that have the chance to bind Hfq in a given period. This lowers the probability that the cognate target mRNA will be bound to one of the limited number of Hfq hexamers and diminishes the robustness of the cognate pair (lower panels, **Fig. 3.13**). Of course, selectively increasing only the affinity of the cognate target mRNA for Hfq would increase Hfq robustness and duplex formation for the cognate

pair, but this would only work if the same sRNA was always induced in response to stress; this is clearly not a realistic scenario.

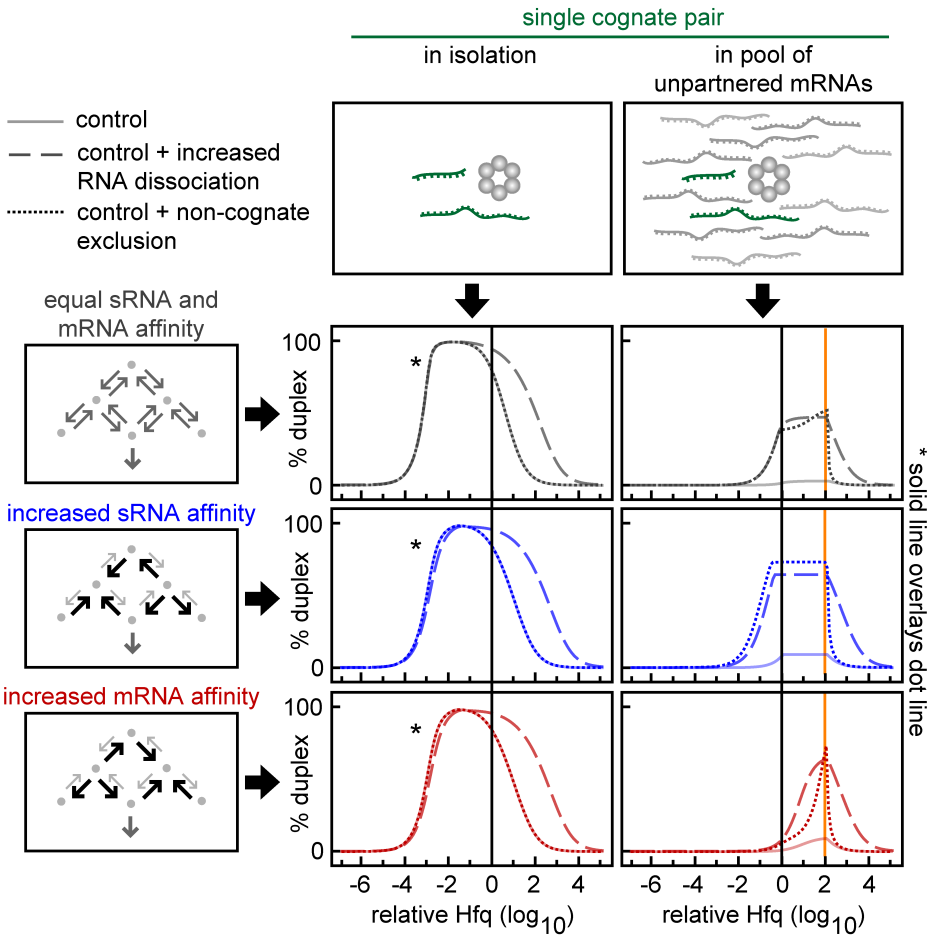

Fig. 3.13 | Duplex formation and Hfq robustness in the presence of excess, unpartnered target mRNAs. Duplex formation for a cognate sRNA-target mRNA pair in the absence (left panels) and presence (right panels) of competing non-cognate mRNAs. The sRNA only forms duplexes with its cognate target mRNA. Each plot shows a control topology with moderate RNA dissociation and no cooperativity (solid line), the control topology with increased RNA dissociation for all reaction steps (dash line) and the control topology with non-cognate exclusion which was created by selectively increasing RNA dissociation from non-cognate duplexes (dot line). The affinity of the cognate sRNA and target mRNA for Hfq was varied by increasing and decreasing two groups of rate constants (k_1, k_4, k^*_4 and k_2, k_3 and k^*_3) so that the ratio $((k_2 \cdot k_3 \cdot k^*_3)/(k_1 \cdot k_4 \cdot k^*_4))^{1/3}$ had values of 10^0, 10^{-2} and 10^2 (equal affinity, high sRNA affinity and high target mRNA affinity respectively). In these simulations, total RNA dissociation which was measured by y_9 had values of 10^0, 10^4 and 10^1 (top, middle and bottom panels respectively). Non-cognate exclusion which was measured by y_8 had values of 10^0, 10^0 and 10^1 (top, middle and bottom panels respectively). The vertical black line indicates where the Hfq concentration equals the total target mRNA concentration in the cognate pair. The vertical yellow line indicates where the Hfq concentration equals the concentration of all competing, non-cognate mRNAs.

In summary, imbalances in sRNA and target mRNA production alter Hfq robustness and maximum duplex formation and some of these effects are only apparent in the context of a network. We demonstrated that an excess of a particular sRNA can increase the efficiency of forming its duplexes but this occurs at the cost of decreased duplex formation and Hfq robustness for other sRNAs. Strategies that reduce non-cognate ternary complex formation such as bias for the sRNA-Hfq or target mRNA-Hfq branch, cooperativity, rapid RNA dissociation, and differences in the affinity of sRNAs and target mRNAs for Hfq can help compensate for imbalanced and unpartnered RNAs.

3.5 DISCUSSION

In this study we investigated a fundamental question; how is Hfq able to efficiently and robustly mediate duplex formation in the internal environment of the cell with many sRNAs and target mRNAs competing for Hfq? The difficulty of this task is exacerbated by the fact that the Hfq concentration varies with environmental cues [112-114] as does the number, types and concentrations of the sRNAs and target mRNAs [118-120]. To address this question we modeled the kinetics of the interactions between Hfq, sRNAs and target mRNAs which are the foundation of sRNA networks.

The model used in this study is very general and the basic qualitative findings are likely to apply to other realistic binding scenarios. For example, if sRNAs and target mRNAs compete for the same sites on Hfq rather than having separate binding sites as we have modeled, then this would simply increase the number of possible non-cognate ternary complexes. In this case there would be $2n^2-n$ non-cognate ternary complexes instead of n^2-n (where n is the number of sRNA-target mRNA pairs) in networks with single-partner pairing. While this reduces Hfq availability, it does not alter the fundamental kinetics of the system except that sRNA-target mRNA pairs acting in isolation would also require a mechanism to prevent non-cognate ternary complexes from forming [*i.e.* $(\text{sRNA})_2$-Hfq and $(\text{target mRNA})_2$-Hfq]. Similarly, the binding of Hfq to DNA [99,100] and polyadenylated RNAs [97,101] will decrease Hfq availability but it will not alter the basic kinetics of duplex formation.

The model was used to systematically explore the possible kinetic mechanisms for increasing the efficiency and robustness of sRNAs. We found two non-mutually exclusive mechanisms that can help achieve these goals: 1. heterotropic cooperativity; and 2. rapid RNA dissociation from Hfq complexes. Both mechanisms can increase

duplex formation and/or robustness for single cognate sRNA-target mRNA pairs acting in isolation, in networks with many competing sRNAs and target mRNAs, and when there are imbalances in sRNA and target mRNA concentrations.

Cooperativity can enhance duplex formation by increasing the proportion of bound Hfq in cognate ternary complexes. This can occur by increasing the formation and/or stability of cognate ternary complexes (*i.e.* cognate selection) or by decreasing the formation and/or stability of non-cognate ternary complexes (*i.e.* non-cognate exclusion) and singly-bound Hfq complexes. The "active" cycling model that was recently proposed [80], where the binding of one RNA to Hfq increases the dissociation of another RNA bound to Hfq, is consistent with what we have termed non-cognate exclusion. There is *in vitro* experimental evidence which is consistent with our definition for cooperativity. Kinetics measurements with synthetic RNA sequences have shown unequal formation of duplexes via the sRNA-Hfq and target mRNA-Hfq branches; this indicates that RNA binding to Hfq is affected by the presence of other RNAs bound to Hfq [74]. Furthermore, sRNAs and target mRNAs dissociate from non-cognate ternary complexes at greater rates than from singly-bound Hfq complexes which is consistent with non-cognate exclusion [31,80]. There is currently no clear evidence for or against cognate selection.

In contrast to cooperativity, high RNA dissociation rates do not alter the relative proportion of Hfq complexes but instead increase the cycling of sRNAs and target mRNAs among cognate and non-cognate ternary complexes and singly-bound Hfq complexes. The increased cycling improves the likelihood that sRNAs and target mRNA are incorporated into a cognate ternary complex before they degrade. Rapid RNA dissociation has been observed *in vitro* for non-cognate ternary complexes [80] and for cognate ternary complexes [74]. The evidence for rapid RNA dissociation from singly-bound Hfq complexes is less clear; singly-bound Hfq complexes have reported half-lives (which is a measure of the rate of RNA dissociation) in the absence of competition which varies from approximately 165 minutes [80] to 0.03 s [74] to between 2.1×10^{-5} and 4.1×10^{-5} s [31].

How sRNAs and target mRNAs interact while bound to Hfq is unclear. It may involve allosteric changes in the Hfq hexamer, electrostatic interactions or direct interactions between the sRNAs and/or target mRNAs bound to Hfq. Whatever the mechanism, *in vivo* experiments suggest that the 5′ sequences of sRNAs and target mRNAs are in at

least some cases sufficient to ensure specific interactions between the sRNAs and the target mRNAs [121]. Furthermore, a comparative analysis of sRNAs across different genomes has shown that the binding sequences, irrespective of their location, are the only regions that are consistently conserved [122].

We demonstrated that duplex formation not only depends on the kinetics of the sRNA-target mRNA pair of interest but it also depends on the kinetics and abundance of competing RNAs. This was highlighted by showing that in isolation sRNA-target mRNA pairs that form stable Hfq complexes can perform better than pairs that form unstable Hfq complexes, whereas in a sRNA network the reverse can occur. Further illustrating the importance of the environment on duplex formation, we showed that excess production of a sRNA can be used to selectively enhance the formation of one type of duplex at the cost of decreased duplex formation for other sRNA-target mRNA pairs. Similarly, we demonstrated that a pool of unpartnered target mRNAs can alter duplex formation for a cognate pair.

It is clear that more *in vitro* and *in vivo* experimental data is needed to understand duplex formation in biologically realistic scenarios where there is competition for Hfq. In particular it is important to ascertain which of the two mechanisms for reducing sequestration (cooperativity or high RNA dissociation) is more important and whether it varies among sRNAs. This could be determined by isolating singly-bound Hfq complexes, cognate ternary complexes and non-cognate ternary complexes and measuring their association and dissociation rate constants via electrophoretic mobility shift assays, surface plasmon resonance and filter binding assays [31,80]. Furthermore, because duplex formation is dependent on the number, types and concentrations of the other RNAs that are competing for Hfq, *in vivo* measurement of these factors under specific physiological conditions is desirable to obtain an accurate quantification of sRNA activity. Genome wide identification and quantification of sRNAs and target mRNAs that compete for Hfq could be obtained by deep sequencing expression analysis [123,124]. Of course there is also a need to determine the availability of Hfq under the same conditions. Direct measurement of Hfq concentration [113,125] may not be helpful because Hfq has the capacity to bind DNA and other proteins and therefore the available fraction of Hfq would be difficult to ascertain. It may be more practical to directly measure Hfq availability and competition using a reporter system [21].

Part 3: Crosstalk among Hfq-dependent small RNAs

It is important to acknowledge that even with mechanisms such as cooperativity and rapid RNA dissociation acting to optimize Hfq function, the sRNA network may not signal effectively under some conditions due to Hfq competition. Recently, it was shown *in vivo* that over-expressing a single sRNA or target mRNA can be sufficient to generally disrupt sRNA signaling due to Hfq competition [21]. The inability of Hfq to mediate sRNA signaling under only certain conditions could be a desirable feature when individual signals (*e.g.* iron deficiency, oxidative stress, nutrient limitation) need to be over-ridden or controlled centrally under specific conditions such as pathogenesis or stress. Cells could globally turn off or tune the activity of large subsets of sRNAs by varying the production of Hfq or the transcription of RNA competitors.

In conclusion, there are simple kinetic mechanisms that can increase the efficiency and robustness of Hfq activity to enable it to mediate multiple sRNA signals in parallel. These mechanisms are cooperativity and/or rapid RNA dissociation which minimize the sequestration of sRNAs, target mRNAs and Hfq. Determining the role of these mechanisms *in vivo* will require further characterization of the composition and kinetics of sRNA networks in cells under physiologically relevant conditions. The general model we have presented is a valuable framework to guide, analyze and interpret these future experiments.

3.6 MODEL

In our model [H], [S_i] and [T_j] represent the concentrations of unbound Hfq hexamer, the *i*th unbound sRNA and the *j*th unbound target mRNA respectively. [HS_i] and [HT_j] denote singly-bound Hfq hexamers with the *i*th sRNA and the *j*th target mRNA respectively. [HS_iT_j] denotes a ternary complex where Hfq is bound to the *i*th sRNA and the *j*th target mRNA. [$D_{i,j}$] denotes the sRNA-target mRNA duplex formed by the combination of the *i*th sRNA and the *j*th target mRNA. In this model, Hfq cannot bind more than one sRNA or target mRNA.

The number of differential equations needed to describe a network depends on number of types of sRNA and target mRNA species. In a network with n types of sRNA species (*i.e.* $i = 1$ to n) and m types of target mRNA species (*i.e.* $j = 1$ to m) there will be: 1. a single equation that describes the dynamics of free Hfq; 2. n equations that describe the dynamics of free sRNAs; 3. m equations that describe the dynamics of free mRNAs; 4. n equations that describe the dynamics of sRNA-Hfq complexes; 5. m equations that

describe the dynamics of target mRNA-Hfq complexes; 6. *n·m* equations that describe the dynamics of ternary Hfq complexes; and 7. *n·m* equations that describe the dynamics of sRNA-target mRNA duplexes.

$$\frac{d[\mathrm{H}]}{dt} = \alpha_{\mathrm{H}} - \beta[\mathrm{H}] - \sum_{i=1}^{n}\left(k_{1(i)}[\mathrm{H}][\mathrm{S}_i] - k_{-1(i)}[\mathrm{HS}_i]\right) - \sum_{j=1}^{m}\left(k_{2(j)}[\mathrm{H}][\mathrm{T}_j] - k_{-2(j)}[\mathrm{HT}_j]\right) + \sum_{i=1}^{n}\sum_{j=1}^{m} k_{5(i,j)}[\mathrm{HS}_i\mathrm{T}_j] \quad [3.1]$$

$$\frac{d[\mathrm{S}_i]}{dt} = \alpha_{\mathrm{S}_i} - \beta[\mathrm{S}_i] - \left(k_{1(i)}[\mathrm{H}][\mathrm{S}_i] - k_{-1(i)}[\mathrm{HS}_i]\right) - \sum_{j=1}^{m}\left(k_{4(i,j)}[\mathrm{HT}_j][\mathrm{S}_i] - k_{-4(i,j)}[\mathrm{HS}_i\mathrm{T}_j]\right) \quad [3.2]$$

$$\frac{d[\mathrm{T}_j]}{dt} = \alpha_{\mathrm{T}_j} - \beta[\mathrm{T}_j] - \left(k_{2(j)}[\mathrm{H}][\mathrm{T}_j] - k_{-2(j)}[\mathrm{HT}_j]\right) - \sum_{i=1}^{n}\left(k_{3(i,j)}[\mathrm{HS}_i][\mathrm{T}_j] - k_{-3(i,j)}[\mathrm{HS}_i\mathrm{T}_j]\right) \quad [3.3]$$

$$\frac{d[\mathrm{HS}_i]}{dt} = -\beta[\mathrm{HS}_i] + \left(k_{1(i)}[\mathrm{H}][\mathrm{S}_i] - k_{-1(i)}[\mathrm{HS}_i]\right) - \sum_{j=1}^{m}\left(k_{3(i,j)}[\mathrm{HS}_i][\mathrm{T}_j] - k_{-3(i,j)}[\mathrm{HS}_i\mathrm{T}_j]\right) \quad [3.4]$$

$$\frac{d[\mathrm{HT}_j]}{dt} = -\beta[\mathrm{HT}_j] + \left(k_{2(j)}[\mathrm{H}][\mathrm{T}_j] - k_{-2(j)}[\mathrm{HT}_j]\right) - \sum_{i=1}^{n}\left(k_{4(i,j)}[\mathrm{HT}_j][\mathrm{S}_i] - k_{-4(i,j)}[\mathrm{HS}_i\mathrm{T}_j]\right) \quad [3.5]$$

$$\frac{d[\mathrm{D}_{i,j}]}{dt} = -\beta[\mathrm{D}_{i,j}] + k_{5(i,j)}[\mathrm{HS}_i\mathrm{T}_j] \quad [3.6]$$

$$\frac{d[\mathrm{HS}_i\mathrm{T}_j]}{dt} = -\beta[\mathrm{HS}_i\mathrm{T}_j] + \left(k_{3(i,j)}[\mathrm{HS}_i][\mathrm{T}_j] - k_{-3(i,j)}[\mathrm{HS}_i\mathrm{T}_j]\right) + \left(k_{4(i,j)}[\mathrm{HT}_j][\mathrm{S}_i] - k_{-4(i,j)}[\mathrm{HS}_i\mathrm{T}_j]\right) - k_{5(i,j)}[\mathrm{HS}_i\mathrm{T}_j] \quad [3.7]$$

The parameters α_{H}, α_{Si} and α_{Tj} indicate the production rates of Hfq, the *i*th sRNA and the *j*th target mRNA respectively. $k_{1(i)}$, $k_{-1(i)}$, $k_{2(j)}$ and $k_{-2(j)}$ describe the formation and dissociation of singly-bound Hfq complexes and $k_{3(i,j)}$, $k_{-3(i,j)}$, $k_{4(i,j)}$, $k_{-4(i,j)}$ describe the formation and dissociation of ternary complexes. The *i* subscript in these rate constants denotes which of the *n* sRNAs is binding or unbinding and the *j* subscript denotes which of the *m* target mRNAs is binding or unbinding. $k_{5(i,j)}$ specifies the formation and release of a sRNA-target mRNA duplex from a Hfq ternary complex and is zero for non-cognate sRNA-target mRNA pairs (*i.e.* when $i \neq j$) unless otherwise stated. As

explained above, the rate constant for degradation and dilution (β) was the same for all species to keep the model as simple as possible; this simplification does not alter the basic qualitative results (**Fig. 3.14**).

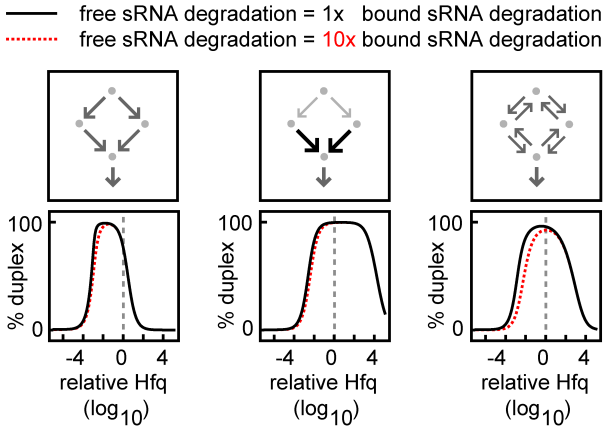

Fig. 3.14 | Varying sRNA degradation minimally affects duplex formation. The effect of the free sRNA degradation rate constant on duplex formation was examined in a system with independent sRNA and target mRNA binding (left panels), positive cooperative association (middle panels) and rapid RNA dissociation (right panels). For native sRNAs the degradation rate is typically greater than for target mRNAs and Hfq. Therefore we increased free sRNA degradation by 10-fold (compared to other simulations). The degradation of the target mRNA, Hfq and Hfq complexes were unchanged. The simulations showed that increasing the degradation of free sRNA compared to the target mRNA and Hfq had minimal affect on duplex formation (Note: the simulated difference is greater than typically occurs physiologically). The grey dash line indicates a 1:1 ratio of [total target mRNA] to [Hfq], where the [total target mRNA] = [T] + [HT] + [HST] + [D]; [T], [HT], [HST] and [D] are the concentrations of free target mRNA, target mRNA-Hfq complex, cognate ternary complex and duplex respectively.

We did not include the possibility of Hfq dodecamers and the exchange of RNAs between Hfq hexamers (which must also form a Hfq dodecameric complex) in the basic model [126,127]. This would have required the creation of additional reaction paths and therefore increased complexity for the basic model. Furthermore, Hfq dodecamers have not been consistently observed *in vivo* and a recent study has shown that single Hfq hexamers are sufficient for duplex formation [128]. Of course the latter does not exclude a role for Hfq dodecamers in duplex formation but current evidence suggests it is not essential.

The parameter values for the simulations are provided in the **Table 3.1**. Many of the kinetic parameters for duplex formation have not been quantified or the reported values vary widely (see **Discussion**). For this reason, the values for the kinetic parameters were given arbitrary values and these were varied over several orders of magnitude to ensure that biologically relevant regimes are likely to fall within the ranges selected.

Part 3: Crosstalk among Hfq-dependent small RNAs

All the data presented in this study is generated from the steady state solutions to the differential equations described above by allowing numerically generated time courses to run to convergence using the NDSolve subroutine of Mathematica 6 (Wolfram Research, Champaign, IL). Simulations were run for at least 100 times the RNA lifetime and convergence was confirmed by ensuring that the solutions at 90% and 100% of the simulation differed by no more 1×10^{-10}.

Table 3.1 | Kinetic parameters used in the simulations

Fig 3.2	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
k_1	conc.$^{-1}$·time^{-1}	$10^{6.5}$	$10^{4.5}$	$10^{2.5}$	$10^{4.5}$	$10^{2.5}$	$10^{0.5}$	$10^{2.5}$	$10^{0.5}$	$10^{-1.5}$
k_2	conc.$^{-1}$·time^{-1}	$10^{2.5}$	$10^{4.5}$	$10^{6.5}$	$10^{0.5}$	$10^{2.5}$	$10^{4.5}$	$10^{-1.5}$	$10^{0.5}$	$10^{2.5}$
k_3	conc.$^{-1}$·time^{-1}	$10^{2.5}$	$10^{4.5}$	$10^{6.5}$	$10^{0.5}$	$10^{2.5}$	$10^{4.5}$	$10^{-1.5}$	$10^{0.5}$	$10^{2.5}$
k_4	conc.$^{-1}$·time^{-1}	$10^{6.5}$	$10^{4.5}$	$10^{2.5}$	$10^{4.5}$	$10^{2.5}$	$10^{0.5}$	$10^{2.5}$	$10^{0.5}$	$10^{-1.5}$
k_5	time^{-1}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}
$k_{-1,-2,-3,-4}$	time^{-1}	0	0	0	0	0	0	0	0	0
$\alpha_{S,T}$	conc.·time^{-1}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}
β	time^{-1}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}

Fig 3.3C	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
k_1	conc.$^{-1}$·time^{-1}	$10^{6.5}$	$10^{4.5}$	$10^{2.5}$	$10^{4.5}$	$10^{2.5}$	$10^{0.5}$	$10^{2.5}$	$10^{-1.5}$	$10^{-1.5}$
k_2	conc.$^{-1}$·time^{-1}	$10^{2.5}$	$10^{4.5}$	$10^{6.5}$	$10^{0.5}$	$10^{2.5}$	$10^{4.5}$	$10^{-1.5}$	$10^{-1.5}$	$10^{2.5}$
k_3	conc.$^{-1}$·time^{-1}	$10^{2.5}$	$10^{0.5}$	$10^{-1.5}$	$10^{4.5}$	$10^{2.5}$	$10^{0.5}$	$10^{6.5}$	$10^{6.5}$	$10^{2.5}$
k_4	conc.$^{-1}$·time^{-1}	$10^{-1.5}$	$10^{0.5}$	$10^{2.5}$	$10^{0.5}$	$10^{2.5}$	$10^{4.5}$	$10^{2.5}$	$10^{6.5}$	$10^{6.5}$
k_5	time^{-1}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}	10^{3}
$k_{-1,-2,-3,-4}$	time^{-1}	0	0	0	0	0	0	0	0	0
$\alpha_{S,T}$	conc.·time^{-1}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}
β	time^{-1}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}

Fig 3.3F	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2,3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	$10^{2.5}$	$10^{2.5}$	$10^{2.5}$	$10^{2.5}$	$10^{2.5}$	$10^{2.5}$	$10^{2.5}$	$10^{2.5}$
k_5	time	10^{0}	10^{3}	10^{6}	10^{0}	10^{3}	10^{6}	10^{0}	10^{3}	10^{6}
$k_{-1,-2,-3,-4}$	time	10^{8}	10^{8}	10^{8}	10^{4}	10^{4}	10^{4}	10^{0}	10^{0}	10^{0}
$\alpha_{S,T}$	conc.·time^{-1}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}	10^{2}
β	time^{-1}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}	10^{0}

Fig 3.4A	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{0.5}$
$k_{3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{4.5}$
k_5	time^{-1}	10^{3}	n/a	10^{3}	n/a	n/a	n/a	10^{3}	n/a	10^{3}
$k_{-1,-2,-3,-4}$	time^{-1}	0	n/a	0	n/a	n/a	n/a	10^{4}	n/a	10^{4}
$\alpha_{S,T}$	conc.·time^{-1}	10^{2}	n/a	10^{2}	n/a	n/a	n/a	10^{2}	n/a	10^{2}
β	time^{-1}	10^{0}	n/a	10^{0}	n/a	n/a	n/a	10^{0}	n/a	10^{0}

Part 3: Crosstalk among Hfq-dependent small RNAs

Fig 3.4B	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{0.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{0.5}$
$k_{3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{4.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{4.5}$
k_5	time^{-1}	10^3	n/a	10^3	n/a	n/a	n/a	10^3	n/a	10^3
$k_{-1,-2}$	time^{-1}	10^4	n/a	10^4	n/a	n/a	n/a	10^6	n/a	10^6
$k_{-3,-4}$	time^{-1}	10^4	n/a	10^4	n/a	n/a	n/a	10^2	n/a	10^2
$\alpha_{S,T}$	conc.·time^{-1}	10^2	n/a	10^2	n/a	n/a	n/a	10^2	n/a	10^2
β	time^{-1}	10^0	n/a	10^0	n/a	n/a	n/a	10^0	n/a	10^0

Fig 3.6	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2,3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{2.5}$
k_5	time^{-1}	10^3	n/a	10^3	n/a	n/a	n/a	10^3	n/a	10^3
$k_{-1,-2,-3,-4}$	time^{-1}	10^0	n/a	10^0	n/a	n/a	n/a	10^0	n/a	10^0
$\alpha_{S,T}$	conc.·time^{-1}	10^2	n/a	10^2	n/a	n/a	n/a	10^2	n/a	10^2
β	time^{-1}	10^0	n/a	10^0	n/a	n/a	n/a	10^0	n/a	10^0

Fig 3.7	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{2.5}$
$k_{3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{3.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{3.5}$
$k^*_{3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{1.5}$	n/a	$10^{1.5}$
k_5	time^{-1}	10^3	n/a	10^3	n/a	n/a	n/a	10^3	n/a	10^3
$k_{-1,-2}$	time^{-1}	10^0	n/a	10^0	n/a	n/a	n/a	10^0	n/a	10^0
$k_{-3,-4}$	time^{-1}	10^0	n/a	10^{-1}	n/a	n/a	n/a	10^0	n/a	10^{-1}
$k^*_{-3,-4}$	time^{-1}	10^0	n/a	10^0	n/a	n/a	n/a	10^1	n/a	10^1
$\alpha_{S,T}$	conc.·time^{-1}	10^2	n/a	10^2	n/a	n/a	n/a	10^2	n/a	10^2
β	time^{-1}	10^0	n/a	10^0	n/a	n/a	n/a	10^0	n/a	10^0

Fig 3.8	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2,3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{2.5}$
$k^*_{3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{2.5}$
k_5	time^{-1}	10^3	n/a	10^3	n/a	n/a	n/a	10^6	n/a	10^6
$k_{-1,-2,-3,-4}$	time^{-1}	10^0	n/a	10^4	n/a	n/a	n/a	10^0	n/a	10^4
$k^*_{-3,-4}$	time^{-1}	10^0	n/a	10^4	n/a	n/a	n/a	10^0	n/a	10^4
$\alpha_{S,T}$	conc.·time^{-1}	10^2	n/a	10^2	n/a	n/a	n/a	10^2	n/a	10^2
β	time^{-1}	10^0	n/a	10^0	n/a	n/a	n/a	10^0	n/a	10^0

Part 3: Crosstalk among Hfq-dependent small RNAs

Fig 3.9‡	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2,3,4}$	conc.$^{-1}$·time^{-1}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k^*_{3,4}$	conc.$^{-1}$·time^{-1}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
k_5	time^{-1}	$10^{1.5}$	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k_{-1,-2,-3,-4}$	time^{-1}	10^{0}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k^*_{-3,-4}$	time^{-1}	10^{0}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$\alpha_{T,S}$	conc.·time^{-1}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
β	time^{-1}	10^{0}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Fig 3.9◊	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2,3,4}$	conc.$^{-1}$·time^{-1}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k^*_{3,4}$	conc.$^{-1}$·time^{-1}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
k_5	time^{-1}	$10^{1.5}$	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k_{-1,-2,-3,-4}$	time^{-1}	10^{4}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k^*_{-3,-4}$	time^{-1}	10^{4}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$\alpha_{T,S}$	conc.·time^{-1}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
β	time^{-1}	10^{0}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Fig 3.10	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2,3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{2.5}$
$k^*_{3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	n/a	n/a	n/a	$10^{2.5}$	n/a	$10^{2.5}$
k_5	time^{-1}	10^{3}	n/a	10^{3}	n/a	n/a	n/a	10^{3}	n/a	10^{3}
k^*_5	time^{-1}	0	n/a	0	n/a	n/a	n/a	10^{3}	n/a	10^{3}
$k_{-1,-2,-3,-4}$	time^{-1}	10^{0}	n/a	10^{4}	n/a	n/a	n/a	10^{0}	n/a	10^{4}
$k^*_{-3,-4}$	time^{-1}	10^{0}	n/a	10^{4}	n/a	n/a	n/a	10^{0}	n/a	10^{4}
$\alpha_{S,T}$	conc.·time^{-1}	10^{2}	n/a	10^{2}	n/a	n/a	n/a	10^{2}	n/a	10^{2}
β	time^{-1}	10^{0}	n/a	10^{0}	n/a	n/a	n/a	10^{0}	n/a	10^{0}

Fig 3.11	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,3}$	conc.$^{-1}$·time^{-1}	10^{6}	10^{-2}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a
$k_{2,4}$	conc.$^{-1}$·time^{-1}	10^{-2}	10^{6}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a
k_5	time^{-1}	10^{1}	10^{1}	10^{1}	n/a	n/a	n/a	n/a	n/a	n/a
$k_{-1,-2,-3,-4}$	time^{-1}	0	0	0	n/a	n/a	n/a	n/a	n/a	n/a
α_S	conc.·time^{-1}	17-83	17-83	17-83	n/a	n/a	n/a	n/a	n/a	n/a
α_T	conc.·time^{-1}	17-83	17-83	17-83	n/a	n/a	n/a	n/a	n/a	n/a
α_H	conc.·time^{-1}	10^{2}	10^{2}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a
β	time^{-1}	10^{0}	10^{0}	10^{0}	n/a	n/a	n/a	n/a	n/a	n/a

Part 3: Crosstalk among Hfq-dependent small RNAs

Fig 3.12[£]	units	TL	TC	TR	ML	MC	MR	BL	BC	BR
$k_{1,2,3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k^*_{3,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
k_5	time^{-1}	$10^{1.5}$	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k_{-1,-2,-3,-4}$	time^{-1}	10^{1}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$k^*_{-3,-4}$	time^{-1}	10^{1}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
α_{S1}	conc.·time^{-1}	10^{2}-10^{4}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
$\alpha_{T1,S2,T2}$	conc.·time^{-1}	10^{2}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a
β	time^{-1}	10^{0}	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Fig 3.13[¶]	units	TL	TC	TR	ML	MC	MR	BL	BC	BR
$k_{1,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{3.5}$	n/a	$10^{3.5}$	$10^{1.5}$	n/a	$10^{1.5}$
$k_{2,3}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{1.5}$	n/a	$10^{1.5}$	$10^{3.5}$	n/a	$10^{3.5}$
k^*_3	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{1.5}$	n/a	$10^{1.5}$	$10^{3.5}$	n/a	$10^{3.5}$
k^*_4	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{3.5}$	n/a	$10^{3.5}$	$10^{1.5}$	n/a	$10^{1.5}$
k_5	time^{-1}	10^{3}	n/a	10^{3}	10^{3}	n/a	10^{3}	10^{3}	n/a	10^{3}
$k_{-1,-4}$	time^{-1}	10^{0}	n/a	10^{0}	10^{-1}	n/a	10^{-1}	10^{1}	n/a	10^{1}
$k_{-2,-3}$	time^{-1}	10^{0}	n/a	10^{0}	10^{1}	n/a	10^{1}	10^{-1}	n/a	10^{-1}
k^*_{-3}	time^{-1}	10^{0}	n/a	10^{0}	10^{1}	n/a	10^{1}	10^{-1}	n/a	10^{-1}
k^*_{-4}	time^{-1}	10^{0}	n/a	10^{0}	10^{-1}	n/a	10^{-1}	10^{1}	n/a	10^{1}
$\alpha_{S1,T1}$	conc.·time^{-1}	10^{2}	n/a	10^{2}	10^{2}	n/a	10^{2}	10^{2}	n/a	10^{2}
$\alpha_{T\ non\text{-}cog.}$	conc.·time^{-1}	0	n/a	10^{4}	0	n/a	10^{4}	0	n/a	10^{4}
β	time^{-1}	10^{0}	n/a	10^{0}	10^{0}	n/a	10^{0}	10^{0}	n/a	10^{0}

Fig 3.13[¥]	units	TL	TC	TR	ML	MC	MR	BL	BC	BR
$k_{1,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{3.5}$	n/a	$10^{3.5}$	$10^{1.5}$	n/a	$10^{1.5}$
$k_{2,3}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{1.5}$	n/a	$10^{1.5}$	$10^{3.5}$	n/a	$10^{3.5}$
k^*_3	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{1.5}$	n/a	$10^{1.5}$	$10^{3.5}$	n/a	$10^{3.5}$
k^*_4	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{3.5}$	n/a	$10^{3.5}$	$10^{1.5}$	n/a	$10^{1.5}$
k_5	time^{-1}	10^{3}	n/a	10^{3}	10^{3}	n/a	10^{3}	10^{3}	n/a	10^{3}
$k_{-1,-4}$	time^{-1}	10^{2}	n/a	10^{2}	10^{1}	n/a	10^{1}	10^{3}	n/a	10^{3}
$k_{-2,-3}$	time^{-1}	10^{2}	n/a	10^{2}	10^{3}	n/a	10^{3}	10^{1}	n/a	10^{1}
k^*_{-3}	time^{-1}	10^{2}	n/a	10^{2}	10^{3}	n/a	10^{3}	10^{1}	n/a	10^{1}
k^*_{-4}	time^{-1}	10^{2}	n/a	10^{2}	10^{1}	n/a	10^{1}	10^{3}	n/a	10^{3}
$\alpha_{S1,T1}$	conc.·time^{-1}	10^{2}	n/a	10^{2}	10^{2}	n/a	10^{2}	10^{2}	n/a	10^{2}
$\alpha_{T\ non\text{-}cog.}$	conc.·time^{-1}	0	n/a	10^{4}	0	n/a	10^{4}	0	n/a	10^{4}
β	time^{-1}	10^{0}	n/a	10^{0}	10^{0}	n/a	10^{0}	10^{0}	n/a	10^{0}

Fig 3.13§	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,4}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{3.5}$	n/a	$10^{3.5}$	$10^{1.5}$	n/a	$10^{1.5}$
$k_{2,3}$	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{1.5}$	n/a	$10^{1.5}$	$10^{3.5}$	n/a	$10^{3.5}$
k^*_3	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{1.5}$	n/a	$10^{1.5}$	$10^{3.5}$	n/a	$10^{3.5}$
k^*_4	conc.$^{-1}$·time^{-1}	$10^{2.5}$	n/a	$10^{2.5}$	$10^{3.5}$	n/a	$10^{3.5}$	$10^{1.5}$	n/a	$10^{1.5}$
k_5	time^{-1}	10^{3}	n/a	10^{3}	10^{3}	n/a	10^{3}	10^{3}	n/a	10^{3}
$k_{-1,-4}$	time^{-1}	10^{0}	n/a	10^{0}	10^{-1}	n/a	10^{-1}	10^{1}	n/a	10^{1}
$k_{-2,-3}$	time^{-1}	10^{0}	n/a	10^{0}	10^{1}	n/a	10^{1}	10^{-1}	n/a	10^{-1}
k^*_{-3}	time^{-1}	10^{2}	n/a	10^{2}	10^{3}	n/a	10^{3}	10^{1}	n/a	10^{1}
k^*_{-4}	time^{-1}	10^{2}	n/a	10^{2}	10^{1}	n/a	10^{1}	10^{3}	n/a	10^{3}
$\alpha_{S1,T1}$	conc. ·time^{-1}	10^{2}	n/a	10^{2}	10^{2}	n/a	10^{2}	10^{2}	n/a	10^{2}
$\alpha_{T\ non\text{-}cog.}$	conc. ·time^{-1}	0	n/a	10^{4}	0	n/a	10^{4}	0	n/a	10^{4}
β	time^{-1}	10^{0}	n/a	10^{0}	10^{0}	n/a	10^{0}	10^{0}	n/a	10^{0}

Fig 3.14	**units**	**TL**	**TC**	**TR**	**ML**	**MC**	**MR**	**BL**	**BC**	**BR**
$k_{1,2}$	conc.$^{-1}$·time^{-1}	n/a	n/a	n/a	n/a	n/a	n/a	$10^{2.5}$	$10^{0.5}$	$10^{2.5}$
$k_{3,4}$	conc.$^{-1}$·time^{-1}	n/a	n/a	n/a	n/a	n/a	n/a	$10^{2.5}$	$10^{4.5}$	$10^{2.5}$
k_5	time^{-1}	n/a	n/a	n/a	n/a	n/a	n/a	10^{3}	10^{3}	10^{3}
$k_{-1,-2,-3,-4}$	time^{-1}	n/a	n/a	n/a	n/a	n/a	n/a	0	0	10^{4}
$\alpha_{S,T}$	conc. ·time^{-1}	n/a	n/a	n/a	n/a	n/a	n/a	10^{2}	10^{2}	10^{2}
$\beta_{S\ free\ (blk)}$	time^{-1}	n/a	n/a	n/a	n/a	n/a	n/a	10^{0}	10^{0}	10^{0}
$\beta_{S\ free\ (red)}$	time^{-1}	n/a	n/a	n/a	n/a	n/a	n/a	10^{1}	10^{1}	10^{1}
β	time^{-1}	n/a	n/a	n/a	n/a	n/a	n/a	10^{0}	10^{0}	10^{0}

Table 3.1 | Kinetic parameters used in the simulations. Abbreviations used to specify the panels in the figures are: top left (TL), top center (TC), top right (TR), middle left (ML), middle center (MC), middle right (MR), bottom left (BL), bottom center (BC) and bottom right (BR). ‡ indicates sRNA-target mRNA pairs that form "stable" complexes with Hfq and ◊ indicates "unstable" complexes with Hfq (**Fig 3.9U**). £ Two sRNA-target mRNA pairs were simulated which have the same kinetic parameters (both panels represent the same set of simulations with varying $sRNA_1$ production). ¶ Control topologies (solid curves). ¥ Topologies with increased RNA dissociation relative to the control (dash curves). § Topologies exhibiting non-cognate exclusion (dot curves). Hfq production (α_H) in all panels is varied from 10^{-5} to 10^{7} concentration·time^{-1} unless otherwise indicated. conc. = concentration.

4. Crosstalk in the CsrA system

Having discussed crosstalk among mRNAs (as well as sRNAs) in the context of the Hfq-dependent sRNA network, we now shift our focus and examine crosstalk among mRNAs in the CsrA regulatory system. This system also constitutes an RNA regulatory network because the non-coding RNAs CsrB and CsrC are shared regulators of the translation of numerous mRNAs.

4.1 ABSTRACT

While signal specificity is often essential for cell survival, many signaling molecules are shared between pathways which creates avenues for unwanted crosstalk [41]. This is particularly true in the growing field of translational regulatory networks (*e.g.* the CsrA network and the Hfq-dependent small RNA network) in which small RNAs and proteins regulate stoichiometrically a fluctuating pool of diverse target mRNAs; crosstalk arises in these networks from competition for shared regulators [16,20,47]. While many strategies have been identified to mitigate crosstalk among biochemical pathways (including protein scaffolding [129], cross-pathway inhibition [130] and kinetic insulation [41]), they largely address crosstalk between receptor proteins or phosphorelays in ways that are not readily adaptable to translational regulatory networks. In this study, we reverse engineered the CsrA regulatory network to show that it can buffer crosstalk among its mRNA targets. Specifically, we show that non-coding RNAs (*e.g.* CsrB) which bind and sequester the system's central regulator (CsrA) can buffer competition for that central regulator similarly to how a pH buffer damps fluctuations in free protons. We show that target mRNA translation is affected far less by expression of a competitor mRNA in cells with high concentrations of both CsrA and CsrB than in cells with lower CsrA and CsrB levels because the competitor mRNA sequesters a much smaller fraction of the total CsrA in the former case than in the latter. Unlike many mechanisms for minimizing crosstalk, buffering by sequestration is inherently scalable in that no additional regulatory components are required to keep new pathways (*i.e.* mRNA targets) insulated as they are added into the network.

4.2 INTRODUCTION

Preventing crosstalk between distinct signaling pathways is a general problem faced by nearly all modes of communication from the electronic to the biochemical. Crosstalk is particularly problematic in biochemical systems since many distinct signaling pathways share common regulatory molecules [41]. While several strategies have been described for mitigating the crosstalk that arises through these shared regulatory components (*e.g.* protein scaffolding [129], cross-pathway inhibition [130] and kinetic insulation [41]) most of these do not scale effectively with the number of pathways to be isolated; in other words, the complexity of these mechanisms (*e.g.* the number of distinct scaffold proteins, cross-inhibitor molecules, or signaling timescales required) grows as the number of pathways to be isolated increases. In this study, we reverse engineered the CsrA regulatory network and show that it possesses a mechanism to mitigate crosstalk among the mRNA targets of CsrA (**Fig. 4.1**). Unlike the above strategies for avoiding crosstalk, the strategy we describe using CsrA regulation scales well for large networks because it relies on the principle of buffering; as such, the effect of adding new mRNA competitors for the central regulator (CsrA) can be arbitrarily diminished by increasing the concentration of that central regulator in tandem with its buffering agent (CsrB).

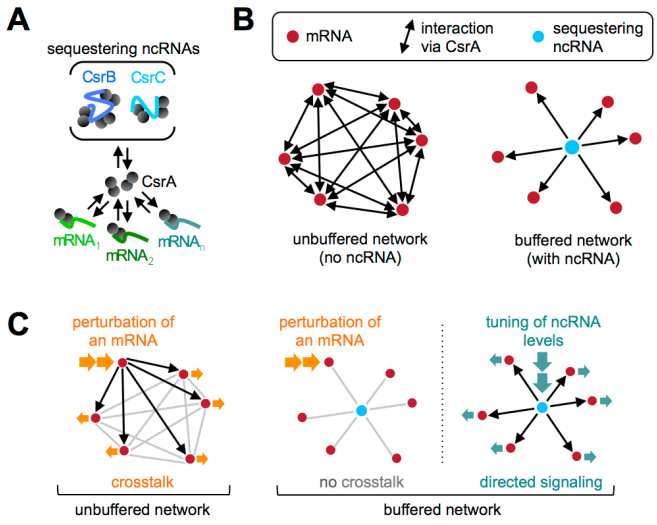

Fig. 4.1 | CsrA regulatory network. (*A*) The CsrA protein binds to and silences numerous target mRNAs. CsrA is sequestered by non-coding RNAs (ncRNAs; *e.g.* CsrB and CsrC) which inhibit that silencing. (*B*) Buffered and unbuffered networks. Left: CsrA 'links' mRNAs by allowing changes in the concentration of one mRNA to influence the translation of others (through competition for CsrA) (**Fig. 4.2, 4.7**). Right: Sequestering ncRNAs can buffer this crosstalk (**Fig. 4.3, 4.7**). (*C*) Effects of perturbations. Left: Perturbing the concentration of one mRNA can affect translation of other mRNAs via "substrate availability crosstalk" [42] (unbuffered network). Center: perturbations of mRNA concentrations are strongly damped which mitigates crosstalk (buffered network). Right: Target

mRNAs remain sensitive to changes in ncRNAs since they compete for CsrA much more effectively than other mRNAs (**Fig. 4.5**) (buffered network).

The CsrA network is of general interest, not only because it plays a central and conserved role in controlling bacterial metabolism, motility, biofilm formation and pathogenesis, but also because it represents a larger class of "translational regulatory network" in which a central translational regulator coordinates protein expression from multiple genetically distinct mRNAs [16,19,33]. All of these networks are potentially susceptible to crosstalk via competition for the central regulator through what has been termed "substrate availability crosstalk" [42] (**Fig. 4.1**). In *Escherichia coli*, the CsrA network consists of the CsrA protein, numerous CsrA target genes (*e.g. cstA*, *pgaABCD*, *glgCAP*, *hfq*, *ycdT* and *ydeH*, among others) and upstream regulators of CsrA which include the sequestering non-coding RNAs (ncRNAs) CsrB and CsrC. CsrA typically regulates its target genes by binding to and blocking the ribosomal binding site on the mRNA, thus silencing translation of the target protein [19,32,33]. Because the ncRNAs CsrB and CsrC contain multiple CsrA binding sites closely resembling those found on the target mRNAs (enough for binding to ~nine and ~three to four CsrA dimers respectively), they are able to compete with target mRNAs for binding to CsrA; in this way, CsrB and CsrC act as inhibitors of CsrA activity [49,50].

This study is divided into two parts. In the first part, we characterize the steady state dose/response relationships between a single target mRNA, CsrA, and the sequestering ncRNAs CsrB and CsrC to lay the foundation for understanding crosstalk, buffering and coordinated signaling. Specifically, we demonstrate (i) that CsrA is sequestered by the target mRNAs it regulates (**Fig. 4.2**), (ii) that the addition of a sequestering non-coding RNA (ncRNA) buffers the target mRNA against changes in the CsrA concentration (**Fig. 4.3**) and (iii) that for ncRNAs to be effective signaling molecules, CsrA levels must reside within a range set by the amount of target mRNA and the amount of ncRNA produced upon induction: too little CsrA will not silence the target mRNAs and too much CsrA will not be sufficiently sequestered when the ncRNA is turned on (**Fig. 4.4**).

In the second part, we introduce additional target mRNAs into our analysis to quantify competition for CsrA and examine crosstalk among mRNA targets. Specifically, we demonstrate that when transcribed from identical promoters, target mRNAs are much less effective at competing for CsrA than are the ncRNAs CsrB and CsrC; this places CsrB and CsrC in a unique position where they can modulate the availability of free

CsrA *in vivo* much more strongly than their mRNA counterparts (**Fig. 4.5**). Additionally, we show that the magnitude of crosstalk between target mRNAs can be tuned by altering the concentrations of CsrA and its sequestering ncRNA partner (**Fig. 4.7**). When CsrA and CsrB levels are low, the translation of individual mRNAs is sensitive to the presence of other mRNAs which can compete for the small pool of CsrA; when CsrA and CsrB levels are high, the system can achieve the same level of 'free' CsrA activity (and thus the same level of target expression) as with low CsrA and CsrB but with diminished sensitivity to the presence of extra mRNA competitors.

4.3 RESULTS

4.3.1 Modeling Summary

To examine crosstalk in the CsrA regulatory network we incorporated competing target mRNAs into our previously described model of the CsrA cascade [14]. In brief, this model used ordinary differential equations (details below) to describe the production, clearance, association, dissociation and/or catalytic activity of the target mRNAs, target proteins, CsrA, sequestering non-coding RNAs (*e.g.* CsrB, CsrC), and the associated complexes. Except where otherwise stated, the parameters used in the simulations in this study correspond to those used in our prior work [14].

4.3.2 Experimental System

Except where otherwise stated, experiments in this study were preformed using fully synthetic gene circuits in strains with chromosomal *csrA*, *csrB*, *csrC*, *csrD*, *glgCAP* and *pgaABCD* deleted; additionally, experiments were preformed in exponentially growing cells at steady state. As described previously [14], deletion of *glgCAP* was necessary for the survival of cells lacking *csrA* [56] and deletion of *pgaABCD* was necessary to allow cells lacking *csrA* to be made competent for transformation; deletion of *csrA*, *csrB*, *csrC* and *csrD* allowed us to remove the known feedback regulation present in the native system.

In this study, we used several synthetic targets for CsrA including *glgC-gfp*, *hfq-gfp*, *glgC-rfp*, *pgaA-rfp*, *hfq-rfp*, *cstA-rfp* and *ydeH-rfp*. To construct each target gene, (i) the 5' untranslated region (UTR) of a native mRNA known to include both CsrA binding sites a ribosomal binding site (RBS) and (ii) several codons of the coding sequence from the native target gene were translationally fused to either the *gfp* or *rfp* (*mCherry*) coding sequence (details below). In each case CsrA silences the translation of GFP or RFP

allowing expression of these "target proteins" to be quantified by fluorescence (**Fig. 4.6**). Full-length transcripts of CsrA targets were not used as competitors because expression of selected full-length transcripts (*e.g. cstA*) dramatically inhibited cell growth (doubling time = τ = 56 +/- 1 min without *cstA* induction versus 120 +/- 15 min with *cstA*; HL6105); in contrast, expressing RFP had affected cell growth only minimally (τ = 53 +/- 1 min without *rfp* induction versus 52 +/- 1 min with *rfp*; HL6055).

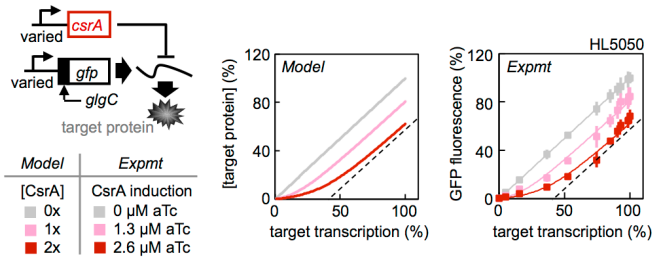

Fig. 4.2 | CsrA can be sequestered by mRNA targets. Error bars are SEM of duplicate measurements. The steady-state concentration of target protein (approximated by GFP fluorescence) is presented as a function of target transcription with three levels of CsrA induction. Center: The model predicts threshold linear behavior in the presence of CsrA. Dashed lines have the slope of the modeled gray line. Right: Target mRNA transcription (from PLlacO-1:*glgC-gfp*) is induced to varying degrees using IPTG; CsrA production (from PLtetO-1:st7:*csrA*) is induced using aTc. The percent of transcription of *glgC-gfp* (relative to the maximum for the promoter) is estimated from the IPTG concentration used and the GFP fluorescence level of PLlacO-1:*glgC-gfp* at zero aTc (HL5050). Modeled curves (from center panel) are reproduced with the *in vivo* data (right panel) as a guide to the eye.

4.3.3 CsrA can be Sequestered by mRNA Targets

To demonstrate that CsrA can be sequestered by the target mRNAs that it regulates, we measured the dose/response relationship between CsrA and one of its target mRNAs. Experimentally, we accomplished this by placing *csrA* and *glgC-gfp* (a target gene) under the control of orthogonal inducible promoters (PLtetO-1 and PLlacO-1 respectively) [54]; we then varied the transcription rate of the target mRNA at selected levels of *csrA* induction and measured GFP fluorescence at steady state (**Fig. 4.2**). Our model predicts that when varying the rate of transcription of the target mRNA, expression of the target protein will exhibit threshold linear behavior in the presence of CsrA (**Fig. 4.2**, red lines) instead of the simple linear behavior exhibited in its absence (**Fig. 4.2**, grey lines). This threshold linear behavior occurs because a fixed quantity of CsrA dimers can only silence a fixed quantity of target transcripts per unit time. Once the number of target transcripts exceeds the amount that the available CsrA can silence (termed the "stoichiometric point"), any further increase in transcription causes an unattenuated increase in translation; in other words, target protein levels would increase with target mRNA transcription along the same slope as if CsrA were absent

(**Fig. 4.2**, dashed black line). Past the stoichiometric point, no additional silencing occurs because the mRNAs that are silenced have sequestered the available CsrA away from other copies of the same transcript. Our *in vivo* results confirm the predicted threshold linear behavior; in this way, we confirm that CsrA is sequestered by its target mRNAs. We note that threshold linear responses have been observed previously in small RNA regulation [47] where target mRNAs compete to be silenced by small RNA molecules rather than by CsrA.

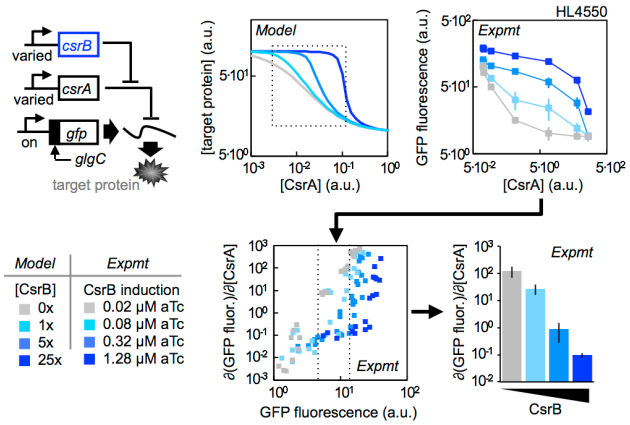

Fig. 4.3 | CsrB buffers the target mRNA against changes in CsrA levels. Error bars are SEM of five (upper right panel) or three or more (lower right panel) measurements. Top Row: The steady-state concentration of target protein (approximated by GFP fluorescence) is presented as a function of the CsrA concentration with four levels of CsrB induction. Center Top: CsrB inhibits CsrA. The dotted box shows the region for comparison with the *in vivo* data. Right Top: CsrA (from PLlacO-1:st7:*csrA*) is induced to varying degrees using IPTG; CsrB production (from PLtetO-1:*csrB*) is induced using aTc. The relative concentration of CsrA ([CsrA]) is estimated from the IPTG concentration used and the GFP fluorescence level of PLlacO-1:st7:*gfp* (HL4510) as described previously [14]. Center Bottom: Scatterplot of the derivative of GFP fluorescence with respect to CsrA concentration (y-axis) as a function of GFP fluorescence (x-axis) and level of CsrB induction; the coordinates for each point were determined by averaging the GFP fluorescence of two samples from the same overnight culture with adjacent values of CsrA induction (to find the 'x' value) and by dividing the difference in GFP fluorescence of those two samples by the difference in the estimated relative [CsrA] for those samples (to find the 'y' value). n = 30 points for each level of CsrB induction. Right Bottom: Average of the y-axis value of the points between the dotted lines in the center bottom panel, sorted by level of CsrB induction.

4.3.4 CsrB Buffers the Target mRNA Against Changes in CsrA Levels

To demonstrate that sequestering ncRNAs can buffer target mRNA translation against changes in the CsrA concentration, we measured the dose/response relationship between CsrA and CsrB. Experimentally, we accomplished this by placing *csrA* and *csrB* under the control of inducible promoters (PLlacO-1 and PLtetO-1 respectively); *glgC-gfp* was expressed constitutively. Using this system, we varied the expression of CsrA at selected concentrations of CsrB and measured target protein levels at steady state; as expected, the model predicts and our experiments confirm that with higher levels of CsrB, more CsrA is required to repress target expression to a comparable level (**Fig. 4.3**,

top row). Additionally, our experiments show that increasing CsrB levels decreases the sensitivity of target expression to changes in CsrA levels (*i.e.* d[target protein]/d[CsrA] which is approximated by d(GFP fluorescence)/d[CsrA]) at steady state) (**Fig. 4.3**, bottom row); by comparing points with roughly same level of GFP fluorescence (*e.g.* points that lie between the two vertical dotted lines) we show that higher levels of CsrB induction correspond to lower levels of sensitivity to CsrA at comparable levels of CsrA activity (**Fig. 4.3** bottom right). In summary, we have shown that a sequestering ncRNA can dampen (or "buffer") the translation of the target mRNA against changes in CsrA, making those changes less influential.

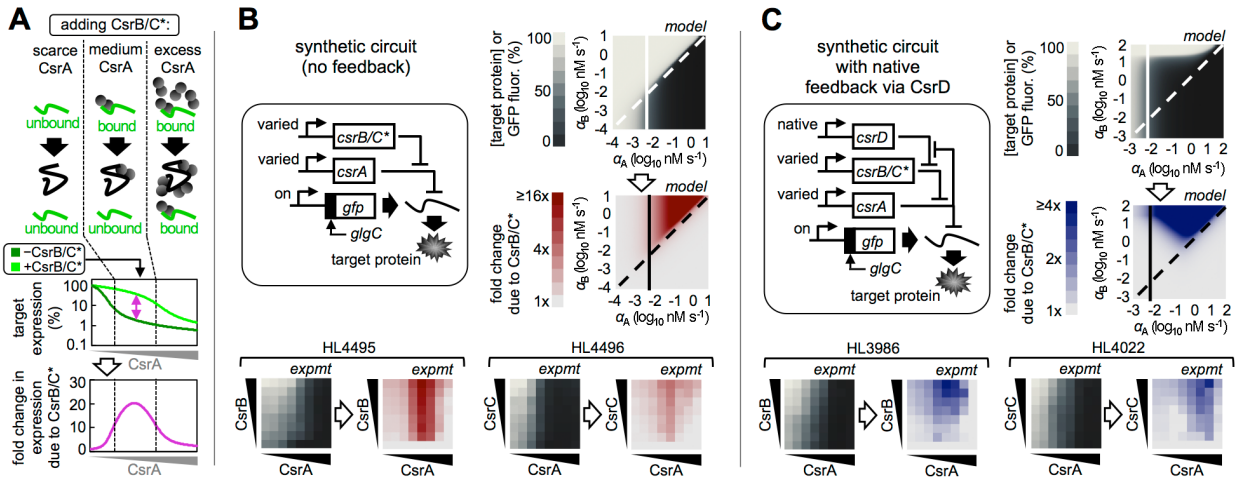

Fig. 4.4 | Signaling via sequestration is sensitive to CsrA levels. Experimental heatmaps are the average of two or more measurements per pixel; data points exhibiting multimodal distributions of cell fluorescence (which occurred sporadically at the highest levels of CsrA induction) were not included in the averaging. (*A*) Schematic and illustrative plots showing that competitors for CsrA have most influence at intermediate CsrA concentrations. *CsrB/C indicates CsrB or CsrC. (*B*) Fully synthetic circuit without feedback. Grayscale heatmaps show either the steady-state concentration of the target protein (model) or steady-state GFP fluorescence (expmt) as a function of the production rates of CsrA and either CsrB or CsrC. Color heatmaps show the ratio of [target protein] (model) or GFP fluorescence (expmt) in the *presence* of non-zero ncRNA induction to those same measures in the *absence* of ncRNA induction. (*C*) Circuit with synthetic *csrA*, synthetic *csrB* or *csrC* and native *csrD* (and thus with negative feedback). Grayscale and color heatmaps are as in *B*.

4.3.5 Signaling via Sequestration is Sensitive to CsrA Levels

Having examined how CsrB affects a target's response to changes in CsrA (**Fig. 4.3**), we now examine how CsrA affect a target's response to changes in CsrB (**Fig. 4.4**). Our model predicts that for CsrB or another ncRNA to alter the expression of a CsrA target, the CsrA levels must be neither too high nor too low. Low CsrA levels (*i.e.* a level of CsrA that is insufficient to repress the available target mRNA) prevent CsrB from signaling because CsrB cannot reduce the silencing of the target mRNAs (since there is

little to no silencing to begin with); in contrast, an excess of CsrA can bind all of the available target mRNA and CsrB simultaneously and CsrB will not stop CsrA from silencing its targets (**Fig. 4.4*A***). In this way CsrB can only play a meaningful regulatory role when the 'off' and 'on' expression levels of CsrB are well adjusted relative to the CsrA concentration so that moving CsrB expression from the 'off' state to the 'on' state will cause the system to cross the stoichiometric point where CsrA and CsrB levels are balanced (dashed line on modeled heatmaps, **Fig. 4.4*B***); additionally, this crossing needs to occur such that CsrA levels are above the stoichiometric point where CsrA and target mRNA levels are balanced (solid line on modeled heatmaps, **Fig. 4.4*B***). Our *in vivo* measurements using CsrB and CsrC confirm the model's prediction that (for a fixed maximum concentration of the ncRNA) there is an optimal level of CsrA for ncRNA signaling (expmt panels, **Fig. 4.4*B***) and thus that CsrA and ncRNA expression levels need to be well coordinated for signaling to function.

Known feedback regulation within the native CsrA network does not significantly alter the above result. Since it has been shown that CsrA levels can indirectly influence the concentration of its sequestering ncRNA partners (*e.g.* CsrA inhibits the production of CsrD which in turn promotes the degradation of CsrB and CsrC [14,51,64] we incorporated *csrD* and its repression by free CsrA into both our model and experimental system to test its effects (**Fig. 4.4*C***). CsrD and its inhibition by CsrA were modeled as described below and were incorporated into our *in vivo* system by leaving the native *csrD* gene intact in the chromosome. As before, *in vivo* measurements using CsrB and CsrC confirm the model's prediction (**Fig. 4.4*C***). Even in the presence of known feedback regulation, CsrA and ncRNA expression levels still need to be tuned appropriately for ncRNA to effectively signal.

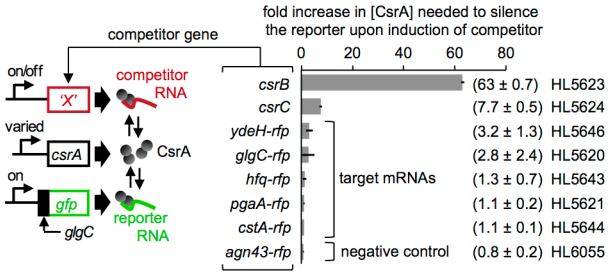

Fig. 4.5 | CsrB strongly outcompetes target mRNAs for CsrA. Error bars are SEM of duplicate measurements. The fold increase in [CsrA] required to silence *glgC-gfp* upon induction of a competing RNA is shown for several RNA competitors; we considered *glgC-gfp* 'silenced' when GFP fluorescence dropped to 10% of its original value upon CsrA induction. The concentration of CsrA required to silence *glgC-gfp* (both in the presence and absence of the competitor) was estimated by varying CsrA induction

as detailed in **Fig. 4.9** We confirmed the expression of each of the target mRNAs and the negative control as well as repression of the target mRNAs by CsrA via RFP fluorescence (**Fig. 4.6**).

4.3.6 CsrB Strongly Outcompetes Target mRNAs for CsrA

We have shown above that, like the ncRNAs CsrB and CsrC, target mRNAs can also act as a competitor for (and thus a means to sequester) CsrA (**Fig. 4.2**); we now quantitatively compare how effectively target mRNAs and sequestering ncRNAs compete with a common reporter mRNA for CsrA (**Fig. 4.5**). Since the concentration of CsrA strongly influences how much a competitor will affect target expression (**Fig. 4.4**), we quantified competition for CsrA *in vivo* by measuring the fold increase in [CsrA] required to silence the reporter mRNA (*glgC-gfp*) upon induction of the competitor gene (**Fig. 4.5**, **Fig. 4.9**). We considered the reporter mRNA to be 'silenced' when GFP fluorescence drops to $1/10^{th}$ of its initial value.

The results are striking: the ncRNAs (in particular CsrB) compete for CsrA much more effectively than do any of the competitor mRNAs. While CsrB required 63 ± 0.7 fold more CsrA to silence the reporter, *ydeH-rfp* (the strongest mRNA competitor tested) required only 3.2 ± 1.3 fold more CsrA. This difference can be understood in terms of several compounding advantages that increase the ability of CsrB (and to a lesser extent CsrC) to sequester CsrA. First, CsrB has a higher affinity for CsrA than do any of the measured target mRNAs (while CsrC has a higher affinity than most): CsrB and CsrC have k_d values for CsrA binding of ~0.5 nM and 8.7 nM respectively; [49] while target mRNAs have k_d values of 2.3 nM (*ydeH*) 64, 22 nM (*pgaABCD*) [57], 38 nM (*hfq*) [131], 39 nM (*glgCAP*) 53, and 40 nM (*cstA*) [132]. Second, CsrB and CsrC both have more binding sites for CsrA than do the target mRNAs: CsrB and CsrC can simultaneously bind ~18 and ~9 CsrA monomers respectively 49. In contrast, mRNAs that are targets for CsrA typically have between one and six known sites for CsrA monomers: *hfq* has one [131], *cstA* and *glgC* have four [53,132], and *pgaA* has six [57]. Third, CsrB and CsrC (in the absence of *csrD*) have longer half-lives than do typical mRNAs; without *csrD*, CsrB and CsrC both have measured half-lives longer than 30 min 51 while typical *E. coli* mRNAs have half-lives on the order of single minutes [133]. Additionally while the degradation rate of CsrB and CsrC are unaffected by CsrA binding [49,63], target mRNAs are more rapidly degraded when bound to CsrA (*e.g.* we estimate from [72] that *glgC* mRNA has a half-life of ~1 and ~2 min in the presence and absence of *csrA* respectively). By having longer half-lives, steady state CsrB and CsrC levels are higher than their competitor mRNA counterparts when produced at the same rate (*e.g.* by the same promoter).

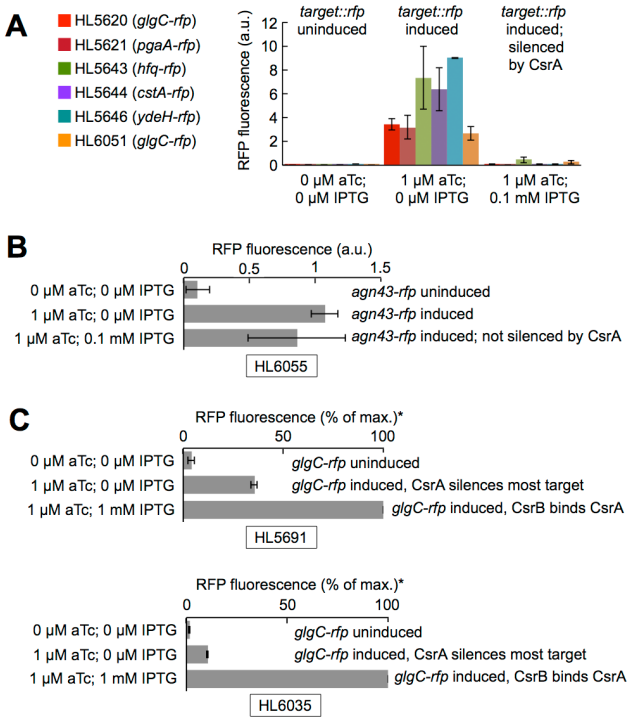

Fig. 4.6 | RFP fluorescence measurements of competitor mRNAs. Error bars are SEM of duplicate measurements. (*A*) Competitor mRNAs (with CsrA target sequence fused to *mCherry*) were both expressed upon their induction and silenced upon the induction of CsrA. (*B*) The negative control "competitor" mRNA (that contained the *mCherry* coding region but did not contain sequence for known CsrA binding sites) was expressed upon its induction but was not silenced by induction of CsrA. (*C*) The *glgC-rfp* competitor mRNA was partially silenced even upon its induction in both HL5691 and HL6035, consistent with constitutive expression of CsrA; this partial silencing was lifted upon induction of CsrB. *Here, the RFP fluorescence for each sample from the same overnight culture is reported relative to the RFP fluorescence of the +aTc, +IPTG sample; this was necessary to prevent differences between overnight cultures from obscuring the repeated pattern of relative fluorescence values within samples from the same overnight culture.

4.3.7 Competitor mRNA Expression and Regulation

Since each competitor mRNA used in this study is a translational fusion with the 5′-UTR of a native CsrA target gene and small segment of its coding sequence fused in frame to the coding region for *mCherry* (a red fluorescent protein), RFP fluorescence was used as a measure of translation of the competitor mRNA. In this way RFP fluorescence measurements were used (i) to confirm that the competitor mRNA genes induced during competition experiments (*e.g.* in **Fig. 4.5**) expressed competitor mRNAs upon induction and (ii) to confirm that those competitor mRNAs were susceptible to silencing by CsrA (**Fig. 4.6**). Each strain used in **Fig. 4.5**, **Fig. 4.9** and **Fig. 4.10** produced RFP fluorescence in the presence of competitor mRNA induction, but only if CsrA was not also induced (**Fig. 4.6*A***). In contrast, the strain (HL6055) with the negative control

'competitor' gene (*agn43-rfp*) produced RFP fluorescence in the presence of *agn43-rfp* induction, independently of whether CsrA was also induced (**Fig. 4.6*B***). For strains in which CsrA expression was constitutive (*e.g.* HL5691 and HL6035), RFP fluorescence increased marginally upon induction of the *glgC-rfp* competitor in the absence of CsrB induction, but much more significantly in the presence of CsrB (**Fig. 4.6*C***). All of these results are consistent with the true competitor mRNAs (*i.e.* all but *agn43-rfp*) being expressed upon induction and binding specifically to CsrA.

4.3.8 Topology of the Network of CsrA-Binding RNAs

The disparity between how effectively ncRNAs and target mRNAs sequester CsrA has important implications for the topology of the CsrA regulatory network. In the absence of a ncRNA, target mRNAs may influence one another's expression by competing with one another for CsrA; there is likely some hierarchy to this crosstalk (*e.g.* as has been observed in sRNA networks [47] but when compared to sequestering ncRNAs, mRNA targets for CsrA appear to be on a comparable footing. This leads to a highly interconnected network where nodes represent RNA partners for CsrA binding and directed edges represent the effect that each RNA can have on another's translation resulting from competition for CsrA (left network, **Fig. 4.1*B***). In contrast, when a strong sequestering ncRNA (*e.g.* CsrB) is incorporated into the network, that ncRNA will dominate competition for CsrA and the level of translation of the target mRNAs will be dictated much more strongly by expression of the ncRNA than by any of the mRNAs themselves. This leads to a strongly star-shaped network with a central hub RNA (*e.g.* CsrB) controlling the level of transcription of the subordinate target mRNAs (right network, **Fig. 4.1*B***). Strictly speaking, all of the original edges from the highly interconnected network are still present, but their magnitude is now dwarfed by regulation from CsrB; for this reason, the inter-mRNA edges from the network without a sequestering ncRNA (left network, **Fig. 4.1*B***) were not drawn in the network with the ncRNA present (right network, **Fig. 4.1*B***). By conceptualizing the CsrA regulatory network in this way, it is suggestive that ncRNAs such as CsrB may have the capacity to buffer crosstalk among mRNA targets. We examine this concept both theoretically and experimentally in the next section.

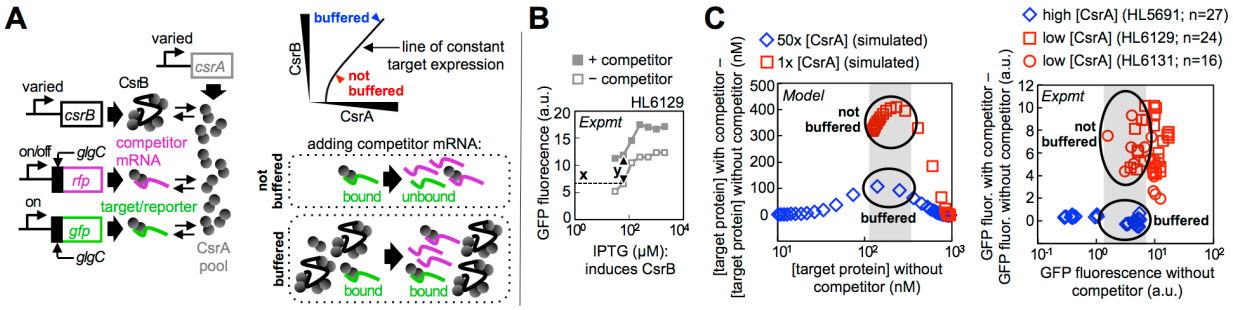

Fig. 4.7 | CsrB buffers crosstalk between target mRNAs. (*A*) Gene circuit schematic (left) and diagram explaining when crosstalk is buffered (right). Low [CsrA] and low [CsrB] (red triangle, "not buffered" scenario) allows competing mRNAs (purple) to rob the target mRNA (green) of CsrA; higher [CsrA] and higher [CsrB] (blue triangle, "buffered" scenario) dampen the impact of the competing mRNA on target expression since the competing mRNA can only sequester what is now a much smaller fraction of the total CsrA. (*B*, *C*) Simulation and *in vivo* measurement of mRNA-to-mRNA crosstalk and buffering. CsrB induction is varied (IPTG between 0 and 2 mM) in strains with fixed 'high' (HL5691) or 'low' (HL6129, HL6131) levels of CsrA expression in order to find conditions in which (i) high [CsrA] and high [CsrB] and (ii) low [CsrA] and low [CsrB] yield the same level of target expression (gray region: left and right plots of panel *C*). In all conditions, competitor mRNA is induced and the effect on the target expression is recorded. (*B*) Example of CsrB and competitor mRNA induction. This data is re-presented in panel *C* with additional data from **Fig. 4.8**. (*C*) In both simulated (left) and experimental (right) scatterplots, individual points represent the difference in target expression caused by induction of the competitor mRNA (y-axis; from panel *B*: distance between the black triangles) as a function of target expression in the absence of competitor (x-axis; from panel *B*: height of the black dashed line).

4.3.9 Buffering of Crosstalk Among CsrA Targets

To describe how CsrB can buffer crosstalk among mRNA targets of CsrA, we turn to an analogy with buffering acid. In this analogy, CsrA can be thought of as analogous to H^+ while CsrB can be thought of as analogous to buffering agent.

With solutions, the concentration of *free* H^+ (or H_3O^+) is often "buffered" by adding a molecule to that solution that will bind H^+. As more of this buffering molecule is added, larger changes in the *total* amount of H^+ are needed to appreciably change the concentration of *free* H^+ in solution (typically measured via pH). Buffering agents are important to diminish the effect that (i) perturbations in the *total* concentration of H^+ or (ii) perturbations in the concentrations of molecules that bind H^+ will have on the concentration of *free* H^+.

In the case of CsrA and CsrB, we can think of the concentration of *free* CsrA as being buffered by CsrB. As shown in **Fig. 4.3**, as more CsrB is added, larger changes in the total amount of CsrA are needed to appreciably change the concentration of free CsrA (approximately reported by GFP fluorescence). We now investigate whether CsrB can be used to decrease the effect that a competitor mRNA has on translation from a distinct CsrA target (much in the way a buffer of H^+ prevents impurities from altering pH).

Part 4: Crosstalk in the CsrA system

Specifically, we want to know whether we can achieve the same intermediate level of translation (and thus GFP fluorescence) from a reporter target mRNA in both a buffered and unbuffered state; this would be analogous to achieving the same pH in solutions exhibiting differing buffering capacities.

As can be seen in the induction spaces in **Fig. 4.4,** the same intermediate level of target expression can be achieved at low levels of induction CsrA and CsrB (**Fig. 4.7*A***, red triangle) and at high levels of induction CsrA and CsrB (**Fig. 4.7*A***, blue triangle). As explained by our analogy with pH buffering, we expect that with low concentrations of CsrA and CsrB the induction of an mRNA competitor would have a notable effect on whether a reporter mRNA would be bound by CsrA, and thus whether or not the reporter protein would be translated (**Fig. 4.7*A***, "not buffered"); in contrast, we would expect that with high concentrations of CsrA and CsrB the induction of an mRNA competitor would have much less of an effect on target mRNA occupancy and target protein translation (**Fig. 4.7*A***, "buffered").

To test these hypotheses, we constructed synthetic circuits as follows (**Fig. 4.7*A***). In each strain (HL5691, HL6129 and HL6131), *glgC-gfp* mRNA (the reporter) was transcribed constitutively; the promoter used was weaker than that used to transcribe the reporter mRNA in **Fig. 4.5** [14] so that the competitor mRNAs would have an increased advantage over the reporter when competing for CsrA. Also in each strain, CsrB and *glgC-rfp* mRNA (the competitor) were transcribed using inducible promoters (PLlacO-1 and PLtetO-1 respectively). CsrA was expressed constitutively in each strain, but using promoters with different strengths (PconNoHind, PconNoHindM8, and PconNoHindM10 in HL5691, HL6129 and HL6131 respectively with PconNoHind much stronger than the other two). Using each strain, CsrB was induced to varying degrees (using 0-2 mM IPTG) while the competitor mRNA was either uninduced or fully induced (using 0 or 1 μM aTc). Varying CsrB induction allowed us to identify conditions that with distinct buffering capacities for CsrA at the similar levels of CsrA activity (**Fig. 4.7*C***, gray regions).

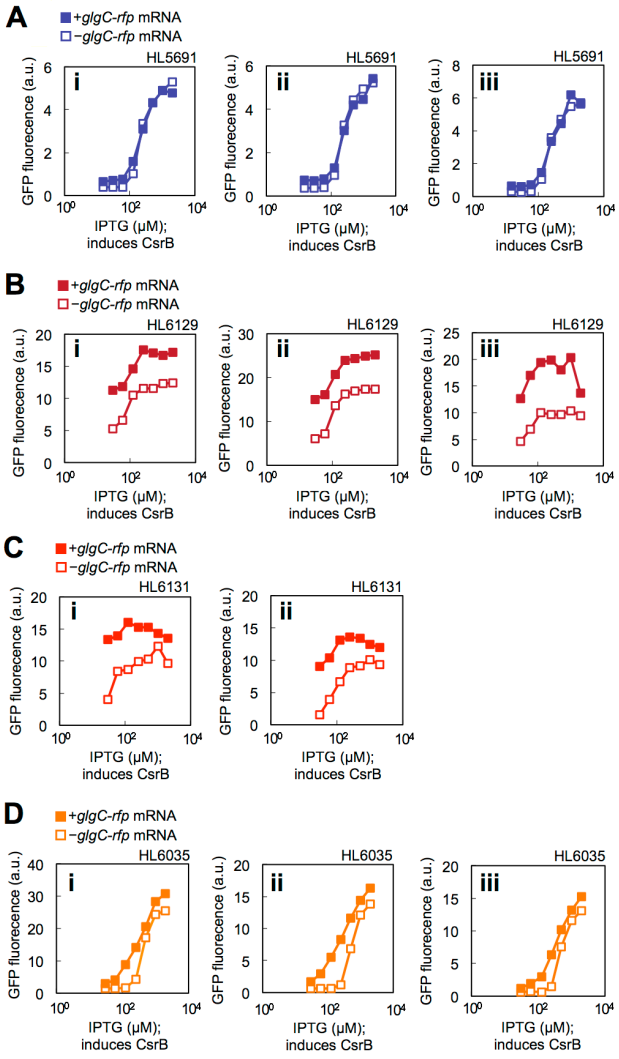

Fig. 4.8 | Raw data for crosstalk scatterplot. Data from experimental replicates are shown individually (*e.g.* as i, ii and iii) and thus plotted points are not shown with error bars. (*A*-*C*) This raw data is re-presented in **Fig. 4.7** to compare the effect of adding a competitor mRNA (*i.e. glgC-rfp*) at similar levels of target expression across different strains with different levels of CsrB induction. In each strain: *glgC-gfp* (the target mRNA) is expressed from PconNoHindM12; CsrA is expressed from a constitutive promoter, each of distinct strength; CsrB is expressed from PLlacO-1 and induced with IPTG as shown; *glgC-rfp* (the target mRNA) is expressed from PLtetO-1 and is either uninduced or induced with 0 or 1 μM aTc respectively. (*A*) Data from HL5691. (*B*) Data from HL6129. (*C*) Data from HL6131. (*D*) Data from HL6035. This raw data was not included in the **Fig. 4.7** scatterplot because HL6035 contains chromosomal rather than synthetic *csrA*; it has been reported that native *csrA* is autoregulated by the CsrA protein, which complicates the analysis of this data. We do observe, however, that crosstalk can occur in systems with chromosomal *csrA,* signifying that crosstalk is not abolished by *csrA* feedback. CsrB, *glgC-gfp* (the target mRNA) and *glgC-rfp* (the competitor mRNA) are expressed as described for *A*-*C*.

Part 4: Crosstalk in the CsrA system

To visually compare the buffering capacities of conditions that produce the same level of target expression, we presented our data using scatterplots (**Fig. 4.7*C***); sensitivity to expression of the competitor was plotted on the y-axis (*i.e.* "target expression with the competitor mRNA induced" – "target expression without the competitor mRNA induced"; shown as the distance between black triangles in **Fig. 4.7*B***) while the level of target expression (in the absence of the competitor mRNA) was plotted on the x-axis (height of black dashed line in **Fig. 4.7*B***). In the model scatterplot, each distinct point represents the comparison of two simulations (one with competitor expression and one without) at the same rate of CsrB production (**Fig. 4.7*C***, left); in the experimental scatterplot, each distinct point represents the comparison of two samples (one with the competitor induced and one without) taken from the same overnight culture and with the same level of CsrB induction (**Fig. 4.7*C***, right).

Our simulations and our *in vivo* experiments agree with our prediction from the analogy with acid buffering: induction of a competitor mRNA had a much greater effect on the reporter when CsrA was "not buffered" (*i.e.* [CsrA] and [CsrB] were low; red points in grey shaded region, **Fig. 4.7*C***) than when CsrA was "buffered" (*i.e.* [CsrA] and [CsrB] were high; blue points in grey shaded region, **Fig. 4.7*C***). Interestingly, GFP fluorescence from strains with low CsrA levels were less reproducible than from buffered strains when comparing replicates of cells with the same level of CsrB induction. This effect is itself consistent with the buffering analogy since we would expect a less buffered system to be more susceptible to stochastic and environmental fluctuations; the effect can be observed both in the scatterplot (**Fig. 4.7*C***, right: y-axis coordinates of blue versus red points) and from the induction curve data that was processed to produce them (**Fig. 4.8*A*** versus **Fig. 4.8*B***, ***C***).

4.4 DISCUSSION

In this study, we reverse engineered and quantitatively modeled the CsrA regulatory network to understand how crosstalk among target mRNAs can be prevented while still allowing for coordinated and graded regulation of those targets. We have shown that the degree of crosstalk between target mRNAs can be tuned by modulating the concentrations of CsrA and CsrB in tandem (**Fig. 4.7**) while the level of translation of those targets can be tuned in a coordinated way by varying the relative concentrations of CsrA and CsrB (**Figs. 4.3-4.10**). We note that even under buffered conditions modulating CsrB induction is still an effective way for continuously tuning target

expression (**Fig. 4.7**, **Fig. 4.8*A***) and that this is in part because CsrB is much more effective at sequestering CsrA than the other RNAs in the CsrA regulatory network (**Fig. 4.5**).

The CsrA system is not the only translational regulatory network facing crosstalk via competition for shared regulatory molecules. For example, in the past two decades the number of identified small regulatory RNAs (srRNAs) that target and regulate mRNA translation has increased dramatically with examples spanning all kingdoms of life [16,134]. Many of these srRNAs have multiple mRNA targets and/or share central regulatory complexes (*e.g.* Hfq, RISC) making them susceptible to the same forms of crosstalk as the CsrA system [16]; because the mechanisms of crosstalk are so similar, understanding how crosstalk can be buffered in the CsrA system has implications for other translational regulatory networks beyond CsrA homologs. Indeed, it has been proposed that within srRNA networks, fluctuations in translation from 'principle' target mRNAs may be buffered by a pool of 'auxiliary' target mRNAs which can bind to a shared central regulator but are not strongly regulated by it [135]. These two examples of mRNA buffering (*i.e.* in the CsrA and srRNA networks) may well represent a general principle of RNA buffering within translational regulatory networks, the regulatory consequences of which we are only beginning to explore.

In conclusion, this study shows that the CsrA regulatory network contains all of the components needed to implement a scalable strategy for preventing mRNA-to-mRNA crosstalk while allowing coordinated and graded regulation of those mRNA targets. The sequestering ncRNAs at the center of this buffering strategy ensure its scalability because the addition of new target mRNA 'nodes' to the network can always be countered by a coordinated increase in CsrA and CsrB expression. This study illuminates yet another example of how CsrB (and regulatory RNAs in general) can enhance the robustness of their surrounding regulatory systems [14]. Finally, since the CsrA system is a member of a larger class of translational regulatory networks present in all kingdoms of life, the principle of RNA buffering described here is of general importance for understanding how gene expression is controlled.

4.5 METHODS

4.5.1 Gene Expression Measurement and Analysis

4.5.1.1 Cell growth protocols

Steady state measurements of target and/or competitor protein expression were performed by inoculating 5-50 mL of overnight culture in 5 mL of LB media with 100 μg/mL ampicillin and/or 50 μg/mL kanamycin as needed to ensure plasmid retention. Cultures were grown for 3-4 h with shaking at 37 °C and 200 rpm, and then 0.5-5 mL of culture was inoculated into 3-5 mL of fresh LB with antibiotics and isopropyl β-D-1-thiogalactopyranoside (IPTG; 0.01 to 1 mM), anhydrotetracycline (aTc; 0.01 to 1 μM), both or neither. Collected cells were placed on ice. Single-cell GFP expression was measured using flow cytometry as previously described [21]. Single-cell RFP expression was measured using microscopy as described below.

For steady-state measurements of RFP fluorescence, single colonies were grown overnight in LB media with 100 μg/mL of ampicillin and/or 50 μg/mL of kanamycin as required for plasmid retention. Overnight cultures were diluted ~ 1/1000 in fresh LB media containing the same antibiotics as the overnight LB as well as IPTG and aTc as described for the specific experiment (**Fig. 4.6**). These newly diluted cultures were grown for ~ 3 to 4 hours at 37°C with shaking to early exponential growth (OD_{600} ~ 0.1); cells were then harvested and placed on ice. To prepare slides for microscopy, ~ 1 mL of LB per sample was centrifuged at 16,100 × g for 1 min to concentrate the cells, these concentrated cells were then resuspended in ~ 4 μL of LB media; the 4 μL of media was then applied to the slide and secured beneath a glass coverslip.

4.5.1.2 Microscopy protocols

Microscopy was performed using a TE2000E microscope (Nikon) with an X-cite 120PC lamp (Exfo) and a 100× objective with phase 3 contrast. The excitation filter, dichroic mirror and emission filter used for microscopy measurements were 575 ± 25 nm, 610 nm, and 640 ± 25 nm respectively. Images were captured using a Pixus 1024×1024 pixel CCD camera (Princeton Instruments) in conjunction with Metamorph 7.0 (Molecular Devices). Metamorph was also used to identify cells within the images and log their average RFP fluorescence. RFP fluorescence was recorded from ~ 30 to 300 cells per sample; each experimental condition was measured in duplicate (*i.e.* two independent samples from distinct overnight cultures were measured per condition). The average

fluorescence for each sample was determined by first averaging the fluorescence of cells within that sample and subsequently subtracting background fluorescence.

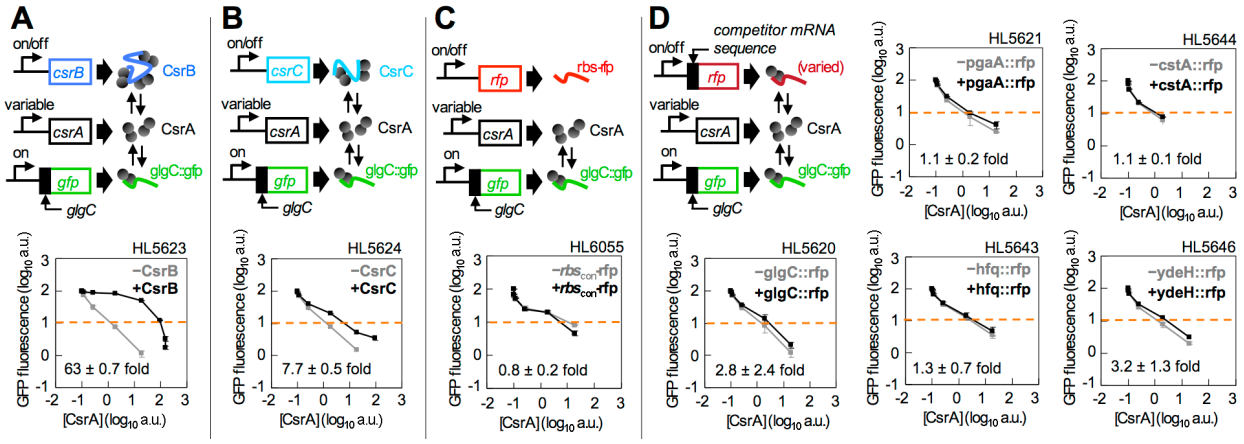

Fig. 4.9 | CsrB strongly outcompetes target mRNAs for CsrA. Experiments to determine the fold increase in CsrA required to silence a target mRNA upon the addition of a competitor RNA (as presented in **Fig. 4.5**). Error bars are SEM of duplicate measurements. In each scenario GFP fluorescence from the target mRNA reporter is presented as a function of CsrA concentration in the presence and absence of an mRNA competitor. CsrA was expressed from PLlac:st7:*csrA* which was induced using between 0 and 1 mM IPTG. The x-axis of each plot is calibrated using induction curve data from PLlacO-1:st7:*gfp* (HL4510). Each competitor was expressed from PLtetO-1 and was either uninduced (0 aTc, gray curves) or induced (1 μM aTc, black curves). (*A*) The ncRNA CsrB strongly competed with *glgC-gfp* mRNA for CsrA. (*B*) The ncRNA CsrC moderately competed with *glgC-gfp* mRNA for CsrA. (*C*) Negative control; the *rfp* mRNA did not compete with the *glgC-gfp* mRNA. (*D*) The five examined target mRNAs (*i.e.* CsrA target sequences translationally fused to *rfp*) competed only marginally with *glgC-gfp* mRNA for CsrA. RFP fluorescence measurements confirm the expression of these mRNA competitors and their repression by CsrA (**Fig. 4.6**).

4.5.1.3 Quantifying competition for CsrA

Because we have shown that the sequestration of CsrA is most observable over a window of CsrA concentrations (**Fig. 4.3**), we measured competition for CsrA *in vivo* over a range of CsrA expression levels. In each case the reporter mRNA, CsrA, and the competitor RNAs are expressed using PconNoHindM2 (a constitutive promoter), PLlacO-1 (induced with between 0 and 200 μM IPTG) and PLtetO-1 (induced with 0 or 1 μM aTc) respectively. For each experiment, the relative intracellular concentration of CsrA was estimated from the amount of IPTG added using the induction curve for PLlacO-1:st7:*gfp* as described previously [14]. For the purposes of quantifying competition, we consider the reporter mRNA to be 'silenced' when GFP fluorescence drops to $1/10^{th}$ of its initial value (*i.e.* $1/10^{th}$ of the fluorescence measurement taken at the lowest level of CsrA achieved, which occurs at 0 μM IPTG). Using this definition, we quantified competition by determining the fold-increase in CsrA required to silence the reporter mRNA in the presence versus in the absence of the RNA competitor (orange line intersecting the black and grey curves respectively, **Fig. 4.9**).

Part 4: Crosstalk in the CsrA system

When the ncRNAs CsrB (**Fig. 4.9*A***) and CsrC (**Fig. 4.9*B***) were induced, substantially more CsrA was required to silence the reporter mRNA (63 ± 0.7 fold and 7.7 ± 0.5 fold respectively) than when the ncRNAs were not transcribed. To show that these numbers specifically reflect competition for CsrA among RNAs with well-defined CsrA binding sites (and would arise upon induction of a generic transcript), we performed a control experiment with the competitor RNA replaced with *rfp* mRNA (**Fig. 4.9*C***). For this negative control the 5′ untranslated region (UTR) of the *rfp* mRNA was taken from the *agn43* gene that does not possess any known CsrA binding sites. As expected this non-competitor did not increase the amount of CsrA required for silencing of the reporter (the amount of CsrA needed 'increased' by 0.8 ± 0.2 fold).

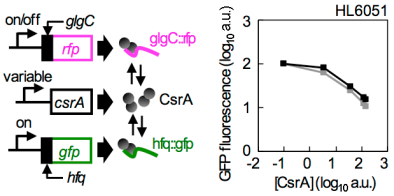

Fig. 4.10 | Alternate target mRNA: *hfq-gfp*. Error bars are SEM of two or more measurements. CsrA is expressed from PLlac:st7:*csrA* which is induced using between 0 and 1 mM IPTG. The x-axis of each plot is calibrated using induction curve data from PLlacO-1:st7:*gfp* (HL4510). The *glgC-rfp* competitor was expressed from PLtetO-1 and was either uninduced (0 aTc, gray curve) or induced (1 μM aTc, black curve). The *glgC-rfp* mRNA competed only marginally with *hfq-gfp* mRNA for CsrA. RFP fluorescence measurements confirm the expression of the *glgC-rfp* mRNA competitor and its repression by CsrA in this strain (**Fig. 4.6*A***).

Having quantified how effectively ncRNAs compete with the *glgC-gfp* reporter mRNA *in vivo*, we now use this same synthetic system to determine the strength of competition from native target mRNA leader sequences (**Fig. 4.9*D***). Since the binding sites for CsrA are largely located in the 5′ UTR of target mRNAs (typically overlapping the ribosomal binding site), we fused parts of the 5′ UTR of known CsrA targets (parts known to include CsrA binding sites) to *rfp* to create several competitor mRNAs. These competitor mRNAs include *glgC-rfp*, *pgaA-rfp*, *hfq-rfp*, *cstA-rfp* and *ydeH-rfp* and their construction is detailed below. When induced, these competitor mRNAs (**Fig. 4.9*D***) increase the amount of CsrA required for silencing only marginally more than the negative control (**Fig. 4.9*C***) and much less than the ncRNAs (**Fig. 4.9*A, B***). To confirm that the apparent weak competition for CsrA on the part of the target mRNAs was not an aberrant result arising from *glgC-gfp* (the reporter mRNA) being a particularly strong mRNA competitor, we performed an additional control with *hfq-gfp* as the reporter mRNA and *glgC-rfp* as the competitor (**Fig. 4.10**). We selected *hfq-gfp* as a likely weaker

competitor than *glgC-gfp* because the *hfq* 5′ UTR only has one known binding site for CsrA while the *glgC* 5′ UTR has four; additionally, the *hfq* and *glgC* mRNAs have comparable affinity for CsrA [53,131]. While *hfq-gfp* was not completely silenced by the addition of CsrA (*i.e.* the reporter only dropped to about $1/8^{th}$ of its initial fluorescence) it is clear from inspection that *glgC-rfp* does not significantly rob *hfq-gfp* of CsrA (**Fig. 4.10**). Note: for each of the strains using *rfp* or an *rfp* fusion mRNA competitor, RFP fluorescence measurements were recorded to ensure that the competitor mRNA was both expressed upon its induction, and repressed by induction of CsrA (**Fig. 4.6*A, B***).

4.5.2 Detailed Model Description

4.5.2.1 Model overview

To describe the CsrA regulatory network, we constructed a model with ordinary differential equations (ODEs) for each of the following species: the target protein ("GFP"), the target mRNA ("m_{GFP}"), the competitor protein ("RFP"), the competitor RNA ("m_{RFP}"), the CsrA dimer ("A"), the CsrB non-coding RNA ("B"), the CsrA-target mRNA complex ("Am_{GFP}"), the CsrA-competitor mRNA complex ("Am_{RFP}") and the CsrA-CsrB complex ("AB"). This model extends our previous model of the CsrA cascade [14] by incorporating additional competitor mRNAs and proteins that were not considered previously. As with our previous model, "B" represents a pair of CsrA binding sites within the CsrB molecule that can be bound by a single CsrA dimer; in this way the CsrA-CsrB complex "AB" corresponds to a CsrA dimer that occupies one pair of CsrA binding sites. Parameter values (discussed below for each specific simulation) are identical to those used in our previous study except where otherwise stated [14].

We now describe the equations and assumptions for the core model of the CsrA network used in this study. In our *in vivo* experiments, the green fluorescent protein GFP was used as the target protein and the red fluorescent protein mCherry (RFP) was used as the competitor protein. In or model, the target protein concentration ([GFP]) is determined by its production rate (the free target mRNA concentration ([m_{GFP}]) multiplied by the rate constant α_G) and its degradation rate ([GFP] multiplied by the rate constant for passive dilution (β_{dil})).

$$\frac{d[\mathrm{GFP}]}{dt} = \alpha_{\mathrm{G}}[\mathrm{m_{GFP}}] - \beta_{\mathrm{dil}}[\mathrm{GFP}] \qquad [4.1]$$

Similarly, the competitor protein concentration ([RFP]) is determined by its production rate (the free competitor mRNA concentration ([m_{RFP}]) multiplied by the rate constant α_R) and its degradation rate ([RFP] multiplied by the rate constant for passive dilution (β_{dil})).

$$\frac{d[\mathrm{RFP}]}{dt} = \alpha_{\mathrm{R}}[\mathrm{m_{RFP}}] - \beta_{\mathrm{dil}}[\mathrm{RFP}] \qquad [4.2]$$

The target mRNA concentration ([m_{GFP}]) is determined by its production rate (α_{mGFP}, which depends on how frequently it is transcribed) and its removal rate (determined by passive dilution, active degradation, and silencing by CsrA). The rates of active degradation and passive dilution are equal to [m_{GFP}] multiplied by the rate constants β_m and β_{dil} respectively. The rate of removal of free target mRNA by formation of the CsrA-target complex (Am_{GFP}) is equal to the product of the target mRNA concentration, the CsrA dimer concentration and the association rate constant (k_1). The rate of release of free target mRNA from the CsrA-target complex is equal to the product of the concentration of the complex ([Am_{GFP}]) and the dissociation rate constant (k_{-1}).

For simplicity, we initially assume that the target and competitor leader sequences (within the target and competitor mRNAs) bind to one CsrA dimer each and that bound CsrA completely silences translation. The appropriateness of this simplifying assumption was validated experimentally both in this study and our previous work [14]. We note that, theoretically, free target mRNA could be generated by active degradation of CsrA from within the CsrA-target complex; however, since we know that the active degradation of CsrA is negligible, we omit reaction this from our model equations.

$$\frac{d[\mathrm{m_{GFP}}]}{dt} = \alpha_{\mathrm{mGFP}} - (\beta_{\mathrm{m}} + \beta_{\mathrm{dil}})[\mathrm{m_{GFP}}] - k_1[\mathrm{A}][\mathrm{m_{GFP}}] + k_{-1}[\mathrm{Am_{GFP}}] \qquad [4.3]$$

The competitor mRNA concentration ([m_{RFP}]) is determined by the same processes as the target mRNA concentration. The rate constants governing the association and dissociation of CsrA dimers from the competitor mRNA are treated as identical to those of the target mRNA (*i.e.* k_1 and k_{-1}) for the purposes of this model; this is reasonable since in the simulated competition experiment (**Fig. 4.7**), the 5′ UTR sequence which contains the CsrA binding sites is the same on the target mRNA (*glgC-gfp*) and the competitor mRNA (*glgC-rfp)*. The rate constant for the active degradation of the

competitor mRNA is assumed to be the same as the target, because most mRNAs have similar half-lives (on the order of single minutes) [133] and because the target and competitor mRNAs (*glgC-gfp* and *glgC-rfp* respectively) have very similar sequences. a_{mRFP} is the production rate of the competitor mRNA. The concentration of the CsrA-competitor complex is $[Am_{RFP}]$.

$$\frac{d[\mathrm{m_{RFP}}]}{dt} = \alpha_{\mathrm{mRFP}} - (\beta_{\mathrm{m}} + \beta_{\mathrm{dil}})[\mathrm{m_{RFP}}] - k_1[\mathrm{A}][\mathrm{m_{RFP}}] + k_{-1}[\mathrm{Am_{RFP}}] \qquad [4.4]$$

The concentration of CsrA-target complex ($[Am_{GFP}]$) is determined by (i) the association and dissociation of the target mRNA with CsrA dimers and (ii) the clearance of the complex, either by passive dilution (rate constant β_{dil}) or by the active degradation of the target mRNA while it is bound to CsrA the complex (rate constant β_{mAm}).

$$\frac{d[\mathrm{Am_{GFP}}]}{dt} = k_1[\mathrm{A}][\mathrm{m_{GFP}}] - (k_{-1} + \beta_{\mathrm{mAm}} + \beta_{\mathrm{dil}})[\mathrm{Am_{GFP}}] \qquad [4.5]$$

The same processes that determine the concentration of the CsrA-target complex, determine the concentration of the CsrA-competitor complex ($[Am_{RFP}]$). The rate constant for active degradation of competitor mRNA that is bound to CsrA is assumed to be the same as the target mRNA bound to CsrA (β_{mAm}); again, this is a reasonable assumption because the mRNA sequences are highly similar and thus CsrA should affect the degradation of the two mRNAs in similar ways.

$$\frac{d[\mathrm{Am_{RFP}}]}{dt} = k_1[\mathrm{A}][\mathrm{m_{RFP}}] - (k_{-1} + \beta_{\mathrm{mAm}} + \beta_{\mathrm{dil}})[\mathrm{Am_{RFP}}] \qquad [4.6]$$

The concentration of free CsrB ([B]) is determined by its production rate (α_B) and its removal rate (determined by passive dilution, active degradation, and sequestration into the CsrA-CsrB complex). To keep the model simple, [B] represents the concentration of CsrB binding sites rather than the concentration of complete CsrB RNA molecules [14]. In other words, the model quantifies the concentration of CsrB in terms of the number of CsrA dimers that it can bind to: one unit of [B] can bind one unit of [A]. We note that while this simplification requires the assumption that each CsrA dimer binds to the full length CsrB molecule independently, we have shown previously that this independent binding model is sufficient to describe the dynamic behavior of the

Part 4: Crosstalk in the CsrA system

CsrA cascade [14]. The rate of passive dilution of free CsrB equals [B] times the rate constant for passive dilution (β_{dil}). The rate of active degradation of free CsrB (in the absence of the CsrD protein) equals [B] times the rate constant β_B. Formation of the CsrA-CsrB complex occurs at a rate equal to the product of [A], [B] and the association rate constant for the CsrA-CsrB complex (k_2). CsrB dissociates from the CsrA-CsrB complex at a rate equal to the product of concentration of the complex ([AB]) and the dissociation constant k_{-2}.

$$\frac{d[\mathrm{B}]}{dt} = \beta_{\mathrm{B}} - (\beta_{\mathrm{B}} + \beta_{\mathrm{dil}})[\mathrm{B}] - k_2[\mathrm{A}][\mathrm{B}] + k_{-2}[\mathrm{AB}] \qquad [4.7]$$

The concentration of the CsrA-CsrB complex is determined by its rate of formation (described above), dissociation (described above), by the rate of active degradation of CsrB while CsrB is bound to CsrA (equal to the concentration of the complex [AB] multiplied by the rate constant β_{BAB}) and by the rate of passive clearance of the complex (equal to [AB] multiplied by the rate constant β_{dil}).

$$\frac{d[\mathrm{AB}]}{dt} = k_2[\mathrm{A}][\mathrm{B}] - (k_{-2} + \beta_{\mathrm{BAB}} + \beta_{\mathrm{dil}})[\mathrm{AB}] \qquad [4.8]$$

The concentration of free CsrA dimer ([A]) is determined by the production rate of the dimer (α_A) by its clearance (by passive dilution), and by CsrA binding to and dissociation from (i) the target mRNA, (ii) the competitor mRNA and (iii) CsrB. To maintain the simplicity of the model CsrA dimers are generated in a single reaction step as described previously [14]. The rate of passive dilution of free CsrA dimers is equal to the product of [A] and the rate constant β_{dil}. Free CsrA dimers form and dissociate from the CsrA-target or CsrA-competitor complexes as described above.

$$\begin{aligned}\frac{d[\mathrm{A}]}{dt} = {} & \alpha_{\mathrm{A}} - [\mathrm{A}](\beta_{\mathrm{dil}} + k_1[\mathrm{m_{GFP}}] + k_1[\mathrm{m_{RFP}}] + k_2[\mathrm{B}]) + \\ & (k_{-1} + \beta_{\mathrm{mAm}})([\mathrm{Am_{GFP}}] + [\mathrm{Am_{RFP}}]) + (k_{-2} + \beta_{\mathrm{BAB}})[\mathrm{AB}]\end{aligned} \qquad [4.9]$$

4.5.2.2 Simulation of dose/response relationships

In **Figs. 4.2 & 4.3**, we simulated the dose/response relationships between a target mRNA, CsrA and CsrB. To perform these simulations, we used the model ODEs described above.

In **Fig. 4.2**, the genetic circuit contained only a CsrA target gene (*glgC-gfp*) and *csrA*. Since *csrB* and the competitor gene were not present in the system, the rate of production of CsrB and the rate of production of the competitor mRNA (α_B and α_{mRFP}, respectively) were set to zero. Since target mRNA transcription was induced over a range of levels in our experiments, the rate of target mRNA transcription (α_{mGFP}) was assigned a range of values between 0 and 5 $nM \cdot s^{-1}$ in our simulations. Since CsrA production was induced to three distinct levels in our experiments, the rate of CsrA dimer production (α_A) was given three distinct values (0, $2 \cdot 10^{-2}$ and $4 \cdot 10^{-2}$ $nM \cdot s^{-1}$) in our simulations. The production rates for the target mRNA and CsrA (α_{mGFP} and α_A) were manually selected to fit the data; all other parameters (including rate constants for binding, unbinding, degradation and dilution) had values as previously described for the two-step cascade [14]. For each simulation (*i.e.* for each combination of parameters), the steady levels of all molecular species were obtained by solving numerically for the dynamic behavior of the system and allowing that simulation to converge as time increased. These simulations were performed by integrating the model's ODEs using the ode15s solver in MATLAB. The initial conditions for all molecular species were set to zero.

In **Fig. 4.3**, the genetic circuit contained a CsrA target gene (*glgC-gfp*), *csrA* and *csrB*. Since the competitor gene was not present in the system, the rate of production of the competitor mRNA (α_{mRFP}) was set to zero. Since target mRNA transcription was not varied in these experiments, the rate of target mRNA transcription (α_{mGFP}) was fixed at $6 \cdot 10^{-2}$ $nM \cdot s^{-1}$ in our simulations as per our prior model [14]. Since CsrA production was induced over a range of levels in our experiments, the rate of CsrA production was assigned a range of values between 0 and 10^{-4} $nM \cdot s^{-1}$ in our simulations. Since CsrB production was induced to four distinct levels in our experiments, the rate of CsrB production (α_B) was given four distinct values (0, $8 \cdot 10^{-3}$, $4 \cdot 10^{-2}$ and $2 \cdot 10^{-1}$ $nM \cdot s^{-1}$) in our simulations. The production rates for CsrA and CsrB were manually selected to fit the data; as for **Fig. 4.2**, all parameters that were not mentioned (including rate constants for binding, unbinding, degradation and dilution) had values as previously described for the two-step cascade [14].

Part 4: Crosstalk in the CsrA system

In **Fig. 4.4*B***, the genetic circuit contains the same components as in **Fig. 4.3**. Without the competitor mRNA gene the rate of production of the competitor mRNA (α_{mRFP}) remained zero. Target mRNA transcription was not varied in these experiments (just as in **Fig. 4.3**) so the rate of target mRNA transcription (α_{mGFP}) was maintained at $6{\cdot}10^{-2}$ $nM{\cdot}s^{-1}$ as above. Since the rates of production of both CsrA and the ncRNA (*i.e.* CsrB or CsrC) were induced over a range of levels in our experiments, the rates of production of both CsrA and the ncRNA were each assigned a range of values between 10^{-4} and 10^{1} $nM{\cdot}s^{-1}$. All parameters that were not mentioned (including rate constants for binding, unbinding, degradation and dilution) had values as previously described for the two-step cascade [14]. Since CsrA and ncRNA levels were varied over a wide range in both our *in vivo* experiments and our simulations, no fitting of parameters was needed to demonstrate agreement between the experimental and simulated results; our model and our experiments (with CsrB and CsrC) both clearly shows that induction of a sequestering RNA has the strongest effect on target mRNA translation at intermediate levels of CsrA expression. We did not distinguish between CsrB and CsrC in our simulation, because the effectiveness of sequestration by CsrC shows the same multiphasic (*i.e.* non-monotonic) behavior as a function of CsrA expression that CsrB does. In this example CsrC operates simply as a less potent version of CsrB.

In **Fig. 4.4*C***, the genetic circuit contains the same components as in **Fig. 4.4*B*** with the addition of chromosomal *csrD*. The *csrD* gene expresses the CsrD protein, which binds to CsrB and CsrC and promotes their active degradation by RNAse E [51]. Additionally, it has been shown that expression of CsrD (from chromosomal *csrD*) is inhibited by free CsrA [14,51,64]. This regulation creates a chain of interactions, the net result of which is a negative feedback loop: CsrA inhibits CsrD expression, CsrD promotes CsrB clearance, and CsrB inactivates CsrA (**Fig. 4.4*C***). As in **Fig. 4.4*B***, we model the system using only the parameters for CsrB since the regulation via CsrB and CsrC showed the same qualitative behavior *in vivo*. To incorporate CsrD into our model, two additional equations are required to keep track of free CsrD and the CsrB-CsrD complex; these equations model CsrD in the same way as our prior study [14]. Free CsrD is produced at a maximum rate described by a_D, but that rate is attenuated by the presence of free CsrA; we account for this inhibition by multiplying the production rate constant for CsrD (α_D) by a hill-type function that has a range from 0 to 1 and reaches its half-maximum when the concentration of free CsrA ([A]) equals the feedback regulation constant k_f. Free CsrD is binds to CsrB at a rate equal to the product of the concentration of free CsrB ([B]), the concentration of free CsrD ([D]), and the rate constant k_{ES}. Free

Part 4: Crosstalk in the CsrA system

CsrD is released from the CsrB-CsrD complex in two ways: by dissociation of CsrB (which occurs at a rate equal to the product of the concentration of the CsrB-CsrD complex ([BD]) and the rate constant k_{-ES}), and by the degradation of CsrB by RNAse E (which occurs at a rate equal to the product of [BD] and the rate constant k_P. CsrD is cleared passively by dilution at a rate equal to [D] multiplied by the rate constant β_{dil}.

$$\frac{d[\mathrm{D}]}{dt} = \alpha_{\mathrm{D}} \left(\frac{k_{\mathrm{f}}}{k_{\mathrm{f}} + [\mathrm{A}]} \right) - \beta_{\mathrm{dil}}[\mathrm{D}] - k_{\mathrm{ES}}[\mathrm{B}][\mathrm{D}] + k_{\text{-ES}}[\mathrm{BD}] \qquad [4.10]$$

The concentration of the CsrB-CsrD complex ([BD]) increases when CsrB binds to CsrD, and decreases by the dissociation of CsrB from CsrD and by the degradation of CsrB by RNAse E as described above. The CsrB-CsrD complex can also be cleared by dilution at a rate equal to [BD] multiplied by the rate constant β_{dil}.

$$\frac{d[\mathrm{BD}]}{dt} = k_{\mathrm{ES}}[\mathrm{B}][\mathrm{D}] - (k_{\text{-ES}} + k_{\mathrm{P}} + \beta_{\mathrm{dil}})[\mathrm{BD}] \qquad [4.11]$$

In addition to incorporating **Eqs. 4.10 & 4.11** above into our model of the CsrA network, **Eq. 4.7** also must be modified to include terms for free CsrB binding to CsrD and CsrB dissociation from the CsrB-CsrD complex. **Eq. 4.7′** below includes these modifications.

$$\frac{d[\mathrm{B}]}{dt} = \beta_{\mathrm{B}} - (\beta_{\mathrm{B}} + \beta_{\mathrm{dil}})[\mathrm{B}] - k_2[\mathrm{A}][\mathrm{B}] + k_{-2}[\mathrm{AB}] - k_{\mathrm{ES}}[\mathrm{B}][\mathrm{D}] - k_{\text{-ES}}[\mathrm{BD}] \qquad [4.7']$$

In summary, the model of the CsrA network presented in **Fig. 4.4*C*** was simulated using **Eqs. 4.1-4.6**, **4.7′** (rather than **4.7**), and **4.8-4.11**. The rates of production of the target mRNA and the competitor mRNA are identical to **Fig. 4.4*B***; all parameters that were not mentioned (including rate constants for binding, unbinding, degradation and dilution) had values as previously described for the three-step cascade with feedback [14].

4.5.2.3 Simulation of mRNA-to-mRNA crosstalk and buffering

Crosstalk and buffering in the CsrA network were simulated in **Fig. 4.7*C***. In the corresponding experiments, the genetic circuit contained both a CsrA target gene (*glgC-gfp*) and a CsrA competitor gene (*glgC-rfp*) in addition to *csrA* and *csrB*. Note: since *csrD* was not present in this system, crosstalk and buffering was modeled using **Eqs. 4.1**

through **4.9** (and not **Eqs. 4.7', 4.10 or 4.11**). Since CsrA production was maintained at high or low levels of expression constitutively in different strains, we divided our simulations into high and low [CsrA] scenarios accordingly. In the 1x [CsrA] scenario (red points, **Fig. 4.7*C***) and the 50x [CsrA] scenario (blue points, **Fig. 4.7*C***), the production rate of CsrA (α_A) was assigned values of $4{\cdot}10^{-2}$ nM·s^{-1} and 2 nM·s^{-1} respectively. Since the rate of production of CsrB was induced over a range of levels in each strain, the rate of production CsrB was assigned a range of values between 10^{-4} and 10^{6} nM·s^{-1} in both the 1x [CsrA] and 50x [CsrA] scenarios. We do not expect the production rate of CsrB in the *in vivo* experiment to have quite this dynamic range, but by simulating this larger window of CsrB production rates the simulations provide more context in which to understand the experimental results. Target mRNA transcription was not varied in these experiments so the rate of target mRNA transcription (α_{mGFP}) was maintained at $1.2{\cdot}10^{-1}$ nM·s^{-1}. The rate of competitor mRNA transcription (α_{mRFP}) was set to either 0 or 3 nM·s^{-1}. The production rates for the target mRNA, competitor mRNA and CsrA (α_{mGFP}, α_{mRFP} and α_A) were manually selected to fit the data; all other parameters (including rate constants for binding, unbinding, degradation and dilution) had values as previously described for the two-step cascade [14].

4.5.3 Bacterial Strains and Plasmids

Strain and plasmid information is provided in **Table 4.1**. Plasmid maps for plasmids first presented here are shown in **Fig. 4.11**. Oligonucleotide sequences are provided in **Table 4.2**. The nucleotides taken from the *E. coli* chromosome to construct each synthetic target and competitor mRNA sequence are as follows: the *glgC-gfp* and *glgC-rfp* mRNAs both contained nucleotides -61 to +8 relative to the *glgC* start codon; the *hfq-gfp* and *hfq-rfp* mRNAs contained nucleotides -71 to +56 relative to the *hfq* start codon; the *cstA-rfp* mRNA contained nucleotides -40 to +25 relative to the *cstA* start codon; the *pgaA-rfp* mRNA contained nucleotides -234 to +23 relative to the *pgaA* start codon; the *ydeH-rfp* mRNA contained nucleotides -37 to +22 relative to the *ydeH* start codon; the *agn43-rfp* negative control 'competitor' contained nucleotides -187 to +49 relative to the *agn43* start codon. The PconNoHind, PconNoHindM2 and PconNoHindM12 promoters were described previously 14 and are variants of Pcon/O3 [55]. The PconNoHindM8 promoter is identical to PconNoHindM12 except that seven nucleotides have been removed between -10 site and the ribosomal binding site. The PconNoHindM10 promoter is identical to PconNoHind except for the removal of a nucleotide between the -10 and -35 sites. The PLlacO-1 promoter, PLtetO-1 promoter and T1 terminator

sequences were PCR amplified from pZ system plasmids [54]. The ribosomal binding sequences st2, st3 and st7 [83] were synthesized by Integrated DNA Technologies. The *gfp* gene and the T1T2 terminator sequence were PCR amplified from the pTAK102 plasmid [54]. The *mCherry* gene was PCR amplified from a plasmid provided by R. Tsien (University of California, San Diego, CA) [84]. The Asp terminator sequence was PCR amplified from the pLex plasmid (Invitrogen). Sources for the *tetR* and *lacIq* gene sequences were reported previously [21,22]. Sequences for the *csrA*, *csrB*, *csrC*, *csrD* and *cstA* genes as well as the 5'UTR and portions of the *cstA*, *glgC*, *hfq*, *pgaA*, *ycdT* and *ydeH* genes were PCR amplified from the chromosome of *E. coli* MG1655 (Yale E. coli Stock Center, CGSC #7740).

Table 4.1 | Plasmids and strains

*ID**	*Description*
pHL600†	PLtetO-1:*csrB*, PLlacO-1:RBS(st7):*csrA*
pHL601	PLtetO-1:*csrC*, PLlacO-1:RBS(st7):*csrA*
pHL662†	PLtetO-1:RBS(st7):*mCherry*, PLlacO-1:RBS(st7):*gfp*
pHL1318†	PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st2):*csrA*
pHL1335†	PconNoHindM12:*glgC*::*gfp*, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st2):*csrA*
pHL1355†	PconNoHindM12:*glgC*::*gfp*, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st3):*csrA*
pHL1488	PconNoHindM10:*glgC*::*gfp*, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st2):*csrA*
pHL1529†	PLlacO-1:*glgC*::*gfp*, PconNoHind:RBS(st3):*tetR*, PconNoHind:RBS(st2):*csrA*
pHL1590	PLtetO-1:RBS(st3):*csrA*
pHL1722	PLtetO-1:*glgClead*::*mCherry*, PLlacO-1:RBS(st7):*csrA*
pHL1724	PLtetO-1:*pgaAleadFull*::*mCherry*, PLlacO-1:RBS(st7):*csrA*
pHL1726	PLtetO-1:*hfqlead*::*mCherry*, PLlacO-1:RBS(st7):*csrA*
pHL1745	PLtetO-1:*cstAlead*::*mCherry*, PLlacO-1:RBS(st7):*csrA*
pHL1747	PLtetO-1:*ydeHlead*::*mCherry*, PLlacO-1:RBS(st7):*csrA*
pHL1756†	PconNoHindM12:*glgC*::*gfp*, PconNoHind:RBS(st3):*tetR*
pHL1757†	PconNoHindM2:*glgC*::*gfp*, PconNoHind:RBS(st3):*tetR*
pHL1790	PLtetO-1:*glgClead*::*mCherry*, PLlacO-1:*csrB*
pHL1914	PconNoHindM2:*hfqlead*::*gfp*, PconNoHind:RBS(st3):*tetR*
pHL1928	PLtetO-1:*Agn43RBS:mCherry*, PLlacO-1:RBS(st7):*csrA*
pHL1940	PLtetO-1:*cstA*‡, PLlacO-1:RBS(st7):*csrA*
pHL1947	PconNoHindM12:*glgC*::*gfp*, PconNoHind:RBS(st3):*tetR*, PconNoHindM8:RBS(st3):*csrA*
pHL1948	PconNoHindM12:*glgC*::*gfp*, PconNoHind:RBS(st3):*tetR*, PconNoHindM10:RBS(st3):*csrA*
HL3796†	HL716§ + *ΔcsrA* + *ΔcsrB* + *ΔcsrC* + *ΔglgCAP* + *ΔpgaABCD*
HL3986	HL3796 + pHL600 + pHL1335
HL4018†	HL716§ + *ΔcsrB* + *ΔcsrC* + *ΔcsrD* + *ΔglgCAP* + *ΔpgaABCD*
HL4022	HL3796 + pHL601 + pHL1335
HL4142†	HL716§ + *ΔcsrA* + *ΔcsrB* + *ΔcsrC* + *ΔcsrD* + *ΔglgCAP* + *ΔpgaABCD*
HL4495†	HL4142 + pHL1335 + pHL600
HL4496	HL4142 + pHL1335 + pHL601
HL4510†	HL4142 + pHL1318 + pHL662

HL4550	HL4142 + pHL1488 + pHL600
HL5050	HL4142 + pHL1529 + pHL1590
HL5620	HL4142 + pHL1722 + pHL1757
HL5621	HL4142 + pHL1724 + pHL1757
HL5623	HL4142 + pHL600 + pHL1757
HL5624	HL4142 + pHL601 + pHL1757
HL5643	HL4142 + pHL1726 + pHL1757
HL5644	HL4142 + pHL1745 + pHL1757
HL5646	HL4142 + pHL1747 + pHL1757
HL5691	HL4142 + pHL1355 + pHL1790
HL6035	HL4018 + pHL1756 + pHL1790
HL6051	HL4142 + pHL1914 + pHL1722
HL6055	HL4142 + pHL1928 + pHL1757
HL6105	HL4142 + pHL1940 + pHL1757
HL6129	HL4142 + pHL1947 + pHL1790
HL6131	HL4142 + pHL1948 + pHL1790

Table 4.1 | Plasmids and strains. Selected plasmids and strains were submitted to the Addgene repository (www.addgene.org). *ID numbers beginning with "pHL" correspond to plasmids; ID numbers beginning with "HL" correspond to strains. [†]These plasmids and strains were reported previously [14]. [‡]This indicates full-length *cstA* gene with its native leader region amplified with the cstAleadSalF and cstAstopApaR oligos (see **Table 4.2**). [§]This strain is MG1655 with $lacI_q$ added to the chromosome [21].

Table 4.2 | Oligonucleotides

name	*description*	*sequence*
csrARBSXmaF	PCR amplifies *csrA* with a synthetic RBS (st7)	CCTCCCGGGTAAGGAGGAAAAAAAATGCTGATTCTGACTCGTCGAGTTG
csrARBS3XmaF	PCR amplifies *csrA* with a synthetic RBS (st3)	CCTCCCGGGTAAGGAGGAAAATGCTGATTCTGACTCGTCGAGTTG
csrAKpnHindR	PCR amplifies *csrA*	GGCCAAGCTTCTTTCAGGTACCTTAGTAACTGGACTGCTGGGATTTTTCAG
csrARBS2NotIF	PCR amplifies *csrA* with a synthetic RBS (st2)	TCCTGCGGCCGCTAAGGAGGAAATGCTGATTCTGACTCGTCGAGTTG
csrARBS3NotIF	PCR amplifies *csrA* with a synthetic RBS (st3)	TCCTGCGGCCGCTAAGGAGGAAAATGCTGATTCTGACTCGTCGAGTTG
csrAApaR	PCR amplifies *csrA*	TAAGGGCCCTTAGTAACTGGACTGCTGGGATTTTTCAG
csrBSalF	PCR amplifies *csrB*	CAAGTCGACGAGTCAGACAACGAAGTGAACATC
csrBApaR	PCR amplifies *csrB*	CATGGGCCCAATAAAAAAAGGGAGCACTGTATTCACAGC
csrBXmaF	PCR amplifies *csrB*	CCTCCCGGGGAGTCAGACAACGAAGTGAACATC
csrBHindR	PCR amplifies *csrB*	GGCCAAGCTTAATAAAAAAAGGGAGCACTGTATTCACAGC
csrCSalF	PCR amplifies *csrC*	CAAGTCGACATAGAGCGAGGACGCTAACAGGAAC
csrC2ApaR	PCR amplifies *csrC*	CATGGGCCCCAGTATAGATTTGCGGCGGAATCT
GFPRBSSalSphF	PCR amplified *gfp* with *Sal*I & *Sph*I sites	TTAGTCGACTAAGGAGGAAAAAGCATGCGTAAAGGAGAAGAACTTTTC
RCYFPNoRBSSphF	PCR synthesis of *mCherry* without an RBS sequence	TACGCATGCTGAGCAAGGGCGAGGAG
RCYApaIR	PCR synthesis of *mCherry*	CATGGGCCCTTACTTGTACAGCTCGTCCATGCC
PconNoHindBamHF	PCR synthesis of Pcon promoter with no *Hin*dIII site	CGCGGATCCTCGAGCACCGTCGTTGTTGACATTTTTATGCTTGGCGGTTATAAT
PconNoHindXmaR	PCR synthesis of Pcon promoter with no *Hin*dIII site	CCTCCCGGGTGTGTGGAATCCATTATAACCGCCAAGCATAAAAATGTCAACAAC
PconNoHindEcoRF	PCR synthesis of Pcon promoter with no *Hin*dIII site	CCGGAATTCTCGAGCACCGTCGTTGTTGACATTTTTATGCTTGGCGGTTATAAT
PconNoHindNotIR	PCR synthesis of Pcon promoter with no *Hin*dIII site	TCCTGCGGCCGCCTGTGTGGAATCCATTATAACCGCCAAGCATAAAAATGTCAACAAC
PconM2NoHindAatF	PCR synthesis of PconM2 promoter with no *Hin*dIII site	CGCGACGTCTCGAGCACCGTCGTTGTTTACATTTTTATGCTTGGCGGTTATGAT
PconM2NoHindSalR	PCR synthesis of PconM2 promoter with no *Hin*dIII site	TTAGTCGACCTGTGTGGAATCCATCATAACCGCCAAGCATAAAAATGTAAACAAC
PconM6NoHindAatF	PCR synthesis of PconM10 promoter with no *Hin*dIII site	CGCGACGTCTCGAGCACCGTCGTTGTTGACATTTTTATGCTTGGCGTATAAT
PconM10NoHindSalR	PCR synthesis of PconM10 promoter with no *Hin*dIII site	TTAGTCGACCTGTGTG GAATCCATTATACGCCAAGCATAAAAATGTCAACAAC
PconM8NoHindAatF	PCR synthesis of PconM8 and M12 promoter with no *Hin*dIII site	CGCGACGTCTCGAGCACCGTCGTTGTTTACATTTTTATGCTTGGCGGTTATGGT
PconM8NoHindSalR	PCR synthesis of PconM8 promoter with no *Hin*dIII site	TTAGTCGACGAATCCACCATAACCGCCAAGCATAAAAATGTAAACAAC
PconM12NoHindSalR	PCR synthesis of PconM12 promoter with no *Hin*dIII site	TTAGTCGACCTGTGTGGAATCCACCATAACCGCCAAGCATAAAAATGTAAACAAC

cstAstopApaR	PCR synthesis of full length *cstA*	CATGGGCCCTTAGTGTGCGCCTTTTGCCTGCGC
glgCleadSalF	PCR amplifies 5'UTR of *glgC* leader for fusion to *gfp*	CCTGTCGACTCTGGCAGGGACCTGCACACGGATTG
glgCleadSphR	PCR amplifies 5'UTR of *glgC* for fusion to *gfp*	TACGCATGCTAACCATGACTAACTCCTTTTTTATCATCTCTGG
cstAleadSalF	PCR amplifies 5'UTR of *cstA* leader for fusion to *gfp*	CCTGTCGACAAATGTAACATCTCTATGGACACG
cstAleadSphR	PCR amplifies 5'UTR of *cstA* leader for fusion to *gfp*	TACGCATGCTCCAGACGAGGTATTTCCCTGATTT
pgaAleadFullSalF	PCR amplifies 5'UTR of *pgaA* leader for fusion to *gfp*	CCTGTCGACAGGCATTGGGATTTATGCCGTATTCC
pgaAleadSphR	PCR amplifies 5'UTR of *pgaA* leader for fusion to *gfp*	TACGCATGCTTTTTCTGCTACTTGAATACATCCTGTATTACTC
HfqleadSalF	PCR amplifies 5'UTR of *hfq* leader for fusion to *gfp*	CCTGTCGACGTATCGTGCGCAATTTTTTCAGAA
HfqleadSphR	PCR amplifies 5'UTR of *hfq* leader for fusion to *gfp*	TACGCATGCGTTCCCGACGCAGTGCGTTCAG
ydeHleadSalF	PCR amplifies 5'UTR of *ydeH* leader for fusion to *gfp*	CCTGTCGACAATAGCGCGCACAAGGAACTGTGA
ydeHleadSphR	PCR amplifies 5'UTR of *ydeH* leader for fusion to *gfp*	TACGCATGCTTTCCGTTGTCTTCTTGATCATTGC
AgdownswitchSalF	PCR synthesis of RBS sequence from the Agn43 system	TTAGTCGACGATAAGCTAATAATAACCTTTGTC
AgmvwSphIR	PCR synthesis of RBS sequence from the Agn43 system	TTTACGCATGCTCATGTGATTCCATACCAG
TetRRBS3XmaF	PCR synthesis of *tetR* with a synthetic RBS (st3)	TATCCCGGGTAAGGAGGAAAATGTCTAGATTAGATAAAAGTAAAG
TetRHindIIIR	PCR synthesis of *tetR*	GGCCAAGCTTAAGACCCACTTTCACATTTAAG

Table 4.2 | Oligonucleotides. Oligonucleotide sequences for selected PCR primers that were used to construct the plasmids (**Table 4.1**, **Fig. 4.11**).

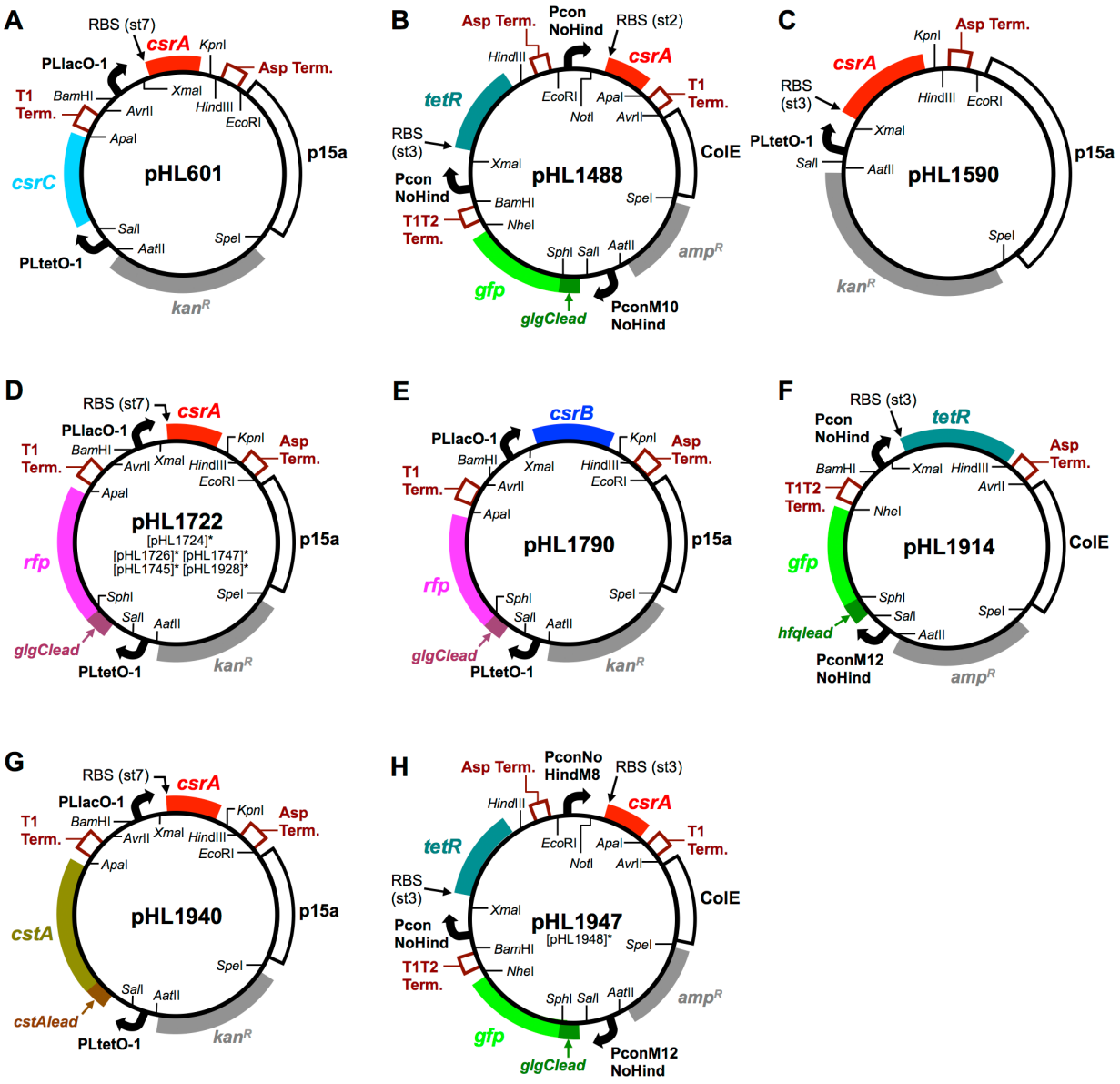

Fig. 4.11 | Plasmid maps. Only new plasmids with maps not drawn elsewhere are depicted here. *Plasmids listed in brackets are similar to the plasmid shown except for the differences described here. Relative expression levels from the different ribosome binding sequence (RBS) used are as follows: st7 > st3 > st2. p15a and ColE are origins of replication. T1 Term, T1T2 Term and Asp Term are terminator sequences. (***A***) pHL601. (***B***) pHL1488. (***C***) pHL1590. (***D***) pHL1722. In pHL1724, *pgaAleadFull* replaces *glgClead*. In pHL1726, *hfqlead* replaced *glgClead*. In pHL1745, *cstAlead* replaced *glgClead*. In pHL1747, *ydeHlead* replaced *glgClead*. In pHL1928, the agn43RBS replaced *glgClead*. (***E***) pHL1790. (***F***) pHL1914. (***G***) pHL1940. (***H***) pHL1947. In pHL1948, PconNoHindM10 replaced PconNoHindM8.

5. Conclusion

Gene regulatory circuits possess a wide variety of functionality, enabling them to control separate sets of genes in distinct ways in accordance with the genes' physiological roles. Understanding how to configure genetic circuits to achieve these distinct modes of functionality is important not only for understanding native biological systems, but also for the rational design and construction of synthetic circuits for industrial, medical and environmental purposes. The three studies presented in this thesis each investigated the functionality and architecture of translational RNA regulatory networks.

The first study demonstrated that the dynamics of signaling in the CsrA regulatory cascade is strongly influenced both (i) by the interplay between slowly-cleared regulatory proteins and upstream regulators that either sequester or degrade them and (ii) by negative feedback that arises from the inhibition of CsrD expression by CsrA. While adjusting molecular turnover and applying negative feedback are both known mechanisms for manipulating the dynamic behavior of biochemical systems, this first study is important because it identified these two mechanisms as key players in the conserved CsrA network which governs bacterial metabolism, motility, virulence, quorum sensing and biofilm formation; additionally, this study helped to establish the operational properties of the CsrB (and CsrC) regulatory RNAs: non-coding RNAs which have stirred significant interest in the microbiology community for their distinctive mechanism of action.

The second study described the functional constraints of the Hfq-dependent sRNA network. Because a large pool of genetically distinct sRNA and mRNA molecules must interact with the shared Hfq chaperone for those sRNAs to regulate their targets, competition for Hfq can easily lead to crosstalk among pathways and potentially the collapse of the entire regulatory network [20]. In order to enable the numerous parallel pathways to utilize Hfq, certain kinetic constraints must be obeyed in order to avoid rampant formation of mismatched sRNA-Hfq-mRNA complexes. This study is important because it provided a general conceptual and mathematical framework for understanding these constraints.

Part 5: Conclusion

The third study also examined how an RNA regulatory network may avoid crosstalk, but this time in the context of the CsrA regulatory system. In this case, sequestering non-coding RNAs (*e.g.* CsrB) are shown to be able to buffer the activity of CsrA; additionally, that buffering is shown to mitigate mRNA-to-mRNA crosstalk by minimizing the influence that changes in mRNA levels has on CsrA activity. This result is important because it demonstrated that the principle of buffering, while well known in the context of pH, can also be applied to understanding genetic regulatory networks.

All told, the three studies presented here systematically examined signaling dynamics and crosstalk within two RNA networks, both of which are highly conserved across species and share common features with other RNA regulatory mechanisms. By understanding both the realized and the potential operational behaviors of these model systems we enhance our comprehension of the space in which all RNA regulatory mechanisms function and evolve.

6. References

1. Goldbeter, A. & Koshland, D. E., An amplified sensitivity arising from covalent modification in biological systems. *Proc Natl Acad Sci* **78** (11), 6840-6844 (1981).
2. Louis, M. & Becskei, A., Binary and graded responses in gene networks. *Sci. STKE* **2002** (143), pe33 (2002).
3. Chen, D. & Arkin, A. P., Sequestration-based bistability enables tuning of the switching boundaries and design of a latch. *Mol Syst Biol* **8** (620) (2012).
4. Gardner, T. S., Cantor, C. R. & Collins, J. J., Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**, 339-342 (2000).
5. Buchler, N. E. & Cross, F. R., Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Mol Syst Biol* **5** (272) (2009).
6. Arias, A. M. & Hayward, P., Filtering transcriptional noise during development: concepts and mechanisms. *Nature Rev Genet* **7**, 34-44 (2006).
7. Viney, M. & Reece, S. E., Adaptive noise. *Proc R Soc B* **280** (1767), 20131104 (2013).
8. Becskei, A. & Serrano, L., Engineering stability in gene networks by autoregulation. *Nature* **405**, 590-593 (2000).
9. Elowitz, M. B. & Leibler, S., A synthetic oscillatory network of transcriptional regulators. *Nature* **403**, 335-338 (2000).
10. Stricker, J. *et al.*, A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516-520 (2008).
11. Ross, J. & Arkin, A. P., Complex systems: From chemistry to systems biology. *Proc Natl Acad Sci* **106** (16), 6433-6434 (2009).
12. Epstein, I. R. & Pojman, J. A., *An introduction to nonlinear chemical dynamics: oscillations, waves, patters and chaos* (Oxford Univeristy Press, New York, 1998).
13. Andrews, S. S. & Arkin, A. P., Simulating cell biology. *Curr Biol* **16** (14), R523-R527 (2006).
14. Adamson, D. N. & Lim, H. N., Rapid and robust signaling in the CsrA cascade via RNA-protein interactions and feedback regulation. *Proc Natl Acad Sci* **110** (32), 13120-13125 (2013).
15. Adamson, D. N. & Lim, H. N., Non-coding RNAs buffer mRNA crosstalk in the CsrA regulatory network. *(manuscript in preparation)* (2013).
16. Jost, D., Nowojewski, A. & Levine, E., Small RNA biology is systems biology. *BMB Reports* **44** (1), 11-21 (2011).
17. De Lay, N., Schu, D. J. & Gottesman, S., Bacterial small RNA-based negative regulation: Hfq and its accomplices. *J Biol Chem* **288** (12), 7996-8003 (2013).
18. Storz, G., Vogel, J. & Wassarman, K. M., Regulation by Small RNAs in Bacteria: Expanding Frontiers. *Mol. Cell* **43** (6), 880-891 (2011).

19. Romeo, T., Vakulskas, C. A. & Babitzke, P., Post-transcriptional regulation on a global scale: Form and function of Csr/Rsm systems. *Environ Microbiol* **15** (2), 313-324 (2013).
20. Adamson, D. N. & Lim, H. N., Essential requirements for robust signaling in Hfq dependent small RNA networks. *PLoS Comput Biol* **7** (8), e1002138 (2011).
21. Hussein, R. & Lim, H. N., Disruption of small RNA signaling caused by competition for Hfq. *Proc Natl Acad Sci* **108** (3), 1110-1115 (2011).
22. Hussein, R. & Lim, H. N., Direct comparison of small RNA and transcription factor signaling. *Nucl Acids Res* **40** (15), 7269-7279 (2012).
23. Chao, Y. & Vogel, J., The role of Hfq in bacterial pathogens. *Curr Opin Microbiol* **13** (1), 24-33 (2010).
24. Sun, X., Zhulin, I. & Wartell, R. M., Predicted structure and phyletic distribution of the RNA-binding protein Hfq. *Nuc Acids Res* **30** (17), 3662-3671 (2002).
25. Chen, S., Zhang, A., Blyn, L. B. & Storz, G., MicC, a second small-RNA regulator of Omp protein expression in Escherichia coli. *J Bacteriol* **186** (20), 6689-6697 (2004).
26. Massé, E., Vanderpool, C. K. & Gottesman, S., Effect of RyhB small RNA on global iron use in Escherichia coli. *J Bacteriol* **187** (20), 6962-6971 (2005).
27. Papenfort, K. & Vogel, J., Multiple target regulation by small noncoding RNAs rewires gene expression at the post-transcriptional level. *Res Microbiol* **160** (4), 278-287 (2009).
28. Salim, N. N., Faner, M. A., Philip, J. A. & Feig, A. L., Requirement of upstream Hfq-binding (ARN)x elements in glmS and the Hfq C-terminal region for GlmS upregulation by sRNAs GlmZ and GlmY. *Nuc Acids Res* **40** (16), 8021-8032 (2012).
29. Majdalani, N., Cunning, C., Sledjeski, D., Elliott, T. & Gottesman, S., DsrA RNA regulates translation of RpoS message by an anti-antisense mechanism, independent of its action as an antisilencer of transcription. *Proc Natl Acad Sci* **95** (21), 12462–12467 (1998).
30. Moon, K. & Gottesman, S., Competition among Hfq-binding small RNAs in Escherichia coli. *Mol Microbiol* **82** (6), 1545-1562 (2011).
31. Olejniczak, M., Despite similar binding to the Hfq protein regulatory RNAs widely differ in their competition performance. *Biochemistry* **50** (21), 4427-4440 (2011).
32. Timmermans, J. & Van Melderen, L., Post-transcriptional global regulation by CsrA in bacteria. *Cell Mol Life Sci* **67** (17), 2897-2908 (2010).
33. Lucchetti-Miganeh, C., Burrowes, E., Baysse, C. & Ermel, G., The post-transcriptional regulator CsrA plays a central role in the adaptation of bacterial pathogens to different stages of infection in animal hosts. *Microbiology* **154**, 16-29 (2008).
34. Yakhnin, A. V. *et al.*, CsrA activates flhDC expression by protecting flhDC mRNA from RNase E-mediated cleavage. *Mol Microbiol* **87** (4), 851-866 (2013).
35. Tan, C., Reza, F. & You, L., Noise-Limited Frequency Signal Transmission in Gene Circuits. *Biophys J* **93** (11), 3753-3761 (2007).

36. Litvak, V. *et al.*, Function of C/EBPdelta in a regulatory circuit that discriminates between transient and persistent TLR4-induced signals. *Nat Immunol* **10** (4), 437-443 (2009).
37. Tyson, J. J., Chen, K. C. & Novak, B., Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr Opin Cell Biol* **15** (2), 221-231 (2003).
38. Knox, B. E., Devreotes, P. N., Goldbeter, A. & Segel, L. A., A molecular mechanism for sensory adaptation based on ligand-induced receptor modification. *Proc Natl Acad Sci* **83** (8), 2345-2349 (1986).
39. Spiro, P. A., Parkinson, J. S. & Othmer, H. G., A model of excitation and adaptation in bacterial chemotaxis. *Proc Natl Acad Sci* **94** (14), 7263-7268 (1997).
40. Yi, T.-M., Huang, Y., Simon, M. I. & Doyle, J., Robust perfect adaptation in bacterial chemotaxis through integral feedback control. *Proc Natl Acad Sci* **97** (9), 4649-4653 (2000).
41. Behar, M., Dohlman, H. G. & Elston, T. C., Kinetic insulation as an effective mechanism for achieving pathway specificity in intracellular signaling networks. *Proc Natl Acad Sci* **104** (41), 16146–16151 (2007).
42. Donaldson, R. & Calder, M., Modelling and Analysis of Biochemical Signalling Pathway Cross-talk. *EPTCS* **19**, 40-54 (2010).
43. Mather, W. H., Hasty, J., Tsimring, L. S. & Williams, R. J., Translational cross talk in gene networks. *Biophys J* **104** (11), 2564-2572 (2013).
44. Rosenfeld, N., Elowitz, M. B. & Alon, U., Negative autoregulation speeds the response times of transcription networks. *J Mol Biol* **323** (5), 785-793 (2002).
45. Hooshangi, S., Thiberge, S. & Weiss, R., Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *Proc Natl Acad Sci* **102** (10), 3581-3586 (2005).
46. Mehta, P., Goyal, S. & Wingreen, N. S., A quantitative comparison of sRNA-based and protein-based gene regulation. *Mol Syst Biol* **4** (221) (2008).
47. Levine, E., Zhang, Z., Kuhlman, T. & Hwa, T., Quantitative characteristics of gene regulation by small RNA. *PLoS Biol* **5** (9), e229 (2007).
48. Shimoni, Y. *et al.*, Regulation of gene expression by small non-coding RNAs: a quantitative view. *Mol Syst Biol* **3** (138) (2007).
49. Weilbacher, T. *et al.*, A novel sRNA component of the carbon storage regulatory system of Escherichia coli. *Mol Microbiol* **48** (3), 657-670 (2003).
50. Liu, M. Y. *et al.*, The RNA molecule CsrB binds to the global regulatory protein CsrA and antagonizes its activity in Escherichia coli. *J Biol Chem* **272** (28), 17502-17510 (1997).
51. Suzuki, K., Babitzke, P., Kushner, S. R. & Romeo, T., Identification of a novel regulatory protein (CsrD) that targets the global regulatory RNAs CsrB and CsrC for degradation by RNase E. *Genes & Dev* **20**, 2605-2617 (2006).
52. Gutiérrez, P. *et al.*, Solution structure of the carbon storage regulator protein CsrA from Escherichia coli. *J Bacteriol* **187** (10), 3496–3501 (2005).

53. Baker, C. S., Morozov, I., Suzuki, K., Romeo, T. & Babitzke, P., CsrA regulates glycogen biosynthesis by preventing translation of glgC in Escherichia coli. *Mol Microbiol* **44** (6), 1599-1610 (2002).

54. Lutz, R. & Bujard, H., Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucl. Acids Res.* **25** (6), 1203-1210 (1997).

55. Lanzer, M. & Bujard, H., Promoters largely determine the efficiency of repressor action. *Proc Natl Acad Sci* **85** (23), 8973-8977 (1988).

56. Timmermans, J. & Van Melderen, L., Conditional essentiality of the csrA gene in Escherichia coli. *J. Bacteriol.* **191** (5), 1722-1724 (2009).

57. Wang, X. *et al.*, CsrA post-transcriptionally represses pgaABCD, responsible for synthesis of a biofilm polysaccharide adhesin of Escherichia coli. *Mol Microbiol* **56** (6), 1648-1663 (2005).

58. Andersen, J. B. *et al.*, New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* **64** (6), 2240-2246 (1998).

59. Gottesman, S., Proteases and their targets in Escherichia coli. *Annu Rev Genet* **30**, 465–506 (1996).

60. Rosenfeld, N. & Alon, U., Response delays and the structure of transcription networks. *J Mol Biol* **329** (4), 645-654 (2003).

61. Grilly, C., Stricker, J., Pang, W. L., Bennett, M. R. & Hasty, J., A synthetic gene network for tuning protein degradation in Saccharomyces cerevisiae. *Mol Syst Biol* **3** (127) (2007).

62. Yakhnin, H. *et al.*, Complex regulation of the global regulatory gene csrA: CsrA-mediated translational repression, transcription from five promoters by Eσ70 and EσS, and indirect transcriptional activation by CsrA. *Mol Microbiol* **81** (3), 689-704 (2011).

63. Gudapaty, S., Suzuki, K., Wang, X., Babitzke, P. & Romeo, T., Regulatory interactions of Csr components: the RNA binding protein CsrA activates csrB transcription in Escherichia coli. *J Bacteriol* **183** (20), 6017-6027 (2001).

64. Jonas, K. *et al.*, The RNA binding protein CsrA controls cyclic di-GMP metabolism by directly regulating the expression of GGDEF proteins. *Mol Microbiol* **70** (1), 236-257 (2008).

65. Overgaard, M., Johansen, J., Møller-Jensen, J. & Valentin-Hansen, P., Switching off small RNA regulation with trap-mRNA. *Mol Microbiol* **73** (5), 790-800 (2009).

66. Figueroa-Bossi, N., Valentini, M., Malleret, L. & Bossi, L., Caught at its own game: regulatory small RNA inactivated by an inducible transcript mimicking its target. *Genes Dev* **23** (17), 2004-2015 (2009).

67. Mukherjee, S. *et al.*, CsrA–FliW interaction governs flagellin homeostasis and a checkpoint on flagellar morphogenesis in Bacillus subtilis. *Mol Microbiol* **82** (2) (2011).

68. Banse, A. V., Chastanet, A., Rahn-Lee, L., Hobbs, E. C. & Losick, R., Parallel pathways of repression and antirepression governing the transition to stationary phase in Bacillus subtilis. *Proc Natl Acad Sci* **105** (40), 15547–15552 (2008).
69. Alba, B. M. & Gross, C. A., Regulation of the Escherichia coli σE-dependent envelope stress response. *Mol Microbiol* **52** (3), 613-619 (2004).
70. Mercante, J., Edwards, A. N., Dubey, A. K., Babitzke, P. & Romeo, T., Molecular geometry of CsrA (RsmA) binding to RNA and its implications for regulated expression. *J Mol Biol* **392** (2), 511-528 (2009).
71. Taniguchi, Y. *et al.*, Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. *Science* **329** (5991), 533-538 (2010).
72. Liu, M. Y., Yang, H. & Romeo, T., The product of the pleiotropic Escherichia coli gene csrA modulates glycogen biosynthesis via effects on mRNA stability. *J Bacteriol* **177** (10), 2663-2672 (1995).
73. Mosteller, R. D., Goldstein, R. V. & Nishimoto, K. R., Metabolism of individual proteins in exponentially growing Escherichia coli. *J Biol Chem* **255** (6), 2524-2532 (1980).
74. Hopkins, J. F., Panja, S. & Woodson, S. A., Rapid binding and release of Hfq from ternary complexes during RNA annealing. *Nuc Acids Res* **39** (12), 5193-5202 (2011).
75. Kime, L., Jourdan, S. S., Stead, J. A., Hidalgo-Sastre, A. & McDowall, K. J., Rapid cleavage of RNA by RNase E in the absence of 5′ monophosphate stimulation. *Mol Microbiol* **76** (3), 590–604 (2009).
76. Redko, Y. *et al.*, Determination of the catalytic parameters of the N-terminal half of Escherichia coli ribonuclease E and the identification of critical functional groups in RNA substrates. *J Biol Chem* **278**, 44001-44008 (2003).
77. Vellanoweth, R. L. & Rabinowitz, J. C., The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. *Mol Microbiol* **6** (9), 1105-1114 (2006).
78. Alon, U., *An Introduction to Systems Biology: Design Principles of Biological Circuits* (Chapman & Hall/CRC, Boca Raton, 2007).
79. Kim, H. S. *et al.*, Distinct binding properties of TIAR RRMs and linker region. *RNA Biol* **392** (2), 511-528 (2009).
80. Fender, A., Elf, J., Hampel, K., Zimmermann, B. & Wagner, G. H., RNAs actively cycle on the Sm-like protein Hfq. *Genes Dev* **24**, 2621-2626 (2010).
81. Law, M. J., Rice, A. J., Lin, P. & Laird-Offringa, I. A., The role of RNA structure in the interaction of U1A protein with U1 hairpin II RNA. *RNA* **12**, 1168-1178 (2006).
82. Datsenko, K. A. & Wanner, B. L., One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci* **97** (12), 6640–6645 (2000).
83. Vellanoweth, R. L. & Rabinowitz, J. C., The influence of ribosome-binding-site elements on translational efficiency in Bacillus subtilis and Escherichia coli in vivo. *Mol Microbiol* **6** (9), 1105–1114 (1992).

84. Shaner, N. C. *et al.*, Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat Biotechnol* **22** (12), 1567-1572 (2004).

85. Repoila, F. & Darfeuille, F., Small regulatory non-coding RNAs in bacteria: physiology and mechanistic aspects. *Biol Cell* **101** (2), 117-131 (2009).

86. Sittka, A., Pfeiffer, V., Tedin, K. & Vogel, J., The RNA chaperone Hfq is essential for the virulence of Salmonella typhimurium. *Mol Microbiol* **63** (1), 193-217 (2007).

87. Liu, Y. *et al.*, Hfq is a global regulator that controls the pathogenicity of Staphylococcus aureus. *PLoS ONE* **5** (9), e13069 (2010).

88. Ding, Y., Davis, B. M. & Waldor, M. K., Hfq is essential for Vibrio cholerae virulence and downregulates σE expression. *Mol Microbiol* **53** (1), 345-354 (2004).

89. Fantappiè, L. *et al.*, The RNA Chaperone Hfq is involved in stress response and virulence in Neisseria meningitidis and is a pleiotropic regulator of protein expression. *Infect Immun* **77** (5), 1842-1853 (2009).

90. Christiansen, J. K., Larsen, M. H., Ingmer, H., Søgaard-Andersen, L. & Kallipolitis, B. H., The RNA-binding protein Hfq of Listeria monocytogenes: role in stress tolerance and virulence. *J Bacteriol* **186** (11), 3355-3362 (2004).

91. Kajitani, M. & Ishihama, A., Identification and sequence determination of the host factor gene for bacteriophage Q beta **19** (5), 1063–1066 (1991).

92. Møller, T. *et al.*, Hfq: A bacterial Sm-like protein that mediates RNA-RNA interaction. *Mol Cell* **9** (1), 23-30 (2002).

93. Link, T. M., Valentin-Hansen, P. & Brennan, R. G., Structure of Escherichia coli Hfq bound to polyriboadenylate RNA. *Proc Natl Acad Sci* **106** (46), 19292–19297 (2009).

94. Zhang, A., Wassarman, K. M., Ortega, J., Steven, A. C. & Storz, G., The Sm-like Hfq protein increases OxyS RNA interaction with target mRNAs. *Mol Cell* **9** (1), 11-22 (2002).

95. Salim, N. N. & Feig, A. L., An upstream Hfq binding site in the fhlA mRNA leader region facilitates the OxyS-fhlA interaction. *PLoS ONE* **5** (9), e13028 (2010).

96. Morita, T., Maki, K. & Aiba, H., RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. *Genes Dev* **19**, 2176-2186 (2005).

97. Mohanty, B. K., Maples, V. F. & Kushner, S. R., The Sm-like protein Hfq regulates polyadenylation dependent mRNA decay in Escherichia coli. *Mol Microbiol* **54** (4), 905-920 (2004).

98. Sukhodolets, M. V. & Garges, S., Interaction of Escherichia coli RNA polymerase with the ribosomal protein S1 and the Sm-like ATPase Hfq. *Biochemistry* **42** (26), 8022–8034 (2003).

99. Updegrove, T. B., Correia, J. J., Galletto, R., Bujalowski, W. & Wartell, R. M., E. coli DNA associated with isolated Hfq interacts with Hfq's distal surface and C-terminal domain. *Biochim Biophys Acta* **1799** (8), 588-596 (2010).

100. Takada, A., Wachi, M., Kaidow, A., Takamura, M. & Nagai, K., DNA Binding properties of the hfq gene product of Escherichia coli. *Biochem Biophys Res Commun* **236** (3), 576-579 (1997).

101. Mikulecky, P. J. *et al.*, Nat Struct Mol Biol **11** (12), 1206-1214 (2004).

102. Sittka, A., Sharma, C. M., 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* **6** (3), 266-275 (2009).

103. Zhang, A. *et al.*, Global analysis of small RNA and mRNA targets of Hfq. *Mol Microbiol* **50** (4), 1111-1124 (2003).

104. Lease, R. A. & Woodson, S. A., Cycling of the Sm-like Protein Hfq on the DsrA small regulatory RNA. *J Mol Biol* **344** (5), 1211-1223 (2004).

105. Geissmann, T. A. & Touati, D., Hfq, a new chaperoning role: binding to messenger RNA determines access for small RNA regulator. *EMBO J* **23** (2), 396-405 (2004).

106. Aiba, H., Mechanism of RNA silencing by Hfq-binding small RNAs. *Curr Opin Microbiol* **10** (2), 134-139 (2007).

107. Soper, T. J. & Woodson, S. A., The rpoS mRNA leader recruits Hfq to facilitate annealing with DsrA sRNA. *RNA* **14**, 1907-1917 (2008).

108. Arluison, V. *et al.*, Spectroscopic observation of RNA chaperone activities of Hfq in post-transcriptional regulation by a small non-coding RNA. *Nucl Acids Res* **35** (3), 999-1006 (2007).

109. Rajkowitsch, L. & Schroeder, R., Dissecting RNA chaperone activity. *RNA* **13**, 2053-2060 (2007).

110. Massé, E., Escorcia, F. E. & Gottesman, S., Coupled degradation of a small regulatory RNA and its mRNA targets in Escherichia coli. *Genes Dev* **17**, 2374-2383 (2003).

111. Altuvia, S., Identification of bacterial small non-coding RNAs: experimental approaches. *Curr Opin Microbiol* **10** (3), 257-261 (2007).

112. Zhang, A. *et al.*, The OxyS regulatory RNA represses rpoS translation and binds the Hfq (HF-I) protein. *EMBO J* **17** (20), 6061-6068 (1998).

113. Ali Azam, T., Iwata, A., Nishimura, A., Ueda, S. & Ishihama, A., Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *J Bacteriol* **181** (20), 6361-6370 (1999).

114. Vytvyska, O. *et al.*, Host factor I, Hfq, binds to Escherichia coli ompA mRNA in a growth rate-dependent fashion and regulates its stability. *Proc Natl Acad Sci* **95** (24), 14118-14123 (1998).

115. Copeland, R. A., *Enzymes: A Practical Introduction to Structure, Mechanism, and Data Analysis, 2nd Edition* (Wiley-VCH, 2000).

116. Hopkins, J. F., Panja, S., McNeil, S. A. N. & Woodson, S. A., Effect of salt and RNA structure on annealing and strand displacement by Hfq. *Nucl Acids Res* **37** (18), 6205-6213 (2009).

117. Altuvia, S., Zhang, A., Argaman, L., Tiwari, A. & Storz, G., The Escherichia coli OxyS regulatory RNA represses fhlA translation by blocking ribosome binding. *EMBO J* **17** (20), 6069-6075 (1998).

118. Massé, E. & Gottesman, S., A small RNA regulates the expression of genes involved in iron metabolism in Escherichia coli. *Proc Natl Acad Sci* **99** (7), 4620-4625 (2002).

119. Altuvia, S., Weinstein-Fisher, D., Zhang, A., Postow, L. & Storz, G., A small, stable RNA induced by oxidative stress: role as a pleiotropic regulator and antimutator. *Cell* **90** (1), 43-53 (1997).

120. Vanderpool, C. K. & Gottesman, S., Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. *Mol Microbiol* **54** (4), 1076-1089 (2004).

121. Papenfort, K., Bouvier, M., Mika, F., Sharma, C. M. & Vogel, J., Evidence for an autonomous 5′ target recognition domain in an Hfq-associated small RNA. *Proc Natl Acad Sci* **107** (47), 20435-20440 (2010).

122. Peer, A. & Margalit, H., Accessibility and evolutionary conservation mark bacterial small-RNA target-binding regions. *J Bacteriol* **193** (7), 1690-1701 (2011).

123. Sittka, A. *et al.*, Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. *PLoS Genet* **4** (8), e1000163 (2008).

124. 't Hoen, P. A. C. *et al.*, Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucl Acids Res* **36** (21), e141 (2008).

125. Carmichael, G. G. & Weber, K., The Host factor required for RNA phage Q beta RNA replication in vitro. *J Biol Chem* **250** (10), 3607-3612 (1975).

126. Arluison, V. *et al.*, Structural Modelling of the Sm-like Protein Hfq from Escherichia coli. *J Mol Biol* **320** (4), 705-712 (2002).

127. Brescia, C. C., Kaw, M. K. & Sledjeski, D. D., The DNA binding protein H-NS binds to and alters the stability of RNA in vitro and in vivo. *J Mol Biol* **339** (3), 505-514 (2004).

128. Hwang, W., Arluison, V. & Hohng, S., Dynamic competition of DsrA and rpoS fragments for the proximal binding site of Hfq as a means for efficient annealing. *Nucl Acids Res* **39** (12), 5131-5139 (2011).

129. Zalatan, J. G., Coyle, S. M., Rajan, S., Sidhu, S. S. & Lim, W. A., Conformational control of the Ste5 scaffold protein insulates against MAP kinase misactivation. *Science* **337** (6099), 1218-1222 (2012).

130. Bardwell, L., Mechanisms of MAPK signalling specificity. *Biochemical Society Transactions* **34**, 837–841 (2006).

131. Baker, C. S. *et al.*, CsrA inhibits translation initiation of Escherichia coli hfq by binding to a single site overlapping the Shine-Dalgarno sequence. *J Bacteriol* **189** (15), 5472-5481 (2007).

132. Dubey, A. K. *et al.*, CsrA regulates translation of the Escherichia coli carbon starvation gene, cstA, by blocking ribosome access to the cstA transcript. *J Bacteriol* **185** (15), 4450-4460 (2003).

133. Bernstein, J. A., Khodursky, A. B., Lin, P.-H., Lin-Chao, S. & Cohen, S. N., Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. *Proc Natl Acad Sci* **99** (15), 9697-9702 (2002).

134. Ghildiyal, M. & Zamore, P. D., Small silencing RNAs: an expanding universe. *Nat Rev Genet* **10** (2), 94-108 (2009).

135. Jost, D., Nowojewski, A. & Levine, E., Regulating the many to benefit the few: role of weak small RNA targets. *Biophys J* **104** (8), 1773-1782 (2013).