Computationally-guided design and selection of high performing ribosomal active site mutants

Camila Kofman[a,b,c], Andrew M. Watkins[d,e], Do Soon Kim[a,b,c,f], Jessica A. Willi [b,c], Alexandra C. Wooldredge[b,c], Ashty S. Karim[a,b,c], Rhiju Das[d,g], Michael C. Jewett[a,b,c,h]

Affiliations

a. Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA
b. Chemistry of Life Processes Institute, Northwestern University, Evanston, IL 60208, USA
c. Center for Synthetic Biology, Northwestern University, Evanston, IL 60208, USA
d. Department of Biochemistry, Stanford University, Stanford, CA 94305, USA
e. Prescient Design, Genentech, South San Francisco, CA 94080, USA
f. Inceptive Nucleics, Inc. Palo Alto, CA 94304, USA
g. Department of Physics, Stanford University, Stanford, CA 94305, USA
h. Robert H. Lurie Comprehensive Cancer Center and Simpson Querrey Institute, Northwestern University, Chicago, IL 60611, USA

Correspondence

Michael C. Jewett, m-jewett@northwestern.edu

Abstract

Understanding how modifications to the ribosome affect function has implications for studying ribosome biogenesis, building minimal cells, and repurposing ribosomes for synthetic biology. However, efforts to design sequence-modified ribosomes have been limited because point mutations in the ribosomal RNA (rRNA), especially in the catalytic active site (peptidyl transferase center; PTC), are often functionally detrimental. Moreover, methods for directed evolution of rRNA are constrained by practical considerations (e.g., library size). Here, to address these limitations, we developed a computational rRNA design approach for screening guided libraries of mutant ribosomes. Our method includes *in silico* library design and selection using a Rosetta stepwise Monte Carlo method (SWM), library construction and *in vitro* testing of combined ribosomal assembly and translation activity, and functional characterization *in vivo*. As a model, we apply our method to making modified ribosomes with mutant PTCs. We engineer ribosomes with as many as 30 mutations in their PTCs, highlighting previously unidentified epistatic interactions, and show that SWM helps identify sequences with beneficial phenotypes as compared to random library sequences. We further demonstrate that some variants improve cell growth *in vivo*, relative to wild type ribosomes. We anticipate that SWM design may serve as a powerful tool for rRNA design.

Introduction

The ribosome is a complex macromolecular machine that has evolved to synthesize proteins by catalyzing peptide bonds between amino acids. Essential for all life, the ribosome is considered to be a ribozyme, as its catalytic active site, the peptidyl transferase center (PTC), is primarily composed of ribosomal RNA (rRNA) (1–4). Consequently, ribosome assembly and function are tightly linked to rRNA folding and stability, and the rRNA sequence has been evolutionarily constrained to enable folding into a structure capable of rapidly and reliably catalyzing peptide bonds (5, 6).

The production of modified ribosomes with biochemical defects (7–10) or altered capabilities (e.g., β-amino acid incorporation (11, 12)) serves as an important tool to better understand molecular translation and enable synthetic biology applications (13). However, designing and engineering mutant ribosomes is far from a trivial undertaking because ribosomal mutations can disrupt translation in ways that are often lethal to cells. For example, previous efforts to create modified ribosomes have shown that making even single point mutations to the PTC sequence can nullify the ribosome's ability to properly assemble or catalyze bond formation (14–17). Not only is it difficult to identify functional small-scale mutations, but some detrimental mutations can also be rescued by synergistic mutations in adjacent or distal regions, highlighting the design challenges of rRNA engineering and importance of considering epistatic relationships between residues when designing libraries (18, 19).

While tethered ribosomes (20–23) or cell-free strategies (24–28) can be used to identify functionally detrimental ribosomal mutations, efforts to build modified ribosomes remain hampered by practical considerations in making and evaluating rRNA libraries. For example, the combinatorial space for rRNA evolution is large, such that random mutagenesis and selection approaches cannot be feasibly used to screen all possible variants. In addition, due to primer bias, randomized libraries constructed using PCR are known to have imbalanced initial populations, skewing assessments of a library's members by overemphasizing the more common ones (29). PCR-based library construction approaches are also difficult to apply to multiple regions of rRNA that are close in three-dimensional space but not primary sequence space, which is common in the structurally complex PTC (30). A further challenge is that DNA libraries are typically propagated in cells, where transformation idiosyncrasies limit library size (9, 11, 31, 32). Alternative approaches are thus needed to test unbiased rRNA libraries of larger size and complexity such that we can explore diverse energy landscapes via large-scale sequence changes and identify mutant ribosomes with significantly modified architectures.

Here, to explore rRNA design rules with high-throughput methods for identifying synergistic mutations, we develop a computationally guided approach for making modified ribosomes. First, we use a stepwise Monte Carlo method (SWM) in Rosetta to score, rank, and select rRNA library members using an all-atom energy score (33) in an unbiased way. While previous methods have used computationally expensive, low-resolution coarse graining or small perturbations to fully build conformations (34), SWM requires much less computational power to reach an equivalent level of atomic accuracy. This approach also allows us to define libraries in three-dimensional space, including any residues that are potentially interacting and could mutate to play compensatory roles. We combine this computational approach with a high throughput *in vitro* ribosome synthesis, assembly, and translation (iSAT) screening platform (13, 35–37) to test the computationally identified mutants, allowing us to rapidly assay promising candidates. The resulting ribosome mutants highlight the flexibility of the PTC to large-scale mutations and elucidate previously unknown epistatic relationships between distal regions of the PTC.

Unexpectedly, many of these highly mutated variants can support life in cells with only minor phenotypic effects. We anticipate that our high-throughput, computationally guided approach will allow for improved studies of complex rRNA libraries to ultimately enable novel ribosomal activity as well as deeper understanding of molecular translation.

Materials & Methods

SWM design simulations

To obtain an initial structure for stepwise Monte Carlo design, a crystal structure of the *Escherichia coli* ribosome (PDB code: 4YBB (38)) was obtained and loaded into PyMOL. The residues of interest, local to a particular site in the ribosome, were selected, and that selection was expanded to include a 25.0 Å sphere of neighboring residues, enough to encompass several shells of indirect interactions. The full selection, including both residues of interest and neighbors, were saved to a 'native' PDB file. This 'native' file was passed to a Python script distributed with the Rosetta application, *tools/rna_tools/pdb_util/pdb2fasta.py*, to obtain a corresponding FASTA-formatted file with appropriate numbering. Finally, the 25.0 Å sphere of neighbors, but omitting the actual residues of interest, were saved to a 'starting' PDB file, ready for stepwise Monte Carlo design.

The sequence positions within the FASTA file that corresponded to the residues of interest were edited with a text editor to ensure the design simulation would sample any nucleotide but the wild-type nucleic acid identity: *a* was changed to *b* (the IUPAC ambiguous single-letter code representing 'anything but adenosine'); *c* was changed to *d*; *g* was changed to *h*; *u* was changed to *v*. Because the region being redesigned would be free to resample its backbone conformation, some 'adaptation' between the fully flexible designed region and the totally rigid crystal context was necessary. To this end, the residues adjacent in primary sequence to any redesigned residue were indicated to the *-extra_min_res* flag: these residues, although not subject to explicit backbone sampling, were subject to quasi-Newtonian energy minimization along with the designed residues during simulation.

Simulations for rRNA helices 73, 75 and 91 were run for 1000 Monte Carlo cycles, while simulations for helix 92 were run for 2000 Monte Carlo cycles due to its structural complexity. At least 10,000 independent trajectories were run for each library. Full code examples for setting up, conducting, and analyzing design simulations are provided at https://github.com/everyday847/ptc_swm_modeling; documentation for stepwise Monte Carlo that includes details on design simulations is available at https://new.rosettacommons.org/docs/latest/application_documentation/stepwise/stepwise_monte_carlo/stepwise.

Sequence alignment and analysis

A dataset consisting of 1,614 pre-aligned and phylogenetically arranged bacterial and archaeal 23S sequences was downloaded and analyzed as previously published (15). Full code examples of the analyses are provided at https://github.com/camilakofman/PTC_SequenceAlignments.

Forward folding with SWM

In a design simulation, different sequences may be the lowest scoring frame of a trajectory at significantly variable frequencies. As a result, it can be difficult to make confident comparisons between the best energy sampled for two sequences or to set a strict threshold to select a small

number of desired variants. Instead of establishing a strict cutoff selecting only a few variants for experimental characterization based on variable quantities of data, we ran individual 'forward-folding' simulations on a larger number of fixed sequences, using the 200 top-scoring sequences from the design simulation. These 'forward-folding' simulations used only 500 cycles and generated exactly 400 models each, ensuring an 'apples to apples' comparison among sequences for the final selection that would be inaccessible to a design simulation alone. We ran these simulations specifically for helix 75, where we were interested in whether lower scores would correlate to superior performance in iSAT so accuracy and fair sampling for the single highest score was paramount. We elected not to repeat the simulations for the other helices due to computational resource constraints. Full code examples for setting up, conducting, and analyzing forward folding simulations are provided at https://github.com/everyday847/ptc_swm_modeling.

Plasmid construction & preparation

Plasmids were ordered from Twist in two backbones: one in the pT7rrnB backbone (36) and one in pAM552G (22). Plasmids used for testing of variants in iSAT were built using pT7rrnB, a 7,311-bp plasmid. This plasmid carries an *E. coli* rRNA operon, rrnB, under the control of the T7 promoter, as well as the ampicillin resistance gene. Constructs from Twist in the pT7rrnB backbone were transformed into chemically competent *E. coli* Dh10β cells and plated on LB plates supplemented with 50 µg/mL Carbenicillin (Cb_{50}). Plates were incubated at 37 ºC overnight. Single colonies were picked and grown overnight at 37 ºC in 50-mL of LB media supplemented with 100 µg/mL Carbenicillin (Cb_{100}). Plasmids were then purified using the ZymoPure II Plasmid Miniprep Kit. The resulting plasmids were further purified via ethanol precipitation using 5 M NH_4OAc for use in cell-free reactions.

Plasmids used for testing of variants in the Squires strain (39, 40) were built using the 7,451-bp pAM552G plasmid in POP2136 cells. Like the pT7rrnB plasmid, pAM552G carries a copy of the rrnB operon as well as an ampicillin resistance gene. However, in pAM552G, expression of rrnB is under control of the phage lambda pL promoter, which is in turn regulated by the temperature-sensitive bacteriophage lambda cI857 repressor (41). Plasmids from Twist were transformed into chemically competent *E. coli* POP2136 cells and plated on LB plates containing Cb_{50}. Plates were incubated at 30 ºC to prevent expression of rRNA from the plasmid. Colonies were picked and grown overnight in 5-mL Cb_{100} cultures and grown at 30 ºC to continue repression of rRNA expression. Plasmids were purified using the ZymoPURE Plasmid Miniprep Kit. Purified plasmids were then transformed into electrocompetent SQ171fg cells (22, 39). Of note, both plasmids contain an A2058G mutation, which endows the resulting ribosome with Erythromycin (Ery) resistance. Ery is used for the *in vivo* selection.

Strain culture & harvest

S150 lysates, total protein of the 70S ribosome (TP70) and T7 RNA Polymerase were prepared as previously reported (25, 36). 10 mL of an overnight culture of *E. coli* (MRE600 strain) cells were added into a liter of 2xYTPG medium (2xYTP with 18 g/L of glucose) and grown at 37 °C with shaking at 250 rpm until OD600 reached 3. Culture was spun down at 5,000xg for 10 minutes and kept on ice between all transfer steps. Supernatant was removed and pellet was resuspended in S30 buffer (10 mM TrisOAc pH=8.2, 14 mM $Mg(OAc)_2$, 60 mM KOAc). Cell suspension was spun at 10,000xg for 3 minutes twice more, removing supernatant between each spin and resuspending in 40 mL of fresh S30 buffer. After the third spin, pelleted were weighed and flash frozen with liquid nitrogen before storing at -80 °C. S30 Buffer was then added at a ratio of 5 mL per 1g of cell mass, and cells resuspended by vortexing until fully thawed. 100 µL of HALT Protease Inhibitor Cocktail was added per 10 mL cell suspension, and 75 µL of Takara Recombinant RNase Inhibitor was added per 4 grams of dry cell mass. Cells were lysed at

~25,000 psi with a C3 Avestin Homogenizer and a second aliquot of Takara Recombinant RNase Inhibitor was added at a ratio of 75 µL per 4 grams of initial pellet. Cell debris were pelleted by centrifugation at 12,000xg at 4 °C for 15 minutes. Supernatant (S12 extract) was recovered for S150 extract preparation and layered on top of an equivalent volume of sucrose cushion buffer (20 mM Tris–HCl (pH 7.2 at 4 °C), 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 37.7% sucrose) in Ti70 tubes. Samples were then ultracentrifuged at 90,000xg for 18 hours, after which the supernatant was transferred into fresh Ti70 tubes and spun at 150,000xg for 3 hours and pellets were gently washed with buffer C (10 mM Tris–OAc (pH 7.5 at 4 °C), 60 mM NH_4Cl, 7.5 mM $Mg(OAc)_2$, 0.5 mM EDTA, 2 mM DTT). Ribosome concentration in the pellets was measured using A_{260} Nanodrop measurements (1 A_{260} unit of 70S = 24 pmol 70S). After the second spin, the top 2/3 of the supernatant was collected and transferred into MWCO=3.5 K dialysis tubing (SnakeSkin) and dialyzed 2 x 1.5 hours x 3 L of S150 Extract Buffer at 4 °C. For the 3rd dialysis, 3L of fresh S150 Extract Buffer was used to dialyze overnight (12-15 hours). S150 extract was concentrated using Centripreps (3 kDa MWCO) until A_{260}=25 and A_{280}=15. Extract was aliquoted and flash frozen in liquid nitrogen. TP70 was prepared as previously described (26).

iSAT reactions

5 µL iSAT reactions were performed in 384-well nunc_267461 plates, with 4 replicates per reaction, and set up as previously described (15, 25, 36). The Echo 525 Acoustic Liquid Handler was used to aliquot reaction components into the wells. Reaction components were prepared in two separate mixtures: 1) the DNA plasmid with a small amount of premix to enable better liquid handling, and 2) the remaining reaction components. Premix was mixed with DNA at a volume ratio of 1:2.2 uL premix:DNA so to enable consistent results by increasing viscosity. Reagent mix 2, containing the S150 extract, was added into the wells initially, and then the DNA plasmid mix was aliquoted into each well. Reactions were run in a plate reader at 37 °C, measuring fluorescence (Excitation: 485 nm, Emission: 528 nm) every 15 minutes and with constant shaking for 15 hours. 40% PEG8000 (Sigma-Aldrich P1458-23ML) was added into the reaction premix for a final volume of 10%; 1M DTT was added at a final volume of 0.2%.

Plasmid replacement and selection in SQ171fg cells

Electrocompetent *E. coli* SQ171fg cells containing a pCSacB plasmid with kanamycin resistance (KanR) (22, 40) were prepared and stored in 50 µL aliquots. The Squires strain is a modified *E. Coli* strain that has all seven rRNA operons deleted from the genome. The pCSacB/KanR plasmid carries the sequence for Ribo-Tv2 (42), which serves as the sole rRNA operon in the cell. When pAM552G plasmids carrying the mutated ribosomal operon of interest as well as an ampicillin resistance gene are transformed into the cell, the original pCSacB -Ribo-Tv2 plasmid can be removed by plating on sucrose and carbenicillin (Cb). The success of the selection is then verified by confirming that the strain is no longer resistant to Kan.

50 ng of purified mutant pAM552G plasmid was transformed into 50 µL of cells. Cells were recovered in 850 µL of SOC in a 1.5-mL microcentrifuge tube at 37 °C for 1 hour, while shaking at 250 rpm. After 1 hour, 270 µL of the cell recovery was added to 2 mL of Super Optimal broth with Catabolite repression (SOC) containing 50 µg/mL Cb (Cb_{50}) and 0.25% sucrose in a 14-mL plastic culture tube. Tubes were incubated at 37 °C overnight, for 16-18 hours. The tubes were then spun down at room temperature for 5 minutes at 4,000xg. 2 mL of clear supernatant was removed, leaving the cell pellet to be concentrated into the remaining 300 µL. The concentrated cell suspension was plated on lysogeny broth (LB) agar plates containing 5% sucrose, 50 µg/mL Cb, and 20 µg/mL Ery. Plates were incubated at 37 °C until colonies appeared. 8 colonies were picked from each plate and spotted onto two LB-agar plates, one Cb_{50} plate and one Kan_{50} plate. Colonies that grew successfully on Cb_{50} but not on Kan_{50} were picked and grown overnight in LB with Cb_{50} to be midiprepped using the ZymoPURE™ II Plasmid Midiprep Kit. Midiprepped

plasmids were then submitted for Sanger sequencing to confirm the presence of the 23S sequence mutations and ensure that no additional mutations had arisen during the selection process. Constructs that did not yield colonies on the Lb-$Suc_{5\%}$-Cb_{50}-Ery_{20} plates were transformed two subsequent times to ensure that the construct did not support life. Constructs that did not yield "clean" colonies, meaning they grew on both antibiotics, were troubleshot by picking and spot plating additional colonies. If this process was again unsuccessful, transformations were attempted a total of three times before concluding that the construct was not able to support life.

As an additional check that the cells were living solely on the mutated ribosomes, we grew up 5-mL overnight cultures of the successfully transformed SQ171fg strains and purified the total RNA using the Qiagen™ RNeasy Mini kit. We then ran RT-PCRs using the Invitrogen™ SuperScript IV One-Step RT-PCR system to amplify regions of rRNA that were mutated in our variants (primers used listed in **Supplementary Table S1** as mutated fragment FP/RP**)**. The products of these RT-PCRs were then submitted for Sanger sequencing.

Spot growth experiment

SQ171fg strains containing the mutated ribosomes of interest were grown overnight in 3-mL cultures, with Cb_{50}. In the morning, the OD_{600} of each culture was measured, and normalized to an OD_{600} of 1. Four ten-fold dilutions of each construct were prepared (OD=0.1, 0.01, 0.001, 0.0001). 3-µL of each dilution was carefully pipetted onto a Cb_{50} plate. Plates were incubated at 30 °C and 37 °C and imaged as soon as a construct at the most dilute concentration showed cell growth. Spot growth experiments were completed three separate times to ensure consistent results.

Cloning and selection of randomized Helix 75 library

Primers were designed with randomized nucleotides at the helix 75 library location. Two PCRs were performed using primers with nucleotides randomized at the correct location (**Supplementary Table S1**). These two fragments were ligated using Gibson assembly and transformed into chemically competent Dh10B cells. The transformation was allowed to recover for one hour at 37 °C before being plated on Cb50 and grown overnight at 37 °C. Fourteen colonies were picked randomly, and plasmids purified as reported above to test in iSAT.

Results

Our goal was to establish a high-throughput, computationally guided approach to identify functionally active mutant ribosomes. As model regions to mutate, we focused on helices within the PTC, as the PTC plays the central role in the dynamic process of peptide bond formation. Specifically, we selected Helix 73 (H73), Helix 75 (H75), Helix 91 (H91), and Helix 92 (H92), and combinations thereof. H73 is in the aminoacyl site (A-site) of the PTC and makes contacts with the dynamic ribosomal (r-)protein L3 (43–45). The average conservation of H73 residues explored in this study is 73.42% across the domains of bacteria and archaea, and three residues (G2046, C2047 and G2621) have >91.45% conservation (**Supplementary Figure S1).** H75 is known to play a role in the assembly of the nascent polypeptide exit tunnel (5) which is essential to proper polypeptide folding; multiple bases in this helix are >90% conserved (**Supplementary Figure S1**). Along with helices 76 and 79, H75 forms the base of the L1 stalk (46), which facilitates binding, movement, and release of deacylated tRNAs (47). H91 and H92 together form one side of the highly conserved "accommodation corridor," where aminoacylated tRNAs are directed into the PTC in a specific orientation. The average conservation value for H91 residues mutated in this work is 92.16%, and the H92 region, which contacts r-protein L14 (45), contains five residues that have greater than 95.54% conservation (**Supplementary Figure S1**).

Establishing Stepwise Monte Carlo library selection on Helix 75

We first established the ability of SWM to computationally design and select mutations in H75 of the 23S rRNA, which sits near at the edge of the PTC but has ≥90% sequence conservation among all bacterial ribosomes (48). In short, SWM searches libraries via an add-and-delete move method with stochastic sampling and outputs a Rosetta all-atom energy score (33). This score is a linear combination of scaled statistical and physics-based energy terms, which serves as a metric to understand and compare sequence stabilities (49). These simulations and resulting scores account for interactions with nearby residues, whether RNA, protein, ion, or water. For H75, we created a library of rRNA variants by selecting eight nucleotides that make up the center of the helix and permitted each residue to be anything other than its identity in the wildtype (WT) ribosome (**Figure 1A**). Using this problem definition, we ran 10,000 SWM design simulations and selected 50 resulting sequences whose scores spanned the energy score range to build (**Supplementary Table S2**).

With these SWM-scored mutant rRNA sequences at hand, we next tested their performance in the high-throughput iSAT platform, a readout for both assembly and translational activity. iSAT co-activates the processes of rRNA synthesis and processing, ribosome assembly, and translation in a one-pot *in vitro* reaction (25). Performance in iSAT was quantified by monitoring superfolder green fluorescent protein (sfGFP) expression. The 50 rRNA mutant sequences were tested in iSAT to see if the SWM conformation score was correlated to iSAT performance. Maximum sfGFP synthesized after a 16-hour iSAT reaction incubated at 37 ºC was measured for each construct and normalized relative to that of the wildtype ribosome control. We observed an inverse correlation ($r=-0.62$, $p=1.3e-06$) between performance in iSAT and SWM score (**Figure 1B**). Lower scores, which indicate a more stable rRNA structural conformation, were more likely to yield functional ribosomes in iSAT, a result consistent with our recent work (50). Given this correlation, we moved forward with using SWM score as a metric for selection of successful mutants.

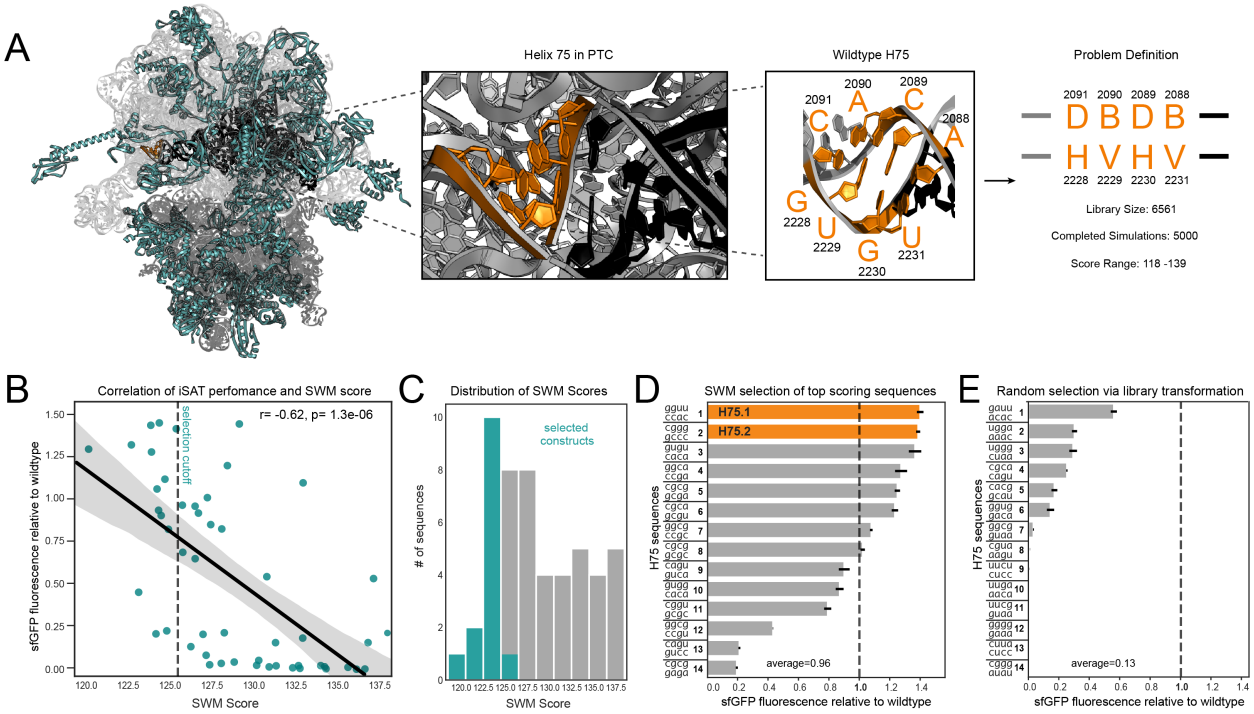

Figure 1. Application of SWM to selection of H75 variants yields high-performing mutants in iSAT. (**A**) Structure and library design for H75. H75, highlighted in orange, sits on the edge of the PTC, which is highlighted in black. Ribosomal proteins are highlighted in teal. Ribosome structure accessed from PDB ID: 4YBB. (**B**) Correlation between iSAT activity of mutants and their SWM scores. Dashed line indicates SWM score cutoff for selection. Gray highlighted region represents the 95% confidence interval. (**C**) Selection scheme for constructs based on their SWM scores. (**D**) Normalized sfGFP expression in iSAT of SWM selected constructs. (**E**) Normalized sfGFP production of randomly selected mutants. Dashed line highlights wild type sfGFP production. Data are presented of means of n = 3 experiments with standard deviation shown.

While we initially tested constructs with a broad range of SWM scores to explore the correlation between score and performance in iSAT, we wanted to confirm that using this relationship as a selection criterion would enable identification of highly successful constructs. We therefore picked sequences with scores in the top 30% (**Figure 1C, Figure 1D**) and selected an equal number of constructs from a randomized library as a negative control using randomized primers (**Figure 1E, Supplementary Table S1**) to test in iSAT. Of the constructs selected using our scoring metric, 7/14 of the selected sequences outperformed the WT sequence in iSAT. The average relative performance of SWM selected sequences was 0.96 (**Figure 1D**) compared to the average of 0.13 for sequences randomly selected from the negative control library (**Figure 1E**). This indicates that our selection method allows for the identification of variants expressing high levels of sfGFP in iSAT and highlights the flexibility of the PTC to mutations when using the SWM design strategy. Additionally, many of the selected sequences were non-trivial solutions, in that they did not maintain the base pairing pattern of the WT H75. For example, H75.1 (cggu,gcgc), which can have only two Watson-Crick (WC) base pairing interactions as opposed to the four WC base pairs in the WT helix, shows near WT sfGFP levels in iSAT. We also see that many of the selected sequences, such as H75.1, H75.4, and H75.5, do not have a WC interaction at the fourth position between residues 2228 and 2091. Although some crystal structures have found this pair to be closely interacting (38), other studies, specifically mapping secondary

structure of the 23S rRNA using base-pairing and stacking interactions (51) show that G2228 is pulled away from C2091 due to a bulging motif at the base of Helix 79. This indicates that our computational modeling approach was able to detect and account for the additional flexibility of this base pair and score sequences that left this region less rigid favorably.

Application of SWM to select for functional mutated helices in the A and P sites

We next sought to use SWM design and selection on additional motifs within the PTC to assess mutational flexibility and find novel ribosomal mutants. We chose H73, H91, and H92, three helices highlighted in **Figure 2A** (52) that are known to play roles of varying importance to the ribosome's dynamic activity. As previously noted, these three helices are highly sequence conserved (**Supplementary Figure S1).** Notably, the sequences of H91 are H92 are more than 90% conserved in the domain of bacteria, and H92 contains three bases that are universally conserved across all domains of life (48). Using a similar approach as for H75, we ran design simulations for H73, H91 and 92, and as above selected 14 sequences from each library that had energy scores in the best 30% of scores to build and test in iSAT (**Supplementary Table S3, Supplementary Table S4**). All three simulations provided us with functional 23S rRNA sequences, and multiple sequences outperformed the WT control in iSAT (**Figure 2B-D**). The observed trends highlighted that the ribosomal mutants do not have to exactly mimic WT base pairing patterns to yield high performing variants in iSAT. For example, variants H73.3, H73.4 and H73.5 (**Figure 2B**), which are all at least as high-performing as WT in iSAT, have at least one base pair that is not a canonical WC base pair. Additionally, our data supported previous findings that variant performance is highly sensitive to even small sequence changes. The identity of even one non-WC base pair can affect sfGFP production considerably; H91.3 and H91.4 are nearly identical, but H91.4 produces 20% less sfGFP due to a single nucleotide change converting a c:c pair to an a:c pair (**Figure 2C, Supplementary Table S3**). This suggests that 2539C is more favorable than 2539A, perhaps due to interactions occurring between 2539 and nearby rRNA or r-protein residues; or the result of slightly stronger hydrogen bonding of a C*C bond than an A*C bond (53). Surprisingly, between H91.3 and H91.6, we observed that changing the first base pair from a:u to c:g and the fourth from c:c to c:g actually leads to an almost 2-fold knockdown of normalized sfGFP expression. The selected sequences from H92 were the least successful in iSAT, indicating that the ribosome is more sensitive to changes in this helix (**Figure 2D**). This is likely due to the important role of the post-transcriptionally modified WT base Um2552, which is known to trigger late-stage ribosome assembly (54, 55) and whose proper modification may be impaired by mutations to bases in the H92 library. Another factor may be that H92 is specifically recognized by DbpA, an RNA helicase that is known to play an important role in ribosome assembly (56, 57), and mutations in H92 affect DpbA's ability to properly recognize its substrate (6). Despite these design challenges, some of these mutants were functional in iSAT. Of note, we find that canonical base pairing does not guarantee high performance; H92.4 has canonical base pairing but yields less than a third of the sfGFP production of H92.1. A modified library design for H92 that includes additional randomized residues may allow for better compensatory mutations to be identified in this region.

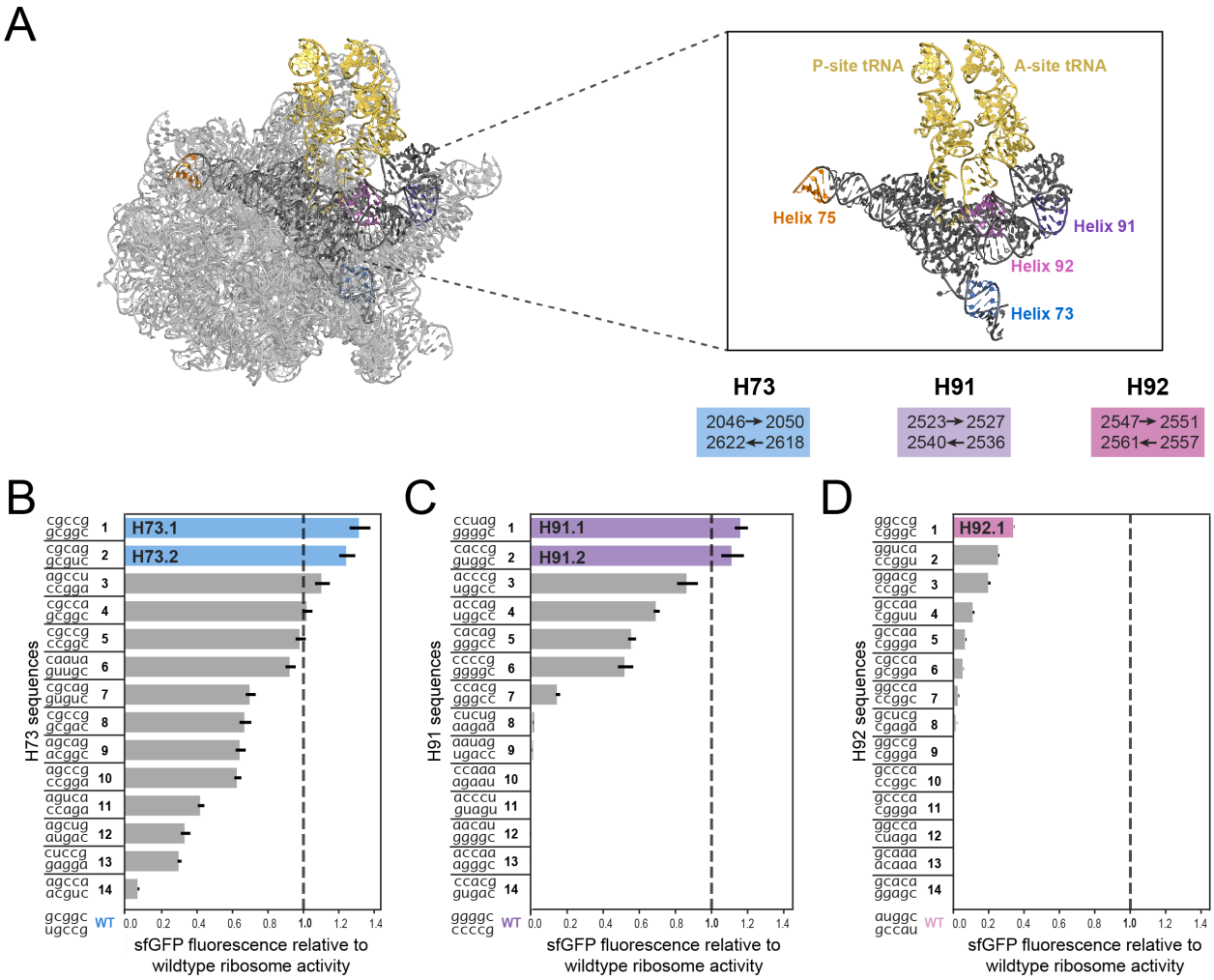

Figure 2. Design of additional helices using SWM yields multiple highly active mutants. (**A**) Structure of the 23S rRNA in gray with PTC highlighted in black. Libraries H73, H75, H91 and H92 shown in blue, orange, purple and pink, respectively. A and P-site tRNAs are highlighted in yellow. Ribosome structure accessed from PDB ID: 7K00. (**B**) iSAT results for selected H73 variants. Dashed line indicates the activity of WT, normalized to 1. Data are presented of means of n = 3 experiments with standard deviation shown. (**C**) iSAT results for selected H91 variants. (**D**) iSAT results for selected H92 variants.

Combinatorial analysis of top performing sequences highlights complex epistatic interactions in the PTC

We next wondered if combining mutations across different helices in the PTC would lead to compensatory, beneficial phenotypes. To test this, we selected top-performing sequences from each library (highlighted in **Figure 1D** and **Figure 2B-D**) and constructed all possible combinations of the four library region sequences including WT, yielding a total of 54 combinatorial rRNA constructs to test in iSAT. While none of the constructs with all four library locations mutated produced detectable levels of sfGFP in iSAT, many of the 3-way combinations were highly functional and even competitive with the WT control (**Figure 3**; **Supplementary Table S5, Supplementary Figure S2**).

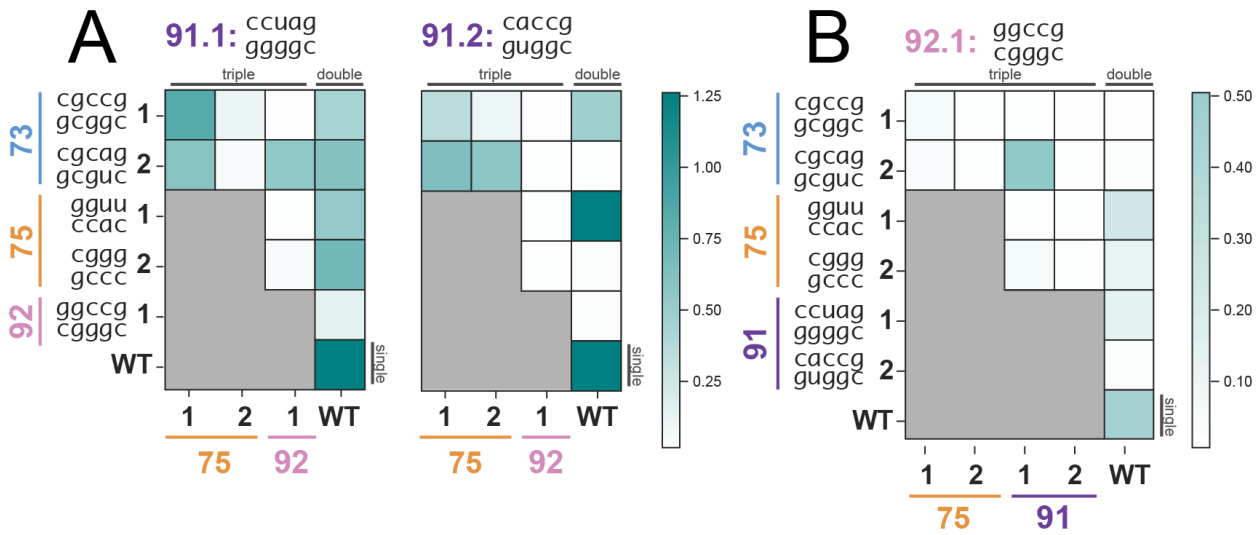

Figure 3. iSAT activities of two and three-way PTC mutant combinations yield constructs with varying sfGFP expression and highlight epistatic interactions. (**A**) H91 mutations kept constant; all constructs shown include a mutated H91, as indicated in title. (**B**) H92.1 mutant included in all combinations. Data represent the average of n = 3 independent experiments.

The results uncovered epistatic interactions between helices that highlight the highly interacting structure and complexity of the PTC. Of note, sfGFP production in iSAT could be recovered by adding additional mutated helices in some cases. For instance, we observed that H91.2 combined with H73.2 (C43) is inactive, but when further combined with either H75.1 or H75.2 (C31, C37) the sfGFP production relative to WT increases to above 43% of WT (**Figure 3A, Supplementary Table S5**). Although H73/H91 and H75 sit at opposite ends of the PTC, the addition of 8 mutated residues at H75 unexpectedly complements mutated H91.2 and H73.2, rendering them compatible. In another example, we showed that the double mutant of H91.2 combined with H73.1 (C20) is moderately functional while H91.2 with H73.2 (C43), which has only one base pair difference from H73.1, does not produce detectable levels of sfGFP. H73's effects on H91 are different when looking at the H91.1 mutant; H91.1 has 0.50 relative protein production with both H73.1 (C22) and H73.2 (C45). Similarly, while H91.1 is functional as a double mutant with H75.1 (C55) or H75.2 (C60), H91.2 yields very high sfGFP expression when paired with H75.1 (C53) but none with H75.2 (C58) -- even though H75.1 and H75.2 perform almost identically in iSAT individually (**Figure 1D**, **Supplementary Table S3**). Of note, most triple mutants including H92.1 are inactive except for in C44, its combination with H73.2 and H91.1 (**Figure 3B, Supplementary Table S5**). In fact, the addition of these 20 mutations from H73.2 and H91.1 endows the ribosome with a greater than 20% increase in normalized sfGFP production compared to any double or single mutant containing H92.1 (**Supplementary Table S5**), indicating a sensitive relationship between the three helices. This could be explained by role of the L3 protein, which acts as a dynamic switch to coordinate binding of elongation factors and has been reported to interact with helices 73, 91 and 92 as amino-acid charged tRNAs are introduced into the A-site and shuttled to the P-site (43). These mutant combinations highlight four key findings: (i) there exist previously unexplored relationships between helices in the PTC, (ii) considering dynamic and distal interactions in rRNA is essential for successful rRNA library design, (iii) ribosome performance in iSAT is sensitive to even single base pair differences in helices, and (iv) the SWM approach enables large-scale mutations in the PTC despite high sequence conservation.

Highly mutated, computationally designed ribosomes support life

We then transformed all combinatorial constructs and single-helix mutants into *E. coli* to test whether these mutant ribosomes could support translation of the *E. coli* proteome. We used a previously described selection scheme (15, 40). In short, pAM552G plasmids conferring carbenicillin resistance and encoding the mutant ribosomes were individually transformed into the *E. coli* SQ171fg strain (58), which lacks chromosomal rRNA alleles and lives on the pCSacB plasmid carrying the RiboT-v2 sequence (42). The pCSacB plasmid also contains a counter-selectable marker (sacB), which confers sucrose sensitivity, and a kanamycin resistance cassette. Thus, by growing the transformed strains in the presence of carbenicillin and sucrose, the pCSacB plasmid can be eliminated, leaving only the pAM552G ribosome mutant plasmid.

While most combinatorial constructs were not able to support life, many of them were successful and enabled cell growth closely resembling that of WT (**Figure 4A**). Notably, strains C16, C33 and C39, which all harbor ribosomes that have >10% of their PTCs mutated from WT, are still able to support cell survival and growth. Additionally, the *in vivo* analysis highlighted epistatic interactions that differed from those identified in iSAT. For example, we found that strain C45, containing H73.2 and H91.1, grew slowly, but when combined with H75.1 to produce strain C33, grew more robustly. Likewise, the combination of H91.1 with either H75.1 (C55) or H75.2 (C60) exhibited faster growth than of H91.1 alone. Of the constructs that had greater than 1/3 relative sfGFP expression to WT in iSAT (**Supplementary Tables S5, S6**), more than 73% were able to support life; thus, our data suggest performance in iSAT above a certain threshold may serve as a predictor of whether a mutant can support life. However, there were some exceptions. For example, C10 (H73.1-H75.1-H91.1) was high-performing in iSAT but did not support life (**Supplementary Table S6)**. Conversely, strain C39 was able to support cell growth despite its low performance in iSAT. These disparities are likely due to differences in the concentrations of the many dozens of assembly factors involved in ribosome biogenesis in either environment (59). We also measured the growth profiles of these strains at a lower temperature (30 °C) to observe any phenotypic changes that may be more pronounced in suboptimal growth conditions (**Figure 4B**). Surprisingly, strains C33 and C39, which showed slightly impaired growth at 37 ºC, grew significantly more robustly than WT at the lower temperature. This suggests that the mutations in strains C33 and C39 may lead to improved folding and assembly at lower temperature in cells.

Of note, plasmid exchange in the Squires strains can lead to the appearance of clones that are sucrose resistant and kanamycin sensitive, but still carry the wild type rRNA operon, possibly due to its integration into the genome or plasmid. These clones may consequently express wild type rRNA, convoluting the assessment of whether the mutant ribosomes are supporting life. To ensure that the ribosomes being expressed in our study were comprised of mutant and not WT rRNA, the entire plasmid-borne rRNA operon sequence was confirmed to contain the desired mutations by Sanger sequencing of miniprepped plasmids. The total RNA was then extracted from eight of the combination strains and RT-PCRs of the mutated regions were carried out, confirming that the rRNA being expressed in the cells matched the mutant sequence as well. No wildtype sequences were detected in any of the strains tested, indicating that the cells were harboring and expressing only the mutated ribosomal operon sequence (**Supplementary Figure S3**).

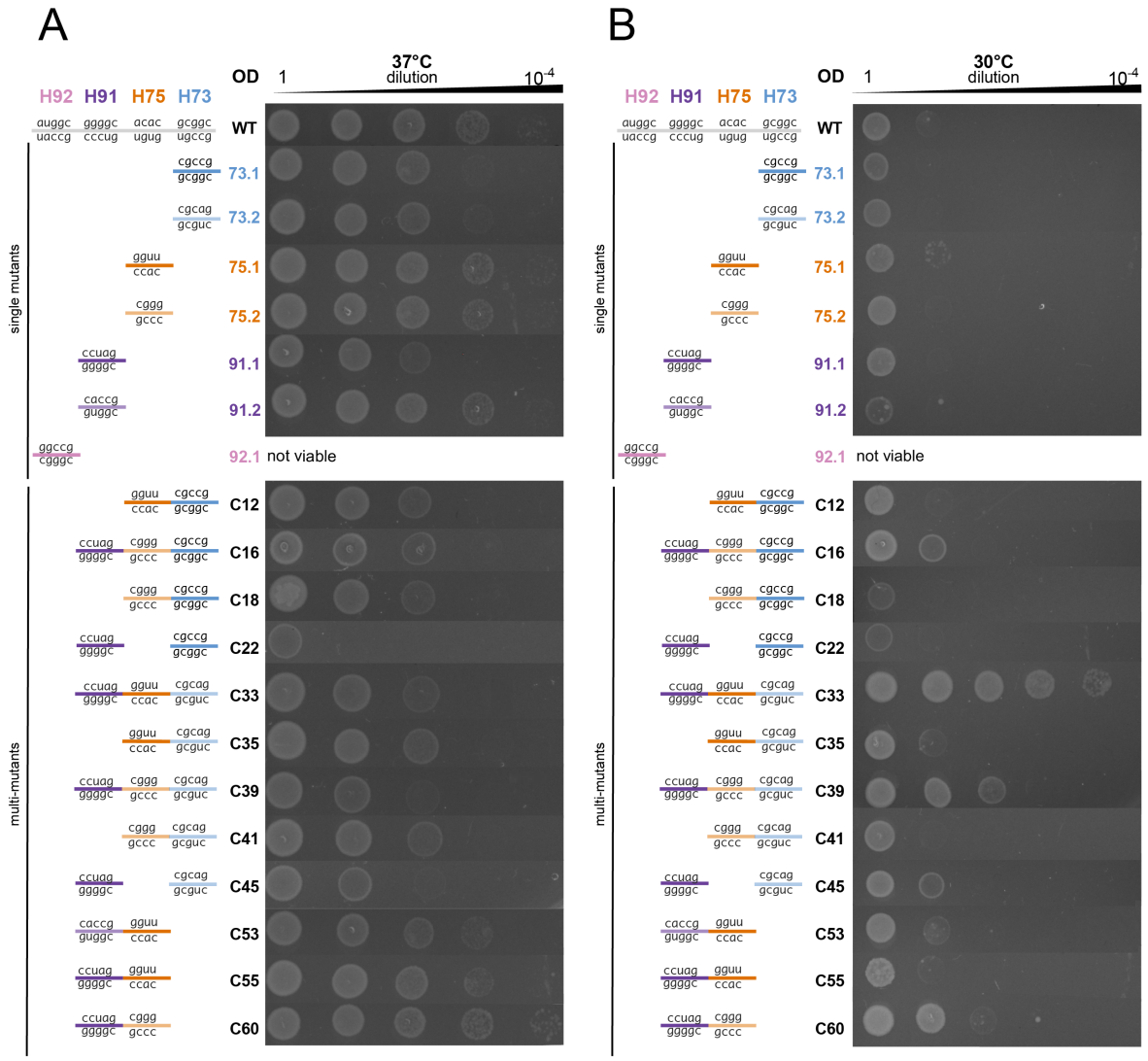

Figure 4. Spot growth assays of controls and combination constructs show that many mutants can support cell growth competitive with wildtype. (**A**) Spot growth assay at 37 °C (**B**) Spot growth assay at 30 °C. Data representative of n=3 independent experiments.

Discussion

Here, we developed a computational rRNA structure prediction method to select for highly active ribosomal PTC mutants from complex libraries and explore previously unidentified epistatic interactions between distal helices in the PTC. Our work has several key features.

First, we showed that we can use SWM to successfully select for high performing mutants using an all-atom energy score, and that this approach can serve as a tool to design significantly mutated variants that are not only functional, but even enable higher protein production yields in the cell-free environment than the wildtype ribosome sequence. This finding is consistent with current understanding of the importance of rapid rRNA folding for ribosome assembly and function (60), which has been understood to be a function of the molecule's minimum free energy (61, 62). By combining helix mutants, we built functional ribosomes with up to 30 mutations, the most highly mutated designed PTCs to our knowledge, showing that the PTC is amenable to this method of computationally vetted mutation. Many of these multi-mutants were able to support life; some

strains showed improved growth phenotypes compared to that of a strain carrying a WT ribosome. Additionally, we observed that constructs that were more successful *in vitro* had a higher probability of being able to support life in cells and identified a general rule for predicting *in vivo* success as a function of iSAT performance, which appears to be agnostic to helix location. By applying this heuristic, future design efforts may be able to test fewer candidates to arrive at successful rRNA sequences.

Second, our approach enabled library design through the lens of three-dimensional structure. We believe this feature is important for rRNA design, as nucleotides that are distant in sequence space are often highly interacting in three-dimensional space such that mutating a single residue can have off-target effects on other rRNA motifs. SWM also allows for unbiased library assessment, which is experimentally challenging due to inherent biases of primer synthesis and template:primer interactions. By computationally investigating large libraries of mutations in the PTC, we were able to explore the folding energy landscape and find alternative minima that retained—and sometimes improved—ribosome function.

Third, select combinatorial mutants in this study highlighted previously unidentified epistatic interactions between helices in the PTC. For example, performance of A-site helix mutants (H91) was strongly affected by mutations in the E-site (H75). This may be attributed to the roles that these helices play in tRNA translocation. H75 forms a three-way helical junction at the base of the L1 stalk, a dynamic feature that travels a path of ~60 Å to aid in releasing tRNAs from the ribosome (46, 63); thus, the sequence of H75 likely affects the bending movement of the stalk. H91, as noted earlier, forms part of the accommodation corridor, which undergoes key conformational changes as the tRNA is moved into the ribosome (64). Mutations in H75 and H91 may therefore be co-dependent via the central roles they play in translocating tRNAs. Additionally, H92, which interacts closely with the A-site tRNA and also makes up part of the accommodation corridor, is thought to act as a dynamic gate that slows the tRNA acceptor stem before permitting its passage into the P-site (63, 65). This activity may be affected by the residue identities of H91, potentially explaining the observed sensitivity of our H92 variants to small changes in H91 sequence. We also found that multi-helix variant performance was highly sensitive to small sequence differences; a single base pair change in H73, for example, rendered the multi-mutant with H91.1/H92.1 incapable of producing sfGFP. This may be attributed to the role of H73 in downstream rRNA folding pathways and assembly of the folding tunnel due to its position in the "central core" of the 23S, from which all other domains extend (51). This 23S rRNA core is speculated to fold independently into its active form and create a base for the A and P-sites to form via interactions with r-protein L3 (51). This interaction may thus serve as the basis of epistatic interactions between H73 and H92, as H92 interacts closely with r-protein L14, which forms a tight cluster with L3 and L19 (45, 66, 67). In addition to the previously discussed Um2552 mutation challenges, changes to H92 also may affect PTC mediated folding of the nascent protein chain as bases A2560 and U2561 have been reported to interact closely with unfolded proteins and play a role in the nucleation of protein folding (68, 69). While H73 and H91 have been reported to be related via their proximity and interactions with the L3/L14/L19 r-protein cluster (43, 66), their relationship with H75, as shown by combinations of H73.1/H91.1 with H75.1/H75.2, has not been previously documented to our knowledge.

Our results suggest that these rRNA motifs are functionally co-dependent, perhaps due to altered mobility of the L1 stalk and tRNA shuttling, PTC-mediated peptide folding, and interactions with r-proteins; and that multi-mutants in these helices, while functional in certain combinations, can render the ribosome inactive if incompatible. Notably, merging highly functional mutants does not guarantee that the resulting variant will be successful; as seen with H75.2 and H91.2 (C58), which individually both outperformed the WT ribosome in iSAT, but when combined abolished

iSAT activity (**Figures 1D, 2C and 3A**). This finding emphasizes the high interconnectivity of the PTC and the need to approach engineering rRNA through a wide lens. It is an oversimplification of the design challenge to identify the most active small-scale mutants to later combine them. Thus, high-throughput screens will likely be required to test diverse mutations in combination to enable discovery of the most promising multi-mutants. The unexpected relationships between distant helices of the PTC underscore the dynamic activity of the ribosomal active site and help improve our understanding of how to account for these kinds of interactions in future ribosome engineering efforts.

Looking forward, we anticipate that energy-based structure predictions such as SWM will be important to facilitate ribosome design. This promises to advance our understanding of rRNA function and molecular translation, as well as accelerate efforts in making modified ribosomes with expanded functions for chemical and synthetic biology.

Data Availability

Methods and input files used to run SWM simulations are available at https://github.com/everyday847/ptc_swm_modeling. Methods and input files used to run sequence conservation analysis are available at https://github.com/camilakofman/PTC_SequenceAlignments. Other data is available in the Supplementary Information.

Acknowledgements

The authors would like to thank Dr. Gyorgy Bagnigg and Argonne National Labs for access to the advanced computing center. The authors are also thankful to Ramya Rangan and Phillip Pham for help beta testing the computational pipeline, as well as to Dr. Joshua Levy for insightful conversations regarding the computational methods used in this work. C.K. was supported by the National Science Foundation Graduate Research Fellowship Program under grant no. DGE-1842165. This work was also supported by the Army Research Office (W911NF-16-1-0372) and Army Contracting Command (W52P1J-21-9-3023). The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the U.S. Government.

Conflicts of Interest

M.C.J. is a cofounder of SwiftScale Biologics, Stemloop, Inc., Design Pharmaceuticals, and Pearl Bio. The interests of M.C.J. are reviewed and managed by Northwestern University in accordance with their conflict-of-interest policies. All other authors declare no competing interests.

Author Contributions

C.K., A.M.W., D.S.K., R.D., and M.C.J. conceived the project idea. A.M.W. and R.D. wrote the stepwise Monte Carlo design code in Rosetta. C.K. and A.M.W. ran computational SWM simulations. C.K. performed sequence conservation analysis. J.A.W. optimized conditions to run iSAT reactions in high-throughput and prepared S150 extract with C.K. C.K. and A.C.W. prepared initial DNA constructs. A.K. provided input in preparing the figures. A.K., M.C.J., A.M.W., J.A.W., D.S.K., and R.D. provided input in experiment planning and edited the paper.

References

1. Isenbarger,T.A., Carr,C.E., Stewart Johnson,S., Finney,M., Church,G.M., Gilbert,W., Zuber,M.T., Ruvkun,G., Carr,C.E., Johnson,S.S., *et al.* (2008) The Most Conserved Genome Segments for Life Detection on Earth and Other Planets. 10.1007/s11084-008-

9148-z.
2. Nissen,P., Hansen,J., Ban,N., Moore,P.B. and Steitz,T.A. (2000) The structural basis of ribosome activity in peptide bond synthesis. *Science*, **289**, 920–930.
3. Steitz,T.A. and Moore,P.B. (2003) RNA, the first macromolecular catalyst: the ribosome is a ribozyme. *Trends Biochem. Sci.*, **28**, 411–418.
4. Cech,T.R. (2000) The Ribosome Is a Ribozyme. *Science (80-.).*, **289**, 878–879.
5. Wilson,D.M., Li,Y., LaPeruta,A., Gamalinda,M., Gao,N. and Woolford,J.L. (2020) Structural insights into assembly of the ribosomal nascent polypeptide exit tunnel. *Nat. Commun. 2020 111*, **11**, 1–15.
6. Kaczanowska,M. and Rydén-Aulin,M. (2007) Ribosome Biogenesis and the Translation Process in Escherichia coli. *Microbiol. Mol. Biol. Rev.*, **71**, 477–494.
7. Cochella,L. and Green,R. (2004) Isolation of antibiotic resistance mutations in the rRNA by using an in vitro selection system. *Proc. Natl. Acad. Sci. U. S. A.*, **101**, 3786–3791.
8. Yassin,A., Fredrick,K. and Mankin,A.S. (2005) Deleterious mutations in small subunit ribosomal RNA identify functional sites and potential targets for antibiotics. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 16620–16625.
9. Sato,N.S., Hirabayashi,N., Agmon,I., Yonath,A. and Suzuki,T. (2006) Comprehensive genetic selection revealed essential bases in the peptidyl-transferase center. *Proc. Natl. Acad. Sci. U. S. A.*, **103**, 15386–15391.
10. Yassin,A. and Mankin,A.S. (2007) Potential new antibiotic sites in the ribosome revealed by deleterious mutations in RNA of the large ribosomal subunit. *J. Biol. Chem.*, **282**, 24329–24342.
11. Dedkova,L.M., Fahmi,N.E., Golovine,S.Y. and Hecht,S.M. (2006) Construction of modified ribosomes for incorporation of D-amino acids into proteins. *Biochemistry*, **45**, 15541–15551.
12. Dedkova,L.M., Fahmi,N.E., Paul,R., Del Rosario,M., Zhang,L., Chen,S., Feder,G. and Hecht,S.M. (2012) β-puromycin selection of modified ribosomes for in vitro incorporation of β-amino acids. *Biochemistry*, **51**, 401–415.
13. D'Aquino,A.E., Kim,D.S. and Jewett,M.C. (2018) Engineered ribosomes for basic science and synthetic biology. *Annu. Rev. Chem. Biomol. Eng.*, **9**, 311–340.
14. Thompson,J., Kim,D.F., O'Connor,M., Lieberman,K.R., Bayfield,M.A., Gregory,S.T., Green,R., Noller,H.F. and Dahlberg,A.E. (2001) Analysis of mutations at residues A2451 and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit. *Proc. Natl. Acad. Sci.*, **98**, 9002–9007.
15. D'Aquino,A.E., Azim,T., Aleksashin,N.A., Hockenberry,A.J., Krüger,A. and Jewett,M.C. (2020) Mutational characterization and mapping of the 70S ribosome active site. *Nucleic Acids Res.*, **48**, 2777–2789.
16. Rakauskaite,R. and Dinman,J.D. (2011) Mutations of highly conserved bases in the peptidyltransferase center induce compensatory rearrangements in yeast ribosomes. *Rna*, **17**, 855–864.
17. Green,R., Samaha,R.R. and Noller,H.F. (1997) Mutations at nucleotides G2251 and U2585 of 23 S rRNA perturb the peptidyl transferase center of the ribosome. *J. Mol. Biol.*, **266**, 40–50.
18. Long,K.S., Munck,C., Andersen,T.M.B., Schaub,M.A., Hobbie,S.N., Böttger,E.C. and Vester,B. (2010) Mutations in 23S rRNA at the peptidyl transferase center and their relationship to linezolid binding and cross-resistance. *Antimicrob. Agents Chemother.*, **54**, 4705–4713.
19. Dutheil,J.Y., Jossinet,F. and Westhof,E. (2010) Base pairing constraints drive structural epistasis in ribosomal RNA sequences. *Mol. Biol. Evol.*, **27**, 1868–1876.
20. Fried,S.D., Schmied,W.H., Uttamapinant,C. and Chin,J.W. (2015) Ribosome Subunit Stapling for Orthogonal Translation in E.coli. *Angew. Chem. Int. Ed. Engl.*, **54**, 12791.

21. Aleksashin,N.A., Szal,T., d'Aquino,A.E., Jewett,M.C., Vázquez-Laslop,N. and Mankin,A.S. (2020) A fully orthogonal system for protein synthesis in bacterial cells. *Nat. Commun.*, 10.1038/s41467-020-15756-1.
22. Orelle,C., Carlson,E.D., Szal,T., Florin,T., Jewett,M.C. and Mankin,A.S. Protein synthesis by ribosomes with tethered subunits. 10.1038/nature14862.
23. Kolber,N.S., Fattal,R., Bratulic,S., Carver,G.D. and Badran,A.H. Orthogonal translation enables heterologous ribosome engineering in E. coli. 10.1038/s41467-020-20759-z.
24. Youngman,E.M. and Green,R. (2005) Affinity purification of in vivo-assembled ribosomes for in vitro biochemical analysis. *Methods*, **36**, 305–312.
25. Jewett,M.C., Fritz,B.R., Timmerman,L.E. and Church,G.M. (2013) In vitro integration of ribosomal RNA synthesis, ribosome assembly, and translation. *Mol. Syst. Biol.*, **9**, 678.
26. Hammerling,M.J., Fritz,B.R., Yoesep,D.J., Kim,D.S., Carlson,E.D. and Jewett,M.C. (2020) In vitro ribosome synthesis and evolution through ribosome display. *Nat. Commun.*, **11**.
27. Huang,S., Aleksashin,N.A., Loveland,A.B., Klepacki,D., Reier,K., Kefi,A., Szal,T., Remme,J., Jaeger,L., Vázquez-Laslop,N., *et al.* (2020) Ribosome engineering reveals the importance of 5S rRNA autonomy for ribosome assembly. *Nat. Commun.*, **11**, 1–13.
28. Yassin,A. and Mankin,A.S. (2007) Potential New Antibiotic Sites in the Ribosome Revealed by Deleterious Mutations in RNA of the Large Ribosomal Subunit *. 10.1074/jbc.M703106200.
29. Neylon,C. (2004) Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res.*, **32**, 1448.
30. Kim,D.S., Watkins,A., Bidstrup,E., Lee,J., Topkar,V., Kofman,C., Schwarz,K.J., Liu,Y., Pintilie,G., Roney,E., *et al.* (2022) 3D-structure-guided evolution of a ribosome with tethered. *Nat. Chem. Biol.*
31. Dedkova,L.M. and Hecht,S.M. (2019) Expanding the Scope of Protein Synthesis Using Modified Ribosomes. *J. Am. Chem. Soc.*, **141**, 6430–6447.
32. Dedkova,L.M., Fahmi,N.E., Paul,R., Del Rosario,M., Zhang,L., Chen,S., Feder,G. and Hecht,S.M. (2012) B-Puromycin Selection of Modified Ribosomes for in Vitro Incorporation of B-Amino Acids. *Biochemistry*, **51**, 401–415.
33. Watkins,A.M., Geniesse,C., Kladwang,W., Zakrevsky,P., Jaeger,L. and Das,R. (2018) Blind prediction of noncanonical RNA structure at atomic accuracy. *Sci. Adv.*, **4**, 1–13.
34. Boniecki,M.J., Lach,G., Dawson,W.K., Tomala,K., Lukasz,P., Soltysinski,T., Rother,K.M. and Bujnicki,J.M. (2016) SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. *Nucleic Acids Res.*, **44**, e63.
35. Liu,Y., Fritz,B.R., Anderson,M.J., Schoborg,J.A. and Jewett,M.C. (2015) Characterizing and alleviating substrate limitations for improved in vitro ribosome construction. *ACS Synth. Biol.*, **4**, 454–462.
36. Fritz,B.R. and Jewett,M.C. (2014) The impact of transcriptional tuning on in vitro integrated rRNA transcription and ribosome construction. *Nucleic Acids Res.*, **42**, 6774–6785.
37. Fritz,B.R., Jamil,O.K. and Jewett,M.C. (2015) Implications of macromolecular crowding and reducing conditions for in vitro ribosome construction. *Nucleic Acids Res.*, **43**, 4774–4784.
38. Noeske,J., Wasserman,M.R., Terry,D.S., Altman,R.B., Blanchard,S.C. and Cate,J.H.D. (2015) High-resolution structure of the Escherichia coli ribosome. *Nat. Struct. Mol. Biol.*, **22**, 336–341.
39. Quan,S., Skovgaard,O., McLaughlin,R.E., Buurman,E.T. and Squires,C.L. (2015) Markerless Escherichia coli rrn deletion strains for genetic determination of ribosomal binding sites. *G3 Genes, Genomes, Genet.*, **5**, 2555–2557.
40. Asai,T., Condon,C., Voulgaris,J., Zaporojets,D., Shen,B., Al-Omar,M., Squires,C. and Squires,C.L. (1999) Construction and initial characterization of Escherichia coli strains with few or no intact chromosomal rRNA operons. *J. Bacteriol.*, **181**, 3803–3809.

41. George,H.J., Watson,R.J., Harbrecht,D.F. and Delorbe,W.J. (1987) A Bacteriophage λ cI857 Cassette Controls λ PL Expression Vectors at Physiologic Temperatures. *Bio/Technology 1987 56*, **5**, 600–603.
42. Carlson,E.D., d'Aquino,A.E., Kim,D.S., Fulk,E.M., Hoang,K., Szal,T., Mankin,A.S. and Jewett,M.C. (2019) Engineered ribosomes with tethered subunits for expanding biological function. *Nat. Commun.*, 10.1038/s41467-019-11427-y.
43. Meskauskas,A. and Dinman,J.D. (2015) Ribosomal Protein L3: Gatekeeper to the A Site. *Mol. Cell*, 10.1016/j.molcel.2007.02.015.
44. Meskauskas,A. and Dinman,J.D. (2008) Ribosomal protein L3 functions as a 'rocker switch' to aid in coordinating of large subunit-associated functions in eukaryotes and Archaea. *Nucleic Acids Res.*, **36**, 6175–6186.
45. Watson,Z.L., Ward,F.R., Méheust,R., Ad,O., Schepartz,A., Banfield,J.F. and Cate,J.H.D. (2020) Structure of the bacterial ribosome at 2 Å resolution. *Elife*, **9**, 1–62.
46. Réblová,K., Šponer,J. and Lankaš,F. (2012) Structure and mechanical properties of the ribosomal L1 stalk three-way junction. *Nucleic Acids Res.*, **40**, 6290–6303.
47. Cornish,P. V., Ermolenko,D.N., Staple,D.W., Hoang,L., Hickerson,R.P., Noller,H.F. and Ha,T. (2009) Following movement of the L1 stalk between three functional states in single ribosomes. *Proc. Natl. Acad. Sci.*, **106**, 2571–2576.
48. Doris,S.M., Smith,D.R., Beamesderfer,J.N., Raphael,B.J., Nathanson,J.A. and Gerbi,S.A. (2015) Universal and domain-specific sequences in 23S-28S ribosomal RNA identified by computational phylogenetics. *Rna*, **21**, 1719–1730.
49. Alford,R.F., Leaver-Fay,A., Jeliazkov,J.R., O'Meara,M.J., DiMaio,F.P., Park,H., Shapovalov,M. V., Renfrew,P.D., Mulligan,V.K., Kappel,K., *et al.* (2017) The Rosetta All-Atom Energy Function for Macromolecular Modeling and Design. *J. Chem. Theory Comput.*, **13**, 3031–3048.
50. Krüger,A., Watkins,A.M., Wellington-Oguri,R., Romano,J., Kofman,C., DeFoe,A., Kim,Y., Anderson-Lee,J., Fisker,E., Townley,J., *et al.* (2021) Community science designed ribosomes with beneficial phenotypes. *bioRxiv*, 10.1101/2021.09.05.458952.
51. Petrov,A.S., Bernier,C.R., Hershkovits,E., Xue,Y., Waterbury,C.C., Hsiao,C., Stepanov,V.G., Gaucher,E.A., Grover,M.A., Harvey,S.C., *et al.* (2013) Secondary structure and domain architecture of the 23S and 5S rRNAs. *Nucleic Acids Res.*, **41**, 7522.
52. Voorhees,R.M., Weixlbaumer,A., Loakes,D., Kelley,A.C. and Ramakrishnan,V. (2009) Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nat. Struct. Mol. Biol.*, **16**, 528–533.
53. Lemieux,S. and Major,F. (2002) RNA canonical and non-canonical base pairing types: a recognition method and complete repertoire. *Nucleic Acids Res.*, **30**, 4250.
54. Arai,T., Ishiguro,K., Kimura,S., Sakaguchi,Y., Suzuki,T. and Suzuki,T. (2015) Single methylation of 23S rRNA triggers late steps of 50S ribosomal subunit assembly. *Proc. Natl. Acad. Sci. U. S. A.*, **112**, E4707–E4716.
55. Wang,W., Li,W., Ge,X., Yan,K., Sekhar Mandava,C., Sanyal,S. and Gao,N. Loss of a single methylation in 23S rRNA delays 50S assembly at multiple late stages and impairs translation initiation and elongation. 10.1073/pnas.1914323117/-/DCSupplemental.
56. Karginov,F. V. and Uhlenbeck,O.C. (2004) Interaction of Escherichia coli DbpA with 23S rRNA in different functional states of the enzyme. *Nucleic Acids Res.*, **32**, 3028.
57. Sharpe Elles,L.M., Sykes,M.T., Williamson,J.R. and Uhlenbeck,O.C. (2009) A dominant negative mutant of the E. coli RNA helicase DbpA blocks assembly of the 50S ribosomal subunit. *Nucleic Acids Res.*, **37**, 6503.
58. Carlson,E.D., d'Aquino,A.E., Kim,D.S., Fulk,E.M., Hoang,K., Szal,T., Mankin,A.S. and Jewett,M.C. (2019) Engineered ribosomes with tethered subunits for expanding biological function. *Nat. Commun.*, **10**.
59. Williamson,J.R. and Davis,J.H. (2017) Structure and dynamics of bacterial ribosome

biogenesis. *Philos. Trans. R. Soc. B Biol. Sci.*, **372**.
60. Woodson,S.A. (2008) RNA folding and ribosome assembly. *Curr. Opin. Chem. Biol.*, **12**, 667.
61. Chen,S.J. and Dill,K.A. (2000) RNA folding energy landscapes. *Proc. Natl. Acad. Sci.*, **97**, 646–651.
62. Doshi,K.J., Cannone,J.J., Cobaugh,C.W. and Gutell,R.R. (2004) Evaluation of the suitability of free-energy minimization using nearest-neighbor energy parameters for RNA secondary structure prediction. *BMC Bioinformatics*, **5**, 1–22.
63. Noller,H.F., Lancaster,L., Zhou,J. and Mohan,S. (2017) The ribosome moves: RNA mechanics and translocation. *Nat. Struct. Mol. Biol.*, **24**.
64. Whitford,P.C., Geggier,P., Altman,R.B., Blanchard,S.C., Onuchic,J.N. and Sanbonmatsu,K.Y. (2010) Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. *RNA*, **16**, 1196–1204.
65. Sanbonmatsu,K.Y., Joseph,S. and Tung,C.S. (2005) Simulating movement of tRNA into the ribosome during decoding. *Proc. Natl. Acad. Sci. U. S. A.*, **102**, 15854–15859.
66. Davies,C., White,S.W. and Ramakrishnan,V. (1996) The crystal structure of ribosomal protein L14 reveals an important organizational component of the translational apparatus. *Structure*, **4**, 55–66.
67. Spahn,C.M.T., Penczek,P.A., Leith,A. and Frank,J. (2000) A method for differentiating proteins from nucleic acids in intermediate-resolution density maps: cryo-electron microscopy defines the quaternary structure of the Escherichia coli 70S ribosome. *Structure*, **8**, 937–948.
68. Das,D., Samanta,D., Hasan,S., Das,A., Bhattacharya,A., Dasgupta,S., Chakrabarti,A., Ghorai,P. and Das Gupta,C. (2012) Identical RNA-Protein Interactions in Vivo and in Vitro and a Scheme of Folding the Newly Synthesized Proteins by Ribosomes. *J. Biol. Chem.*, **287**, 37508.
69. Samanta,D., Mukhopadhyay,D., Chowdhury,S., Ghosh,J., Pal,S., Basu,A., Bhattacharya,A., Das,A., Das,D. and DasGupta,C. (2008) Protein folding by domain V of Escherichia coli 23S rRNA: Specificity of RNA-protein interactions. *J. Bacteriol.*, **190**, 3344–3352.

Computationally-guided design and selection of high performing ribosomal active site mutants

Camila Kofman[a,b,c], Andrew M. Watkins[d,e], Do Soon Kim[a,b,c], Jessica A. Willi [b,c], Alexandra C. Wooldredge[b,c], Ashty S. Karim[a,b,c], Rhiju Das[d,f], Michael C. Jewett[a,b,c,g]

Affiliations

a. Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL 60208, USA
b. Chemistry of Life Processes Institute, Northwestern University, Evanston, IL 60208, USA
c. Center for Synthetic Biology, Northwestern University, Evanston, IL 60208, USA
d. Department of Biochemistry, Stanford University, Stanford, CA 94305, USA
e. Prescient Design, Genentech, South San Francisco, CA 94080, USA
f. Inceptive Nucleics, Inc. Palo Alto, CA 94304, USA
g. Department of Physics, Stanford University, Stanford, CA 94305, USA
h. Robert H. Lurie Comprehensive Cancer Center and Simpson Querrey Institute, Northwestern University, Chicago, IL 60611, USA

Correspondence

Michael C. Jewett, m-jewett@northwestern.edu

Supplemental Information

Supplementary Tables

Table S1. Primers used in this study.

H75 Insert FP	TGAACCTTTACTATAGCTTGBDBDTGAACATTGAGCCTTGATGTGT
H75 Insert RP	CCAGTCAAACTACCCACCAGBDBDTGTCCGCAACCCGGATTA
H75 Backbone FP	TGGTGGGTAGTTTGACTGGGG
H75 Backbone RP	TGGTGGGTAGTTTGACTGGGG
Mutated fragment FP for RT-PCR (514)	ACACTCTTTCCCTACACGACGCTCTTCCGATCTcgtaatccgggttgcggac
Mutated fragment RP for RT-PCR	GACTGGAGTTCAGACGTGTGCTCTTCCGATCTtcgtactaggagcagcccc

Table S2. iSAT activities, SWM scores, and sequences of constructed H75 mutants. Mutants H75.1-14, highlighted in blue, were selected for by SWM based on the score-cutoff metric as shown in Figure 1B-D.

Name	Normalized iSAT Activity	Standard Deviation	Top SWM Score	Sequence
H75.1*	1.400	0.022	123.577	gguu, cacc
H75.2*	1.387	0.013	122.990	cggg,cccg
H75.3	1.367	0.041	124.748	gugu,acac
H75.4	1.275	0.039	121.649	ggca,agcc
H75.5	1.250	0.017	118.696	cgcg,agcg
H75.6	1.234	0.021	123.039	cgca,ugcg
H75.7	1.079	0.008	123.966	ggcg,cgcc
H75.8	1.023	0.014	123.422	cgcg,cgcg
H75.9	0.901	0.035	123.527	cagu,acug
H75.10	0.871	0.026	123.707	gugg,acac
H75.11	0.792	0.023	124.208	cggu,cgcg
H75.12	0.433	0.002	122.151	ggcg,ugcc
H75.13	0.212	0.006	124.093	cagu,ccug
H75.14	0.195	0.007	123.348	cgcg,agag
H75.15	1.394	0.028	129.105	guuu,aacc
H75.16	1.058	0.010	133.509	uaga,uaua
H75.17	0.973	0.006	126.898	acac,gugu
H75.18	0.930	0.031	125.170	cggu,cccc
H75.19	0.924	0.030	126.049	caac,guug
H75.20	0.885	0.005	126.279	gggu,caca
H75.21	0.819	0.005	127.129	aggc,acag
H75.22	0.794	0.024	127.892	cgcg,cccg
H75.23	0.661	0.013	125.199	cgcg,accg
H75.24	0.624	0.015	126.054	ggcg,ugac
H75.25	0.522	0.008	131.004	gugu,caca
H75.26	0.511	0.007	138.373	auag,caca
H75.27	0.201	0.006	128.066	cggc,aguc
H75.28	0.192	0.005	126.597	cagu,cgug
H75.29	0.182	0.027	138.646	gaag,caca
H75.30	0.171	0.005	133.484	uacg,ccua
H75.31	0.145	0.002	137.995	aucg,caca
H75.32	0.145	0.004	131.605	ggcg,uaac
H75.33	0.122	0.002	125.756	cggc,acuc
H75.34	0.073	0.001	126.812	cggu,cgag
H75.35	0.043	0.003	130.445	gagc,ugca
H75.36	0.034	0.001	128.751	cgcg,acag
H75.37	0.026	0.000	127.855	cgcg,agac
H75.38	0.018	0.001	127.055	cagu,cgcg
H75.39	0.015	0.000	130.303	cggu,ccac
H75.40	0.014	0.001	134.727	ggcg,uaaa
H75.41	0.013	0.001	133.192	ugug,acac
H75.42	0.012	0.001	131.646	cggu,cgaa
H75.43	0.007	0.000	134.992	cggu,caac
H75.44	0.006	0.000	131.166	cgcg,acaa
H75.45	0.004	0.000	136.609	ccaa,aacc
H75.46	0.001	0.000	132.792	cgcg,acac
H75.47	-0.002	0.000	133.290	cagu,cgcc
H75.48	-0.004	0.000	135.088	guaa,agcg
H75.49	-0.006	0.002	137.776	caca,ucac
H75.50	-0.007	0.001	137.197	cucg,uccc

Table S3. SWM scores and sequences of H73, H91 and H92 design constructs. Sequences and scores of the 14 constructs selected from each design simulation.

Construct Name	Relative Activity	Standard Deviation	2046-2050	2618-2622	SWM Score
H73-1	1.333	0.059	cgccg	cggcg	-50.677
H73-2	1.261	0.045	cgcag	cugcg	-54.276
H73-3	1.121	0.041	agccu	aggcc	-53.670
H73-4	1.037	0.027	cgcca	cggcg	-55.429
H73-5	0.998	0.028	cgccg	cggcc	-47.495
H73-6	0.941	0.029	caaua	cguug	-40.547
H73-7	0.715	0.028	cgcag	cugug	-54.779
H73-8	0.686	0.033	cgccg	cagcg	-55.794
H73-9	0.659	0.028	agcag	cggca	-38.793
H73-10	0.643	0.020	agccg	aggcc	-40.986
H73-11	0.435	0.018	aguca	agacc	-48.076
H73-12	0.348	0.027	agcug	cagua	-48.787
H73-13	0.314	0.011	cuccg	aggag	-48.627
H73-14	0.082	0.005	agcca	cugca	-48.101

Construct Name	Relative Activity	Standard Deviation	2523-2527	2536-2540	SWM Score
H91-1	1.178	0.036	ccuag	cgggg	-32.745
H91-2	1.130	0.062	caccg	cggug	-33.896
H91-3	0.880	0.058	acccg	ccggu	-33.291
H91-4	0.709	0.016	accag	ccggu	-30.450
H91-5	0.573	0.022	cacag	ccggg	-30.475
H91-6	0.537	0.042	ccccg	cgggg	-28.604
H91-7	0.162	0.012	ccacg	ccggg	-31.576
H91-8	0.024	0.002	cucug	aagaa	-32.275
H91-9	0.017	0.001	aauag	ccagu	-32.944
H91-10	0.010	0.001	ccaaa	uaaga	-30.967
H91-11	0.009	0.002	acccu	ugaug	-27.221
H91-12	0.008	0.002	aacau	cgggg	-27.104
H91-13	0.005	0.001	accaa	cggga	-28.561
H91-14	0.005	0.000	ccacg	cagug	-32.170

Construct Name	Relative Activity	Standard Deviation	2547-2551	2557-2561	SWM Score
H92-1	0.344	0.002	ggccg	cgggc	-10.061
H92-2	0.261	0.003	gguca	uggcc	-8.448
H92-3	0.204	0.007	ggacg	cggcc	-6.888
H92-4	0.115	0.005	gccaa	uuggc	-6.631
H92-5	0.072	0.003	gccaa	agggc	-7.065
H92-6	0.058	0.001	cgcca	aggcg	-11.034
H92-7	0.031	0.002	ggcca	cggcc	-9.743
H92-8	0.022	0.001	gcucg	agagc	-10.205
H92-9	0.008	0.000	ggccg	agggc	-8.210
H92-10	0.006	0.000	gccca	cggcc	-8.778
H92-11	-0.003	0.003	gccca	agggc	-7.389
H92-12	-0.003	0.001	ggcca	agauc	-7.392
H92-13	-0.004	0.001	gcaaa	aaaca	-7.494
H92-14	-0.008	0.000	gcaca	cgagg	-7.942

Table S4. Overall SWM score ranges for each library simulation.

H73 Design Results	
Min Score	-55.794
Max Score	-8.414
Range	64.208

H91 Design Results	
Min Score	-33.896
Max Score	12.121
Range	46.017

H92 Design Results	
Min Score	-11.034
Max Score	15.442
Range	26.476

Table S5. Sequences and iSAT activities of all single and multi-mutants. "Activity" is signal of construct in iSAT relative to the wildtype ribosome's signal.

Name	Mutant Combination	H73 Sequence	H75 Sequence	H91 Sequence	H92 Sequence	Activity	Std. Dev.
C1	73.1,75.3,91.2,92.1	cgccg, cggcg	guuu, aacc	caccg, cggug	ggccg, cgggc	-0.003	0.001
C2	73.1,75.3,91.2,H92-WT	cgccg, cggcg	guuu, aacc	caccg, cggug	auggc, gccau	0.485	0.007
C3	73.1,75.3,91.1,92.1	cgccg, cggcg	guuu, aacc	ccuag, cgggg	ggccg, cgggc	-0.005	0.001
C4	73.1,75.3,91.1,H92-WT	cgccg, cggcg	guuu, aacc	ccuag, cgggg	auggc, gccau	0.228	0.066
C5	73.1,75.3,H91-WT,92.1	cgccg, cggcg	guuu, aacc	ggggc, guccc	ggccg, cgggc	0.010	0.000
C6	73.1,75.3,H91-WT,H92-WT	cgccg, cggcg	guuu, aacc	ggggc, guccc	auggc, gccau	0.397	0.021
C7	73.1,75.1,91.2,92.1	cgccg, cggcg	gguu, cacc	caccg, cggug	ggccg, cgggc	-0.001	0.000
C8	73.1,75.1,91.2,H92-WT	cgccg, cggcg	gguu, cacc	caccg, cggug	auggc, gccau	0.253	0.030
C9	73.1,75.1,91.1,92.1	cgccg, cggcg	gguu, cacc	ccuag, cgggg	ggccg, cgggc	-0.001	0.001
C10	73.1,75.1,91.1,H92-WT	cgccg, cggcg	gguu, cacc	ccuag, cgggg	auggc, gccau	0.651	0.033
C11	73.1,75.1,H91-WT,92.1	cgccg, cggcg	gguu, cacc	ggggc, guccc	ggccg, cgggc	0.038	0.006
C12	73.1,75.1,H91-WT,H92-WT	cgccg, cggcg	gguu, cacc	ggggc, guccc	auggc, gccau	0.854	0.070
C13	73.1,75.2,91.2,92.1	cgccg, cggcg	cggg, cccg	caccg, cggug	ggccg, cgggc	-0.004	0.001
C14	73.1,75.2,91.2,H92-WT	cgccg, cggcg	cggg, cccg	caccg, cggug	auggc, gccau	0.055	0.016
C15	73.1,75.2,91.1,92.1	cgccg, cggcg	cggg, cccg	ccuag, cgggg	ggccg, cgggc	0.004	0.005
C16	73.1,75.2,91.1,H92-WT	cgccg, cggcg	cggg, cccg	ccuag, cgggg	auggc, gccau	0.067	0.014
C17	73.1,75.2,H91-WT,92.1	cgccg, cggcg	cggg, cccg	ggggc, guccc	ggccg, cgggc	0.002	0.002
C18	73.1,75.2,H91-WT,H92-WT	cgccg, cggcg	cggg, cccg	ggggc, guccc	auggc, gccau	0.741	0.057
C19	73.1,H75-WT,91.2,92.1	cgccg, cggcg	acac, gugu	caccg, cggug	ggccg, cgggc	-0.003	0.000
C20	73.1,H75-WT,91.2,H92-WT	cgccg, cggcg	acac, gugu	caccg, cggug	auggc, gccau	0.350	0.033
C21	73.1,H75-WT,91.1,92.1	cgccg, cggcg	acac, gugu	ccuag, cgggg	ggccg, cgggc	-0.006	0.001
C22	73.1,H75-WT,91.1,H92-WT	cgccg, cggcg	acac, gugu	ccuag, cgggg	auggc, gccau	0.328	0.046
C23	73.1,H75-WT,H91-WT,92.1	cgccg, cggcg	acac, gugu	ggggc, guccc	ggccg, cgggc	0.000	0.003
C24	73.2,75.3,91.2,92.1	cgcag, cugcg	guuu, aacc	caccg, cggug	ggccg, cgggc	-0.005	0.001
C25	73.2,75.3,91.2,H92-WT	cgcag, cugcg	guuu, aacc	caccg, cggug	auggc, gccau	0.352	0.003
C26	73.2,75.3,91.1,92.1	cgcag, cugcg	guuu, aacc	ccuag, cgggg	ggccg, cgggc	-0.003	0.000
C27	73.2,75.3,91.1,H92-WT	cgcag, cugcg	guuu, aacc	ccuag, cgggg	auggc, gccau	0.262	0.003
C28	73.2,75.3,H91-WT,92.1	cgcag, cugcg	guuu, aacc	ggggc, guccc	ggccg, cgggc	-0.001	0.000
C29	73.2,75.3,H91-WT,H92-WT	cgcag, cugcg	guuu, aacc	ggggc, guccc	auggc, gccau	0.845	0.044
C30	73.2,75.1,91.2,92.1	cgcag, cugcg	gguu, cacc	caccg, cggug	ggccg, cgggc	-0.001	0.000
C31	73.2,75.1,91.2,H92-WT	cgcag, cugcg	gguu, cacc	caccg, cggug	auggc, gccau	0.484	0.027
C32	73.2,75.1,91.1,92.1	cgcag, cugcg	gguu, cacc	ccuag, cgggg	ggccg, cgggc	-0.001	0.000
C33	73.2,75.1,91.1,H92-WT	cgcag, cugcg	gguu, cacc	ccuag, cgggg	auggc, gccau	0.450	0.019
C34	73.2,75.1,H91-WT,92.1	cgcag, cugcg	gguu, cacc	ggggc, guccc	ggccg, cgggc	0.028	0.001
C35	73.2,75.1,H91-WT,H92-WT	cgcag, cugcg	gguu, cacc	ggggc, guccc	auggc, gccau	1.243	0.023
C36	73.2,75.2,91.2,92.1	cgcag, cugcg	cggg, cccg	caccg, cggug	ggccg, cgggc	0.003	0.001
C37	73.2,75.2,91.2,H92-WT	cgcag, cugcg	cggg, cccg	caccg, cggug	auggc, gccau	0.431	0.039
C38	73.2,75.2,91.1,92.1	cgcag, cugcg	cggg, cccg	ccuag, cgggg	ggccg, cgggc	0.001	0.001
C39	73.2,75.2,91.1,H92-WT	cgcag, cugcg	cggg, cccg	ccuag, cgggg	auggc, gccau	0.035	0.006
C40	73.2,75.2,H91-WT,92.1	cgcag, cugcg	cggg, cccg	ggggc, guccc	ggccg, cgggc	0.006	0.002
C41	73.2,75.2,H91-WT,H92-WT	cgcag, cugcg	cggg, cccg	ggggc, guccc	auggc, gccau	0.340	0.016
C42	73.2,H75-WT,91.2,92.1	cgcag, cugcg	acac, gugu	caccg, cggug	ggccg, cgggc	0.001	0.001
C43	73.2,H75-WT,91.2,H92-WT	cgcag, cugcg	acac, gugu	caccg, cggug	auggc, gccau	0.001	0.001
C44	73.2,H75-WT,91.1,92.1	cgcag, cugcg	acac, gugu	ccuag, cgggg	ggccg, cgggc	0.419	0.029
C45	73.2,H75-WT,91.1,H92-WT	cgcag, cugcg	acac, gugu	ccuag, cgggg	auggc, gccau	0.454	0.051
C46	73.2,H75-WT,H91-WT,92.1	cgcag, cugcg	acac, gugu	ggggc, guccc	ggccg, cgggc	0.010	0.002
C47	H73-WT,75.3,91.2,92.1	gcggc, gccgu	guuu, aacc	caccg, cggug	ggccg, cgggc	0.000	0.001
C48	H73-WT,75.3,91.2,H92-WT	gcggc, gccgu	guuu, aacc	caccg, cggug	auggc, gccau	1.148	0.025
C49	H73-WT,75.3,91.1,92.1	gcggc, gccgu	guuu, aacc	ccuag, cgggg	ggccg, cgggc	-0.001	0.000
C50	H73-WT,75.3,91.1,H92-WT	gcggc, gccgu	guuu, aacc	ccuag, cgggg	auggc, gccau	0.420	0.026
C51	H73-WT,75.3,H91-WT,92.1	gcggc, gccgu	guuu, aacc	ggggc, guccc	ggccg, cgggc	0.051	0.003
C52	H73-WT,75.1,91.2,92.1	gcggc, gccgu	gguu, cacc	caccg, cggug	ggccg, cgggc	0.005	0.001
C53	H73-WT,75.1,91.2,H92-WT	gcggc, gccgu	gguu, cacc	caccg, cggug	auggc, gccau	1.232	0.006
C54	H73-WT,75.1,91.1,92.1	gcggc, gccgu	gguu, cacc	ccuag, cgggg	ggccg, cgggc	0.007	0.002
C55	H73-WT,75.1,91.1,H92-WT	gcggc, gccgu	gguu, cacc	ccuag, cgggg	auggc, gccau	0.410	0.022
C56	H73-WT,75.1,H91-WT,92.1	gcggc, gccgu	gguu, cacc	ggggc, guccc	ggccg, cgggc	0.179	0.002

C57	H73-WT,75.2,91.2,92.1	gcggc, gccgu	cggg, cccg	caccg, cggug	ggccg, cgggc	0.005	0.000
C58	H73-WT,75.2,91.2,H92-WT	gcggc, gccgu	cggg, cccg	caccg, cggug	auggc, gccau	0.004	0.005
C59	H73-WT,75.2,91.1,92.1	gcggc, gccgu	cggg, cccg	ccuag, cgggg	ggccg, cgggc	0.036	0.003
C60	H73-WT,75.2,91.1,H92-WT	gcggc, gccgu	cggg, cccg	ccuag, cgggg	auggc, gccau	0.542	0.023
C61	H73-WT,75.2,H91-WT,92.1	gcggc, gccgu	cggg, cccg	ggggc, guccc	ggccg, cgggc	0.089	0.010
C62	H73-WT,H75-WT,91.2,92.1	gcggc, gccgu	acac, gugu	caccg, cggug	ggccg, cgggc	0.012	0.000
C63	H73-WT,H75-WT,91.1,92.1	gcggc, gccgu	acac, gugu	ccuag, cgggg	ggccg, cgggc	0.101	0.008
H92.1	H73-WT,H75-WT,H91-WT,H92-7	gcggc, gccgu	acac, gugu	ggggc, guccc	ggccg, cgggc	0.344	0.002
H91.1	H73-WT,H75-WT,H91-11,H92-WT	gcggc, gccgu	acac, gugu	ccuag, cgggg	auggc, gccau	1.178	0.036
H91.2	H73-WT,H75-WT,H91-9,H92-WT	gcggc, gccgu	acac, gugu	caccg, cggug	auggc, gccau	1.130	0.062
H73.1	H73-10,H75-WT,H91-WT,H92-WT	cgccg, cggcg	acac, gugu	ggggc, guccc	auggc, gccau	1.333	0.059
H75.2	H73-WT,H75-39,H91-WT,H92-WT	gcggc, gccgu	cggg, cccg	ggggc, guccc	auggc, gccau	1.387	0.013
H73.2	H73-8,H75-WT,H91-WT,H92-WT	cgcag, cugcg	acac, gugu	ggggc, guccc	auggc, gccau	1.261	0.045
H75.3	H73-WT,H75-43,H91-WT,H92-WT	gcggc, gccgu	guuu, aacc	ggggc, guccc	auggc, gccau	1.394	0.028
H75.1	H73-WT,H75-41,H91-WT,H92-WT	gcggc, gccgu	gguu, cacc	ggggc, guccc	auggc, gccau	1.400	0.022
Wildtype	H73-WT,H75-WT,H91-WT,H92-WT	gcggc, gccgu	acac, gugu	ggggc, guccc	auggc, gccau	0.979	0.029

Name	Normalized sfGFP in iSAT	Std. Dev.	Supports life (Y/N)
H75.1	1.400	0.022	Y
H75.2	1.387	0.013	Y
H73.1	1.333	0.059	Y
H73.2	1.261	0.045	Y
C35	1.243	0.023	Y
C53	1.232	0.006	Y
H91.1	1.178	0.036	Y
H91.2	1.130	0.062	Y
C12	0.854	0.070	Y
C18	0.741	0.057	Y
C10	0.651	0.033	N
C60	0.542	0.023	Y
C31	0.484	0.027	N
C45	0.454	0.051	Y
C33	0.450	0.019	Y
C37	0.431	0.039	N
C44	0.419	0.029	N
C55	0.410	0.022	Y
C20	0.350	0.033	N
H92.1	0.344	0.002	N
C41	0.340	0.016	Y
C22	0.328	0.046	Y
C8	0.253	0.030	N
C4	0.228	0.066	N
C56	0.179	0.002	N
C63	0.101	0.008	N
C61	0.089	0.010	N
C16	0.067	0.014	Y
C14	0.055	0.016	N
C51	0.051	0.003	N
C11	0.038	0.006	N
C59	0.036	0.003	N
C39	0.035	0.006	Y
C34	0.028	0.001	N
C62	0.012	0.000	N
C5	0.010	0.000	N
C46	0.010	0.002	N
C54	0.007	0.002	N
C40	0.006	0.002	N
C52	0.005	0.001	N
C57	0.005	0.000	N
C15	0.004	0.005	N
C58	0.004	0.005	N
C36	0.003	0.001	N
C17	0.002	0.002	N
C38	0.001	0.001	N
C42	0.001	0.001	N
C43	0.001	0.001	N
C23	0.000	0.003	Y
C47	0.000	0.001	N
C7	-0.001	0.000	N
C9	-0.001	0.001	N
C28	-0.001	0.000	N
C30	-0.001	0.000	N
C32	-0.001	0.000	N
C49	-0.001	0.000	N
C1	-0.003	0.001	N
C19	-0.003	0.000	N
C26	-0.003	0.000	N
C13	-0.004	0.001	N
C3	-0.005	0.001	N
C24	-0.005	0.001	N
C21	-0.006	0.001	N

Table S6. Analysis of iSAT data for prediction of ability to support life in cells. Constructs were ranked by their normalized sfGFP yield in iSAT in descending order to visualize trends in iSAT performance and ability to support life. Constructs that were able to support life are highlighted in blue. Of the constructs that had a normalized activity greater than 30% in iSAT, 16/22 (73%) were able to support life.

Supplementary Figures

Figure S1. 23S rRNA sequence alignments for regions mutated in this study. 1,614 bacterial and archaeal 23S rRNA sequences were aligned to visualize the sequence conservation in the helices explored in this study.

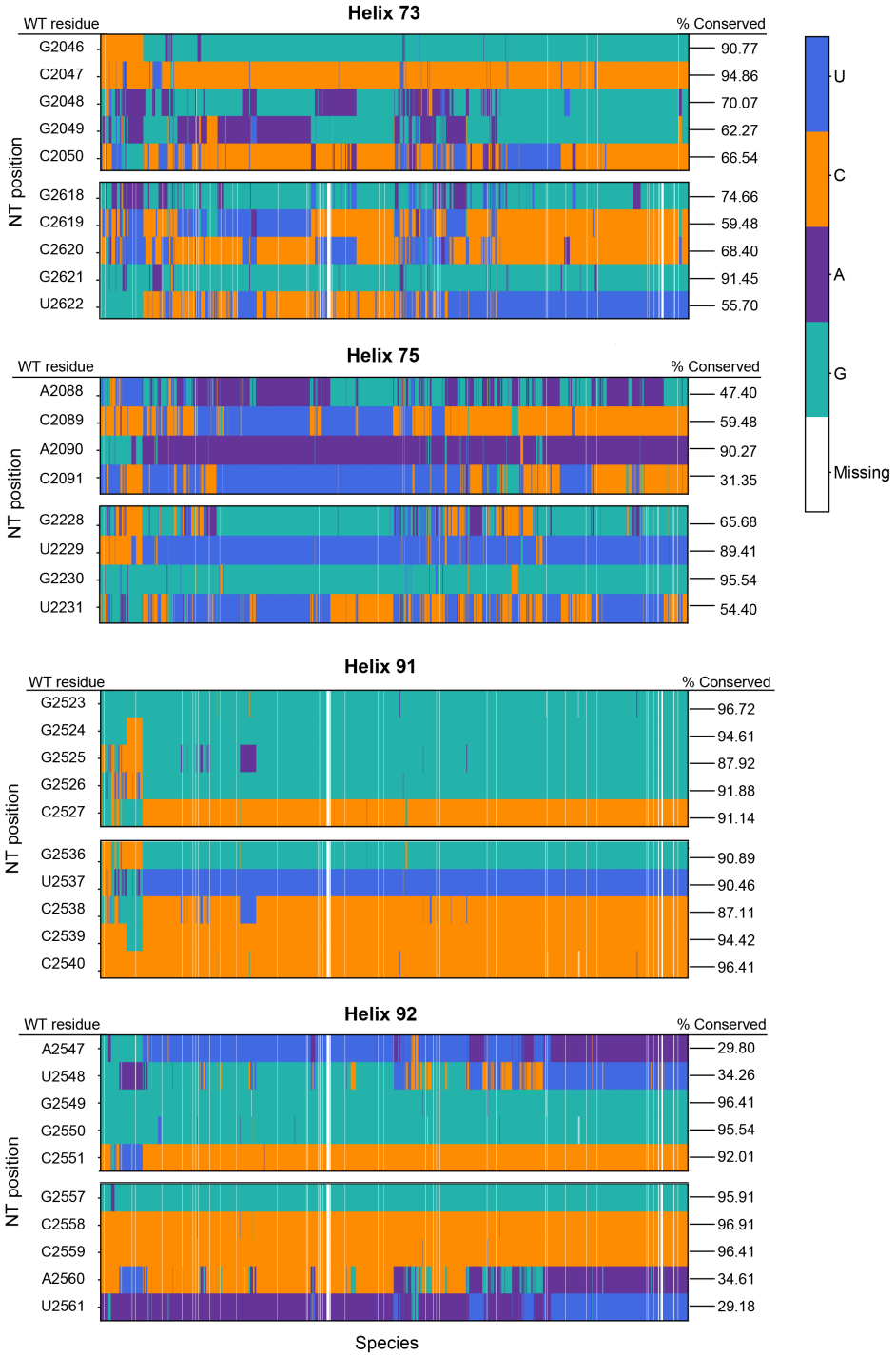

Figure S2. Complete heatmaps of single, double and triple combination construct activities in iSAT. Heatmaps show expression of sfGFP in iSAT normalized by the amount produced by the wildtype control sequence.

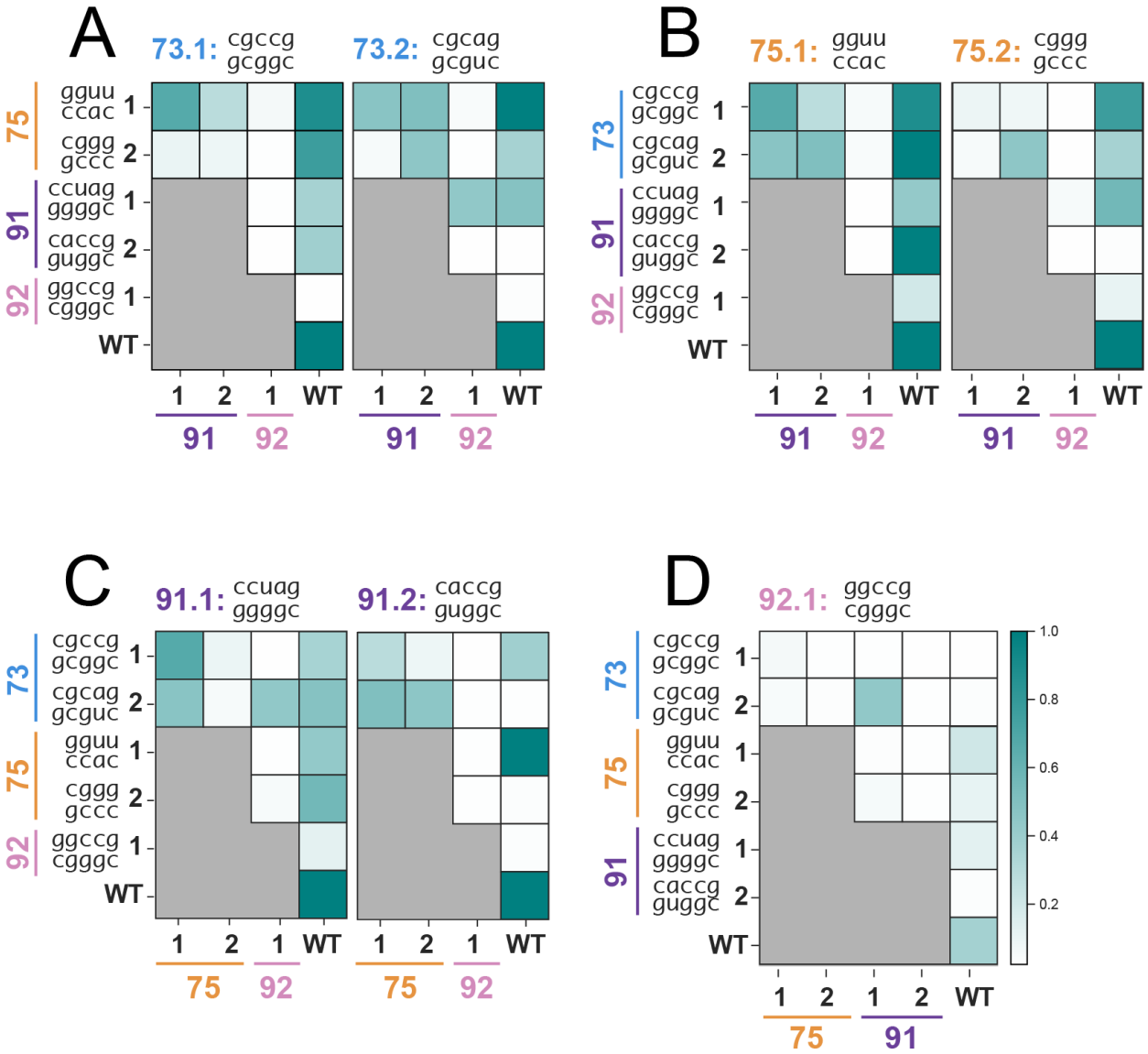

Figure S3. Alignment of Sanger sequencing of RT-PCR products from total RNA extraction of SQ171fg strains carrying mutant ribosomes after selection. Alignment performed using SnapGene (from Insightful Science; available at snapgene.com). Mutations match expected genotypes as shown in Table S5.