

# Ribosome Processivity and Co-translational Protein Folding

Dissertation  
for the award of the degree  
“Doctor rerum naturalium” (Dr. rer. nat.)  
of the Georg-August-Universität Göttingen

within the doctoral program *GGNB – Biomolecules*  
of the Georg-August University School of Science (GAUSS)

submitted by

**Michael Thommen**

from Vevey, Switzerland

Göttingen, 2015





## Members of Examination Board /Thesis Committee

Prof. Dr. Marina Rodnina (1<sup>st</sup> Referee)  
Department of Physical Biochemistry  
Max Planck Institute for Biophysical Chemistry  
Göttingen Germany

Prof. Dr. Heinz Neumann (2<sup>nd</sup> Referee)  
Department of Applied Synthetic Biology  
Institute for Molecular Structural Biology  
Georg-August University  
Göttingen Germany

Prof. Dr. Marina Bennati  
Electron Spin Resonance Spectroscopy group  
Max Planck Institute for Biophysical Chemistry  
Göttingen Germany

## Additional Members of the Examination Board

Prof. Dr. Claudia Höbartner  
Institute for Organic and Biomolecular Chemistry  
Georg-August University  
Göttingen Germany

Prof. Dr. Helmut Grubmüller  
Department of Theoretical and Computational Biophysics  
Max Planck Institute for Biophysical Chemistry  
Göttingen Germany

Prof. Dr. Holger Stark  
3D Cryo-Electron Microscopy group  
Max Planck Institute for Biophysical Chemistry  
Göttingen Germany

Date of oral examination: 3<sup>rd</sup> of November 2015

## *Affidavit*

I hereby declare that my thesis entitled "Ribosome Processivity and Co-translational Protein Folding" has been written independently and with no other sources and aids than quoted. This thesis (wholly or in part) has not been submitted elsewhere for any academic award or qualification.

Michael Thommen

August, 2015

Göttingen, Germany

## Related publications

Rudorf, S., Thommen, M., Rodnina, M.V., and Lipowsky, R\*. (2014). Deducing the kinetics of protein synthesis *in vivo* from the transition rates measured *in vitro*. PLoS Comput Biol 10, e1003909.

Buhr, F.†, Jha, S.†, Thommen, M.†, Mittelstaet, J., Kutz, F., Schwalbe, H. \*, Rodnina, M.V. \*, and Komar, AA\*. (2016). Synonymous codons direct co-translational folding towards different protein conformations. Mol Cell 61, 341-51.

\* corresponding author

† equal contribution

# Table of Contents

Abstract.....	1
1 Introduction.....	2
1.1 The ribosomal elongation cycle.....	2
1.2 Substrate selection in the A site.....	5
1.3 Incorporation of non-canonical amino acids.....	8
1.4 Codon usage bias.....	10
1.5 Co-translation protein folding.....	13
1.6 Aim of the thesis.....	16
2 Results.....	17
2.1 Modeling of translation elongation rates.....	17
2.1.1 Parameters for modeling of elongation rates <i>in vitro</i> .....	21
2.1.2 Modeling of elongation rates <i>in vivo</i> .....	24
2.2 Pausing during translation of natural mRNA.....	28
2.2.1 Translation intermediates.....	28
2.2.2 Quantitative description of pausing during translation of HemK.....	31
2.2.3 Hybridization of peptidyl-tRNA on micro-array chips.....	41
2.2.4 Hybridization of 3'-labeled tRNA on micro-array.....	45
2.3 Co-translational incorporation of fluorescent probes into nascent peptides.....	50
2.3.1 Fluorescent labels at the N-terminus of nascent chains.....	50
2.3.2 Activity of tRNA <sup>Cys</sup> <sub>GCA</sub> and suppressor tRNA <sup>Cys</sup> from <i>E. coli</i> .....	53
2.3.3 Fluorescent labeling of Cys-tRNA.....	56
2.3.4 Incorporation of fluorescent probes at internal positions.....	57
2.3.5 Context dependence for incorporation of fluorescent probes.....	65
2.4 Codon-specific elongation rates modulate co-translational folding.....	73
2.4.1 Harmonization of synonymous codon usage of GBC.....	73
2.4.2 Sequences of U and H yield differential expression of GBC in <i>E. coli</i> ..	76
2.4.3 Dynamics of translation elongation of sequences coding for GBC.....	77

2.4.4	Co-translational folding of GBC is modulated by the codon usage .....	80
2.4.5	Monitoring the folding of the N-terminal domain of GBC by FRET .....	83
3	Discussion .....	87
3.1	Markov model for translation elongation .....	87
3.2	Sequence specific, transient pausing.....	89
3.3	Incorporation of fluorescent labels mediates by Cys-tRNA.....	95
3.4	Translation elongation and co-translational folding.....	101
4	Material and Methods .....	107
4.1	Equipment .....	107
4.2	Software.....	107
4.3	Chemicals and Consumables .....	108
4.4	Strains and Plasmids .....	109
4.5	Buffers and solutions .....	110
4.6	Molecular biology protocols .....	111
4.6.1	Site-directed mutagenesis following the QuikChange protocol .....	111
4.6.2	Mutagenesis by isothermal assembly (Gibson Assembly) .....	111
4.6.3	Transformation .....	111
4.7	DNA constructs.....	112
4.7.1	HemK constructs .....	112
4.7.2	Subcloning of gene for tRNA <sup>Cys</sup> .....	112
4.7.3	Subcloning of gene for adenylate kinase .....	112
4.7.4	Subcloning of gene for phosphoglycerate kinase .....	112
4.7.5	Subcloning of the gene for codon-optimized HemK.....	112
4.7.6	Site-directed mutagenesis for tRNA <sup>Cys</sup> constructs.....	113
4.7.7	Site-direct mutagenesis for HemK constructs .....	113
4.7.8	Site-direct mutagenesis for GBC constructs.....	113
4.8	General RNA protocols .....	116
4.8.1	Ethanol or isopropanol precipitation.....	116
4.8.2	Phenol or phenol:chloroform extraction .....	116



4.8.3	General protocol for <i>in vitro</i> transcription.....	116
4.9	mRNA constructs .....	117
4.9.1	Short synthetic mRNA .....	117
4.9.2	mRNA for single-turnover translation .....	117
4.10	Protein purification .....	118
4.10.1	CysRS-His <sub>6</sub> .....	118
4.10.2	Initiation and elongation factors .....	118
4.11	Protocols related to tRNA .....	119
4.11.1	TCA precipitation of aminoacyl-tRNA.....	119
4.11.2	Formation of EF-Tu•GTP•aa-tRNA ternary complex.....	119
4.11.3	Native PAGE analysis of aminoacylation level .....	119
4.11.4	Analysis of aa-tRNA by Acid Urea PAGE .....	119
4.11.5	Analysis of peptidyl-tRNA by Bis-Tris PAGE.....	120
4.11.6	Aminoacylation of total tRNA .....	120
4.11.7	Purification of tRNA <sup>Cys</sup> transcripts .....	120
4.11.8	Purification of tRNA <sup>Cys</sup> from E. coli total tRNA.....	121
4.11.9	Aminoacylation of tRNA <sup>Cys</sup> variants.....	121
4.11.10	Labeling of Cys-tRNA with thiol-reactive probes.....	121
4.11.11	Labeling of the α-amino group of Met-tRNA <sup>Met</sup> .....	122
4.11.12	Separation of fluorescence-labeled tRNA by HPLC .....	122
4.12	Ribosomal complexes .....	122
4.12.1	Formation of 70S initiation complexes .....	122
4.12.2	Purification of 70S initiation complexes .....	122
4.12.3	Separation of dipeptides .....	123
4.13	Single-turnover <i>in vitro</i> translation .....	123
4.13.1	Sample preparation .....	123
4.13.2	Pulse-proteolysis .....	123
4.13.3	Translation <i>in vitro</i> with fluorescence-labeled Cys-tRNA.....	123
4.13.4	Tris-tricine PAGE.....	124

4.13.5	TCA precipitation of nascent chains .....	124
4.14	Micro-array experiments .....	124
4.14.1	Micro-array design.....	124
4.14.2	Preparation of peptidyl-tRNA samples for micro-array .....	126
4.14.3	Stripping of micro-array slide .....	126
4.14.4	3'-labeling of tRNA.....	126
4.14.5	3'-labeling of tRNA extracted from stalled nascent chains.....	127
4.14.6	3'-labeling of Spike-In probes .....	128
4.14.7	Hybridization of 3'-labeled total tRNA .....	128
4.14.8	Hybridization of 3'-labeled tRNA extracted from LepB 1-35 aa .....	129
	References .....	130
	Appendix .....	152
	Acknowledgment.....	156
	Curriculum vitae .....	161

## Abstract

Synthesis of polypeptide chains on the ribosome occurs at a non-uniform rate. mRNA sequences translated at a high rate are interspersed by segments that lead to translational pausing. The rate of translation elongation is rate-limiting for conformational sampling during co-translational folding of the nascent chain. The global and local rates of translation elongation are important modulators of the efficiency of co-translational folding. For the understanding of co-translational folding, a description of the dynamics of the nascent chain synthesis has to be established in the first place.

To understand how tRNA pools affect local and global translation velocities, we have constructed a mathematical stochastic model based on the availability and competition of tRNA isoacceptors for the A site. We predicted the elemental rate constants for the decoding steps *in vivo*, which allowed us to calculate codon-specific translation elongation rates for any mRNA sequence. We then have analyzed the pausing during translation of natural mRNA *in vitro*. The accumulation of translation intermediates upon synthesis of the model protein HemK could not be explained by the propensity of mRNA to form local stable secondary structures or by the usage frequency of codons. We have established novel experimental approaches that might allow the identification of the site of translational pausing with single codon resolution.

To investigate co-translational folding in real-time, we developed tools to incorporate FRET probes at internal positions of nascent chains mediated by tRNA<sup>Cys</sup>. We observed that peptide bond formation was limited by substrate positioning in the peptidyl-transferase center of aminoacyl-tRNA carrying larger fluorophores. Glycine residues upstream of the incorporation site tuned the incorporation of larger fluorescent probes by increasing the flexibility of the peptidyl-tRNA in the P site.

The interplay between nascent chain elongation and co-translational folding was studied using the model protein  $\gamma$ -B-crystallin encoded by the original bovine sequence and a sequence harmonized for *Escherichia coli* codon usage. We observed differences in local and global translation elongation for both mRNA sequences. By applying a pulse proteolysis approach and the incorporation of FRET pairs in the nascent chain, we detected altered kinetics of co-translational folding for both domains of  $\gamma$ -B-crystallin depending on the mRNA sequence. In summary, we show how changes in local translation elongation can influence the acquisition of folded states in the nascent chain.

# Introduction

## 1 Introduction

### 1.1 The ribosomal elongation cycle

In all living cells, genetic information encoded in DNA is transcribed into RNA sequences by RNA polymerases. RNA molecules can be functionally divided in coding and non-coding RNA that can be either translated into a polypeptide sequence (messenger RNA, mRNA) or fulfill functional roles as major components of ribosomes (ribosomal RNA, rRNA), adaptors (transfer RNA, tRNA), or regulators. On the ribosome, mRNA are translated into polypeptide sequences that can fold into functional proteins. The ribosome is composed of two subunits. In the case of the prokaryotic ribosomes, the small subunit, which has a sedimentation coefficient of 30 Svedberg units (S), and the large 50S subunit form the 70S ribosome.

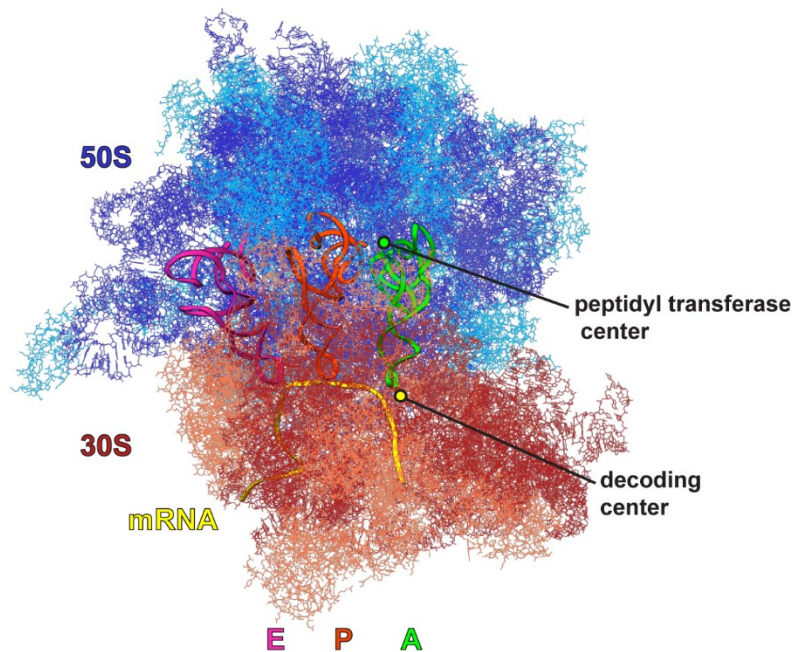


Figure 1-1: Crystal structure of the 70S ribosome containing mRNA and tRNA (Jenner et al., 2010b). tRNA in the A, P, and E site are highlighted in green, orange, and purple, respectively. The mRNA is indicated in yellow. The rRNA is shown in darker blue and red for the 50S and 30S subunit, respectively. Ribosomal proteins are drawn in lighter shades for both subunits.

The 30S subunit, which is composed of 16S rRNA and 21 proteins, harbors the decoding center where mRNA base triplets are read by tRNA as codons. The catalytic center for peptide bond formation is located in the 50S subunit that is composed of 23S and 5S rRNA together with 31 proteins. Because peptide bond formation is

## Introduction

catalyzed exclusively by 23S rRNA in the peptidyl transferase center (PTC), the ribosome is a ribozyme (Ban et al., 2000; Nissen et al., 2000).

tRNA molecules are short non-coding RNA of 73 to 94 nucleotide length that serve as adaptor molecules that bridge the decoding center and PTC on the two ribosomal subunits. Every tRNA contains an anticodon for decoding of mRNA codons and an acceptor end to which an amino acid (aa) can be esterified at the 3'-ribose. The resulting aminoacyl-tRNA (aa-tRNA) can read the codon and deliver the respective aa to the PTC. The L-shaped tertiary structure of tRNA is formed by interaction of its D and T $\Psi$ C loops (Figure 1-2). Aminoacyl-tRNA synthetases (aaRS) recognize identity elements of tRNA molecules involving base composition and shape in order to catalyze aa-specific charging of tRNA molecules with high fidelity (Saks et al., 1994). The ribosome contains three binding sites for tRNA molecules. The binding sites for tRNA molecules are referred to as acceptor site (A site), peptidyl-tRNA binding site (P site), and exit site (E site).

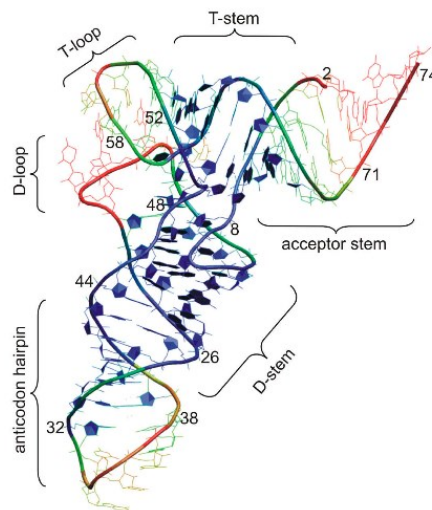


Figure 1-2: Tertiary fold of tRNA. Secondary structure elements of canonical tRNA are indicated. Functional roles and typical post-transcriptional modification of the tRNA sequence to dihydrouracil (D) and T $\Psi$ C (T) to pseudo-uridine ( $\Psi$ ) formed the nomenclature for the secondary structure elements. Nucleotides in the anticodon sequence adapt a flipped out and stacked configuration which is required for the interaction with an mRNA codon. tRNA undergoes conformational changes upon binding to aaRS and substrate selection in the A site. Figure adapted from (Agirrezabala and Valle, 2015).

Protein synthesis on the ribosome can be divided in the four main phases of initiation, elongation, termination, and recycling (Figure 1-3). In the first step during translation initiation, initiation factors (IF) associate with the 30S subunit to form a

## Introduction

30S pre-initiation complex (PIC) to which initiator tRNA fMet-tRNA<sup>fMet</sup> can be recruited (Milon et al., 2012). The mRNA binds to the 30S PIC independently of the recruitment of the IF (Milon et al., 2012). The secondary structures of the mRNA in the ribosome binding region and interactions with the anti-Shine-Dalgarno (anti-SD) sequence in 16S rRNA modulate the recruitment of mRNA to the 30S subunit (Shine and Dalgarno, 1974; Studer and Joseph, 2006). Recognition of the AUG start codon by initiator fMet-tRNA<sup>fMet</sup> results in the formation of the 30S initiation complex (IC) (Milon et al., 2008). Association of the 50S subunit with the 30S IC triggers GTP hydrolysis by IF2, resulting in the dissociation of IF1 and IF2. Following dissociation of IF3, the 70S IC is competent in translation elongation (Milon and Rodnina, 2012).

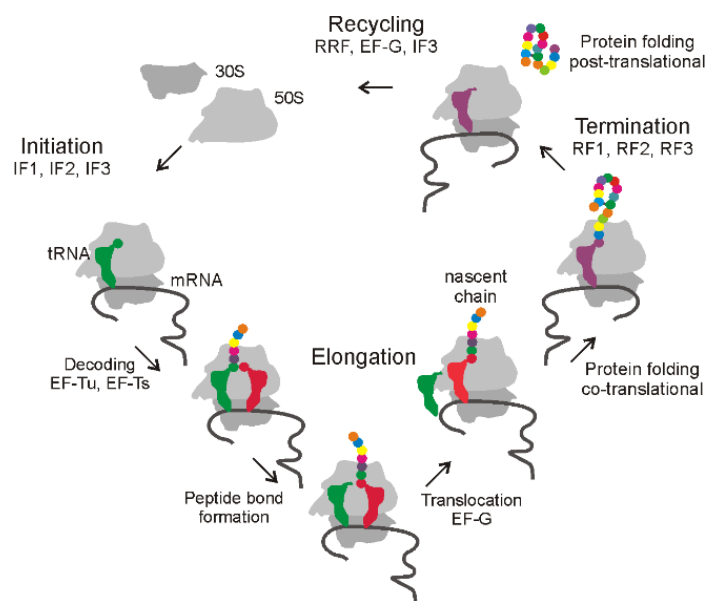


Figure 1-3: Main stages of protein synthesis. During translation initiation, IFs facilitate the assembly of both ribosomal subunits, mRNA and initiator tRNA<sup>fMet</sup> into the 70S IC that is competent to proceed in translation elongation. aa-tRNA are delivered as ternary complex of EF-Tu•GTP•aa-tRNA to the A site. Following peptide bond formation, EF-G catalyzes the translocation of the ribosome on the mRNA by one codon. Upon encounter of stop codon, RFs hydrolyze the peptidyl-tRNA thereby release the nascent chain from the ribosome. After termination of the peptide synthesis, the subunits dissociate with the help of RRF and EF-G in order to reassemble during the following rounds of translation initiation. The nascent chain may sample folded states already during translation elongation.

The cycles of translation elongation are dependent on the elongation factors (EF) EF-Tu (together with EF-Ts) and EF-G. EF-Tu delivers aa-tRNA in a ternary complex of EF-Tu•GTP•aa-tRNA to the A site of the ribosome. GTP hydrolysis by EF-Tu and phosphate release following the codon recognition reduce the affinity of EF-

## Introduction

Tu•GDP for the aa-tRNA and allow aa-tRNA to accommodate in the PTC (Kothe and Rodnina, 2006). Peptidyl transfer leading to elongation of the nascent chain by one residue is accomplished by donation of the peptidyl chain to the A-site tRNA, which leaves a deacylated tRNA in the P site. In order to continue in translation elongation, the ribosome has to move by one codon along the mRNA and bound tRNA. EF-G accelerates the translocation by several orders of magnitude (Rodnina and Wintermeyer, 2011). The binding and GTP hydrolysis by EF-G on the ribosome are thought to bias the energetic landscape of the large-scale conformational fluctuations of the ribosome towards the post-translocation state (Bock et al., 2013; Fischer et al., 2010; Holtkamp et al., 2014b).

Stop codons in the A site are recognized by release factors (RF). The amber UAG and ochre UAA stop codons are recognized by release factor 1 while the opal UGA and ochre UAA codons are recognized by RF2. Stop codons recognition by RF is achieved by tripeptide sequences specific for RF1 and RF2, respectively (Ito et al., 2000). A GGQ motif present in both RF1 and RF2 is crucial for catalysis of peptide release. The Gln residue of the GGQ motif is post-translationally methylated (Nakahigashi et al., 2002). The backbone and side chain of Gln coordinate and optimally position a water molecule for hydrolysis of the peptidyl-tRNA in the P site (Dincbas-Renqvist et al., 2000; Shaw and Green, 2007; Weixlbaumer et al., 2008). Following the release of the nascent chain, GTP hydrolysis by RF3 facilitates the dissociation of RF1/RF2 from the ribosome (Peske et al., 2014). The recycling of ribosomes after termination involves the ribosome recycling factor (RRF), EF-G, and IF3. RRF and EF-G facilitate the dissociation of the ribosomal subunits while IF3 prevents their re-association (Peske et al., 2005) and acts as a bridging factor towards the next round of initiation (Milon and Rodnina, 2012).

### *1.2 Substrate selection in the A site*

The selection of aa-tRNA cognate to the codon in the A site from the pool of different aa-tRNAs is achieved by a two-step kinetic discrimination mechanism (Hopfield, 1974; Pape et al., 1999; Thompson, 1988; Yarus, 1992). A single kinetic proofreading step based on the discrimination of the base pairing energies of the codon anti-codon duplex does not confer sufficient selectivity to maintain an average aa mis-incorporation frequency of  $10^{-4}$  (Kramer and Farabaugh, 2007; Parker, 1989; Rodnina and Wintermeyer, 2001). The overall fidelity is amplified by applying two sequential aa-tRNA selection steps separated by an irreversible GTP hydrolysis reaction: The initial selection step preceding GTP hydrolysis and the proofreading

## Introduction

step following GTP hydrolysis (Bilgin et al., 1992; Thompson, 1988). In addition to the tRNA dissociation rates which depend on the base pairing in the codon-anticodon complex, the rates of the GTPase activation of EF-Tu and the accommodation of aa-tRNA in the peptidyl-transferase center differ for the cognate, near-cognate, and non-cognate aa-tRNAs (Gromadski et al., 2006; Gromadski and Rodnina, 2004; Pape et al., 1999; Rodnina et al., 1996; Wohlgemuth et al., 2011). With respect to the codon in the A site, the tRNAs can be classified as cognate when base pairing in the 1<sup>st</sup> and 2<sup>nd</sup> codon position follows the Watson-Crick geometry, whereas in the 3<sup>rd</sup> codon position wobble decoding is allowed. Near-cognate decoding involves a single nucleotide mismatch at any position. All remaining codon-anticodon interactions are considered as non-cognate. Ternary complexes containing non-cognate aa-tRNA dissociate immediately after initial binding to the ribosomal proteins of the L7/L12 stalk (Diaconu et al., 2005; Gromadski and Rodnina, 2004; Kothe et al., 2004; Rodnina et al., 1994). Both cognate and near-cognate ternary complex proceed at the same rate to the step of codon recognition (Gromadski et al., 2006). Following codon recognition, residues A1492, A1493 and G530 in the decoding center rearrange in order to probe the base-pair independent geometry of the minor groove of codon-anticodon duplex (Ogle et al., 2001). The shape of the minor groove is monitored stringently for the first codon position by A1493 and for the second base-pair by A1492 and G530. The wobble position is only monitored by the coordination of the ribose moiety of the third codon nucleotide by G530 and a Pro residue from ribosomal protein S12. The successful rearrangements in the decoding center initiate domain closure of the 30S subunit (Ogle et al., 2002), distortion of the tRNA bound in the A/T state (Fischer et al., 2015; Stark et al., 1997; Valle et al., 2002), and activation of the GTPase of EF-Tu (Gromadski et al., 2006; Gromadski and Rodnina, 2004; Pape et al., 1999; Rodnina et al., 1996; Voorhees et al., 2010).



# Introduction

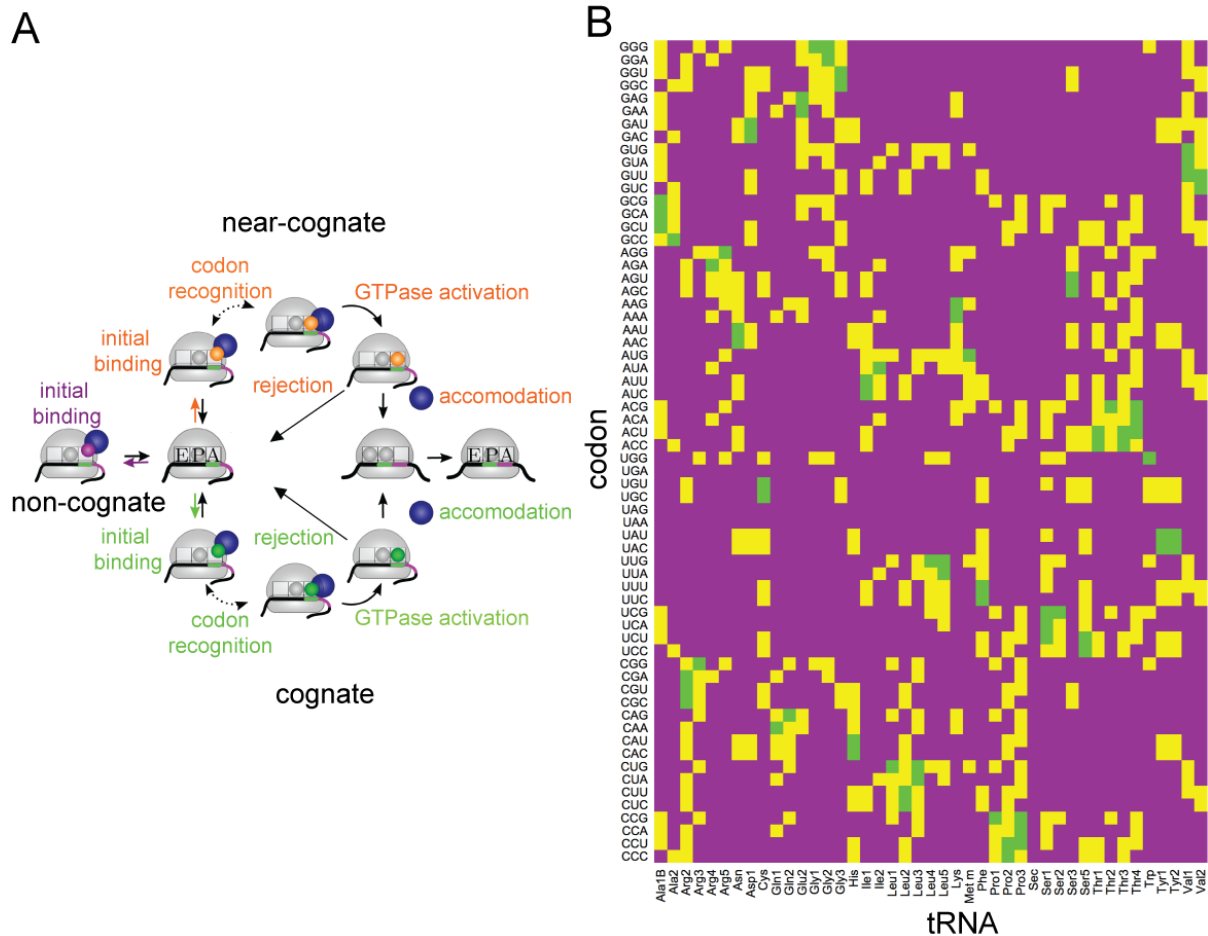


Figure 1-4: Three potential modes of interaction of aa-tRNA with a given codon in the A site. **A** Schematic depiction of substrate selection in the A site. Kinetic steps sampled during cognate (green), near-cognate (orange) and non-cognate (purple) decoding are highlighted. **B** Codon-anticodon matrix for all sense codons and the 43 elongator tRNA in *E. coli*. Cognate codon-anticodon interactions are shown in green, near-cognate in yellow, and non-cognate in purple. Figures were modified from (Rudorf et al., 2014).

The collective conformational changes initiated by the recognition of a cognate codon anti-codon duplex are also referred to as induced fit mechanism that greatly enhances the rate of GTPase activation (Rodnina and Wintermeyer, 2001; Wohlgemuth et al., 2011). Additionally, near-cognate codon-anti codon duplexes dissociate at a much higher rate compared to cognate duplexes leading to a high discrimination against near-cognate ternary complexes at the step of GTPase activation of EF-Tu (Gromadski et al., 2006). Following GTP hydrolysis and dissociation of the aa-tRNA from EF-Tu•GDP, the partitioning between accommodation of aa-tRNA in the PTC and dissociation from the A site ensures proofreading, e.g. rejection of those near-cognate aa-tRNA that passed the initial selection screen. RNA nucleotides in the A-site loop of the PTC are proposed to form

## Introduction

a gate that can delay the accommodation of the aa-tRNA in the A/A state (Kim and Green, 1999). Accommodation in the PTC monitors the conformation of the aa-tRNA that is dependent on the integrity of the codon anti-codon duplex in the decoding center (Gromadski et al., 2006; Jenner et al., 2010a; Whitford et al., 2010; Wohlgemuth et al., 2010).

The rate of translation elongation at a given codon in the A site is proportional to the availability of the aa-tRNA cognate to that codon. In addition, the ratio between the cognate and near-cognate aa-tRNA is crucially important, because near-cognate codon anti-codon duplexes do not dissociate immediately after the step of codon recognition. Addition of increasing concentrations of near-cognate ternary complexes for the A site reduced the rate of GTPase activation in the cognate ternary complexes due to competition (Figure 1-4A) (Wohlgemuth et al., 2010). Because the overall rate of translation is largely limited by the decoding step, a high cellular abundance of competing near-cognate aa-tRNA might significantly reduce the rate of translation elongation at a given codon. The decoding of 61 sense codons by 43 tRNA species with different cellular abundance in *Escherichia coli* (*E.coli*) opens a vast combinatorial space for the prediction of codon-specific elongation rates (Dong et al., 1996).

### *1.3 Incorporation of non-canonical amino acids*

Treatment of Cys-tRNA<sup>Cys</sup><sub>GCA</sub> with Ranley nickel to produce Ala-tRNA<sup>Cys</sup><sub>GCA</sub> yielded the first experimental evidence that aa-tRNA carrying a non-cognate aa could be incorporated into the nascent chain (Chapeville et al., 1962). Similarly, εN-acetyl-Lys-tRNA<sup>Lys</sup><sub>UUU</sub> sustained translation elongation in *E. coli* and rabbit reticulocyte cell-free translation systems (Johnson et al., 1976). Incorporation of crosslinking probes attached to the ε-amino group of Lys-tRNA<sup>Lys</sup><sub>UUU</sub> demonstrated for the first time the direct interaction of the Signal Recognition Particle (SRP) with a signal sequence for transport across the membrane of the endoplasmic reticulum (Kurzchalia et al., 1986). The development of approaches for chemical aminoacylation of pdCpA dinucleotides and ligation to tRNA lacking two nucleotides at the 3'-end by T4 RNA ligase allowed for charging of tRNA with non-canonical amino acids (nc-aa) without the need to derivatize reactive groups of enzymatically charged aa-tRNA (Baldini et al., 1988; Heckler et al., 1984). The combination of chemical aminoacylation with an amber suppressor tRNA allowed for the site-specific incorporation of nc-aa in the absence of concomitant incorporation of canonical aa (Noren et al., 1989). Following this approach, a large set of nc-aa that were sterically restricted, modified in the

## Introduction

backbone, or carried biophysical probes could be successfully incorporated (Cornish et al., 1994; Ellman et al., 1991; Ellman et al., 1992; Mendel et al., 1993). Ribozyme (Flexizyme)-mediated aminoacylation was developed as alternative to chemical and enzymatic approaches for preparation of nc-aa-tRNA *in vitro* (Goto et al., 2011; Ohuchi et al., 2007). The incorporation of nc-aa-tRNA prepared *in vitro* was extended from cell-free translation systems to mammalian cells by micro-electroporation of charged tRNA (Dougherty and Van Arnam, 2014).

A more general approach for incorporation of nc-aa *in vivo* became accessible by the expansion of the genetic code with orthogonal aaRS/tRNA pairs that do not charge endogenous tRNA and are *vice versa* not substrates for the endogenous aaRS, respectively (Wang et al., 2001). Mutagenesis of the orthogonal TyRS from *Methanocaldococcus jannaschii* allowed for the incorporation of a large set of aromatic nc-aa with a high yield during protein expression in *E. coli* (Neumann et al., 2008a; Wang and Schultz, 2004). Additional orthogonal aaRS/tRNA pairs were generated based on LeuRS/tRNA<sup>Leu</sup> and PylRS/tRNA<sup>Pyl</sup> of bacterial or archaeal origin (Anderson and Schultz, 2003; Chen et al., 2009; Mukai et al., 2008; Neumann et al., 2008b; Neumann et al., 2010a). Incorporation of nc-aa by orthogonal aaRS/tRNA pairs was extended to *Saccharomyces cerevisiae* (*S. cerevisiae*) (Chin et al., 2003; Hancock et al., 2010; Wu et al., 2004) and mammalian cells (Chen et al., 2009; Liu et al., 2007; Sakamoto et al., 2002).

Genetic code expansion *in vivo* has primarily focused on directed evolution and optimization of orthogonal aaRS/tRNA pairs (Liu and Schultz, 2010; Neumann, 2012). Only a small number of studies aimed at the characterization and optimization of interactions of nc-aa-tRNA with additional partners of the translational machinery. The affinity of EF-Tu for nc-aa-tRNA was optimized in order to enhance incorporation of nc-aa (Guo et al., 2009). Mutant ribosomes were selected that increased the efficiency of incorporation by amber suppression (Wang et al., 2007) or quadruplet codon suppression (Neumann et al., 2010b). Potential differences in performance between nc-aa-tRNA and canonical aa-tRNA during the steps of decoding, peptide bond formation and translocation on the ribosome remain largely unexplored.

## Introduction

### *1.4 Codon usage bias*

The genetic code is universal with only few exceptions. Many aa are encoded by several different codons thereby rendering the genetic code degenerate. Although the association of codons with their specific aa is conserved, the usage frequency of codons within a degenerate codon family is highly organism-specific (Grantham et al., 1980). Similarly, the number and type of tRNA genes required for decoding of specific codons are not conserved (Ikemura, 1985). The base composition of a genome was proposed to be shaped by the evolutionary mechanisms of mutation and selection (Hershberg and Petrov, 2008). According to the mutational hypothesis, the codon usage bias might be shaped by the context specificity of both the rate at which point mutations accumulate and the repair efficiency at these sites (Kimura, 1980). The genomic GC-content is thought to arise from mutational mechanisms (Chen et al., 2004). The GC-content of bacterial genomes can be predicted with high accuracy based on analysis of non-coding sequences outside of open reading frames (ORF). Codon usage bias conferred by mutational mechanisms is believed to be evolutionary neutral. On the other hand, differences in the codon usage among ORFs within the same organism are attributed to evolutionary selection based on cellular fitness (Bulmer, 1991; Ikemura, 1981; Sharp and Li, 1987). Highly expressed genes tend to contain codons used at a higher frequency (Folley and Yarus, 1989; Ikemura, 1985). Experimental studies established the dependency of the translation elongation rate on the usage frequency of codons (Sorensen et al., 1989; Varenne et al., 1984). As an example, the small set of genes coding for components of the translational machinery, which demands the major metabolic investment of the cell, displays the highest adaption to the codon usage in many genomes (Andersson and Kurland, 1990). For bacteria, the maximal growth rate of organisms could be correlated with a reduced diversity of tRNA genes and a more pronounced codon usage bias (Higgs and Ran, 2008; Rocha, 2004; Sharp et al., 2005). In this sense, the bacterial lifestyle formed a base for selection of the codon usage. In contrast to unicellular organisms, mammalian gene expression is tissue-specific and shows a weak correlation with codon adaption indicating different selective pressures shaping the codon usage bias (Comeron, 2004; Urrutia and Hurst, 2003). Although the correlation of protein expression levels with codon usage could be enhanced by the consideration of extended decoding abilities of tRNA conferred by post-transcriptional modification of the anticodon sequence (Novoa et al., 2012). Several lines of evidence favor the view that the efficiency of translation elongation was the major selective pressure for evolution of the codon usage bias. The translation of mRNA sequences with high

## Introduction

codon adaption holds the potential to enhance the yield of protein per mRNA transcript. On the other hand, expression of genes with low codon adaption are thought to sequester ribosomes from translating better adapted mRNA sequences and thereby decreasing the global cellular fitness (Hershberg and Petrov, 2008). On the other hand, the rate of translation initiation was found to be limiting for translation elongation, rendering the selection based on translation efficiency less straightforward (Andersson and Kurland, 1990; Bulmer, 1991; Shah et al., 2013). Alternatively, codon adaption was proposed not only as a strategy for improving translational efficiency but also for ensuring the accuracy of processes associated with translation. The reduction of the cost of erroneous protein synthesis may provide a selective pressure affecting organisms from bacteria to human (Drummond and Wilke, 2008). Erroneous protein synthesis was proposed to encompass aa misincorporation during protein synthesis as well as misfolding of correctly synthesized polypeptides. Production of inactive proteins would increase the cost of protein synthesis for a given gene while the production of misfolded, potentially cytotoxic proteins is thought to affect the cellular fitness on a global level. Similarly, the selective pressure for reduction of erroneous synthesis would scale with increasing expression levels. Experimental evidence in favor of selection for translational accuracy was provided by studies revealing that the usage of synonymous codons was more restricted at sites that showed evolutionary conservation at the level of aa sequences (Akashi, 1994; Zhou et al., 2009). Additionally, the codon adaption was found to scale with the length of genes consistent with a higher selective pressure for accurate translation (Stoletzki and Eyre-Walker, 2007). On a cellular level both efficiency and accuracy of translation elongation have to be balanced (Johansson et al., 2012; Wohlgemuth et al., 2011). It is likely that the codon usage might reflect a similar trade-off between selection for efficiency and accuracy.

By switching the focus from the genomic level to translation of specific mRNA sequences, several interesting insights into the non-random use of synonymous codons emerged. Expression of a library of green fluorescent protein (GFP) construct carrying synonymous mutations in *E. coli* indicated that the most significant change in expression level was associated with the stability of mRNA structures upstream of the coding sequence in the translation initiation region, while the codon usage in the coding sequence was of minor influence (Kudla et al., 2009). Sequences coding for the N-terminal aa were found to be enriched in rare codons (Gu et al., 2010; Tuller et al., 2010b). Systematic variation of N-terminal codons revealed that rare codons

## Introduction

increased the expression level by lowering the stability of mRNA structures (Goodman et al., 2013). Alternatively, the occurrence of rare codons at the N-terminus was proposed to act as a “ramp” of elongation rates that leads to optimal spacing of ribosomes on the same mRNA (Tuller et al., 2010a). A co-occurrence bias of certain codons in close proximity within the same mRNA transcript was proposed to be associated with more rapid recharging of tRNA by ribosome-associated multi-aaRS complexes compared to decoding with aa-tRNA that need to diffuse from the cytoplasm to the ribosome (Cannarozzi et al., 2010; Shao et al., 2012). On the other hand, certain pairs of adjacent codons are under negative selection pressure and statistically underrepresented (Coleman et al., 2008; Tats et al., 2008). Ribosomal profiling studies could not establish a correlation between codon usage and local translation elongation of mRNA sequences (Ingolia, 2014). The ribosomal profiling technique allows the visualization of ribosome density along mRNA sequences during translation at steady-state (Ingolia et al., 2009). The protocol involves deep-sequencing of RNA fragments that are protected from nuclease digestion by the ribosome while RNA sequences located outside the ribosome are sensitive to degradation. The read count of protected fragments is normalized to the read count obtained from deep-sequencing of ribosome-free total RNA. Higher ribosome density could only be associated with the location of Pro residues (Artieri and Fraser, 2014; Gardin et al., 2014). In addition, the non-uniform rate of translation initiation among different mRNA transcripts renders the comparison of steady-state ribosome density difficult (Balakrishnan et al., 2014; Quax et al., 2013; Woolstenhulme et al., 2015).

Non-uniform usage of synonymous codons may form a second layer of genome information in addition to the sequence of aa in proteins. This additional layer of genetically encoded information was proposed to be involved in optimization and regulation of biological processes on several levels. Under conditions where translation initiation is not limiting, a similar codon usage allows to equalize the expression level of proteins that are functionally interacting (Fraser et al., 2004; Lithwick and Margalit, 2005). Proteins with different codon adaption are expressed at different stages of the cell cycle coupled with cell cycle-specific expression levels of aaRS (Frenkel-Morgenstern et al., 2012). Similarly, a low codon adaption is a requirement for proper expression and activity of the protein FRQ essential for the circadian clock function in *Neurospora* (Zhou et al., 2013). Codon usage was also linked to the induction of stress-response pathways. The gain or loss of post-transcriptional modification of tRNA anticodons upon stress induction was shown to increase the expression level of genes with non-optimal codon usage (Begley et al.,

## Introduction

2007; Chan et al., 2010). The interplay between codon usage and cellular protein folding will be introduced in next section.

### *1.5 Co-translation protein folding*

Anfinsen proposed that the aa sequence contains all information required for a protein to regain its native state after complete denaturation (Anfinsen, 1973). While many proteins were shown to refold rapidly from the denatured state on the millisecond time-scale, a large number of proteins could not be refolded even at very dilute concentrations or refolding required several hours to completion (Jaenicke, 1991). The question was raised how this fraction of the proteome can fold on a biological timescale in the crowded environment of the cell. Molecular chaperones were proposed to reduce the accumulation of unfolded proteins by binding to the unfolded state and shielding it from aggregation with cellular components (Ellis, 1996; Naylor and Hartl, 2001). The presence of chaperones enhanced the yield of folded protein both *in vitro* and *in vivo* (Makino et al., 1997; Mares et al., 2011; Nimmegern and Hartl, 1993). On the other hand, even in presence of chaperons the refolding *in vitro* of especially multi-domain proteins is less efficient than observed for folding *in vivo* (Frydman et al., 1999; Frydman et al., 1994; Kolb et al., 1994). Chaperons are thought to affect the partitioning between aggregated and properly folded states but not to alter the folding pathway of the protein (reviewed in (Bukau et al., 2006; Frydman, 2001)). Co-translational folding on the ribosome has been shown for an extended set of proteins and has the potential to direct protein folding along alternative and potentially more productive pathways than refolding *in vitro* (Fedorov and Baldwin, 1997; Komar, 2009).

The energetic penalty for solvation of the peptide backbone and aa side chains constitutes the major thermodynamic driving force for protein folding by formation of a compact, hydrophobic core shielded from the aqueous environment (Daggett and Fersht, 2003). From an initial collapse ensemble secondary and tertiary structures may arise. The same driving forces are responsible for protein folding upon renaturation in the test tube as well as for co-translational protein folding on the ribosome (Fedorov and Baldwin, 1997). The vectorial emergence of the polypeptide from the ribosomal exit tunnel during ongoing translation constitutes the main difference for protein folding on the ribosome. Translation elongation is determining the length of the polypeptide sequence and thereby the folded states that can be sampled at a given time. Folded states of nascent chain have been detected during ongoing synthesis with various techniques including conformational antibodies,

## Introduction

ligand binding, and gain of catalytic activity (Fedorov and Baldwin, 1999; Hamlin and Zabin, 1972; Kiho and Rich, 1964; Komar et al., 1997). The presence of the complete aa sequence of the protein outside of the ribosome is in many cases not required for the nascent chain to attain its native folded state. Folding processes take place in the microsecond to millisecond time range (McCammon, 1996). Translation elongation with a rate of 15 to 20 aa s<sup>-1</sup> (~ 50 ms per aa) is believed to be the rate-limiting step for a large spectra of co-translational folding events. The different time scales of folding reactions and translational elongation might impose a sequential order on the folding pathway. A more restricted co-translational folding pathway might allow for more efficient and faster folding compared to the vast combinatorial space of folding process that might be sampled during refolding of a full-length protein (Morrissey et al., 2004).

Alteration of the rate-limiting step of translation elongation was found to modulate the efficiency of co-translational folding. An increase in global translation rate during expression of eukaryotic multi-domain proteins in a bacterial cell-free translation system was correlated with the higher accumulation of misfolded proteins (Netzer and Hartl, 1997). Usage of synonymous codons with different elongation rates was proposed to locally fine-tune the timing for the formation of folded states (Komar, 2009). Stretches of infrequent codons were observed to mark the topological borders of multi-domain proteins (Purvis et al., 1987). Slow translation elongation at stretches of codons with low adaption to the codon usage was suggested to enhance domain folding and interactions of domains during ongoing synthesis on the ribosome (Thanaraj and Argos, 1996b). Several bioinformatic studies provide evidence that even secondary structure elements of proteins might be encoded in the codon usage profile along the mRNA sequence (Lopez and Pazos, 2015; Pechmann and Frydman, 2013; Saunders and Deane, 2010; Thanaraj and Argos, 1996a). Substitution of codons with low adaption to the codon usage for frequently used ones decreased the enzymatic activity of chloramphenicol acetyltransferase expressed in *E. coli* and the metabolic enzyme TRP3 expressed in *S. cerevisiae* (Crombie et al., 1994; Komar et al., 1999; Ramachandiran et al., 2002). A single synonymous codon substitution resulted in altered drug and inhibitor interactions with the multi-drug resistance gene MDR-1 coding for P-glycoprotein. The differences in substrate binding were suggested to arise during co-translational folding leading to altered final conformations (Kimchi-Sarfaty et al., 2007). Substitution of clusters of infrequent codons increased the sensitivity of protein Suf1 from *E. coli* to proteolysis (Zhang et al., 2009). Modulation of translation elongation rates interfered with co-translational



## *Introduction*

domain association in the cystic fibrosis transmembrane conductance regulator (Kim et al., 2015). In summary, a growing list of evidence suggests that the local rate of translation elongation might serve as a kinetic guide for co-translational folding. Investigation of the link between translation elongation and co-translational folding in real-time holds the promise to further unravel the coupling of the two processes.

## Introduction

### 1.6 Aim of the thesis

The general aim of this thesis was to experimentally link the timing of translation elongation and folding of the nascent chain on the ribosome. The following questions were addressed:

- To develop a stochastic model for translation elongation which would take into account the competition of tRNA species for the A site and sequence-specific events of translational pausing in order to allow for the accurate description of translational elongation *in vitro*.
- To establish the incorporation of FRET pairs of non-canonical amino acids during the synthesis of the nascent chain and apply this technique to monitoring the kinetics of protein folding that might be influenced by the codon-specific rate of translation elongation.

## 2 Results

### 2.1 Modeling of translation elongation rates

A first, key question for any mathematical modeling of protein synthesis is to which extent the rate constants obtained for translation elongation *in vitro* can be used to describe translation elongation in *E. coli* cells. Rate constants for codon recognition, codon dissociation and GTPase activation are highly dependent on the chosen buffer conditions (Gromadski and Rodnina, 2004; Johansson et al., 2012). Buffers that allowed the measurement of decoding at high fidelity were designed to mimic the cellular concentrations of magnesium and polyamines (Gromadski and Rodnina, 2004; Johansson et al., 2012; Neidhardt, 1987). In the cell, a trade-off between fidelity and efficiency of decoding has to be maintained in order to sustain high growth rates (Johansson et al., 2012; Wohlgemuth et al., 2011). An informative example of the trade-off between speed and accuracy of protein synthesis can be based on *E. coli* strains carrying mutations in ribosomal protein S12 that lead to increased accuracy at the cost of reduced growth rates (Bohman et al., 1984; Gorini, 1971; Sharma et al., 2007). Co-expression of EF-Tu mutants that cause increased aa mis-incorporation could restore the growth rates of the S12 mutant strains (Tubulekas and Hughes, 1993a, b).

An important validation criterion for mathematical models of translation is that they should faithfully reproduce the fundamental parameters of protein synthesis in the cell, e.g. the rate and fidelity of translation. However, it is known that the rate of protein synthesis changes with the cellular growth rate. In 1958, a seminal discovery was made that the macromolecular composition of *Salmonella typhimurium* depends exponentially on the growth rate and does not depend on the absolute supply of nutrition components (Schaechter, 2015; Schaechter et al., 1958). The exponential dependency varied for different types of macromolecules, e.g. DNA, RNA, and proteins. Later studies established the average amount of DNA per cell (Helmstetter, 1968), the ratio of total protein to DNA (Donachie, 1968), and the ratio of RNA to DNA (Maaloe, 1969; Schleif, 1967) parameterized as functions of the bacterial growth rates (Bremer and Dennis, 2008). These constants allowed the calculation of the fraction of actively translating ribosomes and the rate of peptide elongation as a function of the growth rate for *E. coli* (Maaloe, 1969; Schleif, 1967). The calculated rate of peptide elongation increased from 12 aa s<sup>-1</sup> at 0.6 doublings h<sup>-1</sup> (dbl) h<sup>-1</sup> to 21 aa s<sup>-1</sup> at 2.5 dbl h<sup>-1</sup>. Pulse-labeling experiments with radioactive aa

## Results

yielded consistent elongation rates ranging from 13 to 20 aa s<sup>-1</sup> in the range from 0.6 to 2.5 dbl h<sup>-1</sup> (Bremer and Dennis, 2008). Additionally, the abundance of rRNA and elongator tRNA species from *E. coli* were collectively quantified at different growth rates (Dong et al., 1996). On average, the concentration of tRNA molecules increased only by 80% from 0.4 to 2.5 dbl h<sup>-1</sup>. Major isoacceptor tRNA increased up to two-fold, whereas minor isoacceptors did not increase in abundance. On the other hand, the ribosome concentrations increased 2.5-fold from 0.4 to 2.5 dbl h<sup>-1</sup>. In summary, the parameterization in terms of growth rate provides a quantitative frame-work for the description of translation elongation *in vivo*.

In this section, we describe of a stochastic model for translation elongation in *E. coli* that can be used to calculate translation elongation rates of any given mRNA sequence *in vitro* as well as *in vivo* by considering the complete codon-anticodon matrix shown in Figure 1-4 B (Rudorf et al., 2014). In collaboration with Sophia Rudorf and Reinhard Lipowsky (Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces), we constructed a stochastic Markov model for the translation elongation cycle (Figure 2-1 B). The Markov model incorporated the elementary steps of decoding indicated in Figure 2-1 A (Pape et al., 1999; Wohlgemuth et al., 2011). The nomenclature of transition rates used for the Markov model is listed in Table 2-1 side by side with the nomenclature used in Figure 2-1 A, together with the numerical values of the transition rates. A stochastic process has the Markov property if the distribution of a variable depends only on the distribution of the previous variable. For a given codon-anticodon combination, we incorporated a separate description for the cognate, near-cognate and non-cognate interaction. Initial binding of EF-Tu•GTP•aa-tRNA ternary complexes to the L7/L12 stalk is independent of the tRNA species and of the codon in the A site (Gromadski et al., 2006; Kothe et al., 2004; Rodnina et al., 1996). Therefore, we applied same rates of initial binding and dissociation  $\kappa_{\text{on}}$  and  $\omega_{\text{off}}$  for all ternary complexes. For the non-cognate case, the model contained only a transition from state 0 to state 11. For the cognate and near-cognate cases, states 1 and 6 could further engage in the step of codon recognition with the transition rate  $\omega_{\text{rec}}$ . The rate of codon recognition was shown to be insensitive to a single mismatch in the codon-anticodon duplex (Gromadski et al., 2006; Rodnina et al., 1996). Importantly, near-cognate codon-anticodon duplexes were found to dissociate three orders of magnitude faster than cognate duplexes (Gromadski et al., 2006). Therefore, we used different transition rates for step of codon dissociation in the cognate  $\omega_{21}$  and near-cognate case  $\omega_{76}$  and a common rate for codon recognition  $\omega_{\text{rec}}$ .

## Results

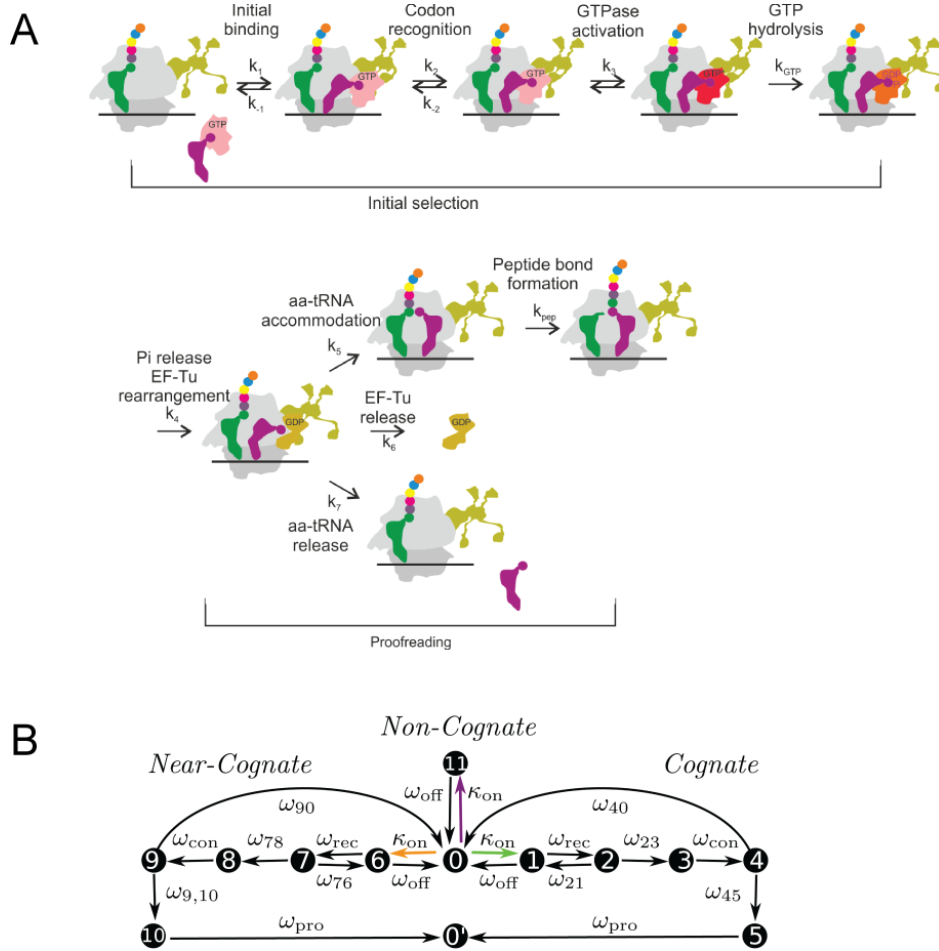


Figure 2-1: Elementary steps of decoding in the A site from the initial binding of the ternary complex until peptide bond formation. **A.** Elementary steps until the irreversible step of GTP hydrolysis contribute to the initial selection. The partitioning between the accommodation and release of the aa-tRNA forms the proofreading step. Measured rate constants are denoted with subscripts. **B** Markov model for translation elongation rates based on *in vitro* rate constants. The model consists of a separate branch for the description of cognate (steps 0 to 5), near-cognate (steps 6 to 10), and non-cognate decoding (step 11). Following peptide bond formation (step 5 resp. 10), the translocation of the ribosome to step 0' takes place with rate  $\omega_{\text{pro}}$ . B was reprinted from (Rudorf et al., 2014).

The cognate codon-anticodon duplex enhances the rate of steps leading to GTPase activation ( $\omega_{23}$  resp.  $\omega_{78}$ ), rendering it 200-fold faster than the near-cognate duplex (Gromadski et al., 2006; Gromadski and Rodnina, 2004; Rodnina and Wintermeyer, 1995). The rate of conformational changes of EF-Tu after GTP hydrolysis and phosphate release are limited by the steps preceding GTPase activation (Kothe and Rodnina, 2006). Therefore, we used a lowest estimate for the rate of conformational changes  $\omega_{\text{con}}$  that was sufficiently high in order not to limit later events.

## Results

Table 2-1: Summary of rate constants used for modeling of elongation rates *in vitro*. The rate constants are listed for both nomenclatures used in Figure 2-1 A and Figure 2-1 B. For published rates the references are indicated.

<u>Step description</u>	<u>Nomenclature</u>		<u>Value at</u>	
	<u>k</u> <sup>a</sup>	<u>ω</u> <sup>b</sup>	<u>20°C</u>	<u>37°C</u>
initial binding TC	k <sub>1</sub>	κ <sub>on</sub>	140 ± 20 μM <sup>-1</sup> s <sup>-1</sup> <sup>c</sup>	175 ± 25 μM <sup>-1</sup> s <sup>-1</sup> <sup>d</sup>
dissociation TC	k <sub>-1</sub>	ω <sub>off</sub>	85 ± 25 s <sup>-1</sup> <sup>c</sup>	700 ± 270 s <sup>-1</sup> <sup>e</sup>
codon recognition	k <sub>2</sub>	ω <sub>rec</sub>	180 ± 30 s <sup>-1</sup> <sup>f</sup>	1500 ± 450 s <sup>-1</sup> <sup>e</sup>
codon dissociation, co	k <sub>-2,co</sub>	ω <sub>21</sub>	0.2 ± 0.03 s <sup>-1</sup> <sup>f</sup>	2 ± 0.6 s <sup>-1</sup> <sup>e</sup>
codon dissociation, nr	k <sub>-2,nr</sub>	ω <sub>76</sub>	140 ± 20 s <sup>-1</sup> <sup>f</sup>	1100 ± 30 s <sup>-1</sup> <sup>e</sup>
GTPase activation, co	k <sub>3,c</sub>	ω <sub>23</sub>	190 ± 30 s <sup>-1</sup> <sup>f</sup>	1500 ± 450 s <sup>-1</sup> <sup>e</sup>
GTPase activation, nr	k <sub>3,nr</sub>	ω <sub>78</sub>	0.6 ± 0.1 s <sup>-1</sup> <sup>f</sup>	7 ± 2 s <sup>-1</sup> <sup>e</sup>
phosphate release	k <sub>4</sub>	ω <sub>con</sub>	50 s <sup>-1</sup> <sup>g</sup>	450 s <sup>-1</sup> <sup>g</sup>
accommodation, co	k <sub>5,co</sub>	ω <sub>45</sub>	22 ± 4 s <sup>-1</sup> <sup>f</sup>	200 ± 40 s <sup>-1</sup> <sup>h</sup>
accommodation, nr	k <sub>5,nr</sub>	ω <sub>9,10</sub>	0.060 ± 0.006 s <sup>-1</sup> <sup>g</sup>	0.26 ± 0.04 s <sup>-1</sup> <sup>h</sup>
aa-tRNA release, co	k <sub>7,co</sub>	ω <sub>40</sub>	0.1 s <sup>-1</sup> <sup>f</sup>	1 s <sup>-1</sup> <sup>e</sup>
aa-tRNA release, nr	k <sub>7,nr</sub>	ω <sub>90</sub>	0.84 ± 0.08 s <sup>-1</sup> <sup>g</sup>	4 ± 0.7 s <sup>-1</sup> <sup>h</sup>
processivity		ω <sub>Pro</sub>	3 ± 1 s <sup>-1</sup> <sup>g</sup>	150 ± 50 s <sup>-1</sup> <sup>g</sup>
average elongation		ω <sub>elo</sub>	0.78 ± 0.02 aa s <sup>-1</sup> <sup>g</sup>	6.9 ± 2.3 aa s <sup>-1</sup> <sup>g</sup>

**a** Nomenclature from (Pape et al., 1999; Wohlgemuth et al., 2011). **b** Nomenclature from (Rudorf et al., 2014) **c** Values from (Gromadski and Rodnina, 2004) **d** Value was extrapolated by using an activation energy of -2.4 kcal mol<sup>-1</sup> (Gromadski and Rodnina, 2004). **e** Values measured in this work. **f** Values from (Gromadski et al., 2006). **g** Values were extrapolated by using an activation energy of -23.4 kcal mol<sup>-1</sup> (Wohlgemuth et al., 2010). **h** Values from (Wohlgemuth et al., 2010).

After dissociation from EF-Tu•GDP, the aa-tRNA can be rejected or move into the PTC on the 50S subunit. For near-cognate codon-anticodon duplexes, the rate of accommodation ω<sub>9,10</sub> is decreased compared to ω<sub>45</sub> for the cognate duplex. The rate of rejection is in a comparable range for both cognate ω<sub>90</sub> and near-cognate duplexes ω<sub>40</sub> (Gromadski et al., 2006). We did not include a separate kinetic step for peptide bond formation because the rate of peptide bond formation is limited by the preceding step of accommodation with the exception of peptidyl transfer involving Pro and Gly (Doerfel et al., 2013; Johansson et al., 2011; Wohlgemuth et al., 2008; Wohlgemuth et al., 2010). In order to continue to the next round of substrate selection in the A site, we have introduced a processing rate ω<sub>pro</sub> for the ribosome to translocate to the next codon. The value of ω<sub>pro</sub> was calculated by equation 1 where ω<sub>elo</sub> is the average elongation rate, t<sub>c</sub> the codon specific elongation time and ρ<sub>c</sub> the codon usage frequency.

## Results

$$\omega_{\text{pro}} = \left[ \frac{1}{\omega_{\text{elo}}} - \sum_c \rho_c t_c \right]^{-1} \quad \text{Equation 1}$$

### 2.1.1 Parameters for modeling of elongation rates *in vitro*

An almost complete parameter set has been already published for the transition rates in HiFi-buffer at 20°C (Table 2-1). The buffer composition and especially the concentration of free Mg<sup>2+</sup> has a profound influence on the rates during initial selection of ternary complexes (Gromadski et al., 2006; Wohlgemuth et al., 2011). In our Markov model, we have only included rate constants measured in HiFi-buffer. The values for  $\omega_{\text{rec}}$ , and  $\omega_{21}$  were reported for the codons UUC and UUU both coding for Phe (Gromadski et al., 2006). Rates of GTP hydrolysis of EF-Tu are on average 2.5-fold higher for duplexes involving only Watson-Crick base pairing than duplexes having a wobble base pair in the third position (Gromadski et al., 2006; Thomas et al., 1988). Therefore, we considered an average of both values for  $\omega_{\text{rec}}$ , and  $\omega_{21}$ . Similarly, the value of  $\omega_{76}$  represented an average of several reported near-cognate interactions (Gromadski et al., 2006). Because only a limited set of parameters was measured at 37°C in HiFi buffer, we used extrapolation based on the Arrhenius relation to calculate the missing transition rates. The rate of  $\omega_{45}$  was previously measured both at 20°C and 37°C (Table 2-1), which allowed us to calculate an activation energy of -23.4 kcal mol<sup>-1</sup> for scaling of the individual transition rates from 20°C to 37°C. The value of  $\omega_{9,10}$  was only reported at 37°C (Wohlgemuth et al., 2010). We measured the value of  $\omega_{9,10}$  at 20°C following the same experimental approach (Wohlgemuth et al., 2010). Peptide bond formation was monitored under multiple turnover conditions by mixing ternary complex containing Phe-tRNA<sup>Phe</sup> and purified 70S IC programmed with the near-cognate CUC codon coding for Leu (Figure 2-2). The value of  $\omega_{9,0}$  was calculated by applying a previously reported value for the selectivity of the proofreading step of 15 (Gromadski and Rodnina, 2004).

## Results

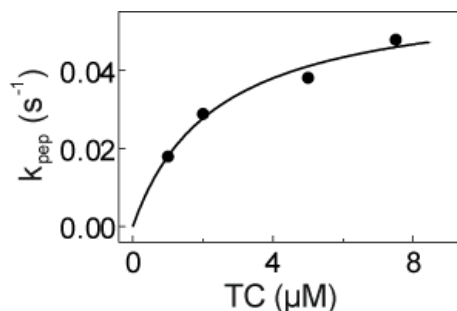


Figure 2-2: Determination of  $\omega_{90}$  and  $\omega_{9,10}$  at 20°C. Michaelis-Menten plot of the initial rate of dipeptide formation of 70S IC initiated with  $f[{}^3\text{H}]\text{-Met-tRNA}^{\text{Met}}$  on mRNA m022CUC (0.14  $\mu\text{M}$ ) and increasing concentrations of  $[{}^{14}\text{C}]\text{Phe-tRNA}^{\text{Phe}}\cdot\text{EF-Tu}\cdot\text{GTP}$ . Fitting to a hyperbolic function yielded  $k_{\text{cat}} = \omega_{9,10}$  of  $0.06 \pm 0.006 \text{ s}^{-1}$  and a  $K_{\text{M}}$  of 2.4  $\mu\text{M}$  for the near-cognate ternary complex.  $\omega_{90}$  was calculated from  $\omega_{90} / (\omega_{9,10} + \omega_{90}) = 1/15$  as reported in (Gromadski and Rodnina, 2004).

In order to constrain the value of  $\omega_{\text{elo}}$  under HiFi buffer conditions, we measured the average rate of peptide synthesis during single-turnover *in vitro* translation of the model protein CspA 1-70 aa (Figure 2-3). The *in vitro* translation was performed in the temperature range between 20 to 37°C and 70S initiation complexes contained fluorescence-labeled initiator tRNA (BodipyFL- $\text{C}_2\text{-Met-tRNA}^{\text{Met}}$ ). A sub-stoichiometric concentration of initiator tRNA with respect to 70S ribosomes limited the translation elongation to a single round. At the endpoint of single-turnover translation, the ribosomes remained bound to the 3' end of the mRNA due to the absence of a stop codon and RFs. Nascent peptides were separated by Tris-tricine PAGE and visualized on a fluorescence scanner. We have noted that at lower temperatures several transient bands accumulated corresponding to nascent peptides with a defined length. In section 2.2, we will focus in more detail on intermediates observed during *in vitro* translation. We have quantified the build-up of full-length CspA over time and calculated the average rate of synthesis  $\omega_{\text{elo}}$  for each temperature (Table 2-2). Analysis of temperature dependence of  $\omega_{\text{elo}}$  by the Arrhenius relation yielded an activation energy  $-22 \pm 3 \text{ kcal mol}^{-1}$  and thereby experimentally validated the extrapolation of the transition rates in Table 2-1 from 20°C to 37°C.



## Results

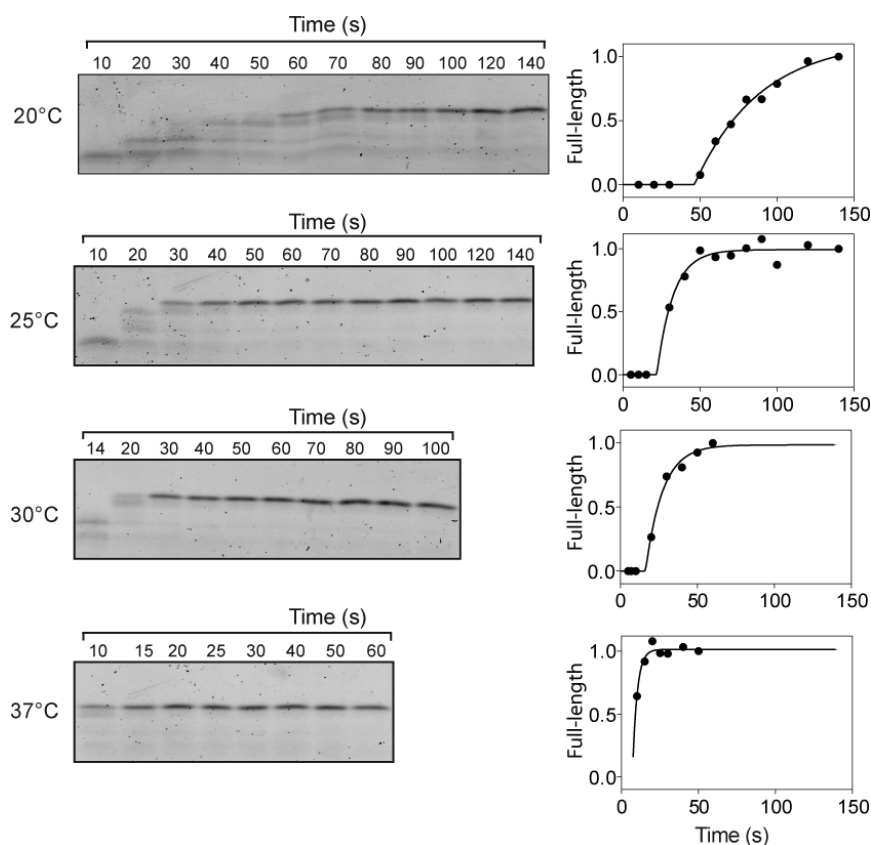


Figure 2-3 Single-turnover *in vitro* translation of CspA 1-70 aa at different temperatures. 70S IC•BodipyFL-C<sub>2</sub>-Met-tRNA<sup>Met</sup> (15 nM) were mixed with factor mix comprising ternary complex (40 μM), EF-Tu in total (100 μM), and EF-G (3 μM). Nascent peptides were separated on Tris-tricine PAGE and imaged by fluorescence scanning. Build-up of full-length CspA was quantified densitometrically and fitted to a model in which a delay precedes a single exponential increase. For the reaction at 37°C, no delay could be resolved by manual mixing and a single exponential model was used. Fitted values for the delay time, exponential rate and calculated elongation time  $\omega_{elo}$  are listed in Table 2-2.

Table 2-2: Parameters obtained by fitting the formation of full-length CspA to a model with a delay followed by a single exponential increase.  $\omega_{elo}$  was calculated by dividing the length of CspA (70 aa) by the sum of the delay and the reciprocal of the exponential rate.

Temperature (°C)	delay (s)	rate (s <sup>-1</sup> )	$\omega_{elo}$ (aa s <sup>-1</sup> )
20	46 ± 2	0.023 ± 0.004	0.8 ± 0.2
25	22 ± 3	0.09 ± 0.03	2.1 ± 0.4
30	16 ± 1	0.09 ± 0.02	2.6 ± 0.3
37 <sup>a</sup>	-	0.010 ± 0.003	6.9 ± 2.3

<sup>a</sup> The time course was described by a single exponential model.  $\omega_{elo}$  was calculated by dividing the length of CspA (70 aa) by the inverse of the exponential rate.

The rate constants listed in Table 2-1 were combined with the relative abundance of all tRNA species as measured by (Dong et al., 1996) for the calculation of the codon specific elongation rates shown in Figure 2-4.

## Results

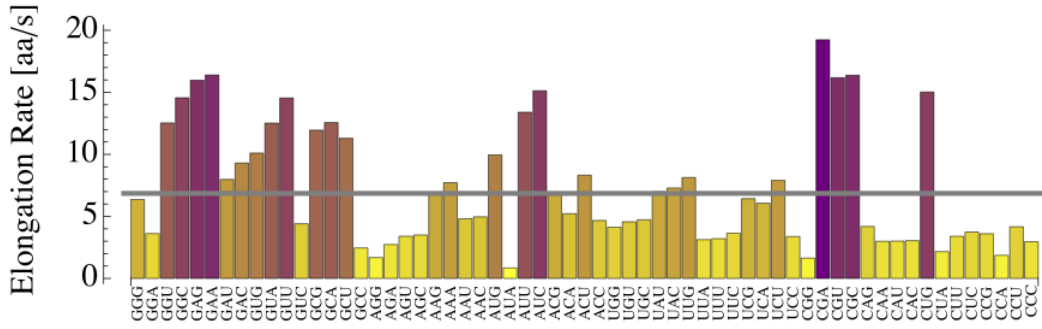


Figure 2-4: Calculated codon specific elongation rates at saturating concentration of all aa-tRNA at 37°C. The experimentally determined average elongation rate  $\omega_{elo}$  is indicated as horizontal line. Calculations were performed by Sophia Rudorf.

### 2.1.2 Modeling of elongation rates *in vivo*

We used the tRNA and ribosome concentrations reported by (Dong et al., 1996) for calculating elongation rates at different growth rates (Table 2-3). In addition, we adjusted the value of  $\kappa_{on}^*$  for molecular crowding *in vivo* with a diffusion coefficient of  $2.57 \times 10^{-12} \text{ m}^2 \text{ s}$  for the ternary complexes as reported in (Fluitt et al., 2007). Sophia Rudorf developed a mathematical approach to simultaneously minimize the distance  $D_{ij}$  for all transition rates  $\omega_{ij}$  obtained *in vitro* and  $\omega_{ij}^*$  *in vivo* (Figure 2-5 A, equation 2).

$$D_{ij} \equiv \ln(\omega_{ij}/\omega_{ij}^*) = \Delta_{ij} \quad \text{Equation 2}$$

By considering equation 3 and 4,  $D_{ij}$  can be regarded as shift of the free energy barrier  $\Delta_{ij}$ , where  $k_b$  is the Boltzmann constant and  $v_{ij}$  and  $\Delta G_{ij}$  are the pre-exponential factor and energy difference for the transition from step  $i$  to  $j$  according to the transition state theory.

$$\omega_{ij} = v_{ij} \exp[-\Delta G_{ij}/k_b T] \quad \text{Equation 3}$$

$$\Delta_{ij} = [\Delta G_{ij}^* - \Delta G_{ij}]/k_b T - \ln(v_{ij}^* - v_{ij}) \quad \text{Equation 4}$$

We did not expect the pre-exponential factors  $v_{ij}$  to change considerably for the *in vivo* rates. Therefore,  $\Delta_{ij}$  is thought to reflect differences in the free energy of the transition states. We obtained small single barrier shifts  $\Delta_{ij}$  for the majority of the transition rates at  $2.5 \text{ dbI h}^{-1}$  (Figure 2-5 B). The single barrier shifts  $\Delta_{ij}$  can be converted to dimensionless scale factor  $\omega_{ij}^*/\omega_{ij}$  by equation 5 (Figure 2-5 C).

## Results

$$\omega_{ij}^*/\omega_{ij} = \exp(-\Delta_{ij}) \quad \text{Equation 5}$$

The largest scale factors in Figure 2-5 C were associated with the dissociation step of near-cognate codon-anticodon duplexes with transition rate  $\omega_{76}$  and the dissociation of ternary complexes  $\omega_{\text{off}}$ . Additionally, the values for rate of codon recognition  $\omega_{\text{rec}}$  and the rate of near-cognate GTPase activation  $\omega_{78}$  changed by a factor of two-fold. The scaled *in-vivo* rates for 0.7, 1.07, 1.6, and 2.5 dbl h<sup>-1</sup> are listed in Table 2-3.

Table 2-3: List of scaled *in vivo* transition rates for different growth rates.

<u><i>in vivo</i> rates</u>	<u>0.7 dbl h<sup>-1</sup></u>	<u>1.07 dbl h<sup>-1</sup></u>	<u>1.6 dbl h<sup>-1</sup></u>	<u>2.5 dbl h<sup>-1</sup></u>
$\kappa_{\text{on}}^*$ (s <sup>-1</sup> μM <sup>-1</sup> )	94	94	94	94
$\omega_{\text{off}}^*$ (s <sup>-1</sup> )	1400	1700	2100	2300
$\omega_{\text{rec}}^*$ (s <sup>-1</sup> )	2100	2500	3000	3300
$\omega_{21}^*$ (s <sup>-1</sup> )	2	2	2	2
$\omega_{23}^*$ (s <sup>-1</sup> )	1600	1600	1700	1700
$\omega_{\text{con}}^*$ (s <sup>-1</sup> )	490	500	530	540
$\omega_{45}^*$ (s <sup>-1</sup> )	270	300	340	350
$\omega_{40}^*$ (s <sup>-1</sup> )	1	1	1	1
$\omega_{76}^*$ (s <sup>-1</sup> )	2700	3100	3900	4300
$\omega_{78}^*$ (s <sup>-1</sup> )	5	5	4	4
$\omega_{9,10}^*$ (s <sup>-1</sup> )	0.27	0.27	0.27	0.27
$\omega_{90}^*$ (s <sup>-1</sup> )	6	6	6	
$\omega_{\text{pro}}^*$ (s <sup>-1</sup> )	190	200	230	230
$\omega_{\text{elo}}^*$ (aa s <sup>-1</sup> ) <sup>a</sup>	15	18	22	23

**a** The values for  $\omega_{\text{elo}}$  were taken from (Bremer and Dennis, 2008).

The adjustment of two transition rates for near-cognate decoding  $\omega_{76}$  and  $\omega_{78}$  increased the efficiency of initial selection  $(\omega_{23}/\omega_{21})(\omega_{76}/\omega_{78})$  at 2.5 dbl h<sup>-1</sup> by a factor of 7.7 in comparison to efficiency calculated from the parameters for *in vitro* translation at 37°C. In addition, a gain in efficiency of proofreading by 2.9-fold could be calculated from the ratio  $(\omega_{45}/\omega_{40})(\omega_{90}/\omega_{9,10})$ . Interestingly, the *in vivo* parameters allowed the calculation of a misincorporation frequency of  $3 \times 10^{-4}$  Lys-tRNA<sup>Lys</sup><sub>UUU</sub> at the near-cognate codon GAA that compared well to the experimental value of  $2 \times 10^{-4}$  determined at the same growth rate of 0.7 dbl h<sup>-1</sup> (Kramer and Farabaugh, 2007).

## Results

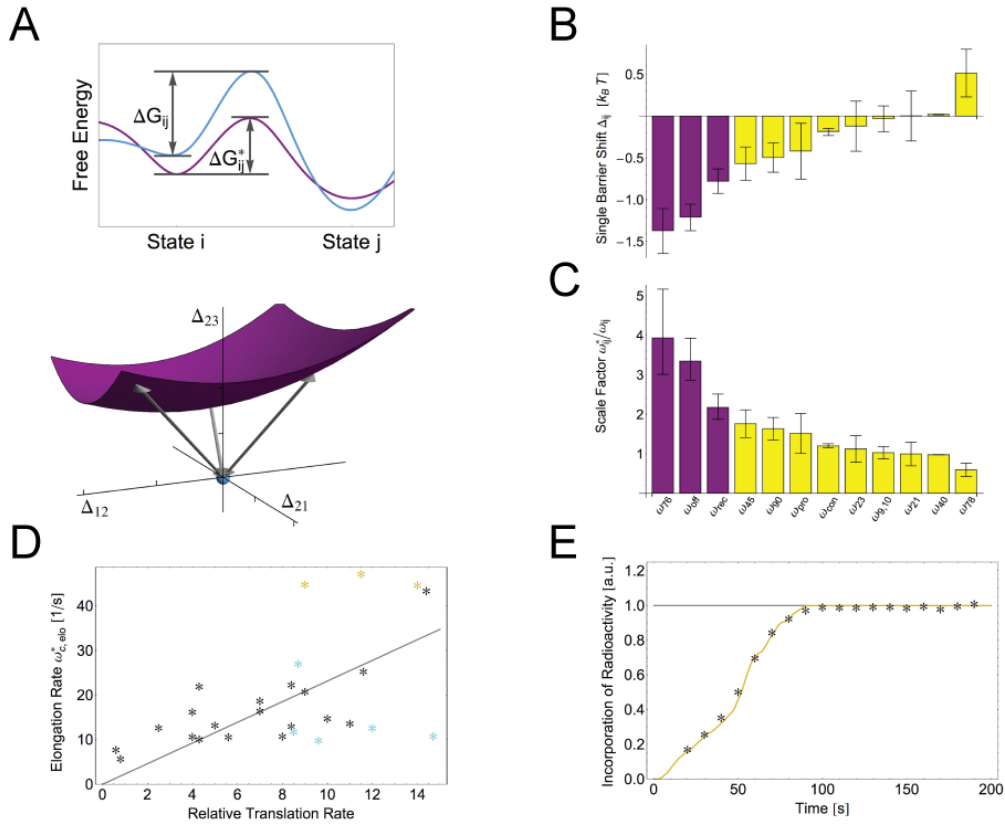


Figure 2-5: Mapping of transition rates *in vivo* and *in vitro*. **A** (Top panel) Free energy barrier  $\Delta G_{ij}^*$  and  $\Delta G_{ij}$  for the transition from state i to state j *in vivo* or *in vitro*. The difference between  $\Delta G_{ij}^*$  and  $\Delta G_{ij}$  corresponds to the logarithmic differences  $\Delta_{ij}$ . (Bottom panel) Subsection of the multi-dimensional hypersurface defined by three coordinates for different barrier shifts and one fixed parameter for translation *in vivo*. The *in vitro* situation is located at the origin. The shortest Euclidean distance from the origin to the surface corresponds to the predicted minimized distance. **B** Minimized energy differences of the transition states  $\Delta_{ij}$  during translation *in vitro* and under cellular conditions at 2.5 dbl h<sup>-1</sup>. **C** The shifts of the transition states in subfigure B can be expressed by a dimensionless scale factor  $\omega_{ij}^*/\omega_{ij}$  with equation 5. Steps with a scaling factor larger than two are shown in purple. **D** Correlation of relative elongation rates *in vivo* from (Curran and Yarus, 1989) with the calculated codon-specific elongation rates at 2.5 dbl h<sup>-1</sup>. The overall Pearson correlation coefficient could be increased from 0.56 to 0.73 when codons decoded by tRNA<sup>Arg2</sup> IGC (gold) and pyrimidine-rich codons (cyan) were excluded. **E** Incorporation of radioactive methionine during pulse-chase labeling of a culture over-expressing LacZ. Symbols correspond to experimental data points measured by (Sorensen and Pedersen, 1991). The simulated incorporation of methionine based on the calculated parameters for translation elongation at 0.7 dbh h<sup>-1</sup> is shown as solid line. The figure and calculations were provided by Sophia Rudorf.

Relative codon-specific elongation rates were determined for 29 sense codons *in vivo* (Curran and Yarus, 1989). The relative elongations rates were calculated from the efficiency of a programmed +1 frame-shift event upon expression of a RF-2/LacZ

## Results

fusion protein. Rapid decoding of codons in the A site decreased the frame-shift efficiency. The calculated codon-specific elongation rates *in vivo* showed a generally good correlation with the reported relative elongation rates (Figure 2-5 D). The correlation could be further enhanced when certain codon families were excluded. The elongation rates of codons decoded by tRNA<sup>Arg2</sup><sub>IGC</sub> were most likely over-estimated in our model, because it assumed a uniform elemental rates of decoding for all cognate tRNA. In contrast, Adenosine 34 in the anticodon of tRNA<sup>Arg2</sup><sub>IGC</sub> is post-transcriptionally deamidated to inosine, allowing the tRNA to read the codons CGA, CGC and CGU, which might occur at a lower rate. Additionally, the relative elongation rates at pyrimidine-rich codons might have been under-estimated by Curran and Yarus due to possible -1 frame-shift events occurring at pyrimidine-rich slippery sites (Figure 2-5 D).

As an additional validation of the values for codon-specific elongation rates *in vivo*, we simulated the time course of incorporation of methionine residues during translation of the *lacZ* mRNA sequence under cellular conditions at 0.7 dbl h<sup>-1</sup> (Figure 2-5 E). The simulated time course of methionine incorporation was compared to the experimental time course obtained by pulse-chase labeling with [<sup>35</sup>S]-methionine of *E. coli* cultures growing at same rate (Sorensen and Pedersen, 1991). The simulated time course of translation of the 649 aa long protein LacZ was in excellent agreement with the experimental work.

## Results

### 2.2 Pausing during translation of natural mRNA

In Section 2.1, we have described a computational model that predicts translation elongation rates during *in vitro* translation. This model considered the concentrations of tRNA species cognate or near-cognate to the codon in the A site as a sole parameter determining the rate of translation elongation. However, the rate of translation elongation can be additionally modulated by mechanisms other than codon usage and availability of isoacceptor tRNA. Stable mRNA secondary structure elements that have to be unwound at the mRNA entry channel are thought to limit translation elongation (Qu et al., 2011; Takyar et al., 2005; Wen et al., 2008). Similarly, the interaction of the anti-SD sequence of the 16S rRNA with complementary SD and SD-like sequences in the mRNA was proposed to lead to translation stalling *in vivo* (Li et al., 2012). Furthermore, at the level of aa, interactions of nascent peptide sequences with specific regions of the ribosomal tunnel were shown to induce unproductive conformations of the PTC (Ito and Chiba, 2013). Similarly, pausing was reported during translation of consecutive Pro residues (Doerfel and Rodnina, 2013; Doerfel et al., 2013; Ude et al., 2013). In this section, we will focus on the characterization of events that result in accumulation of translational intermediates during single-turnover *in vitro* translation of natural mRNA.

#### 2.2.1 Translation intermediates

We observed the accumulation of translation intermediates during the translation of several natural mRNA (Figure 2-6 A to C). Transient as well as more permanent translation pausing was observed at different location along the mRNA. We did not detect any translation intermediates for CspA 1-70 aa (Figure 2-3 A) when translation was performed at 37°C. Strikingly, at lower temperatures translational intermediates started to accumulate. Translational intermediates might accumulate during translation of most natural mRNA and a large variety of decay rates could be associated with the specific translational pauses.

## Results

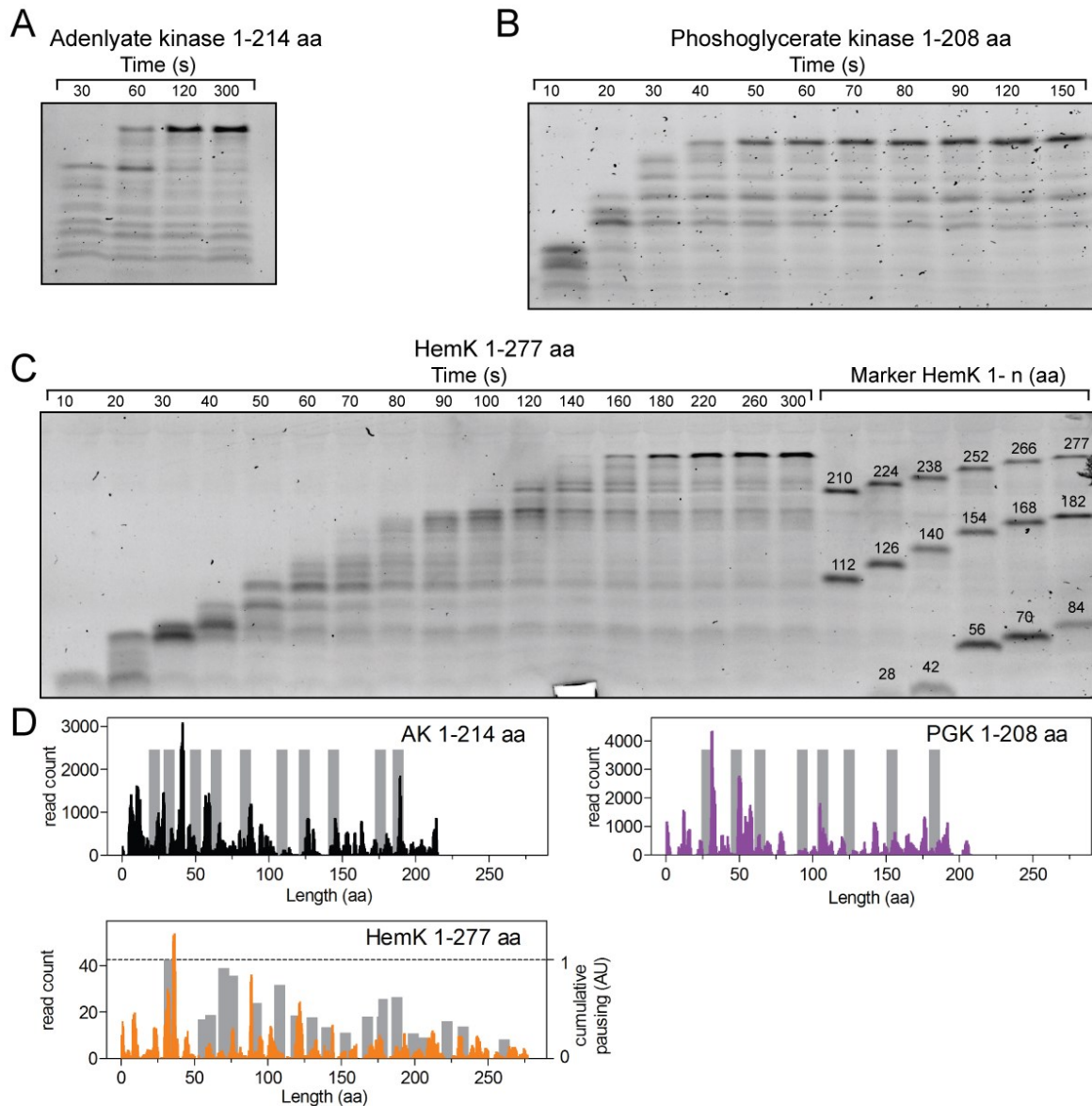


Figure 2-6: Single-turnover translation of natural mRNA **A** of Adenylate kinase 1-214 aa, **B** Phosphoglycerate kinase 1-208 aa and **C** HemK 1-277 aa (from Jörg Mittelstätt, PhD thesis). Single-turnover translation of all natural mRNA in this section was performed under identical conditions: 70S IC formed with BodipyFL- $C_2$ -Met-tRNA<sup>Met</sup> (20 nM) was mixed with saturating concentrations of ternary complexes (40  $\mu$ M) and EF-G (3  $\mu$ M) at 37°C. Marker proteins were prepared by translation of mRNA truncated at the 3'-end as indicated. **D** Reads of ribosome-protected fragment along the sequences of mRNA from *E. coli* strain K12 (Li et al., 2012) that were translated *in vitro* in A to C. The location of pausing sites resolved on gel is indicated in gray within a range of  $\pm 3$  aa. The quantitative characterization of pausing during translation of the HemK 1-277 aa sequence from *E. coli* strain BL21(DE3) in Section 2.2.2 allowed the calculation of a cumulative, steady-state level for the pauses that might be compared to the steady-state ribosome density obtained during ribosome profiling.

## Results

Single-turnover *in vitro* translation has the potential to yield information complementary to the ribosome profiling technique (Ingolia et al., 2009). The most important difference between the two experimental approaches is that only a single elongating ribosome is present per mRNA during single-turnover translation, while ribosome profiling reflects the density of elongating polysomes. The overall ribosome density in the profiling experiment is additionally influenced by the rate of the flanking process of translation initiation and termination. In the profiling experiment, higher occupancy of ribosomes at specific location on the mRNA is inversely proportional to the local translation elongation rate (Dana and Tuller, 2014; Nakahigashi et al., 2014). The mRNA that were used for single-turnover translation in Figure 2-6 A to C were shown to be present at different levels in the cell. In rich-medium, 10 to 100 molecules of the release-factor methyl transferase HemK protein are synthesized in the *E. coli* cell per second (Li et al., 2014). The key enzymes phosphoglycerate (PGK) and adenylate-kinase (AK) that are involved in glycolysis or the balance of the cellular level of AMP, ADP and ATP, are synthesized at a two orders of magnitude higher rate. The cellular mRNA level is also reflected in the read counts of the ribosome profiling data shown for the different mRNA in Figure 2-6 D. The nucleotide sequences of AK and PGK from *E. coli* strain DH5 $\alpha$  are identical to the sequences from strain K12 MG1655 used in the ribosome profiling study (Li et al., 2012). The sequence of HemK from strain BL21(DE3) differed from strain K12 MG1655 at 13 codons (<5% of total sequence) involving five substitutions on the aa level and eight changes to synonymous codons. The complete alignment of the aa and nucleotide sequence of HemK is presented in the Appendix. For the mRNA investigated during translation *in vitro*, ribosome profiling has revealed numerous regions with high ribosome density (Figure 2-6 D). A subset of pauses detected during single-turnover translation clearly overlapped with regions of increased ribosome density in the profiling data. It might be possible that the same molecular event has led to accumulation of ribosomes at specific regions on the mRNA *in vitro* and *in vivo*. On the other hand, several translation intermediates detected during single-turnover translation located to regions with low ribosome density reported by the ribosome profiling approach. Additionally, translation of PGK 1-208 aa *in vitro* showed the very transient accumulation of intermediates within the first 100 aa that might be in contrast to the high steady-state level of ribosome density in the same region provided by ribosome profiling (Figure 2-6 C). It is likely that the translational pausing observed during single-turnover *in vitro* translation might contain additional information on translation elongation that is not covered by ribosome profiling. In-



## Results

depth analysis of the complex pattern of translation intermediates is required for description of basic features of pausing during translation natural mRNA.

### 2.2.2 Quantitative description of pausing during translation of HemK

Jörg Mittelstätt (Department of Physical Biochemistry, Max Planck Institute for Biophysical Chemistry) reported in his PhD thesis for the first time the complex pattern of translational pausing during single-turnover *in vitro* translation of HemK 1-277 aa (Figure 2-6 C). Prompted by the excellent resolution of the translation intermediates and the transient appearance of many translation intermediates compared to the more permanent nature of translational pausing observed for other mRNA (Figure 2-6), we focused on the extraction of quantitative information on the accumulation and decay of the translation intermediates during the time course of synthesis of HemK 1-277 aa.

We quantified the intensity profile along the vertical axis of each individual lane by averaging the intensity across the horizontal axis. Bands observed on gel corresponded to peaks in the intensity profile appearing at a characteristic migration distance. The intensity of individual bands was quantified by numerical integration of the peaks. We normalized the integrated intensity of the peaks to the total intensity of the complete lane after background subtraction. At the endpoint of translation after 5 min, the full-length product accounted for 42% of the total fluorescence intensity (Figure 2-7 A). The average rate of translation elongation leading to the synthesis of full-length HemK 1-277 aa was  $1.42 \pm 0.04$  aa s<sup>-1</sup>. We extrapolated the length of the peptides corresponding to translation intermediates by applying a calibration curve for the electrophoretic mobility based on HemK peptides with defined length (Figure 2-7 B).

## Results

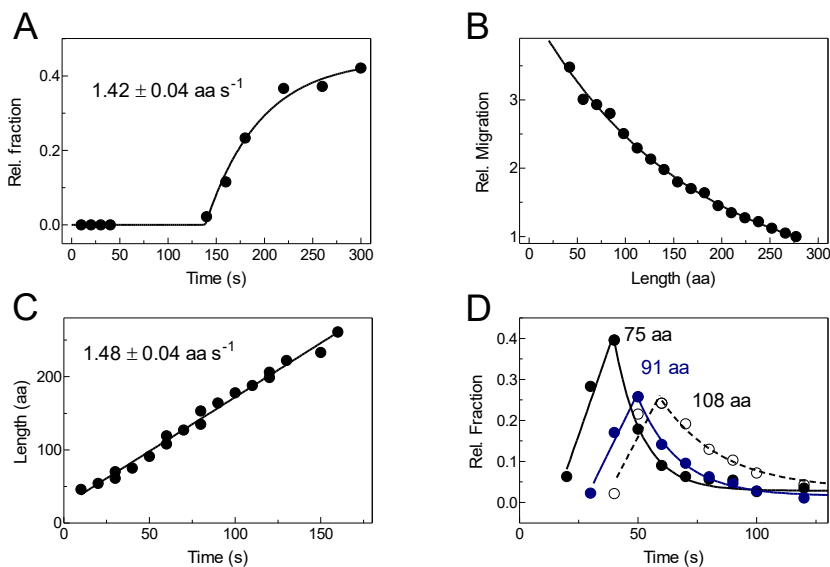
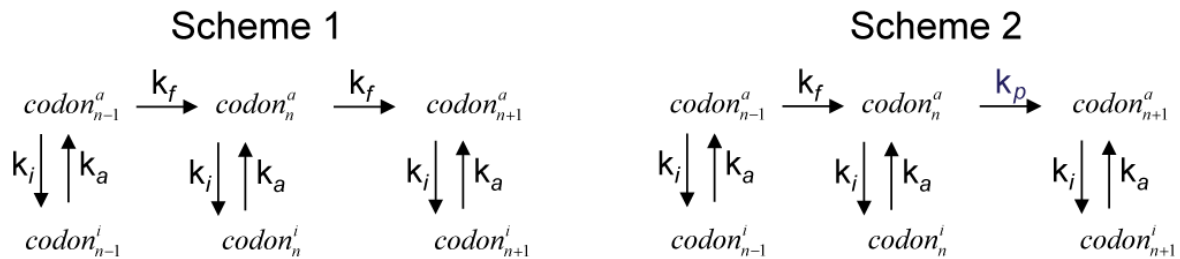


Figure 2-7: Characterization of pausing during translation of HemK 1-277 aa. **A** Accumulation of full-length HemK 1-277 aa. Product formation was evaluated using a model with a delay followed by a single exponential increase. The full-length product started to accumulate after  $138 \pm 3 \text{ s}$ . **B** Electrophoretic migration of the HemK marker peptides with defined length. The migration of shorter marker peptides was normalized to the migration of the HemK 1-277 aa marker peptide. The length of the translation intermediates was estimated from the exponential fit (solid line) of the length dependent migration of the marker peptides. **C** The average translation rate was calculated from the comparison of the length of individual translation intermediates with the time of their maximal accumulation. **D** Quantification of the appearance and decay of individual translation intermediates. The decay of the intermediates was fit to a single exponential model; yielding a rate of  $0.085 \pm 0.006 \text{ s}^{-1}$  for the intermediate at 75 aa,  $0.057 \pm 0.002 \text{ s}^{-1}$  for the intermediate at 91 aa, and  $0.039 \pm 0.003 \text{ s}^{-1}$  for the intermediate at 108 aa. The time resolution did not allow the extraction of rates for the build-up of the intermediates.

We obtained the same value for the average translation rate calculated from the accumulation of the full-length product and from the time of maximal accumulation of a specific intermediate (Figure 2-7 A&C). The good agreement of both values indicates that the formation of the translation intermediates occurred with the same rate as the average translation elongation. The decay rates of the intermediates that were obtained as described in Figure 2-7 D were slow compared to the ongoing synthesis of nascent peptides. This finding raised the question of how the translation intermediates could be sampled sequentially during synthesis of HemK 1-277 aa, while keeping the average rate of synthesis at a high level. Alternatively, in a scenario of stochastic sampling of translation intermediates, a fraction of ribosomes would partition into translation pausing at a given site while the remaining ribosomes

## Results

would continue to elongate. We were not able to describe the sampling of intermediates during translation of HemK by explicit kinetic modeling due to difficulties in constraining the enormous set of parameters required for such a model. To explain the formation and consumption of the intermediates, a Markov modeling approach was developed in collaboration with Sophia Rudorf and Reinhard Lipowsky (Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces). The intrinsic average translation elongation rate was assumed to be  $k_f$  of  $4 \text{ s}^{-1}$  in the absence of pauses (Scheme 1). This estimation for the average rate of translation elongation was based on the observation that the first intermediate of approximately 46 aa length was found after 10 s of translation (Figure 2-6 C). The value of  $k_f$  was thought to represent the rate of translation elongation in the absence of pausing. A codon-unspecific mode of inactivation was introduced with rate  $k_i$  of  $0.48 \text{ min}^{-1}$  for the inactivation and  $k_a$  of  $0.03 \text{ min}^{-1}$  for the reactivation step. A very small fraction of ribosomes would partition into the inactive mode during translation at the rate of  $k_f$ .



The fraction of ribosomes switching to the inactive mode increased at codons elongated with a slower rate  $k_p$  corresponding to sites of translational pausing (Scheme 2). Based on these four parameters, the model was able to describe the global accumulation and consumption of the translation intermediates observed during translation of HemK (Figure 2-8). In addition, the model accounted for the experimentally observed rate of synthesis of full-length HemK. The description of the experimental data was further improved by selecting three different values of  $k_p$  for translation intermediates that decayed with a fast ( $k_p 0.35 \text{ s}^{-1}$ ), intermediate ( $k_p 0.19 \text{ s}^{-1}$ ), or slow rate ( $k_p 0.08 \text{ s}^{-1}$ ), leading to short, intermediate, and long translational pauses.

## Results

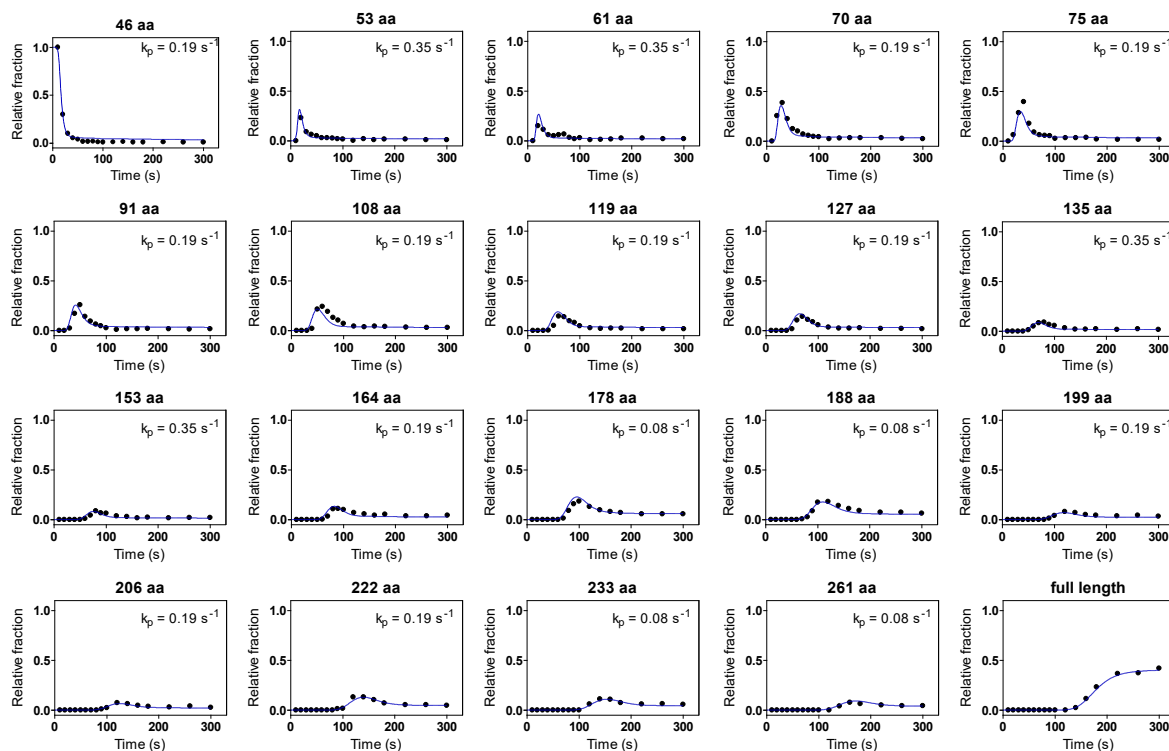


Figure 2-8 Global model for accumulation and decay of translation intermediates during translation of HemK 1-277 aa (courtesy of Sophia Rudorf). The intensity of each translation intermediate was normalized to the fluorescence intensity of the complete lane (black dots). The simulated accumulation and decay of the translation intermediates according to Scheme 2 is shown as blue line. The rate of  $k_p$  used for the description of individual intermediates is indicated.

So far, the exact molecular mechanism for the steps of inactivation and re-activation is not known. On the other hand, we could experimentally validate the assumption that the translation intermediates are sampled in a sequential manner. We observed that bands corresponding to translation intermediates of 54-56 aa and 90-92 aa length accounted for the majority of translation products until 180 s when the translation was performed at 20°C (Figure 2-9). These main translational intermediates were clearly sampled sequentially at lower temperature. Similarly, less pronounced bands continued in synthesis before the next translation intermediate accumulated to highest intensity. The high rate of translation elongation at elevated temperature was most likely masking the sequential accumulation of translation intermediates. However, we could not rule out that certain translation intermediates were selectively stabilized at lower temperatures. The activation energy of for the synthesis of full-length HemK 1-277 aa was  $-24 \pm 2$  kcal mol<sup>-1</sup>, as calculated from the Arrhenius plot in the temperature range from 20°C to 37°C. This value was

## Results

comparable to the activation energy calculated for the translation of CspA 1-70 aa (Section 2.1.1) which showed significantly less translation pausing over the same temperature range. The activation energy calculated from the rate of formation of HemK 1-277 aa was also comparable to the activation energy for the decoding process in isolation as measured from the temperature dependence of the step of GTPase activation (Johansson et al., 2008a) and accommodation (Wohlgemuth et al., 2010). The similarity of the activation energy for decoding and for translation of HemK 1-277 aa might point to the individual decoding steps as rate-limiting barriers and not the translocation steps that are sensitive to mRNA secondary structure elements (Chen et al., 2013; Qu et al., 2011). On the other hand, we could not assess the contribution of individual translational pauses to the temperature dependence of the global translation elongation.

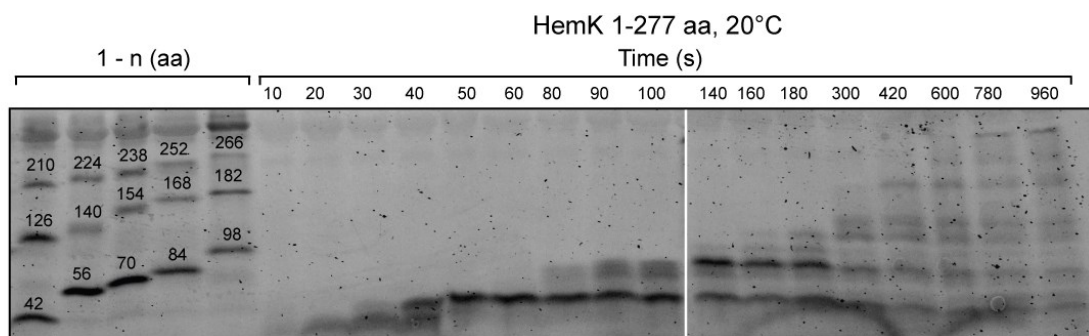


Figure 2-9: Time course of translation of HemK 1-277 aa at 20°C.

Translation of mRNA shortened at the 3'-end led to the accumulation of identical translation intermediates as observed for the translation of the respective region of full-length HemK 1-277 aa (Figure 2-10). This suggests that translational intermediates were unlikely to accumulate due to long-range mRNA secondary structures that need to be unwound by the elongating ribosome. On the other hand, the removal of local mRNA structures would require the truncation of the mRNA to similar length as the location of the translation intermediate. The small difference in the length of translation products derived from the translation intermediate and the final product of the truncated mRNA did not allow separation of both species by Tris-tricine PAGE. For example, the translation product of HemK 1-42 aa yielded only one single band.

## Results

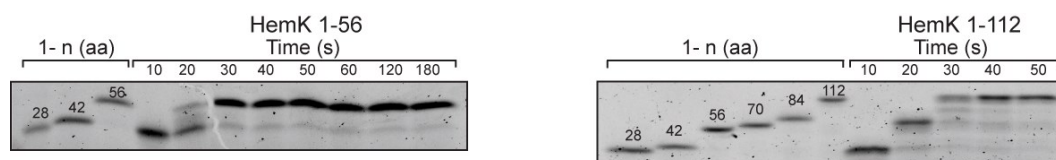


Figure 2-10: Translation of shorter segments of the mRNA coding for HemK.

The same pattern of translation intermediates was observed when the mRNA was heated to 98°C and rapidly cooled prior formation of the IC. Local mRNA secondary structure elements might have reformed during the cooling process or formation of the IC. In an alternative scenario, secondary structure elements could be shaped during the movement of the elongation ribosome along the mRNA. The helicase site of the ribosome acts on the mRNA 11 nucleotides upstream of the P site codon (Qu et al., 2011; Takyar et al., 2005). We used the mFold and Nupack algorithms to predict the local structure along the mRNA (Dirks and Pierce, 2003; Zuker, 2003). We calculated the minimum free energy for a window of 40 nucleotides and shifted the window by one nucleotide along the mRNA sequence. However, we could not detect any significant change in the predicted free energy indicative of secondary structure formation near the position of the pausing sites listed in Table 2-4.

Table 2-4: List of translation intermediates of within the first 100 aa of HemK that were detected during translation of mRNA with different length. Standard deviation were calculated from the length determined in independent experiments.

Intermediate	Length (aa)
1	33 ± 3
2	56 ± 3
3	70 ± 3
4	76 ± 3
5	92 ± 3

We used saturating concentrations of total aa-tRNA for the translation of all mRNA sequences in Section 2.2. The accumulation and consumption of translation intermediates could not be influenced by increasing the concentration of total aa-tRNA (Jörg Mittelstät, PhD thesis). The addition of more aa-tRNA simultaneously increases the concentration of cognate, near-cognate and non-cognate tRNA species for a given codon in the A site (Figure 1-4 B). A higher absolute concentration of the aa-tRNA might not increase the rate of decoding of all codons in the same manner because the

## Results

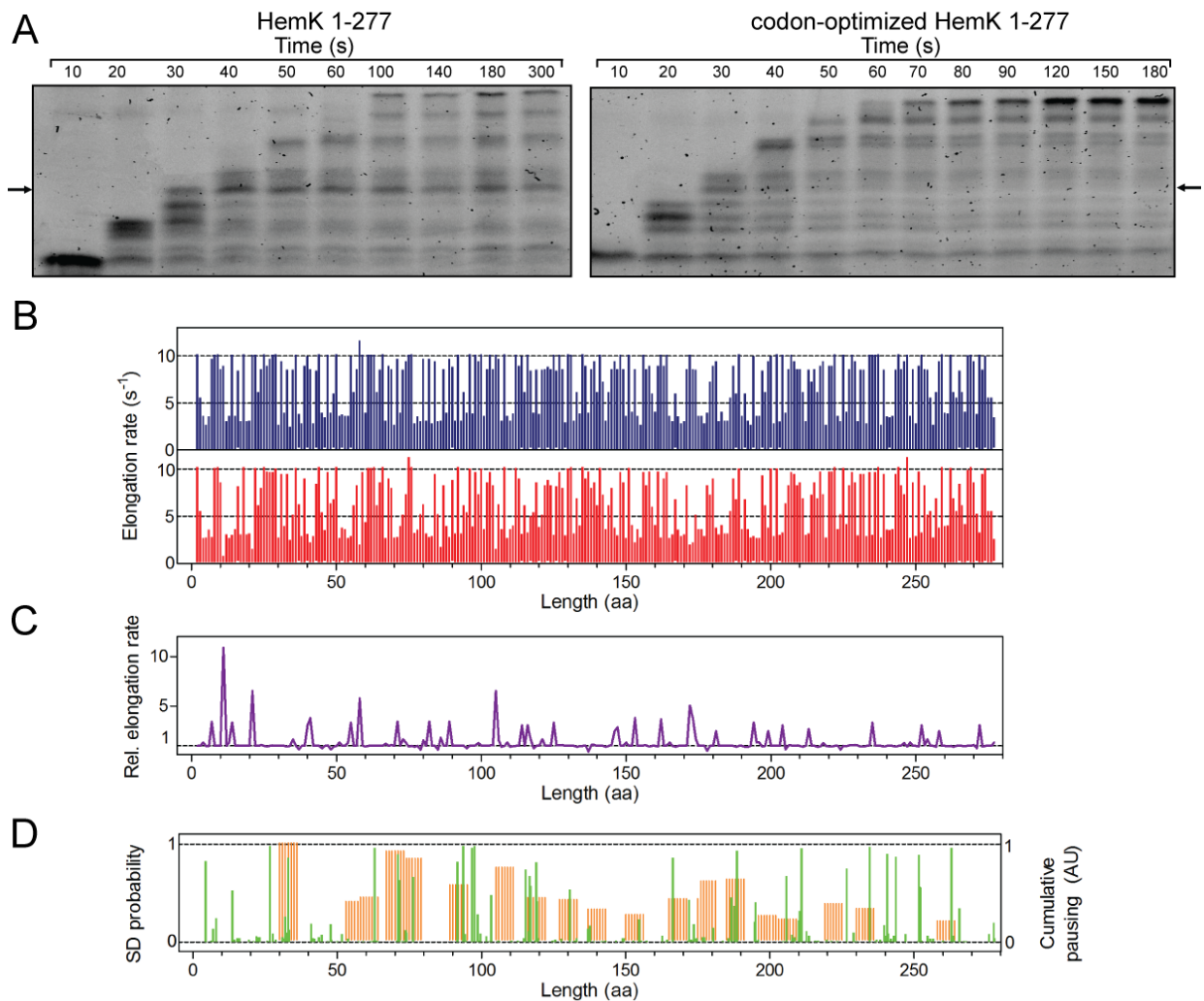


Figure 2-11: Translation of mRNA coding for HemK with the nucleotide sequence from the *E. coli* strain BL21(DE3) and codon-optimized sequence. **A** Single-turnover translation of both mRNA sequences was performed under identical conditions. The position of the translation intermediate of  $108 \pm 3$  aa is indicated by arrows. **B** Codon-specific elongation rates calculated according to (Rudorf et al., 2014) are shown in red for the sequence from strain BL21(DE3) and in blue following codon-optimization. **C** Ratio of the calculated codon-specific elongation rates in C obtained by dividing the elongation rates for the codon-optimized sequence by the rates for the sequence from strain BL21(DE3). **D** Predicted probability of stable interaction of the anti-SD sequence with the nucleotide sequences of HemK from *E. coli* strain K12 in a 10 nt window (green). The prediction was performed in the Genome Wide Information on Protein Synthesis visualized (GWIPS-viz) browser (Michel et al., 2014). The cumulative, steady-state pausing level (orange) was calculated from the sum of the intensities over time for individual intermediates during translation of the BL21(DE3) sequence (Figure 2-8).

relative abundance of the isoacceptor tRNA is not changed. As an example, if the decoding of a codon cognate to a rare aa-tRNA might be limited by the much higher abundance of near-cognate aa-tRNA, then an increase in absolute concentration of

## Results

aa-tRNA would not be expected to increase the rate of decoding. In order to evaluate whether the low relative abundance of isoacceptor tRNA could result in translational pausing at saturating concentration of total aa-tRNA, we analyzed the translation of a codon-optimized variant of HemK. In an attempt to remove codons that are dependent on minor isoacceptor tRNA for decoding, we substituted every codons for the synonymous codon that is most frequently used in *E. coli*. An alignment of the nucleotide sequence from strain *E. coli* strain BL21(DE3) with the codon-optimized variant of HemK is reported in the Appendix together with the codon table used for gene optimization. The codon substitution increased the global translation elongation rate compared to the sequence from BL21(DE3) (Figure 2-11 A & Figure 2-12). Translation intermediates accumulated for both mRNA sequences and several pauses could be mapped to a similar length (Figure 2-12). Interestingly, the translation intermediate in the BL21(DE3) sequence that was mapped to a length of  $108 \pm 3$  aa was absent or accumulated to a much smaller extent for the codon-optimized sequence. Further, the accumulation of a translation intermediates of  $67 \pm 3$  aa significantly delayed the elongation of the sequence from BL21(DE3) compared to the optimized sequence. We calculated the codon specific elongation rates for both mRNA sequences based on the model for elongation rate *in vitro* described in Section 2.1 (Rudorf et al., 2014). The model for translation elongation rates *in vitro* took into account the concentration of all isoacceptor tRNA and the competition of near-cognate tRNA (Figure 2-11 C). The largest difference in calculated elongation rates is predicted for the Ile codon at position 11, while the second-largest difference was observed for the Arg codon CGG a positions 21 and 105 (Figure 2-11 D). Potential translation intermediates at positions 11 and 21 were outside the range of molecular weight that could be resolved by the Tris-tricine gels. The translation intermediate of  $108 \pm 3$  aa length could corresponded to the rare codon CGG at position 105 that is decoded by the minor isoacceptor tRNA<sup>Arg<sub>3</sub></sup><sub>CCG</sub>. The CGG codon was replaced by CGU in the codon-optimized sequence that is decoded by the major isoacceptor tRNA<sup>Arg<sub>2</sub></sup><sub>ICG</sub>. We have analyzed the free energy of folding of mRNA secondary structures in the region from codon 90 to 140 and found that the codon-optimized sequence had a higher propensity to form secondary structure in this region compared to the sequence from BL21(DE3).



## Results

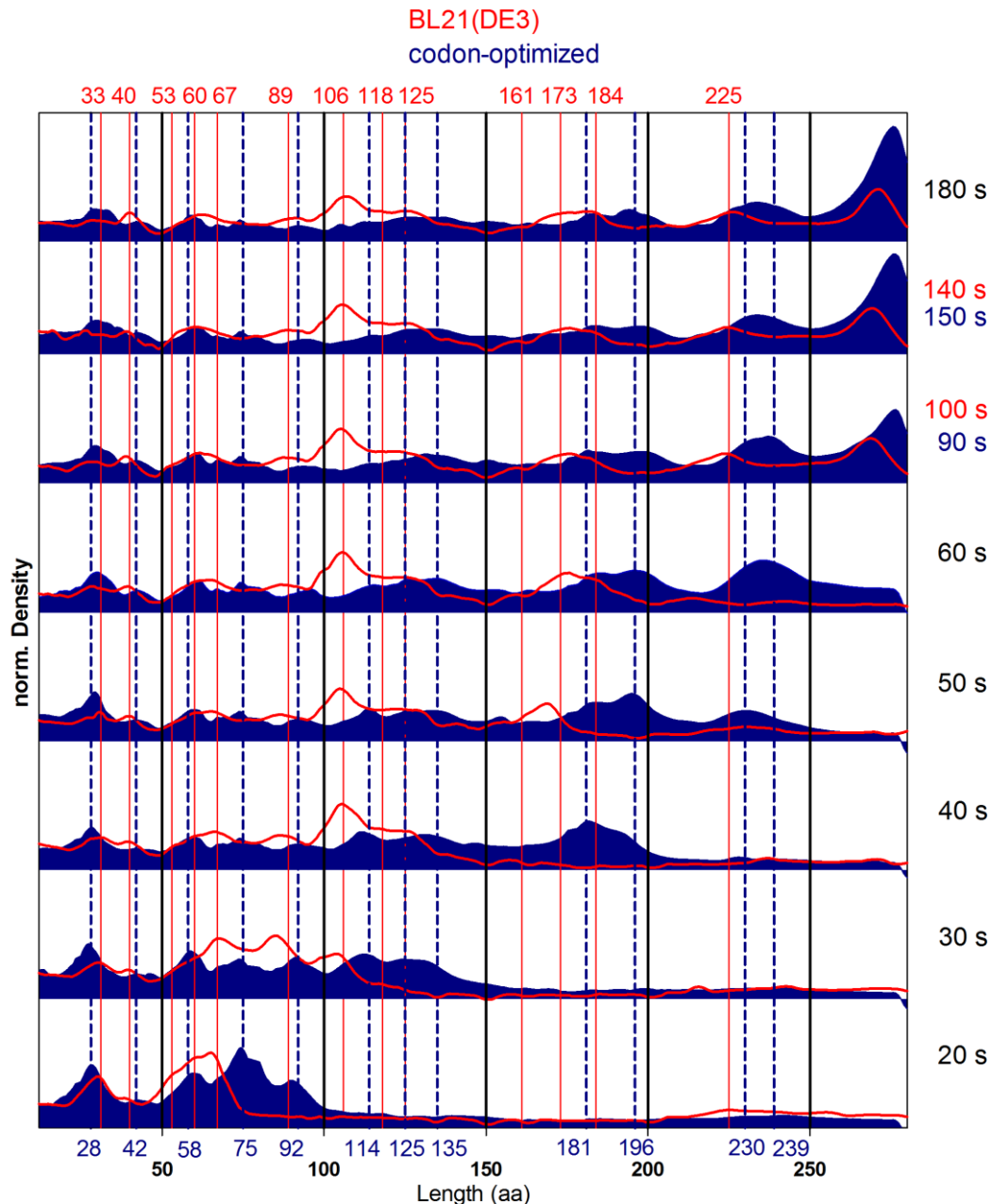


Figure 2-12: Detailed densitometric analysis of the accumulation of nascent peptides during translation of the different HemK nucleotide sequences in Figure 2-11 A. The density of nascent peptides at different time points of *in vitro* translation is outlined in red for the nucleotide sequence from BL21(DE3) or blue for the codon-optimized sequence. The location of translation intermediates is indicated in the same color code for both sequences. Time points of *in vitro* translation are indicated in black when the accumulation of nascent peptides is compared after the same time interval or color coded when the time scale differed.

We concluded the translation intermediate detected at position  $108 \pm 3$  aa was more likely linked to the substitution of individual codons and availability of cognate tRNA than to the local secondary structure of the mRNA. The reduced accumulation of the

## Results

translation intermediate at  $67 \pm 3$  aa might have been linked to the substitution of the CCC codon at position 67 to CCG. The codon CCC is exclusively decoded by tRNA<sup>Pro2</sup><sub>GGG</sub>, while the codon CCG is read by both the major isoacceptor tRNA<sup>Pro1</sup><sub>CGG</sub> and the minor acceptor tRNA<sup>Pro3</sup><sub>UGG</sub>. We could not correlate the accumulation of additional translation intermediates with the codons substituted during the codon-optimization of the HemK sequence. However, it remained possible that accumulation of transitional pauses of HemK 1-277 aa was linked to the usage of a specific set of codons. The codon-optimization resulted in substitution of 49% of the codons, leaving half of the sequence unmodified.

Alternatively, several independent mechanisms not related to the decoding of specific codons could underlie the accumulation of individual translation intermediates. Translational pausing *in vivo* (Li et al., 2012) was attributed to presence of SD-like sequences in the mRNA, which could interact with the anti-SD sequence of 16S rRNA thereby slowing down translation. Although the correlation between the presence of the SD-like sequences and translational pausing was questioned in the subsequent studies (Martens et al., 2015; Woolstenhulme et al., 2013), we have also analyzed potential interactions between the anti-SD sequence and the HemK mRNA. In fact, SD-like sequences in the HemK transcript from *E. coli* strain K12 were predicted in the regions which could account for the mapped translational pauses of approximately 33, 70, 75, 92, 120, 130, 168, or 168 aa length (Figure 2-11 D, Table 2-5). Several SD-like sequences were conserved between the HemK sequences from *E. coli* strains K12 and BL21(DE3) (Appendix). We have also analyzed the predicted interaction energies of the SD-like sequences by using an independent prediction tool (Table 2-5, Figure 2-11 D). The process of codon-optimization of HemK removed some of the SD-like sequences while creating them at novel positions. The removal of a strong SD-like sequence at codons 99-100 could have been related to the lower accumulation of the translation intermediate at  $108 \pm 3$  aa (Figure 2-11 A, Figure 2-12). Potential interactions with the anti-SD sequence might have accounted for the accumulation of a similar number of pauses for the sequence from strain BL21(DE3) and after codon-optimization.

## Results

Table 2-5: Positions of the SD-like sequences and the predicted strength of interactions between the 16S rRNA sequence 5' ACCUCCUUA 3' and nucleotide sequences of the HemK mRNA were predicted using the IntaRNA algorithm (Wright et al., 2014).

<u>BL21(DE3)</u>		<u>Codon-optimized</u>	
<u>codons</u>	<u>(kcal mol<sup>-1</sup>)</u>	<u>codons</u>	<u>(kcal mol<sup>-1</sup>)</u>
1-2	-2.9	28-29	-2.2
34-36	-3.0	78-79	-3.1
72-74	-2.6	107-108	-3.2
75-75	-2.1	114-115	-2.0
95-96	-3.3	172-174	-2.3
99-100	-4.9	230-232	-3.5
116-118	-3.7	252-254	-3.6
236-237	-2.0	272-274	-2.1

In order to unambiguously determine the causes of translational pausing at a specific site, the positions of the translational pauses on the mRNA must be determined with higher resolution. Marker peptides with a defined length allowed us to map the position of pausing sites within a range of six codons. However, for investigation of translational pausing in more detail a single-codon resolution would be desirable. Such high resolution could not be achieved by separating the translation products on Tris-tricine PAGE. We thus decided to develop alternative approaches that could yield localization of pausing sites with single-codon resolution.

### 2.2.3 Hybridization of peptidyl-tRNA on micro-array chips

Analysis of nucleic acid sequences on micro-array chips follows the same principles of hybridization to an oligonucleotide probe as used for Northern and Southern blotting. In the first application of a micro-array chip, DNA sequences were printed by a robot on the chip and fluorescence-labeled cDNA was produced by reverse transcription of mRNA sequences in the presence of fluorescence-labeled nucleotides (Schena et al., 1995). Reverse-transcription in the presence of nucleotides labeled with fluorophores that could be excited and detected at different wavelength allowed the determination of relative mRNA levels between two samples in a two-color experiment. Similarly, fluorescence-labeled RNA can be hybridized directly to a complementary DNA probe. Tao Pan and co-workers pioneered the analysis of tRNA by micro-arrays. The relative abundance of specific tRNA molecules was detected after ligation of a fluorescence-labeled mixed DNA-RNA oligonucleotide sequence to the conserved CCA end of the tRNA (Dittmar et al., 2004). In a later study, misaminoacylation of tRNA by MetRS was analyzed by detection of tRNA aminoacylated with [<sup>35</sup>S]-methionine after hybridization (Netzer et al., 2009).

## Results

With the help of Dr. Gabriela Salinas-Riester (Transcriptom Analyse Labor (TAL), University Medicine Göttingen), we designed an approach for extraction and hybridization of fluorescence-labeled peptidyl-tRNA during the course of *in vitro* translation. Although deacylation of peptidyl-tRNA is slower than that of aa-tRNA (Bresler et al., 1968), we chose hybridization conditions reported for the hybridization of more labile aa-tRNA (Netzer et al., 2009). Peptidyl-tRNA contains the information on both the last aa added to the peptide chain and the codon in the P site. Hybridization of peptidyl-tRNA extracted at specific time points of translation could yield the position of the translational pausing in the mRNA as well as the aa sequence (Figure 2-13 A). A suitable fluorescence label was incorporated into the peptidyl-tRNA by performing translation with initiator tRNA<sup>Met</sup> labeled with Bodipy576/589-C<sub>2</sub> at the  $\alpha$ -amino group of methionine (Section 2.3.1).

In the previous work, DNA probes complementary to the complete sequence of the tRNA were printed on the micro-array (Dittmar et al., 2004; Dittmar et al., 2005; Netzer et al., 2009). In contrast, we have used 60-mer gene expression micro-array chips compatible with the equipment available at the TAL. 60-mer cDNA probes have a different selectivity for nucleotide mismatches depending on the spacing from the surface of the micro-array slide (Hughes et al., 2001). For a 60-mer cDNA probe, the highest selectivity against a single nucleotide mismatch during hybridization is located within the first ten nucleotides at the 5' end of the probe (Figure 2-13 A). We inserted the same sequences previously successfully used for Northern blotting of all isoacceptor tRNA from *E. coli* into the region with highest base-pair discrimination of the 60-mer cDNA probes (Table 4-11) (Dong et al., 1996); no additional nucleotide sequence needed to be added as linker. The probe for tRNA<sup>Phe</sup> formed an exception because the sequence used for Northern blotting could not be inserted at the 5' end of the DNA probe. Additionally, we have selected seven tRNA from *S. cerevisiae* as a control for non-specific hybridization. Each of the 45 probes was present in 140 replicates per sub-array.

## Results

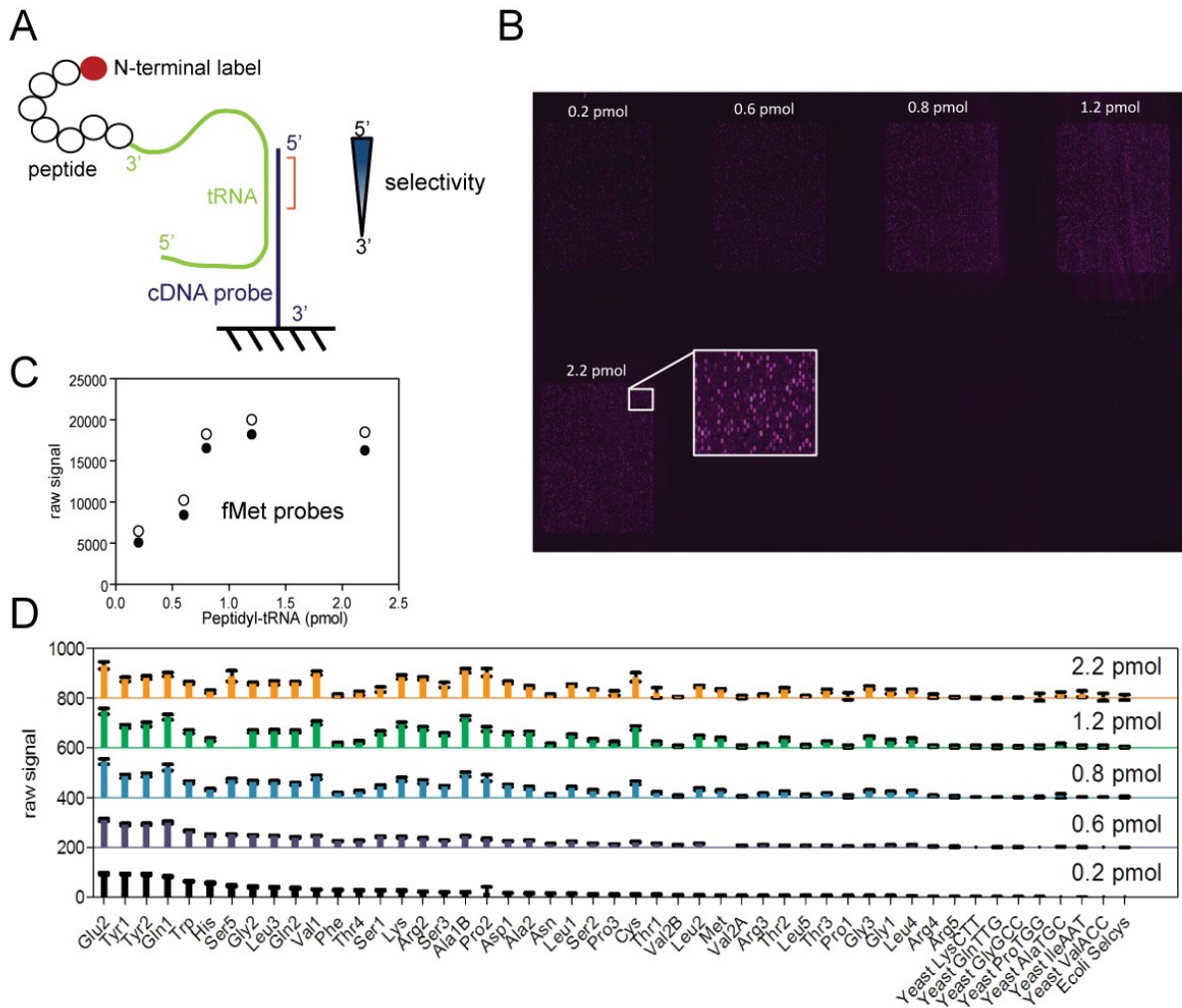


Figure 2-13: Hybridization of peptidyl-tRNA on micro-array chips. **A** Schematic representation of hybridization of peptidyl-tRNA on micro-array slide. Peptidyl-tRNA are extracted during the time course of translation of ribosomes initiated with Bodipy576/589-Met-tRNA<sup>fMet</sup>. Peptidyl-tRNA are hybridized to specific DNA probes attached to the surface of the micro-array chip. Sequences previously used for Northern blotting by (Dong et al., 1996) are incorporated in the most selective region of the cDNA probes. **B** Image of the microarray slide with increasing amounts of peptidyl-tRNA spotted. The amount of peptidyl-tRNA was calculated based on the radioactivity of Bodipy576/589-C<sub>2</sub>-[<sup>3</sup>H]Met. Hybridization was performed for 90 min at 65°C. A magnification to the level of spots for individual cDNA probes is provided for the sub-array hybridized with 2.2 pmol peptidyl-tRNA. **C** Uncorrected fluorescence intensities for the probes fMet1 and fMet2 at increasing amount of peptidyl-tRNA. **D** Intensities for the all cDNA probes with the exception of the fMet probes depicted in C. Probes are ordered according to decreasing intensity as obtained for hybridization of 0.2 pmol peptidyl-tRNA.

We selected the antibiotic viomycin as potent inhibitor for halting translation at chosen time and location on the mRNA. Viomycin inhibits the translocation of the

## Results

peptidyl-tRNA from the A site to the P site by trapping the peptidyl-tRNA in the A site (Modolell and Vazquez, 1977). We verified that the addition of viomycin to ongoing single-turnover translation prevented further nascent chain elongation for a period longer than 30 minutes at 37°C. Arrested ribosomes were concentrated by pelleting through a sucrose cushion and the peptidyl-tRNA was released by resuspension of the pelleted ribosomes in a buffer containing EDTA and SDS. We used commercial buffers provided by Agilent for the hybridization and subsequent washing steps of the micro-array slide. In previous studies, hybridization of aa-tRNA was performed at pH 4.8 (Netzer et al., 2009). We observed strong precipitation of ribosomal proteins at this pH and performed hybridization at pH 5.5. No hydrolysis of peptidyl-tRNA at pH 5.5 was observed during the incubation for 2.5 h at 65°C as verified by the analysis of the peptidyl-tRNA on neutral Bis-tris PAGE.

We determined the sensitivity of the micro-array by spotting increasing amounts of peptidyl-tRNA extracted after 25 s of translation of HemK 1-112 aa (Figure 2-10 & Figure 2-13 B). We could detect hybridization even at the lowest amount of peptidyl-tRNA spotted. The probes fMet1 and fMet2 showed two orders of magnitude higher intensities than any other probe. Fluorescence-labeled initiator tRNA might have been co-purified with the peptidyl-tRNA and exhibited a high efficiency of hybridization to the cDNA probes. Similarly, at low concentration of peptidyl-tRNA, we observed the highest signal for probes that were hybridizing with tRNA carrying short N-terminal peptides (Figure 2-13 D). The intensities were following the order of the N-terminal aa sequence fMEYQHW found in HemK. Short peptides generated from the drop-off of peptidyl-tRNA from the ribosome appeared to have a higher hybridization efficiency than peptidyl-tRNA carrying longer peptides. At higher amounts of peptidyl-tRNA, the signal intensities for cDNA probes hybridizing to presumably longer peptidyl-tRNA started to accumulate (Figure 2-13 D). We did not observe significant hybridization to the negative control probes of tRNA from *S. cerevisiae*. The intensity of most probes followed a similar saturation dependency as observed for the probes fMet1 and fMet2 (Figure 2-13C). On the other hand, a subset of probes did not show an increase in intensity at higher concentration of peptidyl-tRNA comprising the probes for Asn, His, Phe, Ser1, Thr4, Tyr1, Tyr2, Trp, Val2A, and Val2B. For the following experiments, we hybridized 1.5 to 2 pmol of peptidyl-tRNA per sub-array.

We extracted peptidyl-tRNA at 10, 20, and 25 s during translation of HemK 1-112 aa. The signal intensity of the individual cDNA probes was normalized to the intensity of the fMet probes for each time point. Most probes did not exhibit a change

## Results

in intensity over time (not shown), except for Gly2 and Lys, which showed a strong decrease in intensity (Figure 2-14). Interestingly, only one codon for lysine is present in HemK 1-112 aa and two GGG codons that could be read by tRNA<sup>Gly2</sup>. Both codons are found next to each other at positions 34 to 35 with the sequence AAAGGG.

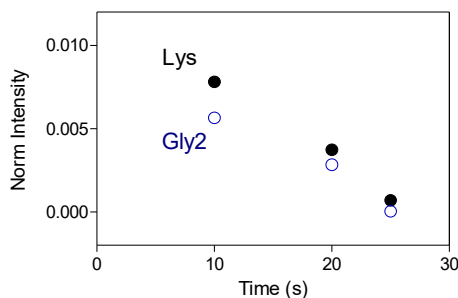


Figure 2-14: Decay of normalized intensities for hybridization of peptidyl-tRNA to probes Lys and Gly2 during translation of HemK 1-112 aa. Translation was inhibited at indicated time points by addition of viomycin and peptidyl-tRNA were extracted. Hybridization was performed for 2.5 h at 65°C. Intensities of the Lys and Gly2 probes were normalized to the intensities of the the probes fMet1 and fMet2 for each time point.

Interestingly, this corresponds to the region where the first translation intermediate was mapped based on the marker peptides. Halting the translation with viomycin has the potential to trap different peptidyl-tRNA species at the same site of translational pausing. Unfortunately, we were not able to detect accumulation of a hybridization signal corresponding to Asp56 or Arg112 at the endpoint of translation of HemK 1-56 or HemK 1-112 aa. Peptides longer than 40 aa might not be detected due to lower hybridization efficiency of the corresponding peptidyl-tRNA.

### 2.2.4 Hybridization of 3'-labeled tRNA on micro-array

In section 2.2.3, we observed that peptidyl-tRNA carrying longer peptide chains showed a lower hybridization efficiency. To circumvent potential limitations imposed by longer peptide chains on the peptidyl-tRNA, we sought to establish a different sample preparation protocol. We removed the nascent peptide from the tRNA after purification of ribosomal complexes containing peptidyl-tRNA. We followed a general approach for modification of the 3'-end of RNA strands involving oxidation of the vicinal diol groups at the 3'-ribose to aldehyde groups and subsequent reaction of the aldehydes with hydrazide-functionalized cyanine (Cy) fluorophores (Figure 2-15) (Zamecnik et al., 1960).

## Results

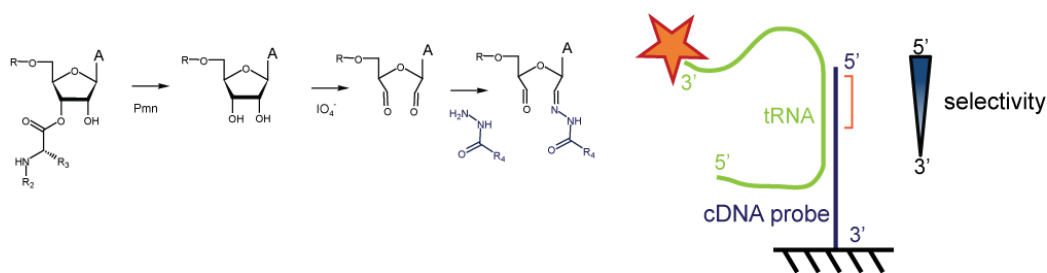


Figure 2-15: Schematic depiction of the sample preparation protocol for hybridization of 3'-labeled tRNA. Translation *in vitro* was halted at specific time point by the addition of viomycin. Ribosome nascent chains complexes were purified through sucrose cushion. Nascent peptides were released from the ribosome by addition of puromycin (Pmn). Ribosomal proteins and mRNA/rRNA were removed. The ribose at the 3'-end of deacylated tRNA was oxidized with periodate and aldehyde groups were labeled with hydrazide-functionalized fluorophores. The 3'-labeled tRNA were hybridized to the same cDNA probes on the micro-array chip as described in section 2.2.3.

Based on a similar concept, several studies analyzed the aminoacylation level of tRNA on micro-array chips (Dittmar et al., 2005; Zaborske and Pan, 2010; Zaborske et al., 2009). In these studies, the ribose at the 3'-end of deacylated tRNA was oxidized with periodate to prevent ligation of a fluorescence-labeled mixed RNA/DNA oligonucleotide to the 3'-end of the tRNA molecule by T4 RNA ligase. Oxidation with periodate requires two hydroxyl groups in vicinal positions and the aminoacyl ester bond prevents oxidization of the ribose. Following treatment with periodate and subsequent deacylation of the aminoacyl-tRNA, only tRNA that was aminoacylated during periodate treatment could be ligated to the fluorescent oligonucleotide. This is in contrast to the sample preparation protocol outlined in Figure 2-15 where only the oxidized ribose groups can be fluorescence-labeled.

We reasoned that the analysis of the aminoacylation level of total tRNA could be an interesting benchmark for the hybridization of 3'-labeled tRNA on micro-array. We subjected aminoacylated and deacylated total tRNA to oxidation with periodate and labeled the oxidized 3'-end of the tRNA with both Cy3-hydrazide and Cy5-hydrazide in separate reactions. We combined aminoacylated and deacylated samples labeled with different cyanine fluorophores for hybridization in a two-color experiment. The ratio of the fluorescence intensity between Cy3 and Cy5 was used to calculate the aminoacylation level. Differences in labeling efficiency between the cyanine fluorophores was assessed in previous studies by the comparison of the aminoacylation level calculated from the inverted combinations of Cy3: Cy5 vs.



## Results

Cy5: Cy3 (Zaborske and Pan, 2010). Similarly, we performed separate two-color experiment using aminoacylated samples labeled with Cy3 or with Cy5 and deacylated samples labeled with the corresponding second fluorophore (Figure 2-16).

Unfortunately, the usage of a higher amount Spike-In probes for detection of the edge-points of the sub-arrays resulted in the unspecific hybridization to several probes complementary to tRNA species. The Spike-In probes, as described in detail the Methods section (4.14.6), consist of 10 polyadenylated transcripts from the Adenovirus E1A gene that were labeled internally with Cy3-CTP and labeled at the 3'-end with Cy5 hydrazide. Spike-In probes were included in the 3'-labeled tRNA samples for the purpose of hybridization to edge points on each sub-arrays that formed the bases of the coordinate grid required for extraction of the intensities of individual cDNA probes from the scanned image of the micro-array chip. Titration of the amount of Spike-In probes during following control experiments yielded a high level of unspecific hybridization to probes Asp1, Cys, Gly2, Ser3, Trp, and Val1 in the Cy3 channel. On the other hand, the Asn, Tyr1, and Tyr2 probes showed unspecific hybridization in the Cy5 channel. All the probes characterized by unspecific hybridization were removed from the analysis in Figure 2-16.

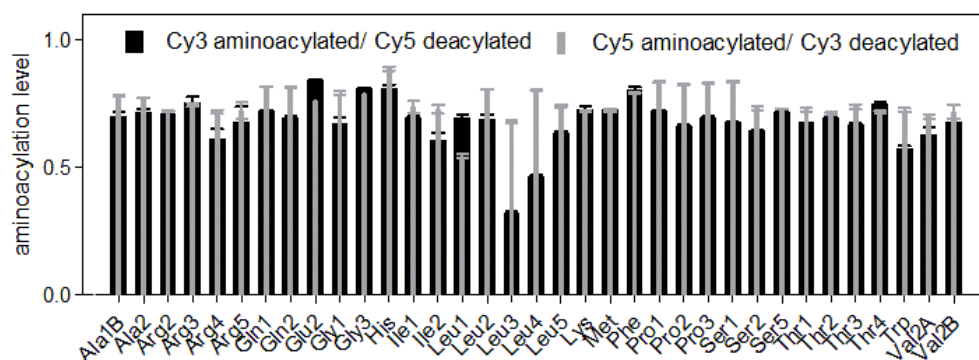


Figure 2-16: Analysis of the aminoacylation level. Equal amounts of total tRNA samples (0.68 pmol) were hybridized on micro-array. The fluorescence intensities in the Cy3 or Cy5 channel were normalized to the fraction of labeled tRNA in the total tRNA sample. Aminoacylated samples were labeled to 6% with either Cy3 or Cy5 and deacylated total tRNA was labeled with 22% Cy3 and 18% Cy5. The aminoacylation level was calculated from the ratio of the raw intensity from both channels normalized to the fraction of fluorescence-labeled total tRNA.

The largest discrepancy in calculated aminoacylation level from both combinations of cyanine fluorophores was observed for tRNA<sup>Leu3</sup> UAG and tRNA<sup>Leu4</sup> CAA. The low abundance of tRNA<sup>Leu3</sup> UAG in *E. coli* total tRNA could have been the reason

## Results

for the variation between the experiments with different combination of cyanine fluorophores (Dong et al., 1996). However, the discrepancy for tRNA<sup>Leu</sup><sub>4 CAA</sub> could not be explained by the abundance of the tRNA isoacceptor. Additional controls will be required to characterize and validate the response of the fluorescence signals detected in both channels of the micro-array scanner that formed the basis for determination of the aminoacylation level. The response of the fluorescence scanner could be assessed by keeping the amount of total tRNA labeled with one fluorescent probe constant and varying the amount of total tRNA labeled with the complementary fluorescent probe.

Encouraged by the measurements of aminoacylation levels, we tested whether we could detect tRNA that is bound in the A or P sites at a given time during translation. We chose to identify peptidyl-tRNA<sup>Gln</sup><sub>2, CUG</sub> that would be located in the P site at the endpoint of translation of leader peptidase LepB 1-35 aa (Bornemann et al., 2008). Ribosomes were allowed to translate until the end of the truncated mRNA and purified. The sample for hybridization on micro-array chip was prepared according to the workflow outlined in Figure 2-15. The peptide was released by incubation with puromycin (Pmn) and the deacylated tRNA was extracted and labeled with Cy3- hydrazide at the 3'-end.

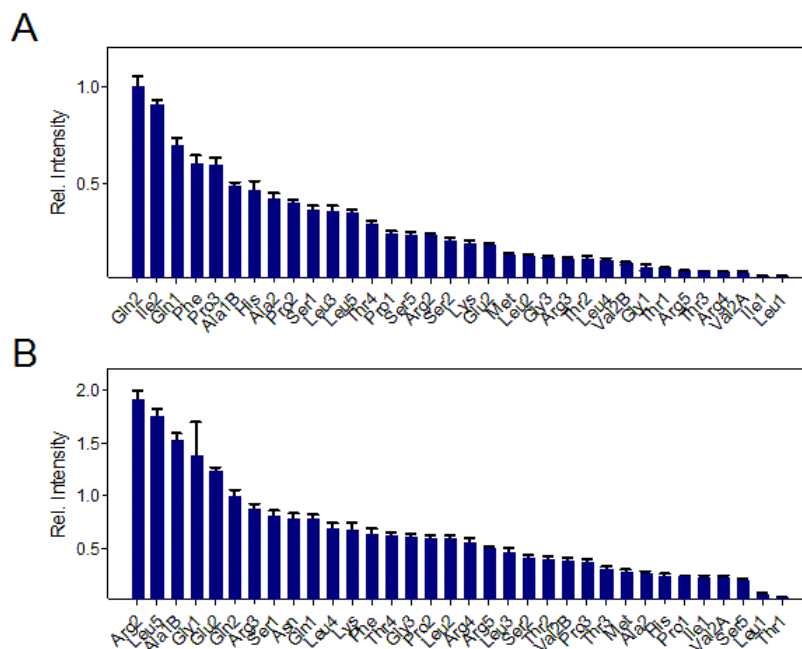


Figure 2-17: Identification of tRNA in the P site of a defined ribosomal complex. tRNA extracted at the endpoint of translation of LepB 1-35 aa and labeled at the 3'-end with Cy3 was hybridized together with 0.68 pmol deacylated total tRNA labeled to 18% with Cy5. **A** Raw intensities of the DNA probes in the Cy3 channel. **B** Intensities corrected by the relative hybridization efficiencies calculated from the total tRNA in the Cy5 channel.

## Results

We hybridized the extracted tRNA<sup>Gln2, CUG</sup> labeled with Cy3 together with deacylated total tRNA labeled with Cy5 as reference standard. The raw intensities of the cDNA probes in the Cy3 channel are presented in Figure 2-17 A. The highest intensity was observed for the Gln2 probe consistent with the hybridization of tRNA<sup>Gln2, CUG</sup>. Additionally, the closely related probe Gln1 showed the third highest intensity. However, several other DNA probes showed significant signal level as well. We calculated the relative hybridization efficiency for all isoacceptor tRNA to the cDNA probes from the signal intensity of the cDNA probes in the Cy5 channel after normalization to the relative abundance of the isoacceptor species as reported by (Dong et al., 1996). We corrected the raw intensities of the tRNA extracted from the RNC in the Cy3 channel with the calculated relative hybridization efficiency (Figure 2-17 B). This correction completely reshuffled the order of intensities for the cDNA probes. Unfortunately, the correction for hybridization efficiency increased the intensities of several cDNA probes over the intensity of the Gln2 probe.

The comparison of the absolute intensities between different cDNA probes on the same micro-array is a general limitation of the micro-array approach. For our application, we were not able to extract information on the absolute amount of tRNA hybridized to individual probes. Information on absolute hybridization of tRNA species might not be required for identification of pausing sites. In a two-color experiment, 3'-labeled tRNA extracted from ribosomal complexes stalled by viomycin at different time points could be normalized to an identical standard e.g. total tRNA labeled with Cy3 or Cy5. Similarly as described for peptidyl-tRNA samples in section 2.2.3, normalized intensities at different time points would allow the identification of pausing sites by the change of intensities for the different cDNA probes over time. The extraction of peptidyl-tRNA bound by the ribosome at a given time of translation, deacylation, and labeling at the 3'-end would allow the identification tRNA species in absence of constraints imposed by the composition and length of the nascent chain on the peptidyl-tRNA.

## Results

### 2.3 Co-translational incorporation of fluorescent probes into nascent peptides

Fluorescence spectroscopy is a key technique for addressing a large number of biological questions. Fluorescence labeling of mRNA, tRNA and ribosomal proteins allowed the in-depth characterization of the translation machinery at various steps of the protein synthesis. aa-tRNA molecules labeled with spectroscopic probes at the aa have been used to monitor events ranging from formation of the ternary complex (Mittelstaet et al., 2013; Nakata et al., 2006), through the steps of translation elongation (Doerfel et al., 2013; Holtkamp et al., 2014a; Indrisiunaite et al., 2015; Mittelstaet et al., 2013; Odom et al., 1990; Samatova et al., 2014), to co-translational and post-translational folding of the synthesized nascent peptide (Ellis et al., 2008; Johnson, 2005; Picking et al., 1991a; Picking et al., 1991b; Tsalkova et al., 1998; Woolhead et al., 2006; Woolhead et al., 2004).

The incorporation of nc-aa is a key challenge for synthetic biology; the efficiency of incorporation is highly dependent on the chemical structure of the nc-aa (Ellman et al., 1991; Hohsaka et al., 1993; Kajihara et al., 2005; Starck et al., 2003). Additionally, nc-aa-tRNA may have altered affinity to EF-Tu (Jeong et al., 2012; Mittelstaet et al., 2013; Nakata et al., 2006). Only a few mechanistic studies characterized the decoding and peptide bond formation involving nc-aa-tRNA in the A site (Jeong et al., 2014; Jeong et al., 2012; Mittelstaet et al., 2013; Wang et al., 2015).

#### 2.3.1 Fluorescent labels at the N-terminus of nascent chains

tRNA labeled at the aa with Bodipy fluorophores were successfully applied to study protein synthesis on the ribosome. BodipyFL-C<sub>2</sub>-Met-tRNA<sup>Met</sup> was used as polarization probe to study the dynamics of nascent chains (Ellis et al., 2008), or as a fluorescence reporter for the characterization of decoding and accommodation of Lys-tRNA<sup>Lys</sup> labeled with Bodipy576/589-C<sub>2</sub>-NHS on the  $\epsilon$ -amino group (Mittelstaet et al., 2013), EF-G catalyzed translocation (Holtkamp et al., 2014a), and finally for detection of nascent chains during single-turnover *in vitro* translation (Doerfel et al., 2013; Samatova et al., 2014). The N-terminal fluorophore did not decrease the overall rate and yield of the single-turnover *in vitro* translation (Doerfel et al., 2013; Holtkamp et al., 2014a; Mittelstaet et al., 2013; Samatova et al., 2014).

## Results

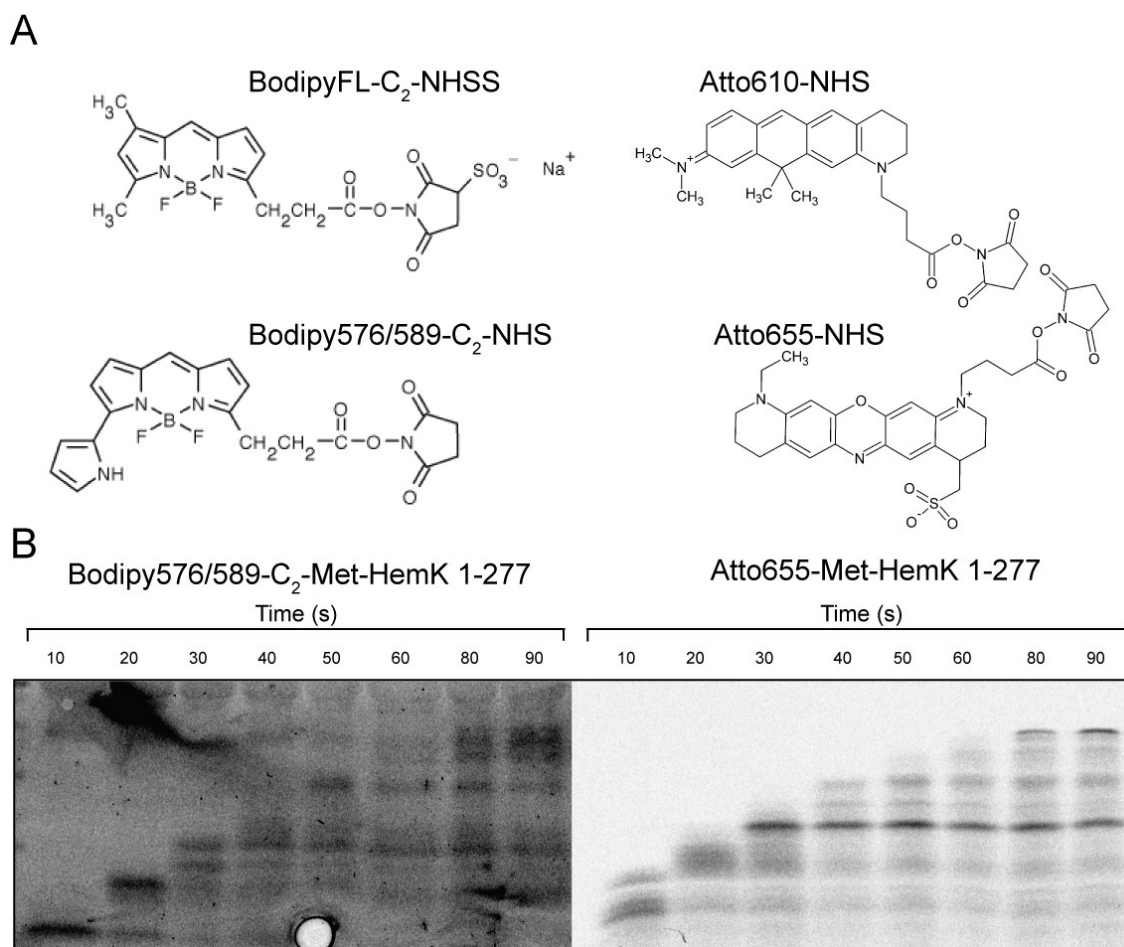


Figure 2-18: Incorporation of fluorescent probes into the N-terminus of nascent chains. **A** Chemical structures of amine-reactive fluorophores used for labeling of Met-tRNA<sup>Met</sup>. **B** Single-turnover translation of the HemK 1-277 aa from *E. coli* strain BL21(DE3) with Bodipy576/589-C<sub>2</sub>-Met-tRNA<sup>Met</sup> (left) or Atto655-Met-tRNA<sup>Met</sup> (right). Single-turnover translation was performed under identical conditions as reported in Section 2.2.2.

Larger fluorophores have been incorporated into the N-terminus during cell-free translation (Miura et al., 2010). Inspired by the properties of BodipyFL-C<sub>2</sub>-Met-tRNA<sup>Met</sup>, we investigated whether fluorophores with increasing size could be incorporated at the N-terminus of nascent peptides and would allow efficient protein synthesis *in vitro*. Bodipy576/589-C<sub>2</sub>-NHS that contains an additional pyrrolyl substituent was successfully incorporated into internal positions of nascent chains (Mittelstaet et al., 2013; Woolhead et al., 2004). We detected the unaltered dynamics of the translation intermediates characteristic for HemK 1-277 aa when single-turnover translation was performed with Bodipy576/589-C<sub>2</sub>-Met-tRNA<sup>Met</sup> compared to translation under identical conditions with BodipyFL-C<sub>2</sub>-Met-tRNA<sup>Met</sup> (Figure 2-11A & Figure 2-18 B). Similarly, N-terminal Bodipy576-C<sub>2</sub>-Met yielded

## Results

accumulation of full-length HemK 1-277 aa at a rate of  $3.42 \pm 0.08$  aa s<sup>-1</sup> that was comparable to rate of formation of full-length HemK of  $> 3$  aa s<sup>-1</sup> with N-terminal BodipyFL-C<sub>2</sub>-Met (Figure 2-11A & Figure 2-18 B). Additionally, single-turnover translation with initiator tRNA labeled with Atto610-NHS and the even larger Atto655-NHS allowed protein synthesis at a high rate of  $3.11 \pm 0.08$  aa s<sup>-1</sup> and  $3.4 \pm 0.5$  aa s<sup>-1</sup>, respectively (not shown for Atto610-NHS, Atto655-NHS Figure 2-18 B). However, we have also noted that the intermediate mapped to a length of  $108 \pm 3$  aa accumulated to a higher level at the expense of full-length product formation when single-turnover translation was performed with N-terminal Atto655-Met (Figure 2-18 B). Similarly, the single-turnover translation with N-terminal Atto610-Met led to a more pronounced accumulation of the intermediate with  $108 \pm 3$  aa length (not shown). Although the large fluorescent probes Atto610 and Atto655 appeared to be compatible with protein synthesis at a high overall rate, it might be possible that the presence of the fluorescent probes significantly altered the local rates of translation elongation, resulting in the more pronounced accumulation of translation intermediates. Notably, the fluorescence of Atto610 was highly sensitive to the presence of reducing agents such as DTT, which limited the feasibility of this fluorophore for translation assays that require a reducing agent. In summary, we have observed that the fluorophores build from aromatic ring systems in a linear, extended arrangement allowed the synthesis of HemK 1-277 aa at a similar rate (Figure 2-11A & Figure 2-18). On the other hand, N-terminal fluorescent probes containing larger and branched aromatic ring systems resulted in the accumulation of translation intermediates and in a reduced rate of translation elongation depending on the dimension of the fluorescent probes (Miura et al., 2010; Ramachandiran et al., 2002; Ramachandiran et al., 2000). It might be possible that N-terminal fluorophores with a linear, extended structure potentially mimicking a polypeptide chain are sterically less restricted during movement from the P site through the polypeptide exit tunnel. Similarly, a higher incorporation level of nc-aa containing linearly arranged aromatic ring systems at internal positions might also reflect constraints imposed during the elongation of the nascent chain through the exit tunnel (Hohsaka et al., 1999; Kajihara et al., 2005).

## Results

### 2.3.2 Activity of tRNA<sup>Cys</sup><sub>GCA</sub> and suppressor tRNA<sup>Cys</sup> from *E. coli*

The fluorophore IA-NBD (Figure 2-23) was incorporated by *in vitro* transcribed Cys-tRNA<sup>Cys</sup><sub>GCA</sub> from *E. coli* and *S. cerevisiae* during cell-free translation in rabbit reticulocyte and wheat germ extract (Gubbens et al., 2010). Additionally, suppressor tRNA decoding the amber UAG, opal UGA, and ochre UAA stop codon based on tRNA<sup>Cys</sup> from *S. cerevisiae* and an opal suppressor tRNA<sup>Cys</sup> from *E. coli* allowed the incorporation of the same fluorescent probe. We considered labeling of Cys-tRNA<sup>Cys</sup> as a promising approach for incorporation of fluorescent probes into nascent chains. A pK<sub>a</sub> value of 8.3 for the unique thiol group of cysteine (Bulaj et al., 1998) allows labeling at lower pH than that required for labeling of primary amino groups and thereby would limit the hydrolysis of the labile aminoacyl ester bond during the labeling procedure. Additionally, a large variety of commercial thiol-reactive fluorophores is commercially available.

We prepared tRNA<sup>Cys</sup><sub>GCA</sub> and the amber, ochre and opal suppressor tRNA<sup>Cys</sup> variants by *in vitro* transcription. The reported protocol for aminoacylation did not allow aminoacylation on a preparative scale (Gubbens et al., 2010). We optimized the protocol starting from conditions published for analytical aminoacylation of tRNA<sup>Cys</sup><sub>GCA</sub> (Komatsoulis and Abelson, 1993). We found that the excess of L-cysteine over tRNA was the key parameter for preparative aminoacylation of tRNA<sup>Cys</sup><sub>GCA</sub> due to a rather high K<sub>M</sub> value (50 μM) of Cysteinyl-tRNA synthase (CysRS) for cysteine (Fersht and Dingwall, 1979). In addition, a cyclization activity of CysRS that transforms cysteine charged on tRNA<sup>Cys</sup> into thiolactone was shown to take place when the substrate L-cysteine was limiting (Jakubowski, 1994). The anticodon sequence of tRNA<sup>Cys</sup><sub>GCA</sub> is a major identity element for recognition of tRNA<sup>Cys</sup><sub>GCA</sub> by CysRS (Komatsoulis and Abelson, 1993). However, high concentrations of recombinant CysRS allow for the aminoacylation of suppressor tRNA<sup>Cys</sup> variants (Gubbens et al., 2010). For all tRNA<sup>Cys</sup> variants, we obtained a specific activity of 700 pmol/OD<sub>260nm</sub>, which corresponded to an aminoacylation level of 45 % (Figure 2-19). Upon interaction with EF-Tu-GTP and separation by size-exclusion chromatography, the ratio of the ternary complex peak to the deacylated tRNA peak was consistent with an aminoacylation efficiency close to 50%.

## Results

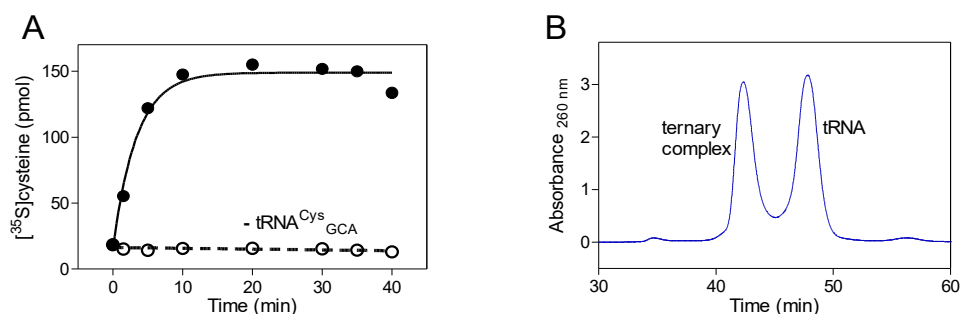


Figure 2-19: Aminoacylation of tRNA<sup>Cys</sup> variants. **A** Time course of aminoacylation of tRNA<sup>Cys</sup><sub>GCA</sub> (280 pmol) with [<sup>35</sup>S]cysteine. High background binding of unbound cysteine was reported previously (Hou et al., 1993). **B** Separation of ternary complex of EF-Tu•GTP•Cys-tRNA<sup>Cys</sup><sub>UCA</sub> from deacylated tRNA on size-exclusion column.

We proceeded to the characterization of the translation efficiency of tRNA<sup>Cys</sup><sub>GCA</sub> and its suppressor variants. As a control for the activity of the tRNA transcripts, which lack post-transcriptional modifications, we purified fully modified tRNA<sup>Cys</sup><sub>GCA</sub> from *E. coli* total tRNA. We measured the rate of peptide bond formation as a combined indicator for the initial selection and proofreading steps. We obtained similar apparent rates of peptide bond formation for the Cys-tRNA<sup>Cys</sup><sub>GCA</sub> transcript and the fully modified Cys-tRNA<sup>Cys</sup><sub>GCA</sub>, about 2 s<sup>-1</sup> (Figure 2-20). GTPase activation during the decoding of the UGC codon by fully modified and *in vitro* transcribed Cys-tRNA<sup>Cys</sup><sub>GCA</sub> in TAKM<sub>7</sub> buffer at 25°C occurs at a rate of 49 s<sup>-1</sup> and 20 s<sup>-1</sup>, respectively (Liu et al., 2011). The relatively small differences observed for the step of initial selection (Liu et al., 2011) and for the combined measure of initial selection and proofreading by the rate of peptide bond formation (Figure 2-20) indicated similar translation efficiency of modified and unmodified Cys-tRNA<sup>Cys</sup><sub>GCA</sub>.



## Results

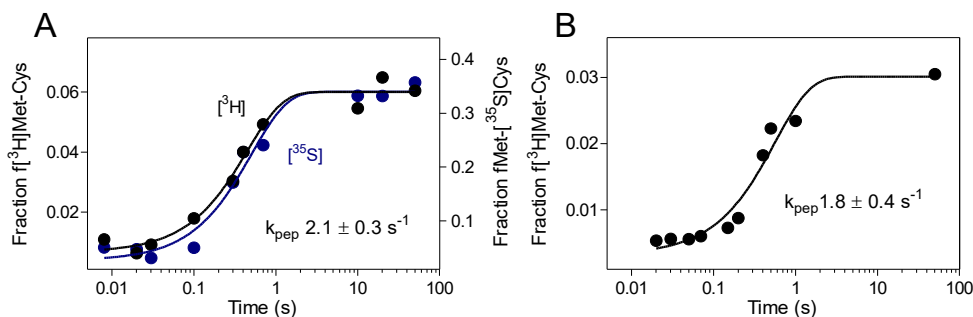


Figure 2-20: Peptide bond formation with the transcript and fully modified Cys-tRNA<sup>Cys</sup>GCA. 70S IC•f<sup>[3H]</sup>Met-tRNA<sup>Met</sup> with the UGC codon in the A site (0.38  $\mu\text{M}$ ) was mixed with  $[^{35}\text{S}]$ Cys-tRNA<sup>Cys</sup>GCA transcript (0.08  $\mu\text{M}$ ) **A** or fully modified Cys-tRNA<sup>Cys</sup>GCA (0.08  $\mu\text{M}$ ) **B**. The reactions were carried out in HiFi buffer at 20°C. Fractions of dipeptides were calculated separately for the  $[^3\text{H}]$  and  $[^{35}\text{S}]$  labels from the ratio of dipeptides formed to the sum of dipeptides and unreacted aa. The apparent rate constants were determined by single-exponential fitting.

We could not detect a significant level of peptide bond formation of ochre suppressor Cys-tRNA<sup>Cys</sup>UUA and opal suppressor Cys-tRNA<sup>Cys</sup>UCA using programmed 70S IC. The lack of translational efficiency of the ochre and opal suppressor tRNA might be explained in terms of the extended anticodon hypothesis suggested by Michael Yarus (Yarus, 1982). The translational efficiency of an anticodon sequence was proposed to be enhanced by the bases in the anticodon loop. The last nucleotide of the anticodon is specifying the combination of bases and post-transcriptional modification within the anticodon loop that allow for efficient decoding in *E. coli*. The extended anticodon hypothesis was experimentally verified by the construction of suppressor tRNA with altered suppression efficiency (Raftery and Yarus, 1987; Yarus et al., 1986a; Yarus et al., 1986b). Nucleotides in the anticodon loop might not directly interact with the decoding center but tune the conformational changes of the tRNA after codon recognition (Ledoux et al., 2009; Olejniczak and Uhlenbeck, 2006). The tRNA is thought to rely the signal for GTPase activation from the decoding center in the 30S subunit to EF-Tu bound at GTPase activation center in the 50S subunit (Piepenburg et al., 2000). Therefore, a single mutation of G34 to U in the anticodon sequence of opal suppressor Cys-tRNA<sup>Cys</sup>UCA was sufficient to abolish its activity in decoding. On the other hand, we could detect peptide bond formation of amber suppressor Cys-tRNA<sup>Cys</sup>CUA. We measured the rate of accommodation of Cys-tRNA by the fluorescence change of BodipyFL-C<sub>2</sub>-Met-tRNA<sup>Met</sup> in the P site (Mittelstaet et al., 2013). For Cys-tRNA<sup>Cys</sup>GCA, the apparent rate constant of accommodation was  $1.4 \text{ s}^{-1}$ , which was comparable to the rate of peptide bond formation in Figure 2-20 A. However, the rate

## Results

of accommodation of the amber suppressor Cys-tRNA<sup>Cys</sup><sub>CUA</sub> was 4-fold lower. The nucleotides at position 32 and 38 at the beginning and end of the anticodon loop, respectively, are important modulators of the translation efficiency of amber suppressor tRNA (Olejniczak and Uhlenbeck, 2006; Yarus et al., 1986a). The highest suppression efficiency was reported for the combination of C32 and with A38. We generated a C32/A38 pair in Cys-tRNA<sup>Cys</sup><sub>CUA</sub> by mutagenesis of U32 to C. The variant Cys-tRNA<sup>Cys</sup><sub>CUA U32C</sub> exhibited a similar apparent rate constant of accommodation as Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (Figure 2-21).

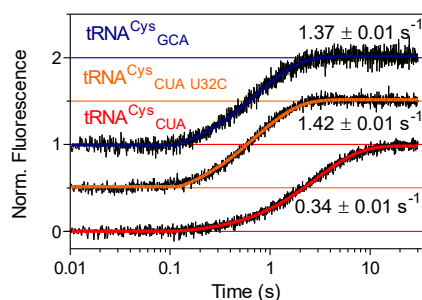


Figure 2-21: A-site accommodation of Cys-tRNA. 70S IC•BodipyFL-C<sub>2</sub>-Met-tRNA<sup>fMet</sup> with codons UGC or UAG in the A site (40 nM) were rapidly mixed with ternary complexes (200 nM) containing Cys-tRNA<sup>Cys</sup><sub>GCA</sub>, Cys-tRNA<sup>Cys</sup><sub>CUA</sub>, or Cys-tRNA<sup>Cys</sup><sub>CUA U32C</sub>. BodipyFL was excited at 473 nm and the emission was recorded after passing a 500 nm cut-off filter. The normalized fluorescence traces are off-set for better visibility and were fit to a model with a delay followed by a single exponential increase. A similar delay time of 0.11-0.12 s was obtained for all traces. Experiments were carried out in HiFi buffer at 20°C.

### 2.3.3 Fluorescent labeling of Cys-tRNA

A protocol for labeling of fully modified Cys-tRNA<sup>Cys</sup><sub>GCA</sub> with BodipyFL-C<sub>2</sub>-maleimide (Figure 2-23) has been published (Lien et al., 2002). Using this protocol with minor changes allowed us to obtain Cys-tRNA<sup>Cys</sup><sub>GCA</sub> labeled with BodipyFL-C<sub>2</sub>-mal in a good yield. We separated labeled tRNA from unlabeled and deacylated tRNA species by HPLC (Figure 2-22 A). The same labeling protocol resulted in a low yield of labeling with BodipyFL-C<sub>1</sub>-iodoacetamide. We analyzed the time course of labeling with BodipyFL-C<sub>1</sub>-IA on urea acid PAGE (Figure 2-22 B). The extent of labeling was linearly depended on the concentration of DMF in the reaction. We obtained a good yield of labeling with various iodoacetamide-functionalized probes when the labeling reaction was performed in 50% DMF and the reaction time was extended to 45 min.

## Results

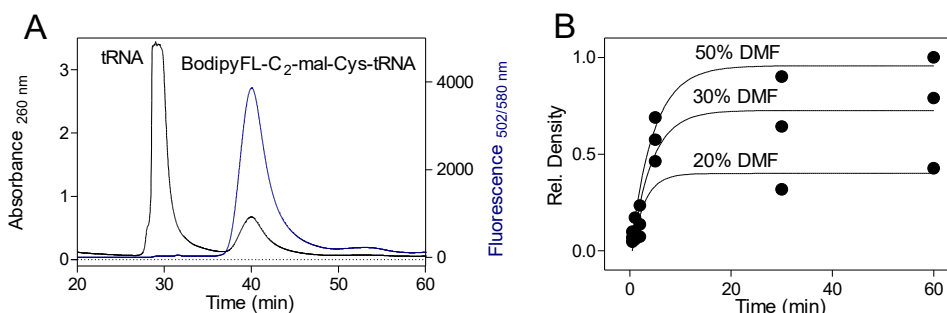


Figure 2-22: Fluorescence labeling and purification of Cys-tRNA. **A** Chromatogram of HPLC separation of BodipyFL-C<sub>2</sub>-mal-Cys-tRNA<sup>Cys<sub>GCA</sub></sup> from unlabeled and deacylated tRNA. **B** Labeling of Cys-tRNA<sup>Cys<sub>GCA</sub></sup> with BodipyFL-C<sub>1</sub>-IA. Time course of labeling was analyzed on a urea acid PAGE and was quantified by fluorescence scanning. Labeling reactions were performed with Cys-tRNA<sup>Cys<sub>GCA</sub></sup> (30  $\mu$ M) and BodipyFL-C<sub>1</sub>-IA (0.6 mM) in HEPES pH 7.5 (15 mM) with increasing concentrations of DMF. The time courses were fit to single-exponential functions.

### 2.3.4 Incorporation of fluorescent probes at internal positions

EF-Tu has a lower affinity for a number of nc-aa than for natural aa (Jeong et al., 2012; Mittelstaet et al., 2013; Nakata et al., 2006). Expansion of the aa binding pocket of EF-Tu increased the affinity for nc-aa-tRNA (Chapman et al., 2012; Doi et al., 2007; Ohtsuki et al., 2010). However, as a consequence of the increase in binding affinity for nc-aa-tRNA, the dissociation of nc-aa-tRNA from EF-Tu•GDP after GTP hydrolysis on the ribosome may become impaired and rate-limiting for peptide-bond formation (Mittelstaet et al., 2013; Schrader et al., 2011). tRNA bodies with increased affinities for EF-Tu•GTP were used as alternative strategy to tune the incorporation of nc-aa-tRNA (Guo et al., 2009; Jeong et al., 2014). tRNA<sup>Cys<sub>GCA</sub></sup> has an intrinsically high affinity for EF•GTP (Asahara and Uhlenbeck, 2005). This favorable property was unlikely to be altered by mutations introduced in the anticodon loop of tRNA<sup>Cys<sub>UCA</sub></sup><sub>U32C</sub> and may partially compensate for the presumed reduced affinity of the aa moiety, but the formation of the ternary complex has to be tested for every nc-aa used.

## Results

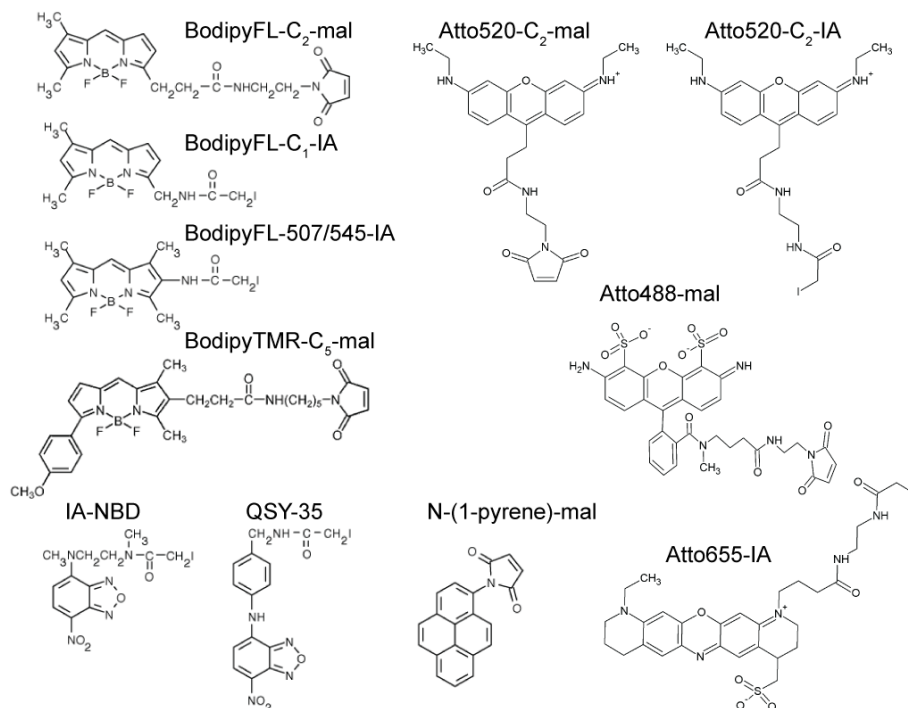


Figure 2-23: Chemical structures of fluorophores used for incorporation at internal positions of nascent chains.

In the light of the potentially reduced affinity for EF-Tu•GTP for fluorescence-labeled Cys-tRNA, we have analyzed the incorporation of a set of maleimide-functionalized fluorescent probes into dipeptides at increasing concentrations of EF-Tu•GTP binary complex (Figure 2-24). In order to limit ternary complex formation to a single round, EF-Tu•GTP was purified from excess GTP using size-exclusion chromatography on NAP-5 column. Incorporation of BodipyFL-C<sub>2</sub>-mal and Atto520 into nascent chains was reported previously (Jeong et al., 2013; Lien et al., 2002). BodipyTMR-C<sub>5</sub>-mal has a similar size as Bodipy579/589-C<sub>2</sub>, which was successfully incorporated into position 34 of HemK (Mittelstaet et al., 2013). L-1-pyrene-alanine that contains the same pyrene group as N-(1-pyrene)-mal strongly reduced the affinity to EF-Tu•GTP when charged on tRNA<sup>Pro</sup> (Doi et al., 2007). Additionally, we have included Atto655-IA as an estimate for the size-limit of fluorophores that can be successfully incorporated. We observed incorporation of all fluorescent probes into dipeptides. The presence of unlabeled Cys-tRNA reduced the amplitude of dipeptides for Cys-tRNA labeled with BodipyTMR-C<sub>5</sub>-mal and N-(1-pyrene)-mal (Figure 2-24 C&D). From the initial slopes of the titrations in Figure 2-24, we have calculated that 10 to 30 molecules of EF-Tu•GTP yielded one molecule of dipeptide involving different fluorescent probes. The small excess of EF-Tu•GTP required in terms of

## Results

concentrations for incorporation of labeled Cys-tRNA allowed us to proceed to the incorporation of fluorescent probes at internal positions during single-turnover translation of the model protein HemK.

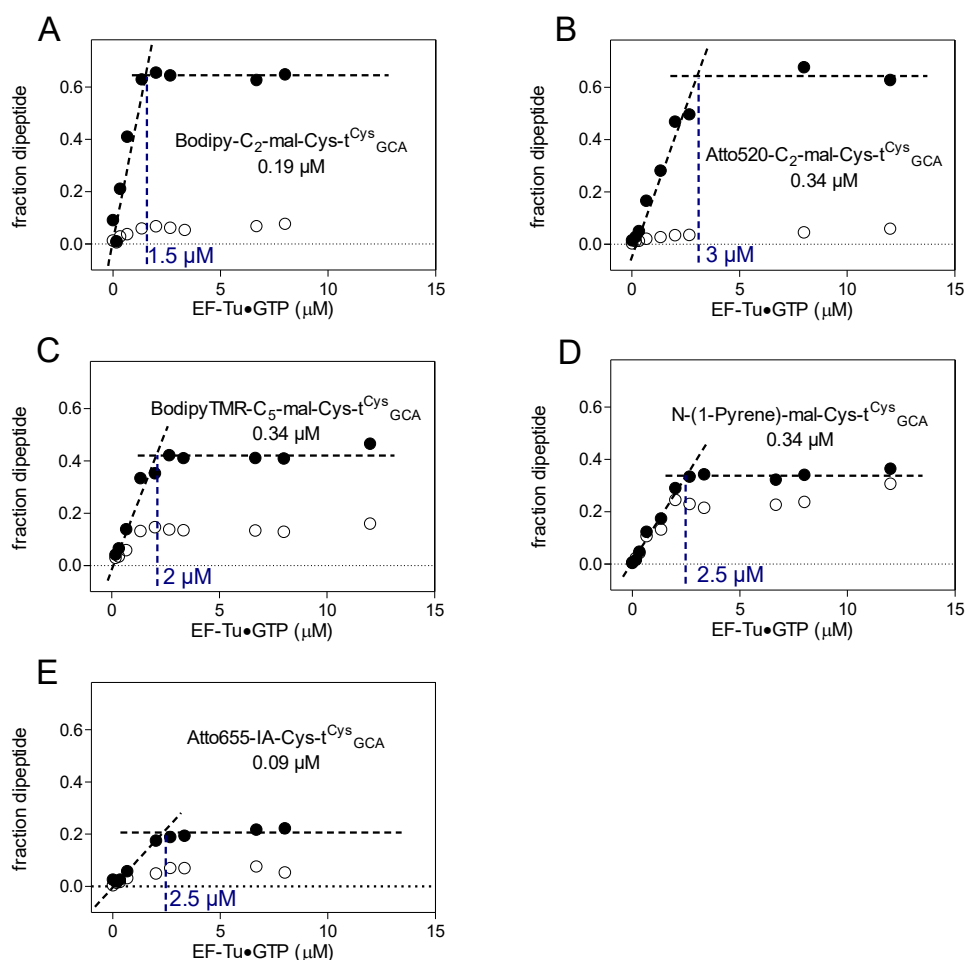


Figure 2-24: Dipeptide formation with fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub> at increasing concentrations of purified EF-Tu•GTP binary complex at 25°C in TAKM<sub>7</sub> buffer. Preformed ternary complexes were mixed with 70S IC programmed with UGC codon in the A site (5 pmol in final volume of 30 μL, or 0.17 μM). Dipeptide formation was allowed to proceed for 2 min. Unreacted f<sup>3</sup>H]Met and both unlabeled and fluorescence-labeled f<sup>3</sup>H]Met-Cys dipeptides could be clearly separated by HPLC on a RP-8 column. The fraction of unlabeled and fluorescence-labeled f<sup>3</sup>H]Met-Cys dipeptides are shown in open symbols and closed symbols, respectively.

Translation reactions were carried out with total tRNA from *E. coli* that was aminoacylated with all aa except for cysteine and supplemented by the respective fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub>. We were able to incorporate BodipyFL-C<sub>2</sub>-mal into position 4 of HemK 1-70 aa (Figure 2-25 A). However, we could not detect any incorporation of the same fluorophore into position 34. On the other hand,

## Results

Bodipy576/589-C<sub>2</sub>-ε-NH<sub>2</sub>-lysine was successfully incorporated into position 34 (Mittelstaet et al., 2013). Upon incorporation of BodipyFL-C<sub>2</sub>-mal into position 4, we could detect the build-up of fluorescence-labeled full-length product with an average rate of  $1.4 \pm 0.3$  aa s<sup>-1</sup> (Figure 2-24 A). Translation of the same mRNA with N-terminal BodipyFL-C<sub>2</sub>-Met and total tRNA aminoacylated with all aa yielded full-length product at a rate of  $1.7 \pm 0.1$  aa s<sup>-1</sup>. An incorporation efficiency of 50% could be estimated from the comparison of the fluorescence intensities between the N-terminal BodipyFL and the same probe incorporated at position 4.

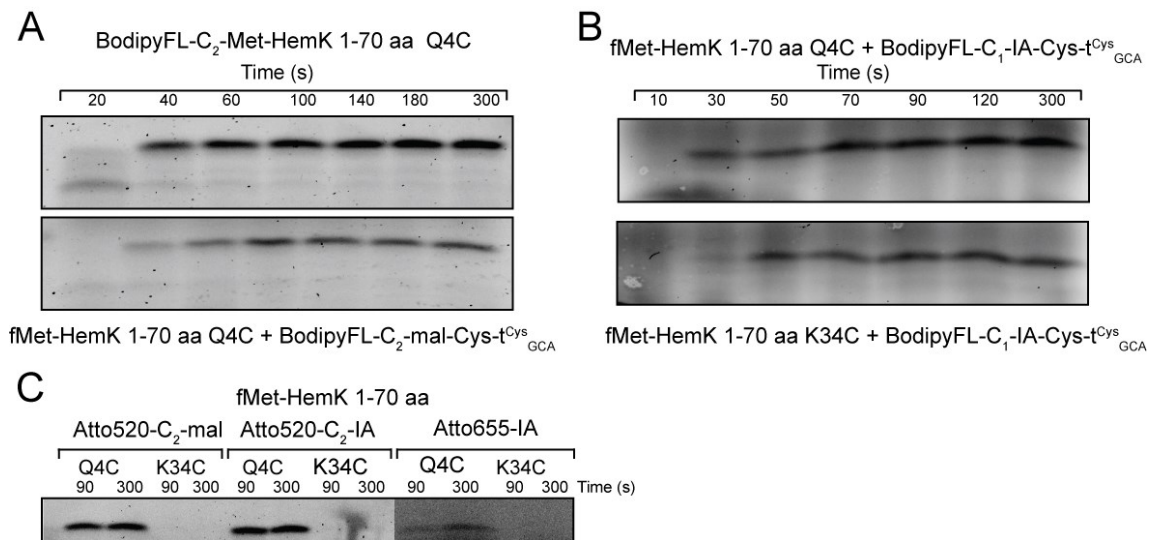


Figure 2-25: Incorporation of fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub> during *in vitro* translation of HemK 1-70 aa at 37°C. For incorporation of fluorescent probes, single-turnover translation was performed with an identical concentration of 70S IC (20 nM), ternary complexes (15 μM), and EF-G (3 μM) at 37°C. The excess of EF-Tu•GTP and fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub> varied for different experiments. In this figure, fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (200 nM) were used together with EF-Tu•GTP (100 μM). **A** (Top panel) 70S IC•BodipyFL-C<sub>2</sub>-Met-tRNA<sup>fMet</sup> initiated on mRNA HemK 1-70 aa Q4C was mixed with aa-tRNA aminoacylated with all aa. (Bottom panel) 70S IC•fMet-tRNA<sup>fMet</sup> initiated on mRNA HemK 1-70 aa Q4C was mixed with aa-tRNA aminoacylated without cysteine and supplemented by BodipyFL-C<sub>2</sub>-mal-Cys-tRNA<sup>Cys</sup><sub>GCA</sub>. **B** 70S IC•fMet-tRNA<sup>fMet</sup> were initiated on mRNA HemK Q4C 1-70 aa (Top panel) or on mRNA HemK K34C 1-70 aa (Bottom panel). 70S IC were mixed with aa-tRNA aminoacylated without cysteine and supplemented by BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub>. All samples in A or B were separated on the same gel in order to allow comparison by densitometry. **C** 70S IC•fMet-tRNA<sup>fMet</sup> initiated on mRNA HemK 1-70 aa Q4C or mRNA HemK 1-70 aa K34C were mixed with aa-tRNA aminoacylated without cysteine and supplemented by Cys-tRNA<sup>Cys</sup><sub>GCA</sub> labeled with the indicated fluorophores.

To our surprise, we found that BodipyFL-C<sub>1</sub>-IA could be incorporated into both position 4 and 34 (Figure 2-25 B). During incorporation of BodipyFL-C<sub>1</sub>-IA at positions 4 and 34, full-length product was formed at an average rate of approximately  $1.5$  aa s<sup>-1</sup> that was consistent with the rate of full-length product

## Results

formation in the presence of total aa-tRNA aminoacylated with all aa (Figure 2-25 B and Figure 2-25 A Top Panel). The fluorescence intensity of the full-length band at the endpoint of translation indicated an incorporation level of 65% for BodipyFL-C<sub>1</sub>-IA at position 34 relative to position 4. Similarly, the incorporation of the sterically more challenging fluorophores Atto520-C<sub>2</sub>-mal, Atto520-C<sub>2</sub>-IA and Atto655-IA was detectable only at position 4 (Figure 2-25 C). Atto520-C<sub>2</sub>-IA and Atto520-C<sub>2</sub>-mal were incorporated to a comparable level. During incorporation of the large probe Atto655-IA, synthesis of full-length product was not completed after 90 s indicating a translation rate of less than 0.8 aa s<sup>-1</sup>.

We investigated the fluorescence change of BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> during incorporation into position 34 of HemK 1-70 aa (Figure 2-26 A). The change of the fluorescence signal over time could be fitted to a model in which a delay precedes a single exponential increase. The reciprocal of the delay time followed a similar concentration dependence as extracted from the endpoint intensities of the full-length band of HemK K34C 1-70 aa translated in the presence of increasing concentrations of BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (Figure 2-26 B, Table 2-6), i.e. the shorter the delay time, the more fluorescence-labeled product was formed. The decrease in the delay time reflected the saturation of ribosomes during decoding at position 34 with the fluorescence-labeled ternary complex. The exponential rate,  $k_{app}$  decreased at increasing concentrations of fluorescence-labeled tRNA (Figure 2-26 C, Table 2-6). The finding that the rate of the exponential phase decreased with increasing concentration of BodipyFL-C<sub>1</sub>-IA was surprising and we were interested to determine the molecular event that was underlying the fluorescence change in Figure 2-26 A.

## Results

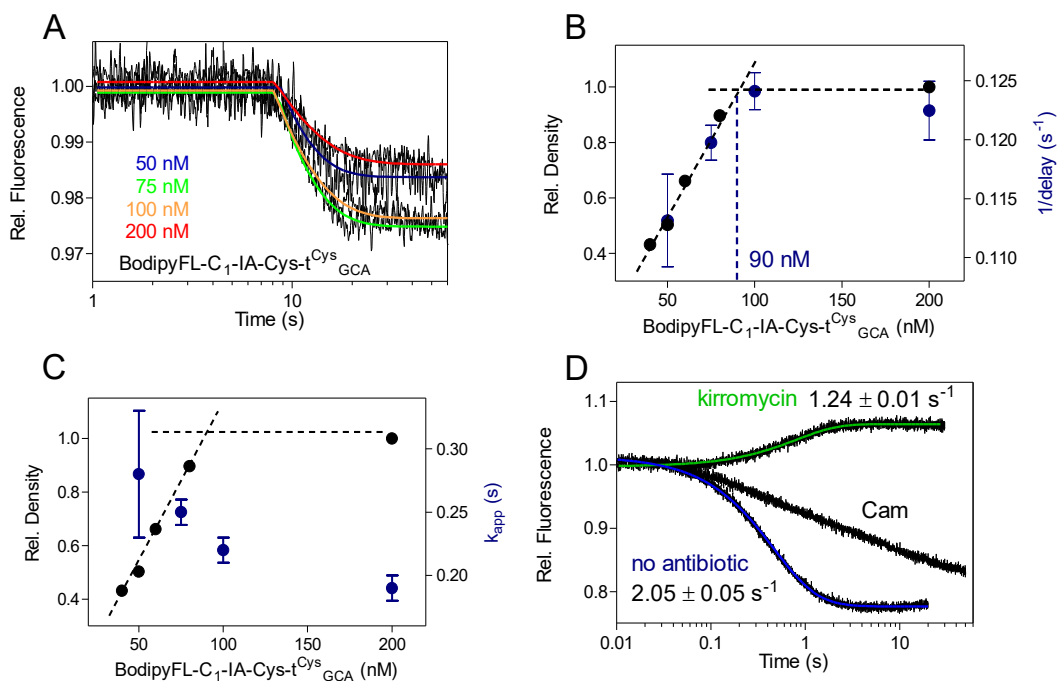


Figure 2-26: Fluorescence change of BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> upon decoding and accommodation in the A site. **A** Change in fluorescence monitored in a stopped-flow apparatus during single-turnover translation of HemK 1-70 aa K34C in the presence of increasing concentrations of BodipyFL-C<sub>1</sub>-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> and constant EF-Tu•GTP (75 μM). The fluorescence signal was fitted to a model with a delay followed by an exponential decrease. **B** Relative incorporation efficiency of BodipyFL-C<sub>1</sub>-IA obtained from densitometric analysis of the full-length band of HemK 1-70 aa K34C at the endpoint of translation (left y-axis). The concentration dependence of the inverse of the delay time in A is shown on the right y-axis. **C**. Comparison of the relative incorporation efficiency with the concentration dependence of the exponential rate  $k_{app}$  from A. **D** Fluorescence change after rapid mixing of 70S IC•fMet-tRNA<sup>fMet</sup> initiated on mRNA fMet-Cys (100 nM) with EF-Tu•GTP•BodipyFL-C<sub>1</sub>-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (20 nM), and EF-Tu in total (10 μM) at 37°C in TAKM<sub>7</sub> buffer. Kirromycin and chloramphenicol (Cam) were used at final concentrations of 0.1 mM and 2 mM, respectively.

Table 2-6: Parameter obtained by fitting the fluorescence change during translation of HemK 1-70 aa K34C in presence of increasing concentrations BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (\*aa-tRNA) (Figure 2-26 A) to a model with a delay preceding a single exponential increase.

*aa-tRNA (nM)	delay (s)	1/delay (s <sup>-1</sup> )	rate (s <sup>-1</sup> )	1/rate (s)
50	8.8 ± 0.3	0.113 ± 0.004	0.28 ± 0.05	3.6 ± 0.5
75	8.4 ± 0.1	0.120 ± 0.001	0.25 ± 0.01	4.1 ± 0.2
100	8.1 ± 0.1	0.124 ± 0.002	0.22 ± 0.01	4.6 ± 0.2
200	8.2 ± 0.2	0.122 ± 0.002	0.19 ± 0.01	5.2 ± 0.3



## Results

To better assign the underlying reactions, we simplified the experimental system to a single decoding event. We prepared 70S IC programmed with mRNA coding for fMet-Cys and added EF-Tu•GTP•BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> to these complexes. Upon binding of these ternary complexes to the A site, we observed a decrease in fluorescence intensity (Figure 2-26 D). The signal change observed in Figure 2-26 A & D could have reflected an event during decoding, accommodation or peptide bond formation. To distinguish between the different steps, we selected the antibiotic kirromycin to inhibit the dissociation of the aa-tRNA from EF-Tu•GDP after GTP hydrolysis and to stabilize the distorted conformation of the tRNA in the A/T state (Rodnina et al., 1995; Rodnina et al., 1994; Stark et al., 1997). In the presence of kirromycin, we observed a fluorescence increase with a similar rate as the decrease of fluorescence in absence of the antibiotic. We concluded that a step during initial selection leading to GTPase activation was limiting the rate of later processes. This interpretation would be consistent with a slow rate of GTP hydrolysis during decoding of Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (Liu et al., 2011). The opposite direction of the fluorescence change in presence of kirromycin indicated that the decrease in fluorescence intensity occurred either during the release of the fluorescence-labeled tRNA from EF-Tu•GDP, accommodation or peptide bond formation. Chloramphenicol (Cam) binds to two separate locations on the 50S subunit (Lessard and Pestka, 1972). A binding site with higher affinity was identified in the A site loop of the PTC (Schlunzen et al., 2001). The lower affinity binding site could be mapped to nucleotides in the ribosomal exit tunnel in close proximity to the PTC (Long and Porse, 2003). The second binding site additionally overlapped with the binding sites of macrolide antibiotics. In the presence of Cam, we observed a much slower rate of fluorescence decrease. We concluded that the fluorescent changes most likely reported on the step of accommodation into the PTC or the subsequent process of peptide bond formation. The ambiguous mode of action of Cam did not allow us to pin-point the fluorescence change to a specific step in the PTC; however, from the kinetic point of view, the fluorescence change of BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> is limited by the early steps of decoding. The decrease in  $k_{app}$  at increasing concentration of BodipyFL-C<sub>1</sub>-IA Cys-tRNA<sup>Cys</sup><sub>GCA</sub> is consistent with kinetic scheme involving a slow step preceding the concentration dependent fluorescence change (Fersht, 1999). According to this scheme, at saturating concentration of BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub>, the value of  $k_{app}$  would be expected to approach the limiting rate of the first kinetic step calculated from  $1/\text{delay}$  of  $0.122 \pm 0.002 \text{ s}^{-1}$  (Figure 2-26 B & C, Table 2-6).

## Results

Alternatively, the hyperbolic dependence of the change in fluorescence change during translation of HemK 1-70 aa K34C (Figure 2-26 C) might be interpreted in the light of de-synchronization of elongating ribosomes during a single-turnover *in vitro* translation. At low concentration of fluorescence-labeled tRNA, the ribosomes translating with the highest speed have a higher chance to recruit the fluorescence-labeled tRNA (first come-first serve). An average rate of translation elongation for the ribosomes participating in substrate selection of the fluorescence-labeled tRNA at position 34 could be calculated from the sum of the delay time and the inverse of the exponential rate (Table 2-6). We obtained an average rate of translation elongation of  $2.7 \pm 0.2$  aa s<sup>-1</sup> at 50 nM of fluorescent tRNA. The rate decreased to  $2.1 \pm 0.4$  aa s<sup>-1</sup> as calculated from the sum of the delay times at saturating concentration of labeled tRNA. This value matched the average translation rate of  $1.7 \pm 0.1$  aa s<sup>-1</sup> determined for HemK 1-70 aa Q4C in the presence of all aa-tRNA (Figure 2-25 A). Importantly, we could not exclude that a step occurring sequential after the step monitored by the fluorescence decrease of Cys-tRNA<sup>Cys<sub>GCA</sub></sup> might be limiting the translation elongation in the presence of BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys<sub>GCA</sub></sup>.

Encouraged by the successful incorporation of BodipyFL-C<sub>1</sub>-IA into position 34 of HemK, we investigated the incorporation of the non-fluorescent acceptor probe QSY-35 into the same position. The N-terminal domain of HemK folds independently of the C-terminal domain (Jörg Mittelstätt & Markus Jäger, unpublished data). We monitored the incorporation of QSY-35 indirectly by the decrease in fluorescence intensity of the N-terminal BodipyFL-C<sub>2</sub>-Met due to fluorescence resonance energy transfer (FRET) to QSY-35. From the crystal structure of HemK, the separation of the N-terminus and positions 34 was approximated to a distance of 15 Å (Graille et al., 2005) and the Förster radius for BodipyFL/QSY-35 was calculated as 42 Å. HemK 1-70 aa K34C was translated in the presence of increasing concentrations of QSY-35-Cys-tRNA<sup>Cys<sub>GCA</sub></sup>, released at the endpoint of the translation reaction with Pmn and the nascent chains were allowed to fold. We recorded the emission spectra of the N-terminal BodipyFL-C<sub>2</sub>-Met in the released chains (Figure 2-27 A).

## Results

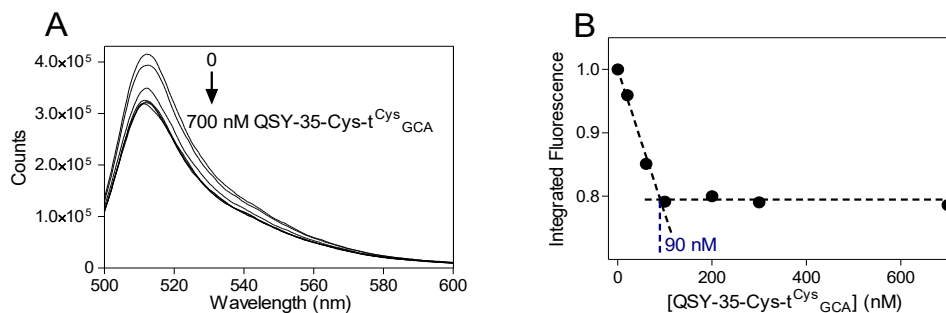


Figure 2-27: Measurement of FRET between the N-terminal BodipyFL-C<sub>2</sub>-Met and QSY-35 at position 34 of released chains of HemK 1-70 aa K34C. Single-turnover translation was performed with increasing concentrations of QSY-35-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> and constant EF-Tu•GTP (75  $\mu$ M) for 5 min at 37°C. Nascent chains were released with Pmn (1.5 mM) and incubated for 5 min at 37°C. **A** BodipyFL-C<sub>2</sub>-Met of released chains was excited at 473 nm and the emissions spectra was recorded. **B** Integrated intensities of the spectra in A as a function of the concentration of QSY-35-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> used during translation.

The incorporation of QSY-35 led to a 20% decrease in fluorescence intensity (Figure 2-27 B). The inflection point of the titration indicated that a similar 4.5-fold excess of fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub> over translating ribosomes was required for optimal incorporation as previously observed for BodipyFL-C<sub>1</sub>-IA (Figure 2-26 B). Based on the assumption that all released chains adopted the folded state and after correction for ratio of BodipyFL-C<sub>2</sub>-Met-tRNA<sup>fMet</sup> that was not incorporated into the 70S IC (19% of BodipyFL-C<sub>2</sub>-Met-tRNA<sup>fMet</sup>), we have calculated a minimal level of incorporation of 30% for QSY-35.

### 2.3.5 Context dependence for incorporation of fluorescent probes

The results presented in section 2.3.4 indicated that fluorescence-labeled Cys-tRNA were efficiently delivered by EF-Tu•GTP to the A site and that the labeled tRNA proceeded through the steps of decoding on the 30S subunit and accommodation into the PTC on the 50S subunit at a rate that was compatible with the overall rate of translation elongation during single-turnover translation of the model protein HemK. However, the finding that several large fluorophores could only be incorporated into position 4 but not position 34 of the HemK pointed to the peptidyl transfer reaction as the limiting step (Figure 2-25 C). The efficiency of incorporation of fluorescent probes appeared to be modulated by the length of the peptide attached to P site tRNA. In addition, we could not detect any significant incorporation of fluorescent probes into position R8C of HemK (not shown). The difficulty for incorporation into position 8 of HemK indicated that not only the length of the peptidyl-tRNA but also the aa

## Results

sequence upstream of the incorporation site might modulate the incorporation efficiency of fluorescent probes. Similarly, the Hohsaka group reported a more efficient incorporation at N-terminal positions for a structurally diverse set of nc-aa (Abe et al., 2010; Matsubara et al., 2013; Shozen et al., 2009; Watanabe et al., 2007). For another set of nc-aa, a strong dependence on the aa sequence upstream of the incorporation was described (Ezure et al., 2014; Hohsaka et al., 2004; Iijima and Hohsaka, 2009; Shozen et al., 2009; Watanabe et al., 2007).

We investigated the influence of the upstream aa residue on the incorporation efficiency of fluorescent probes. We selected HemK 1-70 aa K34C as model system and created a library of all aa substitutions at position 33. This library allowed us to investigate the incorporation of fluorescent probes under experimental conditions where the nascent peptide attached to the P-site tRNA spans the entire exit tunnel. Therefore, a similar dependency on the upstream aa might be expected for incorporation sites even more distant from the N-terminus than position 33. During translation of the codons after the incorporation site, the fluorescent probe would be elongated across the complete ribosomal exit tunnel. In this sense, restrictions imposed by the fluorescent probes on the elongation steps after the successful incorporation might be detected by the accumulation of novel translation intermediates. We observed a different dependency on the upstream aa for the incorporation of fluorophores with increasing size (Figure 2-28 A). Incorporation of the smaller probe IA-NBD into full-length HemK 1-70 aa could be observed for the majority of the aa substitutions. The incorporation efficiency of IA-NBD could be correlated with the sterical properties of the side chains of the upstream aa residue (Figure 2-28 B). Small side chains allowed for highly efficient incorporation. Substitutions at C- $\beta$  led to a decrease in incorporation efficiency from the smaller hydroxyl group of threonine to the bulkier methyl group of valine. No peptide bond formation was observed for upstream Ile, Leu, and Pro. Similarly, the presence of an upstream Leu at position R8C of HemK might have been sufficient to preclude incorporation of fluorescent probes at this N-terminal position. Longer, non-branched aa side chains allowed incorporation at intermediate levels. Surprisingly, all aromatic side chains favored the incorporation of IA-NBD despite their larger substituents.

## Results

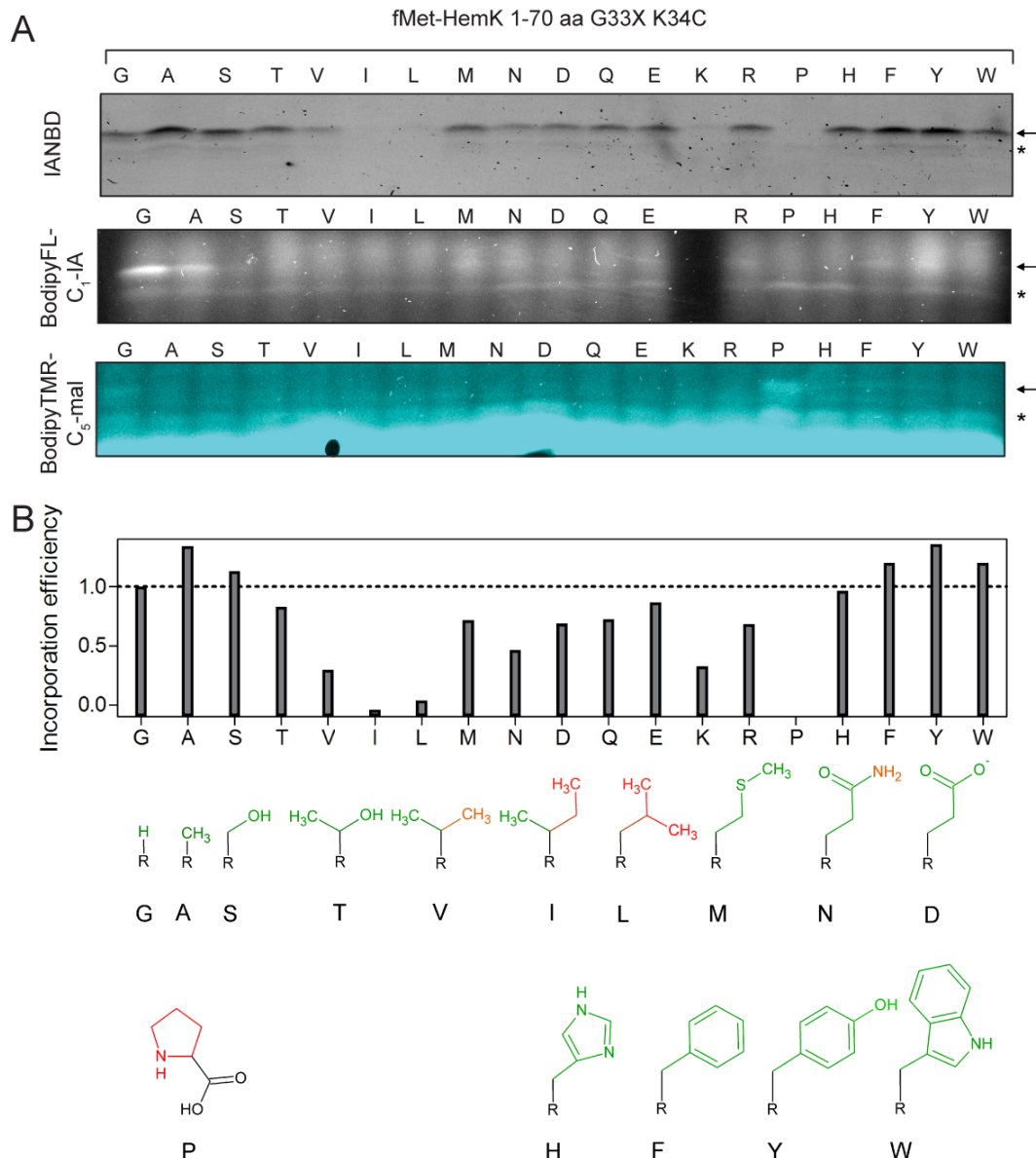


Figure 2-28: Dependence of the incorporation efficiency on the aa upstream of the incorporation site. **A** Single-turnover translation was performed with fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (120 nM) and EF-Tu•GTP (45 μM) for 5 min at 37°C. EF-G labeled at Cys196 with Alexa674-mal (20 nM) was included in the translation reaction as loading control. Full-length products are indicated by arrows and shorter peptides are marked by asterisks. **B** Quantification of the incorporation efficiency of IA-NBD from A. The incorporation was normalized to the efficiency observed for upstream Gly. The substituents of aa side chains are color-coded according to their contribution to the incorporation efficiency.

The incorporation efficiency of the larger BodipyFL-C<sub>1</sub>-IA and BodipyTMR-C<sub>5</sub>-mal exhibited a stronger context dependency (Figure 2-28 A). Unfortunately, high background intensities resulting from hydrolysis products of the fluorescence-labeled Cys-tRNA made quantification difficult. For BodipyFL-C<sub>1</sub>-IA, higher levels of incorporation into full-length product could only be observed for upstream Gly and

## Results

Ala substitutions. Similarly, for BodipyTMR-C<sub>5</sub>-mal, full-length product was only formed for upstream Gly and Pro. In summary, the presence of upstream Gly and in certain cases Ala and Pro residues might allow to increase the incorporation efficiency of larger fluorophores. Additionally, we observed the accumulation of a fluorescent peptide of smaller molecular weight than HemK 1-70 aa for all constructs during incorporation of the Bodipy fluorophores and for certain upstream substitutions with IA-NBD (Figure 2-28 A). These truncated peptides could have resulted from inefficient translocation following incorporation or inefficient formation of the subsequent peptide bond. Notably, for the incorporation of D-amino acids, efficient translocation after incorporation of D-aa but inefficient peptidyl transfer to subsequent L-aa was reported (Englander et al., 2015).

The Hohsaka group reported increased incorporation of large nc-aa into maltose binding protein at two positions with an upstream di-Gly motif (Ezure et al., 2014; Iijima and Hohsaka, 2009; Shozen et al., 2009; Watanabe et al., 2007). Incorporation of Atto520 was reported in a Ser and Gly-rich linker sequence and the large tetramethyl-rhodamine (TAMRA) fluorophore could be incorporated downstream of a repeat of four Gly residues in a membrane protein (Jeong et al., 2013; Ohtsuka et al., 2011). We realized that several small residues upstream of the incorporation site might have the potential to increase the flexibility of the peptidyl-donor substrate and thereby improve the incorporation of sterically more challenging fluorophores. We performed scanning mutagenesis of the residues 30 to 32 of HemK with Gly and prepared all combination of multiple Gly substitutions. We analyzed the incorporation efficiency of fluorophores with increasing size as a function of the number and position of upstream Gly residues (Figure 2-29). The small probe IA-NBD was incorporated to a similar extent into all constructs with upstream Gly substitutions (Figure 2-29 B). The incorporation of BodipyFL-C<sub>1</sub>-IA and Bodipy507/545-IA was enhanced in constructs with an additional Gly at positions -4 to -2 relative to the incorporation site. For the constructs with a single Gly substitution, Gly at position -2 showed the largest increase in incorporation efficiency. The incorporation of the larger probes BodipyTMR-C<sub>5</sub>-mal and Atto520-C<sub>2</sub>-IA was only enhanced when two upstream residues were substituted for Gly.

## Results

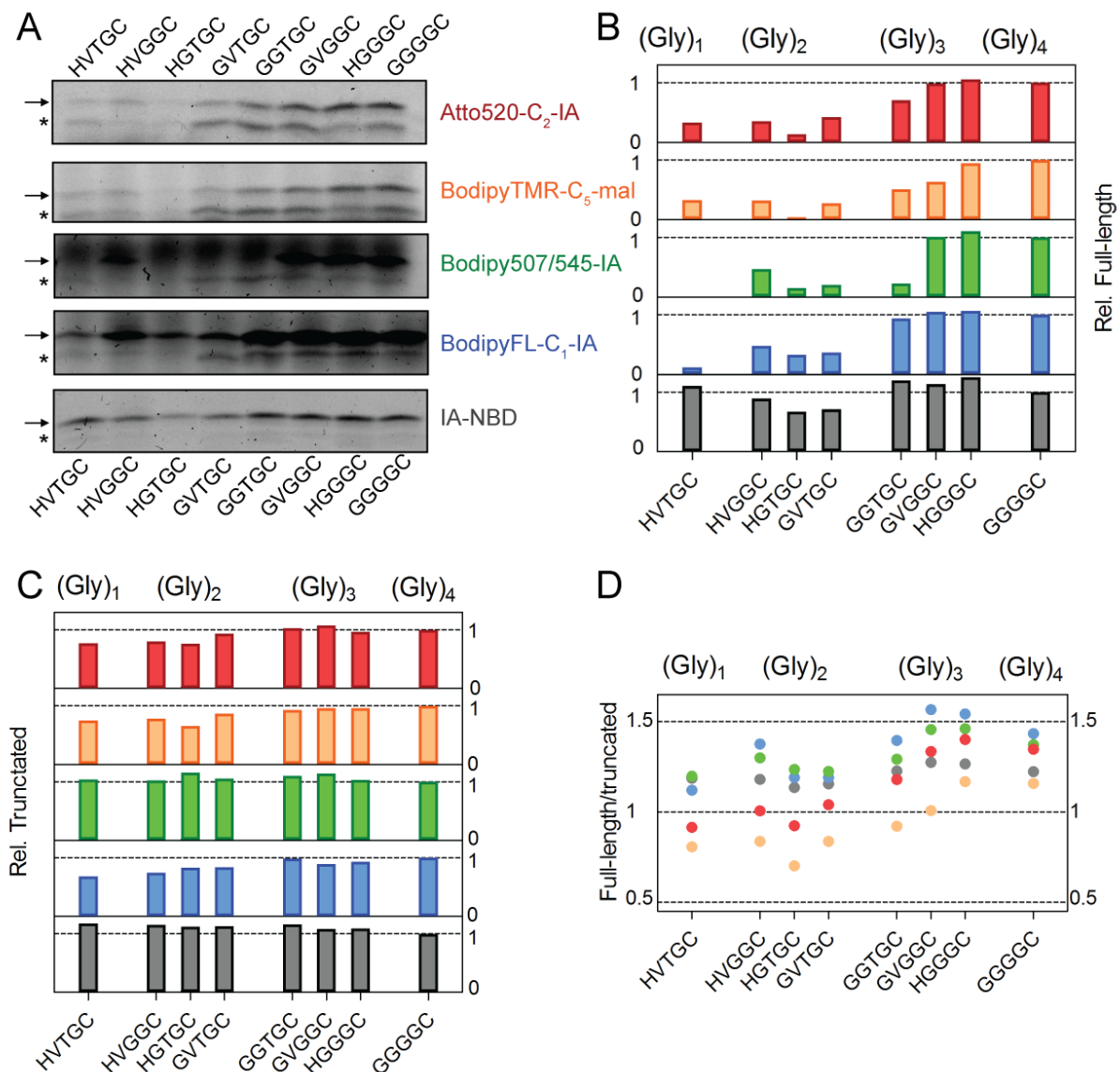


Figure 2-29: Influence of upstream Gly residues on the incorporation of fluorescent probes into HemK 1-70 aa K34C. **A** Single-turnover translation was performed with fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (120 nM) and EF-Tu•GTP (45 μM) for 5 min at 37°C. EF-G Cys196-Alexa674-mal (20 nM) was included as loading control. The sequence HVTGC corresponds to the residues at positions -4 to -1 upstream of the incorporation site of HemK K34C. Full-length product is indicated by arrows and truncated peptides are marked by asterisks. The same color code is used for the fluorophores. **B** Full-length bands were quantified densitometrically and normalized relative to the construct with the sequence GGGGC. **C** Relative quantification of the truncated peptides was performed in the same manner described for the full-length product in B. **D** The ratio of the full-length product to the truncated fragment was calculated from the background corrected raw intensities.

Among the constructs with two Gly substitutions, higher incorporation of all probes larger than IA-NBD was observed for constructs with a longer, continuous stretch of

## Results

Gly residues adjacent to the incorporation site. No additional increase in incorporation efficiency could be detected by the substitution of the last residue from HGGGC to GGGGC. The efficiency of incorporation of the intermediate-sized probes BodipyFL-C<sub>1</sub>-IA and Bodipy507/545-IA was enhanced 10-fold by altering the upstream sequence from HVTG to HGGG. A smaller three-fold increase was observed for the larger fluorophores BodipyTMR-C<sub>5</sub>-mal and Atto520-C<sub>2</sub>-IA (Figure 2-29 B).

The incorporation of fluorescent probes could not be additionally enhanced *in trans* by ribosomes carrying truncations in the loops of ribosomal protein L4 ( $\Delta$ 84-87 aa) and L22 ( $\Delta$ 63-64 aa) for any of the constructs with upstream Gly substitutions (data not shown). The loops of L4 and L22 form the narrowest constriction in the ribosomal exit tunnel (Figure 2-30); the removal of these loops abolished stalling of SecM and ErmCL arrest peptides (Ito and Chiba, 2013).

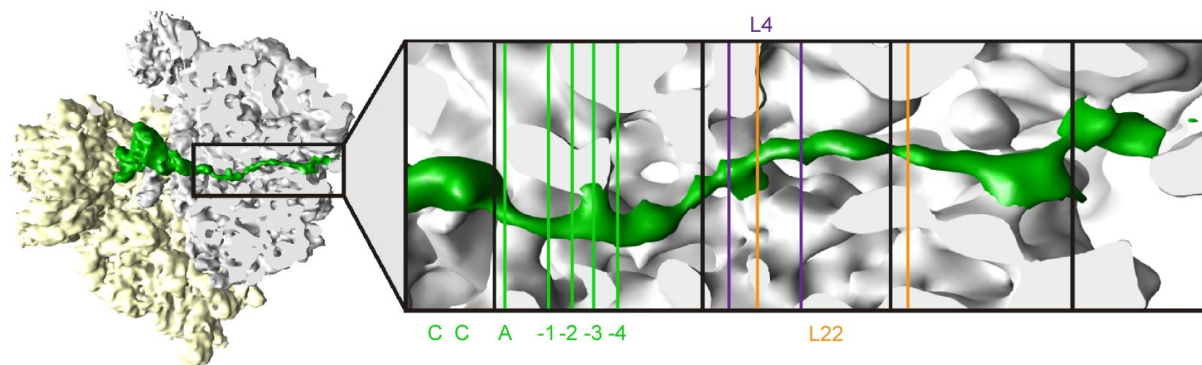


Figure 2-30: Reconstruction of the density of the peptidyl-tRNA in the P site and the nascent chain in the ribosomal exit tunnel at 5.6 Å resolution by cryo-electron microscopy. Figure showing the SecM stalling peptide in the ribosomal exit tunnel was modified from (Bhushan et al., 2011). The approximate location of the CCA end of the tRNA, the residues at position -4 to -1 that modulated the incorporation of fluorescent probes, and the constriction formed by the ribosomal proteins L4 and L22 are indicated.

Depending on the aa sequence of the nascent chain, residues at positions -10 to -16 in the nascent chain counting from the aa in the A contact the L4/L22 constriction site. The finding that the incorporation efficiency could not be enhanced further by ribosomes carrying deletions in L4 and L22 indicated that widening of the constriction site was not a promising strategy to enhance the peptidyl transfer involving nc-aa in the A site.

For the set of constructs with Gly substitutions, we also detected fluorescent peptides of shorter length than 70 aa (Figure 2-28 A, Figure 2-29 A&C). Incorporation of larger fluorophores led to a higher level of truncated products. The presence of Gly



## Results

residues upstream of the incorporation site modulated the level of these shorter fragments, but to a smaller extent than observed for the level of full-length bands (Figure 2-29 C). The fraction of the shorter peptides relative to the full-length product was lower for constructs with two or three Gly residues adjacent to the incorporation site (Figure 2-29 D). Therefore, the partitioning between accumulation of truncated or full-length products seemed to follow a similar dependency on the sequence upstream of the incorporation site. During translation of HemK 1-70 aa HGGG K34C in the presence of BodipyTMR-C<sub>5</sub>-mal-Cys-tRNA<sup>Cys</sup><sub>GCA</sub>, the truncated fragment and full-length product accumulated at the same rate of  $\sim 0.7$  aa s<sup>-1</sup> (data not shown). EF-P was reported to bind to the empty E site and alleviate pausing at Pro and poly-Pro stretches by positioning of the P-site tRNA (Doerfel et al., 2013). We investigated the ability of saturating concentrations of EF-P to enhance the incorporation of fluorescence-labeled Cys-tRNA. We could not detect any difference in incorporation efficiency for fluorescent probes with increasing size (data not shown). Similarly, the truncated fragments accumulated to a similar extent in presence of EF-P.

The addition of a benzoic acid substituent to the xanthene structure found in Atto520 leads to the class of rhodamine fluorophores that are suitable for single-molecule techniques due to their bright fluorescence and high photostability. Fluorescent probes with rhodamine structures were reported to be inefficiently incorporated into nascent chains (Abe et al., 2010; Kang et al., 2007; Traverso et al., 2003). We analyzed the incorporation of the rhodamine fluorophore Atto488-mal and the larger carborhodamine Atto633-IA (structure not published) into position K34C for the set of constructs with upstream Gly substitution (Figure 2-31). We could detect the formation of the truncated products for Atto488-mal and Atto633-IA that had the same length as the truncated fragments shown in Figure 2-29 A. Additionally, the short fragments followed a similar dependency on the upstream sequence as observed for the fluorophores analyzed in Figure 2-29 C.

## Results

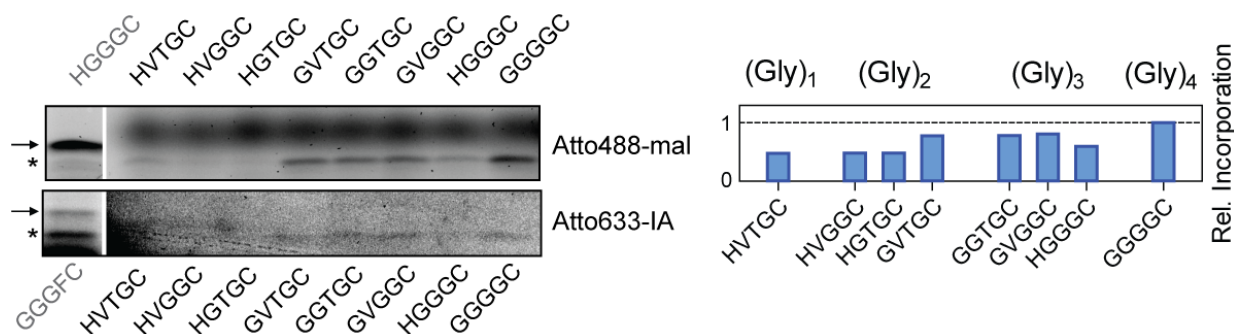


Figure 2-31: Incorporation of Atto488-mal and Atto633-IA into HemK 1-70 aa K34C constructs with upstream Gly substitutions. Single-turnover translation was performed with fluorescence-labeled Cys-tRNA<sup>Cys</sup><sub>GCA</sub> (120 nM) and EF-Tu•GTP (45 μM) for 5 min at 37°C. EF-G Cys196-Alexa674-mal (20 nM) was present in the translation reaction as loading control. Reference samples translated in the presence of BodipyTMR-C<sub>5</sub>-mal-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> that were separated on the same gel are highlighted in gray. Full-length products are indicated by arrows and shorter fragments are marked by asterisks. The intensity of shorter fragments that accumulated in presence of Atto488-mal-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> were normalized to the intensity of the construct with sequence GGGGC. The intensities of bands observed for Atto633-IA-Cys-tRNA<sup>Cys</sup><sub>GCA</sub> was too low for the relative quantification.

Unfortunately, we could not detect accumulation of full-length products labeled with Atto488-mal and Atto633-IA, while for the even larger Atto655-IA, we could not detect accumulation of any translation product. The observation that the fluorophores with rhodamine structures were incorporated into truncated peptides might represent a promising starting point for further optimization of the incorporation of larger fluorophores. The absence of any translation product for incorporation of Atto655-IA at position 34 might mark the size-limit for fluorescent probes that can be incorporated at positions distant from the N-terminus.

## Results

### 2.4 Codon-specific elongation rates modulate co-translational folding

In this section, we report the influence of synonymous codon usage on the co-translational protein folding *in vivo* and *in vitro* for the model protein  $\gamma$ -B-crystallin (GBC) from *Bos taurus* (*B. taurus*). The work was performed in collaboration with Sujata Jha and Anton A. Komar (Cleveland State University) as well as with Florian Buhr and Harald Schwalbe (Johann Wolfgang Goethe University Frankfurt).

#### 2.4.1 Harmonization of synonymous codon usage of GBC

Members of the  $\alpha$ -,  $\beta$ -, or  $\gamma$ -crystallin families form the major structural component of eye lens fiber cells (Harding and Dilley, 1976). After formation of the eye lens during mammalian development, the long-lived fiber cells lose the nucleus and only limited protein turnover is thought to occur during the lifetime of the cell (Wannemacher and Spector, 1968). Members of the  $\alpha$ - and  $\beta$ -crystallin families form homo- and heterodimers, while members of the  $\gamma$ -crystallin family are monomeric (Wistow et al., 1981). The crystal structure of the 175 aa long GBC from *B. taurus* revealed that the protein consists of two symmetric domains formed from entirely  $\beta$ -sheet structures (Figure 2-32 A). The symmetry originated from the presence of two Greek key motifs in each domain (Wistow et al., 1983).

GBC showed high stability against denaturant-induced and thermal unfolding *in vitro* (Rudolph et al., 1990; Sharma et al., 1990). The C-terminal domain (CTD) contains a greater number of positively charged residues than the N-terminal domain (NTD). At pH 2, the CTD could be destabilized to a larger extent due to electrostatic repulsion than the NTD. At this pH value, unfolding and refolding of GBC proceed through an intermediate with a native-like NTD and an unfolded CTD. The domain formation of the CTD, rather than the formation of the domain interface, was rate-limiting during refolding of GBC (Rudolph et al., 1990). The isolated, recombinant CTD was less stable than within the full-length protein. In contrast, the stability of the isolated NTD was independent of the presence of the CTD (Mayr et al., 1997).

## Results

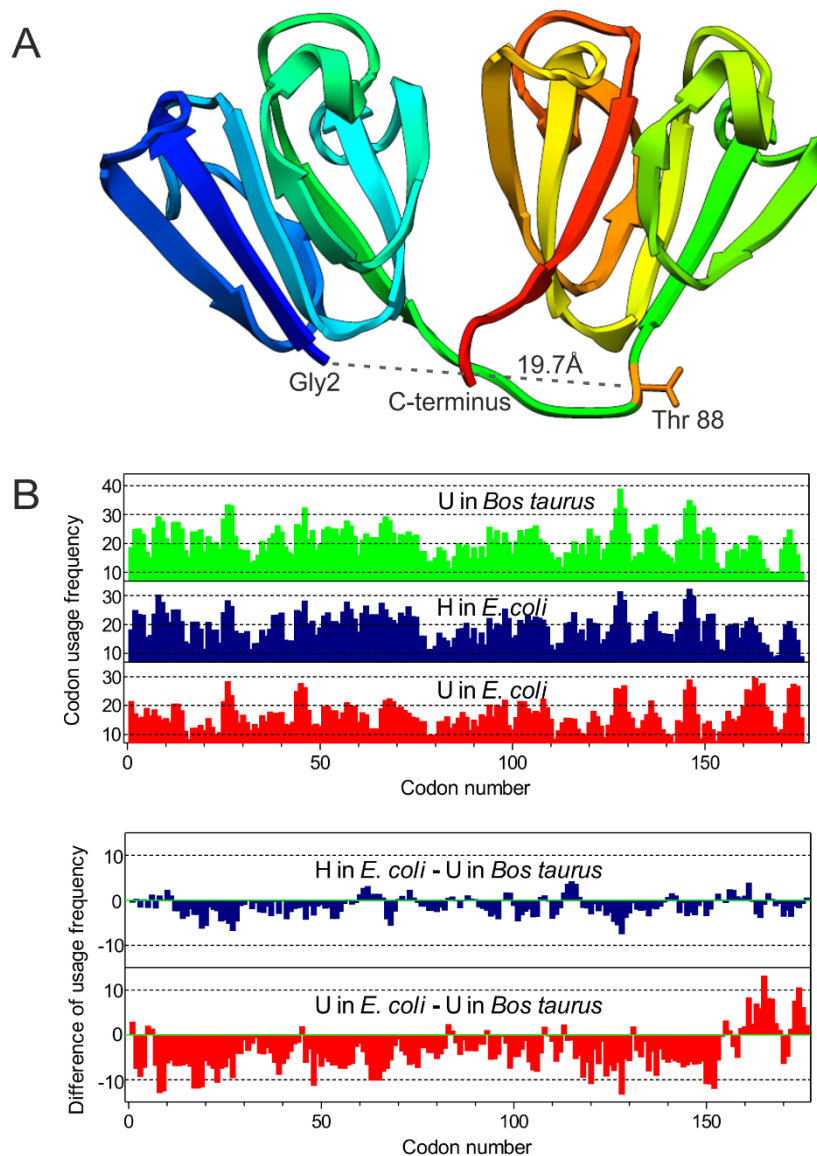


Figure 2-32: Structural properties and codon usage profile of synonymous sequences coding for GBC. **A** Crystal structure showing the domain organization of GBC (pdb 4GCR). In *B. taurus*, the methionine at the N-terminus is cleaved post-translationally; hence the distance between the N-terminus and threonine 88 is approximated by the distance from Gly2 to Thr88. Fluorescent probes were incorporated at the N-terminus and position 88 for monitoring co-translational folding of the NTD by FRET (section 2.4.4). **B** (Top panel) Codon usage frequency along the U sequence of GBC in *B. taurus* (green), and *E. coli* (red), while the codon usage harmonized for *E. coli* in the H sequence is shown in blue. (Bottom panel) Difference in codon usage frequency relative to the U sequence in *B. taurus*.

Mutagenesis of Phe56 at the domain interface significantly destabilized the CTD in the full-length protein (Palme et al., 1997). A stabilization of  $-16 \text{ kJ mol}^{-1}$  was reported for the interdomain interaction compared to only  $-5 \text{ kJ mol}^{-1}$  for the CTD and an overall stability of  $-40 \text{ kJ mol}^{-1}$  of full-length GBC (Mayr et al., 1997). The above

## Results

mentioned rates and stabilities were collectively measured at pH 2. At pH 7, the simultaneous unfolding of both domains occurs in a single transition at high concentration of guanidinium hydrochloride (GdnHCl) (Rudolph et al., 1990; Sharma et al., 1990). Additionally, the recombinant, isolated NTD and CTDs are characterized by a transition midpoint of denaturation at pH 7 of 3 M and 2.2 M GdnHCl, respectively. Under physiological conditions, the NTD is only slightly more stable than the CTD (Mayr et al., 1997). Translation of GBC in a rabbit reticulocyte extract supplemented with bovine tRNA revealed non-uniform rates of translation elongation along the coding sequence. The localization of a set of translation intermediates was correlated with the presence of rare codons in the linker region and the CTD (Komar and Jaenicke, 1995). The hypothesis was formed that co-translational folding of the NTD during the slower synthesis of the CTD might allow for more efficient folding of the CTD by formation of the domain interface (Komar and Jaenicke, 1995).

In the current study, we have prepared a variant of the GBC coding sequence harmonized for the codon usage in *E. coli* denoted as H and compared to the non-harmonized, original bovine sequence denoted as U (Figure 2-32 B). Harmonization of the codon usage between different organisms is trying to select the synonymous codon with closest usage frequency in both organisms (Angov et al., 2008). The process of harmonization maintains both the frequently used and rare codons. Therefore, similar relative elongation rates along the sequence might be retained (Spencer et al., 2012). In contrast, optimization of all codons for the synonymous codon with the highest usage frequency often increased the fraction of insoluble proteins during recombinant expression (Chaney and Clark, 2015; Gustafsson et al., 2004; Rosano and Ceccarelli, 2009). Harmonization of the GBC coding sequence led to an increase in overall codon adaptation index (CAI) for the expression in *E. coli*. The CAI increased from 0.64 to 0.78 upon changing the U to the H sequence; the latter is closer to a CAI value of 0.85 for expression of the U sequence in the native host *B. taurus*.

## Results

### 2.4.2 Sequences of U and H yield differential expression of GBC in *E. coli*

Sujata Jha analyzed the expression of GBC protein from the sequences of U and H in *E. coli* BL21(DE3). The harmonized sequence led a 1.5-fold higher expression level of soluble full-length protein compared to the non-harmonized sequence (Figure 2-33 A). The increase in soluble full-length protein could not be explained by the partitioning of soluble and insoluble protein expressed from both sequence variants. An antibody specific for the C-terminal His<sub>6</sub>-tag of the GBC revealed an increased accumulation of truncated, insoluble fragments during expression of GBC from the U sequence (Figure 2-33 B). The binding of the antibody against the C-terminal His<sub>6</sub>-tag indicated that the fragments at least partially resulted from the degradation of full-length GBC post-synthesis. Analysis by mass spectrometry confirmed that the truncated products originated from GBC and, importantly, that the sequences of U and H led to the expression of GBC with the correct aa sequence without misincorporation. Analysis of the expression of GBC by Western blotting and ELISA applying a polyclonal antibody against GBC revealed a higher level of truncated fragments accumulated in the pellet fraction for the U sequence compared to the H sequence. The results obtained *in vivo* indicated that during expression of GBC from the U sequence a larger fraction of potentially not properly folded protein was subjected to degradation.

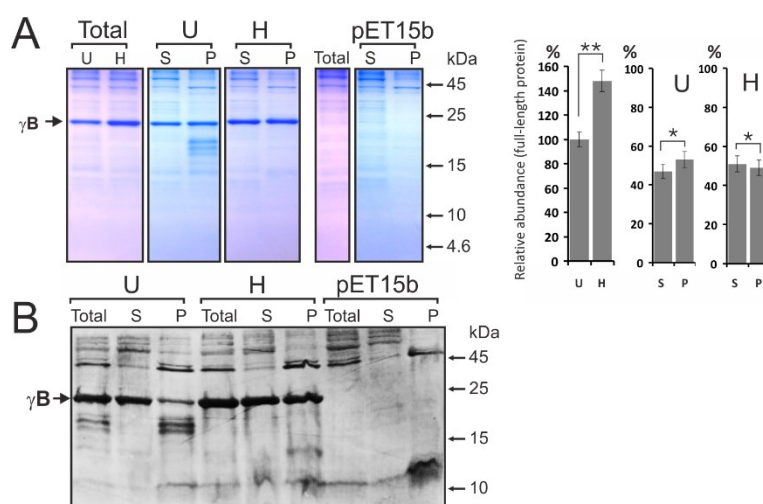


Figure 2-33: Analysis of the recombinant expression of GBC from the sequence of U and H in *E. coli* BL21(DE3). The partition of the total protein into soluble protein (S) and into insoluble protein that can be pelleted (P) was analyzed by Coomassie staining **A** and Western blotting **B** with primary antibody specific for the C-terminal His<sub>6</sub>-tag of GBC. The experiments were provided by Sujata Jha.

## Results

Florian Buhr performed NMR experiments on purified GBC expressed in *E. coli* from the U and H sequences. The NMR spectra of the protein expressed from the H sequence exhibited characteristic chemical shifts of cysteine residues oxidized to disulfides. Upon addition of the reducing agent DTT, the peaks in the spectra of the H protein relaxed to the spectra obtained for GBC expressed from the U sequence. The expression of GBC from the U sequence yielded a protein in a completely reduced state. Cysteine residues at position 15, 18, 22, 32, 41, and 76 in the H protein showed chemical shift differences that indicated the formation of several alternative combinations of disulfide bonds. The heterogeneity in the pattern of disulfides was supported by the analysis of the GBC variants by analytical ion-exchange and reversed-phase chromatography that yielded on single peak for GBC from the U sequence compared to several peaks for the H sequence. GBC extracted from the bovine eye lens also contained disulfide bonds. This notion was supported by the presence of a disulfide bond between residues 18 and 22 in the crystal structure of GBC extracted from the bovine eye lens (Najmudin et al., 1993). On the other hand, the presence of disulfides introduced in GBC during expression of the H sequence in the reducing environment of the bacterial cytoplasm was highly surprising. The results of Florian Buhr raised the question whether the disulfide bonding occurred co-translational, post-translational, or was trapping different conformations during the steps of protein purification.

### 2.4.3 Dynamics of translation elongation of sequences coding for GBC

To study translation and co-translational folding of GBC, we first followed the accumulation of nascent peptides during single-turnover translation of the U and H sequence coding for GBC 1-175 aa. A high concentration of ternary complexes was required for efficient translation of the U sequence. In comparison, the same concentration of ternary complexes allowed translation of the longer HemK 1-277 aa at a maximal rate (Section 2.2.1). We have analyzed the average translation efficiency under the chosen conditions by measuring the incorporation of [<sup>14</sup>C]Phe into the nascent chain by TCA precipitation at the end point of translation. From the ratio of [<sup>14</sup>C]Phe/BodipyFL-C<sub>2</sub>-[<sup>3</sup>H]Met incorporated into the nascent chain, we have calculated an overall translation efficiency of 62% for the U sequence and a higher efficiency of 76% for the H sequence.

## Results

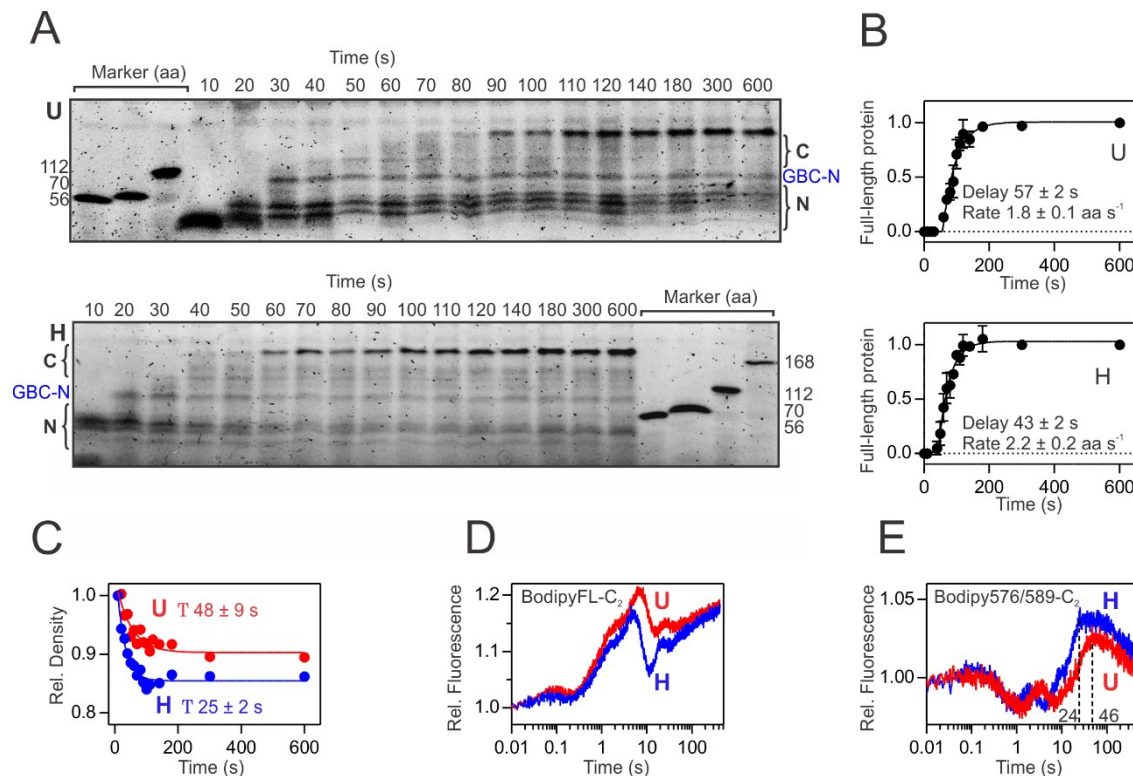


Figure 2-34: Single-turnover translation of the U and H sequences coding for GBC. Translation of both GBC sequences was performed by mixing of 70S IC (15 nM) with a factor mix comprising ternary complexes (40  $\mu$ M), EF-Tu in total (100  $\mu$ M) and EF-G (3  $\mu$ M) at 37°C. **A** Accumulation of nascent peptides during translation of GBC analyzed on Tris-tricine PAGE. Peptides comprising sequences of the NTD or of both domains are indicated by brackets. A translation intermediate closely corresponding in length to the complete NTD is highlighted in blue. **B** Quantification of the rate of synthesis of full-length GBC. Experimental data were fitted to a model with a delay followed by a single exponential increase; standard deviations (n=3) are indicated. **C** Analysis of the cumulative decay of translation intermediates in the NTD marked by brackets in A. The combined density of the N-terminal intermediates was normalized to the time point at 10 s and the cumulative decay was fitted with a single exponential equation. **D** Fluorescence change during single-turnover translation with the N-terminal BodipyFL-C<sub>2</sub>-Met excited at 470 nm and emission recorded after a 500 nm cut-off filter. **E** Same as D using Bodipy576/589-C<sub>2</sub>-Met excited at 540 nm and emission recorded after a 570 nm cut-off filter.

We observed a pronounced accumulation of translation intermediates in the NTD for the U sequence (Figure 2-34 A). Translation of the H sequence led to a lower extent of pausing during translation of the NTD. Additionally, the translation intermediates in the NTD disappeared with a different cumulative decay time of 25 s for the H sequence compared to 48 s for U (Figure 2-34 C). A longer delay time was also observed for the synthesis of the full-length product during translation of the U sequence (Figure 2-34 B). The pausing during translation of the NTD reduced the average rate of synthesis to 1.8 aa s<sup>-1</sup> for the U sequence compared to 2.2 aa s<sup>-1</sup> for



## Results

H. In summary, differences in local translation of the NTD modulate the global rate of translation elongation. We also observed additional translation intermediates in the CTD during translation of the H sequence that might correspond to codons with lower frequency introduced during the process of harmonization (Figure 2-32 B).

We compared the change in fluorescence of the N-terminal fluorophores BodipyFL-C<sub>2</sub>-Met and Bodipy576/589-C<sub>2</sub>-Met during single-turnover translation of the U and H sequences (Figure 2-34 D&E). Close proximity of electron-rich nucleobases such as guanine and aa side chains of Trp and to a lesser extent of Met, His, Tyr results in fluorescence quenching of BodipyFL by the photoinduced electron transfer (PET) mechanism (Boens et al., 2012; Marme et al., 2003; Neuweiler et al., 2003; Olofsson et al., 2006; Torimura et al., 2001). Less is known about potential quenching of Bodipy576/589. The small change in fluorescence of Bodipy576/589 during the time period where the fluorophore is passing through the guanine rich exit tunnel indicated that the fluorescence of Bodipy576/589 might be less subjected to quenching by PET. During translation with N-terminal BodipyFL-C<sub>2</sub>-Met (Figure 2-34 D), we observed a delay in the fluorescence changes for the U sequence compared to H of 2 to 5 s for the pronounced minima and maxima of the fluorescence traces within the first 30 s of translation. The large fluorescence increase at 5 s for H and 7 s for U most likely reported on a change in the environment of the fluorescent probe that occurred when the N-terminal fluorophore was moving through the ribosomal exit tunnel. The ribosome exit tunnel can contain a stretch of 30 to 35 aa (Wilson and Beckmann, 2011). The minimum at 10 s and 14 s for H and U, respectively could have marked the timing when the N-terminus emerged from the exit tunnel. At these time points, a length of the nascent chain of 22 aa for H and 25 aa for U could be approximated based on the average rate of translation elongation for both sequences (Figure 2-35 B). The fluorescence change of BodipyFL-C<sub>2</sub>-Met indicated that translation elongation was already significantly delayed for the U sequence while the nascent chain was still contained within the ribosomal exit tunnel. Translation with Bodipy576/579-C<sub>2</sub>-Met at the N-terminus showed a similar 2 to 4 s delay of the minima and maxima of the fluorescence traces within the first 10 s (Figure 2-34 E). Interestingly, the major increase in fluorescence at around 26 s for the H sequence and 48 s for U revealed a larger relative delay of 22 s that was of same duration as the delay time observed for the cumulative decay of the translation intermediates in the NTD and the relative delay in synthesis of full-length GBC (Figure 2-34 B&C). Therefore, differences in translation elongation for the

## Results

synonymous sequences of U and H accumulated throughout the synthesis of the entire NTD.

### 2.4.4 Co-translational folding of GBC is modulated by the codon usage

The next question was whether the observed differences in local and global translation elongation rates for the sequences of U and H could influence the efficiency co-translational folding of GBC. We took advantage of the high stability of GBC against proteolytic degradation (Mandal et al., 1987). The purified, recombinant proteins expressed from the U and H sequences were resistant to proteolysis with the protease proteinase K (PK). For both the U and H sequence, PK treatment resulted in the formation of a distinct proteolytic fragment derived from ribosome-bound GBC 1-175 aa (Figure 2-35 A).

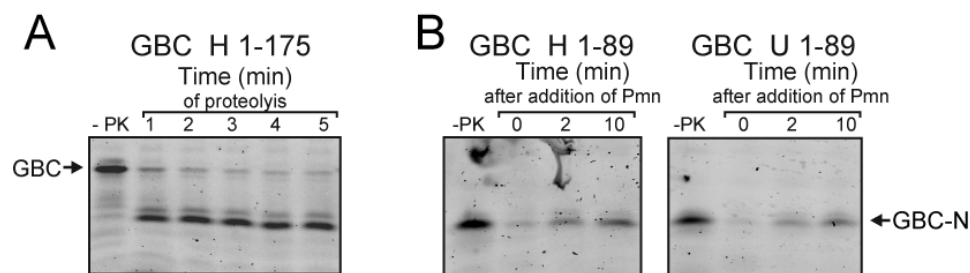


Figure 2-35: Proteolysis of full-length GBC 1-175 aa and GBC 1-89 aa by PK. **A.** Ribosome-bound full-length GBC H 1-175 aa was subjected to proteolysis after 20 min of translation. Full-length GBC and the fragment corresponding to the NTD (GBC-N) are indicated. **B** Probing of the folded state of the NTD before and after the release from the ribosome with Pmn (1.3 mM). The U and H sequences of GBC 1-89 aa were translated for 10 min before PK addition. A proteolysis pulse with PK (5.4 pmol) was applied to all samples for 2 min at 37°C.

The band of full-length GBC was degraded during the time course of proteolysis. The rate of degradation of GBC 1-175 aa was proportional to the concentration of PK. The fragment of smaller molecular weight did not show any degradation over time even at a 6-fold higher concentration of PK. Translation of GBC 1-89 aa comprising the complete NTD and the domain linker confirmed that the shorter fragment corresponded in length to the NTD of GBC (Figure 2-35 B). The fragment of GBC 1-89 aa was previously obtained by limited proteolysis of GBC from *B. taurus* with pepsin and showed a cooperative unfolding transition but could only be partially refolded after denaturation (Sharma et al., 1990). After release from the ribosome with Pmn, the NTD of GBC translated from both the U and H sequence rapidly folded

## Results

into a proteolysis-protected state that resisted a proteolysis pulse by PK (Figure 2-35 B). Co-translational folding of the NTD occurred as soon as the complete sequence emerged from the ribosomal exit tunnel.

We analyzed the formation of protection for the two proteolysis-resistant states (Figure 2-35 A) in a time-resolved manner by subjecting the samples after various translation time to pulse proteolysis (Figure 2-36 A). During translation of GBC 1-175 aa, we observed the accumulation of the N-terminal fragment at a rate of  $\sim 0.5 \text{ min}^{-1}$  for the U sequence and of  $\sim 0.7 \text{ min}^{-1}$  for the H sequence (Figure 2-36 B). The formation of the corresponding proteolysis-resistant state occurred most likely co-translationally at a rate comparable to the synthesis of full-length GBC 1-175 aa. Unfortunately, the pulse proteolysis approach did not allow us to reliably extract differences in the kinetics of the formation of the PK-resistant NTD translated from the U and H sequences. Additionally, we observed a time-dependent increase in the amount of the PK-resistant full-length GBC. Both the rate of formation and the end level of protection of the full-length fragment was lower than for the N-terminal fragment. These findings were in agreement with the lower resistance of full-length GBC against proteolysis by PK (Figure 2-35 A). We observed a significant difference in the rate of protection for the full-length GBC of  $0.03 \pm 0.02 \text{ min}^{-1}$  for the U sequence and  $0.11 \pm 0.07 \text{ min}^{-1}$  for H. We also observed a decrease in the intensity for the N-terminal fragment at longer incubation times, consistent with the sequential gain of protection in both domains. The accumulation of proteolysis-resistant full-length GBC indicated that nascent chains with a proteolysis-protected N-terminal sequence could acquire additional protection in the CTD while still being attached to the ribosome. Formation of a highly proteolysis-resistant folding state within the CTD as observed for the recombinant full-length GBC was most likely precluded by the fact that at the endpoint of translation 30 to 35 aa remain sequestered in the ribosomal exit tunnel. The differences in the rate of appearance of the PK-resistant CTD suggested that the nascent chain translated from the U and H sequences were sampling different conformations.

## Results

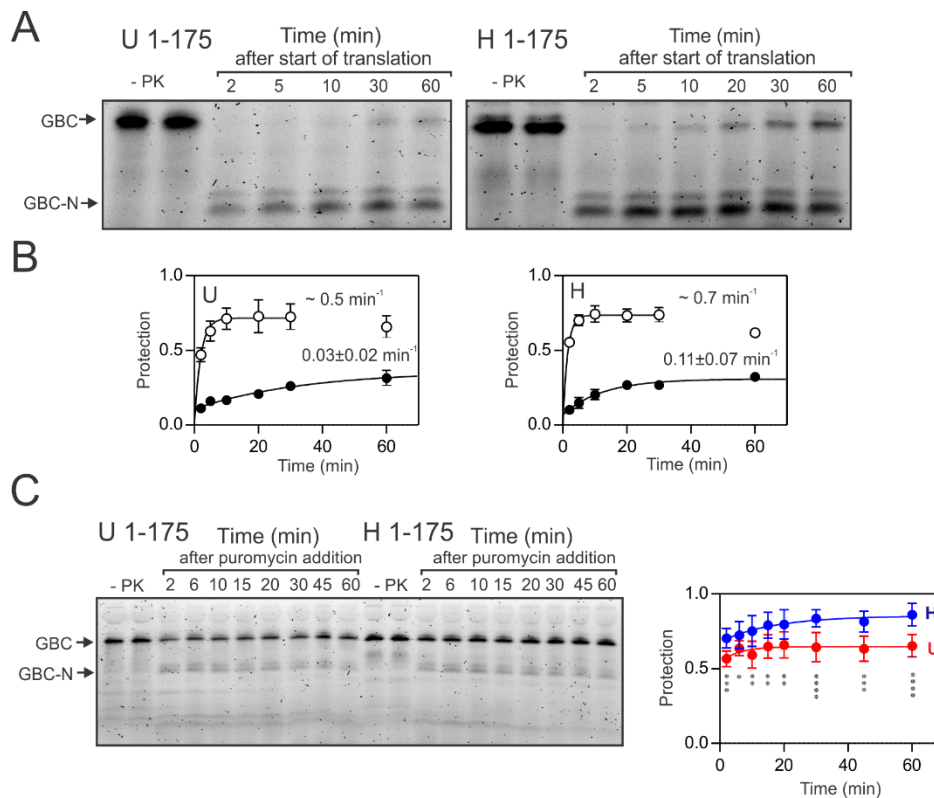


Figure 2-36: Pulse proteolysis of ribosome-bound and ribosome released GBC 1-175 aa. **A** Pulse proteolysis of ribosome-bound chains. Samples at different time points of translation were subjected to a pulse with PK (5.4 pmol for 2 min at 37°C). **B** Quantification of the proteolysis-resistant fragments over time. For each time point, the intensities of the N-terminal and full-length fragment were normalized separately to the intensity of full-length GBC in the absence of PK. Error bars show the SEM for  $n=7$  replicates. **C** Pulse proteolysis of nascent chains released with Pmn (1.3 mM) after 3 min of translation. Ribosome-released nascent chains were subjected to pulse proteolysis as described in A. The protection for full-length GBC was normalized to the intensity of GBC 1-175 aa in the absence of PK. Error bars are standard deviations of  $n=9$  replicates. Significance levels were calculated by Student's two-tailed, unpaired t-test: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.005$ , and \*\*\*\*  $p < 0.0001$ .

To test whether the differences in protection against proteolysis were retained in the nascent chain after release for the ribosome, we released the nascent chains from the ribosome shortly after the synthesis was completed by addition of Pmn and monitored the formation of proteolysis-resistant folded states (Figure 2-36 C). We observed a rapid increase in the amount of full-length GBC in the dead-time of the pulse proteolysis approach. Compared to the proteolysis of ribosome-bound nascent chains, the proportion of N-terminal fragments was small. Significant differences in the protection level of the full-length GBC were observed immediately after the release from the ribosome. The synthesized U protein was to 56% resistant against the proteolysis pulse compared to 70% for H protein. For the H sequence, we also

## Results

observed a larger post-translational increase in protection to a level of 86% within one hour compared to only 65% for U.

### 2.4.5 Monitoring the folding of the N-terminal domain of GBC by FRET

For ribosome-bound nascent chains translated from synonymous mRNA sequences, we have detected altered protection in the CTD of GBC. For the ribosome-released GBC 1-175 aa, we have detected differences in yield of folded protein (Section 2.4.4). Unfortunately, the time resolution of the pulse proteolysis approach did not allow the detection of differences in the kinetics of co-translation folding of the NTD. To test whether in addition to the CTD also the folding of the NTD was affected by the synonymous codon choice, we monitored the co-translational folding of the NTD by FRET between the acceptor fluorophore at the N-terminus and a donor fluorophore incorporated at an internal position. 70S IC containing Bodipy576/589-C<sub>2</sub>-Met-tRNA<sup>Met</sup> were prepared for N-terminal incorporation of the FRET acceptor (Figure 2-18, Figure 2-34 E), whereas the donor probe BodipyFL-C<sub>1</sub>-IA was incorporated at an internal position. Amber suppressor Cys-tRNA<sup>Cys</sup><sub>CUA U32C</sub> that was developed in Section 2.3.2 was applied for site-specific incorporation of the donor fluorophore. We selected the incorporation site based on the upstream sequence that was found to modulate the efficiency of incorporation of fluorescence probes (Section 2.3.5). Position 88 of GBC in the linker between the domains contains an upstream Thr-Gly sequence (Figure 2-32 A) that was identical to the upstream sequence of position 34 of HemK used for establishing the incorporation of fluorescent probes in section 2.3.4. In order not to additionally perturb the linker sequence of GBC, we decided against the substitution of upstream codons for Gly that could increase the incorporation efficiency BodipyFL-C<sub>1</sub>-IA by an order of magnitude (Section 2.3.5). BodipyFL-C<sub>1</sub>-IA was incorporated into GBC 1-175 aa at position 88 during translation of the U and H sequences (Figure 2-37 A). Full-length GBC was synthesized with an identical overall rate compared to single-turnover translation with N-terminal BodipyFL-C<sub>2</sub>-Met (Figure 2-34 A & Figure 2-37 A), demonstrating that the presence of the larger Bodipy576/589-C<sub>2</sub>-Met at the N-terminus and the incorporation of BodipyFL-C<sub>1</sub>-IA at position 88 did not decrease the global rate of translation elongation. We observed a new band corresponding to GBC 1-87 aa indicating that the amber codon at position 88 was only partially suppressed by BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>CUA U32C</sub> (Figure 2-37 A).

## Results

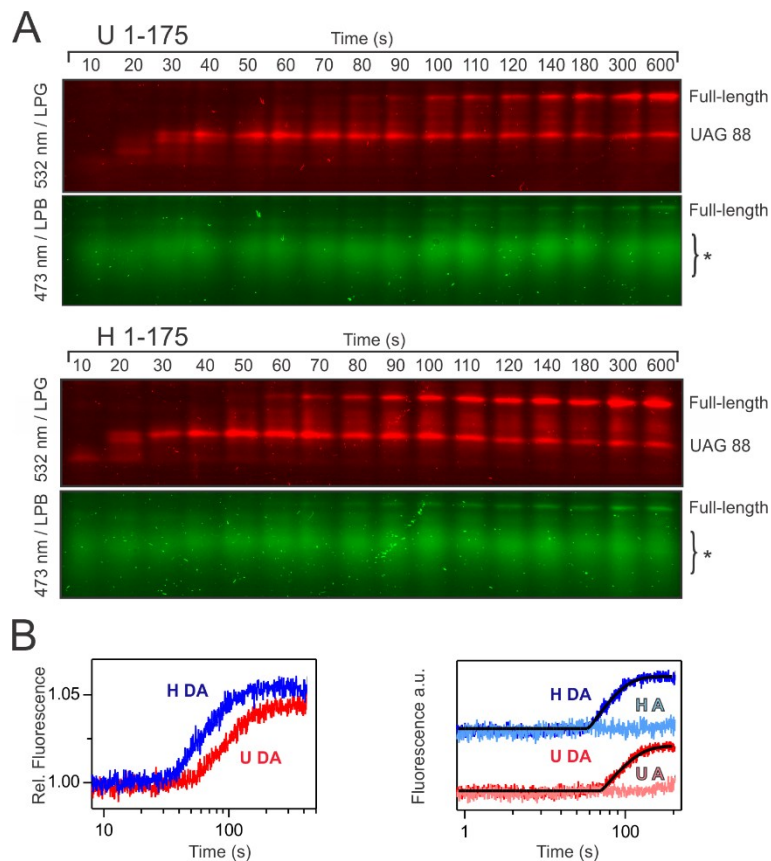


Figure 2-37 Incorporation of BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>CUA</sub>U<sub>32</sub>C (150 nM) into position 88 upon translation of GBC 1-175 aa. **A** Analysis of translation products on Tris-tricine PAGE. In order to separate the fluorescence emission of BodipyFL (green) and Bodipy576/589 (red), the gels were scanned with different lasers (473 nm or 532 nm) and filters (long-pass filter blue LPB 510 nm or long-pass filter green LBG 575 nm) (Mittelstaet et al., 2013). The pause corresponding to the UAG codon at position 88 is indicated and fluorescent hydrolysis products of BodipyFL-C<sub>1</sub>-IA-Cys-tRNA<sup>Cys</sup><sub>UCA</sub>U<sub>32</sub>C are marked by asterisks. **B** (Left panel) To monitor FRET during co-translation folding of the NTD, translation of GBC 1-175 aa was performed in the presence of donor (D) and acceptor (A) fluorophores. The increase in acceptor fluorescence was monitored in a stopped-flow apparatus. The donor BodipyFL was excited at 470 nm and the emission of the acceptor Bodipy576/589 was monitored after passing a 570 nm cut-off filter. (Right panel) Comparison of the FRET signal due to folding of the NTD to the fluorescence change during translation of GBC 1-175 aa in the presence of the acceptor fluorophore and unlabeled Cys-tRNA<sup>Cys</sup><sub>UCA</sub>U<sub>32</sub>C (150 nM). The fluorescence signals for the H sequence are shifted for better visibility. The change in FRET efficiency was fitted to a model with a delay followed by a single exponential increase and the parameters are listed in Table 2-7.

To measure FRET, we monitored the fluorescence emission of the acceptor fluorophore Bodipy576/589 during translation of GBC 1-175 aa in a stopped-flow apparatus upon excitation of the donor fluorophore (Figure 2-37 B). Acceptor fluorescence increased due to FRET between the fluorescent probes after a delay of 35 s for the H sequence compared 51 s for the U. This delay closely corresponded to

## Results

the difference of 14 s calculated for the synthesis of full-length GBC from the U and H sequences, respectively (Figure 2-34 B). The Förster distance for the BodipyFL/Bodipy576/589 FRET pair is 54 Å (Mittelstaet et al., 2013). The distance between the fluorescent probes at the N-terminus and position 88 was approximately 20 to 25 Å based on the crystal structure of GBC (Figure 2-32 A), which should result in a FRET efficiency of almost 100%. The experimentally observed a maximum amplitude change is only 5%, which is much less than one would expect for a 100% FRET efficiency. We noted a significant bleed-through of donor emission into the acceptor channel that might account for an apparently small amplitude of the FRET signal. Additionally, the incorporation efficiency of the acceptor probe Bodipy576/589-C<sub>2</sub>-[<sup>3</sup>H]-Met was about 68% as determined from TCA precipitation of the nascent chains. The investigation of the context-dependence for incorporation of BodipyFL-C<sub>1</sub>-IA indicated that the incorporation level of the donor might be as low as 10% (Section 2.3.5, Figure 2-37 A). The combined incorporation efficiency of the donor and acceptor probes allowed the calculation of theoretical, maximal FRET change of 6-7% that was close to the experimentally observed value (Figure 2-37 B).

The delay time extracted from the FRET traces reflected the timing of initial compaction of the NTD. A translation intermediate closely corresponding in length to the NTD accumulated to maximal intensity after 20 s for the H sequence and 30 s for U (Figure 2-34 A). The difference in time between the synthesis of the NTD and the onset of the FRET change would allow the translation of additional 33 to 36 aa for both sequences. This length matched the number of aa that is required for spanning the ribosomal exit tunnel (Wilson and Beckmann, 2011). We concluded from this calculation that for both sequences the folding of the NTD started as soon as the complete domain sequence emerged outside the ribosomal exit tunnel. Additionally, the exponential increase in FRET occurred at a different rate upon U and H mRNA translation, indicating that the folding of the NTD was altered by the presence of the synonymous codons (Table 2-7). The addition of the delay time and the duration of the increase in FRET yielded a total folding time of 74 s for the H sequence and 109 s for U. A comparison of the folding time with the average time of synthesis of GBC 1-175 aa demonstrated that the folding of the NTD occurred co-translationally for both sequences before the complete protein was synthesized.

## Results

Table 2-7: Parameters obtained by fitting the fluorescence traces in Figure 2-37 B to a model with a delay followed by a single-exponential increase. The duration of the exponential phase was calculated from the reciprocal of the exponential rate.

<u>sequence</u>	<u>delay (s)</u>	<u>rate (min<sup>-1</sup>)</u>	<u>duration (s)</u>
U	50.5 ± 0.5	1.01 ± 0.02	59 ± 1
H	35.1 ± 0.3	1.53 ± 0.02	39 ± 1

In summary, we have observed that co-translational folding of both GBC domains was modulated by the translation elongation of mRNA sequences with synonymous codon substitutions. Increased pausing during translation of the N-terminal codons for the U sequence might be the origin of slower folding of the NTD. The observed differences in the proteolysis-resistance of the CTD in the ribosome-bound chains and the differences in yield of folded full-length GBC after release from the ribosome might arise from the altered folding of the U and H NTD.



## 3 Discussion

### 3.1 Markov model for translation elongation

In collaboration with Sophia Rudorf and Reinhard Lipowsky, we have constructed a Markov model for codon-specific elongation rates that was based on transition rates measured *in vitro* under HiFi buffer conditions. The model allowed the calculation of elongation rates specific for a given codon sequence by considering the complete codon-anticodon matrix (Figure 1-4 B). The simultaneous consideration of the concentration all aa-tRNA species at the same time allows for a more accurate prediction of the rate of translation elongation than would be possible based on the concentration of cognate aa-tRNA only. Additionally, the elongation rates calculated by the Markov model confer a much more detailed insight into the dynamics of translation elongation of a given sequence compared to indirect measures of codon usage, codon adaption index, and number of tRNA genes in the genome.

The comparison of the calculated codon-specific elongation rates of HemK encoded by the nucleotide sequence from *E. coli* strain BL21(DE3) and a codon-optimized sequence led to the prediction of translational pausing at the CGG codon at position 105 that was decoded by the minor isoacceptor tRNA<sup>Arg3</sup><sub>CCG</sub>. This result highlighted the potential of the Markov model for the prediction and rationalization of the non-uniform translation elongation observed during single-turnover translation of natural mRNA.

Sophia Rudorf developed a mathematical approach that allowed us to perform the simultaneous minimization of the differences in the rates of elementary steps during decoding *in vitro* and under cellular conditions. The elementary rate constants measured *in vitro* could be applied with only minor scaling for the description of translation elongation *in vivo*. This result opposed previous criticisms that claimed the selectivity at the step of initial selection to be incompatible with both the low level of aa misincorporation and the fast average rate of peptide synthesis in the cell (Johansson et al., 2008b). The scaling of *in vitro* transition rates to *in vivo* conditions increased the selectivity of the initial selection by only 7.7-fold compared to the two orders of magnitude higher selectivity proposed by the Ehrenberg group. Further criticism concerned the high affinity, thus inhibitory potential, of the codon-independent state of initial binding for non-cognate ternary complexes. The calculated parameters for translation elongation *in vivo* yielded only a minor decrease in affinity for the initial binding step by lowering the rate of ternary complex binding

## Discussion

$\kappa_{on}^*$  due to molecular crowding in the cytoplasm and a slight increase in the rate of dissociation of ternary complexes from the ribosome.

For description of translation elongation both *in vitro* and *in vivo*, the transition rates for steps during cognate and near-cognate decoding represented averaged values measured for the decoding of selected codons. The codons-specific elongation rates resulted exclusively from the rate of initial binding and competition for GTPase activation that was dependent on the concentration of aa-tRNA species. The calculated elongation showed a good correlation with relative elongation rates measured *in vivo*. As an outlook, the predictive power of the Markov model for codon-specific elongation rates and aa misincorporation frequencies could be enhanced by parameterization of the decoding steps for specific codon-anticodon duplexes based on experimentally measured constants.

## Discussion

### 3.2 Sequence specific pausing

Translation elongation proceeds with non-uniform rates *in vivo* and *in vitro*. The recently developed ribosome profiling technique revealed the prevalent character of translational pausing during steady-state translation in the cell. We observed the accumulation of translation intermediates during single-turnover translation of natural mRNA. The number of intermediates that accumulated and the rate of decay of these intermediates differed strongly among mRNA sequences.

With the aim to provide a quantitative insight in the accumulation of translation intermediates and their conversion to longer peptides, we focused on the description of the pausing pattern during translation of HemK 1-277 aa. We have observed that the accumulation of translation intermediates occurred at the same rate of average translation elongation as the full length product was synthesized. Surprisingly, the conversion of translation intermediates into longer peptides, which we hereafter call “decay”, occurred at a lower rate compared to translation elongation and the individual decay rates differed by one order of magnitude. The slower decay of translation intermediates is difficult to reconcile with the notion of obligatory sequential formation of translation intermediates upon synthesis of full-length products. The introduction of a codon-independent, reversible step of inactivation that occurs at a slower rate than the decay of translational intermediates, allowed the description of translational pausing in a global sequential model. So far, the molecular basis of the process of temporal inactivation is unknown. On the other hand, we could follow the sequential interconversion of translation intermediates at lower temperatures. The analysis of the temperature dependence of the average rate of translation elongation of HemK 1-277 aa yielded an activation energy comparable in magnitude as measured for the isolated steps of GTP hydrolysis by EF-Tu during decoding or accommodation in the PTC (Johansson et al., 2008a; Wohlgemuth et al., 2010). Although the process of translation elongation was most likely globally limited by the steps of decoding at any temperature, we could not identify the extent by which individual translational pauses contributed to the observed activation energy.

The helicase activity of the ribosome mediated by the ribosomal proteins S3 and S4 was reported to unwind a complementary duplex of 27 base pairs in length with a predicted melting temperature of 70°C (Takyar et al., 2005). Optical tweezer experiments implicated two modes of helicase activity on the ribosome (Qu et al., 2011). The ribosome may act as a helicase by stabilizing the open state of an RNA duplex. In addition, an active contribution of the energy of peptide bond formation or GTP hydrolysis by EF-G on the ribosome was proposed to be involved in the

## Discussion

unwinding of closed RNA duplexes. In the same study, an increase in the GC content of the mRNA duplex paralleled a decrease in elongation rate. Single-molecule experiments revealed that both the step of translocation and dissociation of the E-site tRNA were slowed down by the presence of secondary structure elements 13 nt downstream of the P site (Chen et al., 2013). For a pseudoknot mRNA structure, the back rotation of the 30S subunit during translocation was proposed to limit both translocation and the E-site tRNA dissociation (Caliskan et al., 2014). Translational pausing during translation *in vitro* was depending on the stability of G-quadruplex structures inserted with correct spacing in respect to the P site (Endoh et al., 2013a, b). Notably, RNA structures that were used in the above mentioned studies were of considerable stability. On the other hand, translation of lacZ sequences containing engineered mRNA secondary structures *in E. coli* did not decrease the rate of translation elongation (Sorensen et al., 1989). We could not detect any correlation between the location of pausing sites and the predicted secondary structure of the mRNA. In contrast, we observed an increased probability of secondary structure formation for the codon-optimized sequence of HemK in the region of codons 90 to 140 but a decreased accumulation of the translation intermediate of  $108 \pm 3$  aa length. Finally, the activation energy calculated from the temperature dependence of the global rate of translation elongation indicated that a step during decoding was rate-limiting. In contrast to the steps of translocation and E-site tRNA dissociation, the decoding in the A site is unlikely to be unaffected by the presence of potential mRNA secondary structures. We favor the view that the translational pausing was most likely not related to the secondary structure content of downstream mRNA sequences. Recent studies have provided a link between codon usage, mRNA structure, and mRNA degradation (Deneke et al., 2013; Pedersen et al., 2011; Presnyak et al., 2015). mRNA regions with higher ribosome density were found to have a longer half-life time in the cell. The location of mRNA secondary structure content was correlated with the usage of frequent codons (Gorochowski et al., 2015). A combination of high secondary structure content with a high rate of translation elongation was proposed to lead to an optimal spacing between the ribosomes in the polysomes (Mao et al., 2014; Mitarai et al., 2008). The optimization for longer half-lifetime of the mRNA as a determinant of the mRNA sequence would make it difficult to separate the influence of codon usage and mRNA secondary structure content on the rate of translation elongation.

Further, we asked the question whether the reading of mRNA codons exclusively cognate to rare isoacceptor tRNA could underlie the accumulation of

## Discussion

translational pauses even at saturating concentration of total aa-tRNA due to competition with near-cognate tRNA. Pausing at CUA codons decoded by the rare tRNA<sup>Leu3</sup><sub>UAG</sub> was detected by using an *in vitro* translation system synchronized by the omission of tRNA aminoacylated with Ile (Endoh et al., 2012). The translation of HemK 1-277 aa encoded by a mRNA sequence containing only the most frequently used codons in *E. coli* led to a decreased accumulation a translation intermediates mapped to the location of  $67 \pm 3$  aa and  $108 \pm 3$  aa while other effects turned out to be more subtle. Based on the computational approach described in Section 2.1 (Rudorf et al., 2014), a six-fold increase in codon-specific elongation at position 105 was predicted for the substitution of the codon CGG decoded by the low abundant tRNA<sup>Arg3</sup><sub>CCG</sub> for CGU read by the major isoacceptor tRNA<sup>Arg2</sup><sub>IGC</sub> during codon optimization of the HemK sequence. Mutagenesis of this codon in the background of the nucleotide sequence from strain BL21(DE3) would need to be performed in order to validate the hypothesis that the low relative abundance of isoacceptor tRNA<sup>Arg3</sup><sub>CCG</sub> among more abundant near-cognate tRNA species might be responsible for the accumulation of the translation intermediate at  $108 \pm 3$  aa. For the translation intermediate at  $67 \pm 3$  aa length, only a modest increase in the codon-specific elongation rate of 20% was predicted for the substitution of the codon CCC at position 67 read by tRNA<sup>Pro2</sup><sub>GGG</sub> to codon CCG decoded by tRNA<sup>Pro1</sup><sub>CGG</sub> and tRNA<sup>Pro3</sup><sub>UGG</sub>. The small magnitude of the increase in codon-specific elongation rate conferred by the codon substitution indicated that decoding of the CCC codon was most likely not limited by the availability of cognate isoacceptor tRNA. Importantly, the subtle changes of the remaining translation intermediates indicated that the low abundance of isoacceptor tRNA was not a general cause of translational pausing. A similar conclusion was supported by the accumulation of several translation intermediates during translation of mRNA sequences with a higher codon adaption (Figure 2-6 A & B). The mRNA coding for AK 1-214 aa with a CAI of 0.64 and PKG 1-208 aa with CAI of 0.74 contained in average more frequent codons than the sequence of HemK 1-277 aa from BL21(DE3) with a CAI of only 0.30. Therefore, the accumulation of translational pauses during *in vitro* translation of mRNA sequences from *E. coli* using total aa-tRNA extracted from the same organism could not be generally explained by the low abundance of tRNA isoacceptors.

The interaction of the anti-SD on the 16S rRNA with mRNA sequences 8-11 nt upstream of the A site was correlated with pausing *in vivo* (Li et al., 2012). However, several studies have questioned this conclusion. The enrichment of ribosome-protected fragments complementary to the anti-SD was suggested as an

## Discussion

artefact during library construction (Martens et al., 2015). The presence of Gly-rich tripeptides, irrespective of the complementarity to the anti-SD sequence, yielded a better correlation with pausing detected in the same profiling data (Woolstenhulme et al., 2013). Despite this criticism, complementarity with the anti-SD sequence could be readily predicted for the sequences of HemK and showed an interesting correlation with the location of the pausing site. The optimization of the codon sequence of HemK to the most frequently used codons introduced a similar number of novel sequences predicted to be complementary to the anti-SD sequence. The role of the anti-SD sequence in translational pausing was experimentally validated by performing a profiling experiment involving orthogonal ribosomes with a mutated anti-SD sequence and complementary mutations in the ribosome binding site on the mRNA (Li et al., 2012; Rackham and Chin, 2005; Wang et al., 2007). The mutagenesis or even the complete deletion of the anti-SD sequence (Vesper et al., 2011) could yield insight into a potential role of SD and SD-like sequences in the accumulation of intermediates during single-turnover translation. Sequences complementary to the anti-SD sequence are an integral part of complex regulatory mechanisms such as frame-shifting into the +1 or -1 register on the mRNA (Larsen et al., 1994; Weiss et al., 1988). In analogy, the SD-like sequences might not be the only effectors of translational pausing and other elements on the mRNA, tRNA or nascent peptide level might be required to trigger accumulation of translation intermediates.

We have developed new experimental approaches for the identification of the site of translational pausing with single-codon resolution. We performed single-turnover translation with N-terminal Bodipy576/579-C<sub>2</sub>-Met, extracted the fluorescence-labeled peptidyl-tRNA and hybridized the intact peptidyl-tRNA to cDNA probes complementary to all *E. coli* tRNA species on a micro-array chip. Unfortunately, we have observed that the hybridization efficiency of peptidyl-tRNA was decreasing with increasing length of the peptide chain. Nevertheless, we were able to identify a potential translation intermediate occurring at codons 34 to 35 of HemK with sequence AAAGGG. The proposed pausing site overlapped with the location of the shortest translation intermediate resolved by Tris-tricine PAGE. Two scenarios could explain the hybridization of peptidyl-tRNA to the Gly2 and Lys probes. In the first model, translational pausing could occur independently at two different sites upstream of both Lys34 and Gly35. Incubation with viomycin would inhibit the subsequent step of translocation after the translational pause was overcome. Alternatively, the detection of hybridization signals for the Gly2 and Lys

## Discussion

probes might be linked to each other. In this case, the translational pause might occur after peptide bond formation of Lys-tRNA<sup>Lys</sup><sub>UUU</sub> in the A site at position 34. A subsequent step leading to transfer of the peptidyl chain to Gly-tRNA<sup>Gly</sup><sub>2 UCC</sub> might be limiting and peptidyl-tRNA<sup>Lys</sup><sub>UUU</sub> would accumulate. Incubation with viomycin would halt the translation after the transfer of the peptidyl chain to Gly-tRNA<sup>Gly</sup><sub>2 UCC</sub> at the step of translocation giving rise to the signal of peptidyl-tRNA<sup>Gly</sup><sub>2 UCC</sub>. Based on the experimental data, we cannot distinguish between the two models. Mutagenesis of the sequence at the proposed pausing site and of upstream residues might validate the location of the pausing site and provide insight into the mechanism of translational pausing. Interestingly, the substitution of K34C preserved the translational pausing at position 33 ± 3 aa. It will be of interest to analyze the extent of translational pausing for all the constructs with upstream aa substitutions that were generated for the purpose of characterizing the context-dependent incorporation of fluorescent probes (Section 2.3.5).

In order to avoid the length-dependent decrease in hybridization efficiency of peptidyl-tRNA carrying longer nascent chains, we have established an alternative sample preparation protocol involving oxidative labeling at the 3'-end of deacylated tRNA. As an additional benefit, deacylated tRNA labeled at the 3'-end allowed hybridization for a longer time period at elevated temperatures compared to peptidyl-tRNA, which should increase the sensitivity and specificity for hybridization to 60-mer cDNA probes. For initial characterization of the new labeling approach and the hybridization of 3'-labeled tRNA, we quantified the aminoacylation level of total aa-tRNA. The two-color hybridization experiments yielded a reasonable estimate of the aminoacylation level of total tRNA of 70% based on the raw intensities obtained from the micro-array scanner. Additionally, we were also able to identify the tRNA species located in the P site of a defined ribosomal complex. On the other hand, the same sample also exhibited hybridization to a number of additional cDNA probes. The origin of those additional hybridization signals will need to be assessed in further experiments. We have observed a high level of unspecific hybridization to a specific set of cDNA probes by the presence of Spike-In probes in the hybridization reaction. The cDNA probes characterized by high levels of unspecific hybridization were removed from the analysis. However, it could have been possible that reduced levels of unspecific hybridization occurred also for a more extended set of cDNA probes complementary to isoacceptor tRNA. Spike-In probes could be omitted in future experiments and thereby the level of unspecific hybridization might be suppressed sufficiently to achieve the low level of unspecific hybridization of tRNA molecules

## Discussion

reported in the literature (Dittmar et al., 2004; Jones et al., 2011; Netzer et al., 2009; Wiltrout et al., 2012). In the near time, an optimized protocol for hybridization of 3'-labeled tRNA extracted from ribosomal complexes during *in vitro* translation might allow the identification of translational pauses by the analysis of the temporal change in the normalized hybridization pattern of cDNA probes. Alternatively to hybridization of tRNA, a recently published study reported the sequencing of tRNA molecules by applying a novel thermostable reverse-transcriptase and demethylase enzymes to remove post-transcriptional modification that would impair reverse-transcription of tRNA (Zheng et al., 2015).



## Discussion

### 3.3 Incorporation of fluorescent labels mediated by Cys-tRNA

We have established protocols for aminoacylation of Cys-tRNA variants and labeling with thiol-reactive fluorescent probes on a preparative scale. We compared the translational performance of Cys-tRNA variants generated by *in vitro* transcription to the fully modified Cys-tRNA<sup>Cys<sub>GCA</sub></sup> purified from *E. coli* total tRNA. Both aa-tRNA formed dipeptides at a similar apparent rate. Mutagenesis of the anticodon sequence for the construction of opal and ochre suppressor tRNA resulted in Cys-tRNA variants inactive in peptide bond formation in an *E. coli* translation system. The apparent rate of the accommodation in the PTC that precedes peptide bond formation was reduced for amber suppressor Cys-tRNA<sup>Cys<sub>CUA</sub></sup> in comparison to Cys-tRNA<sup>Cys<sub>GCA</sub></sup>. Mutagenesis of the nucleotide U32 to C in the anticodon loop restored the rate of accommodation of Cys-tRNA<sup>Cys<sub>UCA</sub> U32C</sup> to same value as measured for Cys-tRNA<sup>Cys<sub>GCA</sub></sup>. Our results are supported by a recent publication reporting translational activity of an amber suppressor tRNA<sup>Cys</sup> from *Bacillus subtilis* that naturally contains C32/A38 in an *E. coli* cell-free translation system (Koubek et al., 2015).

We observed that large and sterically demanding fluorescent probes could be incorporated into the N-terminus of nascent chains by performing translation initiation with fluorescence-labeled Met-tRNA<sup>Met</sup>. Fluorescence-labeled Cys-tRNA enabled the incorporation of fluorescent probes with similar dimensions into dipeptides. Notably, relatively low concentrations of EF-Tu•GTP allowed efficient delivery of the nc-aa-tRNA to the A site. We noted that larger fluorescent probes could only be incorporated into positions close to the N-terminus. BodipyFL was incorporated into positions position 34 of HemK and position 88 of GBC. At positions more distant from the N-terminus, the preference for the linker group of BodipyFL became more pronounced. Only the smaller iodoacetamide linkage of BodipyFL-C<sub>1</sub>-IA was incorporation by ribosome that harbored peptidyl-tRNA with longer nascent chains in the P site; consistent with the more efficient incorporation of several nc-aa at N-terminal positions (Abe et al., 2010; Matsubara et al., 2013; Shozen et al., 2009; Watanabe et al., 2007). The length of the peptide attached to the P site tRNA and the interaction of the nascent chain with nucleotides in the exit tunnel might differentially modulate the ability of the PTC to catalyze peptide bonds with nc-aa-tRNA and canonical aa-tRNA in the A site. Longer nascent chains attached to peptidyl-tRNA increased the rate of peptidyl transfer to the substrate analogue Pmn (Brunelle et al., 2006; Katunin et al., 2002). Longer peptidyl chains also strongly increased the peptidyl transfer from the P site tRNA with a deoxy-ribose substitution at nucleotide A76, while for short chains deoxy-A76-tRNA led to a decrease in the

## Discussion

rate of peptide bond formation by a factor of 100-fold (Huang and Sprinzl, 2011; Koch et al., 2008; Weinger et al., 2004; Zaher et al., 2011). These observations suggest that longer peptidyl chains increase the reactivity of the P site substrate by optimal positioning in the PTC. Such positioning and thereby conformational restriction of the P-site tRNA might not be beneficial for the peptidyl transfer to nc-aa-tRNA. We investigated the ability of all aa to serve as peptide donor substrates to Cys-tRNA labeled with fluorescent probes of increasing size in the A site under the condition that P-site tRNA was carrying a nascent peptide that spanned the complete ribosomal exit tunnel. We observed efficient peptidyl transfer for most aa to the small nc-aa Cys-IA-NBD with the exception of the bulky  $\beta$ -branched aa Leu and Ile and the imino acid Pro. Under the same conditions, only the smallest aa Gly and Ala as P site substrate led to efficient peptidyl transfer to Cys-tRNA labeled with the larger probe BodipyFL-C<sub>1</sub>-IA. For the even larger BodipyTMR-C<sub>5</sub>-mal, we observed incorporation with Gly and Pro as P site substrates despite of the presence of the maleimide group that precluded incorporation of BodipyFL-C<sub>2</sub>-mal with Gly as P site substrate. The longer C<sub>5</sub> linker in BodipyTMR-C<sub>5</sub>-mal might have decreased the negative contribution of the maleimide group on the incorporation efficiency compared to the shorter linker C<sub>2</sub> in BodipyFL-C<sub>2</sub>-mal. The decreased set of P site substrates that allowed efficient incorporation of larger probes indicated an increase in difficulty for proper orientation of donor and acceptor substrates for peptidyl transfer.

Several reports indicated that the incorporation efficiency of structurally diverse nc-aa was enhanced by the presence of small aa upstream of the incorporation site (Abe et al., 2010; Ezure et al., 2014; Jeong et al., 2013; Jeong and Ueda, 2014; Kajihara et al., 2006; Ohtsuka et al., 2011; Shozen et al., 2009; Watanabe et al., 2007). The aa sequence of the nascent peptide close to the P-site tRNA was also found to modulate the peptide bond formation involving a specific set of natural aa. In the absence of EF-P, the synthesis of the XPPP peptide sequences leads to different levels of stalling with XPP in the P site depending on the identity of the residue X (Elgamal et al., 2014; Starosta et al., 2014). Residues with small side chains at position X decreased the level of stalling compared to larger aa side chains. On the other hand, EF-P alleviates stalling at poly-Pro stretches irrespective of the upstream sequence possibly be positioning of the P-site tRNA (Doerfel et al., 2013; Ude et al., 2013). Our results on the dependency of the incorporation efficiency of fluorescent probes with increasing size on the number and location of Gly residues at positions -1 to -4 from the incorporation site suggest that the incorporation of the small probe IA-NBD was unaltered for all constructs with upstream Gly

## Discussion

substitutions. In contrast, for larger fluorophores, the incorporation level increased with the number of upstream Gly residues. Gly substitutions closer to the incorporation site and continuous stretches of Gly residues led to a more pronounced enhancement of incorporation. The highest incorporation level was observed for three consecutive Gly residues flanking the incorporation site. The addition of the fourth Gly did not lead to a further enhancement. The magnitude of the enhancement was dependent on the size of the fluorescent probe and allowed to approximate the absolute incorporation level. The incorporation of Bodipy fluorophores with methyl substituents increased 10-fold when the upstream sequence was changed from HVTG to HGGG. On the other hand, the incorporation of the larger BodipyTMR-C<sub>5</sub>-mal and Atto520-C<sub>2</sub>-IA increased only three-fold. We could not detect any incorporation of the even larger Atto655-IA at position 34 of HemK for the complete set of constructs with Gly substitutions. In contrast, Atto655-IA could be incorporated into dipeptides and positions 4 of HemK. These results may suggest that the substitution of upstream residues with Gly increased the flexibility of the peptidyl-tRNA in the P site, but could not restore the same conformational freedom that is accessible to peptidyl-tRNA at N-terminal positions. Therefore, the constraints imposed by the peptide in the exit tunnel could not be fully alleviated at internal positions. So far, we have analyzed the end level of incorporation for the constructs with Gly substitutions upstream of the incorporation site. In order to gain additional information on the peptidyl transfer to large fluorescent probes in the A site, it would be of interest to investigate the kinetics of incorporation of large nc-aa for set of constructs with upstream Gly substitutions. Further, compared to other P site substrates, the intrinsic reactivity of Gly as P site substrate is reduced as measured by the peptidyl transfer to the substrate analogue Pmn and the reactivity can be enhanced six fold in presence of EF-P (Doerfel et al., 2013). The chemical reactivity of poly-Gly in the P site might differ from the reactivity of a single Gly residue and could be addressed by applying the same experimental approach. The intrinsic chemical reactivity of poly-Gly in the P site could potentially yield interesting insight in the peptidyl-transfer to cognate aa and nc-aa.

We have further noted the accumulation of fluorescence-labeled peptides shorter than the expected full-length product. Shorter peptides also accumulated in cases where we achieved successful incorporation of the fluorescent probes into the full-length product. The accumulation of short peptides was more pronounced during incorporation of larger probes. The number and location of Gly residues upstream of the incorporation site modulated the accumulation of the shorter

## Discussion

peptides. Upstream sequences that allowed high levels of incorporation of fluorescent probes into full-length product also led to increased accumulation of shorter products. Analysis of the time course of incorporation showed that the shorter fragment accumulated with the same rate as the full-length product. The same rate of synthesis and dependence on the sequence upstream of the incorporation site indicated that the formation of the shorter fragment corresponded to a kinetically competent intermediate during synthesis of the full-length product. A fraction of this translation intermediate might not be able to continue in further translation elongation. Upstream sequences with a continuous stretch of Gly residues immediately adjacent to the incorporation site led to a lower ratio of truncated fragments when normalized to the accumulation of full-length product. These findings supported the hypothesis that the positioning of the nc-aa that was translocated into the P site might limit the next round of peptidyl transfer to a canonical aa in the A site. Further experiments will be required to address the molecular events and mechanism leading to the accumulation of shorter fragments.

The Hohsaka group optimized the upstream sequence of an N-terminal tag (ProX-tag) that enables the incorporation of large nc-aa at position 9 when fused to proteins of interest (Abe et al., 2014; Abe et al., 2011; Abe et al., 2010). For the incorporation of large fluorescent probes, we found that three consecutive Gly residues upstream of the incorporation site formed a sequence context that enhanced incorporation at internal positions. By analogy, this upstream sequence context can be regarded as a tag sequence that can be transplanted into proteins of interest. The high degree of context-specificity for the incorporation of larger fluorescent probes limits the number of suitable incorporation sites compared to small nc-aa that are believed to be incorporated at any position of proteins of interest with little dependence on the sequence context (Liu and Schultz, 2010; Neumann, 2012). Although distinct preferences for residues flanking the incorporation site were also detected for small nc-aa (Pott et al., 2014). The incorporation of large fluorescent probes at carefully selected sites yielded key information on the co-translational folding of GBC (Section 2.4.5) and holds further promise for the investigation of co-translational processes in more detail by single-molecule techniques. However, the substitution of the sequence upstream of the incorporation site with three consecutive Gly residues is expected to disrupt secondary structure elements. On the other hand, Gly substitutions might be structurally tolerated in loop and linker regions of proteins. For loop regions, the increase in loop size by insertion of additional residues increases the entropic penalty for loop folding, which results in

## Discussion

increased rates of unfolding and a less pronounced effect on the rate of refolding (Dagan et al., 2013; Ladurner and Fersht, 1997). Therefore, substitution of the native residues in loops might be a more promising approach for insertion of Gly stretches with less consequences on protein stability.

We have observed that the maleimide linker of BodipyFL-C<sub>2</sub>-mal did not allow incorporation into internal positions, in contrast to BodipyFL-C<sub>1</sub>-IA with a shorter linker and an iodoacetamide group. BodipyFL attached to the *para*-amino group of 4-amino-phenylalanine was characterized by a higher incorporation level compared to linkage to the  $\epsilon$ -amino group of Lys (Kajihara et al., 2006). Similarly, an extended set of nc-aa were incorporated more efficiently when linked to 4-aminophenylalanine compared to a Lys linker (Abe et al., 2010; Hohsaka et al., 2003; Shozen et al., 2009; Watanabe et al., 2007). As an example, for nc-aa containing the small fluorescent dansyl group, the linkage to 4-aminophenylalanine allowed for its incorporation *in vitro* with a low degree of context dependence compared to dansyl-lysine (Hohsaka et al., 2004). The inefficient incorporation of dansyl-Lys was supported by reports of independent groups (Cornish et al., 1994; Steward et al., 1997). Interestingly, dansyl-Ala showed good incorporation *in vivo* when charged on an amber suppressor tRNA by an orthogonal aaRS (Summerer et al., 2006); this finding underscores the importance of linker sequences for the efficient incorporation of specific functional groups. As a final example, N-acetyl-galactosamine, a sugar moiety of a rather small molecular weight attached to Thr as linker, could only be incorporated into N-terminal positions (Matsubara et al., 2013). A hypothesis was proposed that an aromatic linker sequence might be more efficient for incorporation of larger nc-aa based on the high efficiency of incorporation of several small aromatic nc-aa (Ellman et al., 1991; Hohsaka et al., 2003). The identity of the aa modulates the rate of dissociation of the aa-tRNA from the A site. Phe was proposed to form additional interactions with the PTC compared to the smaller aa Ala (Dale et al., 2009). Furthermore, Pmn derivatives with different aa side chains exhibited a broad range of binding affinities to the PTC (Bhuta et al., 1981; Starck et al., 2003; Vanin et al., 1974). Interestingly, the affinity of a Pmn compound containing a tyrosine side chain was much higher compared to the compound with a lysine side chain. These reports suggest that the PTC has a different specificity for aa side chains on aa-tRNA. Alternative modes of interaction with the PTC conferred by the linker sequence of nc-aa may allow for more beneficial orientation of larger substituents. It might be possible that the aromatic 4-aminophenylalanine linker leads to a different positioning of large nc-aa in the PTC compared to aliphatic linkers, such as Lys and

## Discussion

Thr. Despite a strong context-dependence, additional upstream aa seemed to be compatible with the incorporation of BodipyFL linked to 4-aminophenylalanine compared to the incorporation of BodipyFL-C<sub>1</sub>-IA by Cys-tRNA (Figure 2-28 A) (Iijima and Hohsaka, 2009; Kajihara et al., 2006). An optimized linker structure for moderate-sized nc-aa might lead to a significant increase in the number of possible incorporation sites in proteins due to a relaxed requirement for the residues upstream of the incorporation site. On the other hand, for successful incorporation of large fluorescent probes, a consistent requirement for the optimization of the upstream sequence context of the incorporation site that involved more than one residue was reported (Ezure et al., 2014; Jeong et al., 2015; Jeong et al., 2013; Jeong and Ueda, 2014; Ohtsuka et al., 2011). These findings indicate that alternative positioning in the PTC conferred exclusively by the linker structure could not restore peptidyl transfer to large, challenging nc-aa in the A site.

## Discussion

### 3.4 Translation elongation and co-translational folding

We investigated the co-translational folding of GBC translated from mRNA sequences differing in the usage of synonymous codons. Harmonization of the bovine sequence for the codon usage in *E. coli* increased the yield of recombinant GBC expressed in *E. coli* compared to the original bovine sequence. Upon expression of the H sequence a smaller amount of truncated fragments of GBC was observed, indicating that the lower yield during expression of the U sequence was associated with cellular degradation of GBC that attained different conformations. Western blotting with an antibody against the C-terminal His<sub>6</sub>-tag demonstrated that at least in part the degradation products retained the C-terminal sequence and were therefore derived from the full-length GBC after completion of synthesis. Analysis of the translation elongation of the synonymous sequences *in vitro* indicated that translation of the H mRNA resulted in the accumulation of full-length GBC at a higher average rate of 2.2 aa s<sup>-1</sup> compared to 1.8 aa s<sup>-1</sup> for the U mRNA. In addition to the altered rate of global translation elongation, specific translation intermediates accumulated upon translation of the two sequences, indicating differences in local translation elongation. During translation of the codons in the NTD, translation intermediates accumulated for both sequences. The U sequence led to more pronounced translational pausing compared to the H sequence resulting in a cumulative delay of ~20 s for the translation of the NTD. The relative delay time during translation of the N-terminal codons was comparable to the delay time observed for the build-up of full-length GBC. Thus, a major contribution to the altered global rate of protein synthesis accumulated due to differences in local translation elongation of N-terminal codons. Based on the average rate of translation elongation, the nascent chain translated from the U and H sequence would be expected to span the complete ribosomal exit tunnel after 19 s and 15 s, respectively. Changes in fluorescence intensity of N-terminal fluorescent probes revealed a significant relative delay of ~5 s during the first 10 s of translation for the U compared to H sequences indicating that a considerable part of the cumulative delay during synthesis of the NTD occurred before the nascent chain emerged from the ribosomal exit tunnel. The remaining part of the relative delay time between the U and H peptide synthesis accumulated during a period where segments of the nascent polypeptide are outside the ribosome and might sample folded states.

The development of a pulse proteolysis approach allowed us to monitor the formation of both domains while the nascent chain of GBC was still bound to the ribosome. An N-terminal fragment comprising the NTD was completely resistant

## Discussion

towards proteolysis with PK for longer time periods and at higher concentrations of PK, which indicated the formation of a folded state with high stability. The accumulation of folded states in both domains of GBC, occurred on a slower time scale which could be adequately measured by pulse proteolysis. Accumulation of a proteolysis-resistant state of full-length GBC occurred at an almost 4-fold slower rate for the U sequence compared to H. At the endpoint of translation 30 to 35 residues of the CTD remained inside the ribosome exit tunnel. The proteolysis-resistant state of full-length GBC most likely corresponded to a stage of initial docking of the CTD on the folded NTD. The formation of a stable CTD required the residues sequestered in the ribosomal tunnel as indicated by the increased PK-resistance of the CTD in the ribosome-released chains. For the H sequence, a higher rate of accumulation of a proteolysis-resistant state in the CTD of ribosome-bound GBC also paralleled a higher yield of folded GBC after the release from the ribosome. Therefore, the observed differences for accumulation of PK-resistant states on the ribosome appeared to have also reported on the formation of on-path folding states leading to more productive folding of GBC released from the ribosome. Similarly, after release from the ribosome, we observed a larger increase in folded protein over time for the H sequence. A sufficiently pre-organized ensemble of folded states might have been formed that allowed more efficient post-translational maturation after the release from the ribosome. We were further wondering whether differences in folding of the CTD might originate from altered folding of the NTD. We monitored the folding of the NTD by FRET between N-terminal Bodipy576/589-Met and BodipyFL-C<sub>1</sub>-IA incorporated at the amber codon at position 88. The comparison of the delay time before the onset of the FRET signal with the time course of nascent peptide synthesis indicated that the NTD started to fold as soon as the complete sequence of the NTD emerged from the ribosome exit tunnel. Importantly, the time required for folding of the NTD as measured by the duration of the increase in FRET demonstrated that during translation of the U sequence additional 20 s were required for folding of the NTD. In summary, our data highlighted the importance of the rate of formation of the NTD for productive folding of GBC on the ribosome and after the release. Slower formation of the NTD during translation of the U sequence decreased the efficiency of CTD formation and docking of the CTD on the NTD. Interestingly, the interplay between increased local translation elongation rates and more efficient protein folding on the ribosome was not documented experimentally until present. The experimental data allowed us to suggest several alternative scenarios linking the



## Discussion

slower translation elongation of the U sequence especially in the N-terminal region to the slower rate of folding of the NTD.

In a first scenario, early folding states that could have a profound influence on the folding pathway of the NTD might be already sampled when the polypeptide sequence is elongated through the exit tunnel and enters the vestibule. The vestibule is located 80 Å from the PTC and forms a larger cavity of about 20 Å in length at the end of the exit tunnel (Lu and Deutsch, 2005b; Voss et al., 2006). Depending on the extend of secondary structure formation of the nascent chain in folding zones upstream and downstream of the L4/L22 constriction site, a different number of residues is required for spanning the exit tunnel from the PTC to the vestibule (Lu and Deutsch, 2005a). Formation of tertiary structure elements e.g.  $\alpha$ -helical hairpins (Kosolapov and Deutsch, 2009; Tu et al., 2014; Tu and Deutsch, 2010) and even domain folding of several proteins was reported to take place in the vestibule (Kowarik et al., 2002; Kudlicki et al., 1995; Makeyev et al., 1996). The translational pausing observed for the U sequence led to more pronounced accumulation of peptides of 30 to 50 aa in length. Under the assumption that 20 to 25 residues are required for spanning the distance between the P site and the vestibule (O'Brien et al., 2010; Wilson and Beckmann, 2011), the cumulative delay of the individual pausing events might have allowed for the formation of an altered folding state in the vestibule. The formation of the cryptic disulfide bonds observed in the NMR spectra of the purified GBC expressed from the H sequence might have monitored the conformational ensemble in this initial nucleus for later folding processes. On the other hand, we could not detect initial folding states in the vestibule based on the change in fluorescence of N-terminal fluorescent probes. Similarly, the position of the FRET probes was not suitable for the monitoring of such initial folded states. The presence of disulfide bonds in GBC extracted from bovine eye lens indicated that disulfide bond formation was compatible with the properly folded structure of GBC. The reduction potential of the *E. coli* cytoplasm disfavors the stable formation of disulfide bonds and the tripeptide glutathione acts as cellular redox buffer (Derman et al., 1993; Stewart et al., 1998). The accessibility of the nascent chain for small molecules increases steeply across the vestibule to a level comparable to that outside of the ribosome (Lu et al., 2007). Low accessibility of cysteine side chains to the redox buffer might have allowed the formation of disulfide bonds that could become subsequently shielded from solvent interactions within the developing folded states. During the translation of the U sequence, disulfide bonds might form less efficiently or might be solvent-exposed for longer time period due to translational pausing. On

## Discussion

the other hand, disulfide bond formation did not constitute a necessary step for folding of the NTD because the presence of the stronger reducing agent DTT in HiFi buffer allowed the folding of the NTD during translation of both sequences. Alternatively, if disulfide bonds were not already formed at the interface of the ribosome exit tunnel with the vestibule, the formation of disulfide bonds outside the ribosome in the cytoplasm would be highly unlikely and the observation of disulfide bonds in the purified proteins might be attributed to the locking of conformations during the procedures of protein purification. Protein expressed from the U sequence might sample conformations in which the probability for disulfide bond formation is lower.

Apart from altered folding at an early state in the vestibule, the folding of the NTD outside the ribosome could have been influenced by any difference in translation elongation during synthesis of the NTD including the linker region and the parts of the CTD as a spacer for spanning the exit tunnel. The folding of the NTD might have integrated the difference in elongation time of a single codon or the cumulative of several decoding events. The finding that the harmonized sequence with a lower degree of translational pausing and a higher rate of overall protein synthesis enhanced the formation of productive folded states was in contrast to the majority of literature published on codon-optimization to enhance the efficiency of co-translational protein folding. The substitution of codons and clusters of codons with low adaption to the codon usage with frequently used codons was correlated with an elevated level of misfolded proteins (Chang et al., 2005; Crombie et al., 1992; Fedyunin et al., 2012; Kim et al., 2015; Komar et al., 1999; Siller et al., 2010). Theoretical studies applying analytical parametrization of translation elongation and the folding probability of protein domains at each codon confirmed that a high rate of translation elongation could significantly delay domain folding (Ciryam et al., 2013; O'Brien et al., 2012). The removal of rare codons from the mRNA sequence coding for the protein Suf1 from *E. coli* was linked to higher sensitivity to proteolytic digestion (Zhang et al., 2009). Coarse-grained simulation of the co-translational folding pathway of Suf1 yielded a higher probability of domain folding when the clusters of rare codons with slower elongation rates were explicitly included in the simulation or alternatively when the average synthesis rate was decreased (Tanaka et al., 2015). In contrast, our experimental data indicated that the co-translational folding of the NTD occurred at a rate that was compatible with the rate of translation elongation measured for the H sequence. Translation of the H sequence also led to the formation of proteolysis-resistant, presumably on-path folded states in both

## Discussion

domains of GBC within a shorter time-scale. Slower translation elongation of the U sequence might have allowed the formation of unproductive folded states that could act as kinetic traps for domain folding. Simulations based on analytical parametrization of the co-translational folding process indicated that a high rate of translation elongation at segments prone to accumulate off-path folding intermediates increased the probability of domain folding provided that later folding events did not become limited by the high rate of translation elongation. A scenario involving the accumulation of off-path intermediates as consequence of altered translation elongation rate *in vitro* would need to be carefully validated against the situation *in vivo* that is characterized by a rate of translation elongation of 13 to 22 aa s<sup>-1</sup> (Bremer and Dennis, 2008). The folding of the NTD occurred co-translationally for both sequences during single-turnover translation *in vitro*. A combination of slower folding of the NTD from the U sequence with a higher rate of synthesis *in vivo* might even shift the folding of the NTD post-translationally in part.

Measurements of the resistance against mechanic unfolding of the ribosome-bound T4 lysozyme by optical tweezers indicated a chaperone activity of the vestibule that prevents the formation of a completely folded state (Kaiser et al., 2011). Electrostatic interactions between the ribosome and charged residues in the nascent chain reduced the rotational freedom of an N-terminal fluorescent probe (Knight et al., 2013). Lattice Monte Carlo simulation of the contact map of proteins with different folding topology indicated that local folding based on short-range interactions was allowed on the ribosomal surface, although long-range interactions involving the frequently observed joining of the termini in small proteins was disfavored (Krobath et al., 2013). Coarse-grained simulation of protein folding within the vestibule region indicated that tertiary interactions were sampled, but only transiently, including the  $\alpha$ -helical hairpin of the Kv1.3 protein that was previously reported to fold in the vestibule based on cross-linking experiments (O'Brien et al., 2010; Tu et al., 2014). An appealing scenario compatible with the slower rate of folding of the NTD during translation of the U sequence might rely completely on the accumulation of differences in local translation elongation times. The folding of the NTD outside the ribosome might be limited by the translation of a sufficiently long polypeptide linker, even including residues topologically associated with CTD. In a further scenario, the folding of the NTD might be delayed by the environment of the vestibule and the proximity of the ribosome. Off-path intermediates might be populated during a different time period and thereby to a different extent depending

## *Discussion*

on the rate of translation elongation for the linker sequence that renders the NTD folding competent outside the ribosome.

The multitude of scenarios that could be consistent with the experimental data underlines both the complexity of co-translational folding processes and the limitation of published experimental strategies to distinguish among the proposed scenarios. The set of approaches described in this thesis, comprising a computational model for codon-specific elongation rates, techniques that might allow the localization of pausing sites, and the successful incorporation of a large set of fluorescent probes, would allow the experimental testing of the above proposed scenarios at the current time or in the near future.

## 4 Material and Methods

### 4.1 Equipment

Table 4-1 List of equipment

Equipment	Supplier
Äkta Purifier Plus	GE Healthcare
Bio-vision imaging system	Peqlab Biotechnologie
Benchtop centrifuge 5415R and 5810R	Eppendorf
Centrifuge Avanti J-26 XP	Beckmann Coulter
Electrophoresis power supply EV261	Peqlab Biotechnologie
Emuliflex	Avestin Europe
Fermenter Biostat B Plus	Sartorius Stedium Biotech
FLA900 fluorescence imager	Fuji Film
Fluorolog spectrofluorometer	Horiba Jobin Yvon
FluoroMax 4	Horiba Jobin Yvon
Incubator shaker series Innova44	New Brunswick
Liquid scintillation counter TriCarb 3110 TR	PerkinElmer
Milli-Q Advantage A10	Millipore
Nanodrop 2000C	Peqlab Biotechnologie
pH electrode	Sentix81
pH meter	inoLAB
Plates incubator INE600	Memmert
PerfectBlue Dual Gel system Twin S / ExWS	Peqlab Biotechnologie
Stopped-flow SX20D spectrometer	Applied Photophysics
Ultracentrifuge Optima L-100 XP	Beckmann Coulter
Vortex Genie 2	Scientific Industries
Water bath RE104 and E100	Lauda
HPLC unit 2695	Waters
HPLC Fluorescence detector 2475	Waters
HPLC UV/VIS detector 2489	Waters
Fraction collector Frac 920	GE Healthcare
Savant SPD111V SpeedVac	ThermoScientific
Lambda Bio+ UV/VIS spectrometer	Perkin Elmer

### 4.2 Software

The software used for data acquisition and analysis is listed in Table 4-2.

Table 4-2 List of software.

Software	Supplier
FluorEssence 3.5	Horiba Scientific
GraphPad Prism 5.0	GraphPad Software
ImageJ 1.48	NIH
R for win 3.2.0	R-Project
Origin 8.0	OriginLab
R Studio 0.98.1103	R Studio
UCSF Chimera 1.10.1	Schrödinger

## Material and Methods

### 4.3 Chemicals and Consumables

Chemicals were obtained from Sigma-Aldrich, Merck, or Roth. Chemicals and consumables from other suppliers are listed in Table 4-3.

Table 4-3 List of Chemicals and Consumables

Product	Supplier
10x Blocking agent	Agilent
2x Hi-RPM hybridization buffer	Agilent
Antarctic Phosphatase	New England Biolabs
Atto488-mal	Atto-Tec
Atto520-C <sub>2</sub> -IA	Atto-Tec
Atto520-C <sub>2</sub> -mal	Atto-Tec
Atto610-NHS	Atto-Tec
Atto633-mal	Atto-Tec
Atto655-IA	Atto-Tec
Atto655-NHS	Atto-Tec
Bodipy507/545-IA	life technologies
Bodipy576/598-C <sub>2</sub> NHS	life technologies
BodipyFL-C <sub>1</sub> -IA	life technologies
BodipyFL-C <sub>2</sub> -mal	life technologies
BodipyFL-C <sub>2</sub> -NHSS	life technologies
BodipyTMR-C5-mal	life technologies
Cellulose nitrate filter, 0.45 $\mu$ m, 25 mm	Sartorius
Centrifugal Concentrators	Millipore
Cy3-hydrazide	GE Healthcare
Cy5-hydrazide	GE Healthcare
DpnI	New England Biolabs
Gene Expression wash buffer 1	Agilent
Gene Expression wash buffer 2	Agilent
IA-NBD	life technologies
Irga-Safe	Perkin Elmer
N-(1pyrene)-mal	life technologies
NdeI	New England Biolabs
NucleoSpin PCR cleanup kit	Machery-Nagel
NucleoSpin Plasmid kit	Machery-Nagel
P-6 spin columns	Biorad
Phusion Polymerase	Thermoscientific
Pyruvate kinase	Roche
Qsy-35	life technologies
QuickLigase	New England Biolabs
QuickZint 361	Zinnser Analytic
RNeasy Midi kit	Quiagen
SmaI	New England Biolabs
RNAsin	Promoga
T5 Exonuclease	New England Biolabs
Taq DNA ligase	Biozymes
total tRNA	Roche
XhoI	New England Biolabs

The columns used in this work are listed in Table 4-4.

## Material and Methods

Table 4-4 List of columns

Column	Supplier
HiTRAP Q 5 mL	GE Healthcare
LiChroCart 250-10, LiChrospher WP300 RP-18	Merck
LiChroCart 250-4, LiChrospher 100 RP-8	Merck
LiChroCart 250-4, LiChrospher WP300 RP-18	Merck
monoQ 5/50 GL	GE Healthcare
NAP-5	GE Healthcare
Superdex 75 10/300 GL	GE Healthcare

### 4.4 Strains and Plasmids

All subcloning and mutagenesis was performed with *E. coli* strain DH5a. Chemical competent DH5a cells were prepared in house or obtained from New England Biolabs. Proteins were expressed in BL21(DE3) cells.

Table 4-5 List of Strains

Genotype	Strain
DH5a	<i>F- endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG <math>\Phi</math>80dlacZ<math>\Delta</math>M15 <math>\Delta</math>(lacZYA-argF)U169, hsdR17(rK- mK+), <math>\lambda</math>-</i>
BL21(DE3)	<i>F- ompT gal dcm lon hsdSB(rB- mB-) <math>\lambda</math>(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5])</i>

Plasmids that were not constructed during this thesis are listed in Table 4-6.

Table 4-6 List of Plasmids

Construct	Vector	Promotor	Resistance	Tag
CspA	pUC18	T7	Amp	None
CysRS	pCA24N	T7	Cam	C-term His <sub>6</sub>
EF-G	pet24	T7	Kan	His-tag
EF-Tu	peT21	T7	Kan	C-term His <sub>6</sub>
GBC_H	pET15b	T7	Amp	C-term His <sub>6</sub>
GBC_U	pET15b	T7	Amp	C-term His <sub>6</sub>
HemK	pET24a	T7	Kan	C-term His <sub>6</sub>
IF-1	pET11a	T7	Amp	None
IF-2	pET11M	T7	Kan	None
IF-3	pET24a	T7	Kan	C-term His <sub>6</sub>

## Material and Methods

### 4.5 Buffers and solutions

Table 4-7: List of buffers and solutions

Buffer	Composition
Acid PAGE 3.5% gel	3.5 % (v/vol) 19:1 acrylamide:bis-acrylamide, 8 M urea, 100 mM NaOAc pH 5.2
Acid PAGE 6.5% gel	6.5 % (v/vol) 19:1 acrylamide:bis-acrylamide, 8 M urea, 100 mM NaOAc pH 5.2
Acid PAGE loading, 2x	150 mM NaOAc pH 5.2, 10 mM EDTA, 8 M urea, 0.05% (w/vol) bromophenol blue 0.05% (w/vol) xylene cyanol
Acid PAGE running, 1x	100 mM NaOAc pH 5.2
Aminoacylation, 5x	500 mM HEPES-KOH pH 7.5, 35 mM MgCl <sub>2</sub> , 75 mM KCl
Bis-tris running, 1x	50 mM Bis-Tris/MES pH 6.33, 1 mM EDTA, 0.1% (w/vol) SDS
DNA loading, 6x	30% (v/vol) glycerol, 0.25% (w/vol) bromophenol blue 0.25% (w/vol) xylene cyanol
HAKM <sub>7</sub> , 1x	50 mM HEPES-KOH pH 7.5, 70 mM NH <sub>4</sub> Cl, 30 mM KCl, 7 mM MgCl <sub>2</sub>
HiFi compensation, 2x	50 mM Tris-HCl pH 7.5, 70 mM NH <sub>4</sub> Cl, 30 mM KCl, 16 mM putrescine, 1 mM spermidine
HiFi, 1x	50 mM Tris-HCl pH 7.5, 70 mM NH <sub>4</sub> Cl, 30 mM KCl, 3.5 mM MgCl <sub>2</sub> , 8 mM putrescine, 0.5 mM spermidine
HPLC salt stock, 5x	100 mM NaOAc, 50 mM MgCl <sub>2</sub> , 2 M NaCl, pH 4.8
HPLC, 40%	40% EtOH (v/vol), 1x HPLC salts, pH 5.0
HPLC, 5%	5% EtOH (v/vol), 1x HPLC salts, pH 5.0
IMAG	25 mM Tris-HCl pH 8.0, 300 mM NaCl, 1 mM MgCl <sub>2</sub> , 3 mM mercapto-ethanol, 10% (v/vol) glycerol
Jagow Anode, 10x	1 M Tris-HCl pH 8.9
Jagow Cathode, 10x	1 M Tris, 1 M Tricine, 1 % SDS
LB	10 g L <sup>-1</sup> NaCl, 10 g L <sup>-1</sup> tryptone, 5 g L <sup>-1</sup> yeast extract
LB Agar	LB, 15 g L <sup>-1</sup> agar
Lysis	30 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM MgCl <sub>2</sub> , 1 mM mercapto-ethanol, 10% (v/vol) glycerol, 0.1 mM PMSF
Methylene blue staining	80 mM NaOAc pH 5.0, 1g/L methylene blue
Native PAGE gel	50 mM Tris, 50 mM boric acid, 5% (v/vol) acrylamide (29:1), 10 mM Mg(OAc) <sub>2</sub> , 1 mM EDTA pH 8.0, 4% (v/vol) glycerol
Native PAGE loading	50% (v/vol) glycerol. 0.05 % bromophenol blue
Native PAGE running	50 mM Tris, 50 mM boric acid, 1 mM EDTA pH 8.0, 10 mM Mg(OAc) <sub>2</sub>
Native PAGE TC, 5x	200 mM HEPES-KOH pH 7.5, 260 mM NaOAc, 40 mM Mg(OAc) <sub>2</sub>
PBS, 10x	100 mM NaH <sub>2</sub> PO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> pH 7.4, 1.37 M NaCl, 25 mM KCl
RNA loading	1xTBE, 8 M urea, 0.05 % (w/vol) bromophenol blue
RNA-Q-A	50 mM NaOAc pH 5.0, 10 mM MgCl <sub>2</sub>
RNA-Q-B	50 mM NaOAc pH 5.0, 10 mM MgCl <sub>2</sub> , 1.1 M NaCl
SDS loading, 4x	150 mM Tris-HCl pH 6.5, 12% (w/vol) SDS, 30% (v/vol) glycerol, 6% (v/vol) mercapto-ethanol
TAE, 50x	242 g L <sup>-1</sup> Tris, 57.1 ml L <sup>-1</sup> glacial acetic acid, 100 ml L <sup>-1</sup> 0.5 M EDTA pH 8
TAKM <sub>7</sub> , 1x	50 mM Tris-HCl pH 7.5, 70 mM NH <sub>4</sub> Cl, 30 mM KCl, 7 mM MgCl <sub>2</sub>
TBE, 10x	107.8 g L <sup>-1</sup> Tris, 55.03 g L <sup>-1</sup> boric acid, 40 ml L <sup>-1</sup> 0.5 M EDTA pH 8
Transcription, 5x	200 mM Tris-HCl pH 7.5, 75 mM MgCl <sub>2</sub> , 50 mM NaCl, 10 mM spermidine,
UREA PAGE, 12% gel	1x TBE, 12% (v/vol) acrylamide (19:1), 8 M Urea



## Material and Methods

### 4.6 Molecular biology protocols

#### 4.6.1 Site-directed mutagenesis following the QuikChange protocol

Site directed mutagenesis was performed following the work-flow of the patented QuikChange protocol (Stratagene, La Jolla) as follows, Phusion polymerase (2 units) in high fidelity buffer, dNTP (0.2 mM), primer (0.2  $\mu$ M) each, and plasmid template (0.6 ng  $\mu$ l<sup>-1</sup>) were used and the PCR program contained the following steps: 3 min 98°C, 35 cycles (30s 98°C, 20 s 57°C, 30 s kb<sup>-1</sup> 72°C), 10 min 72°C. PCR fragments were analyzed on 1.5% agarose gel and digested overnight with DpnI.

#### 4.6.2 Mutagenesis by isothermal assembly (Gibson Assembly)

The isothermal assembly mix was prepared and site-directed mutagenesis was performed as described in the original protocol (Gibson et al., 2009). DNA fragment overlapping by 18 nt were prepared by PCR amplification (Shampo and Kyle, 2002). The PCR fragments were purified with NucleoSpin PCR cleanup kit, the fragment concentration was determined by absorption at 260 nm, and equal amounts of each fragment (0.01-0.05 pmol) were used. Isothermal assembly reaction was performed for 25 min at 55°C and digested with DpnI overnight.

#### 4.6.3 Transformation

Chemically competent cells (50  $\mu$ L) were thawed on ice, DNA solution (2 -7  $\mu$ L) was added, cells were incubated with the DNA for 30 min, and transformed by heat-shock for 30 s at 42°C. Following heat shock, cells were incubated for 5 min at 4°C. LB media (500  $\mu$ L) was added and cells were incubated for 60 min at 190 rpm, 37°C. Cells were pelleted for 5 min at 3'000 rpm in an Eppendorf benchtop centrifuge and resuspended in LB (100  $\mu$ L) and plated on LB agar plates supplemented with the appropriate selection antibiotics.

## Material and Methods

### 4.7 DNA constructs

#### 4.7.1 HemK constructs

The unmodified nucleotide sequence of HemK from *E. coli* strain BL21(DE3) was used for the investigation of translation intermediates in Section 2.2. All constructs used for incorporation of fluorescence-labeled Cysteinyl-tRNA in Section 2.3 were derived from a plasmid with the mutations Y3F / W6F / C52S / C97S / C111S / C134S / C261S prepared by Jörg Mittelstätt.

#### 4.7.2 Subcloning of gene for tRNA<sup>Cys</sup>

The gene coding for tRNA<sup>Cys</sup> with an upstream T7-promotor was generated by primer-overlap PCR with phosphorylated primers tRNA<sub>cys</sub>\_F/R. The PCR fragment was purified with NucleoSpin PCR cleanup kit. Vector pUC19 was digested with SmaI and dephosphorylated with antarctic phosphatase. The insert was blunt-end ligated with QuickLigase. A clone was selected that had the proper orientation to be amplified with standard M13 forward and tRNA<sub>cys</sub>\_trans\_R primer (Table 4-8).

#### 4.7.3 Subcloning of gene for adenylate kinase

The gene coding for Adenylate kinase (*adk*) was amplified from *E. coli* DH5α genomic DNA with primers AK\_sub\_F/R and inserted between the NdeI and XhoI sites of pET24a.

#### 4.7.4 Subcloning of gene for phosphoglycerate kinase

The gene coding for Phosphoglycerate kinase (*pgk*) was amplified from *E. coli* DH5α genomic DNA with primer PGK\_sub\_F /R and inserted between the NdeI and XhoI sites of pET24a.

#### 4.7.5 Subcloning of the gene for codon-optimized HemK

The synthetic gene coding for the codon-optimized HemK was designed by Jörg Mittelstätt and ordered from Eurofins MWG operon in vector pEX-A. For this thesis, the gene was inserted into pET24a by isothermal assembly (Section 4.6.2) using primers HemK\_opt\_sub\_F/R complementary to both HemK opt and pET24a.

## Material and Methods

### 4.7.6 Site-directed mutagenesis for tRNA<sup>Cys</sup> constructs

Anticodon sequences for amber, ochre, and opal suppressor tRNA<sup>Cys</sup> genes were introduced by QuikChange mutagenesis (Section 4.6.1) using the primer pairs tCys\_amber\_m\_F/R, tCys\_ochre\_m\_F/R, or tCys\_opal\_m\_F/R. The mutation of U32C for amber suppressor tRNA<sup>Cys</sup> was introduced by primer pair tCys\_amb\_U32C\_m\_F/R.

### 4.7.7 Site-direct mutagenesis for HemK constructs

The mutations Q4C, R8C and K34C were introduced by QuikChange mutagenesis using primer pairs HemK\_Y3F\_Q4C\_m\_F/R, HemK\_Y3F\_R8C\_m\_F/R or HemK\_K34C\_m\_F/R.

Mutagenesis at position 33 was performed by isothermal assembly with the combination of primer HemK\_Y3F\_W6F\_m\_F with primers HemK\_G33X\_K34C\_m\_R and by the combination of primer HemK\_Y3F\_W6F\_m\_F with primers HemK\_G33X\_K34C\_m\_R.

Substitution of residues at position 30-32 to glycine was performed by isothermal assembly with the combination of primer HemK\_Y3F\_W6F\_m\_F with primers of the family HemK\_30<sup>H</sup>/<sub>G</sub>\_31<sup>V</sup>/<sub>G</sub>\_32<sup>T</sup>/<sub>G</sub>\_K34C\_m\_R and by the combination of primer HemK\_Y3F\_W6F\_m\_R with primers of the family HemK\_30<sup>H</sup>/<sub>G</sub>\_31<sup>V</sup>/<sub>G</sub>\_32<sup>T</sup>/<sub>G</sub>\_K34C.

### 4.7.8 Site-direct mutagenesis for GBC constructs

The mutation of T88Amb was introduced into the U and H sequence of GBC by isothermal assembly by combining primer T7Promotor\_reverse\_comp with the respective forward primer GBC\_H/U\_T88Amb\_m\_F and by the combination of T7\_forward primer with the respective reverse primers GBC\_H/U\_T88Amb\_m\_R.

Table 4-8 List of Primers

Name	Sequence 5'-3'
AK_sub_F	GGG GAT TTT CCA TAT GCG TAT CAT TCT GCT TGG C
AK_sub_R	CGC TTT CCT CGA GGC CGA GGA TTT TTT CCA G
AK_trans_R	GCC GAG GAT TTT TTC CAG ATC
GBC_H_1-89_trans_R	AAA AGT ACC GGT ATG TTG TGG GAT TAA G
GBC_H_T88Amb_m_F	CCC ACA ACA TAC CGG TTA GTT TCG GAT GCG G
GBC_H_T88Amb_m_R	CCG CAT CCG AAA CTA ACC GGT ATG TTG TGG G
GBC_U_1-89_trans_R	GAA AGT GCC GGT GTG TTG CG
GBC_U_T88Amb_m_F	CGC AAC ACA CCG GCT AGT TCA GAA TGA GAA TCT ATG
GBC_U_T88Amb_m_R	CAT AGA TTC TCA TTC TGA ACT AGC CGG TGT GTT GCG
GBC_U_trans_R	ATA AAA ATC CAT CAC CCG TCT TAA AGA ACC
HemK_1_210_trans_R	TGC CGC AAC CAG CGC AGT
HemK_1-112_trans_R	ACG GCA AGG TTG TTC AGG CA

## Material and Methods

HemK_1-126_trans_R	CGC AAG CGC AAT CGC CC
HemK_1-14_trans_R	AAG TTG GCT TAT TGC TTC ACG
HemK_1-140_trans_R	ATC TAC AGC GGT AAT TTC GCA GT
HemK_1-154_trans_R	CTG GGC ATT ACG TTG TGC CAG AG
HemK_1-168_trans_R	CCA GTC GCT TTG CAG AAT GCG
HemK_1-182_trans_R	GCT GAC AAT CAT CGT AAA CTG CT
HemK_1-196_trans_R	TTG TTG AAG ATG TGG GTC CTG CT
HemK_1-224_trans_R	CGA CTG TTC GAT GAT ATG CAC GA
HemK_1-238_trans_R	ATG TTC CAG AAG CAG AAA GCC G
HemK_1-252_trans_R	GAG GAT AAA TGC TTG TCG CAC CG
HemK_1-268_trans_R	ATC ACC ATA GTC ACG GCA GG
HemK_1-277_trans_R	TTG ATA ATA GCG GCC GAG CGT T
HemK_1-28_trans_R	CAG CAG GAT TTC AGC ATC ACG C
HemK_1-42_trans_R	AAA GGC GAG GAT AAA AGT ACG C
HemK_1-56_trans_R	ATC AAG TTG CTG ACA TTG TTC G
HemK_1-70_trans_R	ATG AGC AAT GGG TTC ACC ATC G
HemK_1-84_trans_R	AAC AAA TAA CGG CAA CGA CCA GA
HemK_1-98_trans_R	CAG ACA CTC CGT ATC CGG G
HemK_G33A_K34C_m_F	CTG CTG GAA CAT GTT ACC GCA TGC GGG CGT ACT TTT ATC
HemK_G33A_K34C_m_R	GAT AAA AGT ACG CCC GCA TGC GGT AAC ATG TTC CAG CAG
HemK_G33D_K34C_m_F	CTG CTG GAA CAT GTT ACC GAC TGC GGG CGT ACT TTT ATC
HemK_G33D_K34C_m_R	GAT AAA AGT ACG CCC GCA GTC GGT AAC ATG TTC CAG CAG
HemK_G33E_K34C_m_F	CTG CTG GAA CAT GTT ACC GAA TGC GGG CGT ACT TTT ATC
HemK_G33E_K34C_m_R	GAT AAA AGT ACG CCC GCA TTC GGT AAC ATG TTC CAG CAG
HemK_G33F_K34C_m_F	GCT GGA ACA TGT TAC CTT CTG CGG GCG TAC TTT TAT C
HemK_G33F_K34C_m_R	GAT AAA AGT ACG CCC GCA GAA GGT AAC ATG TTC CAG C
HemK_G33H_K34C_m_F	CTG CTG GAA CAT GTT ACC CAC TGC GGG CGT ACT TTT ATC
HemK_G33H_K34C_m_R	GAT AAA AGT ACG CCC GCA GTG GGT AAC ATG TTC CAG CAG
HemK_G33I_K34C_m_F	CTG CTG GAA CAT GTT ACC ATC TGC GGG CGT ACT TTT ATC
HemK_G33I_K34C_m_R	GAT AAA AGT ACG CCC GCA GAT GGT AAC ATG TTC CAG CAG
HemK_G33K_K34C_m_F	CTG CTG GAA CAT GTT ACC AAA TGC GGG CGT ACT TTT ATC
HemK_G33K_K34C_m_R	GAT AAA AGT ACG CCC GCA TTT GGT AAC ATG TTC CAG CAG
HemK_G33L_K34C_m_F	CTG CTG GAA CAT GTT ACC CTG TGC GGG CGT ACT TTT ATC
HemK_G33L_K34C_m_R	GAT AAA AGT ACG CCC GCA CAG GGT AAC ATG TTC CAG CAG
HemK_G33M_K34C_m_F	CTG CTG GAA CAT GTT ACC ATG TGC GGG CGT ACT TTT ATC
HemK_G33M_K34C_m_R	GAT AAA AGT ACG CCC GCA CAT GGT AAC ATG TTC CAG CAG
HemK_G33N_K34C_m_F	CTG CTG GAA CAT GTT ACC AAC TGC GGG CGT ACT TTT ATC
HemK_G33N_K34C_m_R	GAT AAA AGT ACG CCC GCA GTT GGT AAC ATG TTC CAG CAG
HemK_G33P_K34C_m_F	GCT GGA ACA TAGT TAC CCC GTG CGG GCG TAC TTT TAT C
HemK_G33P_K34C_m_R	GAT AAA AGT ACG CCC GCA CGG GGT AAC ATG TTC CAG C
HemK_G33Q_K34C_m_F	CTG CTG GAA CAT GTT ACC CAG TGC GGG CGT ACT TTT ATC
HemK_G33Q_K34C_m_R	GAT AAA AGT ACG CCC GCA CTG GGT AAC ATG TTC CAG CAG
HemK_G33R_K34C_m_F	CTG CTG GAA CAT GTT ACC CGT TGC GGG CGT ACT TTT ATC
HemK_G33R_K34C_m_R	GAT AAA AGT ACG CCC GCA ACG GGT AAC ATG TTC CAG CAG
HemK_G33S_K34C_m_F	CTG CTG GAA CAT GTT ACC AGC TGC GGG CGT ACT TTT ATC
HemK_G33S_K34C_m_R	GAT AAA AGT ACG CCC GCA GCT GGT AAC ATG TTC CAG CAG
HemK_G33T_K34C_m_F	CTG CTG GAA CAT GTT ACC ACC TGC GGG CGT ACT TTT ATC
HemK_G33T_K34C_m_R	GAT AAA AGT ACG CCC GCA GGT GGT AAC ATG TTC CAG CAG
HemK_G33V_K34C_m_F	CTG CTG GAA CAT GTT ACC GTA TGC GGG CGT ACT TTT ATC
HemK_G33V_K34C_m_R	GAT AAA AGT ACG CCC GCA TAC GGT AAC ATG TTC CAG CAG
HemK_G33W_K34C_m_F	CTG CTG GAA CAT GTT ACC TGG TGC GGG CGT ACT TTT ATC
HemK_G33W_K34C_m_R	GAT AAA AGT ACG CCC GCA CCA GGT AAC ATG TTC CAG CAG
HemK_G33Y_K34C_m_F	CTG CTG GAA CAT GTT ACC TAC TGC GGG CGT ACT TTT ATC
HemK_G33Y_K34C_m_R	GAT AAA AGT ACG CCC GCA GTA GGT AAC ATG TTC CAG CAG
HemK_GGGF_K34C_m_F	CTG GAA GGT GGC GGT TTC TGC GGG CGT AC
HemK_GGGF_K34C_m_R	CCC GCA GAA ACC GCC ACC TTC CAG CAG G
HemK_GGGG_K34C_m_F	CTG GAA GGT GGC GGT GGC TGC GGG CGT ACT TTT ATC C
HemK_GGGG_K34C_m_R	CCC GCA GCC ACC GCC ACC TTC CAG CAG GAT TTC AGC
HemK_GGTG_K34C_m_F	CTG GAA GGT GGC ACC GGC TGC GGG CGT ACT TTT ATC C

## Material and Methods

HemK_GGTG_K34C_m_R	CCC GCA GCC GGT GCC ACC TTC CAG CAG GAT TTC AGC
HemK_GVGG_K34C_m_F	CTG GAA GGT GTT GGT GGC TGC GGG CGT ACT TTT ATC C
HemK_GVGG_K34C_m_R	CCC GCA GCC ACC AAC ACC TTC CAG CAG GAT TTC AGC
HemK_GVTG_K34C_m_F	CTG GAA GGT GTT ACC GGC TGC GGG CGT ACT TTT ATC C
HemK_GVTG_K34C_m_R	CCC GCA GCC GGT AAC ACC TTC CAG CAG GAT TTC AGC
HemK_HGGF_K34C_m_F	CTG GAA CAT GGT GGT TTC TGC GGG CGT ACT TTT ATC C
HemK_HGGF_K34C_m_R	CCC GCA GAA ACC ACC ATG TTC CAG CAG GAT T
HemK_HGGG_K34C_m_F	CTG GAA CAT GGT GGT GGC TGC GGG CGT ACT TTT ATC C
HemK_HGGG_K34C_m_R	CCC GCA GCC ACC ACC ATG TTC CAG CAG GAT TTC AGC
HemK_HGTG_K34C_m_F	CTG GAA CAT GGT ACC GGC TGC GGG CGT ACT TTT ATC C
HemK_HGTG_K34C_m_R	CCC GCA GCC GGT ACC ATG TTC CAG CAG GAT TTC AGC
HemK_HVGF_K34C_m_F	CTG GAA CAT GTT GGT TTC TGC GGG CGT ACT TTT ATC C
HemK_HVGF_K34C_m_R	CCC GCA GAA ACC AAC ATG TTC CAG CAG GAT T
HemK_HVGG_K34C_m_F	CTG GAA CAT GTT GGT GGC TGC GGG CGT ACT TTT ATC
HemK_HVGG_K34C_m_R	GAT AAA AGT ACG CCC GCA GCC ACC AAC ATG TTC CAG
HemK_K34Amb_m_F	GGA ACA TGT TAC CGG CTA GGG GCG TAC TTT TAT CC
HemK_K34Amb_m_R	GGA TAA AAG TAC GCC CCT AGC CGG TAA CAT GTT CC
HemK_K34C_m_F	CTG GAA CAT GTT ACC GGC TGC GGG CGT ACT TTT ATC C
HemK_K34C_m_R	GGA TAA AAG TAC GCC CC AGC CGG TAA CAT GTT CCA G
HemK_opt_1-277_trans_R	ATA ATA ACG ACC CAG GGT AAC ACG TTC AT
HemK_opt_sub_F	CTT TAA GAA GAT ATA CAT ATG GAA TAT CAG CAT
HemK_opt_sub_R	GTG GTG GTG GTG GTG CTC GAG CTG ATA ATA ACG ACC
HemK_Y3F_Q4C_m_F	GAA GGA GAT ATA CAT ATG GAA TTC TGC CAC TTT TTA CGT G
HemK_Y3F_Q4C_m_R	GCT TCA CGT AAA AAG TGG CAG AAT TCC ATA TGT ATA TCT CCT TC
HemK_Y3F_R8C_m_F	GGA GAT ATA CAT ATG GAA TTC CAA CAC TTT TTA TGC G
HemK_Y3F_R8C_m_R	CGC ATA AAA AGT GTT GGA ATT CCA TAT GTA TAT CTC C
HemK_Y3F_W6F_m_F	GGA GAT ATA CAT ATG GAA TTC CAA CAC TTT TTA CGT GAA GC
HemK_Y3F_W6F_m_R	GCT TCA CGT AAA AAG TGT TGG AAT TCC ATA TGT ATA TCT CC
M13_Forward	GTA AAA CGA CGG CCA GT
PGK_1-208_trans_R	AGC GAT TTT AGA CAG GGA GTC CAG AAC
PGK_sub_F	GAG GAT TCC ATA TGT CTG TAA TTA AGA TGA CCG ATC TGG
PGK_sub_R	CCT GTG ATC TCG AGC TTC TTA GCG CGC TCT TC
PGK_trans_R	CCT CCT AGC GCG CTC TTC GA
T7_forward	AGT TGC TGC AGT AAT AC
T7Promotor_reverse_comp	GTA TTA CTG CAG CAA CT
tCys_amb_U32C_m_F	GGT TAT GTA GCG GAC TCT AAA TCC GTC TAG TCC G
tCys_amb_U32C_m_R	CGG ACT AGA CGG ATT TAG AGT CCG CTA CAT AAC C
tCys_amber_m_F	CGG TTA TGT AGC GGA TTC TAA ACT CGT CTA GTC CG
tCys_amber_m_R	CCG GAC TAG ACG GAT TTA GAA TCC GCT ACA TAA CC
tCys_ochre_m_R	CCG GAC TAG ACG GAT TTA AAA TCC GCT ACA TAA CC
tCys_ochre_m_F	CGG TTA TGT AGC GGA TTT TAA ACT CGT CTA GTC CG
tCys_opal_m_F	CGG TTA TGT AGC GGA TTT CAA ATC CGT CTA GTC CG
tCys_opal_m_R	CCG GAC TAG ACG GAT TTG AAA TCC GCT ACA TAA CC
tCys_R_trans	TGG AGG CGC GTT CCG GAG TC
tCys_sub_F	Phosphate-AGT TGC TGC AGT AAT ACG ACT CAC TAT AGG CGC GTT AAC AAA GCG GTT ATG TAG CGG ATT G
tCys_sub_R	Phosphate -TGG AGG CGC GTT CCG GAG TCG AAC CGG ACT AGA CGG ATT TGC AAT CCG CTA CAT AAC CGC T

## Material and Methods

### 4.8 General RNA protocols

#### 4.8.1 Ethanol or isopropanol precipitation

RNA was precipitated by the addition of  $1/5$  volume of 20% (w/v) KOAc pH 5.0 and 2.5 volumes of EtOH. Alternatively, RNA was precipitated with  $1/5$  volume 20% (w/v) KOAc pH 5.0 and one volume of *iso*-propanol.

#### 4.8.2 Phenol or phenol:chloroform extraction

One volume of phenol or phenol:chloroform (50:50) mixture and  $1/5$  volume of 20% (w/v) KOAc pH 5.0 were added to the RNA solution. Phases were mixed using a vortex and separated by centrifugation for 3 min at 4'000 g, 16°C. The aqueous phase was removed and the organic phase was re-extracted with  $1/5$  volume 20% (w/v) KOAc pH 5.0 and  $4/5$  volumes ddH<sub>2</sub>O.

#### 4.8.3 General protocol for *in vitro* transcription

PCR templates were generated by PCR amplification using Phusion polymerase (2 units) in high fidelity buffer, dNTP (0.2 mM), primer (0.2  $\mu$ M) each, and plasmid template (0.6 ng  $\mu$ l<sup>-1</sup>) (Shampo and Kyle, 2002). PCR program: 3 min 98°C, 35 cycles (30 s 98°C, 20 s 57°C, 30 s kb<sup>-1</sup> 72°C), 5 min 72°C. The *in vitro* transcription mixture contained  $1/10$  volume of PCR amplified DNA template, DTT (10 mM), ATP, GTP, CTP, UTP (3 mM each), GMP (5 mM), pyrophosphatase (0.01 U  $\mu$ L<sup>-1</sup>), RNAsin (20  $\mu$ L mL<sup>-1</sup>) and T7 RNA polymerase (1.6 U  $\mu$ L<sup>-1</sup>) in 1xTranscription buffer (Sampson and Uhlenbeck, 1988). Transcription was performed at 37°C for 3 h and products were purified on HiTRAP Q 5 mL columns or by using the RNeasy Midi kit.

## Material and Methods

### 4.9 mRNA constructs

#### 4.9.1 Short synthetic mRNA

The short synthetic mRNA listed in Table 4-9 were ordered from IBA, Göttingen.

Table 4-9 List of synthetic mRNA. The coding sequence is underlined.

mRNA	Sequence
amber	G GCA AGG AGG UAA AUA <u>AUG UAG ACG AUU</u>
m022CUC	G UUA ACA GGU AUA CAU ACU <u>AUG CUC GUU AUU ACU</u>
m022MCF	G UUA ACA GGU AUA CAU ACU <u>AUG UGC UUU AC</u>
ochre	G GCA AGG AGG UAA AUA <u>AUG UAA ACG AUU</u>
opal	G GCA AGG AGG UAA AUA <u>AUG UGA ACG AUU</u>

#### 4.9.2 mRNA for single-turnover translation

DNA templates for *in vitro* transcription were amplified from intact plasmids as described in 4.8.3 with primer pairs listed in Table 4-10.

Table 4-10: Primer pairs for amplification of DNA templates uses for *in vitro* transcription.

Construct	forward Primer	Reverse Primer
AK 1-214	T7_forward	AK_trans_R
CspA 1-70	T7_forward	CspA_trans_R
GBC H 1-175, incl. mutations	T7_forward	GBC_H_trans_R
GBC H 1-89	T7_forward	GBC_H_1-89_trans_R
GBC U 1-175, incl. mutations	T7_forward	GBC_U_trans_R
GBC U 1-89	T7_forward	GBC_U_1-89_trans_R
HemK 1-112	T7_forward	HemK_1-112_trans_R
HemK 1-126	T7_forward	HemK_1-126_trans_R
HemK 1-14	T7_forward	HemK_1-14_trans_R
HemK 1-140	T7_forward	HemK_1-140_trans_R
HemK 1-154	T7_forward	HemK_1-154_trans_R
HemK 1-168	T7_forward	HemK_1-168_trans_R
HemK 1-182	T7_forward	HemK_1-182_trans_R
HemK 1-196	T7_forward	HemK_1-196_trans_R
HemK 1-210	T7_forward	HemK_1-210_trans_R
HemK 1-224	T7_forward	HemK_1-224_trans_R
HemK 1-238	T7_forward	HemK_1-238_trans_R
HemK 1-252	T7_forward	HemK_1-252_trans_R
HemK 1-266	T7_forward	HemK_1-266_trans_R
HemK 1-277	T7_forward	HemK_1-277_trans_R
HemK 1-28	T7_forward	HemK_1-28_trans_R
HemK 1-42	T7_forward	HemK_1-42_trans_R
HemK 1-56	T7_forward	HemK_1-56_trans_R
HemK 1-70, incl. mutations	T7_forward	HemK_1-70_trans_R
HemK 1-84	T7_forward	HemK_1-84_trans_R
HemK 1-98	T7_forward	HemK_1-98_trans_R
HemK opt 1-277	T7_forward	HemK_opt_1-277_trans_R
PGK 1-208	T7_forward	PGK_1-208_trans_R
PGK 1-384	T7_forward	PGK_trans_R
tRNA <sup>Cys</sup> , incl. mutations	M13 fw	tRNACys_trans_R

## Material and Methods

### 4.10 Protein purification

#### 4.10.1 CysRS-His<sub>6</sub>

CysRS-His<sub>6</sub> was expressed and purified by Theresia Uhlendorf (technical assistant, Department of Physical Biochemistry, MPI for Biophysical Chemistry). The plasmid pCA24N carrying the gene for CysRS with a C-terminal His<sub>6</sub> tag was transformed into BL21(DE3). For overexpression, cultures were inoculated to 0.1 OD<sub>600nm</sub> with an overnight culture. At OD<sub>600nm</sub> of 0.8, overexpression was induced with isopropyl β-D-1-thiogalactopyranoside (IPTG, 0.5 mM). CysRS-His<sub>6</sub> was overexpressed for 3h at 37°C and cells were pelleted by centrifugation for 20 min at 6'000 rpm in a Beckmann JLA 8.1000 rotor. Cell pellet was resuspended in Lysis buffer, and opened by pressure with Emusiflex, cell lysate was cleared by centrifugation for 20 min at 25'000 rpm in a JA 25.50 Beckmann rotor and loaded on Ni-IDA resin in a 35 mL gravity flow columns equilibrated with buffer IMAG supplemented with imidazole (10 mM). The column was washed with 40 mL of IMAG buffer supplemented with imidazole (10 mM) and eluted with 20 mL of IMAG buffer supplemented with imidazole (250 mM). The protein was concentrated in a Vivaspin 20 concentrator with a 5 kDa molecular weight cut-off (MWCO) and dialyzed into 1xHAKM7 buffer supplemented with mercapto-ethanol (6 mM) and glycerol (10% (v/vol)).

#### 4.10.2 Initiation and elongation factors

Initiation factors IF1, IF2, and IF3 were purified as described (Milon et al., 2007). EF-Tu was purified as binary complex of EF-Tu•GDP as described (Rodnina and Wintermeyer, 1995). EF-G was purified according to the published protocol (Wilden et al., 2006).



## Material and Methods

### 4.11 Protocols related to tRNA

#### 4.11.1 TCA precipitation of aminoacyl-tRNA

aa-tRNA solution (50  $\mu$ L) was precipitated with cold 10% (w/v) TCA, 50% EtOH (vol/v) (500  $\mu$ L). BSA (10 mg mL<sup>-1</sup>, 10  $\mu$ L) was used as co-precipitant. Samples were incubated for 30 min on ice, spotted on pre-moisten nitrocellulose filters and washed with 5% (w/v) TCA (5 mL). Filters were dissolved in QuickZint 361 (10 mL) and subjected to scintillation counting.

#### 4.11.2 Formation of EF-Tu•GTP•aa-tRNA ternary complex

Ternary complexes were formed in TAKM<sub>7</sub> or HiFi buffer supplemented with GTP (2 mM), Phosphoenolpyruvate (PEP) (6 mM), and pyruvate kinase (1%). In order to keep the desired concentration of free Mg<sup>2+</sup> ions, one eq. of MgCl<sub>2</sub> was added per eq. of GTP. Similarly, 0.5 eq. MgCl<sub>2</sub> were added for each eq. of PEP. An excess of EF-Tu•GTP (2.5 eq) was used for ternary complex formation with aa-tRNA (1 eq). Binary complex of EF-Tu•GTP was formed by incubation of EF-Tu•GDP in presence of PEP and pyruvate kinase for 15 min at 37°C before the aa-tRNA was added in order to insure complete turnover of GDP. Ternary complexes were allowed to form by incubation of EF-Tu•GTP with aa-tRNA for 2 min at 37°C.

#### 4.11.3 Native PAGE analysis of aminoacylation level

Analysis of aminoacylation level by separation of the ternary complex on native PAGE was performed as described (Ohtsuki et al., 2010). EF-Tu•GTP binary complex was formed in 1xNative PAGE TC buffer. For each lane, EF-Tu•GTP (60 pmol, 5  $\mu$ L) was mixed with 5x Native PAGE TC buffer (2  $\mu$ L) and aa-tRNA solution (3  $\mu$ L) followed by incubation for 2 min at 37°C. Native PAGE loading buffer (2  $\mu$ L) was added prior loading. The native PAGE was run for 2h at 4°C at 100 V constant, stained with Commassie and destained with 5% (vol/v) acetic acid, 10% (vol/v) EtOH.

#### 4.11.4 Analysis of aa-tRNA by Acid Urea PAGE

Acid urea PAGE was performed as described (Janssen et al., 2012). 2x Acid PAGE loading buffer was added to samples. Gels were prepared with a separating layer of Acid PAGE 6.5% gel and a stacking layer of Acid PAGE 3.5% gel. Gels were run for 2.5 h at 120 V constant, 4°C with 1xAcid PAGE running buffer.

## Material and Methods

### 4.11.5 Analysis of peptidyl-tRNA by Bis-Tris PAGE

Bis-Tris PAGE was modified from (Chadani et al., 2012). 2x SDS loading buffer was added to samples. Gels were assembled with a separating layer (16.5% (w/vol) acrylamide, 13% (v/vol) glycerol), a spacer layer (10 % (w/vol) acrylamide) and a stacking layer (4 % (w/vol) acrylamide). All layers contained Bis-acrylamide (3% (w/vol)), Bis-Tris-HCl pH 6.4 (0.35 M) and SDS (0.3% (v/vol)). 1xBis-Tris running buffer was used for electrophoresis. Gels were run at 30°C, 40 min at 36 V (by limiting the voltage), 1 h at 75 V and 1 h at 150 V. After electrophoresis, gels were washed in ddH<sub>2</sub>O and imaged by FLA-9000 fluorescence scanner (GE healthcare).

### 4.11.6 Aminoacylation of total tRNA

Total tRNA from *E. coli* MRE-600 (Roche) was phenol extracted and precipitated with EtOH prior to aminoacylation. aa stocks of L-Ala, L-Arg, L-Asn, L-Cys, L-Lys, L-Gln, L-Gly, L-His, L-Met, L-Pro, L-Ser, L-Thr, L-Ile, and L-Val were prepared at a concentration of 200 mM in ddH<sub>2</sub>O. L-Tyr was prepared at a concentration of 140 mM in ddH<sub>2</sub>O. L-Asp, L-Glu, and L-Phe were prepared at a concentration of 180 mM in 100 mM Tris-HCl pH 9.0 at 50°C. L-Trp and L-Leu were prepared in 300 mM Tris-HCl pH 9.0 at 37°C. aa were added to a final concentration of 0.3 mM in the aminoacylation reaction. total tRNA (80 OD<sub>260nm</sub> mL<sup>-1</sup>) was aminoacylated with S-100 enzymes (3% (v/vol)) in TAKM<sub>20</sub> buffer pH 7.5 supplemented with ATP (3 mM) and DTT (2 mM) for 30 min at 37°C. tRNA was extracted with phenol and precipitated with EtOH. The tRNA pellet was resuspended in ddH<sub>2</sub>O. Total aa-tRNA was purified on three serial HiTRAP Q 5 mL columns using an ÄktaPurifier FPLC system. Total aa-tRNA was loaded on columns equilibrated in buffer RNA-Q-A and was eluted with a linear gradient from 0 to 100% RNA-Q-B buffer in 40 min at a flow-rate of 2 mL min<sup>-1</sup>. Fractions containing tRNA were pooled and precipitated with EtOH. The concentration of total aa-tRNA was determined by an extinction coefficient  $\epsilon_{260\text{ nm}}$  of 1'730'000 M<sup>-1</sup> cm<sup>-1</sup>.

### 4.11.7 Purification of tRNA<sup>Cys</sup> transcripts

tRNA<sup>cys</sup> variants were purified on 5 mL HiTRAP Q column using an ÄktaPurifier FPLC system. tRNA was loaded on column equilibrated in buffer RNA-Q-A and eluted with a linear gradient of 12 CV from 0 to 100% buffer RNA-Q-B. Fractions of 0.4 CV were collected, separated on 12% UREA PAGE and visualized by staining with methylene blue. Fractions containing tRNA<sup>Cys</sup> variants were pooled and precipitated with isopropanol.

## Material and Methods

### 4.11.8 Purification of tRNA<sup>Cys</sup> from *E. coli* total tRNA

Total *E. coli* tRNA (Sigma) was fractionated on Sepharose4B column followed by fractionation on Phenyl-Sepharose column (Holmes et al., 1975; Mittelstaet et al., 2013). Fractions were analyzed for aminoacylation with cysteine. Fractions that contained Cys-tRNA after separation on Phenyl-Sepharose column were aminoacylated in a preparative scale and the ternary complex of EF-Tu•GTP•Cys-tRNA<sup>Cys</sup> was purified on a Superdex S-75 size exclusion column. The concentration of ternary complex was determined according to the relation of 2.36  $\mu\text{M}$  ternary complex per OD<sub>260nm</sub> mL<sup>-1</sup>.

### 4.11.9 Aminoacylation of tRNA<sup>Cys</sup> variants

Aminoacylation of tRNA<sup>Cys</sup> variants (35 OD<sub>260nm</sub> mL<sup>-1</sup>) was performed in 1x aminoacylation buffer supplemented with ATP (2 mM), DTT (5 mM) and L-cysteine (150  $\mu\text{M}$ ) for 40 min at 37°C. For aminoacylation of tRNA<sup>Cys<sub>GCA</sub></sup>, S-100 enzymes (5% (v/vol)) were used. Suppressor tRNA<sup>Cys</sup> variants were aminoacylated with recombinant CysRS-His<sub>6</sub> (6  $\mu\text{M}$ ). For aminoacylation with [<sup>35</sup>S]cysteine, fresh stocks with an initial activity of 1000 disintegrations per minute (dpm) /pmol were prepared.

### 4.11.10 Labeling of Cys-tRNA with thiol-reactive probes.

Labeling of Cys-tRNA variants (30  $\mu\text{M}$ ) was performed with maleimide-functionalized fluorescent probes (1 mM) in HEPES pH 7.5 (15 mM) and 80% (v/vol) DMF for 10 min at 25°C. Alternatively, labeling of Cys-tRNA (30  $\mu\text{M}$ ) was performed with iodoacetamide-functionalized fluorescent probes (1 mM) in HEPES pH 7.5 (15 mM) and 50% (v/vol) DMF for 45 min at 25°C. Unbound fluorophore was removed according to different procedures depending on hydrophobicity of the fluorescent label. Moderately hydrophobic fluorescent probes could be removed by buffer exchange on Nap-5 column. Hydrophobic fluorescent probes were efficiently removed by extraction with a 50:50 phenol:chloroform mixture. Following purification from unlabeled and deacylated tRNA by HPLC, the concentration of labeled Cys-tRNA was calculated according to a specific activity of 1600 pmol/OD<sub>260nm</sub> determined experimentally from fMet-Cys dipeptide formation.

## Material and Methods

### 4.11.11 Labeling of the $\alpha$ -amino group of Met-tRNA<sup>fMet</sup>

Labeling of Met-tRNA<sup>fMet</sup> (30  $\mu$ M) with Atto610-NHS (2 mM) and Atto655-NHS (2.2 mM) was performed in NaHCO<sub>3</sub> pH 8.3 (30 mM), 50% (vol/v) DMF for 3 h at 4°C. Labeling of Met-tRNA<sup>fMet</sup> (30  $\mu$ M) with Bodipy576/589-C<sub>2</sub>-NHS (5 mM) was performed in NaOAc pH 5.0 (100 mM), 80% (vol/v) DMSO for 18 h at 4°C. Following the labeling reaction, unbound fluorophore was removed by extraction with 50:50 phenol:chloroform and tRNA was precipitated with isopropanol.

### 4.11.12 Separation of fluorescence-labeled tRNA by HPLC

Fluorescence-labeled tRNA was purified on LiChroCart 250-4, LiChrospher WP300 RP-18 column by using step gradient from buffer A (HPLC 5%) to buffer B (HPLC 40%) at a flow rate of 0.5 mL min<sup>-1</sup>: 15 min 0% B constant, 2 min 0-30% B, 20 min 30% B constant, 30 min 30-100% B.

## 4.12 Ribosomal complexes

### 4.12.1 Formation of 70S initiation complexes

70S ribosomes were purified as described in (Rodnina and Wintermeyer, 1995) and initiator tRNA fMet-tRNA<sup>fMet</sup> was prepared as described (Milon et al., 2007). 70S IC were formed at a concentration of 0.5 to 1  $\mu$ M in TAKM<sub>7</sub> buffer supplemented with DTT (2 mM) and GTP (1 mM) for 60 min at 37°C. For 70S IC•fMet-tRNA<sup>fMet</sup>, [<sup>3</sup>H]Met-tRNA<sup>fMet</sup> (1.5 eq.), IFs (1.5 eq.) each, and mRNA (3 to 4 eq.) and 70S ribosomes (1 eq.) were used. For 70S•IC with fluorescence-labeled [<sup>3</sup>H]Met-tRNA<sup>fMet</sup>, fluorescence-labeled initiator tRNA (1 eq.), IFs (1.5 eq.) each, mRNA (3 to 4 eq.), and 70S ribosomes (1.5 eq.) were used. 70S IC formation was analyzed by spotting of IC (2 to 10 pmol) on nitrocellulose filters, filter were washed with cold TAMK<sub>7</sub> (5 mL), and scintillation counted with QuickZint 361 cocktail.

### 4.12.2 Purification of 70S initiation complexes

70S IC was purified by centrifugation through 1.1 M sucrose cushion in TAKM<sub>7</sub> in a Beckmann TLA-55 rotor for 2 h at 55'000 rpm. Pellets were resuspended for 30 min in TAKM<sub>7</sub> buffer and aliquots were flash-frozen in liquid nitrogen and stored at -80°C. The concentration of IC was re-examined by nitro-cellulose filtration after thawing of aliquots.

## Material and Methods

### 4.12.3 Separation of dipeptides

Dipeptide formation was stopped by the addition of KOH (1 M, 1 volume, or 0.8 M in the case of quenching of reaction time points in the quench-flow apparatus). RNA components were hydrolyzed during incubation at 37°C for 30 min. Samples were acidified by the addition of AcOH (1/10 volume) and separated on LiChroCart 250-4, LiChrospher 100 RP-8 column by applying a gradient of 0 to 65% acetonitrile in 0.1% TFA. Fractions were collected and subjected to liquid-scintillation counting with Irga-Safe scintillation cocktail.

### 4.13 Single-turnover *in vitro* translation

#### 4.13.1 Sample preparation

Single-turnover *in vitro* translation was performed as described (Doerfel et al., 2013; Gu et al., 2005; Mittelstaet et al., 2013). 70S IC were diluted into 1xHiFi buffer and mixed with ternary complexes and EF-G (3  $\mu$ M) at 37°C in 1x HiFi buffer. Time point (25  $\mu$ L) were removed and the reactions were quenched by addition of NaOH (2 M, 5  $\mu$ L), hydrolyzed by incubation at 37°C for 30 min, and neutralized with HEPES acid (2 M, 5.5  $\mu$ L).

#### 4.13.2 Pulse-proteolysis

Aliquots of PK were prepared in ddH<sub>2</sub>O. PK (2.7  $\mu$ M, 2  $\mu$ L) was added to time points (25  $\mu$ L) of *in vitro* translation. Proteolysis was allowed to proceed for exactly 2 min at 37°C, PK was inhibited by the addition of phenylmethylsulfonyl fluoride (PMSF) (150 mM, 2  $\mu$ L), and samples were immediately frozen in liquid nitrogen. Samples were thawed on ice in presence of RNaseA (0.4 mg mL<sup>-1</sup>, 5  $\mu$ L) and incubated for 30 min at 37°C. 2x SDS loading buffer was added and samples were separated on Tris-tricine PAGE.

#### 4.13.3 Translation *in vitro* with fluorescence-labeled Cys-tRNA

Time points of translation (25  $\mu$ L, 15 to 20 nM in 70S IC) with fluorescence-labeled tRNA<sup>Cys</sup> variants (120-200 nM) were stopped by rapid freezing in liquid nitrogen. Samples were thawed on ice in presence RNaseA (0.4 mg mL<sup>-1</sup>, 5  $\mu$ L). In order to avoid overlap of hydrolysis products with peptide bands of interest, fluorescence-labeled Cys-adenylates resulting from the digestion with RNaseA were further hydrolyzed by extended incubation for 16 h at 37°C.

## Material and Methods

### 4.13.4 Tris-tricine PAGE

Tris-tricine PAGE was performed as reported (Schagger, 2006). Gels were assembled with a separating layer (16.5% (w/vol) acrylamide, 13% (v/vol) glycerol), a spacer layer (10 % (w/vol) acrylamide) and a stacking layer (4 % (w/vol) acrylamide). All layers contained Bis-acrylamide (3% (w/vol)), Tris-HCl pH 8.45 (1 M) and SDS (0.3% (w/vol)). A discontinuous buffer system of 1x Cathode buffer and 1x Anode buffer was used for electrophoresis. Gels were run at 30°C, 1.5 h at 30 V (by limiting the voltage) and for an additional 2 to 2.5 h at 85 V. After electrophoresis, gels were washed in ddH<sub>2</sub>O and imaged by FLA-9000 fluorescence scanner (GE healthcare).

### 4.13.5 TCA precipitation of nascent chains

Single-turnover translation reaction (200 µL, 15 - 20 nM in 70S IC) was stopped by the addition of KOH (5M, 20 µL). RNA components were hydrolyzed during incubation for 30 min at 37°C. Nascent chains were precipitated by the addition of 5% TCA (600 µL) and incubation for 30 min at 4°C. Precipitated nascent chains were spotted on nitro-cellulose filters pre-moisten in 5% TCA. Filters were first washed with 5% TCA (5 mL) and then with 30% iso-propanol (5 mL) in order to remove traces of TCA that would quench the [<sup>3</sup>H] signal during scintillation counting. Filters were dissolved in QuickZint 361 scintillation liquid (10 mL) and subjected to scintillation counting.

## 4.14 Micro-array experiments

### 4.14.1 Micro-array design

Gene expression micro-array slides with 8 sub-arrays with 15k features each were ordered from Agilent Technologies. The slides contained 45 probes for *E. coli* tRNA and 7 probes for tRNA from *S. cerevisiae*. All probes were present in 140 replicates corresponding to 50% usage of the sub-arrays. The sequence of the individual probes is reported in Table 4-11.

Table 4-11: DNA probes for hybridization of tRNA on micro-array. Sequences used for Northern blotting by (Dong et al., 1996) are underlined.

Probe	Sequence 5' to 3'
Ala1B	CGAACCG <u>CAGACCTCCTGCGTGCAA</u> AGCAGGCGCTCTCCCAGCTGAGCTATAGCCCCGCA
Ala2	CGAACCG <u>CTGACCTCTTGCCATG</u> CAAGCGCTCTCCCAGCTGAGCTATAGCCCCGCC
Arg2	GAACCT <u>CCGACCGCTCGGTT</u> CGTAGCCGAGTACTCTATCCAGCTGAGCTACGGATGCCGT
Arg3	GAACCTGAGACCTCTGCCT <u>CCGG</u> AGGGCAGCGCTCTATCCAGCTGAGCTACGGGCGCCGG

## Material and Methods

Arg4	GAACCTGCGGCCACGACTTAGAAGGTCGTTGCTCTATCCAAGTGAAGGGGCGCAGA
Arg5	TCGAACTGCAATTAGCCCTTAGGAGGGGCTCGTTATATCCATTTAACTAAGAGGACAGG
Asn	CGAA <del>CC</del> AGTGACATACGGATTAACAGTCCGCCGTTCTACCGACTGAACTACAGAGGAAAC
Asp1	GAA <del>CCC</del> GCGACCCCCTGCGTGACAGGCAGGTATTCTAACCGACTGAACTACCGCTCCGAC
Cys	GTCGAA <del>CC</del> GGACTAGACGGATTGCAATCCGCTACATAACCGCTTTGTTAACGCGCCTGC
Gln1	TCGAA <del>CC</del> AGGGAATGCCGGTATCAAAAACCGGTGCCTTACCGCTTGGCGATACCCACAA
Gln2	TCGAA <del>CC</del> TCGGAATGCCGGAATCAGAATCCGGTGCCTTACCGCTTGGCGATACCCACAG
Glu2	CGAA <del>CCCC</del> TGTTACCGCCGTGAAAGGGCGGTGTCTCTGGCCTCTAGACGAAGGGGACGAA
Gly1	ATCGAA <del>CCCT</del> CGTATAGAGCTTGGGAAGCTCTCGTTCTACCATTGAACTACGCCCGCGGG
Gly2	TCGAA <del>CCCC</del> GATCATCAGCTTGGAAAGGTGAGGTAATAGCCATTATACGATGCCCGCGG
Gly3	CGAA <del>CT</del> C <del>CG</del> GACCCCAGCCTTGGCAAGGTCGTGCTCTACCAACTGAGCTATTCCC <del>CG</del> CGC
His	GAAC <del>CC</del> CAGACAAGTGAATCACAATCCAGGGCTCTACCAACTGAGCTATAGCCACCCAC
Ile1	GAAC <del>CA</del> CCGACCTCACCCCTTATCAGGGGTGCGCTCTAACACCTGAGCTACAAGCCTATC
Ile2	TGAAC <del>CC</del> AGCGACCAAGCGATTATGAGTCGCTGCTCTAACCACTGAGCTAAGGGGCCATG
Leu1	GAA <del>CCCC</del> CACGTCCGTAAGGACACTAACACCTGAAGCTAGCGCGTCTACCAATTCCGCCA
Leu2	GAA <del>CCCC</del> GTAAGCCCTATTGGGCACTACCACCTCAAGGTAGCGTGTCTACCAATTCCACCA
Leu3	GCAC <del>ACCT</del> TGCGGCGCCAGAACCCTAAATCTGGTGCCTCTACCAATTTCGCCACTCCC <del>CG</del> C
Leu4	GAA <del>CC</del> GCGCACGTATTTCTACGGTTGATTTTGAATCAACTGCGTCTACCGATTTCCGCACT
Leu5	GAA <del>CCCC</del> GACAGCGCGAACCGCCGAGGGATTTAAATCCCTTGTGTCTACCGATTCCACCA
Lys	CGAA <del>CC</del> TGCGACCAATTGATTAAAAGTCAACTGCTCTACCAACTGAGCTAACGACCCAAA
fMet1	GAACCGACGATCTT <del>CGGGT</del> TATGAGCCCCAGCAGCTACCAGGCTGCTCCACCCCGCGATG
fMet2	GAACCGACGACCTT <del>CGGGT</del> TATGAGCCCCAGCAGCTACCAGGCTGCTCCACCCCGCGATG
Met	GAAC <del>CT</del> TGTGACCCCATCATTATGAGTGATGTGCTCTAACCAACTGAGCTACGTAGCCATG
Phe	CGAACCAAGGACACGGGGATTTTCAATCCCC <del>TGCTCT</del> ACCGACTGAGCTATCCGGGCTTC
Pro1	GAA <del>CC</del> TCCGACCCCCTTCGTCCCCGAACGAAGTGCCTACCAGGCTGCGCCAATCACCGCCG
Pro2	GAA <del>CC</del> TCCGACCCCCGACACCCCATGACGGTGCCTACCAGGCTGCGCTACGTGCCGCC
Pro3	GAA <del>CC</del> TCCGACCCACTGGTCCCAAACCAGTTGCGCTACCAAGCTGCGTACTCGCCGCCA
Ser1	TGGA <del>ACC</del> CTTTCGGGTGCGCGGTTTTCAAGACCGGTGCCTTCAACCGCTCGGCCACACTT
Ser2	CCG <del>G</del> TAGAGTTGCCCTACTCCGGTTTTTCGAGACCGGTCCGTTACGCCGCTCCGGCATCT
Ser3	GAAC <del>CCCC</del> GGATGCAGCTTTTGACCGCATACTCCCTTAGCAGGGGAGCGCCTTACGCCTCT
Ser5	CCG <del>A</del> TACGTTGCCGTATACACACTTTCCAGGCGTGCCTTACGCCACTCGGACACCTCA
Thr1	AGTCGAACTGGGGACCTCACCCCTTACCAAGGGTGCCTCTACCAACTGAGCCATATCAGC
Thr2	CGAA <del>CC</del> TACGACCTTCGCATTACGAATGCGTGTCTCTACCAACTGAGCTATATCGGCACG
Thr3	CGAA <del>CT</del> GCCGACCTCACCCCTTACCAAGGGTGCCTCTACCAACTGAGCTATATCAGCACC
Thr4	CGAA <del>CT</del> GGTGACCTACTGATTACAAGTCAGTTGCTCTACCTACTGAGCTAAGTCGGCACA
Trp	CGAACT <del>CCCA</del> ACACCCGGTTTTGGAGACCGGTGCTCTACCAATTGAACTACGCCCTTGG
Tyr1	AC <del>CT</del> TTCGAAGTCGATGACGGCAGATTTACAGTCTGCTCCCTTTGGCCGCTCGGGAACCC
Tyr2	AC <del>CT</del> TTCGAAGTCTGTGACGGCAGATTTACAGTCTGCTCCCTTTGGCCGCTCGGGAACCC
Val1	CGAAC <del>CC</del> GCGACCCCCTCCTTGTAAAGGGAGGTGCTCTCCAGCTGAGCTAATCACCCGTA
Val2A	GAAC <del>CA</del> CCGACCCCCACCATGTCAAGGTGGTGTCTCTAACCAACTGAGCTACGGACGCGTC
Val2B	GAAC <del>CA</del> ACGACCCCCACCATGTCAAGGTGGTGTCTCTAACCAACTGAGCTATGAACGCGTC
Ecoli Selcys	TGCCCGGACCGCTGGCGGCCCAACTGGATTTGAAGTCCAGCCGCTCACCGGAGACGA
Yeast Ala <sub>TGC</sub>	AATCGAACCGATGACCTCTTCTTGAAGGGAAGCGCGCTACCAACTGCGCCATGTGCC
Yeast Gln <sub>TTG</sub>	ATTCGAACCGGGGTTGTCCGGATCAAAACCGAAAGTGATAACCACTACACTATAGGACC
Yeast Ile <sub>AAT</sub>	ATCGAACCGCTGATCCCCGCGTTATTAGCACGGTGCCTTAAACCAACTGGGCCAAGAGACC
Yeast Lys <sub>SCTT</sub>	GCTCGAACCCCTAACCTTATGATTAAGAGTCATACGCGCTACCGATTGCGCCAACAAGGC
Yeast Val <sub>ACC</sub>	ATCGAACTGGGGACGTTCTGCGTGTAAAGCAGATGCCATAACCGACTAGACCACGAAACC
Yeast Pro <sub>TGG</sub>	GAATTGAACCCAGGGCCTCTCGCGCCAAAGCGAGAATCATACCACTAGACCACACGCC
Yeast Gly <sub>GCC</sub>	GGAATCGAACCGGGGGCCCAACGATGGCAACGTTGATTTTACCACTAAACCACTTGCGC

## Material and Methods

### 4.14.2 Preparation of peptidyl-tRNA samples for micro-array

Translation of HemK 1-112 aa was performed by mixing of 70S IC•Bodipy576/589-C<sub>2</sub>-Met-tRNA<sup>Met</sup> (20 nM) with factor mix containing ternary complexes (40 μM), EF-Tu in total (100 μM), and EF-G (3 μM). Translation was stopped by the addition of viomycin (200 μM) and samples were incubated for additional 2 min at 37°C. Ribosomes were pelleted through sucrose cushion (1.1 M) in TAKM<sub>7</sub> buffer supplemented with viomycin (200 μM) in a Beckmann TLA-55 rotor at 55'0000 rpm, 4°C for 2 h. Ribosome pellet was resuspended in NaOAc-AcOH pH 4.8 (100 mM), EDTA pH 7.0 (10 mM) and SDS (0.05% (w/vol)). 2x Hi-RPM hybridization buffer (one volume) was added to yield a final pH of 5.5. Peptidyl-tRNA was released during incubation for 1h at 25°C and 10 min at 65°C followed by centrifugation for 5 min at 5'000 rpm. Before spotting on the micro-array slide, 10x blocking agent (5 μL) was added to samples (45 μL volume) followed by incubation for 10 min at 65°C. Hybridization was performed for 2.5 h at 65°C. After hybridization, the array slide was washed with Gene Expression wash buffer 1 for 1 min at 25°C, with Gene Expression wash buffer 2 for 1 min at 37°C, and finally washed with acetonitrile. The micro-array was scanned with a micro-array Scanner G2565, Agilent.

### 4.14.3 Stripping of micro-array slide

After scanning of the micro-array slide, the cDNA probes were stripped from the hybridized sample as described (Poulsen et al., 2008). The micro-array slide was incubate in NaOH (50 N) for 10 min at RT, followed by incubation in SDS (1% (w/vol)) for 10 min at 70°C, and neutralized by incubation in 0.1x PBS for 10 min at 25°C.

### 4.14.4 3'-labeling of tRNA

Total aa-tRNA was deacylated by incubation in Tris-HCl pH 9.0 (1 M) for 2 h at 37°C. Deacylated tRNA was precipitated with EtOH and resuspended in ddH<sub>2</sub>O. Aminoacylated or deacylated total tRNA (10'000 pmol each) was concentrated to a volume of ~30 μL in a 0.5 mL Amicon centrifugal concentrator (Millipore) with a MWCO of 10 kDa for 20 min at 14'000 relative centrifugal force (rcf), 4°C in a Benchtop centrifuge 5415R. NaOAc pH 5.3 (100 mM, 500 μL) supplemented with KIO<sub>4</sub> (5 mM) was added and the sample was concentrated in the 10 kDa centrifugal concentrator for 20 min at 14'000 rcf, 4°C followed by incubation for an additional 10 min at 4°C. Ethylen glycol (10 mM) was added and samples were incubated for 10 min at 4°C. Samples was concentrated in the 10 kDa centrifugal concentrator for 20 min at 14'000 rcf, 4°C and NaOAc pH 5.3 (100 mM, 500 μL) was added. The step



## Material and Methods

of buffer exchange was repeated three times. Samples were adjusted to an equal volume (64  $\mu\text{L}$ ). Each sample was split in two and labeled separately with Cy3-hydrazide (3.4 mM, 4.4  $\mu\text{L}$  in DMF) or Cy5-hydrazide (2.5 mM, 6  $\mu\text{L}$  in DMF) for 16 h at 25°C. Unreacted fluorophore was removed by repeated extraction with 50:50 phenol:chloroform, followed by extraction with diethyl ether to remove traces of phenol and the tRNA was precipitated with isopropanol. The fraction of labeled tRNA was calculated with an extinction coefficient  $\epsilon_{260\text{ nm}}$  of 1'730'000  $\text{M}^{-1}\text{ cm}^{-1}$  for the tRNA, with  $\epsilon_{550\text{ nm}}$  136'000  $\text{M}^{-1}\text{ cm}^{-1}$  for Cy3, and  $\epsilon_{650\text{ nm}}$  250'000  $\text{M}^{-1}\text{ cm}^{-1}$  for Cy5. The fraction of labeling for deacylated tRNA samples was on average 0.22 for both Cy3 and Cy5. For labeling of aminoacylated tRNA, a fraction of labeling ranging from 0.06 to 0.09 was observed for different preparations of total aa-tRNA.

### 4.14.5 3'-labeling of tRNA extracted from stalled nascent chains

Ribosome nascent chain complexes at the endpoint of translation of leader peptidase (LepB) 1-35 aa were purified through a sucrose cushion as described in (Bornemann et al., 2008). The ribosome nascent chain complexes of LepB 1-35 aa were prepared by Anna Pfeiffer (technical assistant, MPI for Biophysical Chemistry, Department of Physical Biochemistry, Göttingen). The purified ribosome nascent chain complexes were diluted to a concentration of 0.33  $\mu\text{M}$  in TAKM<sub>7</sub> buffer supplemented with puromycin (1 mM) and incubated for 60 min at 37°C. The ribosomal proteins were removed by phenol extraction and ddH<sub>2</sub>O (100  $\mu\text{L}$ ) was added to the aqueous phase (30  $\mu\text{L}$ ). The ribosomal RNA was removed by passing the sample through a 0.5 mL Amicon centrifugal concentrator with a 100 kDa MWCO (Millipore) by centrifugation for 20 min at 14'000 rcf, 4°C. ddH<sub>2</sub>O (100  $\mu\text{L}$ ) was applied to the centrifugal concentration in order to recover additional tRNA during a second centrifugation step. The combined flow-through was concentrated in a 0.5 mL Amicon centrifugal concentrator (Millipore) with a MWCO of 10 kDa during centrifugation for 20 min at 14'000 rcf, 4°C. NaOAc pH 5.3 (100 mM, 500  $\mu\text{L}$ ) supplemented with KIO<sub>4</sub> (5 mM) was added, the sample was concentrated again in the centrifugal concentration with 10 kDa cut-off for 20 min at 14'000 rcf, 4°C and incubated for another 40 min at 4°C. Ethylene glycol (10 mM) was added and samples were incubated for 10 min at 4°C. NaOAc pH 5.3 (100 mM, 500  $\mu\text{L}$ ) was added and the sample was concentrated with a centrifugal concentrator with 10 kDa cut-off. The step of buffer exchange was repeated four times in total. The recovered sample (40  $\mu\text{L}$ ) was labeled with Cy3-hydrazide (3.4 mM, 7  $\mu\text{L}$  in DMF) for 16 h at 25°C. Unreacted fluorophore was removed by repeated extraction with 50:50 phenol:chloroform and remaining traces

## Material and Methods

of phenol were removed by extraction of the aqueous phase with diethyl ether. The sample was passed through a P6 gel filtration spin-column (Biorad) that was equilibrated in ddH<sub>2</sub>O. This step was required for removal of traces of organic solvents.

### 4.14.6 3'-labeling of Spike-In probes

The gene expression micro-array described in Section 4.14.1 contained additional cDNA probes that could be used for the assessment of the amplification steps of RNA sequences. The standard workflow for gene expression profiling developed by Agilent involved reverse transcription of mRNA sequences followed by the transcription of the generated cDNA sequences with Cy3 and Cy5 labeled CTP. The One Color RNA Spike-In kit contains 10 polyadenylated transcripts from the Adenovirus E1A gene that can be spiked-in as internal reference into preparations of total RNA. The steps of reverse transcription and transcription in presence of labeled nucleotides are not required for hybridization of fluorescence labeled tRNA or peptidyl-tRNA. The One-color RNA Spike-In kit also contained sequences that hybridized with specific probes at the edges of the subarray on the micro-array slide. These points at the edges are required as reference coordinates for the extraction of the intensities of the individual cDNA probes on the micro-array slide. The RNA sequences of the One Color RNA Spike-In kit were reverse-transcribed and transcribed in presence of labeled Cy3-CTP by Susanne Luthin (technical assistant, TAL, University Göttingen). The One Color RNA Spike-In labeled internally with Cy3 were labeled additionally at the 3'-end with Cy5-hydrazide as described in section 4.14.4. A ratio of labeling of Cy3/Cy5 of 3.2 was obtained.

### 4.14.7 Hybridization of 3'-labeled total tRNA

Samples (20 µL in ddH<sub>2</sub>O) contained equal amounts of total tRNA (0.68 pmol) labeled to a different degree with either Cy3 or Cy5 and Spike-In probes (0.5 pmol) labeled with both Cy3 and Cy5. 10x blocking agent (5 µL) was added and the sample was incubated for 30 min at 65°C. 2x Hi-RPM hybridization buffer (25 µL) was added and the sample was hybridized for 16 h at 65°C. The microarray was washed and scanned as described in Section 4.14.2.

## *Material and Methods*

### 4.14.8 Hybridization of 3'-labeled tRNA extracted from LepB 1-35

aa

Sample (20  $\mu$ L in ddH<sub>2</sub>O) contained tRNA extracted from purified LepB 1-35 aa ribosome nascent chain complex (17  $\mu$ L) that was labeled with Cy3-hydrazide, Cy5-labeled, deacylated total tRNA (0.68 pmol), and Cy3- and Cy5-labeled Spike-In probes (0.5 pmol). The sample was processed and hybridized as described in section 4.14.7.

## References

- Abe, R., Jeong, H.J., Arakawa, D., Dong, J., Ohashi, H., Kaigome, R., Saiki, F., Yamane, K., Takagi, H., and Ueda, H. (2014). Ultra Q-bodies: quench-based antibody probes that utilize dye-dye interactions with enhanced antigen-dependent fluorescence. *Sci Rep* 4, 4640.
- Abe, R., Ohashi, H., Iijima, I., Ihara, M., Takagi, H., Hohsaka, T., and Ueda, H. (2011). "Quenchbodies": quench-based antibody probes that show antigen-dependent fluorescence. *J Am Chem Soc* 133, 17386-17394.
- Abe, R., Shiraga, K., Ebisu, S., Takagi, H., and Hohsaka, T. (2010). Incorporation of fluorescent non-natural amino acids into N-terminal tag of proteins in cell-free translation and its dependence on position and neighboring codons. *J Biosci Bioeng* 110, 32-38.
- Agirrezabala, X., and Valle, M. (2015). Structural insights into tRNA dynamics on the ribosome. *Int J Mol Sci* 16, 9866-9895.
- Akashi, H. (1994). Synonymous codon usage in *Drosophila melanogaster*: natural selection and translational accuracy. *Genetics* 136, 927-935.
- Anderson, J.C., and Schultz, P.G. (2003). Adaptation of an orthogonal archaeal leucyl-tRNA and synthetase pair for four-base, amber, and opal suppression. *Biochemistry* 42, 9598-9608.
- Andersson, S.G., and Kurland, C.G. (1990). Codon preferences in free-living microorganisms. *Microbiol Rev* 54, 198-210.
- Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. *Science* 181, 223-230.
- Angov, E., Hillier, C.J., Kincaid, R.L., and Lyon, J.A. (2008). Heterologous protein expression is enhanced by harmonizing the codon usage frequencies of the target gene with those of the expression host. *PLoS One* 3, e2189.
- Artieri, C.G., and Fraser, H.B. (2014). Accounting for biases in riboprofiling data indicates a major role for proline in stalling translation. *Genome Res* 24, 2011-2021.
- Asahara, H., and Uhlenbeck, O.C. (2005). Predicting the binding affinities of misacylated tRNAs for *Thermus thermophilus* EF-Tu.GTP. *Biochemistry* 44, 11254-11261.
- Balakrishnan, R., Oman, K., Shoji, S., Bundschuh, R., and Fredrick, K. (2014). The conserved GTPase LepA contributes mainly to translation initiation in *Escherichia coli*. *Nucleic Acids Res* 42, 13370-13383.
- Baldini, G., Martoglio, B., Schachenmann, A., Zugliani, C., and Brunner, J. (1988). Mischarging *Escherichia coli* tRNAPhe with L-4'-[3-(trifluoromethyl)-3H-diazirin-3-yl]phenylalanine, a photoactivatable analogue of phenylalanine. *Biochemistry* 27, 7951-7959.
- Ban, N., Nissen, P., Hansen, J., Moore, P.B., and Steitz, T.A. (2000). The complete atomic structure of the large ribosomal subunit at 2.4 angstrom resolution. *Science* 289, 905-920.
- Begley, U., Dyavaiah, M., Patil, A., Rooney, J.P., DiRenzo, D., Young, C.M., Conklin, D.S., Zitomer, R.S., and Begley, T.J. (2007). Trm9-catalyzed tRNA modifications link translation to the DNA damage response. *Mol Cell* 28, 860-870.

## References

- Bhushan, S., Hoffmann, T., Seidelt, B., Frauenfeld, J., Mielke, T., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2011). SecM-stalled ribosomes adopt an altered geometry at the peptidyl transferase center. *Plos Biol* 9.
- Bhuta, A., Quiggle, K., Ott, T., Ringer, D., and Chladek, S. (1981). Stereochemical control of ribosomal peptidyltransferase reaction. Role of amino acid side-chain orientation of acceptor substrate. *Biochemistry* 20, 8-15.
- Bilgin, N., Claesens, F., Pahverk, H., and Ehrenberg, M. (1992). Kinetic properties of *Escherichia coli* ribosomes with altered forms of S12. *J Mol Biol* 224, 1011-1027.
- Bock, L.V., Blau, C., Schroder, G.F., Davydov, II, Fischer, N., Stark, H., Rodnina, M.V., Vaiana, A.C., and Grubmuller, H. (2013). Energy barriers and driving forces in tRNA translocation through the ribosome. *Nat Struct Mol Biol* 20, 1390-1396.
- Boens, N., Leen, V., and Dehaen, W. (2012). Fluorescent indicators based on BODIPY. *Chem Soc Rev* 41, 1130-1172.
- Bohman, K., Ruusala, T., Jelenc, P.C., and Kurland, C.G. (1984). Kinetic impairment of restrictive streptomycin-resistant ribosomes. *Mol Gen Genet* 198, 90-99.
- Bornemann, T., Jockel, J., Rodnina, M.V., and Wintermeyer, W. (2008). Signal sequence-independent membrane targeting of ribosomes containing short nascent peptides within the exit tunnel. *Nat Struct Mol Biol* 15, 494-499.
- Bremer, H., and Dennis, P.P. (2008). Modulation of Chemical Composition and Other parameters of the Cell at Different Exponential Growth Rates. In *EcoSal-Escherichia coli and Salmonella: Cellular and Molecular Biology*, Böck A., Curtis R., Kaper J.B., Karp P.D., and Neidhardt F.C., eds. (Washington D.C.: ASM Press).
- Bresler, S., Grajevskaja, R., Kirilov, S., and Saminski, E. (1968). Stability of peptidyl-tRNA - the intermediate of protein synthesis. *Biochim Biophys Acta* 155, 465-475.
- Brunelle, J.L., Youngman, E.M., Sharma, D., and Green, R. (2006). The interaction between C75 of tRNA and the A loop of the ribosome stimulates peptidyl transferase activity. *RNA* 12, 33-39.
- Bukau, B., Weissman, J., and Horwich, A. (2006). Molecular chaperones and protein quality control. *Cell* 125, 443-451.
- Bulaj, G., Kortemme, T., and Goldenberg, D.P. (1998). Ionization-reactivity relationships for cysteine thiols in polypeptides. *Biochemistry* 37, 8965-8972.
- Bulmer, M. (1991). The selection-mutation-drift theory of synonymous codon usage. *Genetics* 129, 897-907.
- Caliskan, N., Katunin, V.I., Belardinelli, R., Peske, F., and Rodnina, M.V. (2014). Programmed -1 frameshifting by kinetic partitioning during impeded translocation. *Cell* 157, 1619-1631.
- Cannarozzi, G., Schraudolph, N.N., Faty, M., von Rohr, P., Friberg, M.T., Roth, A.C., Gonnet, P., Gonnet, G., and Barral, Y. (2010). A role for codon order in translation dynamics. *Cell* 141, 355-367.
- Chadani, Y., Ito, K., Kutsukake, K., and Abo, T. (2012). ArfA recruits release factor 2 to rescue stalled ribosomes by peptidyl-tRNA hydrolysis in *Escherichia coli*. *Mol Microbiol* 86, 37-50.
- Chan, C.T., Dyavaiah, M., DeMott, M.S., Taghizadeh, K., Dedon, P.C., and Begley, T.J. (2010). A quantitative systems approach reveals dynamic control of tRNA modifications during cellular stress. *PLoS Genet* 6, e1001247.

## References

- Chaney, J.L., and Clark, P.L. (2015). Roles for synonymous codon usage in protein biogenesis. *Annu Rev Biophys* 44, 143-166.
- Chang, H.C., Kaiser, C.M., Hartl, F.U., and Barral, J.M. (2005). De novo folding of GFP fusion proteins: high efficiency in eukaryotes but not in bacteria. *J Mol Biol* 353, 397-409.
- Chapeville, F., Lipmann, F., Von Ehrenstein, G., Weisblum, B., Ray, W.J., Jr., and Benzer, S. (1962). On the role of soluble ribonucleic acid in coding for amino acids. *Proc Natl Acad Sci U S A* 48, 1086-1092.
- Chapman, S.J., Schrader, J.M., and Uhlenbeck, O.C. (2012). Histidine 66 in Escherichia coli elongation factor tu selectively stabilizes aminoacyl-tRNAs. *J Biol Chem* 287, 1229-1234.
- Chen, C., Zhang, H., Broitman, S.L., Reiche, M., Farrell, I., Cooperman, B.S., and Goldman, Y.E. (2013). Dynamics of translation by single ribosomes through mRNA secondary structures. *Nat Struct Mol Biol* 20, 582-588.
- Chen, P.R., Groff, D., Guo, J., Ou, W., Cellitti, S., Geierstanger, B.H., and Schultz, P.G. (2009). A facile system for encoding unnatural amino acids in mammalian cells. *Angew Chem Int Ed Engl* 48, 4052-4055.
- Chen, S.L., Lee, W., Hottes, A.K., Shapiro, L., and McAdams, H.H. (2004). Codon usage between genomes is constrained by genome-wide mutational processes. *Proc Natl Acad Sci U S A* 101, 3480-3485.
- Chin, J.W., Cropp, T.A., Anderson, J.C., Mukherji, M., Zhang, Z., and Schultz, P.G. (2003). An expanded eukaryotic genetic code. *Science* 301, 964-967.
- Ciryam, P., Morimoto, R.I., Vendruscolo, M., Dobson, C.M., and O'Brien, E.P. (2013). In vivo translation rates can substantially delay the cotranslational folding of the Escherichia coli cytosolic proteome. *Proc Natl Acad Sci U S A* 110, E132-140.
- Coleman, J.R., Papamichail, D., Skiena, S., Futcher, B., Wimmer, E., and Mueller, S. (2008). Virus attenuation by genome-scale changes in codon pair bias. *Science* 320, 1784-1787.
- Comeron, J.M. (2004). Selective and mutational patterns associated with gene expression in humans: influences on synonymous composition and intron presence. *Genetics* 167, 1293-1304.
- Cornish, V.W., Benson, D.R., Altenbach, C.A., Hideg, K., Hubbell, W.L., and Schultz, P.G. (1994). Site-specific incorporation of biophysical probes into proteins. *Proc Natl Acad Sci U S A* 91, 2910-2914.
- Crombie, T., Boyle, J.P., Coggins, J.R., and Brown, A.J.P. (1994). The folding of the bifunctional trp3 Protein in yeast is influenced by a translational pause which lies in a region of structural divergence with Escherichia-coli Indoleglycerol-Phosphate Synthase. *Eur J Biochem* 226, 657-664.
- Crombie, T., Swaffield, J.C., and Brown, A.J.P. (1992). Protein folding within the cell is influenced by controlled rates of polypeptide elongation. *J Mol Biol* 228, 7-12.
- Curran, J.F., and Yarus, M. (1989). Rates of aminoacyl-tRNA selection at 29 sense codons in vivo. *J Mol Biol* 209, 65-77.
- Dagan, S., Hagai, T., Gavrillov, Y., Kapon, R., Levy, Y., and Reich, Z. (2013). Stabilization of a protein conferred by an increase in folded state entropy. *Proc Natl Acad Sci U S A* 110, 10628-10633.

## References

- Daggett, V., and Fersht, A.R. (2003). Is there a unifying mechanism for protein folding? *Trends Biochem Sci* 28, 18-25.
- Dale, T., Fahlman, R.P., Olejniczak, M., and Uhlenbeck, O.C. (2009). Specificity of the ribosomal A site for aminoacyl-tRNAs. *Nucleic Acids Res* 37, 1202-1210.
- Dana, A., and Tuller, T. (2014). Properties and determinants of codon decoding time distributions. *BMC Genomics* 15 *Suppl* 6, S13.
- Deneke, C., Lipowsky, R., and Valleriani, A. (2013). Effect of ribosome shielding on mRNA stability. *Phys Biol* 10, 046008.
- Derman, A.I., Prinz, W.A., Belin, D., and Beckwith, J. (1993). Mutations that allow disulfide bond formation in the cytoplasm of *Escherichia coli*. *Science* 262, 1744-1747.
- Diaconu, M., Kothe, U., Schlunzen, F., Fischer, N., Harms, J.M., Tonevitsky, A.G., Stark, H., Rodnina, M.V., and Wahl, M.C. (2005). Structural basis for the function of the ribosomal L7/12 stalk in factor binding and GTPase activation. *Cell* 121, 991-1004.
- Dincbas-Renqvist, V., Engstrom, A., Mora, L., Heurgue-Hamard, V., Buckingham, R., and Ehrenberg, M. (2000). A post-translational modification in the GGQ motif of RF2 from *Escherichia coli* stimulates termination of translation. *EMBO J* 19, 6900-6907.
- Dirks, R.M., and Pierce, N.A. (2003). A partition function algorithm for nucleic acid secondary structure including pseudoknots. *J Comput Chem* 24, 1664-1677.
- Dittmar, K.A., Mobley, E.M., Radek, A.J., and Pan, T. (2004). Exploring the regulation of tRNA distribution on the genomic scale. *J Mol Biol* 337, 31-47.
- Dittmar, K.A., Sorensen, M.A., Elf, J., Ehrenberg, M., and Pan, T. (2005). Selective charging of tRNA isoacceptors induced by amino-acid starvation. *EMBO Rep* 6, 151-157.
- Doerfel, L.K., and Rodnina, M.V. (2013). Elongation factor P: Function and effects on bacterial fitness. *Biopolymers* 99, 837-845.
- Doerfel, L.K., Wohlgemuth, I., Kothe, C., Peske, F., Urlaub, H., and Rodnina, M.V. (2013). EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* 339, 85-88.
- Doi, Y., Ohtsuki, T., Shimizu, Y., Ueda, T., and Sisido, M. (2007). Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system. *J Am Chem Soc* 129, 14458-14462.
- Donachie, W.D. (1968). Relationship between cell size and time of initiation of DNA replication. *Nature* 219, 1077-1079.
- Dong, H., Nilsson, L., and Kurland, C.G. (1996). Co-variation of tRNA abundance and codon usage in *Escherichia coli* at different growth rates. *J Mol Biol* 260, 649-663.
- Dougherty, D.A., and Van Arnem, E.B. (2014). In vivo incorporation of non-canonical amino acids by using the chemical aminoacylation strategy: a broadly applicable mechanistic tool. *ChemBiochem* 15, 1710-1720.
- Drummond, D.A., and Wilke, C.O. (2008). Mistranslation-induced protein misfolding as a dominant constraint on coding-sequence evolution. *Cell* 134, 341-352.
- Elgamal, S., Katz, A., Hersch, S.J., Newsom, D., White, P., Navarre, W.W., and Ibba, M. (2014). EF-P dependent pauses integrate proximal and distal signals during translation. *PLoS Genet* 10, e1004553.

## References

- Ellis, J.P., Bakke, C.K., Kirchdoerfer, R.N., Jungbauer, L.M., and Cavagnero, S. (2008). Chain dynamics of nascent polypeptides emerging from the ribosome. *ACS Chem Biol* 3, 555-566.
- Ellis, R.J. (1996). Revisiting the Anfinsen cage. *Fold Des* 1, R9-15.
- Ellman, J., Mendel, D., Anthony-Cahill, S., Noren, C.J., and Schultz, P.G. (1991). Biosynthetic method for introducing unnatural amino acids site-specifically into proteins. *Methods Enzymol* 202, 301-336.
- Ellman, J.A., Mendel, D., and Schultz, P.G. (1992). Site-specific incorporation of novel backbone structures into proteins. *Science* 255, 197-200.
- Endoh, T., Kawasaki, Y., and Sugimoto, N. (2012). Synchronized translation for detection of temporal stalling of ribosome during single-turnover translation. *Anal Chem* 84, 857-861.
- Endoh, T., Kawasaki, Y., and Sugimoto, N. (2013a). Stability of RNA quadruplex in open reading frame determines proteolysis of human estrogen receptor alpha. *Nucleic Acids Res* 41, 6222-6231.
- Endoh, T., Kawasaki, Y., and Sugimoto, N. (2013b). Translational halt during elongation caused by G-quadruplex formed by mRNA. *Methods* 64, 73-78.
- Englander, M.T., Avins, J.L., Fleisher, R.C., Liu, B., Efraim, P.R., Wang, J., Schulten, K., Leyh, T.S., Gonzalez, R.L., Jr., and Cornish, V.W. (2015). The ribosome can discriminate the chirality of amino acids within its peptidyl-transferase center. *Proc Natl Acad Sci U S A* 112, 6038-6043.
- Ezure, T., Nanatani, K., Sato, Y., Suzuki, S., Aizawa, K., Souma, S., Ito, M., Hohsaka, T., von Heijne, G., Utsumi, T., *et al.* (2014). A cell-free translocation system using extracts of cultured insect cells to yield functional membrane proteins. *PLoS One* 9, e112874.
- Fedorov, A.N., and Baldwin, T.O. (1997). Cotranslational protein folding. *J Biol Chem* 272, 32715-32718.
- Fedorov, A.N., and Baldwin, T.O. (1999). Process of biosynthetic protein folding determines the rapid formation of native structure. *J Mol Biol* 294, 579-586.
- Fedyunin, I., Lehnhardt, L., Bohmer, N., Kaufmann, P., Zhang, G., and Ignatova, Z. (2012). tRNA concentration fine tunes protein solubility. *FEBS Lett* 586, 3336-3340.
- Fersht, A.R. (1999). *Structure and Mechanism in Protein Science*, 2 edn (New York: W.H. Freeman and Company).
- Fersht, A.R., and Dingwall, C. (1979). Cysteinyl-tRNA synthetase from *Escherichia coli* does not need an editing mechanism to reject serine and alanine. High binding energy of small groups in specific molecular interactions. *Biochemistry* 18, 1245-1249.
- Fischer, N., Konevega, A.L., Wintermeyer, W., Rodnina, M.V., and Stark, H. (2010). Ribosome dynamics and tRNA movement by time-resolved electron cryomicroscopy. *Nature* 466, 329-333.
- Fischer, N., Neumann, P., Konevega, A.L., Bock, L.V., Ficner, R., Rodnina, M.V., and Stark, H. (2015). Structure of the *E. coli* ribosome-EF-Tu complex at <3 Å resolution by Cs-corrected cryo-EM. *Nature* 520, 567-570.
- Fluitt, A., Pienaar, E., and Viljoen, H. (2007). Ribosome kinetics and aa-tRNA competition determine rate and fidelity of peptide synthesis. *Comput Biol Chem* 31, 335-346.



## References

- Folley, L.S., and Yarus, M. (1989). Codon contexts from weakly expressed genes reduce expression in vivo. *J Mol Biol* 209, 359-378.
- Fraser, H.B., Hirsh, A.E., Wall, D.P., and Eisen, M.B. (2004). Coevolution of gene expression among interacting proteins. *Proc Natl Acad Sci U S A* 101, 9033-9038.
- Frenkel-Morgenstern, M., Danon, T., Christian, T., Igarashi, T., Cohen, L., Hou, Y.M., and Jensen, L.J. (2012). Genes adopt non-optimal codon usage to generate cell cycle-dependent oscillations in protein levels. *Mol Syst Biol* 8, 572.
- Frydman, J. (2001). Folding of newly translated proteins in vivo: the role of molecular chaperones. *Annu Rev Biochem* 70, 603-647.
- Frydman, J., Erdjument-Bromage, H., Tempst, P., and Hartl, F.U. (1999). Co-translational domain folding as the structural basis for the rapid de novo folding of firefly luciferase. *Nat Struct Biol* 6, 697-705.
- Frydman, J., Nimmesgern, E., Ohtsuka, K., and Hartl, F.U. (1994). Folding of nascent polypeptide chains in a high molecular mass assembly with molecular chaperones. *Nature* 370, 111-117.
- Gardin, J., Yeasmin, R., Yurovsky, A., Cai, Y., Skiena, S., and Futcher, B. (2014). Measurement of average decoding rates of the 61 sense codons in vivo. *Elife* 3.
- Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* 6, 343-345.
- Goodman, D.B., Church, G.M., and Kosuri, S. (2013). Causes and effects of N-terminal codon bias in bacterial genes. *Science* 342, 475-479.
- Gorini, L. (1971). Ribosomal discrimination of tRNAs. *Nat New Biol* 234, 261-264.
- Gorochoowski, T.E., Ignatova, Z., Bovenberg, R.A., and Roubos, J.A. (2015). Trade-offs between tRNA abundance and mRNA secondary structure support smoothing of translation elongation rate. *Nucleic Acids Res* 43, 3022-3032.
- Goto, Y., Katoh, T., and Suga, H. (2011). Flexizymes for genetic code reprogramming. *Nat Protoc* 6, 779-790.
- Graille, M., Heurgue-Hamard, V., Champ, S., Mora, L., Scrima, N., Ulryck, N., van Tilbeurgh, H., and Buckingham, R.H. (2005). Molecular basis for bacterial class I release factor methylation by PrmC. *Mol Cell* 20, 917-927.
- Grantham, R., Gautier, C., Gouy, M., Mercier, R., and Pavé, A. (1980). Codon catalog usage and the genome hypothesis. *Nucleic Acids Res* 8, r49-r62.
- Gromadski, K.B., Daviter, T., and Rodnina, M.V. (2006). A uniform response to mismatches in codon-anticodon complexes ensures ribosomal fidelity. *Mol Cell* 21, 369-377.
- Gromadski, K.B., and Rodnina, M.V. (2004). Kinetic determinants of high-fidelity tRNA discrimination on the ribosome. *Mol Cell* 13, 191-200.
- Gu, S.Q., Jockel, J., Beinker, P., Warnecke, J., Semenov, Y.P., Rodnina, M.V., and Wintermeyer, W. (2005). Conformation of 4.5S RNA in the signal recognition particle and on the 30S ribosomal subunit. *RNA* 11, 1374-1384.
- Gu, W., Zhou, T., and Wilke, C.O. (2010). A universal trend of reduced mRNA stability near the translation-initiation site in prokaryotes and eukaryotes. *PLoS Comput Biol* 6, e1000664.

## References

- Gubbens, J., Kim, S.J., Yang, Z., Johnson, A.E., and Skach, W.R. (2010). In vitro incorporation of nonnatural amino acids into protein using tRNA(Cys)-derived opal, ochre, and amber suppressor tRNAs. *RNA* 16, 1660-1672.
- Guo, J., Melancon, C.E., 3rd, Lee, H.S., Groff, D., and Schultz, P.G. (2009). Evolution of amber suppressor tRNAs for efficient bacterial production of proteins containing nonnatural amino acids. *Angew Chem Int Ed Engl* 48, 9148-9151.
- Gustafsson, C., Govindarajan, S., and Minshull, J. (2004). Codon bias and heterologous protein expression. *Trends Biotechnol* 22, 346-353.
- Hamlin, J., and Zabin, I. (1972). Galactosidase: immunological activity of ribosome-bound, growing polypeptide chains. *Proc Natl Acad Sci U S A* 69, 412-416.
- Hancock, S.M., Uprety, R., Deiters, A., and Chin, J.W. (2010). Expanding the genetic code of yeast for incorporation of diverse unnatural amino acids via a pyrrolysyl-tRNA synthetase/tRNA pair. *J Am Chem Soc* 132, 14819-14824.
- Harding, J.J., and Dilley, K.J. (1976). Structural proteins of the mammalian lens: a review with emphasis on changes in development, aging and cataract. *Exp Eye Res* 22, 1-73.
- Heckler, T.G., Chang, L.H., Zama, Y., Naka, T., Chorghade, M.S., and Hecht, S.M. (1984). T4 RNA ligase mediated preparation of novel "chemically misacylated" tRNAPheS. *Biochemistry* 23, 1468-1473.
- Helmstetter, C.E. (1968). Origin and sequence of chromosome replication in *Escherichia coli* B-r. *J Bacteriol* 95, 1634-1641.
- Hershberg, R., and Petrov, D.A. (2008). Selection on codon bias. *Annu Rev Genet* 42, 287-299.
- Higgs, P.G., and Ran, W. (2008). Coevolution of codon usage and tRNA genes leads to alternative stable states of biased codon usage. *Mol Biol Evol* 25, 2279-2291.
- Hohsaka, T., Abe, R., Shiraga, K., and Sisido, M. (2003). Incorporation of fluorescently labeled nonnatural amino acids into proteins in an *E. coli* in vitro translation system. *Nucleic Acids Res*, 271-272.
- Hohsaka, T., Kajihara, D., Ashizuka, Y., Murakami, H., and Sisido, M. (1999). Efficient incorporation of nonnatural amino acids with large aromatic groups into streptavidin in vitro protein synthesizing systems. *J Am Chem Soc* 121, 34-40.
- Hohsaka, T., Muranaka, N., Komiyama, C., Matsui, K., Takaura, S., Abe, R., Murakami, H., and Sisido, M. (2004). Position-specific incorporation of dansylated non-natural amino acids into streptavidin by using a four-base codon. *FEBS Lett* 560, 173-177.
- Hohsaka, T., Sato, K., Sisido, M., Takai, K., and Yokoyama, S. (1993). Adaptability of nonnatural aromatic amino acids to the active center of the *E. coli* ribosomal A site. *FEBS Lett* 335, 47-50.
- Holmes, W.M., Hurd, R.E., Reid, B.R., Rimerman, R.A., and Hatfield, G.W. (1975). Separation of transfer ribonucleic acid by sepharose chromatography using reverse salt gradients. *Proc Natl Acad Sci U S A* 72, 1068-1071.
- Holtkamp, W., Cunha, C.E., Peske, F., Konevega, A.L., Wintermeyer, W., and Rodnina, M.V. (2014a). GTP hydrolysis by EF-G synchronizes tRNA movement on small and large ribosomal subunits. *EMBO J* 33, 1073-1085.

## References

- Holtkamp, W., Wintermeyer, W., and Rodnina, M.V. (2014b). Synchronous tRNA movements during translocation on the ribosome are orchestrated by elongation factor G and GTP hydrolysis. *Bioessays* *36*, 908-918.
- Hopfield, J.J. (1974). Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc Natl Acad Sci U S A* *71*, 4135-4139.
- Hou, Y.M., Westhof, E., and Giege, R. (1993). An unusual RNA tertiary interaction has a role for the specific aminoacylation of a transfer RNA. *Proc Natl Acad Sci U S A* *90*, 6776-6780.
- Huang, Y., and Sprinzl, M. (2011). Peptide bond formation on the ribosome: the role of the 2'-OH group on the terminal adenosine of peptidyl-tRNA and of the length of nascent peptide chain. *Angew Chem Int Ed Engl* *50*, 7287-7289.
- Hughes, T.R., Mao, M., Jones, A.R., Burchard, J., Marton, M.J., Shannon, K.W., Lefkowitz, S.M., Ziman, M., Schelter, J.M., Meyer, M.R., *et al.* (2001). Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol* *19*, 342-347.
- Ieong, K.W., Pavlov, M.Y., Kwiatkowski, M., Ehrenberg, M., and Forster, A.C. (2014). A tRNA body with high affinity for EF-Tu hastens ribosomal incorporation of unnatural amino acids. *RNA* *20*, 632-643.
- Ieong, K.W., Pavlov, M.Y., Kwiatkowski, M., Forster, A.C., and Ehrenberg, M. (2012). Inefficient delivery but fast peptide bond formation of unnatural L-aminoacyl-tRNAs in translation. *J Am Chem Soc* *134*, 17955-17962.
- Iijima, I., and Hohsaka, T. (2009). Position-specific incorporation of fluorescent non-natural amino acids into maltose-binding protein for detection of ligand binding by FRET and fluorescence quenching. *Chembiochem* *10*, 999-1006.
- Ikemura, T. (1981). 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 Escherichia-coli translational system. *J Mol Biol* *151*, 389-409.
- Ikemura, T. (1985). Codon usage and tRNA content in unicellular and multicellular organisms. *Mol Biol Evol* *2*, 13-34.
- Indrisiunaite, G., Pavlov, M.Y., Heurgue-Hamard, V., and Ehrenberg, M. (2015). On the pH dependence of class-1 RF-dependent termination of mRNA translation. *J Mol Biol* *427*, 1848-1860.
- Ingolia, N.T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. *Nat Rev Genet* *15*, 205-213.
- Ingolia, N.T., Ghaemmighami, S., Newman, J.R., and Weissman, J.S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* *324*, 218-223.
- Ito, K., and Chiba, S. (2013). Arrest peptides: cis-acting modulators of translation. *Annu Rev Biochem* *82*, 171-202.
- Ito, K., Uno, M., and Nakamura, Y. (2000). A tripeptide 'anticodon' deciphers stop codons in messenger RNA. *Nature* *403*, 680-684.
- Jaenicke, R. (1991). Protein folding: local structures, domains, subunits, and assemblies. *Biochemistry* *30*, 3147-3161.

## References

- Jakubowski, H. (1994). Editing function of Escherichia coli cysteinyl-tRNA synthetase: cyclization of cysteine to cysteine thiolactone. *Nucleic Acids Res* 22, 1155-1160.
- Janssen, B.D., Diner, E.J., and Hayes, C.S. (2012). Analysis of aminoacyl- and peptidyl-tRNAs by gel electrophoresis. *Methods Mol Biol* 905, 291-309.
- Jenner, L., Demeshkina, N., Yusupova, G., and Yusupov, M. (2010a). Structural rearrangements of the ribosome at the tRNA proofreading step. *Nat Struct Mol Biol* 17, 1072-1078.
- Jenner, L.B., Demeshkina, N., Yusupova, G., and Yusupov, M. (2010b). Structural aspects of messenger RNA reading frame maintenance by the ribosome. *Nat Struct Mol Biol* 17, 555-560.
- Jeong, H.J., Itayama, S., and Ueda, H. (2015). A signal-on fluorosensor based on quench-release principle for sensitive detection of antibiotic rapamycin. *Biosensors (Basel)* 5, 131-140.
- Jeong, H.J., Ohmuro-Matsuyama, Y., Ohashi, H., Ohsawa, F., Tatsu, Y., Inagaki, M., and Ueda, H. (2013). Detection of vimentin serine phosphorylation by multicolor Quenchbodies. *Biosens Bioelectron* 40, 17-23.
- Jeong, H.J., and Ueda, H. (2014). Strategy for making a superior Quenchbody to proteins: effect of the fluorophore position. *Sensors (Basel)* 14, 13285-13297.
- Johansson, M., Bouakaz, E., Lovmar, M., and Ehrenberg, M. (2008a). The kinetics of ribosomal peptidyl transfer revisited. *Mol Cell* 30, 589-598.
- Johansson, M., Jeong, K.W., Trobro, S., Strazewski, P., Aqvist, J., Pavlov, M.Y., and Ehrenberg, M. (2011). pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA. *Proc Natl Acad Sci U S A* 108, 79-84.
- Johansson, M., Lovmar, M., and Ehrenberg, M. (2008b). Rate and accuracy of bacterial protein synthesis revisited. *Curr Opin Microbiol* 11, 141-147.
- Johansson, M., Zhang, J., and Ehrenberg, M. (2012). Genetic code translation displays a linear trade-off between efficiency and accuracy of tRNA selection. *Proc Natl Acad Sci U S A* 109, 131-136.
- Johnson, A.E. (2005). The co-translational folding and interactions of nascent protein chains: a new approach using fluorescence resonance energy transfer. *FEBS Lett* 579, 916-920.
- Johnson, A.E., Woodward, W.R., Herbert, E., and Menninger, J.R. (1976). N-epsilon-acetyllysine transfer ribonucleic-acid - Biologically-active analog of aminoacyl transfer ribonucleic-Acids. *Biochemistry* 15, 569-575.
- Jones, T.E., Alexander, R.W., and Pan, T. (2011). Misacylation of specific nonmethionyl tRNAs by a bacterial methionyl-tRNA synthetase. *Proc Natl Acad Sci U S A* 108, 6933-6938.
- Kaiser, C.M., Goldman, D.H., Chodera, J.D., Tinoco, I., Jr., and Bustamante, C. (2011). The ribosome modulates nascent protein folding. *Science* 334, 1723-1727.
- Kajihara, D., Abe, R., Iijima, I., Komiyama, C., Sisido, M., and Hohsaka, T. (2006). FRET analysis of protein conformational change through position-specific incorporation of fluorescent amino acids. *Nat Methods* 3, 923-929.
- Kajihara, D., Hohsaka, T., and Sisido, M. (2005). Synthesis and sequence optimization of GFP mutants containing aromatic non-natural amino acids at the Tyr66 position. *Protein Eng Des Sel* 18, 273-278.

## References

- Kang, S.H., Jun, S.Y., and Kim, D.M. (2007). Fluorescent labeling of cell-free synthesized proteins by incorporation of fluorophore-conjugated nonnatural amino acids. *Anal Biochem* 360, 1-6.
- Katunin, V.I., Muth, G.W., Strobel, S.A., Wintermeyer, W., and Rodnina, M.V. (2002). Important contribution to catalysis of peptide bond formation by a single ionizing group within the ribosome. *Mol Cell* 10, 339-346.
- Kiho, Y., and Rich, A. (1964). Induced enzyme formed on bacterial polyribosomes. *Proc Natl Acad Sci U S A* 51, 111-118.
- Kim, D.F., and Green, R. (1999). Base-pairing between 23S rRNA and tRNA in the ribosomal A site. *Molecular Cell* 4, 859-864.
- Kim, S.J., Yoon, J.S., Shishido, H., Yang, Z., Rooney, L.A., Barral, J.M., and Skach, W.R. (2015). Protein folding. Translational tuning optimizes nascent protein folding in cells. *Science* 348, 444-448.
- Kimchi-Sarfaty, C., Oh, J.M., Kim, I.W., Sauna, Z.E., Calcagno, A.M., Ambudkar, S.V., and Gottesman, M.M. (2007). A "silent" polymorphism in the MDR1 gene changes substrate specificity. *Science* 315, 525-528.
- Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide-sequences. *J Mol Evol* 16, 111-120.
- Knight, A.M., Culviner, P.H., Kurt-Yilmaz, N., Zou, T., Ozkan, S.B., and Cavagnero, S. (2013). Electrostatic effect of the ribosomal surface on nascent polypeptide dynamics. *ACS Chem Biol* 8, 1195-1204.
- Koch, M., Huang, Y., and Sprinzl, M. (2008). Peptide-bond synthesis on the ribosome: no free vicinal hydroxy group required on the terminal ribose residue of peptidyl-tRNA. *Angew Chem Int Ed Engl* 47, 7242-7245.
- Kolb, V.A., Makeyev, E.V., and Spirin, A.S. (1994). Folding of firefly luciferase during translation in a cell-free system. *EMBO J* 13, 3631-3637.
- Komar, A.A. (2009). A pause for thought along the co-translational folding pathway. *Trends Biochem Sci* 34, 16-24.
- Komar, A.A., and Jaenicke, R. (1995). Kinetics of translation of gamma B crystallin and its circularly permuted variant in an in vitro cell-free system: possible relations to codon distribution and protein folding. *FEBS Lett* 376, 195-198.
- Komar, A.A., Kommer, A., Krasheninnikov, I.A., and Spirin, A.S. (1997). Cotranslational folding of globin. *J Biol Chem* 272, 10646-10651.
- Komar, A.A., Lesnik, T., and Reiss, C. (1999). Synonymous codon substitutions affect ribosome traffic and protein folding during in vitro translation. *FEBS Lett* 462, 387-391.
- Komatsoulis, G.A., and Abelson, J. (1993). Recognition of tRNA(Cys) by Escherichia coli cysteinyl-tRNA synthetase. *Biochemistry* 32, 7435-7444.
- Kosolapov, A., and Deutsch, C. (2009). Tertiary interactions within the ribosomal exit tunnel. *Nat Struct Mol Biol* 16, 405-411.
- Kothe, U., and Rodnina, M.V. (2006). Delayed release of inorganic phosphate from elongation factor Tu following GTP hydrolysis on the ribosome. *Biochemistry* 45, 12767-12774.

## References

- Kothe, U., Wieden, H.J., Mohr, D., and Rodnina, M.V. (2004). Interaction of helix D of elongation factor Tu with helices 4 and 5 of protein L7/12 on the ribosome. *J Mol Biol* 336, 1011-1021.
- Koubek, J., Chen, Y.R., Cheng, R.P., and Huang, J.J. (2015). Nonorthogonal tRNA<sup>cys</sup> amber for protein and nascent chain labeling. *RNA* 21, 1672-1682.
- Kowarik, M., Kung, S., Martoglio, B., and Helenius, A. (2002). Protein folding during cotranslational translocation in the endoplasmic reticulum. *Mol Cell* 10, 769-778.
- Kramer, E.B., and Farabaugh, P.J. (2007). The frequency of translational misreading errors in *E. coli* is largely determined by tRNA competition. *RNA* 13, 87-96.
- Krobath, H., Shakhnovich, E.I., and Faisca, P.F. (2013). Structural and energetic determinants of co-translational folding. *J Chem Phys* 138, 215101.
- Kudla, G., Murray, A.W., Tollervey, D., and Plotkin, J.B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* 324, 255-258.
- Kudlicki, W., Chirgwin, J., Kramer, G., and Hardesty, B. (1995). Folding of an enzyme into an active conformation while bound as peptidyl-transfer-RNA to the Ribosome. *Biochemistry* 34, 14284-14287.
- Kurzchalia, T.V., Wiedmann, M., Girshovich, A.S., Bochkareva, E.S., Bielka, H., and Rapoport, T.A. (1986). The signal sequence of nascent preprolactin interacts with the 54K polypeptide of the signal recognition particle. *Nature* 320, 634-636.
- Ladurner, A.G., and Fersht, A.R. (1997). Glutamine, alanine or glycine repeats inserted into the loop of a protein have minimal effects on stability and folding rates. *J Mol Biol* 273, 330-337.
- Larsen, B., Wills, N.M., Gesteland, R.F., and Atkins, J.F. (1994). rRNA-mRNA base pairing stimulates a programmed -1 ribosomal frameshift. *J Bacteriol* 176, 6842-6851.
- Ledoux, S., Olejniczak, M., and Uhlenbeck, O.C. (2009). A sequence element that tunes *Escherichia coli* tRNA(Ala)(GGC) to ensure accurate decoding. *Nat Struct Mol Biol* 16, 359-364.
- Lessard, J.L., and Pestka, S. (1972). Studies on the formation of transfer ribonucleic acid-ribosome complexes. 23. Chloramphenicol, aminoacyl-oligonucleotides, and *Escherichia coli* ribosomes. *J Biol Chem* 247, 6909-6912.
- Li, G.W., Burkhardt, D., Gross, C., and Weissman, J.S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cell* 157, 624-635.
- Li, G.W., Oh, E., and Weissman, J.S. (2012). The anti-Shine-Dalgarno sequence drives translational pausing and codon choice in bacteria. *Nature* 484, 538-541.
- Lien, L., Ananda, P., Seneviratne, K., Jaikaran, A.S., and Andrew Woolley, G. (2002). Site-specific biosynthetic incorporation of a fluorescent tag into proteins via cysteine-tRNA(Cys). *Anal Biochem* 307, 252-257.
- Lithwick, G., and Margalit, H. (2005). Relative predicted protein levels of functionally associated proteins are conserved across organisms. *Nucleic Acids Res* 33, 1051-1057.
- Liu, C., Gamper, H., Liu, H., Cooperman, B.S., and Hou, Y.M. (2011). Potential for interdependent development of tRNA determinants for aminoacylation and ribosome decoding. *Nat Commun* 2, 329.

## References

- Liu, C.C., and Schultz, P.G. (2010). Adding new chemistries to the genetic code. *Annu Rev Biochem* 79, 413-444.
- Liu, W., Brock, A., Chen, S., Chen, S., and Schultz, P.G. (2007). Genetic incorporation of unnatural amino acids into proteins in mammalian cells. *Nat Methods* 4, 239-244.
- Long, K.S., and Porse, B.T. (2003). A conserved chloramphenicol binding site at the entrance to the ribosomal peptide exit tunnel. *Nucleic Acids Res* 31, 7208-7215.
- Lopez, D., and Pazos, F. (2015). Protein functional features are reflected in the patterns of mRNA translation speed. *BMC Genomics* 16, 513.
- Lu, J., and Deutsch, C. (2005a). Folding zones inside the ribosomal exit tunnel. *Nat Struct Mol Biol* 12, 1123-1129.
- Lu, J., and Deutsch, C. (2005b). Secondary structure formation of a transmembrane segment in Kv channels. *Biochemistry* 44, 8230-8243.
- Lu, J., Kobertz, W.R., and Deutsch, C. (2007). Mapping the electrostatic potential within the ribosomal exit tunnel. *J Mol Biol* 371, 1378-1391.
- Maaloe, O. (1969). Analysis of bacterial Growth. *Dev Biol* 3 (Suppl), 33-58.
- Makeyev, E.V., Kolb, V.A., and Spirin, A.S. (1996). Enzymatic activity of the ribosome-bound nascent polypeptide. *FEBS Lett* 378, 166-170.
- Makino, Y., Amada, K., Taguchi, H., and Yoshida, M. (1997). Chaperonin-mediated folding of green fluorescent protein. *J Biol Chem* 272, 12468-12474.
- Mandal, K., Chakrabarti, B., Thomson, J., and Siezen, R.J. (1987). Structure and stability of gamma-crystallins. Denaturation and proteolysis behavior. *J Biol Chem* 262, 8096-8102.
- Mao, Y., Liu, H., Liu, Y., and Tao, S. (2014). Deciphering the rules by which dynamics of mRNA secondary structure affect translation efficiency in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 42, 4813-4822.
- Mares, R.E., Melendez-Lopez, S.G., and Ramos, M.A. (2011). Acid-denatured Green Fluorescent Protein (GFP) as model substrate to study the chaperone activity of protein disulfide isomerase. *Int J Mol Sci* 12, 4625-4636.
- Marme, N., Knemeyer, J.P., Sauer, M., and Wolfrum, J. (2003). Inter- and intramolecular fluorescence quenching of organic dyes by tryptophan. *Bioconjug Chem* 14, 1133-1139.
- Martens, A.T., Taylor, J., and Hilser, V.J. (2015). Ribosome A and P sites revealed by length analysis of ribosome profiling data. *Nucleic Acids Res* 43, 3680-3687.
- Matsubara, T., Iijima, K., Watanabe, T., Hohsaka, T., and Sato, T. (2013). Incorporation of glycosylated amino acid into protein by an in vitro translation system. *Bioorg Med Chem Lett* 23, 5634-5636.
- Mayr, E.M., Jaenicke, R., and Glockshuber, R. (1997). The domains in gammaB-crystallin: identical fold-different stabilities. *J Mol Biol* 269, 260-269.
- McCammon, J.A. (1996). A speed limit for protein folding. *Proc Natl Acad Sci U S A* 93, 11426-11427.
- Mendel, D., Ellman, J., and Schultz, P.G. (1993). Protein-Biosynthesis with Conformationally Restricted Amino-Acids. *J Am Chem Soc* 115, 4359-4360.

## References

- Michel, A.M., Fox, G., A, M.K., De Bo, C., O'Connor, P.B., Heaphy, S.M., Mullan, J.P., Donohue, C.A., Higgins, D.G., and Baranov, P.V. (2014). GWIPS-viz: development of a ribo-seq genome browser. *Nucleic Acids Res* *42*, D859-864.
- Milon, P., Konevega, A.L., Gualerzi, C.O., and Rodnina, M.V. (2008). Kinetic checkpoint at a late step in translation initiation. *Mol Cell* *30*, 712-720.
- Milon, P., Konevega, A.L., Peske, F., Fabbretti, A., Gualerzi, C.O., and Rodnina, M.V. (2007). Transient kinetics, fluorescence, and FRET in studies of initiation of translation in bacteria. *Methods Enzymol* *430*, 1-30.
- Milon, P., Maracci, C., Filonava, L., Gualerzi, C.O., and Rodnina, M.V. (2012). Real-time assembly landscape of bacterial 30S translation initiation complex. *Nat Struct Mol Biol* *19*, 609-615.
- Milon, P., and Rodnina, M.V. (2012). Kinetic control of translation initiation in bacteria. *Crit Rev Biochem Mol Biol* *47*, 334-348.
- Mitarai, N., Sneppen, K., and Pedersen, S. (2008). Ribosome collisions and translation efficiency: optimization by codon usage and mRNA destabilization. *J Mol Biol* *382*, 236-245.
- Mittelstaet, J., Konevega, A.L., and Rodnina, M.V. (2013). A kinetic safety gate controlling the delivery of unnatural amino acids to the ribosome. *J Am Chem Soc* *135*, 17031-17038.
- Miura, M., Muranaka, N., Abe, R., and Hohsaka, T. (2010). Incorporation of fluorescent-labeled non-alpha-amino carboxylic acids into the N-terminus of proteins in response to amber initiation codon. *B Chem Soc Jpn* *83*, 546-553.
- Modolell, J., and Vazquez (1977). The inhibition of ribosomal translocation by viomycin. *Eur J Biochem* *81*, 491-497.
- Morrissey, M.P., Ahmed, Z., and Shakhnovich, E.I. (2004). The role of cotranslation in protein folding: a lattice model study. *Polymer* *45*, 557-571.
- Mukai, T., Kobayashi, T., Hino, N., Yanagisawa, T., Sakamoto, K., and Yokoyama, S. (2008). Adding l-lysine derivatives to the genetic code of mammalian cells with engineered pyrrolysyl-tRNA synthetases. *Biochem Biophys Res Commun* *371*, 818-822.
- Najmudin, S., Nalini, V., Driessen, H.P., Slingsby, C., Blundell, T.L., Moss, D.S., and Lindley, P.F. (1993). Structure of the bovine eye lens protein gammaB(gammaII)-crystallin at 1.47 Å. *Acta Crystallogr D Biol Crystallogr* *49*, 223-233.
- Nakahigashi, K., Kubo, N., Narita, S., Shimaoka, T., Goto, S., Oshima, T., Mori, H., Maeda, M., Wada, C., and Inokuchi, H. (2002). HemK, a class of protein methyl transferase with similarity to DNA methyl transferases, methylates polypeptide chain release factors, and hemK knockout induces defects in translational termination. *Proc Natl Acad Sci U S A* *99*, 1473-1478.
- Nakahigashi, K., Takai, Y., Shiwa, Y., Wada, M., Honma, M., Yoshikawa, H., Tomita, M., Kanai, A., and Mori, H. (2014). Effect of codon adaptation on codon-level and gene-level translation efficiency in vivo. *BMC Genomics* *15*, 1115.
- Nakata, H., Ohtsuki, T., Abe, R., Hohsaka, T., and Sisido, M. (2006). Binding efficiency of elongation factor Tu to tRNAs charged with nonnatural fluorescent amino acids. *Anal Biochem* *348*, 321-323.
- Naylor, D.J., and Hartl, F.U. (2001). Contribution of molecular chaperones to protein folding in the cytoplasm of prokaryotic and eukaryotic cells. *Biochem Soc Symp*, 45-68.



## References

- Neidhardt, F.C. (1987). Chemical composition of *Escherichia coli* (Washington D.C.: American Society of Microbiology).
- Netzer, N., Goodenbour, J.M., David, A., Dittmar, K.A., Jones, R.B., Schneider, J.R., Boone, D., Eves, E.M., Rosner, M.R., Gibbs, J.S., *et al.* (2009). Innate immune and chemically triggered oxidative stress modifies translational fidelity. *Nature* 462, 522-526.
- Netzer, W.J., and Hartl, F.U. (1997). Recombination of protein domains facilitated by co-translational folding in eukaryotes. *Nature* 388, 343-349.
- Neumann, H. (2012). Rewiring translation - Genetic code expansion and its applications. *FEBS Lett* 586, 2057-2064.
- Neumann, H., Hazen, J.L., Weinstein, J., Mehl, R.A., and Chin, J.W. (2008a). Genetically encoding protein oxidative damage. *J Am Chem Soc* 130, 4028-4033.
- Neumann, H., Peak-Chew, S.Y., and Chin, J.W. (2008b). Genetically encoding N(epsilon)-acetyllysine in recombinant proteins. *Nat Chem Biol* 4, 232-234.
- Neumann, H., Slusarczyk, A.L., and Chin, J.W. (2010a). De novo generation of mutually orthogonal aminoacyl-tRNA synthetase/tRNA pairs. *J Am Chem Soc* 132, 2142-2144.
- Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J.W. (2010b). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441-444.
- Neuweiler, H., Schulz, A., Bohmer, M., Enderlein, J., and Sauer, M. (2003). Measurement of submicrosecond intramolecular contact formation in peptides at the single-molecule level. *J Am Chem Soc* 125, 5324-5330.
- Nimmegern, E., and Hartl, F.U. (1993). ATP-dependent protein refolding activity in reticulocyte lysate. Evidence for the participation of different chaperone components. *FEBS Lett* 331, 25-30.
- Nissen, P., Hansen, J., Ban, N., Moore, P.B., and Steitz, T.A. (2000). The structural basis of ribosome activity in peptide bond synthesis. *Science* 289, 920-930.
- Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C., and Schultz, P.G. (1989). A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* 244, 182-188.
- Novoa, E.M., Pavon-Eternod, M., Pan, T., and Ribas de Pouplana, L. (2012). A role for tRNA modifications in genome structure and codon usage. *Cell* 149, 202-213.
- O'Brien, E.P., Hsu, S.T., Christodoulou, J., Vendruscolo, M., and Dobson, C.M. (2010). Transient tertiary structure formation within the ribosome exit port. *J Am Chem Soc* 132, 16928-16937.
- O'Brien, E.P., Vendruscolo, M., and Dobson, C.M. (2012). Prediction of variable translation rate effects on cotranslational protein folding. *Nat Commun* 3, 868.
- Odom, O.W., Picking, W.D., and Hardesty, B. (1990). Movement of tRNA but not the nascent peptide during peptide bond formation on ribosomes. *Biochemistry* 29, 10734-10744.
- Ogle, J.M., Brodersen, D.E., Clemons, W.M., Jr., Tarry, M.J., Carter, A.P., and Ramakrishnan, V. (2001). Recognition of cognate transfer RNA by the 30S ribosomal subunit. *Science* 292, 897-902.
- Ogle, J.M., Murphy, F.V., Tarry, M.J., and Ramakrishnan, V. (2002). Selection of tRNA by the ribosome requires a transition from an open to a closed form. *Cell* 111, 721-732.

## References

- Ohtsuka, T., Neki, S., Kanai, T., Akiyoshi, K., Nomura, S.M., and Ohtsuki, T. (2011). Synthesis and in situ insertion of a site-specific fluorescently labeled membrane protein into cell-sized liposomes. *Anal Biochem* 418, 97-101.
- Ohtsuki, T., Yamamoto, H., Doi, Y., and Sisido, M. (2010). Use of EF-Tu mutants for determining and improving aminoacylation efficiency and for purifying aminoacyl tRNAs with non-natural amino acids. *J Biochem* 148, 239-246.
- Ohuchi, M., Murakami, H., and Suga, H. (2007). The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. *Curr Opin Chem Biol* 11, 537-542.
- Olejniczak, M., and Uhlenbeck, O.C. (2006). tRNA residues that have coevolved with their anticodon to ensure uniform and accurate codon recognition. *Biochimie* 88, 943-950.
- Olofsson, M., Kalinin, S., Zdunek, J., Oliveberg, M., and Johansson, L.B. (2006). Tryptophan-BODIPY: a versatile donor-acceptor pair for probing generic changes of intraprotein distances. *Phys Chem Chem Phys* 8, 3130-3140.
- Palme, S., Slingsby, C., and Jaenicke, R. (1997). Mutational analysis of hydrophobic domain interactions in gamma B-crystallin from bovine eye lens. *Protein Sci* 6, 1529-1536.
- Pape, T., Wintermeyer, W., and Rodnina, M. (1999). Induced fit in initial selection and proofreading of aminoacyl-tRNA on the ribosome. *EMBO J* 18, 3800-3807.
- Parker, J. (1989). Errors and alternatives in reading the universal genetic code. *Microbiol Rev* 53, 273-298.
- Pechmann, S., and Frydman, J. (2013). Evolutionary conservation of codon optimality reveals hidden signatures of cotranslational folding. *Nat Struct Mol Biol* 20, 237-243.
- Pedersen, M., Nissen, S., Mitarai, N., Lo Svenningsen, S., Sneppen, K., and Pedersen, S. (2011). The functional half-life of an mRNA depends on the ribosome spacing in an early coding region. *J Mol Biol* 407, 35-44.
- Peske, F., Kuhlenkoetter, S., Rodnina, M.V., and Wintermeyer, W. (2014). Timing of GTP binding and hydrolysis by translation termination factor RF3. *Nucleic Acids Res* 42, 1812-1820.
- Peske, F., Rodnina, M.V., and Wintermeyer, W. (2005). Sequence of steps in ribosome recycling as defined by kinetic analysis. *Mol Cell* 18, 403-412.
- Picking, W., Picking, W.D., and Hardesty, B. (1991a). The use of synthetic transfer-RNAs as probes for examining nascent peptides on Escherichia-coli ribosomes. *Biochimie* 73, 1101-1107.
- Picking, W.D., Odom, O.W., Tsalkova, T., Serdyuk, I., and Hardesty, B. (1991b). The conformation of nascent polylysine and polyphenylalanine peptides on ribosomes. *J Biol Chem* 266, 1534-1542.
- Piepenburg, O., Pape, T., Pleiss, J.A., Wintermeyer, W., Uhlenbeck, O.C., and Rodnina, M.V. (2000). Intact aminoacyl-tRNA is required to trigger GTP hydrolysis by elongation factor Tu on the ribosome. *Biochemistry* 39, 1734-1738.
- Pott, M., Schmidt, M.J., and Summerer, D. (2014). Evolved sequence contexts for highly efficient amber suppression with noncanonical amino acids. *ACS Chem Biol* 9, 2815-2822.
- Poulsen, L., Soe, M.J., Snakenborg, D., Moller, L.B., and Dufva, M. (2008). Multi-stringency wash of partially hybridized 60-mer probes reveals that the stringency along the probe decreases with distance from the microarray surface. *Nucleic Acids Res* 36, e132.

## References

- Presnyak, V., Alhusaini, N., Chen, Y.H., Martin, S., Morris, N., Kline, N., Olson, S., Weinberg, D., Baker, K.E., Graveley, B.R., *et al.* (2015). Codon optimality is a major determinant of mRNA stability. *Cell* 160, 1111-1124.
- Purvis, I.J., Bettany, A.J.E., Santiago, T.C., Coggins, J.R., Duncan, K., Eason, R., and Brown, A.J.P. (1987). The efficiency of folding of some proteins is increased by controlled rates of translation *in vivo* - a hypothesis. *J Mol Biol* 193, 413-417.
- Qu, X., Wen, J.D., Lancaster, L., Noller, H.F., Bustamante, C., and Tinoco, I., Jr. (2011). The ribosome uses two active mechanisms to unwind messenger RNA during translation. *Nature* 475, 118-121.
- Quax, T.E., Wolf, Y.I., Koehorst, J.J., Wurtzel, O., van der Oost, R., Ran, W., Blombach, F., Makarova, K.S., Brouns, S.J., Forster, A.C., *et al.* (2013). Differential translation tunes uneven production of operon-encoded proteins. *Cell Rep* 4, 938-944.
- Rackham, O., and Chin, J.W. (2005). A network of orthogonal ribosome x mRNA pairs. *Nat Chem Biol* 1, 159-166.
- Raftery, L.A., and Yarus, M. (1987). Systematic alterations in the anticodon arm make tRNA(Glu)-Suoc a more efficient suppressor. *EMBO J* 6, 1499-1506.
- Ramachandiran, V., Kramer, G., Horowitz, P.M., and Hardesty, B. (2002). Single synonymous codon substitution eliminates pausing during chloramphenicol acetyl transferase synthesis on *Escherichia coli* ribosomes *in vitro*. *FEBS Lett* 512, 209-212.
- Ramachandiran, V., Willms, C., Kramer, G., and Hardesty, B. (2000). Fluorophores at the N terminus of nascent chloramphenicol acetyltransferase peptides affect translation and movement through the ribosome. *J Biol Chem* 275, 1781-1786.
- Rocha, E.P. (2004). Codon usage bias from tRNA's point of view: redundancy, specialization, and efficient decoding for translation optimization. *Genome Res* 14, 2279-2286.
- Rodnina, M.V., Fricke, R., Kuhn, L., and Wintermeyer, W. (1995). Codon-dependent conformational change of elongation factor Tu preceding GTP hydrolysis on the ribosome. *EMBO J* 14, 2613-2619.
- Rodnina, M.V., Fricke, R., and Wintermeyer, W. (1994). Transient conformational states of aminoacyl-tRNA during ribosome binding catalyzed by elongation factor Tu. *Biochemistry* 33, 12267-12275.
- Rodnina, M.V., Pape, T., Fricke, R., Kuhn, L., and Wintermeyer, W. (1996). Initial binding of the elongation factor Tu.GTP.aminoacyl-tRNA complex preceding codon recognition on the ribosome. *J Biol Chem* 271, 646-652.
- Rodnina, M.V., and Wintermeyer, W. (1995). GTP consumption of elongation factor Tu during translation of heteropolymeric mRNAs. *Proc Natl Acad Sci U S A* 92, 1945-1949.
- Rodnina, M.V., and Wintermeyer, W. (2001). Fidelity of aminoacyl-tRNA selection on the ribosome: kinetic and structural mechanisms. *Annu Rev Biochem* 70, 415-435.
- Rodnina, M.V., and Wintermeyer, W. (2011). The ribosome as a molecular machine: the mechanism of tRNA-mRNA movement in translocation. *Biochem Soc Trans* 39, 658-662.
- Rosano, G.L., and Ceccarelli, E.A. (2009). Rare codon content affects the solubility of recombinant proteins in a codon bias-adjusted *Escherichia coli* strain. *Microb Cell Fact* 8, 41.

## References

- Rudolph, R., Siebendritt, R., Nesslerer, G., Sharma, A.K., and Jaenicke, R. (1990). Folding of an all-beta protein: independent domain folding in gamma II-crystallin from calf eye lens. *Proc Natl Acad Sci U S A* *87*, 4625-4629.
- Rudorf, S., Thommen, M., Rodnina, M.V., and Lipowsky, R. (2014). Deducing the kinetics of protein synthesis in vivo from the transition rates measured in vitro. *PLoS Comput Biol* *10*, e1003909.
- Sakamoto, K., Hayashi, A., Sakamoto, A., Kiga, D., Nakayama, H., Soma, A., Kobayashi, T., Kitabatake, M., Takio, K., Saito, K., *et al.* (2002). Site-specific incorporation of an unnatural amino acid into proteins in mammalian cells. *Nucleic Acids Res* *30*, 4692-4699.
- Saks, M.E., Sampson, J.R., and Abelson, J.N. (1994). The transfer RNA identity problem: a search for rules. *Science* *263*, 191-197.
- Samatova, E., Konevega, A.L., Wills, N.M., Atkins, J.F., and Rodnina, M.V. (2014). High-efficiency translational bypassing of non-coding nucleotides specified by mRNA structure and nascent peptide. *Nat Commun* *5*, 4459.
- Sampson, J.R., and Uhlenbeck, O.C. (1988). Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. *Proc Natl Acad Sci U S A* *85*, 1033-1037.
- Saunders, R., and Deane, C.M. (2010). Synonymous codon usage influences the local protein structure observed. *Nucleic Acids Res* *38*, 6719-6728.
- Schaechter, M. (2015). A brief history of bacterial growth physiology. *Front Microbiol* *6*, 289.
- Schaechter, M., Maaloe, O., and Kjeldgaard, N.O. (1958). Dependency on medium and temperature of cell size and chemical composition during balanced growth of *Salmonella typhimurium*. *J Gen Microbiol* *19*, 592-606.
- Schagger, H. (2006). Tricine-SDS-PAGE. *Nat Protoc* *1*, 16-22.
- Schena, M., Shalon, D., Davis, R.W., and Brown, P.O. (1995). Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* *270*, 467-470.
- Schleif, R. (1967). Control of production of ribosomal protein. *J Mol Biol* *27*, 41-55.
- Schlunzen, F., Zarivach, R., Harms, J., Bashan, A., Tocilj, A., Albrecht, R., Yonath, A., and Franceschi, F. (2001). Structural basis for the interaction of antibiotics with the peptidyl transferase centre in eubacteria. *Nature* *413*, 814-821.
- Schrader, J.M., Chapman, S.J., and Uhlenbeck, O.C. (2011). Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc Natl Acad Sci U S A* *108*, 5215-5220.
- Shah, P., Ding, Y., Niemczyk, M., Kudla, G., and Plotkin, J.B. (2013). Rate-limiting steps in yeast protein translation. *Cell* *153*, 1589-1601.
- Shampo, M.A., and Kyle, R.A. (2002). Kary B. Mullis - Nobel laureate for procedure to replicate DNA. *Mayo Clin Proc* *77*, 606-606.
- Shao, Z.Q., Zhang, Y.M., Feng, X.Y., Wang, B., and Chen, J.Q. (2012). Synonymous codon ordering: a subtle but prevalent strategy of bacteria to improve translational efficiency. *PLoS One* *7*, e33547.
- Sharma, A.K., Minke-Gogl, V., Gohl, P., Siebendritt, R., Jaenicke, R., and Rudolph, R. (1990). Limited proteolysis of gamma II-crystallin from calf eye lens. Physicochemical studies on the N-terminal domain and the intact two-domain protein. *Eur J Biochem* *194*, 603-609.

## References

- Sharma, D., Cukras, A.R., Rogers, E.J., Southworth, D.R., and Green, R. (2007). Mutational analysis of S12 protein and implications for the accuracy of decoding by the ribosome. *J Mol Biol* 374, 1065-1076.
- Sharp, P.M., Bailes, E., Grocock, R.J., Peden, J.F., and Sockett, R.E. (2005). Variation in the strength of selected codon usage bias among bacteria. *Nucleic Acids Res* 33, 1141-1153.
- Sharp, P.M., and Li, W.H. (1987). The rate of synonymous substitution in enterobacterial genes is inversely related to codon usage bias. *Molecular Biology and Evolution* 4, 222-230.
- Shaw, J.J., and Green, R. (2007). Two distinct components of release factor function uncovered by nucleophile partitioning analysis. *Mol Cell* 28, 458-467.
- Shine, J., and Dalgarno, L. (1974). The 3'-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. *Proc Natl Acad Sci U S A* 71, 1342-1346.
- Shozen, N., Iijima, I., and Hohsaka, T. (2009). Site-specific incorporation of PEGylated amino acids into proteins using nonnatural amino acid mutagenesis. *Bioorg Med Chem Lett* 19, 4909-4911.
- Siller, E., DeZwaan, D.C., Anderson, J.F., Freeman, B.C., and Barral, J.M. (2010). Slowing bacterial translation speed enhances eukaryotic protein folding efficiency. *J Mol Biol* 396, 1310-1318.
- Sorensen, M.A., Kurland, C.G., and Pedersen, S. (1989). Codon usage determines translation rate in *Escherichia coli*. *J Mol Biol* 207, 365-377.
- Sorensen, M.A., and Pedersen, S. (1991). Absolute in vivo translation rates of individual codons in *Escherichia coli*. The two glutamic acid codons GAA and GAG are translated with a threefold difference in rate. *J Mol Biol* 222, 265-280.
- Spencer, P.S., Siller, E., Anderson, J.F., and Barral, J.M. (2012). Silent substitutions predictably alter translation elongation rates and protein folding efficiencies. *J Mol Biol* 422, 328-335.
- Starck, S.R., Qi, X., Olsen, B.N., and Roberts, R.W. (2003). The puromycin route to assess stereo- and regiochemical constraints on peptide bond formation in eukaryotic ribosomes. *J Am Chem Soc* 125, 8090-8091.
- Stark, H., Rodnina, M.V., Rinke-Appel, J., Brimacombe, R., Wintermeyer, W., and van Heel, M. (1997). Visualization of elongation factor Tu on the *Escherichia coli* ribosome. *Nature* 389, 403-406.
- Starosta, A.L., Lassak, J., Peil, L., Atkinson, G.C., Virumae, K., Tenson, T., Remme, J., Jung, K., and Wilson, D.N. (2014). Translational stalling at polyproline stretches is modulated by the sequence context upstream of the stall site. *Nucleic Acids Res* 42, 10711-10719.
- Steward, L.E., Collins, C.S., Gilmore, M.A., Carlson, J.E., Ross, J.B.A., and Chamberlin, A.R. (1997). In vitro site-specific incorporation of fluorescent probes into beta-galactosidase. *J Am Chem Soc* 119, 6-11.
- Stewart, E.J., Aslund, F., and Beckwith, J. (1998). Disulfide bond formation in the *Escherichia coli* cytoplasm: an in vivo role reversal for the thioredoxins. *EMBO J* 17, 5543-5550.
- Stoletzki, N., and Eyre-Walker, A. (2007). Synonymous codon usage in *Escherichia coli*: selection for translational accuracy. *Mol Biol Evol* 24, 374-381.

## References

- Studer, S.M., and Joseph, S. (2006). Unfolding of mRNA secondary structure by the bacterial translation initiation complex. *Mol Cell* 22, 105-115.
- Summerer, D., Chen, S., Wu, N., Deiters, A., Chin, J.W., and Schultz, P.G. (2006). A genetically encoded fluorescent amino acid. *Proc Natl Acad Sci U S A* 103, 9785-9789.
- Takyar, S., Hickerson, R.P., and Noller, H.F. (2005). mRNA helicase activity of the ribosome. *Cell* 120, 49-58.
- Tanaka, T., Hori, N., and Takada, S. (2015). How co-translational folding of multi-domain protein is affected by elongation schedule: molecular simulations. *PLoS Comput Biol* 11, e1004356.
- Tats, A., Tenson, T., and Remm, M. (2008). Preferred and avoided codon pairs in three domains of life. *BMC Genomics* 9, 463.
- Thanaraj, T.A., and Argos, P. (1996a). Protein secondary structural types are differentially coded on messenger RNA. *Protein Sci* 5, 1973-1983.
- Thanaraj, T.A., and Argos, P. (1996b). Ribosome-mediated translational pause and protein domain organization. *Protein Sci* 5, 1594-1612.
- Thomas, L.K., Dix, D.B., and Thompson, R.C. (1988). Codon choice and gene expression: synonymous codons differ in their ability to direct aminoacylated-transfer RNA binding to ribosomes in vitro. *Proc Natl Acad Sci U S A* 85, 4242-4246.
- Thompson, R.C. (1988). EFTu provides an internal kinetic standard for translational accuracy. *Trends Biochem Sci* 13, 91-93.
- Torimura, M., Kurata, S., Yamada, K., Yokomaku, T., Kamagata, Y., Kanagawa, T., and Kurane, R. (2001). Fluorescence-quenching phenomenon by photoinduced electron transfer between a fluorescent dye and a nucleotide base. *Anal Sci* 17, 155-160.
- Traverso, G., Diehl, F., Hurst, R., Shuber, A., Whitney, D., Johnson, C., Levin, B., Kinzler, K.W., and Vogelstein, B. (2003). Multicolor in vitro translation. *Nat Biotechnol* 21, 1093-1097.
- Tsalkova, T., Odom, O.W., Kramer, G., and Hardesty, B. (1998). Different conformations of nascent peptides on ribosomes. *J Mol Biol* 278, 713-723.
- Tu, L., Khanna, P., and Deutsch, C. (2014). Transmembrane segments form tertiary hairpins in the folding vestibule of the ribosome. *J Mol Biol* 426, 185-198.
- Tu, L.W., and Deutsch, C. (2010). A folding zone in the ribosomal exit tunnel for Kv1.3 helix formation. *J Mol Biol* 396, 1346-1360.
- Tubulekas, I., and Hughes, D. (1993a). Growth and translation elongation rate are sensitive to the concentration of EF-Tu. *Mol Microbiol* 8, 761-770.
- Tubulekas, I., and Hughes, D. (1993b). Suppression of rpsL phenotypes by tuf mutations reveals a unique relationship between translation elongation and growth rate. *Mol Microbiol* 7, 275-284.
- Tuller, T., Carmi, A., Vestsigian, K., Navon, S., Dorfan, Y., Zaborske, J., Pan, T., Dahan, O., Furman, I., and Pilpel, Y. (2010a). An evolutionarily conserved mechanism for controlling the efficiency of protein translation. *Cell* 141, 344-354.
- Tuller, T., Waldman, Y.Y., Kupiec, M., and Ruppin, E. (2010b). Translation efficiency is determined by both codon bias and folding energy. *Proc Natl Acad Sci U S A* 107, 3645-3650.

## References

- Ude, S., Lassak, J., Starosta, A.L., Kraxenberger, T., Wilson, D.N., and Jung, K. (2013). Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* 339, 82-85.
- Urrutia, A.O., and Hurst, L.D. (2003). The signature of selection mediated by expression on human genes. *Genome Res* 13, 2260-2264.
- Valle, M., Sengupta, J., Swami, N.K., Grassucci, R.A., Burkhardt, N., Nierhaus, K.H., Agrawal, R.K., and Frank, J. (2002). Cryo-EM reveals an active role for aminoacyl-tRNA in the accommodation process. *EMBO J* 21, 3557-3567.
- Vanin, E.F., Greenwel, P., and Symons, R.H. (1974). Structure-activity-relationships of puromycin analogs on Escherichia-coli polysomes. *FEBS Lett* 40, 124-126.
- Varenne, S., Buc, J., Llobes, R., and Lazdunski, C. (1984). Translation Is a non-uniform process - effect of transfer-RNA availability on the rate of elongation of nascent polypeptide-chains. *J Mol Biol* 180, 549-576.
- Vesper, O., Amitai, S., Belitsky, M., Byrgazov, K., Kaberdina, A.C., Engelberg-Kulka, H., and Moll, I. (2011). Selective translation of leaderless mRNAs by specialized ribosomes generated by MazF in Escherichia coli. *Cell* 147, 147-157.
- Voorhees, R.M., Schmeing, T.M., Kelley, A.C., and Ramakrishnan, V. (2010). The mechanism for activation of GTP hydrolysis on the ribosome. *Science* 330, 835-838.
- Voss, N.R., Gerstein, M., Steitz, T.A., and Moore, P.B. (2006). The geometry of the ribosomal polypeptide exit tunnel. *J Mol Biol* 360, 893-906.
- Wang, J., Kwiatkowski, M., and Forster, A.C. (2015). Kinetics of ribosome-catalyzed polymerization using artificial aminoacyl-tRNA substrates clarifies inefficiencies and improvements. *ACS Chem Biol*.
- Wang, K., Neumann, H., Peak-Chew, S.Y., and Chin, J.W. (2007). Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion. *Nat Biotechnol* 25, 770-777.
- Wang, L., Brock, A., Herberich, B., and Schultz, P.G. (2001). Expanding the genetic code of Escherichia coli. *Science* 292, 498-500.
- Wang, L., and Schultz, P.G. (2004). Expanding the genetic code. *Angew Chem Int Ed Engl* 44, 34-66.
- Wannemacher, C.F., and Spector, A. (1968). Protein synthesis in the core of calf lens. *Exp Eye Res* 7, 623-625.
- Watanabe, T., Muranaka, N., Iijima, I., and Hohsaka, T. (2007). Position-specific incorporation of biotinylated non-natural amino acids into a protein in a cell-free translation system. *Biochem Biophys Res Commun* 361, 794-799.
- Weinger, J.S., Parnell, K.M., Dorner, S., Green, R., and Strobel, S.A. (2004). Substrate-assisted catalysis of peptide bond formation by the ribosome. *Nat Struct Mol Biol* 11, 1101-1106.
- Weiss, R.B., Dunn, D.M., Dahlberg, A.E., Atkins, J.F., and Gesteland, R.F. (1988). Reading frame switch caused by base-pair formation between the 3' end of 16s ribosomal-RNA and the messenger-RNA during elongation of protein-synthesis in Escherichia-coli. *EMBO J* 7, 1503-1507.

## References

- Weixlbaumer, A., Jin, H., Neubauer, C., Voorhees, R.M., Petry, S., Kelley, A.C., and Ramakrishnan, V. (2008). Insights into translational termination from the structure of RF2 bound to the ribosome. *Science* *322*, 953-956.
- Wen, J.D., Lancaster, L., Hodges, C., Zeri, A.C., Yoshimura, S.H., Noller, H.F., Bustamante, C., and Tinoco, I. (2008). Following translation by single ribosomes one codon at a time. *Nature* *452*, 598-603.
- Whitford, P.C., Geggier, P., Altman, R.B., Blanchard, S.C., Onuchic, J.N., and Sanbonmatsu, K.Y. (2010). Accommodation of aminoacyl-tRNA into the ribosome involves reversible excursions along multiple pathways. *RNA* *16*, 1196-1204.
- Wilden, B., Savelsbergh, A., Rodnina, M.V., and Wintermeyer, W. (2006). Role and timing of GTP binding and hydrolysis during EF-G-dependent tRNA translocation on the ribosome. *Proc Natl Acad Sci U S A* *103*, 13670-13675.
- Wilson, D.N., and Beckmann, R. (2011). The ribosomal tunnel as a functional environment for nascent polypeptide folding and translational stalling. *Curr Opin Struct Biol* *21*, 274-282.
- Wiltrout, E., Goodenbour, J.M., Frechin, M., and Pan, T. (2012). Misacylation of tRNA with methionine in *Saccharomyces cerevisiae*. *Nucleic Acids Res* *40*, 10494-10506.
- Wistow, G., Slingsby, C., Blundell, T., Driessen, H., De Jong, W., and Bloemendal, H. (1981). Eye-lens proteins: the three-dimensional structure of beta-crystallin predicted from monomeric gamma-crystallin. *FEBS Lett* *133*, 9-16.
- Wistow, G., Turnell, B., Summers, L., Slingsby, C., Moss, D., Miller, L., Lindley, P., and Blundell, T. (1983). X-ray analysis of the eye lens protein gamma-II crystallin at 1.9 Å resolution. *J Mol Biol* *170*, 175-202.
- Wohlgemuth, I., Brenner, S., Beringer, M., and Rodnina, M.V. (2008). Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. *J Biol Chem* *283*, 32229-32235.
- Wohlgemuth, I., Pohl, C., Mittelstaet, J., Konevega, A.L., and Rodnina, M.V. (2011). Evolutionary optimization of speed and accuracy of decoding on the ribosome. *Philos Trans R Soc Lond B Biol Sci* *366*, 2979-2986.
- Wohlgemuth, I., Pohl, C., and Rodnina, M.V. (2010). Optimization of speed and accuracy of decoding in translation. *EMBO J* *29*, 3701-3709.
- Woolhead, C.A., Johnson, A.E., and Bernstein, H.D. (2006). Translation arrest requires two-way communication between a nascent polypeptide and the ribosome. *Mol Cell* *22*, 587-598.
- Woolhead, C.A., McCormick, P.J., and Johnson, A.E. (2004). Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. *Cell* *116*, 725-736.
- Woolstenhulme, C.J., Guydosh, N.R., Green, R., and Buskirk, A.R. (2015). High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. *Cell Rep* *11*, 13-21.
- Woolstenhulme, C.J., Parajuli, S., Healey, D.W., Valverde, D.P., Petersen, E.N., Starosta, A.L., Guydosh, N.R., Johnson, W.E., Wilson, D.N., and Buskirk, A.R. (2013). Nascent peptides that block protein synthesis in bacteria. *Proc Natl Acad Sci U S A* *110*, E878-887.
- Wright, P.R., Georg, J., Mann, M., Sorescu, D.A., Richter, A.S., Lott, S., Kleinkauf, R., Hess, W.R., and Backofen, R. (2014). CopraRNA and IntaRNA: predicting small RNA targets, networks and interaction domains. *Nucleic Acids Res* *42*, W119-123.



## References

- Wu, N., Deiters, A., Cropp, T.A., King, D., and Schultz, P.G. (2004). A genetically encoded photocaged amino acid. *J Am Chem Soc* *126*, 14306-14307.
- Yarus, M. (1982). Translational efficiency of transfer RNA's: uses of an extended anticodon. *Science* *218*, 646-652.
- Yarus, M. (1992). Proofreading, NTPases and translation: constraints on accurate biochemistry. *Trends Biochem Sci* *17*, 130-133.
- Yarus, M., Cline, S., Raftery, L., Wier, P., and Bradley, D. (1986a). The translational efficiency of tRNA is a property of the anticodon arm. *J Biol Chem* *261*, 10496-10505.
- Yarus, M., Cline, S.W., Wier, P., Breeden, L., and Thompson, R.C. (1986b). Actions of the anticodon arm in translation on the phenotypes of RNA mutants. *J Mol Biol* *192*, 235-255.
- Zaborske, J., and Pan, T. (2010). Genome-wide analysis of aminoacylation (charging) levels of tRNA using microarrays. *J Vis Exp*.
- Zaborske, J.M., Narasimhan, J., Jiang, L., Wek, S.A., Dittmar, K.A., Freimoser, F., Pan, T., and Wek, R.C. (2009). Genome-wide analysis of tRNA charging and activation of the eIF2 kinase Gcn2p. *J Biol Chem* *284*, 25254-25267.
- Zaher, H.S., Shaw, J.J., Strobel, S.A., and Green, R. (2011). The 2'-OH group of the peptidyl-tRNA stabilizes an active conformation of the ribosomal PTC. *EMBO J* *30*, 2445-2453.
- Zamecnik, P.C., Stephenson, M.L., and Scott, J.F. (1960). Partial Purification of Soluble RNA. *Proc Natl Acad Sci U S A* *46*, 811-822.
- Zhang, G., Hubalewska, M., and Ignatova, Z. (2009). Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. *Nat Struct Mol Biol* *16*, 274-280.
- Zheng, G., Qin, Y., Clark, W.C., Dai, Q., Yi, C., He, C., Lambowitz, A.M., and Pan, T. (2015). Efficient and quantitative high-throughput tRNA sequencing. *Nat Methods*.
- Zhou, M., Guo, J., Cha, J., Chae, M., Chen, S., Barral, J.M., Sachs, M.S., and Liu, Y. (2013). Non-optimal codon usage affects expression, structure and function of clock protein FRQ. *Nature* *495*, 111-115.
- Zhou, T., Weems, M., and Wilke, C.O. (2009). Translationally optimal codons associate with structurally sensitive sites in proteins. *Mol Biol Evol* *26*, 1571-1580.
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* *31*, 3406-3415.

## Appendix

### Codon-optimized variant of HemK

#### Nucleotide sequence

ATGGAATATCAGCATTGGCTGCGTGAAGCAATTAGCCAGCTGCAGGCAAGCGAAAGCCCCGCGTCGTGATGCAGAAATTCTGCT  
 GGAACATGTTACCGGTAAAGGTCGTACCTTTATTCTGGCATTGGTCAAACCCAGCTGACCGATGAACAGTGTGACGAGCTGG  
 ATGCACTGCTGACCCGTCGTGATGGTGAACCGATTGCACATCTGACCGGTGTTTCGTGAATTTTGGAGCCTGCCGCTGTTT  
 GTTAGCCCGGCAACCCTGATTCCGCGTCCGGATACCGAATGTCTGGTTGAACAGGCACTGGCACGTCTGCCGGAACAGCCGTCG  
 TCGTATTCTGGATCTGGGTACCGGTACCGGTGCAATTGCACTGGCACTGGCAAGCGAACGTCCGGATTGTGAAATTACCGCAG  
 TTGATCGTATGCCGGATGCGATTGCACTGGCACAGCGTAATGCACAGCATCTGGCAATTAATAATATTCGTATTCTGCAGAGC  
 GATTGGTTTAGCGCACTGGCAGGTCAGCAGTTTACCATGATTGTTAGCAATCCGCCGTATATTGATGAACAGGATCCGCATCT  
 GCAGCAGGGTGATGTTTCGTTTTGAACCGCTGACCGCACTGGTTGCAGCAGATAGCGGTATGGCAGATATTGTTTCATATTATTG  
 AACAGAGCCGTAATGCACTGGTTAGCGGTGGTTTTCTGCTGCTGGAACATGGTTGGCAGCAGGGTGAAGCAGTTTCGTCAGGCA  
 TTTATTCTGGCAGGTTATCATGATGTTGAAACCTGTGCTGATTATGGTGATAATGAACGTGTTACCCTGGGTCGTTATTATCA  
 GTAA

#### Codon table used for gene optimization

*Escherichia coli* [gbbct]: 8087 CDS's (2330943 codons)

fields: [triplet] [frequency: per thousand] ([number])							
UUU	24.4 ( 56791)	UCU	13.1 ( 30494)	UAU	21.6 ( 50400)	UGU	5.9 ( 13662)
UUC	13.9 ( 32513)	UCC	9.7 ( 22637)	UAC	11.7 ( 27239)	UGC	5.5 ( 12777)
UUA	17.4 ( 40627)	UCA	13.1 ( 30502)	UAA	2.0 ( 4664)	UGA	1.1 ( 2674)
UUG	12.9 ( 30084)	UCG	8.2 ( 19071)	UAG	0.3 ( 751)	UGG	13.4 ( 31207)
CUU	14.5 ( 33816)	CCU	9.5 ( 22121)	CAU	12.4 ( 28919)	CGU	15.9 ( 37134)
CUC	9.5 ( 22074)	CCC	6.2 ( 14379)	CAC	7.3 ( 17117)	CGC	14.0 ( 32720)
CUA	5.6 ( 12951)	CCA	9.1 ( 21237)	CAA	14.4 ( 33607)	CGA	4.8 ( 11216)
CUG	37.4 ( 87261)	CCG	14.5 ( 33795)	CAG	26.7 ( 62329)	CGG	7.9 ( 18434)
AUU	29.6 ( 68942)	ACU	13.1 ( 30518)	AAU	29.3 ( 68348)	AGU	13.2 ( 30749)
AUC	19.4 ( 45213)	ACC	18.9 ( 44139)	AAC	20.3 ( 47233)	AGC	14.3 ( 33255)
AUA	13.3 ( 31065)	ACA	15.1 ( 35293)	AAA	37.2 ( 86726)	AGA	7.1 ( 16583)
AUG	23.7 ( 55356)	ACG	13.6 ( 31794)	AAG	15.3 ( 35652)	AGG	4.0 ( 9238)
GUU	21.6 ( 50261)	GCU	18.9 ( 44034)	GAU	33.7 ( 78663)	GGU	23.7 ( 55283)
GUC	13.1 ( 30515)	GCC	21.6 ( 50411)	GAC	17.9 ( 41619)	GGC	20.6 ( 47962)
GUA	13.1 ( 30461)	GCA	23.0 ( 53619)	GAA	35.1 ( 81727)	GGA	13.6 ( 31729)
GUG	19.9 ( 46309)	GCG	21.1 ( 49169)	GAG	19.4 ( 45154)	GGG	12.3 ( 28720)

Coding GC 47.30% 1st letter GC 53.83% 2nd letter GC 40.61% 3rd letter GC 47.45%

# Appendix

## Optimization scheme

(Format: <optimal frequency>   <pre-optimisation frequency>   <optimised frequency>)											
<b>D (18)</b>			<b>E (21)</b>			<b>F (9)</b>			<b>G (18)</b>		
gat: 1.00   0.50   1.00	gaa: 1.00   0.71   1.00	ttt: 1.00   0.89   1.00	ggt: 1.00   0.22   1.00	gac: 0.00   0.50   0.00	gag: 0.00   0.29   0.00	ttc: 0.00   0.11   0.00	ggc: 0.00   0.39   0.00	gga: 0.00   0.06   0.00	ggg: 0.00   0.33   0.00		
<b>A (29)</b>			<b>C (5)</b>			<b>* (1)</b>			<b>L (33)</b>		
gct: 0.00   0.21   0.00	tgt: 1.00   0.40   1.00	taa: 1.00   0.00   1.00	tta: 0.00   0.12   0.00	gcc: 0.00   0.10   0.00	tgc: 0.00   0.60   0.00	tga: 0.00   1.00   0.00	ttg: 0.00   0.06   0.00	gca: 1.00   0.21   1.00	ctt: 0.00   0.15   0.00	ctc: 0.00   0.18   0.00	
gcg: 0.00   0.48   0.00							cta: 0.00   0.03   0.00		ctg: 1.00   0.45   1.00		
<b>M (4)</b>			<b>N (5)</b>			<b>H (8)</b>			<b>I (17)</b>		
atg: 1.00   1.00   1.00	aat: 1.00   0.60   1.00	cat: 1.00   0.88   1.00	att: 1.00   0.47   1.00	aac: 0.00   0.40   0.00	cac: 0.00   0.12   0.00		atc: 0.00   0.47   0.00		ata: 0.00   0.06   0.00		
<b>K (2)</b>			<b>T (15)</b>			<b>W (4)</b>			<b>V (14)</b>		
aaa: 1.00   1.00   1.00	act: 0.00   0.20   0.00	tgg: 1.00   1.00   1.00	gtt: 1.00   0.21   1.00	acc: 1.00   0.40   1.00			gtc: 0.00   0.29   0.00		gta: 0.00   0.21   0.00		
	aca: 0.00   0.07   0.00			acg: 0.00   0.33   0.00			gtg: 0.00   0.29   0.00				
<b>Q (22)</b>			<b>P (14)</b>			<b>S (12)</b>			<b>R (21)</b>		
caa: 0.00   0.50   0.00	cct: 0.00   0.14   0.00	tcc: 0.00   0.08   0.00	cgt: 1.00   0.48   1.00	cag: 1.00   0.50   1.00	ccc: 0.00   0.14   0.00	tcg: 0.00   0.25   0.00	cgc: 0.00   0.33   0.00		cga: 0.00   0.10   0.00		
	cca: 0.00   0.14   0.00	agc: 1.00   0.50   1.00	cgg: 0.00   0.10   0.00		ccg: 1.00   0.57   1.00		cgg: 0.00   0.10   0.00				
<b>Y (6)</b>											
tat: 1.00   1.00   1.00											

# Appendix

## Alignment of aa and codon sequences of HemK variants

B121(DE3) K12 MG 1655	M E Y Q H W L R E A I S Q L Q A S E S P R R D A E I L L E H V T G	[33] [33]
B121(DE3) K12 MG 1655	A T G A A T A T C A A C A C T G T T A C G T G A A G C A T A T A G C C A A C T T C A G G C G A G C G A A G C C C G G C G T G A T G C T G A A A T C T G C T G G A A C A T G T T A C C G G C	[99] [99]
optimized	. . . . . G . T . . . . . C . G . . . . . T . . . . . G . . . . . A . . . . . T . . . . .	[99] [99]
B121(DE3) K12 MG 1655	K G R T F I L A F G E T Q L T Q L T D E Q C Q L D A L T R R D G E	[66] [66]
B121(DE3) K12 MG 1655	A A G G G C G T A C T T T A T C C T C G C C T T T G G T G A A C G C A G C T G A C T G A C G A A C A A T G T C A G C A A C T T G A T G C G C T A C T G A C A C G T C G T C G C G A T G G T G A A	[198] [198]
optimized	. . . . . T . . . . . C . . . . . C . . . . . C . . . . . T . . . . . A . . . . . G . . . . . C . . . . .	[198] [198]
B121(DE3) K12 MG 1655	P I A H L T G V R E F W S L P L F V S P A T L I P R P D T E C L V	[199] [199]
B121(DE3) K12 MG 1655	C C A T T G C T C A T T A A C C G G G G T G C G A G A A T T C T G G T C G T T G C C G T T A T T G T T C G C C A G C G A C C T T A A T T C C G C C C G G A T A C G G A G T G T C T G G T G	[297] [297]
optimized	. . . . . A . . . . . C . G . . . . . T . . . . . T . . . . . A G C . . . . . C . G . . . . . T . . . . . C . A . . . . .	[297] [297]
B121(DE3) K12 MG 1655	E Q A L A R L P E Q P C R I L D L G T G T G A I A L A L A S E R P	[132] [132]
B121(DE3) K12 MG 1655	G A G A G C A C T G C C G G T T G C C T G A A C A A C C T T G C C G T A T T C T G A T C T C G G G A C G G T A C C G G G G C G A T T G C G C T T G C G C T G C T A G C G A G C G C C G	[396] [396]
optimized	. . . . . A . . . . . A . . . . . T . . . . . G . . . . . T . . . . . G . . . . . T . . . . . A . . . . . A . . . . .	[396] [396]
B121(DE3) K12 MG 1655	D C E I T A V D R M P D A V A L A Q R N A Q H L A I K N I R I L Q	[165] [165]
B121(DE3) K12 MG 1655	G A C T G C G A A T T A C C G C T G A G A T C G T A T G C C C G A T G C T G T C G C T C T G G C A C A A C G T A A T G C C A G C A T C T G C G G A T C A A A A T A T C G C A T T C T G C A A	[495] [495]
optimized	. . . . . T . . . . . A . . . . . A . . . . . T . . . . . G . . . . . A . . . . . T . . . . . A . . . . . T . . . . .	[495] [495]
B121(DE3) K12 MG 1655	S D W F S A L A G Q Q F T M I V S N P P Y I D E Q D P H L Q Q G D	[198] [198]
B121(DE3) K12 MG 1655	A G C G A C T G G T T A G T G C G C T G C C G G G C A G C A G T T T A C G A T G A T T G T C A G C A A T C C G C C G T A T T G A C G A G C A G C A C C A C A C A T C T T C A A C A A G G C G A T	[594] [594]
optimized	. . . . . T . . . . . C . A . . . . . A . . . . . A . . . . . T . . . . . C . . . . . T . . . . . T . . . . . T . . . . . G . . . . . G . . . . .	[594] [594]
B121(DE3) K12 MG 1655	V R F E P L T A L V A A D S G M A D I V H I I E Q S R N A L V S G	[231] [231]
B121(DE3) K12 MG 1655	G T C C G C T T G A G C C G C T C A C T G C G C T G G T T G C G G C A G A C A G T G G A A T G C A G A C A T C G T G C A T C A T C G A A C A G T C G C G T A C G C G T G G T A T C C G G C	[693] [693]
optimized	. . . . . T . . . . . A . . . . . G . . . . . C . A . . . . . A . . . . . T . . . . . C . . . . . T . . . . . T . . . . . A . . . . . A . . . . . T . . . . . A . . . . .	[693] [693]
B121(DE3) K12 MG 1655	G F L L L E H G W Q Q G E A V R Q A F I L A G Y H D V E T C R D Y	[264] [264]
B121(DE3) K12 MG 1655	G G C T T C T G C T T C T G A A C A T G G C T G G C A G C A G G C G A A G C G T G C G A C A A G C A T T T A T C C T C G C G G G T A T C A T G A C G T C G A A A C C T G C C G T G A C T A T	[792] [792]
optimized	. . . . . T . . . . . G . . . . . T . . . . . T . . . . . A . . . . . T . . . . . A . . . . . T . . . . . T . . . . . T . . . . . T . . . . .	[792] [792]
B121(DE3) K12 MG 1655	G D N E R V T L G R Y Y Q	[277] [277]
B121(DE3) K12 MG 1655	G G T A T A C G A G C G C G T A A C G C T C G G C C G C T A T T A C A A	[831] [831]
optimized	. . . . . T . . . . . A . . . . . T . . . . . T . . . . . C . . . . . G . . . . . T . . . . .	[831] [831]

## Appendix

### List of abbreviations

Abbreviation	Description
A site	acceptor site
aa	amino acid
aaRS	aminoacyl-tRNA synthetase
aa-tRNA	aminoacyl-tRNA
AK	adenylate kinase
anti-SD	anti-Shine-Dalgarno
<i>B. taurus</i>	<i>Bos taurus</i>
CAI	codon adaption index
Cam	chloramphenicol
CTD	C-terminal domain
Cy	cyanine
D	dihydrouracil
dbl	doubling
E site	exit site
<i>E. coli</i>	<i>Escherichia coli</i>
EF	elongation factors
EF-P	elongation factor-P
GBC	$\gamma$ -B-crystallin
GdnHCl	guanidinium HCl
GFP	green fluorescent protein
h	hour
IC	initiation complex
IF	initiation factor
mRNA	messenger RNA
nc-aa	non-canonical amino acid
nt	nucleotide
NTD	N-terminal domain
ORF	open reading frame
P site	peptidyl-tRNA binding site
PEG	polyethylene glycol
PeT	photoinduced electron transfer
PGK	phosphoglycerate kinase
PIC	pre-initiation complex
PK	proteinase K
Pmn	puromycin
PMSF	phenylmethylsulfonylfloride
PTC	peptidyl transferase center
pyl	pyrrolysine
RF	release factor
RRF	ribosome recycling factor
S	Svedberg
s	second
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
SD	Shine-Dalgarno
SDS	sodium dodecyl sulfate
SRP	signal recognition particle
TAL	Transcriptom Analyse Labor
TAMRA	tetramethyl-rhodamine
TCA	trichloroacetic acid
tRNA	transfer RNA

## List of Figures

Figure 1 1: Crystal structure of the 70S ribosome containing mRNA and tRNA.....	2
Figure 1 2: Tertiary fold of tRNA.....	3
Figure 1 3: Main stages of protein synthesis.....	4
Figure 1 4: Three potential modes of interaction of aa-tRNA with a given codon in the A site. ....	7
Figure 2 1: Elementary steps of decoding in the A site.....	19
Figure 2 2: Determination of $\omega_{90}$ and $\omega_{9,10}$ at 20°C.....	22
Figure 2 3 Single-turnover <i>in vitro</i> translation of CspA 1-70 aa at different temperatures.....	23
Figure 2 4: Calculated codon specific elongation rates at saturating concentration of all aa-tRNA.....	24
Figure 2 5: Mapping of transition rates <i>in vivo</i> and <i>in vitro</i> .....	26
Figure 2 6: Single-turnover translation of natural mRNA.....	29
Figure 2 7: Characterization of pausing during translation of HemK 1-277 aa.....	32
Figure 2 8 Global model for accumulation and decay of translation intermediates during translation of HemK 1-277 aa.....	34
Figure 2 9: Time course of translation of HemK 1-277 aa at 20°C.....	35
Figure 2 10: Translation of shorter segments of the mRNA coding for HemK.....	36
Figure 2 11: Translation of mRNA coding for HemK with the nucleotide sequence from the <i>E. coli</i> strain BL21(DE3) and codon-optimized sequence.....	37
Figure 2 12: Detailed densitometric analysis of the accumulation of nascent peptides during translation of the different HemK nucleotide sequences. ....	39
Figure 2 13: Hybridization of peptidyl-tRNA on micro-array chips. ....	43
Figure 2 14: Decay of normalized intensities for hybridization of peptidyl-tRNA to probes Lys and Gly2 during translation of HemK 1-112 aa.....	45
Figure 2 15: Schematic depiction of sample preparation protocol for hybridization of 3'-labeled tRNA. ....	46
Figure 2 16: Analysis of the aminoacylation level. ....	47
Figure 2 17: Identification of tRNA in the P site of a defined ribosomal complex.....	48
Figure 2 18: Incorporation of fluorescent probes into the N-terminus of nascent chains. ....	51
Figure 2 19: Aminoacylation of tRNA <sup>Cys</sup> variants.....	54
Figure 2 20: Peptide bond formation with the transcript and fully modified Cys-tRNA <sup>Cys</sup> <sub>GCA</sub> .....	55
Figure 2 21: A-site accommodation of Cys-tRNA.....	56
Figure 2 22: Fluorescence labeling and purification of Cys-tRNA. ....	57
Figure 2 23: Chemical structures of fluorophores used for incorporation at internal positions of nascent chains. ....	58
Figure 2 24: Dipeptide formation with fluorescence-labeled Cys-tRNA <sup>Cys</sup> <sub>GCA</sub> at increasing concentrations of purified EF-Tu•GTP binary complex. ....	59
Figure 2 25: Incorporation of fluorescence-labeled Cys-tRNA <sup>Cys</sup> <sub>GCA</sub> during into HemK 1-70 aa. ....	60
Figure 2 26: Fluorescence change of BodipyFL-C1-IA-Cys-tRNA <sup>Cys</sup> <sub>GCA</sub> upon decoding and accommodation in the A site. ....	62
Figure 2 27: Measurement of FRET between the N-terminal BodipyFL-C2-Met and QSY-35 at position 34 of released chains of HemK 1-70 aa K34C.....	65
Figure 2 28: Dependence of the incorporation efficiency on the aa upstream of the incorporation site. ....	67

## Appendix

Figure 2 29: Influence of upstream Gly residues on the incorporation of fluorescent probes into HemK 1-70 aa K34C. ....	69
Figure 2 30: Reconstruction of the density of the peptidyl-tRNA in the P site and the nascent chain in the ribosomal exit tunnel at 5.6 Å resolution by cryo-electron microscopy. ....	70
Figure 2 31: Incorporation of Atto488-mal and Atto633-IA into HemK 1-70 aa K34C constructs with upstream Gly substitutions. ....	72
Figure 2 32: Structural properties and codon usage profile of synonymous sequences coding for GBC ..... 74	74
Figure 2 33: Analysis of the recombinant expression of GBC from the sequence of U and H in <i>E. coli</i> BL21(DE3). ....	76
Figure 2 34: Single-turnover translation of the U and H sequences coding for GBC. ....	78
Figure 2 35: Proteolysis of full-length GBC 1-175 aa and GBC 1-89 aa by PK. ....	80
Figure 2 36: Pulse proteolysis of ribosome-bound and ribosome released GBC 1-175 aa. ....	82
Figure 2 37 Incorporation of BodipyFL-C <sub>1</sub> -IA-Cys-tRNA <sup>Cys</sup> <sub>CUA</sub> U32C into position 88 upon translation of GBC 1-175 aa. ....	84

## List of Tables

Table 2 1: Summary of rate constants used for modeling of elongation rates in vitro. ....	20
Table 2 2: Parameters obtained by fitting the formation of full-length CspA to a model with a delay followed by a single exponential increase. ....	23
Table 2 3: List of scaled in vivo transition rates for different growth rates. ....	25
Table 2 4: List of translation intermediates of within the first 100 aa of HemK that were detected during translation of mRNA with different length. ....	36
Table 2 5: Positions of the SD-like sequences and the predicted strength of interactions ..... 41	41
Table 2 6: Parameter obtained by fitting the fluorescence change during translation of HemK 1-70 aa K34C in presence of increasing concentrations BodipyFL-C <sub>1</sub> -IA-Cys-tRNA <sup>Cys</sup> <sub>GCA</sub> ..... 62	62
Table 2 7: Parameters obtained by fitting the fluorescence traces in Figure 2 36 B to a model with a delay followed by a single-exponential increase. ....	86
Table 4 1 List of equipment. ....	107
Table 4 2 List of software. ....	107
Table 4 3 List of Chemicals and Consumables. ....	108
Table 4 4 List of columns. ....	109
Table 4 5 List of Strains. ....	109
Table 4 6 List of Plasmids. ....	109
Table 4 7: List of buffers and solutions. ....	110
Table 4 8 List of Primers. ....	113
Table 4 9 List of synthetic mRNA. ....	117
Table 4 10: Primer pairs for amplification of DNA templates uses for <i>in vitro</i> transcription. ....	117
Table 4 11: DNA probes for hybridization of tRNA on micro-array. ....	124

## Acknowledgment

### Acknowledgment

The 4 years that I have spent in Göttingen were very rich in experiences and moments encompassing the complete spectrum of emotions. I would like to acknowledge the people that were contributing to them and were participating/precipitating in them.

I am very grateful to Marina V. Rodnina for selecting me as PhD student in her department, for all the support during the experimental work and especially during preparation of this thesis.

I would like to thank my thesis committee members Heinz Neumann and Marina Bennati for their support and advises. I would also like to thank the members of the extended examination committee Claudia Höbartner, Helmut Grubmüller, and Holger Stark.

I am exceptionally grateful to Jörg Mittelstätt for the great supervision during the first 1.5 years of my doctoral studies. I am grateful to Jörg and Andrey Konevega for teaching me the largest part of what I currently know about protein translation and especially the tricky little tRNA molecules. In team work, Jörg and Andrey created the best working atmosphere during my entire doctoral studies. I am further really grateful to Lena and Andrey for the perfectly arranged invitation to Saint Petersburg and Gatchina and for being great hosts during the complete trip.

I would also like to thank Marcus Jäger for supervision and for providing in depth insights into fluorescence spectroscopy and protein folding.

I would like to thank Reinhard Lipowsky and especially Sophia Rudolf (Theory and Bio-Systems, Max Planck Institute of Colloids and Interfaces, Golm) for a continuing, interdisciplinary and particularly fruitful collaboration. I have really enjoyed the collaboration with Sophia and I hope that we will be able to solve the puzzle of sequence-specific translational pausing in joined efforts.

I am deeply grateful to the team at the Transcriptome Analyse Labor (University Medicine Göttingen). I would like to thank Gabriela Salinas-Riester for supporting and in principle enabling the projects involving hybridization of tRNA on micro-array chips by giving me full access to the instrumentation, for helpful discussions and



## Acknowledgment

explanations concerning micro-array techniques and analysis, for help with design of the cDNA probes and experimental approaches. I would also like to thank Susanne Luthin for all the help in preparation of samples, spotting and hybridization; and especially for the preparation of the Cy3-labeled Spike-In probes. I would like to further thank Claudia Pommerenke for help with data analysis.

I would like to thank the team of Vladimir Belov (Facility for Synthetic Chemistry, Max Planck Institute f. Biophysical Chemistry) for synthesis of organic molecules.

I would like to thank Ingo Wohlgemuth for all the support in trouble-shooting, providing theoretical background and discussions. In principle, when I had a question then I was first asking Ingo. I also would like to thank Ingo for the creative brain-storming that led to the idea for the peptidyl-tRNA array and for proofreading my thesis.

I am grateful to Ekaterina Samatova and Riccardo Belardinelli for many helpful insights and discussions. I would also like to thank Katya and Ricc for several collaborative experiments and projects.

I would like to thank Lili Dörfel forming a productive and pleasant environment in our corner of the lab and for all the helpful discussions back to back.

I would like to thank Shrutee Jakhanwal for cloning and discovering the increased translational efficiency of Cys-tRNA<sup>Cys</sup><sub>UCA U32C</sub> during her lab rotation.

I am grateful to all members of the department of Physical Biochemistry who design, struggle, and finally create their own experimental approaches or perform their research in a collaborative manner.

I am grateful to Dimitra Papastavrou for all the help in administrative and organizational tasks.

I am grateful to all technical assistants of the department. Anna Pfeifer for helping in RNA preparations and providing me with additional EF-Tu. Tanja Wiles for preparation of components for *in vitro* transcription and for performing transcription. Theresia Uhlendorf for cloning and purification of functional CysRS. Anthony Wiles

## Acknowledgment

for preparation of “lab-made” isothermal assembly aliquots (Gibson assembly). Olaf Geintzer for purification and preparation of tRNA. Sandra Kappler for ribosome preparations. Michael Zimmermann for EF-G. Christina Kothe for initiation factors. I would also like to thank Manuela Beck-Corell and Franziska Hummel.

So far all the acknowledgements were work-related. Yes, I have spent a big part of my day in the lab, but not all. I would also like to acknowledge the people that kept my happy and sane during these years with many “ups and downs”.

I would like to thank Albena for being the best significant other that I could image, for all the support in every situation, and for the good times that we have spent during the last years (road trip to BG, diving, Barcelona and other escapes from the “local minimum”). I am sure that I would have left Göttingen several times without her support. She also gave me courage to look further into the future.

I have also found many new friends here in Göttingen. I was invited during the first weekend to join the grilling in the garden of the MPI settlement in the Immanuel-Kant-Strasse and I felt welcome. Grilling in this gardens started to form a ritual to meet among friends, so that in the end I have also moved into the Kant-Strasse. Rituals in another form started when I was invited to join trips to go skiing in France and Austria. I am really grateful for all that. I have also found friends in the department and in the GGNB Biomolecules program. I am really happy that I have met you, Albena, Sergey, the two Katyas, Jörg, Kathrin, Andrey, Lena, Raffa, Ricc, Dima, Inessa, Tales, Karine, Andreea, Irena, Ingo, Chris, Karin and Liu here in Göttingen.

Finally, I would like to thank my parents and my sister for supporting me during all the years with all their love.

## Curriculum vitae

### Curriculum vitae

#### *Personal information*

Michael Sebastian Thommen

born on the 23<sup>rd</sup> of November 1985

Citizen of Vevey, Vaud, Switzerland

#### *Relevant Education*

09/2011 - 8/2015

Doctoral studies in Biology

Department for Physical Biochemistry

Max Planck Institute for Biophysical Chemistry

Göttingen

09/2009 - 03/2011

M.Sc. studies in Biology

Department of Biology

Eidgenössische Technische Hochschule (ETH)

Zürich

09/2006 - 03/2010

B.Sc. studies in Biology

Department of Biology

Eidgenössische Technische Hochschule (ETH)

Zürich



