ARTICLES
https://doi.org/10.1038/s41589-022-01064-w

nature chemical biology

Check for updates

Three-dimensional structure-guided evolution of a ribosome with tethered subunits

Do Soon Kim[1,2,8,12], Andrew Watkins[3,9,12], Erik Bidstrup[1,2,10], Joongoo Lee[1,2,11], Ved Topkar[3], Camila Kofman[1,2], Kevin J. Schwarz[4], Yan Liu[5], Grigore Pintilie[6], Emily Roney[1,2], Rhiju Das[3,7] and Michael C. Jewett[1,2] ✉

RNA-based macromolecular machines, such as the ribosome, have functional parts reliant on structural interactions spanning sequence-distant regions. These features limit evolutionary exploration of mutant libraries and confound three-dimensional structure-guided design. To address these challenges, we describe Evolink (evolution and linkage), a method that enables high-throughput evolution of sequence-distant regions in large macromolecular machines, and library design guided by computational RNA modeling to enable exploration of structurally stable designs. Using Evolink, we evolved a tethered ribosome with a 58% increased activity in orthogonal protein translation and a 97% improvement in doubling times in SQ171 cells compared to a previously developed tethered ribosome, and reveal new permissible sequences in a pair of ribosomal helices with previously explored biological function. The Evolink approach may enable enhanced engineering of macromolecular machines for new and improved functions for synthetic biology.

Directed evolution of macromolecular machines can explain principles of biological design and generate new catalytic apparatuses for synthetic biology[1–8]. Unfortunately, directed evolution methods can be hindered by practical considerations. The combinatorial space for evolution is immense (for example, in a 300 amino acid protein, there are roughly 20^{300} possible amino acid sequences), and random mutagenesis alone cannot screen all possible variants[9–12]. Furthermore, macromolecular machines often have complex tertiary structures that contribute to their function[13], where residues distant in sequence are close in three-dimensional (3D) space (Fig. 1a). Even given effective selections, performant designs cannot be easily recovered. Such practical limitations are exacerbated in large macromolecular machines, such as the bacterial ribosome, which has three ribosomal RNAs (rRNAs) comprising roughly 4,500 nucleotides (nt) (that is, the 16S rRNA, 23S rRNA and 5S rRNA) and 54 proteins[1–4,8,9,14].

Directed evolution of the ribosome has emerged as a promising opportunity in chemical and synthetic biology[1–5,7–9,14–22]. A major goal of ribosome engineering is to repurpose the ribosome for diverse genetically encoded chemistries to create new classes of enzymes, therapeutics and materials by selectively incorporating noncanonical monomers into peptides and proteins. While the natural ribosome works well for many noncanonical α-amino acids, there is poor compatibility with the natural translation apparatus for numerous classes of non-α-amino acids (for example, backbone-extended amino acids; γ–, δ–, ε– and so on) leading to inefficiencies in incorporation[1–4,23,24].

Methods for engineering ribosomes have been developed to address these inefficiencies[7,16,21,25,26]. In vivo, tethered ribosomes have made possible the first fully orthogonal ribosomal messenger RNA system in cells, where a subpopulation of ribosomes are available for engineering and are independent from wild-type ribosomes supporting cell life[18]. Tethered ribosome systems have two key features. First, the anti-Shine–Dalgarno sequence of the 16S rRNA can be mutated yielding orthogonal ribosomes that selectively initiate translation of orthogonal mRNAs (o-mRNAs) with mutated Shine–Dalgarno sequences[19,27,28]. Second, the small and large ribosomal subunits are covalently linked together (Fig. 1b). In the first tethered ribosome system, Ribo-T, the core 16S and 23S rRNAs were joined together to form a single chimeric molecule via helix h44 of the 16S rRNA and helix H101 of the 23S rRNA[18]. By selecting otherwise dominantly lethal rRNA mutations in the large ribosomal subunit, Ribo-T was evolved to synthesize protein sequences that are inaccessible to the natural ribosome[18]. Since the initial discovery of Ribo-T and a subsequent stapled design[15], new orthogonal Ribo-T–mRNA pairs as well as tether sequences have been optimized[9,14,22]. During optimization, tether residues were randomized in sequence but not in length[9], and mutations to surrounding residues surrounding a fixed RNA linker (the J5/5a junction from the *Tetrahymena* group I intron) were investigated[14]. Despite the improvement, the potential of tethered ribosomes remains limited by their low activity.

The untapped potential and existing inefficiencies of tethered ribosome systems motivate new directed evolution-based approaches to engineer these systems and improve their activity. Previous works were limited in throughput in evaluating designs (for example, 48 and 108 members were evaluated in two previous efforts[9,14]) due to reliance on clonal isolation and functional testing. A possible bottleneck has been that the regions of interest in

[1]Department of Chemical and Biological Engineering, Northwestern University, Evanston, IL, USA. [2]Center for Synthetic Biology, Northwestern University, Evanston, IL, USA. [3]Department of Biochemistry, Stanford University, Stanford, CA, USA. [4]Department of Chemistry, University of Illinois Urbana Champaign, Champaign, IL, USA. [5]Division of CryoEM and Bioimaging, SSRL, SLAC National Accelerator Laboratory, Menlo Park, CA, USA. [6]Department of Bioengineering, Stanford University, Stanford, CA, USA. [7]Department of Physics, Stanford University, Stanford, CA, USA. [8]Present address: Inceptive Nucleics, Inc., Palo Alto, CA, USA. [9]Present address: Prescient Design, Genentech, South San Francisco, CA, USA. [10]Present address: Robert F. Smith School of Chemical and Biomolecular Engineering, Cornell University, Ithaca, NY, USA. [11]Present address: Department of Chemical Engineering, Pohang University of Science and Technology, Pohang, Republic of Korea. [12]These authors contributed equally: Do Soon Kim, Andrew Watkins. ✉e-mail: m-jewett@northwestern.edu

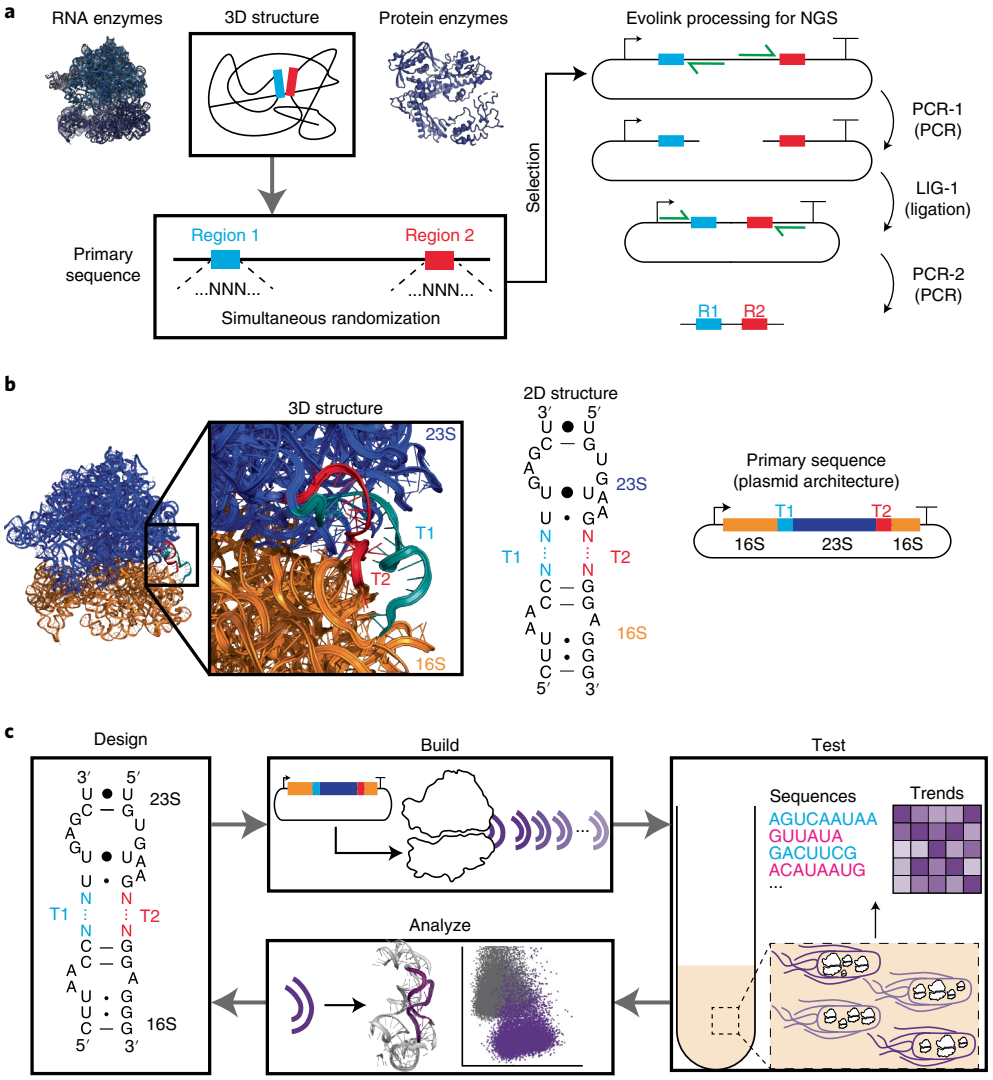


Fig. 1 | Overview of Evolink and tethered ribosome design and evolution. a, RNA- and protein-based enzymes with regions that are distal in primary sequence but proximate in 3D space (regions 1 and 2, blue and red, respectively), and are likely functionally linked. Molecular biology steps of Evolink (PCR-1, LIG-1, PCR-2) to link regions together in a single amplicon that enable overlapping NGS readouts. DNA oligos (green), can be flexibly designed depending on the machine architecture encoded on a plasmid. **b**, Rosetta-predicted structure of a previously reported tethered ribosome showing tethers, denoted T1 and T2, in 3D space as well as likely secondary structure representation. Representative encoding plasmid (right) is shown. **c**, The DBTA evolution scheme. Test includes selection, Evolink and the resulting NGS reads. Analyze involves Rosetta modeling to infer tether structure and predicted stability. Results from each round feed into Design and Build.

the tethered ribosomes are separated by roughly 2,900 nt (the length of the circularly permuted 23S rRNA[18]), and current readily available methods for next-generation sequencing (NGS) are typically limited to overlapping read lengths of roughly 300 nt. Methods that address this[29,30] face limitations that hinder applications to large macromolecular machines such as the ribosome. Briefly, they rely on custom bioinformatic pipelines, barcoding strategies inherent to protein-based machines, or are limited in the distance between regions of interest[29–31].

With these limitations in mind, we present a molecular biology technique called Evolink (evolution and linkage) (Fig. 1a). Evolink connects two or more regions of nucleic acid sequence that are distant in primary sequence but close in 3D structure (in RNA or protein form) to enable NGS readouts of winning phenotypes. Evolink uses low-cost methods familiar to most molecular biology practitioners with flexible design parameters that lower the barrier for powerful directed evolution experiments without requiring advanced knowledge of statistics or computational methods to infer regions of likely coevolution in macromolecules as previous approaches necessitated[29–31]. In this work, we apply Evolink to improve tethered ribosomes and to study the function of rRNA helices implicated in ribosome subunit interaction and function.

The combination of Evolink with computational modeling allows for efficient evolution of macromolecular machines with complex structures, such as the ribosome. Because Evolink generates a plasmid as its endpoint molecule, it can be iterated to link multiple regions together and can be adjusted to different molecules of interest. The limits of Evolink are bound by available sequencing read lengths, and thus we expect its use will scale with advances in sequencing technology. Looking forward, we anticipate the Evolink approach will be valuable for future engineering of ribosomes and other macromolecular machines.

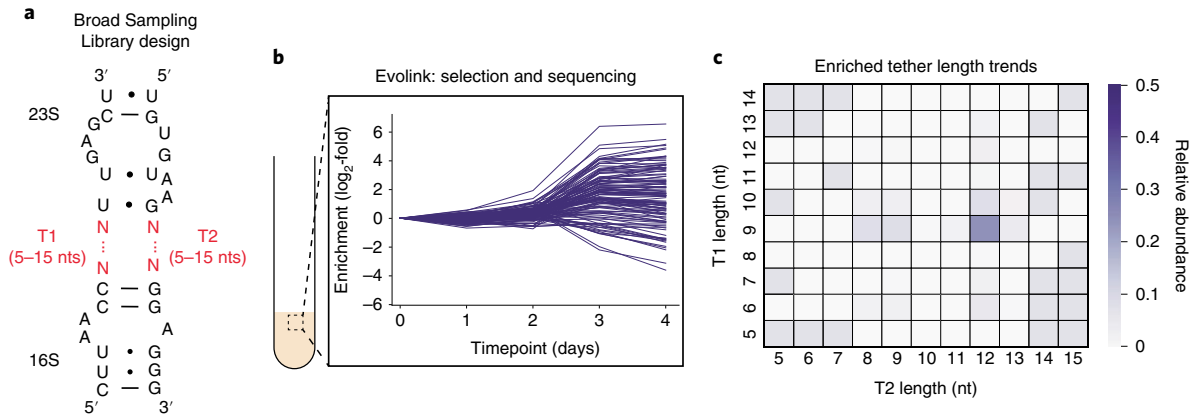


Fig. 2 | Results of the Broad Sampling Library. a, Residues targeted in this library (red) depicted with surrounding residues (black) in native secondary structure. **b**, Fold enrichment (log_2) of tether sequence pairs during selection in liquid culture over four time points (one timepoint per day for 4 days). **c**, Analysis of the NGS results reveals convergence toward 9 and 12 nt for T1 and T2 regions, respectively. Data representative of three independent experiments. On average, 1,789 unique genotypes were observed with sequencing coverage greater than ten reads per genotype. Sequencing of the starting plasmid library revealed a diversity greater than 800,000. The theoretical diversity of this library was roughly 2.1×10^{18}.

Results

Linking of sequence-distant regions on a single NGS read. We aimed to develop a generalizable method guided by computational design for directed evolution of sequence-distant sites of macromolecular machines, first focusing on evolving the tether sequences in tethered ribosomes. We started by developing Evolink, a three-step process that uses polymerase chain reaction (PCR), ligation and a second PCR to bring together formerly separated regions of a plasmid for a continuous NGS read. This process is analogous to amplifying and closing the 'backbone' of a plasmid, where the 'insert' omitted from amplification is the sequence separating the two regions of interest. Because Evolink uses simple, general-purpose molecular biology techniques (PCR and ligation), it can extend to other plasmid-encoded molecular machines (Fig. 1a).

To start, we demonstrated the three key molecular biology steps of Evolink (termed PCR-1, LIG-1, PCR-2) (Fig. 1a, right) on a clonal plasmid sample encoding Ribo-T v2 (9) (Extended Data Fig. 1a). In our architecture for PCR-1, in which T1 is upstream (5′) of T2, the forward primer binds upstream of T2 and the reverse primer binds downstream (3′) of T1, such that the primers for PCR-1 bind 'inside' the two regions of interest (Fig. 1a, top right). For PCR-1 primers, the sequence between each respective primer and region of interest (reverse primer-T1 and forward primer-T2 in this case) defines the NGS amplicon length. Additionally, the primers' sequences can be adapted for restriction enzyme-based, isothermal assembly[31], or blunt-end ligation for the subsequent ligation (LIG-1). We assessed the compatibility of PCR-1 with multiple primer sets with designed overhangs for Type I/II restriction enzyme digestion, 5′ phosphorylation for blunt-end ligation or overlapping complimentary sequences for isothermal assembly. We found PCR-1 to be successful with all four primer sets featuring different 5′ modifications (either phosphorylation or custom sequences) (Extended Data Fig. 1b).

Following PCR-1, the product was cyclized via unimolecular ligation in LIG-1, proximally linking the previously distant regions. PCR-1 products that used primers harboring restriction enzyme cut sites were processed with a corresponding enzymatic digest. Those that used 5′ phosphorylated primers or enzymatic digestion were ligated with T4 ligase, while those featuring overlapping complementary sequences used isothermal assembly for ligation[32].

Finally, we carried out PCR-2 with a new set of primers to amplify the now-linked regions of interest. In PCR-2, the primers are designed with the forward primer upstream of T1 and the reverse primer downstream of T2, such that now the primers are 'outside' of the regions of interest (Fig. 1a, bottom right). The sequences between each respective primer and region of interest (forward primer-T1 and reverse primer-T2 in this case) contribute to the final NGS amplicon. We designed primers such that the final product is roughly 200 nt and can be directly used in NGS library preparation. PCR-2 was performed with products from four different ligation methods (Type I/II restriction enzyme digestion and ligation, blunt-end ligation and isothermal assembly), each with eight different input template amounts into the ligation (1, 2, 5, 10, 20, 30, 40, 50 ng). We observed successful generation of the desired amplicon for NGS for all 32 reactions tested (Extended Data Fig. 1), and moved forward with blunt-end ligation because it did not rely on any specific DNA sequence and proved successful with the minimum amount of template tested.

Applying Evolink to tethered ribosomes. We sought to apply Evolink to develop mutant tethered ribosomes for improved activity (Fig. 1b). Specifically, we looked to improve on the function of Ribo-T v2 (ref. [9]) by simultaneously optimizing residues comprising the tether for length and sequence, leveraging the throughput of Evolink and post facto structural modeling. Central to our efforts was the iterative application of design-build-test-analyze (DBTA) cycles (Fig. 1c), where multiple libraries can be tested, each library building on results and analysis of ones previously. Previous efforts carrying out a single pass of library design, building and selection/screening, were limited by the breadth of the libraries tested. Our iterative process allowed for progressively deeper focus without compromising breadth. Because Evolink uses NGS, our approach allows for larger sampling and screening of the solution space compared to past efforts.

First, we elected to broadly sample possible lengths and sequences of T1 and T2 with a degenerate library of 5–15 nt (Fig. 2a). The library of tether designs was transformed into an *Escherichia coli* strain SQ171 lacking *rrn* operons on the genome[33] and viable cells, which survive solely on tethered ribosomes, were identified by growth on agar[9,18]. Resulting colonies were collected and selection was carried out in liquid culture (Fig. 2b). We passaged cells in liquid culture for roughly 40 generations, hypothesizing that faster growing mutants would become more enriched. Plasmids from the culture were collected daily over four days and subjected to Evolink and NGS. T1 and T2 sequences, which represent the two strands of RNA that make up the tether, were

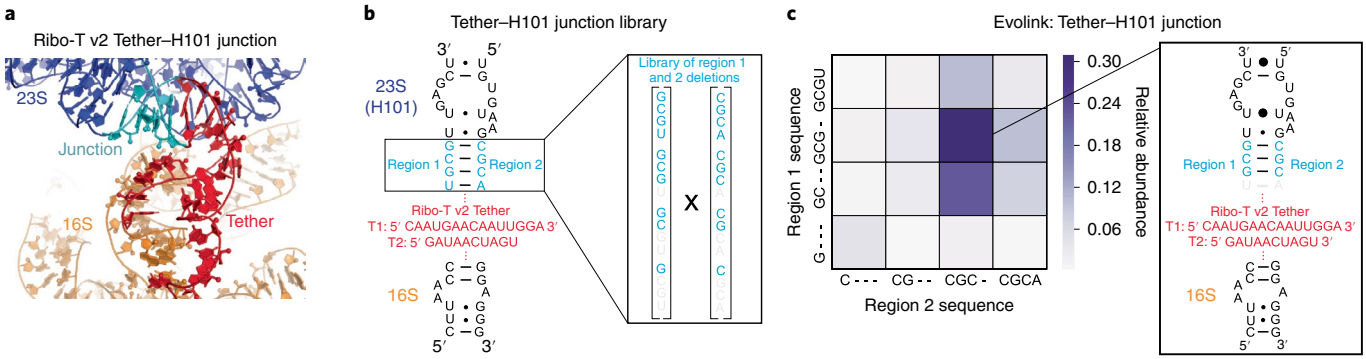


Fig. 3 | Investigation of the Tether–H101 junction. a, Rosetta modeling of the Ribo-T v2 Tether and surrounding residues. The junction (cyan) consists of nucleotides that connect the Tether (red) to the rest of H101 (blue) in the 23S rRNA. **b**, Secondary structure depiction of the library for testing deletion effects in the junction. **c**, Results from Evolink showing convergence toward specific Ribo-T v2 Tether–H101 junction sequence. Heatmap data representative of three independent experiments. The theoretical diversity of this library was 16 and this was observed in the sequencing of the starting plasmid library as well as the genotypes sequenced throughout the Evolink process.

directly linked in a single amplicon for NGS, taking advantage of overlapping reads with high fidelity to improve identification of pairwise interactions between the two regions. Subsequent analysis revealed a range of enrichments for many genotypes observed over the time course (Fig. 2b). Specifically, we observed enrichment ($\log_2$-fold change) values between about 5 and 6, and roughly 1,800 unique genotypes after the LB-agar-based selection converging to roughly 450 unique genotypes over the time course (Fig. 2b and Extended Data Fig. 2). Two key features emerged from these data. First, the same T1 sequences paired with multiple T2 sequences (Supplementary Fig. 1). For example, T1: 5′-CAGGGUACACC-3′ paired with T2: 5′-CCCAUUCA-3′, 5′-AUUCACUUGG-3′ and 5′-CGACGAGCG-3′ to yield enrichment values of 5.69, 2.17 and −1.5, respectively. These data indicate that contributions of the two tether sequences to overall ribosome assembly and function depend on each other and are not simply additive. Second, we observed a trend in the sequencing data toward specific optimal tether lengths, converging on a length of 9 nt for T1 and 12 nt for T2 (Fig. 2c).

Structural fragility of the Tether–H101 junction. Based on previous literature that showed stapled ribosome function is sensitive to the connection between the tether and 23S rRNA residues[14] ('Tether–H101 junction'), we wondered whether the Tether–H101 junction would also be important in the Ribo-T design context[9,18] (Fig. 3a). To explore this question, we next fixed the tether identity according to the Ribo-T v2 sequence[9] and constructed a library that consisted of every possible combination of base deletions in the Tether–H101 junction region (Fig. 3b). This systematic approach avoided the assumption that the wild-type ribosome's secondary structure was preserved when tethered. Similar to the Broad Sampling Library, we used the SQ171 strain for selection. Evolink results on this library converged to 5′-GCG-3′ and 5′-CGC-3′ in regions 1 and 2, respectively, revealing that base changes in the Tether–H101 junction indeed affect tethered ribosome function (Fig. 3c). These results indicated that the folding behavior of this junction may have an important influence on both tethered ribosome structure and function.

To further explore this hypothesis, we turned to computational modeling to gauge structural stability of the Tether–H101 junction (Supplementary Fig. 2). We suspected secondary structure modeling with ViennaRNA[34] and tertiary structure modeling with Rosetta FARFAR2 (ref. [35]) might help to understand possible structural features that may contribute to improved tether RNAs and overall ribosome function, and use those insights to inform subsequent library design. First, we used RNAcofold to conduct secondary structure predictions on the four most prevalent tether sequences that emerged from the Broad Sampling Library (for example, a 10 or 12 nt tether, T1: 5′-AUGACAUGGU-3′ and T2: 5′-CCGGCUUCGGAA-3′) to assess the degree to which tether structure was dependent on its structural context (Supplementary Fig. 2). If the tether's structure is perfectly independent of the surrounding residues, the same base pairing would be observed regardless of surrounding residues included in the RNAcofold analysis. We computed the minimum free energy secondary structure of the tether under two different conditions. The first, 'unconstrained' calculation, allowed the adjacent 23S rRNA junction (Helix 101 in the wild-type ribosomal 23S rRNA) to 're-fold' rather than constraining it to adopt the base pairing observed in experimental structures of the *E. coli* ribosome[36]. In the second, 'constrained' calculation, the 23S rRNA junction residues must adopt native base pairing. For three of four tethers, we observed the same tether base pairs in the constrained and unconstrained structures, but the adjacent 23S junction maintained its wild-type structure in only one case (Supplementary Fig. 2a,c,d). For the remaining tether, significantly different RNA secondary structures were observed between the 'constrained' and 'unconstrained' models (Supplementary Fig. 2b). To further characterize tether flexibility, we also calculated delta entropy at each position, paying particular attention to residues predicted to substantially rearrange in our secondary structure modeling (Supplementary Fig. 2).

We conducted 3D modeling of these tethers to augment our understanding (Fig. 4a–d and Extended Data Fig. 3a–d), using Rosetta's RNA fragment assembly code[35] to model analogous constrained and unconstrained states of the tether with FARFAR2 (Fig. 4b,c, respectively, and Extended Data Fig. 3a–d). For each tether, the constrained and unconstrained simulations resulted in significantly different structures and energy distributions (compare Fig. 4b,c; also Fig. 4d and Extended Data Fig. 3a–d), suggesting that the Tether–H101 junction may not be particularly stable. Our results from investigating the Tether–H101 junction, both experimentally and computationally, led us to reinforce the structure of the Tether–H101 junction, as well as to optimize tether length and sequence together in subsequent rounds of directed evolution. Instead of the naturally occurring 5′-GCG-3′/5′-CGC-3′ stem, we chose a 5′-CUG-3′/5′-CAG-3′ sequence in case a two-register shift might compromise the hypothesis behind our design.

Evolink and validation of a Designed Junction library. With the range of tether lengths informed by the Broad Sampling Library and the designed base pairs at the Tether–H101 junction, we next

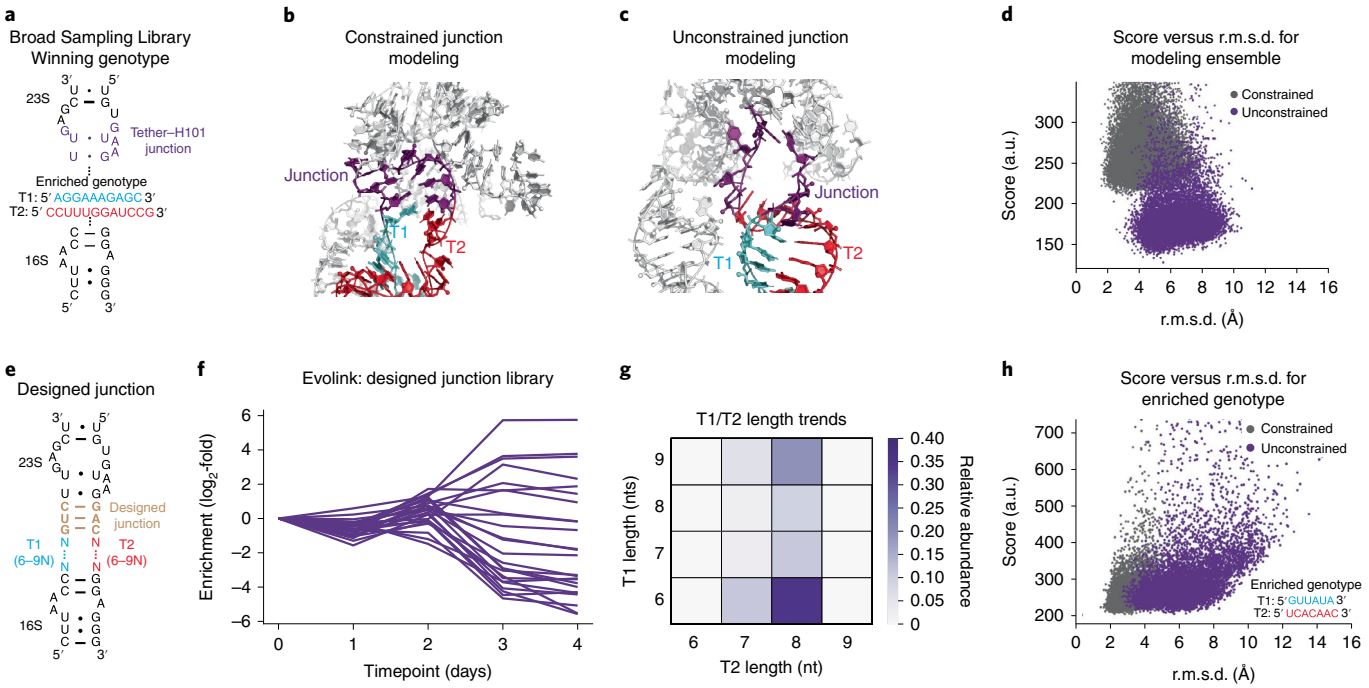


Fig. 4 | Integration and validation of designed junction into library design. a, The sequence of T1 and T2 tethers selected from the Broad Sampling Library. **b**,**c**, Rosetta modeling of the Tether–23S junction (purple) shows large differences between constraining (**b**) or not constraining (**c**) to native base pairing. **d**, Rosetta score versus r.m.s.d. for constrained and unconstrained models of the enriched sequence. a.u., arbitrary units. **e**, Library with the Designed Tether–23S junction, reinforced by three synthetic base pairs (gold). **f**, Representative fold enrichment (log_2) of tether sequences from selection and Evolink on the Designed Tether–23S junction library. Data are representative of three independent experiments, and we observed on average 1,483 unique genotypes with sequencing coverage greater than ten reads per genotype. Sequencing of the starting plasmid library revealed a diversity greater than 800,000. The theoretical diversity of this library was roughly 1.2×10^{11}. **g**, Heatmap of relative abundance of tether lengths showing convergence toward 6 and 8 nt for T1 and T2, respectively. **h**, Rosetta score versus r.m.s.d. for constrained and unconstrained models of an enriched sequence from the designed library.

performed Evolink on the Designed Junction Library followed by 3D-structure analysis, featuring 6 to 9 random nucleotides for both T1 and T2 regions with the addition of three synthetic base pairs at the Tether–H101 junction. This was to increase the independence of tether folding from junction folding (Fig. 4e). Evolink was carried out over four time points/days (Fig. 4f). Tether lengths converged to a length of 6 and 8 nt for T1 and T2, respectively, with the most commonly observed sequence being T1, 5′-GUUAUA-3′ and T2, 5′-AUCCCAGG-3′ (Fig. 4g). Post facto modeling of select highly enriched genotypes as described previously (Methods, Structural fragility of the Tether–H101 junction) revealed improved agreement between constrained and unconstrained models compared to the Broad Sampling Library (Fig. 4h and Extended Data Fig. 3e–h). Notably, modeling revealed predicted base pairing in the designed junction residues in both the constrained and unconstrained models, as well as predicted base pairing in the Tether–H101 junction compared to the Broad Sampling Library winner (compare Extended Data Fig. 4a,b with Fig. 4b,c). To further strengthen our confidence in library design, we measured the growth rates of different libraries, which contain mixed populations of ribosomes, to observe whether there are differences in cellular growth at the population level. This would help us discern whether the predicted gains in structural stability from our models might contribute to improved tethered ribosome function (Supplementary Fig. 3). We observed a substantial improvement in growth rates (doubling times) in the Designed Junction Library as a mixed population relative to the Broad Sampling Library. This supported our working hypothesis that structural stabilization of the Tether-H101 junction may be beneficial for tethered ribosome function.

Isolation of and orthogonal protein synthesis with Ribo-T v3. We carried out a final round of design and Evolink to identify candidates for clonal isolation and characterization of improved tethered ribosomes. The library combined the lessons learned from our three previous libraries. First, tether lengths ranging from 5 to 9 nt for T1 and 6 to 9 nt for T2 were tested based on the previous libraries converging to 6 and 8 nt for T1 and T2, respectively (Fig. 5a). Second, we kept the Designed Tether–H101 junction featuring base pairs that we hoped would contribute to improved structural stability in the tethers. Evolink was carried out to identify enriched genotypes encoding T1 and T2 (Extended Data Fig. 2c).

Next, we were interested in whether T1 and T2 sequences displayed cooperativity as we had previously observed enrichment of specific combinations between T1 and T2 sequences during evolution (Supplementary Fig. 1). To test hypotheses of cooperativity and to isolate a final winning genotype, we tested 16 individual genotypes combining the top four enriched sequences for the T1 and T2 regions from this library for their ability to carry out orthogonal superfolder green fluorescent protein (sfGFP) synthesis compared to a previously improved orthogonal tethered ribosome, oRibo-T v2 (Fig. 5b,c). Orthogonal translation output is a unique application for tethered ribosomes and an important measure of their function. In this experiment, the anti-Shine–Dalgarno of the tethered ribosome's small subunit is mutated to selectively translate o-mRNAs (encoding sfGFP) with a correspondingly mutated Shine–Dalgarno sequence. Of the 16 genotypes, 14 T1/T2 pairs outperformed oRibo-T v2 in orthogonal sfGFP synthesis (Fig. 5c), yield support to our search for improve tethered ribosomes. Further, we observed combinatorial effects among the 16 individual genotypes tested: as an extreme example, depending on the paired T1, the sequence

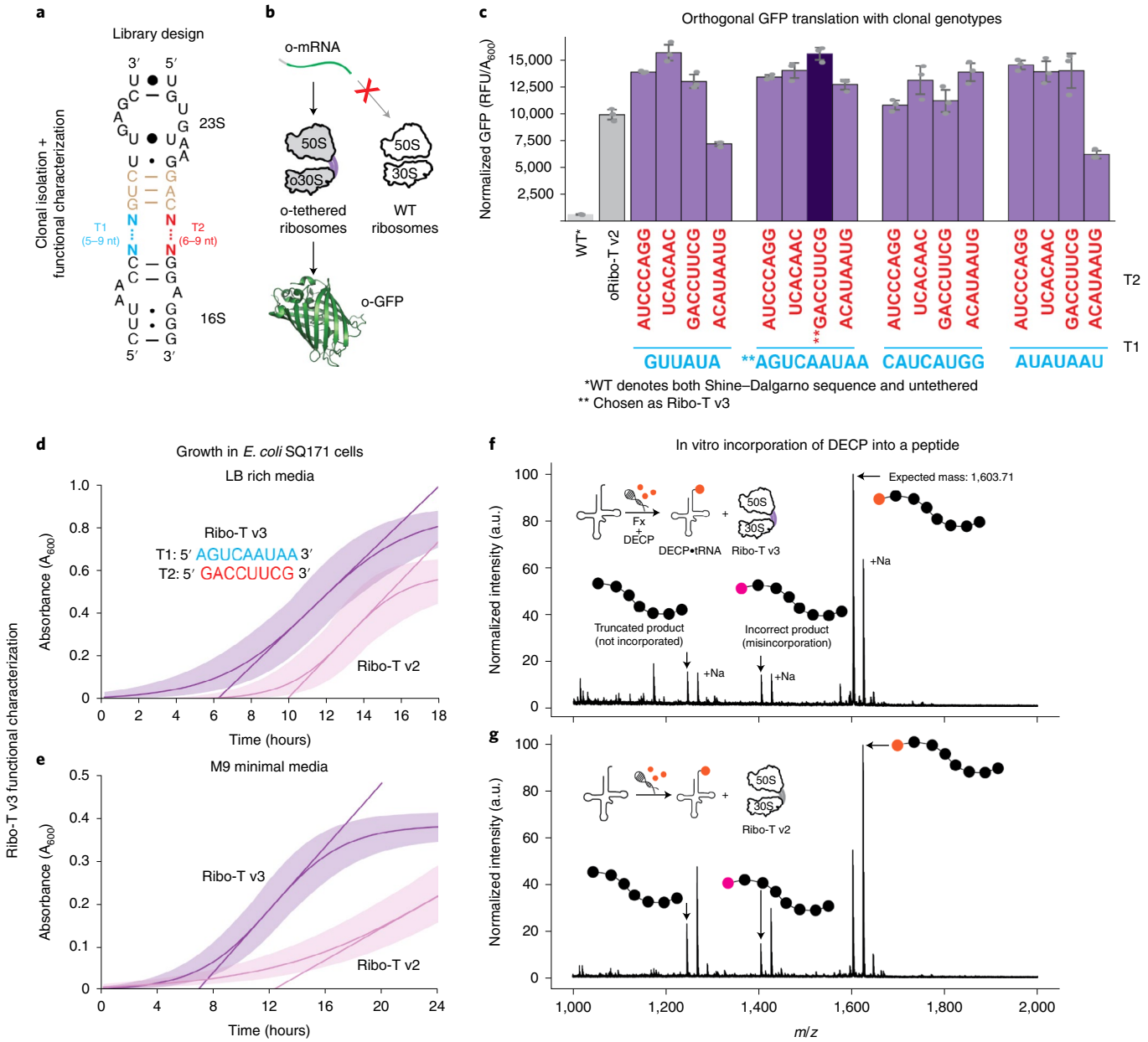


Fig. 5 | Clonal isolation and test of Ribo-T v3 function. a, The final library that combines the Designed Tether–23S junction and lengths informed by Evolink results from the Designed Tether–23S junction library. **b**, Cartoon schematic for orthogonal sfGFP synthesis. **c**, Orthogonal sfGFP synthesis by 16 candidate Ribo-T v3 tether pairs based on the four most popular T1 and T2 genotypes. Data are presented as mean values and are from three biological replicates ($n = 3$) and error bars indicate standard deviation, representative of three independent experiments. **d**,**e**, Growth of SQ171 cells living on Ribo-T v3 and Ribo-T v2 on rich lysogeny (or LB) media (**d**) and minimal M9 media (**e**). Data are presented as mean values and are from 20 biological replicates ($n = 20$) and error shown represents a 95% confidence interval on each estimated parameter in the sigmoid curve fit. Data representative of three independent experiments. **f**,**g**, Incorporation of DECP into a sequence-defined peptide by a purified sample of Ribo-T v3 (**f**) and Ribo-T v2 (**g**) in an in vitro protein synthesis reaction using flexizymes. Mass spectrometry data representative of three independent experiments.

T2: 5′-ACAUAAUG-3′ could perform 30% better than oRibo-T v2 or 30% worse (Fig. 5c), supporting the hypothesis that the tethers interact with functional consequences. The two highest performing genotypes were (1) T1, 5′-GUUAUA-3′ and T2, 5′-UCACAAG-3′ and (2) 5′-AGUCAAUAA-3′ and T2: 5′-GACCUUCG-3′, which each showed increased orthogonal protein synthesis over Ribo-T v2 by 56 and 58%, respectively (Fig. 5c). Of these, we carried out further characterization for T1, 5′-AGUCAAUAA-3′ and T2, 5′-GACCUUCG-3′, which we termed Ribo-T v3. The choice of this genotype was supported by enrichment trends observed during selection, which suggested that a 8-nt T2 was more broadly enriched compared to a 6-nt T1 (Extended Data Fig. 2c, right).

Characterization of Ribo-T v3. With Ribo-T v3, we sought to characterize its biochemical, functional, and structural properties. First, we tested the ability of Ribo-T v3 to support cellular life as a general measure of ribosome function[9,18]. We compared growth rates of SQ171 cells living on Ribo-T v3 and Ribo-T v2 in Luria Bertani (LB) medium (Fig. 5d) and minimal M9 media (Fig. 5e). This revealed improved growth for cells supported by Ribo-T v3

(Fig. 5d,e, Supplementary Table 3). Although doubling times in LB media were equal within error, cells growing on Ribo-T v3 exhibited a 97% improvement in doubling time in M9 media. Additionally, SQ171 cells living on Ribo-T v3 showed 59 and 77% improvement in lag time for LB and M9 media, respectively (Supplementary Table 3). This suggests that differences between Ribo-T v2 and Ribo-T v3 extend beyond ribosome function at the molecular scale, but also has implications at the phenotypic level when considering coordination with other cellular machinery during cellular growth.

Building on improved doubling times and reduced lag times for cells supported by Ribo-T v3 compared to Ribo-T v2 (Fig. 5d), we explored why the new linkage joining the subunits might be better. First, quantitative PCR (qPCR) of rRNA from cells expressing orthogonal Ribo-T ribosomes was performed to assess RNA abundance (Supplementary Table 4). On average, rRNA from oRibo-T v3 was 1.3 ± 0.27 times more abundant than rRNA from oRibo-T v2. Thus, we posit that improved orthogonal protein synthesis with oRibo-T v3 compared to oRibo-T v2 may be in part a result of improved stability or assembly (for example, by minimizing RNA misfolding during maturation of rRNA) into functional ribosomes. Second, we applied dimethyl sulfate (DMS) mutational profiling with sequencing (DMS-MaPseq)[37] to purified tethered ribosomes to assess potential differences in structure probed by reactivity to DMS. For both Ribo-T v2 and Ribo-T v3 polysomes, we observed high chemical reactivity in the tether regions and low reactivity in the helices that connect the designed tethers. This indicates that the tether regions are highly accessible to chemical modification and with the helices less so, consistent with our expectations (Extended Data Fig. 5). Taken together, these data suggest that Ribo-T v3 may achieve better performance through differences in ribosome biogenesis, although this can be explored more comprehensively in the future[38].

Toward the vision of genetic code expansion, we challenged Ribo-T v3 to incorporate a noncanonical amino acid into a peptide. The idea was not to engineer Ribo-T v3 to be better than a natural ribosome at incorporating noncanonical amino acids, but rather to show that oRibo-T was compatible with applications geared toward expanding the chemistry of life[1–4,14,24]. We chose a noncanonical L-α-amino acid ((*R*)-2-amino-3-(7-(diethylamino)-2-oxo-2H-chromene-3-carboxamido)propanoic acid, DECP) featuring a diethylamino coumarin group on its side chain (Fig. 5f and Supplementary Figs. 4 and 5). The monomer, with a bulky side chain, has not yet been shown to be incorporated into a peptide ribosomally, and thus presents an attractive target to showcase Ribo-T v3's ability to expand the chemical biology toolbox of engineered translation machinery. For demonstration purposes, we used a cell-free transcription and translation platform based on the PURExpress system[39–41]. The monomer DECP was charged onto $tRNA^{fMet}(CAU)$ using a flexizyme[23] (Supplementary Fig. 4c), and added to the PURExpress reaction with a purified sample of Ribo-T v3 or Ribo-T v2 (Fig. 5e). Mass spectrometry revealed that DECP was successfully incorporated into the N-terminus of a peptide by both Ribo-T v3 and Ribo-T v2 (Fig. 5f,g and Supplementary Fig. 5). We observed improved incorporation of DECP by Ribo-T v3 compared to Ribo-T v2 based on less prominent peaks of mis-incorporated or truncated products observed in matrix-assisted laser desorption-mass spectrometry. We note that the amount of misincorporation or truncation products produced by Ribo-T v3 is greater than that of the wild-type untethered ribosome (Supplementary Fig. 6). While future opportunities exist for engineering the tethered ribosome to specifically polymerize noncanonical amino acids with high fidelity, Ribo-T v3's improved ribosome function compared to Ribo-T v2 shows its immediate use for genetic code expansion.

Finally, we carried out single-particle cryogenic-electron microscopy (cryo-EM) on Ribo-T v3 polysomes to solve a medium resolution structure of Ribo-T v3 and prove that the tether exists. Raw micrographs of the polysomes confirm a characteristic 'beads on a string' behavior of translating ribosomes (Extended Data Fig. 6), and a ribosome that has dissociated from mRNA demonstrates an 'open clamshell' configuration that would be expected of a tethered ribosome (Extended Data Fig. 6). The final cryo-EM density map had a resolution of 4.18 Å and fit a previously published ribosome structure well, confirming that tethering the two subunits does not substantially change the 3D conformation of the ribosome (Fig. 6a). We were surprised to find that even for this medium-density map, we were able to resolve the tether region. Using DRRAFTER[42], we built the tether sequence into this density and confirmed that the tether sequence fits into this additional electron density in a structurally reasonable way (Fig. 6b). Qualitatively, the structure of the backbone generally agrees with previous simulations of the tether region, recognizing that predicted and solved structures are samples from a dynamic ensemble of possible structures. The resolution of the tether residues supported our working theory that structure-driven tether design is a sound approach for tethered ribosome engineering.

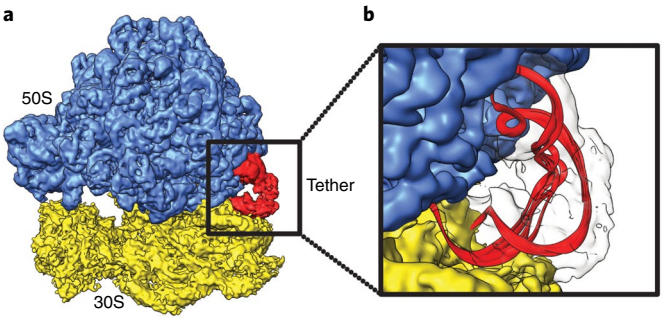


Fig. 6 | Cryo-EM structure of the Ribo-T v3 ribosome at 4.18-Å resolution. A 4.18-Å density of the Ribo-T v3 ribosome was generated through single-particle analysis. **a**, The 4YBB ribosome structure fit into the Ribo-T v3 density. **b**, Enlargement of the density of the Ribo-T v3 tether region with top ten DRRAFTER models of the tether built into the density.

Evolink extension to ribosomal helices. To demonstrate the generalizability of Evolink, we applied Evolink to explore the simultaneous mutability of two ribosomal helices, H71 (23S rRNA) and h44 (16S rRNA). The B3 subunit (residue A1483 of the 16S rRNA) of the ribosome has been implicated as critical in ribosomal subunit association, and was also revealed to be the target of antibiotics such as viomycin that have anti-tuberculosis activity[43–45] (Supplementary Fig. 7a,b). Evolink on H71 and h44 simultaneously resulted in a wide range of $\log_2$-fold enrichment between genotypes (Supplementary Fig. 7d) ranging from −7.95 to 3.47, including genotypes harboring mutations previously reported to be significantly detrimental to ribosome function[43,45]. We recover genotypes that mutate A1483 but are enriched in our selection, revealing potentially compensatory mutations that may rescue ribosome function. This proof-of-concept generalization of Evolink suggests that it may be used to reveal interesting interactions in macromolecular machines.

Discussion

In this work, we present an improved tethered ribosome platform, termed Ribo-T v3. Key to our effort was the development of Evolink, a technique for evolving regions in macromolecular machines far apart in primary sequence but proximal (and potentially functionally linked) in three-dimensional space. Evolink uses widely available molecular biology protocols (PCR and ligation) to link together distant sites of a plasmid in a single NGS read, alleviating previous limitations to ribosome evolution enforced by short NGS read

lengths (roughly 300 nt). We note that although Evolink makes much more tenable linking together distant sequences, it still faces limitations of NGS read length (for example, if all of the regions of interest and accessory sequences required for linking them do not fit within the current read length limits, regions of interest are in a particular difficult region for generating amplicons or when working outside the context of a plasmid). We carried out four iterations of our DBTA directed evolution experiment, featuring library designs informed by NGS results as well as structural modeling. Libraries explored simultaneous variation of tether sequence and length, as well as interaction between the tether and its junction with H101, culminating in design of a library that yielded Ribo-T v3.

Ribo-T v3 features new tether RNA sequences (T1, 5′-AGUCAAUAA-3′ and T2, 5′-GACCUUCG-3′) as well as designed base pairs at the Tether–H101 junction. Ribo-T v3 exhibits up to a 58% improvement in orthogonal sfGFP translation and a 97% improvement in growth rate as well as a 77% improvement in lag time in SQ171 cells growing in M9 minimal media. Cells supported by Ribo-T v3 in rich LB media exhibit comparable growth rates to those living on Ribo-T v2, but a 59% improvement in lag time. This is consistent with our evolution experiments that favor cells that emerge from stationary phase most quickly. Additionally, we showcase Ribo-T v3's potential for expanding the chemical toolbox of orthogonal translation systems through the incorporation of a new noncanonical amino acid featuring a bulky side chain (DECP) into a peptide using an in vitro transcription and translation reaction supplemented with synthetic transfer RNAs charged with flexizyme. Taken together, the improvements in Ribo-T v3 are important because they permit assay improvements (for example, faster growth for RNA extractions; higher density cultures for increased downstream RNA and ribosome characterization) and facilitate rapid prototyping of orthogonal translation systems.

Looking forward, we predict that Ribo-T v3 will accelerate new advances in orthogonal translation systems to expand the palette of genetically encoded chemistries[9,14,16,46]. Further, the demonstration of Evolink applied to H71 and h44 of the rRNA reveal that Evolink could be used to study more complex interactions beyond base pairing. Finally, we expect Evolink will advance directed evolution efforts, especially those for large macromolecular machines, for synthetic biology.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41589-022-01064-w.

Received: 22 January 2021; Accepted: 17 May 2022;
Published online: 14 July 2022

References

1. Dedkova, L. M., Fahmi, N. E., Golovine, S. Y. & Hecht, S. M. Enhanced D-amino acid incorporation into protein by modified ribosomes. *J. Am. Chem. Soc.* **125**, 6616–6617 (2003).
2. Dedkova, L. M., Fahmi, N. E., Golovine, S. Y. & Hecht, S. M. Construction of modified ribosomes for incorporation of D-amino acids into proteins. *Biochemistry* **45**, 15541–15551 (2006).
3. Dedkova, L. M. et al. β-Puromycin selection of modified ribosomes for in vitro incorporation of β-amino acids. *Biochemistry* **51**, 401–415 (2012).
4. Dedkova, L. M. & Hecht, S. M. Expanding the scope of protein synthesis using modified ribosomes. *J. Am. Chem. Soc.* **141**, 6430–6447 (2019).
5. Des Soye, B. J., Patel, J. R., Isaacs, F. J. & Jewett, M. C. Repurposing the translation apparatus for synthetic biology. *Curr. Opin. Chem. Biol.* **28**, 83–90 (2015).
6. Ellefson, J. W. et al. Synthetic evolutionary origin of a proofreading reverse transcriptase. *Science* **352**, 1590–1593 (2016).
7. Hammerling, M. J., Fritz, B. R., Yoesep, D. J., Carlson, E. D. & Jewett, M. C. In vitro ribosome synthesis and evolution through ribosome display. *Nat. Commun.* **11**, 1108 (2020).
8. Maini, R. et al. Protein synthesis with ribosomes selected for the incorporation of β-amino acids. *Biochemistry* **54**, 3694–3706 (2015).
9. Carlson, E. D. et al. Engineered ribosomes with tethered subunits for expanding biological function. *Nat. Commun.* **10**, 3920 (2019).
10. Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **10**, 866–876 (2009).
11. Sailer, Z. R. & Harms, M. J. Molecular ensembles make evolution unpredictable. *Proc. Natl Acad. Sci. USA* **114**, 11938–11943 (2017).
12. Sarkisyan, K. S. et al. Local fitness landscape of the green fluorescent protein. *Nature* **533**, 397–401 (2016).
13. Ramakrishnan, V. Ribosome structure and the mechanism of translation. *Cell* **108**, 557–572 (2002).
14. Schmied, W. H. et al. Controlling orthogonal ribosome subunit interactions enables evolution of new function. *Nature* **564**, 444–448 (2018).
15. Fried, S. D., Schmied, W. H., Uttamapinant, C. & Chin, J. W. Ribosome subunit stapling for orthogonal translation in *E. coli*. *Angew. Chem.* **127**, 12982–12985 (2015).
16. Liu, C. C., Jewett, M. C., Chin, J. W. & Voigt, C. A. Toward an orthogonal central dogma. *Nat. Chem. Biol.* **14**, 103–106 (2018).
17. Neumann, H., Wang, K., Davis, L., Garcia-Alai, M. & Chin, J. W. Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* **464**, 441–444 (2010).
18. Orelle, C. et al. Protein synthesis by ribosomes with tethered subunits. *Nature* **524**, 119–124 (2015).
19. Rackham, O. & Chin, J. W. A network of orthogonal ribosome· mRNA pairs. *Nat. Chem. Biol.* **1**, 159–166 (2005).
20. Wang, K., Neumann, H., Peak-Chew, S. Y. & Chin, J. W. Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat. Biotechnol.* **25**, 770–777 (2007).
21. Liu, Y., Kim, D. S. & Jewett, M. C. Repurposing ribosomes for synthetic biology. *Curr. Opin. Chem. Biol.* **40**, 87–94 (2017).
22. Liu, F., Bratulić, S., Costello, A., Miettinen, T. P. & Badran, A. H. Directed evolution of rRNA improves translation kinetics and recombinant protein yield. *Nat. Commun.* **12**, 5638 (2021).
23. Goto, Y., Katoh, T. & Suga, H. Flexizymes for genetic code reprogramming. *Nat. Protoc.* **6**, 779–790 (2011).
24. Melo Czekster, C., Robertson, W. E., Walker, A. S., Söll, D. & Schepartz, A. In vivo biosynthesis of a β-amino acid-containing protein. *J. Am. Chem. Soc.* **138**, 5194–5197 (2016).
25. Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53–60 (2017).
26. 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**, 678 (2013).
27. Hui, A. & de Boer, H. A. Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in *Escherichia coli*. *Proc. Natl Acad. Sci. USA* **84**, 4762–4766 (1987).
28. Rackham, O. & Chin, J. W. Cellular logic with orthogonal ribosomes. *J. Am. Chem. Soc.* **127**, 17584–17585 (2005).
29. Cho, N. et al. De novo assembly and next-generation sequencing to analyse full-length gene variants from codon-barcoded libraries. *Nat. Commun.* **6**, 8351 (2015).
30. Yoo, J. I., Daugherty, P. S. & O'Malley, M. A. Bridging non-overlapping reads illuminates high-order epistasis between distal protein sites in a GPCR. *Nat. Commun.* **11**, 690 (2020).
31. Borgström, E. et al. Phasing of single DNA molecules by massively parallel barcoding. *Nat. Commun.* **6**, 7173 (2015).
32. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
33. Asai, T., Zaporojets, 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).
34. Lorenz, R. et al. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* **6**, 26 (2011).
35. Watkins, A. M., Rangan, R. & Das, R. FARFAR2: improved de novo Rosetta prediction of complex global RNA folds. *Structure* **28**, 963–976 (2020).
36. Noeske, J. et al. High-resolution structure of the *Escherichia coli* ribosome. *Nat. Struct. Mol. Biol.* **22**, 336–341 (2015).
37. Zubradt, M. et al. DMS-MaPseq for genome-wide or targeted RNA structure probing in vivo. *Nat. Methods* **14**, 75–82 (2017).
38. Aleksashin, N. A. et al. Assembly and functionality of the ribosome with tethered subunits. *Nat. Commun.* **10**, 930 (2019).
39. Lee, J., Schwarz, K. J., Kim, D. S., Moore, J. S. & Jewett, M. C. Ribosome-mediated polymerization of long chain carbon and cyclic amino acids into peptides in vitro. *Nat. Commun.* **11**, 4304 (2020).

40. Lee, J. et al. Expanding the limits of the second genetic code with ribozymes. *Nat. Commun.* **10**, 5097 (2019).
41. Lee, J., Torres, R., Byrom, M., Ellington, A. D. & Jewett, M. C. Ribosomal incorporation of cyclic β-amino acids into peptides using in vitro translation. *Chem. Commun.* **56**, 5597–5600 (2020).
42. Kappel, K. et al. De novo computational RNA modeling into cryo-EM maps of large ribonucleoprotein complexes. *Nat. Methods* **15**, 947–954 (2018).
43. Sun, Q., Vila-Sanjurjo, A. & O'Connor, M. Mutations in the intersubunit bridge regions of 16S rRNA affect decoding and subunit–subunit interactions on the 70S ribosome. *Nucleic Acids Res.* **39**, 3321–3330 (2011).
44. Zhang, L. et al. The structural basis for inhibition of ribosomal translocation by viomycin. *Proc. Natl Acad. Sci. USA* **117**, 10271–10277 (2020).
45. Pulk, A., Maiväli, Ü. & Remme, J. Identification of nucleotides in *E. coli* 16S rRNA essential for ribosome subunit association. *RNA* **12**, 790–796 (2006).
46. Aleksashin, N. A. et al. A fully orthogonal system for protein synthesis in bacterial cells. *Nat. Commun.* **11**, 1858 (2020).

Methods

Library construction. Plasmid libraries of Ribo-T tethers were generated using PCR with the plasmid encoding Ribo-T v2 (ref. [9]), as the template. Oligonucleotides (IDT) encoding degenerate bases (Ns) in place of the tethers were used to amplify the insert that includes both tethers and the 23S rRNA (referred to as the insert) (Fig. 1c). For the Tether–23S junction, oligos encoded deletions in the specified region, not degenerate bases (Fig. 2e). Another pair of oligos amplified the remainder of the plasmid (referred to as the backbone) (Supplementary Table 1). Resulting amplicons were purified using the Omega Cycle-Pure kit (Omega BioTek), then digested with DpnI (NEB) to remove the template. The insert and backbone were ligated using isothermal DNA assembly[32], and transformed into POP2136 cells via electroporation. Posttransformation, the cells were recovered in 800 µl of SOC growth medium at 30 °C for 90–120 min, then plated on LB-agar plates containing 100 µg ml^{-1} carbenicillin. The plates were incubated at 30 °C for 16–18 h until colonies appeared. All colonies were scraped from the agar plates and plasmid extraction was performed using a Zymo-PURE Midiprep II kit (Zymo Research).

Selection of tethered ribosomes. The libraries of Ribo-T tethers were transformed into SQ171 cells lacking chromosomal ribosomes[33]. Next, 100 ng of the plasmid library was transformed into 50 µl of SQ171 cells via electroporation, then recovered with 500 µl of SOC at 37 °C with shaking at 250 r.p.m. for 2 h. After, another 1.5 ml of SOC was added to the cells and the final 2 ml culture was brought to 100 µg ml^{-1} carbenicillin and 0.25% sucrose. These cells were then incubated at 37 °C with shaking at 250 r.p.m. for 16–18 h. After incubation, cells were plated onto LB-agar plates containing: carbenicillin (100 µg ml^{-1}), sucrose (5% w/v) and erythromycin (250 µg ml^{-1}) and incubated at 37 °C for 20–24 h until colonies appeared. Colonies were then washed from the agar plates with LB containing 100 µg ml^{-1} carbenicillin (roughly 5 ml of LB-carbenicillin per 100 mm petri dish) and grown to saturation at 37 °C with 250 r.p.m. shaking. Then 1 ml of the solution was reserved and plasmids were extracted using the Zymo-PURE Miniprep kit (Zymo Research). The saturated culture was then subject to passaging over 4 days in LB containing 100 µg ml^{-1} carbenicillin, and plasmids were extracted each day for sequencing.

Preparation of amplicons for NGS. Plasmids extracted from selection cultures were linearized using PCR, and purified using the Omega Cycle-Pure kit. Then 20 ng of the purified product was used in a 20 µl ligation reaction containing T4 ligase (NEB) and the appropriate accompanying buffer. After incubation at 37 °C for 2 h, 2 µl of the ligation reaction was used directly in a 20 µl PCR with 15 cycles of amplification, which generated the amplicon for NGS. The resulting product was then purified and prepared for NGS using the NEBNext Ultra II DNA Library Prep kit (NEB). The resulting library was run on a MiSeq (Illumina) using a 150-cycle MiSeq Reagent Kit v3 (Illumina).

Analysis of NGS results. Paired-end reads from Illumina sequencing were assembled using PANDASeq[47]. Reads that had coverage (number of redundant reads) of less than ten were filtered and excluded from analysis. Pairs of sequences were then identified, and the following parameters were calculated.

Abundance was calculated using the following formula:

$$\text{Abundance}_{i,\,n} = \frac{\text{reads}_{i,\,n}}{\sum_i^S \text{reads}_{i,n}}$$

for a specific genotype i at timepoint n, and S represents the total number of unique genotypes at timepoint n after filtering as described above.

Fold enrichment was calculated using the following formula:

$$\text{Enrichment}_{i,n} = \log_2 \frac{\text{abundance}_{i,n}}{\text{abundance}i,_0}$$

for a specific genotype i at timepoint n, and abundance$_0$ represents the abundance after selection on agar plates as previously described before any liquid culture.

Post facto computational modeling of tether. For 3D modeling studies, we set up FARFAR2 simulations[35] using a crystal structure of the *E. coli* ribosome[36] (Protein Data Bank (PDB) code 4YBB). Starting from that structure, we truncated the stem-loops 23S rRNA Helix 101 (H101) and 16S rRNA helix 44 (h44), removing the residues that are deleted in all tethered ribosome constructs, and renumbered those residues to facilitate building a continuous RNA chain.

Using that initial structure as a template, we built the remaining residues of the tether using the FARFAR2 algorithm, conducted on 200 CPUs for 24 h, generating several thousand structures. We conducted simulations under two conditions: in one, only tether residues were resampled; in another, a junction on the 23S side of the tether was resampled as well.

All inputs and command files used in setting up computational modeling are available at https://github.com/everyday847/ribotv3_simulations.

Measurement of orthogonal GFP production. Combinations of potentially high-performant tether designs were identified from NGS results and built into a plasmid containing both an orthogonal tethered ribosome gene (oRibo-T) and an orthogonal superfolder GFP (o-sfGFP) coding sequence (mutated Shine–Dalgarno sequence). Then 10 ng of sequence-confirmed plasmids were transformed into 25 µl of BL21(DE3) cells via electroporation, recovered in 1 ml of SOC, and plated on agar plates containing 100 µg ml^{-1} of carbenicillin. Individual colonies were picked ($n = 3$) for inoculation of 100 µl of LB media containing 100 µg ml^{-1} carbenicillin. Cultures were incubated at 37 °C for 14–16 h with 2 mm continuous linear shaking in a plate reader (Agilent BioTek Synergy H1) and absorbance at 600 nm (OD_{600}) was monitored to ensure saturation. After cultures reached saturation, each culture was diluted to roughly 0.01 OD_{600} in fresh LB media containing 100 µg ml^{-1} of carbenicillin and 1 mM isopropyl β-D-1-thiogalactopyranoside to induce transcription of the orthogonal GFP gene. Cultured were incubated at 37 °C for 14–16 h with 2 mm continuous linear shaking in a plate reader and OD_{600} was monitored along with fluorescence (485/528 nm excitation/emission). Orthogonal GFP production (fluorescence) was normalized by OD_{600}. Fluorescence and turbidity measurements were made using the BioTek Synergy H1 Hybrid Reader and BioTek Gen 5 v.2.09 software.

Measurement of relative rRNA amounts by qPCR. Ribosomes from BL21(DE3) cells harboring oRibo-T v2 and oRibo-T v3 were harnessed through a sucrose cushion as described previously[39]. Primers probing regions specific to tethered ribosomes (Supplementary Table 1) were used to perform qPCR, using the iTaq Universal SYBR Green Supermix Kit (Bio-Rad), in 20 µl reactions according to the manufacturer's specifications. Bio-Rad CFX96 machines were used for recording data. The resulting qPCR data were analyzed using CFX Maestro software (Bio-Rad) and Ct values were exported for future analysis.

Growth rate characterization of Ribo-T v3. A plasmid encoding tether sequences corresponding to Ribo-T v3 (named pRTv3), was constructed using Gibson assembly[32]. Next, 10 ng of pRTv3 was transformed into 50 µl of SQ171fg cells[9,18] via electroporation and recovered in 500 µl of SOC at 37 °C for 2 h with shaking at 250 r.p.m.. After recovery, 1.5 ml of SOC was added and supplemented with 100 µg ml^{-1} carbenicillin and 0.25% (w/v) sucrose (final concentrations). After overnight (16–18 h) recovery at 37 °C with 250 r.p.m. shaking, the cells were spun down (4,000*g*, 10 min) and plated on LB-agar plates containing 100 mg m^{-1}: carbenicillin, 5% sucrose and 250 µg ml^{-1} erythromycin. Individual colonies were picked, and resistance to 100 µg ml^{-1} carbenicillin and sensitivity to 50 µg ml^{-1} kanamycin was checked on LB-agar plates to confirm successful swapping of ribosome plasmids in the SQ171fg cells. Colonies that successfully replaced pCSacB[33] with pRTv3 were carried through for analysis.

In a 96-well plate, 100 µl of LB media containing 100 µg ml^{-1} carbenicillin, 5% sucrose and 250 µg ml^{-1} erythromycin was inoculated with a colony from an LB-agar plate containing 100 µg ml^{-1} carbenicillin, 5% sucrose and 250 µg ml^{-1} erythromycin and incubated for 14–16 h at 37 °C with 2 mm lateral shaking in a plate reader (Agilent BioTek Synergy H1). Absorbance at 600 nm was monitored to ensure cultures reached saturation. After incubation, cultures were diluted to an A_{600} of roughly 0.05 (roughly 20-fold) in 100 µl of LB media containing 100 µg ml^{-1} carbenicillin, 5% sucrose and 250 µg ml^{-1} erythromycin and incubated for 18 h at 37 °C with 2 mm of lateral shaking, and absorbance at 600 nm (A_{600}) was monitored. Turbidity measurements were made using the BioTek Synergy H1 Hybrid Reader and BioTek Gen 5 v.2.09 software.

Preparation of DECP-CME. The graphical representation of synthesis is shown in Supplementary Fig. 8. Cyanomethyl-2-amino-3-(7-(diethylamino)-2-oxo-2H-chromene-3-carboxamido) propanoate (DECP-CME, 5) was prepared with three steps using the synthetic methods previously described[36,41]. First, 268 mg (1 mmol) of 7-(diethylamino)-2-oxo-2*H*-chromene-3-carboxylic acid (1) and 162 mg (1 mmol) of carbonyldiimidazole were added to a flask and sealed with a septum. 5 ml of anhydrous *N*,*N*-dimethylformamide was added into the flask using an oven-dried syringe and stirred at room temperature for 2 h. Then 204 mg (1 mmol) of (*R*)-3-amino-2-((*tert*-butoxycarbonyl)amino)propanoic acid (2) was added and stirred overnight. The product was extracted with ethyl acetate after washing the crude reaction mixture with 1 M HCl, water and brine. Second, 38 ml (0.6 mmol) of chloroacetonitrile and 104 ml (0.75 mmol) of triethylamine were added to 223 mg (0.5 mmol) of the purified 2-((*tert*-butoxycarbonyl)amino)-3-(7-(diethylamino)-2-oxo-2*H*-chromene-3-carboxamido) propanoic acid (3) in 1 ml of DCM and stirred overnight. The organic layer was washed with 1 M HCl, water and brine and dried over $MgSO_4$. (3) 1 ml of 50% of trifluoroacetic acid solution in DCM was added to the purified cyanomethyl-2-((*tert*-butoxycarbonyl) amino)-3-(7-(diethylamino)-2-oxo-2*H*-chromene-3-carboxamido)propanoate (4) to deprotect the Boc group. The final product was dried under high vacuum and obtained as pale yellow powder (yield 57%). Confirmation of synthesis was carried out via nuclear magnetic resonance, and the resulting data analyzed using both MNOVA (MestraNova) and Bruker Topspin software packages.

Purification of RTv3 for in vitro translation reactions and biochemical analysis. In brief, SQ171fg cells harboring pRTv3 as the sole source of ribosomes were grown to mid-exponential phase ($A_{600} = 0.3$–0.8) in 500 ml of LB media containing 100 µg ml^{-1} carbenicillin and 250 µg ml^{-1} erythromycin. Cells were spun down,

lysed using homogenization and ribosomes were harvested using a sucrose cushion as described previously[39]. Ribosome pellets were resuspended in Buffer C (10 mM pH 7.5 Tris Acetate, 60 mM ammonium chloride, 7.5 mM magnesium acetate, 0.5 mM ethylenediaminetetraacetic acid and 2 mM dithiothreitol) and brought to a concentration of 15 mM (A_{260} = 625). Resuspended ribosomes were used directly in in vitro translation reactions.

In vitro translation reactions for incorporation of DECP by RTv3. Preparation of DNA templates for RNAs. The DNA templates for flexizymes and tRNA preparation were synthesized as previously described[23,39–41]. Sequences of the final DNA templated used for in vitro transcription by the T7 polymerase are available in Supplementary Table 5.

Preparation of Fxs and tRNAs. Flexizymes (Fxs) and tRNAs were prepared using an in vitro transcription kit (HiScribe T7 High yield RNA synthesis kit, NEB E2040S) and purified by the previously reported methods[39–41].

Charging DECP into tRNA by Fx. The acylation experiment was performed first using flexizyme with three flexizymes (e, d and aFx). The Fx reaction was carried out as follows: 1 µl of 0.5 M HEPES (pH 7.5) or bicine (pH 8.8), 1 µl of 10 µM microhelix (mihx, tRNA mimic) and 3 µl of nuclease-free water were mixed in a PCR tube with 1 µl of 10 µM eFx, dFx and aFx, respectively. The mixture was heated for 2 min at 95 °C and cooled down to room temperature over 5 min. Next, 2 µl of 0.3 M $MgCl_2$ in water was added to the mixture and incubated for 5 min at room temperature. Followed by the incubation of the reaction mixture on ice for 2 min, 2 µl of 25 mM DECP-CME in dimethylsulfoxide was then added to the reaction mixture. The reaction mixture was incubated for 16 h on ice in a cold room. The optimal acylation reaction was determined by measuring the acylation yield using an acidic polyacrylamide gel (pH 5.2). $tRNA^{fMet}$(AUG) was charged with DECP under the condition obtained from the mihx acylation experiment. The charged tRNA was precipitated using ethanol and used for in vitro translation without further purification.

In vitro protein translation reaction. The noncanonical substrate incorporation experiment was performed using the PURExpress (Δribosome, Δaa, ΔtRNA, E3315Z) system. DECP-charged $tRNA^{fMet}$(CAU) was dissolved in 1 µl of 1 mM NaOAc (pH 5.2) and added into 9 µl of solution mixture containing 2 µl of Solution A, 1.2 µl of Factor mix, 1.8 µl of Ribo-T v3 (2.4 µM in final reaction), 1 µl of endogenous tRNA mixture, 1 µl of DNA plasmid (130 ng µl^{-1}), 1 µl of nuclease-free water and 1 µl of 5 mM amino acid mixtures (Trp, Ser, His, Pro, Gln, Phe, Glu, Lys and Thr). The reaction mixture was incubated in 37 °C for 2 h.

The target peptide produced in the PURE reaction was purified by using MagStrep (type 3) XT beads in 5% suspension (IBA Lifesciences), which selectively pull down the target peptide bearing the Strep tag (WSHPQFEK) at the C-terminal region. After pulling down the target peptide, the magnetic beads were washed with the Strep-Tactin XT wash buffer (IBA Lifesciences) and treated with 0.1% SDS solution in water. The beads were heated at 95 °C in a PCR machine to denature the target peptide bound to the beads. The magnetic beads were removed on a magnet rack and the obtained peptide was analyzed by mass spectrometry and resulting data analyzed using Bruker Flexanalysis software.

Chemical mapping of tethers by DMS-MaPseq. Targeted chemical mapping by DMS was carried out for the tether regions of both Ribo-T v2 and Ribo-T v3 ribosomes by a method adapted from DMS-MaPseq. Polysomes were isolated via sucrose cushion and quantified by nanodrop. Polysomes were diluted to 1 µM in Buffer G and treated with DMS (or water for an untreated negative control) at a final concentration of 1.5% for 7.5 min. The DMS modification reaction was quenched by addition of 2 reaction volumes of beta-mercaptoethanol and RNA was extracted by Direct-Zol RNA Miniprep Plus kit (Zymo Research). Then 250 ng of extracted RNA was mixed with 10 pmol reverse transcription primer (either oVT629 or oVT631 for tether 1 or tether 2, respectively) and heated to 65 °C for 5 min, placed on ice for 2 min and then mixed with 4 µl of 5× TGIRT FS buffer (250 mM Tris-HCl pH 8.3, 375 mM KCl, 15 mM $MgCl_2$), 2 µl of 10 mM dNTPs, 1 µl of water, 0.5 µl of TGIRT-III enzyme. Reverse transcription was carried out at 57 °C for 3 h and quenched by addition of 5 µl of 1 N NaOH and 5 µl of 0.5 M EDTA and incubating at 65 °C for 15 min to denature reverse transcriptase and hydrolyze RNA. The complementary DNA was purified via Zymo Oligo Clean and Concentrator kit (Zymo Research) and eluted in 12.5 µl of water. Next, 2 µl of cDNA was amplified in a 20 µl PCR using the NEBNext Ultra II Q5 Master Mix (NEB) with 0.5 µM each of both forward and reverse primers (oVT628/629 for T1, oVT630/631 for T2) and amplified for 20 cycles via two-step PCR (98 °C for 30 s, 20 cycles of 98 °C for 10 s followed by 72 °C for 60 s, 72 °C for 5 min). Reactions were purified using SPRISelect beads (Beckman Coulter), eluted in 25 µl of water and confirmed for appropriate tether amplification by 2% agarose gel stained with SYBRSafe (Invitrogen). Sample-specific index sequences were added to 20 µl of the eluted amplicons via 50 µl of indexing PCR reactions using the NEBNext Ultra II Q5 Master Mix and 0.5 µM each of iTru 5L and iTru 7L indexing primers for five cycles with the following program: 98 °C for 30 s, five cycles of (98 °C for 10 s, 65 °C for 30 s, 72 °C for 30 s), final extension at 72 °C for 5 min. Amplicons were purified by SPRISelect beads, quantified by qPCR and Qubit DNA HS fluorometry (Invitrogen) and sequenced via paired-end sequencing on an Illumina MiSeq using the MiSeq v2 300 cycle kit (200 cycles × 100 cycles).

Chemical mapping data analysis. Analysis of chemical mapping experiments was done using the RNAFramework package[48]. Demultiplexed sample-specific sequencing read fastq files were downloaded from Illumina Basespace. The Read1 fastq file was aligned against the tether 1 or tether 2 sequence using the rf-map command with '-b2 -cqo' flags. Mutations in the alignments were counted using the 'rf-count' command with the '-m' flag and normalized using the 'rf-norm' command using '-sm 4 -nm 2 -rb AC' flags. Normalized reactivity values were superimposed on the predicted RNA structure[18] using VARNA.

Cryo-EM microscopy. Ribo-T-v3 polysomes isolated via sucrose cushion were diluted to 1 µM in Buffer G. Polysomes were blotted onto glow-discharged Quantafoil R 2/1 mesh copper grids, incubated for 1 min at 4 °C at 100% humidity, then flash frozen in liquid ethane using a Vitrobot Mark IV (Thermo Fisher Scientific). Grids were imaged on a 200 kV Talos Arctica electron microscope (Thermo Fisher Scientific) with a pixel size of 0.86 Å per pixel. Then 1,312 40-frame videos were taken with a per-frame duration of 0.0337 s per frame and a per-frame dose of 1.25 e^-/A^2. Further information is available in Supplementary Table 7.

Analysis of cryo-EM data. All steps of single-particle analysis including motion correction, CTF estimation, particle extraction, two-dimensional (2D) classification, ab initio reconstruction and 3D refinement were carried out using CryoSPARC Live[49]. All micrographs were imported and roughly 1,000 particles were manually picked to generate preliminary 2D classifications to guide automated picking. Initially, 102,279 particles were extracted. Then 2D classes were used to subset the extracted particles to preserve only the ones containing ribosome-like features. Two more rounds of 2D classification were run to further remove poor particles. A final count of 34,852 particles were then used for ab initio reconstruction and 3D refinement. A final gold-standard Fourier shell correction map with a resolution of 4.18 Å was achieved.

The refined density map was imported into Chimera, z-flipped and Gaussian blurred with a blur factor of 1.5. The 4YBB ribosome structure was first fit into the blurred density, then the fit was refined into the un-blurred density, using the Chimera 'Fit to Map' command. The fitted density was segmented using the Segger tool[50] using default parameters and segments corresponding to the tether region were selected and extracted for 3D modeling.

The tether sequence was built into a segmented density map using standard DRRAFTER commands in Rosetta v.3.13. In total, 10,000 models were generated and the top ten models visually consistent with the contoured full density map were selected.

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

Data availability

The authors declare that all experimental data supporting the findings of this study are available within the paper and its supplementary files. Publicly available data, such as the 4YBB (PDB) ribosome structure are mentioned explicitly when used. All data related to models are available upon request from the authors. The map and fitted model for the cryo-EM data are reported as Electron Microscopy Data Bank entry no. EMD-26666 and PDB structure 7UPH. Source data are provided with this paper.

Code availability

All inputs and command files used in setting up computational modeling are available at https://github.com/everyday847/ribotv3_simulations.

References

47. Masella, A. P., Bartram, A. K., Truszkowski, J. M., Brown, D. G. & Neufeld, J. D. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinf.* **13**, 31 (2012).
48. Incarnato, D., Morandi, E., Simon, L. M. & Oliviero, S. RNA Framework: an all-in-one toolkit for the analysis of RNA structures and post-transcriptional modifications. *Nucleic Acids Res.* **46**, e97–e97 (2018).
49. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
50. Pintilie, G. D., Zhang, J., Goddard, T. D., Chiu, W. & Gossard, D. C. Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *J. Struct. Biol.* **170**, 427–438 (2010).

Acknowledgements

This work was supported by the National Science Foundation (grant no. MCB-1716766), the Human Frontiers Science Program (grant no. RGP0015/2017), the Army Research Office (grant no. W911NF-16-1-0372), all to M.C.J. R.D. thanks the NIGMS MIRA R35 award for funding. We thank J. Lucks and M. Evans at Northwestern and R. Kretsch

at Stanford for discussions. Some of this work was performed at the Stanford-SLAC Cryo-EM Center (S2C2), which is supported by the National Institutes of Health Common Fund Transformative High-Resolution Cryo-Electron Microscopy program (grant no. U24 GM129541). The US 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 US Government or the National Institutes of Health.

Author contributions

D.S.K., M.C.J., A.W. and R.D. conceived the study and designed experiments. D.S.K., E.B., E.R. and C.K. worked on establishing the Evolink method. D.S.K. and E.B. carried out experiments for tethered ribosome evolution, orthogonal GFP expression and cellular growth rates. J.L. and K.J.S. carried out experiments for noncanonical amino acid incorporation. D.S.K. and A.W. analyzed results and designed the libraries. A.W. carried out computational modeling. V.T. performed chemical mapping experiments. D.S.K., V.T. and Y.L. performed cryo-EM sample preparation and data collection. V.T. and G.P. performed cryo-EM data analysis and structure determination. D.S.K., A.W., M.C.J. and R.D. wrote the manuscript with participation by all authors.

Competing interests

M.C.J. and D.S.K. are coinventors on the US provisional patent application that incorporates discoveries described in this manuscript. M.C.J. has a financial interest in Pearl Bio, and his interests are reviewed and managed by Northwestern University in accordance with their competing interest policies. All other authors declare no competing interests.

Additional information

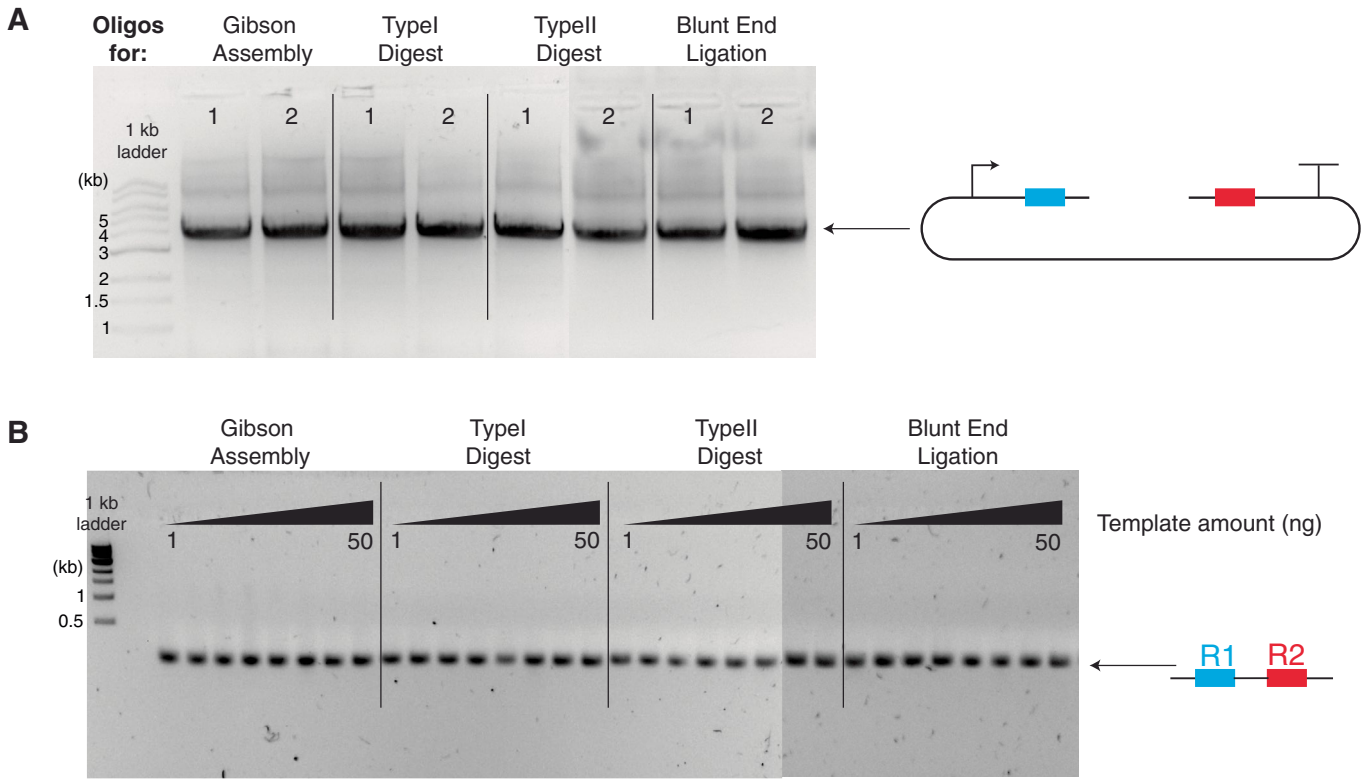
Extended data are available for this paper at https://doi.org/10.1038/s41589-022-01064-w.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41589-022-01064-w.

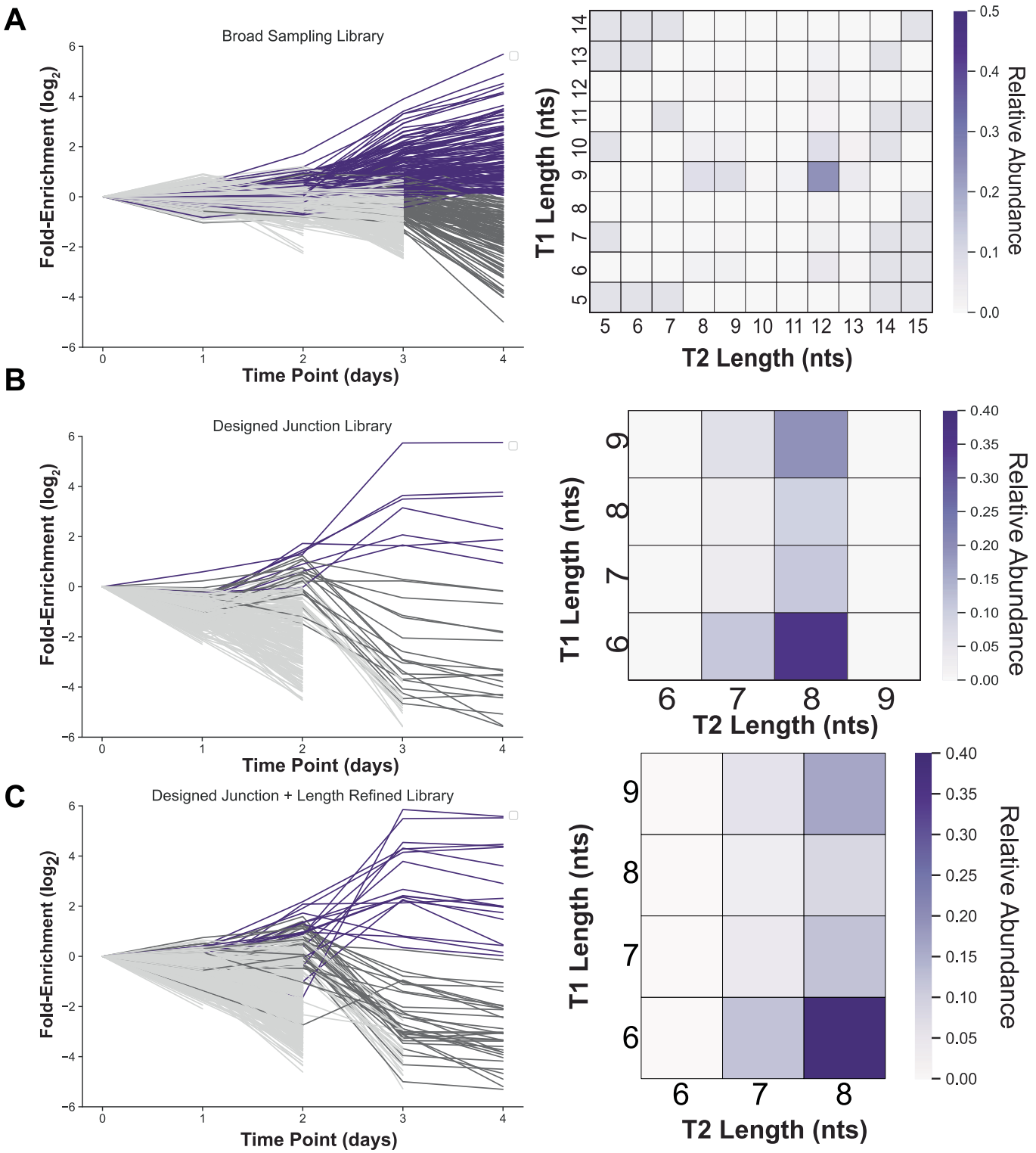
Correspondence and requests for materials should be addressed to Michael C. Jewett.

Peer review information *Nature Chemical Biology* thanks Luc Jaeger, Jérôme Waldispühl and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

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



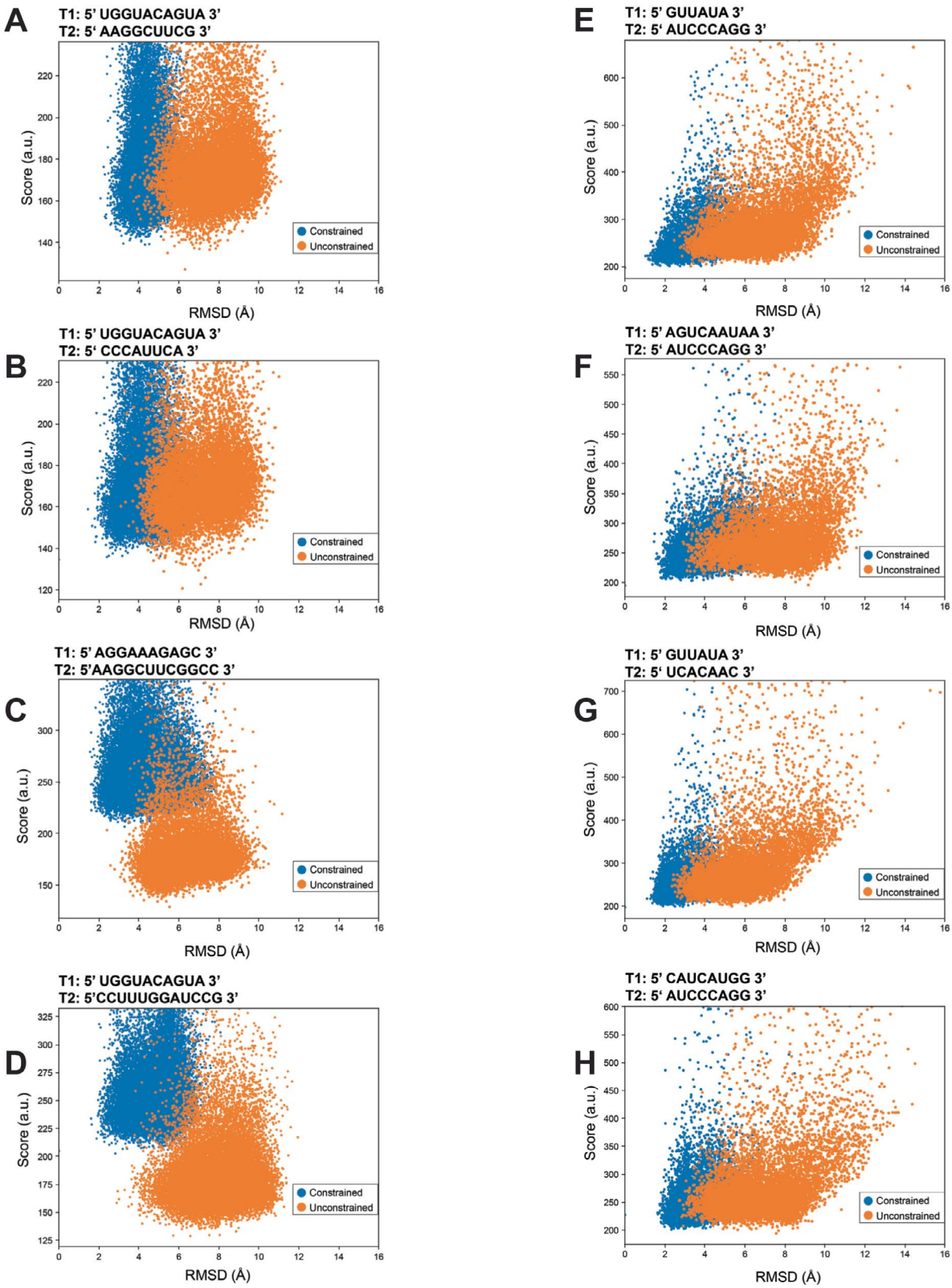
Extended Data Fig. 1 | Optimization of molecular biology steps involved in library preparation workflow of Evolink. a, A clonal sample of the tethered ribosome (Ribo-T v2) is linearized using different oligos compatible with multiple ligation protocols. **b**, From the different ligation products, generation of final amplicon for next-generation sequencing can happen with a wide range of ligation methods and starting template amounts in the PCR. Gel data representative of two independent experiments.



Extended Data Fig. 2 | Enrichment of individual genotypes throughout full Evolink experiment. a–c, Positively enriched genotypes (purple) and negative enriched genotypes (dark gray) can be tracked throughout multiple time points throughout selection. Genotypes that drop out during selection can also be identified (light gray). Corresponding heat maps that reveal trends in selected tether lengths also helped inform designs. Generally, across the three libraries tested in this work, **(a)** the Broad Sampling Library, **(b)** the Designed Junction Library, and **(c)** the Designed Junction + Length Refined Library, log_2-fold enrichment values between -6 to 6 are observed. Enrichment and heatmap data representative of three independent experiments.

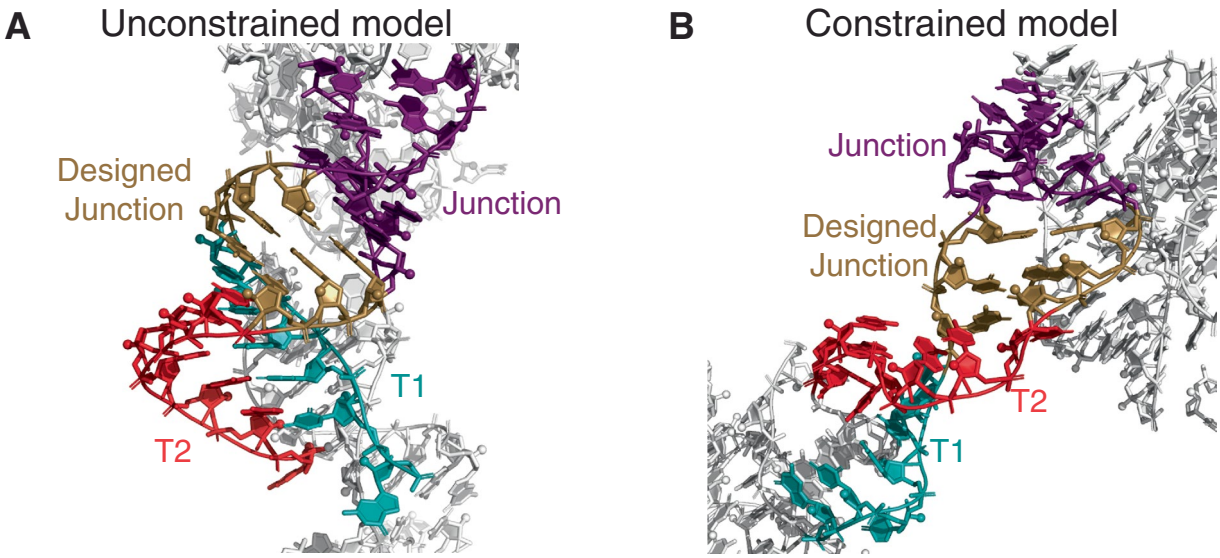
Broad Sampling Library Winners

Designed Junction Library Winners

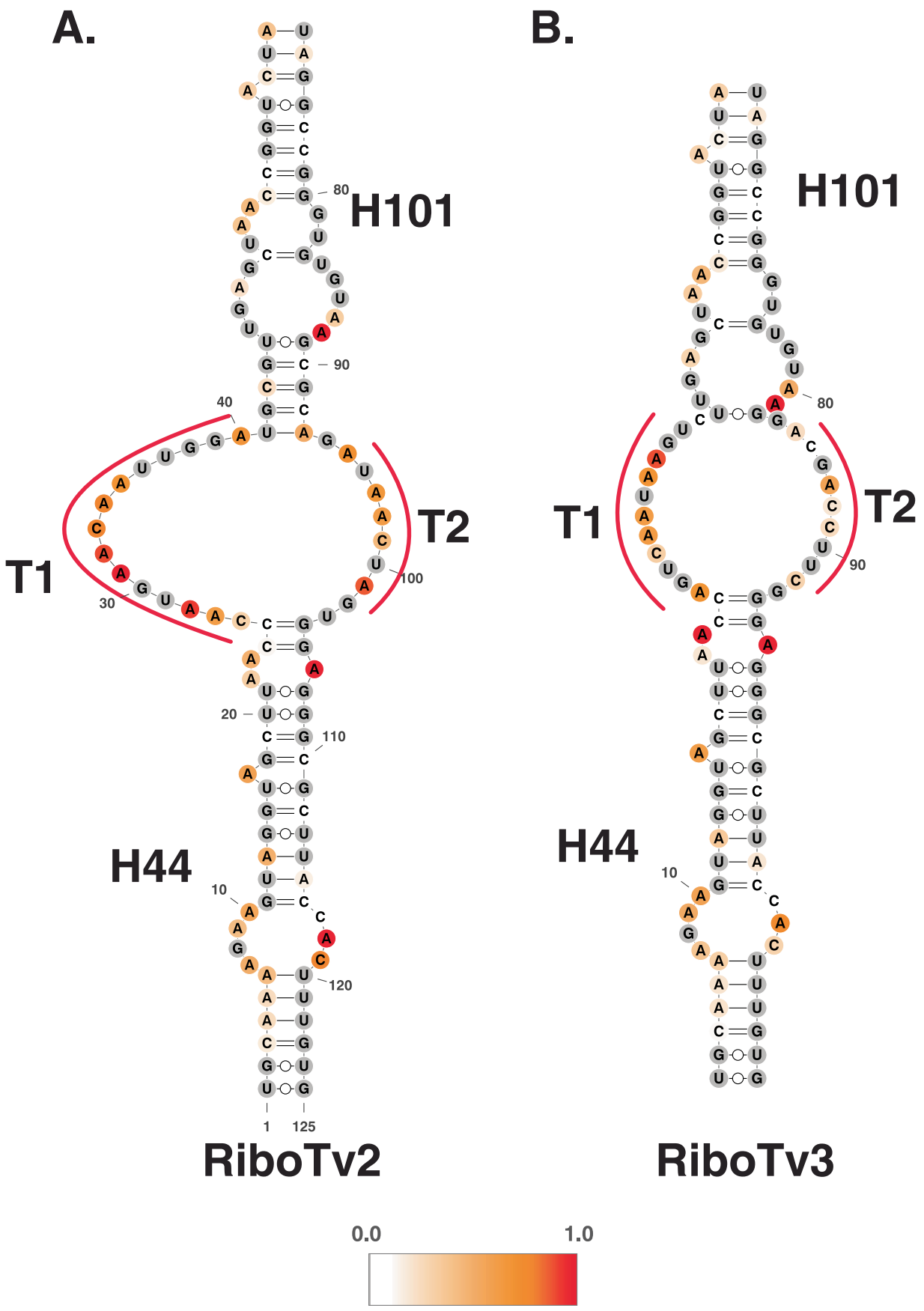


Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Score vs. Root-Mean-Standard-Deviation analysis of FARFAR2 simulations of enriched tether sequences. a–**d**, For the Broad Sampling Library, we observe striking differences between simulations that constrained (blue) or did not constrain (orange) 3D structures of the Tether-H101 junction. Of the four modeled genotypes, two sequence (**c**,**d**) exhibit particularly substantial differences, hinting at structural instability in the Tether-H101 junction. **e**–**h**, When similar simulations are performed with enriched tether sequences from the Designed Junction Library (designed sequences at the Tether-H101 junction), the results of FARFAR2 simulations reach similarly low scores in constrained vs. unconstrained modeling runs.

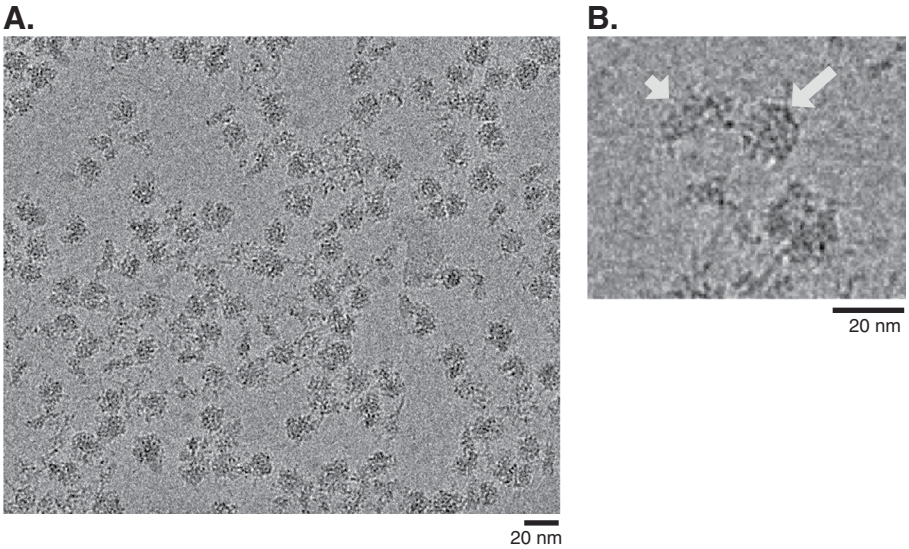


Extended Data Fig. 4 | Representative constrained and unconstrained 3D models of Designed Junction Library winner. The winning genotype from Fig. 4h was modeled using Rosetta, and representative outputs are shown. In both the **(a)** unconstrained and **(b)** constrained model, the Designed Junction residues are predicted to base pair, reinforcing structural stability to this region.



Extended Data Fig. 5 | See next page for caption.

Extended Data Fig. 5 | Chemical reactivity of tethers to DMS in whole polysomes. Targeted structure probing was performed on the tethers of both RiboTv2 and RiboTv3 polysomes via DMS-MaPseq. The per-nucleotide chemical reactivities of the tethers and their adjacent rRNA stems can be seen in the figure for both RiboTv2 (**a)** and RiboTv3 **(b)**. Gray shaded nucleotides represent U and G residues that are not modified by DMS.



Extended Data Fig. 6 | Raw Cryo-EM micrographs of Ribo-Tv3 polysomes. Single-particle Cryo-EM was carried out on RiboTv3 polysomes. **a**, A representative raw micrograph shows that RiboTv3 polysomes look like characteristic 'beads on a string' as expected for actively translating ribosomes. **b**, A tethered ribosome that has dissociated from an mRNA looks like an open clamshell, as would be expected for a tethered ribosome. The large and small subunits are indicated by white arrows.

nature research

Corresponding author(s): Michael C Jewett

Last updated by author(s): Dec 16, 2021

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	All data were collected using stated instruments and associated commercially available or open-source software. Commercial software used includes: Gen5 v. 2.09 (BioTek) for fluorescence measurements, Image Lab 6.1 (Bio-Rad) for gel images, Bruker TopSpin for acquiring NMR data, MNOVA (MestraNova) for acquiring NMR data, Bruker FlexControl and Bruker FlexAnalysis (Bruker) for acquiring MALDI-TOF MS data.
Data analysis	All data were analyzed using commercially-available softwares or in-house data analysis scripts written in python s as described in the manuscript. Commercial software for data analysis include Bruker TopSpin for analyzingNMR data, MNOVA (MestraNova) for analyzing NMR data, and Bruker FlexAnalysis (Bruker) for analyzing MALDI-TOF MS data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All data generated or analyzed during this study are included in the manuscript or supplementary information or are either included in the supplemental information or available from the corresponding author upon request. We report no restrictions on data availability.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	All experiments performed included at least three independent experiments. For in-cell based assays, sample sizes ranging from n=6 to n=18 were employed depending on material availability. For in-solution studies, sample sizes were n=3. Next-generation sequencing reads were allocated until a minimum amount of depth(10 reads) was met for any samples used in analysis.
Data exclusions	None to report.
Replication	All attempts at replication were successful and biological triplicates of each sample were used to verify reproducibility.
Randomization	Samples were analyzed equally, then characterized fully and reported in completion. There was no requirement for randomization.
Blinding	Blinding was not relevant. Animal or human participants were not used in this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging