

ARTICLE

<https://doi.org/10.1038/s41467-020-14705-2>

OPEN

# In vitro ribosome synthesis and evolution through ribosome display

Michael J. Hammerling<sup>1,7</sup>, Brian R. Fritz<sup>1,7</sup>, Danielle J. Yoesep<sup>1</sup>, Do Soon Kim<sup>1</sup>, Erik D. Carlson<sup>1,6</sup> & Michael C. Jewett<sup>1,2,3,4,5\*</sup>

Directed evolution of the ribosome for expanded substrate incorporation and novel functions is challenging because the requirement of cell viability limits the mutations that can be made. Here we address this challenge by combining cell-free synthesis and assembly of translationally competent ribosomes with ribosome display to develop a fully in vitro methodology for ribosome synthesis and evolution (called RISE). We validate the RISE method by selecting active genotypes from a  $\sim 1.7 \times 10^7$  member library of ribosomal RNA (rRNA) variants, as well as identifying mutant ribosomes resistant to the antibiotic clindamycin from a library of  $\sim 4 \times 10^3$  rRNA variants. We further demonstrate the prevalence of positive epistasis in resistant genotypes, highlighting the importance of such interactions in selecting for new function. We anticipate that RISE will facilitate understanding of molecular translation and enable selection of ribosomes with altered properties.

<sup>1</sup>Department of Chemical and Biological Engineering, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA. <sup>2</sup>Center for Synthetic Biology, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA. <sup>3</sup>Simpson Querrey Institute, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA. <sup>4</sup>Chemistry of Life Processes Institute, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA. <sup>5</sup>Robert H. Lurie Comprehensive Cancer Center, Northwestern University, 2145 Sheridan Road, Evanston, Illinois 60208, USA. <sup>6</sup>Present address: Department of Chemical Engineering, Stanford University, Stanford, California 94305, USA. <sup>7</sup>These authors contributed equally: Michael J. Hammerling, Brian R. Fritz. \*email: [m-jewett@northwestern.edu](mailto:m-jewett@northwestern.edu)

The ribosome—the macromolecular machine that polymerizes  $\alpha$ -amino acids into polypeptides (i.e., proteins) according to messenger RNA (mRNA) templates—is the catalytic workhorse of the translation apparatus. The bacterial ribosome is made up of one small and one large subunit, the 30S and 50S, respectively. The 30S subunit is composed of 21 ribosomal proteins (r-proteins) and the 16S ribosomal RNA (rRNA), and is primarily responsible for decoding mRNA. The 50S subunit is composed of 33 r-proteins, the 23S rRNA, and the 5S rRNA, and serves to accommodate tRNA–amino acid monomers, catalyze peptide bond formation, and excrete polypeptides.

The extraordinary synthetic capability of the ribosome has provided the basis for recombinant DNA technology. Recently, engineering mutant ribosomes to elucidate key principles underpinning the process of translation, program cellular function, and generate novel sequence-defined polymers has emerged as a major opportunity in chemical and synthetic biology<sup>1–7</sup>. Although efforts to engineer ribosomal function have achieved some success<sup>8,9</sup>, especially with the recent advance of tethered ribosome systems<sup>10–12</sup>, a significant repurposing of the translation apparatus is constrained by challenges inherent to altering translation in living cells. A completely in vitro approach could overcome cell viability constraints to repurpose the ribosome into a re-engineered machine for understanding, harnessing, and expanding the capabilities of the translation apparatus.

In this study, we present a cell-free ribosome synthesis and evolution method called RISE. RISE combines an in vitro integrated synthesis, assembly, and translation (iSAT) system with ribosome display<sup>13–15</sup> (Fig. 1). Specifically, a library of rRNA sequence variants is transcribed from plasmid DNA and assembled into a library of ribosomes. Meanwhile, a truncated mRNA encoding a selective peptide is also transcribed in the same reaction. Nascent functional ribosomes within the library then translate the selective peptide and stall to form mRNA–ribosome–peptide ternary complexes. These complexes are selectively captured by the peptide, rRNA is isolated and reverse transcribed, and the resulting ribosomal DNA enters another cycle of RISE for further enrichment of active sequences. In addition to showing that RISE can enrich active ribosomal genotypes from a library of  $\sim 1.7 \times 10^7$  members, we evolve ribosome resistance to the antibiotic clindamycin from a more

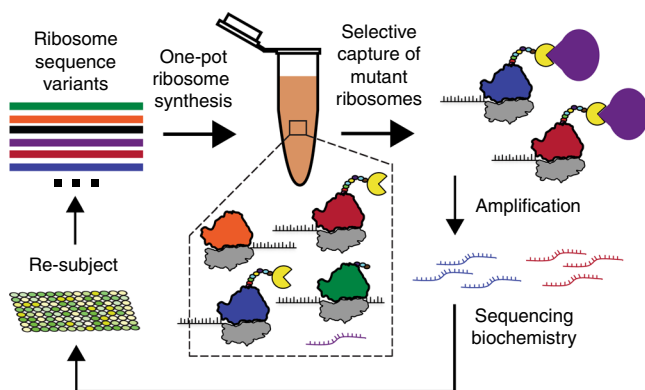
targeted library. Deep sequencing analysis from the clindamycin selection reveals a densely connected network of winning genotypes, which display an overall trend of positive epistasis, suggesting the importance of epistatic interactions in evolving ribosome function.

## Results

**Optimization of ribosome display coupled with iSAT.** We hypothesized that we could develop RISE by linking iSAT with ribosome display. As a model we focused on *Escherichia coli* ribosomes because the translation apparatus of *E. coli* is the best understood and most characterized, both biochemically and genetically<sup>16</sup>. Previously, non-physiological conditions and low reconstitution efficiencies from in vitro-transcribed 23S rRNA represented one of the most serious bottlenecks to constructing and evolving *E. coli* ribosomes in vitro<sup>17</sup>. We recently addressed these bottlenecks by developing an integrated method for the physiological assembly of *E. coli* ribosomes called iSAT<sup>16,18–20</sup>. iSAT enables efficient one-step co-activation of rRNA transcription, assembly of transcribed rRNA with native r-proteins into *E. coli* ribosomes, and synthesis of functional protein by these ribosomes in a ribosome-free S150 crude extract. Importantly, iSAT ribosomes possess  $\sim 70\%$  of the protein synthesis activity of in vivo-assembled *E. coli* ribosomes<sup>18</sup>, which could enable the selection and evolution of ribosomes if the approach were connected to ribosome display.

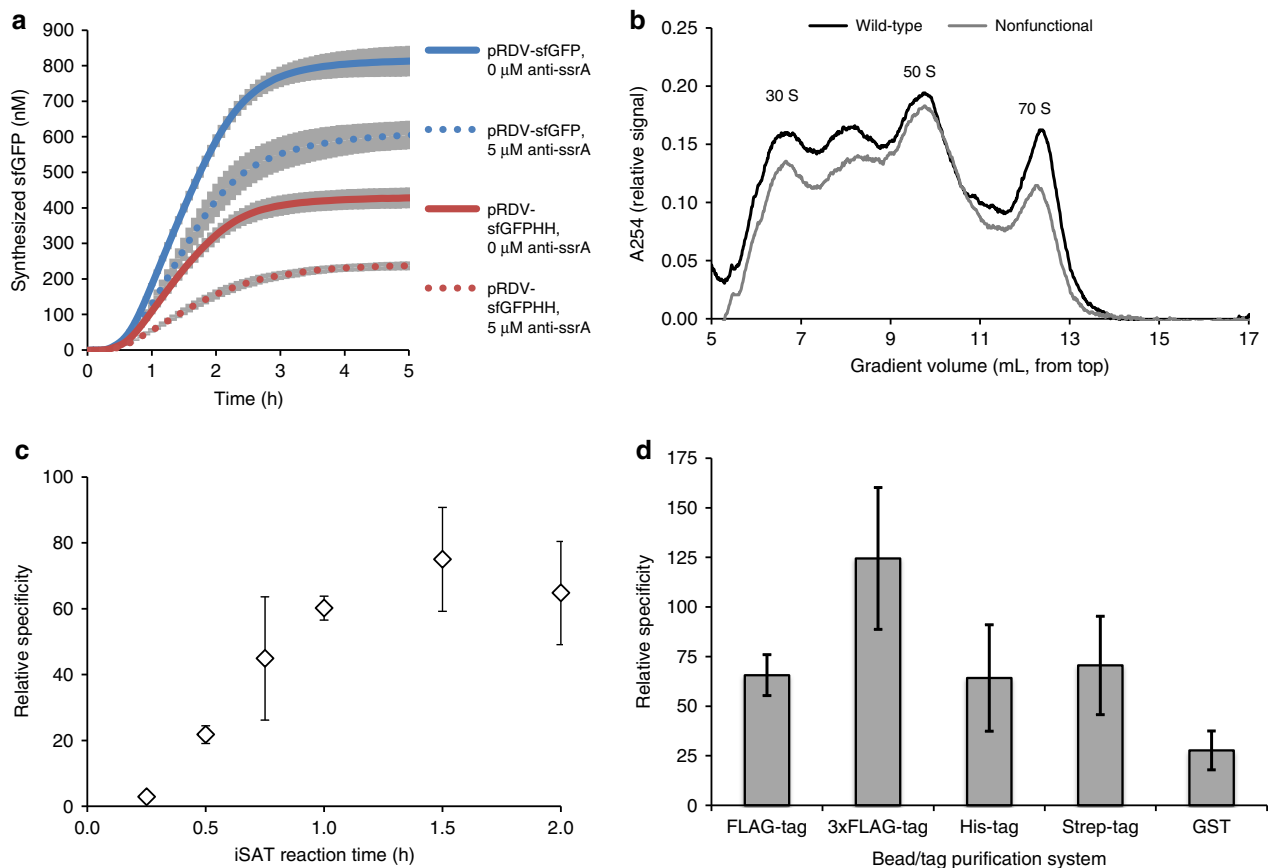
The development of RISE required the optimization of all aspects of the ribosome display method, including stable ribosome stalling, ribosome capture, and recovery of ribosomal coding DNA (cDNA). We began our investigation by attempting to create stalled mRNA–ribosome–peptide ternary complexes. To assess iSAT ribosome stalling, we inserted the gene for superfolder green fluorescent protein (sfGFP) upstream of the spacer sequence in the pRDV vector developed for ribosome display by the Plückthun lab<sup>13,14</sup>. The goal was to gain insight into reaction dynamics and quantify sfGFP production in our iSAT reactions as a proxy for properly stalled ribosomes. Properly stalled iSAT ribosomes should generate one sfGFP molecule for each translating ribosome, with a maximum concentration of 300 nM ribosomes based on the maximal total r-protein concentration in iSAT reactions. Importantly, to ensure that the stop codon of sfGFP did not lead to ribosome release, the pRDV vector was modified to include the self-cleaving hammerhead (HH) ribozyme after the peptide-coding region, which cleaves off the stop codon and generates a truncated mRNA<sup>18,21</sup>. We tested constructs with and without the HH ribozyme in iSAT in the presence and absence of 5  $\mu$ M anti-ssrA oligonucleotide (Fig. 2a). The addition of anti-ssrA oligonucleotides is routinely used in ribosome display to prevent the recycling of stalled ribosomes via the transfer-mRNA encoded by the gene *ssrA*<sup>22,23</sup>. The combination of the HH gene element and the anti-ssrA oligonucleotide reduced translation to below 300 nM of sfGFP, suggesting that ribosomes are efficiently stalling.

To assess the specificity of RISE for selecting functional variants, we next carried out RISE with a mock selection whereby reactions contained either wild-type (WT) or nonfunctional rDNA operons that included lethal point mutations in both the 16S and 23S rRNA ( $\Delta$ C967 and G2252A, respectively)<sup>24,25</sup>. These mutations resulted in an iSAT translation activity of  $<2\%$  of the iSAT activity from WT rDNA operons. From these reactions, we determined relative specificity of different RISE conditions by dividing the relative capture of functional ribosomes by the relative capture of nonfunctional ribosomes, as determined by reverse transcription–quantitative PCR detection of 23S rRNA in the eluates (Supplementary Fig. 1). Sedimentation analysis



**Fig. 1** Diagram of in vitro ribosome synthesis and evolution (RISE).

Starting with a T7-promoted rDNA operon, mutations are introduced to generate an rDNA operon library. The rDNA operon library is then used in an iSAT reaction to assemble a library of ribosomes. Functional ribosomes then translate selective peptides and stall to form ternary complexes, allowing for selection of the peptides and, thus, isolation of functional ribosomes. The rRNA of the functional ribosomes is then purified and reverse transcribed into cDNA for sequencing and reinsertion into the rDNA operon.



**Fig. 2** Development and optimization of RISE selection conditions. **a** iSAT translation of sfGFP under stalling conditions. pRDV-sfGFP plasmids were made with (red) or without (blue) a 3' HH ribozyme for processing of transcribed mRNA to remove the stop codon. Anti-ssrA oligonucleotide was included at either 0  $\mu$ M (solid) or 5  $\mu$ M (dotted). **b** Sedimentation analysis of iSAT reactions with wild-type or nonfunctional rDNA operon. Peak identities are labeled. **c** Relative specificity of capture by anti-FLAG magnetic beads for iSAT reactions displaying FLAG-tag. Reactions were incubated from 15 min to 2 h. **d** Comparison of relative specificity of various bead/tag purification systems for use in RISE. Wash methods were held constant for comparison, and reactions were incubated for 1.5 h, except for reactions with GST, which were incubated for 2 h. For **a**, **c**, and **d**, values represent averages of three independent reactions ( $n = 3$ ), while shading in **a** and bars in **c** and **d** represent 1 SD.

supported this approach by demonstrating that nonfunctional rRNA still formed native-like ribosomal particles to preserve potential nonspecific interactions (Fig. 2b).

With the observations of efficient ribosome stalling from actively translating ribosomes, we set out to optimize the reaction duration, which has been shown to be critical for optimizing capture of stable ternary complexes<sup>26</sup>. We observed that iSAT reactions require ~30 min for rRNA synthesis, ribosome assembly, and translation of detectable levels of sfGFP (Fig. 2a). By varying iSAT reaction times, we observed that relative RISE specificity is highest at 1.5 h (Fig. 2c). Visualization of the recovered nucleic acids from 1.5 h reactions shows that reactions with nonfunctional rDNA operon plasmids do not show visible rRNA capture, whereas reactions with functional WT rDNA operon plasmids show bands representative of 23S and 16S rRNAs (Supplementary Fig. 2).

The key conceptual shift in RISE from conventional ribosome display is that rRNA sequences are selected rather than mRNA sequences. As ribosome mutants are the focus, we next explored conditions for optimal ribosome capture. The idea was to synthesize an N-terminal peptide handle (e.g., the FLAG-tag) to emerge first from the ribosome and be displayed for affinity purification and capture. Several common selective peptide tags were tested, including the FLAG-tag, 3xFLAG-tag, His-tag, Strep-tag, and glutathione S-transferase (GST), in the pRDV vector using commercially available magnetic capture beads. iSAT

reactions expressing each tag or protein were incubated for 1.5 h, except for GST, which was incubated for 2 h to account for the additional translation and folding times associated with expressing a large selective protein. Given the results of this screen, we elected to use the 3xFLAG-tag peptide in RISE reactions, which exhibited a 124-fold specificity for functional over inactive ribosomes (Fig. 2d).

We next optimized binding and wash conditions for the captured ternary complex displaying the 3xFLAG-tag. Addition of bovine serum albumin (BSA), Tween<sup>®</sup> 20, or heparin to the binding and/or wash buffers were tested for their ability to decrease nonspecific binding between the magnetic bead and the 3xFLAG-tag. Of these additives, only the presence of BSA was found to increase relative specificity for active ribosomes by 29% (one-tailed Welch's *t*-test,  $p = 0.042$ ) (Supplementary Fig. 3a, b and Supplementary Table 3). More importantly, we found that RISE specificity was substantially improved by increasing the number of washing steps from 5 to 10 ( $p = 0.048$ ) (Supplementary Fig. 3c, d). Although we observed a decrease in specific capture of 29% by increased wash stringency ( $p = 0.030$ ), nonspecific capture was reduced by 95% ( $p = 0.016$ ), resulting in a relative specificity of >1000-fold for WT vs. inactive ribosomes (Supplementary Fig. 3d).

To complete RISE, we developed a strategy to reverse-transcribe recovered rRNA sequences and reinsert the recovered cDNA into the operon plasmid. Given constraints in reverse

transcribing full-length rRNA, which contains challenging post-transcriptional modifications in some regions<sup>27,28</sup>, initial focus was given to manipulation of only the peptidyl transferase center (PTC) encoded by the 23S rRNA. We used reverse-transcription PCR (RT-PCR) to recover a 660 bp region of the 23S rRNA gene (bases 1962 to 2575, *E. coli* numbering), which contains as many as 13 posttranscriptional modifications, but none of which block reverse transcription by disrupting base pairing. Off-site type IIs restriction digestion and ligation was used to insert the 660 bp fragment from the 23S rRNA gene into pT7rrn $\Delta$ 660, an rDNA operon plasmid lacking the amplified gene fragment. Assembled libraries were transformed into cells to maintain a stable plasmid stock and this plasmid pool was used in the next round of in vitro selection.

**Ribosomal mutants in the peptidyl transferase center.** To demonstrate that RISE can recover translationally competent ribosomes from a large library of variants, we initially performed two cycles of selection on a library consisting of degenerate mutations to bases surrounding the PTC of the ribosome. To improve the likelihood of discovering diverse variants from a large pool, we chose 12 nucleotides of the 23S rRNA spanning positions 2461–2489 (Supplementary Table 4), which are not conserved (12NC) between the bacterial species *E. coli*, *Bacillus stearothermophilus*, and *Thermus aquaticus*: organisms for which in vitro reconstitutions are commonly performed<sup>29–31</sup>. This resulted in a library with a theoretical diversity of  $\sim 1.7 \times 10^7$  members.

RISE was applied to this library over two cycles. iSAT activity tests of the libraries were performed, and recovered operon pools show an increase in protein synthesis after each RISE cycle (Supplementary Fig. 4a). Sequence traces of the initial library show degeneracy of the mutated bases, but we observed rapid convergence toward the WT sequence after two RISE cycles (Supplementary Fig. 5). Isolation and sequencing of rDNA of selected variants from the 12NC library revealed that 20 of 26 evolved sequences of the 12NC pool had reverted to WT. We performed iSAT for the remaining variants, finding that five of the six tested variants showed >50% WT iSAT activity (Supplementary Fig. 4b). The number of point mutations in these variants ranged from 1 to 11 mutations from among the 12 degenerate bases, demonstrating that the PTC is amenable to substantial restructuring of certain regions, while maintaining translational activity (Supplementary Table 4).

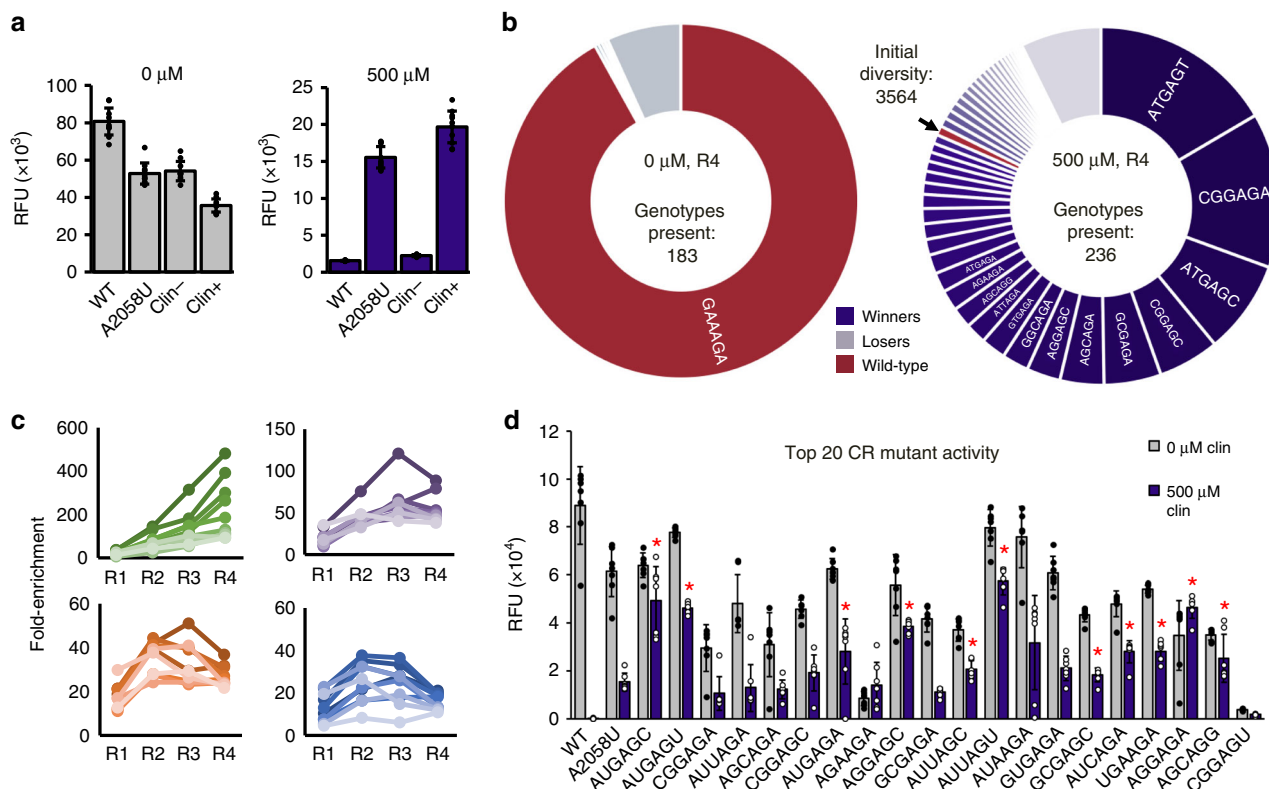
**Directed evolution of clindamycin resistance using RISE.** With the RISE platform at hand, we next sought to use it to select and evolve mutant ribosomes for new functions. We chose resistance to the ribosome-binding antibiotic clindamycin as a model<sup>32</sup>. Previously, Cochella and Green<sup>32</sup> demonstrated that ribosome display can be used for the selection of functional ribosomes from a ribosome library, but their method required in vivo ribosome synthesis and ribosome purification. This method limits the possible library size due to the bottleneck of transformation in cells, biases the library toward only those that assemble and are not dominantly lethal in vivo, and limits throughput due to the requirement to separate mutant ribosomes from WT ribosomes<sup>32</sup>. RISE avoids these bottlenecks, which should facilitate the development of rapid, high-throughput selection strategies for ribosome evolution. We developed a clindamycin-resistance (CR) library by randomizing positions 2057 to 2062 of the 23S rRNA, a region which contributes several binding interactions with clindamycin<sup>33</sup> (Supplementary Fig. 6). Importantly, this contiguous region extends away from the PTC and around the nascent peptide exit tunnel, suggesting mutations to this region may not

disrupt peptide bond formation (Supplementary Fig. 7). We then applied RISE to the CR library in the presence of 500  $\mu$ M clindamycin in iSAT reactions. As a control, we carried out selections in the absence of clindamycin to evaluate the impact of selection pressure on the RISE method. Before selection, the libraries demonstrated no detectable protein synthesis due to the prevalence of inactive ribosomal variants. After four rounds of selection, the library selected in the absence of clindamycin (Clin–) demonstrated 1.5-fold greater bulk protein synthesis than the library selected in the presence of clindamycin (Clin+) (Fig. 3a, gray bars). Conversely, the Clin+ library improved to have 8.7-fold greater bulk protein synthesis activity after the completion of four RISE cycles in the presence of clindamycin than the Clin– library (Fig. 3a, purple bars). These results are commensurate with the conditions under which the libraries were selected and demonstrate the ability of RISE to select ribosomal genotypes that are tailored to functioning optimally under a given selective challenge.

**Evolutionary dynamics of clindamycin-resistance selection.** To achieve high-resolution analysis of selection dynamics, we performed deep sequencing of the initial CR library and at each round of selection. We observed 3935 of 4096 possible genotypes (96%) during the selection but reduced the number considered in our analysis to 3564 using our quality control criteria (Methods and Supplementary Fig. 8). The 0  $\mu$ M selection resulted in a rapid convergence of the library toward the WT 23S rRNA sequence, although 183 genotypes were still present by the end of the selection (Supplementary Fig. 9). In contrast, the 500  $\mu$ M selection resulted in the depletion of WT 23S rRNA sequence and a more heterogeneous overall pool, which converged to 236 genotypes by the final round of selection. From four rounds of RISE on the CR library, the ten most abundant of these genotypes (which comprised 0.38% of the initial library) made up  $\sim 64\%$  of the final population, indicating rapid convergence to a handful of the most active CR genotypes (Fig. 3b and Supplementary Fig. 10). Selection dynamics of the top 40 genotypes show the expected trends of a consistent selection pressure, with the fittest genotypes increasing in frequency throughout the course of the selection (top left panel), whereas moderately fit genotypes increase initially, then decrease in frequency in later rounds as the average fitness in the population increases (Fig. 3c). To re-assess optimal RISE incubation time, but now with finer resolution than before (Fig. 2c), we performed time-course RISE reactions in which CR genotypes were enriched from a degenerate library in 500  $\mu$ M clindamycin. We then compared enrichment of each genotype after four rounds of selection against its enrichment when the first round of selection was collected at time points from 7.5 to 90 min (Supplementary Fig. 11). Supporting our initial bulk data, deep sequencing analysis confirmed efficient selection occurs for 90 min time points.

A simple method for calculating expected final frequency of a genotype in the population that ignores interactions between positions entails multiplying the final frequency at each selected position together to obtain an expected final frequency for the genotype. Many winning sequences differ substantially from their expected frequency. For example, the consensus “winner” 5'-AUGAGA-3' sequence is at only 33% the expected final frequency based on this method, although it is still present in the top ten sequences by fold-change. In contrast, two other top sequences, 5'-GCGAGC-3' and 5'-UGAAGA-3', were found at 7.9-fold and 18.0-fold higher frequency than expected (Supplementary Fig. 12), respectively. These unexpected winners highlight the importance of creating diverse libraries that sample broad sequence space to discover difficult-to-predict





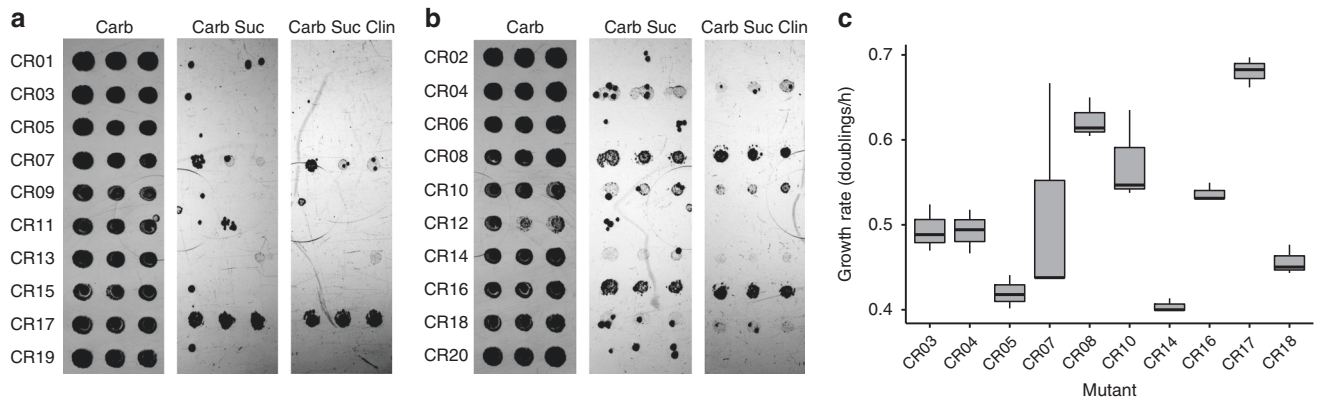
**Fig. 3 Results of selection of CR ribosomes in RISE.** After four rounds of selection in the absence (0  $\mu\text{M}$ ) or presence (500  $\mu\text{M}$ ) of clindamycin, the resulting selected libraries were analyzed. **a** In 0  $\mu\text{M}$  clindamycin (gray bars), WT ribosomes maintain the highest expression of sfGFP in iSAT. The A2058U mutant and Clin<sup>-</sup> selection pool also have high activity, whereas the Clin<sup>+</sup> selection pool has the lowest activity, although all samples express sfGFP. In contrast, in 500  $\mu\text{M}$  clindamycin (purple bars), the WT and Clin<sup>-</sup> conditions show almost no sfGFP expression, whereas the A2058U mutant and Clin<sup>+</sup> selection pool retain robust expression. Notably, the results of the Clin<sup>+</sup> selection pool have a higher activity than the control A2058U mutant. **b** Donut plot depicting the frequency of the top mutants ranked by final frequency after four rounds of selection in 0  $\mu\text{M}$  (left) and 500  $\mu\text{M}$  (right) clindamycin. Although WT has nearly taken over the population in the 0  $\mu\text{M}$  case, the 500  $\mu\text{M}$  case retains genotypic diversity, and WT is being rapidly supplanted by CR genotypes. WT is indicated with an arrow on the 500  $\mu\text{M}$  plot. **c** Selection dynamics of top genotypes 1–10 (green), 11–20 (purple), 21–30 (orange), and 31–40 (blue) in the 500  $\mu\text{M}$  selection as ranked by final fold enrichment. Selection dynamics are consistent with a constant selection pressure where very fit genotypes increase in frequency throughout the course of the selection, whereas moderately fit genotypes increase initially but begin to decrease as the average fitness of the population surpasses their fitness. **d** The activity of WT, A2058U, and the top 20 genotypes from the Clin<sup>+</sup> selection as ranked by fold enrichment at the end of the selection. All but one mutant is active in iSAT, and the majority (11 of 20) have higher activity in the presence of clindamycin than the A2058U mutant discovered in other studies (Welch’s one-sided *t*-test with Benjamini and Hochberg correction). Asterisk (\*) indicates *p*-value < 0.05. Values represent averages of ten iSAT independent reactions (*n* = 10) for **a** and seven independent iSAT reactions (*n* = 7) for **d**. Error bars represent 1 SD. Exact *p*-values for **d** are reported in Supplementary Table 5.

sequences that may have high activity in the given selective challenge.

After completion of the in vitro evolution of the CR library, the top 20 genotypes from the 500  $\mu\text{M}$  condition (ranked by their fold enrichment at the end of the selection) were individually cloned and tested in iSAT. All but one genotype translated sfGFP in the presence of 500  $\mu\text{M}$  clindamycin, indicating successful selection for CR (Fig. 3d). A majority (11 of 20) were more active in iSAT reactions with clindamycin compared to the known A2058U mutation previously demonstrated to confer CR in vivo and also shown in Fig. 3d. Surprisingly, none of our most active CR mutants were observed previously in in vitro CR selections<sup>32</sup>.

**In vivo testing of mutants selected in the RISE platform.** To understand why some of these sequence variants had not been previously observed, we assessed whether the RISE platform discovered variants that were non-viable in vivo. To do so, we exchanged the T7 promoters of the six non-WT evolved 12NC variants and the top 20 CR operon variants with *P<sub>L</sub>* promoters for in vivo transcription by native polymerases. These constructs

were then transformed into the Squires strain, which lacks rRNA sequences on the genome and instead derives rRNA from a plasmid-based operon. The Squires strain serves as a testing platform for whether an rRNA sequence (encoded by the plasmid) can support life, and generally if such rRNA sequences are compatible with cell growth<sup>34</sup>. Transformants were plated on: (i) carbenicillin; (ii) carbenicillin and sucrose; and for CR mutants, (iii) carbenicillin, sucrose, and clindamycin, to assess the ability of these ribosomes to support life (Fig. 4a, b and Supplementary Fig. 13). Carbenicillin selects for the presence of our variant rDNA operon plasmid, sucrose counter-selects the WT rDNA already present in the Squires strain (through the SacB negative selection marker), and clindamycin tests if the CR phenotype also transfers to in vivo settings. The 12NC and CR variants showed no inherent toxicity to the cells when plated on carbenicillin alone, but 5 of 6 12NC variants and only 8 of 20 CR variants were clearly capable of supporting cell growth upon loss of the WT rDNA operon on sucrose and clindamycin agar, suggesting that the other variants do not form viable ribosomes in vivo. However, two mutants (CR03, CR05) that were inviable on plates supported



**Fig. 4 Viability of CR ribosome variants in vivo.** Isolated clindamycin-resistant (CR) rDNA variants were altered to include a native promoter for in vivo expression. **a, b** The top 20 variants were transformed into the Squires strain, and cells were spotted on (left) LB with 100  $\mu\text{g}/\text{mL}$  carbenicillin, (center) LB with 100  $\mu\text{g}/\text{mL}$  carbenicillin and 5% w/v sucrose, and (right) LB with 100  $\mu\text{g}/\text{mL}$  carbenicillin, 5% w/v sucrose, and 350  $\mu\text{g}/\text{mL}$  clindamycin. In the left condition, cells are expected to grow unless the ribosomes are dominant-lethal, and all spots grow as expected. In the center condition, the WT ribosomal operon should be lost, and cells should only grow if the mutant ribosome can support life or if successful escape mutants evolve. Finally, in the right condition, escape mutants are unlikely due to the additional constraint of clindamycin resistance, and spots should only grow if the CR mutant provided can support life. Mutants CR04, CR07, CR08, CR10, CR14, CR16, CR17, and CR18 are all observed to support life on plates. Results are a representative example of three independent experiments. **c** The mutants were also grown in LB with 100  $\mu\text{g}/\text{mL}$  carbenicillin, 5% w/v sucrose, and 350  $\mu\text{g}/\text{mL}$  clindamycin liquid cultures in three independent biological replicates ( $n = 3$ ). An additional two mutants, CR03 and CR05, support cell growth in this liquid medium compared with solid medium, although with a very low growth rate. These mutants are arranged in decreasing order of final fold enrichment in the RISE selection from left to right. Notably, final fold enrichment does not predict growth rate in vivo, supporting the notion that RISE can select ribosome variants that would not be favored in in vivo selections. For box plots, the bottom and top hinges denote the 25% and 75% quantile, whiskers extend to 1.5 times the interquartile range, the diamond represents the mean, the center line is the median, and outliers are black points.

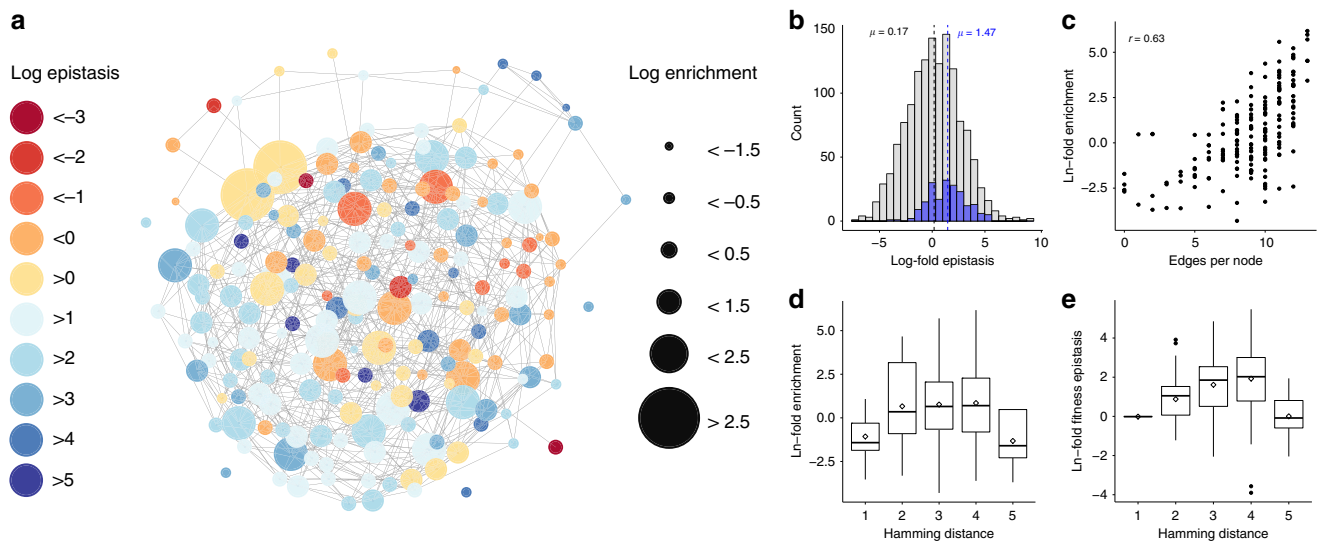
cell growth in liquid culture (Fig. 4c). Of the CR mutants that were viable in vivo, there was no clear relationship between fold enrichment after RISE and cellular growth rate. Interestingly, the non-viable CR variants all contained mutations at the 2062 position, which is consistent with the previous observation that such mutations are lethal despite in vitro ribosome activity<sup>32,35</sup>. The result that ribosomes that are active in iSAT may be unable to support life demonstrates the potential of RISE to select ribosomes that are proficient at translating challenging templates or monomers in vitro without concern for viability in living cells.

**Patterns of fitness and epistasis in evolved clindamycin-resistant ribosomes.** We next sought to understand the role that epistatic interactions played in our most active genotypes. Epistasis—the nonadditive impact of combinations of mutations within or between genes—is broadly believed to play an important role in the evolution of biomolecules and biological systems<sup>36</sup>. Despite this, little computational work<sup>37</sup>, and to our knowledge, no experimental work has been done to measure the degree to which epistatic interactions impact ribosome function. The combination of RISE with deep sequencing analysis allowed us to estimate epistasis values of genotypes for examining general trends and identifying unexpected functional sequences. We calculated an expected enrichment for each genotype assuming an independent contribution of each nucleotide to the fitness of CR mutants (Methods).

To gain both an intuitive and a quantitative understanding of the relationship between fitness, epistasis, and relatedness in our CR fitness landscape, we visualized all surviving CR mutants as a network plot of point-mutant-adjacent genotypes (Fig. 5a). This network has several salient features that suggest that RISE is an efficient selection strategy for enriching genotypes in a manner commensurate with their fitness in the selection. If the impact of each individual residue on overall ribosome fitness is independent, the actual enrichment of each mutant will correlate closely to this expected enrichment, and the log-transformed epistasis

values will be close to zero. The distribution of values for those genotypes for which epistasis could be calculated is approximately normally distributed (Shapiro–Wilk test,  $p = 0.117$ ) with a mean slightly above zero (one-sided Student’s  $t$ -test,  $p = 0.005$ ) (Fig. 5b, gray bars). This slight positive bias may result from the fact that epistasis values could not be calculated for genotypes that dropped out of the population after one round of selection. Average log-fold enrichment is similar for genotypes 2–4 mutations away from WT (i.e., Hamming distance), although the seven fittest variants all have three to four differences (Fig. 5d). Average epistasis increases with greater distance from WT up to a Hamming distance of 4 (Fig. 5e). Mutants with a Hamming distance of 5 do not conform to the general trend on either plot, likely because genotypes containing 5 mutations necessarily contain a mutation at positions 2060 or 2061, which are important for translation function and are nearly completely converged to WT in our selection (Supplementary Fig. 10).

Edges per node is a measure of the relatedness of a winning genotype to other winning genotypes in the final CR population. This measure correlates well with enrichment (Pearson’s  $r = 0.63$ ), indicating that genotypes that are closely related to other genotypes in the final population are most likely to have high fitness (Fig. 5c). The high degree of relatedness between the most enriched genotypes is encouraging, as it implies that RISE efficiently selects on some shared property of these closely related genotypes—presumably translational fitness. This is further evidenced by the high measured translational activity of these genotypes (Fig. 3d). In contrast, edges per node does not correlate with epistasis (Pearson’s  $r = -0.09$ )—indicating that relatedness of genotypes is not predictive of their propensity for epistatic interactions (Supplementary Fig. 14). Finally, we observed a highly significant increase (one-sided Student’s  $t$ -test,  $p = 2.2 \times 10^{-16}$ ) in average epistasis values in the winning CR genotypes compared with the overall population, indicating that positive epistasis plays an important role in determining the fittest genotypes in this selection (Fig. 5b). As positive epistasis confounds the prediction of genotype fitness, this result highlights the value of a selection



**Fig. 5** Quantitative analysis of fitness and epistasis of the network of top clindamycin-resistant genotypes. **a** Network plot of clindamycin-resistant genotypes from RISE selection. In this network plot, each node represents a surviving genotype after four rounds of selection in 500  $\mu\text{M}$  clindamycin. Edges connect nodes that are a single point mutation away from each other. The color of the node represents the calculated epistasis value for that genotype. The size of the node represents the  $\log_{10}$  of the final enrichment at the end of the selection. This transformation was performed to enable depiction in the network plot on a scale appropriate for publication. This network plot may be interpreted as a visual representation of the combined fitness and epistasis landscapes of the top CR genotypes. **b** Average epistasis value of winning genotypes (blue histogram,  $n = 236$  genotypes, which survived four rounds of selection) is substantially more positive ( $\mu = 1.47$ ,  $p = 2.2 \times 10^{-16}$ , Welch's one-sided Student's  $t$ -test) compared with the overall population of mutants (gray histogram,  $n = 1360$  genotypes) that survived after one round of selection ( $\mu = 0.17$ ). **c** Relatedness to other winning genotypes in the population (Edges per node) is correlated with enrichment. **d** Box plots of CR mutants that are 1–5 mutations away from wild-type (Hamming distance) are depicted with the natural log transformation of final fold enrichment after selection in clindamycin or **e** epistasis values. The number of winning genotypes ( $n$ ) used to construct box plots for Hamming distance = 1, 2, 3, 4, or 5 is equal to 11, 31, 72, 70, and 5, respectively, for both sets of box plots. Both values increase with distance from wild-type, except in the case of Hamming distance = 5, which necessarily entails mutations to highly conserved nucleotide positions. For box plots, the bottom and top hinges denote the 25% and 75% quantile, whiskers extend to 1.5 times the interquartile range, the diamond represents the mean, the center line is the median, and outliers are black points.

scheme such as RISE, which can search vast genotype space for a desired activity with limited predictive intervention. Future experiments with other selections will determine whether the importance of positive epistasis in the ribosome is a general phenomenon, or if each selection experiment produces different epistatic trends.

To test the function of some of the most unusual top positive epistasis mutants, we cloned 8 of the mutants with variation at positions 2060 or 2061, the most converged positions in our library, which are generally thought to require WT nucleotides for function<sup>38</sup>. Several of these mutants were functional in iSAT, demonstrating that their measured positive epistasis was a real effect, and WT nucleotides at these positions are not strictly required for translation (Supplementary Fig. 15). These results in epistatic interactions in the ribosome suggest the importance of deep epistasis-scanning experiments to acquire a high-resolution map of the positional interactions that govern ribosome function.

## Discussion

In this study, we present the RISE method as a tool for evolving ribosomes in vitro. This was accomplished by uniting in vitro ribosome synthesis, assembly, and translation with ribosome display. By optimizing reaction conditions, we show that RISE can enrich an active 23S rRNA gene against an inactive variant >1000-fold per round of selection. We applied RISE to the in vitro evolution of ribosomes from a complex pool of variants, rapidly enriching for active ribosomes from a large PTC library ( $\sim 10^7$  members) and clindamycin-resistant mutants from a targeted library ( $\sim 4 \times 10^3$  members). The ability to recover highly

divergent genotypes, which are translationally active and yet may not support cell growth, shows that RISE can access genotypes that would not be recoverable from an in vivo system, allowing us to explore more diverse functional sequence space of the translation machinery.

Looking forward, we anticipate that RISE will serve as a valuable tool to understand fundamental ribosomal biochemistry and explore its mutability and epistatic interactions, as well as broaden efforts to expand the range of genetically encoded chemistry into proteins. This is because RISE removes the restrictions of cell viability, transformation efficiency, and growth delays encountered by in vivo studies. With this approach, ribosomal mutants can be rapidly created and screened, and in conditions that would not be possible in cells (e.g., non-natural pH or redox environment, or in the presence of additives that are toxic to cells). RISE also holds promise to map the mutability and epistatic interactions of the *E. coli* ribosome, as well as dissect in vitro ribosome assembly and function in fine detail. Based on the prevalence of positive epistatic interactions in our winning CR genotypes, we predict that such interactions will be critical to engineering new function in the ribosome. Furthermore, structural studies of ribosomes engineered for  $\beta$ -amino acid polymerization have struggled to resolve the interactions between residues that make this phenotype possible, suggesting that statistical inference of residue interactions using a method such as RISE could prove fruitful<sup>39</sup>. For these reasons, we posit that RISE will facilitate efforts to both understand the fundamental constraints of the ribosome's RNA-based active site and create biopolymers containing mirror-image (D- $\alpha$ -) and backbone-extended ( $\beta$ - and  $\gamma$ -) amino acids<sup>40–42</sup>. Such polymers could



accelerate the discovery of next-generation therapeutics and materials<sup>43,44</sup>.

## Methods

**Plasmid and library construction.** The plasmids pT7rrnB (containing rDNA operon *rrnB*) and the reporter plasmids pK7LUC and pY71sGFP were used in iSAT reactions<sup>18,19</sup>. A variant of pT7rrnB with a 660 bp deletion in the 23S gene, named pT7rrnBA660, was created by inverse PCR (Supplementary Table 1).

Plasmids for ribosome display were constructed from the pRDV plasmid<sup>13</sup>. For selective peptide or protein gene insertion, the gene was first amplified by PCR with primers encoding a 5'-GGTGGT-3' spacer and restriction sites for either NcoI for forward primers or BamHI for reverse primers (Supplementary Table 1). The amplified genes and pRDV were digested with NcoI and BamHI, and the correct fragments were isolated by gel electrophoresis and extracted. Fragments were ligated with Quick Ligase (NEB) and transformed into chemically competent DH5 $\alpha$  cells, plated, and grown overnight. Resulting isolated colonies were grown for plasmid purification and sequencing.

Libraries of rDNA operons were created from the pT7rrnB plasmid through PCR amplification of particular rDNA fragments with phosphorylated primers containing overhangs of degenerate bases (Supplementary Table 1). DNA fragments were ligated and PCR amplified for in vitro insertion into the pT7rrnB plasmid (see below).

**Total protein of 70S ribosomes (TP70) preparation.** Two volumes of glacial acetic acid were added to purified *E. coli* 70S ribosomes in Buffer C (10 mM Tris-OAc (pH = 7.5 at 4 °C), 60 mM NH<sub>4</sub>Cl, 7.5 mM Mg(OAc)<sub>2</sub>, 0.5 mM EDTA, 2 mM dithiothreitol (DTT)) with 0.2 mM spermine and 2 mM spermidine, and 100 mM Mg(OAc)<sub>2</sub> to precipitate rRNA. This sample was mixed well and centrifuged at 16,000  $\times$  g for 30 min to pellet rRNA. Supernatant containing r-proteins was collected and mixed with five volumes of chilled acetone and stored overnight at -20 °C. Precipitated protein was then collected by centrifugation at 10,000  $\times$  g for 30 min, dried, and resuspended in simplified high-salt buffer with urea (10 mM Tris-OAc (pH = 7.5 at 4 °C), 10 mM Mg(OAc)<sub>2</sub>, 200 mM KGlu, 1 mM DTT, 6 M urea). The sample was transferred to midi-size 1000 Da molecular weight cutoff (MWCO) Tube-O-Dialyzers (G-Biosciences) and dialyzed overnight against 100 volumes of simplified high-salt buffer with urea. The sample was then dialyzed against 100 volumes of simplified high-salt buffer without urea three times for 90 min each. This sample was clarified at 4000  $\times$  g for 10 min, and concentration was determined from A230 NanoDrop readings (1 A230 unit of TP70 = 240 pmol TP70). TP70 samples were then aliquoted, flash-frozen, and stored at -80 °C.

**S150 extract preparation.** S30 lysate generated from MRE600 cells collected at OD600 of 0.5 was layered in a 1:1 volumetric ratio on a high-salt sucrose cushion containing 20 mM Tris-HCl pH 7.2, 500 mM NH<sub>4</sub>Cl, 10 mM MgCl<sub>2</sub>, 0.5 mM EDTA, 6 mM  $\beta$ -mercaptoethanol, and 37.7% sucrose in a Ti70 ultracentrifuge tube. Samples were centrifuged at 4 °C and 90,000  $\times$  g overnight. Supernatants were spun at 4 °C and 150,000  $\times$  g for an additional 3 h. The top two-thirds of the supernatant was removed, taking care not to disturb the pellet. Supernatant was dialyzed in Snakeskin (Thermo) dialysis membrane tubing (3500 Da MWCO) against 50 volumes of high-salt S150 extract buffer (10 mM Tris-OAc (pH 7.5 at 4 °C), 10 mM Mg(OAc)<sub>2</sub>, 20 mM NH<sub>4</sub>OAc, 30 mM KOAc, 200 mM KGlu, 1 mM spermidine, 1 mM putrescine, 1 mM DTT). Dialysis buffer volume was exchanged after 2 h for three dialysis steps. A fourth dialysis was performed overnight. Extract was clarified at 4000  $\times$  g for 10 min and concentrated to ~4 mg/mL total protein concentration using 3000 Da MWCO Centriprep concentrators (EMD Millipore) to account for dilution during preparation. Samples were then aliquoted, flash-frozen, and stored at -80 °C.

**iSAT reactions.** iSAT reactions were performed by mixing salts, substrates, and cofactors with 1 nM reporter plasmid and a molar equivalent of pT7rrnB in S150 extract<sup>16,18–20</sup>. For ribosome display reactions, the anti-ssrA oligonucleotide (5'-TTA AGC TGC TAA AGC GTA GTT TTC GTC GTT TGC GAC TA-3') was included at 5  $\mu$ M to prevent dissociation of stalled ribosomes<sup>13</sup>. Then a mix of proteins was added to final concentrations of ~2 mg/mL S150 extract, 300 nM total protein of the 70S ribosome (TP70), and 60  $\mu$ g/mL T7 RNA polymerase. Reactions were mixed gently by pipetting and incubated at 37 °C. Replicate reactions to test translational activity of clones were assembled using the Echo 525 acoustic liquid handling robot to minimize error in pipetting. For sGFP production, the background autofluorescence of reactions without a ribosomal operon plasmid was measured and subtracted from the value of the test cases, and absolute quantification was performed using a standard curve derived from a dilution series of purified recombinant sGFP<sup>19</sup>.

**Sedimentation analysis.** Ribosome profiles were determined from 50  $\mu$ L iSAT reactions by incubating reactions for 2 h at 37 °C, loading them onto a 10–40% sucrose gradient made with Buffer C and ultra-centrifuging in SW32.1 tubes at 35,000 r.p.m. and 4 °C for 18 h. Gradients were then analyzed through spectrophotometry and fractionation (500  $\mu$ L fractions). Ribosome profiles were generated

from absorbance of the gradient at 254 nm, and peaks were determined from comparison with previous traces.

**Ribosome synthesis and evolution.** Fifteen microliter iSAT reactions were performed using pRDV 3xFLAG as the ribosome display template and with the addition of the anti-ssrA oligonucleotide. Reactions were incubated at 37 °C for 90 min for each round of selection or from 7.5 to 90 min for the time course. At completion, reactions were placed at 4 °C and diluted with four volumes (60  $\mu$ L) of binding buffer (50 mM Tris-acetate (pH 7.5 at 4 °C), 50 mM magnesium acetate, 150 mM NaCl, 1% Tween<sup>®</sup> 20, and 0.0–5.0% BSA or 0–20 mg/mL heparin). Meanwhile, for each reaction, 10  $\mu$ L packed gel volumes of magnetic beads with selective markers or antibodies were washed three times with 50  $\mu$ L bead wash buffer (50 mM Tris-acetate (pH 7.5 at 4 °C), 50 mM magnesium acetate, 150 mM NaCl). Diluted iSAT reactions were added to washed beads and incubated at 4 °C for 1 h with gentle rotation to suspend beads in solution. Reactions were then washed 5 or 10 times with wash buffer (50 mM Tris-acetate (pH 7.5 at 4 °C), 50 mM magnesium acetate, 150–1000 mM NaCl, and 0.05–5% Tween<sup>®</sup> 20), with 5 or 15 min incubations of each wash step at 4 °C. Wash buffer was removed from the beads, 50  $\mu$ L elution buffer (50 mM Tris-acetate (pH 7.5 at 4 °C), 150 mM NaCl, 50 mM EDTA (Ambion)) was added, and the beads were incubated at 4 °C for 30 min with gentle rotation. Elution buffer was recovered from the beads, and rRNA was purified using a Qiagen RNeasy MinElute Cleanup Kit for rRNA analysis and/or RT-PCR.

**Reverse-transcription PCR reaction.** For quantitative RT-PCR of 23S rRNA, RNA recovered from ribosome display was diluted 1:100 with nuclease-free water to dilute EDTA in the elution buffer. Diluted samples were used with the iTaq<sup>™</sup> Universal SYBR<sup>®</sup> Green One-Step Kit (Bio-Rad) in 10  $\mu$ L reactions following product literature. For quantification, primers were designed for amplification of 23S rRNA using Primer3 software (Supplementary Table 2)<sup>45</sup>. Reactions were monitored for fluorescence in a CFX96<sup>™</sup> Real-Time PCR Detection System (Bio-Rad). A standard curve was generated from a dilution series of 70S *E. coli* ribosomes (NEB) to ensure linearity of the assay (Supplementary Fig. 1).

For recovery of rRNA from ribosome display, rRNA from captured ribosomes was purified with the RNeasy MinElute Cleanup Kit (Qiagen) and eluted with 14  $\mu$ L nuclease-free water. Purified rRNA was used with the SuperScript<sup>®</sup> III One-Step RT-PCR System with Platinum<sup>®</sup> Taq DNA Polymerase (Invitrogen<sup>™</sup>) in 30  $\mu$ L reactions following product literature. Primers were designed for use in both 23S rRNA recovery and in vitro operon plasmid assembly (see below) (Supplementary Table 2).

**In vitro operon plasmid assembly.** For library construction or rRNA recovery, PCR was performed using Q5 polymerase with primers that amplify the bases 1962–2575 of the 23S rRNA gene. The primers also included unique cut sites for the off-site type IIIs restriction enzyme SapI, and the ~660 bp fragments and pT7rrnBA660 were digested with SapI in 1 $\times$  CutSmart<sup>™</sup> buffer (NEB) for 2 h at 37 °C. DNA was purified, and inserts and plasmid were ligated at a 1:1 molar ratio with Quick Ligase (NEB) for 20 min at room temperature. The resulting plasmids were purified with DNA Clean & Concentrator-5<sup>™</sup> (Zymo Research), eluted with nuclease-free water, and analyzed by NanoDrop to determine concentration.

**Isolation of individual sequence variants from RISE pools.** Upon completion of RISE, reassembled rDNA operon pools were transformed into electrocompetent POP cells for repression of the T7 promoter to prevent lethality of particular rDNA operon variants. Individual colonies were screened by colony PCR and sequencing. Colonies resulting in unclear sequencing reads were discarded. From the sequenced colonies, variants of interest were grown for plasmid preparation, and the purified plasmids were used in iSAT reactions as described above. Libraries, RISE pools, and isolated plasmids were submitted to the Genomics Core Facility at Northwestern University for sequencing.

**Deep sequencing of RISE selection cDNA pools.** Mixed populations or clonal samples were prepared for deep sequencing beginning with a PCR of a 205 bp region containing the six randomized bases of the CR library near the center of the amplicon. These amplicons were processed for deep sequencing using NEBNext Ultra II FS DNA Library Prep Kit for Illumina (E78055) and indexed using NEBNext Multiplex Oligos Index Primers Sets 1 and 2 (E7335S, E7500S). Samples were pooled and sequenced on an Illumina MiSeq v2 Micro flowcell with 2  $\times$  150 bp paired-end reads. Upon completion, .fastq files were trimmed using *sickle* (<https://github.com/najoshi/sickle>), filtered and aligned in Geneious, and the frequency of each genotype was counted and analyzed using a Python script. Of the possible 4096 genotypes, 3935 were observed during the selection. We attribute the <4% genotypes not observed to the limited sequencing depth provided by the MiSeq instrument and to a bias for guanine nucleotides in the oligos used to generate the library, which caused our several-fold excess of transformants to be insufficient to capture all members. Only genotypes that were present with an absolute read count of 5 or greater in all sequenced pools were included in downstream analysis, to lessen issues of increased error associated with observations at the bottom of the deep-sequencing detection limit. This reduced the error



in the dataset and restricted analysis to only those genotypes that were at 1.3% their initial frequency in the population or greater at the end of the selection, excluding most of the completely nonfunctional genotypes from correlation analysis.

**Squires strain analysis.** Isolated variants were cloned into the appropriate *in vivo* vectors. Particularly, the T7 promoter was exchanged with the  $P_L$  promoter using standard molecular biology techniques (PCR, digestion, and ligation). Between 150–200 ng of sequence-verified plasmids were transformed into 50  $\mu$ L of electrocompetent *E. coli* SQ171 cells, which lack chromosomal copies of the rDNA operon<sup>34</sup>. The transformants were recovered in 800  $\mu$ L of SOC for 2 h at 37 °C. SOC (1.2 mL) supplemented with 100  $\mu$ g/mL carbenicillin was added to the recovering culture and the culture was further recovered overnight (16 h) at 37 °C. After the overnight recovery, 1  $\mu$ L of culture was freshly inoculated into 100  $\mu$ L of LB media containing 100  $\mu$ g/mL carbenicillin in a 96-well plate with linear shaking maintained at 37 °C. Growth was monitored via absorbance at 600 nm, and at exponential phase (OD<sub>600</sub> between 0.4 and 0.6), cultures were serially diluted to 10<sup>-3</sup> to 10<sup>-6</sup> OD<sub>600</sub> and 2  $\mu$ L of the dilution were spot plated onto agar plates containing 100  $\mu$ g/mL carbenicillin, 350  $\mu$ g/mL clindamycin, and/or 5% w/v sucrose.

**Estimating epistasis in the CR selection.** The epistasis value for each mutant was calculated using fitness values after the first round of selection in the presence of 500  $\mu$ M clindamycin according to the following equation:

$$\epsilon = \ln \left[ \frac{W_{1,2,\dots,N} \cdot W_0}{W_1 \cdot W_2 \cdot \dots \cdot W_{N-1} \cdot W_N} \right]$$

where  $W_{1,2,\dots,N}$  represents the fitness of the multi-mutant (measured as enrichment after one round of selection) that is a Hamming distance of  $N$  away from WT;  $W_0$  represents the fitness of WT; and  $W_1$ ,  $W_2$ , and beyond represent the fitness of the individual point mutants that make up the multi-mutant. Thus, the numerator represents the observed fitness of the genotype, and the denominator represents its expected fitness, assuming each mutation contributes independently to the fitness of the genotype they comprise. The natural log of this fraction was then taken to obtain the epistasis value  $\epsilon$ .

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

Next-generation sequencing data supporting this document are available at <http://www.ncbi.nlm.nih.gov/bioproject/602158>. Additional data that support the findings of this study and plasmids required to replicate this work are available from the corresponding author upon reasonable request.

## Code availability

Computer code used to analyze the data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 24 June 2019; Accepted: 28 January 2020;

Published online: 28 February 2020

## References

- d'Aquino, A. E., Kim, D. S. & Jewett, M. C. Engineered ribosomes for basic science and synthetic biology. *Annu. Rev. Chem. Biomol. Eng.* **9**, 311–340 (2018).
- Liu, C. C., Jewett, M. C., Chin, J. W. & Voigt, C. A. Toward an orthogonal central dogma. *Nat. Chem. Biol.* **14**, 103–106 (2018).
- Hammerling, M. J., Kruger, A. & Jewett, M. C. Strategies for *in vitro* engineering of the translation machinery. *Nucleic Acids Res.* 1–16 (2019). <https://doi.org/10.1093/nar/gkz1011>.
- Lee, J. et al. Expanding the limits of the second genetic code with ribozymes. *Nat. Commun.* **10**, 1–12 (2019).
- Liu, Y., Kim, D. S. & Jewett, M. C. Repurposing ribosomes for synthetic biology. *Curr. Opin. Chem. Biol.* **40**, 87–94 (2017).
- Dedkova, L. M. & Hecht, S. M. Expanding the scope of protein synthesis using modified ribosomes. *J. Am. Chem. Soc.* **141**, jacs.9b02109 (2019).
- Silverman, A. D., Karim, A. S. & Jewett, M. C. Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* (2019). <https://doi.org/10.1038/s41576-019-0186-3>.
- Maini, R. et al. Protein synthesis with ribosomes selected for the incorporation of  $\beta$ -amino acids. *Biochemistry* **54**, 3694–3706 (2015).
- Melo Czekster, C., Robertson, W. E., Walker, A. S., Söll, D. & Schepartz, A. *In vivo* biosynthesis of a  $\beta$ -amino acid-containing protein. *J. Am. Chem. Soc.* **138**, 5194–5197 (2016).
- Schmied, W. H. et al. Controlling orthogonal ribosome subunit interactions enables evolution of new function. *Nature* **564**, 444–448 (2018).
- Orelle, C. et al. Protein synthesis by ribosomes with tethered subunits. *Nature* **524**, 119–124 (2015).
- Carlson, E. D. et al. Engineered ribosomes with tethered subunits for expanding biological function. *Nat. Commun.* **10**, 1–13 (2019).
- Zahnd, C., Amstutz, P. & Plückthun, A. Ribosome display: selecting and evolving proteins *in vitro* that specifically bind to a target. *Nat. Methods* **4**, 269–279 (2007).
- Binz, H. K. et al. High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat. Biotechnol.* **22**, 575–582 (2004).
- Mattheakis, L. C., Bhatt, R. R. & Dower, W. J. An *in vitro* polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl Acad. Sci. USA* **91**, 9022–9026 (1994).
- Jewett, M. C., Fritz, B. R., Timmerman, L. E. & Church, G. M. *In vitro* integration of ribosomal RNA synthesis, ribosome assembly, and translation. *Mol. Syst. Biol.* **9**, 1–8 (2013).
- Semrad, K. & Green, R. Osmolytes stimulate the reconstitution of functional 50S ribosomes from *in vitro* transcripts of *Escherichia coli* 23S rRNA. *RNA* **8**, 401–411 (2002).
- Fritz, B. R. & Jewett, M. C. The impact of transcriptional tuning on *in vitro* integrated rRNA transcription and ribosome construction. *Nucleic Acids Res.* **42**, 6774–6785 (2014).
- Fritz, B. R., Jamil, O. K. & Jewett, M. C. Implications of macromolecular crowding and reducing conditions for *in vitro* ribosome construction. *Nucleic Acids Res.* **43**, 4774–4784 (2015).
- Liu, Y., Fritz, B. R., Anderson, M. J., Schoborg, J. A. & Jewett, M. C. Characterizing and alleviating substrate limitations for improved *in vitro* ribosome construction. *ACS Synth. Biol.* **4**, 454–462 (2015).
- Shepotinovskaya, I. V. & Uhlenbeck, O. C. Catalytic diversity of extended hammerhead ribozymes. *Biochemistry* **47**, 7034–7042 (2008).
- Keiler, K. C., Waller, P. R. H. & Sauer, R. T. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger. *RNA Adv. Sci.* **271**, 990–993 (1996).
- Hanes, J. & Plückthun, A. *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl Acad. Sci. USA* **94**, 4937–4942 (1997).
- Jemiolo, D. K., Taurence, J. S. & Giese, S. Mutations in 16S rRNA in *Escherichia coli* at methyl-modified sites: G966, C967, and G1207. *Nucleic Acids Res.* **19**, 4259–4265 (1991).
- Gregory, S. T. & Dahlberg, A. E. Mutations in the conserved P loop perturb the conformation of two structural elements in the peptidyl transferase center of 23 S ribosomal RNA. *J. Mol. Biol.* **285**, 1475–1483 (1999).
- Schaffitzel, C., Hanes, J., Jermutus, L. & Plückthun, A. Ribosome display: an *in vitro* method for selection and evolution of antibodies from libraries. *J. Immunol. Methods* **231**, 119–135 (1999).
- Motorin, Y., Muller, S., Behm-Ansmant, I. & Branlant, C. Identification of modified residues in RNAs by reverse transcription-based methods. *Methods Enzymol.* **425**, 21–53 (2007).
- Cantara, W. A. et al. The RNA modification database, RNAMDB: 2011 update. *Nucleic Acids Res.* **39**, 195–201 (2011).
- Green, R. & Noller, H. F. Reconstitution of functional 50S ribosomes from *in vitro* transcripts of *Bacillus stearothermophilus* 23S rRNA. *Biochemistry* **38**, 1772–1779 (1999).
- Khaitovich, P., Tenson, T., Kloss, P. & Mankin, A. S. Reconstitution of functionally active *Thermus aquaticus* large ribosomal subunits with *in vitro*-transcribed rRNA. *Biochemistry* **38**, 1780–1788 (1999).
- Erlacher, M. D., Chirkova, A., Voegelé, P. & Polacek, N. Generation of chemically engineered ribosomes for atomic mutagenesis studies on protein biosynthesis. *Nat. Protoc.* **6**, 580–592 (2011).
- Cochella, L. & Green, R. Isolation of antibiotic resistance mutations in the rRNA by using an *in vitro* selection system. *Proc. Natl Acad. Sci. USA* **101**, 3786–3791 (2004).
- Schlünzen, F. et al. Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* **413**, 814–821 (2001).
- Asai, T., Zaporjets, D., Squires, C. & Squires, C. L. An *Escherichia coli* strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. *Proc. Natl Acad. Sci. USA* **96**, 1971–1976 (1999).
- Koch, M., Willi, J., Pradere, U., Hall, J. & Polacek, N. Critical 23S rRNA interactions for macrolide-dependent ribosome stalling on the ErmCL nascent peptide chain. *Nucleic Acids Res.* **45**, 6717–6728 (2017).
- Starr, T. N. & Thornton, J. W. Epistasis in protein evolution. *Protein Sci.* **25**, 1204–1218 (2016).
- Dutheil, J. Y., Jossinet, F. & Westhof, E. Base pairing constraints drive structural epistasis in ribosomal RNA sequences. *Mol. Biol. Evol.* **27**, 1868–1876 (2010).

38. Sato, N. S., Hirabayashi, N., Agmon, I., Yonath, A. & Suzuki, T. Comprehensive genetic selection revealed essential bases in the peptidyl-transferase center. *Proc. Natl Acad. Sci. USA* **103**, 15386–15391 (2006).
39. Ward, F. R., Watson, Z. L., Ad, O., Schepartz, A. & Cate, J. H. D. Defects in the assembly of ribosomes selected for  $\beta$ -amino acid incorporation. *Biochemistry* **58**, 4494–4504 (2019).
40. Katoh, T., Tajima, K. & Suga, H. Consecutive elongation of D-amino acids in translation. *Cell Chem. Biol.* **24**, 46–54 (2017).
41. Fujino, T., Goto, Y., Suga, H. & Murakami, H. Ribosomal synthesis of peptides with multiple  $\beta$ -amino acids. *J. Am. Chem. Soc.* **138**, 1962–1969 (2016).
42. Arranz-Gibert, P., Vanderschuren, K. & Isaacs, F. J. Next-generation genetic code expansion. *Curr. Opin. Chem. Biol.* **46**, 203–211 (2018).
43. Passioura, T. & Suga, H. A RaPID way to discover nonstandard macrocyclic peptide modulators of drug targets. *Chem. Commun.* **53**, 1931–1940 (2017).
44. d'Aquino, A. E., Azim, T., Aleksashin, N. A., Hockenberry, A. J., Krüger, A. & Jewett, M. C. Mutational characterization and mapping of the 70S ribosome active site. *Nucleic Acids Res.*, gkaa001 (2020).
45. Rozen, S. & Skaletsky, H. in *Bioinformatics Methods and Protocols* (eds Misener, S. & Krawetz, S. A.) 365–386 (Humana Press, 1999). <https://doi.org/10.1385/1-59259-192-2:365>.

### Acknowledgements

We thank Adam Hockenberry for thoughtful review of the manuscript and Jessica Stark for insightful comments on analysis. We are thankful to Andreas Plückthun for use of the pRDV vector. This work was supported by the Army Research Office (W911NF-11-1-0445 and W911NF-16-1-0372) and the National Science Foundation (NSF) grant MCB-1716766. M.C.J. also acknowledges the David and Lucile Packard Foundation and the and the Dreyfus Teacher-Scholar program.

### Author contributions

M.J.H., B.R.F., and M.C.J. conceived this study. M.J.H., B.R.F., D.J.Y., D.S.K., and E.D.C. designed and executed the experiments. M.J.H., B.R.F., and M.C.J. wrote the manuscript and D.S.K. and D.J.Y. edited the manuscript. M.C.J. provided a supervisory role.

### Competing interests

M.C.J. and B.R.F. are co-inventors on the U.S. provisional patent application 62/098,622, which incorporates discoveries described in this manuscript. All other authors declare no competing interests.

### Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41467-020-14705-2>.

Correspondence and requests for materials should be addressed to M.C.J.

Peer review information *Nature Communications* thanks Frederic Pecorari and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020