

## Defining the mutational and conformational space of prokaryotic ribosomes for enhanced function

Wenfa Ng

Department of Chemical and Biomolecular Engineering, National University of Singapore,  
Email: [u026654b@alumni.nus.edu.sg](mailto:u026654b@alumni.nus.edu.sg) ORCID ID: <https://orcid.org/0000-0001-7629-2140>

### Abstract

Ribosomes are macromolecular complexes important to protein translation, and thus, essential to life. Comprising an ensemble of ribosomal proteins and RNA molecules, ribosomes are conserved in structure and function across all domains of life, but recent structural studies have revealed differentiated structures of ribosomes from bacterial, archaea and eukaryotes. Additionally, unique ribosomal protein mass fingerprints have been found for individual species; thereby, indicating that ribosomes are differentiated in structure amongst different species. Given that structure defines function, differentiated function likely exists amongst ribosomes of different species, which could manifest as differences in translation efficiency that could impact on cell growth rate. But ribosomal proteins also hold phylogenetic significance in informing the evolutionary trajectory of each species. Such ribosomal proteins are thus not highly conserved and offers sufficient sequence space for the evolution of differentiated structure and function in different species. Using ribosomal proteins that hold phylogenetic significance as templates, this study sought to understand the mutational and conformational limits that define functional ribosomes. Specifically, ribosomal proteins in *Bacillus subtilis* that hold phylogenetic cues would be mutated through error-prone polymerase chain reaction to generate variants that are subsequently transformed into *Escherichia coli*. To help assess the functional properties of the heterologous ribosomal proteins, endogenous ribosomal protein genes would be inactivated by multiplex CRISPR interference. Since variants in ribosomal proteins would likely impact on ribosome function and translation efficiency, live/dead screening on LB agar would be effective as a preliminary screen for functional mutants. These mutants would subsequently be inoculated into liquid LB medium in 96 well plates to quantify relative growth rates between different strains harbouring different heterologous variants of ribosomal proteins. Plasmids containing different ribosomal protein mutants would be extracted from each functional strain and subjected to Sanger sequencing for determining the specific mutations involved. Collection of such mutations would provide a comprehensive mutational map that define the limits of ribosomal protein sequence space important to ribosome function. Furthermore, biochemical isolation of ribosomal proteins and their structural characterization by X-ray crystallography or cryo-electron microscopy would further illuminate the structural significance of each mutation on ribosome structure and function; thereby, elucidating the structural tolerance space for functional ribosomes. Overall, generating a diverse pool of mutant ribosomal proteins in viability assays followed by sequencing and structural characterization would help define the mutational and conformational limits of a functional and efficient ribosome.

**Keywords:** CRISPR interference, error-prone polymerase chain reaction, ribosomal proteins, ribosome, viability assay, *Bacillus subtilis*, *Escherichia coli*, sequence space, conformational space, mutational map,

**Subject areas:** biochemistry, biotechnology, microbiology, structural biology, synthetic biology,

## 1. Introduction

Ribosomes are essential macromolecular complex synthesising the proteins necessary for cellular function. Thus, they are necessarily conserved in both structure and function.<sup>1,2</sup> Structural biology studies have indeed shown that there is remarkable conservation of ribosome structure at the domain level, where differentiating differences in structures could be seen for ribosomes of the prokaryotic, eukaryotic and archaeal domains.<sup>3,4</sup> Given that structure underpins function, ribosomes from prokaryotic, eukaryotic and archaeal species should be differentiated in function in the context of differing catalytic efficiency for protein translation, but remain conserved in general function, i.e., protein translation.

However, recent studies have illuminated the existence of differentiated structure to ribosomes from different species within the same domain.<sup>5-8</sup> Specifically, analysis of the molecular mass of the set of ribosomal proteins that constitute the ribosomes revealed the existence of unique fingerprints of ribosomal protein molecular mass for each species.<sup>5,9</sup> Known as ribosomal protein mass fingerprints,<sup>5,9,10</sup> existence of unique mass fingerprints for each species hold implications for the existence of differentiated structure and function to ribosomes from individual species. Thus, ribosomes from different species are likely to be differentiated in structure and function, while the general structures and functions of the ribosomes are conserved. Such differentiation in structure and function in ribosomes are anticipated to arise in differing catalytic efficiency of the macromolecular complex, especially during protein translation.

Multiple ribosomal proteins constitute the ribosomes. For the bacterial ribosomes, ribosomal proteins L1 to L36 assemble together with ribosomal RNA to constitute the large ribosome subunit. On the other hand, ribosomal proteins S1 to S21 together with 16S rRNA constitute the small ribosome subunit. Given that multiple ribosomal proteins constitute each of the ribosome subunit in contact with ribosomal RNA, possibilities for co-evolution between ribosomal proteins, and between ribosomal protein and ribosomal RNA exist.<sup>11-14</sup> Specifically, 16S rRNA is the gold standard biomolecule for chronicling the evolutionary trajectories of different species across the three domains of life.<sup>15-17</sup> Thus, possibilities of ribosomal proteins in the large and small ribosome subunit co-evolving with 16S rRNA due to intimately functional and structural contacts within and between the ribosomal subunits implied that some of the ribosomal proteins could also hold phylogenetic significance in documenting the evolutionary histories of different species.<sup>18-21</sup>

To this end, preliminary data emanating from a study seeking to understand the set of ribosomal proteins that could inform the phylogeny of different *Bacillus* species revealed the existence of ribosomal proteins in the large and small ribosome subunit that could reconstruct the evolutionary histories of studied *Bacillus* species.<sup>18</sup> Specifically, ribosomal proteins that

hold phylogenetic significance could reconstruct phylogenetic trees of the *Bacillus* species in a concordant manner similar to the case for 16S rRNA gene. To help chronicle the evolutionary history of a species, a biomolecule could not be highly conserved in sequence space or there would be lack of sequence diversity to document the evolutionary history traversed by the species. Detailed analysis of the 16S rRNA gene across species have revealed the existence of hypervariable region that could document the evolutionary trajectory of the species.<sup>22-24</sup> Similarly, ribosomal proteins that hold phylogenetic significance could harbour hypervariable regions that could help chronicle the evolutionary processes that shape the genetic history of the species. Such ribosomal proteins are likely to react in similar ways to the evolutionary natural selection forces that shape the 16S rRNA gene and ribosomal proteins in a co-evolutionary dance.

From a different perspective, ribosomal proteins with hypervariable regions that chronicle the evolutionary history of the species are also not especially critical to ribosome structure and function. This argument is logical but tentative as there is paucity of information to confirm the hypothesis. However, observations of highly conserved ribosomal proteins leading to the reconstruction of phylogenetic tree with closely-related *Bacillus* species where, in fact they aren't, according to the phylogenetic tree based on 16S rRNA provided confidence about the plausibility of the above hypothesis. On the other hand, ribosomal proteins that are highly variable and could not reproduce the phylogenetic tree based on 16S rRNA gene also existed in the set of ribosomal proteins examined. Such ribosomal proteins are unlikely to be essential to core ribosome structure and function (i.e., translation) and as such could be mutated in diverse ways resulting in the inability to chronicle the evolutionary history of the different species.

## 2. Proposal

Ribosomal proteins that hold phylogenetic significance thus occupy a special place in ribosomal protein biology given that they likely co-evolved with each other and with 16S rRNA, and at the same time, are not highly conserved in sequence due to structural requirement. Given that they are not highly conserved in sequence, these ribosomal proteins could form the basis of mutagenesis experiments aimed at understanding the limits of the mutational and conformation space of ribosomes. Specifically, error-prone polymerase chain reaction (error-prone PCR) could be used to generate mutant libraries of each of the ribosomal proteins that hold phylogenetic significance. These proteins could be ligated into plasmid vectors and transformed into suitable microbial chassis for understanding the functional impact of the mutant ribosomal proteins on *in vivo* cellular function. Clustered regularly interspersed short palindromic repeats interference (CRISPRi)<sup>25,26</sup> would be used to prevent the expression of the corresponding endogenous ribosomal proteins in the genome of the microbial chassis. Functional mutants harbouring an altered ribosomal protein would be selected by a live/dead screen on LB agar, after which glycerol stock cultures of the functional mutants would be made for subsequent characterization of the growth performance of the mutants. Growth rate is chosen as a proxy parameter for ribosome function given that more efficient translation could result in faster growth rate.

Mutants with enhanced growth rates would be sequenced to determine the mutations that confer improved functionalities to the ribosomal protein and ribosome. To help search for the mutational and conformational space that help define enhanced ribosome function, plasmid vectors would be constructed with multiple mutant ribosomal proteins that could facilitate the understanding of how different mutant proteins influence ribosome function. Similarly, multiplex CRISPRi would be used to silence the expression of endogenous native copies of the ribosomal proteins.

Finally, structural biology studies such as X-ray crystallography, cryo-electron microscopy and molecular dynamics simulations would be used to understand how individual mutations in ribosomal proteins that hold phylogenetic significance influence ribosomal proteins function; thereby, helping us to plot the mutational and conformational space for enhanced ribosome function in prokaryotes.

### 3. Conclusions

Ribosomal proteins could be classified into three major classes: (i) those that co-evolved with each other and 16S rRNA, (ii) ribosomal proteins that could hold a large sequence space and are not essential to core ribosome structure and function, and (iii) highly conserved ribosomal proteins essential to ribosome structure and function. By using ribosomal proteins that hold phylogenetic significance as sequence template, this study sought to understand the sequence and conformational space important to enhanced ribosome function and protein synthesis. Such understanding would illuminate deep mysteries concerning how individual ribosomal proteins influence ribosome function and whether mutating individual or a consortium of ribosomal proteins could lead to enhanced ribosome function.

### Conflicts of interest

The author declares no conflicts of interest.

### Funding

No funding was used in this work.

### References

- 1 Agmon, I., Bashan, A. & Yonath, A. On Ribosome Conservation and Evolution. *Israel Journal of Ecology & Evolution* **52**, 359-374, doi:10.1560/IJEE\_52\_3-4\_359 (2006).
- 2 Kallberg, Y., Segerstolpe, Å., Lackmann, F., Persson, B. & Wieslander, L. Evolutionary Conservation of the Ribosomal Biogenesis Factor Rbm19/Mrd1: Implications for Function. *PLOS ONE* **7**, e43786, doi:10.1371/journal.pone.0043786 (2012).

- 3 Melnikov, S., Manakongtreecheep, K. & Söll, D. Revising the Structural Diversity of Ribosomal Proteins Across the Three Domains of Life. *Molecular Biology and Evolution* **35**, 1588-1598, doi:10.1093/molbev/msy021 (2018).
- 4 Fox, G. E. Origin and Evolution of the Ribosome. *Cold Spring Harbor Perspectives in Biology* **2** (2010).
- 5 Ng, W. Existence of theoretical ribosomal protein mass fingerprints in bacteria, archaea and eukaryotes. *PeerJ Preprints* **6**, e26511v26511, doi:10.7287/peerj.preprints.26511v1 (2018).
- 6 Dao Duc, K., Batra, S. S., Bhattacharya, N., Cate, Jamie H. D. & Song, Yun S. Differences in the path to exit the ribosome across the three domains of life. *Nucleic Acids Research* **47**, 4198-4210, doi:10.1093/nar/gkz106 (2019).
- 7 Xu, Z. & Culver, G. M. Differential assembly of 16S rRNA domains during 30S subunit formation. *RNA* **16**, 1990-2001 (2010).
- 8 Gulen, B. *et al.* Ribosomal small subunit domains radiate from a central core. *Scientific Reports* **6**, 20885, doi:10.1038/srep20885 (2016).
- 9 Ng, W. Theoretical ribosomal protein mass distribution of *Pseudomonas aeruginosa* PAO1. *PeerJ Preprints* **6**, e3500v3501, doi:10.7287/peerj.preprints.3500v1 (2018).
- 10 Ng, W. Theoretical ribosomal protein mass fingerprint database. *PeerJ Preprints* **6**, e26878v26871, doi:10.7287/peerj.preprints.26878v1 (2018).
- 11 Dunstan, M. S., GuhaThakurta, D., Draper, D. E. & Conn, G. L. Coevolution of Protein and RNA Structures within a Highly Conserved Ribosomal Domain. *Chemistry & Biology* **12**, 201-206, doi:10.1016/j.chembiol.2004.11.019 (2005).
- 12 Brandman, R., Brandman, Y. & Pande, V. S. Sequence Coevolution between RNA and Protein Characterized by Mutual Information between Residue Triplets. *PLOS ONE* **7**, e30022, doi:10.1371/journal.pone.0030022 (2012).
- 13 Slinger, B. L., Newman, H., Lee, Y., Pei, S. & Meyer, M. M. Co-evolution of Bacterial Ribosomal Protein S15 with Diverse mRNA Regulatory Structures. *PLOS Genetics* **11**, e1005720, doi:10.1371/journal.pgen.1005720 (2015).
- 14 Roberts, E., Sethi, A., Montoya, J., Woese, C. R. & Luthey-Schulten, Z. Molecular signatures of ribosomal evolution. *Proceedings of the National Academy of Sciences* **105**, 13953, doi:10.1073/pnas.0804861105 (2008).
- 15 Yang, B., Wang, Y. & Qian, P.-Y. Sensitivity and correlation of hypervariable regions in 16S rRNA genes in phylogenetic analysis. *BMC Bioinformatics* **17**, 135, doi:10.1186/s12859-016-0992-y (2016).
- 16 Sato, M. & Miyazaki, K. Phylogenetic Network Analysis Revealed the Occurrence of Horizontal Gene Transfer of 16S rRNA in the Genus *Enterobacter*. *Frontiers in Microbiology* **8**, doi:10.3389/fmicb.2017.02225 (2017).
- 17 Rajendhran, J. & Gunasekaran, P. Microbial phylogeny and diversity: Small subunit ribosomal RNA sequence analysis and beyond. *Microbiological Research* **166**, 99-110, doi:<https://doi.org/10.1016/j.micres.2010.02.003> (2011).
- 18 Ng, W. Ribosomal proteins could explain the phylogeny of *Bacillus* species. *bioRxiv*, 2020.2007.2025.221481, doi:10.1101/2020.07.25.221481 (2020).
- 19 Yutin, N., Puigbò, P., Koonin, E. V. & Wolf, Y. I. Phylogenomics of Prokaryotic Ribosomal Proteins. *PLOS ONE* **7**, e36972, doi:10.1371/journal.pone.0036972 (2012).
- 20 Mullis, A. *et al.* Parallel Concerted Evolution of Ribosomal Protein Genes in Fungi and Its Adaptive Significance. *Molecular Biology and Evolution* **37**, 455-468, doi:10.1093/molbev/msz229 (2020).

- 21     Matte-Tailliez, O., Brochier, C., Forterre, P. & Philippe, H. Archaeal Phylogeny Based on Ribosomal Proteins. *Molecular Biology and Evolution* **19**, 631-639, doi:10.1093/oxfordjournals.molbev.a004122 (2002).
- 22     Guo, F., Ju, F., Cai, L. & Zhang, T. Taxonomic Precision of Different Hypervariable Regions of 16S rRNA Gene and Annotation Methods for Functional Bacterial Groups in Biological Wastewater Treatment. *PLOS ONE* **8**, e76185, doi:10.1371/journal.pone.0076185 (2013).
- 23     Kerrigan, Z., Kirkpatrick, J. B. & D'Hondt, S. Influence of 16S rRNA Hypervariable Region on Estimates of Bacterial Diversity and Community Composition in Seawater and Marine Sediment. *Frontiers in Microbiology* **10**, doi:10.3389/fmicb.2019.01640 (2019).
- 24     Barb, J. J. *et al.* Development of an Analysis Pipeline Characterizing Multiple Hypervariable Regions of 16S rRNA Using Mock Samples. *PLOS ONE* **11**, e0148047, doi:10.1371/journal.pone.0148047 (2016).
- 25     Larson, M. H. *et al.* CRISPR interference (CRISPRi) for sequence-specific control of gene expression. *Nature Protocols* **8**, 2180-2196, doi:10.1038/nprot.2013.132 (2013).
- 26     MacLeod, R. S. *et al.* Effective CRISPR interference of an endogenous gene via a single transgene in mice. *Scientific Reports* **9**, 17312, doi:10.1038/s41598-019-53611-6 (2019).