Precise tuning of bacterial translation initiation by non-equilibrium 5'-UTR unfolding observed in single mRNAs

Sujay Ray

University of Michigan-Ann Arbor

Shiba Dandpat

University of Michigan–Ann Arbor

Surajit Chatterjee

University of Michigan-Ann Arbor

Nils Walter (✉ nwalter@umich.edu)

University of Michigan-Ann Arbor   https://orcid.org/0000-0002-7301-1275

Article

Precise tuning of bacterial translation initiation by non-equilibrium 5'-UTR unfolding observed in single mRNAs

Sujay Ray†, Shiba S. Dandpat, Surajit Chatterjee and Nils G. Walter*

Single-Molecule Analysis Group, Department of Chemistry and Center for RNA Biomedicine, University of Michigan, Ann Arbor, MI 48109, USA. *e-mail: nwalter@umich.edu, †Current address: Wyss Institute for Biologically Inspired Engineering at Harvard University.

KEYWORDS: 5'-UTR structures, preQ$_1$ riboswitch, Shine-Dalgarno sequence, ribosome binding, single-molecule kinetics.

ABSTRACT

Noncoding, structured 5'- untranslated regions (5'-UTRs) of messenger RNAs (mRNAs) control translation efficiency by forming structures that can either recruit or repel the ribosome. Here we exploit a bacterial, $preQ_1$-sensing translational riboswitch to probe how binding of a small ligand controls binding of the bacterial ribosome to the Shine-Dalgarno (SD) sequence. Combining single-molecule fluorescence microscopy with mutational analyses, we find that the stability of 30S ribosomal subunit binding is inversely correlated with the free energy needed to unfold the 5'-UTR during mRNA accommodation from the standby site to the binding cleft. Ligand binding stabilizes 5'-UTR structure to both antagonize 30S recruitment and accelerate 30S dissociation. Depletion of small ribosomal subunit protein S1, known to resolve structured 5'-UTRs, further increases the energetic penalty for mRNA accommodation. The resulting model of rapid standby site exploration followed by gated non-equilibrium unfolding of the 5'-UTR during accommodation provides a mechanistic understanding of translation efficiency.

Introduction

5'-Untranslated regions (5'-UTRs) of messenger RNAs (mRNA) are essential for regulating protein expression in all cells[1-5]. Direct interactions between the small, or 30S, ribosomal subunit, and the 5'-UTR allow modulation of initiation as the rate-limiting step of bacterial mRNA translation[5-8]. The earliest stage of initiation involves two steps as major contributors to overall translation efficiency: reversible binding of the mRNA to a loosely defined "standby site", followed by accommodation into the mRNA binding cleft of the 30S subunit[9-11]. The cleft-accommodated 30S ribosomal binding site (RBS) of the mRNA stretches from nucleotides (nt) −18 to +10 relative to the start site (position +1)[12] and often encompasses a purine-rich region around positions -7 to -4 known as the Shine Dalgarno (SD) sequence[13]. This sequence directly engages with the 3'-terminus of the 16S ribosomal RNA (rRNA) of the 30S subunit – the so-called anti-SD sequence – through base pairing. Translation initiation thus requires unfolding of intrinsic mRNA structure near the start codon, imposing a structure-dependent energetic penalty on translation efficiency[13-17]. Consequently, transcriptome-wide studies have shown that an increased translation efficiency of a given mRNA generally correlates with its reduced propensity to form secondary structures near the RBS[15,16,18].

While the absence of local secondary structure permits efficient translation initiation also on mRNAs that lack an SD sequence[19], structured RNA motifs such as riboswitches embedded in the 5'-UTR require an SD sequence for efficient gene expression, enabling dynamic regulation[14-17,20-22]. Primarily found in bacteria, riboswitches are cis-regulatory elements typically controlling either transcription or translation[15-17]. Following ligand binding to its aptamer domain, a typical translational riboswitch changes the secondary structure of the downstream expression platform to sequester (part of) the SD sequence, thus repressing translation initiation[18].

The extent of riboswitch-mediated translational control varies across the many riboswitch classes discovered to date. In general, larger and more complex riboswitches may entirely abrogate SD and/or start codon access, leading to switch-like ON/OFF behavior[23,24]. In contrast, smaller riboswitches, as exemplified by the H-type pseudoknot of the class-I preQ$_1$ (or 7-aminomethyl-7-deazaguanine) riboswitch from *Thermoanaerobacter tengcongensis* (*Tte*), affect gene expression through only partial SD sequestration (Fig. 1)[25-29]. How binding of a small ligand, which yields only little thermodynamic free energy (~10 kcal/mol)[29,30], regulates mRNA accessibility by the much larger bacterial 30S subunit has remained a conundrum. Transcriptome-wide in vivo SHAPE-MaP analysis of general translation initiation supports a model wherein an mRNA first binds to a standby site on the 30S subunit[7], followed by transient unfolding of 5'-UTR structure during accommodation into the mRNA cleft, imposing a non-equilibrium energetic penalty on translation initiation[10,11]. In the context of riboswitches, these findings suggest that the relative non-equilibrium aptamer unfolding versus 30S binding kinetics may play a gateway role in regulation[31,32], however, the mechanism of this ligand-controlled process has not been observed directly.

In this study, we implement a Single-Molecule Kinetic Analysis of Ribosome Binding (SiM-KARB) assay based on a labeled bacterial 30S subunit to probe 5'-UTR accessibility directly as an important control mechanism of translation initiation by the *Tte* preQ$_1$ riboswitch. Repetitive binding and dissociation of the 30S subunit to the riboswitch-hosting mRNA (henceforth referred to as "R-mRNA") reveals short binding events that mutational analyses enable us to assign as non-specific "standby site" interactions, whereas significantly longer binding events represent "cleft-accommodated" interactions where R-mRNA becomes partially unfolded and accommodated into the mRNA binding cleft of the 30S subunit. We find that both standby and cleft-accommodated binding are similarly

affected by preQ$_1$. Strategic mutations in or near the RBS unveil that increasing aptamer-SD distance and SD:anti-SD base pairing strength both favor initial recognition as well as stable cleft-accommodated binding by the 30S subunit. Finally, the effect of RNA secondary structure near the RBS is further verified by depleting ribosomal protein S1, upon which the 30S subunit shows impaired unfolding of the riboswitch as evidenced by an increased energetic penalty of mRNA accommodation. Taken together, our data support a model wherein the stability of 30S binding is inversely related to the non-equilibrium free energy of aptamer unfolding during mRNA accommodation, enabling the ligand to reduce both 30S binding speed and kinetic stability as two distinguishable mechanisms for modulating translation efficiency.

Results

Riboswitch ligand modulates 30S binding to the mRNA 5'-UTR

Prior crystallographic and single-molecule FRET studies of its isolated aptamer domain have suggested that the *Tte* preQ$_1$ riboswitch controls translation initiation by the ligand-induced sequestration of two nucleotides shared between the P2 helix of the pseudoknot and the SD sequence[28,29,33,34] (Fig. 1a, 1b). The 30S subunit (Fig. 1c), however, occupies a total of ~30 nt in the 5'-UTR that unfold when stably accommodated into the mRNA binding cleft and are thus considered part of the expression platform (Fig. 1d). We *in vitro* transcribed a *Tte* gene-1564 sequence that includes this entire expression platform under the influence of the preQ$_1$ riboswitch, either as a full-length mRNA (R-mRNA^{FL}) or as shorter nascent RNA sequences as would emerge during transcription as a platform for binding the first 30S subunit that initiates the pioneering round of translation (Fig. 1d)[27]. Exploiting the high sequence conservation and comparable thermodynamics of the SD:anti-SD interactions between *Tte* and *Escherichia coli* (*Eco*; Fig. 1e), we used purified *Eco* 30S subunits for

comparing full initiation complex (IC) formation efficiencies on R-mRNAs of different lengths. As expected, we found similar efficiencies for R-mRNA^{FL} and the RNA truncated 30 nt downstream of the start codon (R-mRNA^{+30}), but observed no IC formation on the RNA lacking both SD sequence and start codon (R-mRNA^{-11}; Supplementary Fig. 1). To focus on the role of the 5'-UTR in controlling access of the 30S subunit without the undesired potential for binding further downstream, we therefore chose the short, initiation-competent R-mRNA^{+30} for our studies.

To characterize the 30S binding efficiency of R-mRNA^{+30} in response to the addition of preQ$_1$ ligand, we performed a radioactive filter-binding assay[35], wherein the fraction of 30S-bound over total ^{32}P-radiolabeled R-mRNA^{+30} is calculated from the blotted, double-filtered dots (Fig. 1f). A time course of the normalized 30S-mRNA complex formation in the absence and presence of a high concentration (10 μM)[27] of preQ$_1$ revealed an ~2-fold adverse effect of preQ$_1$ on both pre-steady state kinetics and endpoint (Fig. 1f). This level of modulation is very similar to earlier findings that the translation of R-mRNA^{FL} is decreased *in vitro* by ~40% at saturating preQ$_1$ concentration[27], supporting the notion that the ligand significantly affects translation at the initial 30S binding step. The molecular mechanism of this effect, however, is not revealed by such bulk binding assays.

Riboswitch ligand both antagonizes 30S recruitment and accelerates 30S dissociation

To directly evaluate 30S binding to R-mRNA^{+30} as an important gateway to initiating translation, we developed SiM-KARB, wherein a purified *Eco*-30S subunit with an extended 16S rRNA helix-44 is fluorescently tagged[36] by hybridization with a doubly (to mitigate photobleaching) Cy5-labeled DNA oligonucleotide (Fig. 1c). Repeated binding and dissociation of an excess of these 30S subunits in solution (red) to surface-immobilized 3'-Cy3 labeled R-mRNA^{+30} (green) was monitored by

total internal reflection fluorescence (TIRF) microscopy as transient diffraction-limited co-localization events, performed in the 50 mM Tris-polymix buffer routinely used in translation assays[37] (Fig. 2a,b; see Methods for details). A mixture of short and long binding events was observed with the presence of preQ$_1$ modulating the events (Fig. 2b), supporting two main distinguishable classes of 30S-mRNA interactions, consistent with the known standby and cleft-accommodated binding modes, respectively.

To obtain more quantitative, comparative insights, we pursued several layers of further analysis. First, individual time traces obtained by folding and equilibrating R-mRNA^{+30} in the presence of varying concentrations of preQ$_1$ were idealized into two-state Hidden Markov Models (HMM, Fig. 2b). The HMMs were then used to extract the dwell times during which 30S was either unbound ($\tau_{unbound}$) or bound (τ_{bound}), which both were best fit globally with double-exponential functions[38,39] revealing each two rate constants: 30S binding rate constants $k_{on,fast}$ and $k_{on,slow}$, and 30S dissociation rate constants $k_{off,fast}$ and $k_{off,slow}$, respectively. Of these, the slow $k_{on,slow}$ and $k_{off,slow}$ represented ~30-60% and ~70-90%, respectively, of all binding events and were found to be most significantly affected by the addition of preQ$_1$, with $k_{on,slow}$ decreasing and $k_{off,slow}$ increasing (Fig. 2c and Supplementary Table S2 A,B). Of note, while dual Cy5-labeling prolonged the observation of bound 30S subunits, we corrected all rate constants for photobleaching of both fluorophores (Supplementary Fig. 2). The observed ~27% deceleration of 30S binding with increasing preQ$_1$ (from $k_{on,slow}$ = 0.33 ± 0.04 × 10^6 M^{-1}s^{-1} of relative amplitude = 89 ± 4%; to 0.24 ± 0.03 × 10^6 M^{-1}s^{-1} of 96 ± 8%; Fig. 2c, Supplementary Fig. S3 and Supplementary Table S2A) is consistent with a model of partial SD sequestration upon preQ$_1$-induced folding of the riboswitch. By contrast, the ~3-fold acceleration of 30S dissociation (from $k_{off,slow}$ = 0.002 ± 0.001 s^{-1} of relative amplitude 58 ± 8%; to 0.006 ± 0.002 s^{-1} of 41 ± 8%; Fig. 2c,

Supplementary Fig. S3 and Supplementary Table S2A) suggests an additional effect of $preQ_1$ on dissociation of the assembled 30S-R-mRNA^{+30} complex. Global fitting of the two ligand dependencies with non-cooperative Hill equations yielded a half-saturation point of ~48 nM $preQ_1$ (Fig. 2C), in reasonable agreement with measurements of ligand binding to the isolated riboswitch[28,29] and demonstrating that R-mRNA^{+30} is saturated at 1 µM $preQ_1$.

Second, we asked how individual 30S binding kinetics change for individual R-mRNA^{+30} molecules when folded with 1 µM $preQ_1$ present. To this end, individual trace HMMs were arranged into rastergrams sorted by their relative accessibility to the 30S subunit over the entire trace (Fig. 2d). Based on a threshold bound time derived from the $k_{off,fast}$ and $k_{off,slow}$ dissociation rate constants (Materials and Methods), we categorized individual binding events in the rastergram as either short (red) or long (blue) (Fig. 2d). These rastergrams indicate that each R-mRNA^{+30} molecule interacts quite distinctly with 30S subunits and shows long-term retention of its 30S occupancy behavior. As expected from an SD sequestration model, the number of total 30S binding events decreases upon addition of saturating $preQ_1$ ligand, by ~17% (Fig. 2d). Moreover, while the shorter binding events were less affected (~6%), the long binding events reduced by ~33% with the addition of preQ1 (supporting table S3), suggesting that the longer, cleft accommodated binding of the 30S are primarily disfavored by preQ1 induced riboswitch folding.

Lastly, inspired by previous studies equally suggesting that 30S occupancy behaviors of individual mRNA molecules may be retained long-term, in aggregate leading to a significant downregulation of protein expression[27], we plotted probability histograms that bin the total fraction of time single R-mRNA^{+30} molecules remain 30S associated within our experimental time window of 1,000 s (Supplementary Fig. 4B). From the combined histogram of all datasets (Supplementary Fig.

4A), a 30S occupancy ranking of individual molecules as Low (L, fractional 30S bound time <0.10), Medium (M, fractional 30S bound time between 0.10 and 0.20) or High (H, fractional 30S bound time >0.20) was derived. In the absence of $preQ_1$, 30 ± 4 %, 44 ± 3 % and 26 ± 5 % of molecules show L, M, and H occupancy, respectively; whereas upon addition of saturating 1 μM $preQ_1$, a large majority of the molecular population (81 ± 2 %) exhibits L occupancy, at the expense of both the M (13 ± 2 %) and H (6 ± 2 %) occupancies (Fig. 2e).

These findings suggest that individual molecules belong to subpopulations that respond individually to $preQ_1$ addition with a common trend in which 30S binding becomes suppressed, thereby lowering the probability of translation initiation. Our data also support the model of two binding interactions between R-mRNA^{+30} and 30S subunits, with weak standby site binding and more stable (longer-lived) binding events upon cleft accommodation, both of which are similarly impacted by the presence of $preQ_1$. Finally, the ligand not only competes with 30S for binding to the RNA pseudoknot, but also actively accelerates dissociation of 30S subunits already bound to R-mRNA^{+30}.

The riboswitch-hosting mRNA dynamically adapts to changes in $preQ_1$ concentration

The persistence of individual R-mRNA^{+30} in their equilibrium interaction with 30S subunits raises the question to what extent 30S binding to R-mRNA^{+30} changes in response to sudden, non-equilibrium variations in its environment, as may occur in the bacterial cells due to, e.g., environmental stresses[40]. To address this question, we used a non-equilibrium $preQ_1$ "ligand-jump" experiment from zero to saturating ligand while tracking 30S binding (Fig. 3a, Materials and methods)[27,41]. Consistent with our equilibrium experiments, upon $preQ_1$ addition we observed a general reduction in 30S binding frequency (Fig. 3a top and middle) and a decrease in the 30S bound

time (Fig. 3a, bottom). Overall, the resulting 30% decrease in $k_{on,slow}$ and ~ 3 fold increase in $k_{off,slow}$ (Supplementary Fig. 6) suggest that preQ$_1$ binding in situ has similar effects on 30S association of R-mRNA^{+30} as does its equilibrium binding, both antagonizing 30S recruitment and accelerating dissociation.

Next, we analyzed the probability distribution of 30S fractional bound times. In the absence of preQ$_1$, 46 ± 3 % of R-mRNA^{+30} molecules showed L-type 30S occupancy, while 30 ± 3 % and 24 ± 3 % of molecules showed M and H accessibility, respectively (Fig. 3b). By contrast, after addition of saturating preQ$_1$ mRNA molecules predominantly exhibited reduced accessibility (82 ± 2 % L, 12 ± 2 % M and 6 ± 1% L, Fig. 3c). This shift from high to low 30S occupancy of the same molecules in response to preQ$_1$ further supports that dynamic ligand binding induces the same repulsion and destabilization of 30S subunit binding observed under equilibrium conditions.

Finally, a rastergram allowed us to directly compare the mRNA accessibility in the absence and presence of preQ$_1$ (Fig. 3d). 48% of molecules responded to the addition of preQ$_1$ with a significant reduction in 30S occupancy (Supplementary Table S4A). We also found that the short standby site interactions were reduced by ~30%, whereas the longer cleft-accommodated interactions were reduced even more profoundly, by ~60%, underscoring the large effect of preQ$_1$ on particularly cleft-accommodated interactions (Supplementary Table S4B). Still, 46% of molecules remained in the similar 30S occupancy and did not significantly respond to the addition of preQ$_1$. Only a small (~6%) population transitioned from a less to a higher 30S occupancy. The latter observation suggests that, under non-equilibrium conditions, preQ$_1$ in some cases may promote the refolding of an mRNA, as

was observed previously[27]. On the other end of the behavioral spectrum, some 15% of R-mRNA[+30] molecules became completely inaccessible to 30S binding.

Overall, these observations underscore the dynamic nature of the significant response of the riboswitch to ligand binding, rending 30S binding and thus translation efficiency highly adaptive to environmental cues.

Riboswitch pseudoknot proximity and SD sequence strength independently control 30S binding

In bacteria, the SD sequence is not obligatory for translation initiation[42,43]. While only 0.7% of mRNAs are thought to be leaderless in *E. coli*[44], other bacteria harbor up to 72% leaderless mRNAs[45]. The preQ$_1$ riboswitch and the SD sequence overlap such that preQ$_1$ binding promotes P2 helix formation of the aptamer pseudoknot, sequestering two nucleotides of the SD sequence[29] (Fig. 1a,b). Taken together, these observations raise the question of how sequestering only 2 nt can have the profound effect on 30S binding we observe[29]. To address this question, we next created a systematic set of R-mRNA[+30] mutants that systematically decouple the aptamer and SD sequence.

In the first set, the aptamer and SD sequence were increasingly separated by inserting 1-6 nt between them while maintaining the 8-nt long wild-type (WT) SD sequence (Fig. 4a, constructs I1S8-I6S8; Supplementary Table S1 for the list of sequences). A second set of mutations aimed to increasingly weaken the SD:anti-SD interaction by shortening the SD length from 8 to zero nt while maintaining an unstructured 4-nt UAUA insertion between the aptamer and the remaining SD (Fig. 4e,constructs I4S8 to I4S0; Supplementary Table S1 for the list of sequences). All mutants retained all

nucleotides participating in the pseudoknotted aptamer structure, and the WT would be annotated I2S8 in this nomenclature.

Strikingly, using our equilibrium SiM-KARB assay, we found a $k_{on,slow}$ of $0.25 \pm 0.009 \times 10^6$ M^{-1}s^{-1} for I1S8 in the absence of preQ$_1$, within error the same value as the WT displays at saturating preQ$_1$ ($0.24 \pm 0.01 \times 10^6$ M^{-1}s^{-1}; Fig. 4b, Supplementary Table S6). This mutant I1S8, as well as I4S8, showed only a minimal effect of preQ$_1$ on the 30S binding kinetics (Supplementary Fig. 7). That is, when the aptamer is entirely decoupled (removing any overlap) and separated by 1 nt from the canonical 8-nt SD sequence, it represses 30S binding independently of preQ$_1$ as efficiently as the WT does only when fully folded with preQ$_1$. Since the aptamer's P2 helix forms only partially in the absence of preQ$_1$[29], this repression does not depend on occlusion of part of the SD sequence, but instead must arise from the steric hindrance the pre-folded, preQ$_1$-free (or apo) aptamer structure exerts on the 30S subunit. In support of this hypothesis, $k_{on,slow}$ values of $0.34 \pm 0.007 \times 10^6$ M^{-1}s^{-1} and $0.36 \pm 0.006 \times 10^6$ M^{-1}s^{-1} were recovered in the absence of preQ$_1$ once the SD was moved 6 and 4 nt apart from the SD in mutants I6S8 and I4S8, respectively (Fig. 4b, Supplementary Fig. 8 and Table S6), equivalent to the value of $0.33 \pm 0.02 \times 10^6$ M^{-1}s^{-1} for the WT without preQ$_1$. A mutant of intermittent 2-nt distance (I2S8) showed an intermediate increase of $k_{on,slow}$, also in the absence of preQ$_1$ (Fig. 4b, Supplementary Fig. 8 and Table S6), suggesting that the 30S subunit is highly sensitive to the precise distance of the partially formed P2 helix from the SD sequence. A rastergram plot confirmed this trend with an ~240% increase in short standby binding events and an ~7% increase in long binding events of I6S8 relative to I1S8, showing that as the SD sequence moved away from the steric hindrance of aptamer, it became more accessible for 30S to bind (Fig. 4c, Supplementary Table S7). Similarly, the 30S occupancy of single mRNA molecules was low with the aptamer placed proximal to

the SD sequence in I1S8 (11 ± 1 % H, 15 ± 1 % M, and 74 ± 1 % L occupancy; Fig. 4d) and increased

when the aptamer was more spaced out, as in I6S8 (24 ± 2 % H, 33 ± 2 % M and 43 ± 2 % L; Fig. 4H).

Taken together, these observations reveal that the pre-folded apo-aptamer represses 30S

association sterically during the earliest stage of translation initiation, whereas spacing the aptamer's

P2 helix as little as 4 nt from the SD sequence alleviates this steric hindrance of 30S binding entirely.

This finding reveals a surprisingly limited role of the SD sequence itself in effecting riboswitch-

mediated repression of translation initiation. Notably, the dissociation rate constants $k_{off,slow}$ of this

entire set of mutants from I1S8 to I6S8 in the absence of $preQ_1$ remains elevated, closer to that of the

WT in the presence of $preQ_1$ than its absence (Fig. 4b), suggesting that the mRNA-30S complex, once

formed, may be sensitive to secondary structure of the mRNA even further upstream.

To understand the role of the SD sequence, if any, we investigated our second set of mutants,

I4S8 to I4S0 (Fig. 4e, Supplementary Fig. 8 and Table S6). Using our SiM-KARB assay under equilibrium

conditions in the absence of $preQ_1$, we found I4S8's $k_{on,slow}$ value of 0.36 ± 0.006 × 10^6 $M^{-1}s^{-1}$ to

decrease to 0.26 ± 0.008 × 10^6 $M^{-1}s^{-1}$ for I4S0 as the number of complementary SD:anti-SD base pairs

decreased from eight to zero (Fig. 4f, Supplementary Fig. 8 and Table S8). This ~35% decrease in

binding rate constant, and gradual decrease at intermediate SD sequence lengths (Fig. 4f),

underscores that the mRNA's SD sequence is an important 30S recruiting factor. Notably, the 30S

subunit still binds to I4S0 even in the entire absence of the SD sequence, which corroborates the

notion that leaderless mRNAs can function in translation[42,43]. We also observed that the $k_{on,slow}$ for

I4S0 in the absence of $preQ_1$ comes close to that of the WT in its presence (Fig. 4f), providing

evidence that $preQ_1$ binding to the WT aptamer suppresses access to the SD sequence completely,

even though the 30S subunit can still bind elsewhere on our short mRNA. Of note, the dissociation

rate constants $k_{off,slow}$ of this set of mutants from I4S8 to I4S0 in the absence of $preQ_1$ again remains

closer to the elevated value of the WT in the presence of $preQ_1$ (Fig. 4f), consistent with the notion of

mRNA-30S complex sensitivity to proximal mRNA secondary structure. A comparison of the

rastergrams for I4S8 and I4S0 showed that, while the cleft-accommodated (long) binding events were

reduced by ~52%, the short standby bindings were reduced even more, by ~45% (Fig. 4g,

Supplementary Table S7). Finally, the fractional 30S occupancy histograms showed that the mRNA

accessibility progressively decreases from 17 ± 2 % H, 23 ± 2 % M, and 60 ± 2 % L population for I4S8

to 10 ± 2 % H, 8 ± 2 % M, and 82 ± 2 % L for I4S0 (Fig. 4h).

Taken together, our observations from mutations designed to systematically uncouple

aptamer and SD sequence effects indicate that[43] the absence of steric hindrance from proximal

secondary structure and a strong SD:anti-SD interaction are two main, independent 30S recruitment

factors. That is, while $preQ_1$ binding to the aptamer only sequesters two nucleotides of the SD

sequence, the riboswitch affects 30S recruitment through both partial occlusion of the SD sequence

upon ligand binding and, unexpectedly, its inherent secondary structure independent of ligand.

Ribosomal protein S1 unfolds the riboswitch to both facilitate and stabilize 30S binding

Our data so far have shown that $preQ_1$ binding to WT R-mRNA^{+30} reduces SD sequence accessibility

and thereby both slows and destabilizes 30S binding. Ribosomal protein S1 is known to enhance 30S

binding by many mRNAs with structured 5'-UTRs[46,47]. S1's three central domains bind and resolve

RNA structures, including that of the Tte preQ1 riboswitch[33], to help accommodate them into the

mRNA binding cleft (Fig. 5a)[48]. S1 loosely binds to the 30S ribosomal subunit and can be depleted

during purification (ΔS1-30S) and subsequently be reconstituted by adding recombinant S1 (Methods, Supplementary Fig. 9). This approach gives us as tool to test our SD accessibility model further.

Our SiM-KARB assay demonstrated that the ΔS1-30S subunit binds to R-mRNA^{+30} with diminished frequency and stability (Fig. 5b); stable binding was recovered only by reconstituting ΔS1-30S with purified S1 (Supplementary Fig. 10). Accordingly, the mRNA occupancy histograms showed very low accessibility of individual R-mRNA^{+30} molecules to ΔS1-30S, counteracted by S1 reconstitution (Fig. 5c). Plotting the H, M and L population fractions we found that the L population decreased from 70 ± 8 % for ΔS1-30S to 29 ± 4 % for WT-30S, while the M and L population increased from 9 ± 2 % and 21 ± 7 % for ΔS1-30S to 44 ± 3 % and 26 ± 5 %, respectively, for WT-30S (Fig. 5d). These findings suggest that 5'-UTR access is strongly facilitated by S1-mediated RNA unfolding.

Accordingly, we observed a $k_{on,slow}$ of 0.24 ± 0.02 × 10^6 M^{-1}s^{-1} for ΔS1-30S, ~1.4-fold lower than that for WT-30S. In contrast, the $k_{off,slow}$ of 0.01 ± 0.02 s^{-1} for ΔS1-30S was elevated ~2-fold over WT-30S (Fig. 5d). That is, not only does S1 unfold the 5'-UTR to facilitate initial 30S binding (i.e., increases $k_{on,slow}$), but it also stabilizes the mRNA-30S complex once formed (i.e., decreases $k_{off,slow}$). Stoichiometric (1:1) reconstitution of ΔS1-30S with recombinant S1 recovered particularly the $k_{off,slow}$ of WT-30S, whereas $k_{on,slow}$ is restored to a lesser degree (Fig. 5d), possibly because free S1 protein in solution coats the pseudoknot, thereby partially blocking 30S recruitment.

These results underscore, first, the important role of S1 protein in accommodating structured RNAs into the 30S subunit's mRNA binding cleft to initiate translation and, second, further support our model that proximal mRNA structure both antagonizes 30S recruitment and accelerates its dissociation.

Initiation factors only slightly facilitate 30S binding, independent of the 5'-UTR structure

Lastly, we explored how initiation factors (IFs) and initiator fMet-tRNA[fMet], known to be essential for translation initiation *in vivo*, impact 30S binding in the context of 5'-UTR structure[49,50]. While each IF governs a specific function in the initiation process[51,52], we asked to what extent they together influence 30S binding by adding recombinant IF1, IF2, and IF3, and fMet-tRNA[fMet] with the 30S into our SiM-KARB assay (Fig. 5e) and examining the kinetics of 30S binding to *R-mRNA[+30]*. We found that the fractional 30S occupancy in the presence of IFs and tRNA rose from 11 ± 3 % H, 26 ± 2 % M, and 63 ± 4 % L in the absence of preQ$_1$ to 1 ± 1 % H, 4 ± 2 % M, and 95 ± 2 % L at saturating preQ$_1$, not drastically different from that in the absence of IFs and tRNA (compare Figs. 2e and 5f). In the presence of IFs and tRNA[fMet], $k_{on,slow}$ increased to 0.39 ± 0.01 × 10^6 M^{-1}s^{-1} from 0.33 ± 0.04 × 10^6 M^{-1}s^{-1} in their absence; whereas upon addition of preQ$_1$, $k_{on,slow}$ only slightly increased in the presence of IFs and tRNA[fMet] to 0.27 ± 0.04 × 10^6 M^{-1}s^{-1} from 0.24 ± 0.03 × 10^6 M^{-1}s^{-1} in their absence; finally, $k_{off,slow}$ remained within error unchanged (Fig. 5g and Supplementary Table S13B).

In summary, we thus find that IFs and tRNA[fMet] slightly increase $k_{on,slow}$ over the values in their absence, equally in the presence and absence of preQ$_1$, while not significantly affecting $k_{off,slow}$. These observations indicate that the cofactors play a small, supportive role in the earliest stage of translation initiation as probed by 30S binding and exert it equally on the pre-folded apo-riboswitch and the fully folded preQ$_1$-bound riboswitch.

Discussion

By examining 30S-mRNA interactions at the earliest stage of translation initiation, we reveal here how a small, but strategically placed 5'-UTR structure can influence gene expression. We find that 30S

ribosomal subunit binding to the mRNA downstream of an H-type pseudoknotted riboswitch is controlled both by the RNA's secondary structure, independently of ligand binding to the riboswitch aptamer, and by the occlusion of the SD sequence upon ligand binding. Moreover, $preQ_1$ binding itself exerts two distinct effects; first, antagonizing 30S recruitment (i.e., slowing down 30S binding) and, second, destabilizing the 30S-mRNA complex once formed (i.e., accelerating 30S dissociation), giving the mRNA multiple layers of control over translation. Our mutational studies highlight the importance of the SD:anti-SD base pairing to initiate translation[53,54], although it turns out to be only one of several mechanisms by which a translational riboswitch governs gene expression of its hosting mRNA. Such a multi-pronged effect leverages binding of a small ligand into efficient regulation of translation initiation by the much larger bacterial 30S subunit.

Upon decoupling secondary structure and SD sequence, we further find that the precise distance of the aptamer from the SD sequence is a key determinant of initial recruitment of the 30S subunit, with a distance of as small as 4 nt eliminating the aptamer's influence on 30S binding. While the cleft-accommodated mRNA is footprinted from nucleotides (nt) −18 to +10 on the 30S subunit, exceeding the SD sequence by ~10 nt on the 5' side, architecturally the segment of mRNA upstream of the SD sequence is rather solvent exposed[12], rationalizing how steric clashes rapidly diminish on a length scale shorter than the footprint.

Perhaps most strikingly, we find clear evidence for short-lived standby site exploration, sometimes followed by unfolding of the riboswitch for a longer-lived mRNA cleft accommodation, leading to the model in Fig. 6a. Consequently, the riboswitch ligand and ribosomal protein S1 counteract in that $preQ_1$ tightens the pseudoknot fold while S1 unties it to enhance mRNA cleft accommodation (Fig. 6a). Based on our kinetic SiM-KARB measurements, we then can predict a free

17

energy model for the non-equilibrium unfolding of the riboswitch (Fig. 6b, Supplementary Fig. 12). The initial binding energy into the standby site is very similar in the absence and presence of preQ$_1$ (-3.6 kcal/mol). Successive mRNA accommodation by unfolding the mRNA is a spontaneous process yielding -0.5 kcal/mol energy in the absence of preQ$_1$, whereas with preQ$_1$ bound the 5'-UTR is more stably folded and requires +0.9 kcal/mol to unfold. The net energetic penalty for cleft-accommodation of the mRNA in the presence of preQ$_1$ is thus 1.4 kcal/mol. This small penalty – equivalent to a couple of hydrogen bonds – is a reflection of the exchange of two base pairs in helix P2 of the riboswitch for two base pairs of SD:anti-SD interaction (Fig. 1). Due to the multi-pronged effects of the riboswitch on both 30S binding and dissociation, this seemingly small effect suffices for the riboswitch ligand to have significant leverage over translation initiation (Fig. 6a).

The leveraging we discovered here of a local, small-scale binding event into the global, large-scale functional process of translation initiation appears to be a recurring theme in RNA biology. For example, in a recent study of a Mn^{2+} sensing riboswitch, a single metal ion binding at the docking core of an RNA four-way junction was shown to alter the global structure to affect transcription[55]. Furthermore, highly structured motifs around the RBS of an mRNA often restrict 30S recruitment to control translation efficiency[7,11]. The classical model of translation initiation on structured mRNAs invoked an equilibrium thermodynamic process, wherein the 5'-UTR is remodeled into low-energy structures by successive unfolding and refolding that then are accommodated into the mRNA binding cleft of the 30S subunit[56]. Recent transcriptome-wide structural studies have refined this model and shown that a non-equilibrium kinetic competition between mRNA unfolding and 30S dissociation governs translation efficiency[7]. We demonstrate here that this kinetic competition provides leverage to riboswitches to modulate translation efficiency, even in cases where a small ligand (such as preQ$_1$)

only provides limited thermodynamic power. Once an mRNA has repelled a 30S subunit, the competition between mRNAs in the bacterial cell dictates that free subunits initiate translation elsewhere, whereas the unoccupied mRNA may fall prey to ribonucleases[32].

In bacteria, translation efficiency as dictated by secondary structure around the RBS plays a significant role in determining the fate of an mRNA[57-59]. As initiation complexes form, commitment to translation increases while the vulnerability to mRNA decay decreases, as mediated by the bacterial degradosome, Hfq, and Rho[32,60-62]. Consequently, the evolutionary pressure is high to evolve a fine-tuned kinetic modulation of initiation complex formation as a key regulator of global translational activity. The insights presented here will therefore provide an opening for the development of novel antibacterial drugs that derail this finely tuned machinery.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at…

Author contributions

S.R. and N.G.W. designed the study. S.R., S.D. and S.C. performed the experiments and analyzed the data. S.R., S.D., S.C. and N.G.W. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at …

Reprints and permissions information is available at www.nature.com/reprints.

Code availability statement: All custom codes used in this study are available from the corresponding author upon reasonable request.

Data availability statement: The data that support the findings of this study are available from the corresponding author upon reasonable request.

Correspondence and requests for materials should be addressed to N.G.W. (nwalter@umich.edu).

References

1. Araujo, P.R. et al. Before It Gets Started: Regulating Translation at the 5' UTR. *Comp Funct Genomics* **2012**, 475731 (2012).
2. Dvir, S. et al. Deciphering the rules by which 5'-UTR sequences affect protein expression in yeast. *Proc Natl Acad Sci U S A* **110**, E2792-801 (2013).
3. Leppek, K., Das, R. & Barna, M. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat Rev Mol Cell Biol* **19**, 158-174 (2018).
4. Braun, F., Durand, S. & Condon, C. Initiating ribosomes and a 5'/3'-UTR interaction control ribonuclease action to tightly couple B. subtilis hbs mRNA stability with translation. *Nucleic Acids Res* **45**, 11386-11400 (2017).
5. Melnikov, S. et al. One core, two shells: bacterial and eukaryotic ribosomes. *Nat Struct Mol Biol* **19**, 560-7 (2012).
6. Na, D., Lee, S. & Lee, D. Mathematical modeling of translation initiation for the estimation of its efficiency to computationally design mRNA sequences with desired expression levels in prokaryotes. *BMC Syst Biol* **4**, 71 (2010).
7. Mustoe, A.M. et al. Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. *Cell* **173**, 181-195.e18 (2018).
8. Tuller, T. & Zur, H. Multiple roles of the coding sequence 5' end in gene expression regulation. *Nucleic Acids Res* **43**, 13-28 (2015).
9. Andreeva, I., Belardinelli, R. & Rodnina, M.V. Translation initiation in bacterial polysomes through ribosome loading on a standby site on a highly translated mRNA. *Proc Natl Acad Sci U S A* **115**, 4411-4416 (2018).
10. de Smit, M.H. & van Duin, J. Translational Standby Sites: How Ribosomes May Deal with the Rapid Folding Kinetics of mRNA. *Journal of Molecular Biology* **331**, 737-743 (2003).
11. Mustoe, A.M., Corley, M., Laederach, A. & Weeks, K.M. Messenger RNA Structure Regulates Translation Initiation: A Mechanism Exploited from Bacteria to Humans. *Biochemistry* **57**, 3537-3539 (2018).
12. Culver, G.M. Meanderings of the mRNA through the ribosome. *Structure* **9**, 751-8 (2001).
13. Shine, J. & Dalgarno, L. Terminal-sequence analysis of bacterial ribosomal RNA. Correlation between the 3'-terminal-polypyrimidine sequence of 16-S RNA and translational specificity of the ribosome. *Eur J Biochem* **57**, 221-30 (1975).
14. Nudler, E. & Mironov, A.S. The riboswitch control of bacterial metabolism. *Trends Biochem Sci* **29**, 11-7 (2004).
15. Breaker, R.R. Prospects for riboswitch discovery and analysis. *Mol Cell* **43**, 867-79 (2011).
16. Winkler, W.C., Nahvi, A., Sudarsan, N., Barrick, J.E. & Breaker, R.R. An mRNA structure that controls gene expression by binding S-adenosylmethionine. *Nat Struct Biol* **10**, 701-7 (2003).
17. Winkler, W., Nahvi, A. & Breaker, R.R. Thiamine derivatives bind messenger RNAs directly to regulate bacterial gene expression. *Nature* **419**, 952-6 (2002).
18. Barrick, J.E. & Breaker, R.R. The distributions, mechanisms, and structures of metabolite-binding riboswitches. *Genome Biol* **8**, R239 (2007).
19. Scharff, L.B., Childs, L., Walther, D. & Bock, R. Local absence of secondary structure permits translation of mRNAs that lack ribosome-binding sites. *PLoS genetics* **7**, e1002155-e1002155 (2011).

20. Winkler, W.C., Nahvi, A., Roth, A., Collins, J.A. & Breaker, R.R. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281-6 (2004).

21. Winkler, W.C. & Breaker, R.R. Regulation of bacterial gene expression by riboswitches. *Annu Rev Microbiol* **59**, 487-517 (2005).

22. Mandal, M. & Breaker, R.R. Gene regulation by riboswitches. *Nat Rev Mol Cell Biol* **5**, 451-63 (2004).

23. Fuchs, R.T., Grundy, F.J. & Henkin, T.M. S-adenosylmethionine directly inhibits binding of 30S ribosomal subunits to the SMK box translational riboswitch RNA. *Proc Natl Acad Sci U S A* **104**, 4876-80 (2007).

24. Haller, A., Altman, R.B., Souliere, M.F., Blanchard, S.C. & Micura, R. Folding and ligand recognition of the TPP riboswitch aptamer at single-molecule resolution. *Proc Natl Acad Sci U S A* **110**, 4188-93 (2013).

25. Eichhorn, C.D., Kang, M. & Feigon, J. Structure and function of preQ(1) riboswitches. *Biochim Biophys Acta* **1839**, 939-950 (2014).

26. Kim, J.N. & Breaker, R.R. Purine sensing by riboswitches. *Biol Cell* **100**, 1-11 (2008).

27. Rinaldi, A.J., Lund, P.E., Blanco, M.R. & Walter, N.G. The Shine-Dalgarno sequence of riboswitch-regulated single mRNAs shows ligand-dependent accessibility bursts. *Nat Commun* **7**, 8976 (2016).

28. Jenkins, J.L., Krucinska, J., McCarty, R.M., Bandarian, V. & Wedekind, J.E. Comparison of a preQ1 riboswitch aptamer in metabolite-bound and free states with implications for gene regulation. *J Biol Chem* **286**, 24626-37 (2011).

29. Suddala, K.C. et al. Single transcriptional and translational preQ$_1$ riboswitches adopt similar pre-folded ensembles that follow distinct folding pathways into the same ligand-bound structure. *Nucleic Acids Res* **41**, 10462-75 (2013).

30. Liberman, J.A., Bogue, J.T., Jenkins, J.L., Salim, M. & Wedekind, J.E. ITC analysis of ligand binding to preQ$_1$ riboswitches. *Methods Enzymol* **549**, 435-50 (2014).

31. Coppins, R.L., Hall, K.B. & Groisman, E.A. The intricate world of riboswitches. *Current opinion in microbiology* **10**, 176-181 (2007).

32. Ray, S., Chauvier, A. & Walter, N.G. Kinetics coming into focus: single-molecule microscopy of riboswitch dynamics. *RNA Biology* **16**, 1077-1085 (2019).

33. Lund, P.E., Chatterjee, S., Daher, M. & Walter, N.G. Protein unties the pseudoknot: S1-mediated unfolding of RNA higher order structure. *Nucleic Acids Res* (2019).

34. Suddala, K.C. & Walter, N.G. Riboswitch structure and dynamics by smFRET microscopy. *Methods Enzymol* **549**, 343-73 (2014).

35. Rio, D.C. Filter-binding assay for analysis of RNA-protein interactions. *Cold Spring Harb Protoc* **2012**, 1078-81 (2012).

36. Dorywalska, M. et al. Site-specific labeling of the ribosome for single-molecule spectroscopy. *Nucleic Acids Res* **33**, 182-9 (2005).

37. Blanchard, S.C., Kim, H.D., Gonzalez, R.L., Jr., Puglisi, J.D. & Chu, S. tRNA dynamics on the ribosome during translation. *Proc Natl Acad Sci U S A* **101**, 12893-8 (2004).

38. Francis, K. & Kohen, A. Standards for the reporting of kinetic isotope effects in enzymology. *Perspectives in Science* **1**, 110-120 (2014).

39. Li, A., Ziehr, J.L. & Johnson, K.A. A new general method for simultaneous fitting of temperature and concentration dependence of reaction rates yields kinetic and

thermodynamic parameters for HIV reverse transcriptase specificity. *J Biol Chem* **292**, 6695-6702 (2017).

40. Spitzer, J. & Poolman, B. The Role of Biomacromolecular Crowding, Ionic Strength, and Physicochemical Gradients in the Complexities of Life's Emergence. *Microbiology and Molecular Biology Reviews* **73**, 371-388 (2009).

41. Qu, X. et al. Single-molecule nonequilibrium periodic Mg^{2+}-concentration jump experiments reveal details of the early folding pathways of a large RNA. *Proceedings of the National Academy of Sciences of the United States of America* **105**, 6602-6607 (2008).

42. Nakagawa, S., Niimura, Y., Miura, K.-i. & Gojobori, T. Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proceedings of the National Academy of Sciences* **107**, 6382-6387 (2010).

43. Omotajo, D., Tate, T., Cho, H. & Choudhary, M. Distribution and diversity of ribosome binding sites in prokaryotic genomes. *BMC Genomics* **16**, 604 (2015).

44. Romero, D.A. et al. A comparison of key aspects of gene regulation in Streptomyces coelicolor and Escherichia coli using nucleotide-resolution transcription maps produced in parallel by global and differential RNA sequencing. *Mol Microbiol* (2014).

45. Beck, H.J. & Moll, I. Leaderless mRNAs in the Spotlight: Ancient but Not Outdated! *Microbiol Spectr* **6**(2018).

46. Duval, M. et al. Escherichia coli ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. *PLoS Biol* **11**, e1001731 (2013).

47. Studer, S.M. & Joseph, S. Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Mol Cell* **22**, 105-15 (2006).

48. Salah, P. et al. Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic Acids Res* **37**, 5578-88 (2009).

49. Fei, J. et al. Chapter 12 - A Highly Purified, Fluorescently Labeled In Vitro Translation System for Single-Molecule Studies of Protein Synthesis. in *Methods in Enzymology*, Vol. 472 (ed. Walter, N.G.) 221-259 (Academic Press, 2010).

50. Tsai, A. et al. Heterogeneous pathways and timing of factor departure during translation initiation. *Nature* **487**, 390-393 (2012).

51. MacDougall, D.D. & Gonzalez, R.L., Jr. Translation initiation factor 3 regulates switching between different modes of ribosomal subunit joining. *J Mol Biol* **427**, 1801-18 (2015).

52. Wang, J., Caban, K. & Gonzalez, R.L., Jr. Ribosomal initiation complex-driven changes in the stability and dynamics of initiation factor 2 regulate the fidelity of translation initiation. *Journal of molecular biology* **427**, 1819-1834 (2015).

53. Saito, K., Green, R. & Buskirk, A.R. Translational initiation in E. coli occurs at the correct sites genome-wide in the absence of mRNA-rRNA base-pairing. *Elife* **9**(2020).

54. Tollerson, R., 2nd & Ibba, M. Translational regulation of environmental adaptation in bacteria. *J Biol Chem* **295**, 10434-10445 (2020).

55. Suddala, K.C. et al. Local-to-global signal transduction at the core of a Mn(2+) sensing riboswitch. *Nat Commun* **10**, 4304 (2019).

56. Kozak, M. Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **361**, 13-37 (2005).

57. Del Campo, C., Bartholomaus, A., Fedyunin, I. & Ignatova, Z. Secondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA Regulation and Function. *PLoS Genet* **11**, e1005613 (2015).

58. Passalacqua, K.D. et al. Structure and complexity of a bacterial transcriptome. *J Bacteriol* **191**, 3203-11 (2009).

59. Chiaruttini, C. & Guillier, M. On the role of mRNA secondary structure in bacterial translation. *Wiley Interdiscip Rev RNA*, e1579 (2019).

60. Bastet, L. et al. Translational control and Rho-dependent transcription termination are intimately linked in riboswitch regulation. *Nucleic Acids Res* **45**, 7474-7486 (2017).

61. Azam, M.S. & Vanderpool, C.K. Translational regulation by bacterial small RNAs via an unusual Hfq-dependent mechanism. *Nucleic Acids Res* **46**, 2585-2599 (2018).

62. Caron, M.-P. et al. Dual-acting riboswitch control of translation initiation and mRNA decay. *Proceedings of the National Academy of Sciences of the United States of America* **109**, E3444-E3453 (2012).

63. Leontis, N.B. & Westhof, E. Geometric nomenclature and classification of RNA base pairs. *RNA* **7**, 499-512 (2001).
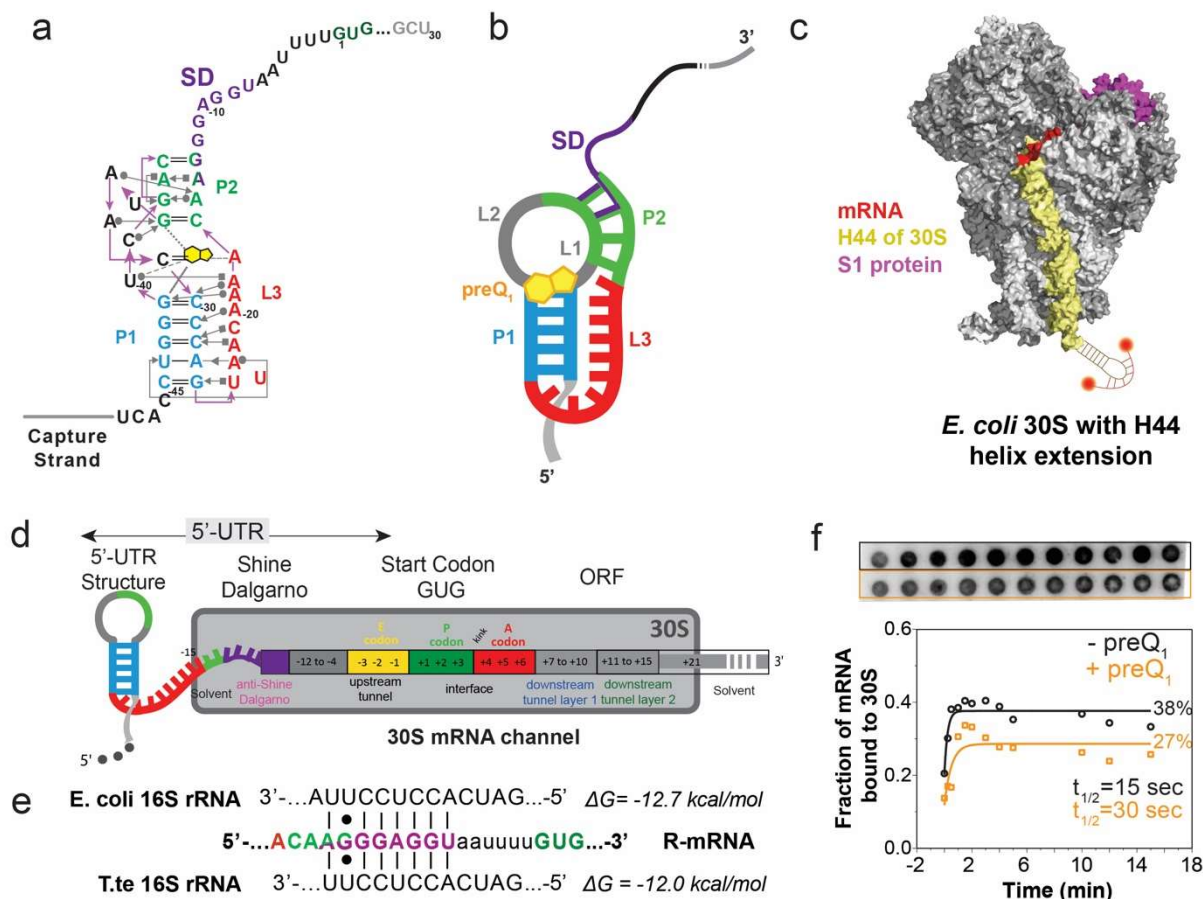
Figure 1: Schematic of a model mRNA with *Tte*-riboswitch at its 5'-UTR to study translation initiation. (a) Structural map of the preQ$_1$

translational riboswitch from *Tte* displayed with Leontis–Westhof notations[63]. The SD sequence (purple) partially overlaps the P2 stem

nucleotides (green). (b) Schematic diagram of the preQ$_1$ bound riboswitch containing mRNA (termed R-mRNA) from *T. tengcongensis* (*Tte*). (c) Schematic representation of the mutated ribosome used here, containing a hairpin extension at the helix-44. The extension is hybridized with a dual-labeled complementary DNA oligonucleotide (red). (d) Schematic diagram of *Tte* R-mRNA and its expected occupancy in the mRNA channel of the 30S subunit during initiation, where +1 is the first nucleotide of the open reading frame. The mRNA channel representation is adapted from Ref. 12. (e) Base-pairing interactions between the Shine-Dalgarno sequence of our *Tte* mRNA transcript and the 16S rRNA of *E. coli* or *Tte*. (f) Autoradiograph of a filter binding membrane measuring the efficiency of 30S-mRNA complex formation as a function of time at zero and 10 μM preQ$_1$ concentrations (top). The resulting fraction of bound mRNA is plotted as a function of time (bottom).
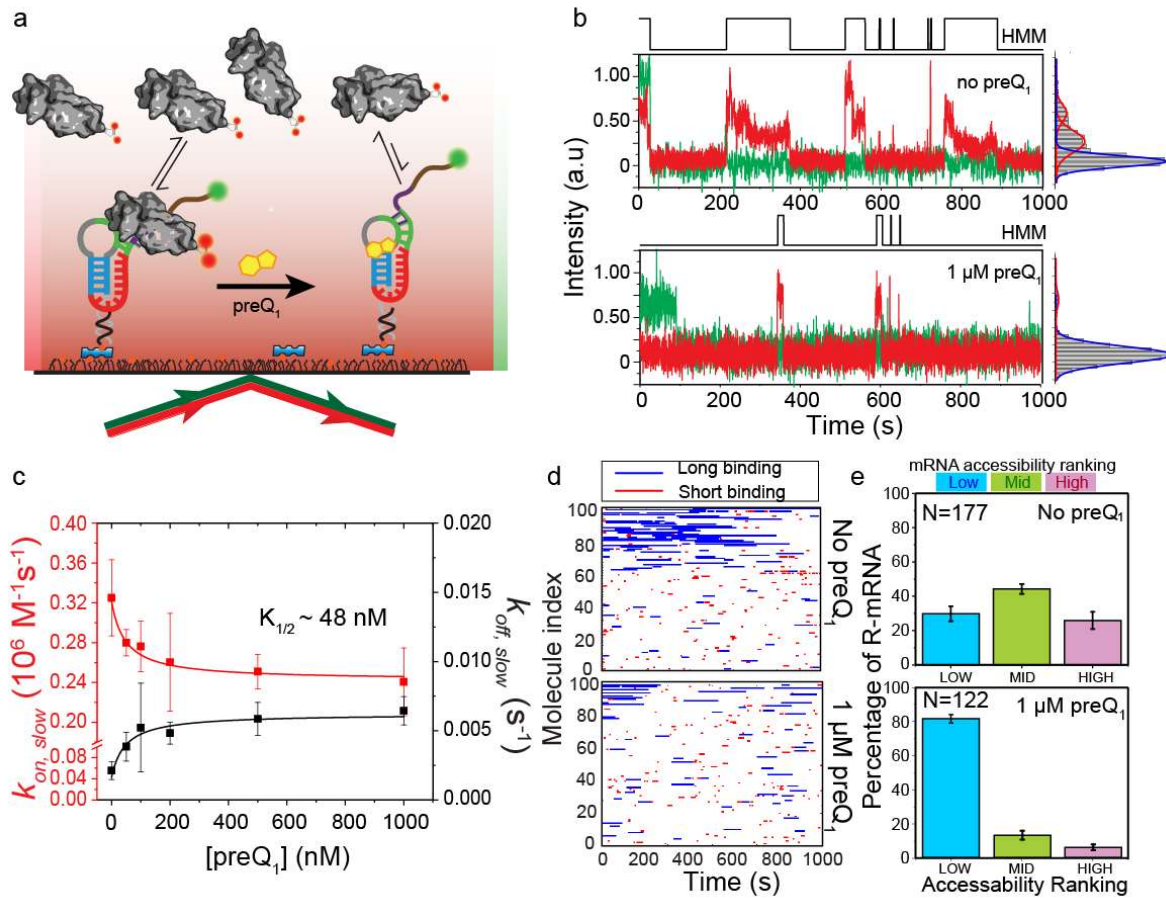
Figure 2: preQ₁-dependent 30S binding to R-mRNA at the single-molecule level. (a) Schematic representation of our SiM-KARB assay. 3'-Cy3 labeled R-mRNA molecules are immobilized on a slide surface. Repeated binding and dissociation of 30S subunits (dual-labeled with Cy5) is monitored through the co-localization of Cy3 and Cy5 fluorescence. (b) Representative R-mRNA^{+30} (green) and 30S (red) fluorescence intensity versus time trajectories and corresponding intensity histograms for a single R-mRNA^{+30} molecule in the absence (top panel) or presence (bottom panel) of preQ₁. HMM idealizations to a two-state model (black, top) are plotted as a function of time. (c) 30S-binding ($k_{on,slow}$, red) and dissociation ($k_{off,slow}$; black) rate constants were determined from exponential fits of

dwell times in the unbound and bound states, respectively, as a function of $preQ_1$ concentration. The corresponding half-saturation ($K_{1/2}$) values were determined by Hill equation fitting with Hill coefficients of -1 and +1, respectively. (d) Rastergram of 100 randomly selected traces of 30S binding to individual R-mRNA^{+30} molecules in the absence (top panel) and presence (bottom panel) of $preQ_1$. 30S binding events on R-mRNA^{+30} are represented in either red (short or standby events) and blue (long or cleft-accommodated events). (e) Histograms of the fractional bound time for which individual R-mRNA^{+30} molecules are occupied by 30S subunits in the absence (top) or presence (bottom) of $preQ_1$. All 30S-bound R-mRNA^{+30} molecules were empirically ranked as low (<0.1 fractional bound time, cyan), mid (0.1 to 0.2 fractional bound-time, green), or high (>0.2 fractional bound time, pink) in 30S occupancy. N represents the total number of molecules included for each condition.
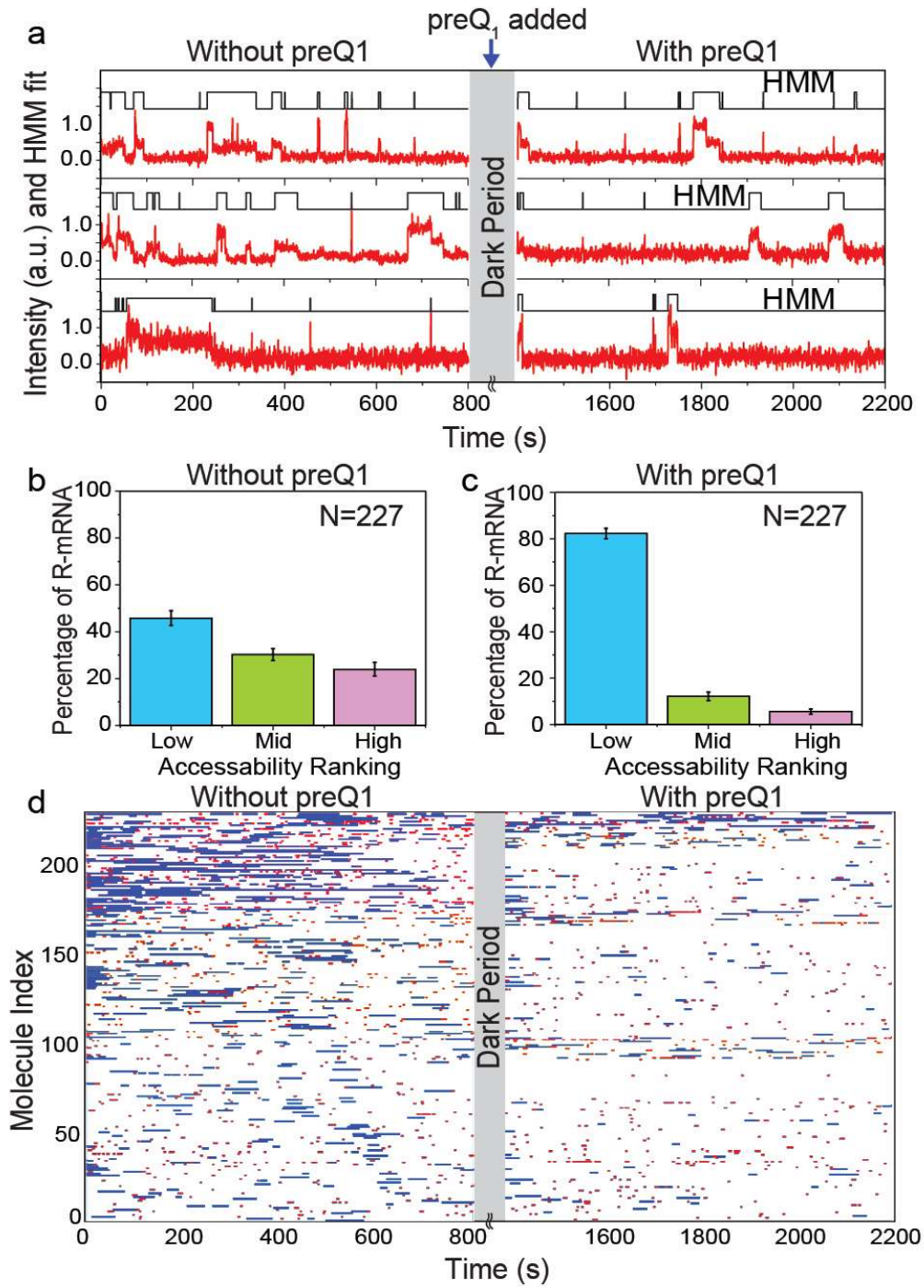
Figure 3: Single mRNA molecules undergo conformational switching upon introduction of preQ₁

ligand in situ. (a) Exemplary single-molecule trajectories from non-equilibrium ligand-jump

experiments composed of segments before and after the jump. 30S binding to the same set of individual R-mRNA^{+30} was monitored first in the absence of preQ$_1$ (left, -preQ$_1$), then in the presence of 1 μM preQ$_1$ (right, +preQ$_1$). The gray axis break represents a 600 sec dark period between segments during which the buffer was exchanged. (b,c) Histograms of the fraction of time individual R-mRNA^{+30} molecules are bound by 30S before (b) and after (b) preQ$_1$ is introduced. Most molecules shift towards lower occupancy upon introduction of preQ$_1$. (d) Rastergram displaying the 30S binding behaviors of individual R-mRNA^{+30} molecules before and after introducing preQ$_1$. 30S binding events on each R-mRNA^{+30} are represented in red (short or standby events) or blue bar (long or cleft-accommodated events).
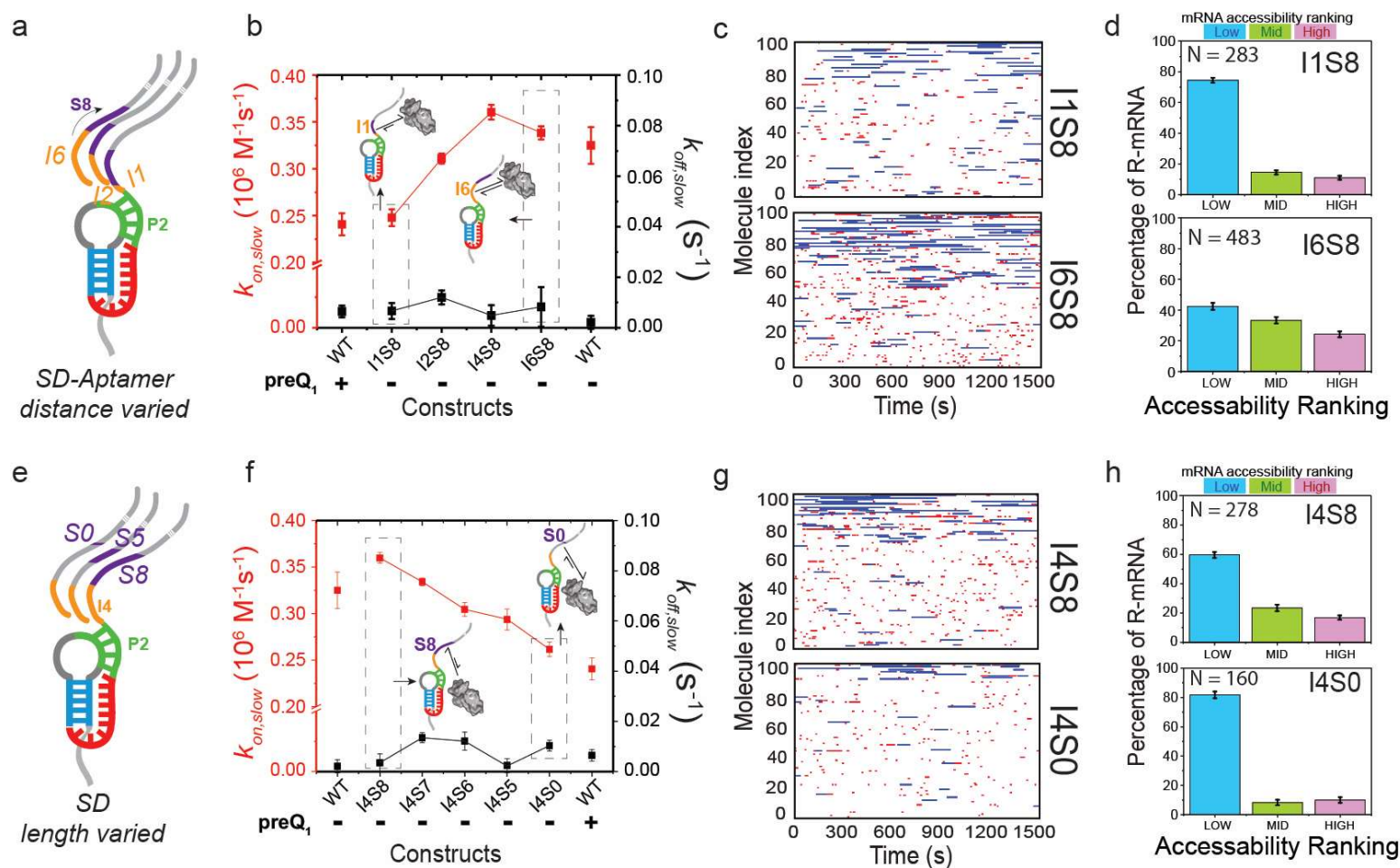
Figure 4: Mutation of the 5'-UTR structure by varying SD-aptamer distance (I) and SD length (S) affects 30S binding. (a) Schematic

representation for varying the SD-aptamer distance from insert length of one base (I1) to six bases (I6). (b) 30S-binding ($k_{on,slow}$, red) and

dissociation ($k_{off,slow}$, black) rate constants for different SD-aptamer distances. Data for WT R-mRNA with and without preQ$_1$ are shown for reference. (c) Rastergram of 100 randomly selected traces of individual mutant R-mRNA^{+30} molecules with a 1-nt (I1S8, top panel) or 6-base insert (I6S8, bottom panel). 30S binding is represented in red (short binding events or standby events) or blue (long binding events or cleft-accommodated events). (d) Histogram of the fractional bound time for individual I1S8 (top) and I6S8 (bottom) mutant R-mRNA molecules. (e) Schematic representation for varying the SD:anti-SD interaction length from full complementarity (S8) to no complementarity (S0) with a fixed number of 4 nt (I4) between the aptamer and SD sequence. (f) 30S binding ($k_{on,slow}$, red) and dissociation ($k_{off,slow}$, black) rate constants for different SD:anti-SD complementarities. (g) Rastergram of 100 randomly selected traces of individual mutant R-mRNA^{+30} molecules with fully available SD (I4S8, top panel) or no SD (I4S0, bottom panel). (h) Histogram of the fractional bound time for individual I4S8 (top) and I4S0 (bottom) mutant R-mRNA molecules.
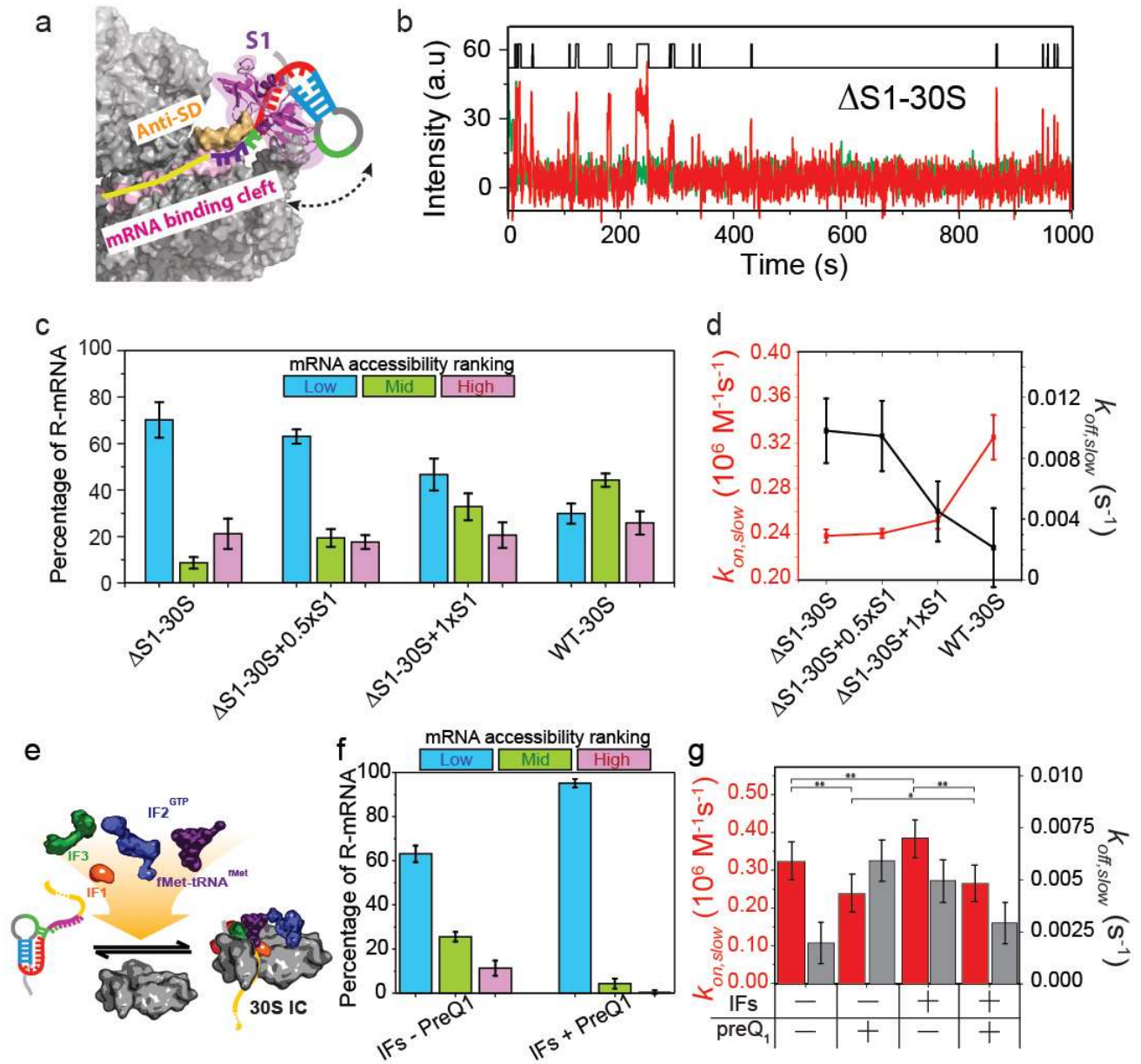
Figure 5: S1 mediates unfolding of the 5'-UTR to enhance and stabilize 30S binding. (a) Schematic of

S1 mediated unfolding of the 5'-UTR of R-mRNA[+30]. (b) Representative R-mRNA[+30] (green) and

associated ΔS1-30S (red) binding time trajectory. An HMM idealization to a two-state model (black,

top) is plotted as a function of time. (c) Histogram of the fractional bound time for individual R-mRNA

molecules occupied by for depleted ΔS1-30S, ΔS1-30S with 50% stoichiometric purified S1 (ΔS1-30S+0.5xS1), ΔS1-30S with fully stoichiometric purified S1 (ΔS1-30S+1xS1), and WT-30S. (d) Corresponding binding ($k_{on,slow}$, red) and dissociation ($k_{off,slow}$, black) rate constants. Data for WT R-mRNA without preQ$_1$ are shown for comparison. (e) Schematic of pre-initiation complex formation in the presence of initiation factors and initiator tRNA. (f) Histogram of the fractional bound-time for individual R-mRNA^{+30} in the presence of all initiation factors and fMet-tRNA^{fMet}, in the absence (top) and presence (bottom) of preQ$_1$. (g) Comparison of binding ($k_{on,slow}$) and dissociation ($k_{off,slow}$) rate constants in the presence or absence of initiation factors fMet-tRNA^{fmet} and preQ$_1$, as indicated. Statistical significance was determined by student t-test as *$p<10^{-2}$; **$p<10^{-4}$.
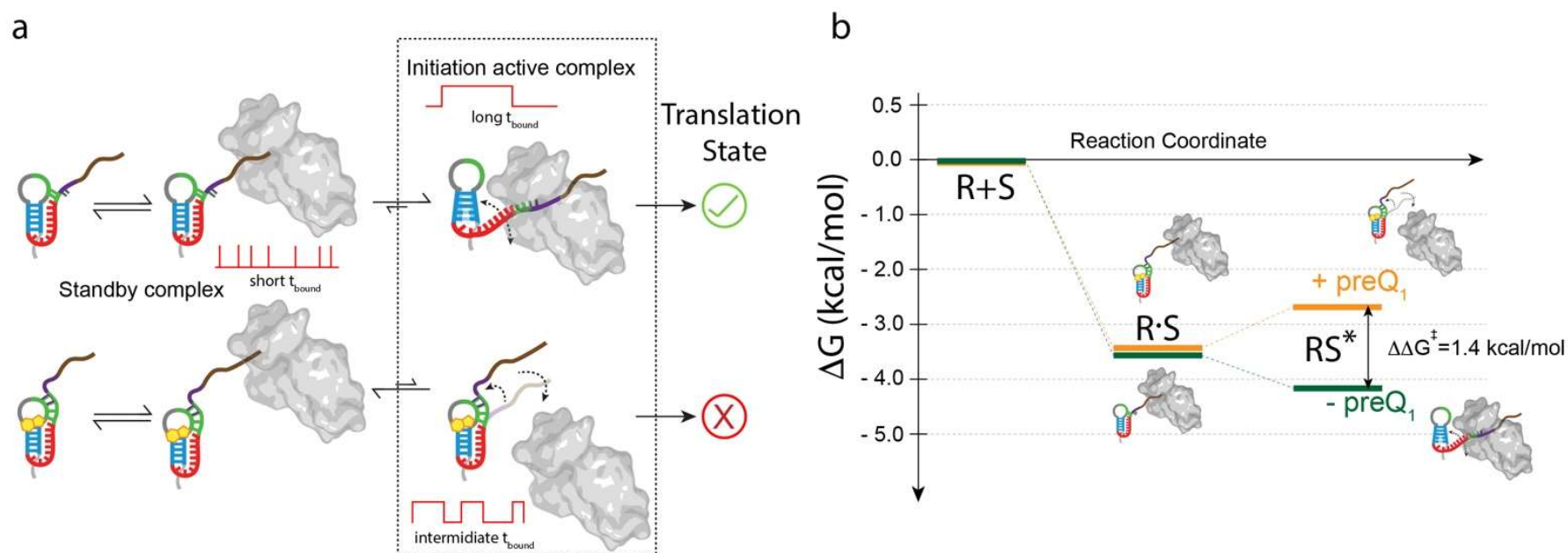
Figure 6: Model for the earliest stage of translation initiation in dependence of an embedded riboswitch structure. (a) Independent of

the preQ$_1$, the 30S dynamically interacts with the R-mRNA scanning for the RBS (and SD sequence) to facilitate translation initiation on the

nascent mRNA. Short, transient 30S binding is categorized as standby binding, whereas long binding events represent mRNA

accommodation with established SD:anti-SD interactions, required to initiate translation. In the absence of preQ$_1$, the SD region is more

accessible, favoring complete accommodation of the mRNA through correct SD:anti-SD interactions and more stable 30S binding. By

contrast, in the presence of preQ$_1$ the SD sequence is partially sequestered, leading to more standby site binding of the mRNA on the 30S subunit and to more frequently disrupted mRNA accommodation, both resulting in unsuccessful translation initiation. (b) Free energy diagram calculated from the 30S binding and dissociation rate constants to R-mRNA^{+30} in the absence or presence of preQ$_1$. An energetic penalty of ~1.4 kcal/mol for formation of the mRNA cleft accommodated complex derives from the dynamic unfolding of the 5'-UTR, necessary for a complete accommodation of the mRNA into the 30S subunit.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- RayetalSupportingInfo.pdf