Title
Structural and Functional Studies of Escherichia coli Ribosomes and Their Polymerization of Unnatural Monomers

Author
Ward, Frederick R.

Publication Date
2021

Peer reviewed|Thesis/dissertation

Structural and Functional Studies of *Escherichia coli* Ribosomes and Their Polymerization of Unnatural Monomers

By

Frederick R. Ward

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in Charge:

Professor Jamie Cate, Chair
Professor Michelle Chang
Professor Matthew Francis
Professor Michael Marletta

Spring 2021

Abstract

Structural and Functional Studies of *Escherichia coli* Ribosomes and Their Polymerization of Unnatural Monomers

by

Frederick R. Ward

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Jamie Cate, Chair

The ribosome catalyzes the synthesis of polypeptides with high efficiency and sequence specificity. Repurposing the ribosome as a platform for manufacturing other sequence defined polymers could access a wide variety of previously unattainable molecules and bulk materials. In this work, we aim to understand both engineered and wild type ribosomes through structural and biochemical analysis. We reveal limitations of previous engineering efforts on the ribosome, highlighting the importance of careful mutation and selection techniques. The ribosome we study is poorly assembled and nonfunctional *in vitro* despite improved polymerization of β-amino acids *in vivo*. We identify key regions of the ribosome that are disrupted by mutations and offer suggestions for more targeted engineering that will preserve efficient ribosome assembly. We also characterize the structure of the wild type ribosome bound to an unnatural, non-α-amino acid monomer for the first time. This monomer is correctly accommodated into the P site of the ribosome, explaining previously observed activity as an initiator substrate. Lastly, we assess a new, short peptide luciferase-complementing reporter in defined *in vitro* translations as a better readout of mutant ribosome activity. Using this assay, we show activity of purified active site mutant ribosomes *in vitro* for the first time, reconciling often observed differences from *in vivo* systems. These projects underscore the importance of detailed characterization of both the input to and results of ambitious bioengineering efforts.

Table of Contents

Acknowledgements

I am overwhelmingly grateful to the host of people who have supported me during my time here at Berkeley. As passionate as I am about science, biology, and the ribosome, it is the people I had the honor of working with that made my PhD memorable, productive, and enjoyable. First, I need to thank Jamie for his 5 years as my advisor and mentor. I have learned so much about the ribosome, structural biology, and scientific thinking in general. Especially appreciated is the endless enthusiasm you can produce after inevitable periods of little progress. It is easy to stay excited and scientifically curious in such an environment! Our labs in Stanley and IGIB were collaborative and collegial places because of the environment you cultivated, which cannot be taken for granted. I am glad someone showed me the mistake I was making by avoiding structural biology (what better place to start than with the ribosome?!). I am also grateful to my committee members, Profs. Chang, Francis, and Marletta, whose extensive experience and insight helped troubleshoot my science and focus my goals during my degree.

Leaving the Cate Lab is going to be hard; there are so many individual people, now friends, who have helped me grow scientifically over the years. Thanks to Wenfei for all the support, collaboration, and patience -- you always had time to look at weird data or commiserate over my latest mistake. I'm wishing you and your the very best as you start your own independent ribosome career. Thanks to Mia and Luisa for countless conversations (of varying degrees of scientific relevance) in the lab and I hope we can all gather to celebrate our graduations properly soon. Thanks to Zoe for being a consummate collaborator and friend. I know we accomplished much more than the sum of two independent efforts could. I appreciate the repeated explanations of cryo-EM physics that I will never understand. Chandrima, I am glad someone so capable and energetic is taking over my often disorganized project structure. I can't wait to see what you, Zoe, and others discover. My undergrads Maddie and Edward have been great, patient students, as I figured out how to be an effective teacher and mentor. Many thanks to the rest of the Cate Lab, I've learned a lot from such a diverse and accomplished group of people. Finally, thanks to the C-GEM team, my MCB cohort, and all my collaborators in and out of Berkeley. I'm lucky to have been part of so many positive spheres at once.

Outside of the lab, I am indebted to my family and friends, especially over the last year when I rarely saw anyone else. Thanks to our COVID pod of Adam, Dan, and Amy, I'm not sure how I would have stayed sane without you all. Much credit to Carl for 24/7 scientific support, much needed cycling and SC2 distractions, and general obligatory identical twin responsibilities. Alice, thanks for the impromptu visits, thoughtful support, and exemplary siblinghood. Thanks to Mom and Dad for making me want to be a scientist on top of everything else you've done for me. Cate, I'm lucky Carl rotated with you in our first year and I am constantly grateful that you're my partner. Knowing I had you in my corner kept me going many times, at least until we could take our frustrations out on the El Toyonal loop. We are in the middle of an overwhelming amount of changes and firsts for us but I know we will take it all in stride as a team. Thank you for everything over the last 4 years and I can't wait to see what we get into next!

1. Introduction

Protein synthesis in all living things is accomplished by the ribosome, a large and highly conserved molecular machine responsible for translating the four letter code of messenger ribonucleic acids (mRNA) into functional proteins. For each position in a growing polypeptide, the ribosome must quickly and accurately select the correct amino acid from 20 possibilities, catalyze the polymerization, and translocate the entire machine one codon for the next round. Nothing else on Earth can synthesize a complex defined polymer with such speed and fidelity.

1.1 Overview of bacterial translation

In bacteria, protein synthesis occurs through the discrete regulated steps of initiation, elongation, and termination (**Figure 1.1**). During initiation, the small ribosomal subunit (30S) associates with initiation factors 1, 2, and 3 (IF1/2/3), mRNA, and formylmethionine-charged initiator transfer RNA (fMet-$tRNA_{fMet}$) through various interactions usually including both base pairing between the 3' end of the 16S ribosomal RNA (rRNA) with a complementary Shine-Dalgarno sequence on the mRNA and recognition of a start codon (usually AUG) by the tRNA (Laursen et al., 2005). Notably, this process can accurately locate a true initiation codon from dozens of other spurious start sites even in the absence of a canonical AUG start codon or even correct rRNA-Shine-Dalgarno pairing (Saito et al., 2020). Following preinitiation complex formation, the large ribosomal subunit (50S) associates, initiation factors release, and the elongation phase begins. For each polymerization of a growing peptide chain the following steps occur on the ribosome: first, elongation factor Tu (EF-Tu) delivers a charged tRNA to the A site of the ribosome. After passing codon-anticodon selection and proofreading steps essential for accurate translation, the aminoacyl-tRNA is accommodated into the peptidyltransferase center (PTC) of the ribosome where it reacts with the P-site tRNA. Polymerization thus transfers the growing peptide chain to the A-site tRNA, necessitating translocation of both tRNAs to free up the A site for the next encoded amino acid. This process, catalyzed by elongation factor G (EF-G), involves intersubunit rotation that moves the P-site tRNA to the E site and the A-site peptidyl-tRNA to the P site, resetting the ribosome. This sequence is repeated, threading the growing protein through the exit tunnel of the 50S subunit until the ribosome reaches a stop codon. Stop codons are recognized by release factors that hydrolyze the ester bond between the polypeptide and tRNA, releasing the complete protein. The ribosome is then dissociated by ribosome recycling factor (RRF) and EF-G, preparing the machinery for the next mRNA (Dunkle & Cate, 2010).

In parallel to the motions of the ribosome, proper synthesis of aminoacyl-tRNAs is equally important for accurate translation. Aminoacyl-tRNA synthetase enzymes (aaRS) are responsible for joining each tRNA species to their corresponding amino acid. Amino acids are first activated by adenylation catalyzed by their cognate aaRS. The enzyme then transfers the monomer to the 2' or 3' end of the terminal adenosine of the proper tRNA, depending on the identity of the enzyme. Like tRNA selection on the ribosome, correct tRNA acylation is challenging and crucial. Many amino acids are similar in size and reactivity, and tRNAs themselves are similar in structure. To overcome this, aminoacylated-tRNAs are proofread through a combination of pre- and post-charging editing and the involvement of several *trans*-acting factors (Shepherd & Ibba, 2015). The ester bond between an amino acid and its

tRNA is often sensitive to hydrolysis, so aminoacyl-tRNAs are protected by association with EF-Tu before delivery to the ribosome (Hentzen et al., 1972; Peacock et al., 2014).

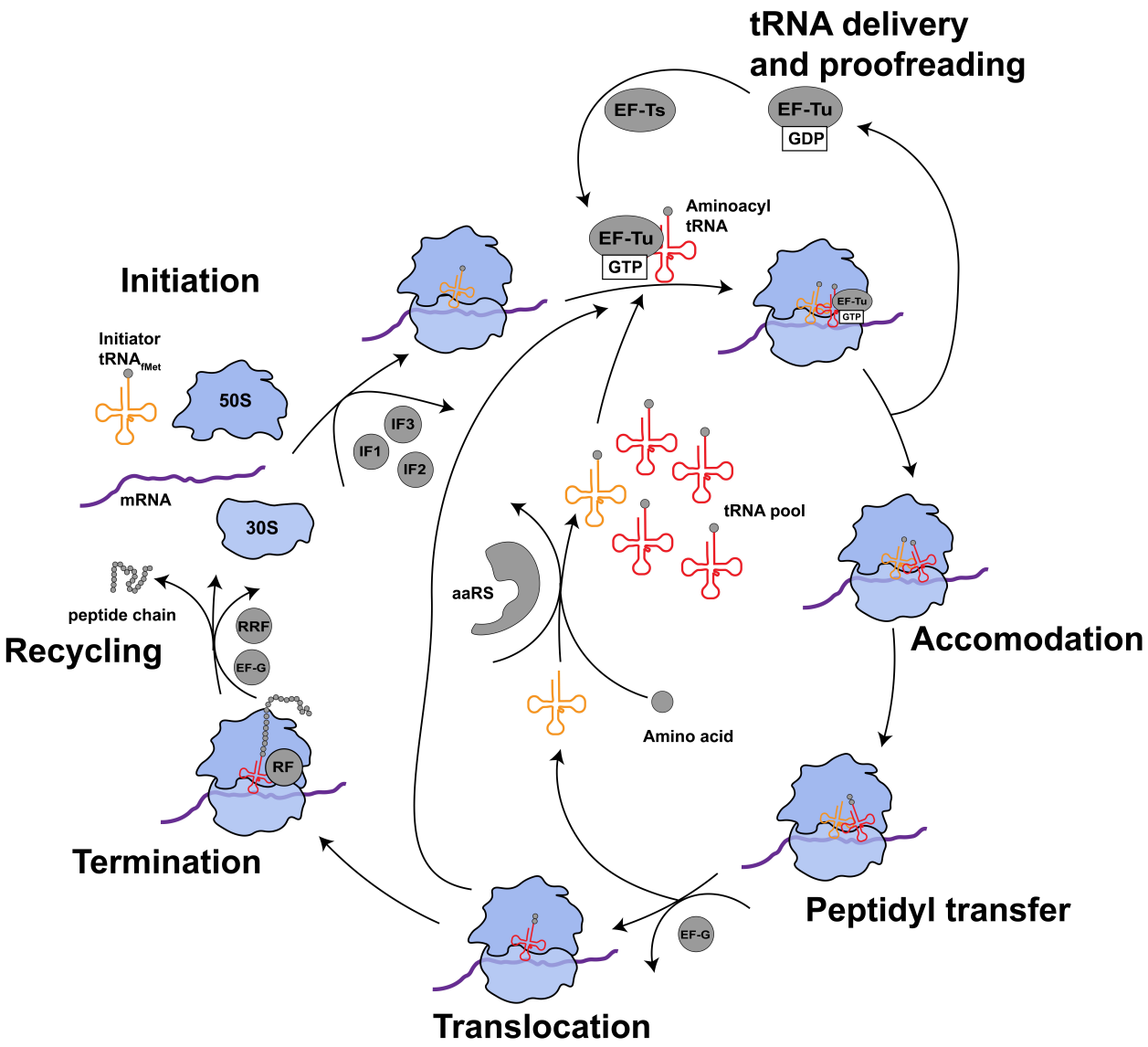


Figure 1.1. Overview of bacterial translation. Translation initiation occurs when initiation factors (IF1/2/3) facilitate association between the ribosomal large (50S) and small (30S) subunits, mRNA, and formylmethionine-charged initiator tRNA. Peptide elongation is then executed by a repeated multistep cycle. First, aminoacyl-tRNA is delivered by EF-Tu-GTP-tRNA ternary complex to the A site of the ribosome. Upon successful proofreading of the tRNA-mRNA anticodon-codon pairing, EF-Tu hydrolyzes its GTP substrate and dissociates. The A-site tRNA is then accommodated into the PTC, followed by peptide bond formation between the P-site substrate and A-site aminoacyl-tRNA. EF-G catalyzes translocation of the newly formed peptidyl-tRNA from the A site to the P site and the now deacylated P-site tRNA to the E site where it dissociates. Translocation also moves the mRNA through the ribosome in a one codon step, preparing for the next tRNA delivery. GDP-bound EF-Tu is recharged with GTP by its nucleotide exchange factor EF-Ts. Deacylated tRNAs are recharged by their cognate aaRSs and reassociate with EF-Tu-GTP. This cycle repeats until a stop codon is reached, where a release factor (RF1/2/3) recognizes the termination codon and hydrolyzes the completed peptide off of its tRNA. EF-G and RRF collaborate to split the ribosome during recycling, releasing the mRNA and both ribosomal subunits.

1.2 Ribosomal incorporation of nonstandard amino acids

The ribosome's unparalleled polymer synthesis abilities have spawned an entire field of research dedicated to repurposing its catalytic functions. Perhaps the most promising feature of the ribosome as an engineering target is not its template-dependent synthesis but its inherent promiscuity. The ribosome has a single active site in which 400 or more different amino acid combinations react during translation, suggesting it has an innately relaxed substrate specificity. Indeed, selenocysteine and pyrrolysine are used in some organisms as the 21st and 22nd amino acids (Ambrogelly et al., 2007), examples of natural genetic code expansion. Furthermore, it was known since the 1960s that the ribosome does not proofread the monomer-tRNA pairing (Chapeville et al., 1962). If a suitable substrate can be delivered to the PTC, it will react with little regard to its linked tRNA identity. Structures of the ribosome bound to aminoacyl-tRNAs all show the A-site side chain sequestered in a pocket, presumably to prevent clashes in the PTC proper and correctly expose the reactive amine (Martin Schmeing et al., 2005; S. Melnikov et al., 2016; S. V. Melnikov et al., 2019; Polikanov et al., 2014). Furthermore, current models of substrate binding in the PTC suggest an evolved induced fit mechanism that allows the PTC to accept large amino acids before closing down on the substrates to promote peptidyl transfer (Lehmann, 2017). This reveals an evolutionary solution for the obligately promiscuous ribosome: accept and remove side chains of all shapes and sizes from the equation. Capitalizing on this in the last three decades, a host of systems have been developed for site-specific incorporation of hundreds of unnatural monomers into ribosomally synthesized proteins (Chin, 2014; Dumas et al., 2015; Liu & Schultz, 2010).

While the ribosome may be accepting of many shapes and sizes of α-amino substrates, traditional genetic code expansion needs to solve the problem of highly specific tRNA charging. Each tRNA is charged by a specific aaRS and the fidelity of these reactions is essential for accurate protein synthesis. Disrupting this balance by mutating an aaRS active site to accept more substrates, for example, would result in mischarging of natural tRNAs with the expanded substrate set and their incorporation into proteins with likely toxic effects. To solve this, orthogonal tRNA/aaRS/substrate pairs were developed that do not interact with natural components. Usually the orthogonal tRNA is assigned to a nonsense codon, often the amber (UAG) stop codon, where it competes with release factor 1's (RF1) termination ability. Because there are no naturally "open" codons, any system for genetic code expansion must overcome competition with natural tRNAs or release factors. The most commonly used systems in bacteria are the *Mj*TyrRs/$tRNA_{CUA}$ (Liu & Schultz, 2010) and PylRS/$tRNA_{CUA}$ (Chin, 2014) pairs, both amber stop codons suppressors. These orthogonal systems have been extensively optimized through design and selection experiments to improve orthogonality and increase yield (Chin, 2014; Liu & Schultz, 2010). Improvements include positive and negative selections for better orthogonality and new substrate recognition (Santoro et al., 2002), knocking out RF1 to improve amber codon suppression (Mukai et al., 2010), removing codons and release factors completely using advanced genome engineering to completely reassign codons (Fredens et al., 2019; Lajoie et al., 2013), developing quadruplet codon decoding tRNAs (Hohsaka et al., 1996; Neumann et al., 2010), and introducing new base pairs in DNA and RNA to generate novel codon-anticodon interactions (Y. Zhang et al., 2017).

Orthogonal tRNA/aaRS pairs have proven extremely useful tools but are fundamentally limited in substrate scope. Active sites of each aaRS are only so malleable through directed evolution, and orthogonal aaRS-tRNA pairs are limited in number. If *in vivo* systems are not

necessary, this problem can be largely circumvented by using a set of RNA aptamers called flexizymes to charge tRNAs. Flexizymes recognize specific leaving groups linked to the carboxyl-ends of substrates and the -CCA end of tRNAs, bringing the two in proximity and promoting charging (Goto et al., 2011; Xiao et al., 2008). There are four different flexizyme RNAs in common use that have different specificities for leaving group and monomer identity, together encompassing a large swath of chemical space (Goto et al., 2011). Pairing flexizymes with defined recombinant *in vitro* translation systems (Shimizu et al., 2001) has allowed exploration of a much wider range of substrates than *in vivo*. The ease of genetic code reprogramming *in vitro* allows, in theory, complete reassignment of every codon to a new monomer. In the context of α-amino acids, this system has, for example, been used to finely probe peptide membrane insertion energetics by systematically varying sidechain properties at a level impossible with the standard 20 amino acids (Öjemalm et al., 2011).

1.3 Ribosomal incorporation of non-L-α-amino monomers

While decades of engineering and optimization have shown the ribosome to be highly tolerant of genetic code expansion with α-amino substrates, backbone-modified monomers pose a more difficult obstacle. With this challenge, though, comes the potential to access new classes of materials and small polymers at specificities and scales previously impossible. Examples of such materials include β-amino acid peptidomimetics with decreased immunogenicity and improved protease resistance (Geueke et al., 2006; Guichard et al., 2000), polyaramids with extreme strength in bulk or potent antibacterial activity as small oligomers (Baumann et al., 2014; Tanner et al., 1989), allyl acrylamide antibacterial oligomers with enhanced serum stability (Porel et al., 2017; Porel & Alabi, 2014), aromatic and peptide foldamers with high-order self assembly properties (Hill et al., 2001; Knight et al., 2015; Rinaldi, 2020), and polyketide or polyketide/peptide hybrid oligomers synthesized with extreme diversity in structure and function across the tree of life (Du et al., 2001; Walsh, 2004). These non-L-α-amino oligomers and polymers are currently synthesized *in vitro* in bulk, which limits control over sequence, or through step-by-step chemical synthesis that allows precise sequence control at the expense of yield. However, the ribosome and accompanying translational apparatus presents a solution to both problems: templated sequence-defined polymerization at scale.

Progress toward ribosomal synthesis of new polymers can be separated into two lines of research: 1) probing the substrate specificity of the wild type ribosome and 2) engineering the PTC to access new chemistry. With limited efficiency and scope, the wild type ribosome can catalyze non-L-α-amino polymerizations. The first example of this identified was ester formation between fMet and a puromycin derivative carrying a hydroxyl group instead of an amine as its nucleophile (Fahnestock et al., 1970). This result was soon extended to show that the wild type *E. coli* ribosome could extend multiple phenyllactic acid monomers, confirming that the ribosome can synthesize polymers beyond peptides (Fahnestock & Rich, 1971). Subsequently, reactions in the PTC have been confirmed with N-methyl amino acids (Subtelny et al., 2008; Jinfan Wang et al., 2014; B. Zhang et al., 2007), peptoids (Kawakami et al., 2008), β-amino acids (Fujino et al., 2016; Katoh & Suga, 2018), cyclic amino acids (Katoh et al., 2020; Katoh & Suga, 2020a; Lee, Schwarz, et al., 2020; Lee, Torres, et al., 2020), long carbon backbone amino acids (Lee, Schwarz, et al., 2020), D-amino acids (Achenbach et al., 2015; Fujino et al., 2013; Katoh, Iwane, et al., 2017; Katoh, Tajima, et al., 2017; Liljeruhm et al., 2019), benzoic acids (Ad et al., 2019; Katoh & Suga, 2020b; Lee et al., 2019), additional hydroxy-acids and malonic acids (Ad

et al., 2019; Ohta et al., 2008), and aromatic foldamers (Rogers et al., 2018). Notably, many substrates are only competent as initiating monomers due to their singly-reactive structure or presumed incompatibility in the A-site (Ad et al., 2019; Goto & Suga, 2009; Rogers et al., 2018).

Of special note are a handful of reports confirming consecutive elongation of non-L-α amino acids, both with D-amino (Katoh, Tajima, et al., 2017) and β-amino acids (Katoh & Suga, 2018). While both monomer classes are not particularly exotic compared to natural proteins, their repeated polymerization is the first foray into new materials from the ribosome. Both results required the use of specially engineered tRNAs to increase affinity for EF-Tu, improving delivery of the acyl-tRNA to the ribosome. Lack of affinity for EF-Tu is an additional problem with many unnatural acyl-tRNAs. EF-Tu-acyl-tRNA affinity interactions are complicated, depending on the identity of the tRNA and charged species (LaRiviere et al., 2001), so mischarging can decrease affinity. Using a $tRNA_{Glu}$ body with high affinity for EF-Tu compensated effectively for a low affinity monomer, allowing unprecedented levels of delivery and incorporation. For β-amino acids, an additional motif from $tRNA_{Pro}$ was grafted on the synthetic tRNA, granting affinity for elongation factor P (EF-P), a translation factor that promotes poly-proline synthesis (Ude et al., 2013) by binding to the E-site of the ribosome. Addition of EF-P to the *in vitro* translation system in combination with the chimeric tRNA allowed synthesis of 7 consecutive β-amino acids (Katoh & Suga, 2018). Together, these results demonstrate that, even with a wild type ribosome, the translational apparatus can be repurposed toward new chemistries.

1.4 Engineering the ribosome to access new chemistry

While research has revealed an impressive catalytic flexibility in the wild type PTC, engineering will be essential to attain useful yields and access all polymers of interest. In tandem with wild type ribosome explorations, several methods for mutating the PTC and screening for new catalytic ability have been developed, resulting in mutant ribosomes with new chemical ability. The primary challenge in PTC engineering has been establishing screens or selections to identify improved mutants. The challenges of codon reassignment and tRNA misacylation with non-L-α-amino monomers *in vivo* hamper the ability to employ a classic selection over a library of PTC mutants. Foundational experiments demonstrating the feasibility of ribosome engineering came from the laboratory of Sidney Hecht where a small region of the 23S rRNA was targeted to improve D-amino acid incorporation (Dedkova et al., 2003). The bases mutated, G2447-A2451, lie next to the A-site cleft, a pocket formed by bases A2451 and C2452, that binds the side chain of the incoming acyl-tRNA. Modeling efforts with early ribosome structures suggested that accommodation of a D-amino acid's aide chain in the A-site cleft would position the monomer's amine incorrectly for peptidyl transfer (Zarivach et al., 2004). All mutated bases but A2450 are mutationally pliable *in vitro* and *in vivo* (d'Aquino et al., 2020; O'Connor et al., 2001; Sato et al., 2006; J. Thompson et al., 2001). Importantly, chloramphenicol selection was used during growth of strains expressing the mutant ribosomes. Chloramphenicol binds the A-site cleft, suggesting that resistant strains might be expressing mutant ribosomes with an altered A-site cleft geometry but otherwise functional PTC. This selection scheme was likely crucial for the experiment's success, as it led to a convergence of clones isolated (Dedkova et al., 2003). S-30 extract *in vitro* translation identified multiple clones with improved D-amino acid incorporation ability. While successful, this engineering scheme did not use a screening or selection technique that tested ribosomes for D-amino acid incorporation. While the potential for PTC engineering was

revealed, more complex screens involving tests for new catalytic activity would likely be necessary for more challenging substrates than D-amino acids.

The next mutant ribosome developed by Hecht and coworkers demonstrated a powerful new selection technique: puromycin derivatives for negative selection. A wider set of 23S rRNA mutations were screened for erythromycin resistance, again using antibiotic resistance as a proxy for PTC remodeling, and sensitivity to β-3-puromycin. Puromycin is an A-site substrate analog antibiotic that prematurely terminates protein synthesis. Mutant ribosomes with enhanced sensitivity to β-3-puromycin, a puromycin analog that resembled a β-amino acid charged tRNA, presumably had a PTC geometry more accepting of the imitated substrate. This selection system circumvented the need for a β-amino-acyl-tRNA charging system *in vivo*, avoiding a challenging hurdle in ribosome engineering. The resulting best isolated mutants showed a ~ 3-5 fold better incorporation of β-puromycin in cell extracts (Dedkova et al., 2012). With the promise of PTC engineering now clear, mutant ribosomes have been developed for the incorporation of dipeptide substrates (Maini, Dedkova, et al., 2015), polyproline stretches (Schmied et al., 2018), and β-amino acids *in vivo* (Melo Czekster et al., 2016). However, there remain significant limitations to all PTC mutants thus far engineered. Primarily, each ribosome is selected and analyzed for mixed α-amino and novel substrate translation. This process returns ribosomes that must still be competent at standard protein synthesis to some degree. Furthermore, each hit from a screen over PTC mutants has been tested only for single incorporation of the new monomer, quite a distance from producing polymers with new backbones. In the context of our studies of a β-amino acid translating mutant PTC ribosome (Ward et al., 2019), these limitations are discussed extensively in chapter two.

1.5 Strategies to enhance ribosome engineering

Currently, the most promising platform for ribosome engineering is the design of a suitable *in vivo* selection system capable of testing billions of PTC rRNA mutants. Even assuming the existence of a suitable method for charging orthogonal tRNAs with exotic monomers *in vivo*, implementing an effective selection on mutant ribosomes is limited by other factors. The primary concern is the centrality of translation to cellular viability. Every protein must be synthesized by the ribosome and the speed and accuracy of this translation is paramount to a cell's survival. Expressing a population of mutant ribosomes with modified PTCs would likely lead to slow, error-prone, or stalled translation on many transcripts, including many essential genes. The accumulation of truncated, misfolded, or incorrect proteins would cause extreme stress to the cell, especially considering protein synthesis can use up to two thirds of all ATP in a growing bacterial cell (Russell & Cook, 1995). In fact, this emergence of inefficient α-amino translation is probably a necessary condition of successfully engineered ribosomes when using current screening and selection techniques, making the intrusion on protein homeostasis unavoidable.

This problem has been greatly ameliorated by the development of orthogonal ribosomes over the last several decades with efforts focused on two key ribosomes interactions: 16S rRNA and Shine-Dalgarno sequence discrimination and subunit association. The Shine-Dalgarno (SD) sequence is a conserved sequence upstream of the AUG start codon in bacterial and archeal mRNA (Shine & Dalgarno, 1975). Base pairing between the SD sequence and the 3' end of the 16S rRNA is a key determinant of correct translation initiation complex localization. Exploiting this simple-to-control molecular interaction, co-mutation of the SD sequence and 16S rRNA was

first attempted by Hui and de Boer, demonstrating that ribosomes with altered 16S rRNAs would preferentially translate mRNAs with complementary SD sequences (Hui & de Boer, 1987). This important result presented a straightforward way to separate pools of translating ribosomes. Later, exhaustive selection over libraries of SD-16S pairs identified sequences with the least cross talk, maximizing orthogonality (Rackham & Chin, 2005).

It is important to note that SD-16S interactions are not the sole determinants of translation initiation. First, SD sequence usage varies both across bacterial and archaeal phyla, with some species appearing to not have SD sequences at all (Chang et al., 2006; Nakagawa et al., 2010). These organisms likely use other forms of translation initiation including leaderless initiation (Balakin et al., 1992; Moll et al., 2001; Udagawa et al., 2004), where 30S or 70S ribosomes bind 5' AUG start codons directly, or ribosomal protein S1 (S1) mediated initiation (Boni et al., 1991; Komarova et al., 2005), where S1 assists in binding mRNA through association with A/U rich tracts (Duval et al., 2013). Interestingly, not all microbes encode a translation-initiation-functional S1, and there is a correlation between lacking S1 and high SD-16S pairing stringency (Farwell et al., 1992; Salah et al., 2009). Finally, even in *E. coli* where SD sequences are common, ribosomes with mutated 16S rRNA 3' ends still initiate at the correct locations genome-wide (Saito et al., 2020). Indeed, even the best orthogonal SD-16S pairs still display some cross talk between mRNA and ribosomes, possibly mediated by S1-mRNA association or other factors. 5' untranslated region engineering beyond the SD site is an avenue for further improving ribosome-mRNA orthogonality.

Even with orthogonal SD-16S sequences directing ribosomes to separate mRNA polls, the subunit architecture of the ribosome presents an additional barrier for PTC engineering. The PTC exists in the large subunit of the ribosome but SD sequences are recognized by the small subunit. Mutant large subunits can associate with any small subunit, orthogonal or wild type, on any mRNA, meaning that altered PTCs cannot be sequestered from normal cellular translation under this regime. Missing is a connection between the small and large subunits that ensure orthogonal 30S subunits only interact with mutant 50S subunits. Recently, this issue has been tackled by physically tethering the two ribosomes together using RNA linkers (Fried et al., 2015; Orelle et al., 2015). Surprisingly, these linkages result in mostly efficient ribosomes that can support *E. coli* growth with only a limited impact on ribosome assembly (Aleksashin et al., 2019), subverting billions of years of ribosome organization. Optimization of these linkers has resulted in improved efficiency and orthogonality (Carlson et al., 2019; Schmied et al., 2018). Notably, these tethered ribosomes have allowed evolution of PTC mutants that can translate challenging peptide sequences by introducing mutations that would be lethal in a centrally-acting ribosome (Aleksashin et al., 2020; Orelle et al., 2015; Schmied et al., 2018). These fully realized orthogonal translation systems offer enormous potential for evolving PTCs capable of new polymerization chemistry *in vivo*.

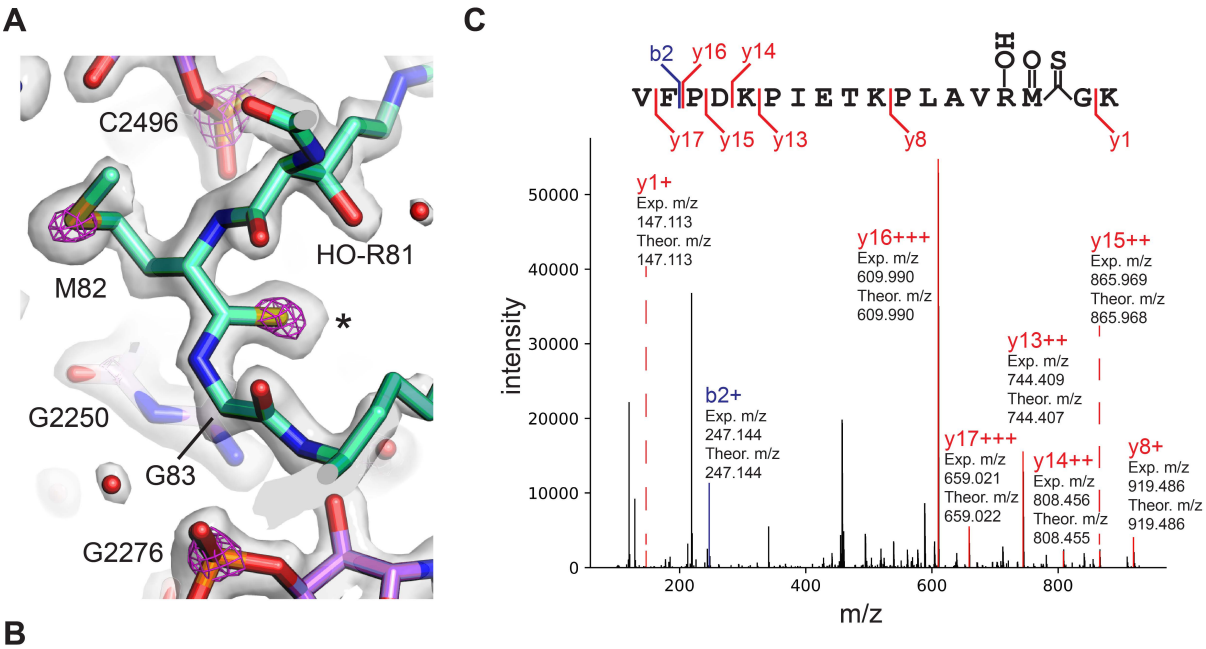
1.6 Understanding ribosome engineering through structure

Any initial efforts to engineer the ribosome should be subjected to a wide array of physical and biochemical probing to reveal how new functions were gained. This endeavour is important from a basic science perspective, where the effect of small mutations on catalysis and specificity can reveal more about how the PTC discriminates, orients, and activates its substrates. Indeed, mutational scanning has been an essential tool for biochemists and structural biologists for decades. Deep analysis of successful PTC mutations is also important for subsequent

engineering steps. Understanding how mutations lead to PTC reorganization and where limitations remain can help collapse the space of possible iterative improvements and accelerate innovation. A valuable approach to understanding PTC engineering is structural biology, especially cryogenic electron microscopy (cryo-EM). Cryo-EM is in the midst of a revolution as improved detectors, techniques, microscopes, and processing strategies have aligned to deliver resolutions often surpassing those from X-ray crystallography (Nogales & Scheres, 2015).

Ribosomes have a long history in the cryo-EM field due to their near-perfect characteristics: large size, asymmetric shape, and high RNA content that deliver excellent contrast and identifiable configurations in micrographs (Brown & Shao, 2018). The first 3D models of the ribosome were generated from early single particle methods (Lake, 1976). Unfortunately, EM was firmly in its 'blobology' era, and detailed structures of the ribosome would come nearly 25 years later after heroic efforts to solve crystal structures by several teams (Ban et al., 1999; Cate et al., 1999; Clemons et al., 1999; Schluenzen et al., 2000). More detailed structures followed of the ribosome in many states of translation (Laurberg et al., 2008; Polikanov et al., 2014; Pulk & Cate, 2013; Schmeing et al., 2009). In the last few years, however, cryo-EM has re-emerged as the critical technique for ribosome structure. Recent advances in sample preparation, data collection, classification, and reconstruction (Dandey et al., 2020; Punjani & Fleet, 2020; S. H. W. Scheres, 2016) have enabled discernment of many separate intermediates of translation at once (Fu et al., 2019; Loveland et al., 2017, 2020). This highlights one of the major advantages of cryo-EM: sample heterogeneity is not necessarily detrimental. Multiple complexes or conformations can simply lead to solving multiple important structures from one experiment, a far cry from the high purity and stability often needed for successful crystallography sessions. Underscoring the power of the technique to understand the chemistry of the ribosome, we recently published the highest resolution ribosome yet achieved at 2 Å. As an example of the high detail achieved, we were able to discover a new post-translational modification in ribosomal protein μL16 from the high resolution density alone (**Figure 1.2**) (Watson et al., 2020). Regularly attaining this level of detail will reveal the precise position of substrates in the PTC and inform engineering efforts to come.

In this work, we first study a ribosome selected for improved incorporation of β-amino acids (Ward et al., 2019), revealing severe assembly defects and complete *in vitro* in activity. Building from structural and biochemical characterization of the poorly functional ribosome, we suggest limitations of current ribosome engineering strategies and possible avenues for improvement. Chapter three describes our efforts to improve ribosome engineering from first principles. We attempt to understand deficiencies in the wild type PTC when non α-amino substrates are bound, with the end goal of identifying more targeted or compensatory mutations to the PTC. Ultimately, this should avoid the assembly problems that arise from large mutations as described in chapter two. Our progress in obtaining high resolution structures of the wild type PTC occupied by non-α-amino monomers is described. Chapter four details progress in restoring or identifying *in vitro* activity in PTC mutant ribosomes. Inspired by the discrepancy in activity of mutant ribosomes *in vivo* versus *in vitro* from chapter two, we explore improved reporter systems that successfully uncover *in vitro* activity in a PTC mutant ribosome for the first time.



B

Peptide	Experimental Mass (Da)	Theoretical Mass (Da)	Mass Diff. (ppm)
VFPDKPITEKPLAVR[O]M[O,O->S]GK	2073.1106	2073.1113	-0.3377
VFPDKPITEKPLAVR[O]	1724.9823	1724.9823	0.0
VFPDKPITEKPLAVR	1708.9865	1708.9874	-0.5266
VFPDKPITEKPLAVRM[O,O->S]GK	2057.1174	2057.1165	0.4375
VFPDKPITEKPLAVR[O]M[O,O]GK (hypothetical)	2073.1106	2073.1290	**-8.8756**

Figure 1.2. High resolution cryo-EM reveals a thioamide modification in uL16. (A) Structural model of thioamide between Met82 and Gly83 in uL16 (mint), with the 50S subunit cryo-EM density map contoured at two levels to highlight sulfur and phosphorus atoms. The lower contour level is shown as a gray surface and the higher contour level is shown as fuchsia mesh. 23S rRNA is shown in purple. Asterisk marks the position of the sulfur in the thiocarbonyl. (B) LC-MS/MS data supporting the presence of a thioamide bond between M82 and G83 of uL16 (Dai et al., 2017). Shown are selected uL16 peptides with designated modifications found in the spectral search and their associated experimental masses, theoretical masses, and mass differences. All peptides were found in multiple fractions and replicates of the experiment. The final row shows a hypothetical peptide identical to the first row, except carrying an oxidation modification instead of O to S replacement. (C) Annotated fragmentation spectra from the LC-MS/MS experiment showing a uL16 peptide with a thioamide bond. Peptide is assigned modifications of: oxidation on M, oxidation on R, and a thiopeptide between M and G. Fragmentation ions are annotated with experimental and theoretical m/z ratios. Adapted from (Watson et al., 2020).

2. Defects in the assembly of ribosomes selected for β-amino acid incorporation

Reproduced with permission from: Ward FR*, Watson ZL*, Ad O, Schepartz A, Cate JHD. Defects in the Assembly of Ribosomes Selected for β-Amino Acid Incorporation. Biochemistry. 2019 Nov 12; 58(45):4494-4504.

* These authors contributed equally

2.1 Introduction

The ribosome is a large molecular machine capable of directing the polymerization of a distinct sequence of amino acids by decoding a messenger RNA (mRNA) template. Its ability to accurately and efficiently select and incorporate the correct monomer from a pool of over 20 substrates makes the ribosome one the most versatile machines for polymer synthesis. No other known methods, from solid-state chemistry to bulk polymerization, can achieve both the specificity and yield of ribosome-catalyzed synthesis, particularly for products longer than ~50 monomers in length (Knight et al., 2015). The ribosome has thus become a target of engineering efforts to eventually develop a ribosome-based platform for the synthesis of new classes of sequence-defined polymers. The large obstacle to these engineering attempts is the several-billion-year optimization of the ribosome, and the rest of the translational apparatus, toward efficient and selective α-amino acid polymerization. New sequence-defined polymerization chemistries will likely require repurposing a suite of enzymes and nucleic acids, including tRNAs, (Katoh & Suga, 2018; Reynolds et al., 2017) aminoacyl tRNA synthetases (aaRSs), (Iqbal et al., 2018; Vargas-Rodriguez et al., 2018) elongation factor Tu (EF-Tu), (Doi et al., 2007; Haruna et al., 2014) in addition to the ribosome itself (Aleksashin et al., 2019; Dedkova et al., 2003, 2012; Melo Czekster et al., 2016; Schmied et al., 2018).

The catalytic core of the ribosome, the peptidyl transferase center (PTC), induces the attack of an A-site aminoacyl-tRNA on the ester bond of the P-site peptidyl-tRNA, transferring and extending the growing polypeptide chain. The PTC is thought to encourage proper polymerization through several mechanisms including orienting the substrates, (Sievers et al., 2004) favoring productive acid-base chemistry, (Polikanov et al., 2014) and protecting the peptidyl-tRNA from dead-end hydrolysis (Martin Schmeing et al., 2005). Like many other enzymes, the PTC undergoes a substantial activating rearrangement upon substrate binding that can accommodate a wide range of monomer side chains. However, the ribosome only weakly catalyzes polymerization of substrates with different backbone chemistries, including peptoids, (Kawakami et al., 2008) N-methyl amino acids, (Subtelny et al., 2008; Jinfan Wang et al., 2014; B. Zhang et al., 2007) D-amino acids, (Englander et al., 2015) aromatic foldamers, (Rogers et al., 2018; Tsiamantas et al., 2019) aramids, (Ad et al., 2019) malonates, (Ad et al., 2019) and β^3-amino acids (Fujino et al., 2016).

β^3-amino acids (referred to here as β-amino acids) are a useful model substrate for novel ribosome polymerization reactions as they introduce new challenges for catalysis while closely resembling the natural α-amino substrates. The extra methylene between the alpha carbon and carboxylic acid groups adds more bulk, rotational degrees of freedom, and diminishes reactivity, (Edsall & Blanchard, 1933) but the nucleophile remains an amino group, as in α-amino acids. The extra methylene group confers advantages to β-amino acid oligomers as peptidomimetics, as they form stable folds, display reduced immunogenicity, and are more resistant to proteolysis

(Daniels et al., 2007; Geueke et al., 2006; Guichard et al., 2000). Importantly, despite reacting poorly in the PTC, under certain conditions β-amino acids can be consecutively polymerized by the ribosome *in vitro* (Katoh & Suga, 2018). Their similarities to natural monomers have made β-amino acids an attractive substrate target for ribosome engineering.

Attempts to engineer new polymerization chemistry into the ribosome have understandably focused on mutations to the PTC. Aggressive mutations of 10+ bases in the PTC have generated ribosomes that can better translate D-amino acids, (Dedkova et al., 2003) dipeptide substrates, (Maini, Dedkova, et al., 2015) polyproline motifs, (Schmied et al., 2018) and β-amino acids (Dedkova et al., 2012; Dedkova & Hecht, 2019; Maini, Chowdhury, et al., 2015; Maini et al., 2013; Melo Czekster et al., 2016). These mutations have been proposed to solve the space, orientation, and reactivity problems inherent to many backbone-modified substrates, but no mechanistic studies of these improved ribosomes have been published. Indeed, most of these mutant ribosomes have only been tested *in vivo* or in lysate-based cell free systems with wild-type ribosomes present. Although these ribosomes represent exciting steps toward new sequence-defined polymerization capabilities, a detailed understanding of how PTC mutations help widen the substrate scope of the ribosome is still lacking. In particular, the PTC is highly conserved across all domains of life and large-scale mutations may have knock-on effects that limit ribosome utility. Therefore, these engineering problems must always be evaluated with a discerning focus on the mechanistic implications of such changes, in order to better understand how these PTC alterations affect ribosome activity.

Here were present structural and biochemical characterization of a β-amino acid translating ribosome, P7A7, discovered via an *in vivo* selection for β-puromycin incorporation (Melo Czekster et al., 2016). The P7A7 PTC carries twelve mutations over two regions near the tRNA A site and exit tunnel. We show that purified P7A7 ribosomes are inactive during *in vitro* translation and do not form stable 70S complexes. A cryo-electron microscopy (cryo-EM) structure of P7A7 50S ribosomal subunits reveals substantial disordering of the PTC and nearby inter-subunit bridge helices when compared with an equivalent wild-type (WT) 50S structure. Analysis of the P7A7 map reveals a depletion of late-assembling ribosomal proteins, which was confirmed using tandem mass tag (TMT) relative quantitation. When compared with existing studies of 50S ribosomal subunit assembly, P7A7 appears trapped as a late assembly intermediate, explaining its poor activity. Our results suggest that the radical PTC mutations often seen in engineered ribosomes may have unintended effects on ribosome assembly and stability that limit the utility of many of these variants.

2.2 Results

Prior analysis of the P7A7 ribosome and its precursor, 040329, was limited to *in vivo* or lysate based characterization in the presence of wild-type ribosomes (Dedkova et al., 2012; Melo Czekster et al., 2016). The complexity of cell lysate based translation and the presence of background WT ribosomes, which can be active in β-amino acid incorporation, (Fujino et al., 2016; Katoh & Suga, 2018) makes it challenging to accurately understand P7A7's altered substrate scope, necessitating the development of a pure P7A7 translation system. Initial attempts at creating *rrn* knockout 'Squires' strains carrying a P7A7-23S rRNA encoding variant of the *rrnb* operon on a plasmid failed, indicating that P7A7 cannot support cellular growth on its own (Asai et al., 1999). While the 12 PTC mutations of P7A7 likely diminish its efficiency at natural protein synthesis, it had been observed to synthesize full-length dihydrofolate reductase *in vivo*,

suggesting a level of translational competency amenable to study. We assayed the levels of P7A7 50S ribosomal subunits in polysome fractions from mid-log phase *E. coli* Mach1 cells using a semi-quantitative reverse transcription-PCR (RT-PCR) assay, recognizing an MS2 RNA tag grafted onto helix 98 of the 23S rRNA (Youngman & Green, 2005). P7A7 ribosomes are depleted, but not absent, in both 70S and polysome fractions (**Figure 2.1A; Figure 2.2A)**, indicating that P7A7 50S subunits are deficient in subunit association**.** To remove background WT ribosomes and develop an isolated P7A7 translation system, we purified P7A7 50S subunits using affinity chromatography of MBP-phage MS2 coat protein fusion bound to a tag on helix 98 **(Figure 2.2B).**

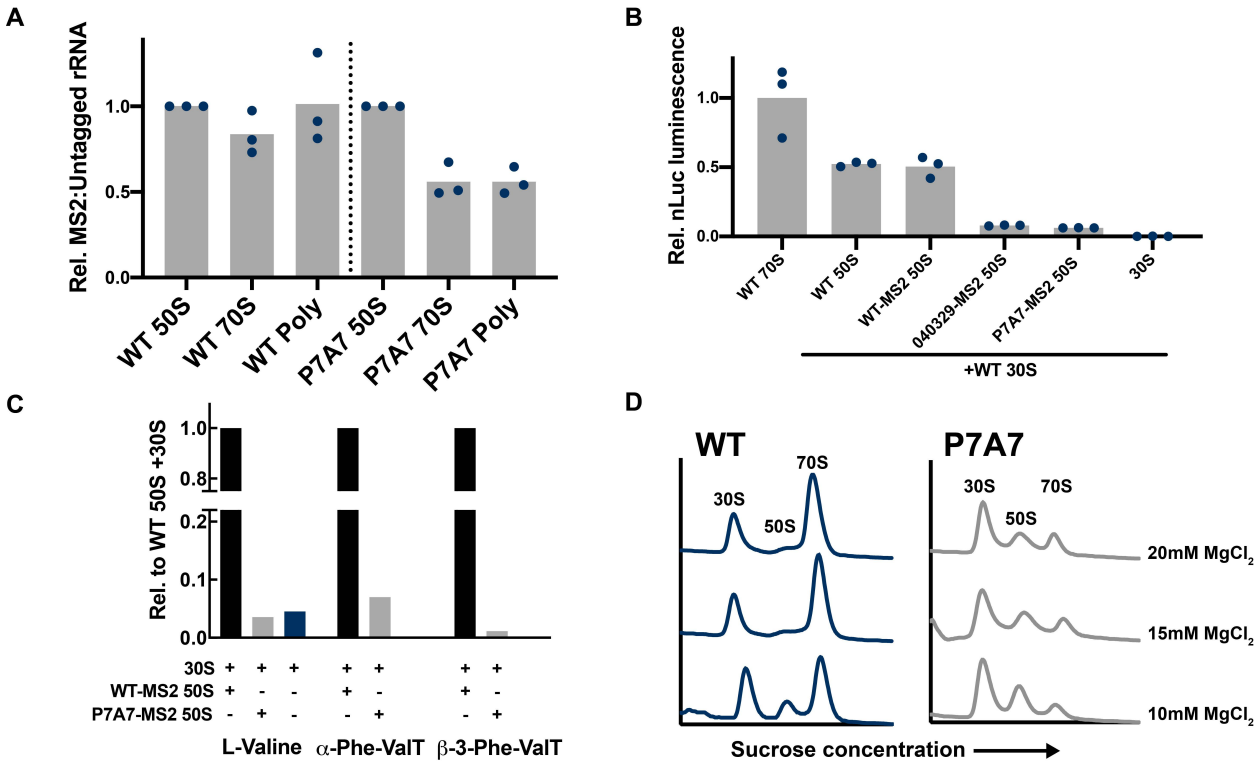


Figure 2.1. WT and P7A7 ribosome activity. A) WT-MS2:WT or P7A7-MS2:WT rRNA ratios in different polysome profile fractions, normalized to 50S ratio per replicate. Relative quantities of rRNA measured by RT-PCR over the MS2-tag region and quantification of bands resolved by polyacrylamide gel electrophoresis. Bars represent mean values. B) Nanoluciferase expression measured from PURExpress Δ-ribosome *in vitro* translation reactions with different purified ribosomes. Values reported relative to WT 70S. '50S' samples represent individually purified 50S and 30S subunits combined at initiation of reaction. Bars represent means. C) LC-MS monitored MXFDYKDDDDK peptide synthesis from PURExpress Δ-ribosome-tRNA-amino-acid *in vitro* translation reactions with varying purified ribosomes. The tRNA for position X was either Val-$tRNA_{val}$, Phe-$tRNA_{val}$ (flexizyme charged), or β-3-Phe-$tRNA_{val}$ (flexizyme charged). Single replicates are shown for each condition. D) *In vitro* apo-ribosome association at differing Mg^{2+} concentrations, measured by sucrose gradient centrifugation and fractionation for WT 50S + WT 30S and P7A7 50S + WT 30S ribosomes. Representative gradients are shown.

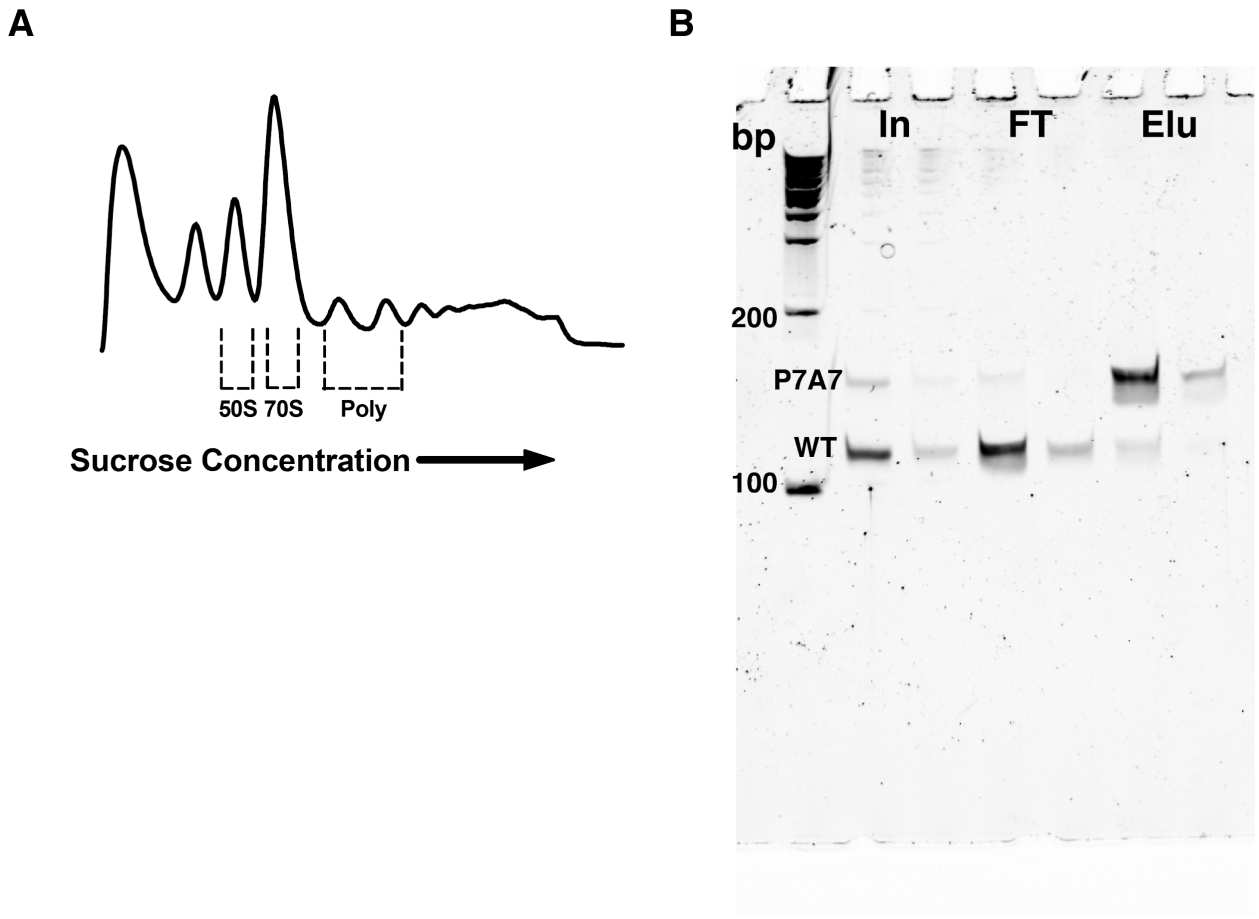


Figure 2.2.2. Distribution of P7A7 50S subunits in polysomes and P7A7 purification.
A) Representative sucrose gradient fractionation of polysomes showing fractions collected for analysis in **Figure 2.1**. The 254 nm absorbance trace is shown. B) P7A7 purification assayed by semi-quantitative RT-PCR over the MS2-tag of isolated rRNA. Input (In), flow through (FT), and elution (Elu) fractions analyzed via two PCRs for each sample,with 0.2 ng and 0.04 ng of input cDNA for the stronger and weaker lanes, respectively.

We tested the activity of purified P7A7 50S subunits with WT 30S subunits in the PURExpress Δ-ribosome *in vitro* transcription-translation system (Shimizu et al., 2001). Translation of a standard α-amino acid nanoluciferase reporter with P7A7 was very low relative to untagged and MS2-tagged WT 50S subunits **(Figure 2.1B)**, especially considering observable WT contamination in MS2-tag purifications **(Figure 2.2B).** Similar results were found with the earlier-identified 040329 ribosome **(Figure 2.1B)**. While *in vitro* translation is known to be far less efficient than seen *in vivo*, the minute activity of purified P7A7 ribosomes was unexpected in light of P7A7's DHFR synthesis (Melo Czekster et al., 2016) and polysome occupancy *in vivo*. After additional purification of P7A7 50S subunits to remove WT contamination, short FLAG-derived α- and β-amino acid containing peptides could not be detected above background levels **(Figure 2.1C; Figure 2.3).** Because subunit association is a necessary step during translation initiation, we compared the magnesium-dependent subunit association of P7A7 and WT ribosomes. Apo-P7A7 ribosomes require high concentrations of magnesium to form 70S complexes, suggesting a defect in association that could impair translation **(Figure 2.1D).**

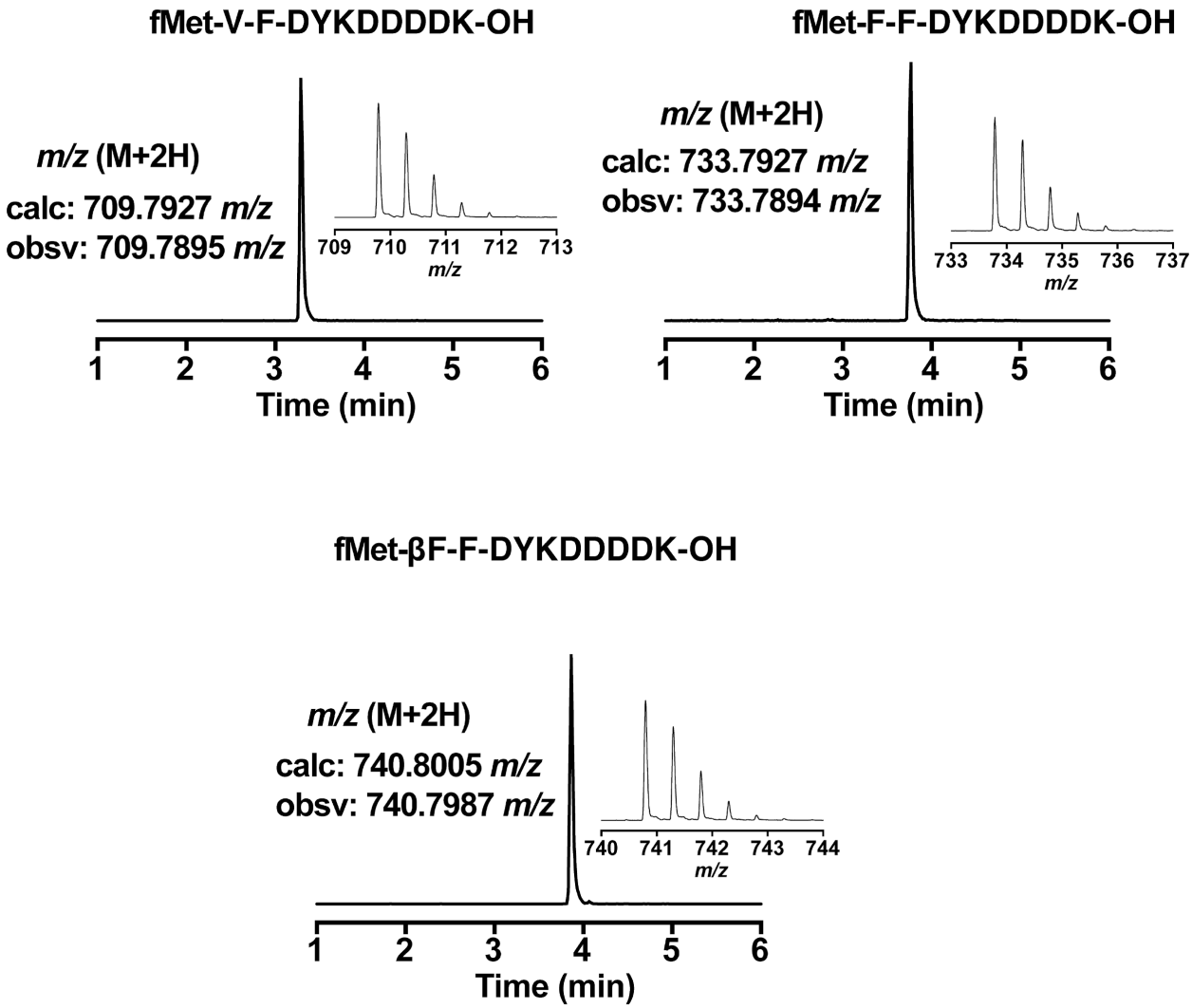


Figure 2.3. Extracted ion chromatograms and mass spectra of FLAG-derived peptides. Representative liquid chromatography and mass spectrometry results for the *in vitro* translation experiments in **Figure 2.2.1C**.

To better understand the basis for P7A7's low activity, we solved the cryo-EM structure of its large subunit, alongside a WT 50S subunit as a control. Images were collected on a 300-kV Titan Krios microscope with a K2 direct electron detector, and processed with RELION (Sjors H. W. Scheres, 2012) and cryoSPARC (Punjani et al., 2017). Global resolutions achieved for the two maps were 3.11 Å and 3.20 Å, respectively (**Figure 2.4; Table 2.1**). Intriguingly, although both maps have substantial regions of the 50S subunit at high resolution, large portions of the ribosome that are resolved in the WT structure are absent in the P7A7 map **(Figure 2.5)**. Heterogeneous ab initio reconstructions of the P7A7 50S subunit revealed two classes of particles that represent the large subunit with two different types of disorder. The major class lacks density for ribosomal proteins uL6, uL10, uL11, uL16, bL33, and bL35 (excluding proteins bL31 and bL36 which are known to disassociate during ribosome isolation), (Nikolay et al., 2018) as well as multiple stretches of the 23S rRNA comprising the PTC, helices H44, H71, H89-H93, and the end of H39. The smaller class further lacks the central protuberance entirely **(Figure 2.6).**

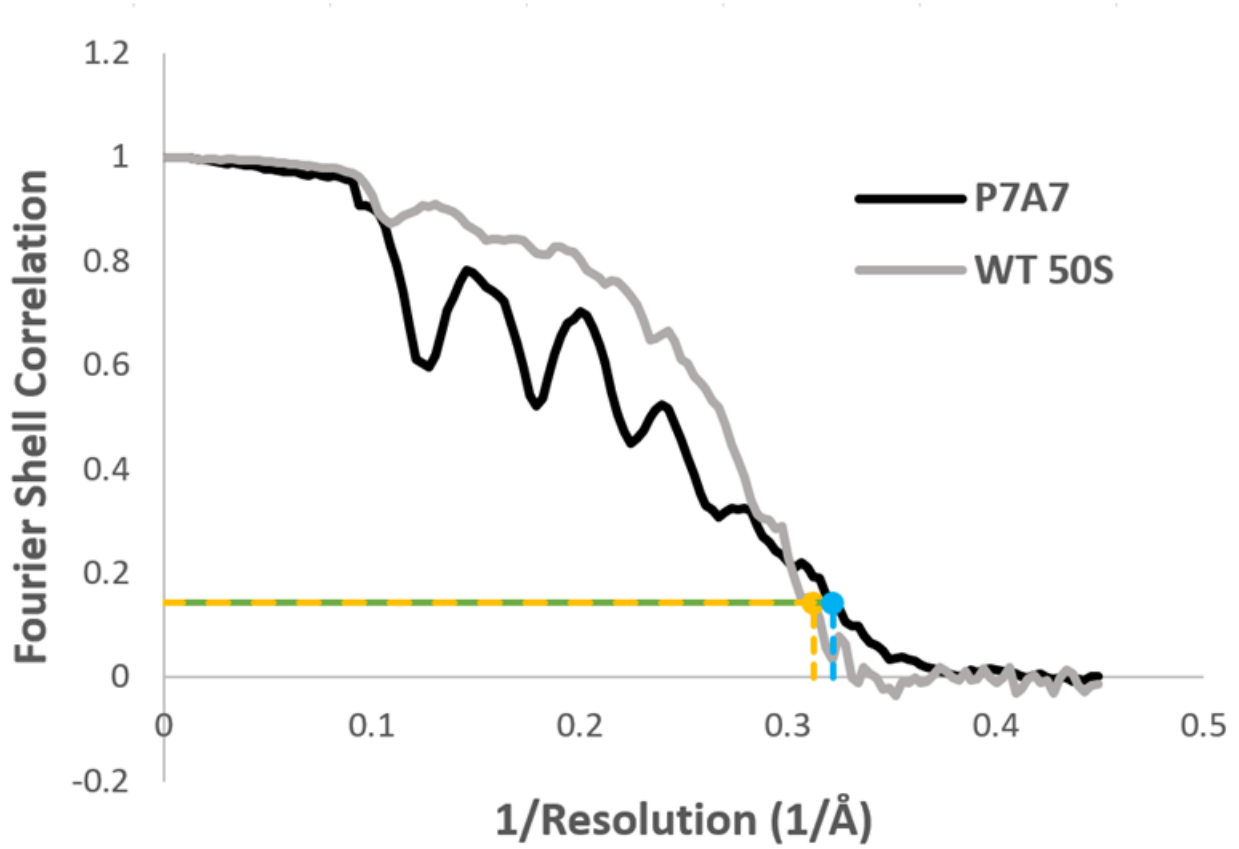


Figure 2.4. Fourier shell correlation (FSC) curves for P7A7 and WT 50S cryo-EM maps. Oscillations in the FSC for P7A7 are an artifact of the focal pair method of data collection, caused by zeroes in the CTF at the predominant defocus value. The same effect is not observed for the WT 50S subunit map due to different targeting of the near-focus value during the data collection session. The gold-standard FSC cutoff value for global resolution (0.143) is marked in gold for WT and blue for P7A7.

Table 2.1. Cryo-EM data collection and refinement information.

	P7A7 large subunit	WT 50S
Microscope	FEI Titan Krios	
Accelerating Voltage (kV)	300	
Detector	GATAN K2 Summit	
Spherical Aberration (mm)	2.7	
Magnification	215,000x	
Defocus targets (μm)	-0.3 and -3.5	-0.5 and -3.6
Micrographs (defocus pairs)	1402	1442

Particles picked	92,647	109,802
Particles refined	37,609	84,372
Resolution achieved (Å)	3.11	3.20

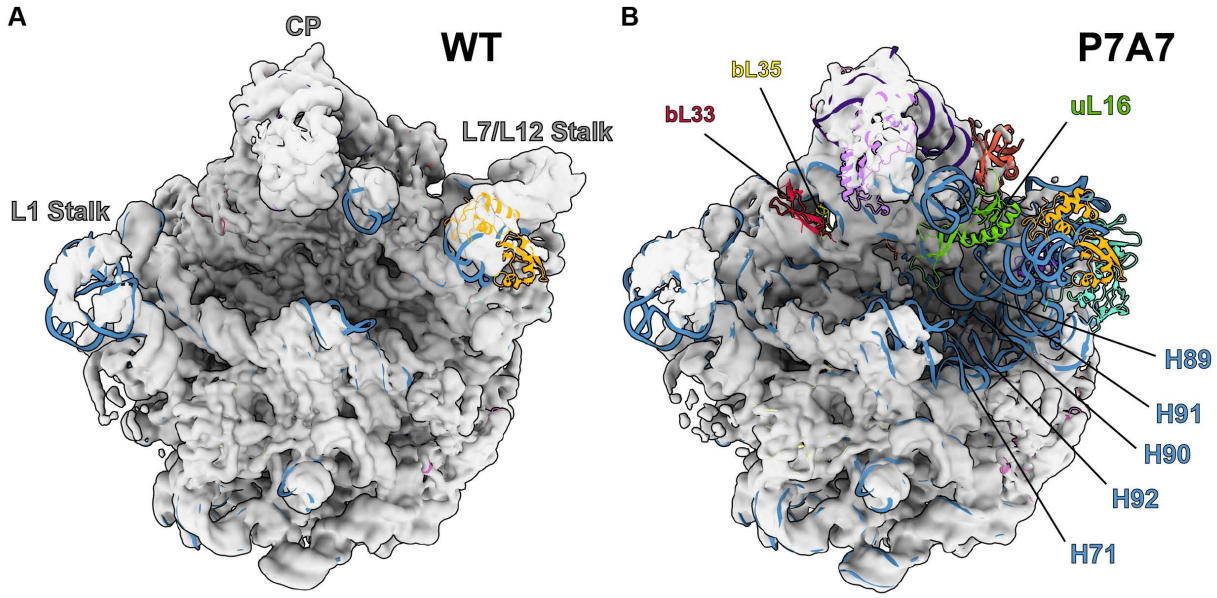


Figure 2.5. Comparison of WT 50S (A) and P7A7 (B) large subunit cryo-EM maps. Maps are low-pass filtered to 8 Å for clarity of structural features, and 50S coordinates from PDB 4YBB(Noeske et al., 2015) are docked to identify missing components. The central protuberance (CP), L1 Stalk, and L7/L12 Stalk are identified on the WT structure, and features confirmed to be missing from P7A7 by combined cryo-EM and LC-MS/MS analysis are labeled on the P7A7 structure.

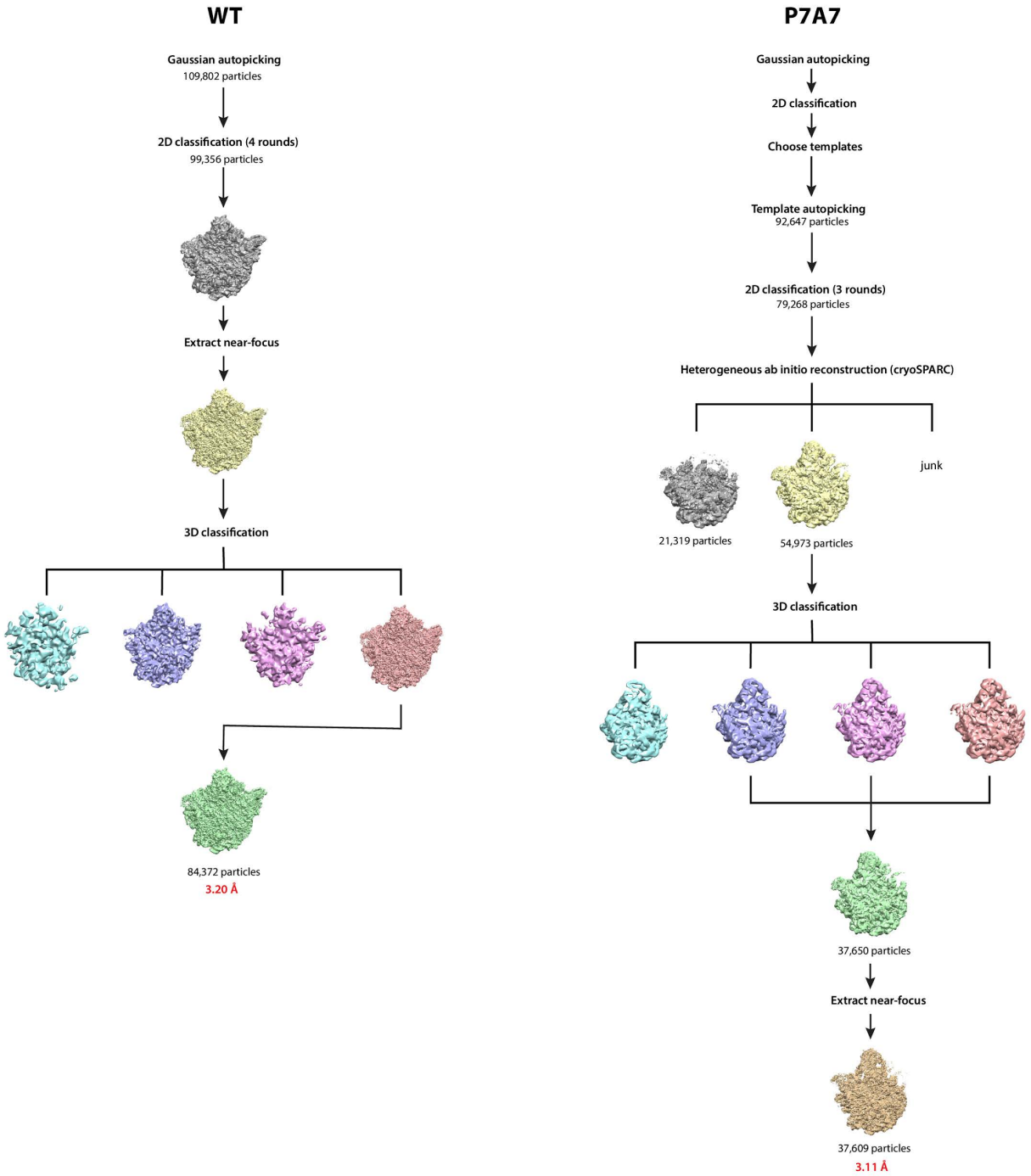


Figure 2.6. Cryo-EM data processing workflow for P7A7 and WT 50S subunits. Most steps were completed in RELION, otherwise in cryoSPARC where noted.The smaller class of P7A7 subunits lacking the CP is shown in grey.

Because cryo-EM often cannot resolve whether features of the ribosome are missing due to their absence from the complex or due to flexibility, we also performed relative quantitation of ribosomal proteins (RPs) using liquid chromatography with tandem mass tag (TMT) spectrometry (LC-MS/MS), with TMT labeling on WT and P7A7 50S subunits (A. Thompson et

al., 2003). Normalized ratios of P7A7:WT RP levels generally matched the structural observations. P7A7 is depleted of RPs uL16, bL33, and bL35 **(Figure 2.7).** We were able to detect stoichiometric levels of RPs uL6, uL10, and uL11, whose cryo-EM density is often weak due to motion in the large subunit arms.(Davis et al., 2016; Nikolay et al., 2018)

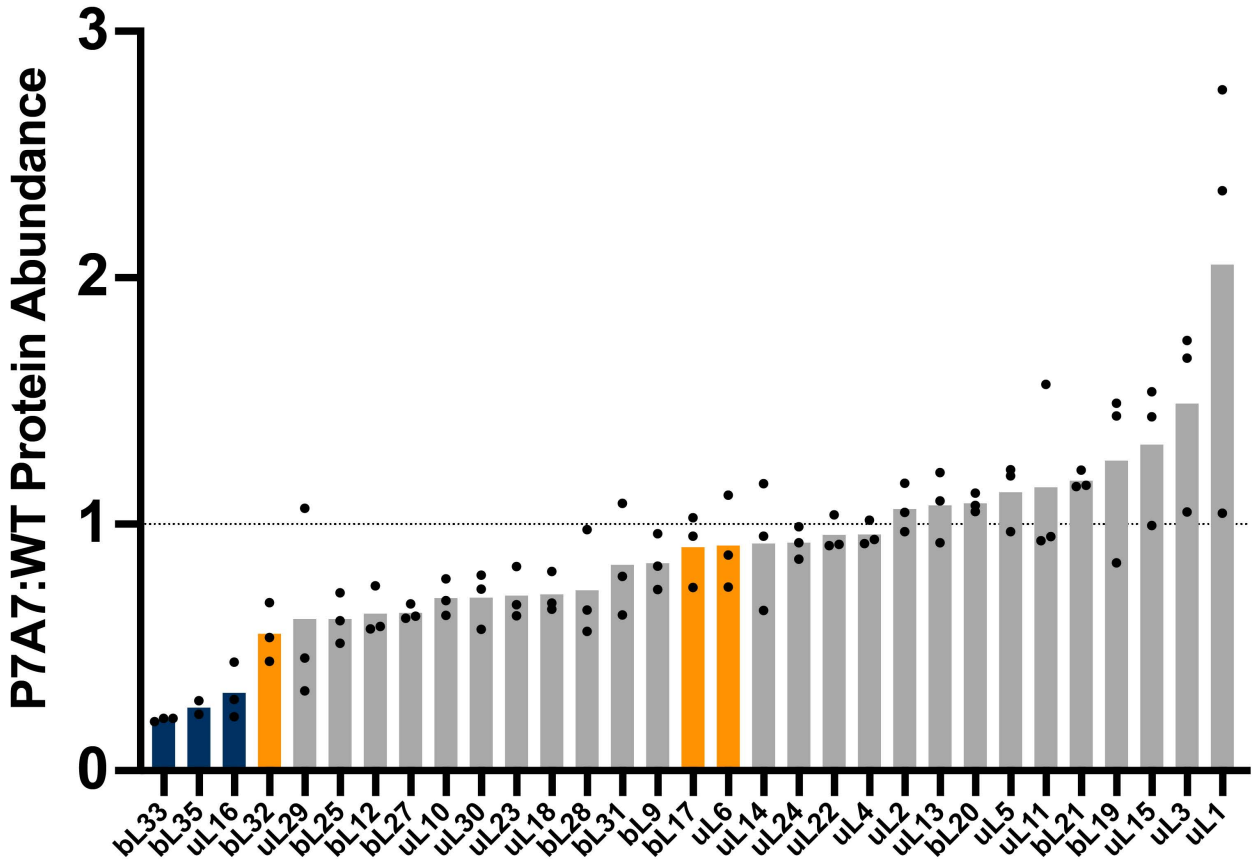


Figure 2.7. 50S ribosomal protein ratios detected by LC-MS/MS and TMT relative quantitation between P7A7 and WT ribosomes. Bars represent means and are colored by absence in final assembly states (E4 and E5) of Davis and Tan, *et al.* (orange) or absence in both Davis and Tan, *et al.* and Nikolay, *et al.* states (IV and V) (blue).(Davis et al., 2016; Nikolay et al., 2018)

In the vicinity of the P7A7 mutations in the PTC, we observe a stark contrast between well-ordered regions seen in the WT structure, with features like base stacking clearly resolved, and adjacent regions where the density is mainly noise **(Figure 2.8).** Comparison of these disordered regions in the P7A7 50S subunit map with the corresponding region in the WT PTC shows several interactions that would be perturbed by some of the mutations **(Figure 2.9)**, (Noeske et al., 2015) and may be responsible for the structural defects we observe. In particular, two base triples–the first between C2063, A2450, and C2501, and the second between C2507, G2553, and G2582–would not be able to form because of the C-A and C-U mutations at positions 2063 and 2507, respectively. Other disrupted interactions include the G2057-C2611 and G2505-C2610 base pairs **(Figure 5)**.

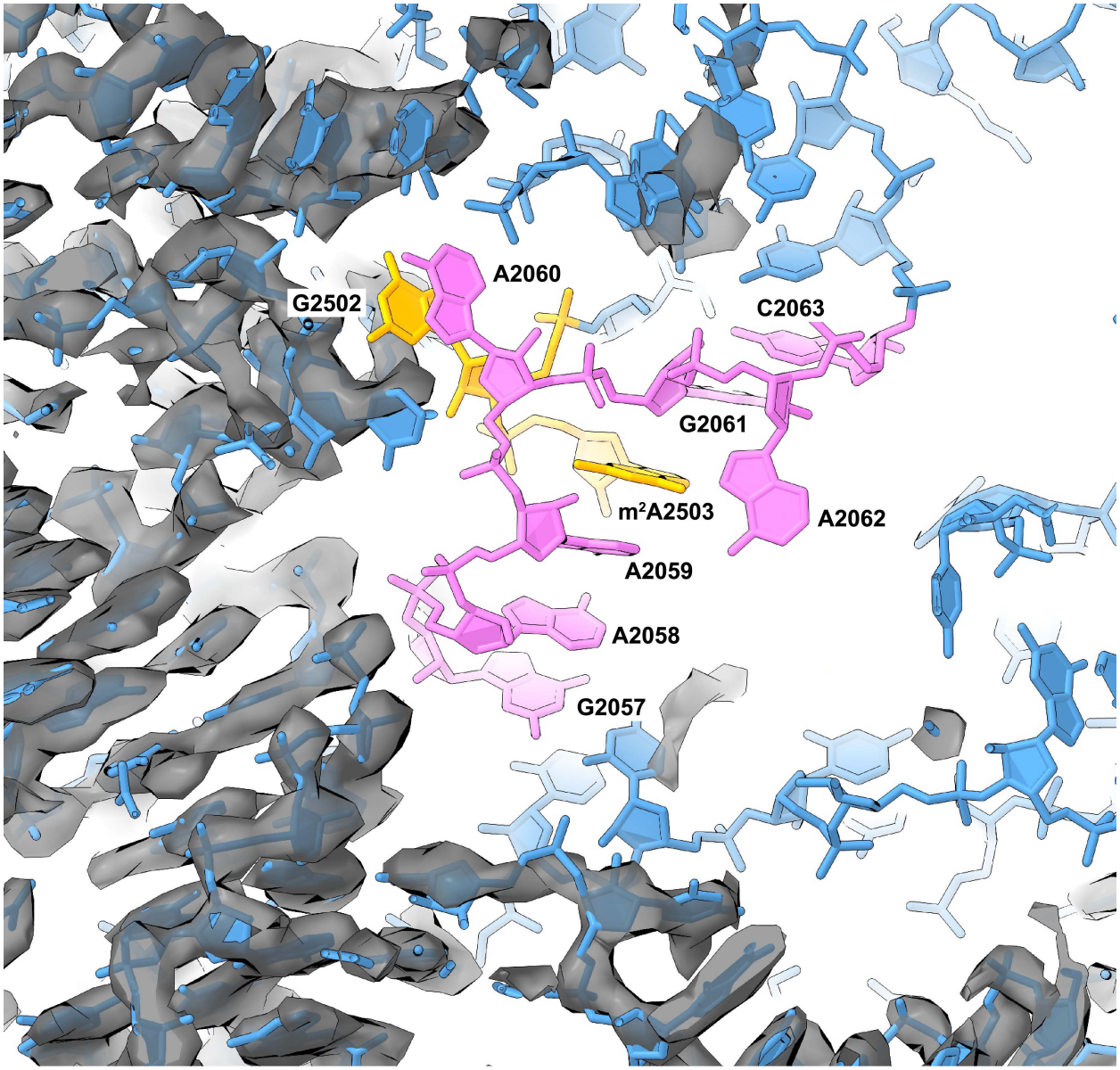


Figure 2.8. Disorder in the PTC of P7A7 50S subunits. Model of the wild-type 23S rRNA in the vicinity of the PTC superimposed on the map of the P7A7 50S subunit. Regions that are well ordered include helices H26, H35, and several loops from domain II. Mutated PTC residues 2057-2063 and 2502-2503 in P7A7 are depicted in pink and gold, respectively. The "hide dust" feature in ChimeraX(Goddard et al., 2018) was used to remove noise for clarity.

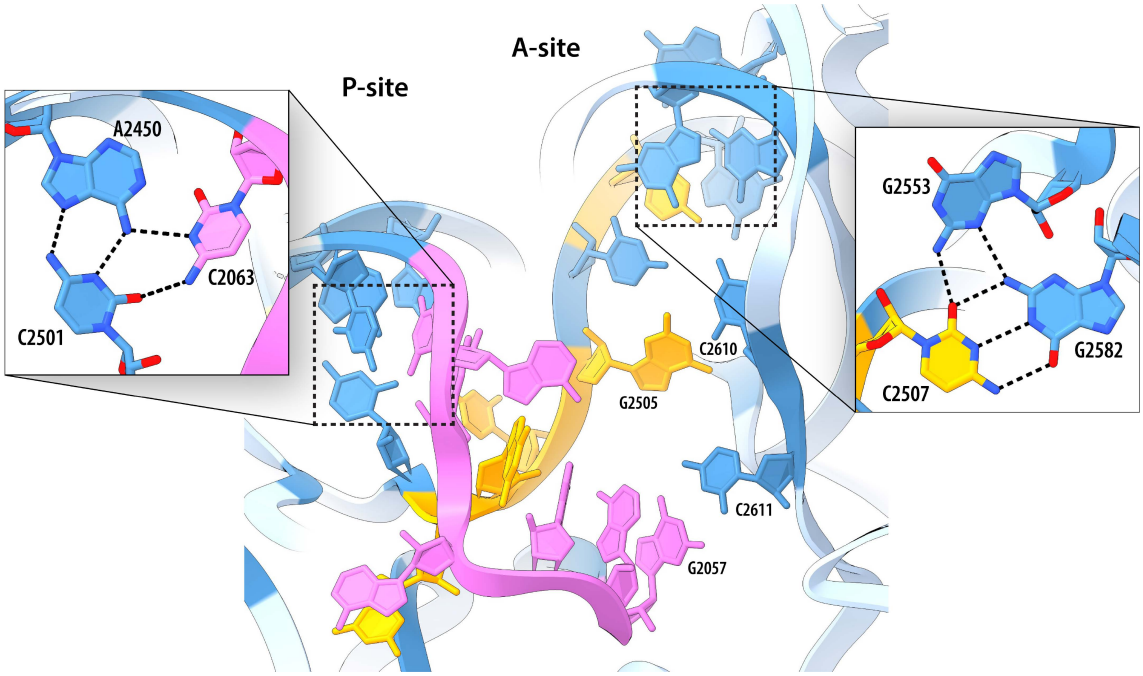


Figure 2.9. WT 23S rRNA residues directly affected by P7A7 mutations. Coordinates shown are from PDB 4YBB.(Noeske et al., 2015) Mutated residues 2057-2063 and 2502-2505, 2507 in P7A7 are depicted in pink and gold, respectively. Unaltered bases within 3 Å of mutations are shown in darker blue, and those further away are lighter blue. Close-ups highlight base triples that would be disrupted by mutations, and two disrupted base pairs are also labeled.

Low-pass filtering of the P7A7 50S subunit map allowed for some interpretation of broader features of its structure that show poor connectivity and are more difficult to discern in the high resolution map. Notably, there appears to be some helical RNA density spanning the PTC, which does not fit with the WT model **(Figure 2.10).** This connectivity is strikingly similar to the non-native PTC organization seen in late intermediates of the 50S assembly pathway both *in vivo* and *vitro* (Davis et al., 2016; Nikolay et al., 2018). While ribosome biogenesis is a complex process, a general order of rRNA helix formation and ribosomal protein association is known (Chen et al., 2012; Herold & Nierhaus, 1987). Despite multiple assembly routes, resulting in a spectrum of assembly intermediates, he PTC is usually observed as the last-assembling 50S motif.

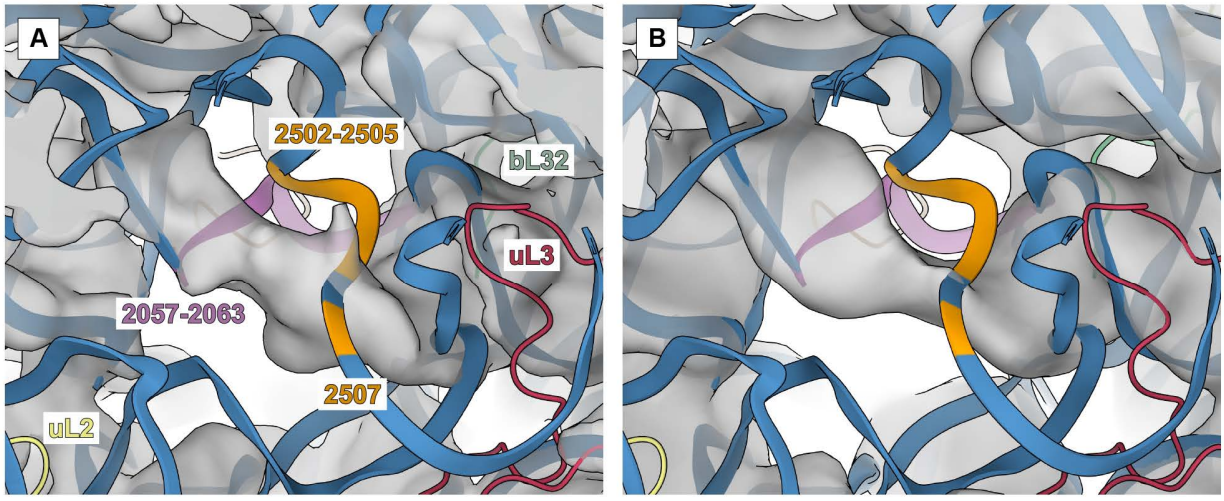


Figure 2.10. Comparison of PTC in the P7A7 map (A) and an assembly intermediate, E4, determined by Davis and Tan, *et al*. (B). Both maps show helical density that does not appear in fully assembled WT structures. P7A7 mutations regions 2057-2063 and 2502-2505, 2507 are colored in pink and orange, respectively. Other 23S rRNA regions are in blue, and nearby proteins uL2, uL3, and bL32 are yellow, red, and green, respectively.

To further explore the relationship of the observed P7A7 structure with assembly pathways, we carried out a detailed tabulation of structural features across the P7A7 and WT 50S subunit EM maps and compared them to known assembly intermediates, including states IV and V from characterization of *in vitro* reconstitution (Nikolay et al., 2018) and states E4 and E5 from *in vivo* intermediates caused by bL17 knockdown (Davis et al., 2016). Using a previously developed calculation method (Davis et al., 2016) we generated a comprehensive list of the cryo-EM occupancies for every ribosomal protein and rRNA helix across all 6 structures **(Figure 2.11A).** Hierarchical clustering grouped P7A7 alongside other PTC-unfolded states, with P7A7 most closely aligned to E4. However, ribosomal protein occupancy, particularly the absence of bL17 and bL32 from E4, revealed P7A7 as a distinct intermediate. Our TMT quantitation results **(Figure 2.7)** correlate well with these analyses. Nevertheless, the large-scale organization of several intermediates is quite similar despite the different methods of assembly perturbation, suggesting that a PTC folding defect is a common roadblock in 50S ribosomal subunit maturation. Interestingly, careful analysis of P7A7's migration on a sucrose gradient reveals that it sediments slower than mature WT 50S subunits, similar to the immature 48S particle seen during *in vitro* reconstitution experiments (Nikolay et al., 2018) **(Figure 2.11B).** Together, our structure, mass spectrometry, and biochemical results suggest that P7A7's minimal activity *in vitro* stems primarily from errors in assembly, leaving the ribosome without a mature PTC.

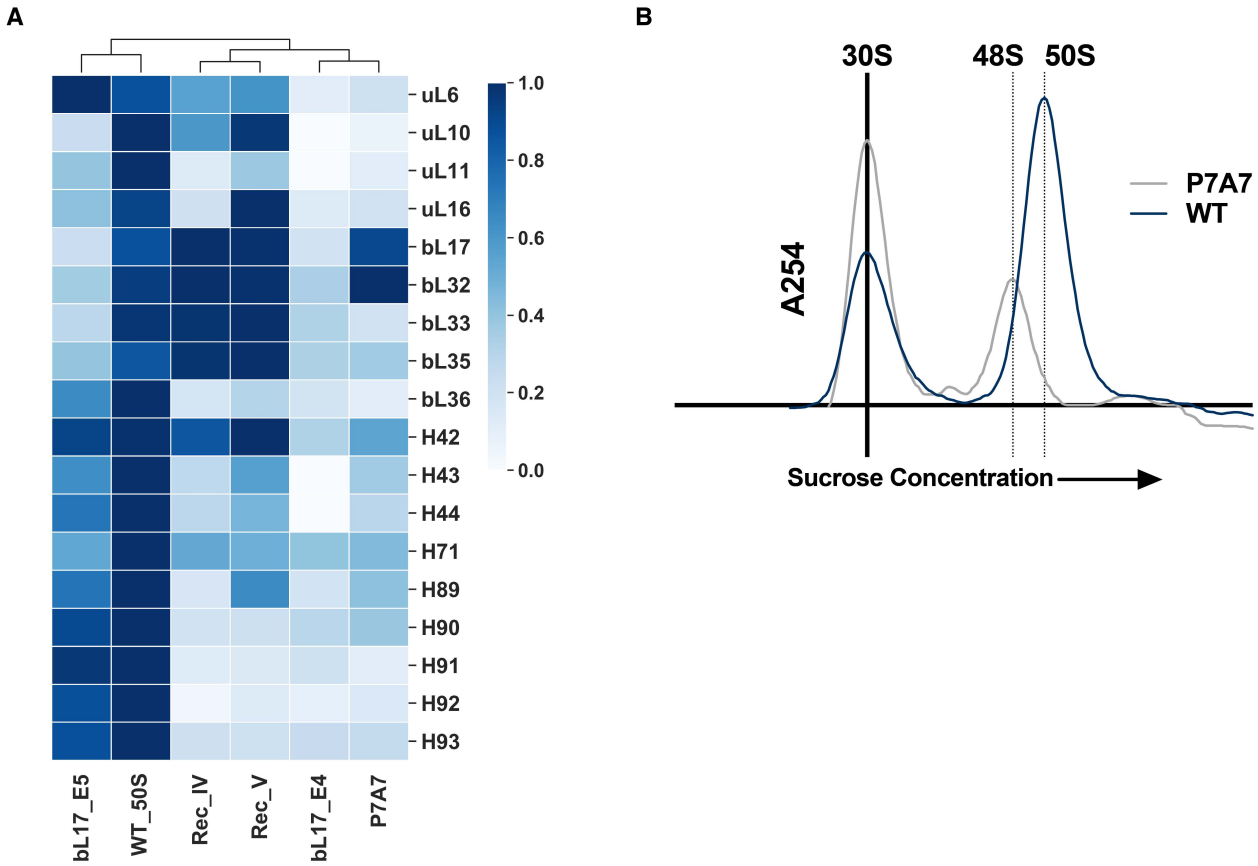


Figure 2.11. Analysis of P7A7 Assembly. A) Heatmap of 50S subunit rRNA helices and protein occupancy across cryo-EM structures of several assembly intermediates, using the method of Davis and Tan, *et al.(Davis et al., 2016)* The heatmap is hierarchically clustered across assembly intermediates. Universally present features have been removed for clarity. B) Sucrose gradient fractionation of MS2-tag purified WT and P7A7 50S subunits run with WT 30S subunits as a standard. The 254 nm absorbance traces are shown, and the estimated sedimentation coefficient for P7A7 is marked.

2.3 Discussion

Ribosome engineering holds promise to expand efficient template-directed polymer synthesis beyond traditional peptide chemistry, unlocking entirely new classes of materials. Toward this end, there have been several attempts at mutating the ribosome PTC to catalyze the synthesis of difficult sequences or accept new substrates, demonstrating some level of success in expanding the chemistry allowed by the ribosome. Many of these ribosomes, including P7A7 studied here, carry a large number of mutations to highly conserved bases in the PTC and nearby exit tunnel. These base mutations are necessary to overcome the strong optimization of the WT PTC toward natural protein synthesis. However, conservation of nucleotides within the PTC is likely not for catalytic ability alone. Ribosome assembly, and particularly the folding of the protein-sparse PTC, is also dependent on rRNA sequence, presenting a multi-parameter optimization problem when engineering ribosomes. While there is evidence that the ribosome possesses an innate tolerance for mutations to the PTC, even to highly conserved bases, (Rakauskaitė & Dinman, 2011-5) many engineered ribosomes likely possess assembly defects that limit their functionality, even if they perform better than WT ribosomes in certain challenging polymerization reactions. In the case of P7A7 ribosomes, the large number of

mutations exceeds the plasticity of the PTC. Further confounding engineering efforts, RNA in general shows less structural stability *in vitro*, (Leamy et al., 2016/ed) possibly exacerbating assembly defects to the point of complete inactivity. These challenges have prevented detailed study of altered-substrate ribosomes *in vitro*, an important step in understanding how new polymerization chemistries are accessed in a remodeled PTC.

A major confounding factor hampering mechanistic analysis of ribosome activity is the large difference between *in vitro* environments and the cellular milieu. Among the chemical and physical differences between these environments, crowding and divalent salt concentration are commonly implicated for their large effects on RNA structure (Leamy et al., 2016/ed). Success in improving conditions for RNA folding by mimicking a crowded macromolecular environment (up to 300 mg/mL (Zimmerman & Trach, 1991)) provides evidence for the importance of excluded volume effects (Kilburn et al., 2010; Strulson et al., 2013; Tyrrell et al., 2013). Notably, *in vitro* translation reactions have been improved by optimizing molecular crowding, reducing agents, and ionic concentrations (Fritz et al., 2015; Jelenc & Kurland, 1979-7; Jewett & Swartz, 2004). Nevertheless, even these reconstituted *in vitro* translation systems are several orders of magnitude less complex than the bacterial cytosol. Missing are many small molecules and proteins that can interact with RNA to ensure functional folding. Even cell-lysate based *in vitro* translation systems used to assay mutant ribosomes still include molecular crowding agents (Dedkova et al., 2012; Maini, Chowdhury, et al., 2015; Maini, Dedkova, et al., 2015; Maini et al., 2013). Additionally, the use of cell lysates suffers from the inability to control the steps of translation critical for determining the activity of mutant or WT ribosomes. For example, WT ribosomes weakly incorporate β-amino acids in cell lysate-based experiments, but are functional in fully reconstituted *in vitro* translation systems with high EF-Tu concentrations, optimized tRNAs, and without competing termination reactions (Fujino et al., 2016; Katoh & Suga, 2018).

There are many plausible hypotheses that explain the effect of PTC mutations on ribosome substrate specificity, including opening up space for bulkier groups (as with β-amino acids and dipeptide substrates), properly orienting flexible substrates (such as β-amino acids), and rearranging suboptimal nucleophiles (i.e. poly-proline stretches, D-amino acids). These mechanistic questions will remain challenging without interpretable structures of these engineered ribosomes and a detailed understanding of conformational coupling of ribosome activities. The general disorder of P7A7's PTC **(Figure 2.8),** in contrast with its *in vivo* activity **(Figure 2.1A),** suggests that disruption of rigid structures in the PTC may be a contributing mechanism to its acceptance of β-amino acids. It is possible that the PTC of P7A7 is more tolerant of a larger β-substrate only because its flexibility allows an accommodating rearrangement. Analogous promiscuity has been observed in protein enzymes, where flexibility in the active site is correlated with wider substrate specificity (Ekroos & Sjogren, 2006; Gatti‐Lafranconi & Hollfelder, 2013; Skopalík et al., 2008). Furthermore, the ribosome is often described as an "entropy trap," in that proper substrate positioning is the primary contributor to highly efficient catalysis (Sievers et al., 2004). The main barrier to β-amino acid polymerization on the ribosome would then be due to improper geometric constraints for β-substrates and not their altered reactivity. An induced-fit mechanism of PTC catalysis has been proposed to be the source of the ribosome's side-chain promiscuity, (Lehmann, 2017; Martin Schmeing et al., 2005) albeit with less disorder than observed in P7A7. As the mutations in P7A7 lead to a less ordered, more flexible PTC fold, we can propose a model in which the additional flexibility confers additional substrate promiscuity. However, this comes at a large cost to overall catalytic efficiency.

Stable RNA folding is particularly important for the PTC, which has a high RNA:protein ratio and folds late in the 50S subunit maturation process (Jomaa et al., 2014; Li et al., 2013; Polacek & Mankin, 2005). With no ribosomal proteins to chaperone or stabilize assembly, the PTC is likely more sensitive to both mutational and environmental disruptions. High magnesium concentrations are often used to stabilize RNA conformations *in vitro*, presumably acting as a stand-in for missing components of the cellular environment. However, *in vivo* RNA structure probing has shown that magnesium can be a poor substitute (Tyrrell et al., 2013). Components of the cytosol may keep the highly mutated P7A7 somewhat functional, but extraction to an *in vitro* environment leads to a nonfunctional PTC. Of particular note are two chaperones, ObgE and EngA, that are proposed to interact with the PTC during assembly and may help direct PTC folding away from inactive intermediates, such as the non-native fold seen in P7A7 **(Figure 2.10)** (Feng et al., 2014; Nikolay et al., 2018; X. Zhang et al., 2014). Thus, due to an unstable PTC, P7A7 ribosomes are not suited for detailed *in vitro* study of improved β-amino acid polymerization reactions by the ribosome. Future efforts toward an understanding could include optimizing a cytosol-mimicking buffer for purification and analysis or mapping *in vivo* RNA structure of P7A7 ribosomes to assess whether it possesses a more stable PTC in cells.

The assembly issues with P7A7 suggest that future ribosome engineering efforts may need to consider three layers of optimization. First, the ribosome must be directed toward a novel function, be it antibiotic resistance, translation efficiency, or new polymerization chemistry. Secondly, the engineered ribosome must interact safely with the rest of the cell, avoiding dominant lethal phenotypes by inhibition of overall translation. Finally, ribosomes must productively assemble and remain stable when deployed for synthesis or study. These three goals may often be in conflict, with the necessary mutation of highly conserved rRNA bases towards new ribosome function threatening to derail cell survival or ribosome assembly. To solve part of this problem, much effort has gone into developing orthogonal ribosomes that do not interfere with normal translation. Mutations to anti-Shine-Dalgarno sequences (Chubiz & Rao, 2008; Hui et al., 1987; Rackham & Chin, 2005) and tethering of ribosomal subunits (Fried et al., 2015; Orelle et al., 2015) have created a platform for powerful ribosome engineering, but there is as of yet no way to select for efficient ribosome assembly. Indeed, some of these engineered orthogonal ribosomes are similarly affected by assembly defects (Aleksashin et al., 2019). Although P7A7 was the top hit in a screen for β-amino acid incorporation (Melo Czekster et al., 2016) and there was little reason to believe it harbored significant assembly defects from initial *in vivo* assays, our results highlight that assembly problems as a result of aggressive mutations can go undetected in traditional cell-based experiments. Future efforts directed to incorporate assembly into selections for new ribosome function could open new opportunities for engineering robust translation systems for sequence-defined polymer synthesis.

2.4 Materials and Methods

Plasmids and Cloning

All PCR reactions were performed using a Q5 DNA polymerase kit (NEB) and manufacturer recommended concentrations of primers, dNTPs, and enzyme. Mach1 cells (ThermoFisher) were used for all rRNA plasmid cloning and expression. See Table 2.2 for primer sequences. Plasmids pKK3535-P7A7 (Tet^R), encoding the P7A7 23S rRNA, and pLK35(Douthwaite et al., 1989) (Amp^R) were used for cloning and rRNA expression. Plasmid pLK35-WT-MS2 was generated by whole-plasmid PCR amplification of pLK35 and blunt end ligation using primers h98_MS2_F/R

(MS2 tag fragments in capitals). Plasmid pLK35-P7A7-MS2 was generated by Gibson assembly of a PCR fragment from pKK3535-P7A7 with primers P7A7_ex_F/R and a fragment from pLK35-WT-MS2 using the corresponding complementary Gibson primers. pKK3535-040329-MS2 was created via iterative whole-plasmid PCR and blunt end ligation using primers h98_MS2_F/R and then primers 040329_F/R.

Table 2.2: Primers used in this study.

h98_MS2_F	CACCCATGTttttgatcagggtcctgaaggaacgttgaag
h98_MS2_R	ATCCTCATGTaaatgatcagggtcagggagaactcatctcg
P7A7_ex_F	aagatgcagtgtacccgcggc Complementary Gibson primer: gccgcgggtacactgcatctt
P7A7_ex_R	gaactgtctcacgacgttctaaaccc Complementary Gibson primer: gggtttagaacgtcgtgagacagttc
MS2_quant_F	cttgccccgagatgagttctccc
MS2_quant_R	gtaccggttagctcaacgcatcgct
040329_F	tggcagggctcatcacatcctggggc
040329_R	gaggtgccaaacaccgcc

Polysome profiling and rRNA quantitation

Polysome profiling was performed as described (Qin & Fredrick, 2013) using Mach1 cells expressing the desired 23S rRNA-expressing plasmid. Fractions were collected and extracted twice with 1 volume phenol-chloroform (pH 4.5), washed with 1 volume chloroform and precipitated with 1/10th vol. 3M sodium acetate (pH 5.2) and 2.5 vol. ethanol at -20˚C overnight. Precipitated RNA was washed with 300 µL cold 70% ethanol and resuspended in water. Ribosomal RNA was reverse transcribed using Superscript III reverse transcriptase (ThermoFisher) with up to 2 µg of input rRNA and primer MS2_quant_R (Table 2.2). RNA was hydrolyzed with 10 M NaOH at 37 ˚C for 30 min and cDNA was precipitated as before with sodium acetate and ethanol. The cDNA pellet was washed with 70% ethanol and resuspended in water, and concentrations were normalized to 1 ng/µL. Quantification of the ratio of MS2-tagged rRNA to WT rRNA was done via semi-quantitative PCR with Q5 DNA polymerase (NEB). Primers MS2_quant_F/R bind outside the MS2 tagged region of 23S cDNA, generating a 146 base pair (bp) band for MS2-tagged molecules and a 114 bp band for WT. Each sample was amplified twice with varying input cDNA amounts or cycle numbers to verify linear PCR amplification. Most samples were optimized around 0.2 ng input cDNA and 8x-12x cycles for a 10 µL PCR reaction. PCR products were run on a 15% acrylamide Tris-Borate-EDTA minigel and stained with SybrSAFE DNA stain. Gel images were quantitated using ImageJ (Schneider et al., 2012).

Untagged WT Ribosome Purification

WT 50S/30S ribosomal subunits were prepared as described previously(Travin et al., 2019) except the 50S/30S peak was isolated alone and subjected to an additional spin over a 15-40% sucrose gradient in buffer C (20 mM Tris-HCl pH 7.5, 60 mM NH_4Cl, 6 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT). Fractions were concentrated and sucrose was removed using 100 kDa cutoff spin filters (millipore). The purified subunits were frozen in aliquots at -80 °C.

MS2 tagged WT, P7A7 and 040329 Ribosome Crude Purification

E. coli Mach1 cells (ThermoFisher) were transformed with plasmid pLK35-WT-MS2, pLK35-P7A7-MS2, or pKK3535-040329-MS2. Overnight cultures of transformants were diluted 1:100 into 3L LB + 100 µg/mL ampicillin (WT and P7A7) or 50 µg/mL tetracycline (040329). Ribosomes were then prepared as described (Travin et al., 2019) up through crude ribosome pelleting. The pellet was resuspended in ~2mL of buffer A (20 mM Tris-HCl pH 7.5, 100 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT) and stored at 4 °C until MS2-tag affinity purification.

MBP-MS2 purification

A pMAL-c2 plasmid (NEB) encoding N-terminally 6xHis tagged MBP-MS2 (Gift from Nadège Liaud, UC-Berkeley) was transformed into BL21 Codon+ RIL chemically competent cells (Agilent). An overnight culture was used to inoculate 3L of ZYM-5052 auto-inducing media (Studier, 2014) with 100 µg/mL ampicillin and grown for 16 h at 37 °C. Cells were cooled and all subsequent steps carried out at 4 ˚C. After pelleting at 4000 *g*, cells were washed and resuspended in 130 mL lysis buffer (20 mM HEPES pH 7.5, 250 mM KCl, 10 mM imidazole, 1 protease inhibitor tablet (Pierce) per 50 mL , 2 mM 2-mercaptoethanol) before lysis by sonication. Lysate was clarified by centrifugation (48,000 *g*, 30 min) and filtered through a 0.2 µM filter. MBP-MS2 was initially purified using an Akta fast protein liquid chromatography (FPLC) system (GE) with a 5 mL HisTrap column (GE). The column was washed with 10 column volumes (CV) of lysis buffer and eluted with a 10 CV linear gradient of lysis buffer with 10-500 mM imidazole. Protein containing fractions were combined and dialyzed against a low salt MS2 buffer (MS2-LS) (20 mM HEPES pH 7.5, 20 mM KCl, 1 mM EDTA, 2 mM 2-mercaptoethanol) in 10 kDa cutoff dialysis cassettes overnight. MBP-MS2 was then purified again on an FPLC using a 5 mL Heparin column (GE). Batches of 50-100 mg protein (estimated by A280, ε=83310 $M^{-1}cm^{-1}$) were bound to the column, washed with 5 CV of MS2-LS, and eluted with a linear gradient of MS2-LS with 20 mM-1 M KCl. MBP-MS2 fractions were combined, glycerol added to 10%, and stored in aliquots at -80 ˚C.

MS2-tagged ribosome purification

All steps were performed at 4 °C. Crude ribosomes (>60 mg) were diluted to ~15 mg/mL in buffer A. MBP-MS2 (10 mg) was diluted to 0.5 mg/mL in MS2-150 buffer (20 mM HEPES pH 7.5, 150 mM KCl, 1 mM EDTA, 2 mM 2-mercaptoethanol) and loaded onto a 5 mL MBPtrap column (GE). The column was washed with 5 CV buffer A, loaded with crude ribosomes, and then attached to an FPLC. Ribosomes were purified with the following program: 5 CV wash with buffer A, 5 CV wash with buffer A with 250 mM NH_4Cl, 10 CV elution with a linear gradient of buffer A with 0-10 mM maltose. Ribosome containing fractions were concentrated and washed with buffer A in 100 kDa cutoff spin filters. Concentration was estimated with the conversion 1 nM 70S ribosomes = 24 A260. MS2-purified ribosomes were stored at -80 ˚C.

For the short peptide *in vitro* translations, P7A7 ribosomes were additionally purified at small scale using the same method over ~100 µL (packed volume) amylose resin.
Ribosomal purity was assayed using semi-quantitative RT-PCR as described in the polysome profiling assay.

Nanoluciferase in vitro translation
In vitro translation (IVT) reactions used the PURExpress Δribosome kit (NEB) and a nanoluciferase reporter plasmid (Promega). IVTs used 0.5-1 µM purified ribosomal subunits, 10 ng/µL reporter plasmid, and all other components as per the manufacturer's recommendation. Reactions (3 µL) were incubated for 60 min at 37 ˚C and quantified with the Nano-glo assay system (Promega).

Formation of acyl-tRNAs used for short peptide synthesis
Synthesis, purification, and aminoacylation of $tRNA^{Val}$ was carried out using the protocol described previously (Ad et al., 2019). Specifically, 10 µL of 250 µM Flexizyme (eFx)(Goto et al., 2011) was added to 10 µL of either 200 mM HEPES with 200 mM KCl (pH 7.5, for α-Phenylalanine) or 500 mM HEPES (pH 7.5 for β-Phenylalanine) and 10 µL of 250 µM $tRNA^{Val}$. The samples were incubated at 95 ºC for 2 min and allowed to reach room temperature in 5 min. 60 µL of 1 M magnesium chloride was then added, followed by 10 µL of a DMSO solution of each cyanomethyl ester amino acid variant (50 mM). Reactions were incubated at 4 ºC for 60 h. The reactions were quenched by addition of sodium acetate (pH 5.2) to a final concentration of 300 mM and ethanol was added to a final volume of 70% (*v/v*). The samples were then incubated at -80 ºC for 1 h and the RNA was pelleted by centrifugation at 21,300 *g* for 30 min at 4 ºC. The supernatant was removed and the pellet was washed with 500 µL of 70% (*v/v*) ethanol (stored at -20 ºC). The sample was then centrifuged at 21,300 *g* for 7 min at 4 ºC and the supernatant was removed. The pellet was air-dried for 2-5 min either at room temperature or on ice. When used immediately, the pellet was resuspended in 1 mM sodium acetate (pH 5.2). If used at a later date, the pellet was stored dry at -80 ºC and resuspended in 1 mM sodium acetate (pH 5.2) before use.

In vitro translation reactions
Templates for *in vitro* translation of short peptides containing a FLAG tag fMet-Val-Phe-Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys (MVFDYKDDDDK, fMVF-Flag) were generated as described previously.(Ad et al., 2019) *In vitro* transcription/translation of (fMVF-Flag was carried out using the combination of PureExpress (ΔtRNA, Δaa (E6840S)) and PureExpress (Δribosome (E3313S)) kits (NEB) with the following modifications. To generate the fMVF-Flag WT peptide reactions contained (25 µL): Solution A (ΔtRNA, Δaa kit), 330 µM methionine, 330 µM valine, solution containing 330 µM tyrosine, 330 µM phenylalanine, and 330 µM lysine, 280 µM aspartic acid (pH 7), tRNA solution (ΔtRNA, Δaa kit), Factors Mix (Δribosome kit), 0.8 µM purified 30S subunits, 0.8 µM purified 50S subunits (either WT or P7A7), 500 ng dsDNA template, and water (to 25 µL). When using $tRNA^{Val}$ (50 µM) charged using eFx, valine was omitted from the reaction mixture. The reactions were then incubated for 6 h at 37 ºC. The reactions were quenched by placing them on ice and adding 25 µL of dilution buffer (10 mM magnesium acetate (Sigma-Aldrich) and 100 mM sodium chloride (Sigma Aldrich)). To remove the proteins and majority of nucleic acid macromolecules, 5 µL of Ni-NTA (Qiagen) slurry was added and the solution was incubated with light agitation at 4 ºC for 50 min.

The Ni-NTA resin was removed by centrifugation at 21,300 *g* for 10 min at 4 ºC. The supernatant was then frozen at -80 ºC for 5 min and centrifuged once more at 21,300 *g* for 10 min at 4 ºC. The supernatant was analyzed on a Poroshell 120 EC-C18 column (2.7 µm, 3.0 × 50 mm, 45 ºC, Agilent) using a linear gradient from 5 to 55% acetonitrile over 6.5 min with 0.1% formic acid as the aqueous mobile phase after an initial hold at 95% 0.1% formic acid for 0.5 min (0.6 mL/min) using a 1290 Infinity II UHPLC (G7120AR, Agilent). Peptides were identified using LC-HRMS as described previously.(Ad et al., 2019)

Magnesium-dependent 70S complex formation
10 pmol WT or P7A7 50S subunits were mixed with 20 pmol WT 30S subunits in 20 µL of buffer C, described above, with 10 mM, 15 mM or 20 mM $MgCl_2$ and incubated for 30 min at 37 ˚C. Reactions were spun over 15-30% sucrose gradients made in buffer C with appropriate $MgCl_2$ concentrations at 178,000 *g* (SW-41 rotor, Beckman Coulter) for 3 hours. A254 traces were measured with an ISCO gradient fractionation system.

EM sample preparation
Samples were deposited onto glow-discharged 300 mesh Quantifoil UltraAuFoil R1.2/1.3 grids with an additional top layer of continuous amorphous carbon floated on. The sample was initially incubated on the grid for approximately 1 minute, after which excess sample was washed off in a buffer containing 20 mM Tris-HCl, pH 7.5, 60 mM NH_4Cl, 6 mM $MgCl_2$, 0.5 mM EDTA, 2mM DTT. A Vitrobot Mark IV was used for plunge-freezing with the settings: 20 °C, 100% humidity, blot force 8, blot time 3 seconds. Inside the Vitrobot, the grid was first side-blotted with filter paper to remove the majority of the solvent, followed by depositing 1.2 µL of buffer prior to blotting and plunging into liquid ethane.

EM Data Collection
Images were collected on an FEI Titan Krios electron microscope operated at 300 keV and with a GIF energy filter. The images were collected on a GATAN K2 Summit camera in super-resolution mode. Magnification was set to 215,000x for a pixel size 0.56 Å (0.28 Å super-resolution size). We collected images using a focal pair approach that we have since determined to have no advantage over a conventional defocus ramp. Briefly, two movies were recorded of each area, the first targeting a defocus level roughly around -0.3 µm (P7A7) or -0.5 µm (WT) with a total dose of 10 e-/ Å2, and the second with a target defocus around -3.5 µm (P7A7) or -3.6 µm (WT) and total dose 20 e-/ Å2. SerialEM(Mastronarde, 2005) was used for automated data collection and Focus (Biyani et al., 2017) was used for real-time monitoring of the data.

Image processing
For all movies, motion correction was performed with MotionCor2. (Zheng et al., 2017) CTF estimation was performed with CTFFind4, (Rohou & Grigorieff, 2015) and poorly fit micrographs were discarded based upon visual inspection. For ease of working with the focal pair data, far-focused movies were first processed to determine particle position and rough orientation before moving to the near-focused movies for high resolution structure determination.

For P7A7, templates for automatic picking were generated in RELION 2.0 (Sjors H. W. Scheres, 2012) by first picking using a Gaussian blob, then classifying the results and choosing the best

classes for templates. Three rounds of 2D classification were performed on 92,647 template-picked particles, reducing the number to 79,268 particles. Ab-initio reconstruction with 3 classes was performed in cryoSPARC v1, (Punjani et al., 2017) resulting in one class that resembled a normal 50S subunit and one class that appeared to be completely missing the central protuberance (CP). These two classes were separately submitted for 3D auto-refinement in RELION. The latter structure (with 21,319 particles from the far-focused data) did not refine to high enough resolution to pursue further. The class with an intact CP was subject to a round of 3D classification without alignment. Three out of four classes were pooled and refined to 4.17 Å from far-focused data. The resulting particle positions and orientations were then applied to near-focus movies with custom python scripts, and those particles were refined starting from local angular searches to 3.11 Å with 37,609 particles.

For WT 50S, all processing was done in RELION 2.0 (Sjors H. W. Scheres, 2012). 109,802 particles were auto-picked from far-focused micrographs with a Gaussian blob. Four rounds of 2D classification were performed to yield 99,356 particles for an initial 3D auto-refinement. Processing was shifted to near-focused micrographs as described for P7A7. 3D classification without alignment was performed after an initial near-focused refinement from local searches. The best class, containing 84,372 particles, was refined again for a final resolution of 3.20 Å. Coordinates for the WT 50S subunit based on PDB entry 4YBB (Noeske et al., 2015) were docked into the density as a rigid body. Molecular graphics were created using ChimeraX (Goddard et al., 2018).

Cryo-EM maps of the WT and P7A7 50S ribosomal subunits were deposited in the EMDB as EMD-20854 and EMD-20853, respectively.

Cryo-EM map occupancy calculation
Map occupancy at each feature was calculated using a previously developed algorithm. (Davis et al., 2016) A suitable contour level for each map was chosen after amplitude scaling and resampling by normalizing to the volume of uL4, a known early assembling ribosomal protein. Complete-linkage clustering was used to group structures by occupancy similarity. Results matched well with visual inspection of map features.

Peptide preparation for LC-MS/MS and TMT quantification
Ribosomal proteins were precipitated in triplicate from WT-MS2 or P7A7-MS2 purified 50S subunits with 20% trichloroacetic acid at 4 °C for 1 hr. Protein pellets were washed 3x with 500 µL 0.01 M HCl in 90% acetone and dried. Protein was digested and TMT labeled using the TMTduplex isobaric mass tagging kit (ThermoFisher).

Mass spectrometry
Mass spectrometry was performed by the Vincent J. Coates Proteomics/Mass Spectrometry Laboratory at UC Berkeley. Peptides were analyzed on a ThermoFisher Orbitrap Fusion Lumos Tribrid mass spectrometry system equipped with an Easy nLC 1200 ultrahigh-pressure liquid chromatography system interfaced via a Nanospray Flex nanoelectrospray source. Samples were injected on a C18 reverse phase column (25 cm x 75 µm packed with ReprosilPur C18 AQ 1.9 µm particles). Peptides were separated by a gradient from 5 to 32% acetonitrile in 0.02% heptafluorobutyric acid over 120 min at a flow rate of 300 nL/min. Spectra were continuously

acquired in a data-dependent manner throughout the gradient, acquiring a full scan in the Orbitrap (at 120,000 resolution with an AGC target of 400,000 and a maximum injection time of 50 ms) followed by 10 MS/MS scans on the most abundant ions in 3 s in the dual linear ion trap (turbo scan type with an intensity threshold of 5000, CID collision energy of 35%, AGC target of 10,000, maximum injection time of 30 ms, and isolation width of 0.7 m/z). Singly and unassigned charge states were rejected. Dynamic exclusion was enabled with a repeat count of 1, an exclusion duration of 20 s, and an exclusion mass width of ±10 ppm. Data was collected using the MS3 method (Ting et al., 2011) for obtaining TMT tag ratios with MS3 scans collected in the orbitrap at a resolution of 60,000, HCD collision energy of 65% and a scan range of 100-500.

LC-MS/MS Data analysis

Protein identification and quantification were done with IntegratedProteomics Pipeline (IP2, Integrated Proteomics Applications, Inc. San Diego, CA) using ProLuCID/Sequest, DTASelect2 and Census (Cociorva et al., 2007; S. K. Park et al., 2008; T. Xu et al., 2006). Tandem mass spectra were extracted into ms1, ms2 and ms3 files from raw files using RawExtractor(McDonald et al., 2004) and were searched against the *E. coli* protein database plus sequences of common contaminants, concatenated to a decoy database in which the sequence for each entry in the original database was reversed (Peng et al., 2003). All searches were parallelized and searched on the VJC proteomics cluster. Search space included all fully tryptic peptide candidates with no missed cleavage restrictions. Carbamidomethylation (+57.02146) of cysteine was considered a static modification; TMT tag masses, as given in the TMT kit product sheet, were also considered static modifications. We required 1 peptide per protein and both tryptic termini for each peptide identification. The ProLuCID search results were assembled and filtered using the DTASelect (Cociorva et al., 2007) program with a peptide false discovery rate (FDR) of 0.001 for single peptides and a peptide FDR of 0.005 for additional peptide s for the same protein. Under such filtering conditions, the estimated false discovery rate was zero for the dataset used. Quantitative analysis on MS3-based MultiNotch TMT data was analyzed with Census 2 in IP2 platform (S. K. R. Park et al., 2014). As TMT reagents are not 100% pure, we referred to the ThermoFisher Scientific TMT product data sheet to obtain purity values for each tag and normalized reporter ion intensities. While identification reports best hit for each peptide, Census extracted all PSMs that can be harnessed to increase accuracy from reporter ion intensity variance. Extracted reporter ions were further normalized by using total intensity in each channel to correct sample amount error.

To account for purity differences between ribosomal protein samples, two additional normalization methods were attempted. First, intensities were normalized to the total 50S ribosomal protein signal for each respective mass tag. Second, intensities were normalized to the averaged signals of early assembling proteins uL4, uL13, bL20, uL22, and uL24. These proteins are expected to be stoichiometrically present in both mature WT 50S and assembly-stalled P7A7 subunits. These two methods produced similar results and the latter was chosen for further analysis. Reported proteins were identified in all three relative quantitation experiments with the exception of L35, which was seen in two.

3. Structural analysis of unnatural monomers in the wild type E. coli ribosome

3.1 Introduction

In light of the assembly defects in the P7A7 ribosome detailed in chapter 2, we sought alternative and parallel methods of improving our ribosome engineering intuition. The total PTC disordering seen in P7A7 limited what we could learn about its catalytic abilities toward β-amino acids. We reasoned that a wild-type ribosome complexed with unnatural substrates of interest could still reveal much about the limitations of the natural PTC. All non-L-α-amino monomers are evidently poor substrates in the PTC; they react slowly, produce low yielding products; and only a handful have been successfully elongated sequentially (Katoh, Tajima, et al., 2017; Katoh & Suga, 2018). Sufficiently high resolution structures of these monomers in the PTC would allow comparison of positioning and orientation relative to α-amino substrates, identifying clashes with PTC residues that prevent polymerization. This style of analysis should be particularly effective because the current consensus around the mechanism of PTC catalysis involves nearly exclusively entropic effects, i.e. proper organization and orientation of substrates and exclusion of solvent (Bieling et al., 2006; Leung et al., 2011; Sievers et al., 2004). The lack of participation in acid-base catalysis by ribosomal residues is particularly encouraging, suggesting that the PTC can be remodeled toward accepting new substrates primarily by changing its shape. There seem to be no precise ribosome-substrate interactions necessary for reactivity that would need to be preserved or mirrored in mutant PTCs. The most important interaction appears to be on a substrate, the 2' OH on A76 of the P-site tRNA is heavily implicated in helping to deprotonate the attacking amine (Bayryamov et al., 2007; Weinger et al., 2004), presumably a moiety that can be maintained in mutant PTC systems.

This type of structural study, complexing the wild type ribosome with unfavored unnatural substrates has been done once before to understand polymerization with D-amino acids. Melnikov and coworkers solved the crystal structure of a D-phenylalanyl-tRNA analog in the A site of the *Thermus thermophilus* ribosome in complex with mRNA and deacylated initiator $tRNA_{fMet}$ (S. V. Melnikov et al., 2019). It is important to note that tRNA analogs, such as CC-puromycin and derivatives, as well as deacylated P-site tRNA have been used regularly to obtain high quality structures of the natural peptidyltransferase reaction (Martin Schmeing et al., 2005; Polikanov et al., 2014; Voorhees et al., 2009). The resulting structure revealed a marked similarity to the wild type substrates. The bases of the tRNA analog were identically positioned against the A loop, indicating the tRNA analog was correctly accommodated into the PTC and rejected the model that the D-amino substrate would prevent the tRNA from inducing an active conformation of the PTC. Furthermore, the D-phenylalanine side chain was well accommodated into the A-site cleft, closely matching the "wildtype" CC-puromycin. The crucial difference is that, as a result of the D-phenylalanine side chain binding to the A-site cleft, its α-amino group is positioned further away from the P-site ester bond, limiting reactivity. Furthermore, the position of the D-phenylalanine's α-amino group precludes it from forming important hydrogen bonding with the P-site A76 2' OH group. Modeling a correct α-amino conformation caused an unavoidable clash between the β-carbon of the D-phenylalanine and PTC base U2506 (*E. coli* 23S rRNA numbering is used). This hypothetical clash immediately suggests possibilities for targeted mutation of the ribosome, namely mutation of U2506 and the A-site cleft (A2451 and C2452), all of which are at least partially mutationally flexible (d'Aquino et al., 2020; Erlacher & Polacek, 2012; Sato et al., 2006; Youngman et al., 2004).

Along similar lines, we have attempted to obtain structures of the wild type ribosome in complex with various non-α-amino charged tRNAs with the hope of identifying precise molecular interactions that limit exotic polymerization chemistries in the PTC. We have obtained structures of a malonic acid methyl ester charged tRNA in the P site but have been hampered by tRNA quality and hydrolysis issues. In parallel, our techniques for ribosome structure determination have been refined and streamlined to incorporate the latest advances in cryo-EM data collection and processing, resulting in the highest resolution ribosome yet published (Watson et al., 2020). Recent efforts to synthesize amide-linked charged tRNAs have delivered more stable substrates and allowed a high resolution view of the PTC.

3.2 Results

Abortive attempts to restore P7A7 ribosome activity

In an attempt to restore P7A7 ribosome activity or identify a crucially destructive single mutation. The P7A7 23S rRNA was revertant-scanned by making individual P7A7 to wild type mutations base-by-base and purifying the resulting mutants. We thought it might be possible that there was a single critical mutation that would restore functional assembly but still maintain improved β-amino acid translation. Unfortunately, none of the purified single base reversions showed restored activity in *in vitro* translation reactions using nanoluciferase templates in the defined PURExpress system (**Figure 3.1**). In light of this result, we chose to focus on WT-ribosome-unnatural-tRNA complexes as our primary source of structural information to guide future engineering.

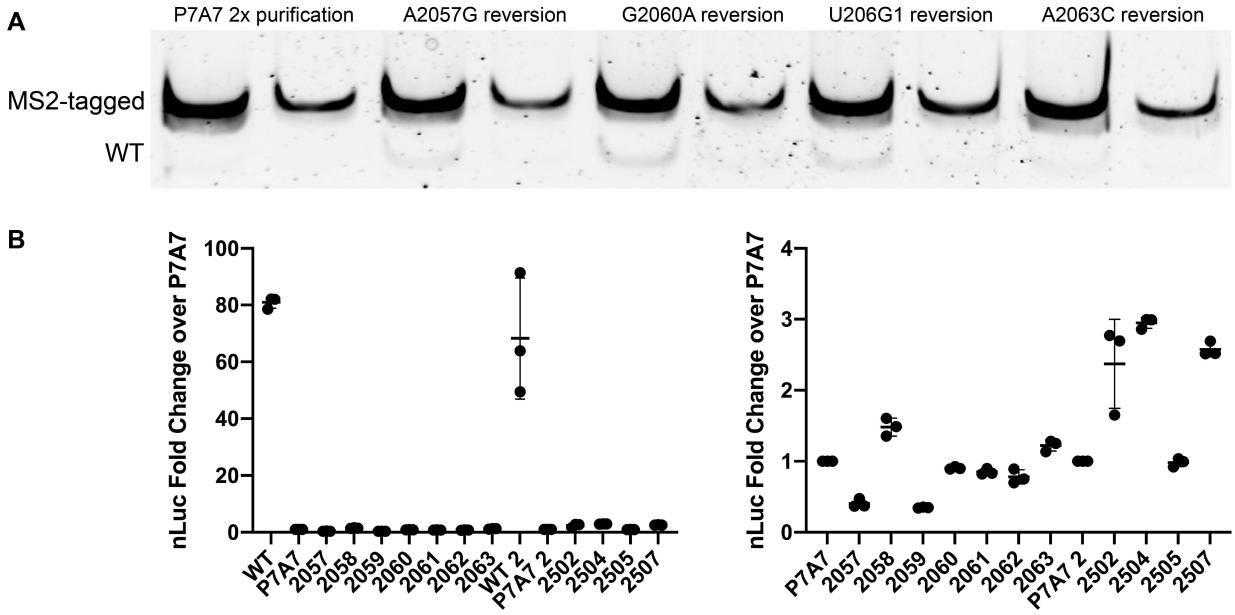


Figure 3.1. P7A7 base revertants purification and activity. A) RT-PCR MS2-tagged ribosome purity assays, as shown in **Figure 2.2.2,** showing purity of double purified P7A7 (5 mL column and 100-200 µL resin) and resin purified single base revertants. WT band corresponds to 2-5% contamination across samples. B) Activity of base revertants in nanoluciferase synthesis in PURExpress *in vitro* translation reaction. Results are presented with and without WT ribosome comparative data. No revertant demonstrated translation improvement that could be separated from low-level WT ribosome contamination as measured in panel A.

Malonic acid methyl ester in the P-site

Our first attempts at WT ribosome complexes with unnatural charged tRNAs used tRNA substrates with already confirmed activity in the WT *E. coli* ribosome (Ad et al., 2019). We used pentafluorobenzoic acid (5F) and Malonic acid methyl ester (MMAL) as $tRNA_{fMet}$ P-site substrates in the complexes alongside a wild type valine-$tRNA_{val}$ in the A site. All tRNAs were prepared using flexizymes (Goto et al., 2011). Multiple data collection sessions resulted in only a single map with recognizable density for MMAL-$tRNA_{fMet}$ in the P-site and no A-site density (**Figure 3.2A**). Model fitting and comparison with a high resolution structure of an fMet-NH_2-$tRNA_{fMet}$ in the P site revealed a similar orientation to the natural initiation monomer (**Figure 3.2B**) (Polikanov et al., 2014), suggesting that MMAL is generally a good fit for the PTC. Aligning the nearby PTC bases with a P-site only late-initiation structure (PDB: 6O9J; (Kaledhonkar et al., 2019) showed similar PTC organization between the two models with disparities at A2602, which had poor density in the MMAL map and U2585, which clashes with the P-site monomer in the 6O9J structure (**Figure 3.2C**). Inspection of the map associated with 6O9J confirmed the modeling of U2585 in direct conflict with the P-site monomer is likely an error in real space refinement. Lastly, our MMAL structure contains a key G:U base pair between G2506 and U2583 that delineates inactive and active PTCs. This pairing is associated with an inactive state of the PTC seen in antibiotic bound structures and initiation complexes before A-site tRNA binding (Kaledhonkar et al., 2019; Martin Schmeing et al., 2005; Osterman et al., 2017), while active complexes see a rotation of G2506 that breaks the pairing (**Figure 3.2D**). Taken together, our MMAL structure agrees well with previous structures of charged tRNA orientation during initiation and PTC organization when only the P-site is occupied. This confirms that MMAL is accommodated naturally into the PTC and supports data that it is competent in translation initiation (Ad et al., 2019).

A final notable feature of the MMAL structure is density matching multiple conformations of the P-site A76 ribose (**Figure 3.3A**) The orientation modeled in Figure 3.2 point the 3'O and MMAL monomer into the PTC, while the "flipped" conformation would point the monomer into a small pocket formed by the tRNA and 23S rRNA base A2602 (**Figure 3.3B**). The high possibility of steric clash of any monomer in this pocket suggests that the "flipped" A76 ribose conformation corresponds to hydrolyzed, uncharged tRNA. Thus, this MMAL structure represented a mixed population of charged and uncharged tRNA, limiting the ultimate resolution of our data.

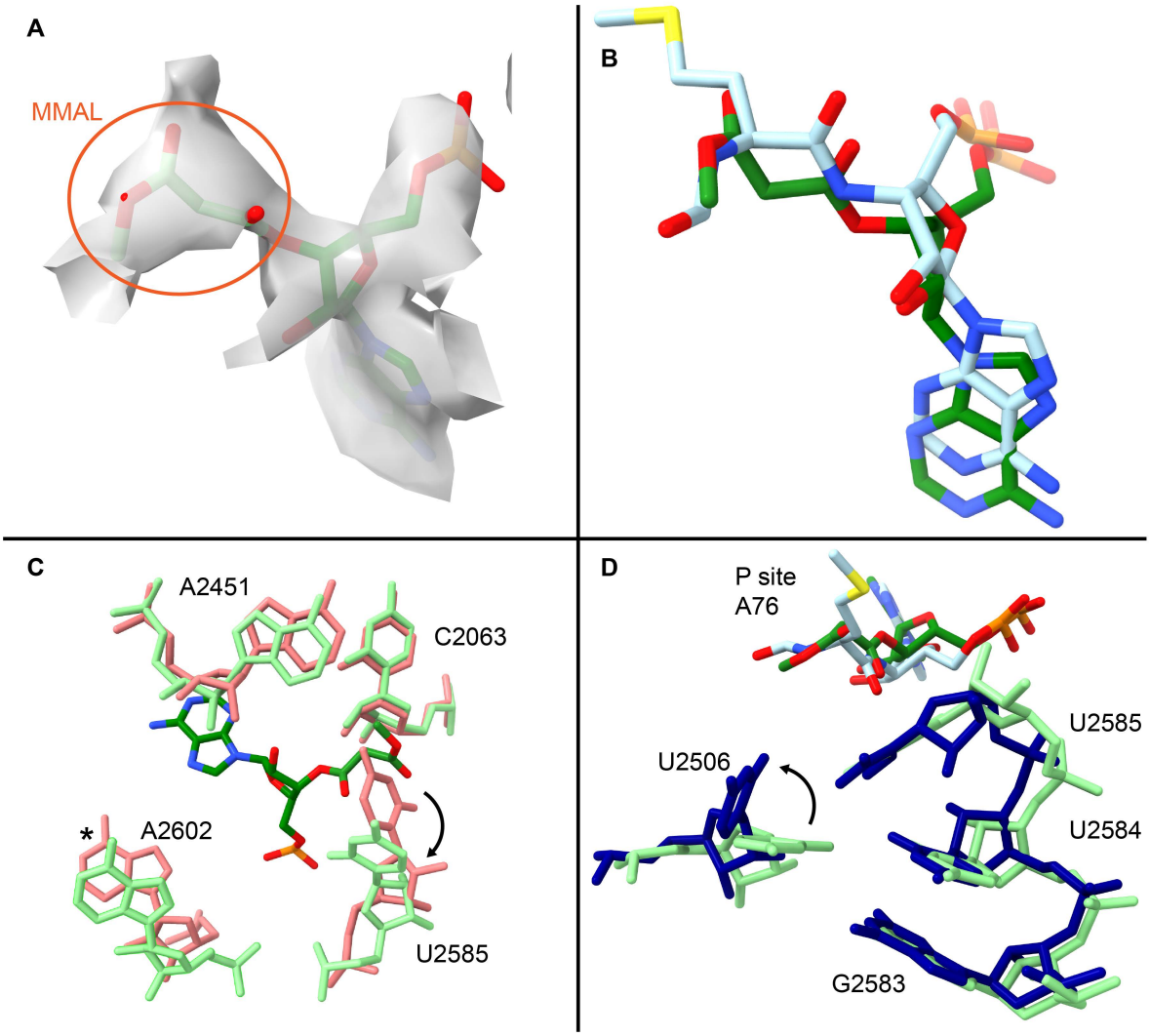


Figure 3.2. MMAL-tRNA$_{fMet}$ in the PTC. A) Cryo-EM charge density map of the A76 of the P-site tRNA and MMAL monomer. MMAL monomer is circled in orange. B) Comparison between fMet-NH$_2$-tRNA$_{fMet}$ from PDB 1VY4 (Polikanov et al., 2014) and MMAL monomers with terminal tRNA adenosines. C) Nearby PTC rRNA bases to the P site compared between MMAL structure and late initiation structure PDB 6O9J (Kaledhonkar et al., 2019). D) U2506-G2583 interactions compared. Colors: MMAL-tRNA$_{fMet}$ in dark green and rRNA in light green. 1VY4 fMet-NH$_2$-tRNA$_{fMet}$ in light blue and rRNA in dark blue, 6O9J rRNA in pink. Asterix (*) indicates a base in the MMAL structure that had poor density, limiting accurate modeling.

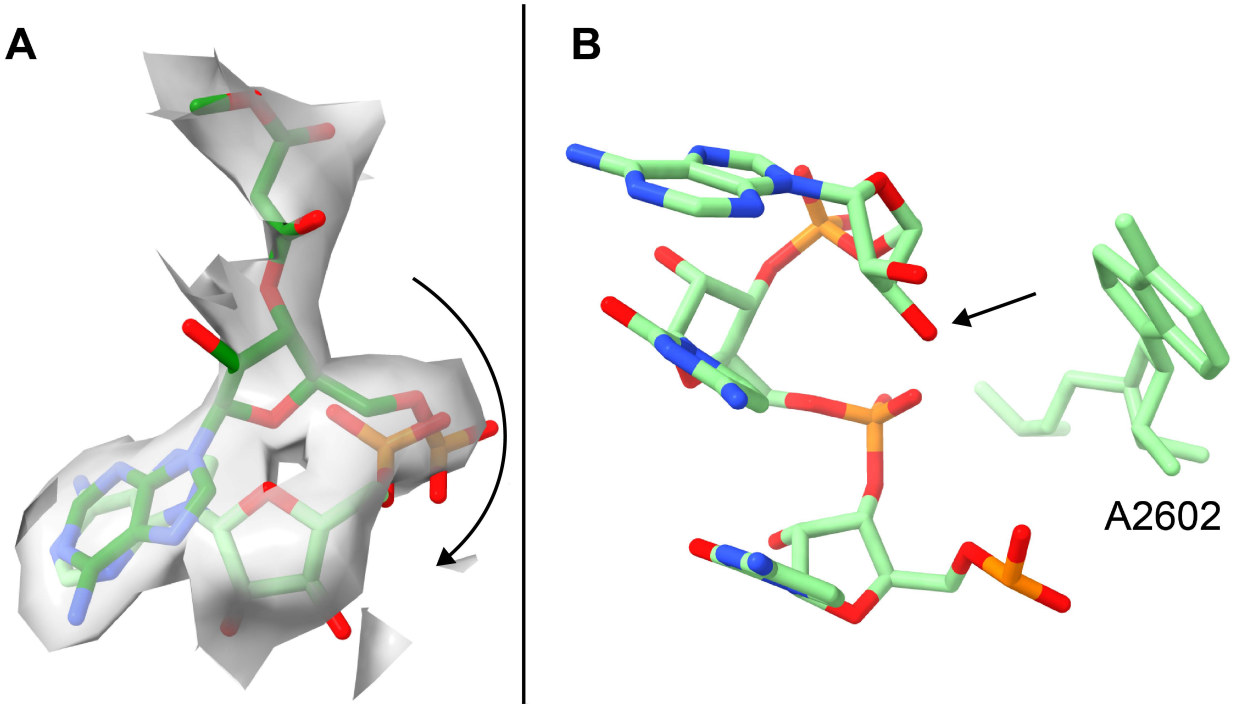


Figure 3.3. Multiple A76 ribosome conformations. A) The MMAL-tRNA$_{fMet}$ model (dark green) and the uncharged "flipped" A76 ribose model (light green) both docked into the map, showing good fit in density for either model. B) The small pocket formed by the P site tRNA and base A2602 into which a hypothetical monomer would project in a "flipped" A76 conformation. 3'OH indicated with an arrow.

Troubleshooting PTC resolution

Several subsequent attempts at obtaining structures with MMAL or 5F in the P site and Val-tRNA$_{val}$ or Gly-tRNA$_{gly}$ in the A site yielded structures with better tRNA occupancies but poor resolution in the PTC. Occupancy of tRNAs was improved by increasing the relative concentration with respect to 70S ribosomes, foregoing sucrose gradient purification steps after complex formation, and adding the antibiotic paromomycin that can increase the affinity of the A-site tRNA for the ribosome (Ogle et al., 2002).

There are multiple possible explanations for our issues in resolving substrates in the PTC. First, molecules like 5F or MMAL might simply not be well accommodated in the PTC. While their reactivity can be detected using sensitive LC-MS/MS techniques to identify monomer-containing peptides at low abundance, any accommodation in the PTC may ultimately be unfavorable, transient, and not captured by equilibrium-driven complex assembly conditions. Secondly, the PTC may accept these unnatural substrates but not bind them in a single stable position. Substantial motion in the PTC would greatly limit resolution in the resulting map. Thirdly, the tRNA-monomer linkages may be unstable and hydrolyze during complex formation and grid preparation. There is substantial evidence that tRNA-monomer ester bonds can be quite unstable and that this stability depends on the identity of the linked monomer (Hentzen et al., 1972; Peacock et al., 2014). Furthermore, the structure of the PTC protects P-site acyl-tRNAs from premature hydrolysis on the ribosome while EF-Tu shields them in solution (Martin Schmeing et al., 2005; Peacock et al., 2014; Zavialov et al., 2002). Our complex preparations do not use EF-Tu to deliver charged tRNAs and it is possible that unusual monomers such 5F or MMAL may not bind in the PTC in a manner that prevents hydrolysis. Fourthly, our

nonenzymatic complex preparation lacking any initiation factors or EF-Tu may result in "broken" complexes with disordered PTCs.

A complex using wild type Met-$tRNA_{fMet}$ as a P-site substrate with Gly-$tRNA_{gly}$ in the A site showed excellent tRNA occupancy, validating our complex assembly approach. However, the PTC in this complex was still disordered and multiple A76 ribose conformations were present, suggesting substantial uncharged tRNA populations even with wild type substrates. While this result did not reveal anything about how unnatural monomers could orient in the PTC, it underscored the necessity of generating stable, intact charged tRNA for any future structural studies.

To improve the stability of our tRNAs against hydrolysis we established a system to synthesize and charge 3'NH_2-tRNAs *in vitro*. The amide-linked charged tRNAs are far more stable in solution and do not react in the PTC, making them ideal substrates for studying pre-attack orientation. We used a previously published system where T7 transcribed N-1 tRNAs lacking A76 are tailed with 3'NH_2-ATP by *Archaeoglobus fulgidus* CCA-adding enzyme (Katoh & Suga, 2019). Notably, complete amino-tailing was accomplished at 1:10th the 3'NH_2-ATP concentration of previously published protocols, conserving valuable 3'NH_2-ATP substrate (**Figure 3.4**).

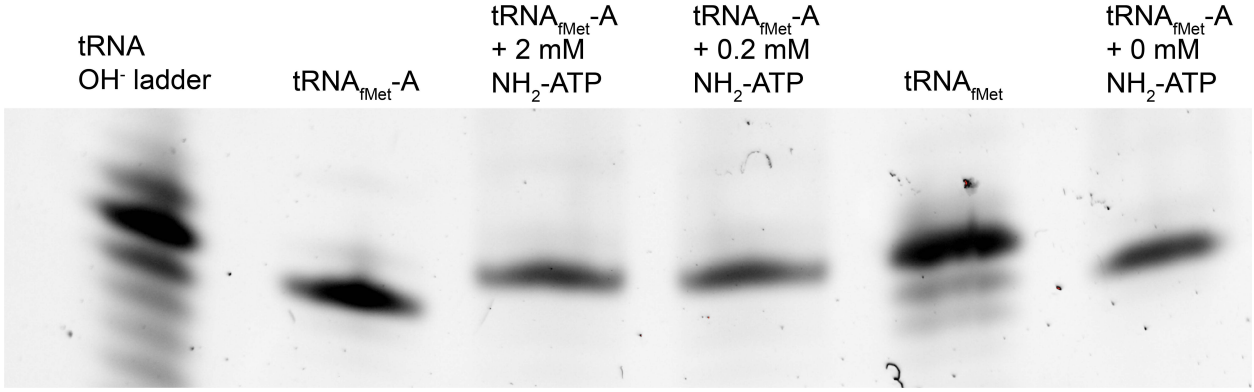


Figure 3.4. tRNA 3' amino tailing reactions. Polyacrylamide tris-borate-EDTA (TBE) urea gel of 3'NH_2-ATP addition onto N-1 $tRNA_{fMet}$ transcripts ($tRNA_{fMet}$-A). Lanes 3,4, and 6 contain 1:1 tRNA:CCA-adding enzyme ratios.

With the amino-tailed tRNAs in hand, we sought to test our hypothesis that amide-linked tRNAs would allow better monomer resolution in the PTC. Tests showed that methionyl-tRNA synthetase (MetRS) could charge methionine onto 3'NH_2-$tRNA_{fMet}$ highly efficiently (**Figure 3.5**). The methionine charging was stable under lengthy treatment with a mild base (TBE running buffer, ~ pH 8.3), confirming a hydrolysis resistant amide linkage. This Met-tRNA was selected as our model amide-tRNA substrate and ribosome complexes were formed with Met-NH_2-$tRNA_{fMet}$ programmed in both the P and A sites.

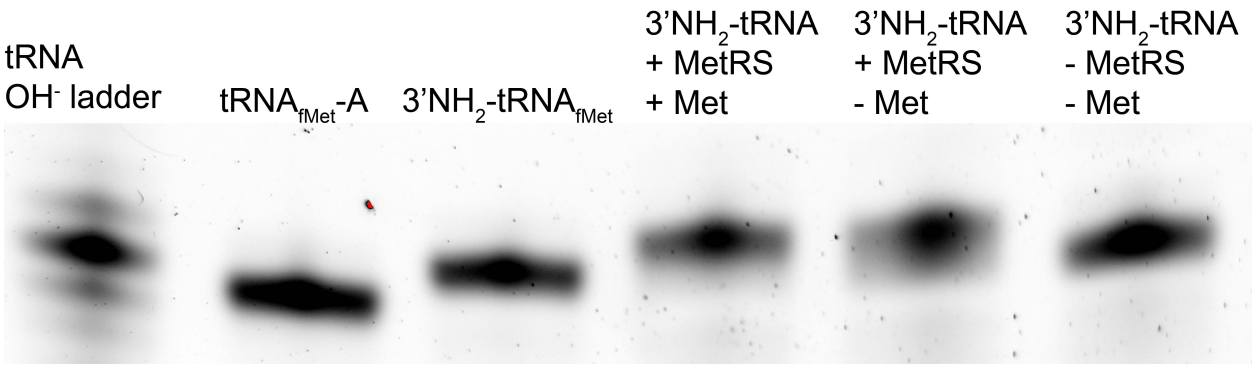


Figure 3.5. Methionine charging of amino-tailed tRNA. Polyacrylamide TBE urea gel of MetRS charging methionine onto 3'NH_2-$tRNA_{fMet}$. methionine charging in lane 4 and 5 results in an upward shift of the tRNA band. Charging in lane 5 in the absence of methionine is attributed to impurities in the purified MetRS.

A Met-Met high resolution PTC structure

We obtained a cryo-EM map of the Met-NH_2-tRNA-Met-NH_2-tRNA-mRNA-ribosome complex that approached the highest resolution yet obtained (Watson et al., 2020). Whereas previous structures had shown excellent detail in nearly every part of the ribosome except the PTC, the amide-linked Met-tRNA ends are well resolved (**Figure 3.6A**), allowing modeling of the monomers. Only the sidechain of the P-site methionine is disordered, understandable given the considerable open space it occupies near the opening of the exit tunnel. Comparison to the 1VY4 structure with fMet-NH_2-$tRNA_{fMet}$ in the P site and Phe-NH_2-$tRNA_{phe}$ in the A site confirms the proper induced PTC state (**Figure 3.6B**). U2506 has rotated away from G2583 and now "squeezes" U2585 against U2584, closing off the PTC and helping to orient the A site substrate. There is no evidence of an alternate A76 ribose conformation in the P-site, confirming a high occupancy of charged tRNA in the complex. The striking improvement in PTC resolution from this experiment identifies amine-tailed tRNA as the crucial process improvement for our cryo-EM workflow.

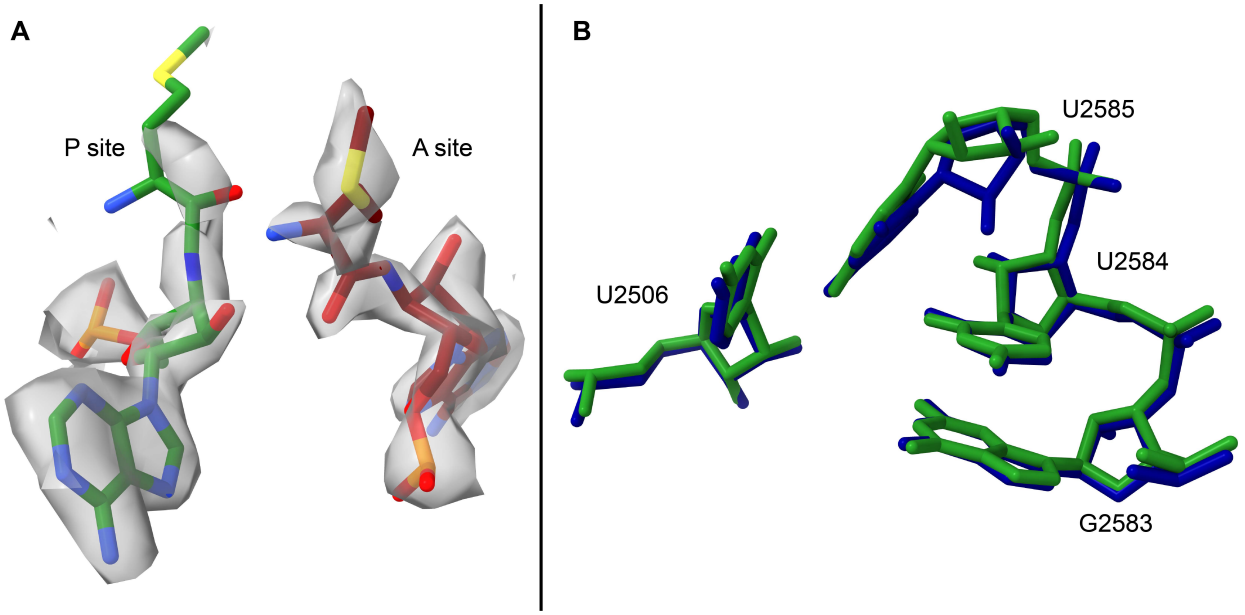


Figure 3.6. A Met-Met high resolution PTC structure. A) Details of the P site and A site Met-NH_2-$tRNA_{fMet}$ terminal NH_2-adenosines and attached methionine monomers. B) Comparison of the met-met (green) induced fit with PDB 1VY4 (blue).

3.3 Discussion

Since the P7A7 ribosome proved largely intractable for structural and functional studies, we have been focusing our efforts on understanding the polymerization chemistries allowed by the wild type ribosome. We have established a system capable of resolving monomers in the PTC to high resolution, allowing unambiguous modeling of their positions. We have also solved the structure of one unnatural monomer in the P site of the PTC, showing that MMAL can be accommodated similarly to natural initiation processes. While this structure is encouraging and, to the best of our knowledge, the only structure of a non-α-amino substrate in the PTC, many questions still remain. The P site itself is likely more permissive of various exotic monomers used as initiation substrates than the A site for elongation, as evidenced by the wide array of substrates already successfully used (Ad et al., 2019; Lee, Schwarz, et al., 2020; Tsiamantas et al., 2019). Initiation monomers also need only react once and thus have none of the reactivity and accommodation concerns of A-site substrates. The P-site ester bond is relatively tightly constrained by its proximity to the tRNA -CCA end, while the reactive group in the A site must be activated and pointed toward the accepting ester by the PTC. Processive polymerization with new backbone chemistry will require engineering to optimize the full cycle of translation on the ribosome. Complexes with new A-site substrates will deliver much needed information about the positioning of reactive groups and the accommodation of bulky or inflexible moieties into the PTC. Ultimately, complexes with novel substrates in both sites will offer more information still.

When evaluating our string of poor PTC structures we proposed four hypotheses for the observed disorder: 1) unnatural monomers cannot enter the PTC, 2) these monomers can enter but do not adopt a stable conformation, 3) charged tRNAs are hydrolyzed during complex formation, and 4) nonenzymatic complex formation is too inefficient to get high occupancy complexes. A control structure with Met- and Gly-tRNAs substantially ruled out hypothesis 4

and we reasoned that hypotheses 1 and 2 were downstream of 3. That is, hydrolysis of charged tRNA would need to be ruled out before monomer accomodation in the PTC could be. Furthermore, evidence for multiple orientations of the P-site tRNA A76 ribose suggested a large population of deacylated tRNA was present. Our decision to pursue hydrolysis resistant amino-tailed tRNAs was validated by the high resolution of the Met-Met PTC structure, setting the stage for observing new chemistry in the ribosome.

While our NH_2-tRNA synthesis and cryo-EM workflow may be ready to deliver these structures, there are a handful of parallel issues that we are working to solve. Most importantly, techniques for charging non-α-amino monomers onto tRNAs are rarely 100% efficient. With first hand knowledge of how a subpopulation of deacylated tRNAs will contaminate structural data, we are pursuing improvements in charging protocols and methods for purifying charged tRNAs. The most versatile charging system, flexizyme chemistry, routinely cannot proceed past ~ 50% yield, necessitating an effective purification protocol. Chromatographic solutions such as reverse phase HPLC provide some promise towards a solution (Blanchard et al., 2004; Mesters et al., 1994). When pure tRNA substrates can be obtained, we anticipate the resulting structures will reveal the limits of wild type ribosome catalyzed polymerization chemistry and guide engineering toward ribosomal synthesis of new classes of materials.

3.4 Materials and Methods

Small scale MS2-tagged ribosome preparation
Purification was performed essentially as described in chapter 2 (See "*MS2 tagged WT, P7A7 and 040329 Ribosome Crude Purification"* and "*MS2 tagged ribosome purification"*) but on a smaller scale. Instead of a 5 mL FPLC column, ~ 200 µL (wet) amylose resin was used to bind MBP-MS2 and then crude ribosomes. Approx. 2 mg of MBP-MS2 was used in each purification. Important note: All buffers were adjusted from chapter 2 to 1 mM Mg^{2+} to reduce WT ribosome contamination. Resin was washed 5x with 1 mL buffer A, at least 5x with buffer A with 250 mM NH_4Cl, and eluted with 3x 200 µL of buffer A with 10 mM maltose. Ribosomes were washed with buffer A in 100,000 Da cutoff 0.5 mL spin filters, concentrated, and stored at -80 ˚C.

Ribosome purity assessment
Analysis of ribsome purity was performed as described in chapter 2 "*Polysome profiling and rRNA quantitation"*

Nanoluciferase in vitro translation
In vitro translation reactions were performed as described in chapter 2 "*Nanoluciferase in vitro translation"*

tRNA synthesis and charging (MMAL and 5F)
MMAL-$tRNA_{fMet}$ and 5F-$tRNA_{fMet}$ were a gift from Omer Ad (Schepartz lab, UC Berkeley). The tRNAs were prepared and charged as in (Ad et al., 2019).

MMAL-tRNA-mRNA-ribosome complex formation
70S ribosomes, mRNA, and tRNAs were prepared as described (Watson, eLife 2020). Ribosome–mRNA–tRNA complexes were formed on ice in a 20 µL reaction with 20 pmol ribosomes, 200 pmol each tRNA, and 200 pmol mRNA in buffer AC (20 mM Tris–HCl pH 7.5,

100 mM NH4Cl, 15 mM $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 2 mM spermidine, 0.5 mM spermine). The complex was incubated for 30 min at 37 ˚C and loaded onto a 12 mL 15–30% sucrose gradient in buffer AC. The gradient was spun at 16,500 rpm for 15.75 h at 4 ˚C (SW-41 rotor, Beckman, 46500 g). The gradient was fractionated and the peak corresponding to 70S ribosomes was collected, concentrated in 0.5 mL 100,000 Da cutoff spin filters (Millipore), and washed at ~20:1 ratio in the filter with buffer AC to dilute sucrose. Concentrated complex was flash frozen and stored at –80 ˚C.

MMAL complex cryo-EM grid preparation
300 mesh 1.2/1.3 UltraAuFoil grids from Quantifoil were topped with a layer of amorphous carbon to support the sample. Immediately prior to grid preparation, the ribosome complex was diluted to 100 nM in buffer AC. 4 μL aliquots of sample were deposited onto the glow-discharged grids and incubated for ~1 min. before plunge freezing into liquid ethane.

MMAL complex image processing
Briefly, raw movies were motion-corrected from within RELION 3.0 (Zivanov et al., 2018) using MotionCor2 v. 1.2.2 (Zheng et al., 2017). CTF estimation was performed using CTFFind4 (Rohou & Grigorieff, 2015), and poor micrographs were rejected based on visual inspection. To generate class-average template images for autopicking particles from the full dataset, a round of autopicking with the Laplacian-of-Gaussian method was first performed on a subset of micrographs. The particles picked from this method were then input for 2D classification, and well-resolved 2D classes were used for autopicking templates. 950,634 template-picked particles were input for two rounds of 2D classification, which yielded 532,595 particles. These were migrated to cryoSPARC v2 (Punjani et al., 2017) for 3D heterogeneous refinement with four classes, two of which used an initial reference volume generated from PDB 4V9D (Dunkle et al., 2011) assembly 2 (70S ribosome with tRNA in P site) and the other two using a volume generated from 4YBB assembly 2 (apo 70S ribosome) (Noeske et al., 2015). This yielded a class with largely 50S subunit particles, an apo 70S class with a relatively disordered 30S subunit, a 70S class with some tRNA density, and a junk class. The 319,013 particles in the class with tRNA density were migrated back to RELION for further refinement, involving two rounds of CTF refinement, one round of Bayesian polishing, and 3D classification without alignment, focused on the region of tRNA density. The final refinement was performed with a 50S subunit mask, and the structure was resolved to 3.01 Å by gold-standard Fourier shell correlation (GS-FSC). A charge density map was calculated in Chimera (Pettersen et al., 2004) as previously described (Jimin Wang, 2017).

MMAL complex modeling into the map
An initial atomic model for the MMAL moiety was generated from cyclo-[2]-octylmalonate from the Cambridge Structural Database. Atoms in this model common to the MMAL group were aligned in coot to the terminal O3′ position of a P-site tRNA model, which was derived from PDB 1VY4 (with the N-formylmethionine residue removed), to generate the charged tRNA model for real-space refinement. The cif file for MMAL was generated with phenix.elbow using the geometry from the initial model, and an additional cif file was created manually for the correct geometry of the linkage between MMAL and A76 of the tRNA. The large ribosomal subunit for the model was from PDB 7K00 (Watson et al., 2020), and real-space refinement of

the complex was performed in PHENIX (Liebschner et al., 2019). Additionally, manual adjustment in coot (Casañal et al., 2020) was used to model the conformation of A76 in the deacylated form.

Preparation of $tRNA_{Gly}$
Due to issues with standard tRNA *in vitro* transcription procedures, $tRNA_{gly}$ was synthesized exactly following (Melo Czekster et al., 2016) up until gel purification. The gel was purified as described but on a 12% polyacrylamide gel.

Preparation of $tRNA_{fMet}$ "standard tRNA prep"
Templates for *in vitro* transcription were prepared using a double PCR amplification method. Base C1 was mutated to G for optimal T7 RNA polymerase initiation. First, long overlapping primers (see _temp primers in **Table 3.1**) were PCR amplified using Q5 DNA polymerase (NEB). Products were gel purified and amplified using short primers (see _amp primers in **Table 3.1**). The 2nd base of each reverse primer in this step was modified with 2'OMe to prevent nonspecific addition by T7 RNA polymerase (Kao et al., 1999). Amplification PCRs were 1x phenol chloroform extracted, 2x chloroform washed, and precipitated with 3x volumes of EtOH. DNA quantity was measured on an agarose gel using a known standard.

T7 *in vitro* transcription was performed in a buffer containing 50 mM Tris-HCl, pH 7.5, 15 mM $MgCl_2$, 5 mM dithiothreitol (DTT), 2 mM spermidine. Reactions contained 2.5 mM each NTP, 1:40 NEB murine RNase inhibitor, 25 μg T7 RNA polymerase (gift from Bruno Martinez, Cate lab), 0.0005 U/uL PPase, and ~ 1000 ng of DNA template per 100 μL of transcription reaction. Reactions were incubated for 16 h at 37 ˚C, treated with 1/20th volume RQ1 DNase (1U/uL stock) for 30 min. at 37 ˚C, and precipitated with 1/10th volume 3M NaOAc pH 5.2 and 3x volumes EtOH. Precipitated tRNA was pelleted, washed once with 70% EtOH and resuspended in loading dye (95% formamide, 40 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol).

Gel purification was performed using ~ 20 cm long 12% polyacrylamide 1X TBE 7M urea gels poured roughly 2 mm thick. Bands were excised using UV shadowing, crushed, frozen on dry ice briefly (with elution buffer), and eluted at 4 ˚C overnight in 300 mM NaOAc pH 5.2, 1 mM EDTA, 0.5% w/v SDS. Approximately 2 mL buffer was used per 500 μL transcription reaction. Eluted tRNA was pipetted off the gel debris and precipitated with 1uL of glycoblue coprecipitant and 3x volumes of EtOH. Pelleted tRNA was resuspended in water and stored at -80 ˚C.

Synthesis of N-1 $tRNA_{fMet}$
See primers in **Table 3.1**. Synthesis was otherwise identical to full length $tRNA_{fMet}$.

Purification of A. fulgidus CCA-adding enzyme
The full CCA-adding enzyme gene was purchased from Twist Bioscience and was cloned into a plasmid carrying a T7 promoter and C-terminal 6x histidine tag. Plasmid was transformed into BL21 (DE3) Rosetta2 pLysS cells for expression. 2x 1 L of LB broth + 100 μg/mL ampicillin was induced with a 1:100 dilution of overnight culture (grown at 37 ˚C). Cells were grown at 37 ˚C until OD600 was ~ 0.5 and induced with 0.5 mM IPTG for 3 h. Cells were pelleted and washed with lysis buffer (20 mM HEPES pH 7.5, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 20

mM imidazole) and stored at -80 ˚C. For lysis, cells were resuspended in ~30 mL lysis buffer with a tablet of Pierce EDTA free protease inhibitor. Cells were lysed using a sonicator to deliver ~ 8000 J of energy and lysate was clarified at 18000 rpm in a JA-20 rotor (Beckman) for 30 min at 4 ˚C. Supernatant was applied to a 1 mL HisTrap column recirculating for ~ 30 min. The bound lysate was fractionated on an FPLC with 5 column volumes (CV) of lysis buffer and 10 CV of linear elution gradient from lysis buffer to lysis buffer with 500 mM imidazole. 0.5 mL fractions were collected and analyzed using SDS-PAGE. CCA-adding enzyme containing fractions were combined and concentrated in a 30000 Da MWCO spin filter and buffer exchanged into MS2-1 buffer. Protein was applied to a 1mL heparin column and processed on an FPLC as in chapter 2 "*MBP-MS2 purification.*" Protein containing fractions were combined and dialyzed against 20 mM HEPES, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, 20% glycerol and stored at -80 ˚C. Final concentration was 54 μM in ~ 1 mL (extinction coefficient estimated 55350 M^{-1} cm^{-1}).

3'NH_2-ATP-tailing of N-1 tRNAs

Amino-tailing of tRNAs was adapted from (Katoh & Suga, 2019). Equimolar amounts of N-1 tRNA and *A. fulgidus* CCA-adding enzyme (usually 2 μM) were combined in a reaction containing 100 mM glycine pH 9, 10 mM $MgCl_2$, 1 mM DTT, 0.002 U/uL PPase, and 0.5 mM NH_2-ATP (purchased from Axxora). Reactions were incubated at 37 ˚C for 2 h. Degradation of tRNAs was seen if reactions were incubated past completion. NH_2-tRNAs were extracted with 1:10 volume 3M NaOAc pH 5.2 and 1 volume acidic phenol chloroform, cleaned twice with 1 volume chloroform, and precipitated with 3 volumes EtOH. NH_2-tRNAs were resuspended in water and reaction yield was analyzed on a 10% acrylamide 7 M urea TBE gel (20 cm). 1-2 pmol of tRNA was loaded per lane and the gel was stained with Sybr Green II after running. NH_2-tRNAs were stored in aliquots at -80 ˚C.

Purification of aaRSs and other translation proteins

Expression plasmids of 6xHis-tagged translation enzymes were a gift from Patrick Ginther (Schepartz Lab, UC Berkeley). Plasmids were transformed into BL21 (DE3) Codon+ RIL cells (T7 promoter) or NEB Express Iq cells (T5 promoter) depending on the promoter controlling gene expression. Culture overnights were diluted into ZYM-5052 autoinducing media (Studier, 2014) and expressed overnight at 37 ˚C. Cells were pelleted and resuspended in lysis buffer (20 mM Tris pH 7.8, 150 mM NaCl, 5 mM imidazole, 0.5 mM EDTA) (~ 35 mL). Cells were lysed with a sonicator on ice until ~ 8000 J had been delivered to the sample. Lysate was clarified by centrifugation at 18000 rpm (JA-20 rotor; Beckman) at 4 ˚C for 30 min. Supernatant was applied to a 5 mL HisTrap column and the column was attached to an FPLC. Protein was purified on the FPLC by washing with 5 CV of lysis buffer with 23 mM imidazole and eluting with a linear gradient of 20 CV from 23-500 mM imidazole. Fractions containing the desired protein were pooled, dialyzed overnight against 50mM HEPES pH 7.5, 100 mM KCl, 10 mM $MgCl_2$, 7 mM BME, 30% glycerol, and concentrated in appropriately sized spin filters. Protein was stored at -80 ˚C in aliquots.

Aminoacylation of tRNAs

All aminoacylation reactions were done in a buffer containing 50 mM HEPES pH 7.5, 20 mM $MgCl_2$, 10 mM KCl, 2 mM DTT, 10 mM ATP, 1:40 volume RNase inhibitor (murine, NEB), 5-10 mM amino acid. Aminoacylation enzymes were usually used at 1 μM and tRNAs at 2-5 μM

depending on stock concentration. Generally, a 1:5 enzyme:tRNA ratio was not exceeded. Reactions were incubated at 37 ˚C for 30 min and then extracted and precipitated as in "*3'NH_2-ATP-tailing of N-1 tRNAs*."

Met-Met ribosome complex formation, cryo-EM, and processing

All steps toward the Met-Met ribosome structure were performed essentially as described for the complexes in (Watson et al., 2020). Briefly, the complex was formed by incubating 100 nM 70S ribosomes, 5 μM mRNA, ~ 1.6 μM Met-NH_2-$tRNA_{fMet}$ (2 μM ideally but sample was limited), and 100 μM paromomycin in buffer AC (20 mM Tris pH 7.5, 100 mM NH_4Cl, 15 $MgCl_2$, 0.5 mM EDTA, 2 mM DTT, 2 mM spermidine, 0.05 mM spermine) for 30 min at 37 ˚C. The mRNA sequence GUAUAA**GGAGG**UAAAAAUGAUGUAACUA was synthesized by IDT. (Shine Dalgarno sequence in bold, 2x AUG codons underlined.) Grid preparation, freezing, and data collection were performed as described. During 3D classification in RELION (Zivanov et al., 2018), particles with occupied A+P sites were selected by classification on the A site alone, because A site binding requires an occupied P site. Figures were prepared in ChimeraX (Goddard et al., 2018).

Table 3.1 Primers used ('mG' indicates 2' O-methylguanosine)

tRNA-fMet-C1G_temp F	AATTCCTGCAGTAATACGACTCACTATAGGCGGGGTGGAGCAGCCTGGTAGCTCGTCGGGCTCATA
tRNA-fMet-C1G_temp R	TGGTTGCGGGGGCCGGATTTGAACCGACGACCTTCGGGTTATGAGCCCGACGAGCTA
tRNA-fMet-C1G_amp F	AATTCCTGCAGTAATACGACTCAC
tRNA-fMet-C1G_amp R	TmGGTTGCGGGGGCC
tRNA-fMet-C1G-A_temp R	GGTTGCGGGGGCCGGATTTGAACCGACGACCTTCGGGTTATGAGCCCGACGAGCTA
tRNA-fMet-C1G-A_amp R	GmGCCCCCGCAACC

4. Short peptide reporters for evaluating mutant ribosomes *in vitro*

4.1 Introduction

Imperative to any study of engineered ribosomes is a sensitive method for detecting activity. Whether the experiment is a selection *in vivo* or in depth characterization *in vitro*, there must be an appropriate readout for polymerization activity. Traditionally, variants of the "fragment reaction" have been employed for their high sensitivity and simple scheme. Only a single peptide bond is formed in this reaction, allowing for precise measurement of kinetics. The original fragment reaction required only 50S subunits, ethanol or methanol, puromycin, and a fragment of formylmethionine charged tRNA (Monro & Marcker, 1967). Radiolabeled formylmethionine was transferred to puromycin by the PTC, and this adduct could be extracted and transfer efficiency quantified by scintillation counting. Over the years, this assay has been optimized to improve sensitivity to remarkable levels (Green & Noller, 1996) and has been employed in foundational experiments. For example, key experiments in validating the entropic mechanism of ribosome catalysis were variations of the fragment reaction (Sievers et al., 2004; Youngman et al., 2004). There are, however, limitations to the fragment assay: its readout of peptidyl transfer requires multiple purification and separation steps, increasing the likelihood of errors, there are noisy side reactions that can occur in the reaction mixture, and its reliance on radioactivity is often an insurmountable obstacle as many labs move away from the labeling technique. Simplifications to this assay have been developed, but it has been impossible to avoid radiolabeling to detect the minute quantities of single peptidyl transfer adducts formed (Polacek et al., 2002).

On a larger scale, reporter proteins such as fluorescent proteins, luciferases, or other enzymes can read out protein synthesis activity. These systems can be used *in vivo* or *in vitro* and are straightforward to execute and quantify. However, they require synthesis of long α-amino acid containing proteins, a chemical process that may not be supported by engineered ribosomes. We have previously used *in vitro* luciferase translation assays to characterize mutant ribosomes, revealing perplexing differences to *in vivo* or lysate-based systems (Ward et al., 2019). Another method for detecting short translational products is the use of liquid chromatography and mass spectrometry (LC-MS) to detect short peptides after *in vitro* translation. This technique is especially powerful because precise mass analysis can be used to infer the incorporation of unnatural monomers by ribosomes. This technique is in widespread use by researchers probing the chemical space wild type ribosomes can polymerize (Ad et al., 2019; Katoh & Suga, 2018; Lee, Schwarz, et al., 2020). This technique suffers from throughput problems, though, because products must often be purified out of translation reactions and run in series on LC-MS instruments, expensive machines in limited supply in academic laboratories.

Here, we aimed to develop a middle-ground assay that combines the sensitivity and simple measurement of enzymatic readout with the minimal polymerization requirement of short peptide translation. Adapting a split luciferase system for *in vitro* translation conditions using the PURExpress defined system, we successfully rescued activity in a PTC mutant ribosome. This result begins to reconcile the marked differences in translation activity between *in vivo*, *in vitro* lysate-based, and *in vitro* defined systems, as detailed in chapter 2. From the early development progress, we anticipate that this assay will be useful in the development and understanding of mutant ribosomes due to its simple design, small translation product, and improved throughput.

4.2 Results

Selection of the HiBiT split nanoluciferase system

A split protein system was a rational choice when we set out to improve translation reporter assays used in our laboratory. In these systems, a protein is separated into two chains such that its activity is dependent on the reassociation of the split pieces. These reporters have been a boon for studying protein-protein interactions, especially when the affinity of the split pieces is so low that meaningful reassociation is contingent on the interaction of attached partners (Shekhawat & Ghosh, 2011). For our assay, we sought split proteins with several characteristics. Most importantly, the split site needed to be near an end of the protein, leaving only a small peptide as one of the interacting pieces. Mutant ribosomes could then be programmed with an mRNA encoding this peptide with the larger piece supplied as an additional reagent. Translation of only the short peptide would restore detectable enzyme activity. Secondly, the affinity between the two pieces must be high enough to make the assay sensitive during *in vitro* translation. The average protein yield of the PURExpress system is ~ 1-10 μM (Tuckey et al., 2014), with a small peptide likely translated to higher efficiency. Contingent on the dynamic range of the protein's readout, we sought a split system with a K_d of < 1 μM, but preferably significantly lower to accurately detect weak translation. Thirdly, the maturation time of the two pieces should be short to ensure a rapid assay, possibly allowing real-time measurement of translation. We considered short peptide complementation systems from three different reporters: β-galactosidase (Nishiyama et al., 2015), superfolder GFP (Cabantous et al., 2005), and nanoluciferase (nLuc) (Dixon et al., 2016). The β-galactosidase fragment could only be shortened to 44 amino acids, which we deemed too long for mutant PTC ribosomes. The split GFP's C-terminal fragment was only 17 amino acids, but maturation times for the protein were hours long. The nanoluciferase system, however, consisted of a 156 amino acid N-terminal and a 12 amino acid C-terminal fragment, the shortest complementation piece yet developed (Dixon et al., 2016). Complementation of enzyme activity was fast and the large dynamic range of nLuc was preserved. Furthermore, the short peptide had been engineered for both low and high affinity to the large substrate, ranging from 700 pM (named HiBiT) to 190 μM (named SmBiT). With its small size, fast complementation, high affinity, and exceptional report dynamic range, this split nLuc system was ideal for our assay development.

A split nLuc system for PURExpress *in vitro* translation

We purified the large N-terminal nLuc fragment (called 11S) and prepared DNA templates of the SmBiT and HiBiT fragments. We first sought to find an appropriate 11S concentration to detect peptides translated *in vitro*. PURExpress *in vitro* translation systems were programmed with the DNA fragments and wild type ribosomes, followed by addition of nLuc substrate and varying concentrations of 11S protein. Initial assay development was performed with the SmBiT low affinity peptide. Switching to the HiBiT high affinity peptide showed improvement in sensitivity but the overall dynamics of the assay were similar between the two. Background in the absence of complementary peptide was low and attributable to the 11S fragment, which retains low but detectable luciferase activity (**Figure 4.1**). Activation of luciferase was immediate, many orders of magnitude over baseline, and linear with respect to 11S concentration.

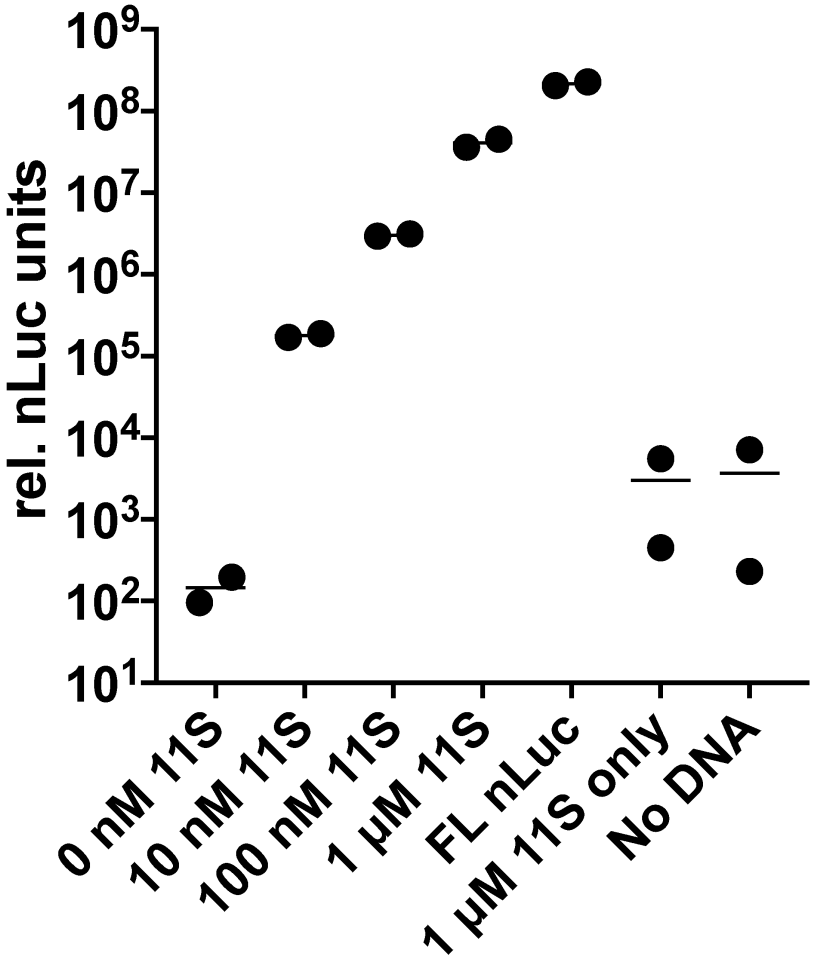


Figure 4.1. SmBiT peptide *in vitro* translation titration with 11S protein. Titration of 11S concentrations into a translation reaction programmed with SmBiT DNA. FL nLuc: Full Length nanoluciferase positive control. 11S only: no translation reaction, only 11S protein in buffer. No DNA: translation reaction not programmed with DNA. Variances in 11S only and No DNA negative control attributable to crosstalk from adjacent highly luminescent wells.

We then tested a suite of PTC mutant ribosomes purified by MS2 affinity tagging, all previously inactive in full length nLuc translation, for activity in this assay. While most ribosomes were just as inactive, the 040329 ribosome engineered for lysate-based *in vitro* translation of β-amino acids was significantly above baseline (**Figure 4.2**) (Dedkova et al., 2012). To our knowledge, this is the first time a pure PTC mutant ribosome has shown activity in the PURExpress system. Previously, the same ribosome unpurified from WT background was used to weakly incorporate aminocyclobutane-carboxylic acid, only detected by LC-MS (Lee, Schwarz, et al., 2020).

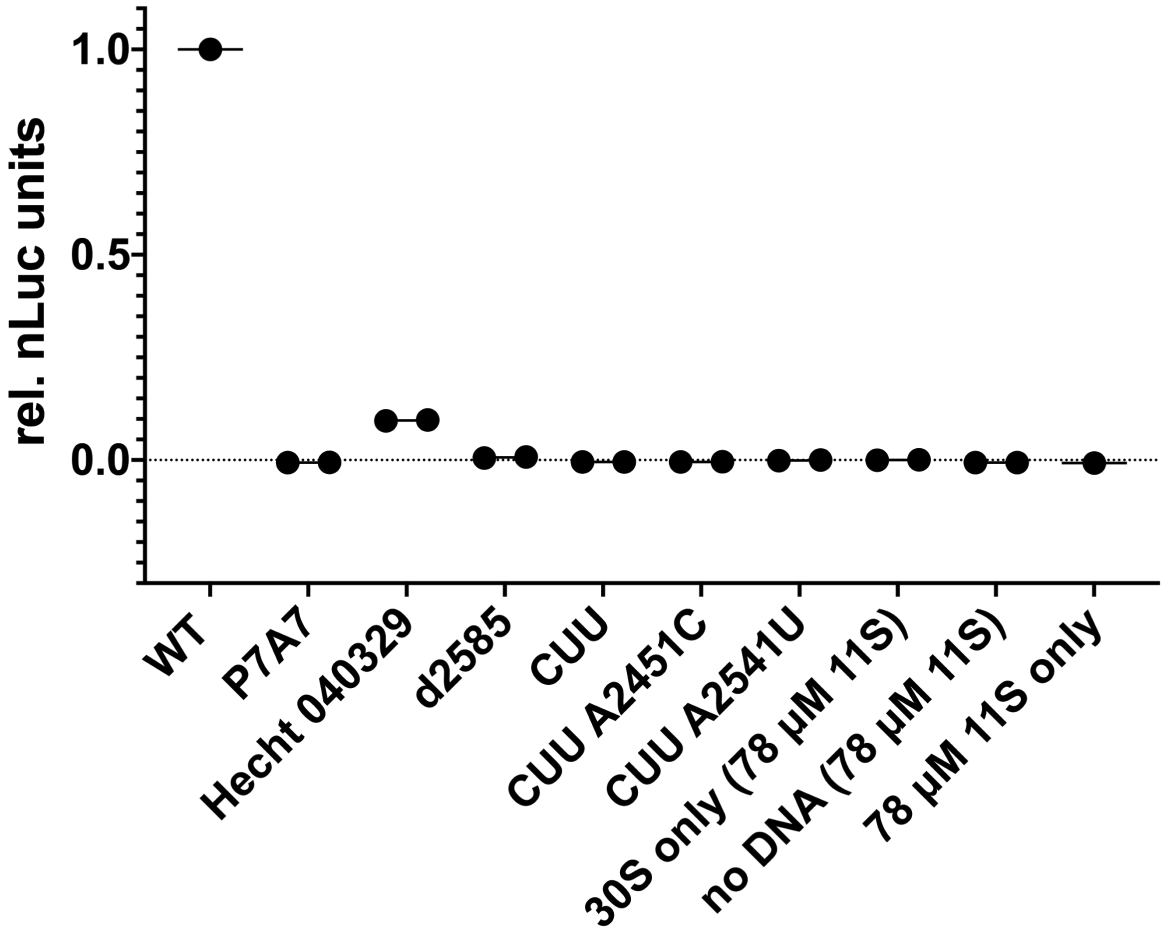


Figure 4.2. Mutant ribosomes in the short peptide assay. Mutant ribosome SmBiT translation after 2 hours incubation. Luminescence normalized to WT signal and 30S only signals as 1.0 and 0.0, respectively. Mutant ribosomes: d2585, deletion of U2585; CUU, mutations A2450U and C2501U. No DNA and 11S only as in Figure 4.1. All reactions had 78 μM 11S added at time of quantification. Reactions run in duplicate except WT and 11S only which were single reactions.

Next, we established a real-time translation assay using our short peptide reporters. In this assay, nLuc luminescence is continuously monitored in a 384-well plate reader at 37 ˚C. Careful preparation of the translation mixtures on ice produces stable baselines, with activity appearing several minutes after incubation begins in the plate reader (**Figure 4.3A**). Accurate pipetting and experiment design allowed using reactions of only 1.8 μL per replicate, spotted in the corner of 384 well plates, conserving valuable PURExpress reagents. This assay can more directly report on translation rates than an endpoint assay, with the initial slope of the luminescence curve reflecting translation of the complementing peptide until nLuc substrate is consumed. We confirmed the activity of the 040329 (Hecht) ribosome using this time course assay (**Figure 4.3A**). The initial slope of the 040329 translation curve was 22% as steep as the equivalent reaction with the wild type ribosome (**Figure 4.3B**). Re-analysis of the purity of the 040329 ribosome preparation revealed only 5% contamination with WT ribosomes (**Figure 4.3C**). To verify that the increased signal was not due to unexpected nonlinearity in our translation system, we included the same translation reactions at varying concentrations of WT ribosomes. The

assay responded linearly to ribosome concentration, suggesting that translation by the 040329 ribosome represents true peptide synthesis by a PTC mutant that cannot be explained by WT contamination or assay dynamics (**Figure 4.3A**).

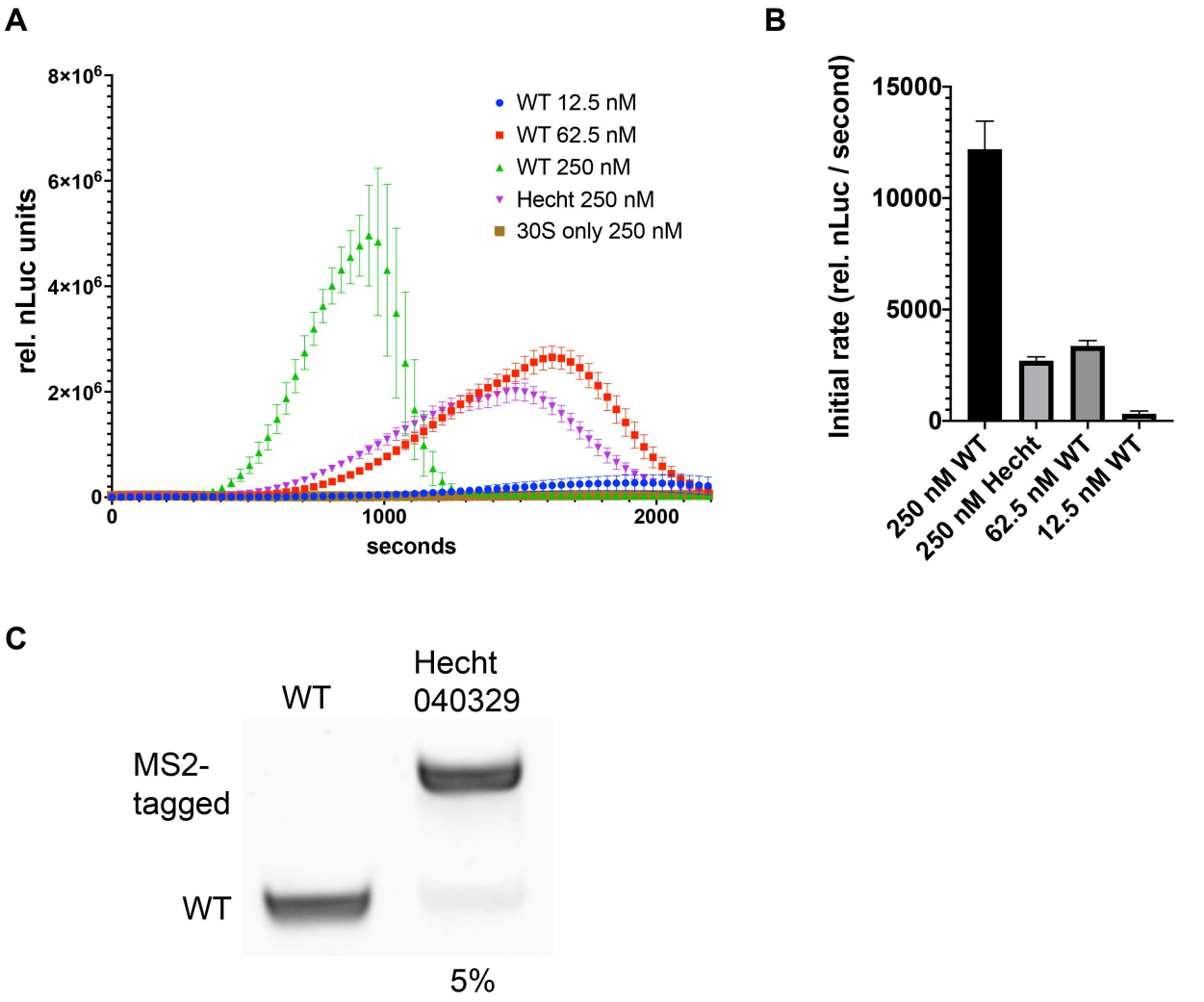


Figure 4.3. PTC mutant ribosome 040329 is active in PURExpress with SmBiT reporter. A) Time course luminescence curves for translation reactions with varying concentrations of WT ribosomes and Hecht 040329. B) Linear fits of the initial slope of each curve in A) (except 30S only negative control). Error bars represent 95% CI of linear fit. C) RT-PCR quantification of MS2-tagged fraction in a WT ribosome preparation and Hecht 040329 preparation. Quantification of WT/total band intensity listed below Hecht lane.

Switching our assay from the low affinity SmBiT peptide (K_D= 190 μM) to the high affinity HiBiT (K_D = 700 pM) reduced background by allowing lower quantities of 11S to return the same signal. Using 100 times less 11S gave the same initial signal during short peptide translation with HiBiT than standard SmBiT conditions (**Figure 4.4A**). Averaging the stable baseline at the start of each reaction (5 min) showed markedly less background signal for the 0.1 μM-11S-HiBiT combination that matched the SmBiT signal (**Figure 4.4B**). Together, these data support using the HiBiT peptide in the assay, with the potential to tune the dynamic range by adding more 11S as needed.

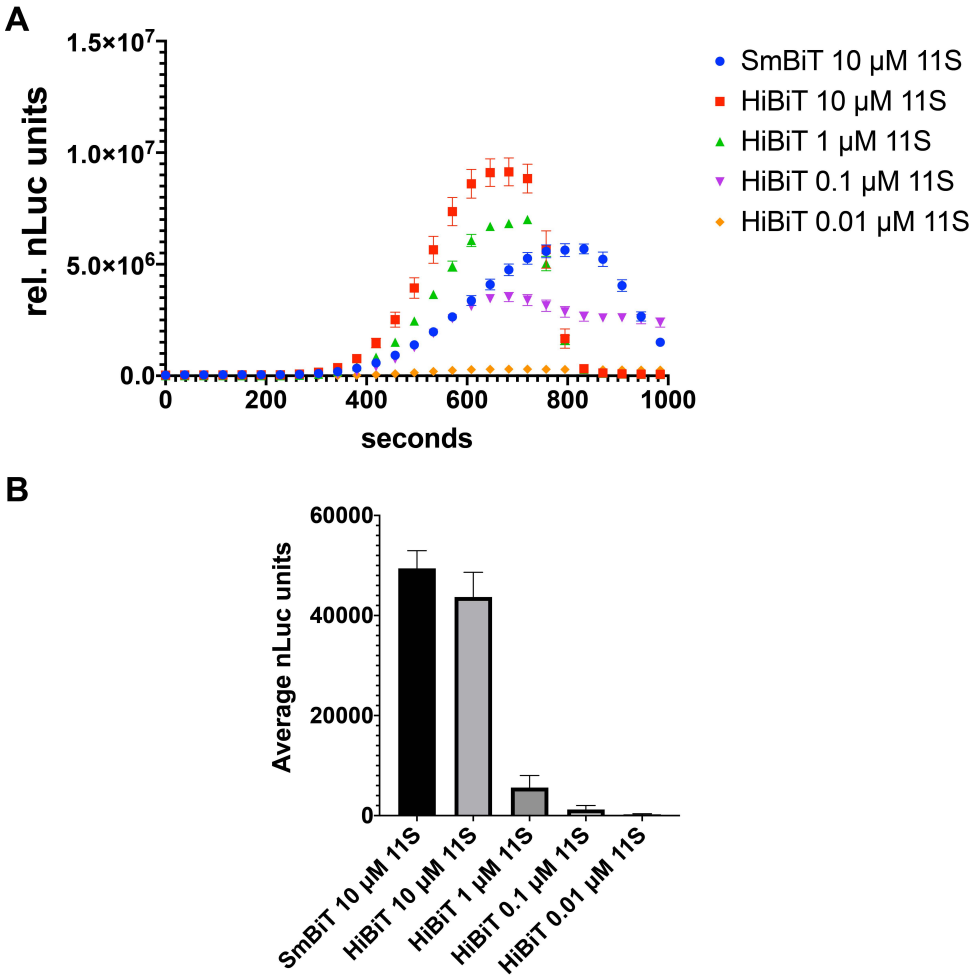


Figure 4.4. SmBit versus HiBiT short peptide. A) Time course curves of SmBiT peptide translation reactions at standard 10 μM 11S concentration versus HiBiT peptide reactions at varying 11S concentrations. B) Average of the first 5 minutes of baseline luminescence levels in each translation reaction. Error bars represent standard error of the mean.

Finally, we examined the ease of opening codons in this assay. The HiBiT peptide contains 10 different amino acids, each of which we withheld during a separate *in vitro* translation. We expected that withholding an amino acid should greatly slow the translation rate of the HiBiT peptide, but instead we saw very little deviation from the all amino acid control as we withheld each substrate (**Figure 4.5A**). This may be due to small amounts of contamination in the PURExpress Δ-AA-tRNA-ribosome kit that we used, so we incubated the *in vitro* translation reactions for 2 h and added excess 11S and nLuc substrate in an endpoint measurement. There, we saw decreased translation when withholding the first amino acid, methionine, but not the second, valine (**Figure 4.5B**). It seems likely there is a combination of contamination and miscoding/mischarging occurring in these amino acid starved reactions. Especially because there are no ribosome rescue mechanisms in the PURExpress mixture, ribosomes will remain stalled on an empty A site for longer than *in vivo*, increasing the chance of translational errors. Further investigation is needed to elucidate the source of contamination or the cause of translation errors that allow robust translation in the absence of necessary amino acids. This is of particular importance when introducing suppressor tRNAs charged with unnatural monomers. Even a small

amount of natural contamination would be charged to the available cognate tRNA and hinder suppression at the codon.

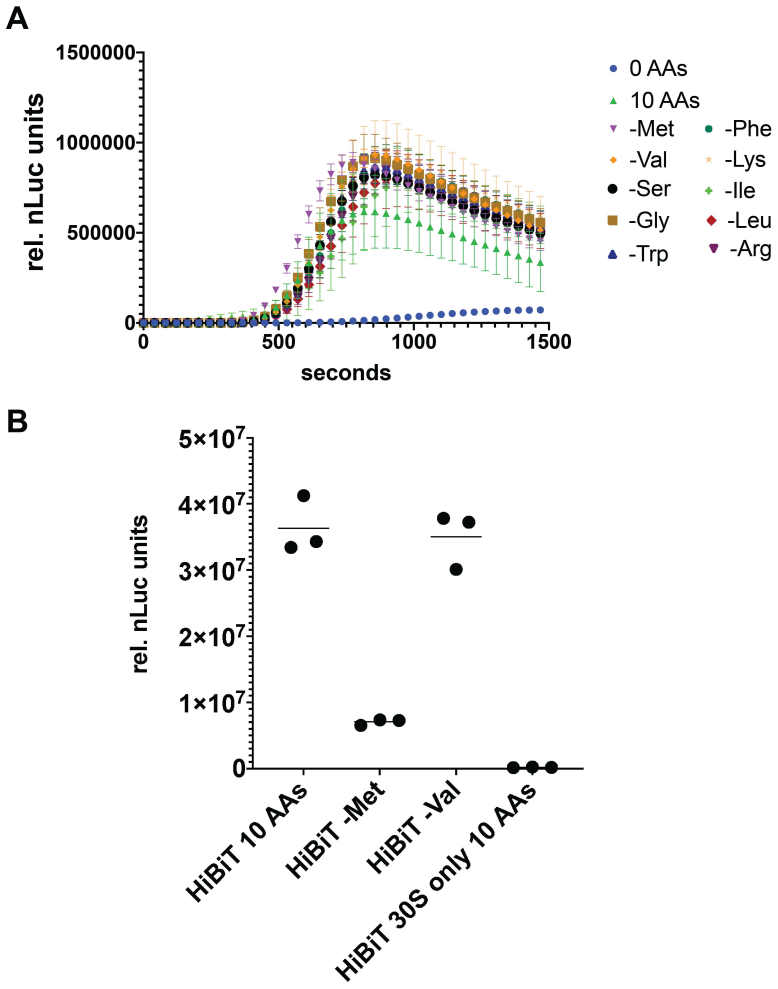


Figure 4.5. Withdrawn amino acids in the HiBiT short peptide translation system. A) Time course luminescence curves of HiBiT translation in the PURExpress Δ-AA-tRNA-ribosome system. 0 AAs, negative control with no amino acids added to translation. 10 AAs, all 10 amino acids in the HiBiT peptide added. B) Endpoint measurement after 2 h of translation.

4.3 Discussion

The HiBiT reporter peptide explored in this chapter is sensitive and short, making it ideal for mutant ribosome characterization. With this assay, we have uncovered *in vitro* PTC mutant activity in the PURExpress system for the first time, confirming our hypothesis that the length of common reporter proteins was a primary contributor to observed inactivity. Furthermore, this system has many advantages that should allow screening of new monomers and other types of genetic code expansion. First, the peptide can tolerate both N- and C-terminal extensions and still retain complementation activity. Both the primary publication and a host of later studies have used the tag in a variety of configurations and with linkers of several sizes (Dixon et al., 2016; Oh-Hashi et al., 2017; Pereira et al., 2019; Schwinn et al., 2018). Our HiBiT assay itself appends an N-terminal methionine as the start codon to no ill effects. These results suggest that appending

suppressed codons to the start of the peptide should be feasible, allowing incorporation of unnatural monomers before the peptide. Such a system could be used to rapidly screen ribosomes or monomer sets for activity, with hits confirmed for accurate incorporation using traditional LC-MS techniques. The HiBiT peptide itself has been trimmed to even smaller size, albeit with commensurate loss in activity (Dixon et al., 2016). However, given the high dynamic range of nLuc, even 1 or 2 orders of magnitude less activity could be tolerated for a less translationally demanding reporter peptide.

The most perplexing result acquired was the apparent contamination of the PURExpress kits with residual amino acid or charged tRNA. Small contaminations pose a significant threat to researchers working near the detection limit of instruments, as they often are when investigating new chemistry on the ribosome. Strong codon suppression is a necessary characteristic for a useful *in vitro* translation assay. Timepoint data suggests that, at least for some amino acids or codons, withdrawn substrates do eventually affect translation efficiency. We are pursuing LC-MS characterization of peptides produced under withdrawn amino acid conditions in the hope of understanding these results.

This leads into a second problem with the HiBiT assay: time course experiments can only be primed with enough substrate for ~ 10-20 minutes of translation. This problem cannot be easily solved by adding more substrate due to solubility and dilution concerns. Thus, there may be benefit in combining time course with endpoint analyses to observe initial rates and then quantify total product formation with excess reagent and 11S afterwards.

Lastly, there are useful applications of this system to *in vivo* ribosome engineering efforts. Current methods use negative selection with puromycin analogs or positive selection with full length proteins with suppressor-tRNA-incorporated unnatural monomers, usually fluorescent reporters or antibiotic resistance enzymes. As we observed *in vitro*, requiring synthesis of a long α-amino template during a selection for new chemical ability may be too demanding of mutant ribosomes. These experiments run the risk of missing successful mutants because they cannot effectively translate the α-amino portion of the selective protein. If the HiBiT systems can be adapted to accommodate N- or C-terminal unnatural monomer extensions, it could serve as a useful *in vivo* screening tool. Recent developments in *in vivo* bioluminescent resonance energy transfer (BRET) reporters show that nLuc systems can be used reliably in *E. coli* (Dippel et al., 2020). Furthermore, BRET technology can combine the high intensity of fluorescent proteins with the low background of luciferases, allowing cell sorting of ribosome libraries using converted luminescence signals (Schaub et al., 2015). While more interrogation and optimization is required, this HiBiT assay is already a promising tool for ribosome engineers.

4.4 Materials and Methods

11S fragment purification

The sequence of the 11S fragment (Dixon et al., 2016) was ordered from IDT with an N-terminal 8x histidine tag. 8xHis-11S DNA was cloned into vector 2b (Macrolab, UC Berkeley) using isothermal DNA assembly after a T7 promoter. Successfully assembled clones were transformed into BL21 (DE3) Rosetta2 pLysS cells for expression. An overnight culture of the cells was diluted 1:100 into 2x 1L of LB + 100 μg/mL ampicillin and grown at 37 ˚C. OD600 was monitored until cells reached ~ 0.4, when cells were induced with 1 mM IPTG for 3 hours. Cells were pelleted, washed in lysis buffer (20 mM HEPES pH 7.5, 50mM KCl, 10% glycerol, 10 mM imidazole), pelleted again and stored at -80 ˚C. Thawed cells were resuspended in ~ 20 mL of

lysis buffer and lysed on ice with a sonicator (~ 3500 J of energy delivered). Lysate was clarified by centrifugation at 18k rpm for 30 min. at 4 ˚C (JA-20 rotor, Beckman). Supernatant was removed and applied to a 1 mL HisTrap column recirculating for 30 min at 4 ˚C. The column was loaded onto an FPLC, washed with 10 mL of lysis buffer, and eluted with a linear gradient of lysis buffer from 20-500 mM imidazole. Fractions containing the large elution peak were collected, analyzed by SDS-PAGE, and 11S containing fractions were dialyzed against lysis buffer without imidazole in a 2000 Da cutoff dialysis cassette overnight at 4 ˚C. The next day, 11S protein was frozen in aliquots at -80 ˚C.

SmBiT and HiBiT template preparation
Templates for both peptides were ordered from IDT. DNA templates were amplified using T7_smbit_F/R primers and Q5 DNA polymerase (**Table 4.1**). PCR reactions were phenol chloroform extracted and precipitated. DNA was quantified on an agarose gel against a known standard.

MS2-tagged ribosome preparation
See methods in chapter 2.

Endpoint in vitro translation reactions
The PURExpress Δ-ribosome or Δ-AA-tRNA-ribosome (NEB) kits were used for all *in vitro* translation reactions. If amino acids were withdrawn in the experiment, the Δ-AA-tRNA-ribosome kit was used. Translation reactions were assembled according to the manufacturer's protocols except reactions were significantly scaled down from 25 µL to 1.8 µL. All reactions were run in triplicate unless otherwise specified. 250 nM 50S MS2-tagged ribosomes (MS2-tagged WT or MS2-tagged mutant) was the default concentration unless otherwise specified. MS2-tagged ribosomes were used even for wild type reactions as a control for the tag and purification conditions. 500 nM 30S ribosomes were included as a two-fold excess. If a ribosome was not included, an equivalent amount of ribosome prep buffer A with 1 mM $MgCl_2$ was used (chapter 2). DNA templates were included at 10 ng/uL. Amino acid mixtures were prepared at 3 mM each amino acid and included at 300 µM final concentration each amino acid. Only the 10 amino acids in the HiBiT peptide, or withdrawn compositions thereof, were included in Δ-AA-tRNA-ribosome PURExpress reactions. Reactions were incubated for 30 min at 37 ˚C (Figures 4.1 and 4.2) or 2 h at 37 ˚C (Figure 4.5B). After incubation, reactions were diluted with 30 µL of the indicated concentration of 11S (10 µM in Figure 4.5) and either a 1:50 dilution of nLuc substrate (Promega) (Figures 4.1 and 4.2) or 200 µM furimazine (Figure 4.5) and transferred to 384 well plates. When possible, empty wells were left between conditions to minimize crosstalk between wells. Luminescence was measured on a Tecan Spark plate reader.

Time course in vitro translation reactions.
Reactions were assembled similarly to endpoint reactions, with extra care to keep preparations on ice as often as possible. 10 µM 11S was included in SmBiT peptide reactions and 0.1 µM 11S in HiBiT peptide reactions. For all time courses, furimazine was included at 0.1 µM from a DMSO stock. Mixtures were carefully spotted in triplicate into the lower left corner of wells on a cold 384 well plate. The plate was transferred to a preheated (37 ˚C) Tecan Spark plate reader and luminescence of the wells was monitored continuously for 1 h or until all curves returned to

baseline. As before, empty wells were left between conditions to minimize crosstalk when possible.

Data analysis
Graphing, linear fitting, and other analysis was performed with Graphpad Prism software.

Table 4.1 Primers and sequences used

8x-His-11S sequence	atgcaccatcaccatcaccatcaccatggttcttctatggtgtttaccctggaagattttgtgggc gattgggaacagaccgcggcgtataacctggatcaggtgctggaacagggcggcgtgagc agcctgctgcagaacctggcggtgagcgtgaccccgattcagcgcattgtgcgcagcggcg aaaacgcgctgaaaattgatattcatgtgattattccgtatgaaggcctgagcgcggatcagat ggcgcagattgaagaagtgtttaaagtggtgtatccggtggatgatcatcattttaaagtgattct gccgtatggcaccctggtgattgatggcgtgaccccgaacatgctgaactattttggccgccc gtatgaaggcattgcggtgtttgatggcaaaaaaattaccgtgaccggcaccctgtggaacg gcaacaaaattattgatgaacgcctgattaccccggatggcagcatgctgtttcgcgtgaccat taacagctaa
SmBiT T7 template	GCGAATTAATACGACTCACTATAGGGTTAACTTTAACAAG GAGAAAAACATGgtgaccggctatcgcctgtttgaagaaattctgtaaCTAGC ATAACCCCTCTCTAAACGGAGGGGTTT
HiBiT T7 template	GCGAATTAATACGACTCACTATAGGGTTAACTTTAACAAG GAGAAAAACATGgtgagcggctggcgcctgtttaaaaaaattagctaaCTAGC ATAACCCCTCTCTAAACGGAGGGGTTTAGTCA
T7_smbit_F	GCGAATTAATACGACTCACTATAG
T7_smbit_R	AAACCCCTCCGTTTAGAG

5. Conclusion and Future Outlook

The ribosome is the most capable polymer synthesis machine known, routinely synthesizing hundred-mer peptides with high efficiency and exquisite sequence definition. Repurposing this platform towards other polymer chemistries would dramatically improve over current techniques and unlock new classes of materials. To this end, we have presented research in understanding first-generation engineered ribosomes, obtaining structures that reveal the limitations of wild type ribosomes, and developing new assays for *in vitro* translation with PTC mutant ribosomes. First, we uncovered deep assembly defects in a ribosome engineered for better incorporation of β-amino acids (Ward et al., 2019), highlighting the importance of screens and selections that optimize ribosome assembly as well as new catalytic ability. Second, we obtained the first structure of a non-amino acid monomer in the PTC. Its position in the PTC explains its activity as an initiator substrate. Along the way, we were able to push ribosome structural biology to record levels of detail (Watson et al., 2020). Third, we have developed a less translationally demanding, high sensitivity assay for defined *in vitro* translation that can rescue activity in PTC mutant ribosomes that previously only functioned *in vivo* or in lysate. These results lay the groundwork for our platform of ribosome design, characterization, and analysis crucial to achieving our long-term goals in ribosome engineering.

However, there are still many open questions, both in direct relation to the work presented here and for ribosome engineering as an endeavour in general, that must be addressed before repurposed ribosomes are accessible. We hypothesize that ribosome assembly will remain a limiting factor for the efficiency of PTC mutants for the time being. It is unfortunate that that PTC assembly only occurs at the end of the entire tightly orchestrated large subunit assembly process. This necessitates that remodeled PTCs do nothing to inhibit the complex network of assembly factors, ribosomal proteins, and rRNA that work to assemble every other piece of the ribosome first. One possible solution to this problem is to build a model of ribosome assembly to extreme detail. Understanding precisely which contacts, regions, and even individual residues are important for efficient ribosome assembly and function could identify “no-go” zones for engineers, allowing maximal mutational exploration with minimal impact to assembly (Walker et al., 2020). Furthermore, structural studies of ribosome assembly will continue to improve our “movie” of ribosome construction (Davis et al., 2016; Nikolay et al., 2018, 2021). A second solution is to extricate the PTC and its function from the rest of the ribosome, distilling the ribosome to its core catalytic activity. The PTC is considered one of the oldest elements that make up the modern ribosome, likely evolving from one or more small RNA fragments (Petrov et al., 2014). There is preliminary evidence that isolated sections of PTC-encoding RNA can catalyze the formation of a few peptide bonds, suggesting that an isolated PTC molecule may be accessible *in vitro* for simplified directed evolution (D. Xu & Wang, 2021). Thirdly, it may be that well-designed selections in systems with strong unnatural monomer tRNA charging, ribosome orthogonality, and multiple selective modes may discover robust repurposed ribosomes. Maybe there just has not been a suitable platform developed to explore at the depth and with the selective stringency required.

Synthesis and purification of non-L-α-amino tRNAs is a continual challenge for nearly any experiment in the field. Mutant aaRSs and flexizymes are large steps forward but neither of them is structure-agnostic nor routinely highly efficient. This is of particular importance for structural studies, where uncharged tRNA may be preferentially incorporated into the PTC if monomers are especially bulky or inflexible. Purification techniques, usually some form of anion-exchange

or reverse phase chromatography will be essential for obtaining substrates useful in cryo-EM. Despite their inefficiencies, flexizymes coupled with chromatographic purification have the potential to offer "one-size-fits-all" tRNA substrate preparation.

A final challenge is bridging the gap between α-amino acid translation and new chemistry. The platforms that allow selections over large libraries are usually *in vivo* systems where mutant ribosomes must coexist in an α-amino world. Orthogonal tethered ribosomes (Carlson et al., 2019; Schmied et al., 2018) prevent ribosome libraries from interfering with normal cellular function, but these mutant ribosomes must still synthesize a readout molecule. In the context of a living cell, this molecule is usually a reporter protein or an enzyme that confers a fitness advantage, all of which are made of α-amino acids. As an example, a screen might ask mutant ribosomes to incorporate one or more unnatural monomers into GFP using a suppressor tRNA. There are 238 amino acids in GFP, so most of the bonds these library members will form will be normal peptide bonds. Such a selection will necessarily return ribosomes at least marginally capable of standard peptide synthesis. This obligate promiscuity will obscure the library members capable of new chemistry but incapable of making standard proteins. Our work presented here on small peptide reporters suggests a way out of this problem. Limiting the α-amino translational demand on engineered ribosomes prevents unwanted selection for wild type function. Future improvements could involve developing enzyme complements or other short signal-producing oligomers, such as receptor binding mimics and other foldamers, with no α-amino components as a way to remove standard peptide synthesis from the engineering equation entirely. The chapters presented here highlight the importance of careful engineering and the power of structural understanding in synthetic biology.

References

Achenbach, J., Jahnz, M., Bethge, L., Paal, K., Jung, M., Schuster, M., Albrecht, R., Jarosch, F., Nierhaus, K. H., & Klussmann, S. (2015). Outwitting EF-Tu and the ribosome: translation with d-amino acids. *Nucleic Acids Research*, *43*(12), 5687–5698.

Ad, O., Hoffman, K. S., Cairns, A. G., Featherston, A. L., Miller, S. J., Söll, D., & Schepartz, A. (2019). Translation of Diverse Aramid- and 1,3-Dicarbonyl-peptides by Wild Type Ribosomes in Vitro. *ACS Central Science*, *5*(7), 1289–1294.

Aleksashin, N. A., Leppik, M., Hockenberry, A. J., Klepacki, D., Vázquez-Laslop, N., Jewett, M. C., Remme, J., & Mankin, A. S. (2019). Assembly and functionality of the ribosome with tethered subunits. *Nature Communications*, *10*(1), 930 1–13.

Aleksashin, N. A., Szal, T., d'Aquino, A. E., Jewett, M. C., Vázquez-Laslop, N., & Mankin, A. S. (2020). A fully orthogonal system for protein synthesis in bacterial cells. *Nature Communications*, *11*(1), 1858.

Ambrogelly, A., Palioura, S., & Söll, D. (2007). Natural expansion of the genetic code. *Nature Chemical Biology*, *3*(1), 29–35.

Asai, T., Zaporojets, D., Squires, C., & Squires, C. L. (1999). An Escherichia coli strain with all chromosomal rRNA operons inactivated: Complete exchange of rRNA genes between bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(5), 1971–1976.

Balakin, A. G., Skripkin, E. A., Shatsky, I. N., & Bogdanov, A. A. (1992). Unusual ribosome binding properties of mRNA encoding bacteriophage lambda repressor. *Nucleic Acids Research*, *20*(3), 563–571.

Ban, N., Nissen, P., Hansen, J., Capel, M., Moore, P. B., & Steitz, T. A. (1999). Placement of protein and RNA structures into a 5 A-resolution map of the 50S ribosomal subunit. *Nature*, *400*(6747), 841–847.

Baumann, S., Herrmann, J., Raju, R., Steinmetz, H., Mohr, K. I., Hüttel, S., Harmrolfs, K., Stadler, M., & Müller, R. (2014). Cystobactamids: myxobacterial topoisomerase inhibitors exhibiting potent antibacterial activity. *Angewandte Chemie* , *53*(52), 14605–14609.

Bayryamov, S. G., Rangelov, M. A., Mladjova, A. P., Yomtova, V., & Petkov, D. D. (2007). Unambiguous evidence for efficient chemical catalysis of adenosine ester aminolysis by its 2'/3'-OH. *Journal of the American Chemical Society*, *129*(18), 5790–5791.

Bieling, P., Beringer, M., Adio, S., & Rodnina, M. V. (2006). Peptide bond formation does not involve acid-base catalysis by ribosomal residues. *Nature Structural & Molecular Biology*, *13*(5), 423–428.

Biyani, N., Righetto, R. D., McLeod, R., Caujolle-Bert, D., Castano-Diez, D., Goldie, K. N., & Stahlberg, H. (2017). Focus: The interface between data collection and data processing in cryo-EM. *Journal of Structural Biology*, *198*(2), 124–133.

Blanchard, S. C., Kim, H. D., Gonzalez, R. L., Jr, Puglisi, J. D., & Chu, S. (2004). tRNA dynamics on the ribosome during translation. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(35), 12893–12898.

Boni, I. V., Isaeva, D. M., Musychenko, M. L., & Tzareva, N. V. (1991). Ribosome-messenger recognition: mRNA target sites for ribosomal protein S1. *Nucleic Acids Research*, *19*(1), 155–162.

Brown, A., & Shao, S. (2018). Ribosomes and cryo-EM: a duet. *Current Opinion in Structural Biology*, *52*, 1–7.

Cabantous, S., Terwilliger, T. C., & Waldo, G. S. (2005). Protein tagging and detection with

engineered self-assembling fragments of green fluorescent protein. *Nature Biotechnology*, *23*(1), 102–107.
Carlson, E. D., d'Aquino, A. E., Kim, D. S., Fulk, E. M., Hoang, K., Szal, T., Mankin, A. S., & Jewett, M. C. (2019). Engineered ribosomes with tethered subunits for expanding biological function. *Nature Communications*, *10*(1), 3920.
Casañal, A., Lohkamp, B., & Emsley, P. (2020). Current developments in Coot for macromolecular model building of Electron Cryo-microscopy and Crystallographic Data. *Protein Science: A Publication of the Protein Society*, *29*(4), 1069–1078.
Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N., & Noller, H. F. (1999). X-ray crystal structures of 70S ribosome functional complexes. *Science*, *285*(5436), 2095–2104.
Chang, B., Halgamuge, S., & Tang, S.-L. (2006). Analysis of SD sequences in completed microbial genomes: non-SD-led genes are as common as SD-led genes. *Gene*, *373*, 90–99.
Chapeville, F., Lipmann, F., Von Ehrenstein, G., Weisblum, B., Ray, W. J., Jr, & Benzer, S. (1962). On the role of soluble ribonucleic acid in coding for amino acids. *Proceedings of the National Academy of Sciences of the United States of America*, *48*, 1086–1092.
Chen, S. S., Sperling, E., Silverman, J. M., Davis, J. H., & Williamson, J. R. (2012). Measuring the dynamics of E. coli ribosome biogenesis using pulse-labeling and quantitative mass spectrometry. *Molecular bioSystems*, *8*(12), 3325–3334.
Chin, J. W. (2014). Expanding and reprogramming the genetic code of cells and animals. *Annual Review of Biochemistry*, *83*, 379–408.
Chubiz, L. M., & Rao, C. V. (2008). Computational design of orthogonal ribosomes. *Nucleic Acids Research*, *36*(12), 4038–4046.
Clemons, W. M., Jr, May, J. L., Wimberly, B. T., McCutcheon, J. P., Capel, M. S., & Ramakrishnan, V. (1999). Structure of a bacterial 30S ribosomal subunit at 5.5 A resolution. *Nature*, *400*(6747), 833–840.
Cociorva, D., L Tabb, D., & Yates, J. R. (2007). Validation of tandem mass spectrometry database search results using DTASelect. *Current Protocols in Bioinformatics / Editoral Board, Andreas D. Baxevanis ... [et Al.]*, *Chapter 13*, Unit 13.4 1–14.
Dandey, V. P., Budell, W. C., Wei, H., Bobe, D., Maruthi, K., Kopylov, M., Eng, E. T., Kahn, P. A., Hinshaw, J. E., Kundu, N., Nimigean, C. M., Fan, C., Sukomon, N., Darst, S. A., Saecker, R. M., Chen, J., Malone, B., Potter, C. S., & Carragher, B. (2020). Time-resolved cryo-EM using Spotiton. *Nature Methods*, *17*(9), 897–900.
Daniels, D. S., Petersson, E. J., Qiu, J. X., & Schepartz, A. (2007). High-resolution structure of a beta-peptide bundle. *Journal of the American Chemical Society*, *129*(6), 1532–1533.
d'Aquino, A. E., Azim, T., Aleksashin, N. A., Hockenberry, A. J., Krüger, A., & Jewett, M. C. (2020). Mutational characterization and mapping of the 70S ribosome active site. *Nucleic Acids Research*, *48*(5), 2777–2789.
Davis, J. H., Tan, Y. Z., Carragher, B., Potter, C. S., Lyumkis, D., & Williamson, J. R. (2016). Modular Assembly of the Bacterial Large Ribosomal Subunit. *Cell*, *167*(6), 1610–1622.e15.
Dedkova, L. M., Fahmi, N. E., Golovine, S. Y., & Hecht, S. M. (2003). Enhanced D-amino acid incorporation into protein by modified ribosomes. *Journal of the American Chemical Society*, *125*(22), 6616–6617.
Dedkova, L. M., Fahmi, N. E., Paul, R., del Rosario, M., Zhang, L., Chen, S., Feder, G., & Hecht, S. M. (2012). β-Puromycin Selection of Modified Ribosomes for in Vitro Incorporation of β-Amino Acids. *Biochemistry*, *51*(1), 401–415.
Dedkova, L. M., & Hecht, S. M. (2019). Expanding the Scope of Protein Synthesis Using

Modified Ribosomes. *Journal of the American Chemical Society*, *141*(16), 6430–6447.
Dippel, A. B., Anderson, W. A., Park, J. H., Yildiz, F. H., & Hammond, M. C. (2020). Development of Ratiometric Bioluminescent Sensors for in Vivo Detection of Bacterial Signaling. *ACS Chemical Biology*, *15*(4), 904–914.
Dixon, A. S., Schwinn, M. K., Hall, M. P., Zimmerman, K., Otto, P., Lubben, T. H., Butler, B. L., Binkowski, B. F., Machleidt, T., Kirkland, T. A., Wood, M. G., Eggers, C. T., Encell, L. P., & Wood, K. V. (2016). NanoLuc Complementation Reporter Optimized for Accurate Measurement of Protein Interactions in Cells. *ACS Chemical Biology*, *11*(2), 400–408.
Doi, Y., Ohtsuki, T., Shimizu, Y., Ueda, T., & Sisido, M. (2007). Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system. *Journal of the American Chemical Society*, *129*(46), 14458–14462.
Douthwaite, S., Powers, T., Lee, J. Y., & Noller, H. F. (1989). Defining the Structural Requirements for a Helix in 23 S Ribosomal RNA that Confers Erythromycin Resistance. *Journal of Molecular Biology*, *209*(4), 655–665.
Du, L., Sánchez, C., & Shen, B. (2001). Hybrid peptide-polyketide natural products: biosynthesis and prospects toward engineering novel molecules. *Metabolic Engineering*, *3*(1), 78–95.
Dumas, A., Lercher, L., Spicer, C. D., & Davis, B. G. (2015). Designing logical codon reassignment - Expanding the chemistry in biology. *Chemical Science* , *6*(1), 50–69.
Dunkle, J. A., & Cate, J. H. D. (2010). Ribosome structure and dynamics during translocation and termination. *Annual Review of Biophysics*, *39*, 227–244.
Dunkle, J. A., Wang, L., Feldman, M. B., Pulk, A., Chen, V. B., Kapral, G. J., Noeske, J., Richardson, J. S., Blanchard, S. C., & Cate, J. H. D. (2011). Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. *Science*, *332*(6032), 981–984.
Duval, M., Korepanov, A., Fuchsbauer, O., Fechter, P., Haller, A., Fabbretti, A., Choulier, L., Micura, R., Klaholz, B. P., Romby, P., Springer, M., & Marzi, S. (2013). Escherichia coli ribosomal protein S1 unfolds structured mRNAs onto the ribosome for active translation initiation. *PLoS Biology*, *11*(12), e1001731.
Edsall, J. T., & Blanchard, M. H. (1933). The Activity Ratio of Zwitterions and Uncharged Molecules in Ampholyte Solutions. The Dissociation Constants of Amino Acid Esters. *Journal of the American Chemical Society*, *55*(6), 2337–2353.
Ekroos, M., & Sjogren, T. (2006). Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proceedings of the National Academy of Sciences*, *103*(37), 13682–13687.
Englander, M. T., Avins, J. L., Fleisher, R. C., Liu, B., Effraim, P. R., Wang, J., Schulten, K., Leyh, T. S., Gonzalez, R. L., Jr, & Cornish, V. W. (2015). The ribosome can discriminate the chirality of amino acids within its peptidyl-transferase center. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(19), 6038–6043.
Erlacher, M. D., & Polacek, N. (2012). Probing functions of the ribosomal peptidyl transferase center by nucleotide analog interference. *Methods in Molecular Biology* , *848*, 215–226.
Fahnestock, S., Neumann, H., Shashoua, V., & Rich, A. (1970). Ribosome-catalyzed ester formation. *Biochemistry*, *9*(12), 2477–2483.
Fahnestock, S., & Rich, A. (1971). Ribosome-catalyzed polyester formation. *Science*, *173*(3994), 340–343.
Farwell, M. A., Roberts, M. W., & Rabinowitz, J. C. (1992). The effect of ribosomal protein S1 from Escherichia coli and Micrococcus luteus on protein synthesis in vitro by E. coli and Bacillus subtilis. *Molecular Microbiology*, *6*(22), 3375–3383.
Feng, B., Mandava, C. S., Guo, Q., Wang, J., Cao, W., Li, N., Zhang, Y., Zhang, Y., Wang, Z.,

Wu, J., Sanyal, S., Lei, J., & Gao, N. (2014). Structural and Functional Insights into the Mode of Action of a Universally Conserved Obg GTPase. *PLoS Biology*, *12*(5), e1001866 1–14.
Fredens, J., Wang, K., de la Torre, D., Funke, L. F. H., Robertson, W. E., Christova, Y., Chia, T., Schmied, W. H., Dunkelmann, D. L., Beránek, V., Uttamapinant, C., Llamazares, A. G., Elliott, T. S., & Chin, J. W. (2019). Total synthesis of Escherichia coli with a recoded genome. *Nature*, *569*(7757), 514–518.
Fried, S. D., Schmied, W. H., Uttamapinant, C., & Chin, J. W. (2015). Ribosome Subunit Stapling for Orthogonal Translation in E. coli. *Angewandte Chemie* , *54*(43), 12791–12794.
Fritz, B. R., Jamil, O. K., & Jewett, M. C. (2015). Implications of macromolecular crowding and reducing conditions for in vitro ribosome construction. *Nucleic Acids Research*, *43*(9), 4774–4784.
Fujino, T., Goto, Y., Suga, H., & Murakami, H. (2013). Reevaluation of the D-amino acid compatibility with the elongation event in translation. *Journal of the American Chemical Society*, *135*(5), 1830–1837.
Fujino, T., Goto, Y., Suga, H., & Murakami, H. (2016). Ribosomal Synthesis of Peptides with Multiple β-Amino Acids. *Journal of the American Chemical Society*, *138*(6), 1962–1969.
Fu, Z., Indrisiunaite, G., Kaledhonkar, S., Shah, B., Sun, M., Chen, B., Grassucci, R. A., Ehrenberg, M., & Frank, J. (2019). The structural basis for release-factor activation during translation termination revealed by time-resolved cryogenic electron microscopy. *Nature Communications*, *10*(1), 2579.
Gatti‐Lafranconi, P., & Hollfelder, F. (2013). Flexibility and Reactivity in Promiscuous Enzymes. *Chembiochem: A European Journal of Chemical Biology*, *14*(3), 285–292.
Geueke, B., Heck, T., Limbach, M., Nesatyy, V., Seebach, D., & Kohler, H.-P. E. (2006). Bacterial β-peptidyl aminopeptidases with unique substrate specificities for β-oligopeptides and mixed β,α-oligopeptides. *The FEBS Journal*, *273*(23), 5261–5272.
Goddard, T. D., Huang, C. C., Meng, E. C., Pettersen, E. F., Couch, G. S., Morris, J. H., & Ferrin, T. E. (2018). UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Science: A Publication of the Protein Society*, *27*(1), 14–25.
Goto, Y., Katoh, T., & Suga, H. (2011). Flexizymes for genetic code reprogramming. *Nature Protocols*, *6*(6), 779–790.
Goto, Y., & Suga, H. (2009). Translation initiation with initiator tRNA charged with exotic peptides. *Journal of the American Chemical Society*, *131*(14), 5040–5041.
Green, R., & Noller, H. F. (1996). In vitro complementation analysis localizes 23S rRNA posttranscriptional modifications that are required for Escherichia coli 50S ribosomal subunit assembly and function. *RNA* , *2*(10), 1011–1021.
Guichard, G., Zerbib, A., Le Gal, F. A., Hoebeke, J., Connan, F., Choppin, J., Briand, J. P., & Guillet, J. G. (2000). Melanoma peptide MART-1(27-35) analogues with enhanced binding capacity to the human class I histocompatibility molecule HLA-A2 by introduction of a beta-amino acid residue: implications for recognition by tumor-infiltrating lymphocytes. *Journal of Medicinal Chemistry*, *43*(20), 3803–3808.
Haruna, K.-I., Alkazemi, M. H., Liu, Y., Söll, D., & Englert, M. (2014). Engineering the elongation factor Tu for efficient selenoprotein synthesis. *Nucleic Acids Research*, *42*(15), 9976–9983.
Hentzen, D., Mandel, P., & Garel, J. P. (1972). Relation between aminoacyl-tRNA stability and the fixed amino acid. *Biochimica et Biophysica Acta*, *281*(2), 228–232.

Herold, M., & Nierhaus, K. H. (1987). Incorporation of six additional proteins to complete the assembly map of the 50 S subunit from Escherichia coli ribosomes. *The Journal of Biological Chemistry*, *262*(18), 8826–8833.

Hill, D. J., Mio, M. J., Prince, R. B., Hughes, T. S., & Moore, J. S. (2001). A field guide to foldamers. *Chemical Reviews*, *101*(12), 3893–4012.

Hohsaka, T., Ashizuka, Y., Murakami, H., & Sisido, M. (1996). Incorporation of Nonnatural Amino Acids into Streptavidin through In Vitro Frame-Shift Suppression. *Journal of the American Chemical Society*, *118*(40), 9778–9779.

Hui, A., & de Boer, H. A. (1987). Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, *84*(14), 4762–4766.

Hui, A., Jhurani, P., & de Boer, H. A. (1987). Directing ribosomes to a single mRNA species: a method to study ribosomal RNA mutations and their effects on translation of a single messenger in Escherichia coli. *Methods in Enzymology*, *153*, 432–452.

Iqbal, E. S., Dods, K. K., & Hartman, M. C. T. (2018). Ribosomal incorporation of backbone modified amino acids via an editing-deficient aminoacyl-tRNA synthetase. *Organic & Biomolecular Chemistry*, *16*(7), 1073–1078.

Jelenc, P. C., & Kurland, C. G. (1979-7). Nucleoside triphosphate regeneration decreases the frequency of translation errors. *Proceedings of the National Academy of Sciences of the United States of America*, *76*(7), 3174–3178.

Jewett, M. C., & Swartz, J. R. (2004). Mimicking the Escherichia coli cytoplasmic environment activates long-lived and efficient cell-free protein synthesis. *Biotechnology and Bioengineering*, *86*(1), 19–26.

Jomaa, A., Jain, N., Davis, J. H., Williamson, J. R., Britton, R. A., & Ortega, J. (2014). Functional domains of the 50S subunit mature late in the assembly process. *Nucleic Acids Research*, *42*(5), 3419–3435.

Kaledhonkar, S., Fu, Z., Caban, K., Li, W., Chen, B., Sun, M., Gonzalez, R. L., Jr, & Frank, J. (2019). Late steps in bacterial translation initiation visualized using time-resolved cryo-EM. *Nature*, *570*(7761), 400–404.

Kao, C., Zheng, M., & Rüdisser, S. (1999). A simple and efficient method to reduce nontemplated nucleotide addition at the 3 terminus of RNAs transcribed by T7 RNA polymerase. *RNA* , *5*(9), 1268–1272.

Katoh, T., Iwane, Y., & Suga, H. (2017). Logical engineering of D-arm and T-stem of tRNA that enhances d-amino acid incorporation. *Nucleic Acids Research*, *45*(22), 12601–12610.

Katoh, T., Sengoku, T., Hirata, K., Ogata, K., & Suga, H. (2020). Ribosomal synthesis and de novo discovery of bioactive foldamer peptides containing cyclic β-amino acids. *Nature Chemistry*, *12*(11), 1081–1088.

Katoh, T., & Suga, H. (2018). Ribosomal Incorporation of Consecutive β-Amino Acids. *Journal of the American Chemical Society*, *140*(38), 12159–12167.

Katoh, T., & Suga, H. (2019). Flexizyme-catalyzed synthesis of 3'-aminoacyl-NH-tRNAs. *Nucleic Acids Research*, *47*(9), e54.

Katoh, T., & Suga, H. (2020a). Ribosomal Elongation of Cyclic γ-Amino Acids using a Reprogrammed Genetic Code. *Journal of the American Chemical Society*, *142*(11), 4965–4969.

Katoh, T., & Suga, H. (2020b). Ribosomal Elongation of Aminobenzoic Acid Derivatives.

Journal of the American Chemical Society, *142*(39), 16518–16522.
Katoh, T., Tajima, K., & Suga, H. (2017). Consecutive Elongation of D-Amino Acids in Translation. *Cell Chemical Biology*, *24*(1), 46–54.
Kawakami, T., Murakami, H., & Suga, H. (2008). Ribosomal Synthesis of Polypeptoids and Peptoid−Peptide Hybrids. *Journal of the American Chemical Society*, *130*(50), 16861–16863.
Kilburn, D., Roh, J. H., Guo, L., Briber, R. M., & Woodson, S. A. (2010). Molecular Crowding Stabilizes Folded RNA Structure by the Excluded Volume Effect. *Journal of the American Chemical Society*, *132*(25), 8690–8696.
Knight, A. S., Zhou, E. Y., Francis, M. B., & Zuckermann, R. N. (2015). Sequence Programmable Peptoid Polymers for Diverse Materials Applications. *Advanced Materials* , *27*(38), 5665–5691.
Komarova, A. V., Tchufistova, L. S., Dreyfus, M., & Boni, I. V. (2005). AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in Escherichia coli. *Journal of Bacteriology*, *187*(4), 1344–1349.
Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H.-R., Haimovich, A. D., Kuznetsov, G., Mercer, J. A., Wang, H. H., Carr, P. A., Mosberg, J. A., Rohland, N., Schultz, P. G., Jacobson, J. M., Rinehart, J., Church, G. M., & Isaacs, F. J. (2013). Genomically recoded organisms expand biological functions. *Science*, *342*(6156), 357–360.
Lake, J. A. (1976). Ribosome structure determined by electron microscopy of Escherichia coli small subunits, large subunits and monomeric ribosomes. *Journal of Molecular Biology*, *105*(1), 131–139.
LaRiviere, F. J., Wolfson, A. D., & Uhlenbeck, O. C. (2001). Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science*, *294*(5540), 165–168.
Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S., & Noller, H. F. (2008). Structural basis for translation termination on the 70S ribosome. *Nature*, *454*(7206), 852–857.
Laursen, B. S., Sørensen, H. P., Mortensen, K. K., & Sperling-Petersen, H. U. (2005). Initiation of protein synthesis in bacteria. *Microbiology and Molecular Biology Reviews: MMBR*, *69*(1), 101–123.
Leamy, K. A., Assmann, S. M., Mathews, D. H., & Bevilacqua, P. C. (2016/ed). Bridging the gap between in vitro and in vivo RNA folding. *Quarterly Reviews of Biophysics*, *49*, 1–26.
Lee, J., Schwarz, K. J., Kim, D. S., Moore, J. S., & Jewett, M. C. (2020). Ribosome-mediated polymerization of long chain carbon and cyclic amino acids into peptides in vitro. *Nature Communications*, *11*(1), 4304.
Lee, J., Schwieter, K. E., Watkins, A. M., Kim, D. S., Yu, H., Schwarz, K. J., Lim, J., Coronado, J., Byrom, M., Anslyn, E. V., Ellington, A. D., Moore, J. S., & Jewett, M. C. (2019). Expanding the limits of the second genetic code with ribozymes. *Nature Communications*, *10*(1), 5097.
Lee, J., Torres, R., Kim, D. S., Byrom, M., Ellington, A. D., & Jewett, M. C. (2020). Ribosomal incorporation of cyclic β-amino acids into peptides using in vitro translation. *Chemical Communications* , *56*(42), 5597–5600.
Lehmann, J. (2017). Induced fit of the peptidyl-transferase center of the ribosome and conformational freedom of the esterified amino acids. *RNA* , *23*(2), 229–239.
Leung, E. K. Y., Suslov, N., Tuttle, N., Sengupta, R., & Piccirilli, J. A. (2011). The mechanism of

peptidyl transfer catalysis by the ribosome. *Annual Review of Biochemistry*, *80*, 527–555.
Liebschner, D., Afonine, P. V., Baker, M. L., Bunkóczi, G., Chen, V. B., Croll, T. I., Hintze, B., Hung, L. W., Jain, S., McCoy, A. J., Moriarty, N. W., Oeffner, R. D., Poon, B. K., Prisant, M. G., Read, R. J., Richardson, J. S., Richardson, D. C., Sammito, M. D., Sobolev, O. V., … Adams, P. D. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Crystallographica. Section D, Structural Biology*, *75*(Pt 10), 861–877.
Liljeruhm, J., Wang, J., Kwiatkowski, M., Sabari, S., & Forster, A. C. (2019). Kinetics of d-Amino Acid Incorporation in Translation. *ACS Chemical Biology*, *14*(2), 204–213.
Li, N., Chen, Y., Guo, Q., Zhang, Y., Yuan, Y., Ma, C., Deng, H., Lei, J., & Gao, N. (2013). Cryo-EM structures of the late-stage assembly intermediates of the bacterial 50S ribosomal subunit. *Nucleic Acids Research*, *41*(14), 7073–7083.
Liu, C. C., & Schultz, P. G. (2010). Adding new chemistries to the genetic code. *Annual Review of Biochemistry*, *79*, 413–444.
Loveland, A. B., Demo, G., Grigorieff, N., & Korostelev, A. A. (2017). Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature*, *546*(7656), 113–117.
Loveland, A. B., Demo, G., & Korostelev, A. A. (2020). Cryo-EM of elongating ribosome with EF-Tu•GTP elucidates tRNA proofreading. *Nature*, *584*(7822), 640–645.
Maini, R., Chowdhury, S. R., Dedkova, L. M., Roy, B., Daskalova, S. M., Paul, R., Chen, S., & Hecht, S. M. (2015). Protein Synthesis with Ribosomes Selected for the Incorporation of β-Amino Acids. *Biochemistry*, *54*(23), 3694–3706.
Maini, R., Dedkova, L. M., Paul, R., Madathil, M. M., Chowdhury, S. R., Chen, S., & Hecht, S. M. (2015). Ribosome-Mediated Incorporation of Dipeptides and Dipeptide Analogues into Proteins in Vitro. *Journal of the American Chemical Society*, *137*(35), 11206–11209.
Maini, R., Nguyen, D. T., Chen, S., Dedkova, L. M., Chowdhury, S. R., Alcala-Torano, R., & Hecht, S. M. (2013). Incorporation of β-amino acids into dihydrofolate reductase by ribosomes having modifications in the peptidyltransferase center. *Bioorganic & Medicinal Chemistry*, *21*(5), 1088–1096.
Martin Schmeing, T., Huang, K. S., Strobel, S. A., & Steitz, T. A. (2005). An induced-fit mechanism to promote peptide bond formation and exclude hydrolysis of peptidyl-tRNA. *Nature*, *438*(7067), 520–524.
Mastronarde, D. N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. *Journal of Structural Biology*, *152*(1), 36–51.
McDonald, W. H., Tabb, D. L., Sadygov, R. G., MacCoss, M. J., Venable, J., Graumann, J., Johnson, J. R., Cociorva, D., & Yates, J. R., 3rd. (2004). MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. *Rapid Communications in Mass Spectrometry: RCM*, *18*(18), 2162–2168.
Melnikov, S., Mailliot, J., Rigger, L., Neuner, S., Shin, B.-S., Yusupova, G., Dever, T. E., Micura, R., & Yusupov, M. (2016). Molecular insights into protein synthesis with proline residues. *EMBO Reports*, *17*(12), 1776–1784.
Melnikov, S. V., Khabibullina, N. F., Mairhofer, E., Vargas-Rodriguez, O., Reynolds, N. M., Micura, R., Söll, D., & Polikanov, Y. S. (2019). Mechanistic insights into the slow peptide bond formation with D-amino acids in the ribosomal active site. *Nucleic Acids Research*, *47*(4), 2089–2100.
Melo Czekster, C., Robertson, W. E., Walker, A. S., Söll, D., & Schepartz, A. (2016). In Vivo Biosynthesis of a β-Amino Acid-Containing Protein. *Journal of the American Chemical*

Society, *138*(16), 5194–5197.
Mesters, J. R., Vorstenbosch, E. L. H., de Boer, A. J., & Kraal, B. (1994). Complete purification of tRNA, charged or modified with hydrophobic groups, by reversed-phase high-performance liquid chromatography on a C4/C18 column system. *Journal of Chromatography. A*, *679*(1), 93–98.
Moll, I., Huber, M., Grill, S., Sairafi, P., Mueller, F., Brimacombe, R., Londei, P., & Bläsi, U. (2001). Evidence against an Interaction between the mRNA downstream box and 16S rRNA in translation initiation. *Journal of Bacteriology*, *183*(11), 3499–3505.
Monro, R. E., & Marcker, K. A. (1967). Ribosome-catalysed reaction of puromycin with a formylmethionine-containing oligonucleotide. *Journal of Molecular Biology*, *25*(2), 347–350.
Mukai, T., Hayashi, A., Iraha, F., Sato, A., Ohtake, K., Yokoyama, S., & Sakamoto, K. (2010). Codon reassignment in the Escherichia coli genetic code. *Nucleic Acids Research*, *38*(22), 8188–8195.
Nakagawa, S., Niimura, Y., Miura, K.-I., & Gojobori, T. (2010). Dynamic evolution of translation initiation mechanisms in prokaryotes. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(14), 6382–6387.
Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., & Chin, J. W. (2010). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature*, *464*(7287), 441–444.
Nikolay, R., Hilal, T., Qin, B., Mielke, T., Bürger, J., Loerke, J., Textoris-Taube, K., Nierhaus, K. H., & Spahn, C. M. T. (2018). Structural Visualization of the Formation and Activation of the 50S Ribosomal Subunit during In Vitro Reconstitution. *Molecular Cell*, *70*(5), 881–893.e3.
Nikolay, R., Hilal, T., Schmidt, S., Qin, B., Schwefel, D., Vieira-Vieira, C. H., Mielke, T., Bürger, J., Loerke, J., Amikura, K., Flügel, T., Ueda, T., Selbach, M., Deuerling, E., & Spahn, C. M. T. (2021). Snapshots of native pre-50S ribosomes reveal a biogenesis factor network and evolutionary specialization. In *Molecular Cell*. https://doi.org/10.1016/j.molcel.2021.02.006
Nishiyama, K., Ichihashi, N., Kazuta, Y., & Yomo, T. (2015). Development of a reporter peptide that catalytically produces a fluorescent signal through α-complementation. *Protein Science: A Publication of the Protein Society*, *24*(5), 599–603.
Noeske, J., Wasserman, M. R., Terry, D. S., Altman, R. B., Blanchard, S. C., & Cate, J. H. D. (2015). High-resolution structure of the Escherichia coli ribosome. *Nature Structural & Molecular Biology*, *22*(4), 336–341.
Nogales, E., & Scheres, S. H. W. (2015). Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. *Molecular Cell*, *58*(4), 677–689.
O'Connor, M., Lee, W. M., Mankad, A., Squires, C. L., & Dahlberg, A. E. (2001). Mutagenesis of the peptidyltransferase center of 23S rRNA: the invariant U2449 is dispensable. *Nucleic Acids Research*, *29*(3), 710–715.
Ogle, J. M., Murphy, F. V., Tarry, M. J., & Ramakrishnan, V. (2002). Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell*, *111*(5), 721–732.
Oh-Hashi, K., Furuta, E., Fujimura, K., & Hirata, Y. (2017). Application of a novel HiBiT peptide tag for monitoring ATF4 protein expression in Neuro2a cells. *Biochemistry and Biophysics Reports*, *12*, 40–45.
Ohta, A., Murakami, H., & Suga, H. (2008). Polymerization of alpha-hydroxy acids by

ribosomes. *Chembiochem: A European Journal of Chemical Biology*, *9*(17), 2773–2778.
Öjemalm, K., Higuchi, T., Jiang, Y., Langel, Ü., Nilsson, I., White, S. H., Suga, H., & von Heijne, G. (2011). Apolar surface area determines the efficiency of translocon-mediated membrane-protein integration into the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(31), E359–E364.
Orelle, C., Carlson, E. D., Szal, T., Florin, T., Jewett, M. C., & Mankin, A. S. (2015). Protein synthesis by ribosomes with tethered subunits. *Nature*, *524*(7563), 119–124.
Osterman, I. A., Khabibullina, N. F., Komarova, E. S., Kasatsky, P., Kartsev, V. G., Bogdanov, A. A., Dontsova, O. A., Konevega, A. L., Sergiev, P. V., & Polikanov, Y. S. (2017). Madumycin II inhibits peptide bond formation by forcing the peptidyl transferase center into an inactive state. *Nucleic Acids Research*, *45*(12), 7507–7514.
Park, S. K. R., Aslanian, A., McClatchy, D. B., Han, X., Shah, H., Singh, M., Rauniyar, N., Moresco, J. J., Pinto, A. F. M., Diedrich, J. K., Delahunty, C., & Yates, J. R., 3rd. (2014). Census 2: isobaric labeling data analysis. *Bioinformatics* , *30*(15), 2208–2209.
Park, S. K., Venable, J. D., Xu, T., & Yates, J. R., 3rd. (2008). A quantitative analysis software tool for mass spectrometry-based proteomics. *Nature Methods*, *5*(4), 319–322.
Peacock, J. R., Walvoord, R. R., Chang, A. Y., Kozlowski, M. C., Gamper, H., & Hou, Y.-M. (2014). Amino acid–dependent stability of the acyl linkage in aminoacyl-tRNA. *RNA* , *20*(6), 758–764.
Peng, J., Elias, J. E., Thoreen, C. C., Licklider, L. J., & Gygi, S. P. (2003). Evaluation of Multidimensional Chromatography Coupled with Tandem Mass Spectrometry (LC/LC−MS/MS) for Large-Scale Protein Analysis: The Yeast Proteome. *Journal of Proteome Research*, *2*(1), 43–50.
Pereira, G. C., Allen, W. J., Watkins, D. W., Buddrus, L., Noone, D., Liu, X., Richardson, A. P., Chacinska, A., & Collinson, I. (2019). A High-Resolution Luminescent Assay for Rapid and Continuous Monitoring of Protein Translocation across Biological Membranes. *Journal of Molecular Biology*, *431*(8), 1689–1699.
Petrov, A. S., Bernier, C. R., Hsiao, C., Norris, A. M., Kovacs, N. A., Waterbury, C. C., Stepanov, V. G., Harvey, S. C., Fox, G. E., Wartell, R. M., Hud, N. V., & Williams, L. D. (2014). Evolution of the ribosome at atomic resolution. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(28), 10251–10256.
Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, *25*(13), 1605–1612.
Polacek, N., & Mankin, A. S. (2005). The ribosomal peptidyl transferase center: structure, function, evolution, inhibition. *Critical Reviews in Biochemistry and Molecular Biology*, *40*(5), 285–311.
Polacek, N., Swaney, S., Shinabarger, D., & Mankin, A. S. (2002). SPARK--a novel method to monitor ribosomal peptidyl transferase activity. *Biochemistry*, *41*(39), 11602–11610.
Polikanov, Y. S., Steitz, T. A., & Innis, C. A. (2014). A proton wire to couple aminoacyl-tRNA accommodation and peptide-bond formation on the ribosome. *Nature Structural & Molecular Biology*, *21*(9), 787–793.
Porel, M., & Alabi, C. A. (2014). Sequence-defined polymers via orthogonal allyl acrylamide building blocks. *Journal of the American Chemical Society*, *136*(38), 13162–13165.
Porel, M., Thornlow, D. N., Artim, C. M., & Alabi, C. A. (2017). Sequence-Defined Backbone Modifications Regulate Antibacterial Activity of OligoTEAs. *ACS Chemical Biology*, *12*(3),

715–723.
Pulk, A., & Cate, J. H. D. (2013). Control of ribosomal subunit rotation by elongation factor G. *Science*, *340*(6140), 1235970.
Punjani, A., & Fleet, D. J. (2020). 3D Variability Analysis: Directly resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM images. In *bioRxiv* (p. 2020.04.08.032466). https://doi.org/10.1101/2020.04.08.032466
Punjani, A., Rubinstein, J. L., Fleet, D. J., & Brubaker, M. A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nature Methods*, *14*(3), 290–296.
Qin, D., & Fredrick, K. (2013). Chapter Eight - Analysis of Polysomes from Bacteria. In J. Lorsch (Ed.), *Methods in Enzymology* (Vol. 530, pp. 159–172). Academic Press.
Rackham, O., & Chin, J. W. (2005). A network of orthogonal ribosome·mRNA pairs. *Nature Chemical Biology*, *1*(3), 159–166.
Rakauskaitė, R., & Dinman, J. D. (2011-5). Mutations of highly conserved bases in the peptidyltransferase center induce compensatory rearrangements in yeast ribosomes. *RNA* , *17*(5), 855–864.
Reynolds, N. M., Vargas-Rodriguez, O., Söll, D., & Crnković, A. (2017). The central role of tRNA in genetic code expansion. *Biochimica et Biophysica Acta, General Subjects*, *1861*(11 Pt B), 3001–3008.
Rinaldi, S. (2020). The Diverse World of Foldamers: Endless Possibilities of Self-Assembly. *Molecules* , *25*(14). https://doi.org/10.3390/molecules25143276
Rogers, J. M., Kwon, S., Dawson, S. J., Mandal, P. K., Suga, H., & Huc, I. (2018). Ribosomal synthesis and folding of peptide-helical aromatic foldamer hybrids. *Nature Chemistry*, *10*(4), 405–412.
Rohou, A., & Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *Journal of Structural Biology*, *192*(2), 216–221.
Russell, J. B., & Cook, G. M. (1995). Energetics of bacterial growth: balance of anabolic and catabolic reactions. *Microbiological Reviews*, *59*(1), 48–62.
Saito, K., Green, R., & Buskirk, A. R. (2020). Translational initiation in E. coli occurs at the correct sites genome-wide in the absence of mRNA-rRNA base-pairing. *eLife*, *9*. https://doi.org/10.7554/eLife.55002
Salah, P., Bisaglia, M., Aliprandi, P., Uzan, M., Sizun, C., & Bontems, F. (2009). Probing the relationship between Gram-negative and Gram-positive S1 proteins by sequence analysis. *Nucleic Acids Research*, *37*(16), 5578–5588.
Santoro, S. W., Wang, L., Herberich, B., King, D. S., & Schultz, P. G. (2002). An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nature Biotechnology*, *20*(10), 1044–1048.
Sato, N. S., Hirabayashi, N., Agmon, I., Yonath, A., & Suzuki, T. (2006). Comprehensive genetic selection revealed essential bases in the peptidyl-transferase center. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(42), 15386–15391.
Schaub, F. X., Reza, M. S., Flaveny, C. A., Li, W., Musicant, A. M., Hoxha, S., Guo, M., Cleveland, J. L., & Amelio, A. L. (2015). Fluorophore-NanoLuc BRET Reporters Enable Sensitive In Vivo Optical Imaging and Flow Cytometry for Monitoring Tumorigenesis. *Cancer Research*, *75*(23), 5023–5033.
Scheres, S. H. W. (2012). RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *Journal of Structural Biology*, *180*(3), 519–530.
Scheres, S. H. W. (2016). Processing of Structurally Heterogeneous Cryo-EM Data in RELION.

Methods in Enzymology, *579*, 125–157.
Schluenzen, F., Tocilj, A., Zarivach, R., Harms, J., Gluehmann, M., Janell, D., Bashan, A., Bartels, H., Agmon, I., Franceschi, F., & Yonath, A. (2000). Structure of functionally activated small ribosomal subunit at 3.3 angstroms resolution. *Cell*, *102*(5), 615–623.
Schmeing, T. M., Voorhees, R. M., Kelley, A. C., Gao, Y.-G., Murphy, F. V., 4th, Weir, J. R., & Ramakrishnan, V. (2009). The crystal structure of the ribosome bound to EF-Tu and aminoacyl-tRNA. *Science*, *326*(5953), 688–694.
Schmied, W. H., Tnimov, Z., Uttamapinant, C., Rae, C. D., Fried, S. D., & Chin, J. W. (2018). Controlling orthogonal ribosome subunit interactions enables evolution of new function. *Nature*, *564*(7736), 444–448.
Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675.
Schwinn, M. K., Machleidt, T., Zimmerman, K., Eggers, C. T., Dixon, A. S., Hurst, R., Hall, M. P., Encell, L. P., Binkowski, B. F., & Wood, K. V. (2018). CRISPR-Mediated Tagging of Endogenous Proteins with a Luminescent Peptide. *ACS Chemical Biology*, *13*(2), 467–474.
Shekhawat, S. S., & Ghosh, I. (2011). Split-protein systems: beyond binary protein-protein interactions. *Current Opinion in Chemical Biology*, *15*(6), 789–797.
Shepherd, J., & Ibba, M. (2015). Bacterial transfer RNAs. *FEMS Microbiology Reviews*, *39*(3), 280–300.
Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., & Ueda, T. (2001). Cell-free translation reconstituted with purified components. *Nature Biotechnology*, *19*(8), 751–755.
Shine, J., & Dalgarno, L. (1975). Determinant of cistron specificity in bacterial ribosomes. *Nature*, *254*(5495), 34–38.
Sievers, A., Beringer, M., Rodnina, M. V., & Wolfenden, R. (2004). The ribosome as an entropy trap. *Proceedings of the National Academy of Sciences of the United States of America*, *101*(21), 7897–7901.
Skopalík, J., Anzenbacher, P., & Otyepka, M. (2008). Flexibility of Human Cytochromes P450: Molecular Dynamics Reveals Differences between CYPs 3A4, 2C9, and 2A6, which Correlate with Their Substrate Preferences. *The Journal of Physical Chemistry. B*, *112*(27), 8165–8173.
Strulson, C. A., Yennawar, N. H., Rambo, R. P., & Bevilacqua, P. C. (2013). Molecular Crowding Favors Reactivity of a Human Ribozyme Under Physiological Ionic Conditions. *Biochemistry*, *52*(46), 8187–8197.
Studier, F. W. (2014). Stable expression clones and auto-induction for protein production in E. coli. *Methods in Molecular Biology* , *1091*, 17–32.
Subtelny, A. O., Hartman, M. C. T., & Szostak, J. W. (2008). Ribosomal Synthesis of N-Methyl Peptides. *Journal of the American Chemical Society*, *130*(19), 6131–6136.
Tanner, D., Fitzgerald, J. A., & Phillips, B. R. (1989). The Kevlar story—An advanced materials case study. *Angewandte Chemie International Edition in English*, *28*(5), 649–654.
Thompson, A., Schäfer, J., Kuhn, K., Kienle, S., Schwarz, J., Schmidt, G., Neumann, T., Johnstone, R., Mohammed, A. K. A., & Hamon, C. (2003). Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical Chemistry*, *75*(8), 1895–1904.
Thompson, J., Kim, D. F., O'Connor, M., Lieberman, K. R., Bayfield, M. A., Gregory, S. T., Green, R., Noller, H. F., & Dahlberg, A. E. (2001). Analysis of mutations at residues A2451

and G2447 of 23S rRNA in the peptidyltransferase active site of the 50S ribosomal subunit. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(16), 9002–9007.

Ting, L., Rad, R., Gygi, S. P., & Haas, W. (2011). MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nature Methods*, *8*(11), 937–940.

Travin, D. Y., Watson, Z. L., Metelev, M., Ward, F. R., Osterman, I. A., Khven, I. M., Khabibullina, N. F., Serebryakova, M., Mergaert, P., Polikanov, Y. S., Cate, J. H. D., & Severinov, K. (2019). Phazolicin – a Novel Thiazole/Oxazole-Modified Peptide Inhibiting the Bacterial Ribosome in a Species-Specific Way. In *bioRxiv* (p. 705970). https://doi.org/10.1101/705970

Tsiamantas, C., Kwon, S., Douat, C., Huc, I., & Suga, H. (2019). Optimizing aromatic oligoamide foldamer side-chains for ribosomal translation initiation. *Chemical Communications* , *55*(51), 7366–7369.

Tuckey, C., Asahara, H., Zhou, Y., & Chong, S. (2014). Protein synthesis using a reconstituted cell-free system. *Current Protocols in Molecular Biology / Edited by Frederick M. Ausubel ... [et Al.]*, *108*, 16.31.1–22.

Tyrrell, J., McGinnis, J. L., Weeks, K. M., & Pielak, G. J. (2013). The Cellular Environment Stabilizes Adenine Riboswitch RNA Structure. *Biochemistry*, *52*(48), 8777–8785.

Udagawa, T., Shimizu, Y., & Ueda, T. (2004). Evidence for the Translation Initiation of Leaderless mRNAs by the Intact 70 S Ribosome without Its Dissociation into Subunits in Eubacteria*. *The Journal of Biological Chemistry*, *279*(10), 8539–8546.

Ude, S., Lassak, J., Starosta, A. L., Kraxenberger, T., Wilson, D. N., & Jung, K. (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science*, *339*(6115), 82–85.

Vargas-Rodriguez, O., Sevostyanova, A., Söll, D., & Crnković, A. (2018). Upgrading aminoacyl-tRNA synthetases for genetic code expansion. *Current Opinion in Chemical Biology*, *46*, 115–122.

Voorhees, R. M., Weixlbaumer, A., Loakes, D., Kelley, A. C., & Ramakrishnan, V. (2009). Insights into substrate stabilization from snapshots of the peptidyl transferase center of the intact 70S ribosome. *Nature Structural & Molecular Biology*, *16*(5), 528–533.

Walker, A. S., Russ, W. P., Ranganathan, R., & Schepartz, A. (2020). RNA sectors and allosteric function within the ribosome. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(33), 19879–19887.

Walsh, C. T. (2004). Polyketide and nonribosomal peptide antibiotics: modularity and versatility. *Science*, *303*(5665), 1805–1810.

Wang, J. (2017). Experimental charge density from electron microscopic maps. *Protein Science: A Publication of the Protein Society*, *26*(8), 1619–1626.

Wang, J., Kwiatkowski, M., Pavlov, M. Y., Ehrenberg, M., & Forster, A. C. (2014). Peptide Formation by N-Methyl Amino Acids in Translation Is Hastened by Higher pH and tRNAPro. *ACS Chemical Biology*, *9*(6), 1303–1311.

Ward, F. R., Watson, Z. L., Ad, O., Schepartz, A., & Cate, J. H. D. (2019). Defects in the Assembly of Ribosomes Selected for β-Amino Acid Incorporation. *Biochemistry*, *58*(45), 4494–4504.

Watson, Z. L., Ward, F. R., Méheust, R., Ad, O., Schepartz, A., Banfield, J. F., & Cate, J. H. (2020). Structure of the bacterial ribosome at 2 Å resolution. *eLife*, *9*. https://doi.org/10.7554/eLife.60482

Weinger, J. S., Parnell, K. M., Dorner, S., Green, R., & Strobel, S. A. (2004). Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nature Structural & Molecular Biology*, *11*(11), 1101–1106.

Xiao, H., Murakami, H., Suga, H., & Ferré-D'Amaré, A. R. (2008). Structural basis of specific tRNA aminoacylation by a small in vitro selected ribozyme. *Nature*, *454*(7202), 358–361.

Xu, D., & Wang, Y. (2021). Protein-free ribosomal RNA scaffolds can assemble poly-lysine oligos from charged tRNA fragments. *Biochemical and Biophysical Research Communications*, *544*, 81–85.

Xu, T., Venable, J. D., Park, S. K., Cociorva, D., & III. Yates, J. R. (2006). ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program. *Molecular & Cellular Proteomics: MCP*, *5*(10), S174–S174.

Youngman, E. M., Brunelle, J. L., Kochaniak, A. B., & Green, R. (2004). The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. *Cell*, *117*(5), 589–599.

Youngman, E. M., & Green, R. (2005). Affinity purification of in vivo-assembled ribosomes for in vitro biochemical analysis. *Methods* , *36*(3), 305–312.

Zarivach, R., Bashan, A., Berisio, R., Harms, J., Auerbach, T., Schluenzen, F., Bartels, H., Baram, D., Pyetan, E., Sittner, A., Amit, M., Hansen, H. A. S., Kessler, M., Liebe, C., Wolff, A., Agmon, I., & Yonath, A. (2004). Functional aspects of ribosomal architecture: symmetry, chirality and regulation. *Journal of Physical Organic Chemistry*, *17*(11), 901–912.

Zavialov, A. V., Mora, L., Buckingham, R. H., & Ehrenberg, M. (2002). Release of peptide promoted by the GGQ motif of class 1 release factors regulates the GTPase activity of RF3. *Molecular Cell*, *10*(4), 789–798.

Zhang, B., Tan, Z., Dickson, L. G., Nalam, M. N. L., Cornish, V. W., & Forster, A. C. (2007). Specificity of Translation for N-Alkyl Amino Acids. *Journal of the American Chemical Society*, *129*(37), 11316–11317.

Zhang, X., Yan, K., Zhang, Y., Li, N., Ma, C., Li, Z., Zhang, Y., Feng, B., Liu, J., Sun, Y., Xu, Y., Lei, J., & Gao, N. (2014). Structural insights into the function of a unique tandem GTPase EngA in bacterial ribosome assembly. *Nucleic Acids Research*, *42*(21), 13430–13439.

Zhang, Y., Ptacin, J. L., Fischer, E. C., Aerni, H. R., Caffaro, C. E., San Jose, K., Feldman, A. W., Turner, C. R., & Romesberg, F. E. (2017). A semi-synthetic organism that stores and retrieves increased genetic information. *Nature*, *551*(7682), 644–647.

Zheng, S. Q., Palovcak, E., Armache, J.-P., Verba, K. A., Cheng, Y., & Agard, D. A. (2017). MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nature Methods*, *14*(4), 331–332.

Zimmerman, S. B., & Trach, S. O. (1991). Estimation of macromolecule concentrations and excluded volume effects for the cytoplasm of Escherichia coli. *Journal of Molecular Biology*, *222*(3), 599–620.

Zivanov, J., Nakane, T., Forsberg, B. O., Kimanius, D., Hagen, W. J., Lindahl, E., & Scheres, S. H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife*, *7*. https://doi.org/10.7554/eLife.42166