Rate and Accuracy of Bacterial Protein Synthesis with Natural and Unnatural Amino Acids

KA-WENG IEONG
Abstract


This thesis addresses different questions regarding the rate, efficiency, and accuracy of peptide bond formation with natural as well as unnatural amino acids: Which step is rate-limiting during peptide bond formation? How does the accuracy vary with different transfer RNAs (tRNAs) and codons and how is it relevant to the living cells? Does proofreading selection of codon reading occur in a single- or multi-step manner as theoretically suggested? How does the E. coli translation system discriminate unnatural amino acids? Based on that, how to improve the incorporation efficiencies of unnatural amino acids?

Based on the study on pH dependence of peptide bond formation, we show that the rate of the chemistry of peptidyl transfer to aminoacyl-tRNA (AA-tRNA) Gly-tRNA\textsubscript{Gly} or Pro-tRNA\textsubscript{Pro} limits the rate of peptide bond formation at physiological pH 7.5, and this could possibly be true for peptidyl transfer to all natural AA-tRNAs at physiological condition.

By studying the efficiency-accuracy trade-off for codon reading by seven AA-tRNA containing ternary complexes, we observe a large variation on the accuracy of initial codon selection and identify several error hot-spots. The maximal accuracy varied 400-fold from 200 to 84000 depending on the tRNA identity, the type and position of the mismatches.

We also propose a proofreading mechanism that contains two irreversible steps in sequence. This could be highly relevant to the living cells in relation to maintaining both high accuracy and high efficiency in protein synthesis.

Finally, we show that peptide bond formation with small and large non-N-alkylated L-unnatural amino acids proceed at rates similar to those with natural amino acids Phe and Ala on the ribosome. Interestingly, the large side chain of the bulky unnatural amino acid only weakens its binding for elongation factor Tu (EF-Tu) but not slows down peptidyl transfer on the ribosome. Our results also suggest that the efficiency of unnatural amino acid incorporation could be improved in general by increasing EF-Tu concentration, lowering the reaction temperature and / or using tRNA bodies with optimal affinities for EF-Tu in the translation system.

Keywords: Ribosome, protein synthesis, translation, efficiency-accuracy trade-off, kinetics, unnatural amino acids

© Ka-Weng Ieong 2014

ISSN 1651-6214
ISBN 978-91-554-9103-1
urn:nbn:se:uu:diva-235534 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-235534)
To my family
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.
## Contents

Introduction.........................................................................................................................9

Methods of kinetic studies .................................................................................................11
   Experimental setup ..................................................................................................11
   Measurement of GTP hydrolysis..............................................................................12
   Measurement of dipeptide formation ..................................................................13

Rate of ribosomal peptidyl transfer ..............................................................................16
   Rate-limiting step? ..............................................................................................16
   Measurement of tRNA accommodation ............................................................17
   Measurement of the chemistry of peptidyl transfer ...........................................18
   pH dependence of peptidyl transfer (Paper I) ..................................................20

Accuracy of tRNA selection in translation .......................................................................23
   In vivo numbers ....................................................................................................23
   Mechanism of decoding .........................................................................................23
   Kinetics of tRNA selection ..................................................................................24
   Trade-off between efficiency and accuracy .........................................................27
   Large variation in accuracy of initial codon selection (Paper II) .......................28
   Two proofreading steps of tRNA selection in translation (Paper III) ...............31

Translation with unnatural amino acids .......................................................................36
   What we know and what we don’t know ............................................................36
   Aminoacylation of tRNAs with unnatural amino acids ........................................37
   Rate and efficiency (Paper IV) ............................................................................38
   The ‘mysterious’ slow kinetics ..............................................................................41
   Choice of tRNA body (Paper V) ...........................................................................43
   Discrimination between natural and unnatural amino acids ............................44

Sammanfattning på svenska ............................................................................................46

Acknowledgements ..........................................................................................................48

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

70S  The complete bacterial ribosome
50S  Large ribosomal subunit
30S  Small ribosomal subunit
IC   Initiation complex
PTC  Peptidyl transfer center
EF   Elongation factor
tRNA Transfer RNA
T3   Ternary complex
AA   Amino acid
Phe  L-Phenylalanine
Ala  L-Alanine
aG   L-allyl-glycine
mS   L-methyl-serine
bK   L-biotinyl-lysine
NMF  N-methyl-L-phenylalanine
I    Initial codon selection
F    Proofreading selection
A    Overall accuracy
A_{nf} Accuracy of AA-tRNA selection in absence of EF-Tu
Introduction

A correct set of proteins is required to maintain the health and function of all living cells. The diversity of proteins is primarily due to their sequence of amino acids, which is determined by the DNA sequence of their genes. The DNA information is transcribed into messenger RNA (mRNA) through a process called transcription. The sequence of the mRNA is translated into the sequence of amino acids that makes up the proteins. This process, called translation, is one of the most fundamental and complicated biological processes in all living cells. Accurate and efficient protein synthesis is required for all organisms to maintain their maximal fitness of their growth.

Ribosome is the translation machinery found in all organisms, it is the most complicated RNA-protein complex in the cell. The bacterial 70S ribosome is made up of two subunits: the large 50S subunit and small 30S subunit. The 50S subunit, where peptidyl transfer occurs, is composed of 23S and 5S ribosomal RNA (rRNA) and along with 33 proteins. The small 30S subunit, where the mRNA is decoded, is composed of 16S rRNA and 21 proteins (Cate et al. 1999; Yusupov et al. 2001).

The process of protein synthesis (translation) is divided into four major stages (reviewed in Schmeing and Ramakrishnan 2009): initiation, elongation, termination and ribosome recycling. During initiation, 16S rRNA of the 30S subunit recognizes the Shine-Dalgarno (SD) sequence in the mRNA and initiator fMet-tRNA^{fMet} is bound to the P (peptidyl) site of the ribosome which displays the start codon of AUG. A functional 70S ribosome is then assembled with the help of initiation factors (IF1, IF2, and IF3) at the correct place on the mRNA. During elongation phase, aminoacyl-tRNA (AA-tRNA) in the form of ternary complex with elongation factor Tu (EF-Tu) and GTP binds to the A site of the 70S initiation complex. After the decoding of AA-tRNA, GTP is hydrolyzed on EF-Tu, EF-Tu:GDP then leaves the ribosome and a peptide bond is formed. This is followed by translocation with the help of elongation factor G (EF-G). The peptidyl-tRNA is then moved from the A site to the P site and next codon to be read is displayed in the empty A site. When the A site displays a stop codon (UAG, UAA, or UGA), only release factor 1 or 2 (RF1 or RF2) can bind and the polypeptide is released from the P site peptidyl-tRNA, leaving the deacylated tRNA in the P site. After that, RF1 or RF2 dissociates from the mRNA bound 70S ribosome with help of release factor 3 (RF3). Finally, the 70S ribosome is split by the ribosome...
recycling factor (RRF) and EF-G before a new round of protein synthesis starts.

The study in this thesis is focused on the elongation phase of protein synthesis, in particular from decoding of AA-tRNA all the way to peptidyl transfer. Mechanistic studies on the rate, efficiency and accuracy of protein synthesis with natural as well as unnatural amino acids will be discussed.
Methods of kinetic studies

Experimental setup

We use the purified *in vitro* translation system developed in our laboratory (Pavlov et al. 1997) to study kinetic properties of various reactions of protein synthesis. 70S ribosomes are purified from *E. coli* MRE600 by sucrose gradient centrifugation. For ease of purification, many of the protein components (initiation factors, EF-Ts, AA-tRNA synthetases) are recombinants with polyhistidine-tags, overexpressed in *E. coli* and purified on nickel columns. However, in order to compare our *in vitro* measurements to *in vivo* data, wild type EF-Tu from *E. coli* MRE600 is used in all my studies since EF-Tu is involved in many steps during protein elongation. Native *E. coli* tRNAs are either purified from *E. coli* MRE600 cells in our laboratory or purchased from companies. mRNAs and unmodified tRNAs are prepared by *in vitro* transcription by T7 polymerase from dsDNA templates that are made by PCR from designed oligoes. Unnatural AA-tRNAs were prepared by ligation of N-nitroveratryloxycarbonyl-AA-pdCpA (*N*-NVOC-AA-pdCpA, synthesized in our laboratory) and 3’-CA-truncated tRNA transcripts by T4 RNA ligase, followed by photolysis of the NVOC protecting group.

Experiments were conducted in polymix buffer (Jelenc and Kurland 1979) for Paper I to III and polymix-like buffer (Pavlov et al. 2008), in which HEPES is used instead of phosphate, for Paper IV and V. Polymix buffer is a near-physiological buffer that contains the polyamines spermidine and putrescine. It is optimized for high rate and fidelity of translation. Also ATP, GTP, phosphoenolpyruvate (PEP), pyruvate kinase and myokinase are always present for energy supply and regeneration.

Translation experiments start by mixing equal volumes of two mixtures, one containing 70S initiation complexes and the other ternary complexes, followed by quenching with either formic acid or KOH (only for dipeptide formation measurements) to stop the reactions at specific incubation times. Fast reactions (average time smaller than 40s) were carried out in a quench-flow instrument (time resolution 2 ms) where independent reactions from two mixtures were quenched at specific time points that we programmed. Slow reactions were performed manually where aliquots from a master reaction mixture were taken out and quenched at different time points. In order to detect the products, [*H]* labeled fMet and GTP were used in the reaction, and the samples were analyzed by HPLC equipped with an on-line radio-
detector, either on a reverse phase C18 column (for separation of unreacted $[^3\text{H}]\text{Met}$ and $[^3\text{H}]\text{dipeptide}$ formed) or an ion exchange monoQ column (for $[^3\text{H}]\text{GDP/GTP}$ separation).

Measurement of GTP hydrolysis

In order to increase the precision of the measurements, in Papers I, II, III, and V, we designed experiments so that a large fraction of hydrolyzed GTP was obtained. AA-tRNAs were in excess over EF-Tu in ternary complex mixture, so that the concentration of EF-Tu limits the concentration of ternary complex and the maximal fraction of GTP hydrolyzed; ribosomes were in excess over ternary complexes, so that all GTP in ternary complex were hydrolyzed when the reaction was finished. The GTP hydrolysis reaction can be described by Scheme 1:

$$R + T3 \xrightleftharpoons[k_d]{k_a} R \cdot T3 \xrightarrow{k_c} R \cdot T3[^3\text{H}]\text{GDP}.$$  

Here R and T3 represent ribosomes and ternary complexes, respectively; $k_a$ and $k_d$ are the association and dissociation rate constants for binding of ternary complexes to ribosomes, respectively; $k_c$ is the first order rate constant for GTP hydrolysis on EF-Tu. The kinetic efficiency $k_{cat}/K_m$ for GTP hydrolysis is defined as the association rate constant $k_a$ multiplied by the probability that GTP is hydrolyzed on EF-Tu,

$$(k_{cat}/K_m)_{GTP} = k_a \cdot \frac{k_c}{k_c + k_d}.$$  

Experimentally, we used $[^3\text{H}]\text{GTP}$ to form ternary complexes and the fraction of hydrolyzed GTP was analyzed by HPLC. We monitored the time evolution of hydrolyzed GTP from which total time $\tau_{GTP}$ was estimated.

$$\tau_{GTP} = \frac{1}{k_{GTP}} = \frac{1}{(k_{cat}/K_m)_{GTP} [R]} + \frac{1}{k_c},$$  

$\tau_{GTP}$ is the total time for binding of ternary complexes to ribosomes and GTP hydrolysis on EF-Tu. For cognate reactions, the rate constant for GTP hydrolysis on EF-Tu, $k_c$, is very fast (>500 s$^{-1}$; (Pape et al. 1998; Johansson et al. 2008a)). At low ribosome concentration, $\tau_{GTP}$ is dominated by the term
\[(k_{cat}/K_m)[R]^{-1},\] which reflects the time for the binding of ternary complexes to ribosomes and \((k_{cat}/K_m)_{GTP}\) can be estimated from

\[
(k_{cat}/K_m)_{GTP} = k_{GTP}/[R].
\]

In fact, we observed such a linear dependence between \(k_{GTP}\) and ribosome concentration \([R]\) for both the cognate and non-cognate GTP hydrolysis reactions in Papers II and III.

The maximal rate of the GTP hydrolysis reaction \((k_{cat})\) was obtained by extrapolation to saturating ribosome concentration \([R] \to \infty, k_{cat} = k_c\).

### Measurement of dipeptide formation

We measure the time evolution of dipeptide formation starting from 70S initiation complexes containing \(f[3^H]\text{Met-tRNA}^{\text{Met}}\) in the A site and ternary complexes. The reaction can be described by Scheme 5, which is an extended version of Scheme 1:

\[
R + T3 \xrightleftharpoons[k_d]{k_a} R \cdot T3_{\text{GTP}} \xrightarrow{k_e} R \cdot T3_{\text{GDP}} \xrightarrow{k_{\text{tu}}} R \cdot AA - tRNA \xrightarrow{k_{ac/pt}} [3^H]\text{dip}
\]

Here \(k_{Tu}\), \(k_{ac/pt}\), and \(q_d\) are the first order rate constants for dissociation of A site tRNA from EF-Tu:GDP after GTP hydrolysis, A site tRNA accommodation / peptidyl transfer, and rejection of AA-tRNA, respectively.

\(k_{cat}/K_m\) for dipeptide formation is in this case defined as the association rate constant \(k_a\) multiplied by the probability that a dipeptide is formed:

\[
(k_{cat}/K_m)_{\text{dip}} = k_a \cdot \frac{k_c}{k_{d} + k_{c}} \cdot \frac{k_{ac/pt}}{k_{ac/pt} + q_d} = (k_{cat}/K_m)_{GTP} \cdot \frac{1}{f},
\]

where the proofreading factor \(f\) (Bilgin et al. 1992) is defined as
\[ f = \frac{k_{\text{ac/pt}} + q_d}{k_{\text{ac/pt}}} \quad (7) \]

\( f \) is the number of GTPs hydrolyzed per peptide bond formation. For cognate reactions, \( q_d \to 0 \) (Gromadski et al. 2006) and \( f_c \approx 1 \),

\[ (k_{\text{cat}} / K_m)^c_{\text{dip}} = k_d \cdot \frac{k_c}{k_c + k_d} = (k_{\text{cat}} / K_m)^c_{\text{GTP}} \quad (8) \]

Therefore, for cognate reactions in Paper III, \( (k_{\text{cat}}/K_m)_{\text{dip}} \) is approximated by \( (k_{\text{cat}}/K_m)_{\text{GTP}} \).

In dipeptide formation experiments, we measure the time evolution of \([\text{^3H}]\text{dipeptide formation}\) and estimate the total time for dipeptide formation, \( \tau_{\text{dip}} \), which includes the total time for GTP hydrolysis (as discussed above) and also the average times for the subsequent steps in Scheme 5. For measurements in Papers I, IV and V, ribosomes were in excess over ternary complexes, so that the rate \( k_{\text{dip}} \) was limited by ribosome concentration \([R]\),

\[ \tau_{\text{dip}} = \frac{1}{k_{\text{dip}}} = \frac{1}{(k_{\text{cat}} / K_m)_{\text{GTP}} \cdot [R]} + \frac{1}{k_{\text{Tu}}} + \frac{1}{f \cdot k_{\text{ac/pt}}} \quad (9) \]

For cognate reactions, \( f \approx 1 \),

\[ \tau_{\text{dip}} = \frac{1}{k_{\text{dip}}} = \frac{1}{(k_{\text{cat}} / K_m)_{\text{GTP}} \cdot [R]} + \frac{1}{k_{\text{Tu}}} + \frac{1}{k_{\text{ac/pt}}} \quad (10) \]

Measuring dipeptide formation for cognate reactions at saturating ribosome concentration \([R] \to \infty\) will give us the minimal time of \( \tau_{\text{dip}} \), \( \tau_{\text{dip}}^\text{min} \), which is the total time for all subsequent steps after binding of ternary complexes to ribosomes.

By measuring GTP hydrolysis and dipeptide formation in the very same experiment as done in Papers I and V, we can estimate both \( \tau_{\text{dip}} \) and \( \tau_{\text{GTP}} \) simultaneously. Subtraction of \( \tau_{\text{GTP}} \) (Eq. 3) from \( \tau_{\text{dip}} \) (Eq. 10) gives the total time \( \tau_{\text{pep}} \) \( (\tau_{\text{acc,pep}}) \) for all subsequent steps after GTP hydrolysis which includes the time for release of EF-Tu:GDP \( (1/k_{\text{Tu}}) \) and accommodation / peptidyl transfer \( (1/k_{\text{ac/pt}}) \).
\[ \tau_{\text{pep}} = \tau_{\text{dip}} - \tau_{GTP} = \frac{1}{k_{Tn}} + \frac{1}{k_{\text{ac}/\text{pt}}} . \quad (11) \]

\( \tau_{\text{pep}} \) could be approximated by \( \tau_{\text{dip}}^{\text{min}} \) given that the time for GTP hydrolysis (1\( /k_c \)) is negligible. For example, in Paper V, measuring GTP hydrolysis is technically difficult for incorporation of the bulky biotinyllysine, so we measured \( \tau_{\text{dip}}^{\text{min}} \) to approximate \( \tau_{\text{pep}} \).

For non-cognate measurements in Paper III, ternary complexes were in excess over ribosomes, so the rate \( k_{\text{dip}} \) was limited by the concentration of ternary complex \([T3]\). The total time estimated for dipeptide formation, \( \tau_{\text{dip}} \), can be expressed as:

\[ \tau_{\text{dip}} = \frac{1}{k_{\text{dip}}} = f' \cdot \left( \frac{1}{(k_{\text{cat}} / K_m)_{GTP} \cdot [T3]} + \frac{1}{k_c} + \frac{1}{k_{Tn}} \right) + \frac{1}{k_{\text{ac}/\text{pt}}} . \quad (12) \]

To estimate \( k_{\text{cat}}/K_m \) for non-cognate dipeptide formation, \( k_{\text{dip}} \) was measured at varying concentration of ternary complex \([T3]\), the data was fitted to a Michaelis–Menten kinetics model:

\[ k_{\text{dip}} = \frac{[T3] \cdot k_{\text{cat}}}{K_m + [T3]} = \frac{(k_{\text{cat}} / K_m)_{\text{dip}}^{\text{nc}} \cdot [T3]}{1 + [T3] / K_m} . \quad (13) \]
Rate of ribosomal peptidyl transfer

Rate-limiting step?

*In vivo*, the speed of protein elongation varies with growth rate, growth medium, and mRNA codon sequence (e.g. (Sorensen and Pedersen 1991)). During exponential phase of bacterial growth in rich medium at 37 °C, protein elongation of *E. coli* proceeds at average rates around 22 amino acids per second (Bremer and Dennis 1987; Sorensen and Pedersen 1991; Liang et al. 2000). This speed, which may reflect one or more slow steps during the elongation cycle, must have been evolutionally optimized to maintain both high speed and high fidelity of translation.

During the elongation phase of protein synthesis, AA-tRNAs in the form of ternary complex with EF-Tu and GTP have to interact with the 70S initiation complexes and go through several steps before a peptide bond is formed. 1) Binding of ternary complexes to ribosomes, 2) codon-anticodon recognition, 3) hydrolysis of GTP on EF-Tu, 4) release of esterified CCA end from EF-Tu:GDP, 5) movement of the CCA end into the peptidyl transfer center (PTC) of the 50S subunit (A site tRNA accommodation), and 6) peptidyl transfer reaction. Step 1 depends on the ribosome or ternary complex concentration. In the living cell, EF-Tu is abundant in the living cell (Bremer and Dennis 1987), but ternary complex binding could be slow due to the low expression level of some AA-tRNAs (Dong et al. 1996) or due to insufficient amino acid supply. Steps 2 and 3 are reported to be very fast (~100 and 500 s⁻¹, respectively; (Pape et al. 1998)). Step 4 depends on the affinities between EF-Tu and A site AA-tRNA, this step may become rate limiting for incorporation of some mis-acylated AA-tRNAs having very high affinity for EF-Tu (Schrader et al. 2011). In contrast, correctly acylated AA-tRNAs have near-uniform affinities for EF-Tu by thermodynamic compensation (LaRiviere et al. 2001) and their release from EF-Tu:GDP after GTP hydrolysis on EF-Tu is therefore not expected to limit the rate of peptide bond formation (Schrader et al. 2011). tRNA accommodation (Step 5) was reported to be around 10 s⁻¹ at 20 °C (Rodnina et al. 1994; Pape et al. 1998) as well as at 37 °C (Bieling et al. 2006) and it limits the rate for peptide bond formation. However, those measurements are not compatible with *in vivo* data implying that the average elongation rate is around 22 s⁻¹ and any single step during the elongation cycle cannot be slower than that (see further discussion below). The chemistry of peptidyl transfer (Step 6) has only been
measured for incorporation of puromycin analogs and the rates vary with different analogs and the length and identity of P site peptide or amino acid (Wohlgemuth et al. 2008) (see further discussion below). Although direct measurement for the chemistry of peptidyl transfer with natural AA-tRNA has never been carried out, it was proposed to be the rate-limiting step in translation based on the very similar activation energies between the subsequent steps after GTP hydrolysis (until peptidyl transfer) and the peptidyl transfer to puromycin for which accommodation is not required (Johansson et al. 2008a).

Measurement of tRNA accommodation

RNA accommodation was originally measured using fluorescence labeled tRNAs. Ternary complexes of yeast (S. cerevisiae) tRNA\textsubscript{Phe} labeled with the fluorescent dye proflavine in position 16/17 (D-loop) was reacted with 70S ribosomes programmed with poly-(U), and changes of the fluorescence signal were measured in a stopped-flow instrument (Rodnina et al. 1994; Pape et al. 1998). The increase of the fluorescence signal reflects the binding of ternary complexes to the ribosome and all subsequent steps up to GTP hydrolysis and the subsequent decrease of the signal reflects only tRNA accommodation in the A site. Rates of tRNA accommodation measured in a stopped-flow experiment were estimated as 8 to 10 s\(^{-1}\) and were identical to the rates of dipeptide formation as estimated from quench-flow experiment. Therefore, it was concluded that the rate of peptide bond formation is limited by slow accommodation of AA-tRNA in the A site.

However, as mentioned above that the measurements were not compatible with \textit{in vivo} data. The reasons could be 1) the translation system is heterogeneous since tRNA\textsubscript{Phe} was from yeast while other components were from \textit{E. coli}, the tRNA from yeast may not have the same kinetic property as \textit{E. coli} tRNAs; 2) the Tris-based buffer used in those experiments did not optimally display \textit{in vivo}-like properties as obtained in earlier experiments based on poly-(U) translation system in polymix buffer (Jelenc and Kurland 1979) which showed \textit{in vivo} compatible rates (Bilgin et al. 1992).

In a more recent study from the Rodnina laboratory (Wohlgemuth et al. 2010), the authors re-measured tRNA accommodation with proflavin labeled yeast tRNA\textsubscript{Phe} (in stopped-flow) as well as peptide bond formation (in quench-flow). First they found that their measurements with \textit{E. coli} tRNA\textsubscript{Phe} in the new Tris-based \textit{High Fidelity (HiFi)} buffer (Gromadski et al. 2006) were comparable with those in polymix buffer, and the maximal rates (\(k_{cat}\)) of dipeptide formation were around 200 s\(^{-1}\) at 37 \(^{\circ}\)C, which was then compatible with \textit{in vivo} measurements. The rate of tRNA accommodation, measured with yeast tRNA\textsubscript{Phe} was around 40 s\(^{-1}\), which is the same as the rate of peptide bond formation measured in the parallel quench-flow experiment at
the same condition, so tRNA accommodation was concluded again to be the rate-limiting step during dipeptide formation. However, given that the $k_{cat}$ for dipeptide formation is 200 s$^{-1}$, the rate-limiting step cannot be slower than that. One explanation could be that the 40 s$^{-1}$ mainly reflected the binding of ternary complexes to the ribosome rather than tRNA accommodation since the measurements were done at sub-saturating substrate concentration (2 µM ribosome, which was in excess over ternary complex; $K_m$ was around 4 to 5 µM).

Measurement of the chemistry of peptidyl transfer

One way to measure the chemistry of peptidyl transfer reaction is to use substrate analogs that rapidly bind to the A side and do not require accommodation before peptide bond formation. Such experiments were carried out with puromycin and its analogs, which structurally mimic the 3’-end of the AA-tRNA. The rate of peptidyl transfer to puromycin depended on the length and the C-terminal amino acid of the peptidyl moiety of the P site tRNA: the rate at 37 °C varied from around 1 s$^{-1}$ with fMet-tRNA$^{fMet}$, 10 to 20 s$^{-1}$ with fMet-Phe-tRNA$^{Phe}$ and fMet-Phe-Phe-tRNA$^{Phe}$ (Katunin et al. 2002), to 100 s$^{-1}$ with fMet-Lys-tRNA$^{Lys}$ (Wohlgemuth et al. 2008) in the P site. Moreover, the reaction rates at 20 °C were about the same as at 37 °C for reactions with fMet-tRNA$^{fMet}$ or fMet-Phe-tRNA$^{Phe}$ in the P site (Youngman et al. 2004; Brunelle et al. 2006), which contradicts the temperature dependence of peptide bond formation observed with natural Phe-tRNA$^{Phe}$ (Johansson et al. 2008a).

![Figure 1. The proposed mechanism of peptidyl transfer reaction. A nucleophilic attack of the α-NH$_2$ group of the A site AA-tRNA (blue, here Phe-tRNA$^{Phe}$) on the ester carbonyl carbon of the P site peptidyl-tRNA (red).]
The mechanism of peptidyl transfer suggests that there is no acid-base catalysis involved by the ribosomal residues (Schmeing et al. 2005; Trobro and Aqvist 2005; Bieling et al. 2006) and the chemistry requires that the α-amino group of the amino acid on the A site tRNA is in its neutral NH$_2^-$ form rather than in protonated NH$_3^+$ form (Figure 1) (Green and Lorsch 2002). Therefore, peptidyl transfer should be pH sensitive and reach its maximal rate when the α-amino group of the amino acid on the A site tRNA is deprotonated, and the pH dependence of the rate of the reaction should reflect the titration of a single proton. Such experiments were preformed with puromycin (Pmn) analogs as well as natural AA-tRNAs. The rate of dipeptide formation with natural Phe-tRNA$^{\text{Phe}}$ at 37 ºC showed no significant pH dependence (fluctuating rate from 2 to 7 s$^{-1}$) in the pH range from 6 to 8 (Bieling et al. 2006), and the rates of tRNA accommodation measured from FRET experiments with fluorescein labeled fMet-tRNA$^{\text{fMet}}$ and quencher QSY labeled Phe-tRNA$^{\text{Phe}}$ also showed similar rates and pH insensitivity. The authors (Bieling et al. 2006) further confirmed that the measurement of peptide formation was completely masked by the rate-limiting pH independent accommodation step, and that it was impossible to study the chemistry of peptidyl transfer with natural AA-tRNAs. However, the rates for dipeptide formation may only reflect the binding of ternary complexes to ribosomes at low substrate concentration, and that could be the reason why no pH dependence was observed for these reactions by these authors (Bieling et al. 2006).

Measurement of peptidyl transfer to Pmn showed complicated pH dependence with two protons being titrated (Katunin et al. 2002; Brunelle et al. 2006; Beringer and Rodnina 2007b). However, the peptidyl transfer to C-Pmn, a supposedly better substrate mimicking the 3’ CA end of A site AA-tRNA, showed titration of only one proton at 20 ºC (Brunelle et al. 2006), while in another study showed titration of two protons at 37 ºC (Beringer and Rodnina 2007a). Moreover, peptidyl transfer to C-Pmn displayed weaker pH dependence than peptidyl transfer to Pmn (Brunelle et al. 2006), but in another study a similar pH dependence was found for these two substrates (Beringer and Rodnina 2007a). CC-Pmn, an even better substrate mimicking the full 3’ -CCA end of A site AA-tRNA, however displayed an even weaker pH dependence than Pmn and C-Pmn (Beringer and Rodnina 2007a).

These experiments suggested that puromycin analogs may not be suitable analogs for AA-tRNAs in the study of the chemistry of peptidyl transfer and previous studies with natural AA-tRNAs did not show good correlation with in vivo data. Therefore, in Paper I, we revisited the pH dependence of peptidyl transfer to Phe-tRNA$^{\text{Phe}}$ and extended the study to other natural AA-tRNAs in our in vivo compatible translation system.
pH dependence of peptidyl transfer (Paper I)

In a previous study of the pH dependence of peptidyl transfer to Phe-tRNA<sup>Phe</sup> (Bieling et al. 2006), only the total time of dipeptide formation, \( \tau_{\text{dip}} \), was measured (the time includes the binding of ternary complexes to ribosomes up to peptide bond formation). In principle, \( \tau_{\text{dip}} \) could be dominated by the initial binding of ternary complex to the ribosome at low substrate concentration. Therefore, in Paper I, we measured the total time for GTP hydrolysis, \( \tau_{\text{GTP}} \), as well as \( \tau_{\text{dip}} \) in the very same experiments. From the difference between \( \tau_{\text{dip}} \) and \( \tau_{\text{GTP}} \) we calculated the time \( \tau_{\text{pep}} \) for all subsequence steps after GTP hydrolysis up to dipeptide formation.

\[
\tau_{\text{pep}} = \frac{1}{k_{\text{pep}}} = \tau_{\text{dip}} - \tau_{\text{GTP}}
\]

\( \tau_{\text{pep}} \) includes the total time for release of AA-tRNA from EF-Tu:GDP after GTP hydrolysis, A site accommodation and the chemistry of the peptidyl transfer reaction. We measured \( k_{\text{pep}} \), as the inverse of \( \tau_{\text{pep}} \), for peptidyl transfer to six natural AA-tRNAs at different pH values at 20 °C. Our results displayed clear pH dependence in all six cases, Phe-tRNA<sup>Phe</sup>, Ile-tRNA<sup>Ile</sup>, Gly-tRNA<sup>Gly</sup>, Pro-tRNA<sup>Pro</sup>, Ala-tRNA<sup>Ala</sup> and Asn-tRNA<sup>Asn</sup> (Figure 3 in Paper I). These results contradict those of a previous study with Phe-tRNA<sup>Phe</sup> (Bieling et al. 2006). Although our measurements for Phe-tPhe<sup>Phe</sup> displayed clear pH dependence, the difference in \( k_{\text{pep}} \) value between the lowest and highest pH was only 2-fold (Figure 3B in Paper I), which could explain the apparent absence of pH dependence in the previous study (Bieling et al. 2006) as due to imprecise data not accurate enough to observe the small variation of the rate of dipeptide formation with varying pH.

In the cases of Gly-tRNA<sup>Gly</sup> and Pro-tRNA<sup>Pro</sup> there were very strong pH effects with pH change in the range from 6 to 8 units and the \( \log_{10}(k_{\text{pep}}) \) values increased linearly with increasing pH and to reach maximal values at high pH. The slopes for \( \log_{10}(k_{\text{pep}}) \) versus pH (in the linear range) were very close to 1 in both cases (Figure 3H in Paper I), indicating that only one proton in a reaction-essential group was titrated with pH. This, we suggest is titration of the \( \alpha \)-amino group of the amino acid in the A site AA-tRNA. By inference, we suggested that variation of \( k_{\text{pep}} \) with pH for all six cases of AA-tRNA was due to titration of a single proton, so the data were accordingly fitted to the model:
\[ k_{\text{pep}} = \frac{k_{\text{pep}}^{\text{max}}}{1 + 10^{(pK_a^{\text{obs}} - \text{pH})}}. \]

From this we estimated the two parameters, \( k_{\text{pep}}^{\text{max}} \) and \( pK_a^{\text{obs}} \), for each AA-tRNA. \( k_{\text{pep}}^{\text{max}} \) is the maximal value of \( k_{\text{pep}} \) when the \( \alpha \)-amino group of the amino acid in the A site tRNA is completely deprotonated; \( pK_a^{\text{obs}} \), the pH-value at which \( k_{\text{pep}} \) attains half of its maximal value, \( k_{\text{pep}}^{\text{max}} \), corresponds to the observed \( pK_a \) value for the amino acids of the AA-tRNA bound to the ribosome.

The observed \( pK_a \) values (\( pK_a^{\text{obs}} \)) correlated well with the \( pK_a \) values of the free amino acids in bulk water (\( pK_a^{\text{aq}} \)) for all cases of AA-tRNA (Table 1 in Paper I). However, all \( pK_a^{\text{obs}} \) values were lower than the corresponding \( pK_a^{\text{aq}} \) values even though \( pK_a^{\text{aq}} \) values were calibrated according to temperature and their amino acid esters. The downshifts of \( pK_a \) values varied from 0.4 to 1.7 pH units. These downshifts could be due to differences in \( pK_a \) in free solution and on the ribosome or to the influence of a pH independent step, tRNA accommodation, preceding peptidyl transfer.

In order to understand the meaning of the downshifts in \( pK_a \) values observed for AA-tRNAs, MD simulations (by the Åqvist laboratory; see Supporting Information in Paper I), based on crystal structures of the peptidyl transfer center (Schmeing et al. 2005; Selmer et al. 2006) and the Linear Interaction Energy (LIE) method (Carlsson et al. 2008), predicts that the \( pK_a \) values of the \( \alpha \)-amino groups of ribosome-bound AA-tRNAs are downshifted in relation to those of the free AA-tRNAs in an amino acid dependent way. The predicted (or calculated) downshift values in \( pK_a \) (\( \Delta pK_a^{\text{calc}} \)) correlated strongly with our experimentally observed downshift values \( \Delta pK_a^{\text{obs}} \) (Figure 2), suggesting that the chemistry of peptidyl transfer may dominate the time for the steps leading to peptidyl transfer, \( \tau_{\text{pep}} \), subsequent to GTP hydrolysis for all six AA-tRNAs including Phe-tRNAphe.

In the cases of peptidyl transfer to Gly-tRNA\textsuperscript{Gly} and Pro-tRNA\textsuperscript{Pro}, our data show that at physiological pH around 7.5, where \( k_{\text{pep}} \) values have big variations with pH, the chemistry of peptidyl transfer, rather than its preceding steps following GTP hydrolysis on EF-Tu, is rate limiting for peptide bond formation. Our results further suggest that direct study of the chemistry of peptidyl transfer to natural AA-tRNAs is possible in the pH range from 6 to 8 units.
Figure 2. Observed $pK_a$ shifts ($\Delta pK_a^{\text{obs}}$), vs. $pK_a$ shifts from MD simulation ($\Delta pK_a^{\text{calc}}$).
Accuracy of tRNA selection in translation

In vivo numbers

Estimates of in vivo missense error frequencies of bacterial protein synthesis are in the range between $10^{-4}$ to $10^{-3}$ (Parker 1989; Kramer and Farabaugh 2007; Drummond and Wilke 2009). These measurements reflect errors accumulated in several reaction steps including aminoacylation, transcription, and translation. In recent study by Farabaugh and collaborators reported translation error measurements down to $10^{-5}$ and even $10^{-6}$ frequencies (Manickam et al. 2014). This method is based on the identifications of the essential amino acids in the active site of β-galactosidase. By introducing mutations of the β-galactosidase gene at a position encoding one of the functionally essential amino acids, one can estimate the misreading error frequencies of the corresponding cognate tRNA on all possible mismatches, based on the recovered enzyme activity of the mutants. The background levels could vary for different mutants depending on their background β-galactosidase activity. The limitation of the in vivo measurements also depends on the errors of aminoacylation and transcription although average error frequencies have been reported to be extremely small: $10^{-6}$ for both aminoacylation (Soll 1990; Schulman 1991) and transcription (Gout et al. 2013).

Mechanism of decoding

The delivery of AA-tRNAs to the ribosome is a multistep process containing irreversible chemical steps (Pape et al. 1998; Marshall et al. 2008; Voorhees and Ramakrishnan 2013). AA-tRNA in the form of ternary complex with EF-Tu and GTP binds to the ribosome via interaction between EF-Tu and ribosomal protein L7/12. This step is suggested to be independent of codon sequences (Kothe et al. 2004; Diaconu et al. 2005). During codon reading, AA-tRNA in the ternary complex undergoes a large conformational change, which results in a more open and ‘bent’ tRNA structure. The distortion of tRNA allows it to adapt its A/T state conformation and to further scan the codon in the decoding site of the 30S subunit (Blanchard et al. 2004; Geggier et al. 2010). During this time AA-tRNA remains bound EF-Tu in the 50S subunit (Moazed and Noller 1989; Stark et al. 1997). The codon recog-
nition process involves the mRNA bound 30S subunit and the anticodon stem-loop (ASL) of the A site tRNA (Ogle et al. 2001). The universally conserved residues in 16S ribosomal RNA (rRNA), A1492, A1493 (in helix 44), and G530 (in loop 530), undergo conformational changes to interact with the minor groove of the codon-anticodon helix (Ogle et al. 2001). The interactions of these three residues depend on the Watson-Crick geometry of the codon-anticodon base pairs at the first two codon positions but allow wobble base pairs G:U or U:G at the third position (Ogle et al. 2001). A non-cognate tRNA has a high probability to be rejected at this point in the form of ternary complex (initial selection) due to the high free energies of the interactions with the three residues of 16S rRNA (Almlof et al. 2007). These interactions further lead to a large-scale domain closure in the 30S subunit (Ogle et al. 2002), stabilization of the distorted AA-tRNA, GTPase activation of EF-Tu and irreversible GTP hydrolysis (Valle et al. 2003; Schuette et al. 2009; Villa et al. 2009; Voorhees et al. 2010). The hydrolysis of GTP leads to release of Pi and switches EF-Tu from the GTP- to the GDP-bound form by extensive conformational changes (Berchtold et al. 1993) and eventually EF-Tu is released from the ribosome. Kinetic studies suggested that the release of EF-Tu:GDP could occur in parallel with the following events: AA-tRNA dissociates from the ribosome (proofreading) or moves into the peptidyl transferase center (PTC) on the 50S subunit (accommodation) and further participates in peptidyl transfer reaction (Pape et al. 1998; Gromadski and Rodnina 2004; Gromadski et al. 2006).

Kinetics of tRNA selection

Ribosomal discrimination of AA-tRNAs occurs in two stages: initial selection and proofreading (Figure 3). The overall accuracy \( A \) of codon selection is the multiplication of the accuracy of initial selection \( I \) and proof-reading factor \( F \):

\[
A = I \times F
\]

16)
Accuracy of protein synthesis is defined as the ratio of the efficiency $(k_{\text{cat}}/K_m)$ for peptidyl transfer between cognate and non-cognate reactions:

$$A = \frac{(k_{\text{cat}} / K_m)^c_{\text{dip}}}{(k_{\text{cat}} / K_m)^nc_{\text{dip}}}$$

Accuracy of initial selection $(I)$ is defined as the corresponding ratio for GTP hydrolysis on EF-Tu and is related to rate constants in Figure 3:

$$I = \frac{(k_{\text{cat}} / K_m)^c_{\text{GTP}}}{(k_{\text{cat}} / K_m)^nc_{\text{GTP}}} = \frac{k_a^c}{k_a^{nc}} \times \frac{1 + k_d^{nc} / k_c^{nc}}{1 + k_d^c / k_c^c}$$

A proofreading factor, $F$, can be derived from Eq. 16; it is the ratio between the probabilities of a peptide bond being formed after GTP is hydrolyzed for a cognate and a non-cognate reaction.

$$F = A / I = \frac{(k_{\text{cat}} / K_m)^{nc}_{\text{GTP}}}{(k_{\text{cat}} / K_m)^nc_{\text{dip}}} \times \frac{(k_{\text{cat}} / K_m)^c_{\text{GTP}}}{(k_{\text{cat}} / K_m)^c_{\text{dip}}} = \frac{1 + q_d^{nc} / k_{\text{pep}}^{nc}}{1 + q_d^c / k_{\text{pep}}^c}$$

Initial selection, proofreading as well as overall accuracy could be estimated by measuring the rate constants from ternary complex binding to the ribosome ($k_a$ and $k_d$) all the way to peptidyl transfer. Fast kinetics studies from Rodnina and collaborators (Gromadski and Rodnina 2004; Gromadski et al. 2006) showed that during the initial selection stage, non-cognate AA-tRNA in ternary complex dissociates rapidly from the ribosome prior to GTP hydrolysis (large rate constant $k_d^{nc}$ in Figure 3) and GTPase activation ($k_c$ in
Figure 3) is 100 to 400-fold slower for a non-cognate than for a cognate codon. Their results also showed that the dissociation rate constant $k_d$ for ternary complexes is dramatically and uniformly reduced around 1000-fold by single mismatches, regardless of the position and type of the mismatches. Furthermore, the uniform dissociation by single mismatches has also been observed in the proofreading stage ($q_d$ in Figure 3), suggesting a uniform translational accuracy for all codons. However, these results were based on fluorescence measurements with labeled homologous or heterologous tRNAs, fluorescent mant-GTP, non-hydrolysable GTP analogues and GTPase deficient mutants of EF-Tu at 20 ºC. Therefore, it could be difficult to generalize to the optimal conditions for rapid *E. coli* translation at 37 ºC.

A recent approach developed in our laboratory to more directly study initial selection is based on the concept of general Michaelis-Meten kinetics and takes advantage of the trade-off between cognate efficiency ($k_{cat}/K_m$) and accuracy ($I$) to estimate the intrinsic accuracy ($d$-value) for the selection of cognate in relation to non-cognate ternary complexes ([Johansson et al. 2012]; see discussion below). The results from our approach ([Johansson et al. 2012]) with native *E. coli* translation components at 37 ºC reveal an up to 17-fold variation from 1500 to 25000 in $d$-values of initial selection for all possible single base mismatch codons discriminated by the tRNA$^{1^{-ys}}$ ternary complex. In paper II, we extended this approach to measure the $d$-values of initial selection by six different tRNA-containing ternary complexes, in order to further understand the variation of accuracy in protein synthesis by different tRNAs and types of mismatches.

Initial selection of a ternary complex ends upon GTP hydrolysis on its EF-Tu, and is followed by a proofreading step. The concept of kinetic proofreading was first introduced by Hopfield ([Hopfield 1974] and Ninio ([Ninio 1975]), and it was identified experimentally for aminoacylation ([Hopfield et al. 1976]) and translation ([Thompson and Stone 1977; Ruusala et al. 1982]) in bacterial protein synthesis. The latter observed in a poly-(U) translation system that the number of GTPs hydrolyzed for non-cognate ternary complexes was much larger than the number of dipeptides formed, concluding that non-cognate AA-tRNAs could be rejected by the ribosome after the GTP hydrolysis that provides the driving force for the proofreading mechanisms. During the proofreading stage, non-cognate AA-tRNAs have high probability to dissociate from the ribosome rather than accommodate in the A site. However, little is known about the mechanistic details of proofreading. Structural studies are based on ribosomal complexes with already accommodated non-cognate anti-codon loops (mimics of the tRNAs) in the A site ([Demeshkina et al. 2012; Demeshkina et al. 2013]). However, since proofreading may occur in different states after GTP hydrolysis such experiments may not fully explain the proofreading mechanism. For example, in which step does the discrimination occur? Could non-cognate AA-tRNA dissociate during its release from EF-Tu:GDP and before moving to the peptidyl trans-
fer center (accommodation)? Does the ribosome discriminate against non-cognate AA-tRNA after it is fully accommodated but before peptidyl transfer? We tried to answer some of these unknowns in Paper III.

Trade-off between efficiency and accuracy

The universal trade-off between efficiency and accuracy of enzymatic reactions can be described from general Michaelis-Menten kinetics as (Kurland and Ehrenberg 1984; Johansson et al. 2008b; Johansson et al. 2012):

\[
\frac{k_{\text{cat}}}{K_m}^c = \frac{k_a (d - A)}{(d - 1)}
\]  

(20)

Where the \( \frac{k_{\text{cat}}}{K_m}^c \) is the efficiency for the cognate product formation, \( k_a \) is the association rate constant for the substrate, \( A \) is the normalized accuracy, and \( d \) is the discrimination parameter. \( A \) is defined as the ratio of efficiency \( \frac{k_{\text{cat}}}{K_m} \) between cognate and non-cognate reactions. In this relationship (Eq. 20), cognate efficiency \( \frac{k_{\text{cat}}}{K_m}^c \) reaches its maximal value \( k_a \) when accuracy \( A \) equals to one, meaning no discrimination between correct or incorrect substrates. \( \frac{k_{\text{cat}}}{K_m}^c \) linearly decreases with increasing accuracy \( A \), and it goes to zero when accuracy \( A \) reaches its maximal possible value \( d \), meaning no functional activity of the enzyme system.

During initial selection of protein synthesis, accuracy and efficiency parameters can be expressed in relation to the rate constants of the steps in Figure 3:

\[
I = \frac{(k_{\text{cat}} / K_m)^c}{(k_{\text{cat}} / K_m)^{nc}} = \frac{1 + d \cdot a}{1 + a},
\]

(21)

where \( \frac{k_{\text{cat}}}{K_m}^c = \frac{k_a}{1 + a} \), \( \frac{k_{\text{cat}}}{K_m}^{nc} = \frac{k_a}{1 + d \cdot a} \), and \( a = \frac{k_a}{k_d} / \frac{k_c}{k_e} \).

By tuning the \( a \)-parameter, accuracy of initial selection \( I \) can be varied from its lowest value 1 to its maximal value \( d \).

Such an efficiency-accuracy trade-off has been recently observed in biochemical experiments in our laboratory (Johansson et al. 2012). By varying concentration of free Mg\(^{2+} \) in the reaction, cognate codon reading efficiency \( \frac{k_{\text{cat}}}{K_m}^c \) for Lys-tRNA\(^{\text{Lys}} \) in ternary complexes with EF-Tu and GTP decreases linearly with increasing accuracy of initial codon selection in response to decreasing concentration of free Mg\(^{2+} \) in the reaction. The linear relationship between cognate efficiency and accuracy means that varying concentration of free Mg\(^{2+} \) in the reactions will only change \( a \)-parameter but
not the association rate constant $k_a$. Then a plot of $(k_{\text{cat}}/K_m)^c$ versus $A$ measured at different concentrations of free Mg$^{2+}$ gives the association rate constant $k_a$ for ternary complexes binding to the ribosome and the maximal accuracy $d$-values of initial selection at its y- and x-intercept, respectively.

Large variation in accuracy of initial codon selection (Paper II)

In Paper II, by using the same approach to measure the trade-off between cognate efficiency and accuracy, we extended the measurement of $d$-value for initial selection by six AA-tRNA containing ternary complexes: Cys-tRNA$^{\text{Cys}}$, Phe-tRNA$^{\text{Phe}}$, Glu-tRNA$^{\text{Glu}}$, His-tRNA$^{\text{His}}$, Tyr-tRNA$^{\text{Tyr}}$ and Asp-tRNA$^{\text{Asp}}$. We measured the accuracy of initial selection as the ratio of $k_{\text{cat}}/K_m$ between cognate and non-cognate reactions for the six tRNAs selected their fully matched cognate codons in relation to their all possible 48 near-cognate codons with single base mismatches including the 6 wobble base pairs at 3$^{rd}$ codon position. In all cases, the cognate efficiency $(k_{\text{cat}}/K_m)^c$ decreased linearly with the accuracy of initial selection $I$ (Figure 4A), as observed previously for tRNA$^{\text{Lys}}$ (Johansson et al. 2012). The association rate constants $k_a$-values were estimated from the linear trade-off plots for six tRNAs, they varies from 75 to 200 (Figure 4B). The maximal intrinsic accuracy, $d$-parameter, were also estimated by the same approach for all misreading codons with single mismatch and wobble reading at 3$^{rd}$ positions by six tRNAs (Table 1 in Paper II).
Figure 4. The linear rate-accuracy trade-off in initial selection for different tRNAs. (A) Rate-accuracy trade-off lines in plots of the efficiency of cognate GTP hydrolysis, \( (k_{\text{cat}}/K_m)^c \), versus the accuracy (calculated as the ratio of \( (k_{\text{cat}}/K_m)^c \) over efficiency of non-cognate GTP hydrolysis \( (k_{\text{cat}}/K_m)^{\text{nc}} \)) for different tRNAs reading single-mismatch codons as indicated in the figure (mismatch codon positions are underlined). For each tRNA misreading case, cognate and non-cognate \( (k_{\text{cat}}/K_m) \) values were measured at different Mg\(^{2+}\) concentrations as shown in Figure 2. The x-intercept gives the maximal accuracy, \( d \), for each misreading case. Insert: rate-accuracy trade-off for Glu-tRNA\(^{\text{Glu}}\) misreading codon GCA (black triangles and black lines), for which the \( d \)-value is estimated as 31000. (B) The association rate constant, \( k_a \), for binding of different aa-tRNAs in ternary complexes to ribosomes. \( k_a \)-values were estimated from intercepts of the trade-off lines with the y-axis, as shown in (A). That is, when the accuracy, \( A \), = 1 is identical to one (meaning no discrimination between cognate and non-cognate reactions), the efficiency reaches its maximal value, \( (k_{\text{cat}}/K_m)^{\text{GTP}}_{\text{max}} = k_a \). tRNA\(^{\text{Lys}}\) data were from (Johansson et al. 2012).

The \( d \)-value data set (graphically illustrated in Figure 5) from seven tRNAs (including previous measurements for tRNA\(^{\text{Lys}}\) (Johansson et al. 2012)) displays large variation in accuracy of initial codon selection on the ribosome, depending on the identity of the misreading tRNA, the type of the mismatch, and the codon position of the mismatch. Overall \( d \)-values of initial selection vary 400-fold in the range from 200 to 84000 (Table 1 in Paper II). In consistence with previous observation (Johansson et al. 2012), for the same type of mismatch \( d \)-values are general highest for mismatches in the second codon positions and lowest in the third codon positions. For most of cases regardless of the tRNA identity and the mismatch positions, U-C mismatches have the highest \( d \)-values while wobble base pairs U-G have the lowest. Furthermore, hot-spots for high initial selection errors have been identified for tRNA\(^{\text{Glu}}\) and tRNA\(^{\text{His}}\) misreading codons with U-G mismatches in second or third codon base positions. The very same error hot-spots for tRNA\(^{\text{Glu}}\) have also been identified in vivo (Manickam et al. 2014), confirming the functional relevance of our in vitro biochemistry measurements. Interestingly, those error hot-spots with the lowest \( d \)-values have G-C / C-G base pair in the first or / and third positions (Table 1 in Paper II), e.g. tRNA\(^{\text{Glu}}\) misreading GGA \( (d = 200) \) and tRNA\(^{\text{His}}\) misreading CGC \( (d = 250) \).
Figure 5. Maximal accuracy variation dependent on codon position, mismatch identity and AA-tRNA. Maximal accuracy values, $d$, for single-mismatch readings by different tRNAs, summarized with respect to mismatch codon position (columns) and mismatch identities (rows for the anticodon bases and colors for the codon bases as indicated in the figure).

This suggests that in general G-C base pairs enhance codon stability. Moreover, also in consistent with previous finding (Johansson et al. 2012), wobble base pairs at 3rd codon positions are less efficient in translation com-
paring to Watson-Crick base pairs. $d$-values are range from 1.3 to 9.0 for those have wobble base pairs at 3\textsuperscript{rd} positions. This could possibly explain a general pattern that the Watson-Crick base pair codons are more frequently used in the genome compared to the codons with wobble base pair at 3\textsuperscript{rd} position (Figure 6).

The experimental study on the efficiency-accuracy trade-off by seven AA-tRNA containing ternary complexes makes it possible to understand the variation on accuracy of codon reading in the living cell and how efficiency and accuracy have been evolutionary tuned to maintain growth and adaptation.

![Figure 6](image.png)

Figure 6. Discrimination of wobble base pairs at 3\textsuperscript{rd} codon base positon. $d$-values for different tRNAs reading their cognate codons with wobble base pairs instead of their fully matched codons. All $d$-values were estimated as described in Figure 4A. tRNA\textsubscript{Lys} data were from (Johansson et al. 2012).

Two proofreading steps of tRNA selection in translation (Paper III)

Based on a single-step selection mechanism (i.e. initial selection), the linear efficiency-accuracy trade-off shown in Figure 4A suggests that living cell needs to sacrifice efficiency a lot in order to maintain high accuracy in translation. For some misreading hot-spots like tRNA\textsubscript{Glu} misreading GAU against its cognate codon GAA, the accuracy is far below the level that required for a functional cell even it reaches its maximal $d$-value, 200. In fact, living organisms have been evolved to have two-step selection. In addition to initial selection, the living organisms use one more selection step, proofreading (Thompson and Stone 1977; Ruusala et al. 1982). The level of overall accu-
racy in protein synthesis is then greatly increased by proofreading without losing the efficiency (Johansson et al. 2008b). Furthermore, theoretical studies also demonstrated that how the biological system could be the benefited from multi-step proofreading (Ehrenberg and Blomberg 1980; Freter and Savageau 1980; Murugan et al. 2012). In Paper III, we propose a new two-step proofreading mechanism that EF-Tu is involved in the first proofreading step (Figure 7A).

Figure 7. Kinetics scheme of tRNA selection in bacterial protein synthesis. (A) EF-Tu-dependent AA-tRNA binding to the A site of the initiation complex. Nocognate tRNA can be rejected during initial selection and proofreading steps. Initial selection ends after GTP hydrolysis on EF-Tu. Here a two-step proofreading mechanism is suggested; first, AA-tRNA can be rejected in the form of EF-Tu:GDP:AA-tRNA after GTP hydrolysis (Proofreading I); later on, AA-tRNA alone can be rejected after its dissociation from EF-Tu:GDP on the ribosome (Proofreading II). (B) Dipeptide synthesis with AA-tRNA only (EF-Tu independent). The suggested common steps in schemes (A) and (B) were marked with dashed boxes in red.
We first measured the proofreading $F$ of the selection by ternary complexes (EF-Tu:GTP:AA-tRNA), and compared it to the accuracy ($A_{nf}$) of peptide bond formation from the same selection by AA-tRNA only in absence of EF-Tu. Proofreading $F$ was estimated by dividing overall accuracy $A$ (from peptide bond formation) by initial selection $I$ (from GTP hydrolysis) ($F = A / I$; Eq. 19). Overall accuracy $A$ or initial selection $I$ was obtained by the ratio of $k_{cat}/K_m$-parameters between cognate and non-cognate reactions for peptide bond formation or GTP hydrolysis reaction, respectively. The accuracy $A_{nf}$ was obtained by the ratio of $k_{cat}/K_{nr}$-parameters between congate and non-cognate peptide bond formation reaction by AA-tRNA only (in the absence of EF-Tu) (Figure 7B).

We compared $F$ and $A_{nf}$-values for tRNA$_{Glu}$ misreading codons GAU, GGA, and GAC (Figure 8) at different concentration of Mg$^{2+}$. We observed that proofreading $F$ is generally larger than the accuracy $A_{nf}$ of EF-Tu free selection. Surprisingly, $A_{nf}$ remained unchanged at [Mg$^{2+}$] varying from 1.3 mM to 3.4 mM (Figure 8), while the proofreading factor $F$ decreased with increasing [Mg$^{2+}$]. For all the tested cases, the differences between $F$ and $A_{nf}$ became smaller with higher [Mg$^{2+}$] and they even converged at 3.4 mM Mg$^{2+}$ for the case of tRNA$_{Glu}$ misreading GAC (Figure 8).

These results suggest a two-step proofreading mechanism. The first step involves EF-Tu while the second step does not. We propose, as shown in Figure 7, that in the first step, a non-cognate AA-tRNA could be rejected by the ribosome in complex with EF-Tu and GDP. The second step starts after the dissociation of EF-Tu:GDP from the ribosome, AA-tRNA alone could be rejected during its accommodation in the A site. The first proofreading step
(F₁) decreases with increasing [Mg²⁺] and could be eliminated at high Mg²⁺ concentration, while the second proofreading step (F₂) is insensitive to Mg²⁺ concentration. Additionally, after GTP hydrolysis, EF-Tu:GDP leaves the ribosome and AA-tRNA is in the pre-accommodation state. When EF-Tu is absent, AA-tRNA alone enters the A site of ribosome and it should appear in the same pre-accommodation state. The prediction from this hypothesis would be proofreading F increases from its minimal value F₂ (accuracy of second proofreading step) with increasing affinity between EF-Tu and AA-tRNA, and the accuracy of the second proofreading step F₂ equals to the accuracy Aₙ of AA-tRNA selection in absence of EF-Tu.

Figure 9. Engineering tRNA mutants with altered EF-Tu affinities. tRNA$_{\text{Glu}}$ (A) and tRNA$_{\text{Phe}}$ (B) mutants with T-stem mutated sequences (positions 51 to 53 and 61 to 63, shown in red box). T-stem sequences of wild type (WT) and different mutants with their corresponding $K_d$ values (in µM) for EF-Tu binding were shown. All tRNAs (WT and mutants) were unmodified and based on native E. coli tRNA$_{\text{Glu}}$ (A) or tRNA$_{\text{Phe}}$ (B) (black with purple anti-codon; tRNA modifications are in green) with changes in blue.

In order to test our hypothesis, we designed tRNA mutants with altered EF-Tu affinities based on E. coli tRNA$_{\text{Glu}}$ and tRNA$_{\text{Phe}}$ (Figure 9). All tRNA mutants were unmodified and prepared by in vitro transcription by T7 RNA polymerase, with the mutations in the T-stem (red box in Figure 9; position 51-53 and 61-63; (Schrader et al. 2009)) to cover a wide range of EF-Tu affinities. Unmodified tRNA$_{\text{Glu}}$ and tRNA$_{\text{Phe}}$ transcripts with wild type (WT) sequence were also prepared. We measured proofreading factor $F$-values and compared them to the accuracy $A_{\text{nf}}$ of the AA-tRNA selection in absence of EF-Tu for 6 misreading cases: tRNA$_{\text{Glu}}$ mutants misread codons GAU, GGA, and GAC, and tRNA$_{\text{Phe}}$ mutants misread codons CUC, UCC, and UUA. All tRNA$_{\text{Glu}}$ or tRNA$_{\text{Phe}}$ mutants had very similar initial selection (I) values for the same misreading cases (Table S2 in Paper III), indicating that the mutations in the T-stem of the tRNA and its EF-Tu affinity had no effect on the initial selection step during tRNA selection. In line with the prediction from our hypothesis, proofreading factor $F$ from the tRNA mutant increased
with increasing EF-Tu affinity while $A_{nf}$-values remained unchanged (Table S3 in Paper III) for all misreading cases. $F$ was plotted against the association constant ($1/K_d$) binding of AA-tRNA to EF-Tu (Figure 10) and displayed a strong positive linear correlation for all misreading cases. Furthermore, the y-intercepts from the plots of $F$ versus $1/K_d$ were indistinguishable from the $A_{nf}$-values for all six cases. It suggested that the second proofreading step ($F_2$) is identical to the accuracy $A_{nf}$ of EF-Tu free selection, which again was perfectly in line with our prediction on the two-step proofreading mechanism.

Our results suggest that living organisms have evolved to use two proofreading steps to enhance the overall accuracy in protein synthesis. The reason could be related to the large variation of initial codon selection as we observed in Paper II, in particular for the error hot-spots where small $d$-values were estimated in initial codon reading by tRNA$^{\text{Glu}}$ and tRNA$^{\text{His}}$.

Figure 10. Comparing proofreading factor ($F$, ■) of selection from ternary complex and accuracy ($A_{nf}$, ○) of selection from AA-tRNA only for tRNA mutants with altered EF-Tu affinities. $F$ and $A_{nf}$ from different tRNA mutants were plotted against the inverse of the tRNA affinities ($1/K_d$). Measurements were shown for Glu-tRNA$^{\text{Glu}}$ mutants misreading GAA (A), GGA (B), and GAC (C) and Phe-tRNA$^{\text{Phe}}$ mutants misreading CUC (D), UCC (E), and UUA (F) from their ternary complexes (F, ■) or AA-tRNA only ($A_{nf}$, ○). For each case of misreading, proofreading, $F$, versus $1/K_d$ were fitted by linear regression, the y-intercept shows the value of second proofreading step ($F_2$); the accuracy of tRNA only selection, $A_{nf}$, versus $1/K_d$ were fitted into a constant, which shows the average value of $A_{nf}$ from different tRNA mutants. $F_2$ and average $A_{nf}$ values from the fit are listed for each case respectively, (A) 8.2 ± 5.0 and 12 ± 0.8; (B) 23 ± 8 and 25 ± 2; (C) 12 ± 5 and 16 ± 1; (D) 34 ± 4 and 39 ± 2; (E) 14 ± 2 and 12 ± 0.3; (F) 27 ± 6 and 14 ± 0.4.
Translation with unnatural amino acids

What we know and what we don’t know

Ribosomal incorporation of unnatural amino acids (AAs) into peptides and proteins has become a powerful tool to study protein structure and function (Xie and Schultz 2005; Doi et al. 2007), and discover new therapeutic molecules (Yamagishi et al. 2011; Guillen Schlippe et al. 2012). However, translation with unnatural amino acids is often inefficient which severely hampers its ability (Forster 2009; Watts and Forster 2012). A lot of effort has been made during the last two decades to understand and improve their incorporation.

Incorporation of unnatural amino acids could be impaired in three main stages during protein synthesis: aminoacylation, ternary complex formation, and peptide elongation on the ribosome. In vivo studies have mainly been focused on the first and second stages (Park et al. 2011). Since unnatural AA-tRNAs could be synthesized artificially (see later discussion), we focused our in vitro studies on the effect on unnatural amino acid incorporation of ternary complex formation and later stage leading to peptide bond formation.

Regarding the mechanism of incorporation deficiency: A D-amino acid is found to be incorporated a thousand times more slowly than the L-amino acid (Yamane et al. 1981) because of the lower EF-Tu affinity and slower accommodation / peptidyl transfer at an unphysiological reaction temperature of 0 °C. Dipeptide formation with N-methyl-Phe was $10^4$ x slower than with natural Phe, and even undetectable with N-butyl-Phe due to the very slow steps occurring after GTP hydrolysis on EF-Tu (Pavlov et al. 2009). A recent study indicates that the chemistry of peptidyl transfer with N-methyl-AAs is limiting the peptide bond formation (Wang et al. 2014). However, the reasons for the inefficient incorporation of unnatural non-N-alkylated L-amino acids into peptides remain unclear (Forster et al. 2003; Gao and Forster 2010). Therefore, in Paper IV, we have investigated the kinetics of peptide bond formation with small and large, non-N-alkylated unnatural L-amino acids at physiological condition to further understand and improve the incorporation of unnatural amino acids.

Incorporation of unnatural amino acids could be improved in several ways: using engineered EF-Tu mutants with higher affinities for the corresponding unnatural AA-tRNAs can improve incorporation of O-
Phosphoserine \textit{in vivo} (Park et al. 2011) and some bulky unnatural amino acids \textit{in vitro} (Doi et al. 2007). However, EF-Tu mutants may not be optimized for incorporating natural amino acids \textit{in vivo} since natural AA-tRNAs have evolved to maintain uniform affinities for EF-Tu (LaRiviere et al. 2001) and EF-Tu mutants with too high affinities could decrease the rates of peptide bond formation (Schrader et al. 2011). Moreover, using EF-Tu mutants with very high affinity for bulky unnatural AA-tRNA may impair peptide bond formation by the slow release of the bulky AA-tRNA from EF-Tu:GDP after GTP hydrolysis (Mittelstaet et al. 2013). A proper choice of tRNA bodies for different unnatural amino acids has been shown to improve their incorporation (Cload et al. 1996; Zhang et al. 2007; Wang et al. 2014), but the mechanism awaits to be investigated. Natural amino acids should match to proper tRNA bodies in order to have an optimized EF-Tu affinity for the AA-tRNA since otherwise AA-tRNAs with too weak or too strong affinities for EF-Tu results in poor ternary complex formation (Asahara and Uhlenbeck 2005) or slow dipeptide formation (Schrader et al. 2011) respectively, but very little is known with unnatural amino acids. Therefore, in Paper V, we have investigated which steps in translation are affected by the matching between unnatural amino acids and tRNA bodies.

Furthermore, the studies on unnatural amino acid incorporation may also answer important questions regarding the mechanism of protein synthesis. Previous studies have shown that ribosomes do not discriminate cognate tRNAs with misacylated amino acids (Effraim et al. 2009), and ribosomes have a ‘kinetic safety gate’ to control delivery of bulky unnatural amino acids by impairing the release of the AA-tRNA from EF-Tu:GDP after GTP hydrolysis (Mittelstaet et al. 2013). Our studies also address if the ribosome can discriminate between small and large unnatural amino acids in different steps during translation.

Aminoacylation of tRNAs with unnatural amino acids

The first challenge for incorporating unnatural amino acids into polypeptides is to acylate them into tRNAs. There are several ways to achieve this purpose: 1) Chemical acylation. This is a classical method that we use in our studies to prepare misacylated tRNAs with natural and unnatural amino acids. We ligate the chemically synthesized $N$-nitroveratryloxy carbonyl protected aminoacyl-dinucleotides (pdCpA-AA-NVOC) to 3’-CA-truncated artificial tRNAs by T4 RNA ligase (Hecht et al. 1978; Bain et al. 1989; Robertson et al. 1989), followed by photolysis of the amino protecting group $N$-NVOC. This method generates at least two changes on the tRNA bodies: lack of tRNA modifications, substitution of the penultimate C with dC (although a very recent research shows that this substitution is not necessary (Kwiatkowski et al. 2014) (epub ahead of print)), and optional mutations in
the acceptor stem to enhance *in vitro* transcription efficiency (for example, CG swaps at positions 3 and 70 of tRNA$^{\text{Phe}}$ that were used in Paper IV). Among these changes, the dC (at position 75) on the tRNA has been found to hamper the incorporation of multiple unnatural amino acids for the still unclear reasons (Forster 2009), but all changes have only minor effect on the kinetics of dipeptide formation (Pavlov et al. 2009). This method has the highest flexibility of the choice between amino acids and tRNAs, it can be used to synthesize any type of AA-tRNA and generate amounts in *mg* range that is feasible for kinetic studies. 2) Synthetase evolution. This method uses aminoacyl tRNA synthetases (AARS) which are selected to accept unnatural amino acids (Hartman et al. 2006) or uses engineering AARS mutants (Santoro et al. 2002; Josephson et al. 2005). It is widely used for *in vivo* studies and produces natural tRNAs acylated with unnatural amino acids in the living cell. However, this method may only be adaptable to a certain set of AARSes and unnatural amino acids, which limits the studies *in vitro*. 3) Ribozyme-based acylation. This method uses artificial ribozymes, called flexizymes, to charge chemically synthesized activated amino acids (could be natural or unnatural) to the tRNA bodies (Murakami et al. 2006). It is a highly flexible method that tolerates a wide range of unnatural amino acids and tRNAs. Compared to the conventional ‘chemical acylation’ this method does not require any changes in the tRNA bodies, which allows direct assessments on the effect of the unnatural amino acid itself rather than on its corresponding AA-tRNA. However, the flexizyme system requires longer reaction time (typical several hours to 1 day (Murakami et al. 2006)) and the amount of AA-tRNA produced is often insufficient for *in vitro* kinetic experiments.

**Rate and efficiency (Paper IV)**

*In vitro* kinetic studies have shown that incorporation of unnatural amino acids is controlled by the translation system in different stages depending on the identity of the amino acids. In Paper IV, we investigated the single incorporation kinetics of non-\(N\)-alkyl \(L\)-AAs, allyl-glycine (aG), methyl-serine (mS), and biotinyl-lysine (bK) (Figure 11), which were the most commonly used class of unnatural amino acids, in order to understand in which steps their incorporations were affected.
Unnatural amino acids as well as natural Phe wereacylated to tRNA<sub>Phe</sub><sup>C3G,G70C</sup> (tRNA<sub>Phe</sub><sup>B</sup>) in this study, and natural Phe-tRNA<sub>Phe</sub> was also prepared by charging by aminoacyl-tRNA synthetase (AARS) as control. Experiments were conducted in near-physiological (polymix) buffer at 37 °C. We observed biphasic (fast and slow) kinetics for dipeptide formation with Phe, small and large unnatural AA-tRNA<sub>Phe</sub><sup>B</sup> as well as natural Phe-tRNA<sub>Phe</sub> (Figure 12A).

**Figure 12.** Effects of EF-Tu concentration on the kinetics of dipeptide synthesis from fMet-tRNA<sub>fMet</sub> and natural Phe-tRNA<sub>Phe</sub>. (A) Time course for normalized fraction of f<sup>3</sup>H]Met-Phe formed at different EF-Tu concentrations. Data were fitted non-linearly to a 2-exponential association model. (B) Curve representing non-linear fitting of the normalized fast phase fractions estimated from the kinetics of dipeptide synthesis (A) at different EF-Tu concentrations. $K_d$ for the binding of Phe-tRNA<sub>Phe</sub> to EF-Tu:GTP was estimated to be 0.26 μM from the fit.
The fraction of the fast phase increased with increasing EF-Tu concentration, but its rate remained unchanged. The fast kinetics represented the rapid peptide bond formation starting from active ternary complex (EF-Tu:GTP:AA-tRNA; T3) binding to 70S initiation complex (IC). Its fraction corresponded to the amount of total AA-tRNAs that was in active ternary complex with EF-Tu and GTP before. The slow kinetics represented ternary complex formation from free EF-Tu:GTP and AA-tRNA in the reaction. From the changes of the fast phase fraction in relation to the EF-Tu concentration, we estimated the dissociation constant, $K_d$, for binding of AA-tRNA to EF-Tu:GTP during translation (Figure 12B).

From the kinetics of dipeptide formation, we estimated the rates of the fast phases for small unnatural amino acids (aG and mS) incorporated from tRNA^Phe^B to be as fast as that for natural Phe incorporated from the same tRNA body, and they are also similar to that in the incorporation from natural Phe-tRNA^Phe^ (Table 1 in Paper IV). Which indicated that incorporation of unnatural amino acids proceeds at similar rates as that of natural amino acids in translation. aG-, mS-, and Phe-tRNA^Phe^B have similar $K_d$ values for EF-Tu, which are 2.5-fold higher comparing to natural Phe-tRNA^Phe^ (Table 1 in Paper IV). Therefore, small unnatural amino acids have similar binding affinity for EF-Tu as natural Phe, but their acylation to the artificial tRNA^Phe^B resulted in a weaker EF-Tu affinity than acylation to natural Phe-tRNA^Phe^. In further studies we by compared Phe acylated to an unmodified tRNA^Phe^ (only lack of tRNA modifications and AA- charged by PheRS) and to tRNA^Phe^B (by chemical acylation), we found that it is the lack of modifications in the tRNA^Phe^B body that causes the decrease in EF-Tu affinity rather than other changes (dC at position 75 and CG swap at postions 3 and 70). This is line with studies from the Uhlenbeck Laboratory (LaRiviere et al. 2001). Our results showed that ribosomal peptide bond formation with the small, unnatural amino acids aG and mS had similar rates as with natural Phe but the unnatural AA-tRNAs were inefficiently delivered to the ribosome by EF-Tu due to the unmodified tRNA body. Our results also suggested that the incorporation efficiency could be improved by increasing EF-Tu concentration in the translation system for small and non-N-acylated unnatural amino acids.

However, for incorporation of the bulky bK, the kinetics was slow for all EF-Tu concentrations below 10 µM, and a fast phase with barely detectable amplitude was observed in experiments at 10 µM EF-Tu. The (slow) rates were similar to the rates of the slow kinetics for the incorporation of small amino acids; reflecting the slow ternary complex formation. Not only fast phase amplitude for incorporation of bK at 10 µM was uncertain also the rate was uncertain (17 ± 12 s⁻¹; Table 1 in Paper IV), the $K_d$ of bK-tRNA^Phe^B for EF-Tu was estimated as 50 µM (Table 1 Paper IV). The high $K_d$ value and the presence of a fast phase implied that the inefficient incorporation of bK (Forster et al. 2003) was due to poor ternary complex formation but not...
to slow peptide bond formation on the ribosome. By measuring the $K_d$ of Lys-tRNA$^{\text{PheB}}$ for EF-Tu binding (0.57 µM) under the same experimental condition, we found that the bulky biotinyl- side chain was responsible for the decrease of the EF-Tu affinity of the AA-tRNA, which led to the low incorporation efficiency for bK.

Given the uncertainty of the fast phase for bK incorporation, we performed the experiments at lower temperature (20 ºC), at which a higher affinity for EF-Tu was expected (Pingoud et al. 1977; Ott et al. 1989). The fast phase in the kinetics of bK incorporation at 20 ºC was clearly visible and even became dominant at EF-Tu concentrations higher than 6 µM. The rate of the fast phase was comparable to the rates from dipeptide formation with natural amino acids at 20 ºC (Johansson et al. 2011), and $K_d$ was 3.3 µM. These suggested that lowering the reaction temperature improved the incorporation efficiency of unnatural, bulky amino acids. The results further confirmed the notion that the bulky side chain of bK only affected ternary complex formation but not peptide bond formation.

The ‘mysterious’ slow kinetics

We observed slow phases for the incorporation of natural and unnatural amino acids, and the slow rates (between 0.5 to 1 s$^{-1}$) were similar for all amino acids, indicating the existence of a common mechanism of on ternary complex formation that could be relevant to protein synthesis in general. However, the mechanism behind the slow phase remains unclear for natural and unnatural amino acids alike. Here I discuss the mechanism for the slow kinetics observed in our measurements.

First, to further confirm that the slow kinetics does reflect slow ternary complex formation, we measure the rates of ternary complex formation from EF-Tu:GTP and AA-tRNA (Supplemental Fig. S1 in Paper IV). The reaction was started by mixing 70S initiation complex mixture containing EF-Tu:GTP and an AA-tRNA mixture lacking EF-Tu. Then EF-Tu and AA-tRNA formed ternary complexes prior to binding to the ribosome. We found that the rate for dipeptide formation was much slower than the rate in experiments with from pre-formed ternary complexes where EF-Tu was pre-incubated with AA-tRNA. This means that T3 formation was rate-limiting in the reaction and that this rate determined the rate of peptide bond formation. The estimated rate of ternary complex formation was, in fact very similar to the rate of the slow phase kinetics where EF-Tu was pre-incubated with AA-tRNA.

Interestingly, the rates of the slow phases remain unchanged with varying EF-Tu concentration (Table 1 in Paper IV). This contradicts to the prediction from a simple single step binding model that the slow rates should increase linearly with increasing EF-Tu concentration. We then measured the rate of
ternary complex formation (as described above) with varying EF-Tu concentrations, and we found that the rates did not change significantly with increasing EF-Tu concentrations (Data not shown). Accordingly, the kinetic data suggested a two-step model where binding of AA-tRNA to EF-Tu:GTP was followed by a slow complex activation step. This slow step could for example be a conformational change of the ternary complex. This means, in other words, that binding of AA-tRNA to EF-Tu itself is much faster than the rate of the conformational change, so that the slow rates we observed are saturated already at low EF-Tu concentration and therefore remain unchanged. bK was an exception: the slow rates of the incorporation kinetics at 37 °C clearly increased with increasing EF-Tu concentration, but the rate of the slow phase saturated at around 1 s⁻¹ and did not depend linearly on EF-Tu concentration.

Another observation was that the slow phase kinetics did not disappear even at saturating EF-Tu concentration for binding of AA-tRNA to EF-Tu (e.g. Fig. 12A curve at 10 µM EF-Tu), fractions of the remaining slow phase varies from 10 to 25% with different natural or unnatural AA-tRNAs. One explanation could be that ribosomes or ternary complexes were heterogeneous, and the slow kinetics would then correspond to one of the populations that displayed fast binding of ternary complex to the ribosome but slow subsequent steps. However, this seems to unlikely since the experiments have been repeated with different batches of ribosomes and tRNAs and in different experimental setups like ribosomes in excess over ternary complexes, and vice versa, and, yet, invariably result in the same fraction of the slow phase at saturating EF-Tu concentration. Therefore, we inspect the problem in the light of theoretical model (Supplementary material in Paper IV), suggesting that the slow kinetics amplitude is related to the forward and backward rate constants for transforming an inactive to an active ternary complex.

One explanation to the remaining slow phase could be the isomerization of the aminoacyl residue between 2' and 3' of the P site tRNA. It is known that the 2',3'-isomerization of AA-tRNA occurs in bulk solution (Griffin et al. 1966). The transition rate constants have been measured by nuclear magnetic resonance (NMR) analysis under physiological condition (at pH 7.3 and 37 °C) and estimated as 3 to 11 s⁻¹ for trans-acylation (migration) of the aminoacyl residue from 2' → 3' and 1 to 4 s⁻¹ from 3' → 2' of the AA-tRNA (Taiji et al. 1983). This means that the equilibrium constant (for 2' → 3') varies between 1 to 10. Assuming the isomerization to occur also for ribosome bound AA-tRNAs, and that only the 3'-isomers can participate in translation. 70S ribosomes can only be initiated by 3'-fMet-tRNA^{fMet}, with the help of initiation factors. However, after initiation complexes are formed, if the fMet residue could move from 3' to 2'- on tRNA^{fMet} with similar rate and equilibrium constants (for 2' → 3') as in bulk solution, then the initiation complex could actually exist in two populations containing 2'- or 3'-
fMet-tRNA\(_{\text{fMet}}\) in P site, from which the active 3’- form of fMet-tRNA\(_{\text{fMet}}\) may occupy up to 90 % of the ribosome at equilibrium (calculated for an equilibrium constant for 2’ → 3’ of 10) before translation starts. This would mean that the dipeptide formation kinetics first displays a fast phase (up to 90 %) by peptide bond formation with the 3’ form of fMet-tRNA\(_{\text{fMet}}\), and a slow phase (10 % or more) with the rates limited by the transacylation rate for the 2’ → 3’ form of fMet-tRNA\(_{\text{fMet}}\), which is in the 1 to 4 s\(^{-1}\) range (in bulk solution). The type of kinetics would correlate very well the rate we observe for the slow phase. On the other hand, EF-Tu brings only 3’ form of AA-tRNA to the ribosome, and it could only be isomerized to 2’ form after its release from EF-Tu:GDP, but given that the subsequent steps after GTP hydrolysis are too fast compared to the rates of isomerization, it is unlikely that isomerization occurs in the A site tRNA.

The behavior of slow kinetics of dipeptide synthesis is ubiquitous in all measurements involving natural as well as unnatural amino acids, revealing an unknown rate-limiting step during ternary complex formation. This may not be important only for unnatural amino acid incorporation, but also for understanding a basic mechanism of protein synthesis in the living cell.

Choice of tRNA body (Paper V)

A proper choice of tRNA body has shown to improve the incorporation efficiency of unnatural amino acids (Zhang et al. 2007; Forster 2009; Wang et al. 2014), but very little has been known about the mechanism of this effect. In order to investigate which steps in translation are affected by the change of tRNA body, in Paper V, we measure the kinetics of dipeptide formation with unnatural as well as natural amino acids using the tRNA\(_{\text{AlaB}}\) body; an unmodified tRNA that has almost the same sequence as natural \(E.\ coli\) tRNA\(_{\text{Ala}}\) albeit with a penultimate dC. It is expected to have higher affinity for EF-Tu than most of others tRNA bodies (Asahara and Uhlenbeck 2002). We compare incorporation from tRNA\(_{\text{AlaB}}\) with that from the tRNA\(_{\text{PheB}}\) body (Paper IV).

We found the incorporation efficiency to improve significantly for the small unnatural amino acids aG and mS, and dramatically for the bulky bK by a swap from the tRNA\(_{\text{PheB}}\) to the tRNA\(_{\text{AlaB}}\) body (Table 2 in Paper V), due to the 10-fold decrease in \(K_d\) value (Table 1 in Paper V) for AA-tRNA binding to EF-Tu caused by the swap. The rates for the fast and slow kinetics for incorporation from tRNA\(_{\text{AlaB}}\) are very similar to those from tRNA\(_{\text{PheB}}\). This means that the efficiency increase is caused by improved delivery of AA-tRNA by EF-Tu to the ribosome.

Now, I discuss the effect of the individual steps on the ribosome for incorporation with unnatural amino acids by tRNA\(_{\text{AlaB}}\) body, and only focus on the fast phase of the incorporation kinetics. Measurement of dipeptide for-
information gives the average times, $\tau_{\text{dip}}$, which includes the total time for the following steps in sequence: 1) binding of ternary complex to the 70S initiation complex, 2) hydrolysis of GTP, 3) dissociation of A site AA-tRNA from EF-Tu:GDP, and 4) accommodation and peptidyl transfer. At low ribosome concentration (like 1 µM in our experiments), $\tau_{\text{dip}}$ is limited by the time for binding of ternary complex to the ribosome rather than the subsequence steps on the ribosome (steps 2 to 4) (see section ‘Experimental setup’ for detail). Therefore, measurements of dipeptide formation with natural and unnatural amino acids in Paper IV do not reflect their differences in rates of the steps on the ribosome.

With the advantage of the higher EF-Tu affinity of the tRNA$^{\text{AlaB}}$ body the dominance of the fast phase in the incorporation kinetics increases so that we are able to measure the average time for GTP hydrolysis, $\tau_{\text{GTP}}$, which includes the total time for step 1 and 2 above. Given the very fast GTP hydrolysis (500 s$^{-1}$; (Pape et al. 1998)), $\tau_{\text{GTP}}$ reflects the time for binding of ternary complex to the initiation complex. We then measured $\tau_{\text{dip}}$ and $\tau_{\text{GTP}}$ in the very same experiment and the difference of these two average times gives the passage time from A/T site to peptidyl transfer for an AA-tRNA (step 3 and 4), $\tau_{\text{acc,pep}} = \tau_{\text{dip}} - \tau_{\text{GTP}}$ ($\tau_{\text{acc,pep}}$ is the same parameter as $\tau_{\text{pep}}$ which is used in Paper I).

However, measuring $\tau_{\text{GTP}}$ for incorporation from bK-tRNA$^{\text{AlaB}}$ is technically difficult due to its low EF-Tu affinity (2.8 µM; at least 10 x higher than other AA-tRNA$^{\text{AlaB}}$s). Therefore, we measure the minimum time of $\tau_{\text{dip}}, \tau_{\text{dip-min}}$, to approximate $\tau_{\text{acc,pep}}$ for bK incorporation (see ‘Experimental setup’ for detail). The results are summarized in Table 4 in Paper V.

We found that peptidyl transfer to Phe-tRNA$^{\text{AlaB}}$ has the longest $\tau_{\text{acc,pep}}$, which correlates with its highest affinity for EF-Tu (Schrader et al. 2011). Ala-tRNA$^{\text{AlaB}}$ and natural Phe-tRNA$^{\text{Phe}}$, the ones having AAs cognate to their tRNA bodies, have shortest $\tau_{\text{acc,pep}}$, and the unnatural aG-, mS- and bK-tRNA$^{\text{AlaB}}$ have intermediate $\tau_{\text{acc,pep}}$ times. The results show that with proper choice of tRNA body, ribosomal incorporation with unnatural amino acids with small and even bulky side chains can attain similar rates of accommodation / peptidyl transfer as natural amino acids.

### Discrimination between natural and unnatural amino acids

During bacterial translation, the ribosome accepts 22 natural amino acids. These have a variety of side chains with different sizes and physical properties. For the incorporation it does not matter if tRNAs are aminoacylated with their cognate or non-cognate amino acids (Effraim et al. 2009). This implies that the ribosome does not strictly discriminate against to the side
chains that do not match the tRNAs. Now I will discuss how the ribosome discriminates between two classes of $L$-unnatural amino acids: unnatural non-$N$-alkylated amino acids with small, or large side chains and $N$-alkylated amino acids.

In our studies in Papers IV and V, we have shown that indeed the ribosome does not discriminate against amino acids even when they have unnatural small or bulky side chains. For the unnatural amino acids with small side chains, like $aG$ and $mS$, this is not surprising since crystallographic structures of the ribosome in complex with P and A site tRNAs (Selmer et al. 2006) and simulation of A site tRNA accommodation (Sanbonmatsu et al. 2005) do not suggest that having the side chains smaller than the big natural side chains would impede any steps on the ribosome. However, for bulky amino acids like $bK$, one would expect that accommodation and the chemistry of peptidyl transfer would be inhibited due to the movement of the bulky side chain. Furthermore, the ‘chemical reactivity hypothesis’ of translation (Zhang et al. 2007) proposes that the chemistry of peptidyl transfer is slowed down by increased steric bulk on or near the $N$ nucleophile and by increased $pK_a$ of this $N$ nucleophile. On the other hand, $bK$ has a very weak affinity for EF-Tu, it may lead to a very fast dissociation of the its AA-tRNA from EF-Tu:GDP after GTP hydrolysis.

Dipeptide formation has been found to be 2 to 4 orders of magnitude slower with $N$-alkylated than with natural AAs (Pavlov et al. 2008; Leong et al. 2014; Wang et al. 2014) depending on the choice of tRNA body and experimental condition (e.g. buffer and pH). The involvement of the $\alpha$-NH$_2$ group during peptidyl transfer and the strong pH dependence of the kinetics of dipeptide synthesis with $N$-methyl AAs (Wang et al. 2014) indicate that the ribosome discriminates the $N$-alkylated AAs by the very slow chemistry of peptidyl transfer, which further leads to the rejection of the accommodated AA-tRNA before peptide bond formation. A competition experiment (Wang et al. 2014), where equal amounts of Pro-tRNA$^{Pro}$ and $N$-methyl-Phe-tRNA$^{Pro}$ were reacted to the 70S ribosome with fMet in the P site and a Pro codon in the A site, shows that the initially rapid incorporation of Pro is followed by slow incorporation, where the slow phase reflects the rejection of the $N$-methyl-Phe-tRNA$^{Pro}$ and re-binding of free Pro-tRNA$^{Pro}$ to the ribosome. The authors calculates that 5 to 8 cycles of delivery and rejection of $N$-methyl-Phe-tRNA$^{Pro}$ are required to form one dipeptide, and this could be considered as a proofreading step that controls the incorporation of unnatural $N$-alkylated amino acids by the ribosome.
Jag anser att som forskare har man ansvar för att (1) bekräfta det som är känt och (2) att besvara frågor om det som är okänt. Därför har jag i min avhandling kritiskt diskuterat vilka vedertagna fakta om proteinsyntesen och vilka mer kontroversiella påståenden som har lett fram till de nya upptäckter vi har gjort i artiklarna.


I artikel 1 studerar vi pH-beroendet för peptidbindningsreaktionen. Det har tidigare föreslagits att denna begränsas av den långsamma tRNA-ackommoderingen (då aminoacyl-tRNA flyttar sig in i ribosomens peptidyltransferascenter), men inte av själva peptidyltransfersteget då den kemiska reaktionen sker, så att mätningar av peptidbindningsreaktionen skulle vara oberoende av pH. I vår studie ser vi ett klart pH-beroende för peptidbindningsreaktionen samt visar att hastigheten för det kemiska peptidyltransferssteget för glycin och prolin begränsar hastigheten för bildningen av en peptidbindning vid fysiologiska förhållanden (pH 7.5). Detta kan vara sant även för samtliga övriga aminoacyl-tRNA vid dessa förhållanden.

I artikel 2, studerar vi avvägningen mellan effektivitet och noggrannhet för initialselektionen i kodonläsningen för sju ternärkomplex innehållande olika aminoacyl-tRNA. Det har tidigare föreslagits att selektionen är likvärdig för olika aminoacyl-tRNA och att felfrekvensen är cirka $10^{-4}$. I motsats till detta, ser vi stor variation i nogrannheten för initialselektionen. Vi indentifierar också flera ”hot spots” för fellläsning i kodonöversättningen, varav vissa också har indentifierats i levande celler. Den maximala nogrannheten varierar med upp till 400 gånger från 200 till 84000, beroende på vilket tRNA...
som studeras samt typen av felläsning och i vilken position i kodonet falläsning sker.

I artikel 3 föreslår vi en mekanism för proofreadingselektionen som innehåller två på varandra följande irreversibla steg som aldrig har observerats tidigare. Denna mekanism kan vara mycket viktig för cellens förmåga att upprätthålla både hög noggrannhet och hög effektivitet i proteinsyntesen.

I artikel 4 och 5 visar vi att bildning av peptidbindningar till små och stora onaturliga L-aminosyror som saknar N-alkylering sker på ribosomen med hastigheter liknande de som uppmätts för de naturliga aminosyrorna fenylalanin och alanin. Intressant nog så påverkar den stora sidokedjan på stora onaturliga aminosyror endast bindningen till EF-Tu genom att göra den svagare, men den påverkar inte hastigheten för peptidyltransferreaktionen på ribosomen. Våra resultat visar också att effektiviteten för inkorporering av onaturliga aminosyror generellt kan förbättras genom att öka EF-Tu koncentrationen, minskareaktionstemperaturenoch/eller genom att tRNAn med högre affinitet för EF-Tu används i proteinsyntessystemet.
Acknowledgements

Måns: Thanks for everything you shared with me, both for science and for life, especially for your passion in science and the critical way that you look at problems. I enjoy working with you. You make me realize that I can actually handle much more things in a time than I thought before….Thank you!

Misha: Thanks for everything that you teach me, especially for the experimental techniques and for knowing so much…..

Tony: It is so nice to work with you and I learnt a lot from you when we were writing the manuscripts together.

Suparna: For all the discussion and scientific inputs.

Ülkü: Thanks for performing the very first experiments with tRNA mutants for our proofreading project, it could have been holding for very long time due to many reasons if you never come and show us the convincing results.

Jinfan: Not only thanks for the proofreading of my thesis, but also all the things that we shared, the conversation and the dinners.

Anneli: Thanks for helping me with the Swedish Summary.

Magnus and Jingji: Thanks for the collaborations.

All past and current members at the ‘Molecular Biology’ Program

I would like to thank my parents for their endless supports.

Kongjing and Max,
You are my support! I don’t have to say much, you mean so much to me!
Much more than I can say!
References


Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 1203

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”.)