In Vitro Kinetics of Ribosomal Incorporation of Unnatural Amino Acids

JINFAN WANG
Ribosomal incorporation of unnatural amino acids (AAs) into peptides or proteins has found broad applications in studying translation mechanism, discovering potential therapeutics, and probing protein structure and function. However, such applications are generally limited by the low incorporation efficiencies of the unnatural AAs.

With in vitro kinetics studies using a purified E. coli translation system, we found that the natural N-alkyl AA carrier, tRNA\(^{\text{Pro}}\), could hasten the incorporation of N-methyl AAs. Also, the incorporation rate increased remarkably with increasing pH in the range of 7 to 8.5, suggesting the rate was limited by peptidyl transfer, not accommodation. Competition experiments revealed that several futile cycles of delivery and rejection of the A site N-methyl AA-tRNA were required per peptide bond formation, and the incorporation yield could be increased by using a higher Mg\(^{2+}\) concentration.

Kinetics of ribosomal polymerization, using AA-tRNA substrates prepared from the standard N-NVOC-AA-pdCpA chemoenzymatic ligation method, clarified that the inefficiency of incorporation was due to the penultimate dC. This dC prompted faster peptidyl-tRNA drop-off, leading to loss of processivities along consecutive incorporations. Circumventing the penultimate dC by using our N-NVOC-AA-pCpA chemoenzymatic ligation or the flexizyme charging method to prepare the AA-tRNA substrates was able to improve the efficiencies of ribosomal consecutive incorporations of unnatural AAs.

By studying the translation steps after aminoacylation of tRNA\(^{\text{Pyl}}\), the favored carrier for unnatural AAs in vivo, we demonstrated surprisingly slow biphasic kinetics of tRNA\(^{\text{Pyl}}\)-mediated amber suppression in vitro. The fast phase amplitude increased with increasing EF-Tu concentration, allowing measurement of \(K_d\) of EF-Tu binding. Results revealed ~25-fold weaker EF-Tu binding affinity of the tRNA\(^{\text{Pyl}}\) body than that of E. coli tRNA\(^{\text{Phe}}\). The fast phase rate was ~30-fold slower than that of native substrates, and this rate was limited by the ~10-fold less efficient AA-tRNA\(^{\text{Pyl}}\):EF-Tu:GTP ternary complex binding to the ribosome. The incorporation was so slow that termination by RF2 mis-reading of the amber codon became a significant competing reaction. The processivity was unexpectedly impaired as ~40% of the dipeptidyl-tRNA\(^{\text{Pyl}}\) could not be elongated to tripeptide. This new overall understanding opens a window of improving unnatural AA incorporation both in vitro and in vivo.

Keywords: ribosome, protein synthesis, unnatural amino acids, kinetics
To Prof. Knud H. Nierhaus.

To Alf, to love, and to life.
List of Papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


Reprints were made with permission from the respective publishers.
Related Publications by the Author

(Not included in the thesis)


V Wang, J., Kwiatkowski, M., Forster, A.C. An unexpected rate-limiting step in ribosomal peptide synthesis. *Manuscript*
Contents

Introduction .................................................................................................................. 9
  The central dogma and the genetic code ................................................................. 9
Bacterial protein synthesis ....................................................................................... 11
  Initiation .................................................................................................................. 11
  Elongation ............................................................................................................... 11
  Termination ............................................................................................................ 13
  Recycling .................................................................................................................. 14
Unnatural amino acids .............................................................................................. 14
  Ribosomal incorporation of unnatural AAs in vivo .............................................. 16
  Ribosomal incorporation of unnatural AAs in vitro .............................................. 18
Dissertation overview ............................................................................................... 21

The present work ...................................................................................................... 22
  Experimental setup .................................................................................................. 22
    In vitro translation system ..................................................................................... 22
  Experimental procedures ....................................................................................... 23
Single incorporation of N-methyl AAs (Paper I) ...................................................... 25
  tRNA$^{ProB}$ accelerates the incorporation rates of N-methyl AAs ....................... 26
  The rate-limiting step ............................................................................................. 27
  The ribosome “proofreads” N-methyl AAs ............................................................ 28
Ribosomal polymer synthesis (Paper II) .................................................................. 31
  Kinetics assay for studying polymer synthesis ...................................................... 32
  Overcoming the kinetic defect of the penultimate dC .......................................... 32
  tRNA$^{Pyl}$-mediated amber suppression (Paper III) ................................................ 35
    Weak EF-Tu binding of AA-tRNA$^{PylB}$ ................................................................. 37
    Inefficient ribosomal binding of the ternary complex ........................................ 38
    Competing reaction from the non-cognate RF2 .................................................. 38
    Impaired elongation processivity of tRNA$^{PylB}$ .................................................. 39

Conclusions and future outlook .............................................................................. 41

Sammanfattning på Svenska ..................................................................................... 44

Acknowledgements ................................................................................................. 46

References ............................................................................................................... 49
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>adenosine</td>
</tr>
<tr>
<td>AA</td>
<td>amino acid</td>
</tr>
<tr>
<td>aaRS</td>
<td>aminoacyl-tRNA synthetase</td>
</tr>
<tr>
<td>AA-tRNA</td>
<td>aminoacyl-tRNA</td>
</tr>
<tr>
<td>aG</td>
<td>allyl-glycine</td>
</tr>
<tr>
<td>A site</td>
<td>aminoacyl-tRNA site</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine 5’-triphosphate</td>
</tr>
<tr>
<td>C</td>
<td>cytidine</td>
</tr>
<tr>
<td>dC</td>
<td>deoxy cytidine</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>E site</td>
<td>exit site</td>
</tr>
<tr>
<td><em>E. coli</em></td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>EF</td>
<td>elongation factor</td>
</tr>
<tr>
<td>fMet</td>
<td>formylmethionine</td>
</tr>
<tr>
<td>G</td>
<td>guanosine</td>
</tr>
<tr>
<td>GDP</td>
<td>guanosine 5’-diphosphate</td>
</tr>
<tr>
<td>GTP</td>
<td>guanosine 5’-triphosphate</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>IF</td>
<td>initiation factor</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>mS</td>
<td>O-methyl-serine</td>
</tr>
<tr>
<td>NMA</td>
<td>N-methyl-alanine</td>
</tr>
<tr>
<td>NMF</td>
<td>N-methyl-phenylalanine</td>
</tr>
<tr>
<td>NGM</td>
<td>N-methyl-glycine</td>
</tr>
<tr>
<td>NVOC</td>
<td>nitroveratryloxycarbonyl</td>
</tr>
<tr>
<td>P site</td>
<td>peptidyl-tRNA site</td>
</tr>
<tr>
<td>PEP</td>
<td>phosphoenol pyruvate</td>
</tr>
<tr>
<td>Pi</td>
<td>inorganic phosphate</td>
</tr>
<tr>
<td>PTH</td>
<td>peptidyl-tRNA hydrolase</td>
</tr>
<tr>
<td>Pyl</td>
<td>pyrrolysine</td>
</tr>
<tr>
<td>PylRS</td>
<td>pyrrolysyl-tRNA synthetase</td>
</tr>
<tr>
<td>RF</td>
<td>release factor</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>rRNA</td>
<td>ribosomal RNA</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>U</td>
<td>uridine</td>
</tr>
</tbody>
</table>
Introduction

The central dogma and the genetic code

Protein synthesis (translation) is a fundamental process in all living organisms. Genetic information stored in deoxyribonucleic acid (DNA) is transcribed into messenger ribonucleic acid (mRNA), which serves as the template for protein synthesis. This is the central dogma of molecular biology. Within the living organisms, proteins perform various functions, including assembling the cell skeleton, responding to stimuli, transporting molecules and most importantly catalyzing different metabolic reactions. Much of the cell mass and energy is devoted for protein synthesis. Fast growing Escherichia coli (E. coli) cells need to duplicate the complete set of proteins that makes up about 55% of the cell dry mass within some 20 min (1, 2), as such the protein synthesis machinery accounts for half of the cell dry mass (3).

Ribosome, the translation machinery, reads the genetic code on the mRNA and catalyzes the polymerization of amino acids (AAs) to produce proteins. The mRNA is composed of four types of nucleotides, adenosine (A), guanosine (G), cytidine (C) and uridine (U), and genetic code is in form of three-letter word (triplet codon). There are $4^3 = 64$ codons, each encodes for one AA or the termination of protein synthesis. In contrast, proteins are made of only 20 kinds of AAs. In most of the cases, one AA can be encoded by more than one codon. The codons that encode for the same AA are called synonymous codons. Adaptor molecules are needed to posit the AAs according to the sequence of the codons on the mRNA. Such adaptor molecules are transfer RNAs (tRNAs). Although different in primary sequences, tRNAs generally have secondary structures in a cloverleaf shape and tertiary structures in a compact L shape (E. coli tRNA$^{Phe}$ as an example shown in Figure 1A,B). In E. coli, the tRNA structure consists of the following characteristics (Figure 1): i) a 5’-terminal monophosphate group (5’ p in Figure 1A); ii) a 3’ CCA end (in orange); iii) an acceptor stem (in magenta); iv) a D-stem-loop (in red) that contains dihydrouridine; v) an anticodon stem-loop (in blue) that contains the anticodon (in light blue) to read the genetic code on the mRNA; vi) a variable loop (in black); vii) a ТΨС stem-loop (in green) that contains the conserved thymine (T)-pseudouridine (Ψ)-cytidine (C) motif; and viii) posttranscriptional modifications of some nucleotides (with gray background). AAs are linked to tRNAs via ester bond before being delivered to
the ribosomes programmed with mRNAs. One type of AA can be linked to one or more types of tRNAs. The tRNAs that can deliver the same type of AA are called isoacceptors. One tRNA isoacceptor may decode one or more types of codons of the corresponding AA.

![Figure 1](image-url)

**Figure 1.** Secondary (A) and tertiary (B) structures of tRNA. *E. coli* tRNA^{Phe} is shown as an example. Color scheme: acceptor stem, magenta; D-stem-loop, red; anticodon stem-loop, blue; anticodon, light blue; variable loop, black; TΨC stem-loop, green; 3'-CCA end, orange; post-transcriptional modification, grey background. Primary sequence and secondary structure were depicted according to (4). Tertiary structure was produced with UCSF Chimera based on PDB 1OB2.

Normally, there are twenty aminoacyl-tRNA synthetases (aaRSs) in cells, and each is responsible for linking one kind of AA to the corresponding tRNA (or tRNAs). This reaction is termed aminoacylation, or charging. The aaRS first binds the cognate AA and ATP to form aminoacyl-AMP, releasing the pyrophosphate (PPi):

\[
\text{AA + ATP } \longrightarrow \text{ aminoacyl-AMP + PPi.}
\]

This is followed by binding of the proper tRNA and transferring the AA from aminoacyl-AMP to the 3’-terminal adenosine, either on its 2’- or 3’-OH:

\[
\text{aminoacyl-AMP + tRNA } \longrightarrow \text{ aminoacyl-tRNA + AMP.}
\]

The resulted aminoacyl-tRNA (AA-tRNA) can act as the substrate for protein synthesis.
Bacterial protein synthesis

The ribosome is the protein synthesis machinery in all living cells. The 70S bacterial ribosome is assembled with two subunits, one 50S large subunit and one 30S small subunit (5). In *E. coli*, the 50S subunit is about 1.5 MDa in molecular weight and has a 5S ribosomal RNA (rRNA) (~120 nucleotides), a 23S rRNA (~2900 nucleotides) and 33 proteins; while the 30S subunit is about 0.85 MDa and is composed of a 16S rRNA (~1500 nucleotides) and 21 proteins (6-8). There are three different tRNA binding sites in the 70S ribosome: the A site for AA-tRNA binding, the P site for peptidyl-tRNA binding and the E site for the exit of the deacylated tRNA (9). Generally, the ribosome initiates the polypeptide elongation process at the start codon (initiation) and reads the genetic code, one triplet codon at a time, along the mRNA until it reaches the stop codon (elongation); the polypeptide is then released and folds into a mature protein (termination), while the ribosome will be recycled for the next round of protein synthesis (recycling) (Figure 2A) (5).

Initiation

In the initiation phase, mRNA and initiator formyl-methionine-tRNA\textsuperscript{fMet} (fMet-tRNA\textsuperscript{fMet}) will bind to the 30S subunit. This step is facilitated by three initiation factors (IFs): IF1, IF2 and IF3 (10). IF3 keeps the 50S subunit from association before the binding of mRNA and tRNA. IF2 can recognize the formyl group of fMet-tRNA\textsuperscript{fMet} and hasten its binding to the 30S subunit (10). In many of the cases in *E. coli*, the interaction between a Shine-Dalgarno sequence upstream of the start codon on the mRNA with the anti-Shine-Dalgarno sequence on the 16S rRNA guides the assembly directly with the AUG start codon in the P site (11). An elongation-competent 70S initiation complex is formed after 50S subunit joining, with the dissociation of the IFs (12). In this initiation complex (Figure 2A), fMet-tRNA\textsuperscript{fMet} occupies the P site that is programmed with the AUG start codon and the vacant A site is available for ternary complex binding.

Elongation

In the elongation phase (Figure 2A), the AA-tRNA that reads the A site codon is delivered to the ribosome in ternary complex with elongation factor thermo-unstable (EF-Tu) and GTP. The AA-tRNA is in its A/T state, with the anticodon stem loop in the decoding center in the A site of the 30S ribosomal subunit, while the acceptor stem remaining EF-Tu:GTP bound (13). Upon the formation of correct codon-anticodon interaction, the GTPase activity of EF-Tu is activated and GTP will be hydrolyzed (14). The 3’ CCA end of the AA-tRNA is then released from EF-Tu:GDP and the AA-tRNA
Figure 2. Overview of different functional phases in bacterial protein synthesis. (A) During initiation phase, the 70S initiation complex is formed with fMet-tRNA\textsuperscript{fMet} programmed at the ribosomal P site, which interacts with the AUG start codon of the mRNA. The initiation complex enters the elongation cycle with the binding of AA-tRNA:EF-Tu:GTP ternary complex. After formation of correct codon-anticodon interaction, GTP is hydrolyzed on EF-Tu, followed by the release of EF-Tu:GDP and inorganic phosphate (P\textsubscript{i}) from the ribosome and the accommodation of the AA-tRNA into the A site. The nascent peptide chain (or fMet for the first elongation event) is transferred from the P site tRNA to the AA-tRNA, extending the peptide chain by one amino acid in peptidyl transfer reaction. This is followed by EF-G catalyzed translocation, in which the P and A site tRNAs are translocated to E and P sites, respectively. Meanwhile, the ribosome moves by one codon along the mRNA, displaying the next codon in the A site. The ribosomal complex is then ready to enter the next elongation cycle. When the A site codon is one of the stop codons, translation termination will occur to release the newly synthesized protein. And the ribosome will be recycled to allow the next round initiation of translation. Figure modified from (15). (B) The proposed peptidyl transfer mechanism: a six-membered transition state is formed when the α-amino group of the A site AA-tRNA (green) nucleophilically attacks the carbonyl carbon of the ester bond that links fMet (or peptide) to the corresponding P site tRNA (blue) (16). R represents the side chain of the AA. Figure modified from (17).

Fully accommodates to the A site, and accept the fMet (or peptide from the second elongation cycle) from the P site tRNA by forming a peptide bond in the peptidyl transfer reaction (Figure 2). In the peptidyl transfer reaction, the α-amino group of the A site AA-tRNA nucleophilically attacks the carbonyl carbon of the ester bond that links fMet (or peptide) to the corresponding P site tRNA (Figure 2B) (17-20). Ribosome accelerates this reaction by ~10\textsuperscript{9}-to 10\textsuperscript{7}-fold, compared to the reaction in solution (21). After peptide bond formation, elongation factor G (EF-G) in complex with GTP will catalyze the movement of the P site deacylated tRNA to the E site and the A site newly synthesized peptidyl-tRNA to the P site, with the consumption of the energy of GTP hydrolysis (Figure 2A) (22). This step, named translocation, is coupled with the ribosome moving one codon in respective to the mRNA, displaying the next codon in the A site. After translocation, the ribosome is ready to repeat such elongation cycles to extend the nascent polypeptide chain by one amino acid per codon until the A site displays the stop codon. Interestingly, recent studies revealed that under certain circumstances, catalyzed by another GTPase (LepA, or EF4), the translocation step could be reversed (23). This back-translocation mechanism was proposed to be important for resolving stalled ribosomes (15).

Termination
When stop codon is programmed in the ribosomal A site, depending on the type of the stop codon, one of the two class-1 release factors (RF1 or RF2) will bind to the ribosomal A site and catalyze the hydrolysis of the ester
bond that links the nascent polypeptide and the P site tRNA. RF1 recognizes stop codons UAG and UAA, while RF2 recognizes UGA and UAA (24). Subsequently, the class-2 release factor, RF3, will accelerate the dissociation of the class-1 release factor from the ribosomal complex (25). The resulted ribosomal complex is called the post-termination complex, in which the P site is bound with one deacylated tRNA.

Recycling

After the post-termination complex is formed, the ribosome needs to be recycled to initiate the next round protein synthesis. Ribosome recycling factor (RRF) and EF-G will bind to the post-termination complex, and trigger the dissociation of the 50S subunit from the complex in a GTP-hydrolysis-dependent manner (26). IF3 will then bind to the 30S subunit to release the deacylated tRNA and prevent the subunits from re-association (27). The mRNA will be free to dissociate from the ribosomal subunit after the dissociation of the tRNA. Recent evidence also indicate that new initiation complex can be formed by the post-termination 70S ribosome scanning along the mRNA to the start codon, triggered by fMet-tRNA$^\text{fMet}$ with the help of IFs (28). This would allow protein synthesis via bypassing the recycling pathway.

Unnatural amino acids

Among the 20 natural AAs, except for proline (Pro) that is a cyclic imino acid, all the others are α-AAs; except for glycine (Gly) that has no stereoisomer, all the others are L-AAs (Figure 3A). In nature, proteins usually require chemistries beyond the 20 natural AAs for their functions. Posttranslational modifications by covalently modifying the AAs via their side chains are required for proper structural, conformational and physicochemical properties of proteins (29). For instance, acetylation and deacetylation of lysine residues in histones are parts of gene expression regulation (30); serine or tyrosine phosphorylation plays an important role in signaling pathways (31); hypusination of eukaryotic translation initiation factor 5A (eIF5A) and β-lysylatation of bacterial elongation factor P (EF-P) at a key lysine residue are essential for the proteins to prompt ribosomal incorporations of consecutive prolines (32–34). If such modifications could be incorporated co-translationally with modified AAs, it would allow site-specific modification of the proteins for probing the function of the modifications and for studying the function of the proteins.
**Figure 3.** Natural and unnatural AAs. R represents different chemical groups. In N-alkyl-AAs, R₂ is methyl group for N-methyl AAs.

Furthermore, apart from the 20 natural AAs, there are many other types of naturally occurring AAs, such as L-α-AAs with different side chains (as the modified AAs mentioned above; when R is unnatural side chain in Figure 3A), D-AAs, N-alkyl-AAs and β-AAs (Figure 3B). These AAs are not involved in natural ribosomal protein synthesis, and are referred as unnatural AAs in this thesis. A better understanding of the mechanism by which the ribosome discriminates unnatural AAs from the natural ones will provide insights into why evolution selected the 20 natural AAs for ribosomal protein synthesis. Interestingly, many of these unnatural AAs even exist in natural peptide products, but are incorporated by nonribosomal peptide synthetases. Some of these nonribosomal peptides are of important pharmaceutical values (35). For example, the widely used immunosuppressant drug, cyclosporine A (Figure 4A), is a cyclic nonribosomal peptide of eleven AAs and contains several N-methyl AAs and one D-AA (36); vancomycin (Figure 4B), one antibiotic that is used for the treatment of infections caused by Methicillin-resistant *Staphylococcus aureus* (MRSA) (37), is also a nonribosomal peptide with unnatural AA moieties. Randomized libraries of such peptidomimetics are invaluable compounds in drug screening (38-41). The unnatural AA moieties may improve the pharmacological properties of the peptides by increasing affinity to the target, resistance to proteolysis and membrane permeability. Efficient co-translational incorporation of unnatural AAs into peptidomimetics will enable creating large libraries since the sequences of the AAs are encoded in the mRNA sequences, which can have enormous variants. Such libraries can be selected against drug targets to discover potential drug leads.
Ribosomal incorporation of unnatural AAs \textit{in vivo}

**Genetic encoding of unnatural AAs and tRNA$^{\text{Pyl}}$**

Genetically encoding one unnatural AA in living cells requires: i) the unnatural AA is cell permeable and metabolically stable; ii) one “open” codon to encode; iii) one tRNA can and only can read this codon; iv) one aaRS to charge this tRNA with the AA; and v) the unnatural AA-tRNA is compatible with the translation factors and ribosomes. Orthogonal AA/tRNA/aaRS pairs (Figure 5) are developed for incorporating unnatural AAs into protein synthesis \textit{in vivo} (43). Orthogonal signifies that the unnatural AA/tRNA/aaRS can function independently from their natural counterparts with cross-reaction. Due to requirement ii), the orthogonal system usually recodes one stop codon for the unnatural AA. Naturally, two AAs other than the 20 natural ones were found to be genetically encoded via suppression of stop codons (44, 45). Selenocysteine, the 21st natural AA, is incorporated by suppression of the UGA stop codon with a complicated translation mechanism that requires special protein factors and RNA elements (46). This natural stop codon suppression system is hence difficult to be manipulated. In contrast, the 22nd natural AA, pyrrolysine (Pyl), is encoded in a very similar way to those for the other 20 AAs (47, 48). It was originally discovered in archaeal species \textit{Methanosarcina barkeri} (44). In those organisms and others (like \textit{Desulffitobacterium}), tRNA$^{\text{Pyl}}$ is aminoacylated with Pyl by the Pyl-tRNA synthetase (PylRS) and can suppress the UAG amber stop codon to encode Pyl.
Figure 5. Orthogonal AA/tRNA/aaRS pair. The unnatural AA, its orthogonal tRNA (O-tRNA) and orthogonal aaRS (O-aaRS) should not cross-react with any of their natural counterparts (indicated by dashed arrows). Figure modified from (49).

Selection of an orthogonal tRNA/synthetase pair
The most straightforward way to create one orthogonal system is by importing a heterologous tRNA/aaRS pair from another life kingdom to the host cell, given that tRNA modifications and tRNA recognition by the aaRS may differ among domains or species (50). To do this, the genes for the aaRS and the tRNA are usually cloned to a plasmid vector and transformed into the host cells. The anticodon of the tRNA is mutated to enable recoding the “open” codon. Mutagenesis is introduced to the aaRS to enable aminoacylation of the tRNA with the unnatural AA. The orthogonality of the tRNA/aaRS pair in the host cells can be verified with two rounds of selections, one positive selection followed by a negative selection (reviewed in (49)). In the positive selection, the “open” codon is introduced to an essential gene of the host cells. Only if the introduction of the unnatural AA/tRNA/aaRS pair could lead to recoding of this codon can the cells express the essential gene and survive the selection. To ascertain that the survival of the host cells is not due to the cross-reactions of this pair with the native ones, in the negative selection, the unnatural AA is omitted from the culture while the tRNA/aaRS pair is transformed into the host cells. At the same time, the “open” codon is introduced to a toxic gene. Any cells that could survive the negative selection will indicate cross-reactions, and the others should have the plasmid that contains the orthogonal variant of the tRNA/aaRS pair selected for encoding the unnatural AA.

Orthogonal ribosome and quadruplet codons
The Pyl/tRNA^Pyl/PylRS pair was proved to be orthogonal not only in bacteria, but also in eukaryotic cells and animals (49). So far, the majority of the
orthogonal systems are based on recoding the amber stop codon with the suppressor tRNA<sub>Pyl</sub>. By directed evolution of the PylRS, a broad spectrum of unnatural AAs was encoded in living cells through tRNA<sup>Pyl</sup>-mediated amber suppression (50, 51). Orthogonal ribosome/mRNA pairs are selected to enhance the incorporation of unnatural AAs in vivo (52). The anti-Shine-Dalgarno sequence of the 16S rRNA in the orthogonal ribosome is mutated, such that the orthogonal ribosome will favor the translation of the mRNA with the mutated Shine-Dalgarno sequence. Together with the mutations in the decoding center, the orthogonal ribosome can enhance incorporation of unnatural AAs into recombinant proteins by decreasing the competition reaction of RF1 (53). The effect was further strengthened by knocking out RF1 from the host cells (54, 55).

In order to simultaneously encode two or more unnatural AAs in vivo, which is limited by the requirement ii) mentioned above, the genetic code is expanded to contain four-letter words, the quadruplet codons (56). Meanwhile, two or multiple orthogonal tRNA/aaRS pairs are needed for encoding such quadruplet codons. The anticodons of the orthogonal tRNAs are extended to four nucleotides. In theory, applying quadruplet codons will add $4^4 = 256$ new codons to the genetic code. To achieve efficient quadruplet decoding, a library in the decoding center of the orthogonal ribosome was created and selected (57). The resulted orthogonal ribosome variant, in combination with the orthogonal quadruplet codon or amber codon suppressor tRNA/aaRS pairs, enabled ribosomal simultaneous incorporations of two or more unnatural AAs in vivo.

**Ribosomal incorporation of unnatural AAs in vitro**

The in vitro translation system was reconstituted with cell extracts (reviewed in (58)) or with hexa histidine-tagged purified components (59, 60). Purified components render the system with more flexibility to be engineered, since the system can be free of constraint from other unwanted components. To reconstitute a purified translation system, ribosomes, individual wild-type or tagged and overexpressed protein factors and aaRSs, and native tRNAs can be purified from cells. In addition, unmodified tRNAs can be easily synthesized and purified from in vitro transcription with T7 RNA polymerase and most of them were shown to be active in translation (61). The production of mRNA can be done in the same way, or be coupled in situ with the translation reaction (in vitro transcription and translation system). “Open” codons can be either stop codons by omitting one of the class-I RFs, or be sense codons by omitting the AAs/tRNA isoacceptors/aaRSs that cells use to decode these codons. Several methods are developed to acylate tRNAs with unnatural AAs in vitro:
Synthetase charging method

Some aaRSs can aminoacylate the cognate tRNAs with not only native substrates, but also their analogs (62). Also, some of them, such as AlaRS, can charge their cognate tRNAs even with swapped anticodons (63). Besides, as described above, aaRSs can be evolved to enable aminoacylation of tRNAs with various unnatural AAs. But this is very laborious and the resulting aaRSs generally have a loss of catalytic efficiency by a factor ~1000 (51). And one evolved aaRS variant can only be applied to one kind of unnatural AA. This method is thus not flexible.

Flexizyme charging method

By in vitro directed evolution, the Suga lab selected some ribozymes, termed Flexizymes, that are able to catalyze the aminoacylation of any tRNAs with any kinds of AAs activated with varying leaving groups (Figure 6) (64). This method has enabled charging tRNAs with unnatural L-AAs, D-AAs, β-AAs, N-alkyl-AAs, hydroxyl acids and even peptides (Figure 6B) (39). Different versions of the flexizymes have preferences for different leaving groups and AA properties. The tRNAs can be native tRNAs, as well as unmodified synthetic tRNAs. In general, flexizymes catalyze the acylation of the 3’-OH of the 3’-terminal A of tRNAs (65). One conserved U in the flexizyme sequences becomes the apex of a hairpin turn (marked with an asterisk in Figure 6B) (66). The 3’-terminal GGU of the flexizyme is then available to form base pairing interaction with the 3’-terminal ACCA of the tRNA, while the 3’-terminal A of the tRNA is positioned close to the activated AA for the catalysis (Figure 6B) (66). The application of the flexizymes has been extended by the creation of an orthogonal ribosome/tRNA system (67). In this orthogonal system, penultimate C of the tRNA is mutated to a G, and compensatory mutations (G2553C, G2251C) are introduced to the 23S rRNA of the E. coli ribosome at the nucleotides that base pair with this tRNA nucleotide in the ribosomal A and P sites. Accordingly, the flexizymes are mutated to enable the charging of the mutant tRNAs. Although very flexible, this method usually requires optimization of the aminoacylation reaction to achieve maximal yield and in some cases the yields are relatively low.
Figure 6. Flexizyme ribozymes. (A) Sequence alignment of different versions of flexizymes (68). Those bases originated from random sequences during the in vitro selection are in bold and the conserved U is marked with an asterisk. (B) Mechanism of flexizyme ribozymes catalyzed aminoacylation of tRNA and the chemical structures of their substrates (69). Abbreviations: LG, leaving group; DBE, dinitrobenzyl ester; CME, cyanomethyl ester; CBT, chlorobenzyl thioester; ABT, amino-derivatized benzyl thioester. R represents various side chains, and X indicates the substrates can be different kinds of AAs, hydroxyl acids and peptides.

Chemoenzymatic ligation charging method

This method takes advantage of the conserved 3′-CCA end of the tRNAs and relies on the reactivity of T4 RNA ligase (70, 71). In chemical synthesis, dinucleotide 5′-pCpA (p stands for the phosphate backbone) is acylated with natural or unnatural AAs at the 2′ or 3′-OH of the A. During the synthesis, the α-amino group of the AA is usually protected by a photo-labile protecting group, nitroveratryloxy carbonyl (NVOC) group. The resulted N-NVOC-AA-pCpA is then ligated to the 3′-CA truncated unmodified tRNA body by T4 RNA ligase to produce the N-NVOC-AA-tRNA. After photolytic removal of the NVOC group, the resulted AA-tRNA is ready to function as the substrate for translation reaction.

\[
\text{tRNA}_{\text{minus 3′-CA}} + N\text{-NVOC-pCpA} \xrightarrow{\text{T4 RNA ligase}} N\text{-NVOC-AA-tRNA} \xrightarrow{\text{UV light}} \text{AA-tRNA}
\]

For the ease of the chemical synthesis, pdCpA (d signifies that the C here is the deoxy cytidine) instead of pCpA was generally used in this method, as
RNA might be degraded in the basic conditions required for the synthesis (72). We recently improved this technology by developing a facile AA-pCpA synthesis method such that the unnatural dC can be circumvented (paper IV, (73)). This method allows charging any unmodified tRNA with any natural and unnatural AAs. However, it is difficult to charge native tRNAs in this way as removing the 3’-terminal CA from tRNA is not readily feasible.

**Dissertation overview**

In general, the efficiencies of ribosomal incorporation of unnatural AAs are low, and hence their applications are limited. Implementing *in vitro* kinetics approaches, this thesis strives towards a better understanding of the mechanism of protein synthesis with unnatural AAs, in particular during the elongation phase. Our results suggest ways to improve the ribosomal incorporation efficiencies of unnatural AAs to facilitate their applications. The thesis is based on three scientific publications with the following contents:

- **Paper I**  Kinetic studies of ribosomal single incorporation of *N*-methyl AAs were performed. The rate-limiting step in the reaction was determined and results suggested ways to improve the incorporation efficiencies.

- **Paper II**  Kinetics of ribosomal consecutive incorporations of non-*N*-alkylated, unnatural L-AAs was studied. Different tRNA-aminoacylation methods were evaluated. Results were compared with those from batch translation reactions to clarify the inefficiencies and improvements.

- **Paper III**  The most widely used tool for genetically encoding unnatural AAs *in vivo*, tRNA<sup>Pyl</sup>-mediated amber suppression, was studied *in vitro*. Kinetics studies revealed unexpected limiting steps and competing reactions.

I performed all the biochemical experiments in these three publications, except for the chemical synthesis of the *N*-NVOC-AA-pdCpA, *N*-NVOC-AA-pCpA and AA-DBE analogs. I was also heavily involved in writing the papers.
The present work

Experimental setup

*In vitro* translation system

Bacterial protein synthesis is a complicated process that involves many factors and multiple steps. Studying the mechanisms of translation *in vivo* is challenging. Different forms of *E. coli in vitro* translation system were developed to study ribosomal incorporation of unnatural AAs, among which S30 extract system (74), PURE (Protein synthesis Using Recombinant Elements) system (60) and its variants are the most popular systems used by others. Such systems usually are coupled transcription-translation systems. In most of the cases, a DNA template that contains the recoded codons for incorporation of unnatural AAs is added to the reaction mixture along with pre-charged and purified unnatural AA-tRNAs. The reaction mixture is incubated for an hour or so. The reactions occurring in the system includes transcription of mRNA, translation initiation, elongation, termination, ribosome recycling, as well as energy regeneration and most likely some competing reactions. The products, labeled peptides or proteins, are analyzed with polyacrylamide gel electrophoresis. With such systems, the translation compatibilities of different unnatural substrates were evaluated, such as translation initiation with non-methionine AAs (75), N-acyl AAs (75), D-AAs (76) and exotic peptides (77), and translation elongation with N-alkyl AAs (78-81), cyclic N-alkyl AAs (82), D-AAs (83), β-AAs (84, 85), dipeptidomimetics (86) and hydroxy acids (87). Also, mutants of ribosomes and elongation factors were compared for improved ribosomal incorporation of unnatural AAs (88, 89). In such batch translation systems, however, it is unclear at which stage(s) the unnatural substrates are discriminated by the natural translation system, making the studies less informative. Hence, one system that could allow interrogating, dissecting and understanding the mechanism of the complex translation reaction is needed.

To this end, in this thesis, we applied a purified *E. coli* translation system (90) to study the kinetics of ribosomal incorporation of unnatural AAs. Wild-type *E. coli* ribosomes and total tRNAs were purified from MRE600 cells (91). Each of the translation initiation, elongation and release factors was hexa histidine-tagged, overexpressed and purified from *E. coli* cells (91,
Tritium labeled fMet-tRNA\textsuperscript{fMet} was prepared by charging tRNA\textsuperscript{fMet} with \[^{3}\text{H}]\text{Met} by MetRS and subsequently formylated by Met-tRNA \textit{N}-formyltransferase (93). Synthetic mRNAs and tRNAs were prepared by \textit{in vitro} transcription with T7 RNA polymerase from DNA templates that were made from primer extension with designed oligos or restriction enzyme digested plasmids (81). The methods used for aminoacylation of tRNAs will be further specified below. Reactions were performed in a polymix buffer containing 5 mM potassium phosphate (for Paper II and III, replaced by 30 mM HEPES in Paper I), 95 mM KCl, 5 mM NH\textsubscript{4}Cl, 5 mM Mg(OAc)\textsubscript{2} (unless otherwise specified), 0.5 mM CaCl\textsubscript{2}, 8 mM putrescine, 1 mM spermidine and 1 mM dithierythritol, supplemented with 10 mM (for Paper II and III, 2 mM for Paper I) phosphoenolpyruvate (PEP), 1 µg/mL pyruvate kinase, 0.1 µg/mL myokinase and 2 mM ATP+GTP (ribosomal mix, ATP : GTP = 1 : 1; ternary complex mix, only ATP for GTP hydrolysis experiments, while ATP : GTP = 1 : 1 for dipeptide/tripeptide/polypeptide formation experiments) for energy regeneration (94, 95). Assuming that one ATP or GTP molecule chelates one Mg\textsuperscript{2+} ion, and the dissociation constant, \(K_d\), for PEP chelating Mg\textsuperscript{2+} ion is 6 mM (96), the free Mg\textsuperscript{2+} in the polymix buffer is around 1.3 mM (for Paper II and III, and 2.4 mM for Paper I). By strategic combination and pre-incubation of the components, our \textit{in vitro} translation system has been used to study the kinetics of translation initiation (10), tRNA selection (97), peptide bond formation (4, 91, 95, 98), translocation (99), termination (100) and recycling (26).

**Experimental procedures**

Kinetics studies were performed as described (4, 91, 95). With radioactive labeled AAs and GTP, we could monitor the peptide bond formation in translation elongation and the associated GTP hydrolysis. In general, two reaction mixtures were prepared in polymix buffer: one ribosomal mix containing vacant 70S ribosomes, mRNA, \[^{3}\text{H}]\text{fMet}-tRNA\textsuperscript{fMet} and initiation factors; one ternary complex mix containing EF-Tu, EF-Ts (omitted for GTP hydrolysis experiments, but \[^{3}\text{H}]\text{GTP was added}), AA-tRNA(s) (or AA/tRNA/aaRS for \textit{in situ} charging), with or without EF-G (without if only measure up to dipeptide formation). The two mixtures were pre-incubated at 37 °C for 15 min to pre-form initiation complex and ternary complex. In this way, we could investigate the kinetics of only the elongation phase. Reactions were done by rapidly combining equal volumes of the two mixtures. Reactions were quenched by mixing with a quencher (final 17% formic acid here) at designed time points. When the reactions were slow enough, the mixing and quenching of the reactions could be performed by hand. When fast reactions were measured, temperature-controlled quench-flow apparatus was used to quench the reactions at very short time points. The resulted samples contained peptidyl-tRNA products labeled with \[^{3}\text{H}]\text{fMet residue}
(and $[^3]$H]GDP in GTP hydrolysis experiments). Since acid precipitates RNA, after centrifuging the samples, the $[^3]$H]peptidyl-tRNAs and the unreacted $[^3]$H]fMet-tRNA are in the pellet, while the $[^3]$H]GDP and the unreacted $[^3]$H]GTP are kept in the supernatant. By running the supernatant on a Mono Q HPLC column coupled with a radioactive detector, we could quantify the fraction of $[^3]$H]GTP that was hydrolyzed at each time point. To the pellet, potassium hydroxide was added to hydrolyze the $[^3]$H]peptides and unreacted $[^3]$H]fMet from tRNAs. This was followed by addition of formic acid to final 17% and subsequent centrifugation. The supernatant from this step was analyzed by C18 reversed-phase HPLC equipped with a radioactive detector to quantify the fraction of the remaining $[^3]$H]fMet and the produced $[^3]$H]peptides. The reaction rate ($k$), which is the inverse of the average time ($\tau$) of the reaction, could be estimated from the time-evolution of the products by fitting the data to a suitable kinetic model.
Single incorporation of N-methyl AAs (Paper I)

With batch translation reactions, different research groups have evaluated the compatibilities of N-methyl AAs on the natural translation machinery and results showed some discrepancies (78-81). Nevertheless, what limited the generally low ribosomal incorporation efficiencies of N-methyl AAs was unclear. By kinetics study, our group demonstrated for the first time that N-methyl-Phe (NMF) and N-butyl-Phe incorporated much slower into dipeptide compared with Pro (95). Similar reactivities were observed in a chemical model system (101), providing a reasonable explanation why Pro is the only N-alkyl AA naturally encoded even though several other N-alkyl AAs are major products of synthetic prebiotic experiments and meteorite analyses (95). Furthermore, the rates of GTP hydrolysis ($k_{\text{GTP}} = 1/\tau_{\text{GTP}}$) on EF-Tu were similar, indicating the reaction was not limited by the delivery of NMF-tRNA to the ribosomal A site, EF-Tu GTPase activation or GTP hydrolysis (Figure 7) (95). Rather, the reaction might be limited by retardation of release of EF-Tu:GDP, accommodation of NMF-tRNA from A/T state to A/A state, and/or peptidyl transfer reaction (Figure 7). Interestingly, Pro incorporated slower when delivered by the non-cognate synthetic tRNA$_{\text{PheB}}$ (tRNA$_{\text{PheB}}$-based synthetic tRNA) than when delivered by the cognate tRNA$_{\text{Pro}}$ isoacceptors (95). This leads to the hypothesis that tRNA$_{\text{Pro}}$ might improve the incorporation efficiency of N-methyl AAs compared with tRNA$_{\text{Phe}}$.

Figure 7. Schematic illustration of translational steps measured in the experiments. During pre-incubation of the ribosomal mix and ternary complex mix, 70S initiation complexes and ternary complexes are formed. In the ternary complex mix, there is an equilibrium of AA-tRNA binding and dissociating from EF-Tu:GTP. $K_d$ is the dissociation constant. When this $K_d$ is small and the concentration of either AA-tRNA or EF-Tu:GTP is in large molar excess over the other, the equilibrium will be kept with nearly complete ternary complex formation. The reaction starts by mixing initiation complexes with ternary complexes. Upon correct codon-anticodon interaction, the GTPase activity of EF-Tu will be activated and GTP will be hydrolyzed on EF-Tu. There is also chance for the ternary complex being rejected by the ribosome. Normally, for cognate substrates, this chance is very small (102). $\tau_{\text{GTP}}$ is the average time spent from the start of the reaction to GTP is hydrolyzed on EF-Tu. $\tau_{\text{dip}}$ is the average time spent from the start of the reaction to dipeptide is formed. $\tau_{\text{pep}}$ is the average time for all the steps subsequent to GTP hydrolysis that lead to peptide bond formation.
tRNA$^{\text{Pro}}$B accelerates the incorporation rates of N-methyl AAs

We therefore charged three N-methyl AAs, NMF, NMA (N-methyl Ala) and NMG (N-methyl Gly), each onto three different synthetic tRNAs, tRNA$^{\text{Phe}}$B, tRNA$^{\text{Ala}}$B and tRNA$^{\text{Pro}}$B (Figure 8), with the N-NVOCA-APdCpA chemoenzymatic ligation method. Kinetics of dipeptide formation from fMet-tRNA$^{\text{Met}}$ and N-methyl AA-tRNAs at 37 °C pH 7.8 revealed that tRNA$^{\text{Pro}}$B was able to accelerate the dipeptide formation rates ($k_{\text{dip}} = 1/\tau_{\text{dip}}$) of all the surveyed three N-methyl AAs (Figure 8B,C,D). This could not be explained simply by increased EF-Tu binding affinities, as the rate of NMF-tRNA$^{\text{Pro}}$B incorporation were independent of EF-Tu concentrations (Supplementary Figure 1 in Paper I, similar results were observed for NMF-tRNA$^{\text{Phe}}$B and NMF-tRNA$^{\text{Ala}}$B (98)) and the reaction rate was not limited by the steps up to GTP hydrolysis on EF-Tu (Figure 7) (95). We postulate that tRNA$^{\text{Pro}}$ body can posit the $N$ nucleophile of the $N$-methyl AAs in a better way to allow peptidyl transfer happening (18).

---

**Figure 8.** Kinetics of dipeptide synthesis with different N-methyl AA-tRNAs. (A) Synthetic, unmodified tRNAs based on natural *E. coli* tRNA$^{\text{Phe}}$, tRNA$^{\text{Ala}}$ and tRNA$^{\text{Pro}}$ sequences (black with purple anticodon, post-transcriptional modifications are in green), with changes in blue. Post-transcriptional modifications for the studied tRNA$^{\text{Pro}}$ isoacceptor are unknown. Kinetics of dipeptide synthesis from fMet-tRNA$^{\text{Met}}$ with NMF (B), NMA (C) and NMG (D) delivered by tRNA$^{\text{Phe}}$B, tRNA$^{\text{Ala}}$B and tRNA$^{\text{Phe}}$B at pH 7.8 were compared. Representative plots are shown.
The rate-limiting step

We also noticed even for the fastest reaction among the N-methyl AA incorporations we measured, the reaction with NMA-tRNA\textsuperscript{ProB}, the rate was nearly 100-fold slower compared with the rate of the reaction with Ala-tRNA\textsuperscript{AlaB} \cite{98}. Finding out the rate-limiting step subsequent to GTP hydrolysis might be the key to further improve the incorporation efficiencies of N-methyl AAs. The overall rate of the steps subsequent to GTP hydrolysis (Figure 7) that lead to peptide bond formation, $k_{\text{pep}}$, can be calculated as $k_{\text{pep}} = 1/\tau_{\text{pep}} = 1/(\tau_{\text{dip}} - \tau_{\text{GTP}})$. It has been shown for natural AAs, including the natural N-alkyl AA Pro, that the $k_{\text{pep}}$ value is pH-dependent, which can be rationalized by the result of titrating a reaction essential $\alpha$-amino group of the A site AA (Figure 2B) \cite{17}. Such pH-dependence of the $k_{\text{pep}}$ value supports a rate-limiting peptidyl transfer reaction, rather than accommodation. Here, we performed pH-dependence experiments with NMF, NMA and NMG delivered by tRNA\textsuperscript{ProB}. Since $\tau_{\text{GDP}}$ was negligible from $\tau_{\text{dip}}$ in dipeptide formation reaction with N-methyl AAs \cite{95}, we measured the $k_{\text{dip}}$ values to approximate the $k_{\text{pep}}$ values (Figure 7). The incorporations of all the three N-methyl AAs showed striking pH-dependence: the reaction rates increased with increasing pH in the range of 7.5 to 8.5 (Figure 9A,B,C). Particularly, when the log\textsubscript{10}(rate) was plotted against pH, the initial slopes for NMF, NMA and NMG were very close to 1 (Figure 9D), indicating the increase in rates with increasing pH was due to the deprotonation of the unreactive, charged, $\alpha$-N group of the N-methyl AAs in a rate-limiting peptidyl transfer reaction (Figure 2B), similar to that reported for Pro and other AAs \cite{17}.
The ribosome “proofreads” N-methyl AAs

The delivery of NMF-\textit{tRNA}^{ProB} to the ribosome by EF-Tu:GTP is very efficient \((95)\). The accommodation of NMF-\textit{tRNA}^{ProB} most likely is also very efficient, since the methyl group is not expected to prevent the release of the AA-tRNA from EF-Tu:GDP or significantly hinder the movement of the much larger AA-tRNA inside the ribosome. For natural substrates, the probability for the ribosome to reject (proofread) the accommodated AA-tRNA after decoding the cognate codon is very low due to a very fast peptidyl transfer reaction and a very small rejection rate \((97, 102)\). The peptidyl transfer reaction with NMF is however \(~3\) orders of magnitude slower than that with Phe \((95)\). Therefore, after decoding the cognate Pro codon, the chance for the accommodated NMF-\textit{tRNA}^{ProB} to be rejected by the ribosome might be high enough to be detected. Indeed, the NMF-\textit{tRNA}^{ProB} bound ribosomes could be chased by the Pro-\textit{tRNA}^{ProB}:EF-Tu:GTP ternary complex in a stag-
gered competition experiment, indicating the ribosome did reject the accommodated NMF-tRNA\textsuperscript{ProB} (Figure 10).

![Figure 10. Staggered competition experiment.](image)

**Figure 10.** Staggered competition experiment. In the reaction buffer, NMF-tRNA\textsuperscript{ProB}:EF-Tu:GTP ternary complex (final 1 \(\mu\)M) was added to initiated 70S ribosomes (final 0.25 \(\mu\)M) at time zero and Pro-tRNA\textsuperscript{ProB}:EF-Tu:GTP ternary complex (final 1 \(\mu\)M) was then added to the reaction mixture at time 10 s. Dipeptide formation of fMet-NMF (blue open triangles) and fMet-Pro (blue filled triangles) were monitored and plotted against time. In this staggered competition reaction, the addition of Pro-tRNA\textsuperscript{ProB}:EF-Tu:GTP ternary complex at time 10 s had the same effect as addition of formic acid on quenching the fMet-NMF synthesis (note that the intersection of the curves with open circles and open triangles is at 10 s). Control experiments were done by replacing with the reaction buffer either NMF-tRNA\textsuperscript{ProB}:EF-Tu:GTP (black filled squares) or Pro-tRNA\textsuperscript{ProB}:EF-Tu:GTP (black open circles) ternary complexes. Representative plots are shown.

From direct competition experiments between NMF-tRNA\textsuperscript{ProB} and Pro-tRNA\textsuperscript{ProB} ternary complexes (Figure 5 in Paper I), we estimated that about 8 cycles of delivery and rejection of NMF-tRNA\textsuperscript{ProB} were required per peptide bond formation (Supplementary Appendix and Figure 6 in Paper I). We could also estimate the rate of the rejection to be 0.19 s\(^{-1}\). By increasing the Mg\textsuperscript{2+} concentration in the direct competition experiment, the rejection rate decreased to about 0.13 s\(^{-1}\), the number of futile cycles of delivery and rejection was reduced by 3 cycles, and the fMet-NMF dipeptide yield increased accordingly. More importantly, our N-methyl AA-tRNA substrates might be tools for studying the proofreading mechanisms of ribosomal protein synthesis (97). Direct determination of the rejection rates of natural AA-tRNAs in
the proofreading step of cognate reactions is of great experimental challenge (97, 102). We show here that slowing down the peptidyl transfer reaction enables determining the rejection rate of NMF-tRNA_{ProB}. This rate might be very similar to that of Pro-tRNA_{ProB} when reading the same cognate codon. In the same way, by charging different tRNAs with NMF, we should be able to estimate the rejection rates during proofreading of any cognate reactions. All together, we demonstrated that the efficiency of ribosomal single incorporation of $N$-methyl AAs could be improved by using the tRNA_{Pro} body, increasing the pH value and using a higher Mg$^{2+}$ concentration.
Ribosomal polymer synthesis (Paper II)

Ribosomal synthesis of polymers of unnatural AAs is the basis of directed evolution of peptidomimetic drug leads with ribosome display and mRNA display techniques (103-105). The standard N-NVOC-AA-pdCpA chemoenzymatic ligation method for charging tRNAs with unnatural AAs has been used for decades for incorporation unnatural AAs in translation (70-72). Ribosomal single incorporation of small non-N-alkyl AAs with substrates prepared by this method usually can display quantitative yields in both crude and purified systems (106, 107), and can even achieve near wild-type kinetics (4, 98). For example, when allylglycine (aG) and methyl-serine (mS) (Figure 11) were charged by the N-NVOC-AA-pdCpA chemoenzymatic ligation method to the tRNA^{AlaB} body, they could be incorporated into dipeptides with similar rates as that of Ala (98). Also, in 30 min-incubation translation reactions using purified translation system and the same types of substrates, single incorporation of aG, mS or Ala into tetrapeptides had similar yields as those of modified or unmodified Ala-tRNA^{Ala} substrates prepared from synthetase charging method (107). However, when incorporated five in a row in the batch translation reactions, the yield for Ala-tRNA^{AlaB} (with dC) decreased by about half (107). This was not due to the lack of tRNA modifications, as the yield for modified and unmodified Ala-tRNA^{Ala} substrates were similar. Rather, the penultimate dC introduced to the tRNA body from the standard N-NVOC-AA-pdCpA chemoenzymatic ligation method inhibited the incorporations. Changing the AA to aG compensated partially this defect, but not mS (107). Besides, when Ala was charged by the AlaRS to tRNA^{Ala} (without dC) with anticodon swapped to that of tRNA^{Thr}, the yield of five consecutive incorporations in the batch translation reaction decreased by 3-fold (107). Changing the anticodon to the anticodon of tRNA^{Asn} even completely abolished the reaction (107). The inhibitory effects of the dC and the anticodon swap could not be overcome by prolonging the reaction time, indicating the reactions competing with the incorporation of the AAs would occur very fast. We hence set off to study the kinetics of these reactions to investigate the mechanisms limiting polymer synthesis.

Figure 11. Chemical structures of Ala, aG and mS.
Kinetics assay for studying polymer synthesis

We first developed a kinetics assay that could monitor each incorporation step along the five consecutive incorporations to obtain detailed kinetic information of the reactions. On a quench-flow apparatus, a limited amount of 70S initiation complex programmed with five Ala codons after the AUG start codon was rapidly mixed with excess of modified AA-tRNA^{Ala}:EF-Tu:GTP ternary complex and EF-G:GTP (5 elongation cycles involved, Figure 2A). The products from each quenching time point were analyzed by C18 reversed-phase HPLC. The HPLC running condition was optimized to get the best resolution of each intermediate product along the hexapeptide synthesis (Figure 12). We therefore could follow the time-evolution of each of the intermediate product in the reaction. The rates and processivities of the incorporation steps could be deduced from the time curves with our data-fitting model (Supplementary Appendix in Paper II). In this model, the loss of processivity at each incorporation step was taken into account, as the processivity was not 100% at all the steps (Figure 12B). Anyhow, results demonstrated fast kinetics with high processivity for each incorporation step in fMet-(Ala)\textsubscript{5} (fMA\textsubscript{5}) synthesis with wild-type modified Ala-tRNA\textsuperscript{Ala} (termed Ala-tRNA_{Ala\_modified} in Paper II, Figure 2b,c and Supplementary Table 1 in Paper II).

![Figure 12](image12.png)

**Figure 12.** Kinetic assay of fMA\textsubscript{5} synthesis. (A) An example of HPLC separation of each intermediate product in fMA\textsubscript{5} synthesis. (B) Representative plots of normalized time-evolution of each intermediate in the kinetic assay of five consecutive incorporations Ala when delivered by the native tRNA\textsuperscript{Ala}.

Overcoming the kinetic defect of the penultimate dC

We next assayed the kinetic effects of the changes in the AA-tRNAs on the five consecutive incorporations: the lack of tRNA modifications, the penultimate dC and the unnatural AAAs. In line with the batch translation reactions (107), tRNA modifications had only minor effect on the rates and processivity (Figure 2b,c in Paper II). When the penultimate dC was introduced, the rates decreased by ~2-fold and the processivity decreased ~10-20% at each
incorporation step, resulting in a yield reduction by half (Figure 2 in Paper II). We found that the penultimate dC would induce faster drop-off of the peptidyl-tRNAs along the hexapeptide synthesis (Figure 13), not frameshifting (Supplementary Figure 2 in Paper II). This explains the loss of processivity during the incorporations.

![Diagram](image)

**Figure 13.** Assay for detecting intermediate peptidyl-tRNAs drop-off along fMA₅ hexapeptide synthesis. (A) Reaction procedure. Peptidyl-tRNA hydrolase (PTH, which only cleaves the peptide from tRNA while the peptidyl-tRNA is off the ribosome) was added to the reaction mixtures for fMA₅ synthesis from [³H]fMet-tRNAfMet and from Ala-tRNAAla (with or without dC). After quenching the reaction with formic acid and centrifugation of the samples, the drop-off fraction of the peptides will stay in supernatant, while the ribosome bound fraction will be in the pellet and thus been separated. (B-E) Representative results of HPLC analyses are shown for duplicate experiments. Drop-off fraction (B) and ribosome bound fraction (C) for assay with Ala-tRNAAla (without dC). Drop-off fraction (D) and ribosome bound fraction (E) for assay with Ala-tRNAAla (with dC).

When the AA was changed to aG or mS, the incorporations were faster. The loss of processivities and yield were compensated partially by aG, but not mS (Figure 2 in Paper II). All these together showed remarkable congruence.
of our fast kinetics studies and the earlier batch translation studies (107). However, in our kinetic system, the anticodon swaps had only minor effects on the yields of hexapeptide synthesis, which differed from the major effects in the batch translation reactions (Figure 2 in Paper II) (107). Nevertheless, the trend was the same: when delivered by the tRNA\textsuperscript{AlaB} with the more detrimental anticodon swap (from Ala to Asn anticodon) as determined in the batch translation system, slower incorporation rates were observed than those when delivered by the tRNA\textsuperscript{AlaB} with the less inhibitory anticodon swap (from Ala to Thr anticodon). It is plausible that our kinetic translation system is more optimal for consecutive incorporations of AAs than the 30 min incubation batch translation system.

We then demonstrated that improved consecutive incorporations could be achieved by using tRNA charging methods that could circumvent the penultimate dC (Figure 3 in Paper II). When aG or mS was charged to the tRNA\textsuperscript{AlaB} by our new \textit{N}-NVOC-AA-p\textit{CpA} chemoenzymatic ligation method (73) or the flexizyme charging method (108), the five consecutive incorporations displayed near wild-type rates and yields in our kinetic translation system (Supplementary Table 1 in Paper II).
tRNA$^{\text{Pyl}}$-mediated amber suppression (Paper III)

The orthogonal tRNA$^{\text{Pyl}}$/PylRS pair has been engineered to enable genetically encoding a broad range of unnatural AAs through suppression of the UAG amber codon (50). It is obvious that RF1-catalyzed termination at UAG codon is a competing reaction. Therefore, efforts have been spent to eliminate this competing reaction by knocking out RF1 (54, 55). Nevertheless, the efficiency of incorporation of unnatural AAs via tRNA$^{\text{Pyl}}$-mediated amber suppression is low. The reaction might be limited by the intracellular concentration of the unnatural AA, tRNA$^{\text{Pyl}}$, and/or PylRS, which further limits the amount of charged AA-tRNA$^{\text{Pyl}}$. Or there might be some defects in protein synthesis with the exogenous AA-tRNA$^{\text{Pyl}}$ substrate. To study individual steps of translation in vivo is difficult. Here, we performed in vitro kinetics studies of tRNA$^{\text{Pyl}}$-mediated amber suppression, under conditions where the reaction was not limited by the amount of charged AA-tRNA$^{\text{Pyl}}$ or by competition with RF1, to elucidate the other competing reactions and limiting steps in translation.

We charged unnatural AAs, aG and mS (Figure 11), to the unmodified synthetic tRNA$^{\text{PylB}}$ with our N-NVOC-AA-pCpA chemoenzymatic ligation method (Figure 14A,B). The two unnatural AAs were selected since they can be incorporated with comparable rates and processivities to those of the natural Ala when delivered by the tRNA$^{\text{AlaB}}$ body (as shown in Paper II), indicating the ribosome does not discriminate their unnatural side chains. Also, when charged to tRNA$^{\text{Phe}}$ or tRNA$^{\text{Ala}}$ body, they displayed similar EF-Tu binding affinities to that of the natural Phe (4, 98). Therefore, any kinetic defect detected here should be attributed to the use of the tRNA$^{\text{PylB}}$ body. Unmodified tRNA$^{\text{PylB}}$ is a good experimental model to study the kinetics, as there are relatively few post-transcriptional modifications in this archaeal tRNA (47), the modification pattern should differ in different life domains (this orthogonal pair is used in bacterial and eukaryotic cells), and it can be produced and purified from in vitro transcription with T7 RNA polymerase in sufficient quantity.
Figure 14. Measurement of EF-Tu binding affinities of the AA-tRNA$^{PylB}$ substrates. (A) The secondary structure of the native *Methanosarcina* species tRNA$^{Pyl}$. Known post-transcriptional modifications are in green and the possibility of pseudouridinylation cannot be ruled out (47). The four big differences between this tRNA and the standard tRNA architecture are marked with red boxes: (i) the connection of the D-stem loop and the acceptor stem has only one nucleotide instead of two, (ii) the D-loop is shorter by three nucleotides, (iii) the anticodon stem has 6 base pairs instead of 5, and (iv) the variable region is short. Red circles indicate missing nucleotides, while nucleotides present are numbered from the 5' end. Nucleotides conserved with *E. coli* tRNA$^{Lys}$ have a grey background. (B) The tRNA$^{Pyl}$-based synthetic tRNA$^{PylB}$ used in this work. The sequence, including the anticodon (underlined), is the same as in (A) and aminoacylation was by ligation of a transcript to N-NVOC-aminoacyl-pCpA followed by photodeprotection. Biphasic kinetics of dipeptide formation from fMet-tRNA$^{fMet}$ and aG-tRNA$^{PylB}$ (C) or mS-tRNA$^{PylB}$ (E) at varying EF-Tu concentrations at 20°C. Representative plots are shown, and the same curves without the data points are shown in Supplementary Figure 1 in Paper III to aid direct visualization of the increase in fast phase amplitudes with increasing EF-Tu. (D) and (F) Measurement of $K_d$ for EF-Tu binding. Curves show the fitting of the fast phase amplitude at different EF-Tu concentrations to a hyperbolic equation for estimation of the $K_d$ values. Bars are standard deviations of duplicate experiments.

Weak EF-Tu binding of AA-tRNA$^{PylB}$

Dipeptide formation from fMet-tRNA$^{fMet}$ and aG- or mS-tRNA$^{PylB}$ at 37 °C with 2.5 µM EF-Tu concentration showed biphasic kinetics: a merely ~15% fast phase with an unexpectedly sluggish rate of ~1.5 s$^{-1}$, and a slow phase with the rate of ~0.1 s$^{-1}$ (Figure 2 in Paper III). The fraction of the fast phase could be increased by lowering the reaction temperature to 20 °C and by increasing the EF-Tu concentration in the reaction (Figure 14). Similar scenarios have previously been observed for dipeptide formation with natural and small non-$N$-alkyl unnatural AA-tRNA substrates under the same reaction conditions (4, 98). The fast phase was attributed to the ribosomal reaction with the fraction of AA-tRNA that was already in ternary complex with EF-Tu:GTP, whereas the slower phase was due to the reaction from the AA-tRNA that needed to form ternary complex before reaction on the ribosome could occur (4). The ratio between the fast and slow phases thus reflected the equilibrium of the ternary complex formation during pre-incubation of the ternary complex mix before the reaction (Figure 7). The $K_d$ for AA-tRNA binding to EF-Tu could be estimated by fitting the different fast phase fractions ($P_{fast}$) at varying EF-Tu concentrations ([EF-Tu]) to a the hyperbolic function: $P_{fast} = P_{fast,max} [EF-Tu]/(K_d + [EF-Tu])$ (4). The $K_d$ values for EF-Tu binding of aG-tRNA$^{PylB}$ and mS-tRNA$^{PylB}$ at 20 °C were determined to be 0.45 and 0.33 µM, respectively (Figure 14B,D). At 37 °C, the $K_d$ value was estimated as ~15 µM for both the substrates based on the ~15% fast phase when [EF-Tu] = 2.5 µM (Figure 2 in Paper III). This is ~25-fold higher than when the two unnatural AAs are charged to tRNA$^{Phe}$ (4). Therefore, the ex-
ogenous tRNA$_{Pyl}^{B}$ has much weaker *E. coli* EF-Tu binding affinity than *E. coli* tRNA$_{Phe}^{B}$.

**Inefficient ribosomal binding of the ternary complex**

The inefficient ternary complex formation could not explain the sluggish fast phase rate of dipeptide formation with AA-tRNA$_{Pyl}^{B}$. We then simultaneously measured the rates of GTP hydrolysis on EF-Tu and dipeptide formation at two ribosome concentrations (ribosome in excess over ternary complex in the reactions) at 20 °C with the two unnatural AA-tRNA$_{Pyl}^{B}$ substrates. The rates for GTP hydrolysis on EF-Tu were about the same as the rates for dipeptide formation, indicating the steps subsequent to GTP hydrolysis up to peptide bond formation were not inhibited by the tRNA$_{Pyl}^{B}$ body (Figure 2). The $k_{cat}/K_M$ values for GTP hydrolysis were ~1 µM$^{-1}$s$^{-1}$, reflecting surprisingly inefficient binding of the ternary complexes to the ribosomes (the same value for Phe-tRNA$_{Phe}^{B}$ was ~13 µM$^{-1}$s$^{-1}$ (91)). This is unprecedented for cognate reactions. Recent bulk and single molecule studies revealed that a disrupted codon-anticodon interaction by chemical modification of the codon on the mRNA could lead to inefficient tRNA selection and translation elongation (109). It is highly possible that the codon-anticodon interaction when tRNA$_{Pyl}$ reads the amber codon is less effective than natural cognate interactions, as the archaean tRNA$_{Pyl}$ has some special features compared with *E. coli* elongator tRNAs (Figure 14A).

**Competing reaction from the non-cognate RF2**

We noticed that, at intracellular concentrations, the rate for the non-cognate RF2 misreading UAG codon should be ~0.1 s$^{-1}$ (110). This is comparable with the rate of the dominant slow phase in the dipeptide formation reaction with AA-tRNA$_{Pyl}^{B}$ at 37 °C. We therefore added RF2 to the dipeptide formation reaction with aG-tRNA$_{Pyl}^{B}$ to test whether RF2-catalyzed termination at UAG codon could be a competing reaction. As shown in Figure 15, the more RF2 added, the more fMet was released from the P site tRNA$_{fMet}^{B}$ and the less fMet-aG dipeptide was formed. Thus, RF2 indeed could efficiently compete with AA-tRNA$_{Pyl}^{B}$ in reading the UAG codon, leading to truncation products.
Figure 15. Competition at the UAG codon between tRNA\textsuperscript{PyIIB}\textsuperscript{-}mediated suppression and RF2 termination at 37°C. To a limited amount of initiation complexes, excess of aG-tRNA\textsuperscript{PyIIB}\textsuperscript{-}EF-Tu\textsuperscript{-}GTP ternary complexes were added, along with 0-12 µM RF2ala (a mutant version of RF2 which allows overexpression, see Paper III). fMet-aG dipeptide formed from fMet-tRNA\textsubscript{fMet} and aG-tRNA\textsubscript{PyIIB} is shown in green, and the fMet released by RF2ala is shown in blue. Mean values and standard deviations were calculated from three replicates.

Impaired elongation processivity of tRNA\textsuperscript{PyIIB}

The peptidyl-tRNA\textsuperscript{PyIIB} needs to be translocated to the P site for transfer of the peptide to the next A site substrate to extend the nascent polypeptide (Figure 2). In order to study the processivity of this elongation step, we measured the kinetics of the fMet-aG (or mS)-Phe tripeptide formation, where aG (or mS) was delivered by tRNA\textsuperscript{PyIIB} to the AUG codon and Phe was delivered by \textit{E. coli} tRNA\textsuperscript{Phe} to the cognate UUC codon. Surprisingly, nearly 40% of fMet-aG or fMet-mS dipeptides failed to be extended to tripeptides (Figure 16A). The loss of processivity could not be overcome by lowering the reaction temperature from 37 °C to 20 °C (Figure 16B), indicating such defect was not due to inefficient ternary complex formation. Presumably, the special features (Figure 14A) of the tRNA\textsubscript{PyI} compared with \textit{E. coli} elongator tRNAs would affect its conformation on the ribosome and lead to impaired elongation processivity, as have been observed for D-AA-tRNA (111) and AA-tRNAs with 3′-terminal deoxy A (112). However, such conformational difference might depend on the sequence context and/or buffer conditions. Therefore, further evidence is needed to demonstrate the relevance of this observation to the \textit{in vivo} situation.
Figure 16. Tripeptide formation reactions. Open triangles and diamonds represent the synthesis of fMet-aG-Phe (fMaGF) and fMet-mS-Phe (fMmSF) tripeptides respectively, where aG and mS were delivered by tRNA$^{PylB}$ and Phe by native E. coli tRNA$^{Phe}$ at 37°C (A) and 20°C (B). Filled symbols show incorporation efficiencies of aG (filled triangles) and mS (filled diamonds) into both dipeptide and tripeptide in the same reaction (calculated by the disappearance fraction of fMet). Representative plots are shown.
Conclusions and future outlook

We applied *in vitro* kinetics approaches to protein synthesis with unnatural AAs. In Paper I, we found that N-methyl AAs could be incorporated faster when delivered by tRNA\textsuperscript{Pro}, but the mechanism remained elusive. It is very likely that this is the result of natural evolution, since the tRNA\textsuperscript{Pro} sequence is selected for delivering the only natural N-alkylated AA, Pro (95). We also showed the rate of dipeptide formation with N-methyl AAs could be increased significantly by increasing pH from 6.5 to 8.5. The striking pH-dependence of the rate supports a rate-limiting peptidyl transfer reaction during single incorporation of N-methyl AAs (17). We further demonstrated that several futile cycles of delivery and rejection of the N-methyl AA-tRNA were required per peptide bond formation, and increased Mg\textsuperscript{2+} concentration could reduce the number of the futile cycles. However, the translation steps after the peptide bond formation with N-methyl AAs need to be studied in the future. Interesting questions include: By what mechanism does tRNA\textsuperscript{Pro} hasten the peptide formation with N-methyl AAs? What is the effect of the C-terminal N-methyl AA on the elongation of the peptidyl-tRNA? How efficient is consecutive incorporation of N-methyl AAs? Can tRNA\textsuperscript{Pro}, higher pH and higher Mg\textsuperscript{2+} concentration improve the efficiency of ribosomal polymerization with N-methyl AAs? Can EF-P improve the kinetics of consecutive incorporations of N-methyl AAs as it does for Pro (33)? Is the tRNA\textsuperscript{Pro} body required for the function of EF-P or is the N-alkylated AA moiety the only necessity for the function of EF-P?

In Paper II, we studied the kinetics of consecutive incorporations of unnatural AAs into polymers. We demonstrated in our kinetic translation system that the penultimate dC introduced from the standard N-NVOC-AA-pdCpA chemoenzymatic ligation method to the AA-tRNA caused kinetic defects, in terms of rates and processivities, during the five consecutive incorporations. The dC prompted fast drop-off of the peptidyl-tRNA along the polymerization, and this was the reason why a prolonged incubation time could not overcome the defects in an earlier batch translation study (107). Circumventing the penultimate dC by using our N-NVOC-AA-pCpA chemoenzymatic ligation method (73) or the flexizyme charging method (108), we showed efficient five consecutive incorporations of unnatural AAs. We therefore suggest using these two methods to prepare substrates for ribosomal synthesis of randomized peptidomimetics libraries. Importantly, we showed for the
first time that small non-\(N\)-alkyl AAs, like aG and mS, would not hinder the translocation step during protein synthesis, as their incorporations displayed near wild-type kinetics.

Here, ways to improve ribosomal incorporation of unnatural AAs in vitro, suggested by the data from our kinetics studies in Papers I and II and from earlier published kinetics results, are summarized in Table 1. Future work can be done to explore the mechanisms by which the natural translation machinery discriminates other types of unnatural AAs, like D-AAs and \(\beta\)-AAs. Such mechanistic understanding should guide the engineering of the translation system to facilitate the application of the unnatural AAs.

### Table 1. Ways to improve ribosomal incorporation of unnatural AAs in vitro

<table>
<thead>
<tr>
<th>Types of unnatural AAs</th>
<th>Ref. 4, 98</th>
<th>Single Incorporation</th>
<th>Consecutive Incorporations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Improving AA-tRNA delivery</td>
<td>Reducing ribosomal rejection</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increased [EF-Tu]</td>
<td>Lower temperature</td>
</tr>
<tr>
<td>small L-AAs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>bulky L-AAs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>N-methyl AAs</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

In paper III, we investigated all the steps subsequent to tRNA charging required for incorporation of unnatural AAs by tRNA\(^{\text{Pyl}}\)-mediated amber suppression. Dipeptide formation with AA-tRNA\(^{\text{PylB}}\) displayed biphasic kinetics, with the fast phase rate \(\sim\)30-fold slower than those of native substrates (98). The fast phase fraction increased with increasing EF-Tu concentration, allowing the demonstration of the \(\sim\)25-fold weaker \(E.\ coli\) EF-Tu binding affinity of the heterologous tRNA\(^{\text{PylB}}\) body than that of the \(E.\ coli\) tRNA\(^{\text{Phe}}\). The weak EF-Tu binding might be overcome by increasing the intracellular concentration of EF-Tu, lowering temperature, and/or altering the T-stem sequence of the tRNA\(^{\text{Pyl}}\) body. It has been shown that three base pairs in the T-stem are the key determinant for EF-Tu binding of the tRNA (113, 114). Therefore, altering the T-stem sequence to a stronger EF-Tu binding sequence might increase the EF-Tu binding affinity of tRNA\(^{\text{Pyl}}\). Indeed, a recent directed evolution work selected a tRNA\(^{\text{Pyl}}\) variant with mutations in the T-stem sequence as well as the adjacent base pair in the acceptor stem (115). The evolved tRNA\(^{\text{Pyl}}\) improved the efficiency of double incorporations of \(N^0\)-acetyllysine into a protein by \(\sim\)5-fold, presumably by increasing the EF-Tu binding affinity of the tRNA as speculated by the authors.

Furthermore, our kinetics studies revealed that the incorporation of unnatural AAs delivered by tRNA\(^{\text{PylB}}\) was rate-limited by the \(\sim\)10-fold less efficient binding of the AA-tRNA\(^{\text{PylB}}\):EF-Tu:GTP ternary complexes to the ribosomes. This suggests even if the AA-tRNA\(^{\text{Pyl}}\) is engineered to have wild-
type binding affinity to E. coli EF-Tu, the incorporation rate would still be limited by the inefficient binding of the ternary complexes to the ribosomes. Also, the slow incorporation of unnatural AAs via tRNA\textsuperscript{PylB}-mediated amber suppression was significantly competed by the non-cognate RF2-catalyzed termination at the amber codon. We propose to address these issues by introducing mutations to the tRNA\textsuperscript{Pyl} sequence at the region involved in the distortion of the tRNA body after being delivered to the ribosomal A site by EF-Tu. Structural studies revealed a distorted, or bended, conformation of the AA-tRNA in the A/T state on the ribosome to allow interaction with both the codon on the mRNA and EF-Tu (13). The distorted regions of the tRNA were mainly in the D-stem and anticodon arm. There is evidence showing that changes in these regions can tune the decoding efficiency of the tRNA, presumably by altering its bendability (116, 117). Hence, it is very plausible that mutating these regions in the tRNA\textsuperscript{Pyl} could lead to improved binding efficiency of the ternary complex to the ribosome. However, any changes in the tRNA\textsuperscript{Pyl} sequence might lead to defects in aminoacylation efficiency of the PylRS and/or the fidelity of translation. Therefore, co-evolution of the tRNA\textsuperscript{Pyl} and PylRS is suggested for improved amber suppression efficiency \textit{in vivo}.

We also found that ~40\% of the dipeptidyl-tRNA\textsuperscript{PylB} failed to be elongated to produce tripeptide in our kinetics assay. However, it is uncertain if such scenario exists \textit{in vivo}, as the elongation processivity may be context dependent. If this was relevant \textit{in vivo}, we would predict that a portion of the truncation protein products was due to the inefficient unnatural AA incorporation and the RF2-mediated termination at UAG codon, while another portion was due to the peptidyl-tRNA\textsuperscript{Pyl} failing to be elongated. Analyzing the composition of the \textit{in vivo} expressed proteins might clarify this uncertainty. This issue may also be addressed by directed evolution of the orthogonal pair.

All in all, our results further elucidated the mechanisms of ribosomal incorporation of unnatural AAs, clarified the inefficiencies and suggested ways to improve the incorporations.
Ribosomkatalyserad inkorporering av onaturliga aminosyror (AS) i peptider eller proteiner har ett brett användningsområde inom studier av translationsmekanismen, upptäckandet av potentiella läkemedel samt för undersökning av proteiner struktur och funktion. Dessa applikationer begränsas dock generellt av den låga inkorporeringseffektiviteten för onaturliga aminosyror på ribosomen. Här undersöker vi, med hjälp av kinetiska metoder in vitro, proteinsyntes med onaturliga substrat för att undersöka de mekanismer som begränsar inkorporeringen av onaturliga aminosyror.

I Artikel I demonstrerade vi först att en syntetisk tRNA kropp, baserad på den naturliga bäraren av N-alkylerade aminosyror, tRNA\textsuperscript{Pro}, i motsats till tRNA-kroppar baserade på tRNA\textsuperscript{Phe} och tRNA\textsuperscript{Ala}, kunde öka den ribosomala inkorporeringshastigheten av tre i studien utvalda N-metyl-aminosyror. Därefter visade vi att N-metyl-aminosyronas inkorporeringshastighet var starkt pH-beroende i pH-intervall 7 till 8,5, vilket stödjer teorin att det hastighetsbegränsande steget är peptidyltransfersteget som kräver deprotonering av N-metylaminosyrans N-nukleofil. Tävlingsexperiment visade att det krävs flera misslyckade cykler av association och dissociation av N-metyl-AS-tRNA till och från ribosomens A-site per bildad peptidbindning och att en ökad Mg\textsuperscript{2+}-koncentration ökar inkorporeringsutbytet. Således kan bildningen av peptidbindningar med N-metyl-AS påskyndas genom att använda tRNA\textsuperscript{Pro}, högre pH och högre Mg\textsuperscript{2+}-koncentrationer.

I Artikel III studerade vi kinetiken för ribosomal inkorporering av onaturliga aminosyror via tRNA$^{\text{Pyl}}$-beroende “amber suppression” (stop-kodons undertryck), vilket är det mest populära verktyget för att genetiskt koda för onaturliga aminosyror in vivo. Vad som var känt var att de konstruerade PylRS varianterna hade ca 1000 gånger lägre tRNA-aminoacyleringsaktivitet. Ingenting var känt rörande inkorporeringsstegen efter laddningen av tRNA:t.

Vi fann att dipeptidbildning med initiatör fMet-tRNA$^{\text{fMet}}$ och tRNA$^{\text{Pyl}}$, laddat med en liten onaturlig L-aminosyra (genom kemoenzymatisk ligering med $N$-NVOC-AS-pCpA), var överraskande trög ca. 30 gånger långsammare än med vildtyp-substrat. Bindningsaffiniteten för tRNA$^{\text{Pyl}}$ till E. coli EF-Tu var ca 25 gånger lägre än för E. coli tRNA$^{\text{Phe}}$, men affiniteten kunde ökas 30-faldigt genom att sänka reaktionstemperaturen. Vi demonstrerade också att hastigheterna i den snabba fasen var begränsade av den överraskande låga, 10 gånger sämre, effektiviteten i bindningen av EF-Tu:GTP:AS-tRNA$^{\text{Pyl}}$-ternärkomplexet till ribosomen, och den konkurrantera termineringsreaktionen, orsakad av felläsning av UAG-kodonet av icke-kognat RF2, blev väldigt signifikant. Vidare försämrade tRNA$^{\text{Pyl}}$-kroppen oväntat elongerings-processiviteten i translationen, då ca. 40% av dipeptidyl-tRNA$^{\text{Pyl}}$ inte kunde elongeras till en tripeptid. Samevolution av tRNA$^{\text{Pyl}}$ och PylRS föreslås därför för att åtgärda alla de defekter som identifierades i vårt arbete.
Acknowledgements

Just a week before submitting this thesis for printing, we sadly learnt that Prof. Knud H. Nierhaus, who kindly agreed to be my dissertation opponent, passed away very suddenly and unexpectedly. A good heart has stopped beating, a good soul ascended to heaven. We have lost a real giant in the field. May God give him eternal rest and the family the strength to bear the great pain.

Firstly, I would like to thank my supervisor, Dr. Anthony C. Forster, for his unwavering support during my time at the department, on work and on personal life. I’m grateful for the numerous opportunities he afforded me, through the projects in the lab, conferences and introductions. Thanks for treating me with the first turkey in my life.

I thank Dr. Måns Ehrenberg for all the help, collaboration and support. Thanks for teaching me the kinetics and being so helpful with my projects. Our meetings and discussions always let me feel encouraged and excited about my work.

I’m also grateful to my co-supervisor, Dr. Suparna Sanyal, who was also my first teacher at Uppsala University. Thanks for teaching me so well and thanks for taking care of me all the time in the lab.

Dr. Marek Kwiatkowski, you made my life much easier. Thanks for all the chemical syntheses you did for my projects. Scientific discussions with you were always inspiring. I learnt a lot from you. And I enjoyed your stories about mountains, lakes and mushrooms.

Dr. Michael Pavlov, thanks for always being there for all the help whenever I asked for. It’s my great pleasure to collaborate with you and learn from you.

I would like to thank the current and passed members from the D9:2 corridor, for fostering a nice atmosphere for doing science: Ka-weng, my bro, who not only taught me everything in the lab, but together with Kongjing and Max shared much of the happy time during these years. Anneli, for always being so kind and patient to me. Thanks for teaching me the “Borg
“experiment” and your help with my other projects. You are a good friend! Doreen, for spending the weekend time in the lab for teaching me how to use the Äkta and how to purify ribosomes. Magnus, for setting a good example for me to learn from and for your help with my fellowship application. Tyson, for taking care of the Forster group. Thank your for all the interesting discussions and your help with my fellowship application. Sweet memories were shared together with Julie and Toby. Hope you guys are doing well now! Sofia, you are like my sister. Thanks for your continuously encouraging me and helping me. It’s a lot of fun to discuss about zodiac and baking with you. Josefine, for shouldering of responsibilities for teaching while I was on vacation and for great help with my Swedish summary in this thesis. Ray, for providing me all kinds of components. Liang, for helping me with protein purifications, gels and many other things inside and outside the lab. Mikael, for teaching me the stopped-flow and for a lot of interesting discussions. I appreciate your hard work for maintaining the HPLCs working. And I’m impressed by how much you know about the field. Wish you success! Chandu, for all your help. I still remember this from my first lab course in Uppsala: “My name is not Sir, my name is Chandu”. Ravi, for interesting discussions and a nice memory at the Cold Spring Harbor. Petar, for sharing the responsibility for arranging the lab cleanings and for keeping smiling to me all the time. Harriet, for being so kind and helpful to me when I needed. Yanhong, for your help and for interesting discussions. Gabriele, for sharing your knowledge about the eukaryotic translation system. Gürkan, it was so nice to meet you on the same train after my escaping death. Changil and Hakmin, for interesting discussions about the traditions in your country. Sandra, Felix, Jan, Ram, Kristin, Sunanda, Onur, Rasa, Atia and Matteo for interesting discussions and chats.

I’m also grateful to colleagues at the Department of Cell and Molecular Biology who helped and supported me. Dr. Leif Kirsebom, for being my teaching mentor and for being so supportive. Dr. Staffan Svärd, a very nice program coordinator for my master study at Uppsala and a very helpful prefect during my PhD study. Drs. Lars Liljas and Maria Selmer, for being nice teachers and being helpful. Drs. Samuel Flores, Johan Åqvist, Sebastian Deindl and Kalle Kipper for discussions. Dr. Hugo Gutierrez, for providing a backup computer when needed. Ms. Terese Bergfors, for introduction to the lab safety and for help with the lab equipment.

I thank all my classmates and friends in Uppsala who enriched my life during my time here. Mao, for being the best and reliable friend. Lu, Xiaofen, Ting, Jingjing and Yang, for all the laughs and tears and sweet memories. Xin and Zheyi, for providing accommodation when I had nowhere to stay. Ji, I’m glad you found your Mr. Right and wish you a happy life. Star, you
are awesome and special. Chuan and Di, Rui and Henrik, Tong and Xin, great time at the dinner tables. Jia, good luck with your life in the USA.

I was fortunate to learn from and work with Dr. Ching Song during my time in Beijing. Thanks for always keeping the door of Room 111, Lihua Building open. It was a satisfying and memorable experience to work together with so many awesome friends in a supportive, pleasant atmosphere.

My special gratitude goes to Alf (who is living in the heaven now) and Gunilla. Thanks for taking care of me as a family member. I really appreciate it. You kept me warm and happy over the years. Helena and Ulla, please take care of yourselves and take care of Gunilla. Maria and Ronny, good luck with the babies. Anthony, my bro, thanks for challenging me with Wing Chun and climbing, and wish you a happy life. Jenny, I respect you a lot for your environmentalism, and you should be proud of the little Alma.

I would like to thank all my good friends from my hometown. As is said, “A new friend is like new wine; when it has aged you will drink it with pleasure.” I feel most relaxed while being together with you guys. Thanks for being friends. You are amazing. See you at the same old place!

Finally, I thank all my relatives and family for their dedication and support during my studies. Yujie, thank you for your love and support, and thank you for being You. My dear parents, grandparents, sister, brother in law and the little niece, you are my source of continuous strength and love. I appreciate your unconditionally support and thank you for fueling my ambitions in life.
References


Acta Universitatis Upsaliensis

*Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1369*

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology*. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)

Distribution: publications.uu.se

urn:nbn:se:uu:diva-282023