Fitness landscapes for an FMN riboswitch in *S. pneumoniae* reveals complex interplay between antibiotic inhibitors and the native ligand as well as insights into the mechanism of action

Rebecca J. Korn and Michelle M. Meyer*

Boston College Biology Department, Chestnut Hill MA 02467

*corresponding author: m.meyer@bc.edu

ABSTRACT
RNA-based control of gene expression is common across all domains of life, yet shows extraordinary variety in terms of environmental cues recognized, mechanisms of action, and magnitude of gene expression changes. Furthermore, the fitness benefit conferred to an organism by RNA-induced regulation is often subtle and challenging to measure. In this work we measure the relative fitness conferred by all single mutants of an FMN riboswitch to the opportunistic pathogen *Streptococcus pneumoniae* under five distinct conditions. We find that in standard culture conditions the RNA sequence is not under strong selection, however the dynamic conditions encountered during mouse lung infection reveal an element that is mutationally robust, but also under weak purifying selection. Furthermore, fitness landscapes measured in the presence of two different targeting antibiotics reveal suites of mutations that confer constitutive activation or repression, which in turn provide insights into the mechanism of action for this RNA with lessons broadly applicable to interpretation of bacterial transcriptomics.

INTRODUCTION
The interplay between genotype and phenotype is central to understanding how organisms adapt to their environments, evolve new traits, and respond to external challenges. In bacteria, regulation that occurs after transcription initiation plays an important role in allowing rapid adjustment to environmental changes. Regulatory RNAs such as trans-acting small RNAs (sRNAs) and cis-active elements, including riboswitches and RNA thermometers, modulate gene expression in response to environmental cues, tuning metabolic processes, inducing stress responses, and activating resistance mechanisms[1]. Riboswitches, RNAs that respond to specific small molecule metabolites to alter gene expression, often control essential metabolic pathways and transporters. These RNAs act as molecular sensors that directly integrate metabolic cues with regulatory features to allow conditional intrinsic transcription termination or Shine-Dalgarno access. Riboswitches have emerged as potential targets for antimicrobials[1–8], and due to their dynamic nature and relatively small size, have also proven to be valuable biophysical model systems for studying the relationships between RNA sequence, structure, and function *in vitro*[9,10]. However, our understanding of how sequence changes in such RNAs drive functional changes in regulation within the cell, and how these regulatory shifts ultimately affect an organism's phenotype and fitness in diverse or dynamic environments, remains limited.

To better understand how changes in ncRNA sequence influence regulatory function and organism phenotype, we generated fitness landscapes for a *Streptococcus pneumoniae* flavin mononucleotide (FMN) riboswitch[11]. *S. pneumoniae* is a significant global health concern as the leading cause of lower respiratory infection worldwide, contributing to over 1 million deaths globally in 2016[12]. In the United States, antibiotic-resistant pneumococcal diseases cause around 1.2 million illnesses and 7,000 deaths annually[13]. There are two FMN riboswitches in *S. pneumoniae* that function as cis-regulatory "off" switches; one preceding the biosynthesis operon (SP_0178-SP_0175, *rib B/A,D,E,H*) and the other preceding the gene SP_0488, which encodes the riboflavin transport protein RibU. Riboflavin (vitamin B_2) is an essential nutrient, and its derivatives, FMN and flavin adenine dinucleotide (FAD), serve as critical cofactors in cellular metabolism, catalyzing redox reactions[1]. While FMN biosynthesis is dispensable for *S. pneumoniae* during *in vivo* infection due to the availability of FMN in the environment, the fitness of strains with transposon insertions into *ribU* is severely attenuated with *in vivo* mouse models[14]. In addition, several different antimicrobial compounds specifically target FMN riboswitches, ultimately inhibiting both FMN synthesis and transport[15–21].

Fitness landscapes measured for other RNAs have yielded invaluable insights into the RNA folding landscape[22], the effects of epistasis[23], the degree to which fitness peaks are isolated from one another[24–26], and the impacts of environmental perturbations on RNA function[27,28]. However, typically such studies were conducted *in vitro* to enable assessment of large number of

variants[22,24–28] or utilized engineered genetic selection schemes to partition functional and non-functional sequences in the context of the cell[23]. In this work, we focus on a large, dynamic RNA in its native context. We assess not the fitness of the RNA sequence, but rather the fitness it confers upon *S. pneumoniae*. By forgoing the insights to be obtained from larger libraries or stronger selective pressures, we focus on those that can only be gained from connecting RNA sequence to organismal phenotype in both a native-like infection setting and within antibiotic perturbed conditions.

In this study, we employed a *S. pneumoniae* FMN riboswitch mutant strain library containing all single-mutations to analyze the relationship between primary sequence, regulatory function, and organism fitness in five distinct environments. We measured the fitness of all point mutations within *in vitro* culture in the presence and absence of riboflavin, as well as in the context of an *in vivo* mouse infection model to demonstrate the robustness of *S. pneumoniae* fitness to single mutations within the RNA, as well as observe weak purifying selection throughout the sequence. Furthermore, antibiotic perturbation of the *in vitro* environment with antibiotics identified suites of mutations with constitutively active and repressed behavior, which in turn enable insights into the RNA mechanism of action. Our study uniquely assesses a regulatory RNA function within its native context to highlight the complex interplay between environment, selective pressure on the organism from the environment, and potential phenotypes accessible via single mutations.

RESULTS

Single mutations are neutral in chemically defined medium without riboflavin

The *S. pneumoniae* FMN riboswitch preceding *ribU* spans 372 nucleotides, from the transcription start site to a transcription termination site just upstream of the translational start[29]. The switch includes a canonical six-stem FMN-binding aptamer with an extended P3 stem, followed by an extended stem following the aptamer that appears to enable intrinsic transcription termination, although the base-pairing pattern required to connect ligand binding to termination is unclear **(Fig. 1F)**. To assess the impact of each of the 1,116 possible single-point mutants on *S. pneumoniae* fitness, we created a library of *S. pneumoniae* strains. We divided the riboswitch sequence into six segments, each of which was synthesized as a doped oligo where each position contained 1% of each alternative base. Individual doped oligos were assembled with unmutated 5' and 3' homology arms and an antibiotic cassette, and the six resulting libraries combined. This strategy enables the use of commercially available doped oligos, while enhancing the percentage of the population made up of single-point mutants, and minimizing the proportion of wild-type sequence in the final population **(Fig. S1A)**. This library was transformed into *S. pneumoniae* TIGR4 and integrated into the native locus via homologous recombination to create a library of strains carrying these mutations. Past work has shown that integration of the antibiotic cassette alone at this site has minimal impact on *S. pneumoniae* fitness[30].

With this library, we assessed the relative fitness of each strain in a chemically defined medium (CDM) in the absence of riboflavin by growing our mutant strain library to mid-log phase (~3 doublings), harvesting a pair of T1 samples, and subsequently inoculating cultures in CDM lacking riboflavin (technical triplicate) and growing for an additional ~10 doublings before harvesting cells (T2). Genomic DNA was extracted from both T1 and T2 samples, and the FMN riboswitch preceding *ribU* amplified via PCR with unique molecular identifiers (UMIs) to enable accurate quantification of each variant. The resulting amplicons were sequenced on a NextSeq2000, and 0.6 - 7.3 million reads were obtained for each sample **(Fig. S1B, Fig. S2A)**.

In examining the distribution of mutation frequencies from T1 and T2 samples in the absence of riboflavin, we observe that the distribution of wildtype versus mutants with 1-6 mutations is similar in both datasets, with about 60% of reads corresponding to the wildtype sequence **(Fig. 1A)**. We also find that all single point mutations are represented in the datasets

at both the T1 and T2 conditions indicating that we assessed the fitness of all single mutants **(Fig. S1C)**. As the number of double mutants for the FMN riboswitch is large (1,245,456), and our construction methodology does not create all double mutants, we did not examine fitness of individual double mutants.

To quantify the fitness of each single mutant, we used the DiMSum pipeline[31]. This software package extracts the frequencies of each single mutant variant before (T1) and after (T2) selection, and calculates fitness based on these frequencies, where a fitness of 0 is equivalent to the wild-type sequence, a fitness >0 is increased frequency in T2, and a fitness <0 is a decreased frequency in T2. DiMSum analyzed 1.8 – 12.5 million deconvoluted reads for each condition (**Fig. S2B**). We find that in CDM without riboflavin, no individual single-point mutation exhibited a fitness greater than 1 (beneficial) or less than -1 (detrimental) (**Table S1, Fig. 1B-E**). Thus, single-point mutations do not confer a substantial fitness advantage or disadvantage to *S. pneumoniae* in culture, and the FMN riboswitch is not under substantial selective pressure. This is consistent with the lack of change in the distribution in our mutation frequency profile **(Fig. 1A)** under these conditions. The intact riboflavin biosynthetic operon (SP_0178 - SP_0175) appears to provide sufficient riboflavin for growth in culture lacking riboflavin and constitutive expression of the transporter at levels accessible by single mutations is not detrimental under these conditions.

Although no individual mutations showed a strong beneficial or detrimental impact on fitness, we assessed the population of mutations to determine whether any underlying patterns might be detectable in the low-magnitude changes observed. Looking closer at specific regions of the riboswitch (ligand-binding aptamer, extended P3 stem (ext. P3 stem), extended terminal stem (ext. term. stem), and terminator) **(Fig. 1F)**, we find that in the aptamer and the ext. P3 stem, transversions have a statistically significant positive impact on fitness compared to transitions **(Fig. 2A)**. However, the difference in fitness median (Δmedian) from transversions to transitions in the aptamer is 0.043, and the Δmedian for the extended P3 stem (ext. P3 stem) is 0.053. Thus, while the difference in fitness between transversions and transitions is statically significant over this large number of mutations, the magnitude of the shift in medians is very small. Whether or not the mutants are in positions expected to base-pair or remain unpaired, and whether the mutations are disruptive or compatible with respect to the predicted base-pairing, does not cause a significant difference in fitness in any region of the riboswitch **(Fig. 2B)**. Given the lack of selective pressure apparent in this condition, going forward we will consider this condition our negative control.

Single mutations are neutral in chemically defined medium containing riboflavin

To assess whether addition of riboflavin to the medium would alter the fitness benefit conferred by expression changes to RibU, the riboflavin transporter, we grew the library for ~10 generations in the presence of riboflavin, and harvested cells for amplicon sequencing as described above. Similar to T1 and the -riboflavin condition, about 60% of the reads corresponded to the wildtype sequence, suggesting that there is not a strong selection for mutants in CDM containing riboflavin **(Fig. 1A)**. Additionally, no mutants had a fitness >1 or <-1 in the presence of riboflavin, implying that there is not a strong fitness advantage or disadvantage to *S. pneumoniae* for any individual mutations under this condition **(Fig. 1B-E)**.

When comparing all regions of the riboswitch to the condition lacking riboflavin, there is no significant difference in fitness compared to the -riboflavin condition for any of the regions **(Fig. 1B-E)**. However, in all individual regions, transversion mutations are slightly higher fitness than transitions, but the Δmedians are again minimal **(Fig. 2C)**. In the aptamer there are statistically significant differences between compatible mutations (those do not break Watson-Crick pairing due to the G•U wobble) at paired bases compared with both mutations disruptive of base-pairing or those at unpaired positions but again the Δmedian is substantially less than one **(Fig. 2D)**. Similarly, in the terminator region there are statistically significant differences in fitness between

mutations disruptive to base-pairing and both mutations to unpaired bases and mutations compatible with the base-pairing, but the Δmedians are small **(Fig. 2D)**. Thus, the addition of riboflavin to the medium does not substantially change the lack of selective pressure on the FMN riboswitch preceding RibU in *S. pneumoniae*.

Mutations to the aptamer domain are beneficial in the presence of ribocil-C

The antibiotic ribocil-C has been shown to interact directly with FMN riboswitches[17,18]. To determine what mutations to the FMN riboswitch preceding RibU confer fitness changes in the presence of this antibiotic, the mutant strain library was grown in CDM (lacking riboflavin) in the presence of 10μM ribocil-C. These cultures were unable to reach OD_{600} >0.4, and were harvested at 12.5 hours. We estimate approximately eight to nine doublings were achieved. In contrast to our non-selective conditions, the distribution of wild-type and mutated variants changed substantially. Most of the reads analyzed in this condition contained 0-2 mutations, with wildtype making up only 30% of the total reads. This corresponds to about half the number of reads attributed to wildtype compared to the conditions with and without riboflavin, indicating that some mutants are more beneficial than wildtype in the presence of ribocil-C **(Fig. 1A)**.

Analysis of individual single-mutations by region of the riboswitch shows that in the presence of ribocil-C, mutations conferring a fitness advantage in ribocil-C were substantially enriched in the aptamer compared to the -riboflavin condition **(Fig. 1B)**. The difference in fitness in the aptamer is both significant and has a Δmedian of 1.442 indicating that mutations to the aptamer confer a fitness advantage in the presence of ribocil-C. Visualizing the locations of beneficial and detrimental mutations on the riboswitch secondary structure shows that mutations to most positions in the aptamer region (83.3%) show positive effects on fitness **(Fig. 3A)**, with many beneficial single-point mutations exhibiting fitness values greater than three (**Fig. 3B**). Since DiMSum calculates fitness based on the natural log of the ratio between variant counts before and after selection compared to wildtype[31], fitness score increases at an exponential rate. Thus, a fitness score of three is noteworthy because it indicates a performance level substantially greater than lower scores, rather than representing a linear increase.

The majority of mutations that confer positive fitness effects in the aptamer are transversions (Δmedian = 2.469) (**Fig. 3C**). In assessing all aptamer positions with respect to their base-pairing status, we see that mutations disruptive to Watson-Crick pairing within the stems are much more beneficial than those that are compensatory (Δmedian = 2.407) or those that are not Watson-Crick paired (Δmedian = 1.900) (**Fig. 3D**). Furthermore, regions that are not paired in the secondary structure, but may be involved in tertiary interactions or ligand binding, show a wider distribution of effects, and are generally positively selected compared to the mutations compatible with base pairing (Δmedian = 0.507), but not significantly so. On balance this pattern suggests that mutations impairing the ability of the aptamer to bind ligand are strongly enriched in this population.

We also observe that mutations to other regions of the switch display significant difference compared to the -riboflavin condition, however these differences are much smaller in magnitude (Δmedian = 0.138 to 0.516) (**Fig. 1C-E**). Furthermore, we see that while transversions are significantly more positive than transitions in the ext. P3 stem, they are more negative than transitions in the ext. term. stem **(Fig. 3C)**, and the Δmedians between the transition and transversion mutations are relatively small (0.226 and 0.343). Mutations that break predicted base-pairing are significantly more positive than mutations at unpaired locations in the ext. P3 stem, but the Δmedian is only 0.231, and mutations that are disruptive to the predicted pairing in extended terminator stem are slightly more deleterious than those predicted to be compatible with the paring (Δmedians = 0.357) (**Fig. 3D**). Thus, mutations to other regions of the riboswitch are potentially slightly deleterious to fitness, but this effect is generally of small magnitude, especially when compared to the positive effect of mutations to the aptamer in this condition.

Mutations to the terminator region are beneficial in the presence of roseoflavin

Although both ribocil-C and roseoflavin target the FMN riboswitch, they are not similar chemically, and structural studies show different modes of interaction[17,18,20,32]. Furthermore, roseoflavin is thought to be imported via the riboflavin transporter (RibU), under control of this riboswitch, and roseoflavin resistance mutants have been raised in FMN riboswitches preceding the transporter in *S. aureus*[20,33–35]. Thus, we hypothesized that an alternative set of mutations might confer a fitness defect or benefit on *S. pneumoniae* in the presence of roseoflavin. To identify these mutants, the mutant strain library was grown with 256nM roseoflavin. Cells grew in the presence of roseoflavin for ~10 doublings before harvest and gDNA extraction. Most of the reads in this condition contained 1-3 mutations, with wildtype only making up ~10% of the total reads **(Fig. 1A)**. While the proportion of wildtype reads in both ribocil-C and roseoflavin conditions is much less than that in the presence or absence of riboflavin, even fewer reads are attributed to wildtype in the roseoflavin condition than in the presence of ribocil-C, suggesting that the wildtype sequence is less fit compared to the positively selected mutants in the presence of roseoflavin than in the presence of ribocil-C.

Looking at mutations that are beneficial in the presence of roseoflavin, mutations to the terminator region are significantly more positive than in the -riboflavin condition, with a high ∆median of 2.067 (**Fig. 1E**). Individual single-point mutations located in the terminator region exhibited fitness values greater than four (**Fig. 4A, B**) suggesting a strong fitness benefit. Transversions in the terminator are significantly more beneficial than transitions in the terminator, with a high ∆median of 3.102 **(Fig. 4C)**. Mutations disruptive to predicted base-pairing in the terminator region result in a significant positive impact on fitness and the highest ∆medians seen in this dataset, 3.128 compared to unpaired nucleotides and 3.125 compared to mutants that are compatible with base-pairing **(Fig. 4D)**. This suggests that disrupting the putative terminator stem is beneficial in the presence of roseoflavin.

In addition to the strong effect observed in the terminator region, we also see that mutations to the ext. P3 stem in the presence of roseoflavin confer a fitness disadvantage when compared to the -riboflavin, with a ∆median of 0.715 **(Fig. 1C, Fig. 4A)**. This indicates that the wildtype sequence of the P3 stem is more beneficial than a mutated P3 stem in the presence of roseoflavin. Transversion substitutions in this region conferred a significantly more negative impact on the fitness of *S. pneumoniae* than transitions with a ∆median of 1.294 **(Fig. 4C)**, and mutations compatible with the predicted base-pairing conferred a significantly more positive impact on fitness than mutations at unpaired regions or those disruptive of predicted pairing, with ∆medians of 0.679 and 0.846 **(Fig. 4D)**. Thus, it appears that maintaining the structure and wildtype sequence of the extended P3 stem is important for *S. pneumoniae* fitness in the presence of roseoflavin.

Finally, we also observe that mutations to the extended terminator stem (ext. term stem) and aptamer are both statistically different compared to our -riboflavin negative control, with small magnitude ∆medians (**Fig. 1B, D**). Assessing these mutations more carefully we see that transitions and transversions in the aptamer are not statistically different in their fitness, but that transversions to the ext. term. stem are significantly more negative than transitions, with a small ∆median of 0.335 (**Fig. 4C**). Mutations that compatible with base-pairing in the ext. term. stem are significantly beneficial compared to mutations at unpaired regions or disruptive mutations, with minimal ∆medians of 0.294, and 0.286 (**Fig. 4D**). Thus, while these differences are statistically significant, they are not of high magnitude.

FMN riboswitch is under mild purifying selection during *in vivo* infection

To understand the role that the FMN riboswitch plays in *S. pneumoniae* survival in an infection model, we inoculated 24 mice with $2.5*10^7$ cells each of the FMN riboswitch mutant strain library, allowed the infection to progress for 20-24 hours, and subsequently harvested and homogenized the lung and plated on blood plates to collect the living cells. For each mouse > 10^5

CFU were ultimately collected following infection due to the large bottleneck that occurs upon *S. pneumoniae* infection of the lung[36]. Genomic DNA was then extracted from colonies collected off these plates and amplicon sequencing performed similarly to our experiment above. A pre-inoculation sample was collected as separate T1 sample for this experiment ($T1_{mouse}$). Close to 70% of the reads coming from both the $T1_{mouse}$ and the mice are the wildtype sequence, similar to the + and -riboflavin conditions **(Fig. 1A)**. This suggests that mutants are neither strongly beneficial or detrimental in the *in vivo* environment.

When compared to the -riboflavin condition, mutations to all regions of the riboswitch displayed significantly negative impacts on fitness, with Δmedians of 0.158 to 0.248. These are substantially higher than the Δmedians of the + or -riboflavin populations, but small in magnitude compared to those from the roseoflavin or ribocil-C challenged populations. Transversions were significantly more detrimental to fitness than transitions in all regions with Δmedians ranging from 0.246 to 0.620 **(Fig. 5A)**. Additionally, mutations that are compatible with predicted base-pairing in the P3 stem have a significantly more positive impact on fitness (Δmedian = 0.280), and mutations that break base-pairing in the terminator have a significantly negative impact on fitness (Δmedian = 0.461) **(Fig. 5B)**. Together, this data suggests that no mutations have strong negative or positive impacts on the fitness of *S. pneumoniae* in mice, but that many mutations are slightly deleterious suggesting purifying selection on the riboswitch sequence where disruptive transversions are selected against more compared to less disruptive transitions.

A biological replicate reinforces the amplicon sequencing results

The reported data is from one FMN riboswitch doped oligo library (Lib1). However, a second library was made in the same manner, as a biological replicate (Lib2). The same amplicon sequencing experiment was repeated for Lib2, and two replicates from each *in vitro* condition were sequenced on the NextSeq2000. To determine if the biological replicates Lib1 and Lib2 report similar results from the amplicon sequencing experiment, we compared the conditions tested between each library. For the four *in vitro* conditions tested, the r values were 0.03717 for -riboflavin, 0.004272 for +riboflavin, 0.6119 for ribocil-C, and 0.6673 for roseoflavin **(Fig. 6A)**. Considering that no mutation in the + or -riboflavin condition has a fitness > |1| for Lib1, and very few do for Lib2, the low r values are not surprising, as the fitness changes likely largely correspond to noise under such conditions where there is minimal selection.

The main difference between the sequenced libraries is that the error on the fitness of Lib2 is greater than the error on the fitness of Lib1 in each condition **(Fig. S3)**. This is likely due to the two vs three replicates sequenced from Lib2 and Lib1 respectively. The Lib2 data predominantly serves to reaffirm the Lib1 results **(Fig. S4, S5, and S6, Table S6)**. In Lib1 in the presence of ribocil-C, four mutants confer a fitness defect <-4 in the ext. term. stem. However, in Lib2 these same mutants in ribocil-C do not have a fitness > |1|. Since these fitness values were not replicated, we did not assess them further.

In the ribocil-C and roseoflavin conditions, the more positive the fitness, the smaller the error, and the more negative the fitness, the larger the error. This is due to the nature of the measurement, where positive fitness values correspond to a higher read count and negative fitness values a lower read count in the raw data. Thus, in our experiment we are more confident in mutants with positive fitness, than those with negative fitness. Going forward, we focus our attention on mutants that are beneficial, rather than detrimental under our selective conditions as these values are higher confidence.

Selected mutants largely confirm behaviors anticipated from high-throughput experiments.

To validate the findings from our high-throughput experiment, we selected a panel of mutations for further analysis. Three mutants beneficial in the presence of roseoflavin (RoF1, RoF2, RoF3), four mutants beneficial in the presence of ribocil-C (RibC1, RibC2, RibC3, RibC4),

and two negative control mutations in non-selected sites (NC1, NC2) were chosen **(Fig. 6B)**. The mutants were selected to sample the different regions of the riboswitch found to be beneficial in the presence of these antibiotics. *S. pneumoniae* strains carrying each individual mutation were constructed and validated (see Methods) **(Fig. 6B)**. To enable a more comprehensive comparison, we also assessed five additional strains: the *S. pneumoniae* TIGR4 wildtype strain (316), and previously published mutants: a wildtype containing a chloramphenicol resistance cassette (WT^c), a ligand-insensitive mutant (mLI), an "always on" mutant (mON), and an "always off" mutant (mOFF)[30]. None of the previously constructed mutants are singe point mutations, but rather represent much more severe changes to the sequence that range from a double mutant (mLI) to large scale deletion of the aptamer (mOFF), or terminator stem (mON).

These fourteen strains were first assessed against a panel of roseoflavin concentrations ranging from 8nM to 16μM. The IC_{50}s for the RibC strains, the NC strains, 316, WT^c, mLI, and MON are ~32nM roseoflavin. The IC_{50}s for RoF1, RoF2, RoF3, and mOFF are nearly 30x higher, around 1μM roseoflavin **(Fig. 6C)**. Thus, the roseoflavin resistant mutants and the mOFF mutant have IC_{50}s an order of magnitude higher than the RibC and control strains. This demonstrates that the mutants positively selected in the presence of roseoflavin identified via the high-throughput experiment are resistant to roseoflavin

To determine if the RibC mutants displayed a similar resistance to ribocil-C, the same panel of fourteen *S. pneumoniae* strains was assessed for growth across a concentration range of 250nM to 4μM ribocil-C. The IC_{50} for all fourteen strains is 1μM – 2μM ribocil-C **(Fig. 6D)**. Although RibC4 and WT^c have a significantly higher IC_{50} than unmodified *S. pneumoniae* TIGR4 (316), the magnitude of this change is small and the concentration is still much lower than that in our selection (10μM). Upon careful inspection of the maximum OD_{600}s at moderate concentrations of ribocil-C, the roseoflavin-resistant mutants RoF1, RoF2, RoF3, and the mOFF mutant, showed significantly decreased maximum OD_{600}s at 4μM compared to wildtype 316 (**Fig. 6E**). This suggests that disrupting the terminator region potentially makes strains simultaneously more sensitive to ribocil-C at high concentrations, and resistant to roseoflavin.

Roseoflavin and ribocil-C selected mutants display contrasting gene expression levels

Based on the results above we hypothesized that mutants positively selected in the presence of roseoflavin are likely constitutively repressed mutants similar to the mOFF mutant. To confirm this, we examined the gene expression of our panel of mutants in the presence and absence of riboflavin via beta-galactosidase reporter assays, similar to how we previously assessed our mLI, mON, and mOFF mutants[30]. We find that the RoF mutants, while still displaying some riboflavin dependency, express the RibU-LacZ translational fusion at lower levels than the riboflavin repressed WT^c, regardless of whether riboflavin is present in the medium **(Fig. 6F)**. In contrast we see that the RibC mutants all constitutively express the RibU-LacZ fusion at higher levels than the riboflavin repressed WT^c, and often higher than WT^c in the absence of riboflavin **(Fig. 6F)**. Thus, in the presence of roseoflavin, constitutive repression of the transporter is beneficial, while in the presence of ribocil-C, constitutive activation is beneficial.

Roseoflavin resistant mutants reveal a translational repression mechanism of action

Previous 3'-end sequencing data identified a termination site associated with this riboswitch just upstream of the translational start[29]. However, how precisely this terminator enabled FMN-dependent termination was unclear from predicted base-pairing. The most positively selected mutants in the presence of roseoflavin cluster to this region of the transcript, with mutations to positions in the base of the aptamer and the region connecting the aptamer to the ext. term. stem displaying the strongest positive impacts on fitness in roseoflavin **(Fig. 4A)**. Based on our validation experiments, these mutations likely all decrease gene expression. However, they also uniformly decrease the stability of the predicted base-pairings, and specifically destabilize the formation of the potential intrinsic terminator stem close to the site of termination.

This finding is surprising, as destabilization of the terminator should result in activation of gene expression rather than repression. This unexpected finding motivated us to look further at the entire transcript sequence to determine if there were any potential interactions with sequence following the terminator. Investigation of the downstream region revealed the presence of extensive potential base-pairing between the Shine-Dalgarno sequence and the region of the RNA constituting one side of the intrinsic terminator and the polyU sequence itself **(Fig. 7B).** The mutations positively selected in the presence of roseoflavin also serve to destabilize the alternative pairing that allows the ribosome to access the Shine-Dalgarno sequence **(Fig. 7A)**, ultimately resulting in constitutive repression of gene expression as demonstrated by our panel of mutants **(Fig. 6F).** Consistent with this, mutations to the terminator in and immediately following the aptamer (bases 224-235) confer a higher fitness than those at the base of the terminator closest to the termination site (bases 355-367) as mutation of bases closer to the aptamer only serves to destabilize the structure that exposes the Shine-Dalgarno sequence, while mutation of bases closer to the site of termination may serve to destabilize both potential structures.

This finding strongly suggests that the mechanism of action for this riboswitch is via regulation of translation initiation rather than transcription attenuation, despite the close association of an intrinsic terminator proximal to the aptamer. To assess this, we performed qPCR to determine *ribU* transcript levels for our WT^C control strain grown in the absence of riboflavin and in the presence of roseoflavin, ribocil-C, or riboflavin. As an internal control, we also assessed levels of *ribD*, which is regulated by the other FMN riboswitch and believed to be transcriptionally regulated. Our results show that levels of *ribU* transcript are not substantially changing in the presence vs. absence of riboflavin, and may very modestly increase in the presence of ribocil-C or roseoflavin **(Fig. 7C)**. In contrast, our beta-galactosidase assays show approximately a 6-fold change between the presence and absence of riboflavin **(Fig. 6F)**. Conversely, levels of *ribD* decrease approximately 4-fold in riboflavin, ribocil-C, and roseoflavin compared the absence of riboflavin, indicating that the altered conditions do impact FMN riboswitch functionality within *S. pneumoniae* and are altering transcript levels of the biosynthesis operon **(Fig. 7D)**. These results support the translational regulation of *ribU*, and our proposed mechanism of action based on the beneficial mutations selected by roseoflavin.

DISCUSSION

By monitoring the growth of a strain library of FMN riboswitch variants in five different environments, we demonstrated that high-throughput techniques can illuminate entire suites of point mutants displaying decreased or increased expression. In the presence of roseoflavin, strongly beneficial mutations result in constitutively repressed expression **(Fig. 4, Fig. 6F)**. This phenotype is likely beneficial in this environment because it reduces roseoflavin transport into the cell. Notably, these mutations also result in increased sensitivity to moderate concentrations of ribocil-C **(Fig. 6E, F)**. In contrast, in the presence of ribocil-C, advantageous mutations occur predominantly within the aptamer. These mutations are likely disruptive to aptamer structure, with transversions showing much greater effect on fitness than transitions **(Fig. 4C)**. These mutations likely impede ligand binding, whether that ligand is ribocil-C, roseoflavin, or FMN, and ultimately result in increased gene-expression **(Fig 6F)** due to destabilization of the ligand-bound aptamer.

The assessment of many mutations in a single experiment further allowed us to make insights into the mechanism of action for this riboswitch. While riboswitch aptamers are typically well-conserved and recognizable via sequence analysis alone, the features that determine how ligand is mediated into genetic regulation or “expression platforms” are not well conserved across species and can be challenging or impossible to deduce from sequence analysis alone[37]. The presence of an intrinsic terminator identified via 3’-end sequencing in close proximity to the aptamer biased our initial hypotheses toward assuming a transcriptional mechanism of action. However, the suite of strongly positively selected mutants in the presence of roseoflavin lead us determine that this FMN riboswitch regulates via a translational repression mechanism. This

finding underscores the challenges associated with prediction of riboswitch mechanism of action in the absence of detailed experiments, even when coupled to high-throughput 3'-end sequencing data.

In addition to the mechanistic insights gained from assessment of our mutations, the fitness landscapes can also shed light on biological forces acting on non-standard features. The extended P3 stem of this FMN riboswitch in *S. pneumoniae* is much longer than the canonical version present in FMN riboswitches[38]. Although this sequence is common amongst *S. pneumoniae* strains, with very few mutations identified in analysis of this sequence across 300 clinical isolates[29], this feature is not shared in amongst other Streptococcus species[39]. This sequence consists of ~120 additional nucleotides, many of them predicted to pair with each other. Similar such extensions are common in many structured RNAs, and their origins, the evolutionary pressures upon them, and what if any biological function they may have are not well understood[40]. In the presence of roseoflavin, mutations to the P3 stem confer a significantly decreased fitness in *S. pneumoniae*, meaning that the wildtype sequence is more beneficial in these conditions. The majority of the deleterious mutations are transversions, and mutations disruptive of predicted base-pairing display lower fitness values than mutations compatible with the pairing **(Fig. 4C, D)**, while mutations to positions not predicted to be paired are similarly deleterious to those that disrupt base-pairing. We see a similar pattern, although with a smaller magnitude change in fitness, in data collected from *in vivo* lung infection. Our findings suggest that this deviation from the canonical FMN aptamer is under selection to maintain structure, and this pairing provides fitness benefits in specific conditions, including during *in vivo* infection.

Beyond specific insights into RNA structure/function, our study also reveals the complex relationship between drugs targeting metabolic control, the native metabolite ligands, and the processes inhibited by such compounds. While the mechanism of resistance to roseoflavin appears to be relatively straightforward, one curious finding is the lack of explicit ribocil-C resistance for mutants strongly positively selected in the presence ribocil-C **(Fig. 4 and Fig. 6D**). There are several potential explanations for this phenotype. First, when measuring IC_{50}, we are measuring resistance to ribocil-C in monoculture rather than in a mixed strain environment. It is possible that there is a more measurable benefit in the mixed strain condition, where some individuals benefit from others surrounding them.

A second possibility is that the positively selected mutants, which display a higher gene expression in the presence of riboflavin, express higher levels of the transporter in the presence of ribocil-C and are thus better able to scavenge FMN or riboflavin from the environment. This could give such mutants an advantage in the presence of ribocil-C. Ribocil-C inhibits both FMN biosynthesis and transport, and is a competitive inhibitor with FMN for riboswitch binding[17,18,20]. Thus, any additional FMN present inside the cell would potentially have substantial impacts on drug efficacy. Indeed, the most frequent mechanism of resistance to roseoflavin is constitutive expression of the biosynthetic operon[15]. However, based on the fact that our ribocil-C selected cultures failed to achieve 10 doublings, and explicit resistance is not apparent in our positively selected mutants, we speculate there are no mutants in our population that confer substantial resistance rather than increased fitness in ribocil-C.

Our study also shows the extent of natural selection on the riboswitch under conditions the organism is more likely to encounter. We find that single mutations are overwhelmingly neutral in the absence of riboflavin, with minimal or no distinction between transitions and transversions across most regions of the riboswitch. This indicates that the switch is not under strong selection under these conditions. When riboflavin is added to medium, a similar, but not identical, portrait emerges. However, when the library is assessed within an *in vivo* lung infection model, weak purifying selection is apparent, with most mutations displaying only small changes in fitness **(Fig. 5)**. However, in aggregate it is clear that many mutants are mildly deleterious across all regions of the sequence (**Fig. 1B-E**) with transversions always appearing more deleterious than transitions. This stands in contrast to the *in vitro* data of antibiotic perturbed conditions where

transversions are positively selected to result in altered behavior of the riboswitch. This result was somewhat surprising, as our mOFF truncation mutant displays substantial defects compared to the wildtype *S. pneumoniae* in the lung infection model[30]. However, the positively selected roseoflavin resistant mutants we validated show 5 to 10-fold greater gene-expression than our truncation mOFF mutant, which is only incrementally above background **(Fig 6F)**. The combination of these results suggests that there are no single mutants capable of conferring such a deleterious phenotype to *S. pneumoniae*, and thus the sequence is inherently robust to deleterious mutations, as has been demonstrated repeatedly for natural sequences[41–44], but is not necessarily a general trait of all functional sequences[24].

MATERIALS AND METHODS

Doped oligo library construction

Six doped oligos, with 5' and 3' constant regions, were ordered from IDT to create a SP_0488 FMN riboswitch mutant library that would be mutated at each nucleotide. The second strand of each doped oligo was synthesized to make the DNA double stranded by combining tris-HCl (pH 8.0), $MgCl_2$, primer, and doped oligo library for each of the six doped oligo libraries. Each doped oligo was PCR amplified, and purified using the Zymo Clean & Concentrator kit. 3' and 5' homology arms were also amplified via PCR and purified using the Zymo Clean & Concentrator kit. The three regions for each doped oligo DNA library were then assembled in two Gibson Assemblies: the first with the doped region and the 3' homology arm, and the second with the 5' homology arm. The six doped oligo libraries were then mixed together in an equimolar pool and transformed into *S. pneumoniae*[45]. Primers used in this study are listed in **Table S5**.

Amplicon sequencing culture conditions and sample collection

A doped oligo library starter culture was washed three times and used to inoculate 10mL of semi-defined minimal media (SDMM), and grown until mid-log phase. This culture was then washed three times, and used to inoculate three replicates of each experimental condition at an OD_{600} of 0.001. The experimental conditions were in the presence of riboflavin, in the absence of riboflavin, 256nM roseoflavin, and 10μM ribocil-C, grown in chemically defined media (CDM) without riboflavin. After eight hours of growth, each culture was spun down, the supernatant pipetted off, and the media replaced with fresh media and ligand. Each culture was grown until it reached an OD_{600} of ~1, or until it had been 12.5 hours. At this point, the cultures were pelleted, the supernatants pipetted off, and the pellets stored at -80°C to use for amplicon sequencing preparation.

In vivo FMN riboswitch mutant strain library challenge

Fitness experiments were performed by challenging mice with the FMN riboswitch mutant strain library. The library was initially grown to mid-log phase, at which point it was washed and resuspended in 1x PBS. Mice were inoculated with $2.5*10^7$ cells via intranasal injection, and sacrificed 24 hours later, where the lungs were removed and homogenized in 1x PBS as described previously[30]. 10μl of homogenized lungs were used for serial dilutions to estimate the bacterial burden. The remainder of the homogenized lungs were plated on chloramphenicol blood agar plates, grown overnight, and the resulting colonies collected for amplicon sequencing.

Amplicon sequencing preparation

The cell pellets were taken out of the -80°C freezer, and their gDNA was isolated. The quantity was determined using a nanodrop, and each sample was diluted to 50ng/μl. 800ng of gDNA was then used in an initial 400μl 5 cycle PCR, spread over 8 PCR tubes, used to add the unique molecular identifiers (UMIs). The products were column purified using the Zymo Clean &

Concentrator kit, and the quantity was determined using the nanodrop. 400ng was then used in a 400µl 30-cycle PCR, spread out over 8 PCR tubes, for amplification of the libraries. These products were then purified using the Zymo Gel Purification kit, and quantified using the Qubit. Samples were mixed together to create an equimolar pool. The pool was quantified using the Qubit, and run on either the NextSeq 2000, using the P1 2x300bp kit, or the MiSeq using the v3 reagent kit. Primers used in this study are listed in **Table S5**.

Identification of antibiotic resistant mutants

After receiving the Fastq files from the NextSeq 2000 or MiSeq, the paired end reads were initially merged using Vsearch[46]. PCR bias was then corrected for by using AmpUMI twice, the first time using the UMIs attached on the 5' end, the second time using the UMIs attached on the 3' end[47]. The deep mutational scanning software DiMSum was then used to identify mutants that are either beneficial, detrimental, or neutral in each condition[31].

FMN RNA mutant strain construction

Oligos ordered from IDT containing the targeted mutations with additional overhangs for Gibson Assembly were used in an initial 30-cycle PCR to add the mutant and overhang to both halves of the DNA containing the 5' and 3' homology arms, the chloramphenicol resistant cassette, and the FMN riboswitch. The PCR products were purified using the Zymo Clean & Concentrator kit, and were used for Gibson Assembly. The Gibson Assembly products were then diluted 1:4, and used in a 30-cycle PCR for amplification. The PCR products were purified using the Zymo Clean & Concentrator kit and quantified using the nanodrop. 50ng DNA was used to transform into wildtype, 316 (TIGR4)[45]. Transformants were plated on chloramphenicol blood plates, and 12 colonies from each mutant were picked, grown for six hours, and analyzed via colony PCR. Glycerol stocks were made for each mutant. PCR products amplifying the genome region of interest from promising mutants were sent to Eton Biosciences for Sanger Sequencing, and one from each mutant was selected to make starters with and use for experiments. Primers used in this study are listed in **Table S5**.

FMN RNA mutant growth assays

One starter per strain was washed three times in 1xPBS, and then used to inoculate 7.5ml of chemically defined media (CDM) without riboflavin. Cultures were grown for ~2 hours. The OD_{600} was determined, and cultures were diluted to an OD_{600} of 0.015 in 7ml CDM-riboflavin. Culture and ligand were mixed to obtain the desired ligand concentration, and 200µl of each condition was plated into three wells of a 96-well plate. The plates were incubated in a BioTek BioSpa 8, where their OD_{600} was read every 30 minutes. IC_{50}s were determined by identifying the antibiotic concentration at which growth is inhibited >50% compared to growth in the -riboflavin condition at 8.5 hours.

LacZ reporter strain construction and beta-galactosidase assays

Genomic DNA from strains carrying the RibC, RoF, and NC mutant strains was PCR amplified with primers SP0647+cat_F and lacZ_SP0488_R to generate a product containing the chloramphenicol resistance cassette and the riboswitch through the first three codons of *ribU* under the control of its native promoter. For each mutant these fragments were combined via Gibson Assembly with flanking sequences (amplified by primers SP0647_F/CAT+SP0647_R, and SP0488+lacZ_F/lacZ_R respectively) that correspond to a soluble *lacZ* sequence, and SP_0647, to result in a translational fusion between *ribU* and *lacZ*. These full-length products were PCR amplified and transformed into a previously reported strain wherein the endogenous beta-galactosidase in *S. pneumoniae* TIGR4 (SP_0648) was replaced with a portion of a soluble *lacZ* and a kanamycin resistance marker[30]. Transformants were screened for correct integration of the

Chl cassette via PCR and presence of the mutation confirmed via Sanger sequencing of the PCR product. Primers used in this study are listed in **Table S5**.

The resulting reporter strains were grown in CDM with and without riboflavin to an OD_{600} of ~0.5 and harvested. Cells were resuspended in Z buffer (50 mM Na2HPO4, 40 mM NaH2PO4, 10 mM KCl, 1 mM MgSO4, 50 mM 2-mercaptoethanol) and beta-galactosidase activity assay was performed as described previously[48].

qPCR analysis

RNA was isolated from cultures, with a starting OD_{600} of 0.1, grown for 3 hours under -riboflavin, +riboflavin, 256nM roseoflavin, and 10μM ribocil-C conditions, using the Qiagen RNeasy kit (Qiagen). The resulting RNA was treated with DNase (TURBO DNA Free Kit (Ambion)). The DNase treated RNA was used to generate cDNA with iScript reverse transcriptase supermix for RT-qPCR (BioRad). Quantitative PCR was performed using a Bio-Rad MyiQ. Each sample was normalized against the 50S ribosomal gene, SP_2204 and were measured in technical triplicates. No-reverse transcriptase controls were included for all samples. Primers used in this study are listed in **Table S5**.

Acknowledgements:

We thank Dr. Indu Warrier and Dr. Elizabeth Gray for their technical advice in conducting this research. This work was supported by NIH grants R01GM134259.

References

1. Panchal, V. & Brenk, R. Riboswitches as Drug Targets for Antibiotics. *Antibiotics (Basel)* **10**, (2021).
2. Ellinger, E. *et al.* Riboswitches as therapeutic targets: promise of a new era of antibiotics. *Expert Opin Ther Targets* **27**, 433–445 (2023).
3. Giarimoglou, N. *et al.* A Riboswitch-Driven Era of New Antibacterials. *Antibiotics (Basel)* **11**, 1243 (2022).
4. Hong, W., Zeng, J. & Xie, J. Antibiotic drugs targeting bacterial RNAs. *Acta Pharm Sin B* **4**, 258–265 (2014).
5. Rekand, I. H. & Brenk, R. Ligand design for riboswitches, an emerging target class for novel antibiotics. *Future Med Chem* **9**, 1649–1663 (2017).
6. Mehdizadeh Aghdam, E., Hejazi, M. S. & Barzegar, A. Riboswitches: From living biosensors to novel targets of antibiotics. *Gene* **592**, 244–259 (2016).
7. Machtel, P., Bąkowska-Żywicka, K. & Żywicki, M. Emerging applications of riboswitches - from antibacterial targets to molecular tools. *J Appl Genet* **57**, 531–541 (2016).
8. Deigan, K. E. & Ferré-D'Amaré, A. R. Riboswitches: discovery of drugs that target bacterial gene-regulatory RNAs. *Acc Chem Res* **44**, 1329–1338 (2011).
9. Chauvier, A. & Walter, N. G. Regulation of bacterial gene expression by non-coding RNA: It is all about time! *Cell Chem Biol* **31**, 71–85 (2024).
10. Chauvier, A. & Walter, N. G. Beyond ligand binding: Single molecule observation reveals how riboswitches integrate multiple signals to balance bacterial gene regulation. *Curr Opin Struct Biol* **88**, 102893 (2024).
11. Winkler, W. C., Cohen-Chalamish, S. & Breaker, R. R. An mRNA structure that controls gene expression by binding FMN. *Proc Natl Acad Sci U S A* **99**, 15908–15913 (2002).
12. GBD 2016 Lower Respiratory Infections Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis* **18**, 1191–1210 (2018).

13. CDC. Antibiotic Resistance Threats in the United States, 2013. 1–114 (2013).
14. van Opijnen, T. & Camilli, A. A fine scale phenotype-genotype virulence map of a bacterial pathogen. *Genome Res* **22**, 2541–2551 (2012).
15. Lee, E. R., Blount, K. F. & Breaker, R. R. Roseoflavin is a natural antibacterial compound that binds to FMN riboswitches and regulates gene expression. *RNA Biol* **6**, 187–194 (2009).
16. Blount, K. F. *et al.* Novel riboswitch-binding flavin analog that protects mice against Clostridium difficile infection without inhibiting cecal flora. *Antimicrob Agents Chemother* **59**, 5736–5746 (2015).
17. Howe, J. A. *et al.* Selective small-molecule inhibition of an RNA structural element. *Nature* **526**, 672–677 (2015).
18. Howe, J. A. *et al.* Atomic resolution mechanistic studies of ribocil: A highly selective unnatural ligand mimic of the E. coli FMN riboswitch. *RNA Biol* **13**, 946–954 (2016).
19. Ott, E., Stolz, J., Lehmann, M. & Mack, M. The RFN riboswitch of Bacillus subtilis is a target for the antibiotic roseoflavin produced by Streptomyces davawensis. *RNA Biol* **6**, 276–280 (2009).
20. Wang, H. *et al.* Dual-Targeting Small-Molecule Inhibitors of the Staphylococcus aureus FMN Riboswitch Disrupt Riboflavin Homeostasis in an Infectious Setting. *Cell Chem Biol* **24**, 576-588.e6 (2017).
21. Pedrolli, D. B. *et al.* A highly specialized flavin mononucleotide riboswitch responds differently to similar ligands and confers roseoflavin resistance to Streptomyces davawensis. *Nucleic Acids Research* **40**, 8662–8673 (2012).
22. Andreasson, J. O. L., Savinov, A., Block, S. M. & Greenleaf, W. J. Comprehensive sequence-to-function mapping of cofactor-dependent RNA catalysis in the glmS ribozyme. *Nat Commun* **11**, 1663 (2020).
23. Li, C., Qian, W., Maclean, C. J. & Zhang, J. The fitness landscape of a tRNA gene. *Science* **352**, 837–840 (2016).
24. Jiménez, J. I., Xulvi-Brunet, R., Campbell, G. W., Turk-MacLeod, R. & Chen, I. A. Comprehensive experimental fitness landscape and evolutionary network for small RNA. *Proc Natl Acad Sci U S A* **110**, 14984–14989 (2013).
25. Pressman, A. D. *et al.* Mapping a Systematic Ribozyme Fitness Landscape Reveals a Frustrated Evolutionary Network for Self-Aminoacylating RNA. *J Am Chem Soc* **141**, 6213–6223 (2019).
26. Bendixsen, D. P., Collet, J., Østman, B. & Hayden, E. J. Genotype network intersections promote evolutionary innovation. *PLoS Biol* **17**, e3000300 (2019).
27. Peri, G., Gibard, C., Shults, N. H., Crossin, K. & Hayden, E. J. Dynamic RNA Fitness Landscapes of a Group I Ribozyme during Changes to the Experimental Environment. *Mol Biol Evol* **39**, msab373 (2022).
28. Hayden, E. J. & Wagner, A. Environmental change exposes beneficial epistatic interactions in a catalytic RNA. *Proc Biol Sci* **279**, 3418–3425 (2012).
29. Warrier, I. *et al.* The Transcriptional landscape of Streptococcus pneumoniae TIGR4 reveals a complex operon architecture and abundant riboregulation critical for growth and virulence. *PLoS Pathog* **14**, e1007461 (2018).
30. Warrier, I. *et al.* RNA cis-regulators are important for Streptococcus pneumoniae in vivo success. *PLoS Genet* **20**, e1011188 (2024).
31. Faure, A. J., Schmiedel, J. M., Baeza-Centurion, P. & Lehner, B. DiMSum: an error model and pipeline for analyzing deep mutational scanning data and diagnosing common experimental pathologies. *Genome Biology* **21**, 207 (2020).
32. Serganov, A., Huang, L. & Patel, D. J. Coenzyme recognition and gene regulation by a flavin mononucleotide riboswitch. *Nature* **458**, 233–237 (2009).

33. Mansjö, M. & Johansson, J. The riboflavin analog roseoflavin targets an FMN-riboswitch and blocks Listeria monocytogenes growth, but also stimulates virulence gene-expression and infection. *RNA Biol* **8**, 674–680 (2011).
34. Gutiérrez-Preciado, A. *et al.* Extensive Identification of Bacterial Riboflavin Transporters and Their Distribution across Bacterial Species. *PLOS ONE* **10**, e0126124 (2015).
35. Pedrolli, D. *et al.* The ribB FMN riboswitch from Escherichia coli operates at the transcriptional and translational level and regulates riboflavin biosynthesis. *FEBS J* **282**, 3230–3242 (2015).
36. Liu, X. *et al.* Exploration of Bacterial Bottlenecks and Streptococcus pneumoniae Pathogenesis by CRISPRi-Seq. *Cell Host Microbe* **29**, 107-120.e6 (2021).
37. Olenginski, L. T., Spradlin, S. F. & Batey, R. T. Flipping the script: Understanding riboswitches from an alternative perspective. *J Biol Chem* **300**, 105730 (2024).
38. McCown, P. J., Corbino, K. A., Stav, S., Sherlock, M. E. & Breaker, R. R. Riboswitch diversity and distribution. *RNA* **23**, 995–1011 (2017).
39. Ontiveros-Palacios, N. *et al.* Rfam 15: RNA families database in 2025. *Nucleic Acids Res* **53**, D258–D267 (2025).
40. Srivastava, A., Cai, L., Mrázek, J. & Malmberg, R. L. Mutational patterns in RNA secondary structure evolution examined in three RNA families. *PLoS One* **6**, e20484 (2011).
41. Zheng, J., Guo, N. & Wagner, A. Selection enhances protein evolvability by increasing mutational robustness and foldability. *Science* **370**, eabb5962 (2020).
42. Guo, H. H., Choe, J. & Loeb, L. A. Protein tolerance to random amino acid change. *Proc Natl Acad Sci U S A* **101**, 9205–9210 (2004).
43. Tokuriki, N. & Tawfik, D. S. Stability effects of mutations and protein evolvability. *Curr Opin Struct Biol* **19**, 596–604 (2009).
44. Wagner, A. Mutational robustness accelerates the origin of novel RNA phenotypes through phenotypic plasticity. *Biophys J* **106**, 955–965 (2014).
45. Bricker, A. L. & Camilli, A. Transformation of a type 4 encapsulated strain of Streptococcus pneumoniae. *FEMS Microbiology Letters* **172**, 131–135 (1999).
46. Mathon, L. *et al.* Benchmarking bioinformatic tools for fast and accurate eDNA metabarcoding species identification. *Molecular Ecology Resources* **21**, 2565–2579 (2021).
47. Clement, K., Farouni, R., Bauer, D. E. & Pinello, L. AmpUMI: design and analysis of unique molecular identifiers for deep amplicon sequencing. *Bioinformatics* **34**, i202–i210 (2018).
48. Miller, J. H. *A Short Course in Bacterial Genetics : A Laboratory Manual and Handbook for Escherichia Coli and Related Bacteria*. (Plainview, N.Y. : Cold Spring Harbor Laboratory Press, 1992).
49. Gruber, A. R., Lorenz, R., Bernhart, S. H., Neuböck, R. & Hofacker, I. L. The Vienna RNA websuite. *Nucleic Acids Res* **36**, W70-74 (2008).

Figure 1:

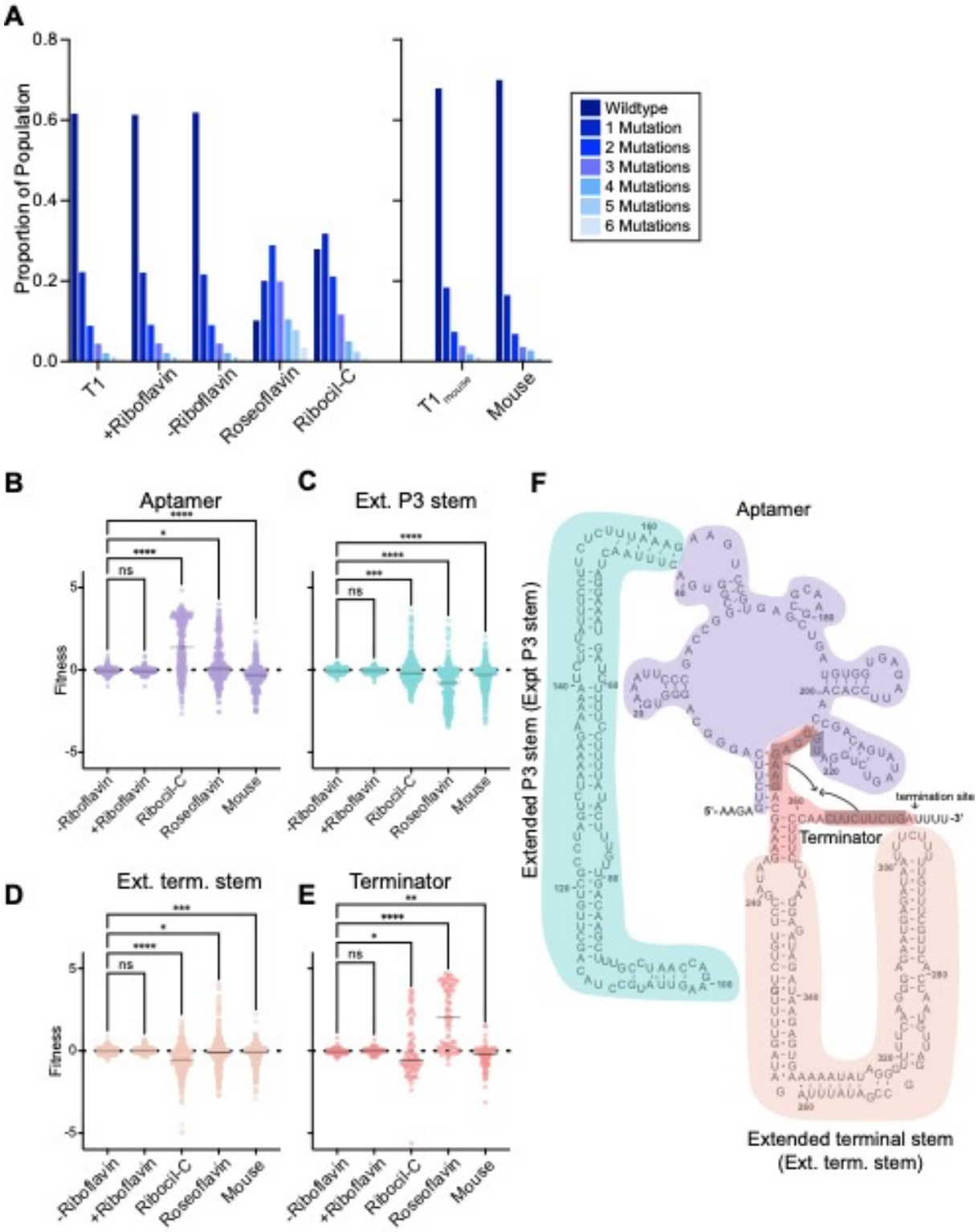
A
Proportion of Population
0.8
0.6
0.4
0.2
0.0
T1
+Riboflavin
-Riboflavin
Roseoflavin
Ribocil-C
T1 mouse
Mouse
Wildtype
1 Mutation
2 Mutations
3 Mutations
4 Mutations
5 Mutations
6 Mutations
B
Aptamer
C
Ext. P3 stem
D
Ext. term. stem
E
Terminator
Fitness
ns
F
Aptamer
Extended P3 stem (Expt P3 stem)
termination site
Terminator
Extended terminal stem
(Ext. term. stem)

Figure 1: The mutant strain library undergoes selection when challenged with antibiotics. A) The proportion of processed reads that have 0 (wildtype sequence) to 6 mutations for each condition and their respective T1. **B-E)** Fitness of mutations to different regions of the riboswitch color coded with diagram **F** relative to the wild-type sequence in five different environments: chemically defined media (CDM) with no riboflavin (-Riboflavin), CDM+ 140 ng/mL riboflavin (+Riboflavin), CDM+ 10μM ribocil-C (Ribocil-C), CDM+ 256nM roseoflavin (Roseoflavin), and within a mouse lung infection (Mouse). For each region, the fitness of the population of single mutants is compared to the fitness in the -Riboflavin condition using Kruskal-Wallis test with Dunn's test for multiple comparisons (adj. $p>0.05$ (ns), adj. $p<0.05$ (*), adj. $p<0.01$ (**), adj. $p<0.001$ (***), $p<0.0001$ (****)). Population medians and exact p-values for each comparison on **Table S2. F)** The SP_0488 FMN riboswitch secondary structure colored according to region. Secondary structure derived from alignment to FMN aptamer consensus structure[39], and folding extended stems with the RNAfold webserver[49].

Figure 2

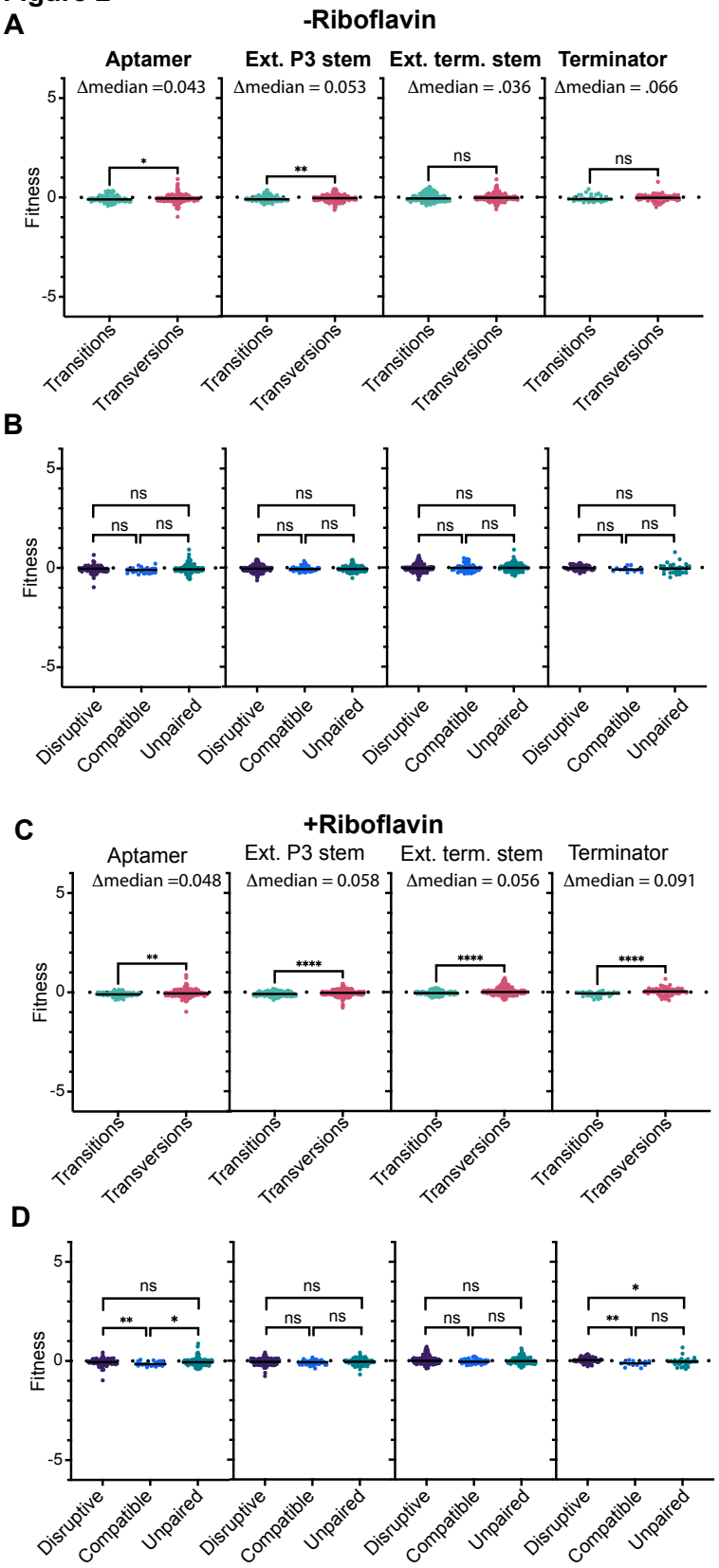
A
-Riboflavin
Aptamer
Ext. P3 stem
Ext. term. stem
Terminator
Δmedian =0.043
Δmedian = 0.053
Δmedian = .036
Δmedian = .066
Fitness
Transitions
Transversions
B
Disruptive
Compatible
Unpaired
C
+Riboflavin
Δmedian =0.048
Δmedian = 0.058
Δmedian = 0.056
Δmedian = 0.091
D

Figure 2: In the presence and absence of riboflavin most mutants are neutral. A) Fitness of transition versus transversion mutations by region of the riboswitch in the absence of riboflavin show only small magnitude differences between population medians. Populations compared using Mann-Whitney test ($p>0.05$ (ns), $p< 0.05$ (*), $p<0.01$ (**), $p<0.001$ (***), $p<0.0001$ (****)). Exact p-values for each comparison on **Table S2. B)** Fitness of mutations in the absence of riboflavin disruptive of predicted base-pairing (see Fig. 1F for predicted secondary structure), compatible with predicted base-pairing, or not predicted to be Watson-Crick paired (unpaired) do not show significant differences. Populations compared using a Kruskal-Wallis test with Dunn's test for multiple comparisons ($p>0.05$ (ns), $p< 0.05$ (*), $p<0.01$ (**), $p<0.001$ (***), $p<0.0001$ (****)). Population medians and exact p-values for each comparison on **Table S2. C)** Fitness of transition versus transversion mutations by region in the presence of riboflavin show only small magnitude differences between population medians. Statistical significance determined as in A. **D**) Fitness of mutations in the absence of riboflavin predicted to be disruptive, or compatible with predicted Watson-Crick base-pairing, or unpaired do not show large magnitude differences. Statistical significance determined as in B.

Figure 3

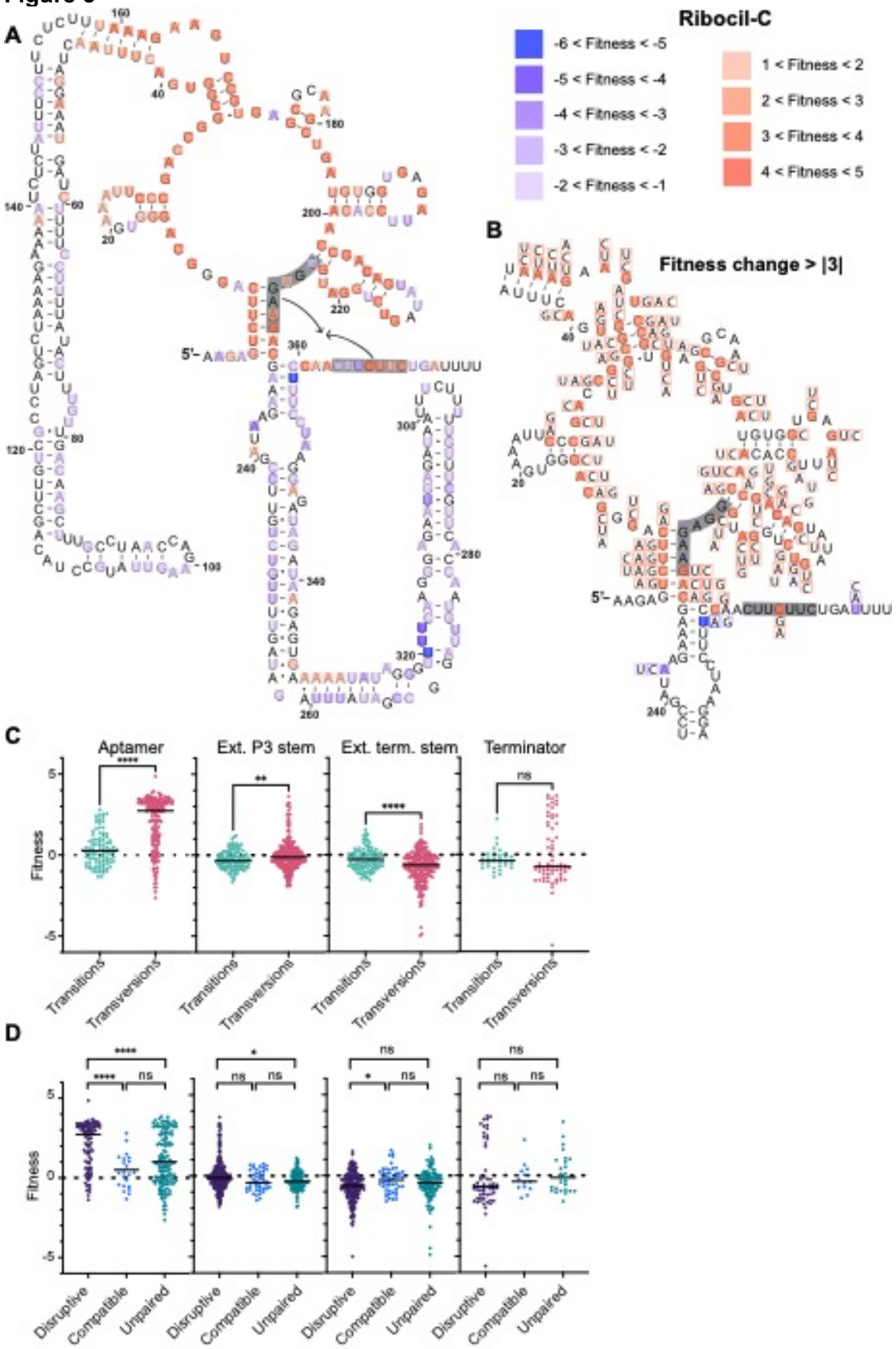
A
Ribocil-C
-6 < Fitness < -5
-5 < Fitness < -4
-4 < Fitness < -3
-3 < Fitness < -2
-2 < Fitness < -1
1 < Fitness < 2
2 < Fitness < 3
3 < Fitness < 4
4 < Fitness < 5
5'-AAGAG
B
Fitness change > |3|
5'-AAGAG
C
Aptamer
Ext. P3 stem
Ext. term. stem
Terminator
Fitness
Transitions
Transversions
D
Fitness
Disruptive
Compatible
Unpaired

Figure 3: Mutations to the aptamer are beneficial in the presence of ribocil-C. A) Heatmap of mutation fitness in ribocil-C relative to wild-type mapped to riboswitch secondary structure. Colors displayed correspond to the fitness of the mutation with largest magnitude change from wild-type. Blue: negative fitness, red: positive fitness. **B)** Inset of aptamer and proximal areas highlighting only positions with fitness changes > |3|, all mutations with a fitness change > |1| for each position are boxed in the color corresponding to the fitness change. **C)** Fitness of transition versus transversion mutations by region of the riboswitch in the presence of ribocil-C shows that transversions are positively selected compared to transitions in the aptamer and ext. P3 stem, and negatively selected in the ext. term. stem. Statistics as on **Fig. 2A. D)** Fitness of mutations in the aptamer disruptive of predicted base-pairing (see Fig. 1F for predicted secondary structure) are significantly more positive compared to those of compatible mutations or mutations to unpaired bases. Statistics and analysis as in **Fig. 2B**.

Figure 4

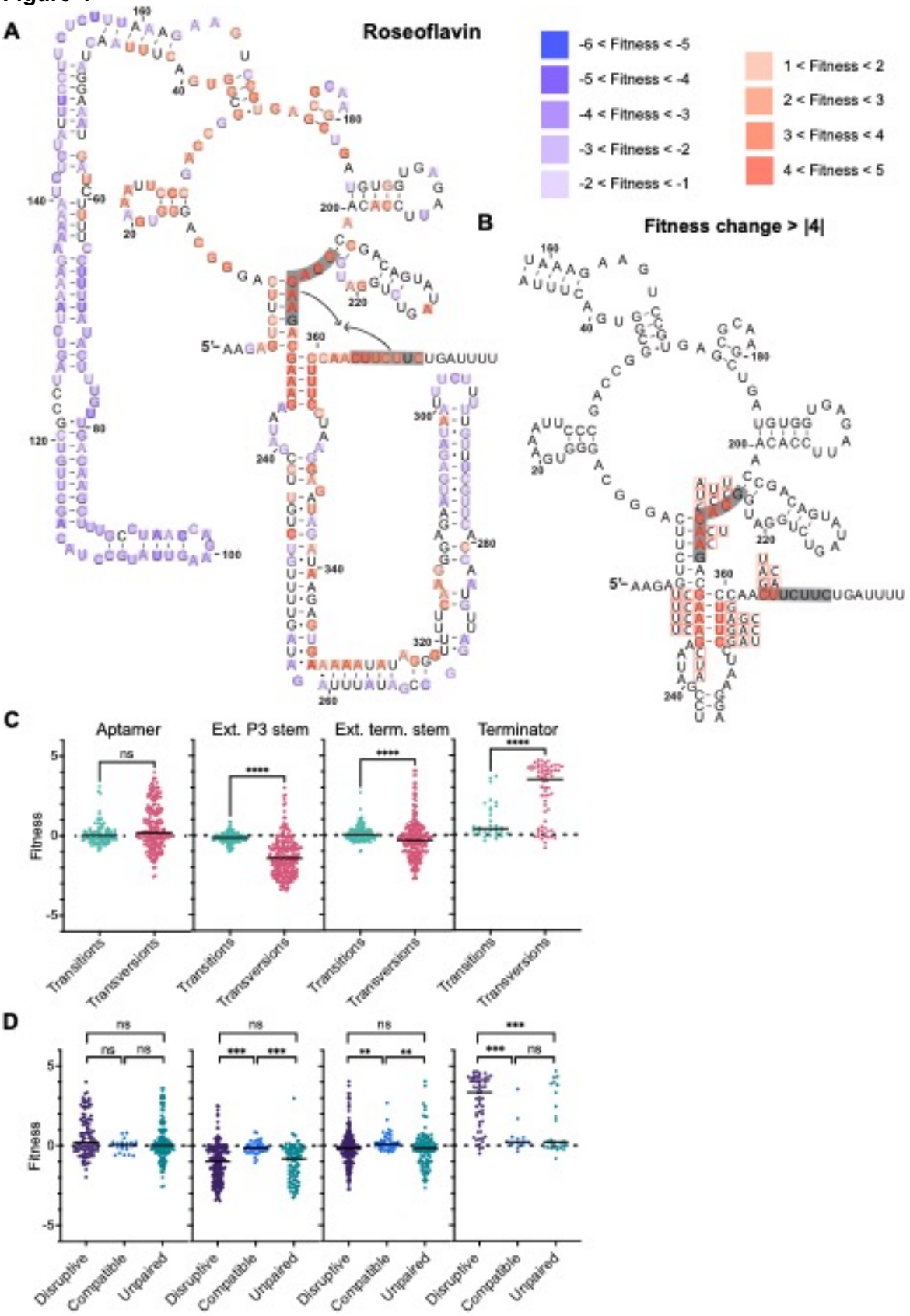
A
Roseoflavin
-6 < Fitness < -5
-5 < Fitness < -4
-4 < Fitness < -3
-3 < Fitness < -2
-2 < Fitness < -1
1 < Fitness < 2
2 < Fitness < 3
3 < Fitness < 4
4 < Fitness < 5
B
Fitness change > |4|
C
Aptamer
Ext. P3 stem
Ext. term. stem
Terminator
Fitness
Transitions
Transversions
D
Disruptive
Compatible
Unpaired

Figure 4: Mutations to the terminator region display the strongest benefit in the presence of roseoflavin. A) Heatmap of mutation fitness in roseoflavin relative to wild-type mapped to riboswitch secondary structure. Colors displayed correspond to the fitness of the mutation with largest magnitude change from wild-type. Blue: negative fitness, red: positive fitness. **B)** Inset of aptamer and proximal areas highlighting only positions with fitness changes > |4|, all mutations with a fitness change > |1| for each position are boxed in the color corresponding to the fitness change. **C)** Fitness of transition versus transversion mutations by region of the riboswitch in the presence of ribocil-C shows that transversions are negatively selected compared to transitions in both extended stems, and positively selected in the terminator region. Statistics as on **Fig. 2A**. **D)** Fitness of mutations in the terminator region disruptive of predicted base-pairing (see Fig. 1F for predicted secondary structure) are significantly more positive compared to those of compatible mutations or mutations to unpaired bases. Mutations disruptive to base-pairing or to unpaired bases display lower fitness compared to compatible mutations in both extended stems. Statistics and analysis as in **Fig. 2B**.

Figure 5

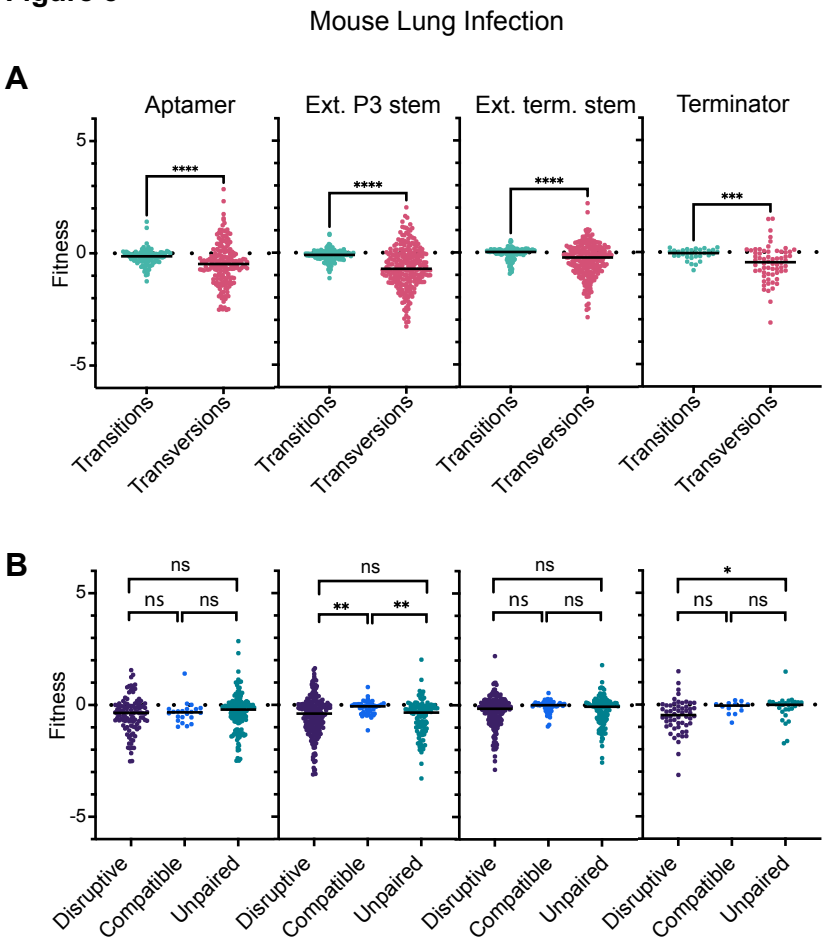

Figure 5: Mutations to all regions of the FMN riboswitch are deleterious during mouse *in vivo* lung infection. A) Fitness of transition versus transversion mutations by region of the riboswitch selected via mouse lung infection shows that transversions are negatively selected compared to transitions in all regions of the RNA. Statistics as on **Fig. 2A**. **B)** Fitness of mutations in the ext. P3 stem that are disruptive to predicted secondary structure or unpaired is significantly lower than that of compatible mutations Fitness of mutations that disrupt predicted base-pairing in the terminator is significantly lower than that for positions predicted to be unpaired. Statistics and analysis as in **Fig. 2B**.

Figure 6

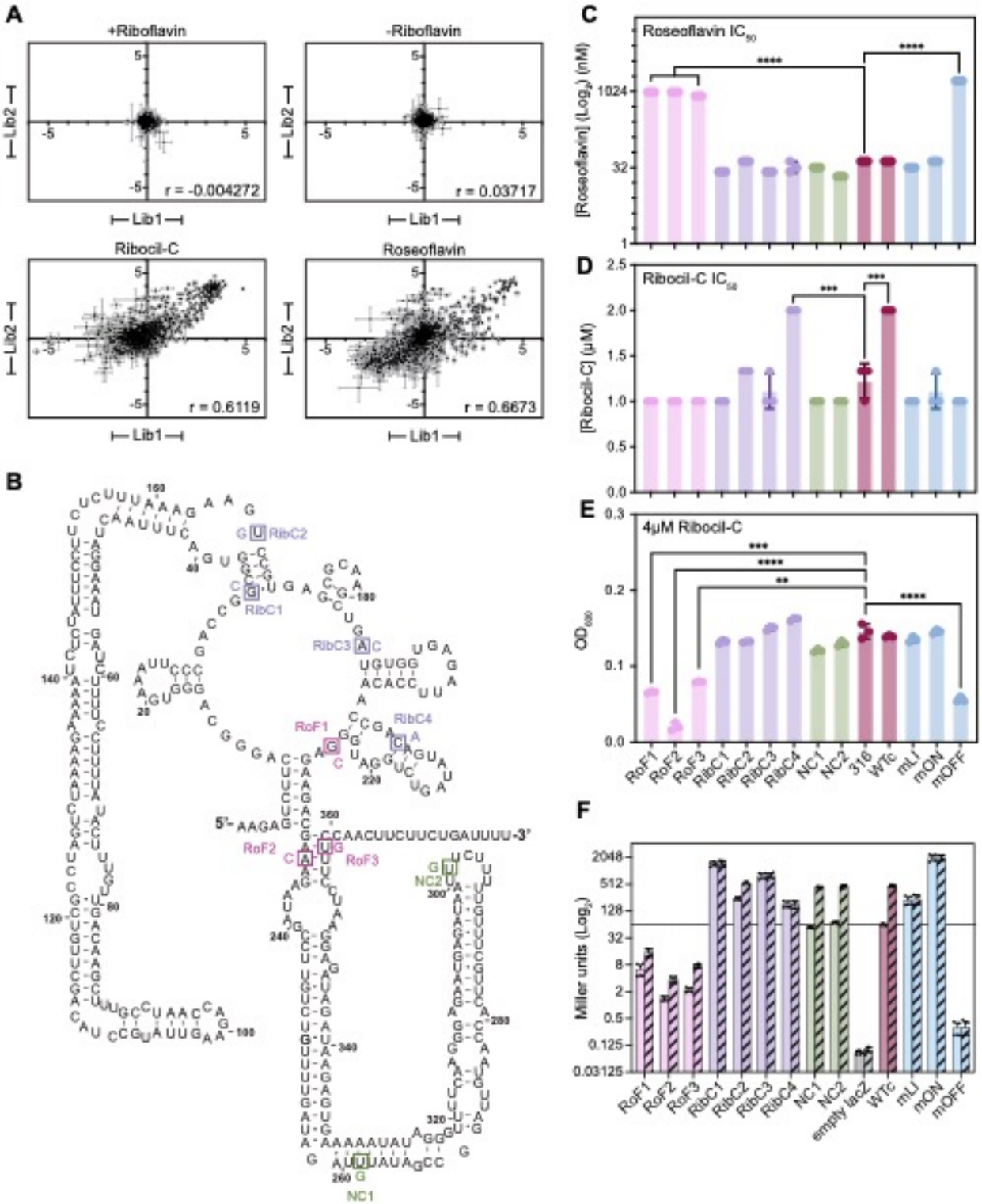
A
+Riboflavin
-Riboflavin
Ribocil-C
Roseoflavin
r = -0.004272
r = 0.03717
r = 0.6119
r = 0.6673
Lib1
Lib2
B
RibC1
RibC2
RibC3
RibC4
RoF1
RoF2
RoF3
NC1
NC2
5'-AAGAGCCAACUUCUUCUGAUUUU-3'
C
Roseoflavin IC50
[Roseoflavin] (Log2) (nM)
D
Ribocil-C IC50
[Ribocil-C] (µM)
E
4µM Ribocil-C
OD600
RoF1
RoF2
RoF3
RibC1
RibC2
RibC3
RibC4
NC1
NC2
316
WTc
mLI
mON
mOFF
F
Miller units (Log2)
empty lacZ

Figure 6: A biological replicate and assessment of selected mutants validates fitness values. A) Fitness values obtained from two distinct libraries, Lib1 and Lib2, plotted on x-y axis with reported error. In + and -riboflavin conditions, most mutations display fitness <|1|, and there is no significant correlation between values in the replicates. For +Ribocil-C and +Roseoflavin there is strong correlation between values obtained for Lib1 and Lib2 (Spearman's correlation co-efficient r shown on graph, $p < 0.0001$), with lower error on the positively selected values. **B)** Roseoflavin selected (RoF), Ribocil-C selected (RibC), and negative control (NC) mutations shown on the secondary structure of the SP_0488 FMN riboswitch. **C)** IC_{50} of roseoflavin for RoF, RibC, NC, and additional control strains shows that RoF mutants are resistant to roseoflavin. Pink: RoF selected mutants, purple: RibC selected mutants, green: negative controls, red: parental TIGR4 or WT^C control containing only Chl cassette, blue: previously published SP_0488 FMN riboswitch mutants[30]. Each point represents mean of technical triplicates. Data analyzed via a nested one-way ANOVA, compared to wildtype 316 (unmodified strain). All strains without significance indicated are not significant (ns) when compared to wildtype 316, ($p>0.05$ (ns), $p< 0.05$ (*), $p<0.01$ (**), $p<0.001$ (***), $p<0.0001$ (****)) **(See Table S3)**. **D)** IC_{50} of ribocil-C shows that positively selected strains are ultimately not strongly resistant. Replicates and data analysis as in part C above. **E)** Maximum OD_{600} for strains when grown with 4μM ribocil-C for 8.5 hours shows that RoF mutants, and mOFF are more susceptible than the unmodified strain to ribocil-C. Replicates and data analysis as in part C above. **F)** ß-galactosidase assays measure gene expression in the presence (solid bars) and absence (stripped bars) of riboflavin. Error bars represent the standard error of the mean (n=6). Empty lacZ (gray) indicates background levels of ß-galactosidase expression. In the presence of riboflavin, RoF mutations display lower gene expression compared to WT (dashed line at WT^C + riboflavin) and RibC mutations display higher gene expression. NC mutants do not display behavior statistically distinct from WT^C, but all other mutants display gene expression in both + and -riboflavin that is statistically different from WT^C via a Brown-Forsythe ANOVA followed by Dunnett's test of multiple comparisons (**Table S4**).

Figure 7

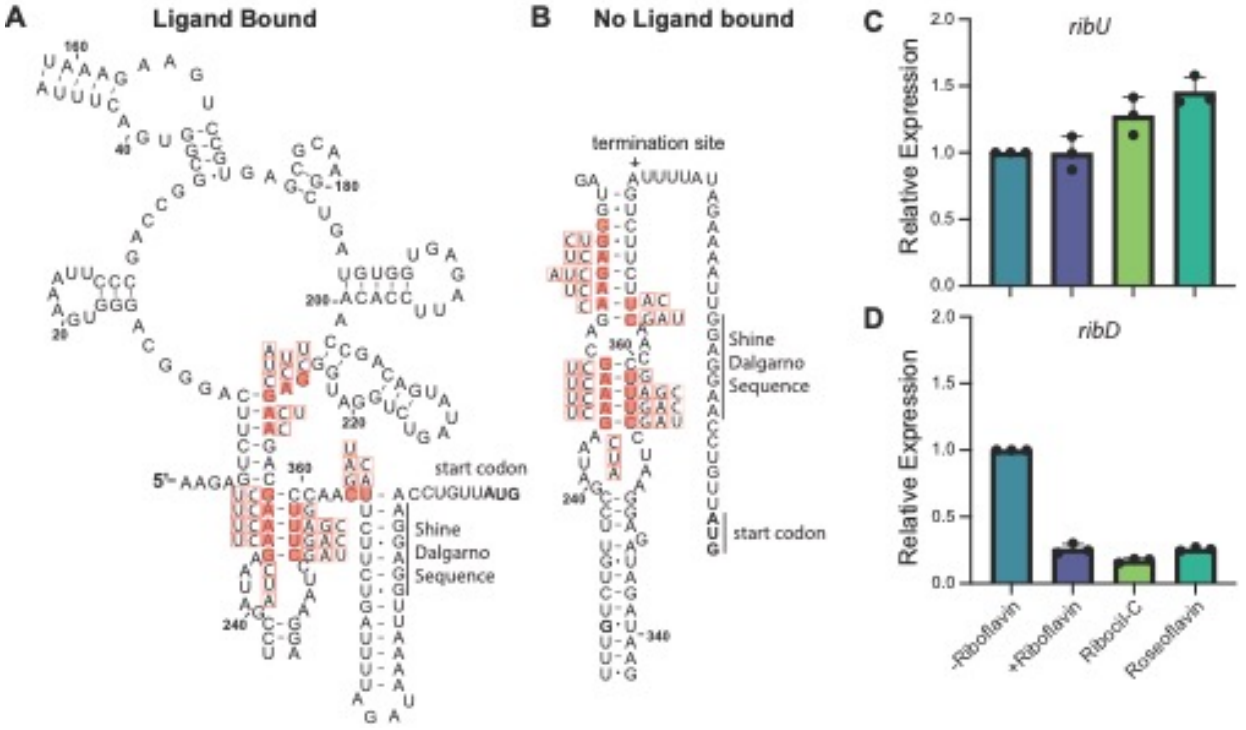

Figure 7: Constitutively repressed roseoflavin resistant mutants suggest a translation inhibition mechanism. A) Proposed structure for the ligand-bound form of the aptamer and downstream sequence including the Shine-Dalgarno and start codon indicated. Positions where a mutation results in fitness >4 in roseoflavin that likely display constitutively repressed expression are indicated. **B)** Proposed structure of the unbound form of the RNA where the prospective terminator forms, but the Shine-Dalgarno sequence is released. Positions where a mutation results in fitness >4 in roseoflavin that likely display constitutively repressed expression are indicated. All such mutations destabilize putative terminator stem, resulting in increased formation of the competing structure. **C)** qPCR analysis of *ribU* transcript in CDM -riboflavin, +riboflavin, +ribocil-C, and +roseoflavin shows minimal change in response to riboflavin and a slight increase in the presence of roseoflavin and ribocil-C. **D)** qPCR analysis of the *ribD* transcript, also regulated via an FMN riboswitch shows the anticipated repression of transcript in the presence of ligands for transcriptional regulation.