Max-Planck-Institut
für Kolloid- und Grenzflächenforschung

Universität Potsdam

Protein Synthesis by Ribosomes

Dissertation

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften (Dr. Rer. Nat.)
in der Wissenschaftsdisziplin Theoretische Physik

eingereicht an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Potsdam

angefertigt in der
Abteilung Theorie und Bio-Systeme
des Max-Planck-Instituts
für Kolloid- und Grenzflächenforschung

von
Sophia Rudorf

Potsdam, den 25.11.2014

Abstract

The synthesis of proteins is a fundamental process in all living cells. Every protein consists of at least one chain of amino acids. The concatenation of individual amino acids into a peptide chain is achieved by biomolecular machines called ribosomes. During protein synthesis, a ribosome moves along a messenger RNA (mRNA). It reads the mRNA codon by codon and takes up the corresponding ternary complexes, each of which contains an aminoacylated transfer RNA (tRNA). This process is called translation elongation.

Translation elongation is a nonuniform process with codon-specific elongation rates and error frequencies. Which mechanisms influence the global and local speed and accuracy of protein synthesis is a long-lasting puzzle. In this thesis, we address this question and develop an analytical theory describing translation as a Markov process. The theory takes major aspects of *in vivo* translation into account. We distinguish non-cognate, near-cognate and cognate ternary complexes that compete for binding to the ribosome in a concentration-dependent manner. The concentrations of free ternary complexes are calculated from the total tRNA concentrations considering the details of the tRNA life cycle. These concentrations determine the codon-specific elongation rates and error frequencies. Furthermore, we elucidate how the concentrations of free ternary complexes are influenced by the mechanism of tRNA release from the ribosomal E site.

Like many other cellular processes, translation is a complex multistep process that includes numerous individual transitions of the involved molecules. Our theory incorporates new experimental data on the rates of these transitions. All transition rates are measured *in vitro*, because so far it is not possible to determine these rates *in vivo*. Therefore, we develop a general method to predict *in vivo* rates from their *in vitro* counterparts. The prediction of the *in vivo* rates is based on a constrained minimization of the kinetic distance, which is a new measure that we introduce to quantitatively compare the kinetics of a process in different environments.

We apply our theory to *Escherichia coli* and study the dependence of the speed and accuracy of protein synthesis on a variety of parameters. We find that the overall speed of protein synthesis strongly depends on individual tRNA concentrations as well as on the abundance of active ribosomes and exhibits a non-equilibrium phase transition at low tRNA or high ribosome concentrations. Furthermore, we show that codon-specific elongation rates and error frequencies are considerably influenced by the overall codon usage in the cell. In general, our theory of protein synthesis is applicable to *in vitro* systems as well as to other prokaryotic or eukaryotic cells.

Contents

Abbreviations and Symbols

Abbreviations

co	Cognate.
no	Non-cognate.
nr	Near-cognate.
dbl/h	Doublings per hour, unit of cellular growth rate.
DNA	Deoxyribonucleic acid.
EF-Tu	Elongation factor thermo unstable.
GTP	Guanosine-5'-triphosphate.
OERC	Overall elongation rate curve.
RNA	Ribonucleic acid.
mRNA	Messenger ribonucleic acid.
tRNA	Transfer ribonucleic acid.
TC	Ternary complex.

Concentrations

$\mathcal{E}$	Total concentration of EF-Tu molecules.
$\mathcal{E}^{\text{fr}}$	Concentration of free EF-Tu molecules.
$\mathcal{R}$	Total concentration of actively translating ribosomes.

X_a	Total concentration of tRNA species a.
$\hat{\mathrm{X}}_a$	Concentration of free ternary complexes containing tRNA species a.
$\hat{\mathrm{X}}_a^{(i)}$	Concentration of ternary complexes of species a bound to ribosomes in internal state i.
$\mathrm{X}_a^{\mathrm{A}}$	Concentration of tRNAs of species a accommodated in ribosomal A sites.
$\mathrm{X}_a^{\mathrm{P}}$	Concentration of tRNAs of species a accommodated in ribosomal P sites.
$\mathrm{X}_a^{\mathrm{E}}$	Concentration of tRNAs of species a accommodated in ribosomal E sites.
$\mathrm{X}_a^{\mathrm{re}}$	Concentration of tRNAs of species a in the process of recharging with a new amino acid.
$\mathrm{X}_a^{\mathrm{ch}}$	Concentration of tRNAs of species a recharged with an amino acid but not part of a ternary complex.
$\mathrm{X}_a^{\mathrm{th}}$	Threshold concentration of tRNA species a.

Rates and Times

κ_{on}	Binding rate constant of ternary complexes initially binding to ribosomes.
ω_{ij}	Transition rate from state i to state j.
ω_{con}	Ribosomal transition rate including GTP hydrolysis and conformational changes within the ribosome upon cognate or near-cognate ternary complex recognition.
ω_{off}	Unbinding rate of initially bound ternary complexes.
ω_{pro}	Processing rate of tRNAs including peptide bond formation and translocation.
ω_{rec}	Rate of cognate and near-cognate ternary complex recognition.
$\omega_{c,\mathrm{elo}}$	Codon-specific elongation rate, inverse of the codon-specific elongation time $\mathrm{t}_{c,\mathrm{elo}}$.
ω_{elo}	Overall elongation rate expressing the number of peptide bonds formed per unit of time and ribosome, inverse of the average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$.

$\mathrm{J}_{\mathrm{elo}}$	Total elongation flux expressing the total number of peptide bonds formed per unit of time.
$\mathrm{t}_{(c\|i)}$	Average total dwell time in state i per elongation cycle on codon c.
$\mathrm{t}_{c,\mathrm{elo}}$	Codon-specific elongation time, inverse of the codon-specific elongation rate $\omega_{c,\mathrm{elo}}$.
$\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$	Average elongation time, inverse of the overall elongation rate ω_{elo}.

Sets

$\mathbf{A}_{\mathrm{co}}$	Set of cognate tRNAs.
$\mathbf{A}_{\mathrm{no}}$	Set of non-cognate tRNAs.
$\mathbf{A}_{\mathrm{nr}}$	Set of near-cognate tRNAs.
$\mathbf{C}_{\mathrm{co}}$	Set of cognate codons.
$\mathbf{C}_{\mathrm{no}}$	Set of non-cognate codons.
$\mathbf{C}_{\mathrm{nr}}$	Set of near-cognate codons.

Other Symbols

$(c\|i)$	State in the Markov model of translation elongation indicating the codon c in the ribosomal A site and the internal state i.
π_{ij}	Transition probability for the transition from state i to state j.
$\mathcal{D}_{ij}$	Kinetic distance for transition from state i to state j.
p_c	Codon usage of codon c.
$\mathcal{P}_{c,\mathrm{co}}$	Probability of cognate accommodation, a measure of the codon-specific fidelity.
$\mathrm{P}_{c,i}$	Steady state probability to find a ribosome on codon c in state i.

$\mathcal{P}_{c,\mathrm{nr}}$	Probability of near-cognate accommodation, also named *near-cognate missense error frequency*.
$\mathbf{S}_c$	State space for a given codon c.

1 Introduction

Without ribosomes, life does not make sense.

Harry Noller, UC Santa Cruz

Cells of all organisms from every kingdom of life consist largely of proteins. Proteins govern almost every cellular process from cell growth and division to metabolic activity, cell motility or signaling. All of these proteins need to be synthesized. As a consequence, the process of protein synthesis is pivotal for a proper functioning of the cells.

Protein synthesis must comply with high requirements concerning speed and accuracy. It must be fast enough to ensure doubling of protein mass within the time scale of cell division. At the same time, the synthesis of proteins must be very precise, because erroneous proteins are often dysfunctional or even harmful to the cell. Therefore, perturbations that hamper or dysregulate protein synthesis can lead to all kinds of cellular defects up to cell death. Indeed, a well-known strategy to fight bacterial infections is based on antibiotics that block protein synthesis in bacteria, hence killing the microorganisms. On the other hand, many diseases in multicellular organisms arise from genetic mutations that affect protein synthesis. Thus, it is of great general interest to understand what determines the speed and accuracy of protein synthesis. In addition, understanding the dynamics of protein synthesis is of technological importance. The production of protein-based materials and medical drugs, like vaccines or human insulin, relies on a sufficient supply with appropriate proteins. Often, the desired proteins are not chemically synthesized. Instead, genetically modified microbes or cell-free expression systems are used. To optimize these organisms and systems with respect to synthesis efficiency and functionality of the protein product, it is again necessary to understand what controls the speed and accuracy of protein synthesis.

In this thesis, we develop and apply a new theoretical framework for the process of protein synthesis. We start with an introduction that gives an overview over the most important components involved in the synthesis of proteins, followed by a brief introduction to Markov processes. The introductory chapters are succeeded by two chapters on our theory of protein synthesis and its experimental background. In the subsequent chapter, we introduce the *kinetic distance* and describe a method to predict protein synthesis in living bacteria from data obtained by *in vitro* experiments. In the last three

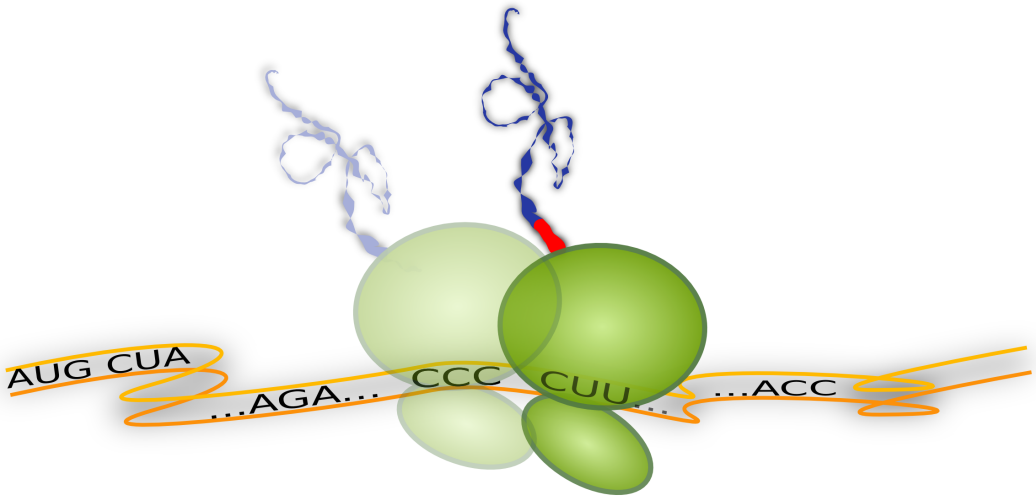


Figure 1.1: **Ribosome translating one codon.** The ribosome (green) moves along the messenger RNA (orange) and translates it codon by codon (... CCC, CUU, ...) into a sequence of amino acids. For each codon, a new amino acid (red) is attached to the growing peptide chain (blue).

chapters, we use our theory to study the dependence of the synthesis process on several experimentally accessible parameters. Revealing these dependencies is crucial to understand protein synthesis under non-physiological conditions related to cancer and other diseases.

1.1 Translation: From mRNA to Protein

Proteins are polymers that consist of amino acids. In spite of the very different tasks these molecules have to fulfill, all proteins have basically the same origin. According to the central dogma of molecular biology, the genetic information of an organism is stored in its DNA (*D*eoxyribo*N*ucleic *A*cid). Those parts of the DNA that encode blueprints for proteins are transcribed into mRNA (*m*essenger *R*ibo*N*ucleic *A*cid) if needed. Similar to DNA, mRNA is built up from four different species of nucleotides, adenine (A), cytosine (C), guanine (G), and uracil (U). A set of three subsequent nucleotides defines a *codon* and each codon corresponds to a specific amino acid. Therefore, a mRNA is a sequence of codons that serves as a template for the assembly of a certain chain of amino acids. A complex molecular machine called *ribosome* moves along the mRNA while it reads codon by codon and constitutes a corresponding peptide chain that folds into a

functional protein, see the cartoon in fig. 1.1. This process is called *translation* and can be divided into the three phases initiation, elongation, and termination.

During initiation of translation, the ribosome binds to the start codon of the mRNA [1]. The subsequent codon-wise movement of the ribosome along the mRNA is called elongation and consists roughly of the following steps. First, an aminoacylated tRNA (*t*ransfer RNA) binds to the ribosome. If the tRNA matches the currently read codon, it is accommodated within the ribosome. The peptide chain is then extended by forming a peptide bond with the amino acid that is attached to the tRNA. Finally, the ribosome moves to the next codon. In this work, we will focus on the elongation phase of translation, for which a more detailed description is given in chapter 3.
Translation is terminated as soon as the ribosome encounters one of the stop codons. Release factors unbind the peptide chain from the last tRNA and the ribosome dissociates from the mRNA [2].

1.2 Amino Acids, tRNAs and Ternary Complexes

To translate a mRNA, a ribosome does not take up amino acids directly. Instead, every amino acid is attached to a tRNA molecule, which contains a sequence of three nucleotides called the *anti-codon*. Only if the anti-codon of the tRNA matches the codon that is read by the ribosome, a new peptide bond is formed between the accompanying amino acid and the elongated peptide chain.

In general, an anti-codon matches a codon when all three pairs of nucleotides follow the Watson-Crick rules of base pairing, i.e., A is paired with U, and C is paired with G. In this case, the tRNA is *cognate* to the codon. However, there are exceptions to this rule and some tRNAs are cognate to codons even though their anti-codon exhibits one or more Watson-Crick mismatches. In fact, it turns out the number of tRNAs is always less than the number of sense codons, implying that the same tRNA can be cognate to more than one codon, and that in turn also codons can be cognate to more than one tRNA. For the bacterium *Escherichia coli* (*E. coli*), the cognate tRNAs have been determined experimentally for each sense codon [3].

Generally, a tRNA that is not cognate to a given codon is called *non-cognate* and will not be accepted by a ribosome translating this codon. There are also exceptions to this rule, as some tRNAs that are not cognate to a given codon are still treated almost as such by the ribosome, though with a lower probability than cognate tRNAs. These tRNAs are referred to as *near-cognate*. Their acceptance by a ribosome can lead to errors in the translated message if the near-cognate tRNA carries a different amino acid than the cognate tRNA. A formal definition of the property *near-cognate* is given by Kramer and Farabaugh [4]: A tRNA is near-cognate to a codon, if the codon differs in a single base

Figure 1.2: **Cognate (blue), near-cognate (purple) and non-cognate tRNAs (white)** for all sense codons in *E. coli* following the definitions in [3] and [4].

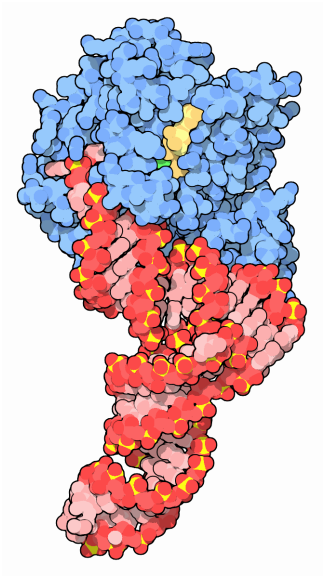


Figure 1.3: **Ternary complex** consisting of an elongation factor EF-Tu (blue), a tRNA (red), and a GTP molecule (yellow) [6].

from any of the cognate codons of this tRNA. However, it is not yet fully understood what determines the probability of acceptance for a near-cognate tRNA, i.e., the degree of similarity to a cognate tRNA, and the precise definition of the term near-cognate is still discussed in the literature [5]. The cognate and near-cognate tRNAs of all sense codons in *E. coli* following the definition of Kramer and Farabaugh [4] are shown in fig. 1.2.

Aminoacylated tRNAs bind to the ribosome as part of a *ternary complex* (TC) consisting of *elongation factor EF-Tu*, aminoacyl tRNA (aa-tRNA) and a guanosine-5'-triphosphate (GTP) molecule, see fig. 1.3. GTP is hydrolyzed to guanosine diphosphate (GDP) during the accommodation of a cognate or near-cognate tRNA in a translating ribosome. After tRNA accommodation the EF-Tu·GDP complex unbinds from the ribosome and GDP is replaced by a new GTP molecule. A used tRNA has lost its amino acid. Before it can again become part of a TC, it needs to be recharged with a new amino acid by an enzyme called aminoacyl tRNA synthetase (aaRS). In general, the cell contains one species of aaRS for each of the twenty different amino acids. Some examples of aaRS are shown in fig. 1.4.

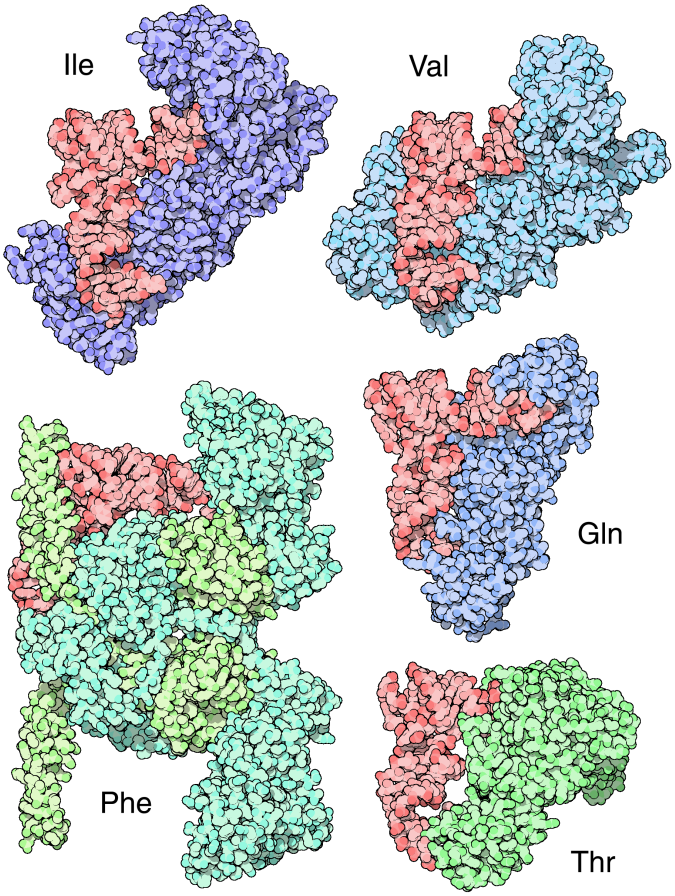


Figure 1.4: **Aminoacyl tRNA synthetases [7].** The cell contains one species of aminoacyl tRNA synthetases for each of the twenty different amino acids. The figure shows five different aminoacyl tRNA synthetases (isoleucyl (Ile), valyl (Val), glutamyl (Gln), phenylalanine (Phe), and threonyl (Thr) tRNA synthetase) in green or blue, with corresponding tRNAs in red.

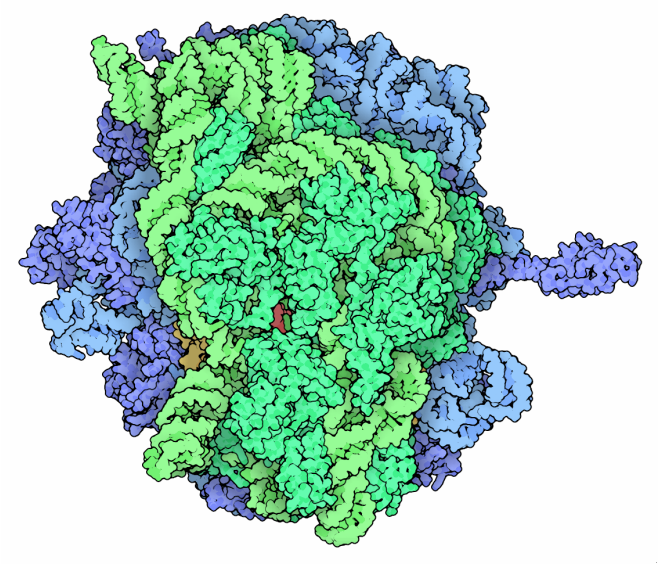


Figure 1.5: **The functional ribosome** consists of a small (green) and a large (blue) subunit containing three ribosomal RNAs and 55 ribosomal proteins [8].

1.3 Structure and Dynamics of Ribosomes

In the past decades, the molecular structure of the ribosome and the ribosomal translation pathways were intensively studied [9, 10]. Each ribosome consists of a small and a large subunit, both of which consist of RNA and protein molecules, see fig. 1.5. In *E. coli*, the small subunit *30S* contains one ribosomal RNA (rRNA) and 21 ribosomal proteins, whereas the large subunit *50S* is assembled from two rRNAs and 34 ribosomal proteins [11]. During the initiation of translation, the 30S subunit binds along with an initiator tRNA to the mRNA upstream of the coding sequence and is complemented by a 50S subunit to form a functional *70S* ribosome at the start codon of the mRNA [1].

The ribosome has three tRNA binding sites, the A, P, and E site. A new cognate or near-cognate aa-tRNA is accommodated in the A site. The P site contains the tRNA that was taken up in the previous elongation step and to which the synthesized peptide chain is attached. After the peptide chain is transferred from the P site to the A site tRNA, the ribosome translocates to the next codon. During translocation, the A site tRNA is pushed to the P site, and the P site tRNA moves to the E site from where it leaves the ribosome. The A and the P site are related to the core tasks of a ribosome, i.e., decoding and peptide bond formation. The function of the E site is not yet fully understood but it is believed that it supports translocation of the ribosome [10]. It is

widely assumed that the E site tRNA is released from the ribosome before a new TC binds to the A site. This was named *2-1-2* pathway of E site tRNA release and has been corroborated by several *in vitro* experiments [12–15]. However, for the early cycles of elongation corresponding to a relatively short length of the nascent peptide chain, the alternative *2-3-2* pathway has also been observed [14]. In this pathway, the binding of a new TC to the A site precedes the release of the E site tRNA [16].

Decoding, peptide bond formation and translocation are the three different phases of an elongation cycle. During the decoding phase, the ribosome distinguishes cognate tRNAs from near- and non-cognate tRNAs. After a cognate tRNA has been recognized, it is accommodated in the ribosomal A site under hydrolysis of the accompanying GTP. In this state, the aminoacylated end of the A site tRNA is in the peptidyl-transferase center (PTC) of the ribosome, pointing towards the P site tRNA and the synthesized peptide chain. Then, the PTC catalyses the peptide bond formation resulting in a transfer of the peptide chain from the P site to the A site tRNA [17]. It is an open question how the ribosome catalyses the peptidyl-transferase reaction though it was suggested that arranging the reaction partners is the key mechanism [10]. Following peptide bond formation, an elongation factor called EF-G binds to the ribosome and, fueled by hydrolysis of a GTP molecule, facilitates its translocation to the next codon.

1.4 Towards a Comprehensive Model of Translation Elongation

Although each elongation cycle involves basically the same steps, the speed of translation is not uniform. For two decades, it is known that *in vivo* elongation rates are codon-specific, i.e., the local velocity of the ribosome depends on the codon that is translated [18, 19]. Many mechanisms have been proposed to explain this variability in the speed of translation. For example, mRNA is often folded into a secondary structure which the ribosome has to open up before it can translate the message [20, 21]. Furthermore, special *Shine-Dalgarno* codon sequences interact with the ribosome and slow it down [22]. Also, the peptide chain can interact with the ribosome, especially the part that is still in the ribosomal exit tunnel [23]. In addition, ribosomes can be slowed down by preceding ribosomes translating the same message if they come into close contact [24]. Equivalently, codon or tRNA dependent variations in the ribosomal processing, and the presence of co-translational processes, like chaperone interactions or co-transcriptional translation, could affect the efficiency of the elongation process [25, 26].

One additional factor is believed to determine the local speed of translation, namely the concentration of the TCs. In contrast to all of the other influencing factors mentioned above, the concentration of TCs represents a universal control of the speed of translation,

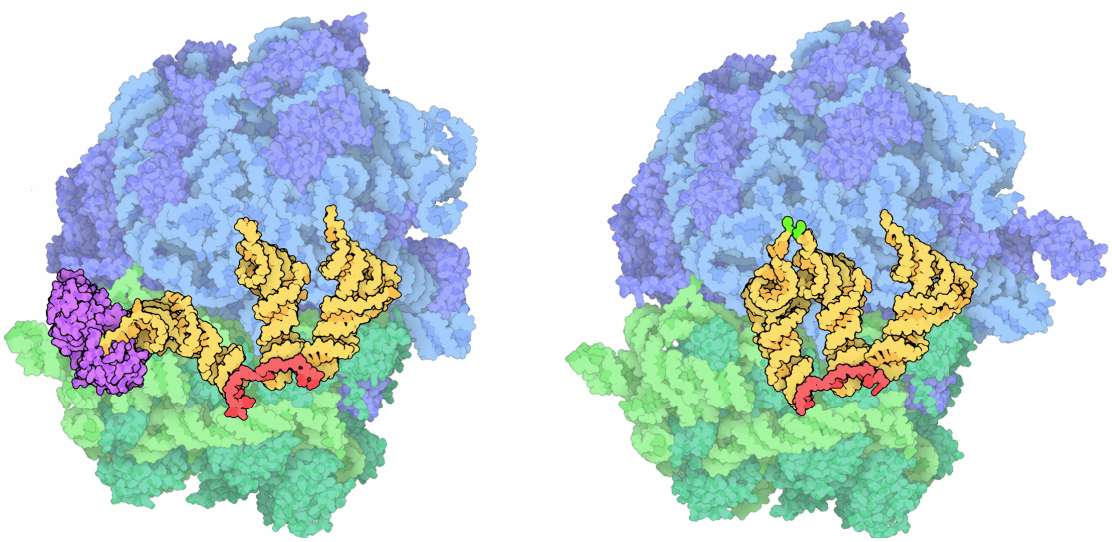


Figure 1.6: **Accommodation of a new tRNA and peptide bond formation, adapted from [8].** Left: A new tRNA (yellow, left) is accommodated in the ribosomal A site while it is still attached to the accompanying elongation factor EF-Tu (purple). Right: A peptide bond is formed and the peptide chain (green) is transferred from the P site tRNA (yellow, middle) to the A site tRNA. In addition, the E site tRNA (yellow, right) and a piece of the translated mRNA (red) are shown.

independent of the mRNA that is translated or the organism that is studied. One can think of the TCs as being the fuel of translation. Because the ribosome has to wait for a cognate TC to bind before it can proceed, it is plausible that translation is faster when more TCs are present. In the literature, this idea was used to define quantities proportional to the concentration of cognate TCs and to interpret these quantities as a measure for the actual codon-specific elongation rate [27, 28].

From a theoretical point of view, not only the cognate, but also the non- and near-cognate TCs should have an effect on the elongation rate of the corresponding codon, as they can bind to the ribosome as well. Experimentally, their influence has still to be clarified [29–31]. The competition between cognate, near- and non-cognate tRNAs was studied theoretically and was found to strongly affect codon-specific elongation rates [32, 33].

In addition to the effect of near- and non-cognate aa-tRNAs, the competition between the active ribosomes for TCs also has a strong influence on the elongation speed. Each *E. coli* cell contains about twenty to thirty thousand ribosomes actively translating at the same time. As they always keep one tRNA in their P site and sometimes a tRNA in their E site, active ribosomes dramatically reduce the concentrations of tRNAs that can form TCs. Therefore, the simultaneous action of ribosomes can turn abundant into rare TCs, and calculations of elongation rates should always take into account the concentrations of available TCs rather than the total amount of tRNAs.

In the course of this work, we develop a theoretical framework that allows to determine the codon-specific speed of translation based on tRNA concentrations. Previously published theoretical work on tRNA concentration dependent elongation rates has several limitations [32–35]. In part, the predictive power of these theoretical models is limited because they omit fundamental aspects of translation. In most studies, the TC concentrations have not been estimated systematically although the calculation of codon-specific elongation rates is based on these concentrations. Instead of considering that the amount of available TCs is affected by binding of tRNAs to ribosomal A, P, and E sites and by the need to recharge processed tRNAs, the measured absolute tRNA concentrations were either used without any corrections or corrected by taking into account only one of these mechanisms. Furthermore, ribosomal transition rates for tRNA accommodation, peptide bond formation and translocation that were measured *in vitro* were not adequately adapted to *in vivo* translation, although it is known that these rates are strongly influenced by buffer conditions [30, 36]. In addition, the impeding competition of cognate with near- and non-cognate TCs for ribosomes is often neglected, leading to an overestimate of the elongation rates [33].

Our theory on translation elongation takes into account all of the aforementioned mechanisms that influence the concentrations of TCs, as well as the competition of TCs,

and a systematical mapping of measured *in vitro* ribosomal kinetics to protein synthesis *in vivo*.

1.5 List of Publications

Research performed in the course of this thesis was partly published in the following peer-reviewed journal publications:

- Sophia Rudorf, Michael Thommen, Marina V. Rodnina, and Reinhard Lipowsky. Deducing the kinetics of protein synthesis *in vivo* from the transition rates measured *in vitro*. *PLoS Computational Biology*, 2014.
- Carlus Deneke, Sophia Rudorf, Angelo Valleriani. Transient Phenomena in Gene Expression after Induction of Transcription. *PLoS One*, 7(4):e35044, 2012.

2 Brief Review of Continuous-Time Markov Processes

In this chapter, we introduce the mathematical tools that are necessary to develop our theoretical description of translation elongation. Translation is a stochastic process governed by thermal fluctuations and diffusive motions, as well as stochastic chemical reactions and mechanical displacements of its protagonists. We describe translation elongation as a Markov process to capture its stochastic nature. Markov processes are common in stochastic modeling of biochemical processes and have been extensively described in the literature, e.g., by Taylor and Karlin [37]. Therefore, in this chapter we confine ourselves to first briefly give a general introduction to Markov processes and then shortly explicate a method to calculate absorption probabilities and mean first passage times.

2.1 Markov Process Description of Biochemical Systems

When a biochemical process is studied experimentally, usually several distinct states of the system of interest are found. These states may, for example, represent different molecular conformations or various steps of a chemical reaction. How many states are observed and how they are distinguished depends on the specific experimental setup and its resolution.

These distinct states define the discrete *state space* of the biochemical process. They are connected by molecular *transitions* such as association and dissociation of molecules, chemical reactions, or mechanical transitions. These transitions occur with certain *transition rates* which can generally as well be experimentally determined.

Often, the state space and the possible transitions between the different states of a process are represented by a directed graph. Each node of the graph corresponds to a state in the state space and directed edges indicate transitions between two states with finite transition rates. A graph representation for a Markov process with three distinct states is shown in fig. 2.1 (A).

A series of states, that are visited by the Markov process over time after starting in a certain initial state, is called a *trajectory*. In other words, a trajectory indicates the

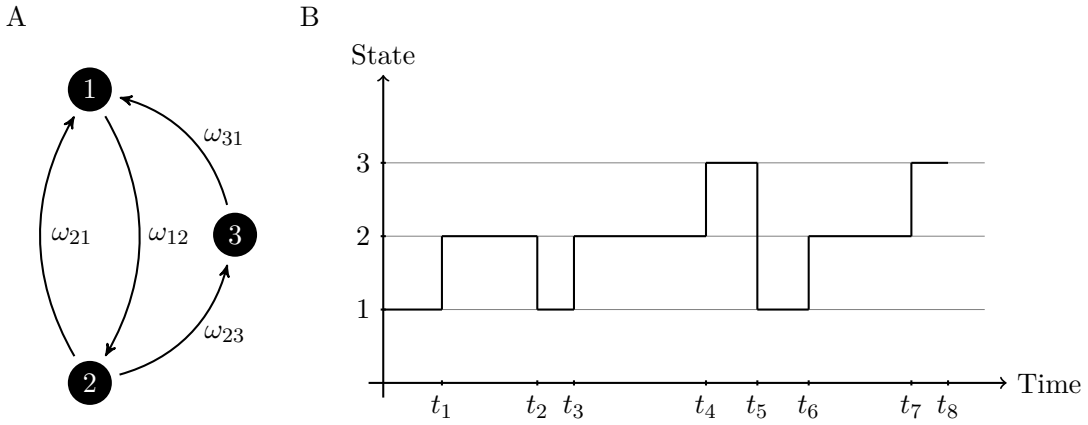


Figure 2.1: **Example of a three-state Markov process.** (A) Graph representation of a three-state Markov process. The rate of transitions from state i to state j is denoted by ω_{ij}. (B) Example of a trajectory. The process starts at $t = 0$ in state 1, dwells in state 1 until $t = t_1$ and makes a transition to state 2. It dwells in state 2, before the process returns to state 1 at $t = t_2$. After dwelling in state 1, another transition to state 2 occurs at $t = t_3$. This time, the process makes a transition to state 3 at $t = t_4$. From state 3, the process returns to state 1 at $t = t_5$, and so on.

state of the process at time t. An individual trajectory of the simple three-state Markov process in fig. 2.1 (A) is shown in fig. 2.1 (B). Due to the stochastic nature of a Markov process, individual trajectories generally differ from each other.

A Markov process can be described by the probabilities $\mathrm{P}_i(t)$ to find the process in state i at time t. The time evolution of these probabilities $\mathrm{P}_i(t)$ is completely determined by the initial state and the transition rates ω_{ij} and can be written in form of *loss-and-gain* equations

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\mathrm{P}_i(t) = \sum_j \left(\mathrm{P}_j(t)\,\omega_{ji} - \mathrm{P}_i(t)\,\omega_{ij}\right) . \tag{2.1}$$

In the physical literature, these loss-and-gain equations are usually called *master equations*. From the master equations it becomes obvious that the Markov process description of a biochemical system incorporates the experimentally gained knowledge on states and transition rates and that no further information is needed.

2.2 Infinitesimal Probabilities and Master Equations

In this section, we introduce the Markov process from a more formal point of view. In particular, we briefly summarize how the master equations of a Markov process are derived from the transition rates.

We consider a Markov process with a discrete state space and continuous-time dynamics. The state space $\mathbf{S} \equiv \{\sigma_1, \sigma_2, ..., \sigma_N\}$ consists of N discrete states $\sigma = \sigma_1, \sigma_2, ..., \sigma_N$. The state that the process has attained at time t is a stochastic variable $X(t)$ and the probability to find the process in state σ_j is expressed as $\mathrm{P}_j(t) \equiv Pr\{X(t) = \sigma_j\}$. A Markov process is a stochastic process without memory. This means that the probability $\mathrm{P}_j(t)$ depends only on the last known state that the process is in before or at time t:

$$\begin{aligned} \mathrm{P}_j(t) &= Pr\{X(t) = \sigma_j | X(t_T) = \sigma_i,\, X(t_{T-1}) = \sigma_h,\, ...,\, X(t_1) = \sigma_a\} \\ &= Pr\{X(t) = \sigma_j | X(t_T) = \sigma_i\}\,, \end{aligned} \tag{2.2}$$

where $Pr\{A|B\}$ is the probability of event A under the condition of event B, $0 \leq t_1 \leq ... \leq t_{T-1} \leq t_T \leq t$ is a series of points in time, and $\sigma_a, ..., \sigma_h, \sigma_i \in \mathbf{S}$. The time evolution of a Markov process is defined by the *infinitesimal probabilities* [37]:

$$Pr\{X(t+\tau) = \sigma_j | X(t) = \sigma_i\} = \omega_{ij}\tau + \mathrm{o}(\tau)\,, \tag{2.3}$$

$$Pr\{X(t+\tau) = \sigma_i | X(t) = \sigma_i\} = 1 - \sum_k \omega_{ik}\tau + \mathrm{o}(\tau)\,, \tag{2.4}$$

where $\sigma_i, \sigma_j \in \mathbf{S}$, ω_{ij} is the rate of transition from state σ_i to state σ_j with $\omega_{ii} = 0$, τ represents an infinitesimal short time scale, and $\mathrm{o}(\tau)$ contains all higher order terms of τ. Note that the transition rates ω_{ij} are related to the average time $\langle s_i \rangle$ that a Markov process dwells in state σ_i and to the time-independent transition probabilities π_{ik} of proceeding from state σ_i to state σ_k

$$\langle s_i \rangle = \frac{1}{\sum_k \omega_{ik}}\,, \tag{2.5}$$

$$\pi_{ik} = \langle s_i \rangle\, \omega_{ik}\,, \tag{2.6}$$

where the summation runs over all transitions from state σ_i to any other state. For a Markov process, the dwell times s_i are exponentially distributed [37]

$$Pr\{s_i \leq t\} = \int_0^t \frac{1}{\langle s_i \rangle} \exp\left[-t'/\langle s_i \rangle\right] \mathrm{d}t' = 1 - \exp\left[-t/\langle s_i \rangle\right]\,. \tag{2.7}$$

Figure 2.2: **Markov process with absorbing states.** States 1 and 4 are absorbing states, whereas states 2 and 3 are transient states. The rate of transitions from state i to state j is denoted by ω_{ij}.

The master equations of a Markov process can be derived from the law of total probability

$$\begin{aligned} \mathrm{P}_j(t+\tau) &= Pr\{X(t+\tau) = \sigma_j | X(0) = x_0\} \\ &= \sum_k Pr\{X(t+\tau) = \sigma_j | X(t) = \sigma_k,\, X(0) = x_0\}\, \mathrm{P}_k(t) \end{aligned} \tag{2.8}$$

by exploiting the Markov property and the definitions (2.3) and (2.4) of the infinitesimal probabilities:

$$\begin{aligned} \mathrm{P}_j(t+\tau) &= \sum_k Pr\{X(t+\tau) = \sigma_j | X(t) = \sigma_k\}\, \mathrm{P}_k(t) \\ &= \sum_{k \neq j} \left(\mathrm{P}_k(t)\, \omega_{kj} - \mathrm{P}_j(t)\, \omega_{jk}\right) \tau + \mathrm{P}_j(t) + \mathrm{o}\left(\tau\right) . \end{aligned} \tag{2.9}$$

In the limit of $\tau \to 0$, eq. (2.9) can be transformed into

$$\frac{\mathrm{d}}{\mathrm{dt}} \mathrm{P}_j(t) = \sum_{k \neq j} \left(\mathrm{P}_k(t)\, \omega_{kj} - \mathrm{P}_j(t)\, \omega_{jk}\right) , \tag{2.10}$$

with the initial condition $X(0) = x_0$.

2.3 Absorbing States

A state of a Markov process is an *absorbing state* if it can be reached from other states of the process but cannot be left. Examples of absorbing states are given by states 1 and 4 in fig. 2.2, whereas states 2 and 3 are *transient* states. In this section, we shortly summarize how the probabilities to absorb in a specific absorbing state are computed and how absorbing states can be used to obtain mean first passage times.

2.3.1 Absorption Probabilities

If a Markov process has more than one absorbing state, the probability of absorption in one particular absorbing state can be computed by a method often called *first step analysis* [37]: The process starts in the initial state σ_i, which is a transient state as, for example, states 2 and 3 in fig. 2.2. Then, with probability π_{ij} (2.6) the process makes a first transition to the next state σ_j. From here, a second transition to state σ_k occurs with probability π_{jk}, and so on.

In summary, the probability u_{if} to reach the absorbing state σ_f from state σ_i depends only on the time-independent transition probabilities π_{ij} from the initial state σ_i to all other states σ_j (2.6) and the absorption probabilities u_{jf} of absorption in σ_f from the states σ_j [37]. For a Markov process with N transient states, a set of N equations has to be solved to compute the absorption probabilities u_{if}. Each equation is of the form

$$u_{if} = \sum_j \pi_{ij}\, u_{jf} \quad \text{for all } i \neq f, \text{ and } u_{ff} = 1. \tag{2.11}$$

Note that by definition $u_{if} = 0$ if σ_i is an absorbing state itself and $i \neq f$. The concept of absorption probabilities can also be applied to calculate the probability that, starting from state σ_i, a specific state σ_f of a Markov process is visited before another state σ_j is attained. To do so, one simply has to consider an auxiliary Markov process in which σ_j and σ_f are absorbing states and then compute the probability u_{if} (2.11) of absorption in state σ_f.

2.3.2 Mean First Passage Times

The mean first passage time ν_{if} is the average time that it takes a process to visit state σ_f after starting in state σ_i. It can be computed using first step analysis [37]. First, one has to define an auxiliary Markov process where σ_f is the only absorbing state. More precisely, in the auxiliary process all transitions from state σ_f to other states are removed and only the subset of states from where state σ_f can be reached is considered. This auxiliary Markov process dwells in the initial state σ_i with an average dwell time $\langle s_i \rangle$. Then, with probability π_{ij} (2.6) it makes a first transition to the next state σ_j. Here, the process dwells again with an average dwell time $\langle s_j \rangle$ before the next transition occurs. This is repeated until the final state σ_f is reached.

Therefore, the mean first passage time ν_{if} is simply the sum of the average dwell times of all states weighted with the probabilities to visit these states prior to σ_f. In particular, the mean first passage time ν_{if} can be expressed implicitly by the average dwell time $\langle s_i \rangle$ in state σ_i (2.5), the time-independent transition probabilities π_{ij} from state σ_i to

state σ_j (2.6), and the mean first passage times ν_{jf} of all states σ_j. For an auxiliary Markov process with a state space composed of N transient states, a set of N equations has to be solved, whereof each is of the form

$$\nu_{if} = \langle s_i \rangle + \sum_j \pi_{ij}\, \nu_{jf} \quad \text{for all } i \neq f, \text{ and } \nu_{ff} = 0. \tag{2.12}$$

3 Theoretical Framework for Translation Elongation

In this chapter, we will develop a comprehensive theoretical framework of translation elongation that will shed light on the intricate relationships between the internal dynamics of ribosomes, the life cycle of tRNAs, and the codon composition of mRNAs.

3.1 Codons and tRNAs

We denote by $\mathbf{C}$ the set of the 61 sense codons which are labeled by $c = 1, 2, ... 61$. The different species of tRNA form the set $\mathbf{A}$. These tRNA species will be distinguished by the label $a = 1, 2, ... |\mathbf{A}|$. The total number $|\mathbf{A}|$ of different tRNA species depends on the organism. In *E. coli*, there are $|\mathbf{A}| = 43$ different elongator tRNA molecules [3]. For each codon c, there is a set $\mathbf{A}_{\rm co}(c)$ of cognate tRNAs. All other tRNAs belong either to the set of near-cognate $\mathbf{A}_{\rm nr}(c)$ or non-cognate tRNAs $\mathbf{A}_{\rm no}(c)$ of the codon c. Therefore, each codon $c \in \mathbf{C}$ leads to a unique partition $\{\mathbf{A}_{\rm co}(c), \mathbf{A}_{\rm nr}(c), \mathbf{A}_{\rm no}(c)\}$ of the set $\mathbf{A}$ such that

$$\mathbf{A} = \mathbf{A}_{\rm co}(c) \cup \mathbf{A}_{\rm nr}(c) \cup \mathbf{A}_{\rm no}(c) \quad \text{for all } c \in \mathbf{C}. \tag{3.1}$$

In some cases, a certain tRNA species is cognate to more than one codon. The set of codons that are cognate to the tRNA with index a is denoted by $\mathbf{C}_{\rm co}(a)$, whereas the near- and non-cognate codons are contained in the complementary sets $\mathbf{C}_{\rm nr}(a)$ and $\mathbf{C}_{\rm no}(a)$, respectively. Therefore, each tRNA $a \in \mathbf{A}$ leads to a unique partition $\{\mathbf{C}_{\rm co}(a), \mathbf{C}_{\rm nr}(a), \mathbf{C}_{\rm no}(a)\}$ of the set $\mathbf{C}$ such that

$$\mathbf{C} = \mathbf{C}_{\rm co}(a) \cup \mathbf{C}_{\rm nr}(a) \cup \mathbf{C}_{\rm no}(a) \quad \text{for all } a \in \mathbf{A}. \tag{3.2}$$

An overview of all cognate, near- and non-cognate tRNAs and codons is given in fig. 1.2 for *E. coli*.

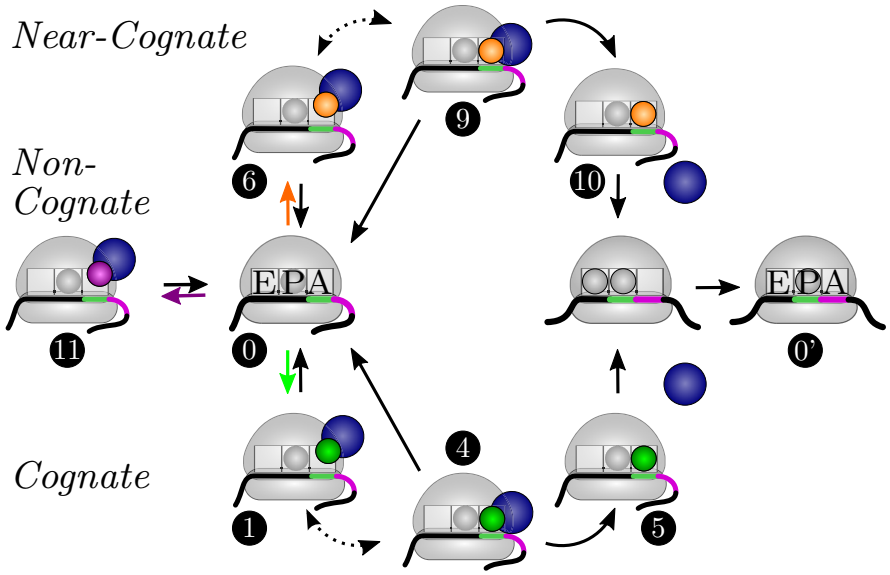


Figure 3.1: **Translation elongation cycle.** The ribosome has three tRNA binding sites, the A, P, and E site. A ribosome that has just arrived at a new (green) codon of a mRNA has an empty A site, whereas the P site is occupied by a tRNA (here shown as gray sphere) that is cognate or near-cognate to the preceding codon. Elongation factor EF-Tu (blue spheres) and tRNAs form ternary complexes (TCs). Cognate, near-cognate and non-cognate TCs bind to the ribosome with concentration dependent rates (green, orange and purple arrows), respectively. Since the initial binding is not codon-specific, all kinds of TCs unbind again from the ribosome with the same dissociation rate. Alternatively, a cognate or near-cognate TC can be recognized by the ribosome before it is either completely released, brought back to the initial binding state, or its tRNA is accommodated in the ribosomal A site. Along with tRNA accommodation, EF-Tu leaves the ribosome. The new A site tRNA is then further processed and shifted to the P site, while the ribosome translocates to the next (purple) codon. The former P site tRNA is now in the E site. Depending on the assumed pathway of tRNA release, the E site tRNA either very rapidly dissociates from the ribosome (2-1-2 pathway), or stays until the next tRNA has accommodated in the ribosomal A site (2-3-2 pathway). Numbers indicate corresponding states of the codon-specific Markov process shown in fig. 3.2.

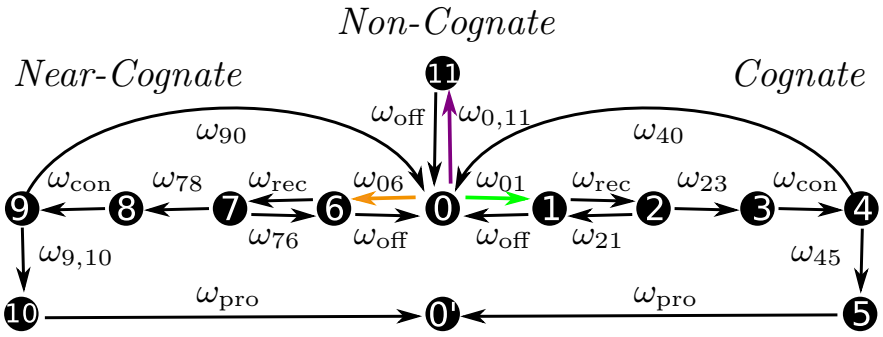


Figure 3.2: **Representation of translation elongation as a Markov process.** Each of the substeps of the elongation cycle is represented by one state in the Markov model, which leads to twelve states per sense codon. State $(c|0)$ indicates the state assumed by a ribosome reading codon c when it is not bound to a TC. State $(c'|0')$ is attained by a ribosome after translocation to the next codon c'. All rates of transitions between states are taken to be codon-independent, except for the binding rates of cognate, near-, and non-cognate TCs (green, orange and purple arrows).

3.2 Theoretical Description of the Elongation Cycle

Our description of translation elongation as depicted in fig. 3.1 is based on the chemical kinetics as determined for *in vitro* assays by Marina Rodnina and her co-workers over the past decades [38–44]. We model translation as a continuous-time Markov process by representing each substep of the elongation cycle with one state in the Markov model as shown in fig. 3.2. In particular, the state of a translating ribosome is characterized by the codon in its A site and by the binding of ternary complexes (TCs) or tRNAs to the ribosomal A, P, and E sites. A ribosome that has just moved to a codon c has a free A site, but its P site is occupied by a tRNA that is cognate or near-cognate to the preceding codon. The corresponding ribosomal state is denoted by $(c|0)$.

When a TC binds to the ribosome, it can be cognate, near-cognate, or non-cognate to the codon in the ribosomal A site. If a cognate TC binds, the ribosome is in state $(c|1)$; the binding of a near-cognate TC leads to state $(c|6)$; if the TC is non-cognate to c, the ribosome is in state $(c|11)$. A non-cognate TC always dissociates from the ribosome which then goes back to state $(c|0)$. After initial binding, a cognate TC has to be recognized by the ribosome, which happens in state $(c|2)$. The recognition is followed by an activation of GTPase and GTP hydrolysis, leading to ribosomal state $(c|3)$, and release of phosphate and rearrangements of the EF-Tu molecule during transition to state $(c|4)$. From here, usually the tRNA gets accommodated and the EF-Tu is released, leading

to state $(c|5)$. However, in rare cases the TC is released from the ribosome which then returns to state $(c|0)$. We will call the states $(c|1)$ to $(c|5)$ the cognate branch of the Markov process.
Besides the cognate branch, the Markov process also involves an analogous near-cognate branch of another five states: a near-cognate TC is recognized in state $(c|7)$ and gets usually rejected during this initial selection step, bringing the ribosome back to state $(c|0)$. Alternatively, GTPase is activated, GTP is hydrolyzed and the ribosome/near-cognate TC complex transits to state $(c|8)$. After release of phosphate and EF-Tu rearrangements during transition to state $(c|9)$, the ribosome usually rejects a near-cognate TC which is why this step is called *proofreading*. If not, the ribosome/near-cognate TC complex moves on to state $(c|10)$, where the near-cognate tRNA gets accommodated in the ribosomal A site and EF-Tu leaves the ribosome, resulting in a misreading error.
After cognate or near-cognate tRNA accommodation, the A site tRNA is processed which includes peptide bond formation and translocation of the ribosome to the next codon along with a transfer of the A site tRNA to the P site and of the P site tRNA to the E site. A new elongation cycle on the following codon c' can begin, indicated by the ribosomal state $(c'|0')$ in fig. 3.2.
Processed tRNAs leave the ribosome from the E site. Afterwards, they bind to synthetases that recharge them with amino acids, before the recharged tRNAs bind again to EF-Tu to form new TCs, see fig. 3.3.

Note that some transitions from one state of the described Markov process to another state are reversible, which is depicted in fig. 3.2 by pairs of opposing arrows. In contrast, several other transitions are taken to be irreversible as indicated by single arrows. These assumptions of irreversibility are based on experimental findings by Marina Rodnina and her co-workers, i.e., these transitions are irreversible within the limits of experimental resolution.

We restrict our theoretical analysis to elongation under low densities of translating ribosomes on mRNAs, i.e., elongation without ribosome-ribosome interactions. However, when translation is studied by stochastic simulations, possible interactions of ribosomes will be taken into account, see chapter B. In fact, experimental findings suggest that under physiological conditions most ribosomes do not interfere [45].

3.3 Binding Rates and Internal Transition Rates

A ribosomal transition rate from state i to state j will be denoted by ω_{ij}. Accordingly, the rate ω_{01} governs the transition from $(c|0)$ to $(c|1)$. The rates ω_{06} and $\omega_{0,11}$ govern the transitions from $(c|0)$ to $(c|6)$ and $(c|11)$, respectively. The rates ω_{01}, ω_{06}, and $\omega_{0,11}$ depend on the molar concentrations $\hat{\mathrm{X}}_a$ of available, free TCs containing tRNA species

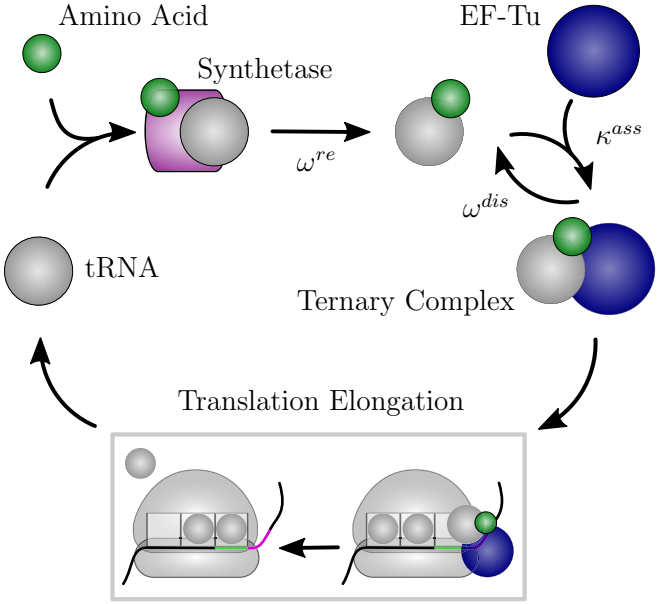


Figure 3.3: **Recharging cycle of tRNAs.** During translation elongation a tRNA delivers its amino acid to the elongating peptide chain. After the processed tRNA is released from a ribosomal E site, it binds to an aminoacyl tRNA synthetase. The synthetase recharges the tRNA with a new amino acid, and the recharged tRNA binds to elongation factor EF-Tu to form a new ternary complex.

a and have the form

$$\omega_{01} = \kappa_{\text{on}} \sum_{a \in \mathbf{A}_{\text{co}}(c)} \hat{\text{X}}_a \,, \tag{3.3}$$

$$\omega_{06} = \kappa_{\text{on}} \sum_{a \in \mathbf{A}_{\text{nr}}(c)} \hat{\text{X}}_a \,, \tag{3.4}$$

and

$$\omega_{0,11} = \kappa_{\text{on}} \sum_{a \in \mathbf{A}_{\text{no}}(c)} \hat{\text{X}}_a \,, \tag{3.5}$$

where κ_{on} is the binding rate constant that we take to be identical for all (cognate, near-cognate, and non-cognate) TCs in agreement with experimental observations [46].

In general, transition rates in the cognate branch of the Markov process are different from their counterparts in the near-cognate branch. As an exception to this rule, based on experimental findings [38, 41] we assume the transition rates to state $(c|2)$ or $(c|7)$ and the transition rates to state $(c|4)$ or $(c|9)$ to be identical for all cognate and near-cognate TCs. According to our experimental collaborator Marina Rodnina from the Max Planck Institute for Biophysical Chemistry, the same assumption holds for the rate of transition to the free state after translocation $(c'|0')$

$$\omega_{12} = \omega_{67} \equiv \omega_{\text{rec}} \,, \tag{3.6}$$

$$\omega_{34} = \omega_{89} \equiv \omega_{\text{con}} \,, \tag{3.7}$$

$$\omega_{50'} = \omega_{10,0'} \equiv \omega_{\text{pro}} \,. \tag{3.8}$$

Furthermore, the dissociation rates from state $(c|1)$, $(c|6)$, or $(c|11)$ back to state $(c|0)$ are taken to be identical for all cognate, near-cognate, and non-cognate TCs

$$\omega_{10} = \omega_{60} = \omega_{11,0} \equiv \omega_{\text{off}} \,, \tag{3.9}$$

because dissociation happens before a physical contact between the A site codon and the anti-codon of the tRNA is made and, thus, a distinction between cognate, near-cognate, and non-cognate TCs is possible. In summary, the 20 transitions within the Markov process defined above can be characterized by three codon dependent binding rates ω_{01}, ω_{06}, and $\omega_{0,11}$, and 12 codon independent internal transition rates, see fig. 3.2.

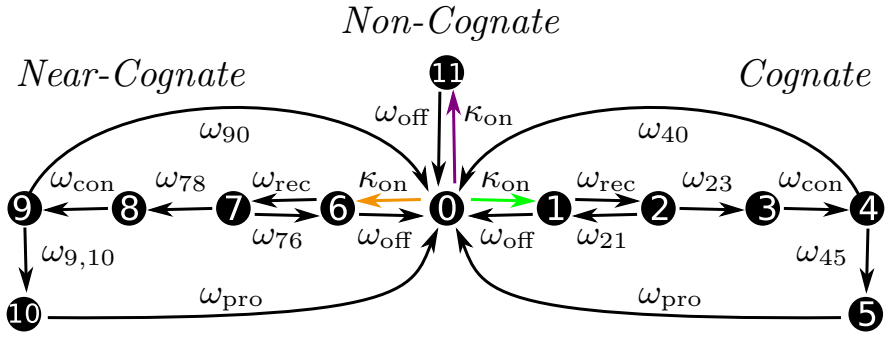


Figure 3.4: **Auxiliary Markov process for the computation of dwell times.** The auxiliary Markov process is almost identical to the Markov process of codon-specific elongation depicted in fig. 3.2, but the absorbing state $(c'|0')$ is omitted. The states $(c|5)$ and $(c|10)$ are directly connected back to the initial state $(c|0)$ with the same transition rate ω_{pro} as in the original process.

3.4 Codon-Specific Elongation Rates and Dwell Times

We define the elongation time $\mathrm{t}_{c,\text{elo}}$ of codon c as the average time that is needed to finish a complete elongation cycle on codon c. According to the Markov process described in the previous sections, this time is the average time to absorption in state $(c'|0')$ from state $(c|0)$. We define the codon-specific elongation rate $\omega_{c,\text{elo}}$ as the inverse of the elongation time $\mathrm{t}_{c,\text{elo}}$. The codon-specific elongation time can be computed as a mean first passage time by first step analysis as is described in section 2.3.2. Alternatively, we can express the elongation time by the sum of the average dwell times $\mathrm{t}_{(c|i)}$ in states $(c|i)$ with $i = 0, 1, \ldots, 11$ per elongation cycle on codon c. This leads to

$$\omega_{c,\text{elo}} \equiv \mathrm{t}_{c,\text{elo}}^{-1} \equiv \left(\sum_{i=0}^{11} \mathrm{t}_{(c|i)} \right)^{-1} . \tag{3.10}$$

The dwell times $\mathrm{t}_{(c|i)}$ are determined using a method introduced by Terrell L. Hill [47]. We consider an auxiliary Markov process that is almost identical to the process depicted in fig. 3.2 with one modification: In the auxiliary process, the absorbing state $(c'|0')$ is omitted, see fig. 3.4. Instead of the transitions $(c|5)$ to $(c'|0')$ and $(c|10)$ to $(c'|0')$, the states $(c|5)$ and $(c|10)$ are connected back to the initial state $(c|0)$ using the same rate ω_{pro} as in the original process. In other words, the state $(c'|0')$ is identified with the state $(c|0)$. Thus, the states $(c|i)$ are decoupled from all other states that the ribosome visited prior to codon c under consideration. Such a decoupling is possible because the Markov process has no memory. The time dependence of the probabilities $\mathrm{P}_{c,i}^{(\mathrm{a})}(t)$ to attain the

states $(c|i)$ in the auxiliary process are described by a set of master equations

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\mathrm{P}^{(\mathrm{a})}_{c,i}(t) = \sum_j \left(\mathrm{P}^{(\mathrm{a})}_{c,j}(t)\,\omega_{ji} - \mathrm{P}^{(\mathrm{a})}_{c,i}(t)\,\omega_{ij}\right)\,, \tag{3.11}$$

where the transition rates ω_{ij} are defined in the previous section and fig. 3.4. Furthermore, the time dependent probabilities $\mathrm{P}^{(\mathrm{a})}_{c,i}(t)$ fulfill the normalization condition

$$\sum_i \mathrm{P}^{(\mathrm{a})}_{c,i}(t) = 1\,. \tag{3.12}$$

The dwell times $\mathrm{t}_{(c|i)}$ can be calculated from the steady state of the auxiliary Markov process. In steady state, the left side of eq. (3.11) vanishes and this set of equations together with eq. (3.12) can be solved for $\mathrm{P}^{(\mathrm{a})}_{c,0}$, $\mathrm{P}^{(\mathrm{a})}_{c,1}$, …, $\mathrm{P}^{(\mathrm{a})}_{c,11}$ where $\mathrm{P}^{(\mathrm{a})}_{c,i}$ denotes the steady state, i.e., $\mathrm{P}^{(\mathrm{a})}_{c,i} \equiv \lim_{t\to\infty} \mathrm{P}^{(\mathrm{a})}_{c,i}(t)$. The steady state probabilities $\mathrm{P}^{(\mathrm{a})}_{c,i}$ of the auxiliary process are identical to the fractions of time that the original process spends in average in the states $(c|i)$ with initial state $(c|0)$ and absorbing state $(c'|0')$ [47], and, thus, yield the dwell times $\mathrm{t}_{(c|i)}$

$$\mathrm{P}^{(\mathrm{a})}_{c,i} = \frac{\mathrm{t}_{(c|i)}}{\mathrm{t}_{c,\mathrm{elo}}}\,, \tag{3.13}$$

with the codon-specific elongation time $\mathrm{t}_{c,\mathrm{elo}}$ defined in eq. (3.10).

It is convenient to express the resulting dwell times in terms effective cognate and near-cognate accommodation rates

$$\omega_{c,\mathrm{co}} \equiv \omega_{01}\,\frac{\pi_{12}\pi_{23}\pi_{45}}{1-\pi_{12}\pi_{21}}\,, \tag{3.14}$$

$$\omega_{c,\mathrm{nr}} \equiv \omega_{06}\,\frac{\pi_{67}\pi_{78}\pi_{9,10}}{1-\pi_{67}\pi_{76}}\,. \tag{3.15}$$

with the probability $\pi_{ij} = \omega_{ij}\,/\sum_k \omega_{ik}$ (2.6) of transition from state $(c|i)$ to state $(c|j)$. Note that the effective accommodation rates are equivalent to cognate or near-cognate binding rates ω_{01} and ω_{06} weighted by the respective probabilities that the tRNA of a TC that has bound to a ribosome successfully accommodates in the A site instead of dissociating from the ribosome before accommodation has happened. These probabilities can be computed by first step analysis, see section 2.3.1.
The dwell time $\mathrm{t}_{(c|0)}$ that the ribosome spends in the free state $(c|0)$ during one complete

elongation cycle at codon c is then given by

$$\mathrm{t}_{(c|0)} = \frac{1}{\omega_{c,\mathrm{co}} + \omega_{c,\mathrm{nr}}} \,. \tag{3.16}$$

The dwell time $\mathrm{t}_{(c|11)}$ with a bound non-cognate TC is

$$\mathrm{t}_{(c|11)} = \mathrm{t}_{(c|0)} \, \frac{\omega_{0,11}}{\omega_{\mathrm{off}}} \,, \tag{3.17}$$

where the dissociation rate ω_{off} is defined in eq. (3.9).
The dwell times $\mathrm{t}_{(c|1)}$ to $\mathrm{t}_{(c|5)}$ in states of the cognate branch can be expressed by

$$\mathrm{t}_{(c|1)} = \mathcal{P}_{c,\mathrm{co}} \, \frac{1}{\omega_{\mathrm{rec}}} \frac{1}{\pi_{23}} \frac{1}{\pi_{45}} \,, \tag{3.18}$$

$$\mathrm{t}_{(c|2)} = \mathcal{P}_{c,\mathrm{co}} \, \frac{1}{\omega_{23}} \frac{1}{\pi_{45}} \,, \tag{3.19}$$

$$\mathrm{t}_{(c|3)} = \mathcal{P}_{c,\mathrm{co}} \, \frac{1}{\omega_{\mathrm{con}}} \frac{1}{\pi_{45}} \,, \tag{3.20}$$

$$\mathrm{t}_{(c|4)} = \mathcal{P}_{c,\mathrm{co}} \, \frac{1}{\omega_{45}} \,, \tag{3.21}$$

$$\mathrm{t}_{(c|5)} = \mathcal{P}_{c,\mathrm{co}} \, \frac{1}{\omega_{\mathrm{pro}}} \,, \tag{3.22}$$

where $\mathcal{P}_{c,\mathrm{co}}$ is the probability of cognate accommodation

$$\mathcal{P}_{c,\mathrm{co}} \equiv \frac{\omega_{c,\mathrm{co}}}{\omega_{c,\mathrm{co}} + \omega_{c,\mathrm{nr}}} \,, \tag{3.23}$$

and ω_{rec}, ω_{con}, and ω_{pro} are defined in eqs. (3.6) to (3.8). Likewise, the dwell times spend in near-cognate states $(c|6)$ to $(c|10)$ are given by

$$\mathrm{t}_{(c|6)} = \mathcal{P}_{c,\mathrm{nr}} \, \frac{1}{\omega_{\mathrm{rec}}} \frac{1}{\pi_{78}} \frac{1}{\pi_{9,10}} \,, \tag{3.24}$$

$$\mathrm{t}_{(c|7)} = \mathcal{P}_{c,\mathrm{nr}} \, \frac{1}{\omega_{78}} \frac{1}{\pi_{9,10}} \,, \tag{3.25}$$

$$\mathrm{t}_{(c|8)} = \mathcal{P}_{c,\mathrm{nr}} \, \frac{1}{\omega_{\mathrm{con}}} \frac{1}{\pi_{9,10}} \,, \tag{3.26}$$

$$\mathrm{t}_{(c|9)} = \mathcal{P}_{c,\mathrm{nr}} \, \frac{1}{\omega_{9,10}} \,, \tag{3.27}$$

$$\mathrm{t}_{(c|10)} = \mathcal{P}_{c,\mathrm{nr}} \, \frac{1}{\omega_{\mathrm{pro}}} \,, \tag{3.28}$$

with the near-cognate accommodation probability

$$\mathcal{P}_{c,\mathrm{nr}} \equiv \frac{\omega_{c,\mathrm{nr}}}{\omega_{c,\mathrm{co}} + \omega_{c,\mathrm{nr}}} \,. \tag{3.29}$$

The probabilities of cognate and near-cognate accommodation $\mathcal{P}_{c,\mathrm{co}}$ and $\mathcal{P}_{c,\mathrm{nr}}$ fulfill the normalization condition $\mathcal{P}_{c,\mathrm{co}} + \mathcal{P}_{c,\mathrm{nr}} = 1$. The accommodation probability $\mathcal{P}_{c,\mathrm{co}}$ is a measure of the codon-specific fidelity: if $\mathcal{P}_{c,\mathrm{co}} = 1$, the ribosome translates codon c without any errors and, thus, with maximal fidelity. Likewise, $\mathcal{P}_{c,\mathrm{nr}}$ represents the codon-specific infidelity, which is also named *near-cognate missense error frequency*.

3.5 Codon Usages and Overall Elongation Rate

One quantity that has been determined experimentally for *E. coli* under various growth conditions is the average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$. It is an average of the codon-specific elongation times $\mathrm{t}_{c,\mathrm{elo}}$ of all codons c that are translated in the cell. This average involves the codon usages p_c, which represent the normalized frequencies or probabilities that a randomly chosen sense codon in the genome is equal to c. The quantities p_c satisfy the obvious relations

$$0 \leq \mathrm{p}_c \leq 1 \quad \text{and} \quad \sum_{c=1}^{61} \mathrm{p}_c = 1 \,. \tag{3.30}$$

The average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$ and its inverse, the overall elongation rate ω_{elo}, are then given by

$$\omega_{\mathrm{elo}} \equiv \langle \mathrm{t}_{c,\mathrm{elo}} \rangle^{-1} = \left(\sum_{c=1}^{61} \mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}} \right)^{-1} \,. \tag{3.31}$$

3.6 Probabilities to Attain Ribosomal States

In a cell or in an *in vitro* translation assay, many mRNAs may exist at the same time with a total number of M molecules. In general, these mRNAs are composed of different codon sequences, although some of them might encode the same protein and, thus, have the same sequence of codons. To identify the state space of the translation process, we first consider the translation of an individual mRNA m of length L_m that is composed of a sequence of codons c_1, c_2, c_3,... c_{L_m}, see fig. 3.5. For the translation process of this

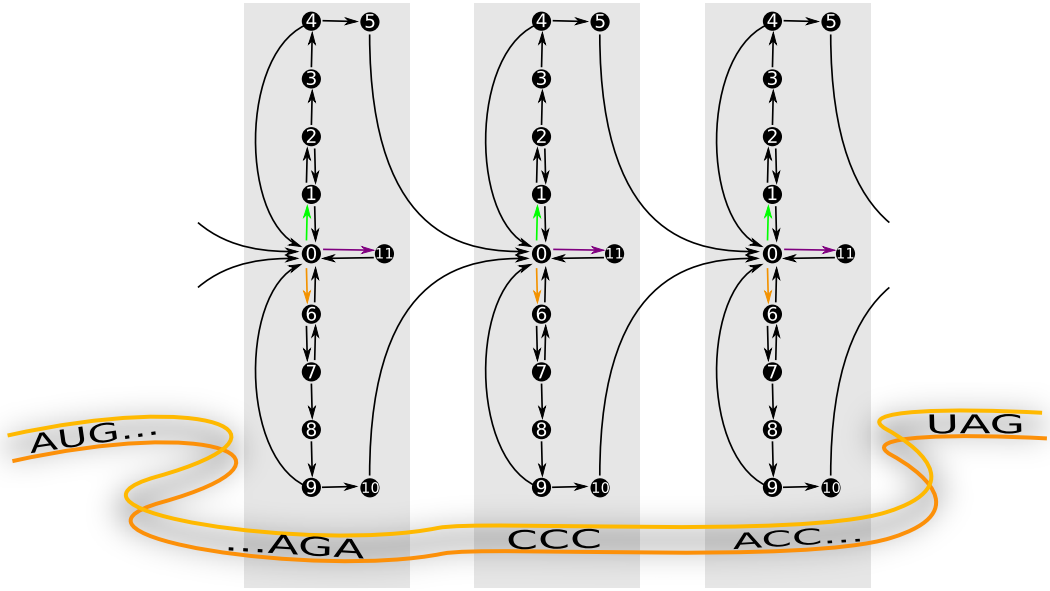


Figure 3.5: **Markov process for the translation of an individual mRNA.** The mRNA (orange band) consists of a sequence of codons, where in this example AUG and UAG indicate the start and the stop codon, respectively. The codons are translated sequentially from start to stop. Hence, the Markov process for the translation of a mRNA is given by the concatenation of individual codon-specific Markov processes as defined in section 3.2 and fig. 3.2.

mRNA, the ribosomal states described in section 3.2 form the state space

$$\mathbf{S}_m \equiv \bigcup_{x,i} \{(c_x|i)\} \tag{3.32}$$

where $1 \leq x \leq L_m$ refers to the position of the codon $c_x \in \mathbf{C}$ and $i = 0, 1, \ldots, 11$ indicates the internal state of the ribosome as introduced in section 3.2. Thus, the state space $\mathbf{S}_m$ consists of $12 \times L_m$ states.
For each state $(c_x|i) \in \mathbf{S}_m$, we can define the time dependent probability $\mathrm{P}_{c_x,i}(t)$ that a ribosome dwells on codon c_x and attains ribosomal state i at time t. For the Markov process described above and in fig. 3.5, the time dependent probabilities fulfill the following master equations

$$\frac{\mathrm{d}}{\mathrm{dt}} \mathrm{P}_{c_x,0}(t) = \sum_{j=1,6,11} (\mathrm{P}_{c_x,j}(t)\, \omega_{\mathrm{off}} - \mathrm{P}_{c_x,0}(t)\, \omega_{0j}) + \sum_{j=5,10} \mathrm{P}_{c_{x-1},j}(t)\, \omega_{\mathrm{pro}}\,, \tag{3.33}$$

for the free states $(c_x|0)$ and $x \geq 2$, and for all other states

$$\frac{\mathrm{d}}{\mathrm{dt}} \mathrm{P}_{c_x,i}(t) = \sum_j (\mathrm{P}_{c_x,j}(t)\, \omega_{ji} - \mathrm{P}_{c_x,i}(t)\, \omega_{ij})\,, \tag{3.34}$$

where the transition rates ω_{ij} are defined in the previous section and fig. 3.2. Furthermore, the time dependent probabilities $\mathrm{P}_{c_x,i}(t)$ fulfill the normalization condition

$$\sum_{x,i} \mathrm{P}_{c_x,i}(t) = 1\,. \tag{3.35}$$

A full Markov model describing translation elongation on all mRNAs in the entire cell or in an *in vitro* translation assay has a state space $\mathbf{S}$ that is composed of all mRNA-specific state spaces $\mathbf{S}_m$

$$\mathbf{S} \equiv \bigcup_m \mathbf{S}_m = \bigcup_m \bigcup_{x,i} \{(c_{x,m}|i)\} \tag{3.36}$$

where the index m runs over all individual mRNAs $1, \ldots, M$ and $c_{x,m} \in \mathbf{C}$ indicates the codon found at position x in mRNA m. Therefore, the full state space of translation elongation consists of $12 \times \sum_{m=1}^{M} L_m$ states, where L_m is the length of mRNA m. The average length of mRNAs in bacteria is about 300 codons [48] and there are more than 1000 mRNA molecules in each *E. coli* cell [49]. Thus, for translation in *E. coli* the state space $\mathbf{S}$ contains more than 3.6×10^6 states. Due to this enormous number of states, in general a detailed time-dependent analysis of translation elongation in the entire cell is not feasible.

However, we can analyze the entire Markov process of cell-wide translation elongation under steady state conditions. In steady state, the flux of ribosomes entering mRNAs is equal to the flux of ribosomes terminating translation and to the flux of ribosomes elongating on mRNAs. In particular, for every codon in the cell the flux of arriving ribosomes balances the flux of leaving ribosomes, and this flux is the same on every codon. This means that the steady state probabilities to attain a certain state only depend on the species c of a codon, but not on its position x. In other words, all codons of one species $c = 1 \ldots, 61$ can be treated identically. Thus, in steady state we can describe translation by a reduced ribosomal state space that consists of $12 \times 61 = 732$ states $(c|i)$, each of which corresponds to a certain codon species c in the ribosomal A site and one of the twelve ribosomal states $i = 0, \ldots, 11$.

Then, in steady state the marginal probability P_c to find a ribosome on codon c is simply given by the relative codon-specific elongation time

$$\mathrm{P}_c = \frac{\mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}}}{\langle \mathrm{t}_{c,\mathrm{elo}} \rangle} \,, \tag{3.37}$$

with codon-specific elongation time $\mathrm{t}_{c,\mathrm{elo}}$ (3.10), codon usage p_c (3.30), and average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$ (3.31). Furthermore, in steady state the joint probability $\mathrm{P}_{c,i}$ to find a ribosome on codon c in state $(c|i)$ is by definition

$$\mathrm{P}_{c,i} = \mathrm{P}_{i|c} \, \mathrm{P}_c \,, \tag{3.38}$$

where $\mathrm{P}_{i|c}$ is the conditional probability to find a ribosome in state $(c|i)$ under the condition that codon c is in the ribosomal A site. The steady state conditional probabilities $\mathrm{P}_{i|c}$ can be computed by using again the auxiliary Markov process introduced in section 3.4 and fig. 3.4. In particular, we can identify the steady state probabilities $\mathrm{P}^{(\mathrm{a})}_{c,i}$ of the auxiliary process with the steady state conditional probabilities $\mathrm{P}_{i|c}$

$$\mathrm{P}^{(\mathrm{a})}_{c,i} = \mathrm{P}_{i|c} \,. \tag{3.39}$$

Thus, in steady state the joint probability $\mathrm{P}_{c,i}$ to find a ribosome in state $(c|i)$ on codon c can be expressed by the dwell times (3.13) derived in section 3.4, and the marginal probability (3.37)

$$\mathrm{P}_{c,i} = \frac{\mathrm{t}_{(c|i)}}{\mathrm{t}_{c,\mathrm{elo}}} \frac{\mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}}}{\langle \mathrm{t}_{c,\mathrm{elo}} \rangle} = \frac{\mathrm{p}_c \, \mathrm{t}_{(c|i)}}{\langle \mathrm{t}_{c,\mathrm{elo}} \rangle} \,. \tag{3.40}$$

3.7 Ribosomal and tRNA Subpopulations

It is important to note that the codon-specific elongation rates (3.10) and, thus, the expression for the overall elongation rate ω_{elo} (3.31) involve the binding rates ω_{01}, ω_{06}, and $\omega_{0,11}$ for cognate, near-cognate, and non-cognate TCs. The latter rates are given by expressions (3.3) to (3.5), which depend on the concentrations $\hat{\mathrm{X}}_a$ of *free* TCs.
Therefore, in this section we will derive implicit equations to compute these concentrations from measured *total* tRNA concentrations. In order to do so, we first define several subpopulations of ribosomes and tRNAs. We then describe the time dependence of their concentrations. Finally, we derive expressions for all of these concentrations in steady state.

3.7.1 Time Dependent Concentrations of Ribosomal and tRNA Subpopulations

We define the ribosomal and tRNA subpopulations as follows. First, we consider translation in the cell or *in vitro* with fixed time-independent values of the total concentration $\mathcal{R}$ of actively translating ribosomes, of the total concentrations X_a of all tRNA species a, and of the total concentration of all EF-Tu molecules $\mathcal{E}$. Second, the total ribosome concentration $\mathcal{R}$ can be decomposed as

$$\mathcal{R} = \sum_{c=1}^{61} \mathcal{R}_c(t)\,, \tag{3.41}$$

where $\mathcal{R}_c(t)$ is the concentration of all ribosomes with codon c in their A sites at time t. The concentration $\mathcal{R}_c(t)$ can be expressed by the marginal probability $\mathrm{P}_c(t)$ to find a ribosome with codon c in its A site

$$\mathcal{R}_c(t) = \mathrm{P}_c(t)\,\mathcal{R}\,. \tag{3.42}$$

Similarly, tRNA molecules belonging to species a can be divided into several subpopulations. The time dependence of all of these tRNA subpopulations is tightly connected to the ribosomal dynamics described by the Markov process that is defined in the previous sections. We denote the concentrations of free TCs by $\hat{\mathrm{X}}_a$ and of cognate, near-cognate, and non-cognate TCs after initial binding to ribosomes as $\hat{\mathrm{X}}_a^{(1)}$, $\hat{\mathrm{X}}_a^{(6)}$, and $\hat{\mathrm{X}}_a^{(11)}$, respectively. Accordingly, the concentrations of cognate TCs that are bound to ribosomes in state $(c|2)$, and of near-cognate TCs bound to ribosomes in state $(c|7)$ are denoted by $\hat{\mathrm{X}}_a^{(2)}$ and $\hat{\mathrm{X}}_a^{(7)}$, respectively. The time dependent concentrations of the initially bound

TCs then fulfill

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(1)}(t) = \kappa_{\mathrm{on}}\,\hat{\mathrm{X}}_a(t) \sum_{c\in\mathbf{C}_{\mathrm{co}}(a)} \mathrm{P}_{0|c}(t)\,\mathcal{R}_c(t) + \hat{\mathrm{X}}_a^{(2)}(t)\,\omega_{21} - \hat{\mathrm{X}}_a^{(1)}(t)\,(\omega_{\mathrm{off}} + \omega_{\mathrm{rec}})\ , \tag{3.43}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(6)}(t) = \kappa_{\mathrm{on}}\,\hat{\mathrm{X}}_a(t) \sum_{c\in\mathbf{C}_{\mathrm{nr}}(a)} \mathrm{P}_{0|c}(t)\,\mathcal{R}_c(t) + \hat{\mathrm{X}}_a^{(7)}(t)\,\omega_{76} - \hat{\mathrm{X}}_a^{(6)}(t)\,(\omega_{\mathrm{off}} + \omega_{\mathrm{rec}})\ , \tag{3.44}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(11)}(t) = \kappa_{\mathrm{on}}\,\hat{\mathrm{X}}_a(t) \sum_{c\in\mathbf{C}_{\mathrm{no}}(a)} \mathrm{P}_{0|c}(t)\,\mathcal{R}_c(t) - \hat{\mathrm{X}}_a^{(11)}(t)\,\omega_{\mathrm{off}}\,, \tag{3.45}$$

with transition rates ω_{ij} as defined in fig. 3.2 and the conditional probability $\mathrm{P}_{0|c}(t)$ that a ribosome dwelling on codon c attains the free state $(c|0)$ with empty A site. The concentration $\mathcal{R}_c(t)$ of all ribosomes with codon c in their A sites is defined above (3.42).

After entering the cognate branch of the codon-specific Markov process shown in fig. 3.2, the TC/ribosome complex visits the states $(c|2)$ to $(c|4)$. The dynamics of the concentrations of the associated TC subpopulations follow

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(2)}(t) = \hat{\mathrm{X}}_a^{(1)}(t)\,\omega_{\mathrm{rec}} - \hat{\mathrm{X}}_a^{(2)}(t)\,(\omega_{21} + \omega_{23})\ , \tag{3.46}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(3)}(t) = \hat{\mathrm{X}}_a^{(2)}(t)\,\omega_{23} - \hat{\mathrm{X}}_a^{(3)}(t)\,\omega_{\mathrm{con}}\,, \tag{3.47}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(4)}(t) = \hat{\mathrm{X}}_a^{(3)}(t)\,\omega_{\mathrm{con}} - \hat{\mathrm{X}}_a^{(4)}(t)\,(\omega_{40} + \omega_{45})\ . \tag{3.48}$$

Similarly, ribosomes in the near-cognate branch attain states $(c|7)$ to $(c|9)$ and the concentrations of the TC subpopulations bound to these ribosomes are denoted by $\hat{\mathrm{X}}_a^{(7)}$, $\hat{\mathrm{X}}_a^{(8)}$, and $\hat{\mathrm{X}}_a^{(9)}$, respectively. Their time dependence is determined by

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(7)}(t) = \hat{\mathrm{X}}_a^{(6)}(t)\,\omega_{\mathrm{rec}} - \hat{\mathrm{X}}_a^{(7)}(t)\,(\omega_{76} + \omega_{78})\ , \tag{3.49}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(8)}(t) = \hat{\mathrm{X}}_a^{(7)}(t)\,\omega_{78} - \hat{\mathrm{X}}_a^{(8)}(t)\,\omega_{\mathrm{con}}\,, \tag{3.50}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\hat{\mathrm{X}}_a^{(9)}(t) = \hat{\mathrm{X}}_a^{(8)}(t)\,\omega_{\mathrm{con}} - \hat{\mathrm{X}}_a^{(9)}(t)\,(\omega_{90} + \omega_{9,10})\ . \tag{3.51}$$

Furthermore, concentrations of tRNAs of species a that got fully accommodated in ribosomal A sites are denoted by $\mathrm{X}_a^{\mathrm{A}}$, of tRNAs bound to ribosomal P sites by $\mathrm{X}_a^{\mathrm{P}}$, and

of tRNAs bound to ribosomal E sites by $\mathrm{X}_a^{\mathrm{E}}$. Their dynamics are described by

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\mathrm{X}_a^{\mathrm{A}}(t) = \hat{\mathrm{X}}_a^{(4)}(t)\,\omega_{45} + \hat{\mathrm{X}}_a^{(9)}(t)\,\omega_{9,10} - \mathrm{X}_a^{\mathrm{A}}(t)\,\omega_{\mathrm{pro}}\,, \tag{3.52}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\mathrm{X}_a^{\mathrm{P}}(t) = \mathrm{X}_a^{\mathrm{A}}(t)\,\omega_{\mathrm{pro}} - \mathrm{X}_a^{\mathrm{P}}(t)\,\omega_{\mathrm{elo}}\,, \tag{3.53}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\mathrm{X}_a^{\mathrm{E}}(t) = \mathrm{X}_a^{\mathrm{P}}(t)\,\omega_{\mathrm{elo}} - \mathrm{X}_a^{\mathrm{E}}(t)\,\frac{\omega_{\mathrm{pro}}\omega_{\mathrm{elo}}}{\omega_{\mathrm{pro}} - \omega_{\mathrm{elo}}}\,. \tag{3.54}$$

As shown in fig. 3.3, de-aminoacylated tRNAs released from ribosomes are recharged by synthetases with new amino acids with rate ω^{re}. The concentration of tRNAs being in this process of recharging is called $\mathrm{X}_a^{\mathrm{re}}$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\mathrm{X}_a^{\mathrm{re}}(t) = \mathrm{X}_a^{\mathrm{E}}(t)\,\frac{\omega_{\mathrm{pro}}\omega_{\mathrm{elo}}}{\omega_{\mathrm{pro}} - \omega_{\mathrm{elo}}} - \mathrm{X}_a^{\mathrm{re}}(t)\,\omega^{\mathrm{re}}\,. \tag{3.55}$$

Note that eqs. (3.54) and (3.55) are given for the 2-3-2 pathway of E site tRNA release, in which the E site tRNA does not leave the ribosome until a new tRNA has accommodated in the ribosomal A site. In contrast, when the 2-1-2 pathway is considered, tRNAs are released rapidly from the E site so that always $\mathrm{X}_a^{\mathrm{E}}(t) \approx 0$ and the first term on the right hand side of eq. (3.55) has to be replaced by the last term of eq. (3.53).

After being recharged, a tRNA needs to bind an EF-Tu molecule to form a new TC, see upper right of fig. 3.3. The concentration $\mathrm{X}_a^{\mathrm{ch}}$ refers to aminoacylated tRNAs that are not bound to EF-Tu molecules. Furthermore, we denote by $\mathcal{E}^{\mathrm{fr}}$ the concentration of free EF-Tu molecules . Free Ef-Tu binds to charged tRNAs with binding rate constant κ^{ass} to form a TC that in turn decays with rate ω^{dis}

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\mathrm{X}_a^{\mathrm{ch}}(t) = \mathrm{X}_a^{\mathrm{re}}(t)\,\omega^{\mathrm{re}} + \hat{\mathrm{X}}_a(t)\,\omega^{\mathrm{dis}} - \kappa^{\mathrm{ass}}\mathcal{E}^{\mathrm{fr}}(t)\mathrm{X}_a^{\mathrm{ch}}(t)\,, \tag{3.56}$$

$$\frac{\mathrm{d}}{\mathrm{dt}}\,\mathcal{E}^{\mathrm{fr}}(t) = \hat{\mathrm{X}}_a^{(4)}(t)\,\omega_{45} + \hat{\mathrm{X}}_a^{(9)}(t)\,\omega_{9,10} + \hat{\mathrm{X}}_a(t)\,\omega^{\mathrm{dis}} - \kappa^{\mathrm{ass}}\mathcal{E}^{\mathrm{fr}}(t)\mathrm{X}_a^{\mathrm{ch}}(t)\,, \tag{3.57}$$

where the first two terms in the latter equation arise from the fact that a TC disintegrates and its EF-Tu leaves the ribosome as soon as its tRNA gets fully accommodated in the ribosomal A site, see fig. 3.1.

Finally, because the number of molecules of each tRNA species a is taken to be conserved, the tRNA subpopulations always add up to the total concentration X_a

$$\begin{aligned}\mathrm{X}_a =& \hat{\mathrm{X}}_a(t) + \sum_{i=1}^{4}\hat{\mathrm{X}}_a^{(i)}(t) + \sum_{i=6}^{9}\hat{\mathrm{X}}_a^{(i)}(t) + \hat{\mathrm{X}}_a^{(11)}(t)\\ &+ \mathrm{X}_a^{\mathrm{A}}(t) + \mathrm{X}_a^{\mathrm{P}}(t) + \mathrm{X}_a^{\mathrm{E}}(t) + \mathrm{X}_a^{\mathrm{re}}(t) + \mathrm{X}_a^{\mathrm{ch}}(t)\,.\end{aligned} \tag{3.58}$$

3.7.2 Steady State Concentrations of tRNA Subpopulations

If we assume that translation is in steady state, the solution of the Markov process for cell-wide translation discussed in section 3.6 leads to expressions for the different tRNA subpopulations in terms of the concentrations of all free TCs. In steady state, the left hand sides of eqs. (3.43) to (3.57) vanish, and we can use eqs. (3.16), (3.40) and (3.42) to obtain

$$\mathrm{P}_{0|c}(t)\,\mathcal{R}_c(t) = \mathrm{P}_{c,0}\,\mathcal{R} = \frac{\omega_{\mathrm{elo}}\,\mathrm{p}_c}{\omega_{c,\mathrm{co}} + \omega_{c,\mathrm{nr}}}\mathcal{R} \tag{3.59}$$

where the effective cognate and near-cognate accommodation rates $\omega_{c,\mathrm{co}}$ and $\omega_{c,\mathrm{nr}}$ are defined in eqs. (3.23) and (3.29), respectively. In steady state, expression (3.59) replaces the corresponding terms in eqs. (3.43) to (3.45).
We focus on one particular tRNA species b. The steady state solutions for cognate bound TCs of species b can be expressed by the free concentrations $\hat{\mathrm{X}}_b$ and $\hat{\mathrm{X}}_a$ of TC species b and all other TC species a, respectively,

$$\hat{\mathrm{X}}_b^{(1)} = \frac{\omega_{\mathrm{elo}}}{\omega_{\mathrm{rec}}\pi_{23}\pi_{45}} \sum_{c\in\mathbf{C}_{\mathrm{co}}(b)} \frac{\mathcal{P}_{c,\mathrm{co}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a}\,\mathcal{R}\,\hat{\mathrm{X}}_b\,, \tag{3.60}$$

$$\hat{\mathrm{X}}_b^{(2)} = \frac{\omega_{\mathrm{elo}}}{\omega_{23}\pi_{45}} \sum_{c\in\mathbf{C}_{\mathrm{co}}(b)} \frac{\mathcal{P}_{c,\mathrm{co}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a}\,\mathcal{R}\,\hat{\mathrm{X}}_b\,, \tag{3.61}$$

$$\hat{\mathrm{X}}_b^{(3)} = \frac{\omega_{\mathrm{elo}}}{\omega_{\mathrm{con}}\pi_{45}} \sum_{c\in\mathbf{C}_{\mathrm{co}}(b)} \frac{\mathcal{P}_{c,\mathrm{co}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a}\,\mathcal{R}\,\hat{\mathrm{X}}_b\,, \tag{3.62}$$

$$\hat{\mathrm{X}}_b^{(4)} = \frac{\omega_{\mathrm{elo}}}{\omega_{45}} \sum_{c\in\mathbf{C}_{\mathrm{co}}(b)} \frac{\mathcal{P}_{c,\mathrm{co}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a}\,\mathcal{R}\,\hat{\mathrm{X}}_b\,, \tag{3.63}$$

where the probability of cognate accommodation $\mathcal{P}_{c,\mathrm{co}}$ is defined in eq. (3.23). Similarly, tRNAs of species b that are contained in near-cognate bound TCs yield

$$\hat{\mathrm{X}}_b^{(6)} = \frac{\omega_{\mathrm{elo}}}{\omega_{\mathrm{rec}}\pi_{78}\pi_{9,10}} \sum_{c\in\mathbf{C}_{\mathrm{nr}}(b)} \frac{\mathcal{P}_{c,\mathrm{nr}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a}\,\mathcal{R}\,\hat{\mathrm{X}}_b\,, \tag{3.64}$$

$$\hat{\mathrm{X}}_b^{(7)} = \frac{\omega_{\mathrm{elo}}}{\omega_{78}\pi_{9,10}} \sum_{c\in\mathbf{C}_{\mathrm{nr}}(b)} \frac{\mathcal{P}_{c,\mathrm{nr}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a}\,\mathcal{R}\,\hat{\mathrm{X}}_b\,, \tag{3.65}$$

$$\hat{\mathrm{X}}_b^{(8)} = \frac{\omega_{\mathrm{elo}}}{\omega_{\mathrm{con}}\pi_{9,10}} \sum_{c\in\mathbf{C}_{\mathrm{nr}}(b)} \frac{\mathcal{P}_{c,\mathrm{nr}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a}\,\mathcal{R}\,\hat{\mathrm{X}}_b\,, \tag{3.66}$$

and

$$\hat{\mathrm{X}}_b^{(9)} = \frac{\omega_{\mathrm{elo}}}{\omega_{9,10}} \sum_{c \in \mathbf{C}_{\mathrm{nr}}(b)} \frac{\mathcal{P}_{c,\mathrm{nr}}\, \mathrm{p}_c}{\sum\limits_{a \in \mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a} \mathcal{R}\, \hat{\mathrm{X}}_b \,. \tag{3.67}$$

For non-cognate bound TCs of species b the following expression holds

$$\hat{\mathrm{X}}_b^{(11)} = \omega_{\mathrm{elo}} \left(\frac{1}{\omega_{\mathrm{rec}} \pi_{23} \pi_{45}} + \frac{1}{\omega_{\mathrm{off}} \pi_{45}} \right) \sum_{c \in \mathbf{C}_{\mathrm{no}}(b)} \frac{\mathcal{P}_{c,\mathrm{co}}\, \mathrm{p}_c}{\sum\limits_{a \in \mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a} \mathcal{R}\, \hat{\mathrm{X}}_b \,. \tag{3.68}$$

The steady state concentration $\mathrm{X}_b^{\mathrm{P}}$ of tRNAs of species b accommodated in ribosomal P sites is given by

$$\mathrm{X}_b^{\mathrm{P}} = \left(\sum_{c \in \mathbf{C}_{\mathrm{co}}(b)} \frac{\mathcal{P}_{c,\mathrm{co}}\, \mathrm{p}_c}{\sum\limits_{a \in \mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a} + \sum_{c \in \mathbf{C}_{\mathrm{nr}}(b)} \frac{\mathcal{P}_{c,\mathrm{nr}}\, \mathrm{p}_c}{\sum\limits_{a \in \mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a} \right) \mathcal{R}\, \hat{\mathrm{X}}_b \,, \tag{3.69}$$

which is related to the steady state concentrations $\mathrm{X}_b^{\mathrm{A}}$, $\mathrm{X}_b^{\mathrm{E}}$, $\mathrm{X}_b^{\mathrm{re}}$, and $\mathrm{X}_b^{\mathrm{ch}}$ of tRNAs of species b in ribosomal A and E sites as well as in de-aminoacylated and aminoacylated form

$$\mathrm{X}_b^{\mathrm{A}} = \frac{\omega_{\mathrm{elo}}}{\omega_{\mathrm{pro}}}\, \mathrm{X}_b^{\mathrm{P}} \,, \tag{3.70}$$

$$\mathrm{X}_b^{\mathrm{E}} = \left(1 - \frac{\omega_{\mathrm{elo}}}{\omega_{\mathrm{pro}}} \right) \mathrm{X}_b^{\mathrm{P}} \,, \tag{3.71}$$

$$\mathrm{X}_b^{\mathrm{re}} = \frac{\omega_{\mathrm{elo}}}{\omega^{\mathrm{re}}}\, \mathrm{X}_b^{\mathrm{P}} \,, \tag{3.72}$$

$$\mathrm{X}_b^{\mathrm{ch}} = \frac{\omega^{\mathrm{dis}}\, \hat{\mathrm{X}}_b}{\kappa^{\mathrm{ass}}\, \mathcal{E}^{\mathrm{fr}}} + \frac{\omega_{\mathrm{elo}}\, \mathrm{X}_b^{\mathrm{P}}}{\kappa^{\mathrm{ass}}\, \mathcal{E}^{\mathrm{fr}}} \,. \tag{3.73}$$

Note that eq. (3.71) only holds for the 2-3-2 pathway of E site tRNA release, as in case of the 2-1-2 pathway the concentration of E site bound tRNA $\mathrm{X}_b^{\mathrm{E}}$ is always zero.

3.7.3 Steady State Concentrations of Free Ternary Complexes

In the previous section, the steady state solutions of the concentrations $\hat{\mathrm{X}}_b^{(1)}, \ldots \mathrm{X}_b^{\mathrm{ch}}$ of the various subpopulations of tRNA species b are presented in terms of the concentrations $\hat{\mathrm{X}}_a$ and $\hat{\mathrm{X}}_b$ of free TCs, see eqs. (3.60) to (3.73). In this section, we will derive an implicit expression for the free concentration $\hat{\mathrm{X}}_b$ of TC species b in steady state.

The total steady state concentration X^{ch} of all charged tRNAs depends on the steady state concentration $\mathcal{E}^{\mathrm{fr}}$ of free EF-Tu molecules, the total concentration X^{P} of P site accommodated tRNAs, and the total concentration $\hat{\mathrm{X}} = \sum_a \hat{\mathrm{X}}_a$ of free TCs

$$\mathrm{X}^{\mathrm{ch}} = \sum_a \mathrm{X}_a^{\mathrm{ch}} = \frac{\omega^{\mathrm{dis}}\,\hat{\mathrm{X}}}{\kappa^{\mathrm{ass}}\,\mathcal{E}^{\mathrm{fr}}} + \frac{\omega_{\mathrm{elo}}\,\mathrm{X}^{\mathrm{P}}}{\kappa^{\mathrm{ass}}\,\mathcal{E}^{\mathrm{fr}}}\,, \tag{3.74}$$

where $\mathrm{X}^{\mathrm{P}} = \sum_a \mathrm{X}_a^{\mathrm{P}}$ equals the total concentration $\mathcal{R}$ of ribosomes as the P site of every active ribosome is occupied by a tRNA. Using eq. (3.74), the concentration $\mathcal{E}^{\mathrm{fr}}$ of free EF-Tu molecules can be deduced from the assumption that the total concentration $\mathcal{E}$ of EF-Tu molecules is conserved, i.e., that the sum of all concentrations of TCs and $\mathcal{E}^{\mathrm{fr}}$ must equal the total concentration $\mathcal{E}$ of EF-Tu

$$\begin{aligned}\mathcal{E}^{\mathrm{fr}} &= \mathcal{E} - \left(\hat{\mathrm{X}} + \sum_{i=1}^{4} \hat{\mathrm{X}}^{(i)} + \sum_{i=6}^{9} \hat{\mathrm{X}}^{(i)} + \hat{\mathrm{X}}^{(11)}\right) \\ &= \mathcal{E} - \left(\mathrm{X} - \mathrm{X}^{\mathrm{A}} - \mathrm{X}^{\mathrm{P}} - \mathrm{X}^{\mathrm{E}} - \mathrm{X}^{\mathrm{re}} - \mathrm{X}^{\mathrm{ch}}\right)\,,\end{aligned} \tag{3.75}$$

where $\mathrm{X} = \sum_a \mathrm{X}_a$ is the sum of all total tRNA concentrations (3.58). The steady state total concentration $\mathrm{X}^{\mathrm{A}} + \mathrm{X}^{\mathrm{E}} + \mathrm{X}^{\mathrm{re}} = \sum_a (\mathrm{X}_a^{\mathrm{A}} + \mathrm{X}_a^{\mathrm{E}} + \mathrm{X}_a^{\mathrm{re}})$ of A and E site accommodated tRNAs and tRNAs before recharging is equal to $\left(\frac{\omega_{\mathrm{elo}}}{\omega_{\mathrm{pro}}} + \frac{\omega_{\mathrm{elo}}}{\omega^{\mathrm{re}}}\right)\mathcal{R}$ and $\left(1 + \frac{\omega_{\mathrm{elo}}}{\omega^{\mathrm{re}}}\right)\mathcal{R}$ for the 2-1-2 and the 2-3-2 pathway, respectively, see eqs. (3.70) to (3.72) in the previous section. Thus, the concentration of free EF-Tu molecules can be written as

$$\mathcal{E}^{\mathrm{fr}} = \frac{\beta}{2} + \sqrt{\frac{\beta^2}{4} + \frac{\omega^{\mathrm{dis}}}{\kappa^{\mathrm{ass}}}\,\hat{\mathrm{X}} + \frac{\omega_{\mathrm{elo}}}{\kappa^{\mathrm{ass}}}\,\mathcal{R}}\,, \tag{3.76}$$

with

$$\beta \equiv \begin{cases} \mathcal{E} - \mathrm{X} + \left(1 + \frac{\omega_{\mathrm{elo}}}{\omega_{\mathrm{pro}}} + \frac{\omega_{\mathrm{elo}}}{\omega^{\mathrm{re}}}\right)\mathcal{R} & \text{for the 2-1-2 pathway,} \\ \mathcal{E} - \mathrm{X} + \left(2 + \frac{\omega_{\mathrm{elo}}}{\omega^{\mathrm{re}}}\right)\mathcal{R} & \text{for the 2-3-2 pathway.} \end{cases} \tag{3.77}$$

Finally, the free concentration $\hat{\mathrm{X}}_b$ of TC species b can be deduced from the steady state concentrations of all tRNA subpopulations (3.60) to (3.73) together with constraint (3.58) that the total number of molecules of tRNA species b is conserved. As a result, we obtain the free concentration $\hat{\mathrm{X}}_b$ of TC species b as an implicit function of all free concentrations $\hat{\mathrm{X}}_a$ of TC species a with $a = 1, \ldots, 43$, the concentration $\mathcal{E}^{\mathrm{fr}}$ of free EF-Tu molecules as determined by eq. (3.76), the concentration $\mathcal{R}$ of ribosomes, the codon-dependent probabilities $\mathcal{P}_{c,\mathrm{co}}$ and $\mathcal{P}_{c,\mathrm{nr}}$ of cognate and near-cognate accommoda-

tion (3.23) and (3.29), and the codon usages p_c

$$\hat{\mathrm{X}}_b = \mathrm{X}_b\Bigg(1 + \frac{\omega^{\mathrm{dis}}}{\kappa^{\mathrm{ass}}\,\mathcal{E}^{\mathrm{fr}}} + \mathcal{R}\Bigg(\Phi_{\mathrm{co}} \sum_{c\in\mathbf{C}_{\mathrm{co}}(b)} \frac{\mathcal{P}_{c,\mathrm{co}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a} + \Phi_{\mathrm{nr}} \sum_{c\in\mathbf{C}_{\mathrm{nr}}(b)} \frac{\mathcal{P}_{c,\mathrm{nr}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a} + \\ + \omega_{\mathrm{elo}}\tau_{\mathrm{no}} \sum_{c\in\mathbf{C}_{\mathrm{no}}(b)} \frac{\mathcal{P}_{c,\mathrm{co}}\,\mathrm{p}_c}{\sum\limits_{a\in\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a}\Bigg)\Bigg)^{-1} . \tag{3.78}$$

For the 2-3-2 pathway of tRNA release from the ribosomal E site, the dimensionless constants Φ_{co} and Φ_{nr} assume

$$\Phi_{\mathrm{co}} = 2 + \omega_{\mathrm{elo}}\left(\tau_{\mathrm{co}} + \frac{1}{\omega^{\mathrm{re}}} + \frac{1}{\kappa^{\mathrm{ass}}\,\mathcal{E}^{\mathrm{fr}}}\right) , \tag{3.79}$$

$$\Phi_{\mathrm{nr}} = 2 + \omega_{\mathrm{elo}}\left(\tau_{\mathrm{nr}} + \frac{1}{\omega^{\mathrm{re}}} + \frac{1}{\kappa^{\mathrm{ass}}\,\mathcal{E}^{\mathrm{fr}}}\right) , \tag{3.80}$$

and in case of the 2-1-2 pathway

$$\Phi_{\mathrm{co}} = 1 + \omega_{\mathrm{elo}}\left(\tau_{\mathrm{co}} + \frac{1}{\omega_{\mathrm{pro}}} + \frac{1}{\omega^{\mathrm{re}}} + \frac{1}{\kappa^{\mathrm{ass}}\,\mathcal{E}^{\mathrm{fr}}}\right) , \tag{3.81}$$

$$\Phi_{\mathrm{nr}} = 1 + \omega_{\mathrm{elo}}\left(\tau_{\mathrm{nr}} + \frac{1}{\omega_{\mathrm{pro}}} + \frac{1}{\omega^{\mathrm{re}}} + \frac{1}{\kappa^{\mathrm{ass}}\,\mathcal{E}^{\mathrm{fr}}}\right) , \tag{3.82}$$

with constant time scales

$$\tau_{\mathrm{co}} = \frac{1}{\omega_{\mathrm{rec}}\pi_{23}\pi_{45}} + \frac{1}{\omega_{23}\pi_{45}} + \frac{1}{\omega_{\mathrm{con}}\pi_{45}} + \frac{1}{\omega_{45}} , \tag{3.83}$$

$$\tau_{\mathrm{nr}} = \frac{1}{\omega_{\mathrm{rec}}\pi_{78}\pi_{9,10}} + \frac{1}{\omega_{78}\pi_{9,10}} + \frac{1}{\omega_{\mathrm{con}}\pi_{9,10}} + \frac{1}{\omega_{9,10}} , \tag{3.84}$$

$$\tau_{\mathrm{no}} = \frac{1}{\omega_{\mathrm{rec}}\pi_{23}\pi_{45}} + \frac{1}{\omega_{\mathrm{off}}\pi_{45}} , \tag{3.85}$$

where π_{ij} represents the probability of transition from state $(c|i)$ to state $(c|j)$, see definition (2.6).
Since *E. coli* contains 43 different elongator tRNA species b, there are 43 different equations (3.78).

4 Specification of Elongation Parameters and *In Vitro* Measurements

In general, the different parameters that enter our theory on translation elongation introduced in the previous chapter take on different values *in vitro* and *in vivo*. Furthermore, the *in vitro* parameters depend on the experimental conditions such as temperature and buffer composition, whereas the *in vivo* parameters depend on the environment in which the cells are growing [3, 49]. In this work, we will focus on four *in vivo* growth conditions for *E. coli* cells leading to growth rates of 0.7, 1.07, 1.6 and 2.5 doublings/hour (dbl/h), as well as on two *in vitro* conditions with different temperatures, 20 °C and 37 °C, but identical buffers.

4.1 Total tRNA, Ribosome and EF-Tu Concentrations

In vitro, the concentrations of ribosomes and total ternary complexes (TCs) are controlled by the experimenter, see section 4.5. Furthermore, for an excess of TCs, the difference between total and free TCs can be neglected, and we assume that the ratios of concentrations of the individual TCs are the same as measured for tRNAs *in vivo* at a growth rate of 2.5 doublings/hour.

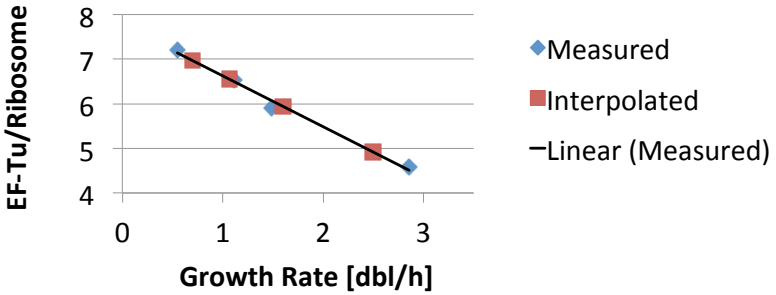


Figure 4.1: **Ratios of EF-Tu to ribosome concentrations.** For consistency, the measured ratios for different growth rates of *E. coli* cells [51] are interpolated in all used data sets to obtain ratios for the same growth rates as studied in [3].

Table 4.1: ***In vivo* concentrations of all tRNAs, actively translating ribosomes $\mathcal{R}$, and EF-Tu molecules $\mathcal{E}$ in *E. coli* for four different growth rates.** Concentrations of tRNAs from table 5 in ref. [3], with $tRNA^{Val2A}$ and $tRNA^{Val2B}$ added (Val2), and individual concentrations of $tRNA^{Gly1}$ and $tRNA^{Gly2}$ as well as $tRNA^{Ile1}$ and $tRNA^{Ile2}$ obtained by using corresponding ratios given in [50]. All concentrations in [µM].

	E. coli growth rates [dbl/h]					*E. coli* growth rates [dbl/h]			
	0.7	1.07	1.6	2.5		0.7	1.07	1.6	2.5
Ala1B	11.73	14.06	17.52	20.97	Lys	6.8	7.35	8.73	10.43
Ala2	2.12	2.33	3.19	3.57	Met m	2.59	2.91	4.10	4.43
Arg2	14.54	15.54	23.77	25.57	Phe	3.6	4.29	4.69	5.11
Arg3	2.61	1.45	2.26	2.3	Pro1	2.44	3.51	2.75	2.67
Arg4	2.35	2.64	3.26	3.52	Pro2	2.51	2.26	4.01	3.75
Arg5	1.57	1.61	2.46	2.2	Pro3	1.89	2.22	2.55	2.56
Asn	3.86	4.35	6.1	7.29	Sec	0.86	0.96	1.05	1.04
Asp1	8.13	8.42	12.04	15.46	Ser1	5.56	5.47	6.98	7.36
Cys	4.88	5.23	7.04	7.07	Ser2	1.04	1.17	1.37	1.45
Gln1	2.72	3.63	3.17	4.38	Ser3	4.39	4.53	5.4	5.67
Gln2	3.08	3.47	5.07	6.27	Ser5	2.6	2.87	3.68	4.03
Glu2	15.58	16.71	24.12	29.35	Thr1	0.41	0.54	0.56	0.67
Gly1	2.87	3.10	4.38	4.43	Thr2	2	2.11	2.67	3.12
Gly2	4.31	4.64	6.57	6.65	Thr3	3.73	3.87	4.86	5.54
Gly3	15.21	16.75	19.84	24.96	Thr4	3.17	3.25	4.99	6.89
His	2.19	2.63	3.35	4.38	Trp	2.78	3.35	4.15	5.02
Ile1	11.29	12.61	18.02	23.56	Tyr1	2.41	2.7	4.61	4.19
Ile2	0.56	0.63	0.90	1.18	Tyr2	3.86	3.75	5.22	5.04
Leu1	14.91	16.76	21.32	22.2	Val1	12.07	11.07	18.99	20.39
Leu2	3.47	4.04	4.72	5.93	Val2	4.39	5.02	6.31	7.21
Leu3	2.49	2.62	3.19	3.17					
Leu4	6.33	6.97	9.66	9.3	$\mathcal{R}$	18.57	23.57	33.85	44.56
Leu5	3.47	4.07	3.65	3.78	$\mathcal{E}$	152.22	181.57	236.69	257.88

The *in vivo* concentrations of total tRNA in *E. coli* are taken from [3]. The corresponding *in vivo* concentrations $\mathcal{R}$ of active ribosomes are calculated from table 3 in [3] by taking into account that only 85% of all ribosomes in the cell are active [52]. Furthermore, the total *in vivo* concentration of EF-Tu can be estimated by interpolating the measured ratios of EF-Tu and ribosome concentrations for different growth rates, see fig. 4.1 and [51], and multiplying these ratios by the ribosome concentration $\mathcal{R}$. All concentrations of total tRNAs, ribosomes and EF-Tu are summarized in table 4.1.

4.2 Codon Usages

In vitro, the codon usages simply correspond to the frequencies of codons in the chosen mRNA. *In vivo*, if the relative abundances of the mRNAs in the cells are known, the codon usages of the sense codons can be derived from the mRNA sequences, see for example [53], by normalizing the frequencies of the different codons. To measure relative mRNA abundances in *E. coli* at a growth rate of about 2.5 dbl/h, the deep sequencing method *RNA-seq* was used [54]. From the RNA-seq data of 4215 different genes, we computed the codon usages as displayed in table 4.2. They exhibit a high correlation to values published in [3] that were determined from relative abundances of 140 abundant proteins. For all other growth rates, we use codon usages published in [3]. Table 4.2 summarizes the codon usages of all sense codons in *E. coli* for the four different growth rates.

4.3 Ribosomal Transition Rates and Overall Elongation Rate

The measurement of the *in vitro* overall elongation rate ω_{elo} is described in section 4.5. The *in vivo* overall elongation rate $\omega^{\star}_{\text{elo}}$ has been determined in [52] for several growth conditions. In three out of four cases, we use the values as stated in [52], because the growth rates studied by these authors are approximately equal to those in [3]. In one case, for the growth rate of 0.7 dbl/h, we interpolate the overall elongation rates for growth rates of 0.6 and 1.0 dbl/h and obtain $\omega^{\star}_{\text{elo}} = 15\,\text{aa/s}$.

The internal *in vitro* rates ω_{ij} are based, to a large extent, on previous measurements as explained in the following paragraph. In addition, new experiments were performed by our collaborators from the Max Planck Institute for Biophysical Chemistry to measure the rates ω_{90} and $\omega_{9,10}$ at 20 °C as well as the overall elongation rate ω_{elo}, both at 20 °C and 37 °C, using the experimental protocols described in [30, 44, 55] and section 4.5.

Table 4.2: ***In vivo* codon usages p_c in percent for all sense codons c in *E. coli*.** For the growth rate of 2.5 dbl/h, the codon usages were determined by deep sequencing. In all other cases, data from [3] were used and renormalized to exclude the stop codons.

	E. coli growth rates [dbl/h]					*E. coli* growth rates [dbl/h]			
	0.7	1.07	1.6	2.5		0.7	1.07	1.6	2.5
AAA	4.67	4.94	5.24	5.50	GAA	5.43	5.54	5.72	4.46
AAC	2.84	2.88	2.93	2.72	GAC	2.98	3.11	3.26	2.63
AAG	1.28	1.38	1.50	1.96	GAG	1.69	1.71	1.75	1.74
AAU	0.89	0.78	0.65	0.96	GAU	2.36	2.25	2.13	2.35
ACA	0.33	0.30	0.27	0.42	GCA	2.23	2.25	2.33	2.50
ACC	2.69	2.73	2.77	2.09	GCC	1.86	1.69	1.48	1.51
ACG	0.70	0.62	0.52	1.01	GCG	2.97	2.86	2.75	2.37
ACU	1.52	1.69	1.85	1.97	GCU	3.05	3.26	3.51	3.74
AGA	0.10	0.08	0.07	0.54	GGA	0.25	0.22	0.18	0.63
AGC	1.15	1.08	1.00	1.38	GGC	3.58	3.58	3.58	2.97
AGG	0.01	0.01	0.00	0.34	GGG	0.43	0.36	0.28	0.92
AGU	0.36	0.30	0.24	0.61	GGU	3.94	4.07	4.26	3.62
AUA	0.09	0.08	0.06	0.20	GUA	1.72	1.88	2.01	2.06
AUC	3.80	3.94	4.17	3.18	GUC	1.06	0.98	0.89	1.08
AUG	2.25	2.25	2.25	2.08	GUG	2.05	1.90	1.79	1.79
AUU	2.06	1.94	1.79	1.80	GUU	3.33	3.58	3.85	3.19
CAA	0.97	0.90	0.81	1.14	UAC	1.65	1.68	1.70	1.15
CAC	1.40	1.40	1.42	1.20	UAU	1.00	0.90	0.79	0.65
CAG	2.90	2.85	2.79	2.81	UCA	0.36	0.31	0.26	0.41
CAU	0.88	0.82	0.73	0.76	UCC	1.16	1.22	1.24	1.21
CCA	0.65	0.64	0.61	0.68	UCG	0.54	0.46	0.38	0.50
CCC	0.28	0.21	0.14	0.43	UCU	1.36	1.42	1.50	1.77
CCG	2.94	2.91	2.92	2.31	UGC	0.51	0.48	0.44	0.68
CCU	0.49	0.48	0.47	0.72	UGG	0.93	0.87	0.81	0.81
CGA	0.12	0.10	0.08	0.35	UGU	0.40	0.37	0.33	0.49
CGC	2.25	2.25	2.19	2.02	UUA	0.55	0.47	0.36	0.75
CGG	0.15	0.12	0.09	0.53	UUC	2.27	2.26	2.29	2.13
CGU	3.37	3.68	3.99	3.16	UUG	0.63	0.58	0.50	0.62
CUA	0.19	0.15	0.11	0.13	UUU	1.16	1.04	0.88	1.19
CUC	0.59	0.56	0.51	0.84					
CUG	6.10	6.17	6.21	5.28					
CUU	0.53	0.47	0.40	0.94					

The *in vitro* value κ_{on} of the binding rate constant was previously measured at 20 °C [40]. Its value at 37 °C was obtained assuming an Arrhenius temperature dependence and using the previously determined activation energy of $2.4\,\mathrm{kcal\,mol^{-1}}$ for initial binding [38]. The dissociation rate ω_{off} at 20 °C is taken from [40]. The decoding rates at 20 °C were obtained by averaging over previously published values as measured for different codons of $\mathrm{tRNA^{Phe}}$. In particular, we averaged the rates given in table 1 of ref. [41] for cognate as well as for near-cognate codons to obtain the rates ω_{rec}, ω_{21}, ω_{23}, ω_{76}, and ω_{78} [56]. The rate ω_{con} has not been measured but estimated under the assumption that it is not rate-limiting. The rate $\omega_{9,10}$ at 37 °C was published previously [30] and was used to determine the rate $\omega_{90} = \omega_{9,10}\,(1 - 0.06)\,/0.06$, i.e., using an error frequency of 0.06 for the proofreading step [30]. The rate ω_{45} has been measured both for 20 °C and for 37 °C [30, 41]. The rate ω_{pro} was calculated for both temperatures from the measured overall elongation rate ω_{elo} using eq. (3.31) and solving it for the unknown ω_{pro}.

We assumed an Arrhenius temperature dependence to estimate some of the *in vitro* rates at 37 °C from their values as measured at 20 °C. These estimates are based on the following considerations. We express the rate ω_{ij} in terms of the activation free energy ΔG_{ij} and the attempt frequency ν_{ij}, and use the decomposition $\Delta G_{ij} = \Delta H_{ij} - \mathrm{T}\Delta S_{ij}$ of the activation free energy ΔG_{ij} into the activation enthalpy ΔH_{ij} and the activation entropy ΔS_{ij}, which leads to

$$\omega_{ij}(\mathrm{T}) = \nu_{ij} \exp\left[\Delta S_{ij}/\mathrm{k_B}\right] \exp\left[-\Delta H_{ij}/\mathrm{k_B T}\right] \tag{4.1}$$

$$\equiv \nu_0 \exp\left[\Delta \tilde{S}_{ij}/\mathrm{k_B}\right] \exp\left[-\Delta H_{ij}/\mathrm{k_B T}\right], \tag{4.2}$$

where the last expression involves the attempt frequency $\nu_0 = \mathrm{k_B T}/h$ as obtained from transition-state theory [57]. In this way, any state-dependence of the attempt frequency ν_{ij} has been absorbed into the activation entropies $\Delta \tilde{S}_{ij} = \Delta S_{ij} + \mathrm{k_B} \ln\left(\nu_{ij}/\nu_0\right)$. If one plots the logarithms $\ln\left(\omega_{ij}/\nu_0\right)$ of the measured rates ω_{ij} as a function of the inverse temperature 1/T, one finds linear relationships [31, 38, 58], which imply that the two unknown parameters in eq. (4.2), ΔH_{ij} and $\Delta \tilde{S}_{ij}$, do not depend on temperature over the experimentally studied temperature range. However, the activation entropies $\Delta \tilde{S}_{ij}$ as obtained from the behavior of $\ln\left(\omega_{ij}/\nu_0\right)$ for small 1/T vary significantly with the ribosomal states i and j [31, 38, 58]. Possible molecular mechanisms for this variation have been recently discussed based on atomistic molecular dynamics simulations [59].
Using the expression in eq. (4.2) with T-independent enthalpies ΔH_{ij} and entropies $\Delta \tilde{S}_{ij}$, we now consider the ratios $\omega_{ij}(\mathrm{T})/\omega_{45}(\mathrm{T})$ at the two temperatures of interest, $\mathrm{T_1} = 20\,^{\circ}\mathrm{C} = 293.15\,\mathrm{K}$ and $\mathrm{T_2} = 37\,^{\circ}\mathrm{C} = 310.15\,\mathrm{K}$. We take the cognate incorporation rate ω_{45} as a reference rate because the value of this rate has been measured at both temperatures. For each individual transition, we then obtain two equations, corresponding to the two temperatures $\mathrm{T_1}$ and $\mathrm{T_2}$, which can be combined to eliminate the enthalpy ΔH_{ij}. As a

result, we obtain the relation

$$\frac{\omega_{ij}(\mathrm{T}_2)}{\omega_{45}(\mathrm{T}_2)} = \left(\frac{\omega_{ij}(\mathrm{T}_1)}{\omega_{45}(\mathrm{T}_1)}\right)^{\mathrm{T}_1/\mathrm{T}_2} \exp\left[\frac{\Delta\tilde{S}_{ij} - \Delta\tilde{S}_{45}}{\mathrm{k_B}} \frac{\mathrm{T}_2 - \mathrm{T}_1}{\mathrm{T}_2}\right]. \tag{4.3}$$

At present, the entropy differences $\Delta\tilde{S}_{ij} - \Delta\tilde{S}_{45}$ are difficult to estimate for all individual transitions from the available experimental data. However, these differences are multiplied by the relative temperature difference $(\mathrm{T}_2 - \mathrm{T}_1)/\mathrm{T}_2 \simeq 0.055$, which is rather small. Therefore, we used the approximate relation

$$\frac{\omega_{ij}(37\,^\circ\mathrm{C})}{\omega_{45}(37\,^\circ\mathrm{C})} \simeq \left(\frac{\omega_{ij}(20\,^\circ\mathrm{C})}{\omega_{45}(20\,^\circ\mathrm{C})}\right)^{293.15/310.15}, \tag{4.4}$$

to estimate the rates ω_{off}, ω_{rec}, ω_{21}, ω_{23}, ω_{76}, ω_{78}, ω_{con}, and ω_{40} at 37 °C from their values as measured at 20 °C.

The prediction of the *in vivo* ribosomal transition rates from the *in vitro* rates is the scope of chapter 5 and will not be discussed here.

4.4 tRNA Recharging and Ternary Complex Formation

Recharging of tRNAs and formation of TCs can be neglected *in vitro*, because TCs are added to the assay in large excess, see section 4.5.
To our knowledge, the binding of de-aminoacylated tRNAs to synthetases and the rate ω^{re} of their subsequent recharging with amino acids has not been measured *in vivo* for each kind of tRNAs and each of the twenty different synthetases. Therefore, we estimated $1/\omega^{\mathrm{re}} = 10\,\mathrm{ms}$ to be the same for all tRNA species as in [60, 61].

The disassembly of TCs into EF-Tu and aa-tRNA is characterized by the dissociation rate ω^{dis}. Likewise, TC assembly from EF-Tu and aa-tRNA is governed by the association rate constant κ^{ass} and the concentration $\mathcal{E}^{\mathrm{fr}}$ of free EF-Tu molecules. We use the values $\kappa^{\mathrm{ass}} = 1/\mu\mathrm{Ms}$ and $\omega^{\mathrm{dis}} = 0.01/\mathrm{s}$ as in [44] and assume that these rates also hold for *in vivo* translation. The free EF-Tu concentration depends on the growth condition and is calculated from solving eq. (3.75). The results are $\mathcal{E}^{\mathrm{fr}} = 5.4\,\mu\mathrm{M}$ for 0.7 dbl/h, $14\,\mu\mathrm{M}$ for 1.07 dbl/h, $20\,\mu\mathrm{M}$ for 1.6 dbl/h, and $23\,\mu\mathrm{M}$ for 2.5 dbl/h.

Table 4.3: ***In vitro* rates of ribosomal transitions.** Apart from the processing rate ω_{pro}, all individual rates in fig. 3.2 and the overall elongation rate ω_{elo} have been measured *in vitro* at 20 °C and/or 37 °C. The processing rate ω_{pro} was calculated from the overall elongation rate ω_{elo} using eq. (3.31).

Rates	20 °C	37 °C	Units
ω_{elo}	0.8 ± 0.2	6.9 ± 2.3	aa/s
κ_{on}	140 ± 20	175 ± 25	$\mu\text{M}^{-1}\,\text{s}^{-1}$
ω_{off}	85 ± 25	700 ± 270	s^{-1}
ω_{rec}	180 ± 30	1500 ± 450	s^{-1}
ω_{21}	0.2 ± 0.03	2 ± 0.6	s^{-1}
ω_{23}	190 ± 30	1500 ± 450	s^{-1}
ω_{con}	50	450	s^{-1}
ω_{45}	22 ± 4	200 ± 40	s^{-1}
ω_{40}	0.1	1	s^{-1}
ω_{76}	140 ± 20	1100 ± 330	s^{-1}
ω_{78}	0.6 ± 0.1	7 ± 2	s^{-1}
$\omega_{9,10}$	0.06 ± 0.006	0.26 ± 0.04	s^{-1}
ω_{90}	0.84 ± 0.08	4 ± 0.7	s^{-1}
ω_{pro}	3 ± 1	150 ± 50	s^{-1}

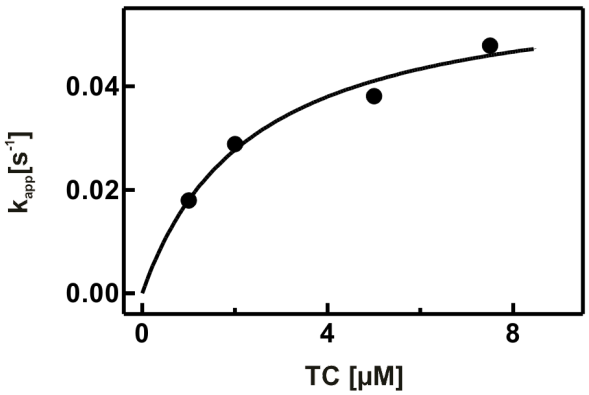


Figure 4.2: ***In vitro* values of the rates ω_{90} and $\omega_{9,10}$ for near-cognate rejection and incorporation at 20 °C.** As described in [30] for 37 °C, the apparent rate k_{app} of dipeptide formation was measured for increasing concentrations of near-cognate TC at 20 °C, revealing $\omega_{9,10} = 0.060 \pm 0.006\,\mathrm{s}^{-1}$ for the rate of near-cognate incorporation. In a second step, $\omega_{90} = 0.84 \pm 0.08\,\mathrm{s}^{-1}$ for the rate of near-cognate rejection was calculated from the proofreading efficiency [40]. The experiments were performed and analyzed by Michael Thommen, Max Planck Institute for Biophysical Chemistry, Göttingen, who also created this figure.

4.5 Experimental Methods

For the experimental part of this work, we collaborated with Marina V. Rodnina, head of the Physical Biochemistry department at the Max Planck Institute for Biophysical Chemistry, and Michael Thommen, a PhD student in her group. In this section, we describe the experiments performed by our collaborators.

4.5.1 *In Vitro* Transition Rates ω_{90} and $\omega_{9,10}$ at 20 °C

Following an experimental protocol described in [30], our collaborators monitored the formation of radioactively labeled f[^{3}H]Met[^{14}C]Phe dipeptides under multiple-turnover conditions. They used initiation complexes comprised of 70S ribosomes, a short mRNA made of the start codon and one CUC codon, and labeled initiator tRNA (70S·mRNA(AUGCUC)·f[^{3}H]Met-tRNA$^{\mathrm{fMet}}$, 0.14 μM) as well as varying concentrations of ternary complexes with labeled tRNA$^{\mathrm{Phe}}$ (EF-Tu·GTP·[^{14}C]Phe-tRNA$^{\mathrm{Phe}}$), which is near-cognate to the CUC codon. From the linear slopes of the dipeptide formation

time courses, the experimenters determined the apparent rate of dipeptide formation k_{app} for each concentration of ternary complexes, see fig. 4.2. From the resulting hyperbolic concentration dependence of k_{app}, they calculated $\omega_{9,10} = 0.060 \pm 0.006\,\mathrm{s}^{-1}$ for the rate of near-cognate incorporation. Using the previously measured efficiency $\omega_{9,10}/(\omega_{9,10} + \omega_{90}) = 1/15$ of the proofreading step [40], our collaborators then obtained the rate $\omega_{90} = 0.84 \pm 0.08\,\mathrm{s}^{-1}$ for near-cognate rejection.

4.5.2 *In Vitro* Overall Elongation Rate ω_{elo} at 20 °C and 37 °C

Our collaborators prepared mRNA encoding the protein CspA from *E. coli* by T7 RNA-polymerase transcription. Then, they synchronized ribosomes by forming initiation complexes consisting of a 70S ribosome, CspA mRNA and an initiator $\mathrm{tRNA}^{\mathrm{fMet}}$ carrying the fluorophore BODIPY-FL at the α-amino group of methionine in the presence of initiation factors (IF1, IF2, and IF3) and GTP.

The experimenters carried out translation in a fully reconstituted translation system by adding initiation complexes (15 nM) to a mixture of EF-Tu·GTP·aminoacyl-tRNA (40 µM aminoacyl-tRNA, 100 µM EF-Tu in total), EF-G (3 µM), GTP (2 mM), phosphoenol pyruvate (6 mM), and pyruvate kinase (0.1 mg mL^{-1}) in HiFi buffer (50 mM Tris-HCl, pH 7.5, 30 mM KCl, 70 mM NH_4Cl, 3.5 mM free $MgCl_2$, 0.5 mM spermidine, and 8 mM putrescine) at either 20 °C or 37 °C [44]. Due to the absence of translation termination and ribosome recycling factors, translation was limited to a single round, i.e., at most one copy of the CspA protein was synthesized per ribosome.

Our collaborators stopped the reactions at specific time intervals and separated the translation products, i.e., peptides of different lengths, by 16.5 % Tris-Tricine-PAGE (polyacrylamide gel electrophoresis). On the gel, the peptides were visualized by the fluorescent reporter BODIPY-FL attached to the N-terminus of the peptides [55], see fig. 4.3. Each band on the gel carries peptides from a specific range of lengths, where smaller peptides are further away from the top and the full length product is found in the upmost band. The intensities of the different bands were quantified and normalized with the image analysis software *ImageJ* to obtain information on the relative abundances of the peptides at the different time intervals after initiation of translation. Finally, from the time course of relative full length product abundance, the overall elongation rate ω_{elo} was determined (right panel of fig. 4.3). For translation at 20 °C and 37 °C, our collaborators obtained $\omega_{\mathrm{elo}} = 0.8 \pm 0.2\,\mathrm{s}^{-1}$ and $\omega_{\mathrm{elo}} = 6.9 \pm 2.3\,\mathrm{s}^{-1}$, respectively.

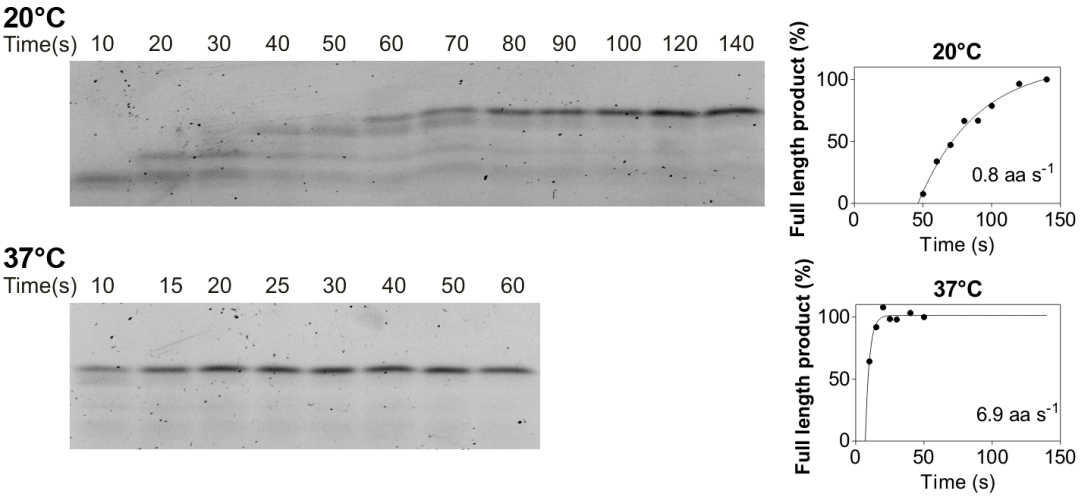


Figure 4.3: ***In vitro* kinetics of CspA translation at** 20 °C **and** 37 °C**.** Left panels: CspA peptides separated by gel electrophoresis and visualized by a fluorescent reporter at different time intervals after initiation of translation. The normalized intensities of the bands of each time interval represent the relative abundances of the corresponding peptides. Full length peptide (70 amino acids) appears in the upmost band, shorter peptides are found below.
Right panels: Deducing the overall elongation rate ω_{elo} by exponential fitting of the time course of the relative amount of full length peptide. Experiments were performed and analyzed by Michael Thommen, Max Planck Institute for Biophysical Chemistry, Göttingen, who also created this figure.

5 From *In Vitro* to *In Vivo*

Over the past two decades, the various substeps of the translation process have been studied in much detail using *in vitro* systems [38–44], but it is often questioned to what extend these results can be applied to living cells [62]. Although *in vivo* translation should proceed via essentially the same steps as *in vitro*, the overall protein synthesis rate turns out to be much faster [52]. This difference in synthesis rate has led to a long, controversial, and fruitless debate, because – in contrast to *in vitro* systems – it has not been possible to study individual conformational transitions of the translating ribosome *in vivo*. In addition, a physical measure to quantitatively compare the effects of different environments on the kinetics of translation was missing.
In this chapter, we address this long-lasting puzzle of *in vitro* and *in vivo* comparability and introduce such a measure in terms of a distance metric for transition rates that we call the *kinetic distance*.

We use the codon-specific Markov process defined in section 3.2, together with the sets of *in vitro* rates at 37 °C in table 4.3, and determine *in vivo* rates for the internal transitions of the ribosome. We do so by minimizing the kinetic distance between the known *in vitro* and the unknown *in vivo* rates, taking the overall speed of the ribosomes into account, see section 5.2. The deduced transition rates allow the prediction of local translation speeds, translational error frequencies, and gene specific translation dynamics. As we show in section 5.4, the predictions are in good agreement with three independent sets of experimental *in vivo* data, without adjusting any fit parameter.
Our results imply that translation in the cell can be accurately described by transition rates that deviate from those measured *in vitro* by a factor of 8 at most, which can be interpreted as shifts of free energy barriers of about $2\,k_BT$ under the assumption that the attempt frequencies do not change. The predicted codon- and mRNA-specific translation rates could be used, e.g., to investigate translational pauses and ribosome traffic. The concept of a distance metric for kinetic rates and the minimization method introduced here are quite general and could be used to derive *in vivo* rates for other biomolecular processes for which we have detailed information about the *in vitro* kinetics.

We published the method presented in this chapter in [63] where it was applied to a simplified model of bacterial translation. In this chapter, we will consider the full theory of translation developed in the previous chapters for both the 2-1-2 and the 2-3-2 pathway of tRNA release from the ribosomal E site.

5.1 Fundamental Differences of Translation *In Vivo* and *In Vitro*

In vitro and *in vivo* translation are distinct in some fundamental aspects, rendering the prediction of *in vivo* translation rates from *in vitro* experiments difficult. Going from an *in vitro* translation system to translation *in vivo* is like changing the buffer conditions, plus introducing a crowded environment and adding additional molecules, such as chaperones, that might take an active role in the translation process. In particular, we cannot simply assume that the transition rates measured *in vitro* and listed in table 4.3 apply to the *in vivo* system as well. On the contrary, at least some of the rates must be different from their *in vitro* counterparts. This difference becomes obvious if one compares the overall *in vitro* elongation rate to the overall *in vivo* elongation rate: Whereas *in vitro* translation proceeds at an overall speed of about 7 amino acids per second (aa/s), see table 4.3, the overall *in vivo* elongation rate is larger by at least a factor of two and can get – depending on the growth conditions – as high as 22 aa/s [52].

Another major difference between translation *in vitro* and *in vivo* is the concentration of free ternary complexes (TCs): *In vitro*, the concentration of TCs can be controlled and can in particular be chosen to exceed the concentration of translating ribosomes by several orders of magnitude. Therefore, the difference between total tRNAs and free TCs can be neglected in *in vitro* experiments. In contrast, *in vivo* the total concentration of tRNAs exceeds the concentration of ribosomes only by a factor of about 10 [49]. Because every ribosome can bind up to three tRNAs, the fraction of bound tRNAs can become large. As long as a tRNA is bound to a ribosome, it is not available for TC formation. This difference between total and available tRNAs is quite substantial for the 2-3-2 pathway of E site tRNA release but is also important for the 2-1-2 pathway. A consequence of the lower tRNA to ribosome ratio is that the recharging of tRNAs with amino acids and the subsequent formation of new TCs becomes important *in vivo*, whereas *in vitro* systems often do not include the components necessary for re-aminoacylation of tRNAs, such as synthetases and additional ATP. Because naked tRNAs do not participate in the elongation cycle and recharging of tRNAs is performed with finite speed, the *in vivo* availability of TCs is reduced. In addition, the *in vivo* concentrations of TCs are also limited by the concentration of EF-Tu.

5.2 *In Vivo* Binding Rates

In the following two sections, we will determine the *in vivo* values of the binding rates and the internal transition rates of the codon-specific Markov process defined in section 3.2. As a convention of notation, the *in vivo* values are indicated by star (*) superscripts.

In vivo, the diffusion of TCs and, thus, their binding to ribosomes is slowed down by molecular crowding, i.e, the dense packing of molecules in the cytosol. Therefore, the *in vivo* value of the binding rate constant $\kappa_{\text{on}}^{\star}$ can be estimated as follows:
First, a TC has to reach the ribosome by diffusive motion. We denote the average time that a certain ribosome has to wait until a TC of species a comes close to it by t_{dif}. *In vivo*, this average time depends on the cellular concentration $\hat{\text{X}}_a$ of TCs, the diffusion constant D of a TC, and the radius r of a ribosome. This diffusion time t_{dif} can be calculated under the simplifying assumption that a TC in a cell that is much bigger than a ribosome undergoes diffusion or Brownian motion with reflecting boundaries provided by the cell walls, until the TC reaches the chosen ribosome. This diffusion time has the form [64]

$$t_{\text{dif}} = \frac{1}{4\pi\, D\, r\, \hat{\text{X}}_a}\,. \tag{5.1}$$

Then, the diffusion related binding rate constant κ_{dif} is simply given by

$$\kappa_{\text{dif}} = \left(t_{\text{dif}}\, \hat{\text{X}}_a\right)^{-1} = 4\pi\, D\, r\,. \tag{5.2}$$

Using a diffusion constant of $2.57\,\mu\text{m}^2/\text{s}$ [32, 65] for a TC in the cytosol and a ribosome radius of 10.5 nm [66], the diffusion related binding rate constant is found to be

$$\kappa_{\text{dif}} = 0.34\,\mu\text{m}^3/\text{s} = 204/(\mu\text{M}\,\text{s})\,.$$

In a second step, the TC binds to the ribosome with the same binding rate constant as *in vitro* $\kappa_{\text{on}} = 175/(\mu\text{M}\,\text{s})$. Note that due to very fast diffusion of ternary complexes in the aqueous buffer used for *in vitro* translation experiments, the diffusive contribution to the *in vitro* binding rate constant κ_{on} can be neglected. The total effective *in vivo* binding rate constant $\kappa_{\text{on}}^{\star}$ is given by

$$\kappa_{\text{on}}^{\star} = \frac{\kappa_{\text{dif}}\kappa_{\text{on}}}{\kappa_{\text{dif}} + \kappa_{\text{on}}} = 94/(\mu\text{M}\,\text{s})\,, \tag{5.3}$$

which is only about 54% of the *in vitro* value κ_{on}.

5.3 *In Vivo* Transition Rates

After the initial binding of a cognate, near-cognate, or non-cognate TC, the kinetics of the elongation process are determined by the 12 internal transition rates defined in section 3.3. In the following we will describe a method to predict the *in vivo* values of these internal transition rates from their *in vitro* counterparts. Because *in vivo* experi-

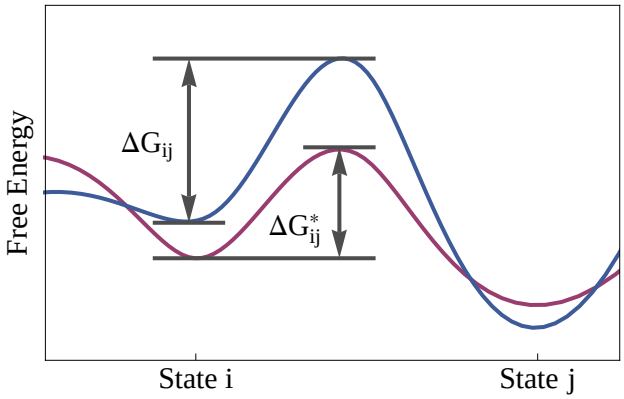


Figure 5.1: **Physical interpretation of the logarithmic difference Δ_{ij}.** Schematic free energy landscape under *in vitro* (blue) and *in vivo* (purple) conditions with the *in vitro* free energy barrier ΔG_{ij} and the *in vivo* barrier $\Delta G^{\star}_{ij}$ for the transition from state i to state j. Under the assumption that the attempt frequency ν_{ij} does not change, the logarithmic difference Δ_{ij} is approximately proportional to the single barrier shift $\Delta G^{\star}_{ij} - \Delta G_{ij}$ as stated in eq. 5.7.

ments with *E. coli* cells are typically performed at 37 °C, we start from the corresponding *in vitro* set in table 4.3. In order to specify a mapping of these 12 internal rates from *in vitro* to *in vivo* in a unique manner, we first introduce a metric to describe the similarity of two sets of kinetic rates that we will call *kinetic distance*. We then look for *in vivo* rates that have the smallest kinetic distance from the *in vitro* rates at 37 °C but at the same time reproduce the measured *in vivo* value of the overall elongation rate.

5.3.1 Kinetic Distance

In the following, we introduce a new metric that we call *kinetic distance*, which describes the similarity of two sets of kinetic rates. First, we define the kinetic distance $\mathcal{D}_{ij}(\omega_{ij}, \omega^{\star}_{ij})$ of the *in vitro* and the *in vivo* value ω_{ij} and $\omega^{\star}_{ij}$ of an individual transition rate in such a way that it only depends on the ratio of these two values. Furthermore, we require the kinetic distance to be symmetric with respect to exchanging ω_{ij} and $\omega^{\star}_{ij}$. The simplest choice for the kinetic distance that fulfills these requirements is given by

$$\mathcal{D}_{ij}(\omega_{ij}, \omega^{\star}_{ij}) \equiv \left|\ln\left(\omega^{\star}_{ij}/\omega_{ij}\right)\right| = |\Delta_{ij}|\,, \tag{5.4}$$

with the logarithmic difference

$$\Delta_{ij} \equiv \ln \omega_{ij} - \ln \omega_{ij}^{\star} = -\ln\left(\omega_{ij}^{\star}/\omega_{ij}\right) \tag{5.5}$$

between the *in vitro* and the *in vivo* value of the individual transition rate.
Note that apart from its simplicity, this expression also has a direct physical interpretation for all 12 internal transition rates of the ribosome: A rate ω_{ij} which describes the transition from state i to state j is governed by the associated free energy barrier ΔG_{ij} between these states and can be written in the form

$$\omega_{ij} = \nu_{ij} \exp(-\Delta G_{ij}/\mathrm{k_B T}) \,, \tag{5.6}$$

where the prefactor ν_{ij} represents the attempt frequency and the thermal energy $\mathrm{k_B T}$ provides the basic free energy scale. The logarithmic difference between the *in vitro* and the *in vivo* value of a transition rate is then given by

$$\Delta_{ij} = [\Delta G_{ij}^{\star} - \Delta G_{ij}]/(\mathrm{k_B T}) - \ln(\nu_{ij}^{\star}/\nu_{ij}) \,, \tag{5.7}$$

which is dominated by the first term, i.e., by the shift of the free energy barrier between state i and state j, if the attempt frequencies ν_{ij} and $\nu_{ij}^{\star}$ are approximately the same. Therefore, in the following we refer to Δ_{ij} as the *single barrier shift* instead of using the less intuitive term *logarithmic difference*. Note that despite this terminology, the quantity Δ_{ij} always contains both the shift of the free energy barrier and the attempt frequency as in eq. (5.7).
Second, we define the overall kinetic distance $\mathcal{D}(\Omega, \Omega^{\star})$ between the *in vitro* and the *in vivo* kinetics Ω and $\Omega^{\star}$ by

$$\mathcal{D}(\Omega, \Omega^{\star}) \equiv \sqrt{\sum \Delta_{ij}^2} = \sqrt{\left[\ln\left(\omega_{ij}^{\star}/\omega_{ij}\right)\right]^2} \,, \tag{5.8}$$

where the summation under the square root runs over all internal transitions. Note that in case of the translation process discussed here $\Omega = (\omega_{\mathrm{off}}, \omega_{\mathrm{rec}}, \omega_{21}, ...)$, $\Omega^{\star} = (\omega_{\mathrm{off}}^{\star}, \omega_{\mathrm{rec}}^{\star}, \omega_{21}^{\star}, ...)$, and the sum has to be taken over 17 internal transitions of which only 12 are independent, as is discussed in section 3.3.
The kinetic distance provides a genuine metric in the mathematical sense such that

$$\mathcal{D}(\Omega_1, \Omega_2) \geq 0 \,, \tag{5.9}$$

$$\mathcal{D}(\Omega_1, \Omega_2) = 0 \iff \Omega_1 = \Omega_2, \tag{5.10}$$

$$\mathcal{D}(\Omega_1, \Omega_2) = \mathcal{D}(\Omega_2, \Omega_1) \,, \tag{5.11}$$

$$\mathcal{D}(\Omega_1, \Omega_3) \leq \mathcal{D}(\Omega_1, \Omega_2) + \mathcal{D}(\Omega_2, \Omega_3) \,, \tag{5.12}$$

where Ω_1, Ω_2, and Ω_3 represent different kinetics of a certain multistep process.

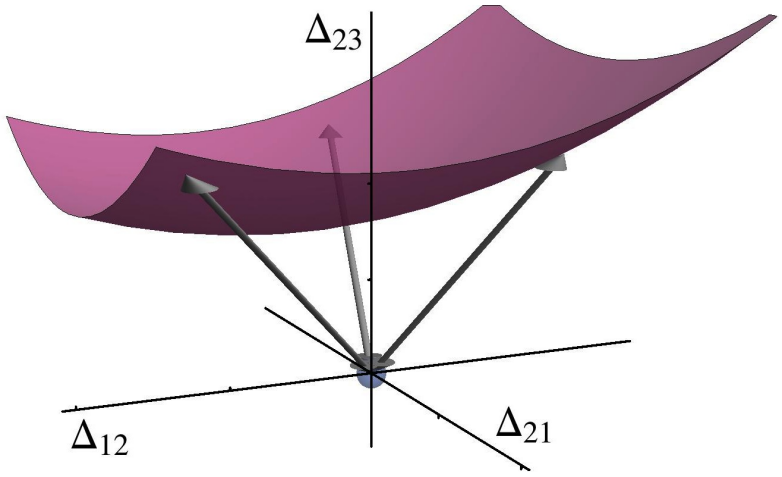


Figure 5.2: **Cartoon of a hypersurface in the multi-dimensional barrier space.** Three-dimensional section of the multi-dimensional barrier space with coordinates Δ_{21}, Δ_{12}, and Δ_{23}. The origin of this space (light blue dot) corresponds to the *in vitro* system. The surface (purple) represents a two-dimensional section of the hypersurface described by eq. (5.13), corresponding to a fixed *in vivo* value for the overall elongation rate. Each point on this surface has a certain kinetic distance that is equal to the Euclidean distance of this point from the origin, as indicated by the three double arrows. The point with the shortest kinetic distance determines the predicted *in vivo* rates $\omega^\star_{ij,\min}$.

5.3.2 Constrained Minimization of the Kinetic Distance

In this section, we predict the *in vivo* values of the internal transition rates from the measured *in vitro* rates. We do so by minimizing the distance $\mathcal{D}$ as given by eq. (5.8) between the known *in vitro* rates and a predicted set of *in vivo* rates under the constraint that the deduced *in vivo* rates reproduce the measured overall elongation rate $\omega^\star_{\text{elo}}$ *in vivo*. In other words, we determine those *in vivo* rates that are consistent with the overall elongation rate $\omega^\star_{\text{elo}}$ measured in living cells and have the smallest kinetic distance from the *in vitro* rates.

In order to impose this constraint, we express the overall elongation rate in terms of the individual transitions, see eq. (3.31). The *in vivo* values $\omega^\star_{ij}$ of the individual transition rates must reproduce the experimentally measured value $\omega^\star_{\text{elo}}$ of the overall elongation rate. This requirement implies the equation

$$\omega_{\text{elo}}\left(\omega^\star_{\text{off}}, \omega^\star_{\text{rec}}, \omega^\star_{21}, ...\right) = \omega^\star_{\text{elo}}\,, \tag{5.13}$$

which represents a constraint on the unknown *in vivo* values $\omega^\star_{ij}$. This constraint can be expressed in terms of the single barrier shifts using the inverse coordinate transformation $\omega^\star_{ij} = \omega_{ij} \exp(-\Delta_{ij})$ with the known *in vitro* values ω_{ij}. As a result, the constraint in eq. (5.13) defines a hypersurface in the multi-dimensional barrier space as illustrated in fig. 5.2. Each point on this hypersurface is compatible with the measured value $\omega^\star_{\mathrm{elo}}$ of the overall elongation rate. This means in particular, that generally every set of rates corresponding to a point on this hypersurface represents *a priori* a possible set of hypothetical *in vivo* translation rates. The Euclidean distance of such a point from the origin is equal to the kinetic distance between the hypothetical *in vivo* and the measured *in vitro* values of the transition rates. Our prediction for the *in vivo* values $\omega^\star_{ij,\mathrm{min}}$ is then obtained by minimizing this kinetic distance, i.e, by the point on the hypersurface that has the shortest distance from the origin, which is motivated as follows.

The kinetic distance minimization approach for the prediction of *in vivo* translation rates is based on the following assumptions and motivations. First, we make the usual assumption that the states of the translation system that have been identified *in vitro* are also present *in vivo*, because the *in vitro* assay is functional and has been optimized to resemble *in vivo* translation over the past decades. Second, it is then plausible to assume that the *in vitro* and *in vivo* values of the individual transition rates do not differ by many orders of magnitude, which implies that the point in the multi-dimensional barrier space that represents the true *in vivo* rates is located in some neighborhood of the origin of this space and, thus, is characterized by a rather small kinetic distance $\mathcal{D}$. More precisely, we assume that well performed *in vitro* experiments enable us to choose from all hypothetical sets of *in vivo* rates that one, which is most likely to reflect the true *in vivo* situation. This is a fundamental assumption and justification for all *in vitro* experiments aiming at elucidating processes in living cells. However, this assumption might not always hold, for example when *in vitro* experiments are not performed under optimal conditions. Therefore, the validity of *in vivo* rates predicted by the minimization of the kinetic distance can not be assessed *a priori* but needs to be checked *a posteriori*, see section 5.4.

The smallest kinetic distance that is compatible with the *in vivo* constraint as given by eq. (5.13) defines the radius of a sphere around the origin that touches the hypersurface. This radius is equal to the Euclidean distance of the hypersurface from the origin of the coordinates. The associated contact point between the minimal sphere and the hypersurface represents the predicted *in vivo* point, and its coordinate values lead to the predicted *in vivo* rates $\omega^\star_{ij,\mathrm{min}}$ based on the known *in vitro* rates ω_{ij}.
In practice, the minimization is performed numerically by using the software *Mathematica* by Wolfram Research, Inc., and its build-in function *NMinimize*. The results are shown in fig. 5.3 and table 5.1, with the latter summarizing all *in vivo* ribosomal transition rates as well as the overall elongation rate, both for the 2-1-2 and the 2-3-2 pathway

of E site tRNA release. For both pathways, most *in vivo* rates do not differ much from their *in vitro* counterparts. The largest differences are found for the unbinding rate $\omega^\star_{\text{off}}$, and the recognition rate $\omega^\star_{76}$, and in particular for the near-cognate rejection rate $\omega^\star_{76}$ that is increased by a factor of 8 compared to its *in vitro* value. However, even the latter change corresponds to a barrier shift Δ_{76} of only about -2. If we could assume that the attempt frequency is not changed, this barrier shift would correspond to a change in the free energy barrier of $-2\,\mathrm{k_B T}$, see eq. (5.7), which could easily arise for example from changes in the hydrogen bond networks around the ribosome/TC complex. However, at present it is not possible to estimate changes in attempt frequencies for the ribosomal transitions from the available experimental data.

5.4 Validation of *In Vivo* Rates

The location of the *in vivo* point in the multi-dimensional barrier space cartooned in fig. 5.2 will in general depend on the rates of the chosen *in vitro* assay. More precisely, the change from one *in vitro* assay to another corresponds to a Euclidean translation of the coordinate system while the shape of the hypersurface defined by eq. (5.13) remains unchanged. Therefore, as we stated above, the validity of the *in vivo* rates $\omega^\star_{ij,\min}$ predicted by the minimization of the kinetic distance is difficult to assess *a priori* but needs to be checked *a posteriori* in a self-consistent manner.

To do so, we compute the codon-specific elongation rates $\omega^\star_{c,\text{elo}}$ as given by eq. (3.10) from the *in vivo* transition rates displayed in table 5.1, see fig. 5.4 and tables D.1 and D.2 in the appendix. In fig. 5.5 (A), we compare the calculated *in vivo* rates $\omega^\star_{c,\text{elo}}$ for *E. coli* cells with a growth rate of 2.5 dbl/h and a 2-1-2 E site tRNA release pathway to *relative translation rates* as obtained in [19] from measured +1 frameshifting vs. readthrough frequencies of different codons. More precisely, the authors of this paper measured how often a specific codon induces a +1 frameshift, i.e., a slight downstream shift of the translating ribosome by one nucleotide leading to an alternative reading frame of the message and, finally, resulting in a completely different peptide chain. They argue that the probability of a codon to induce a frameshift should be proportional to the time that the ribosome dwells at this codon and, thus, be inversely proportional to the translation rate of that particular codon.

For both E site release pathways, we find a reasonable overall agreement between both data sets with a squared Pearson correlation coefficient of 0.25, where the deviations reflect both uncertainties in the experimental method and limitations of the theory as follows from a more detailed analysis of the data. First, the calculated elongation rates for CGA, CGC, and CGU appear to be overestimated. These codons are all read by $\mathrm{tRNA^{Arg2}}$, which does not form a Watson-Crick base pair with any of its cognate codons

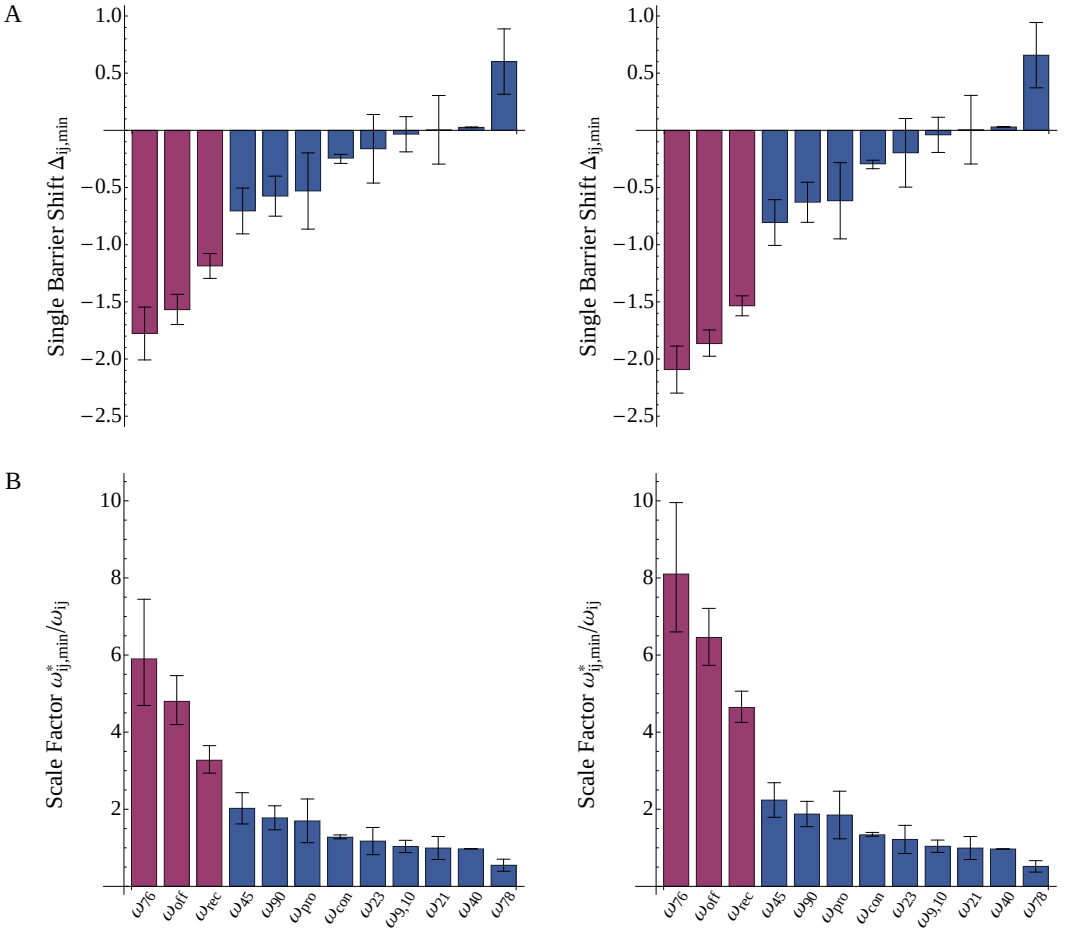


Figure 5.3: **Single barrier shifts $\Delta_{ij,\min}$ and scale factors $\omega^\star_{ij,\min}/\omega_{ij}$ as predicted by the minimization procedure, assuming either a 2-1-2 (left column) or a 2-3-2 (right column) pathway of tRNA release from the E site.** (A) Single barrier shifts $\Delta_{ij,\min}$ that minimize the kinetic distance between *in vitro* and *in vivo* translation. Under the assumption of approximately constant attempt frequencies, the $\Delta_{ij,\min}$ are equal to shifts in free energy barriers in units of k_BT, see eq. 5.7. (B) Scale factors $\omega^\star_{ij,\min}/\omega_{ij} = \exp(-\Delta_{ij,\min})$ for all internal transitions of the ribosome/TC complex. The $\Delta_{ij,\min}$ and, thus, the scale factors were calculated using the *in vitro* values ω_{ij} in table 4.3 for 37 °C. A single barrier shift < 0 leads to a scale factor > 1 and implies that the *in vivo* rate is increased compared to the *in vitro* rate. Transition rates with scaling factors well above 2 are highlighted in purple. Error bars reflect uncertainties in the *in vitro* rates as is explained in section 5.5.

Table 5.1: ***In vivo* rates of ribosomal transitions.** The values of the overall elongation rate $\omega^\star_{\text{elo}}$ for the four growth rates 0.7, 1.07, 1.6, and 2.5 dbl/h were obtained from the data in [52]. These growth rates have been chosen because the total tRNA concentrations have been measured for these conditions [3] as well. The *in vivo* rates of ribosomal transitions were obtained under the assumption of a 2-1-2 (top) or a 2-3-2 (bottom) pathway of tRNA release from the ribosomal E site. Relative standard deviations (RSDs) were obtained from the uncertainties of the *in vitro* rates in table 4.3, see section 5.5.

E site release	Rates	*E. coli* growth rates [dbl/h]				RSD	Units
		0.7	1.07	1.6	2.5		
	$\omega^\star_{\text{elo}}$	15	18	22	22		aa/s
	$\kappa^\star_{\text{on}}$	94	94	94	94	0.1	$\mu\text{M}^{-1}\,\text{s}^{-1}$
2-1-2	$\omega^\star_{\text{off}}$	2000	2100	2900	3400	0.4	s^{-1}
	$\omega^\star_{\text{rec}}$	3000	3100	4200	4900	0.3	s^{-1}
	$\omega^\star_{21}$	2	2	2	2	0.3	s^{-1}
	$\omega^\star_{23}$	1600	1600	1700	1800	0.3	s^{-1}
	$\omega^\star_{\text{con}}$	510	520	560	570		s^{-1}
	$\omega^\star_{45}$	300	320	390	400	0.2	s^{-1}
	$\omega^\star_{40}$	1	1	1	1		s^{-1}
	$\omega^\star_{76}$	3900	4000	5500	6500	0.3	s^{-1}
	$\omega^\star_{78}$	4	4	4	4	0.3	s^{-1}
	$\omega^\star_{9,10}$	0.27	0.27	0.27	0.27	0.2	s^{-1}
	$\omega^\star_{90}$	6	6	7	7	0.2	s^{-1}
	$\omega^\star_{\text{pro}}$	200	220	250	260	0.5	s^{-1}
2-3-2	$\omega^\star_{\text{off}}$	2100	2200	3300	4500	0.4	s^{-1}
	$\omega^\star_{\text{rec}}$	3200	3400	4900	7000	0.3	s^{-1}
	$\omega^\star_{21}$	2	2	2	2	0.3	s^{-1}
	$\omega^\star_{23}$	1600	1700	1800	1800	0.3	s^{-1}
	$\omega^\star_{\text{con}}$	510	530	570	600		s^{-1}
	$\omega^\star_{45}$	310	330	410	450	0.2	s^{-1}
	$\omega^\star_{40}$	1	1	1	1		s^{-1}
	$\omega^\star_{76}$	4100	4300	6300	8900	0.3	s^{-1}
	$\omega^\star_{78}$	4	4	4	4	0.3	s^{-1}
	$\omega^\star_{9,10}$	0.27	0.27	0.27	0.27	0.2	s^{-1}
	$\omega^\star_{90}$	6	7	7	8	0.2	s^{-1}
	$\omega^\star_{\text{pro}}$	210	220	260	280	0.5	s^{-1}

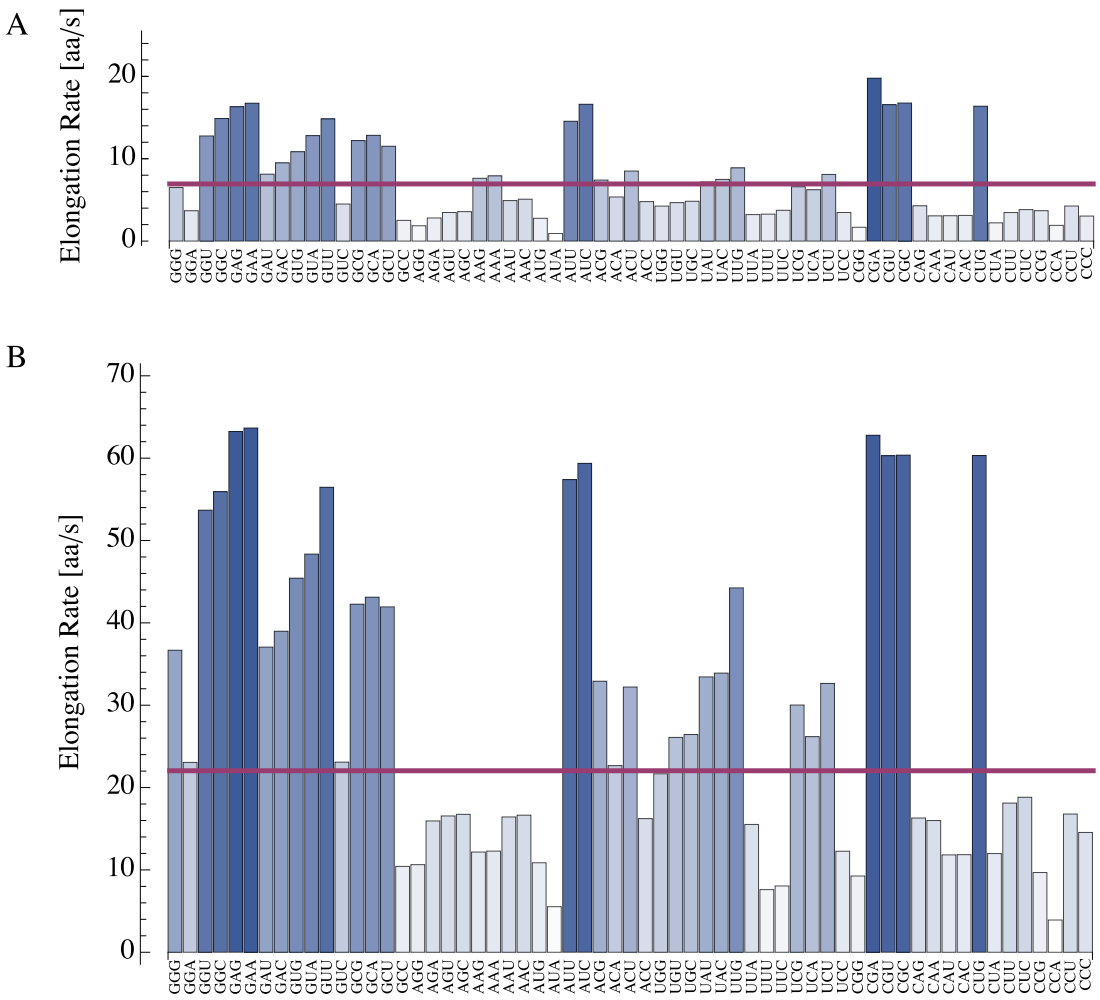


Figure 5.4: **Codon-specific elongation rates *in vitro* and *in vivo*.** (A) *In vitro* values $\omega_{c,\mathrm{elo}}$ at 37 °C, obtained from eq. (3.10) and the individual rates in table 4.3; (B) *In vivo* values $\omega^{\star}_{c,\mathrm{elo}}$ for *E. coli* at a growth rate of 2.5 dbl/h, calculated from eq. (3.10) and the individual rates in table 5.1. The horizontal lines (purple) indicate the values of the corresponding overall elongation rates ω_{elo} and $\omega^{\star}_{\mathrm{elo}}$, respectively. Note that the ratios $\omega^{\star}_{c,\mathrm{elo}}/\omega_{c,\mathrm{elo}}$ are codon-specific as well: Translation of some codons is much less sped up *in vivo* compared to others, for example $\omega^{\star}_{\mathrm{AAA,elo}}/\omega_{\mathrm{AAA,elo}} = 1.6$ and $\omega^{\star}_{\mathrm{AAU,elo}}/\omega_{\mathrm{AAU,elo}} = 3.4$.

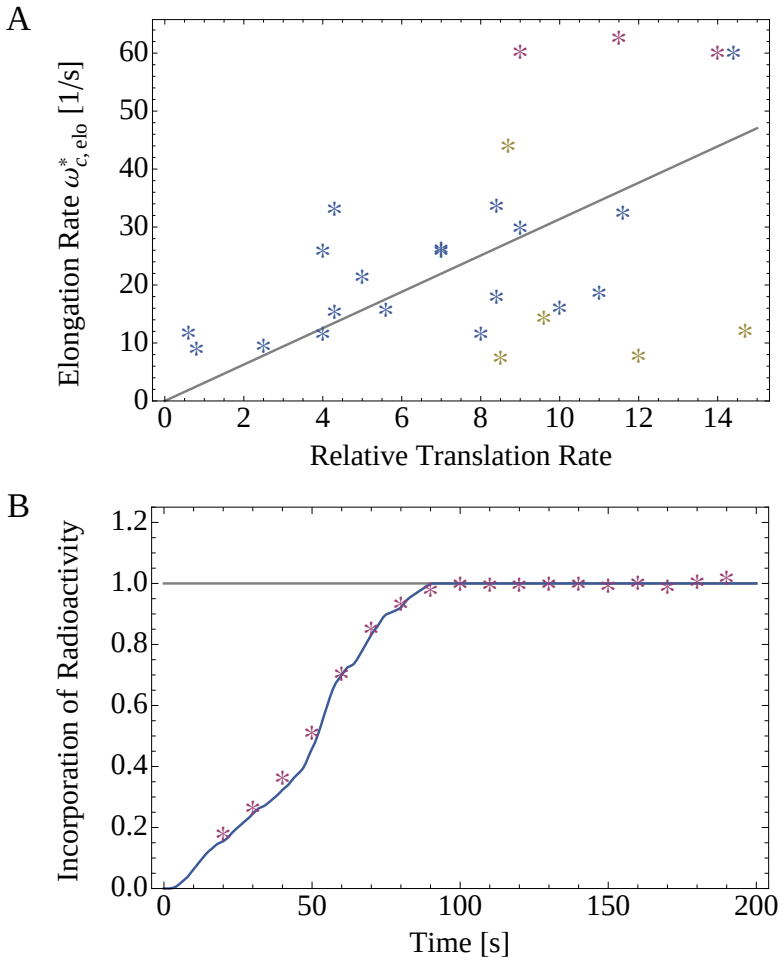


Figure 5.5: **Comparison of theoretical results to *in vivo* experiments.** (A) Codon-specific elongation rates $\omega^{\star}_{c,\mathrm{elo}}$ as determined from the complete set of individual transition rates for *E. coli* at a growth rate of 2.5 dbl/h and a 2-1-2 E site release pathway, see table D.1 in the appendix and fig. 5.4, compared to relative translation rates as measured in [19] for 29 codons; highlighted symbols indicate the codons CGA, CGC, and CGU (purple) as well as the codons UUU, UUC, UUG, UCC, and CCC (yellow). The squared Pearson correlation coefficient is 0.25 for all codons and 0.5 when the highlighted codons are excluded (linear fit in gray). (B) For the incorporation of radioactively labeled amino acids as a function of time, we find very good agreement between experimental data from [18] (purple) and the calculated curve (blue) based on the *in vivo* transition rates for 0.7 dbl/h (table 5.1), without adjusting any fit parameter.

because it carries inosine at the wobble position of its anticodon ICG. The corresponding reductions in the decoding rates are not included in our model because we use only two different sets of values for these rates, corresponding to an average over all cognate and an average over all near-cognate TCs, respectively, see section 4.3. In principle, we could expand our model to include codon-anticodon dependent differences in the decoding rates. However, to limit the number of parameters as well as due to a limited availability of experimental data, we neglect these differences when calculating the elongation rates $\omega^{\star}_{c,\mathrm{elo}}$.
Second, for the experimental setup in [19], the codons UUU, UUC, UUG, UCC, and CCC, when located between a preceding CUU codon and a subsequent CXX codon, generate potential slippery sequences, which can lead to -1 frameshifting events. During a -1 frameshift the ribosome experiences an upstream shift by one nucleotide. At the time when [19] was published, the existence of -1 frameshifting was unknown, whereas nowadays it is supposed to occur even more frequently than $+1$ frameshifting. Therefore, -1 frameshifting events were not taken into account in [19], which implies that the frameshifting rates were underestimated and consequently the translation rates were overestimated for the respective codons. When we exclude these two particular sets of codons, we obtain an increased squared correlation coefficient of 0.5, see fig. 5.5 (A). Thus, we conclude that our reduced set of decoding rates leads to a reliable prediction of codon-specific translation rates for the majority of codons.

To further validate the predicted *in vivo* transition rates $\omega^{\star}_{ij,\mathrm{min}}$, we performed stochastic simulations of the translation of the lacZ gene in *E. coli* and checked whether the calculated elongation rates reflect real *in vivo* situations by comparing the simulation results to available experimental data. We performed simulations to take possible ribosome-ribosome interactions into account that are not part of our analytic framework of translation elongation but could in principle alter the dynamics of *in vivo* translation. A description of the simulation method is given in appendix B. Stochastic simulations of the translation process are based on the elongation rates $\omega^{\star}_{c,\mathrm{elo}}$ of all codons for a growth rate of 0.7 dbl/h.
Experimentally, protein synthesis was studied using radioactively labeled amino acids, as described in [18]. Briefly, in these experiments cells with a growth rate of about 0.7 dbl/h and overexpressing the lacZ gene have been exposed to a 10-seconds pulse of radioactively labeled S-methionine, and the radioactivity of the protein product was measured over time [18]. The simulated time course is in excellent agreement with the experimental data, see fig. 5.5 (B).

Another quantity that can be used to validate the predicted *in vivo* rates is the missense error frequency arising from the accommodation of near-cognate TCs with an incorrect amino acid. The calculated error frequency depends on all individual transitions for the accommodation of a cognate or near-cognate aa-tRNA and, in particular,

on the concentrations of free cognate and near-cognate TCs, whereas it is independent of the concentrations of non-cognates, see eq. (3.29) in section 3.4. Using the predicted *in vivo* rates as given in table 5.1 and calculating the free TC concentrations from the total tRNA concentrations for 0.7 dbl/h, see section 3.7, we obtain for both pathways of tRNA release from the E site an average missense error frequency of 2×10^{-4} for $\mathrm{tRNA^{Lys}}$ misreading codons, in perfect agreement with the measured value [4]. The concentrations of free TCs are listed in tables D.3 and D.4 for the 2-1-2 and the 2-3-2 pathway of E site tRNA release, respectively.

5.5 Estimation of Uncertainties

The rates ω_{ij} of the *in vitro* assay are only known with a certain accuracy. As a consequence, the predicted *in vivo* rates $\omega^\star_{ij}$ have some uncertainty as well. In this section we describe how these uncertainties are estimated.

The Markov process for ribosomal elongation considered here involves 12 distinct transition rates, see fig. 3.2, which implies that the corresponding barrier space has 12 dimensions. We first determined the coordinate values $\Delta_{ij,\min}$ of the *in vivo* point as predicted from the *in vitro* rates ω_{ij}. The largest coordinate values $\Delta_{ij,\min}$ of the predicted *in vivo* point were found for the three transition rates ω_{76}, ω_{off}, and ω_{rec}, see fig. 5.3. We then focused on the errors ϵ_{ij} of these three *in vitro* rates. The minimal and maximal values $\omega_{ij}(1\pm\epsilon_{ij})$ of these three rates define $2^3 = 8$ corners of an *error cuboid* around the origin of the Δ_{ij}-coordinates. For each of these corners, we determined the closest point on the hypersurface. We then estimated the errors of the barrier shifts $\Delta_{ij,\min}$ and eventually of the predicted *in vivo* rates $\omega^\star_{ij,\min}$ from the largest and smallest results as obtained for the 8 different corners. The error ϵ_{ij} of the *in vitro* rate and the error $\epsilon^{\mathrm{bs}}_{ij}$ of the barrier shift are related to the error $\epsilon^\star_{ij}$ of the predicted *in vivo* rate via

$$\omega^\star_{ij,\min}\left(1\pm\epsilon^\star_{ij,\min}\right) = \omega_{ij}\left(1\pm\epsilon_{ij}\right)\exp\left[-\Delta_{ij}\left(1\pm\epsilon^{\mathrm{bs}}_{ij}\right)\right] \approx \omega^\star_{ij,\min}\left[1\pm\left(\epsilon_{ij}+\epsilon^{\mathrm{bs}}_{ij}\right)\right] \tag{5.14}$$

for small $\epsilon^{\mathrm{bs}}_{ij}$. It follows

$$\epsilon^\star_{ij} = \epsilon_{ij} + \epsilon^{\mathrm{bs}}_{ij}\,. \tag{5.15}$$

For practical reasons, we did not precisely estimate the errors of all other rates. Instead, because these rates exhibit similar *in vitro* and *in vivo* values, we assumed that the relative errors of the predicted *in vivo* rates are approximately the same as those of the *in vitro* rates. This is the more accurate the flatter the hypersurface is in the correspond-

ing dimension close to the origin, i.e., the less the overall elongation rate depends on the specific transition rate. A precise estimation of the errors of all 12 predicted *in vivo* rates is computationally costly, because the kinetic distances of all $2^{12} = 4096$ corners of the 12-dimensional *error polyhedron* from the hypersurface need to be determined.

5.6 Conclusions

Our mapping from *in vitro* to *in vivo* leads to a significant increase of three internal rates and essentially no change for the remaining nine rates, see fig. 5.3 (B). The largest difference between the *in vivo* and the *in vitro* value is found for the near-cognate rejection rate ω_{76}, which is increased *in vivo* by a factor of about 6 and 8 for the 2-1-2 and 2-3-2 pathways of tRNA release from the E site, respectively. The other two rates, which are found to be larger *in vivo* than *in vitro*, are the initial dissociation rate ω_{off} and the recognition rate $\omega_{\text{rec}} = \omega_{12} = \omega_{67}$. The derived *in vivo* rates lead to good agreement with three different sets of *in vivo* data without adjusting any fit parameter.

For all internal transition rates, we find that the changes between *in vivo* and *in vitro* correspond to rather small barrier shifts, see fig. 5.3 (A). In fact, the absolute values of all barrier shifts are found to be equal or smaller than about 2. Under the assumption that the attempt frequencies of the ribosomal transitions do not change, this would correspond to changes in free energy barriers of about $2\,k_BT$. Because the cytosol represents a rather complex buffer, such small changes in the free energy barriers can be easily envisaged, arising for example from changes in the hydrogen bond networks around the ribosome/TC complex or from changes in the flexibility of some parts of this complex. However, it is not possible to draw conclusions about changes in attempt frequencies from the experimental data available at present. Molecular Dynamics simulations have been recently applied to explore the free energy landscape of tRNA translocation through the ribosome [67, 68]. From such simulations, one could in principle estimate the attempt frequencies.

Comparing the *in vivo* transition rates in table 5.1 for the 2-1-2 and the 2-3-2 pathway of E site tRNA release, we find that the differences are small for most growth rates. Therefore, our results are robust with respect to the chosen pathway for the release of tRNA from the E site.

Furthermore, our results also show that the high-fidelity buffer at 37 °C, used for the experiments described in section 4.5 and developed in [30, 40, 41], represents a good approximation to the cytosol as far as ribosomal kinetics are concerned, in contrast to earlier estimates [62].
But even though the predicted barrier shifts are relatively small, the associated changes of the transition rates have an interesting consequence for the relative importance of

initial selection and proofreading for the error frequency of protein synthesis. For the codon-specific Markov process introduced in the previous chapters and depicted in fig. 3.2, the efficiency of initial selection and proofreading are described by the coefficients $(\omega_{23}/\omega_{21})(\omega_{76}/\omega_{78})$ and $(\omega_{45}/\omega_{40})(\omega_{90}/\omega_{9,10})$, respectively. Compared to the corresponding *in vitro* value, the *in vivo* value of the initial selection coefficient for a growth rate of 2.5 dbl/h is increased by a factor of 12.4 and 17 for the 2-1-2 and the 2-3-2 pathway of E site release, respectively, whereas the proofreading coefficient is increased by a factor of 3.4 and 4.3, respectively. Thus, *in vivo* translation is less error prone, which is primarily achieved by the improved initial selection of bound TCs.

In our theory of translation elongation, the codon-dependence of the elongation rates arises solely from the initial binding rates that depend on the concentrations of available cognate, near-cognate, and non-cognate TCs, see eqs. (3.3) to (3.5). In particular, we use the same internal rates for all cognate as well as for all near-cognate tRNAs. Thus, the values of the internal rates ω_{rec}, ω_{21}, ω_{23}, ... of the cognate branch represent average values, obtained by averaging over all cognate tRNAs of all codons, and likewise for the internal rates ω_{76}, ω_{78}, ... of the near-cognate branch. *In vitro*, the decoding rates of different cognate codons were observed to be rather similar [41, 69] whereas the near-cognate GTPase activation rate ω_{78} was found to vary between 0.06/s and 1.3/s for different near-cognate codons of $\mathrm{tRNA}^{\mathrm{Phe}}$ [41]. Likewise, recent *in vivo* experiments provided evidence that 4 out of 14 near-cognate codons of $\mathrm{tRNA}^{\mathrm{Gly3}}$ have a much larger missense error frequency than the remaining 10 near-cognates [70]. Theoretically, it is straightforward to study the Markov process introduced in section 3.2 and depicted in fig. 3.2 with codon-specific decoding and processing rates. However, experimentally it is quite challenging to separately measure these rates for each codon and tRNA *in vitro*.

Finally, the method of minimizing the kinetic distance to map *in vitro* onto *in vivo* rates is very general and should be useful for other biomolecular processes, for which we have detailed information about the *in vitro* kinetics but such information is lacking *in vivo*. Examples are provided by the kinetics of protein folding, the activity of enzymes, and the motility of molecular motors.

6 Dependence of Overall Elongation Rate on tRNA Concentrations

Transfer RNAs are key players in translation. Because a sufficient pool of tRNA molecules is crucial to ensure stable protein synthesis, misregulation of tRNA concentrations lead to many human diseases. Recently, tRNA over-expression was found to be related to breast cancer and multiple myeloma [71–73]. Furthermore, diverse metabolic disorders as well as the MERRF (myoclonic epilepsy and ragged-red fiber) disease are caused by mutations of mitochondrial tRNAs [74–77] that can lead to a defective secondary structure of these tRNAs, to their inhibited binding of elongation factors or to their inefficient aminoacylation [78], i.e., to a reduced availability of functional tRNAs. It has been proposed to treat these diseases by targeted import of cytosolic tRNAs into mitochondria [79].
In fact, regulation of the cellular tRNA content is an important mechanism that enables cells to adjust to changing environmental conditions. The synthesis of tRNAs and ribosomes is governed by the *TOR signaling pathway*, which depends on nutrient supply and other stress conditions [80]. Furthermore, tRNAs in the cytosol can be cleaved as part of stress responses [81].

In this chapter, we study how these altered tRNA concentrations influence the translational state of the entire cell. In particular, we investigate the change of the overall elongation rate ω_{elo} under variation of the concentration of specific tRNAs. We find that upon variation of a single tRNA concentration the overall elongation rate ω_{elo} is changed in a non-trivial manner, giving rise to three distinct regimes of translation. In this chapter, we focus on the 2-3-2 pathway of tRNA release from the ribosomal E site, see section 1.3, whenever numerical results are presented. However, all results are qualitatively the same for the 2-1-2 pathway.

6.1 Three Distinct Regimes of Translation

In the cell, the total concentration X_a of the tRNA species a can be affected by the growth conditions or by genetic modifications that alter, for instance, the gene copy numbers of the tRNA species a. To understand how these changes affect protein synthe-

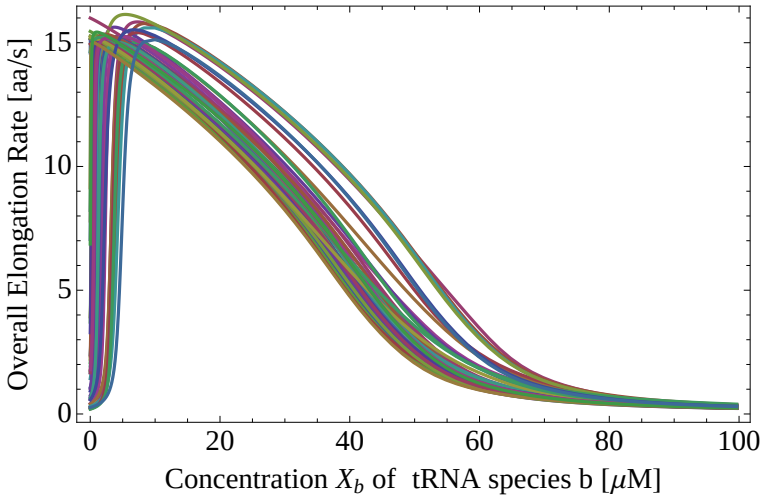


Figure 6.1: **Influence of individual tRNA concentrations on overall elongation rate.** The concentrations of all 43 tRNA species were individually varied, while leaving the concentrations of the respective other tRNA species at their *in vivo* values as given in table 4.1 for a growth rate of 0.7 dbl/h.

sis, we calculate the overall elongation rate ω_{elo} according to eq. (3.31) for varying cellular contents of all 43 individual tRNAs. More precisely, for each computation, we vary the concentration of a single tRNA species but keep all other tRNAs at their concentrations as measured *in vivo* corresponding to the chosen growth conditions. The results are shown in fig. 6.1. In general, the overall elongation rate curves – that we will abbreviate in the following by *OERCs* – of almost all tRNAs have the same shape, revealing three distinct regimes of translation: first, a regime of strongly inhibited translation for very small tRNA concentrations which is followed by a second regime of rapidly increasing overall elongation rate ω_{elo} beyond a specific threshold tRNA concentration; and third, after ω_{elo} has reached a maximum close to the physiological or unperturbed tRNA concentration, the OERCs monotonically decrease with increasing tRNA concentration and asymptotically approach zero.

The different regimes arise from the interplay of three molecular mechanisms. For very low concentrations of a certain tRNA species b, the concentration of the corresponding ternary complex (TC) of species b is very low. Thus, the cognate codons $c \in \mathbf{C}_{\text{co}}(b)$ (see section 3.1) become very slow and their translation is almost always performed through the very inefficient incorporation of near-cognate TCs. Therefore, before the maximum of the OERC is reached, the average speed of elongation is dominated by the elongation rates of individual codons which are limited by the availability of their cognate tRNA.

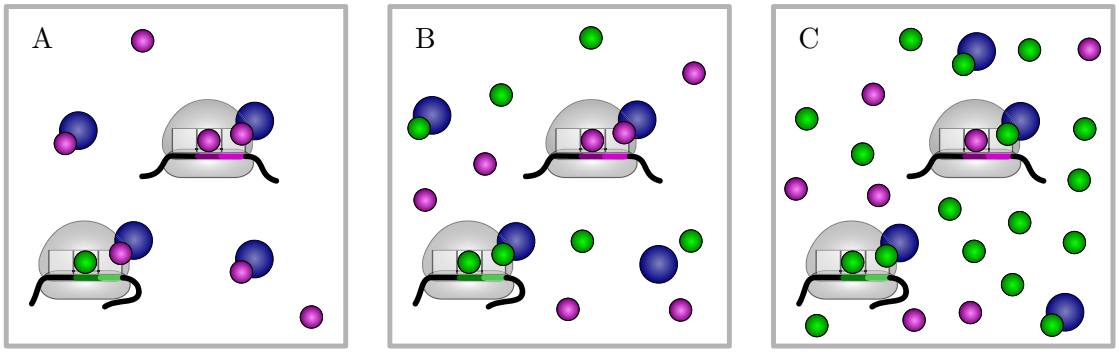


Figure 6.2: **Molecular mechanisms leading to three different regimes of the overall elongation rate.** For simplicity, we illustrate a translation system with only two species of tRNAs (green and purple small spheres) and corresponding cognate codons (green and purple lines) that are translated by ribosomes (gray domes). To translate a codon, a ribosome needs to bind a cognate ternary complex (TC) consisting of elongation factor EF-Tu (large blue sphere) and a tRNA molecule. TCs can associate and dissociate, and tRNA molecules compete for EF-Tu to form TCs. (A) Without free *green* tRNAs, the corresponding TCs cannot form and, thus, ribosomes cannot translate *green* codons. (B) When the ribosomes always find enough TCs containing cognate (either *green* or *purple*) tRNAs, translation proceeds at a significant speed. (C) When the *green* tRNAs are highly abundant, almost all TCs contain *green* tRNA molecules. Thus, translation of *purple* codons becomes very slow. In addition, ribosomes translating *purple* codons are inhibited through binding of the highly abundant non-cognate *green* TCs.

In contrast, beyond the maximum of the OERC, the increasingly high abundance of one individual tRNA of species b leads to a decrease of TCs comprising tRNAs of all other species. This is due to competition for free EF-Tu molecules that are necessary to form TCs, see section 1.2 and fig. 3.3. In addition, the many TCs formed with the highly abundant tRNA inhibit translation by binding mostly as non-cognates to ribosomes waiting for other TCs. A simplified illustration of the mechanisms leading to the three different regimes of the overall elongation rate is given in fig. 6.2.

In the following two sections we investigate the three mechanisms and how they determine the shape of the OERC in detail. First, in section 6.1.1 we qualitatively analyze the features of the OERC by varying different parameters of the translation system. Second, in section 6.1.2 we search for quantitative descriptions of the different regimes of the OERC based on approximations of the analytical expression of the overall elongation rate (3.31).

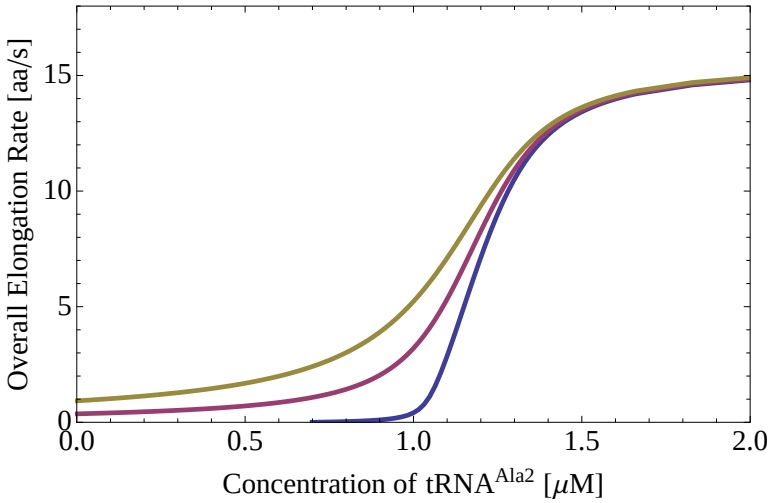


Figure 6.3: **Phase transition at the threshold concentration.** For three different choices of the near-cognate incorporation rate $\omega_{9,10}$ ($0.27\,\mathrm{s}^{-1}$, purple; $10^{-15}\mathrm{s}^{-1}$, blue; 0, yellow), the concentration of $\mathrm{tRNA}^{\mathrm{Ala2}}$ is varied. The concentrations of the respective other tRNA species are left at their *in vivo* values as given in table 4.1 for a growth rate of 0.7 dbl/h. For a vanishing near-cognate incorporation rate, the overall elongation rate becomes zero at the threshold concentration $\mathrm{X}^{\mathrm{th}}_{\mathrm{Ala2}} = 0.69\,\mu\mathrm{M}$, which represents a non-equilibrium phase transition.

6.1.1 Qualitative Features of the Overall Elongation Rate

In the previous section, we briefly mentioned three molecular mechanisms that give rise to the special shape of the OERC. These mechanisms are illustrated in fig. 6.2. In fact, all of the molecular mechanisms are always simultaneously present, but the extend of their influence on the overall elongation rate varies for the different concentrations of the perturbed tRNA. This interference of different mechanisms makes it hard to clearly observe and distinguish them in the OERC. By adjusting different parameters relevant for the translation system it is possible to make the underlying mechanisms visible and to understand in which regime of translation they become important. In this section, we give a qualitative discussion of the influence of these parameters on the OERC, whereas in the following section we elaborate the discussion in a quantitative manner. For the sake of clarity, we concentrate on one specific OERC, namely the OERC of the randomly chosen $\mathrm{tRNA}^{\mathrm{Ala2}}$.

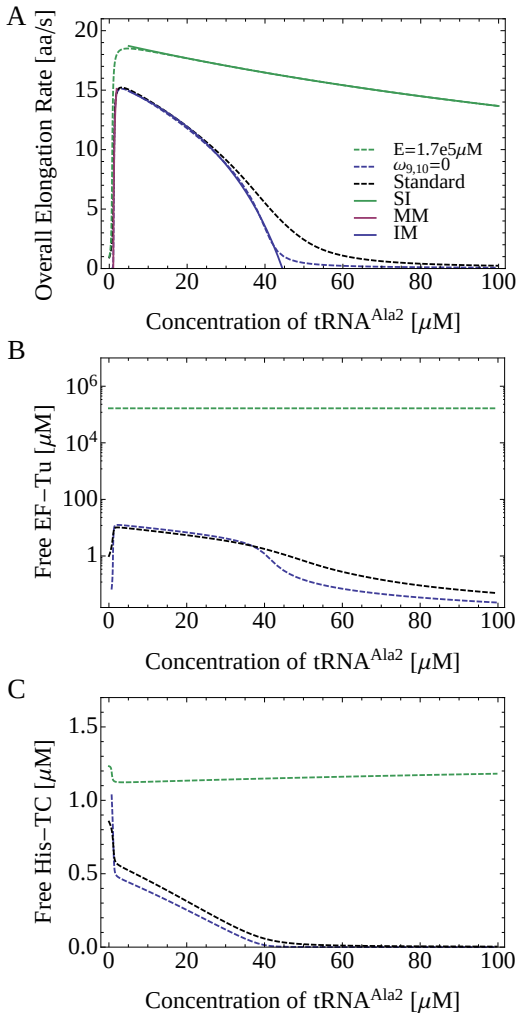


Figure 6.4: **Effects of tRNA$^{\text{Ala2}}$ abundance.** (A) Exact OERCs (dashed) of tRNA$^{\text{Ala2}}$ for different choices of parameters (standard *in vivo* parameters, black; $\omega_{9,10} = 0$, blue; $\mathcal{E} = 0.17\,\text{M}$, green). Approximations (solid): Michaelis-Menten like $\omega_{\text{elo}}^{\text{MM}}$ (6.11) (purple), inverse Michaelis-Menten $\omega_{\text{elo}}^{\text{IM}}$ (6.19) (blue), and non-cognate substrate inhibition $\omega_{\text{elo}}^{\text{SI}}$ (6.14) (green). Concentrations of (B) free EF-Tu molecules, and (C) free TCs comprising tRNA$^{\text{His}}$ under variation of tRNA$^{\text{Ala2}}$ (same colors as dashed lines in (A)). If EF-Tu is highly abundant, increasing concentrations of tRNA$^{\text{Ala2}}$ cause a decrease of the OERC due to non-cognate substrate inhibition (6.14). Otherwise, competition for EF-Tu suppresses all other TCs, e.g., tRNA$^{\text{His}}$, and the OERC decreases in an inverse Michaelis-Menten fashion (6.19), especially when near-cognate incorporation is slow, i.e., $\omega_{9,10}$ is small.

The overall elongation rate becomes very small if we strongly decrease the concentration of an individual tRNA, because the reduction of this tRNA species reduces the concentration of the corresponding TC. Due to the lack of this specific TC, translation of the corresponding cognate codons is then almost always performed by incorporation of near-cognate TCs, which is in general a very slow process. The speed and probability of near-cognate incorporation depends on the near-cognate incorporation rate $\omega_{9,10}$. Figure 6.3 shows that for small concentrations of $\mathrm{tRNA}^{\mathrm{Ala2}}$ the overall elongation rate gets smaller when the near-cognate incorporation rate is reduced. In fact, the overall elongation rate vanishes at a certain threshold concentration $\mathrm{X}^{\mathrm{th}}_{\mathrm{Ala2}}$ for $\omega_{9,10} = 0$. Just above the threshold, the slope of the OERC becomes steeper for smaller $\omega_{9,10}$. For a near-cognate incorporation rate $\omega_{9,10} = 0$, the threshold concentration defines a genuine phase transition far from equilibrium. It separates the phase of tRNA depletion with vanishing overall elongation rate from the phase of finite translation with an overall elongation rate that increases with the concentration of the perturbed tRNA in a Michaelis-Menten like fashion, as we will learn in section 6.1.2. The phase transition at the threshold concentration is discussed in more detail in section 6.2.

The Michaelis-Menten like increase of the OERC is followed by a monotonic decrease, see black dashed line in fig. 6.4 (A). This is caused by the competition of tRNAs for free EF-Tu molecules to form TCs. When the concentration of one tRNA species is increased, more TCs containing this particular tRNA are generated, causing the concentrations of free EF-Tu molecules to decline, see fig. 6.4 (B). As a result, less TCs consisting of EF-Tu and any other tRNA can form whereof an example is given in fig. 6.4 (C) for TCs containing $\mathrm{tRNA}^{\mathrm{His}}$. Consequently, the elongation rates of almost all codons drop, unless they are cognate to the promoted tRNA. As mentioned earlier, if only a few cognate TCs are present, translation of the affected codons proceeds via incorporation of near-cognates. For low near-cognate incorporation rates $\omega_{9,10}$, the compensation of lacking cognate TCs by near-cognate TCs is hampered and the OERC drops faster, see blue dashed line in fig. 6.4 (A). However, unlike in the low concentrations regime, there is no phase transition for $\omega_{9,10} = 0$ at high concentrations, because here the concentrations of free TCs of all species are always finite.
The regime of decreasing overall elongation rate can be divided into a negative and a positive curvature part. We will see in the next section that for vanishing $\omega_{9,10}$, the negative curvature part of the decreasing OERC can be very well described by an *inverse* Michaelis-Menten function. Note, that the strength of the negative curvature also depends on the extend to which the codon-specific elongation rates are governed by the availability of cognate TCs. The OERC is less steep if one of the codon-independent internal ribosomal transition rates is rate limiting (and not the binding of TCs). The right panel of fig. C.1 in the appendix exemplifies the influence of the processing rate ω_{pro} on the overall elongation rate and on the concentrations of free EF-Tu molecules, free Ala2-TCs, and free His-TCs. In particular, it is shown that the negative curvature

of the OERC is very weak for a rate limiting small processing rate ω_{pro}.
To study the effect of competition for EF-Tu molecules on the OERC, it is also interesting to vary the concentration of EF-Tu molecules $\mathcal{E}$ itself. As can be understood from the figures in the right panel of C.2 in the appendix, a small EF-Tu concentration strongly increases the competition effect and yields an OERC that has a significantly steeper decrease compared to OERCs for high EF-Tu concentrations. Furthermore, the impact of EF-Tu competition is also determined by the parameters governing the formation kinetics of TCs. Specifically, these are the association rate constant κ^{ass} and the dissociation rate ω^{dis}. Slow association and fast dissociation can both relieve the competition of tRNAs for EF-Tu as they lead to higher concentrations of free EF-Tu molecules, see fig. C.3 in the appendix.

The competition for EF-Tu molecules is not the only cause for a decreasing OERC. At the same time, the many TCs containing the highly abundant tRNA bind as non-cognates to ribosomes that are waiting for other TCs if they are not translating codons cognate to the promoted tRNA. In other words, a strongly increased TC acts mainly as a substrate inhibiting its own enzyme, the ribosome. In the next section, we will derive an expression for the OERC under non-cognate substrate inhibition. Under physiological conditions, the effect of non-cognate substrate inhibition is much smaller than the competition for EF-Tu molecules mentioned in the last paragraph. However, we can study non-cognate substrate inhibition by varying the concentrations of EF-Tu molecules $\mathcal{E}$, the initial dissociation rate ω_{off}, and the concentrations of actively translating ribosomes $\mathcal{R}$. For large concentrations of EF-Tu, competition for EF-Tu is negligible and the influence of non-cognate substrate inhibition becomes clearly visible, see green dashed line in fig. 6.4 (A). Furthermore, for a small initial dissociation rate ω_{off}, TCs bind stably to ribosomes and therefore a non-cognate bound TC can block a ribosome for a considerable amount of time. The OERC is smoother and resembles more a substrate inhibition curve for small ω_{off} than for large ω_{off}, see fig. C.4 in the appendix. Similarly, small concentrations of ribosomes lead to smoother, substrate inhibition like OERCs and, thus, make the inhibiting nature of non-cognate TCs observable, see fig. C.2 in the appendix.

In summary, by varying different parameters of the translation system we developed a qualitative understanding of the mechanisms that define the shape of the OERC. We showed that translation exhibits a Michaelis-Menten like behavior for small concentrations of the varied $\text{tRNA}^{\text{Ala2}}$. For intermediate and large $\text{tRNA}^{\text{Ala2}}$ concentrations, the competition for EF-Tu leads to a suppression of the other TCs. In combination with non-cognate substrate inhibition caused by the highly abundant TC, the competition for EF-Tu decreases the overall elongation rate.

6.1.2 Quantitative Analysis of the Overall Elongation Rate

In the previous sections, we qualitatively discussed the characteristic features of the OERCs. In this section, we take a closer look at the analytical form of the overall elongation rate as given by eq. (3.31). Based on approximations of the overall elongation rate, we find analytical expressions for the three mechanisms identified in the previous sections, i.e., the Michaelis-Menten like dependence for small tRNA concentrations and the decrease of the overall elongation rate by non-cognate substrate inhibition and EF-Tu competition for intermediate and large concentrations of a particular tRNA.

For our analysis, we focus on the concentration variation of one specific tRNA which we will denote by the label b. First, we can rewrite eq. (3.31) for the inverse of the overall elongation rate, i.e., the average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$

$$\begin{aligned}
\langle \mathrm{t}_{c,\mathrm{elo}} \rangle &= \sum_{c=1}^{61} \mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}} \\
&= \sum_{c \in \mathbf{C}_{\mathrm{co}}(b)} \mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}} + \sum_{c \in \mathbf{C}_{\mathrm{nr}}(b)} \mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}} + \sum_{c \in \mathbf{C}_{\mathrm{no}}(b)} \mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}} \\
&= \sum_{c \in \mathbf{C}_{\mathrm{co}}(b)} \mathrm{p}_c \frac{A_c(\hat{\mathrm{X}}_b)}{B_c(\hat{\mathrm{X}}_b)} + \sum_{c \in \mathbf{C}_{\mathrm{nr}}(b)} \mathrm{p}_c \frac{C_c(\hat{\mathrm{X}}_b)}{D_c(\hat{\mathrm{X}}_b)} + \sum_{c \in \mathbf{C}_{\mathrm{no}}(b)} \mathrm{p}_c \frac{E_c(\hat{\mathrm{X}}_b)}{F_c} + \frac{1}{\omega_{\mathrm{pro}}} \,, \qquad (6.1)
\end{aligned}$$

where $A_c(\hat{\mathrm{X}}_b)$, $B_c(\hat{\mathrm{X}}_b)$, …, $E_c(\hat{\mathrm{X}}_b)$ are linear functions of $\hat{\mathrm{X}}_b$ as given by

$$A_c(\hat{\mathrm{X}}_b) = \tau_{\mathrm{co}} \alpha_{\mathrm{co}} \left(\hat{\mathrm{X}}_b + \sum_{\mathbf{A}_{\mathrm{co}}(c)}{}' \hat{\mathrm{X}}_a \right) + \tau_{\mathrm{nr}} \alpha_{\mathrm{nr}} \sum_{\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a + 1 + \frac{\kappa_{\mathrm{on}}}{\omega_{\mathrm{off}}} \sum_{\mathbf{A}_{\mathrm{no}}(c)} \hat{\mathrm{X}}_a \,, \qquad (6.2)$$

$$B_c(\hat{\mathrm{X}}_b) = \alpha_{\mathrm{co}} \left(\hat{\mathrm{X}}_b + \sum_{\mathbf{A}_{\mathrm{co}}(c)}{}' \hat{\mathrm{X}}_a \right) + \alpha_{\mathrm{nr}} \sum_{\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a \,, \qquad (6.3)$$

and

$$C_c(\hat{\mathrm{X}}_b) = \tau_{\mathrm{co}} \alpha_{\mathrm{co}} \sum_{\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a + \tau_{\mathrm{nr}} \alpha_{\mathrm{nr}} \left(\hat{\mathrm{X}}_b + \sum_{\mathbf{A}_{\mathrm{nr}}(c)}{}' \hat{\mathrm{X}}_a \right) + 1 + \frac{\kappa_{\mathrm{on}}}{\omega_{\mathrm{off}}} \sum_{\mathbf{A}_{\mathrm{no}}(c)} \hat{\mathrm{X}}_a \,, \qquad (6.4)$$

$$D_c(\hat{\mathrm{X}}_b) = \alpha_{\mathrm{co}} \sum_{\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a + \alpha_{\mathrm{nr}} \left(\hat{\mathrm{X}}_b + \sum_{\mathbf{A}_{\mathrm{nr}}(c)}{}' \hat{\mathrm{X}}_a \right) \,, \qquad (6.5)$$

and

$$E_c(\hat{\mathrm{X}}_b) = \tau_{\mathrm{co}}\alpha_{\mathrm{co}} \sum_{\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a + \tau_{\mathrm{nr}}\alpha_{\mathrm{nr}} \sum_{\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a + 1 + \frac{\kappa_{\mathrm{on}}}{\omega_{\mathrm{off}}} \left(\hat{\mathrm{X}}_b + \sum_{\mathbf{A}_{\mathrm{no}}(c)}{}' \hat{\mathrm{X}}_a \right) , \tag{6.6}$$

$$F_c = \alpha_{\mathrm{co}} \sum_{\mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a + \alpha_{\mathrm{nr}} \sum_{\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a \, , \tag{6.7}$$

where the primed sum indicates summation over all TCs of species a in the given set except for TC species b. The time scales τ_{co} and τ_{nr} were defined in eqs. (3.83) and (3.84), and

$$\alpha_{\mathrm{co}} \equiv \kappa_{\mathrm{on}} \frac{\pi_{12}\pi_{23}\pi_{45}}{1-\pi_{12}\pi_{21}} \, , \tag{6.8}$$

$$\alpha_{\mathrm{nr}} \equiv \kappa_{\mathrm{on}} \frac{\pi_{67}\pi_{78}\pi_{9,10}}{1-\pi_{67}\pi_{76}} \, , \tag{6.9}$$

with the transition probabilities π_{ij} defined in eq. (2.6). Now, by taking the inverse of eq. (6.1), the overall elongation rate ω_{elo} can be formally written as an explicit function of the concentration $\hat{\mathrm{X}}_b$ of free TC species b. This function has the form

$$\begin{aligned}
\omega_{\mathrm{elo}}(\hat{\mathrm{X}}_b) = \prod_{\mathbf{C}_{\mathrm{co}}(b)} B_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{nr}}(b)} D_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{no}}(b)} F_c \Bigg(& \\
& \sum_{d \in \mathbf{C}_{\mathrm{co}}(b)} \mathrm{p}_d A_d(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{co}}(b)}{}' B_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{nr}}(b)} D_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{no}}(b)} F_c \\
& + \sum_{d \in \mathbf{C}_{\mathrm{nr}}(b)} \mathrm{p}_d C_d(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{co}}(b)} B_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{nr}}(b)}{}' D_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{no}}(b)} F_c \\
& + \sum_{d \in \mathbf{C}_{\mathrm{no}}(b)} \mathrm{p}_d E_d(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{co}}(b)} B_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{nr}}(b)} D_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{no}}(b)}{}' F_c \\
& + \frac{1}{\omega_{\mathrm{pro}}} \prod_{\mathbf{C}_{\mathrm{co}}(b)} B_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{nr}}(b)} D_c(\hat{\mathrm{X}}_b) \prod_{\mathbf{C}_{\mathrm{no}}(b)} F_c \Bigg)^{-1}
\end{aligned} \tag{6.10}$$

where primed products are taken over all codons $c \neq d$ in the indicated set. Although this expression is not very transparent, we can conclude that the numerator on the right hand side of eq. (6.10) can be written as an explicit polynomial for the concentration $\hat{\mathrm{X}}_b$ of free TCs of species b, i.e, TCs containing tRNAs of the species b that we are focusing on. The degree n of this polynomial is determined by the total number of cognate and

near-cognate codons of b-TC. Similarly, the denominator is a polynomial with a degree of $n+1$.

Of course, we should keep in mind that the concentrations $\hat{\mathrm{X}}_a$ of all other free TC species a implicitly depend on the concentration $\hat{\mathrm{X}}_b$ of free b-TC as is shown in chapter 3, eq. (3.78). Therefore, both the numerator and the denominator in (6.10) are in fact highly non-linear implicit functions of the concentration $\hat{\mathrm{X}}_b$ of free b-TC.

However, we can ignore this interdependence of TC concentrations for a moment and assume that all other TC concentrations are independent of $\hat{\mathrm{X}}_b$. This assumption is especially valid for the limit of low concentrations $\hat{\mathrm{X}}_b$, for which all higher order terms of $\hat{\mathrm{X}}_b$ in eq. (6.10) can be neglected. Furthermore, most TCs are essential in the sense that their cognate codons do not have further cognate TCs, see fig. 1.2. Therefore, $\sum_{a\in\mathbf{A}_{\mathrm{co}}(c)}\hat{\mathrm{X}}_a = \hat{\mathrm{X}}_b + \sum'_{a\in\mathbf{A}_{\mathrm{co}}(c)}\hat{\mathrm{X}}_a = \hat{\mathrm{X}}_b$ for all $c\in\mathbf{C}_{\mathrm{co}}(b)$, where the primed sum is taken over all tRNA species $a\neq b$. If we assume in addition that the incorporation of near-cognate TCs is very unlikely because of the very small rate $\omega_{9,10}$, such that $\alpha_{\mathrm{nr}}\approx 0$ (6.9), the dependence of the overall elongation rate ω_{elo} on the free concentration $\hat{\mathrm{X}}_b$ attains the Michaelis-Menten form as given by

$$\omega_{\mathrm{elo}}(\hat{\mathrm{X}}_b) \approx \omega_{\mathrm{elo}}^{\mathrm{MM}}(\hat{\mathrm{X}}_b) \equiv \frac{V_b^{\mathrm{MM}}\hat{\mathrm{X}}_b}{K_b^{\mathrm{MM}}+\hat{\mathrm{X}}_b} \quad \text{for small } \hat{\mathrm{X}}_b, \tag{6.11}$$

with rate V_b^{MM} and concentration K_b^{MM}

$$\begin{aligned} V_b^{\mathrm{MM}} &= \left(\tau_{\mathrm{co}} + \sum_{c\in\mathbf{C}_{\mathrm{nr}}(b)\cup\mathbf{C}_{\mathrm{no}}(b)} \mathrm{p}_c \frac{1+\tau_{\mathrm{nr}}\alpha_{\mathrm{nr}}\sum'_{\mathbf{A}_{\mathrm{nr}}(c)}\hat{\mathrm{X}}_a + \frac{\kappa_{\mathrm{on}}}{\omega_{\mathrm{off}}}\sum'_{\mathbf{A}_{\mathrm{no}}(c)}\hat{\mathrm{X}}_a}{\alpha_{\mathrm{co}}\sum_{\mathbf{A}_{\mathrm{co}}(c)}\hat{\mathrm{X}}_a} + \frac{1}{\omega_{\mathrm{pro}}}\right)^{-1} \\ &= \left(\sum_{i=1}^{5}\mathrm{t}_i + \sum_{c\in\mathbf{C}_{\mathrm{nr}}(b)\cup\mathbf{C}_{\mathrm{no}}(b)} \mathrm{p}_c\left(\mathrm{t}_{(c|0)} + \sum_{i=6}^{9}\mathrm{t}'_{(c|i)} + \mathrm{t}'_{(c|11)}\right)\right)^{-1}, \end{aligned} \tag{6.12}$$

$$K_b^{\mathrm{MM}} = V_b^{\mathrm{MM}} \sum_{c\in\mathbf{C}_{\mathrm{co}}(b)} \frac{\mathrm{p}_c}{\alpha_{\mathrm{co}}}\left(1+\tau_{\mathrm{nr}}\alpha_{\mathrm{nr}}\sum_{\mathbf{A}_{\mathrm{nr}}(c)}\hat{\mathrm{X}}_a + \frac{\kappa_{\mathrm{on}}}{\omega_{\mathrm{off}}}\sum_{\mathbf{A}_{\mathrm{no}}(c)}\hat{\mathrm{X}}_a\right), \tag{6.13}$$

where the average dwell times $\mathrm{t}_{(c|i)}$ are defined in eq. (3.13), $\mathrm{t}_i = \mathrm{t}_{(c|i)}$ for $\mathcal{P}_{c,\mathrm{co}} = 1$ for all codons c, and the prime at a summation sign indicates that $\hat{\mathrm{X}}_b$ is excluded from the sum of near- or non-cognate TCs.

Similarly, we can examine eq. (6.10) in the limit of high concentrations of free TC species b if we assume that all other TC concentrations are independent of $\hat{\mathrm{X}}_b$. For high concentrations $\hat{\mathrm{X}}_b$, the latter assumption only holds if the concentration of free EF-Tu molecules is not limiting the formation of TCs. This is true, for example, in the case

of high concentrations of total EF-Tu. The highly concentrated TCs of species b act on most ribosomes as competitive, non-cognate substrate inhibitors which leads to the asymptotic behavior

$$\omega_{\mathrm{elo}}(\hat{\mathrm{X}}_b) \approx \omega_{\mathrm{elo}}^{\mathrm{SI}}(\hat{\mathrm{X}}_b) \equiv \frac{W_b}{Q_b + \frac{\kappa_{\mathrm{on}}}{\omega_{\mathrm{off}}}\hat{\mathrm{X}}_b} \quad \text{for large } \hat{\mathrm{X}}_b, \tag{6.14}$$

with rate W_b and dimensionless constant Q_b

$$W_b = \left(\sum_{c \in \mathbf{C}_{\mathrm{no}}(b)} \mathrm{p}_c \mathrm{t}_{(c|0)} \right)^{-1}, \tag{6.15}$$

$$Q_b = \left(\sum_{i=1}^{4} \mathrm{t}_i \sum_{c \in \mathbf{C}_{\mathrm{co}}(b)} \mathrm{p}_c + \sum_{i=6}^{9} \mathrm{t}_i \sum_{c \in \mathbf{C}_{\mathrm{nr}}(b)} \mathrm{p}_c + \mathrm{t}_{5/10} \right) \Bigg/ \sum_{c \in \mathbf{C}_{\mathrm{no}}(b)} \mathrm{p}_c \mathrm{t}_{(c|0)} \ . \tag{6.16}$$

Here again, the average dwell time $\mathrm{t}_{(c|0)}$ is defined in eq. (3.13), $\mathrm{t}_i = \mathrm{t}_{(c|i)}$ for $\mathcal{P}_{c,\mathrm{co}} = 1$ and $\mathcal{P}_{c,\mathrm{nr}} = 1$, respectively, and $\mathrm{t}_{5/10} \equiv \mathrm{t}_5 = \mathrm{t}_{10}$.

Next, we will consider the case of high concentrations of TC species b under EF-Tu limited conditions. In this case, we cannot simply treat the concentrations of all other TCs as independent of TC species b, because tRNAs charged with amino acids have to compete for free EF-Tu molecules in order to form new TCs, see fig. 3.3. Instead, we will focus on the specific TC species k, which is the most affected by depletion through competition of its tRNA with charged tRNA species b for EF-Tu. For example, in *E. coli* for a growth rate of 0.7 dbl/hour, this is the TC containing $\mathrm{tRNA}^{\mathrm{His}}$. The blue and black lines in fig. 6.4 (C), as well as the bottom rows of figs. C.1 to C.4 in the appendix, show that we can very well approximate the range of intermediate concentrations of this most affected TC species k as linearly dependent on the concentration $\hat{\mathrm{X}}_b$ of highly abundant TC species b

$$\hat{\mathrm{X}}_k \approx a_b(k) \left(\mathrm{Y}_b(k) - \hat{\mathrm{X}}_b \right), \tag{6.17}$$

with some positive coefficient $a_b(k)$ and concentration $\mathrm{Y}_b(k)$.
If TC species k gets strongly depleted, the overall elongation time (6.1) is domintated by the elongation times of codons that are exclusively cognate to TC species k and belong to the set of non-cognate codons of TC species b, as codons near-cognate to TC species b could compensate the loss of their cognate TC species k by incorporating near-cognate and abundant TC species b. For the sake of clarity, we consider the case where only

codon c falls into this category, i.e., $\mathbf{C}_{\mathrm{co}}(k) \cup \mathbf{C}_{\mathrm{no}}(b) = \{c\}$. Then, eq. (6.10) reduces to

$$\omega_{\mathrm{elo}}(\hat{\mathrm{X}}_b) \approx \omega_{\mathrm{elo}}^{\mathrm{EC}}(\hat{\mathrm{X}}_b) \equiv \mathrm{p}_c^{-1} \frac{\alpha_{\mathrm{co}} \hat{\mathrm{X}}_k + \alpha_{\mathrm{nr}} \sum_{\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a}{\tau_{\mathrm{co}} \alpha_{\mathrm{co}} \hat{\mathrm{X}}_k + \tau_{\mathrm{nr}} \alpha_{\mathrm{nr}} \sum_{\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a + 1 + \frac{\kappa_{\mathrm{on}}}{\omega_{\mathrm{off}}} \left(\hat{\mathrm{X}}_b + \sum'_{\mathbf{A}_{\mathrm{no}}(c)} \hat{\mathrm{X}}_a \right)} \,. \tag{6.18}$$

The overall elongation rate under EF-Tu competition conditions $\omega_{\mathrm{elo}}^{\mathrm{EC}}(\hat{\mathrm{X}}_b)$ can be further simplified for small near-cognate incorporation probabilities, i.e., for a very small rate $\omega_{9,10}$ such that $\alpha_{\mathrm{nr}} \approx 0$, and large non-cognate dissociation constants $\omega_{\mathrm{off}}/\kappa_{\mathrm{on}}$. These assumptions lead to a mirrored or inverse Michaelis-Menten expression after approximation (6.17) is inserted into eq. (6.18)

$$\omega_{\mathrm{elo}}^{\mathrm{EC}}(\hat{\mathrm{X}}_b) \approx \omega_{\mathrm{elo}}^{\mathrm{IM}}(\hat{\mathrm{X}}_b) \equiv \frac{V_k^{\mathrm{IM}} \left(\mathrm{Y}_b(k) - \hat{\mathrm{X}}_b \right)}{K_k^{\mathrm{IM}} + \left(\mathrm{Y}_b(k) - \hat{\mathrm{X}}_b \right)} \,, \tag{6.19}$$

with rate V_k^{IM} and concentration K_k^{IM}

$$V_k^{\mathrm{IM}} = 1 \,/(\mathrm{p}_c \tau_{\mathrm{co}}) \ , \tag{6.20}$$

$$K_k^{\mathrm{IM}} = \frac{\tau_{\mathrm{nr}} \alpha_{\mathrm{nr}}}{\tau_{\mathrm{co}} \alpha_{\mathrm{co}} a_b(k)} \sum_{\mathbf{A}_{\mathrm{nr}}(c)} \hat{\mathrm{X}}_a + \frac{1}{\tau_{\mathrm{co}} \alpha_{\mathrm{co}} a_b(k)} \,. \tag{6.21}$$

To test the validity of the approximations $\omega_{\mathrm{elo}}^{\mathrm{MM}}(\hat{\mathrm{X}}_b)$, $\omega_{\mathrm{elo}}^{\mathrm{SI}}(\hat{\mathrm{X}}_b)$, and $\omega_{\mathrm{elo}}^{\mathrm{IM}}(\hat{\mathrm{X}}_b)$, we apply them to the OERC of $\mathrm{tRNA}^{\mathrm{Ala2}}$. As is shown in fig. 6.4 (A), the OERC is indeed very well described by Michaelis-Menten and inverse Michaelis-Menten approximations (6.11) and (6.19) over a broad range of $\mathrm{tRNA}^{\mathrm{Ala2}}$ concentrations if the rate of near-cognate incorporation $\omega_{9,10}$ is small. Furthermore, for high concentrations of EF-Tu molecules $\mathcal{E}$ the decrease of the OERC is well explained by non-cognate substrate inhibition (6.14). The fitted parameters are $V_{\mathrm{Ala2}}^{\mathrm{MM}} = 19\,\mathrm{s}^{-1}$, $K_{\mathrm{Ala2}}^{\mathrm{MM}} = 0.24\,\mu\mathrm{M}$, $W_{\mathrm{Ala2}} = 337\,\mathrm{s}^{-1}$, and $Q_{\mathrm{Ala2}} = 18$, $V_k^{\mathrm{IM}} = 25\,\mathrm{s}^{-1}$, $K_k^{\mathrm{IM}} = 28\,\mu\mathrm{M}$, $\mathrm{Y}_b(k) = 44\,\mu\mathrm{M}$.

6.2 Phase Transition at the Threshold Concentration

In this section, we focus on low concentrations of a specific tRNA species b. As we learned in the previous sections, the overall elongation rate is very small if the concentration X_b of tRNA species b is smaller than a certain threshold value $\mathrm{X}_b^{\mathrm{th}}$. In particular, for a vanishing near-cognate incorporation rate $\omega_{9,10}$, the overall elongation rate becomes zero at the threshold concentration $\mathrm{X}_b^{\mathrm{th}}$.

This break-down of the overall elongation rate at the threshold concentration represents a non-equilibrium phase transition. The concentrations of the 43 tRNA species span a concentration space with 43 dimensions. Our results imply that for vanishing near-cognate incorporation rate $\omega_{9,10}$, this concentration space is divided up into two distinct regions, in which the overall elongation rate is finite and zero, respectively. At the boundary or separatrix of these two regions, the translation system undergoes a non-equilibrium phase transition. In the following, we want to investigate what determines the threshold concentration $\mathrm{X}_b^{\mathrm{th}}$.

6.2.1 Asymptotic Behavior of the Overall Elongation Rate

By definition, the threshold is characterized by a vanishing overall elongation ω_{elo} or an infinite average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle = 1/\omega_{\mathrm{elo}}$. The average elongation time is given by eq. (3.31)

$$\langle \mathrm{t}_{c,\mathrm{elo}} \rangle = \sum_{c=1}^{61} \mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}} \,,$$

which implies that at least one codon-dependent elongation time $\mathrm{t}_{c,\mathrm{elo}}$ must diverge at the threshold. The codon-dependent elongation times have the form (3.10)

$$\mathrm{t}_{c,\mathrm{elo}} = \sum_{i=0}^{11} \mathrm{t}_{(c|i)}$$

with average dwell times $\mathrm{t}_{(c|i)}$ in states $i = 0, 1, \ldots, 11$ per elongation cycle defined in eq. (3.13). Note that the probability of cognate accommodation $\mathcal{P}_{c,\mathrm{co}} = 1$ and the dwell time in the near-cognate accommodated state $\mathrm{t}_{(c|10)} = 0$ for a vanishing near-cognate incorporation rate $\omega_{9,10}$. All other average dwell times attain finite, concentration independent values, apart from the average dwell times in states $(c|0)$ and $(c|11)$

$$\mathrm{t}_{(c|0)} = \frac{1 - \pi_{12}\pi_{21}}{\omega_{01}\, \pi_{12}\pi_{23}\pi_{45}} = \frac{1 - \pi_{12}\pi_{21}}{\kappa_{\mathrm{on}}\, \pi_{12}\pi_{23}\pi_{45}} \frac{1}{\sum_{a \in \mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a} \tag{6.22}$$

$$\mathrm{t}_{(c|11)} = \frac{1 - \pi_{12}\pi_{21}}{\omega_{\mathrm{off}}\, \pi_{12}\pi_{23}\pi_{45}} \frac{\sum_{a \in \mathbf{A}_{\mathrm{no}}(c)} \hat{\mathrm{X}}_a}{\sum_{a \in \mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a} \,. \tag{6.23}$$

If we consider the simplest case of a codon with only one cognate tRNA species $a \in \mathbf{A}_{\mathrm{co}}(c) = \{b\}$, the concentration of free cognate TCs is given by $\sum_{a \in \mathbf{A}_{\mathrm{co}}(c)} \hat{\mathrm{X}}_a = \hat{\mathrm{X}}_b$.

Therefore, the codon-dependent elongation time $\mathrm{t}_{c,\mathrm{elo}}$ diverges as

$$\mathrm{t}_{c,\mathrm{elo}} \approx \frac{s}{\kappa_{\mathrm{on}} \hat{\mathrm{X}}_b} \quad \text{for } \hat{\mathrm{X}}_b \approx 0, \tag{6.24}$$

with dimensionless scale

$$s \equiv \frac{1 - \pi_{12}\pi_{21}}{\pi_{12}\pi_{23}\pi_{45}} \left(1 + \frac{\kappa_{\mathrm{on}} \sum_{a \in \mathbf{A}_{\mathrm{no}}(c)} \hat{\mathrm{X}}_a}{\omega_{\mathrm{off}}} \right) . \tag{6.25}$$

If the codon-dependent elongation time $\mathrm{t}_{c,\mathrm{elo}}$ becomes infinite, the average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle = 1/\omega_{\mathrm{elo}}$ becomes infinite as well and diverges as

$$\langle \mathrm{t}_{c,\mathrm{elo}} \rangle \approx \sum_{c \in \mathbf{C}_{\mathrm{co}}(b)} \mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}} \approx \mathrm{p}_b \frac{s}{\kappa_{\mathrm{on}} \hat{\mathrm{X}}_b} \tag{6.26}$$

with

$$\mathrm{p}_b \equiv \sum_{c \in \mathbf{C}_{\mathrm{co}}(b)} \mathrm{p}_c \, . \tag{6.27}$$

In summary, $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$ can only become infinite if at least one concentration $\hat{\mathrm{X}}_b$ vanishes.

6.2.2 Ribosomal Concentrations at the Threshold

The concentration $\mathcal{R}_c$ of ribosomes with their A sites located at codon c is given by

$$\mathcal{R}_c = \frac{\mathrm{p}_c \, \mathrm{t}_{c,\mathrm{elo}}}{\langle \mathrm{t}_{c,\mathrm{elo}} \rangle} \mathcal{R} \, , \tag{6.28}$$

see eq. (3.37). As we approach the threshold at which the concentration $\hat{\mathrm{X}}_b$ of free b TCs vanishes, the average elongation time $\langle \mathrm{t}_{c,\mathrm{elo}} \rangle$ is dominated by the elongation times of codons $c \in \mathbf{C}_{\mathrm{co}}(b)$ cognate to tRNA species b. Thus, the ribosomal concentrations behave as

$$\mathcal{R}_c \approx \begin{cases} \mathrm{p}_c / \mathrm{p}_b \, \mathcal{R} & \text{for } c \in \mathbf{C}_{\mathrm{co}}(b) \, , \\ 0 & \text{for } c \notin \mathbf{C}_{\mathrm{co}}(b) \, , \end{cases} \tag{6.29}$$

with p_b defined in (6.27). These relations imply

$$\mathcal{R} \approx \sum_{c \in \mathbf{C}_{\mathrm{co}}(b)} \mathcal{R}_c \,. \tag{6.30}$$

Therefore, at the threshold where the concentration $\hat{\mathrm{X}}_b$ vanishes, all ribosomes stall on codons $c \in \mathbf{C}_{\mathrm{co}}(b)$ that are cognate to tRNA species b as one would expect intuitively. The stalled ribosomes attain states $(c|i)$ with steady state probabilities $\mathrm{P}^{\mathrm{th}}_{c,i}$

$$\mathrm{P}^{\mathrm{th}}_{c,i} = \mathrm{t}^{\mathrm{th}}_{(c|i)} \, \frac{\kappa_{\mathrm{on}} \hat{\mathrm{X}}_b}{\mathrm{p}_b \, s} \quad \text{for all } i = 0, \ldots, 11, \tag{6.31}$$

with dwell times $\mathrm{t}^{\mathrm{th}}_{(c|i)}$ at the threshold concentration $\mathrm{X}^{\mathrm{th}}_b$, and dimensionless scale s and summed codon usage p_b defined in eqs. (6.25) and (6.27), respectively. The expressions (6.31) are finite only for states $(c|0)$ and $(c|11)$, see eqs. (6.22) and (6.23) and discussion thereof. Thus, at the threshold all ribosomes dwell on codons $c \in \mathbf{C}_{\mathrm{co}}(b)$ with finite probability $\mathrm{P}^{\mathrm{th}}_{c,0}$ in the free A site state $(c|0)$ and with finite probability $\mathrm{P}^{\mathrm{th}}_{c,11}$ in the non-cognate bound state.

6.2.3 Concentrations of tRNA Subpopulations at the Threshold

The total population of tRNA species b consists of several subpopulations of molecules. As is explained in section 3.7, fig. 3.3, and is expressed in eq. (3.58), only one part of tRNA species b is actually present in the cell in form of freely diffusing TCs that are available for binding to ribosomes. The threshold $\mathrm{X}^{\mathrm{th}}_b$ represents the minimal total concentration of tRNA species b that is necessary to obtain a finite concentration $\hat{\mathrm{X}}_b$ of free TCs b.

As is discussed above, at the threshold concentration $\mathrm{X}^{\mathrm{th}}_b$ robust steady state translation is no longer possible and the overall elongation rate ω_{elo} vanishes. Then, from eqs. (3.60) to (3.73) it is easily understood that almost all subpopulations of tRNA species b vanish as well, except for the concentration of P site bound tRNAs of species b which becomes

$$\mathrm{X}^{\mathrm{P,th}}_b = \mathrm{p}_b \, \mathcal{R} \,, \tag{6.32}$$

where p_b is defined in eq. (6.27).

In addition, in case of the 2-3-2 pathway of tRNA release from the ribosomal E site, the concentration of E site bound tRNAs of species b at the threshold follows from eq. (3.71)

$$\mathrm{X}_b^{\mathrm{E,th}} = \mathrm{X}_b^{\mathrm{P,\,th}} = \mathrm{p}_b\,\mathcal{R}. \tag{6.33}$$

Note that $\mathrm{X}_b^{\mathrm{E}}$ is by definition always zero in case of the 2-1-2 pathway of E site release.

We can now calculate the threshold concentration $\mathrm{X}_b^{\mathrm{th}}$ from the conservation of tRNA molecules of species b (3.58) which also holds at the threshold

$$\mathrm{X}_b^{\mathrm{th}} = \begin{cases} \mathrm{p}_b\,\mathcal{R} & \text{for the 2-1-2 pathway,} \\ 2\mathrm{p}_b\,\mathcal{R} & \text{for the 2-3-2 pathway.} \end{cases} \tag{6.34}$$

These relations imply that if the total concentration X_b of tRNA species b is relatively small in comparison to the abundance of the corresponding cognate codons, all of these tRNA molecules are bound by ribosomes which act as absorbing *sponges* for tRNA molecules. Accomodated tRNAs do not leave a ribosome spontaneously, but only after the ribosome has further translocated. Therefore, at the threshold no tRNA molecule can be released from the ribosomes which all stall on codons cognate to tRNA species b, as we described in the previous section.

6.2.4 Redundant tRNA Species

To take a closer look at the phase transition, we calculated the OERCs of all tRNAs for a vanishing near-cognate incorporation rate $\omega_{9,10}$, see fig. 6.5. Inspection of fig. 6.5 shows that most tRNA species a exhibit a nonzero threshold concentration $\mathrm{X}_a^{\mathrm{th}}$, at which the overall elongation rate becomes zero. Closer inspection reveals that some tRNA species a do not have a threshold concentration $\mathrm{X}_{\mathrm{a}}^{\mathrm{th}}$. We call these tRNA species *redundant*. Consequently, in contrast to the redundant tRNAs, all *non-redundant* tRNA species a have a nonzero threshold concentration $\mathrm{X}_a^{\mathrm{th}}$. A vanishing threshold $\mathrm{X}_{\mathrm{a}}^{\mathrm{th}} = 0$ can only occur if two conditions are satisfied. First, at least one other tRNA species, a', must be cognate to the same codon as the redundant tRNA species a. Second, the combined concentration of the other cognate tRNA species a' must be sufficiently large to fully replace the redundant tRNA species a when its concentration X_a becomes small.

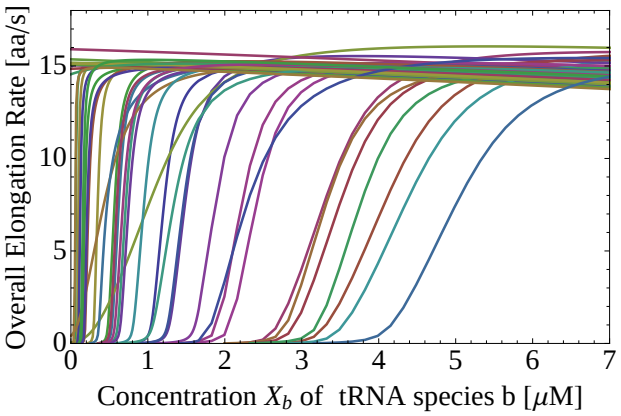


Figure 6.5: **OERCs of all tRNAs for a vanishing near-cognate incorporation rate.** For a near-cognate incorporation rate $\omega_{9,10} = 0$, the concentrations of all 43 tRNA species were individually varied, while leaving the concentrations of the respective other tRNA species at their *in vivo* values as given in table 4.1 for a growth rate of 0.7 dbl/h. *Redundant* tRNAs do not exhibit a threshold concentration, i.e., the overall elongation rate is finite for all concentrations of a redundant tRNA.

6.3 Conclusions

In this chapter, we studied the dependence of the overall elongation rate ω_{elo} on individual tRNA concentrations X_a. We found that this dependence is governed by a universal relationship that applies to all 43 different tRNA species. In all cases, the overall elongation rate curve (OERC) exhibited three different regimes of translation: a depletion regime at relatively low tRNA concentrations, a Michaelis-Menten like regime of rapidly increasing translational speed, and a regime of hampered elongation due to competition for EF-Tu and non-cognate substrate inhibition at relatively high concentrations.

We showed that the rapidly increasing part of the OERC can be well approximated by Michaelis-Menten kinetics with tRNA-specific Michaelis-Menten constants K_a^{MM}, see eq. (6.11) and fig. 6.4. In addition, the the decreasing part of the OERC is well described by inverse Michaelis-Menten kinetics with a Michaelis-Menten constant K_k^{IM}, see eq. (6.19), where k denotes the species of TCs that gets the most inhibited by competition for EF-Tu. If the binding to EF-Tu is not limiting the formation of new TCs, e.g. if the concentration of EF-Tu is very high, the decreasing regime of the OERC follows simple substrate inhibition kinetics, see eq. (6.14). Here, TCs containing the highly abundant tRNA act as non-cognates on most ribosomes, and thereby inhibit overall translation.

Furthermore, for a vanishing near-cognate incorporation rate $\omega_{9,10}$, we found nonzero threshold values $\mathrm{X}_a^{\mathrm{th}}$ for the concentrations of most tRNA species a, see fig. 6.5. Steady state protein synthesis is only possible if the tRNA concentrations exceed these threshold values. In other words, for $\omega_{9,10} = 0$ the concentration space spanned by the 43 individual tRNA concentrations is divided up into two regions, in which the overall elongation rate ω_{elo} is finite and zero, respectively. At the boundary of these regions, the translation system undergoes a non-equilibrium phase transition.

Our theoretical predictions about the limits of robust protein synthesis are accessible to experimental *in vivo* studies. For *E. coli*, artificial depletion of a certain tRNA species can be induced, for example, by minigenes that sequester this tRNA species and by knock-down of the corresponding tRNA genes [82–85]. As one decreases the concentration X_a of tRNA species a, the overall elongation rate is reduced by a factor of two at the concentration $\mathrm{X}_a^{\mathrm{th}} + K_a^{\mathrm{MM}}$ and vanishes at the threshold concentration $\mathrm{X}_a^{\mathrm{th}}$. This sensitivity of protein synthesis to depletion of a *single* tRNA species might also be a relevant property of human diseases such as breast cancer and multiple myeloma. These latter diseases are characterized by an overexpression of the whole translational machinery [71–73, 86], which leads to an increased growth of the cancer cells. Our study implies that the depletion or sequestering of even a single tRNA within these cells would lead to a strong reduction of their protein synthesis and, thus, of their growth rate.

7 Dependence of Total Elongation Flux on Ribosome Concentration

The rate with which a cell population doubles in size is related to the timescale by which cell content and cell volume duplicates. This timescale is tightly linked to the rate of protein synthesis [87, 88], which should be maximal according to basic growth and selection arguments.

In general, as growth requires a high translational capacity, ribosomes are kept in high concentrations in cells. The higher the ribosomal concentration, the higher the rate of protein synthesis and thus the growth rate of the cell population. However, it was observed experimentally that under poor growth conditions *E. coli* cells activate a diversified set of mechanisms that *reduce* the number of actively translating ribosomes. Recently, it was shown how various proteins mediate ribosomal hibernation in cells entering the stationary phase [89] and that free subunits of ribosomes are even actively degraded during starvation [90, 91]. Furthermore, when the speed of the ribosomes is slow due to a poor growth medium, the synthesis of new ribosomes is inhibited [87, 88].

In this chapter, we focus on cell-wide protein synthesis. We show that it is necessary for cells to keep the ribosomal concentration at a seemingly suboptimal level to sustain a reasonable overall rate of protein synthesis in a fluctuating environment. By considering different growth conditions, we find that *in vivo* steady state translation always lies in a region characterized by a ribosome limited protein synthesis rate that is robust against fluctuations of the availability of charged tRNAs.
In this chapter, we assume a 2-3-2 pathway of tRNA release from the ribosomal E site, see section 1.3, whenever numerical results are presented. However, all results are qualitatively the same for the 2-1-2 pathway.

7.1 Total Elongation Flux

The speed of the overall protein synthesis in a cell can be quantified by the total elongation flux J_{elo} which is the number of new peptide bonds that are formed per unit of time and volume. Here, we study the total elongation flux as a function of the concentration

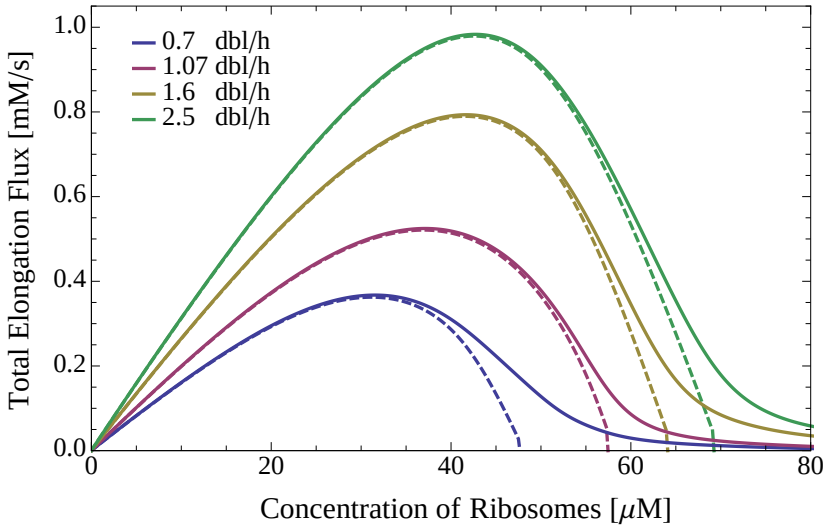


Figure 7.1: **Total elongation flux curves for four different growth conditions in *E. coli*** corresponding to measured growth rates from 0.7 to 2.5 dbl/h. The maxima of the total elongation flux curves define a separation between two regimes: In the first regime, protein synthesis is limited by the concentration of ribosomes. In the second regime, the superabundance of ribosomes depletes the concentrations of TCs and, thus, impedes protein synthesis. Dashed lines were obtained for a near-cognate incorporation rate $\omega_{9,10} = 0$.

of actively translating ribosomes $\mathcal{R}$

$$\mathrm{J_{elo}}(\mathcal{R}) = \omega_{\mathrm{elo}}(\mathcal{R})\,\mathcal{R}\,, \tag{7.1}$$

where $\omega_{\mathrm{elo}}(\mathcal{R})$ is the overall elongation rate defined in eq. (3.31) in section 3.5, which is interpreted as a function of the ribosomal concentration $\mathcal{R}$. Using the *in vivo* parameters for four different growth conditions, see tables 4.1, 4.2, and 5.1, we compute the total elongation flux $\mathrm{J_{elo}}(\mathcal{R})$ for a range of ribosomal concentrations $\mathcal{R}$. The results are plotted in fig. 7.1 (solid lines). For each growth condition, the total elongation flux $\mathrm{J_{elo}}(\mathcal{R})$ exhibits a maximum at a certain ribosomal concentration $\mathcal{R}^{(\mathrm{max})}$ and rapidly decreases for $\mathcal{R} > \mathcal{R}^{(\mathrm{max})}$, before it asymptotically approaches zero for large concentrations of ribosomes.

As one would expect intuitively, at first the total elongation flux $\mathrm{J_{elo}}(\mathcal{R})$ increases with an increasing concentration of ribosomes, because more ribosomes can make more protein in the same amount of time. This is reflected by the explicit linear dependence of $\mathrm{J_{elo}}(\mathcal{R})$ on the ribosomal concentration $\mathcal{R}$ in eq. (7.1).
However, translating ribosomes need to be constantly supplied with ternary complexes (TCs). Each ribosome can bind up to three tRNA molecules, which are then not available for the formation of TCs until they leave the ribosome. Thus, the more ribosomes are present, the less TCs are formed. The increasing competition for TCs slows down all ribosomes and leads to a decreasing overall peptide synthesis rate $\omega_{\mathrm{elo}}(\mathcal{R})$. This competition induced decrease of the overall elongation rate counteracts the benefits of having more ribosomes work in parallel.

For a sufficiently high concentration of ribosomes at least one species k of tRNA molecules gets almost completely absorbed to ribosomal P and E sites. For example, in *E. coli* for a growth rate of 0.7 dbl/hour, this is the TC containing $\mathrm{tRNA^{His}}$, see fig. C.5 in the appendix. Then, translation of codons that are cognate to the depleted tRNA species k proceeds almost entirely via the very slow incorporation of near-cognate TCs, leading to low elongation rates of these codons. Thus, the overall elongation rate ω_{elo} and the total elongation flux $\mathrm{J_{elo}}$ approach zero at high ribosomal concentrations. In fact, for a vanishing near-cognate incorporation rate $\omega_{9,10}$, translation breaks down at the ribosomal threshold concentration $\mathcal{R}^{\mathrm{th}}$, see dashed lines in fig. 7.1. The occurrence of such a non-equilibrium phase transition was discussed in section 6.2 in the light of variable tRNA concentrations. We can reinterpret the result (6.34) derived therein to obtain an expression for $\mathcal{R}^{\mathrm{th}}$

$$\mathcal{R}^{\mathrm{th}} = \begin{cases} \mathrm{X}_k/\mathrm{p}_k & \text{for the 2-1-2 pathway of E site release,} \\ \mathrm{X}_k/2\mathrm{p}_k & \text{for the 2-3-2 pathway of E site release,} \end{cases} \tag{7.2}$$

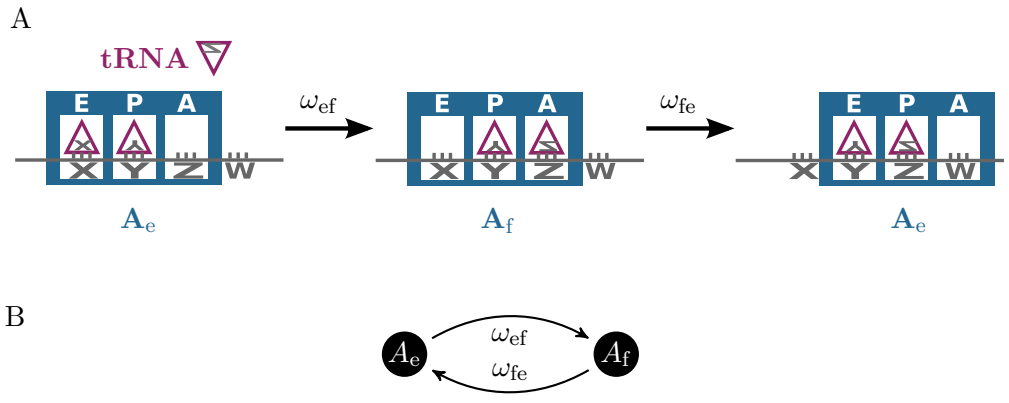


Figure 7.2: **A simplified model of translation elongation.** In the simplified model of translation elongation discussed in this chapter, the ribosome can attain only two states A_e and A_f with empty and tRNA-filled A site, respectively. (A) A ribosome reads a mRNA consisting of a sequence of codons (X, Y, Z, W). During transition from A_e to A_f a tRNA binds to the ribosomal A site. During transition from A_e to A_f, a new peptide bond is formed and the ribosome translocates to the next codon. The corresponding transition rates are referred to by ω_{ef} for transition from state A_e to A_f and by ω_{fe} for transition from state A_f to A_e. The rate ω_{ef} is proportional to the concentration of tRNA. (B) Simplified Markov process of translation elongation for a translation system with only one species of tRNAs and one species of codons.

where p_k denotes the summed codon usage of all codons that are cognate to the depleted tRNA species k (6.27)

In summary, fig. 7.1 shows that there are two regimes of translation separated by the maximum of the total elongation flux for each growth condition. In one regime, protein synthesis is limited by the availability of ribosomes, but cells can well tolerate fluctuations in tRNA concentrations. In the other regime, protein synthesis crucially depends on a reliable supply of tRNAs due to a superabundance of ribosomes. Any stress perturbing the tRNA availability can cause a strong inhibition or, in case of disabled near-cognate incorporation, even a complete halt of protein production.

7.1.1 Analysis of Total Elongation Flux

In this section, we aim to explain the shape of the total elongation flux curve $J_{elo}(\mathcal{R})$ that was introduced in the previous section. In particular, we investigate how the maximum

and the steepness of $J_{elo}(\mathcal{R})$ are determined. To do so, we first develop a theory of translation that is strongly reduced in complexity compared to the extensive theory developed in chapter 3. In this reduced theory, the mechanisms that influence the total elongation flux are much more transparent.

The simplified model is depicted in fig. 7.2 (A). Here, the ribosome can attain *only two* different states on each codon. We denote by A_e those ribosomes that have an empty A site and by A_f those ribosomes that have a filled A site. In addition, no distinction will be made between a tRNA and a TC comprising this tRNA. With this notation, an abstract translation process in which all codons and all tRNAs are treated as identical, i.e, a process with only one species of tRNAs and one species of codons that are cognate, corresponds to the two-state Markov process depicted in fig. 7.2 (B). In this two-state Markov process, a transition from state A_e to state A_f occurs with the forward (left to right) rate ω_{ef}. The opposite transition from state A_f to state A_e occurs with the backward (right to left) rate ω_{fe}. The forward rate ω_{ef} represents the tRNA binding rate to ribosomal A sites and is taken to be proportional to the concentration of free tRNAs

$$\omega_{ef} = \kappa \hat{X} \,, \tag{7.3}$$

with binding rate constant κ. The backward rate ω_{fe} can be identified with the inverse of the average time that a ribosome needs to form a new peptide bond and to move to the next codon. The concentrations $\hat{X}$ of free tRNAs and $\mathcal{R}_e$ and $\mathcal{R}_f$ of ribosomes in state A_e and A_f, respectively, obey the additional constraints

$$\mathcal{R} = \mathcal{R}_e + \mathcal{R}_f \,, \tag{7.4}$$

where $\mathcal{R}$ is the total concentration of actively translating ribosomes, and

$$X = \begin{cases} \hat{X} + \mathcal{R} & \text{for the 2-1-2 pathway of E site release,} \\ \hat{X} + 2\mathcal{R} & \text{for the 2-3-2 pathway of E site release,} \end{cases} \tag{7.5}$$

where X is the total concentration of tRNAs, i.e., the combined concentration of bound and free tRNAs. The factor of 2 on the right hand side of eq. (7.5) arises from the fact that ribosomes always contain two tRNAs if we consider the 2-3-2 pathway of E site release, see fig. 7.2 (A). In steady state, the forward and backward currents in the two-state Markov process depicted in fig. 7.2 (B) balance. Therefore, the steady state concentrations of ribosomes in states A_e and A_f are given by

$$\mathcal{R}_e \, \omega_{ef} = \mathcal{R}_f \, \omega_{fe} \,. \tag{7.6}$$

As in the previous section, we quantify the speed of protein synthesis by the total elongation flux J_{elo}. For the two-state Markov process, this total elongation flux is identical to the backward current, i.e., the number of transitions from state A_f to state A_e per unit of time, multiplied by the total concentration $\mathcal{R}$ of ribosomes. Thus, in steady state and under the constraints (7.4) and (7.5), the total elongation flux is given by

$$J_{elo} = \omega_{fe} \frac{1 - r}{\alpha + 1 - r} \mathcal{R}, \tag{7.7}$$

where r is the dimensionless stoichiometric concentration of active ribosomes

$$r \equiv \begin{cases} \mathcal{R}/X & \text{for the 2-1-2 pathway of E site release,} \\ 2\mathcal{R}/X & \text{for the 2-3-2 pathway of E site release,} \end{cases} \tag{7.8}$$

while α is the ratio of backward and forward transition rates

$$\alpha \equiv \frac{\omega_{fe}}{\omega_{ef}} = \frac{\omega_{fe}}{\kappa X}. \tag{7.9}$$

Once we have fixed the concentration X of total tRNA molecules and expressed the ribosome concentration $\mathcal{R}$ by the dimensionless stoichiometric concentration r, we can interpret (7.7) as a function of r. The total elongation flux curve $J_{elo}(r)$ is plotted for three different values of α in fig. 7.3. The results obtained for the simplified translation model introduced in this section remarkably resemble the total elongation flux curves for the detailed theory of translation elongation with a vanishing near-cognate incorporation rate, see dashed lines in fig. 7.1. In particular, $J_{elo}(r)$ exhibits a non-equilibrium phase transition at the threshold stoichiometric concentration $r^{th} = 1$, which corresponds exactly to the ribosomal threshold concentration (7.2) of the detailed model under the assumption that only one species of tRNAs and one species of codons exist. Furthermore, the function $J_{elo}(r)$ has a maximum at an intermediate stoichiometric concentration $r^{(max)}$, again as in the case of the full translation model. The value of $r^{(max)}$ is determined by the ratio α of backward and forward reaction rates with

$$r^{(max)} = \alpha + 1 - \sqrt{\alpha(\alpha + 1)}. \tag{7.10}$$

The position $r^{(max)}$ of the maximum is shifted towards its upper limit 1 if $\alpha \ll 1$, i.e., if $\kappa X \gg \omega_{fe}$, and towards its lower limit 1/2 if $\alpha \gg 1$, i.e., if $\kappa X \ll \omega_{fe}$, see fig. 7.4. Because the threshold concentration $r^{th} = 1$ is independent of the ratio α of backward and forward reaction rates, α does not only determine the position of the maximum but also the steepness of the total elongation flux curve, see fig. 7.3.

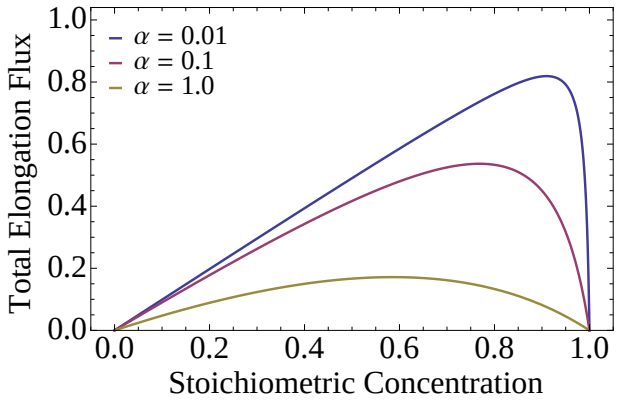


Figure 7.3: **Dependence of the total elongation flux J_{elo} on the stoichiometric concentration of ribosomes for a simplified translation process** treating all codons and tRNAs as identical. The total elongation flux is expressed relative to $\omega_{fe}\mathcal{R}/r$ such that it only depends on the stoichiometric concentration r of ribosomes and the ratio α of the backward and forward rates, see eq. (7.7). The position $r^{(max)}$ of the maximum of the total elongation flux curve is defined by α, see eq. (7.10).

Figure 7.4 and eq. (7.7) reveal how the concentrations of tRNAs and ribosomes confine protein synthesis. For ribosomal concentrations smaller than $r^{(max)}$, the total elongation flux is limited by the amount of actively translating ribosomes, whereas for $r > r^{(max)}$ protein synthesis is impeded by a lack of free tRNAs. In the latter case, too many tRNAs are bound to ribosomal P and E sites and the cell must *decrease* the concentration of ribosomes to *increase* protein synthesis.

We cannot directly apply the results of the simplified theory introduced here to the detailed translation system discussed in the previous chapters. However, we can interpret α as being the ratio of the processing rate ω_{pro} and the effective accommodation rate $\omega_{acc} \equiv 1/\left(\langle t_{c,elo}\rangle - \omega_{pro}^{-1}\right)$, where the average elongation time $\langle t_{c,elo}\rangle$ is defined in (3.31). As the processing rate ω_{pro} is not a limiting rate in the elongation cycle under physiological conditions, this ratio should be small compared to one. Thus, the concentration of ribosomes leading to a maximal total elongation flux should be rather close to the threshold concentration $\mathcal{R}^{th}$. This is indeed true, see fig. 7.1. We conclude that the simplified model introduced in this section provides some insight into the qualitative features of the dependence of the total elongation flux on the ribosome concentration.

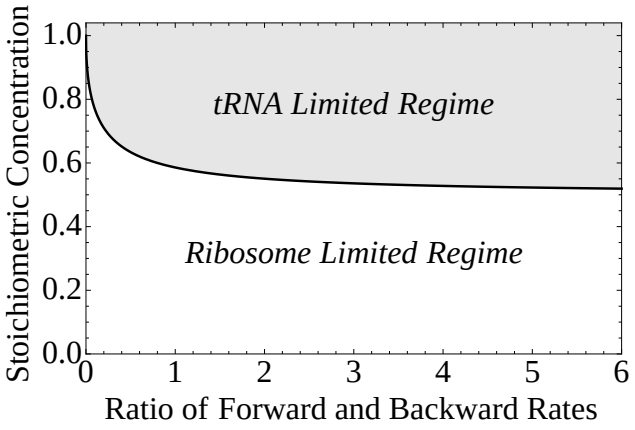


Figure 7.4: **Regimes of protein synthesis.** The black line defines the position $r^{(\max)}$ of the maximum (7.10) solely as function of the ratio α of the backward and forward rates. For stoichiometric ribosomal concentrations $r < r^{(\max)}$, the total elongation flux is limited by the amount of actively translating ribosomes, whereas for $r > r^{(\max)}$ a lack of free tRNAs inhibits protein synthesis.

7.1.2 *In Vivo* Elongation Flux is Seemingly Submaximal

Based on the relationship between protein synthesis rate and population growth rate, cells should optimize their ribosomal concentration to attain a maximal rate of peptide bond formation. To test this conjecture, we compare the actual concentrations of ribosomes in *E. coli* for different growth conditions to the predicted ribosomal concentrations $\mathcal{R}^{(\max)}$ which lead to maximal total elongation fluxes $\mathrm{J}_{\mathrm{elo}}(\mathcal{R}^{(\max)})$.
We find that for almost all growth conditions the *in vivo* concentrations of active ribosomes are considerably smaller than $\mathcal{R}^{(\max)}$, i.e., under physiological conditions peptide synthesis is limited by the availability of ribosomes. As an exception, under optimal growth conditions allowing for a growth rate of 2.5 dbl/h, the *in vivo* concentration of ribosomes is very close to the predicted $\mathcal{R}^{(\max)}$, see fig. 7.5.

This might be surprising at first glance, and one could ask why cells in suboptimal growth conditions do not increase their ribosomal concentration to maximize their protein expression. Recall from section 7.1 that the upper ribosomal concentration $\mathcal{R}^{\mathrm{th}}$ for which the total elongation flux vanishes or approaches zero is determined by the availability of TCs. This limit gets lowered by stress conditions that diminish tRNA or TC concentrations. Thus, if the *in vivo* concentration of ribosomes was very close to $\mathcal{R}^{(\max)}$, even a relatively small reduction of the cellular tRNA content could dramatically impede protein synthesis due to the steep decline of the total elongation flux curve $\mathrm{J}_{\mathrm{elo}}(\mathcal{R})$.

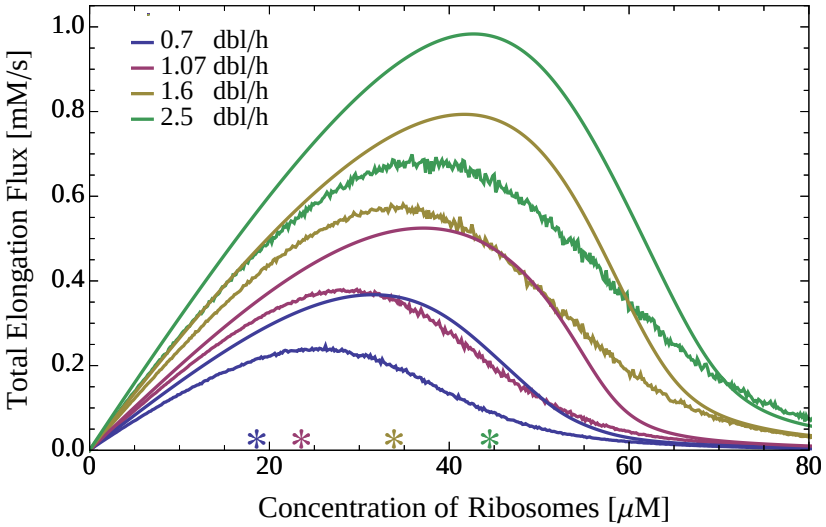


Figure 7.5: **Total elongation flux under fluctuating tRNA concentrations**. The smooth lines show the total elongation flux curves from fig. 7.1 for four different growth conditions corresponding to measured growth rates from 0.7 to 2.5 dbl/h. They are compared to the average total elongation flux obtained for fluctuating tRNA concentrations with a relative variance of $\langle \epsilon^2 \rangle = 0.05$. Each rough line represents an average over 1000 individual computations. The measured *in vivo* concentrations of active ribosomes are highlighted by star symbols. In conditions leading to slower cell growth, the *in vivo* concentrations of ribosomes are optimized for translation under more strongly fluctuating tRNA concentrations. Only at a growth rate of 2.5 dbl/h the *in vivo* concentration of ribosomes is optimized for a maximal total elongation flux in steady environments.

In the following, we further analyze the possibility that keeping ribosomes at a seemingly suboptimal concentration may offer some advantage against small fluctuations in the delivery of TCs. We consider a time dependent tRNA concentration

$$\mathrm{X}(t) = \mathrm{X}_0(1 + \epsilon(t))\,, \tag{7.11}$$

where $\epsilon(t)$ is the relative fluctuation with average zero and variance $\langle\epsilon^2\rangle$. Then, we exploit the fact that the simplified model introduced in the previous section leads to a qualitatively faithful description of the total elongation flux for the detailed theory of translation elongation. Under the assumption of time scale separation between the environmental fluctuations and the cellular response, we can substitute (7.11) in (7.7), thus obtaining

$$\mathrm{J}_{\mathrm{elo}}\,(t) = \omega_{\mathrm{fe}} \frac{1 + \epsilon(t) - r_0}{\alpha_0 + 1 + \epsilon(t) - r_0}\,\mathcal{R}\,, \tag{7.12}$$

where

$$r_0 \equiv \begin{cases} \mathcal{R}/\mathrm{X}_0 & \text{for the 2-1-2 pathway of E site release,} \\ 2\mathcal{R}/\mathrm{X}_0 & \text{for the 2-3-2 pathway of E site release,} \end{cases} \tag{7.13}$$

and

$$\alpha_0 \equiv \frac{\omega_{\mathrm{fe}}}{\kappa\,\mathrm{X}_0}\,. \tag{7.14}$$

We expand the total elongation flux to second order in ϵ and average over the noise. Since first order terms in ϵ vanish in average, the expansion to second order yields an average total elongation flux of

$$\langle \mathrm{J}_{\mathrm{elo}} \rangle = \mathrm{J}_{\mathrm{elo},0} - \omega_{\mathrm{fe}} \frac{\alpha_0 \langle\epsilon^2\rangle}{(\alpha_0 + 1 - r_0)^3}\,\mathcal{R} + \mathcal{O}(\langle\epsilon^3\rangle)\,, \tag{7.15}$$

where $\mathrm{J}_{\mathrm{elo},0}$ is the current in eq. (7.12) for $\epsilon = 0$, and $\mathcal{O}(\langle\epsilon^3\rangle)$ contains all higher order terms. Now, inspection of (7.15) shows that the ribosomal concentration $r^{(\max)}$ leading to a maximal average total elongation flux decreases as the variance of the tRNA concentration $\langle\epsilon^2\rangle$ increases, see fig. C.6 in the appendix.

To check that this decrease of $\mathcal{R}^{(\max)}$ is not only a property of the simplified model, we numerically computed the average total elongation flux as a function of ribosomal concentration $\langle \mathrm{J}_{\mathrm{elo}} \rangle(\mathcal{R})$ under fluctuations of tRNA concentrations using the detailed translation model. Indeed, we find a shift of the position of the maximum $\mathcal{R}^{(\max)}$ towards lower concentrations as predicted by the simplified model, see fig. 7.5. This supports our

conclusions that cellular ribosome concentrations in *E. coli* are optimized to operate in fluctuating environments.

7.1.3 Conclusions

We found that under all measured *in vivo* conditions, *E. coli* cells are in the ribosome limited regime of protein synthesis. The cells keep their total elongation flux far from the theoretical maximum achievable under steady environments, although this deviation decreases for larger growth rates, i.e., under better growth conditions. In the ribosome limited translation regime cells are protected against large fluctuations of the availability of the most rare TC. In fact, a ribosome concentration required to achieve a maximal total elongation flux under unperturbed conditions would lead to a strong decrease of protein synthesis upon reduction of cellular TC content. Further analysis of the translation system under small perturbations showed that the maximal average total elongation flux is indeed achieved at smaller ribosomal concentrations than in the absence of fluctuations in the tRNA concentrations. Therefore, a seemingly suboptimal concentration of ribosomes enables a cell to adjust its translational state to changing environmental conditions before protein synthesis is strongly hampered. We conclude that *E. coli* seems to have adapted its protein synthesis machinery such that it can handle unsteady environments and an unsure supply of nutrients, which these single cell organisms certainly have to cope with.

8 Dependence of Codon-Specific Elongation on Codon Usage

For many applications in research and biotechnology it is necessary to let cells synthesize large amounts of a specific protein. For that purpose, the cells are usually manipulated so that they contain high copy numbers of the mRNA that encodes the desired protein. To optimize protein yield and translation efficiency, instead of wild type mRNA a mutated codon sequence can be used. In the mutated mRNA, slow codons are generally replaced by synonymous codons that do not alter the amino acid sequence of the final protein product but have a higher elongation rate.
In this brief chapter we will learn that the optimization strategy of synonymous codon replacement can actually lead to the opposite rather than the intended effect. In particular, we will investigate how the overexpression of individual genes hampers translation of some codons by changing the cell-wide codon usage.

The codon usage p_c as defined in (3.30) determines how likely it is that a ribosome encounters codon c when it translocates down the mRNA. In the previous chapters, we learned that ribosomes bind up to three tRNAs that are cognate or near-cognate to codons in the ribosomal A, P, and E sites. Thus, the codon usages determine how many of the corresponding cognate or near-cognate tRNAs are bound to ribosomes. This means that the concentrations of available ternary complexes (TCs) and ultimately the codon-specific elongation rates as well as the missense error frequencies depend on the codon usages, see chapter 3. Strong expression of artificially induced genes leads to a shift of the codon usages towards the frequencies that the codons have in the induced gene. Therefore, overexpressing a gene might notably change codon-specific elongation rates.

To test this hypothesis, we calculated the codon-specific elongation rates using codon usages from the β-galactosidase gene lacZ. Figure 8.1 (A) shows the results compared to codon-specific elongation rates obtained from wild type codon usages of *E. coli* cells growing at a growth rate of 2.5 dbl/h. A direct comparison of both sets of codon usages is given in fig. 8.1 (B). Some codons have very similar codon usages in lacZ and in wild type *E. coli* cells, but most codons have significantly different frequencies in lacZ. The lacZ codon usages of AAA and UGG are changed by factors of about 0.3 and 5, respectively, thus differing the most from their corresponding wild type values.

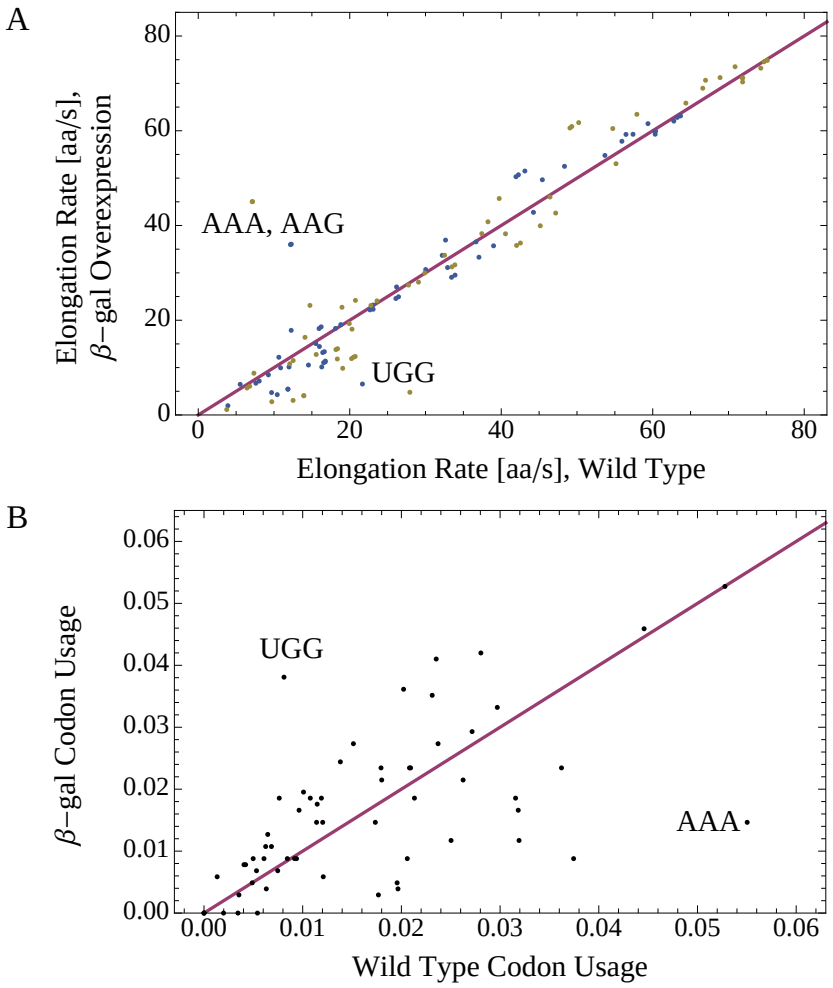


Figure 8.1: **Influence of β-galactosidase overexpression on codon-specific elongation rates.** (A) Elongation rates of all 61 sense codons in wild type *E. coli* cells growing at a rate of 2.5 dbl/h compared to codon-specific elongation rates in *E. coli* cells overexpressing β-galactosidase encoded by the lacZ gene, calculated for the 2-1-2 pathway (blue) and the 2-3-2 pathway (yellow) of tRNA release from the ribosomal E site. Except for the codon usages, all parameters used to calculate the elongation rates were identical in all four calculations, in particular the total concentrations of all tRNAs. (B) Codon usages of all sense codons in wild type *E. coli* cells growing at a rate of 2.5 dbl/h compared to codon usages in *E. coli* cells overexpressing β-galactosidase. Codon usages of AAA and UGG change the most when shifting from wild type to β-galactosidase overexpression.

The codon UGG is rare in wild type cells but abundant in lacZ. Due to the resulting reduction of TCs cognate to UGG in cells overexpressing β-galactosidase, the elongation rate of UGG is decreased by a factor of about two, given that E site tRNA is released via the 2-1-2 pathway. This codon usage related suppression of the elongation rate is even more pronounced if one assumes a 2-3-2 pathway of E site tRNA release: Here, the elongation rate of UGG is decreased by a factor of about six when shifting from wild type to lacZ codon usage, see fig. 8.1 (A).

In contrast, the codon AAA is highly abundant in wild type cells but is not a frequent codon in lacZ. Therefore, the AAA elongation rate is increased in β-galactosidase overexpressing cells by factors of about two and six for the 2-1-2 and the 2-3-2 pathway of E site tRNA release, respectively. In addition, the elongation rate of AAG is increased in β-galactosidase overexpressing cells, although its codon usage is comparable to its wild type value. Like AAA, AAG is cognate to $\mathrm{tRNA}^{\mathrm{Lys}}$ and, thus, an increase in free Lys-$\mathrm{tRNA}^{\mathrm{Lys}}$ TC also increases the elongation rate of AAG as is shown in fig. 8.1 (A).

To elucidate the interdependence of codon usage, TC concentration, elongation rate and missense error frequency, we plotted the concentration of available Lys-$\mathrm{tRNA}^{\mathrm{Lys}}$ TC in fig. 8.2 (A), the codon-specific elongation rates of the corresponding cognate codons AAA and AAG in fig. 8.2 (B), and the missense error frequencies of both codons in fig. 8.3 for a varying codon usage of AAA in the light of the two different pathways of E site release. The ratios of the codon usages of all other codons were kept as in wild type cells and were renormalized to fulfill eq. (3.30). As expected, the concentration of Lys-$\mathrm{tRNA}^{\mathrm{Lys}}$ TC and the elongation rates of AAA and AAG decrease for an increasing codon usage of AAA. In contrast, the missense error frequencies of both codons are strongly elevated for high codon usages of AAA.

In summary, we can conclude that overexpressing a gene can notably change codon-specific elongation rates and missense error frequencies in a cell. This has potential consequences not only for the expression efficiency of the desired protein. For example, folding of the nascent peptide chain might be affected or stalling of ribosomes could be caused, possibly leading to a misfolded or incomplete protein product.

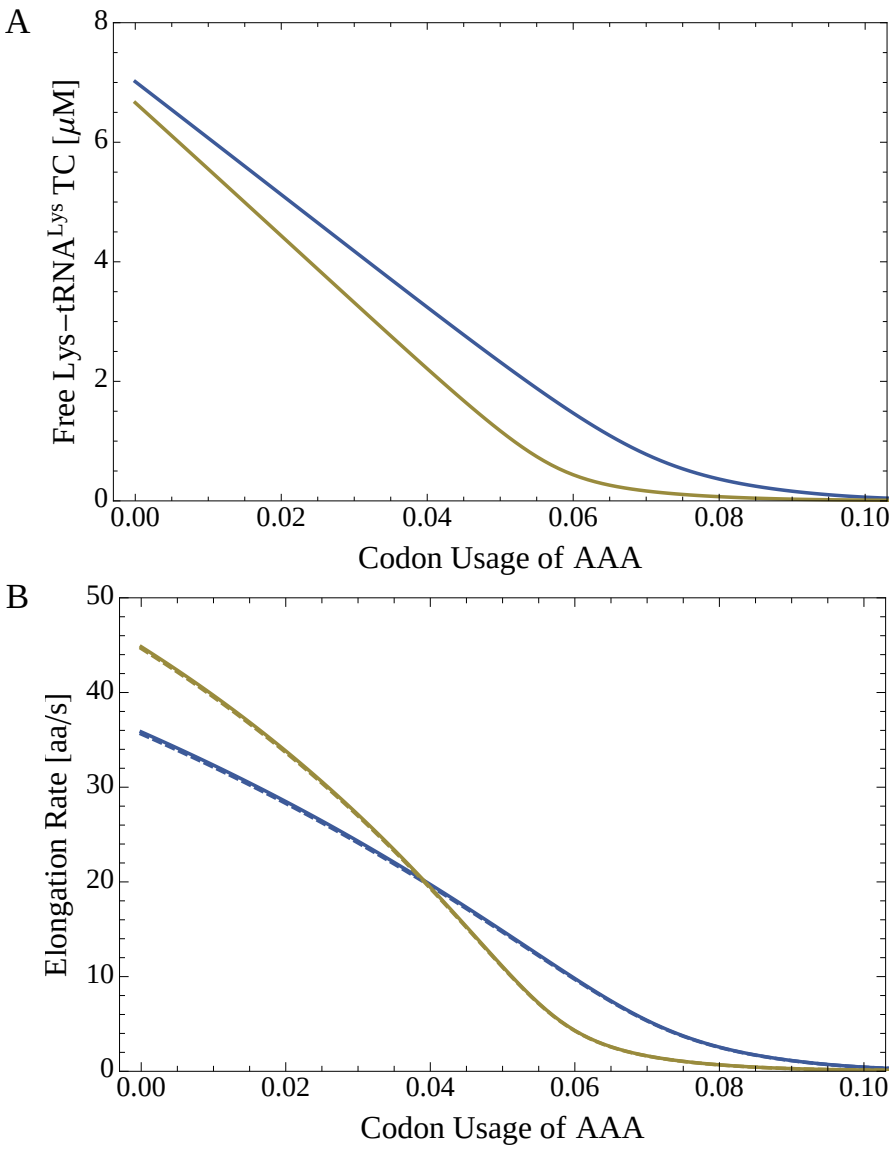


Figure 8.2: **Influence of codon usage on free TC concentration and codon-specific elongation rates.** (A) The concentration of free Lys-tRNA^{Lys} TC decreases when the codon usage of one of its cognate codons, AAA, increases. (B) The decrease in free Lys-tRNA^{Lys} TC concentration leads to decreasing elongation rates of the codons AAA (solid line) and AAG (dashed line), which are both cognate to tRNA^{Lys}. Results are shown for the 2-1-2 pathway (blue) and the 2-3-2 pathway (yellow) of tRNA release from the ribosomal E site.

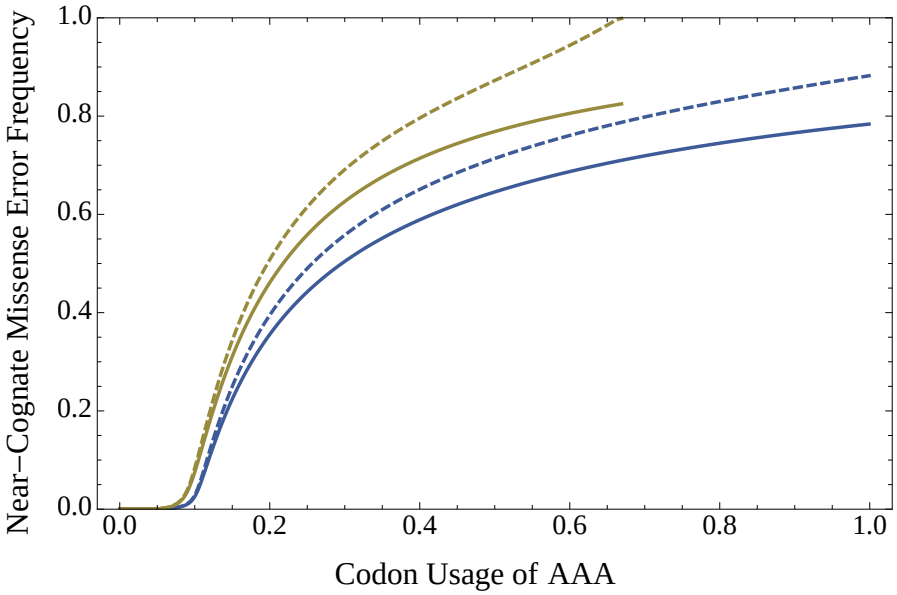


Figure 8.3: **Influence of codon usage on near-cognate missense error frequencies.** The near-cognate missense error frequencies of both codons AAA (solid line) and AAG (dashed line) rise when the codon usage of AAA increases. This results from the depletion of free cognate Lys-tRNA$^{\mathrm{Lys}}$ TCs (see previous fig. 8.2), which increases the probability that a near-cognate TC gets accommodated in the ribosomal A site. Results are shown for the 2-1-2 pathway (blue) and the 2-3-2 pathway (yellow) of tRNA release from the ribosomal E site.

9 Summary, Conclusions and Outlook

It is of great relevance for medical therapies and for the biotechnological production of protein-based substances to understand what determines the speed and accuracy of protein synthesis. To address this long-lasting puzzle, we developed a comprehensive theoretical framework on protein synthesis by ribosomes in bacteria. We described translation elongation as a Markov process with 12 different ribosomal transition rates and considered two alternative pathways of tRNA release from the ribosomal E site, see chapter 3. A fundamental ingredient of our modeling is the dependence of codon-specific elongation rates and missense error frequencies on the concentrations of ternary complexes (TCs) available for uptake by translating ribosomes. We analyzed the steady state of the translation process and derived analytic solutions for the concentrations of TCs, see section 3.7.

Our theoretical description of translation elongation is based on current biochemical knowledge on *in vitro* ribosomal kinetics obtained by our collaborators Marina Rodnina and her co-workers over the past decades, see chapter 3. In addition, we used new experimental data on the overall speed of *in vitro* translation that were obtained by Michael Thommen from the Rodnina laboratory as part of our cooperation, see chapter 4. In particular, we now have a complete set of individual transition rates and overall elongation rates *in vitro* for two different temperatures which are summarized in table 4.3.

However, the *in vitro* ribosomal transition rates must be different from those *in vivo* because the overall elongation rate in living cells is substantially faster than *in vitro*. Therefore, we developed a theoretical method to deduce *in vivo* rates from their *in vitro* values, see chapter 5. We first introduced the *kinetic distance* as a measure for differences in kinetic rates. We then identified a set of rates that is compatible with the *in vivo* overall elongation rate and has a minimal kinetic distance to the measured *in vitro* rates, see table 5.1. We finally validated this set of predicted *in vivo* transition rates with three different sets of experimental data on speed and fidelity of translation in *E. coli* cells from the literature. As a main result, we found that 9 out of 12 *in vivo* transition rates are similar to the measured *in vitro* rates, whereas the other three have considerably increased *in vivo* values. As a consequence, initial selection and proofreading are much more reliable *in vivo*, i.e., *in vivo* translation is less error-prone.

Using the deduced *in vivo* transition rates, we studied the dependence of overall translation on the concentrations of the different tRNAs, see chapter 6. The relationships between the various tRNA concentrations and the overall elongation rate are rather nonlinear, see fig. 6.1, and, thus, very different from previous models in which the overall elongation rate was taken to grow linearly with the tRNA concentrations, see for example [28]. For low tRNA concentrations, we found a dramatic reduction of the overall elongation rate close to a certain threshold concentration. This reduction becomes a genuine non-equilibrium phase transition when the near-cognate incorporation rate $\omega_{9,10}$ is zero. Furthermore, we showed that the dependence of the overall elongation rate on tRNA concentrations can be approximated by a Michaelis-Menten like function for small concentrations and an inverse or mirrored Michaelis-Menten function for larger concentrations. The latter behavior is a consequence of the competition of tRNAs for EF-Tu molecules during the formation of TCs. In contrast, if this EF-Tu competition can be neglected, the promotion of one species of tRNA leads to a decreasing overall elongation rate by non-cognate substrate inhibition.

In addition to the tRNA concentrations, the abundance of actively translating ribosomes also determines the amount of protein that a cell can produce per unit of time. In chapter 7, we found that in contrast to naive intuition the protein output of a cell is reduced when the ribosomal concentration exceeds a certain value. This is a result of TC depletion by active ribosomes, because tRNA molecules get bound to ribosomal A, P, and E sites and, thus, are not available for the formation of new TCs. The ribosomal concentration leading to optimal protein synthesis depends on the ratio of the time that it takes a ribosome to accommodate a new tRNA and the time to process it and translocate to the next codon, as we could show by means of a simplified translation model in section 7.1.1. In addition, we found that in *E. coli* the concentration of ribosomes is adjusted in such a way that these unicellular organisms can tolerate fluctuations of TC content caused, for example, by variable nutrient supply.

Finally, we investigated the influence that the codon usages have on elongation rates and missense error frequencies. We showed in chapter 8 that the overexpression of a gene leads to altered codon-specific elongation rates such that codons that are slow in wild type cells can become fast in overexpressing cells and *vice versa*. This behavior is only predicted to occur if the codon usages in the overexpressed gene are different from those in wild type cells.

The codon-specific Markov process introduced here can be used to study the dependence of the speed and accuracy of *in vivo* and *in vitro* protein synthesis on a variety of parameters which have not been considered in this work. For example, we may study how these quantities vary with changes in the overall TC composition or how changes in internal transition rates, arising for example from protein or rRNA mutagenesis, affect the synthesis rate.

In the study presented here, we chose to focus on translation in *E. coli* because of the extensive data base available for these cells and because of the expertise of our experimental collaborators from the Max Planck Institute in Göttingen. However, our theory is quite general and should be applicable to all prokaryotic or eukaryotic cells.

Protein synthesis in mitochondria is to some extend comparable to that in bacteria, reflecting their supposedly endosymbiotic origin. Consequently, as a next step we plan to apply our theoretical framework of protein synthesis to translation in human mitochondria. This is especially interesting, because our theory predicts that protein synthesis is very sensitive to the depletion of a *single* tRNA species and many mitochondrial diseases are related to deficiencies in individual tRNA species caused by mutations in the mitochondrial genome [74–77, 92]. Starting from these observations, we will use our theory to investigate the long-standing question why these diseases often become evident only in later years even though the underlying disastrous mutations are present already before birth. In addition, we intend to apply our theoretical framework to protein synthesis in eukaryotic cells. We aim to predict mutated codon sequences of individual genes for optimized gene expression in human cells and yeast. In particular, we will apply this codon optimization method in the context of protein-based vaccine development and production.

A Pauses in Protein Synthesis – Analysis of Translation Intermediates

For several weeks in March and April 2013, the author of this thesis visited the Max Planck Institute for Biophysical Chemistry in Göttingen to work in the laboratory of Marina Rodnina. The aim of this research project was to find a theoretical description for data on pauses occurring during the translation of certain mRNAs. The results are not included in the main part of thesis as this project is thematically distinct from the topics presented in the other chapters.

A.1 Brief Introduction

During translation elongation, the ribosome does not always proceed with the same speed. Instead, the speed of translation is inhomogeneous. One reason for this inhomogeneity is the difference in concentrations of ternary complexes, as is studied in detail in this thesis. Apart from this *global* influence of ternary complex abundance on translation, other *local*, or gene-specific, phenomena may play a role. Examples are provided by mRNA secondary structure, codon sequences that can perturb the translating ribosome, or physicochemical properties of the synthesized peptide chain that lead to retarding interactions with the ribosome. Although many ideas exist to explain the occurrence of pauses during translation elongation, the relevance and detailed nature of these inhibiting processes are still poorly understood.

Marina Rodnina and her co-workers developed a method to uncover pauses during *in vitro* translation of individual genes. Briefly, the synchronized *in vitro* translation of identical mRNAs is stopped at defined time points and the peptides that have been synthesized so far are analyzed. If the studied mRNA contains a region of slow translation, many peptides of identical length will be found at certain time points. These peptides are called *translation intermediates*. Their length in amino acids is identical to the codon position of the region of slow translation on the mRNA.
We helped to analyze the time dependence of the concentrations of the various intermediate peptides by developing a simple theoretical model of *in vitro* translation as is explained below.

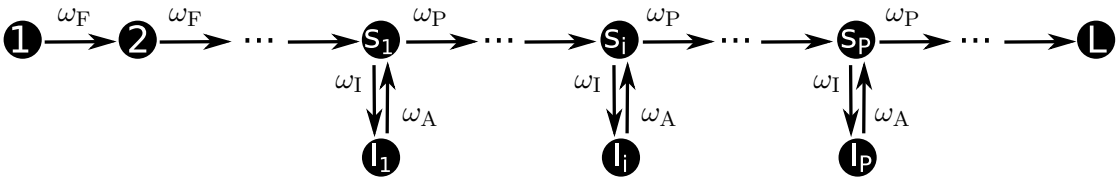


Figure A.1: **Markov model of *in vitro* peptide synthesis.** A mRNA of L codons is represented by a Markov chain of L states. Every translation process starts in state 1. Transitions from one state to the next proceed with forward rate ω_{F}, backward transitions are not allowed. If the mRNA is known to have P slowly translated codons that cause pauses during the translation process, these are represented by slow codon states s_i with $i = 1, ..., \mathrm{P}$. Transitions from slow codon states proceed with forward rate ω_{P}, which is smaller than ω_{F}. In addition, for every slow codon state s_i one inactive ribosome state I_i is introduced. Transitions to inactive ribosome states are only possible from the corresponding slow codon state and occur with rate ω_{I}. Active translation is continued by transition from the inactive to the slow codon state with activation rate ω_{A}.

A.2 Data Processing

As the original data are yet unpublished and confidential, they are not shown here. This does not restrict our presentation of the results of the theoretical analysis of the data. We studied the synthesis of four different peptides labeled A, B, C, and D. The corresponding mRNAs were prepared and translated at 37 °C in the Rodnina laboratory as is described in section 4.5.2 and fig. 4.3. As a result, images of gels reporting on the temporal progress of the synthesis were obtained, comparable to that of mRNA CspA in fig. 4.3.

We processed the images of the gels as follows. We first generated intensity profiles of the different lanes visible on the gels with the image processing tool *Fiji*. We rescaled the position axis of the images from pixels to peptide length in amino acids by using marker peptides of defined lengths that were loaded on the same gels. Then, using the *Peak Analyzer* tool of the software *Origin*, we corrected the profiles for background, identified peaks corresponding to the various intermediate peptides and integrated them using a fixed integration range width for all peaks in all intensity profiles. Afterwards, we normalized the obtained integrated intensities for each time point such that they represented the relative abundances of the different intermediates at a given time. As a result, we obtained for each intermediate its relative abundance as a function of time.

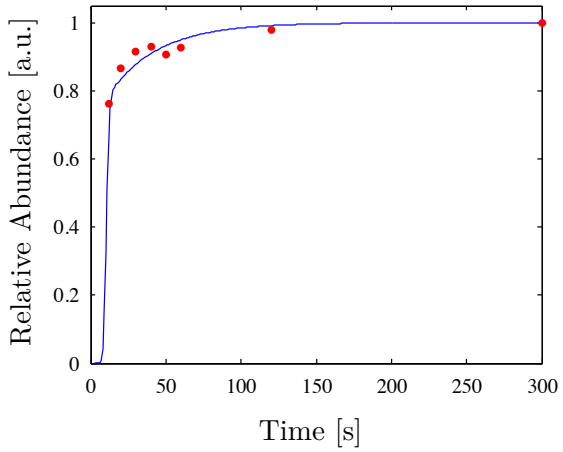


Figure A.2: **Relative abundance vs. time of full length product of peptide A.** For peptide A, no intermediate was observed. Red dots represent experimental data, the blue solid line is a theoretical curve resulting from the Markov model introduced in section A.3. The Markov model for peptide A was built such that it does not contain slow codon states. Thus, one of the fast codon states had to be chosen to be connected to an inactive ribosome state. Rates as obtained by fitting are: $\omega_{\mathrm{F}} = 4/s$, $\omega_{\mathrm{I}} = 1.128/s$, $\omega_{\mathrm{A}} = 0.03/s$.

A.3 Data Analysis

To theoretically describe the time dependent abundances, we modeled the translation of mRNAs as a simple Markov process depicted in fig. A.1. Every codon of a mRNA of length L is represented by one state in the Markov chain. The translation process always starts in state 1 and proceeds subsequently through states 2, 3..., L, with identical rate ω_{F} for each forward transition and no backward transitions. Pauses are introduced to the model by replacing the forward rate ω_{F} by a much smaller rate ω_{P} at certain *slow codon* states s_i corresponding to positions of slow codons in the mRNA.
For every slow codon state s_i, there is one additional *inactive ribosome* state in the Markov model. The inactive ribosome states are linked to the slow codon states by inactivation rates ω_{I} and activation rates ω_{A}, see fig. A.1. A ribosome that is inactive has to become active again, before it can continue to translate the next codon. The inactive ribosome states were introduced to map the experimental observation that the relative abundance of intermediate peptides increases faster than the relative abundance of preceding intermediates decreases.

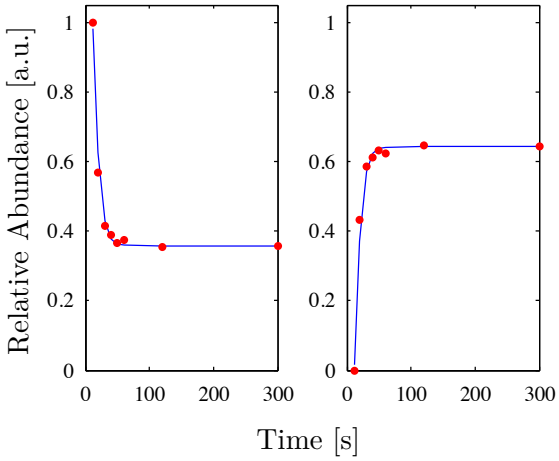


Figure A.3: **Relative abundance vs. time of one intermediate peptide and full length product of peptide B.** For peptide B, only one intermediate was observed. The time course of its abundance is depicted on the left, whereas the build-up of full length peptide is shown on the right. Red dots represent experimental data, blue solid lines are theoretical curves according to fitted parameters. Rates as obtained by fitting are: $\omega_{\mathrm{F}} = 4/s$, $\omega_{\mathrm{P}} = 0.08/s$, $\omega_{\mathrm{I}} = 0.0445/s$, $\omega_{\mathrm{A}} = 0/s$. The vanishing activation rate means that within the time scale of the experiment, inactivated ribosomes did not continue translation.

We implemented the Markov model described above as a *MatLab* script and fitted the rates ω_{F}, ω_{P}, ω_{I}, and ω_{A} to the experimental data for all four peptides. The fits are shown in figs. A.2 to A.5 and the resulting rates are stated in the captions thereof.

A.4 Conclusions

Despite being very simple, the Markov model introduced in the previous section is able to describe *in vitro* translation dynamics of four different peptides. By fitting this model to experimental data on the synthesis of these peptides, we found that slow regions are translated up to 50 times slower than usual codons. In addition, the fitting revealed that a substantial part of the ribosomes gets inactivated before being able to translate a slow codon. The fraction of inactivated ribosomes f_{I} depends on the slow codon forward rate ω_{P} and the inactivation rate ω_{I}

$$f_{\mathrm{I}} = \frac{\omega_{\mathrm{I}}}{\omega_{\mathrm{I}} + \omega_{\mathrm{P}}} \tag{A.1}$$

and differs for the four peptides. The fraction f_{I} of ribosomes that enter the inactivated state at a pause codon is 22 % for peptide A, 36 % for peptide B, 13 % for peptide C, and 2 %, 4 %, and 9 % for the different types of pauses of peptide D.
In future, we aim to elucidate the molecular mechanisms that lead to strong pauses during protein synthesis and to understand why some of the translating ribosomes get inactivated.

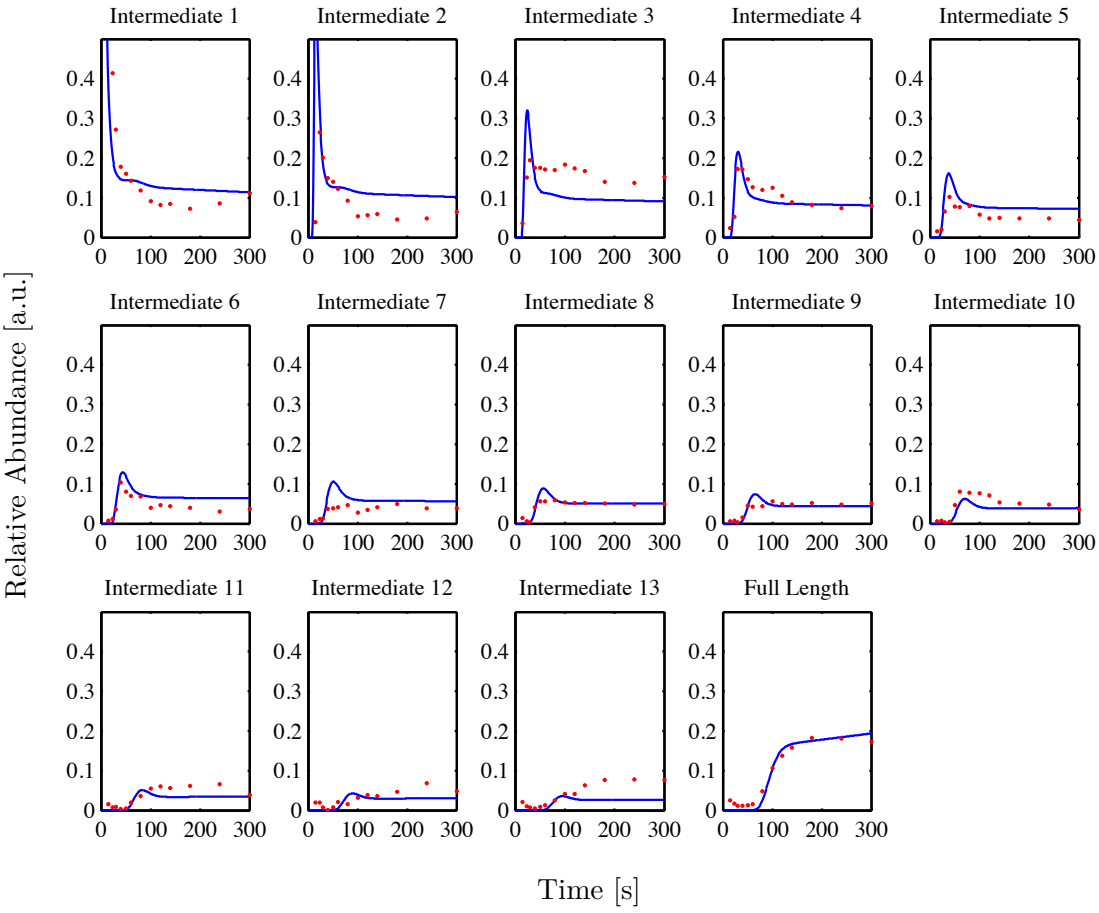


Figure A.4: **Relative abundance vs. time of 13 intermediate peptides and full length product of peptide C.** Red dots represent experimental data, blue solid lines are theoretical curves resulting from the Markov model introduced in section A.3 using fitted parameters. Rates as obtained by fitting are: $\omega_{\mathrm{F}} = 5/s$, $\omega_{\mathrm{P}} = 0.2/s$, $\omega_{\mathrm{I}} = 0.0302/s$, $\omega_{\mathrm{A}} = 0.0005/s$. Only these four parameters were adjusted to fit simultaneously all 14 data sets.

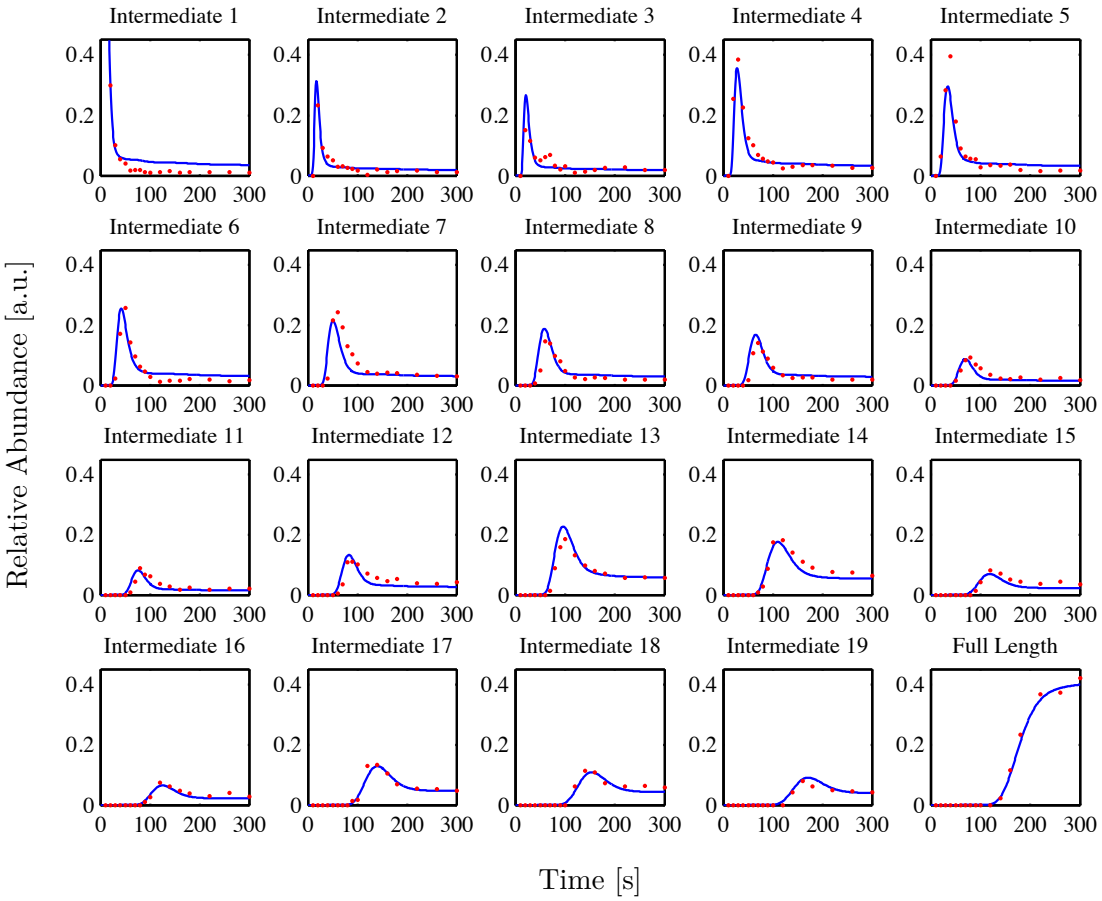


Figure A.5: **Relative abundance vs. time of 19 intermediate peptides and full length product of peptide D.** Red dots represent experimental data, blue solid lines are theoretical curves obtained by fitting the Markov model defined in section A.3 with a slight variation: The higher quality of the experimental data for this peptide allowed a more detailed analysis than in the other cases. Thus, we divided the slow codons of mRNA D into three groups: pause codons with a "standard" slow translation rate ω_{P}, with a faster rate $\omega_{\mathrm{P,fst}}$, and codons with a very slow forward rate $\omega_{\mathrm{P,slw}}$. Rates as obtained by fitting are: $\omega_{\mathrm{F}} = 4/s$, $\omega_{\mathrm{P}} = 0.19/s$, $\omega_{\mathrm{P,fst}} = 0.35/s$ $\omega_{\mathrm{P,slw}} = 0.08/s$, $\omega_{\mathrm{I}} = 0.008/s$, $\omega_{\mathrm{A}} = 0.0005/s$. Fast pause codons correspond to intermediates 2, 3, 10, 11. Slow pause codons correspond to intermediates 13, 14, 17, 18, 19. Note that only these six parameters were adjusted to fit simultaneously all 20 data sets.

B Computer Simulations

To check the consistence of our analytic theory with experimental findings, we performed stochastic simulations of the elongation process based on the codon-specific elongation rates as given by (3.10). As we described in [93], we simulated the motion of ribosomes along a mRNA as a simple stepping of self excluding extended objects on a linear, inhomogeneous chain. New ribosomes enter the initially empty mRNA only if enough space is provided. That means that the A site of the ribosome closest to the start codon must be at least one ribosome size (12 codons) away from it. The time between two of these initiation events is exponentially distributed with an initiation rate that is low enough to avoid strong ribosome queuing. The ribosomes dwell on a codon before they move to the next one provided that it is not covered by the preceding ribosome. The dwell times correspond to the codon-specific elongation times as given by (3.10). A blocked ribosome moves on immediately after the preceding ribosome has gone and the dwell time has passed.

C Additional Figures

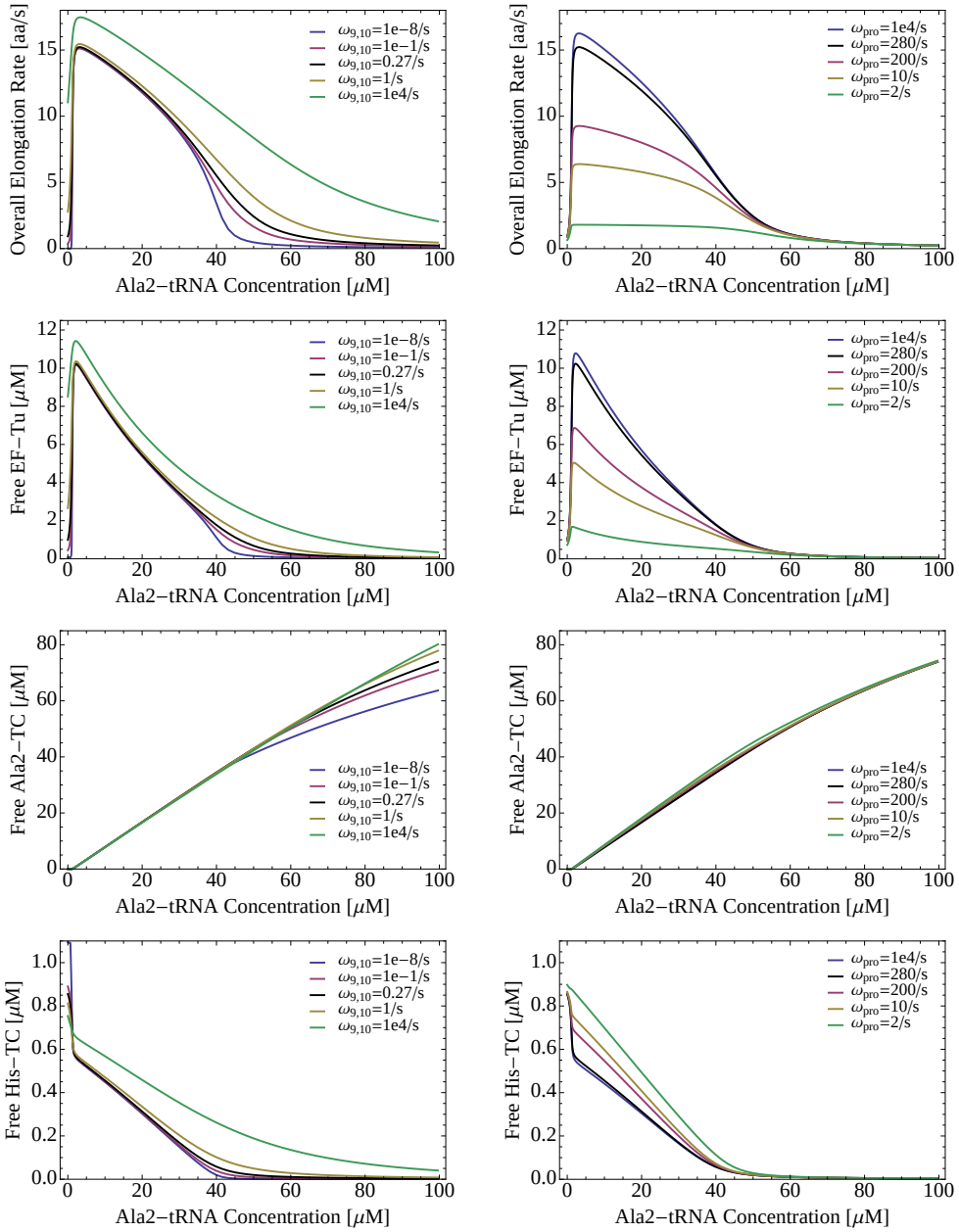


Figure C.1: **Ribosomal kinetics.** The near-cognate accomodation rate $\omega_{9,10}$ (left panel) determines how much the OERC resembles a Michaelis-Menten function, whereas ω_{pro} (right panel) defines the flatness of the OERC.

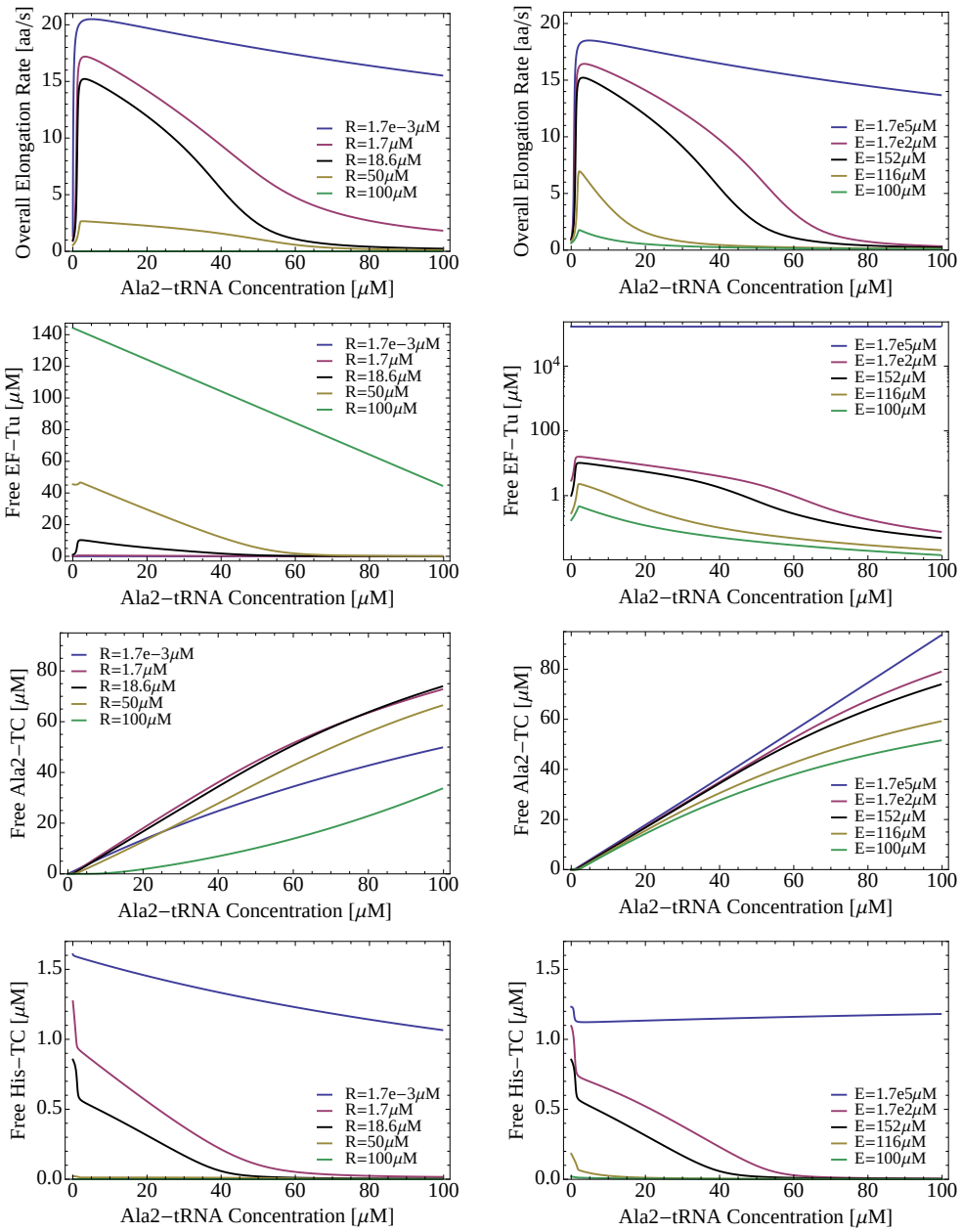


Figure C.2: **Factor concentrations.** Small ribosome concentrations $\mathcal{R}$ (left panel) as well as large EF-Tu concentrations $\mathcal{E}$ (right panel) lead to substrate inhibition of non-cognates, whereas for small EF-Tu concentrations the competition for EF-Tu dominates.

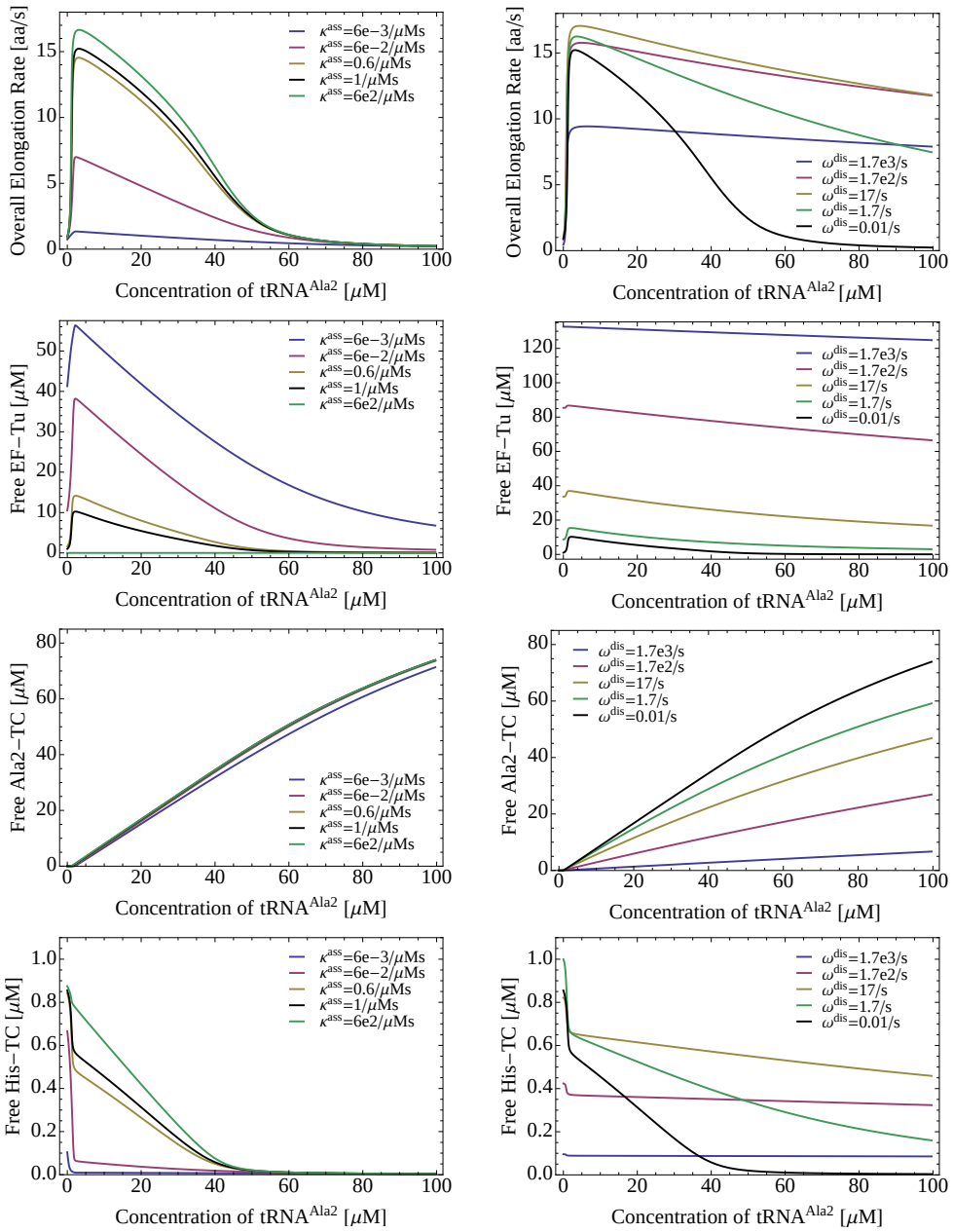


Figure C.3: **TC formation kinetics.** Both slow formation (small κ^{ass}, left panel) and fast dissociation (large ω^{dis}, right panel) of TCs decrease the effect of EF-Tu competition such that substrate inhibition becomes the dominating effect for large tRNA concentrations.

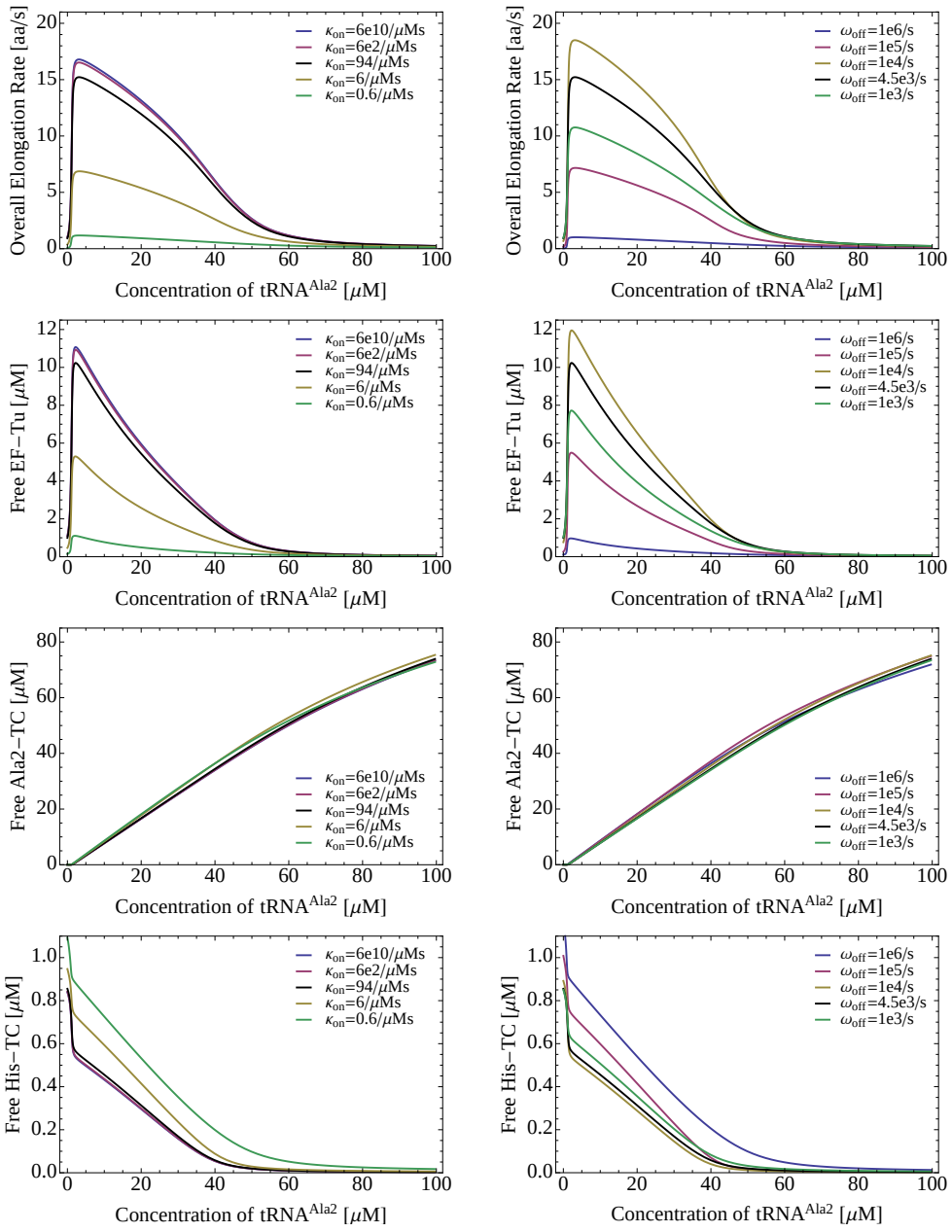


Figure C.4: **Initial TC binding and unbinding.** A small κ_{on} weakens both the substrate inhibition effect of non-cognates and the effect of competition for EF-Tu molecules, therefore the OERC remains almost unchanged (left panel). A decreased unbinding rate ω_{off} just increases the substrate-inhibition effect and leads to a smoother OERC (right panel).

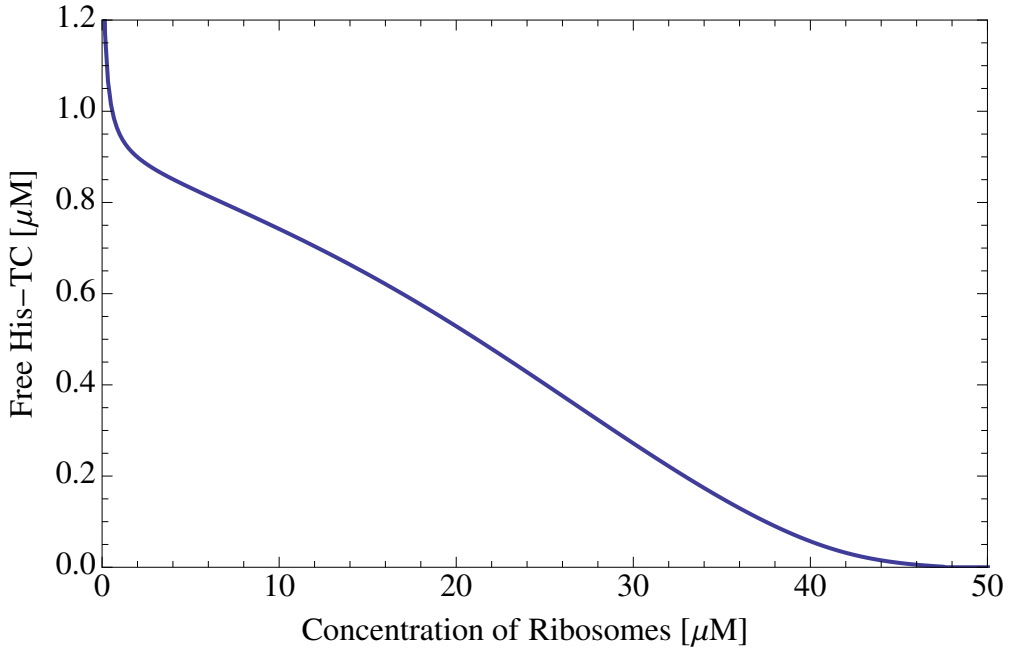


Figure C.5: **Dependence of the concentration of free TCs containing $tRNA^{His}$ on the concentration of ribosomes in *E. coli* for a growth rate of 0.7 dbl/hour.** Each ribosome can bind up to three tRNA molecules, which are then not available for the formation of TCs until they leave the ribosome. Thus, the more ribosomes are present, the less TCs are formed. In *E. coli* for a growth rate of 0.7 dbl/hour, TCs containing $tRNA^{His}$ are the most affected by depletion through active ribosomes.

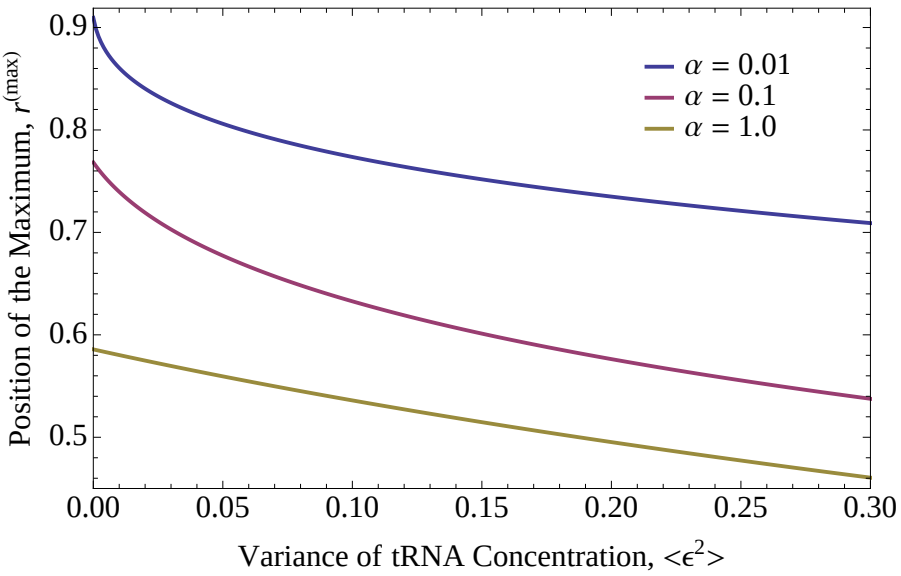


Figure C.6: **Dependence of the position of the maximal total elongation flux on variance in tRNA concentration.** The ribosomal concentration $r^{(\max)}$ leading to a maximal average total elongation flux $\langle \mathrm{J}_{\mathrm{elo}} \rangle (r^{(\max)})$ (7.15) decreases as the variance of the tRNA concentration $\langle \epsilon^2 \rangle$ increases. Results are shown for three different choices of the ratio α of the backward and forward rates (7.9).

D Additional Tables

Table D.1: **Codon-specific elongation rates $\omega^{\star}_{c,\mathrm{elo}}$ for all sense codons c in *E. coli*, assuming a 2-1-2 pathway of tRNA release from the E site.** All rates in $[s^{-1}]$.

	E. coli growth rates [dbl/h]					*E. coli* growth rates [dbl/h]			
	0.7	1.07	1.6	2.5		0.7	1.07	1.6	2.5
AAA	15.3	16.4	14.6	12.3	GAA	38.9	40.1	53.8	63.7
AAC	7.9	10.2	14.8	16.6	GAC	20.9	21.1	30.0	39.0
AAG	15.3	16.3	14.5	12.2	GAG	38.6	39.9	53.3	63.2
AAU	7.8	10.1	14.6	16.4	GAU	19.6	19.9	28.3	37.1
ACA	11.8	11.5	17.7	22.7	GCA	26.7	32.5	37.5	43.1
ACC	10.6	11.1	12.3	16.2	GCC	5.1	6.4	9.9	10.4
ACG	19.6	19.2	27.4	32.9	GCG	25.9	31.5	36.6	42.3
ACU	19.2	19.4	25.6	32.2	GCU	25.7	31.3	36.6	41.9
AGA	12.2	12.8	16.2	15.9	GGA	17.1	17.4	24.9	23.1
AGC	15.8	15.7	19.2	16.7	GGC	36.8	39.1	44.2	55.9
AGG	8.7	8.3	13.0	10.6	GGG	26.9	27.5	38.2	36.7
AGU	15.6	15.5	19.0	16.6	GGU	35.0	37.3	42.1	53.7
AUA	3.0	3.1	4.7	5.5	GUA	31.7	28.0	45.4	48.4
AUC	31.9	34.3	46.8	59.4	GUC	14.8	16.0	20.8	23.1
AUG	5.7	7.0	10.2	10.9	GUG	29.2	25.6	42.2	45.4
AUU	30.5	33.1	44.9	57.4	GUU	38.6	36.4	53.5	56.5
CAA	10.9	14.2	12.4	16.0	UAA		*stop*		
CAC	4.0	6.3	7.6	11.9	UAC	22.5	22.2	33.5	33.9
CAG	6.4	8.2	13.0	16.3	UAG		*stop*		
CAU	4.0	6.3	7.6	11.8	UAU	22.0	21.8	33.0	33.4
CCA	3.9	5.6	4.7	3.9	UCA	22.0	20.7	26.6	26.2
CCC	11.3	9.8	17.7	14.5	UCC	9.0	9.6	11.9	12.3
CCG	9.8	15.3	10.7	9.7	UCG	24.3	23.3	29.8	30.0
CCU	13.7	13.8	20.2	16.8	UCU	26.4	25.8	32.5	32.6
CGA	40.9	41.9	58.1	62.8	UGA		*stop*		
CGC	38.9	39.9	56.0	60.4	UGC	19.6	19.8	27.1	26.4
CGG	12.3	6.6	10.7	9.3	UGG	11.5	13.6	17.2	21.7
CGU	38.8	39.8	55.9	60.3	UGU	19.2	19.5	26.7	26.1
CUA	9.1	9.3	11.6	12.0	UUA	15.5	17.3	16.5	15.5
CUC	13.1	14.7	17.7	18.8	UUC	7.4	10.7	9.8	8.1
CUG	43.7	46.5	56.6	60.3	UUG	34.1	36.1	44.1	44.3
CUU	12.6	14.2	16.9	18.1	UUU	6.9	10.1	9.2	7.6

Table D.2: **Codon-specific elongation rates $\omega^{\star}_{c,\mathrm{elo}}$ for all sense codons c in *E. coli*, assuming a 2-3-2 pathway of tRNA release from the E site.** All rates in $[\mathrm{s}^{-1}]$.

	E. coli growth rates [dbl/h]					*E. coli* growth rates [dbl/h]			
	0.7	1.07	1.6	2.5		0.7	1.07	1.6	2.5
AAA	15.2	15.4	12.0	7.2	GAA	40.0	41.9	58.1	75.1
AAC	7.6	9.6	14.8	18.4	GAC	21.3	21.4	31.9	47.2
AAG	15.1	15.4	11.9	7.1	GAG	39.7	41.7	57.6	74.7
AAU	7.5	9.5	14.6	18.2	GAU	20.0	20.2	30.2	45.2
ACA	12.2	12.0	19.3	27.8	GCA	27.1	33.2	39.2	50.2
ACC	10.6	10.8	12.1	19.0	GCC	5.1	6.4	10.6	12.5
ACG	20.3	20.4	30.3	40.6	GCG	26.3	32.2	38.2	49.3
ACU	19.6	19.6	26.7	38.2	GCU	26.1	32.1	38.4	49.1
AGA	12.9	14.1	18.8	20.8	GGA	18.0	19.2	28.6	29.9
AGC	16.4	16.8	21.6	20.7	GGC	37.7	40.7	47.2	66.6
AGG	9.2	9.2	15.2	14.1	GGG	28.2	30.0	43.2	46.4
AGU	16.3	16.7	21.4	20.5	GGU	36.0	38.9	45.2	64.4
AUA	3.1	3.4	5.5	7.4	GUA	32.6	28.8	48.9	57.9
AUC	32.8	35.8	50.6	70.9	GUC	15.3	16.9	23.1	29.1
AUG	5.7	6.8	10.4	12.5	GUG	30.0	26.3	45.6	54.7
AUU	31.3	34.5	48.6	68.9	GUU	39.6	37.7	57.7	67.0
CAA	11.3	15.4	13.9	20.3	UAA		*stop*		
CAC	3.8	6.0	7.3	13.9	UAC	23.2	23.5	37.1	42.5
CAG	6.3	7.8	13.3	19.1	UAG		*stop*		
CAU	3.8	5.9	7.3	13.9	UAU	22.7	23.0	36.5	42.0
CCA	3.8	5.5	4.3	3.8	UCA	22.8	22.1	29.6	32.5
CCC	11.8	10.6	20.2	18.4	UCC	9.3	9.9	12.7	14.7
CCG	9.8	15.4	10.0	9.7	UCG	25.3	24.9	33.2	37.4
CCU	14.1	14.4	22.2	20.2	UCU	27.4	27.3	35.6	39.7
CGA	42.0	43.7	62.7	74.3	UGA		*stop*		
CGC	40.0	41.8	60.6	71.9	UGC	20.5	21.6	30.8	33.9
CGG	13.0	7.3	12.5	12.1	UGG	11.9	14.5	19.4	27.9
CGU	39.9	41.6	60.5	71.8	UGU	20.1	21.2	30.3	33.5
CUA	9.5	10.0	12.9	15.6	UUA	16.2	18.8	18.8	19.9
CUC	13.7	15.9	20.0	23.6	UUC	7.2	10.4	9.1	6.8
CUG	45.0	48.8	61.2	71.9	UUG	35.5	38.9	49.3	55.2
CUU	13.1	15.4	19.2	22.8	UUU	6.7	9.8	8.5	6.4

Table D.3: **Concentrations of free TCs in *E. coli* for four different growth rates, assuming a 2-1-2 pathway of tRNA release from the E site.** All concentrations in [µM].

	E. coli growth rates [dbl/h]					*E. coli* growth rates [dbl/h]			
	0.7	1.07	1.6	2.5		0.7	1.07	1.6	2.5
Ala1B	5.70	7.89	8.77	9.81	Leu5	2.69	3.28	2.86	2.41
Ala2	0.83	1.15	1.68	1.65	Lys	2.62	3.05	2.46	1.88
Arg2	9.74	10.57	16.20	16.77	Met m	1.03	1.38	1.90	1.86
Arg3	2.27	1.24	1.94	1.49	Phe	1.24	2.03	1.70	1.26
Arg4	2.09	2.34	2.87	2.54	Pro1	1.10	2.07	1.11	0.95
Arg5	1.43	1.45	2.20	1.58	Pro2	1.86	1.71	3.10	2.26
Asn	1.28	1.84	2.57	2.74	Pro3	0.62	1.00	0.77	0.58
Asp1	4.14	4.51	6.42	8.60	Sec	0.79	0.88	0.95	0.92
Cys	3.88	4.22	5.65	4.98	Ser1	3.95	3.96	4.89	4.37
Gln1	1.87	2.73	2.17	2.68	Ser2	0.89	1.01	1.17	1.16
Gln2	1.08	1.53	2.31	2.74	Ser3	3.08	3.28	3.82	3.01
Glu2	9.85	10.87	15.46	19.27	Ser5	1.44	1.65	1.95	1.82
Gly1	2.52	2.73	3.85	3.51	Thr1	0.17	0.24	0.21	0.28
Gly2	3.63	3.95	5.62	4.70	Thr2	1.62	1.73	2.19	2.32
Gly3	9.33	10.76	11.53	15.38	Thr3	1.55	1.73	1.85	2.31
His	0.64	1.12	1.29	1.91	Thr4	1.94	2.05	3.05	3.73
Ile1	6.66	7.93	11.19	15.38	Trp	2.00	2.56	3.11	3.64
Ile2	0.46	0.52	0.75	0.83	Tyr1	1.56	1.80	3.14	2.82
Leu1	10.35	12.03	14.68	14.77	Tyr2	2.49	2.50	3.56	3.39
Leu2	2.46	3.04	3.46	3.42	Val1	7.24	6.46	11.62	11.69
Leu3	1.64	1.80	2.12	1.99	Val2	2.78	3.26	4.09	4.35
Leu4	5.48	6.05	8.31	7.60					

Table D.4: **Concentrations of free TCs in *E. coli* for four different growth rates, assuming a 2-3-2 pathway of tRNA release from the E site.** All concentrations in [µM].

	E. coli growth rates [dbl/h]					*E. coli* growth rates [dbl/h]			
	0.7	1.07	1.6	2.5		0.7	1.07	1.6	2.5
Ala1B	5.43	7.14	7.70	8.48	Leu5	2.66	3.20	2.79	2.26
Ala2	0.77	1.00	1.49	1.42	Lys	2.43	2.50	1.65	0.74
Arg2	9.54	10.00	15.36	15.90	Met m	0.96	1.18	1.62	1.53
Arg3	2.27	1.22	1.92	1.40	Phe	1.14	1.74	1.31	0.74
Arg4	2.08	2.33	2.84	2.41	Pro1	1.04	1.87	0.87	0.68
Arg5	1.42	1.44	2.19	1.52	Pro2	1.83	1.65	3.00	2.07
Asn	1.16	1.52	2.13	2.15	Pro3	0.58	0.87	0.58	0.39
Asp1	3.97	4.03	5.74	7.82	Sec	0.79	0.87	0.95	0.92
Cys	3.84	4.12	5.53	4.78	Ser1	3.89	3.79	4.65	4.02
Gln1	1.84	2.64	2.05	2.46	Ser2	0.88	0.99	1.15	1.14
Gln2	0.99	1.28	1.96	2.29	Ser3	3.03	3.15	3.65	2.70
Glu2	9.61	10.19	14.44	17.98	Ser5	1.39	1.51	1.75	1.57
Gly1	2.52	2.70	3.82	3.44	Thr1	0.16	0.21	0.17	0.23
Gly2	3.61	3.90	5.55	4.51	Thr2	1.60	1.69	2.14	2.23
Gly3	9.08	10.06	10.53	14.33	Thr3	1.46	1.47	1.50	1.94
His	0.57	0.93	1.03	1.61	Thr4	1.89	1.89	2.78	3.33
Ile1	6.47	7.38	10.41	14.56	Trp	1.96	2.48	3.00	3.50
Ile2	0.46	0.52	0.74	0.79	Tyr1	1.52	1.70	2.98	2.68
Leu1	10.17	11.51	13.94	14.00	Tyr2	2.44	2.36	3.38	3.23
Leu2	2.42	2.93	3.33	3.12	Val1	7.04	5.92	10.78	10.72
Leu3	1.60	1.72	2.00	1.88	Val2	2.71	3.05	3.84	4.03
Leu4	5.46	5.99	8.21	7.47					

Bibliography

[1] Assen Marintchev and Gerhard Wagner. Translation initiation: structures, mechanisms and evolution. *Quarterly Reviews of Biophysics*, 37(3-4):197–284, 2004.

[2] Sabine Petry, Albert Weixlbaumer, and Venki Ramakrishnan. The termination of translation. *Current Opinion in Structural Biology*, 18(1):70–77, 2008.

[3] Hengjiang Dong, Lars Nilsson, and Charles G. Kurland. Co-variation of tRNA abundance and codon usage in Escherichia coli at different growth rates. *Journal of Molecular Biology*, 260(5):649–663, 1996.

[4] Emily B. Kramer and Philip J. Farabaugh. The frequency of translational misreading errors in E. coli is largely determined by trna competition. *RNA-a Publication of the RNA Society*, 13(1):87–96, 2007.

[5] Witold Szaflarski, Oliver Vesper, Yoshika Teraoka, Bealta Plitta, Daniel N. Wilson, and Knud H. Nierhaus. New features of the ribosome and ribosomal inhibitors: Non-enzymatic recycling, misreading and back-translocation. *Journal of Molecular Biology*, 380(1):193–205, 2008.

[6] David Goodsell. Elongation factors September 2006 molecule of the month. doi: 10.2210/rcsb_pdb/mom_2006_9. URL `http://www.rcsb.org/pdb/101/motm.do?momID=81`.

[7] David Goodsell. Aminoacyl-tRNA synthetases April 2001 molecule of the month. doi: 10.2210/rcsb_pdb/mom_2001_4. URL `http://www.rcsb.org/pdb/101/motm.do?momID=16`.

[8] David Goodsell. Ribosome January 2010 molecule of the month. doi: 10.2210/rcsb_pdb/mom_2010_1. URL `http://www.rcsb.org/pdb/101/motm.do?momID=121`.

[9] Venki Ramakrishnan. Ribosome structure and the mechanism of translation. *Cell*, 108(4):557–572, 2002.

[10] T. Martin Schmeing and Venki Ramakrishnan. What recent ribosome structures have revealed about the mechanism of translation. *Nature*, 461(7268):1234–1242, 2009.

[11] Denis L. J. Lafontaine and David Tollervey. The function and synthesis of ribosomes. *Nature Reviews Molecular Cell Biology*, 2(7):514–520, 2001.

[12] Yuri P. Semenkov, Marina V. Rodnina, and Wolfgang Wintermeyer. The "allosteric three-site model" of elongation cannot be confirmed in a well-defined ribosome system from Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America*, 93:12183 – 12188, 1996.

[13] Sotaro Uemura, Colin E. Aitken, Jonas Korlach, Benjamin A. Flusberg, Stephen W. Turner, and Joseph D. Puglisi. Real-time tRNA transit on single translating ribosomes at codon resolution. *Nature*, 464:1012 – 1017, 2010.

[14] Chunlai Chen, Benjamin Stevens, Jaskiran Kaur, Zeev Smilansky, Barry S. Cooperman, and Yale E. Goldman. Allosteric vs. spontaneous exit-site (E-site) tRNA dissociation early in protein synthesis. *Proceedings of the National Academy of Sciences of the United States of America*, 108(41):16980 – 16985, 2011.

[15] Jin Chen, Alexey Petrov, Albert Tsai, Sean E. O'Leary, and Joseph D. Puglisi. Coordinated conformational and compositional dynamics drive ribosome translocation. *Nature Structural & Molecular Biology*, 20(6):718 – 727, 2013.

[16] Daniel N. Wilson and Knut H. Nierhaus. The E-site story: the importance of maintaining two tRNAs on the ribosome during protein synthesis. *Cellular and Molecular Life Sciences Cellular and Molecular Life Sciences*, 63:2725 – 2737, 2006.

[17] Malte Beringer and Marina V. Rodnina. The ribosomal peptidyl transferase. *Molecular Cell*, 26(3):311 – 321, 2007.

[18] Michael A. Sørensen and Steen Pedersen. Absolute in vivo translation rates of individual codons in Escherichia coli - the 2 glutamic-acid codons GAA and GAG are translated with a threefold difference in rate. *Journal of Molecular Biology*, 222 (2):265–280, 1991.

[19] James F. Curran and Michael Yarus. Rates of aminoacyl-trans-RNA selection at 29 sense codons *in vivo*. *Journal of Molecular Biology*, 209(1):65–77, 1989.

[20] Jin-Der Wen, Laura Lancaster, Courtney Hodges, Ana-Carolina Zeri, Shige H. Yoshimura, Harry F. Noller, Carlos Bustamante, and Ignacio Tinoco. Following translation by single ribosomes one codon at a time. *Nature*, 452(7187):598–U2, 2008.

[21] Chunlai Chen, Haibo Zhang, Steven L Broitman, Michael Reiche, Ian Farrell, Barry S Cooperman, and Yale E Goldman. Dynamics of translation by single ribosomes through mrna secondary structures. *Nature Structural & Molecular Biology*, 20(5):582–588, 2013.

[22] Gene-Wei Li, Eugene Oh, and Jonathan S. Weissman. The anti-shine-dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature*, 484 (7395):538–U172, 2012.

[23] Catherine A. Charneski and Laurence D. Hurst. Positively charged residues are the major determinants of ribosomal velocity. *PLoS Biology*, 11(3), 2013.

[24] Namiko Mitarai, Kim Sneppen, and Steen Pedersen. Ribosome collisions and translation efficiency: Optimization by codon usage and mRNA destabilization. *Journal of Molecular Biology*, 382(1):236–245, 2008.

[25] Günter Kramer, Thomas Rauch, Wolfgang Rist, Sonja Vorderwülbecke, Holger Patzelt, Agnes Schulze-Specking, Nenad Ban, Elke Deuerling, and Bernd Bukau. L23 protein functions as a chaperone docking site on the ribosome. *Nature*, 419 (6903):171–174, 2002.

[26] Gü)nter Kramer, Daniel Boehringer, Nenad Ban, and Bernd Bukau. The ribosome as a platform for co-translational processing, folding and targeting of newly synthesized proteins. *Nature Structural & Molecular Biology*, 16(6):589–597, 2009.

[27] Mario dos Reis, Renos Savva, and Lorenz Wernisch. Solving the riddle of codon usage preferences: a test for translational selection. *Nucleic Acids Research*, 32(17): 5036–5044, 2004.

[28] Tamir Tuller, Asaf Carmi, Kalin Vestsigian, Sivan Navon, Yuval Dorfan, John Zaborske, Tao Pan, Orna Dahan, Itay Furman, and Yitzhak Pilpel. An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell*, 141(2):344–354, 2010.

[29] Nese Bilgin, Måns Ehrenberg, and Charles Kurland. Is translation inhibited by noncognate ternary complexes. *FEBS Letters*, 233(1):95–99, 1988.

[30] Ingo Wohlgemuth, Corinna Pohl, and Marina V. Rodnina. Optimization of speed and accuracy of decoding in translation. *EMBO Journal*, 29(21):3701–3709, 2010.

[31] Magnus Johansson, Elli Bouakaz, Martin Lovmar, and Måns Ehrenberg. The kinetics of ribosomal peptidyl transfer revisited. *Molecular Cell*, 30(5):589–598, 2008.

[32] Aaron Fluitt, Elsje Pienaar, and Hendrik Vijoen. Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis. *Computational Biology and Chemistry*, 31(5-6):335–346, 2007.

[33] Hermioni Zouridis and Vassily Hatzimanikatis. Effects of codon distributions and tRNA competition on protein translation. *Biophysical Journal*, 95(3):1018–1033, 2008.

[34] Hermioni Zouridis and Vassily Hatzimanikatis. A model for protein translation: Polysome self-organization leads to maximum protein synthesis rates. *Biophysical Journal*, 92(3):717–730, 2007.

[35] Chris A. Brackley, M. Carmen Romano, and Marco Thiel. The dynamics of supply and demand in mRNA translation. *PLoS Computational Biology*, 7(10):e1002203, 2011.

[36] Magnus Johansson, Jingji Zhang, and Måns Ehrenberg. Genetic code translation displays a linear trade-off between efficiency and accuracy of trna selection. *Proceedings of the National Academy of Sciences of the United States of America*, 109 (1):131–136, 2012.

[37] Howard M. Taylor and Samuel Karlin. *An Introduction to Stochastic Modeling*. Academic Press, San Diego, 3rd edition, 1998.

[38] Marina V. Rodnina, Tillmann Pape, Rainer Fricke, Lothar Kuhn, and Wolfgang Wintermeyer. Initial binding of the elongation factor Tu·GTP·aminoacyl-tRNA complex preceding codon recognition on the ribosome. *Journal of Biological Chemistry*, 271(2):646–652, 1996.

[39] Tillmann Pape, Wolfgang Wintermeyer, and Marina V. Rodnina. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the E-coli ribosome. *EMBO Journal*, 17(24):7490–7497, 1998.

[40] Kirill B. Gromadski and Marina V. Rodnina. Kinetic determinants of high-fidelity trna discrimination on the ribosome. *Molecular Cell*, 13(2):191–200, 2004.

[41] Kirill B. Gromadski, Tina Daviter, and Marina V. Rodnina. A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity. *Molecular Cell*, 21:369–377, 2006.

[42] Ute Kothe and Marina V. Rodnina. Delayed release of inorganic phosphate from elongation factor Tu following GTP hydrolysis on the ribosome. *Biochemistry*, 45 (42):12767–12774, 2006.

[43] Niels Fischer, Andrey L. Konevega, Wolfgang Wintermeyer, Marina V. Rodnina, and Holger Stark. Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature*, 466(7304):329–333, 2010.

[44] Jörg Mittelstaet, Andrey L. Konevega, and Marina V. Rodnina. A Kinetic Safety Gate Controlling the Delivery of Unnatural Amino Acids to the Ribosome. *Journal of the American Chemical Society*, 135:17031 – 17038, 2013.

[45] Nicholas T. Ingolia, Liana F. Lareau, and Jonathan S. Weissman. Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. *Cell*, 147(4):789–802, 2011.

[46] Ingo Wohlgemuth, Corinna Pohl, Jörg Mittelstaet, Andrey L. Konevega, and Marina V. Rodnina. Evolutionary optimization of speed and accuracy of decoding on the ribosome. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1580):2979–2986, 2011.

[47] Terrell L. Hill. Interrelations between random walks on diagrams (graphs) with and without cycles. *Proceedings of the National Academy of Sciences of the United States of America*, 85(9):2879–2883, 1988.

[48] Jianzhi Zhang. Protein-length distributions for the three domains of life. *Trends in Genetics*, 16(3):107 – 109, 2000.

[49] Hans Bremer and Patrick P. Dennis. *Modulation of Chemical Composition and Other Parameters of the Cell by Growth Rate*, volume 2, pages 1553–1569. ASM Press, Washington, 2. edition, 1996.

[50] Toshimichi Ikemura. Correlation between the abundance of Escherichia coli transfer RNAs and the occurrence of the respective codons in its protein genes: A proposal for a synonymous codon choice that is optimal for the E. coli translational system. *Journal of Molecular Biology*, 151(3):389 – 409, 1981.

[51] Frederick C. Neidhardt, Philip L. Bloch, Steen Pedersen, and Solvejg Reeh. Chemical measurement of steady-state levels of 10 aminoacyl transfer ribonucleic-acid synthetases in Escherichia coli. *Journal of Bacteriology*, 129(1):378–387, 1977.

[52] Sung-Tzu Liang, Ying-Chun Xu, Patrick P. Dennis, and Hans Bremer. mRNA composition and control of bacterial gene expression. *Journal of Bacteriology*, 182 (11):3037–3044, 2000.

[53] Dennis A. Benson, Ilene Karsch-Mizrachi, David J. Lipman, James Ostell, and Eric W. Sayers. GenBank. *Nucleic Acids Research*, 39:D32–D37, 2011.

[54] Gong Zhang, Ivan Fedyunin, Sebastian Kirchner, Chuanle Xiao, Angelo Valleriani, and Zoya Ignatova. FANSe: an accurate algorithm for quantitative mapping of large scale sequencing reads. *Nucleic Acids Research*, 40(11), 2012.

[55] Lili K. Dörfel, Ingo Wohlgemuth, Christina Kothe, Frank Peske, Henning Urlaub, and Marina V. Rodnina. EF-P Is Essential for Rapid Synthesis of Proteins Containing Consecutive Proline Residues. *Science*, 339(6115):85–88, 2013.

[56] Ute Kothe and Marina V. Rodnina. Codon reading by tRNA with modified uridine in the wobble position. *Molecular Cell*, 25:167–174, 2007.

[57] Athel Cornish-Bowden. *Fundamentals of Enzyme Kinetics*. Wiley-Blackwell, January 2012.

[58] Annette Sievers, Malte Beringer, Marina V. Rodnina, and Richard Wolfenden. The ribosome as an entropy trap. *Proceedings of the National Academy of Sciences of the United States of America*, 101(21):7897–7901, 2004.

[59] Paul C. Whitford, Peter Geggier, Roger B. Altman, Scott C. Blanchard, Jose N. Onuchic, and Karissa Y. Sanbonmatsu. Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. *RNA-a Publication of the RNA Society*, 16(6):1196–1204, 2010.

[60] Igor Cestari and Kenneth Stuart. A spectrophotometric assay for quantitative measurement of aminoacyl-tRNA synthetase activity. *Journal of Biomolecular Screening*, 18(4):490–497, 2013.

[61] Nevena Cvetesic, John J. Perona, and Ita Gruic-Sovulj. Kinetic partitioning between synthetic and editing pathways in class I aminoacyl-tRNA synthetases oc-

curs at both pre-transfer and post-transfer hydrolytic steps. *Journal of Biological Chemistry*, 287(30):25381–25394, 2012.

[62] Magnus Johansson, Martin Lovmar, and Måns Ehrenberg. Rate and accuracy of bacterial protein synthesis revisited. *Current Opinion in Microbiology*, 11(2):141–147, 2008.

[63] Sophia Rudorf, Michael Thommen, Marina V. Rodnina, and Reinhard Lipowsky. Deducing the kinetics of protein synthesis *in vivo* from the transition rates measured *in vitro*. *PLoS Computational Biology*, 10(10):e1003909, 2014.

[64] S. Condamin, Olivier Benichou, and Michel Moreau. Random walks and brownian motion: A method of computation for first-passage times and related quantities in confined geometries. *Physical Review E*, 75(2), 2007.

[65] Michael B. Elowitz, Michael G. Surette, Pierre-Etienne Wolf, Jeffry B. Stock, and Stanislas Leibler. Protein mobility in the cytoplasm of Escherichia coli. *Journal of Bacteriology*, 181(1):197–203, 1999.

[66] Barbar S. Schuwirth, Maria A. Borovinskaya, Cathy W. Hau, Wen Zhang, Anton Vila-Sanjurjo, James M. Holton, and Jamie H. D. Cate. Structures of the bacterial ribosome at 3.5 angstrom resolution. *Science*, 310(5749):827–834, 2005.

[67] Paul C. Whitford, Scott C. Blanchard, Jamie H. D. Cate, and Karissa Y. Sanbonmatsu. Connecting the Kinetics and Energy Landscape of tRNA Translocation on the Ribosome. *PLoS Computational Biology*, 9(3), 2013.

[68] Lars V. Bock, Christian Blau, Gunnar F. Schroeder, Iakov I. Davydov, Niels Fischer, Holger Stark, Marina V. Rodnina, Andrea C. Vaiana, and Helmut Grubmueller. Energy barriers and driving forces in tRNA translocation through the ribosome. *Nature Structural & Molecular Biology*, 20(12):1390–1396, 2013.

[69] Sarah Ledoux and Olke C. Uhlenbeck. Different aa-tRNAs are selected uniformly on the ribosome. *Molecular Cell*, 31:114–123, 2008.

[70] Nandini Manickam, Nabanita A.G Nabanita, Aleeza Abbasi, Kishan Patel, and Philip J. Farabaugh. Studies of translational misreading in vivo show that the ribosome very efficiently discriminates against most potential errors. *RNA*, 20: 9–15, 2014.

[71] Mariana Pavon-Eternod, Suzanna Gomes, Renaud Geslain, Qing Dai, Marsha R. Rosner, and Tao Pan. tRNA over-expression in breast cancer and functional consequences. *Nucleic Acids Research*, 37(21):7268–7280, 2009.

[72] Ying Zhou, Jeffrey M. Goodenbour, Lucy A. Godley, Amittha Wickrema, and Tao Pan. High levels of tRNA abundance and alteration of tRNA charging by bortezomib in multiple myeloma. *Biochemical and Biophysical Research Communications*, 385(2):160–164, 2009.

[73] Shellt Mahlab, Tamir Tuller, and Michal Linial. Conservation of the relative tRNA composition in healthy and cancerous tissues. *RNA-a Publication of the RNA Society*, 18(4):640–652, 2012.

[74] John M. Shoffner, Marie T. Lott, Angela M. S. Lezza, Peter Seibel, Scott W. Ballinger, and Douglas C. Wallace. Myoclonic epilepsy and ragged-red fiber disease (MERRF) is associated with a mitochondrial-DNA transfer RNALys mutation. *Cell*, 61(6):931–937, 1990.

[75] Jose A. Enriquez, Anne Chomyn, and Giuseppe Attardi. MtDNA mutation in MERRF-syndrome causes defective aminoacylation of tRNA(Lys) and premature translation termination. *Nature Genetics*, 10(1):47–55, 1995.

[76] Takehiro Yasukawa, Tsutomu Suzuki, Norie Ishii, Shigeo Ohta, and Kimitsuna Watanabe. Wobble modification defect in tRNA disturbs codon-anticodon interaction in a mitochondrial disease. *EMBO Journal*, 20(17):4794–4802, 2001.

[77] Frederick H. Wilson, Ali Hariri, Anita Farhi, Hongyu Y. Zhao, Kitt F. Petersen, Hakan R. Toka, Carol Nelson-Williams, Khalid M. Raja, Michael Kashgarian, Gerald I. Shulman, Steven J. Scheinman, and Richard P. Lifton. A cluster of metabolic defects caused by mutation in a mitochondrial tRNA. *Science*, 306(5699):1190–1194, 2004.

[78] Jiqiang Ling, Herve Roy, Daoming M. Qin, Mary A. T. Rubio, Juan D. Alfonzo, Kurt Fredrick, and Michael Ibba. Pathogenic mechanism of a human mitochondrial tRNA(Phe) mutation associated with myoclonic epilepsy with ragged red fibers syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, 104(39):15299–15304, 2007.

[79] Geng Wang, Eriko Shimada, Jin Zhang, Jason S. Hong, Geoffrey M. Smith, Michael A. Teitell, and Carla M. Koehler. Correcting human mitochondrial muta-

tions with targeted RNA import. *Proceedings of the National Academy of Sciences of the United States of America*, 109(13):4840–4845, 2012.

[80] Jaehoon Lee, Robyn D. Moir, Kerri B. McIntosh, and Ian M. Willis. TOR signaling regulates ribosome and tRNA synthesis via LAMMER/Clk and GSK-3 family kinases. *Molecular Cell*, 45(6):836–843, 2012.

[81] Debrah M. Thompson and Roy Parker. Stressing out over tRNA cleavage. *Cell*, 138(2):215–219, 2009.

[82] Luis Delgado-Olivares, Efrain Zamora-Romo, Gabriel Guarneros, and Javier Hernadez-Sanchez. Codon-specific and general inhibition of protein synthesis by the tRNA-sequestering minigenes. *Biochimie*, 88(7):793–800, 2006.

[83] Ming Gong, Feng Gong, and Charles Yanofsky. Overexpression of tnaC of Escherichia coli inhibits growth by depleting $tRNA_2^{Pro}$ availability. *Journal of Bacteriology*, 188(5):1892–1898, 2006.

[84] Renaud Geslain, Franck Martin, Alain Camasses, and Gilbert Eriani. A yeast knockout strain to discriminate between active and inactive tRNA molecules. *Nucleic Acids Research*, 31(16):4729–4737, 2003.

[85] Dale Lindsley, Jonathan Gallant, and Gabriel Guarneros. Ribosome bypassing elicited by tRNA depletion. *Molecular Microbiology*, 48(5):1267–1274, 2003.

[86] Igor B. Rosenwald. The role of translation in neoplastic transformation from a pathologist's point of view. *Oncogene*, 23(18):3230–3247, 2004.

[87] Måns Ehrenberg, Hans Bremer, and Patrick P. Dennis. Medium-dependent control of the bacterial growth rate. *Biochimie*, 95:643–658, 2013.

[88] Patrick P. Dennis, Måns Ehrenberg, and Hans Bremer. Control of rRNA synthesis in Escherichia coli: a systems biology approach. *Microbiology and Molecular Biology Reviews*, 68:639–668, 2004.

[89] Yury S. Polikanov, Gregor M. Blaha, and Thomas A. Steitz. How hibernation factors RMF, HPF, and YfiA turn off protein synthesis. *Science*, 336:915–918, 2012.

[90] Akira Wada, Yukiko Yamazaki, Nobuyuki Fujita, and Akira Ishihama. Structure and probable genetic location of a ribosome modulation factor associated with 100S

ribosomes in stationary-phase Escherichia-coli-cells. *Proceedings of the National Academy of Sciences of the United States of America*, 87:2657–2661, 1990.

[91] Michael A. Zundel, Georgeta N. Basturea, and Murray P. Deutscher. Initiation of ribosome degradation during starvation in Escherichia coli. *RNA-a Publication of the RNA Society*, 15:977–983, 2009.

[92] Raquel Moreno-Loshuertos, Gustavo Ferrin, Rebeca Acin-Perez, M. Esther Gallardo, Carlo Viscomi, Acisclo Perez-Martos, Massimo Zeviani, Patricio Fernandez-Silva, and Jose Antonio Enriquez. Evolution meets disease: Penetrance and functional epistasis of mitochondrial tRNA mutations. *PLoS Genetics*, 7(4), 2011.

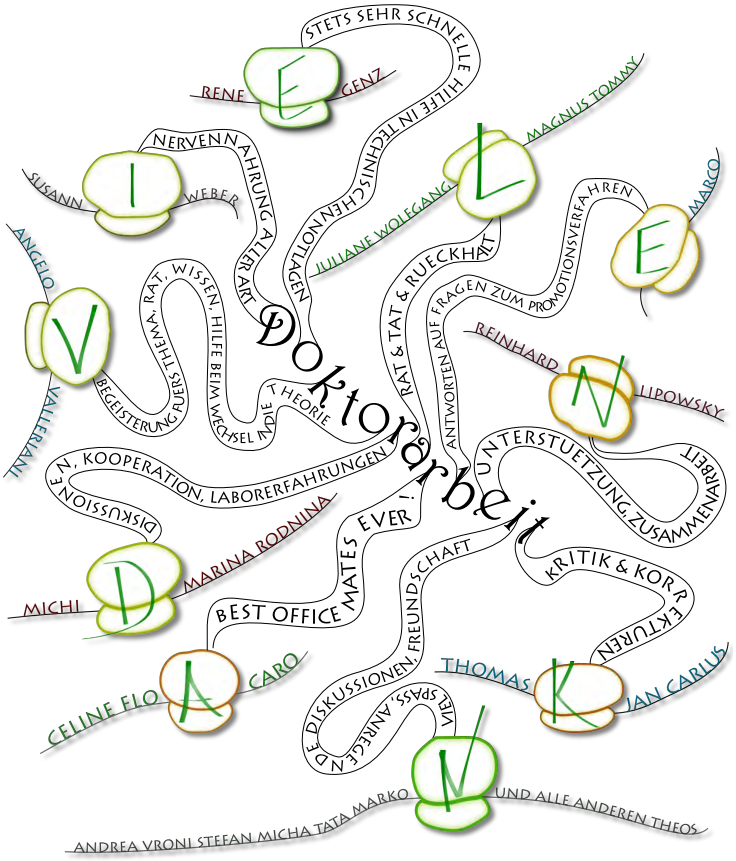
[93] Carlus Deneke, Sophia Rudorf, and Angelo Valleriani. Transient phenomena in gene expression after induction of transcription. *PLoS One*, 7(4), 2012.

List of Figures

List of Tables

Danksagung

Ich möchte allen sehr herzlich danken, die mich auf dem mitunter auch steinigen und langwierigen Weg bis hierher begleitet haben. An erster Stelle danke ich meinem Doktorvater, Prof. Dr. Reinhard Lipowsky, für seine großzügige Unterstützung und die enge Zusammenarbeit. Ich bin ihm nicht nur für die wissenschaftliche Begleitung und die vielen Möglichkeiten, die er mir im Rahmen meiner Promotion eröffnete, sondern auch für Vertrauen und Geduld, Ermutigungen und Ratschläge äußerst dankbar. Mein herzlichster Dank gilt auch allen Freunden, Kollegen und meiner Familie, die alle zum Gelingen dieser Arbeit beigetragen haben:



Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit bisher an keiner anderen Hochschule eingereicht, sowie selbständig und ausschließlich mit den angegebenen Mitteln angefertigt habe.

Potsdam, den 25.11.2014 Sophia Rudorf