# Cotranslational folding of HemK C-terminal domain

**Dissertation**

for the award of the degree

“Doctor rerum naturalium”

of the Georg-August-Universität Göttingen

within the doctoral program

GGNB Biomolecules: Structure – Function – Dynamics

of the Georg-August University School of Science (GAUSS)

submitted by

**Jan Daberger**

from Bad Mergentheim, Germany

Göttingen 2022

**Thesis Committee**

Prof. Dr. Marina Rodnina

Physical Biochemistry, Max Planck Institute for Multidisciplinary Sciences

Prof. Dr. Jörg Enderlein

III. Institute of Physics – Biophysics, Georg-August University Göttingen

Dr. Alexander Stein

Membrane Protein Biochemistry, Max Planck Institute for Multidisciplinary Sciences

**Members of the Examination Board**

Referee: Prof. Dr. Marina Rodnina

Physical Biochemistry, Max Planck Institute for Multidisciplinary Sciences

2nd Referee: Prof. Dr. Jörg Enderlein

III. Institute of Physics – Biophysics, Georg-August University Göttingen

**Further members of the Examination Board**

Dr. Alexander Stein

Membrane Protein Biochemistry, Max Planck Institute for Multidisciplinary Sciences

Prof. Dr. Kai Tittmann

Molecular Enzymology, Georg-August University Göttingen

Dr. Sonja Lorenz

Ubiquitin signaling specificity, Max Planck Institute for Multidisciplinary Sciences

Dr. Alex Faesen

Biochemistry of Signal Dynamics, Max Planck Institute for Multidisciplinary Sciences

Date of oral examination: February 13th, 2023.

# Affidavit

I hereby declare that the presented dissertation entitled “Cotranslational folding of HemK C-terminal domain” has been written independently and with no other sources and aids than quoted.

Göttingen, 12.12.2022

Jan Daberger

## Related publications

Samatova, E., Daberger, J., Liutkute, M., and Rodnina, M.V. (2021). Translational Control by Ribosome Pausing in Bacteria: How a Non-uniform Pace of Translation Affects Protein Production and Folding. Frontiers in Microbiology 11, 619430.

# Table of content

# 1. Abstract

The correct folding of proteins is essential for the survival of the cell. Despite the importance of the question, it remains largely unclear how proteins, in particular large multidomain proteins with mixed α-helical and β-strand structures, fold cotranslationally. In this study, we chose the C-terminal domain of *E. coli* enzyme HemK as the model protein to study folding in a fully reconstituted *in vitro* translation system. Here, we present the folding pathway of the protein step-by-step employing force profile assay using the SecM arrest peptide in combination with time-resolved and steady-state photo-induced electron transfer and Förster resonance energy transfer measurements. This combination allows us to analyze the folding pattern both inside and outside of the polypeptide exit tunnel of the ribosome and to assess the stepwise compaction of the nascent chain. The N-terminal β-hairpin of the C-terminal domain starts to fold inside the polypeptide exit tunnel and undergoes structural rearrangements before emerging from the ribosome. As the nascent peptide emerges from the ribosome, it folds sequentially, starting from the first β-hairpin, which could serve as the folding nucleus for the following C-terminal part adopting a Rossmann fold. The folded domain remains highly dynamic on the ribosome and the final compaction into the native structure occurs after its release from the ribosome.

# 2. Introduction

Translation of the messenger RNA (mRNA) into the sequence of amino acids leads to the production of proteins in all living cells. The key player in mRNA translation is the ribosome, the ribozyme that reads the mRNA codons by selecting the aminoacylated tRNA (aa-tRNA) with the correct amino acid (aa) and facilitates the peptide bond formation. mRNA translation is a cyclic, multistep and multifactorial process that can be separated in four distinct steps (Figure 1).

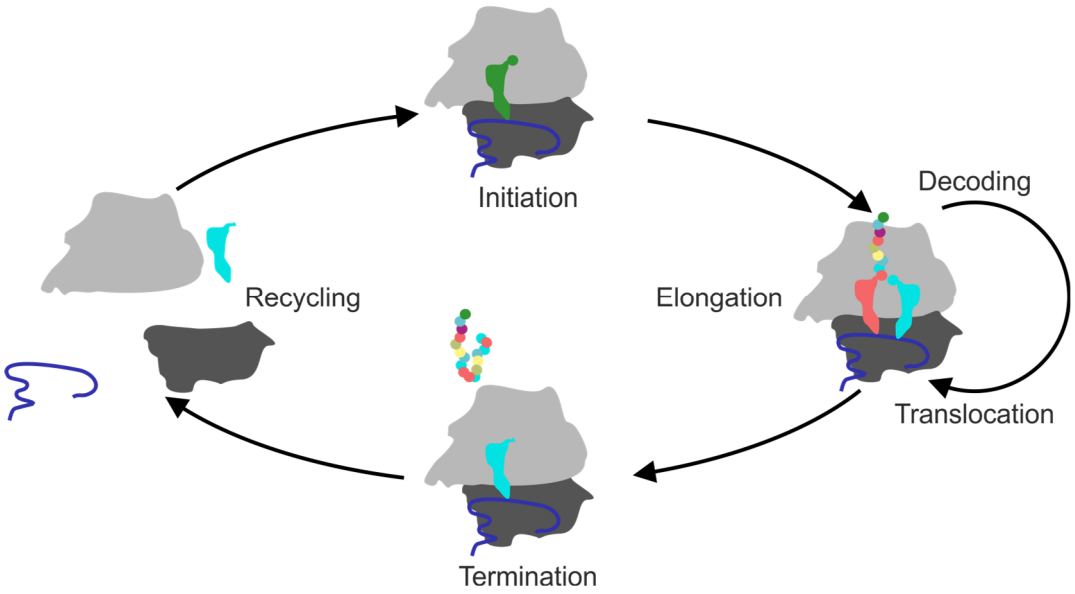



Figure 1. Overview of the four major steps involved in mRNA translation.

During the initiation the ribosome and the initiator tRNA are positioned at the start codon of the open reading frame on the mRNA. During the subsequent elongation the mRNA codons are decoded, the correct aa-tRNA is selected, peptide bond formation takes place, and the ribosome is translocated from codon to codon to synthesize the protein step-by-step. When the ribosome encounters a stop codon, the nascent protein is released with the help of termination (release) factors. Finally, at the recycling step the ribosome is split into subunits to become available for a new round of translation. The next chapter will introduce the structural features of the ribosome and all four steps of translation in detail, with the main focus on the processes in bacteria and specifically in *E. coli*.

## 2.1. Translation

### 2.1.1. Ribosome

The bacterial ribosome (70S) consists of the small 30S subunit and the large 50S subunit. The 30S subunit comprises the 16S ribosomal RNA (rRNA) and 21 ribosomal proteins and the 50S subunit of the 23S rRNA, the 5S rRNA, and 33 proteins. Structurally, the ribosome has four distinct sites, namely the mRNA tunnel, the decoding center, the peptidyl-transfer center (PTC) and the polypeptide exit tunnel that all play different roles during translation (reviewed in Melnikov et al., 2012; Schmeing and Ramakrishnan, 2009).

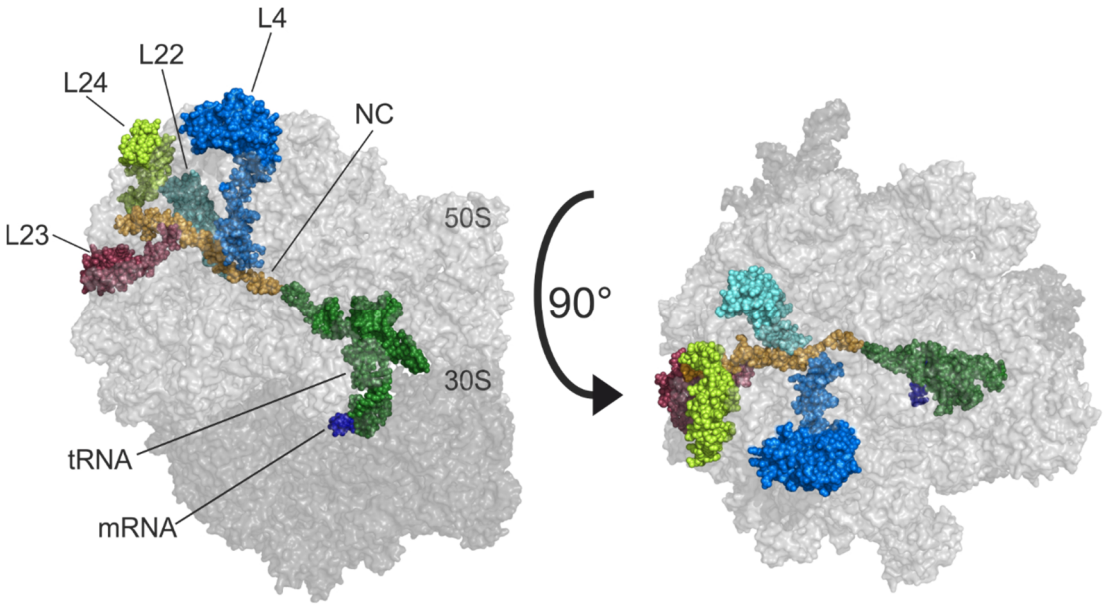



Figure 2. Cryo-EM structure of *E. coli* ribosome (PDB: 7OT5; (Agirrezabala et al., 2022)). Shown in dark grey is the 30S subunit and in light grey the 50S subunit. Both surfaces are shown with 60% transparency to visualize the inner parts of the ribosome, i.e., mRNA (dark blue), P-site tRNA (dark green), the polypeptide exit tunnel, indicated by the 70 aa long nascent chain of CspA (NC; orange) and the ribosomal proteins that line the polypeptide exit tunnel L4 (blue), L22 (cyan), L23 (red) and L24 (light green).

The mRNA tunnel is part of the 30S subunit. During the initiation the mRNA is recruited there and during translation the mRNA is threaded through. The mRNA tunnel ensures the correct positioning of the mRNA at the decoding center. The decoding center is also part of the 30S subunit and hosts the site of the mRNA–tRNA interaction and the correct aa-tRNA selection during the elongation cycle. The PTC is located on the 50S subunit and is the catalytic center of the ribosome where it facilitates the peptide bond formation between the peptidyl-tRNA and the accommodated aa-tRNA. Lastly, the polypeptide exit tunnel leads the peptide from the PTC

to the outside of the ribosome (reviewed in Melnikov et al., 2012; Schmeing and Ramakrishnan, 2009). In Figure 2 the polypeptide exit tunnel is completely occupied by the nascent chain of CspA. The exit tunnel is approximately 100 Å long and varies in width between 10 and 20 Å (Nissen et al., 2000). It is mainly lined with negatively charged rRNA, apart from the four 50S ribosomal proteins L4, L22, L23 and L24. L4 and L22 form the narrow 10 Å wide constriction site that is located roughly 30 Å away from the PTC and the proteins L23 and L24 build the wider vestibule at the end of the tunnel with a diameter of roughly 20 Å (Figure 2) (Ban et al., 2000; Nissen et al., 2000). In general, the polypeptide exit tunnel is highly conserved between the PTC and the constriction site in all three domains of life and deviates closer to the outside of the tunnel (Dao Duc et al., 2019). For example in eukaryotes, the protein L23 that is located at the vestibule has a truncation, which is substituted by L39e (reviewed in Melnikov et al., 2012). Additionally, the polypeptide exit tunnel of eukaryotes possess a second constriction site roughly 50 Å away from the PTC, contributing to the overall smaller size of the eukaryotic polypeptide exit tunnel (Dao Duc et al., 2019).

### 2.1.2. Initiation

Initiation comprises the correct assembly and positioning of the ribosome and the initiator tRNA on the mRNA. Most commonly in *E. coli*, initiation takes place on mRNA that contains a Shine-Dalgarno (SD) sequence upstream of the coding region and the start codon AUG. The initiation occurs on the 30S subunit and involves mRNA, the initiator tRNA (fMet-tRNA$^{fMet}$) and initiation factors IF1, IF2, and IF3 (reviewed in Rodnina, 2018).

In the first step, the pre-initiation complex (30S PIC) is assembled by recruiting IF3 and the the GTPase IF2 to the 30S subunit, followed by the binding of IF1 and fMet-tRNA$^{fMet}$. The mRNA itself can bind any time during the assembly of the 30S PIC and is positioned by the interaction of the SD sequence with the anti-SD sequence in the 16S rRNA of the 30S subunit (Milon et al., 2012). IF3 controls the fidelity of fMet-tRNA$^{fMet}$ selection (reviewed in Gualerzi and Pon, 2015) and is able to reject the fMet-tRNA$^{fMet}$ in case that structured mRNA binds to the 30S PIC (Chen and Wen, 2022).

Start codon selection proceeds via codon–anticodon interaction of the mRNA with fMet-tRNA$^{fMet}$, forming the 30S IC (reviewed in Gualerzi and Pon, 2015; Milon and Rodnina, 2012; Rodnina, 2018). Subsequent joining with the 50S subunit leads to GTP hydrolysis by IF2 and the displacement of IF3 from the subunit interphase (Goyal et al., 2017). This is accompanied by a conformational change of IF2, release of the GTPase reaction product

inorganic phosphate from IF2, the subsequent disruption of the IF2–fMet-tRNA$^{fMet}$ interaction, and the dissociation of IF1 and IF2 (Goyal et al., 2015). Recent findings suggest that after GTP hydrolysis, IF2 stays bound to the ribosome, interacting with fMet-tRNA$^{fMet}$ as IF2–GDP. Phosphate release and consequent dissociation of IF2 leads to a conformational change that positions the fMet-tRNA$^{fMet}$ in the P site and hence, in the elongation competent form of the initiation complex (Kaledhonkar et al., 2019).

### 2.1.3. Elongation

Elongation is the central part of the mRNA translation process. It involves the decoding of the codon in the A site, the correct accommodation of the aa-tRNA, peptide bond formation and translocation by which the ribosomes moves to the next codon. In the first step the A-site codon is decoded by the incoming aa-tRNA that is delivered by the EF-Tu–GTP–aa-tRNA complex (ternary complex; cyan and orange in Figure 3).

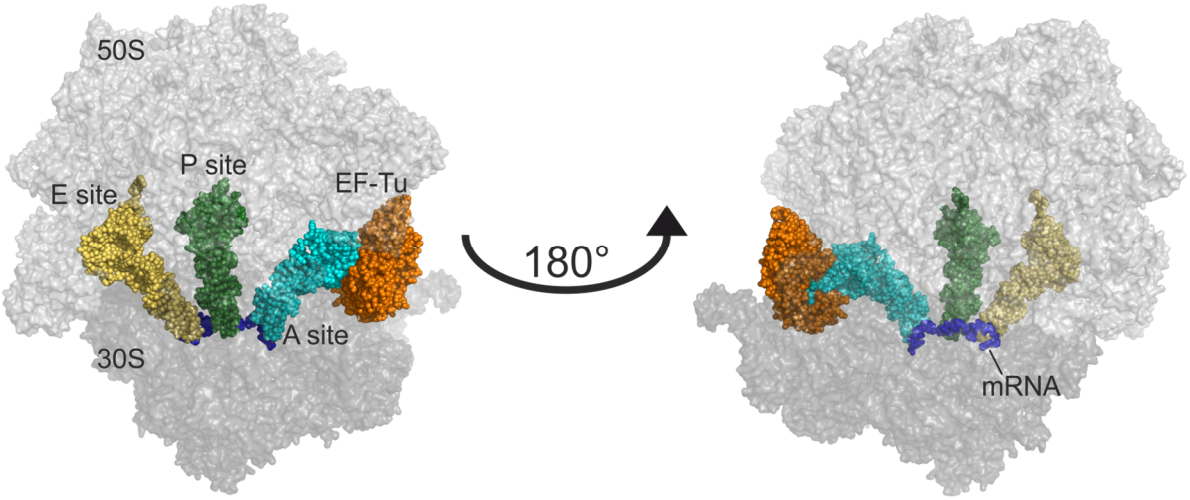



Figure 3. Elongation complex of *E. coli* ribosome before accommodation of the A-site tRNA (PDB: 5AFI; (Fischer et al., 2015)). The 50S subunit is shown in light grey and the 30S subunit in dark grey. Both surfaces are shown with 60% transparency to visualize the inner parts of the ribosome, i.e., the mRNA (blue), E-site tRNA (yellow), P-site tRNA (green), A/T-site tRNA (cyan) with the anticodon positioned in the A site and the CCA end carrying the amino acid complexed by EF-Tu (orange).

The first two base pairs of a correct codon–anticodon complex acquire a Watson-Crick geometry, which is sensed by the universally conserved bases A1492 and A1493 of the 30S subunit in their so-called flipped-out conformation (Ogle et al., 2001). This conformational change enables the nucleotide G530 to change its conformation from the syn to anti conformation interacting with the second and third base pair (Ogle et al., 2001) and eventually

leads to closing the domain of the 30S subunit around the A site (Ogle et al., 2002). This in turn induces GTP hydrolysis in EF-Tu. Figure 3 shows the *E. coli* elongation complex right after GTP hydrolysis, trapped by the antibiotic kirromycin (Fischer et al., 2015). The subsequent phosphate release after GTP hydrolysis triggers a conformational change in EF-Tu, resulting in the release of the aa-tRNA from EF-Tu and the correct accommodation of the aa-tRNA into the A site, positioning the amino acid in the PTC (reviewed in Rodnina et al., 2017). The EF-Tu–GDP complex dissociates from the ribosome and exchanges GDP for GTP with the help of the nucleotide exchange factor EF-Ts, thereby rendering EF-Tu active for the formation of a new ternary complex (reviewed in Maracci and Rodnina, 2016). After aa-tRNA is accommodated in the A site, the PTC catalyzes the peptide bond formation by which the amino group of the A-site amino acid nucleophilically attacks the ester bond of the peptidyl-tRNA (reviewed in Rodnina, 2018). The peptide is transferred from the P-site tRNA to the amino acid of the A-site tRNA. This induces a relative rotation of the subunits, which brings the tRNAs in the hybrid states P/E and A/P, where P/E means that the tRNA anticodon is still located in the P site on the 30S subunit, whereas the tRNA CCA end moves into the E site, thereby moving the tRNAs into the direction of translocation (reviewed in Schmeing and Ramakrishnan, 2009). The rotated form is greatly accelerated and stabilized by EF-G binding and the consequent GTP hydrolysis. Ultimately, the release of the phosphate from EF-G and the conformational change allows to reinstate the non-rotated state, but now with the two tRNAs in the E and P site, respectively (Petrychenko et al., 2021). Subsequent dissociation of EF-G leaves the A site free for the next cycle of decoding and translocation.

### 2.1.4. Termination

Termination is the process that promotes the release of the nascent peptide by hydrolysis of the peptidyl-tRNA ester bond. During a functional translation cycle, termination occurs at the end of the protein coding region of the mRNA when one of the three possible stop codons UAG, UUA or UGA enters the A site, where release factor (RF) 1 interacts with the stop codons UAG and UAA and RF2 with UGA and UAA (reviewed in Korostelev, 2011). RF1 and RF2 have a conserved Gly-Gly-Gln (GGQ) motif that facilitates peptidyl-tRNA hydrolysis at the PTC (reviewed in Youngman et al., 2008). The GGQ is positioned analogously to the aminoacyl group of the A-site tRNA in the PTC (reviewed in Korostelev, 2011). Structural data suggest that the Gln interacts with the ribose of the 3´adenosine of the P-site tRNA to stabilize a reaction-active conformation (Laurberg et al., 2008). The Gln of the GGQ motif is methylated,

which greatly increases the rate of peptide release by up to 70-fold for RF1 and 30-fold for RF2, depending on the peptide context (Pierson et al., 2016). RF3 promotes the dissociation of RF1 and RF2 from the ribosome (reviewed in Rodnina, 2018; Youngman et al., 2008).

Termination also occurs as part of the ribosome rescue mechanism after ribosome stalling or translation on non-stop mRNA. In bacteria, the transfer-messenger (tm)RNA/ SmpB system eventually displays a stop codon to trigger canonical termination (reviewed in Buskirk and Green, 2017; Huter et al., 2017b; Muller et al., 2021). Alternative release factor (Arf)A acts as a rescue system if tmRNA is not available. ArfA binds to the empty A site and provides an interaction platform for the recognition by RF2, which leads to the hydrolysis of the peptidyl-tRNA (Huter et al., 2017c). Alternatively, ArfB can perform the hydrolysis itself and does not require to recruit the canonical release factors (reviewed in Huter et al., 2017b; Muller et al., 2021). ArfB is able to compete with short mRNA stretches that reaches into the A site, in the event of ribosome stalling (Carbone et al., 2020), but displays the highest catalytic efficiency when the A site is completely devoid of mRNA (Chan et al., 2020).

#### 2.1.5. Recycling

After peptide release, the two ribosomal subunits are still joined as 70S ribosome and contain mRNA and the deacylated P-site tRNA. Splitting of the ribosome into subunits is catalyzed by the ribosome recycling factor (RRF) and EF-G. The order of binding seems to be important, with RRF binding first to the ribosome, preferentially in its rotated state, followed by EF-G. The subsequent GTP hydrolysis by EF-G splits the subunits (reviewed in Rodnina, 2018; Seely and Gagnon, 2022).

### 2.2. Rate of translation

All steps that are involved in translation are fundamental, highly conserved mechanisms. However, there are factors that lead to small deviations from these processes in a non-uniform fashion that can tune mRNA translation at every step. In particular the factors that modulate the elongation step are of interest, as they have a direct influence on the codon-specific rate of translation and therefore the timing the protein emerges from the polypeptide exit tunnel.

It is long known that the same protein sequence can be encoded in the mRNA in various ways using synonymous codons, which specify the insertion of the same amino acid, but differ in their nucleotide composition. Synonymous codons can be read by a single tRNA, such as

$tRNA^{Lys}$ or $tRNA^{Phe}$, which in *E. coli* decode their codons AAA/AAG and UUU/UUC, respectively. However, for most four- and six-codon families there are several tRNAs that deliver the same amino acid, but differ in their decoding preferences; such tRNAs are called isoacceptor tRNAs. Isoacceptor tRNAs can usually recognize several synonymous codons through non-Watson-Crick 'wobble' base pairing at the third codon position. How many and which codons can be decoded by one particular tRNA is controlled by post-transcriptional tRNA modification (Grosjean and Westhof, 2016). The rate of decoding a particular codon depends not only on the complementarity of the codon–anticodon complex, but also on the concentration and properties of the respective cognate aa-tRNA in the total tRNA pool.

One important question is whether rare codons are decoded more slowly than the abundant codons, thereby locally slowing down the progression of ribosomes along the mRNA, which is called ribosome pausing. This question should be readily answered by ribosome profiling, a method that enables the detailed measurements of translation *in vivo* at single codon resolution (Ingolia et al., 2009). While in mammalian and yeast cells ribosome profiling demonstrated a robust anti-correlation between codon optimality and ribosome occupancy (Weinberg et al., 2016; Wu et al., 2019), for bacteria, and *E. coli* in particular, the method has been more problematic due to technical caveats of stopping translation and alignment of ribosome-protected mRNA fragments (Oh et al., 2011; Subramaniam et al., 2014). Recent improvements of sample preparation for the ribosome profiling in *E. coli*, including rapid cell harvesting and methods to arrest translation, confirmed that also in bacteria the ribosome density correlates with the codon-adaptation index, consistent with the expectation that rare codons are decoded by lower-abundance tRNAs more slowly than more abundant codons by their respective tRNAs (Mohammad et al., 2019), although the correlation is not as strong as in yeast cells (Weinberg et al., 2016).

Some aspects of the codon usage bias go beyond the simple distribution of rare and abundant codons. Bioinformatics analysis of the *E. coli* genome suggests that certain codon pairs are overrepresented (i.e., codon pairs observed more frequently than predicted) in the genes coding for non-abundant proteins and underrepresented in highly expressed genes (Boycheva et al., 2003; Guo et al., 2012; Moura et al., 2005; Tats et al., 2008). The order of codons in the codon pairs is crucial; for example, the codon pair AAG-UUA is overrepresented in highly expressed genes and UUA-AAG in poorly expressed genes (Boycheva et al., 2003). Although no experimental data are available on how codon combinations may regulate the elongation rate in bacteria, translation inhibition by specific codon pairs has been demonstrated in yeast

(Gamble et al., 2016). In yeast, translation of such inhibitory pairs is slower than expected for the sum of translation times for each codon individually; inhibition is abolished by replacing the sequence with synonymous codons or by reversing the codon order (Gamble et al., 2016; Tesina et al., 2020). The structure of the ribosome stalled on CGA–CCG and CGA–CGA inhibitory codon pairs shows that in both cases the mRNA conformation in the A site disturbs the aa-tRNA binding or accommodation, which causes the ribosome stalling (Tesina et al., 2020) .

Another interesting phenomenon is “codon clustering”, which is found in both pro- and eukaryotes. Rare codons often occur in clusters, rather than being randomly distributed along the mRNA (Chartier et al., 2012; Clarke and Clark, 2008; Makhoul and Trifonov, 2002). In bacteria, including *E. coli*, codon pairs formed by rare codons are overused, whereas pairs formed by common codons are underused, which is not the case in eukaryotes, where this tendency is reversed (Buchan et al., 2006). Finally, some reports suggest “co-occurrence” in eukaryotes of synonymous codons that use the same tRNA in close proximity on the mRNA coding sequence (Cannarrozzi et al., 2010; Shao et al., 2012), although the functional significance of such clustering is not understood. The complexity and strong species-specificity of overall codon distribution suggests that cells adapt their genetic programs to the needs for the protein production beyond amino acid selection.

### 2.2.1. mRNA structure

The nucleotide sequence of the mRNA determines the potential secondary structure, which may create obstacles when the ribosome moves along the mRNA during the translocation phase. The coding regions of mRNA generally appear more structured than non-coding ones, but the structures are dynamic (Del Campo et al., 2015). Ribosome pausing at mRNA structures is likely transient due to the intrinsic helicase activity of the ribosome that can unwind thermodynamically stable mRNA structures (Qu et al., 2011; Takyar et al., 2005). The translation efficiency of open reading frames within a polycistronic mRNAs negatively correlates with increasing levels of mRNA structure (Burkhardt et al., 2017). However, probing of RNA secondary structure with the so-called SHAPE (Selective 2′ Hydroxyl Acylation analyzed by Primer Extension) method in conjunction with parallel analysis of RNA structure (PARS) shows that only the mRNA structures at the initiation site are important and there is no significant correlation between mRNA structure in the coding region and translation efficiency (Del Campo et al., 2015; Mustoe et al., 2018).

### 2.2.2. aa-tRNA abundance

aa-tRNA is delivered to the A site of the ribosome in a ternary complex with EF-Tu and GTP. The rate of decoding is determined by the rate of selection of an aa-tRNA cognate to a given codon from the total pool of aa-tRNAs, followed by GTP hydrolysis by EF-Tu and accommodation of aa-tRNA in the PTC of the ribosome (reviewed in Rodnina, 2018). For every single codon, there is a different distribution of cognate, near-cognate and non-cognate aa-tRNAs, with their different decoding properties, such as the number and geometry of mismatches in the codon-anticodon complex, and aa-tRNA concentrations. The competition between the aa-tRNAs results in variations in decoding times for different codon–aa-tRNA pairs (Dykeman, 2020; Rudorf et al., 2014; Vieira et al., 2016), even though the rates of reactions on the decoding pathway, such as binding, GTP hydrolysis, and tRNA accommodation, are similar for different cognate aa-tRNAs (Ledoux and Uhlenbeck, 2008; Rodnina et al., 2005). Although in bacteria global codon usage matches the tRNA abundance, transient changes in the transcriptome composition due to transcriptional responses may shift mRNA codon bias relative to tRNA concentrations, which is predicted to have strong effects on decoding and may lead to additional, unexpected ribosome pauses (Dykeman, 2020; Rudorf, 2019). Under conditions of rapid growth, tRNAs are almost fully charged (Dittmar et al., 2005). Depletion of amino acids, e.g., during starvation, results in selective aa-tRNA charging levels for some tRNA isoacceptors, with some becoming low and others remaining high (Elf et al., 2003). Codons read by isoacceptors that retain high charging can be used for efficient translation of genes that are essential during amino acid starvation (Dittmar et al., 2005).

In this context, modifications at the tRNA anticodon region play an important role, and their absence can lead to ribosome stalling at the respective codons (Ranjan and Leidel, 2019). In particular, the positions 34 and 37 in the tRNA anticodon stem-loop (ASL) affect several elongation steps, including aa-tRNA selection and translocation. Modifications at both positions affect the structure and conformational dynamics of the ASL and provide chemical groups for the non-canonical interactions with the mRNA (Agris, 2008). The modifications at position 37 are mainly associated with reading frame maintenance (Agris et al., 2018; Hou et al., 2019). The N6-threonylcarbamoyladenosine-(t6A) modification, which is commonly found at position 37 of Arg-, Asn-, Ile-, Lys- Met-, and Ser-tRNAs, is essential for the viability of *E. coli* (Juhling et al., 2009; Thiaville et al., 2015); t6A-deficient yeast strains are viable, but have increased protein aggregation due to global mistranslation (Pollo-Oliveira et al., 2020). The modifications at position 34 are particularly important as modulators of wobble-position

decoding. For example, the 5-oxyacetic acid (cmo5) modification at U34 expands the decoding capacity of tRNA, allowing to decode all four synonymous codons for Ala, Ser, and Val (Agris et al., 2018) and expands the codon recognition by wobble interaction towards unusual base pairing (Agris et al., 2007; Grosjean and Westhof, 2016; Lim and Curran, 2001). Ser-tRNA(UGA), carrying the cmo5 modification at U34 can form stable complexes with all four codons, adopting specific conformations to stabilize each type of interaction (Weixlbaumer et al., 2007). Another cmo5-modified tRNA, Ala-tRNA(UGC), can read the non-Watson Crick synonymous codon GCC correctly, using a U:C base pairing at the 3rd codon position, albeit slower than codons with a conventional wobble base pairing (Kothe and Rodnina, 2007). The lack of the 5-methoxycarbonylmethyl-2-thiouridine (mcm5s2U34) modification in $tRNA^{Lys}$, $tRNA^{Glu}$ and $tRNA^{Gln}$ in yeast results in ribosome queuing at the respective codons, affects expression of a subset of genes enriched for AAA, CAA, and GAA codons and leads to protein aggregation (Nedialkova and Leidel, 2015; Ranjan and Rodnina, 2016; Rezgui et al., 2013; Zinshteyn and Gilbert, 2013). Although some of these specific examples are derived from yeast, it is likely that in *E. coli* the effect of these modifications is similar because the basic mechanisms of codon–anticodon recognition are universally conserved in pro- and eukaryotes. In summary, the decoding properties, concentrations, and chemical modifications of aa-tRNA play an important role during decoding and may contribute to translational pausing.

### 2.2.3. Amino acid composition of the nascent peptide

Experiments with model substrates show that the rate of peptide bond formation depends on the nature of the amino acids acting as peptide donors and acceptors (Wohlgemuth et al., 2008). However, for many aa-tRNAs, the chemistry step is rate-limited by the preceding step of aa-tRNA accommodation in the PTC on the large ribosomal subunit, which is similar for various aa-tRNAs (Wohlgemuth et al., 2008). One exception is the peptide bond formation involving consecutive Pro residues, which is much slower than with other amino acids. The rate of Pro codon translation is context-dependent with the time of pausing modulated by the preceding or following amino acid (Mohammad et al., 2019; Peil et al., 2013; Woolstenhulme et al., 2015). In bacteria, peptide bond formation between consecutive Pro residues is accelerated by EF-P (Doerfel et al., 2013; Ude et al., 2013), which helps to orient and stabilize the reactants in the PTC (Doerfel et al., 2015; Huter et al., 2017a). Nevertheless, even in the presence of EF-P, ribosomes pause on sequences presenting a combination of Pro, Gly and Asp codons (Mohammad et al., 2019). Ribosome pauses can also be caused by the interactions of

the newly synthesized nascent peptide with the walls of the polypeptide exit tunnel via distortion of the optimal geometry of the codon–anticodon pair at the PTC (Ito and Chiba, 2013; Wilson et al., 2016). This phenomenon is best illustrated by the so-called 'arrest peptides' such as SecM, MifM, VemP, ErmCL (Ito and Chiba, 2013). These sequences are about 20 amino acids long. They cause PTC distortion through the interaction with the polypeptide exit tunnel of the ribosome, which results in the reduced rate of the peptide bond formation and ribosome stalling when slowly reacting amino acids are in the A and P sites of the PTC (such as Pro and Gly) (Liutkute et al., 2020b; Mohammad et al., 2019; Seip et al., 2018; Wilson et al., 2016; Woolstenhulme et al., 2013).

The PTC also becomes distorted during translation of poly(A) sequences into stretches of poly-Lys both in *E. coli* and eukaryotes (Chandrasekaran et al., 2019; Koutmou et al., 2015). The ribosomes tend to stall and shift the reading frame when they encounter poly(A) sequences longer than 9 nt. The inhibition of translation, as shown for mammalian ribosomes, is due to conformational changes in both PTC and the decoding center (Chandrasekaran et al., 2019). Poly-Lys in the polypeptide exit tunnel stabilizes the PTC in a conformation that is inhibitory to peptide bond formation. In parallel, the poly(A) stretch of the mRNA adopts a single-stranded helix conformation in the decoding center, which is stabilized by the interactions with the rRNA. The reconfigured decoding center disfavors aa-tRNA binding to the A site, thereby hindering the elongation even further (Chandrasekaran et al., 2019; Tesina et al., 2020). Given the high degree of evolutionary conservation of the functional centers of the ribosome and the consistent stalling of *E. coli* ribosomes by poly-Lys nascent peptide, the same mechanism is likely to operate in bacteria as well. In addition to the specialized stalling sequences, also shorter patches of amino acids may have an impact on the translation efficiency. Statistical analysis of ~ 6000 genes from different organisms expressed in *E. coli* showed that the amino acid identity has a significant impact on protein expression. Bioinformatics analysis of *E. coli* proteome revealed that several amino acid triplets (for example Cys-Met-Tyr, Met-Trp-Cys, Gly-Pro-Pro and Trp-Met-Cys) and thousands of quadruplets (for example Cys-Met-Tyr-Trp), are either completely absent or are several-fold less abundant than expected for random sequences (Navon et al., 2016). Single-molecule Förster resonance energy transfer (FRET) experiments suggested that the presence of such sequences in the mRNA increases the elongation time (Navon et al., 2016). This indicates that ribosomes tend to stall while translating these codon combinations, but the exact reason for stalling is not yet known.

The amino acid charge distribution along the nascent chain may affect translation elongation as well (Lu and Deutsch, 2008; Tuller et al., 2011). Positively charged amino acids upstream of the A-site codon slow down the ribosome irrespective of the codon identity or codon distribution (Charneski and Hurst, 2013; Lu and Deutsch, 2008; Sabi and Tuller, 2015), and there is a linear correlation between the extent of ribosome pausing and the length of positively charged segments (Charneski and Hurst, 2013). It is possible that the positive charge exerts a force on the PTC that increases the distance between P-and A-site tRNA and hence increases the barrier for peptide bond formation (Leininger et al., 2021). The effect of charged amino acids on translation also depends on its location in the nascent peptide. Codons encoding for positively and negatively charged amino acids are non-randomly distributed along the mRNA with higher frequency of positive charges at the beginning and negative charges at the end (Duc and Song, 2018). This observation led to the suggestion that the incorporation of positively charged amino acids at the beginning of elongation helps the N-terminus of the nascent polypeptide to move towards and enter the negatively charged part of the exit tunnel (Duc and Song, 2018). Once the N-terminus passed the exit tunnel, the mean elongation rate increases with the hydrophobicity of the nascent chain (Duc and Song, 2018). While some of these studies have been conducted using yeast as a model organism (Charneski and Hurst, 2013; Duc and Song, 2018; Leininger et al., 2021; Sabi and Tuller, 2015), the tendencies observed in yeast may be common for bacteria as well because residues lining the exit tunnel are highly conserved in the zone proximal to the PTC and diverge only around the vestibule zone (Duc et al., 2019).

Some combinations of amino acids in the nascent peptide may cause even more dramatic effects, for example destabilize ribosomes translating the respective mRNA. *In vitro* translation studies showed that the ribosome becomes destabilized during the synthesis of mainly negatively charged segments, in particular when preceded by positively charged segments (Chadani et al., 2017). This leads to the abortion of translation, a process named intrinsic ribosome destabilization (IRD). The synthesis of a model protein containing 10 consecutive glutamine, but also the sigma-70 subunit of the RNA polymerase RpoD that contains a DDDEDEDEED sequence, is counterbalanced *in vivo* by the ribosomal protein bL31. Stronger IRD sequences that occur naturally are usually constituted of alternating Pro and negatively charged residues. These peptide sequences are suggested to have regulatory impact on the translation, but also transcription of mRNA during transcription-translation coupling. For example the *E. coli* magnesium transporter MgtA is expressed only when the preceding IRD

sequence induces translation abortion under conditions where the magnesium concentration is too low to stabilize the ribosome (Chadani et al., 2017). It was shown that the abortion is greatly reduced if the IRD sequence is located in the middle of the coding region and further away from the N-terminus. In case that the IRD sequence is preceded by up to 30 aa, the property and identity of the aa had no large impact. The translation abortion by IRD sequences that have short preceding segments (5 aa) is reduced with large, bulky amino acids, whereas with small amino acids it is more pronounced (Chadani et al., 2021).

Mechanistically, the effects of interactions of the nascent peptide with the polypeptide exit tunnel are still poorly understood. Nascent peptide interactions with the ribosome may provide a new regulatory mechanism for sequence-specific modulation of translation speed.

### 2.3. Interplay between translation rate and protein stability and activity

The local translation rate determines the timing of peptide emergence from the ribosome and regulates folding of the nascent chain. In particular slowly translating mRNA segments seem to have an important impact on the activity and stability of the nascent protein.

Analysis of how the local translation rate affects folding on the *E. coli* protein chloramphenicol acetyltransferase suggests that proteins variants expressed with perturbed local rates are less stable in solution (Walsh et al., 2020). Similarly, the *E. coli* protein Sufl contains several clustered codons that are decoded by lowly abundant tRNA with a slower translation rate. Substituting these codons with synonymous codons encoded by highly abundant tRNA or directly supplementing rare tRNA increases the translation rate, resulting in protein that is more susceptible to protease digestion on the ribosomes. Control experiments that reduce the translation rate globally by lowering the temperature can restore the stability of the destabilized Sufl variants, clearly indicating that slow translating regions are necessary for protein stability (Zhang et al., 2009). Further examples emphasize the detrimental effect of synonymous codon substitution by more common codons on the solubility caused by misfolding (Cortazzo et al., 2002) or temperature-dependent loss of activity (Xu et al., 2013). The *Neurospora* circadian clock protein FRQ losses its rhythmicity upon codon optimization. Although the protein was more abundant after codon optimization, it acquired an altered conformation leading to the loss of function (Zhou et al., 2013). Overexpressing human single-chain antibodies in *E. coli* with a variety of different synonymous codons results in proteins that differ in solubility and epitope binding (Hu et al., 2013). In general, globally slower translation seems to be beneficial for protein synthesis. Expression of eukaryotic proteins is inherently poor in *E. coli*. However, by

using mutant ribosomes in *E. coli* that are characterized by globally slow translation, the expression of several eukaryotic enzymes was restored without compromising the native *E. coli* proteome (Siller et al., 2010). Likewise, insertion of rare codons into a fluorescent reporter provides more time to fold and form the active fluorescent complex (Sander et al., 2014). Alternatively, restoring the local translation pattern improves protein expression. Structural analysis of the bovine protein gamma-B-crystallin indicate that upon expression in *E. coli* the protein has a reduced stability and solubility. Harmonization of the codon usage, i.e., introducing synonymous codon substitutions to reinstate the mammalian translation pattern, improves the yield and the stability of the protein. The authors suggested that slow translating segments at critical protein junctions caused this improvement (Buhr et al., 2016).

There are also several examples that indicate that fast translating regions are important for the function of membrane proteins. A variety of synonymous codon mutations in the human multidrug receptor 1, substituting common with rare codons, changed the activity of the receptor in regard to affinity and specificity (Kimchi-Sarfaty et al., 2007). Furthermore, the cystic fibrosis transmembrane conductance regulator (CFTR), carrying a single nucleotide exchange positioned close to the membrane-spanning domain 2, acquired a lower activity for chloride transmission than the wt. The point mutation changes the threonine codon from one that is decoded by highly abundant tRNA to a codon that is decoded by lowly abundant tRNA. The phenotype could be rescued by additionally providing the decoding tRNA during the expression (Kirchner et al., 2017). It is possible that rapidly translating regions are important for the topogenesis of membrane proteins. Overall, the topology of a membrane protein is driven by charges at the N-terminus, for example by positive charges directing the N-terminus towards the cytoplasmic side of the membrane (von Heijne, 1989). However, there are indications that additional or prolonged pauses at critical junctions could change the topology (Mercier et al., 2020). The folding of the first nucleotide-binding domain of CFTR was not affected by slower translation; however, accelerating translation at critical segments leads to increased aggregation of the protein (Kim et al., 2015). It needs to be seen whether this is caused by the fact that the domain itself is cytosolic and therefore follows the tendency of the above presented examples of cytosolic proteins.

## 2.4. Protein folding

The process of protein folding describes the pathway that leads to the formation of the final 3D structure of a polypeptide. Historically, protein folding was studied in solution, i.e., using purified proteins that were unfolded in denaturant, pH or temperature and refolded in their native condition, in order to determine the number of formed intermediates, their respective energy levels and the overall rate of folding and unfolding (reviewed in Bhatia and Udgaonkar, 2022; Dill and MacCallum, 2012; Englander and Mayne, 2014). The energy level of each step, and mainly the free energy difference between the unfolded and the native state, was thought to be the main driving force of protein folding, which is widely following the thermodynamic hypothesis by Anfinsen (Anfinsen, 1973). It is worth mentioning that mostly relatively small, globular, cytosolic proteins with approximately 100 aa in length were analyzed in such experiments, because instead of refolding, large multidomain proteins tend to aggregate or misfold (reviewed in Braselmann et al., 2013). Attempts to generalize refolding in solution show that the protein complexity correlates very well with the folding rates in native condition, i.e., less complex proteins with predominantly local rather than long-range contacts fold very rapidly (Broom et al., 2015; Plaxco et al., 2000; Wang and Panagiotou, 2022). However, other models suggest that the natively folded state of a protein may not represent the global free energy minimum, but may rather be trapped in a local minimum with the free energy that may even be positive, rendering the native state unfavorable. As a result, the protein stability is limited over time, but ensured kinetically by the energy barrier established by cotranslational folding of the protein on the ribosome. This may be an active, energy-demanding process assisted by the ribosome, which may act as a chaperone (reviewed in Sorokina et al., 2022). A recent study where the entire *E. coli* proteome was unfolded in denaturant and then refolded in native conditions indicates that up to 30% of all proteins in *E. coli* are non-refoldable and even in the presence of chaperones 15% remain non-refoldable (To et al., 2021; To et al., 2022). These non-refoldable proteins have to follow the cotranslational folding pathway in order to acquire the native state.

### 2.4.1. Cotranslational protein folding inside the ribosome

Folding during ongoing translation depends on the position of the nascent peptide in the polypeptide exit tunnel. Secondary structures and small tertiary structures can form already inside the ribosome. An early FRET and crosslinking study showed that transmembrane helices can form deep inside the polypeptide exit tunnel (Woolhead et al., 2004). Furthermore, the

folding seems to be induced by the ribosome itself, as the helix is not stable in solution (Woolhead et al., 2004). Also the helix formation of consecutive Ala inside the polypeptide exit tunnel depends on their location within. The Ala segments do not fold with the same probability at every position inside the tunnel, but the propensity to fold is strongest close to the PTC and at the part of the tunnel that is closer to the outside. The authors suggested that the nascent chain may need to unfold and refold in order to pass through the narrow section of the constriction site, explaining the low fraction of formed helix at this area (Lu and Deutsch, 2005). The cryo-EM structures of the dipeptidylaminopeptidase B on eukaryotic ribosomes show α-helices can indeed form inside the polypeptide exit tunnel with a strong preference for folding after the constriction site (Bhushan et al., 2010). Both the VemP stalling peptide (Su et al., 2017) and the topoisomerase subunit of the T4 bacteriophage DNA polymerase (Agirrezabala et al., 2017) were resolved with two α-helices inside the polypeptide exit tunnel of *E. coli* ribosomes, one before and one after the constriction site (Su et al., 2017). VemP-mediated translational arrest is induced by the helix that is located at the PTC by inducing a translation-incompetent conformation of the ribosome (Su et al., 2017).

Crosslinking studies indicate that also β-hairpins and small tertiary structures, such as α-helical hairpins, can fold early during translation, as the folding occurs close to the vestibule of the polypeptide exit tunnel (Kosolapov and Deutsch, 2009). For the folding of an α-helical hairpin of the T1 domain of the human voltage-gated potassium channel Kv1.3, the presence of the N-terminal β-hairpin seems to be important for the correct folding and stability because N-terminal truncated versions did not fold as efficiently as in the presence of the β-hairpin. The N-terminal β-hairpin folded more robustly and independently of the C-terminal α-helical hairpin (Kosolapov and Deutsch, 2009). Interestingly, the formation of a β-hairpin inside the polypeptide exit tunnel was also observed by cryo-EM with the *E. coli* cold-shock protein A (CspA), which acquires a fully β-stranded structure in native conditions (Agirrezabala et al., 2022). Several cryo-EM structures clearly show that CspA forms an α-helix shortly after the constriction site, which then starts to rearrange into a β-hairpin close to the outside of the polypeptide exit tunnel at a length of roughly 70 aa (Agirrezabala et al., 2022). Even small tertiary structures can fold inside the polypeptide exit tunnel, as was shown for the small Zinc-finger domain ADR1a (29 aa), which folds roughly 60 Å away from the PTC, but only in the presence of $Zn^{2+}$ (Nilsson et al., 2015; Wruck et al., 2021). In general, there is a tendency that small proteins and domains begin to fold close to the end of the exit tunnel and finalize their native fold after they fully emerged from the exit tunnel (Farias-Rico et al., 2018), or after the

release from the ribosome (Agirrezabala et al., 2022). For larger proteins the folding towards the native state progresses cotranslationally after the nascent peptide reaches the outside of the ribosome.

### 2.4.2. Cotranslational protein folding outside the ribosome

At the vestibule and outside of the ribosome, proteins can start compacting step-by-step, with increasing compaction as more of the nascent chain is exposed outside the polypeptide exit tunnel. Folding of the small α-helical N-terminal domain (NTD) of HemK occurs in such a stepwise manner, as indicated by the analysis of several high FRET intermediates throughout the translation of the mRNA (Holtkamp et al., 2015; Liutkute et al., 2020a; Mercier and Rodnina, 2018). With progressing translation, proteins can also acquire their native structure on the ribosome as shown for the immunoglobulin-like domain of filamin FLN5 in an NMR study (Ahn et al., 2022; Cabrita et al., 2016; Chan et al., 2022). Folding of this domain is delayed until an additional peptide sequence is synthesized that fills the polypeptide exit tunnel and allows the domain to emerge and move away from the ribosome. Only when the domain moves away from the ribosome upon synthesis of the subsequent domain FLN6, FLN5 acquires its native state. FLN6 itself is not folded, but compacts outside of the ribosome while interacting with the ribosome itself. FLN5 is supposed to form this so-called ‘compact disordered state’ as well (Cabrita et al., 2016).

In case of the multidomain protein EF-G, the N-terminal G domain adopts a molten globule structure on the ribosome, which prevents its misfolding together with the following domain II (Liu et al., 2019). The chaperone trigger factor (TF) can sequester parts of the domain II nascent chain, thereby preventing the unfolding of the G domain (Liu et al., 2019). It is possible that also the interaction with the ribosome is important to stabilize folding intermediates. FLN5 has different populations of distinct intermediate states depending on its exposure from the polypeptide exit tunnel, as the ribosome destabilizes the native state, but stabilizes the intermediates (Chan et al., 2022). This destabilization is also shown for HemK NTD when it is positioned close to the ribosome surface (Liutkute et al., 2020a). It is the interaction of the nascent chain with the ribosomal proteins L23 and L24 at the vestibule of the polypeptide exit tunnel that can destabilize peptides and delay the folding (Ahn et al., 2022; Kudva et al., 2018) and is evident by slower folding rates and faster unfolding rates of the nascent protein compared to natively folded protein in solution as was shown for the N-terminal G domain of EF-G by optical tweezer experiments (Liu et al., 2019; Liu et al., 2017). A follow-up study revealed that

the N-terminal G domain of EF-G stays unfolded and forms intermediates and the native structure only after the entire domain has emerged from the ribosome. Based on these observations the authors suggested that the folding of this domain occurs strictly from the C- to the N-terminus (Chen et al., 2020).

Domain-wise folding of proteins on the ribosome is also supported by an NMR study of an SH3 domain (Eichmann et al., 2010; Guinn et al., 2018). SH3 stays mainly unfolded until the entire sequence has emerged from the polypeptide exit tunnel, but unlike EF-G, this appears to happen largely independent of the ribosome, without any interactions that prevent folding or induce the formation of secondary structures (Eichmann et al., 2010). This was confirmed by a later study showing that the folding pathways of this protein on and off the ribosome is identical, following a two-state folding pathway after the SH3 domain is exposed with a GS linker further outside the ribosome (Guinn et al., 2018). Another example of a protein that folds to the native state on the ribosome is FRB (Han et al., 2012). The proteins FRB and FKBP can only interact with each other after they are natively folded and use rapamycin as a bridging agent. Pull-down experiments linked with RiboSeq revealed that FRB is folded as it emerges from mammalian ribosomes because of it pronounced interaction with FKBP and rapamycin (Han et al., 2012). In summary, these examples underscore the complexity of protein folding on the ribosome and emphasize the necessity for additional studies to identify general rules of how proteins fold during translation.

### 2.5. HemK

HemK is a structurally conserved protein that is found in many different bacteria and mitochondria (Yang et al., 2004). It catalyzes the methylation of RF1 and RF2 at the Gln residue of the universally conserved GGQ motif. The methyl group is donated by the cofactor S-adenosyl-methionine (SAM), which subsequently reacts to S-adenosyl-homocysteine (SAH) (Heurgue-Hamard et al., 2002; Nakahigashi et al., 2002). HemK in *E. coli* is not essential, but deletion or mutation causes severe growth effect due to aberrant peptide release from the ribosome after translation, which increases non-stop readthrough of UAG and UGA stop codons and subsequent increase of tmRNA mediated rescue (for tmRNA rescue see chapter 2.1.4) (Nakahigashi et al., 2002).

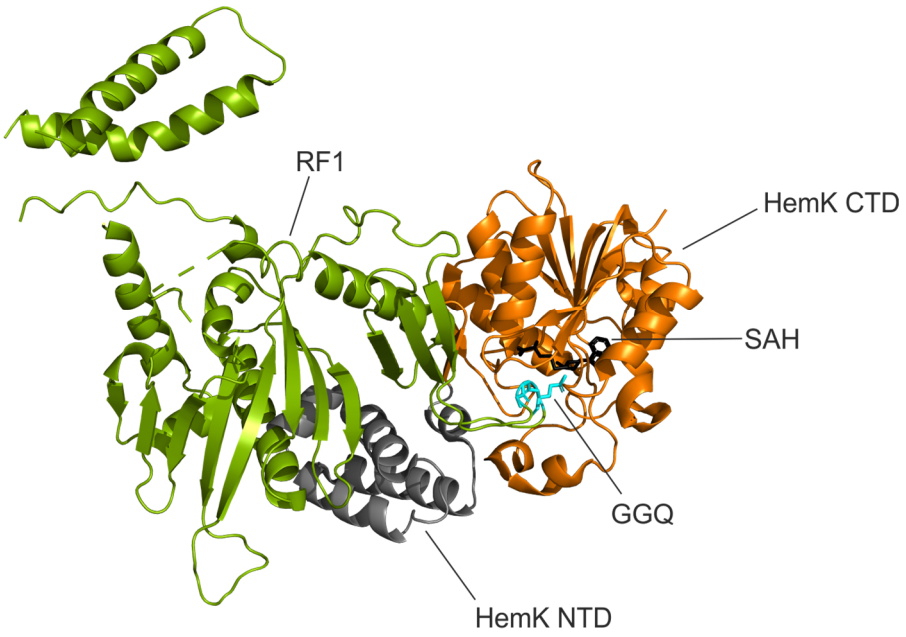



Figure 4. HemK in complex with RF1 (PDB: 2B3T (Graille et al., 2005)). RF1 (green) with the GGQ motif (cyan) pointing towards the catalytic center of HemK CTD (orange). The complex is trapped in the non-reactive state by using the S-adenosyl-homocysteine (SAH; black) as a cofactor. The HemK NTD is shown in (grey).

HemK is a two-domain protein. The NTD binds the release factor and together with the C-terminal domain (CTD) positions the RF1 domain that contains the GGQ motif in the catalytic center of HemK, exposing it to SAM (an unreactive analog SAH in the crystal structure; Figure 4). HemK NTD comprises a 5-helix bundle, whereas the CTD is a Rossmann fold (alternating α-helices and β-strands with a central β-sheet) (Graille et al., 2005; Yang et al., 2004). The folding of HemK NTD was extensively studied both on and off the ribosome. In solution, HemK NTD folds rapidly into its native structure in a two-state fashion and the folding occurs independently of the HemK CTD (Holtkamp et al., 2015). On the ribosome, HemK NTD starts folding already close to the vestibule of the ribosome with a stepwise compaction through the formation of intermediates (Holtkamp et al., 2015; Liutkute et al., 2020a; Mercier and Rodnina, 2018). A native-like intermediate state is reached after the domain has fully emerged from the ribosome, which remains destabilized by the ribosome as long as it is in close proximity to it (Liutkute et al., 2020a). The precise folding pathway of the CTD of HemK remains unknown, but is of high interest because of its complex α/β fold.

## 2.6. Scope of the thesis

During protein synthesis on the ribosome, the nascent chain can start folding inside the polypeptide exit tunnel. With ongoing translation, the nascent chain can adopt α-helical conformations and undergo structural rearrangements. The protein can compact stepwise towards the native state upon emerging from the ribosome or stay unfolded until the entire domain folds in a concerted fashion. Although the folding pathway of some proteins is known, there is still too little information to find general characteristics of protein folding. In order to shed light on the heterogeneity of folding pathways, larger and more complex domain structures have to be analyzed in detail. Here we analyze the cotranslational folding pathway of HemK CTD, an α/β-type protein with a Rossmann fold. The alternating appearance of α-helices and β-strands and the potential for forming the central parallel β-sheet in a stepwise manner makes HemK CTD an attractive model for studying cotranslational protein folding. An *E. coli in vitro* translation system was used in combination with PET and FRET techniques to analyze the protein folding of the nascent peptide. PET was measured between an N-terminal reporter and a Trp at the N-terminus in order to observe structural rearrangements of the peptide inside the polypeptide exit tunnel of the ribosome. FRET was measured between two fluorophores in the nascent chain that were incorporated site-specifically at different positions in the CTD sequence in order to report the folding of different parts of the protein. Combining these techniques on both translating and on stalled ribosome complexes provides insights into the folding pathway of HemK CTD on the ribosome.

# 3. Results

## 3.1. Nomenclature

During the progress of this study, many different mRNA constructs were translated with truncations at both the N- and the C-terminus. Additionally, point mutations were introduced at various positions of the protein. The positions of these point mutation matter for the analysis of the data. In the following, we use the following nomenclature.

Table 1. Nomenclature.

| **Term used** | **Definition** |
|---|---|
| HemK | Full-length 277 aa-long wild-type protein sequence, including NTD and CTD |
| HemK NTD | Residue 1-73 of HemK |
| HemK CTD | Residue 73-277 of HemK with a total length of 205 aa. The residue 73 of HemK was mutated to the start codon AUG |
| Residue or residue number (residue no.) | Residue or residue number are written as a single value (for example '78') or range (73-152). The number always corresponds to the aa number in the HemK sequence. For example, residue 78 in HemK CTD (73-277) is NOT the 78$^{th}$ aa of this specific construct, but represents the 6$^{th}$ amino acid in HemK CTD. This means that residue 78 in HemK and HemK CTD refers to exactly the same aa. |
| HemK 73-152 (as an example) | C-terminal truncation of HemK CTD including residue 73-152 of HemK, yielding an 80 aa translation product |
| Point mutation | All aa of the point mutations in either HemK or HemK CTD are labeled with the residue number (see above).<br>Example:<br>HemK T120K<br>HemK CTD T120K |

## 3.2. Experimental approach to study cotranslational folding of HemK CTD

### 3.2.1. Establishing *in vitro* translation system

To investigate the folding pathway of the HemK CTD, efficient *in vitro* translation was established using a fully reconstituted *E. coli* translation system (Holtkamp et al., 2015). The *in vitro* translation system allows to adapt its composition to the experimental needs. Additionally, it is important to make the protein of interest visible in the crowded environment of elongation factors, ribosomal proteins, rRNA and tRNA. In order to observe the protein of interest during translation and folding, fluorophores were introduced into the nascent peptide.

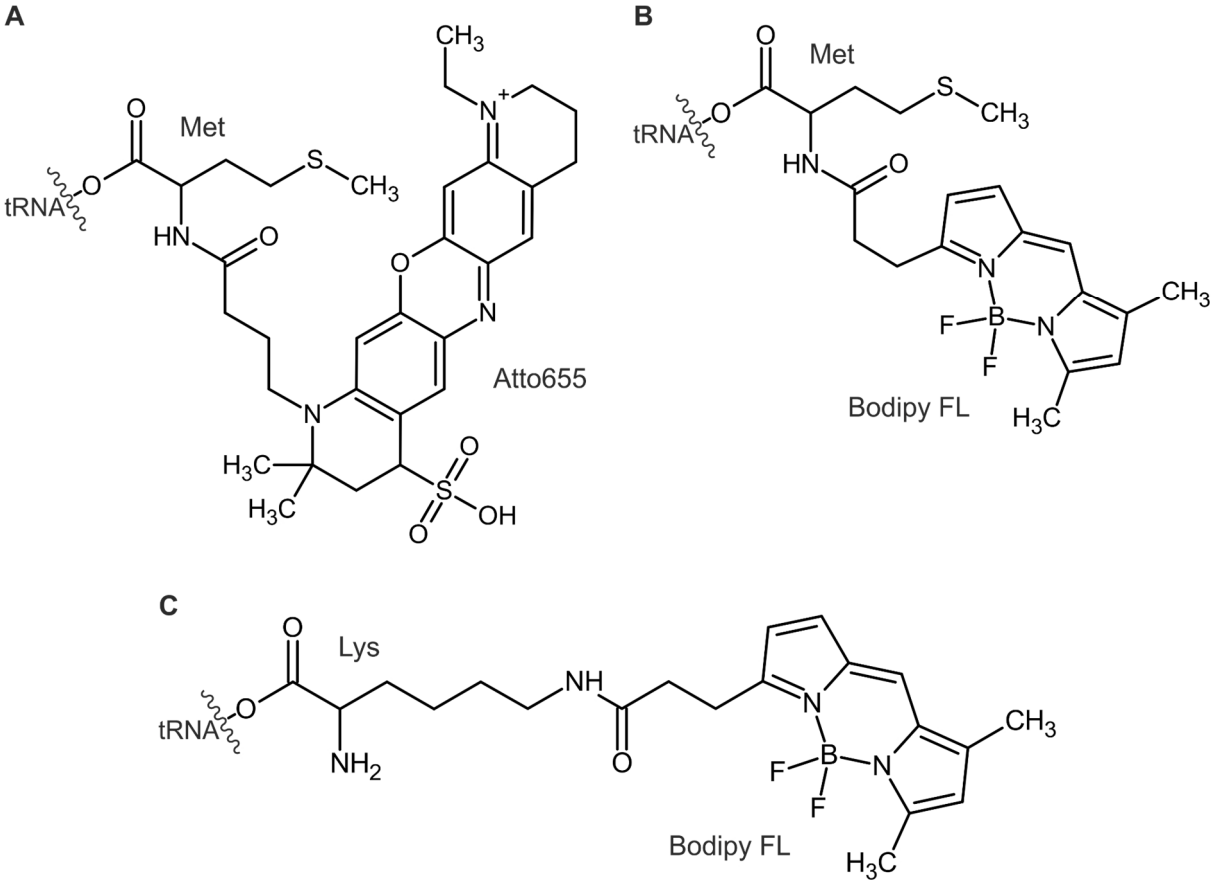



Figure 5. Schematics of the fluorescence reporters used for aa-tRNA labeling. Initiator Met-tRNA$^{fMet}$ is labeled with either (A) Atto655 or (B) Bodipy FL (BOF) at the α-amine of Met, introducing a fluorescent reporter instead of the formyl group. (C) Lys-tRNA$^{Lys}$ is labeled with BOF at the ε-amine of Lys.

The nascent peptide can be labeled at the N-terminal Met with Atto655 (Figure 5A) or Bodipy FL (BOF) (Figure 5B). The fluorescence-labeled Met is delivered to the protein by the initiator BOF/Atto655-Met-tRNA$^{fMet}$ (Mittelstaet et al., 2013).

Additionally, BOF-Lys-tRNA$^{Lys}$ (Figure 5C) was used in this study to introduce BOF-labeled Lys (BOF-Lys) site-specifically into the polypeptide chain at the position of the Lys codon. To achieve efficient mRNA translation, the elongation factors and aa-tRNA need to be supplied in sufficient amounts to avoid the accumulation of truncated protein products due to ribosome stalling. Additionally, we were aiming to adjust the translation system for efficient cotranslational incorporation of BOF-Lys. Therefore, as a first step, the amounts of the elongation factors EF-Tu and EF-Ts and aa-tRNA were optimized to achieve maximum translation of HemK mRNA. Second, several positions along the peptide were screened for the efficient incorporation of BOF-Lys. And third, different concentrations of BOF-Lys-tRNA$^{Lys}$ were tested for maximum incorporation efficiency.

#### 3.2.1.1. Titration of EF-Tu and aa-tRNA

To optimize the translation of HemK CTD mRNA, we used purified EF-Tu and aa-tRNA to optimize the peptide yield in the fully reconstituted *E. coli* translation system (chapter 5.8). In general, to have efficient incorporation of fluorescence-labeled Lys, a higher concentration of EF-Tu than aa-tRNA needs to be present (Mittelstaet et al., 2013). This ensures that all of the aa-tRNAs are bound as EF-Tu–GTP–aa-tRNA ternary complex. This prevents the competition for EF-Tu binding between the total aa-tRNA and the fluorescence-labeled tRNA species (Mittelstaet et al., 2013). The translation products were separated by SDS-PAGE and visualized by the N-terminal fluorescent reporter.

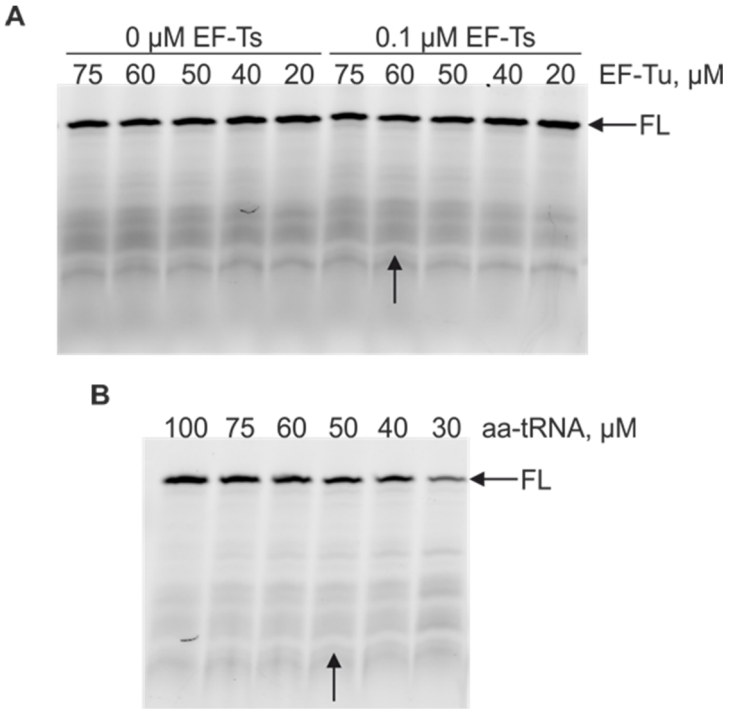



Figure 6. Optimization of HemK CTD *in vitro* synthesis. SDS-PAGE of HemK CTD translation products visualized by fluorescence of N-terminal BOF. (A) Titration of EF-Tu and EF-Ts in the presence of saturating concentrations of total aa-tRNA (100 µM). (B) Testing the effect of aa-tRNA concentration at fixed EF-Tu and EF-Ts condition (arrow in A). The aa-tRNA concentration taken for the future experiments is indicated with an arrow. FL: full-length product.

mRNA translation results in a major top band corresponding to the full-length protein (FL) and minor amounts of smaller peptide fragments due to ribosome stalling (Figure 6). The titration with EF-Tu shows very similar yields in full-length and stalled product formation (below FL in Figure 6) with all tested concentrations. Addition of EF-Ts does not influence the peptide pattern. The condition with 60 µM EF-Tu and 0.1 µM EF-Ts was chosen (arrow in Figure 6A) to test different aa-tRNA concentrations because 60 µM is a sufficiently high concentration to test a wide range of aa-tRNA concentrations. EF-Ts was kept in the translation mixture to ensure full activation of EF-Tu during the formation of the ternary complex for later experiments, i.e., the incorporation of fluorescence-labeled Lys-tRNA.

Contrarily to the EF-Tu titration, the reduction of aa-tRNA has an immediate effect on the translation efficiency, resulting in a lower amount of the full-length product with lower aa-tRNA concentration (Figure 6B). The pattern of the distinct stalling products is unchanged upon aa-tRNA reduction and only the lowest aa-tRNA concentration causes pronounced accumulation of truncated peptides. The aa-tRNA concentration of 50 µM (indicated by the

arrow in Figure 6B) was chosen for future experiments, as it is sufficiently high to ensure effective full-length product formation, and also provides an excess of EF-Tu to ensure the efficient formation of the ternary complex with BOF-Lys-tRNA$^{Lys}$.

#### 3.2.1.2. Screening for BOF-Lys incorporation sites in the polypeptide chain

In order to study the compaction of the nascent chain on the ribosome, we introduced two fluorophores site-specifically into the nascent peptide and used FRET as a readout. The first dye Atto655, was delivered by initiator Atto655-Met-tRNA$^{fMet}$. The second BOF dye was delivered by BOF-Lys-tRNA$^{Lys}$. The first aim of using this FRET system was to observe the folding of the N-terminal sequence of HemK CTD. There is one Lys in the wild-type HemK CTD sequence at residue 159, but this position seemed too far away from the N-terminus of the CTD to report the initial steps of nascent chain folding, as it would report only on later folding events. Hence, the natural Lys159 codon was mutated to Arg and the Lys codon was introduced at different positions closer to the N-terminus of the CTD. It was previously shown that the incorporation efficiency for unnatural fluorescence-labeled amino acids is highly dependent on the preceding amino acid in the protein sequence. Small aa like Gly and Ala preceding the incorporation site facilitate the accommodation of labeled aa (Thommen et al., 2022). Therefore, three positions with a preceding Gly (residue no. 118, 120, 122) and two with preceding Ala (residue no. 127, 129) were selected to be mutated to the Lys codon AAA and tested for BOF-Lys incorporation.

The efficiency of BOF-Lys incorporation at these positions were examined with the previously established translation condition (chapter 3.2.1.1) in the presence of BOF-Lys-tRNA$^{Lys}$ and unlabeled fMet-tRNA$^{fMet}$. The translation products were analyzed on SDS-PAGE.

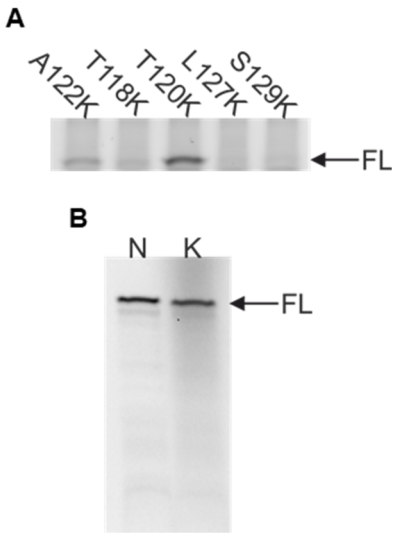



Figure 7. Screening for optimal BOF-Lys incorporation sites. SDS-PAGE of the translation product of (A) HemK CTD K159R variants in the presence of BOF-Lys-tRNA$^{Lys}$ (0.4 µM) and (B) HemK CTD K159R T120K with N-terminal BOF (N) and BOF-Lys (K). FL: full-length product.

BOF-Lys is found in the full-length protein when incorporated at residue T120K and to a smaller extent at A122K (Figure 7A). There is no additional accumulation of shorter truncation products upon addition of BOF-Lys (K in Figure 7B) instead of N-terminal BOF (N in Figure 7B). This implies that the incorporation of the fluorescent dye at the chosen position in the peptide does not lead to ribosome stalling or product dissociation.

#### 3.2.1.3. Titration of BOF-Lys-tRNA$^{Lys}$ for efficient incorporation

In the next step, BOF-Lys-tRNA$^{Lys}$ was titrated to achieve the maximum incorporation efficiency. Translation was performed with Atto655-Met-tRNA$^{fMet}$ to label the nascent peptide at its N-terminus in addition to the internal BOF label. The N-terminal Atto655 was used to assess the total translation efficiency and the internal BOF signal to estimate the incorporation efficiency of BOF-Lys. Both dyes were visualized on SDS-PAGE.

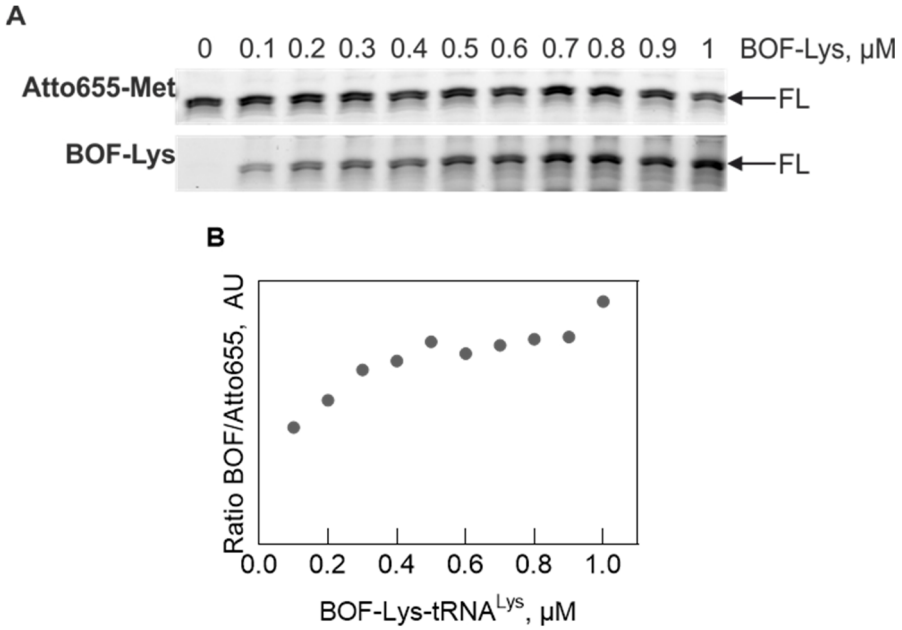



Figure 8. BOF-Lys-tRNA$^{Lys}$ titration. (A) SDS-PAGE of the translation product of HemK CTD K159R T120K in the presence of increasing concentrations of BOF-Lys-tRNA$^{Lys}$ visualized by the N-terminal Atto655 and internal BOF fluorescence. FL: full-length product. (B) Quantification of the intensity of BOF-Lys signal in FL by normalization by the corresponding Atto655 intensity. Ratio is reported in arbitrary unit (AU).

The internal BOF signal of the full-length product (FL with BOF-Lys in Figure 8A) was normalized by the corresponding N-terminal Atto655 signal of the full-length product of the same lane (FL of Atto655-Met in Figure 8A) to account for loading errors. The normalized BOF signal indicates that the incorporation is saturated at 0.5 µM BOF-Lys-tRNA$^{Lys}$ (Figure 8B). The higher ratio at 1 µM is considered an outlier. Therefore, whenever incorporation of BOF-Lys is indicated, BOF-Lys-tRNA$^{Lys}$ was used at a concentration of 0.5 µM.

### 3.2.2. Testing effect of N-terminal dye on translation rate

In order to study the folding of the CTD by using FRET between the N-terminal Atto655 and the internal BOF during ongoing translation, it was crucial to test that the incorporation of fluorophores in the polypeptide chain does not change the rate of mRNA translation. To test the effect of different dyes on the translation kinetics, we performed experiments with both Atto655-Met- and BOF-Met-tRNA$^{fMet}$. The translation reaction was stopped at the indicated time points and the products visualized by SDS-PAGE. Data processing and normalization was performed according to procedures described in chapter 5.8.3.1.

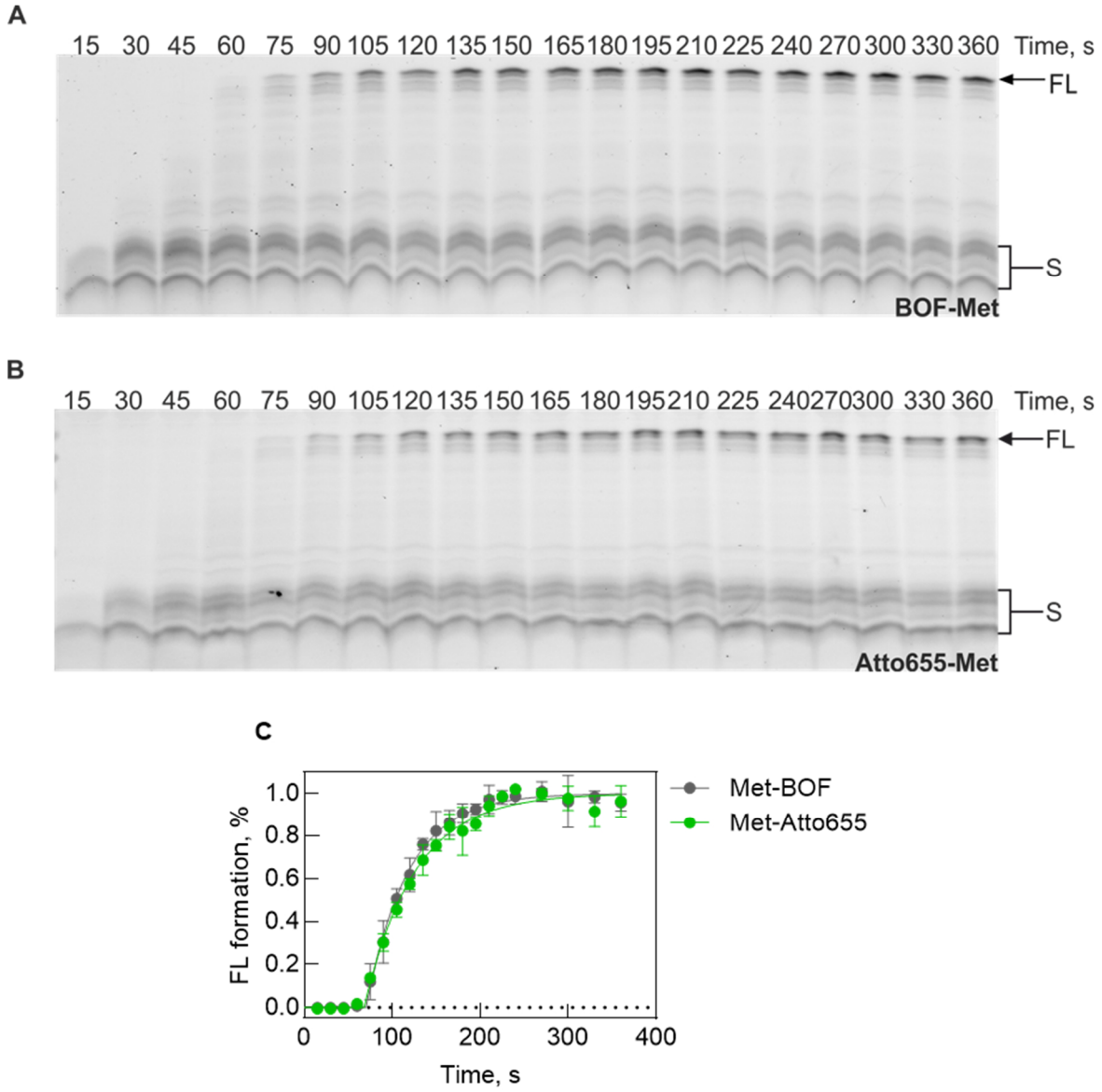



Figure 9. Effect of different N-terminal dyes on the translation kinetics. Representative SDS-PAGE of translation time courses for HemK CTD wt translation products with N-terminal BOF (A) and Atto655 (B). All time courses were performed in triplicates. (C) Quantification of the full-length product as mean ± SD (N=3). Time courses were normalized to 1 according to the amplitude of a delay-exponential function as described in 5.8.3.1. FL: full-length product; S: stalling products.

Over time, mRNA translation leads to the accumulation of a main high-molecular weight translation product corresponding to the full-length protein (FL) that appears when mRNA is translated to the end of the coding sequence and a small amount of lower molecular weight (S) products that appear at some point during the translation due to ribosome stalling (Figure 9A and B). Analysis of the translation time courses indicates that with either dye the full-length product appears after 75 s (FL in Figure 9A, B). Quantification of the full-length product bands

at different time points shows that the translation time courses are identical (Figure 9C), and hence incorporation of different dyes at the N-terminus does not change the rate of translation.

### 3.2.3. Establishing FRET observables to monitor cotranslational protein folding

FRET is a powerful tool to measure any kind of interaction between proteins, RNA, etc. It is well known as a molecular ruler in the distance range of 10-100 Å. Because the FRET efficiency is strongly distance-dependent, it often can be reliably utilized to report changes in distances. However, in many cases, including the experiments performed here, the determination of precise distances is challenging. First, the FRET efficiency depends on the orientation and alignment of the fluorophore dipoles, which depends on the nature of the dye and its mobility when attached to the macromolecule. The orientation of the dyes relative to each other is expressed by the orientation factor $\kappa^2$ (Lakowicz, 2006). Usually, the orientation factor $\kappa^2$ is assumed to be 2/3 and describes two dyes that can freely rotate (Lakowicz, 2006). In the complex environment of a peptide, and depending on the position in the peptide, the mobility of the dye may be considerably more restricted (Buskiewicz et al., 2005). Second, in our case, it is difficult to estimate the FRET efficiency by using the decrease of donor fluorescence compared to a donor-only control, as excess amounts of donor-labeled Lys-tRNA used for translation generate a high fluorescence background that renders this approach very error-prone. Third, most approaches that calculate the FRET efficiency, also from acceptor fluorescence, rely on a labeling efficiency of 100% or that the actual labeling efficiency is taken into account during the calculation (reviewed in Wu and Brand, 1994). For this project, the labeling efficiency and more importantly the cotranslational incorporation efficiency of both the donor and the acceptor dye was not closely monitored.

However, despite these limitations of measuring the exact distances, we can use time-resolved changes in FRET to assess whether the distance changes during protein synthesis. The FRET pair that is used in this project is Bodipy FL (BOF; donor) and Atto655 (acceptor). This FRET pair has a theoretical Förster radius of 46.83 Å (FPbase FRET calculator for $\kappa^2 = 2/3$). FRET was measured between the N-terminal Atto655-Met and BOF-Lys that was incorporated site-specifically during ongoing translation of the mRNA (Figure 10).

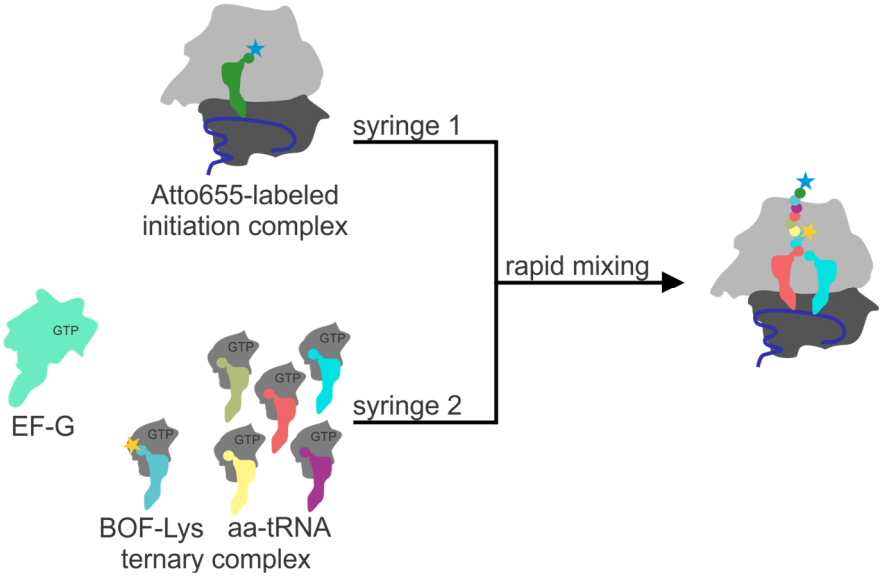



Figure 10. Schematic of a translation experiment monitoring the cotranslational nascent protein compaction by FRET. Syringe 1 of the stopped-flow machine contains purified initiation complex labeled with Atto655-Met. Syringe 2 contains EF-G and EF-Tu–GTP and aa-tRNA. The $tRNA^{Lys}$ is added as the BOF-Lys-$tRNA^{Lys}$. Mixing the solutions of the two syringes starts translation and BOF-Lys is incorporated site-specifically when the Lys codon is reached. When the nascent chain compacts, the two dyes come closer to each other and FRET can be observed.

In order to select the correct excitation wavelength and filters for cotranslational stopped-flow experiments, the excitation and emission spectra of Atto655-Met-$tRNA^{fMet}$ and BOF-Lys-$tRNA^{Lys}$ were measured in the fluorimeter.

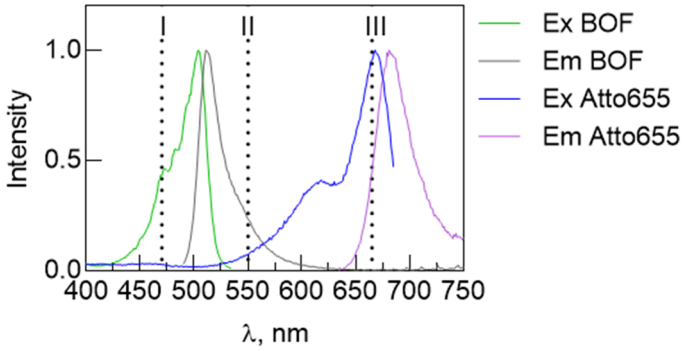



Figure 11. Excitation and emission spectra of BOF-Lys-$tRNA^{Lys}$ and Atto655-Met-$tRNA^{fMet}$ (40 nM each) in $HAKM_7$. Excitation spectrum of BOF-Lys-$tRNA^{Lys}$ (Ex BOF) was measured from 200 nm to 535 nm ($\lambda_{Em}$= 550 nm). Emission spectrum of BOF-Lys-$tRNA^{Lys}$ (Em BOF) was measured from 490 nm to 750 nm ($\lambda_{Ex}$= 480 nm). Excitation spectrum of Atto655-Met-$tRNA^{fMet}$ (Ex Atto655) was measured from 200 nm to 685 nm ($\lambda_{Em}$ = 700 nm). Emission spectrum of Atto655-Met-$tRNA^{fMet}$ (Em Atto655) was measured from 635 nm to 750 nm ($\lambda_{Ex}$= 625 nm). Bandpass was set between 3 nm and 5 nm. I: Excitation wavelength of donor BOF at 470 nm. II: Emission cut-off by KV550 filter to observe donor fluorescence in FRET experiment. III: Emission cut-off by RG665 filter to observe acceptor fluorescence in FRET experiment.

In a stopped-flow experiment, the best signal-to-noise ratio was achieved by excitation with a 470 nm light-emitting diode (LED) (Applied Photophysics; line I in Figure 11) at 2 mA and by using the emission filter KV550 (Schott; line II in Figure 11) for the donor channel and a RG665 filter (Schott; line III in Figure 11) for the acceptor channel (chapter 5.9).

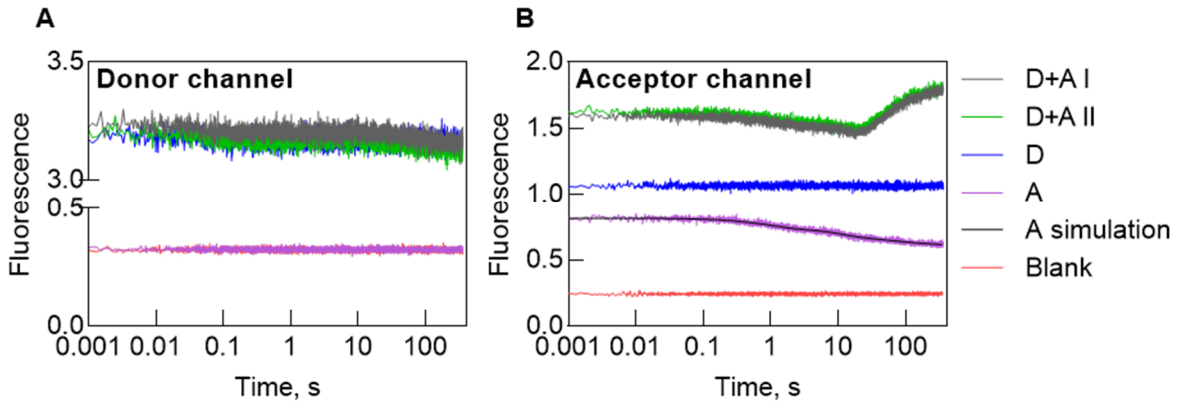



Figure 12: Example of the raw data of a cotranslational folding experiment observing FRET with HemK CTD K159R T120K. FRET experiments (D+A) were performed with two biological replicates consisting of 6 technical replicates each. D-only, A-only and blank control were collected with one biological experiment with 6 technical replicates each. The technical replicates were averaged. Simulation of A-only was calculated in TableCurve2D by model-free fitting to the best-fit equation. D: donor (BOF; internal); A: acceptor (Atto655; N-terminal).

The signal decrease in the donor channel that could originate from FRET (D+A) is marginal (Figure 12A). This is because the donor dye BOF was used in excess (500 nM BOF-Lys vs 40 nM Atto655-IC) and thus, FRET is mostly hidden in the background noise (Figure 12A). Therefore, the signal in the donor channel is not considered for folding experiments. However, because Atto655-Met is present in exactly one copy per nascent protein, we could use the change in acceptor fluorescence as an estimate for FRET. In order to ensure that the signal increase of the D+A sample in the acceptor channel (Figure 12B) originates from FRET, the donor-only, acceptor-only and additionally, the blank control needed to be collected using the same stopped-flow settings (detector sensitivity, excitation power, etc.) as was used for the D+A sample. The initial intensity of the acceptor-only sample (A in Figure 12B) is higher than the blank signal, arising from direct acceptor excitation. In Figure 11 the wavelength marked I is the ideal excitation wavelength of BOF at 470 nm. At this wavelength, it is also possible to excite Atto655 to some extent, especially by using the 470 nm LED that has a bandwidth of 20 nm. Also the donor-only sample has a higher intensity than the blank, originating from bleed-through of the donor emission into the acceptor channel (D in Figure 12B). According

to the spectra in Figure 11, the emission of BOF reaches a stable baseline at 665 nm (III) with approximately 0.2% of the emission peaks intensity. Using a RG665 filter should prevent any detection of donor emission in the acceptor channel. Nevertheless, the donor signal can bleed-through because the absolute intensity that is measured is concentration dependent and the donor is used in excess in this experiment. Additionally, the value 665 of the RG665 filter only indicates the wavelength at which the transmission of light is reduced to 50%. Thus, also lower wavelengths, which are present with higher intensities, can pass through the filter, albeit to a lower degree. In order to extract the fluorescence of the acceptor that originates from FRET, the fluorescence in the acceptor channel was corrected for all non-FRET fluorescence according to equation 8 (chapter 5.9).

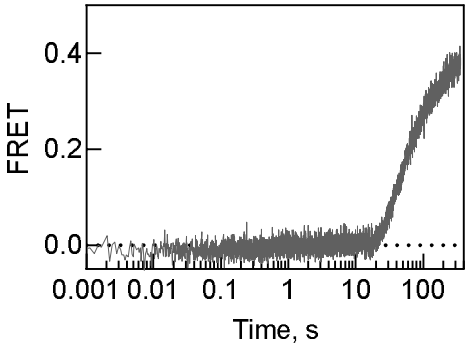



Figure 13. Example of the corrected fluorescence time course that originates from FRET extracted from a cotranslational folding experiment with HemK CTD K159R T120K. FRET: the acceptor fluorescence due to FRET.

The FRET time course in Figure 13 shows that the acceptor fluorescence that originates from FRET appears after a ~10 s delay. This part comprises the time until BOF-Lys is incorporated into the nascent chain and any folding events prior to nascent peptide compaction upon which the dyes come into a FRET-efficient distance range. After the delay, FRET increases steadily during peptide synthesis until the protein is folded. All FRET time courses shown in the following are derived from the full set of replicates and controls and are corrected identically as described above.

### 3.2.1. Identification of the structural elements comprising HemK CTD

As the aim of this project is to study the cotranslational folding pathway of the HemK CTD independently of the NTD, it was necessary to define the exact domain boundaries to include all structural elements that contribute to the folding and stability of the CTD. Structurally,

HemK consists of a 5-helix bundle NTD comprising residues 1-73, a short β-hairpin linker (residues 74-86) and the CTD (residue 87-276) that adopts a Rossmann fold (Figure 14) (Graille et al., 2005; Yang et al., 2004).

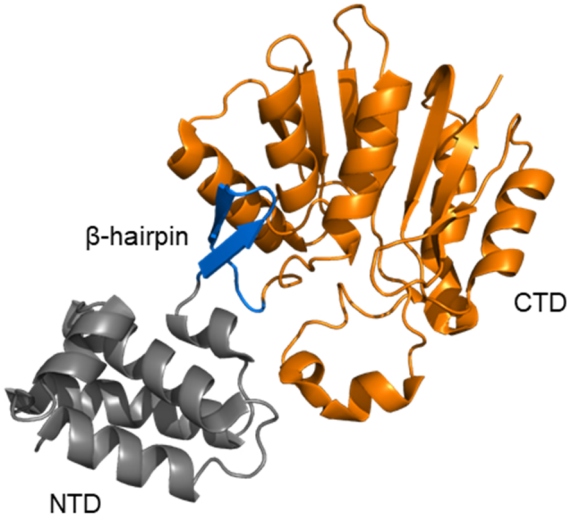



Figure 14. Crystal structure of HemK (PDB: 2B3T; (Graille et al., 2005)). According to Yang et al. (2004) and Graille et al. (2005) HemK consist of the NTD in grey, linker in blue and CTD in orange.

The NTD is able to fold to its native state independently of the CTD and the linker region (Holtkamp et al., 2015). The effect of the β-hairpin linker on the stability of the CTD is yet unknown, but has to be taken into account if we want to study the folding pathway of the domain. In terms of folding, one of the main driving force for protein folding is the hydrophobic effect and its corresponding entropy gain by water displacement upon compaction (Nick Pace et al., 2014). In order to obtain a better idea of the hydrophobic core of HemK CTD, the distribution of hydrophobic amino acids along the protein sequence was analyzed according to Kyte and Doolittle (1982).

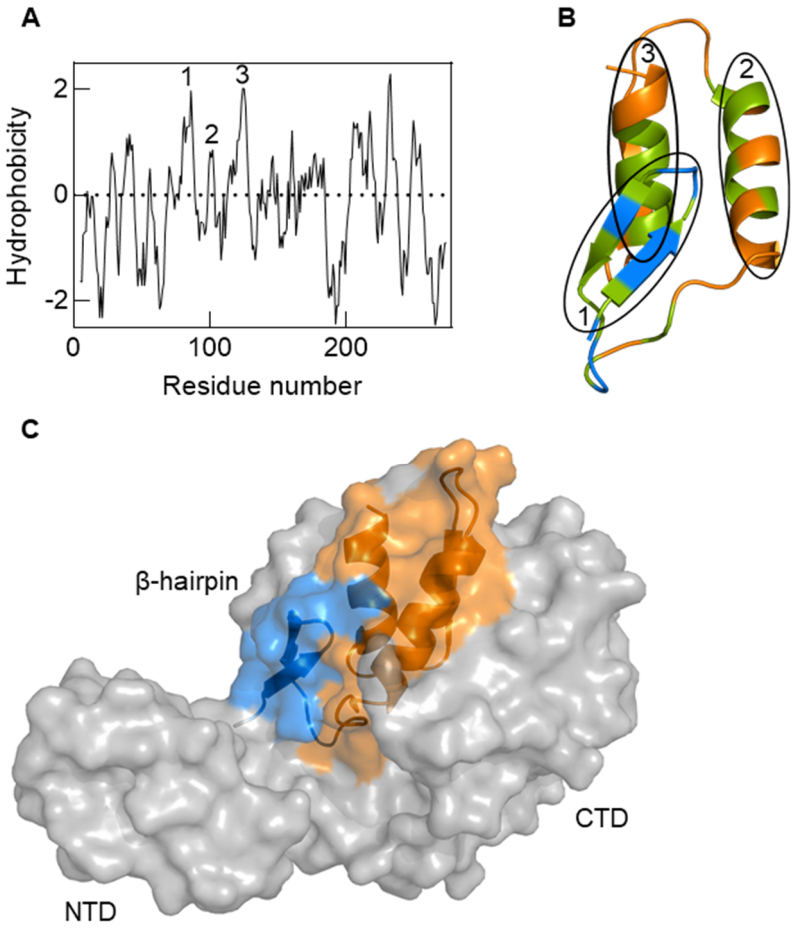



Figure 15. Analysis of the hydrophobicity of HemK primary sequence. (A) Hydrophobicity plot of HemK according to Kyte and Doolittle (1982) calculated by ProtScale (Expasy) using the default setting, i.e., averaging hydrophobicity of 9 aa at every amino acid (±4 aa). Hydrophobic clusters at the N-terminal part of the CTD are labeled with numbers 1-3 and assigned to (B) the respective secondary structures (PDB: 2B3T). The β-hairpin is colored blue and the helices orange. Additionally, all hydrophobic amino acids according to Kyte and Doolittle (1982) are colored green, including Ile, Val, Leu, Phe, Cys, Met and Ala. (C) Crystal structure of HemK in the surface representation. Surface was calculated with PyMol using the default solvent radius of 1.4 Å. The electron density of RF1 was removed for the calculation.

The β-hairpin linker belongs to a hydrophobic cluster and therefore is most likely not an exposed protein region (Figure 15A; peak 1). Additionally, it appears that the β-hairpin that is assigned to the linker region (1 in Figure 15B) is packed against the α-helices 2 and 3 of the CTD and is enclosed in the surface of the CTD (Figure 15C), with no solvent exposed surface between the β-hairpin and α-helices, closing the partially exposed hydrophobic patches of the α-helices. Because the β-hairpin emerges from the ribosome before the rest of the CTD and could potentially act as a folding nucleus for the CTD, we suggested that the β-hairpin is a part of the CTD. In order to test this hypothesis, the chaperone TF was used as a hydrophobic sensor

to analyze whether the linker contributes to shielding the hydrophobic patches of the HemK CTD.

### 3.2.2. Effect of TF on the folding of HemK CTD with and without the N-terminal β-hairpin

Chaperones assist protein folding by partially shielding the exposed hydrophobicity of a protein. By keeping the protein partially unfolded, chaperones can prevent aggregation by delaying the formation of local folding minima and allow folding only at later time points. Chaperones usually bind to exposed hydrophobic patches (reviewed in Kramer et al., 2019). The chaperone TF binds stably to nascent chains with a length of approximately 100 aa (Oh et al., 2011).

Figure 16. Crystal structure of HemK CTD (PDB: 2B3T). Colored according to Yang et al. (2004) and Graille et al. (2005) with the N-terminal β-hairpin in blue and the CTD in orange. Shown are residues 73-277.

Here we compared whether HemK CTD constructs with (residue 73-277) or without (residue 93-277) the N-terminal β-hairpin recruit TF in a cotranslational folding experiment (chapter 5.9).

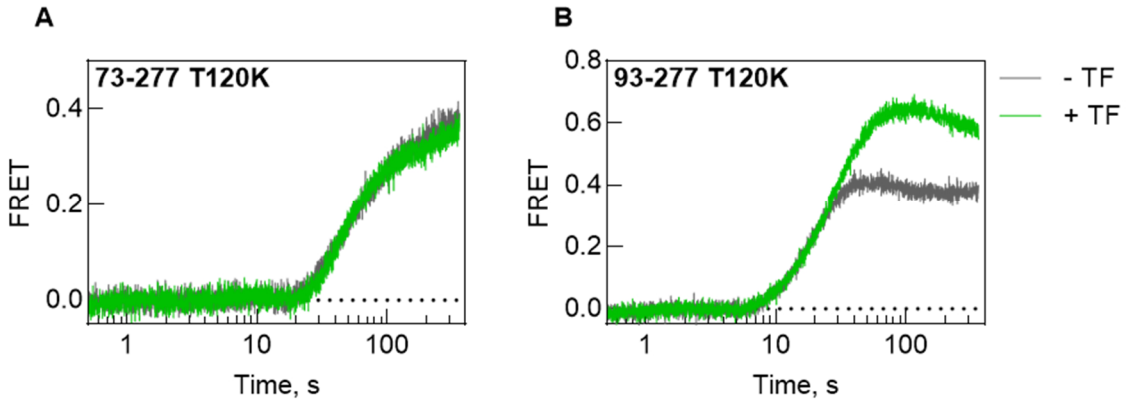



Figure 17. Effect of TF on the cotranslational folding of HemK CTD with and without the β-hairpin. FRET was observed between N-terminal Atto655-Met and internal BOF-Lys in the absence and presence of TF (2 µM) during translation of the mRNA coding for (A) HemK 73-277 K159R T120K or (B) HemK 93-277 K159R T120K. Real-time FRET measurements are performed as detailed in chapter 5.9. Each FRET time course was corrected for non-FRET fluorescence according to equation 8.

TF does not affect the compaction of the longer CTD construct HemK 73-277 that includes the β-hairpin, whereas compaction of the shorter CTD HemK 93-277 is altered upon TF interaction. One likely explanation is that removing the β-hairpin at the N-terminal part of HemK CTD exposes the hydrophobic amino acids of the CTD, which displays the usually hidden hydrophobic patches to TF. This implies that the β-hairpin is in direct contact with the CTD structural elements. Therefore, in the following, we define the β-hairpin as part of the HemK CTD domain and use the 73-277 construct to analyze the folding pathway of HemK CTD inside and outside the polypeptide exit tunnel.

### 3.3. CTD folding pathway on the ribosome

To study the structural rearrangements and sequential folding pathway of HemK CTD inside and outside of the polypeptide exit tunnel, a combination of force profile assay (FPA), photo-induced electron transfer (PET) and FRET measurements were utilized. The FPA was used to estimate the timing of cotranslational folding events as additional CTD sequence emerges from the PTC. Time-resolved and equilibrium PET measurements in the stopped-flow device and by fluorescence correlation spectroscopy (FCS), respectively, were used to get a closer insight into the folding events inside the polypeptide exit tunnel. Finally, time-resolved FRET measurements were employed to investigate the stepwise folding events outside of the polypeptide exit tunnel.

### 3.3.1. Monitoring folding by FPA

FPA is a relatively simple method to estimate the number of folding events in the nascent chain during translation. It can answer the question whether or not folding intermediates are formed by employing the stalling peptide of the protein SecM. The 17 aa-long peptide FSTPVWISQAQGIRAGP induces ribosome stalling by preventing the correct accommodation of the A-site Pro-tRNA. Interactions of the Arg residue of the peptide and the P-site Gly-tRNA with the 23S rRNA surrounding the PTC prevent the conformational changes of the PTC that favors the tRNA binding (Bhushan et al., 2011; Zhang et al., 2015). Pulling by the nascent chain that applies force on the PTC can lift the stalling and resume translation. The pulling force or tension can originate from the interaction of the nascent chain with another protein or from the folding of the nascent chain inside and outside of the polypeptide exit tunnel (Goldman et al., 2015; Ismail et al., 2012; Marino et al., 2018; Nilsson et al., 2015). Any sequence can be screened for these tension and relief events by cloning the protein of interest upstream of a SecM stalling peptide and a reporter peptide (Figure 18).


Figure 18. Scheme of an mRNA construct for the SecM-based FPA. The sequence coding for the protein of interest is colored black. The protein of interest is truncated C-terminally at different chain length, probing the tension at particular steps of translation. The protein of interest is followed by the 17 aa arrest peptide of SecM (orange) and the reporter that is produced when the ribosome stalling is alleviated (green).

We produced mRNAs that code for HemK truncated from the C-terminus approximately every two amino acids. Thereby, the entire sequence of the protein is stepwise screened for tension events by using the corresponding mRNAs in a translation experiment. If a certain length of nascent chain does not create a tension, the translation will abort after producing the SecM stalling sequence and only the stalled product (S) is formed. When there is a tension event, the stalling is alleviated, the translation resumes and the reporter sequence in the mRNA is translated, forming a longer full-length product (FL). The translation products are analyzed by SDS-PAGE and visualized by the N-terminal BOF label. The relative intensities of the stalled and full-length product are used to make conclusion about the magnitude of the tension events. In this work, a detailed FPA was carried out for two sequences, that of the HemK, including the N-terminal domain (1-277), and HemK CTD (73-277) without the NTD. For the HemK protein, only the CTD part (beginning from residue 93) was analyzed because the FPA of the

NTD is already published (Liutkute et al., 2020a), but the entire sequence of HemK CTD was analyzed.

In the following, the SDS-PAGE analysis of the HemK FPA (NTD was translated for all constructs; Figure 19) and the HemK CTD FPA (Figure 20) is presented.

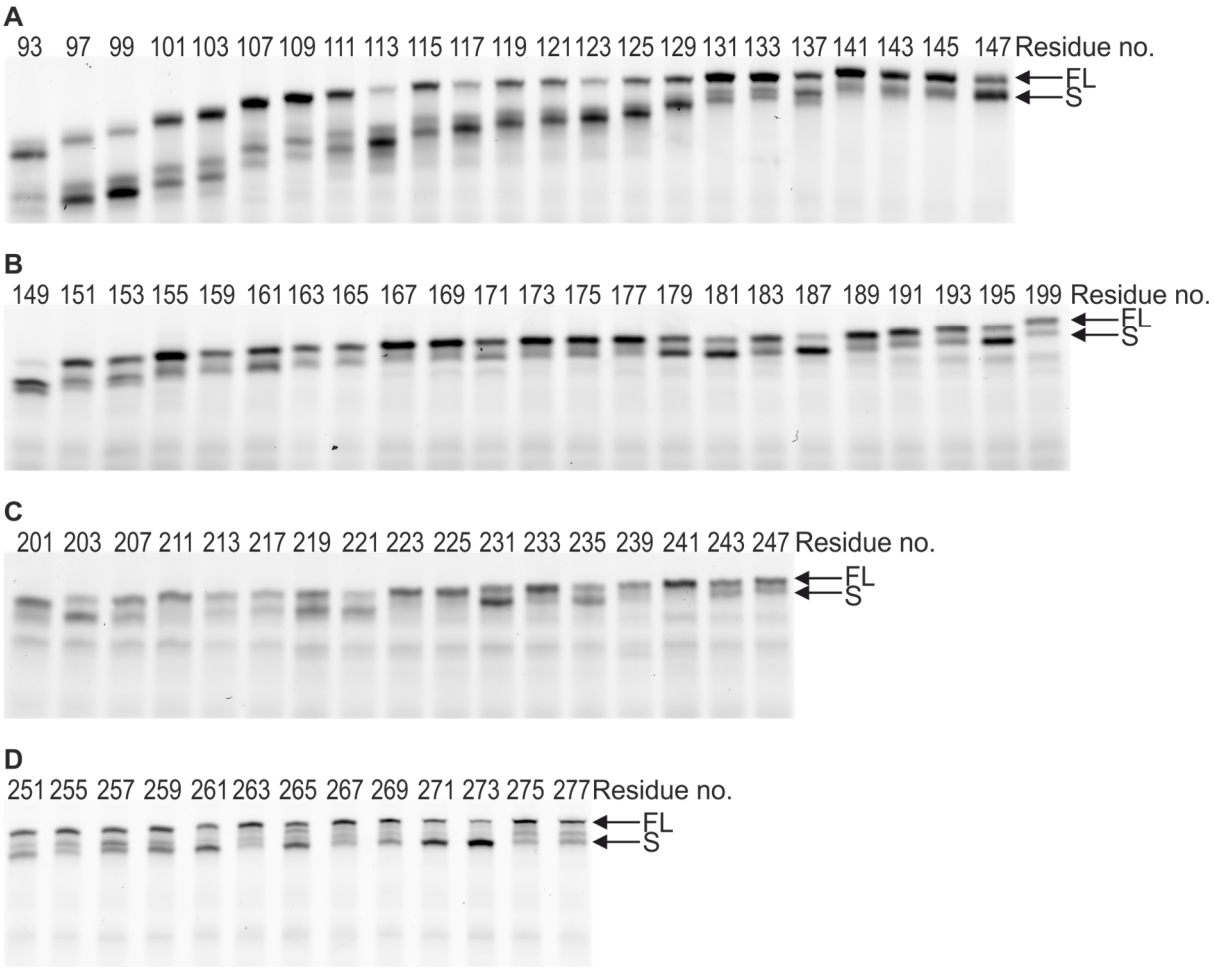



Figure 19. Representative SDS-PAGE of the HemK (1-277) FPA. The lanes are labeled with the residue number of HemK that is still included, i.e., 93 corresponds to HemK 1-93. The shorter product 'S' refers to the translation product that is stalled after translation of the HemK sequence together with the 17 aa SecM arrest peptide. The full-length product 'FL' corresponds to translation products with the additional reporter CspA due to the tension event. Translation products were visualized by N-terminal BOF. In (A)/(B)/(C) the reporter consist of the first 20 aa of CspA and in (D) of 50 aa of CspA. All experiments were carried out in triplicates (only one is shown).

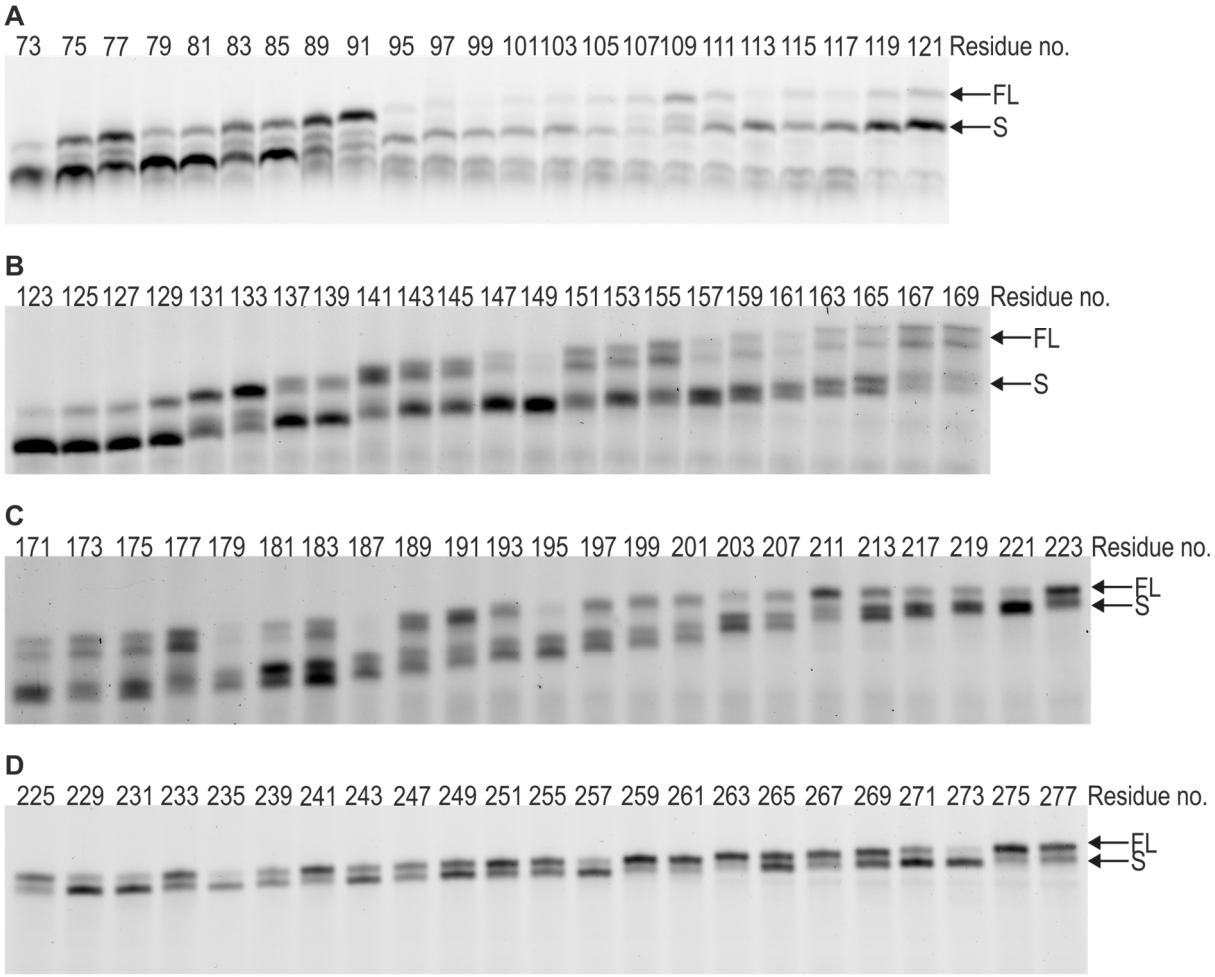



Figure 20. Representative SDS-PAGE of HemK CTD (73-277) FPA. The lanes are labeled with the residue number of the last amino acid of HemK CTD that is still included in the profile, i.e., 73 indicates the construct has only one aa (residue 73) from HemK CTD followed by the SecM and reporter sequences. The shorter product 'S' corresponds to the stalled translation product, 'FL' corresponds to the formation of the product with the additional CspA sequence due to the tension event. Translation products were visualized by N-terminal BOF. All experiments were carried out in triplicates (only one is shown).

The fraction of the full-length product is calculated from the band intensities on the SDS-PAGE according to

$$f_{FL} = \frac{I_{FL}}{I_{FL}+I_S}$$

where $I_{FL}$ is the band intensity of the full-length product and $I_S$ the band intensity of the stalled product. In case that there are stalling intermediates between FL and S (e.g., Figure 19D; sample '277'), the gel intensity of the intermediate is considered as part of the FL intensity because it is peptide that resumed translation due to tension. The translation of the HemK CTD constructs with truncations at residue 147-177 resulted in a double band for both FL and S,

which is inferred by the general trajectory of S and FL of the preceding and following lanes (Figure 20B and C). Therefore, the sum of the double band intensities was used to calculate the fraction.

The sample labeled with 73 of the HemK CTD FPA (Figure 20A) only encodes for the start codon at residue 73 (first amino acid) of HemK CTD upstream of the SecM stalling peptide, and thus, only translates the SecM stalling peptide. This sample represents the tension baseline because it originates from the force the SecM stalling peptide generates at the PTC and is used as the cut-off for tension events.

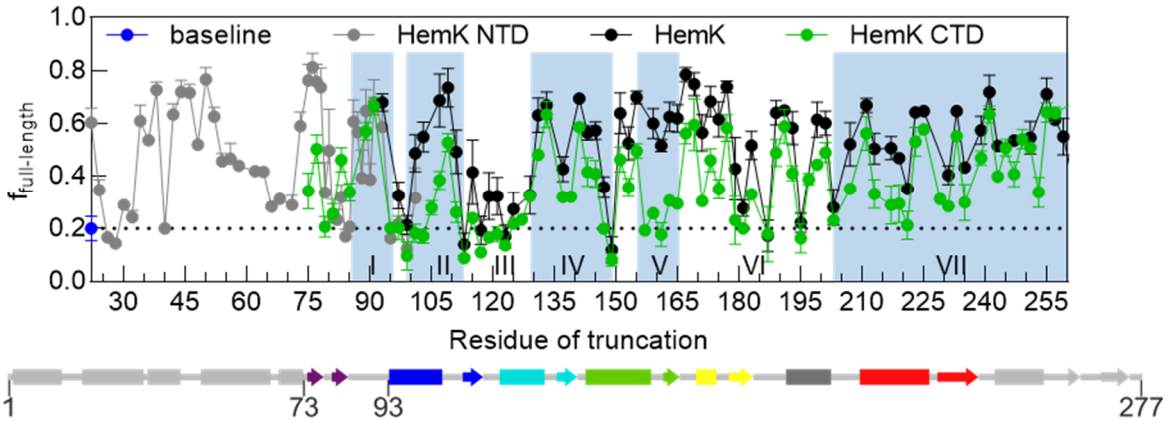



Figure 21. FPA of HemK and HemK CTD. Plotted are the fractions of full-length product on SDS-PAGE against the respective residue of truncation, i.e., the last aa that is still included in the sequence. The FPA of the HemK NTD (grey dots) is considered as part of the HemK FPA and was taken from Liutkute et al. (2020a). High values indicate strong tension. The scheme of HemK secondary structure elements is aligned with the profile to indicate the parts of HemK that were translated upstream of the stalling peptide. Values are plotted as mean ± SD (N=3). Highlighted in pale blue and numbered from I-IV are high tension areas.

The force profile of HemK (black/ grey for NTD part) and HemK CTD (green) indicate that there are many high tension events during the CTD synthesis, interrupted by short periods of translation without tension. The first tension event starts at residue 85-95 and peaks at residue 91 at the very beginning of the CTD (I in Figure 21). The second tension event reaches its maximum at residue 109 (II; 99-113). Here the tension is somewhat lower when HemK CTD is produced, but the peak is found at the same positions (II). In both profiles, this is followed by a low tension stretch from residue number 113-129 (III) and a short high tension stretch between 129 and 149 (IV). Section V is located at residue 155-165 and marks the largest deviation between the FPAs with the HemK CTD FPA exerting almost no tension. The force profile resynchronizes between residues 167-203 with pronounced fluctuations in tension in

both profiles (VI). Starting after residue 203 (VII), the tension is fluctuating, but with a tendency to increase. This continues until the end of the HemK construct at residue 260. Here, it is noteworthy that in addition to the part that encodes HemK and HemK CTD, the 17 aa long SecM peptide is translated as well. In case of HemK this means that at the residue number 260 that are indicated in the plot, 277 amino acids have been synthesized that amount to the full-lengths translation product of HemK.

The force profiles indicate that removing the NTD has only minor effects on the local folding events, in particular, in area II and V. As the FPA reports on folding inside and outside the polypeptide exit tunnel, it is possible that area II and V reflect structural rearrangements involving the NTD. In general, the folding of the CTD seems to occur largely independently of the NTD, which makes it a suitable model protein to study cotranslational folding of large mixed-structured proteins.

### 3.3.2. Folding pathway inside the polypeptide exit tunnel

During translation the growing peptide has to pass through the approximately 100 Å long polypeptide exit tunnel to reach the outside of the ribosome. The polypeptide exit tunnel is relatively narrow with the width ranging between 10 and 20 Å (Nissen et al., 2000). Only limited folding events can happen in its confined space, such as the formation of α-helices (Bhushan et al., 2010; Lu and Deutsch, 2005; Woolhead et al., 2004), small β-hairpins (Agirrezabala et al., 2022; Kosolapov and Deutsch, 2009) and small tertiary structures (Nilsson et al., 2015). Apart from the constriction site, α-helices are able to form along the entire tunnel starting from the PTC, whereas β-hairpins and small tertiary structures are capable to form closer to the exit port, where the exit tunnel becomes wider (Agirrezabala et al., 2022; Bhushan et al., 2010; Kosolapov and Deutsch, 2009; Lu and Deutsch, 2005; Nilsson et al., 2015; Woolhead et al., 2004). The propensity to form these structures inside the ribosome are largely defined by space restriction imposed by the polypeptide exit tunnel. Sensing mobility restriction can be investigated by using PET, a phenomenon that is sensitive to short range protein dynamics. PET quenches fluorescence upon van der Waals contacts between a PET donor and a PET acceptor. After excitation of the fluorophore, both the quencher and the excited fluorophore can act as PET donor or acceptor, depending on the redox potential, but ultimately an electron is transferred if both are in close proximity, leading to the loss of the excited state and hence, quenching (reviewed in Doose et al., 2009). The naturally occurring Trp in the HemK sequence at residue number 78 ($6^{th}$ aa in the CTD) can be used to monitor

PET because it is a very efficient PET donor. PET as a measure for direct physical contact between Trp and the N-terminal dye provides information about the compaction and dynamics of the nascent chain (Neuweiler et al., 2003).

#### 3.3.2.1. Cotranslational PET measurements in stopped-flow device

Trp78 (which is 5 aa away from the N-terminus of the CTD) is in close proximity to the N-terminal BOF and could potentially quench the dye if the dynamics of the protein allows dynamic fluctuations bringing Trp78 in a van der Waals radius of the dye. Changes in fluorescence intensity during translation may indicate compaction / relaxation events of the nascent peptide inside and outside the polypeptide exit tunnel (Agirrezabala et al., 2022; Holtkamp et al., 2015). For HemK CTD (73-277), where Trp78 is solvent-exposed in the native protein structure, the PET pair Trp78/N-terminal BOF-Met73 reports conformational rearrangements of the first strand in the β-hairpin (Figure 22) as it moves from the PTC towards the tunnel exit of the ribosome.


Figure 22. Crystal structure of N-terminal β-hairpin of HemK CTD (PDB: 2B3T). Shown are residues 73-83 of HemK. BOF-Met is marked in red and Trp represented as stick model.

One complication of PET experiments on the ribosome is that during translation the dye passes through the polypeptide exit tunnel and interacts with the rRNA and proteins of the exit tunnel

wall. rRNA is rich in guanine. Guanine is also a highly efficient PET donor. In order to make conclusions about the folding of the peptide itself and thus extract intramolecular Trp-PET, the information about the polypeptide exit tunnel quenching has to be excluded. The most straightforward approach is to use an equivalent sample that does not contain Trp in the nascent peptide in order to estimate the quenching by the ribosome surface. This sample is used to calculate by how much the Trp additionally reduces the fluorescence, hence intramolecular Trp-PET. Therefore, for every peptide studied by PET, a construct with a Phe substitution of Trp (W78F) was studied as well. HemK CTD mRNAs encoding both variants were translated in the stopped-flow machine and fluorescence was measured with the excitation of the N-terminal BOF dye (Figure 23A, B).

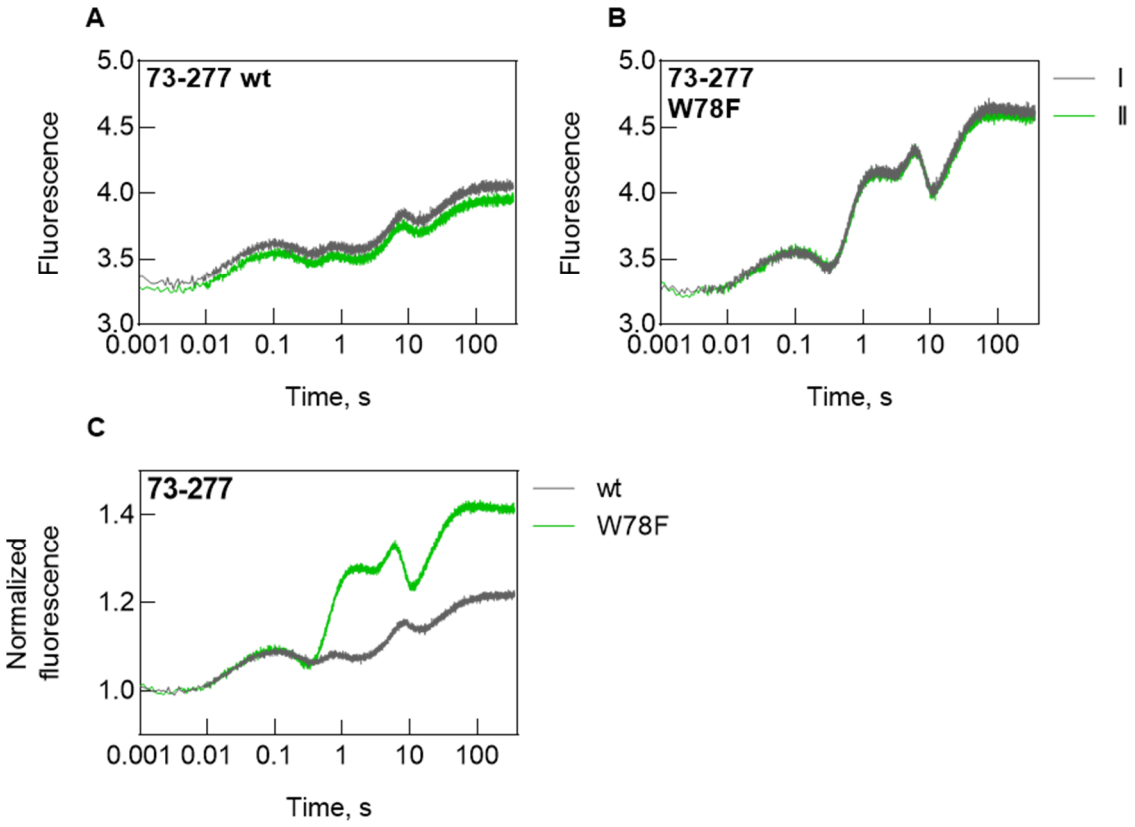



Figure 23. Example of the raw data of cotranslational folding experiment using PET as observable. Fluorescence changes of N-terminal BOF in (A) wt and (B) W78F sequence variants were measured in two independent experiments with 6 technical replicates each. The technical replicates were averaged. (C) Comparison of the normalized and averaged fluorescence time courses. Each biological replicate was normalized to 1 at the initial intensity by using the average fluorescence between 0.001 and 0.005 s, followed by calculating the average of the biological replicates (chapter 5.11).

The initial small signal increase (observed at 0.1 s in Figure 23C) is identical in the wt and W78F experiments and serves as a normalization control. This was checked for all constructs and was necessary to make sure that the calculation of the Trp-PET efficiency can be done correctly according to equation 13 (chapter 5.11), resulting in a Trp-PET efficiency time course (Figure 24).

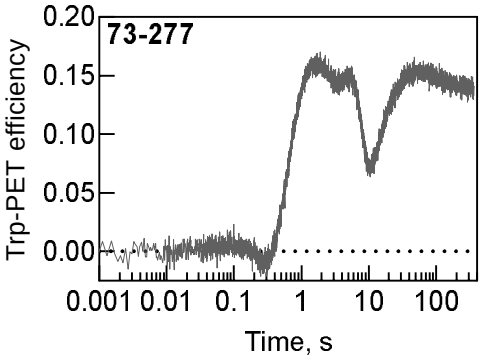



Figure 24. Intramolecular Trp-PET efficiency measured in real time during the synthesis of HemK 73-277.

Trp-PET time course of HemK CTD with different length were obtained in the same way. By using C-terminally truncated constructs, it is possible to relate the early folding events of the nascent peptide with the nascent peptide length. Taking into account that the polypeptide exit tunnel has a length of approximately 100 Å, it takes between 30 aa of a fully extended and 60 aa of an α-helical peptide to span the entire tunnel. Thus, the longest truncation that was analyzed together with the HemK CTD (73-277) coded 60 aa (HemK 73-132) of HemK CTD (Figure 25A).

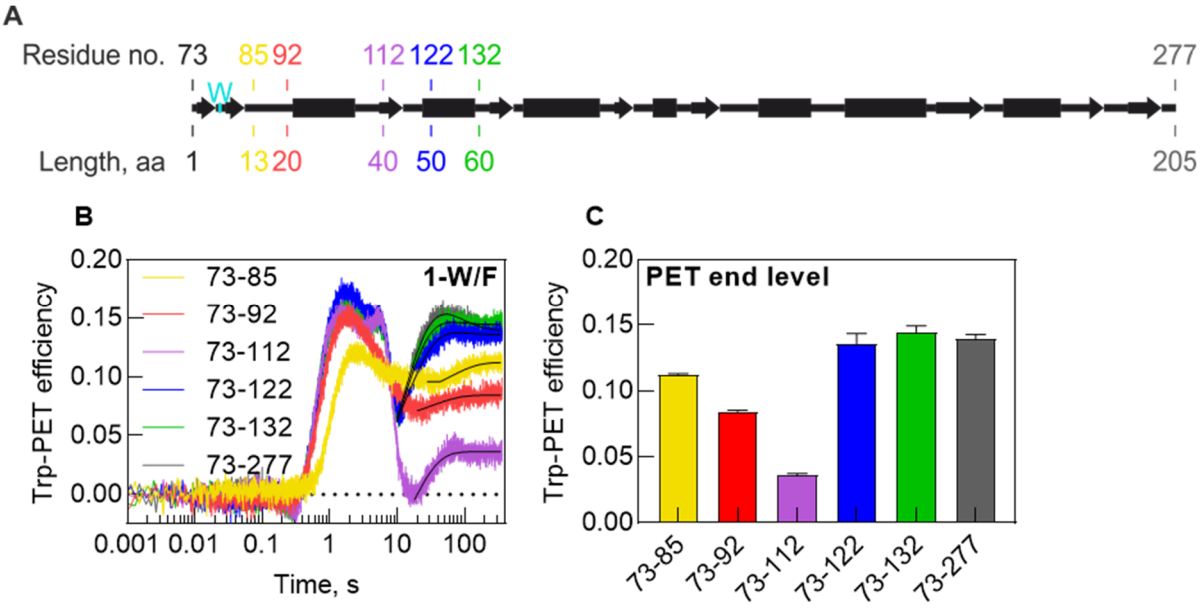



Figure 25. Cotranslational folding experiment monitoring PET. (A) Scheme of HemK CTD secondary structure, labeled with the residue numbers of HemK and the corresponding length (aa) of the respective CTD peptide. Trp78 is indicated as 'W' in cyan. (B) Time-resolved Trp-PET efficiency is calculated according to equation 13 (chapter 5.11) from fluorescence time courses with N-terminal BOF-Met73 in HemK CTD with W78 or W78F constructs of different length as indicated (2 biological replicates with 6 technical replicates each; chapter 5.11). (C) PET end levels extracted from fitting of (B) with a delay-single exponential function (73-85; 73-92; 73-112; equation 9) or delay-double exponential function (73-122; 73-132; 73-277; equation 10), only considering the values passed the minima at around 10 s. Error bars reported as SEM.

The Trp-PET time courses of all studied HemK CTD lengths have similar features (Figure 25B). First, after a delay of about 0.3-0.6 s, Trp-PET efficiency increases rapidly with a peak at approximately 1 s (Figure 25B). Trp-PET can only occur after the incorporation of Trp into the peptide chain, hence, it is likely that this increase reflects the incorporation of Trp78 and the initial compaction of the nascent peptide. Second, there is a drop in Trp-PET efficiency at approximately 10 s, the amplitude depends on the peptide length (Figure 25B). The narrow section of the constriction site would allow passing through of mainly unfolded and extended peptide and therefore limits the mobility that is necessary for PET between the dye and Trp. Finally, PET increases over the time range of 10-100 s, reaching the plateau that is plotted in Figure 25C. The Trp-PET end levels of the short CTD peptides represent the various structural rearrangements of the nascent peptide at different positions inside the polypeptide exit tunnel. The three longest constructs, HemK 73-277, 73-132 and 73-122, have similar Trp-PET end levels. This indicates that beginning from a length of roughly 50 aa, the first β-strand is compacted to a similar degree. In contrast, shorter nascent chains exhibit

different PET end levels, pointing at structural rearrangements of the N-terminal peptide during the translation.

The differences in the final Trp-PET efficiencies can be interpreted in two ways (Neuweiler et al., 2003). First, there can be two sets of stable confirmations, one that is completely fluorescent and the other one that is completely quenched and there are no transitions from one to the other. In this case, the Trp-PET efficiency reflects the ratio of these populations. The second option is that the peptide fluctuates between the two conformations. In this case, the Trp-PET efficiency reflects the position of the equilibrium. In order to distinguish between these two possibilities and test how dynamic the CTD nascent peptide conformations are inside and outside the polypeptide exit tunnel, the steady-state dynamics of the truncated and full-length HemK CTD variants were measured by PET-FCS.

#### 3.3.2.2. Equilibrium dynamics measured by PET-FCS

PET-FCS is a single-molecule technique that reports on fluctuations of fluorescence intensity in a steady-state system and therefore on the structural dynamics in equilibrium (Neuweiler et al., 2003). Here it is utilized as a method complementary to ensemble cotranslational PET measurements and is used to test whether the peptide dynamics inside the polypeptide exit tunnel correlate with the Trp-PET efficiencies measured in the kinetic experiment (chapter 3.3.2.1). In order to make conclusions about the nascent chain dynamics PET-FCS experiments were carried out as described in chapter 5.12.

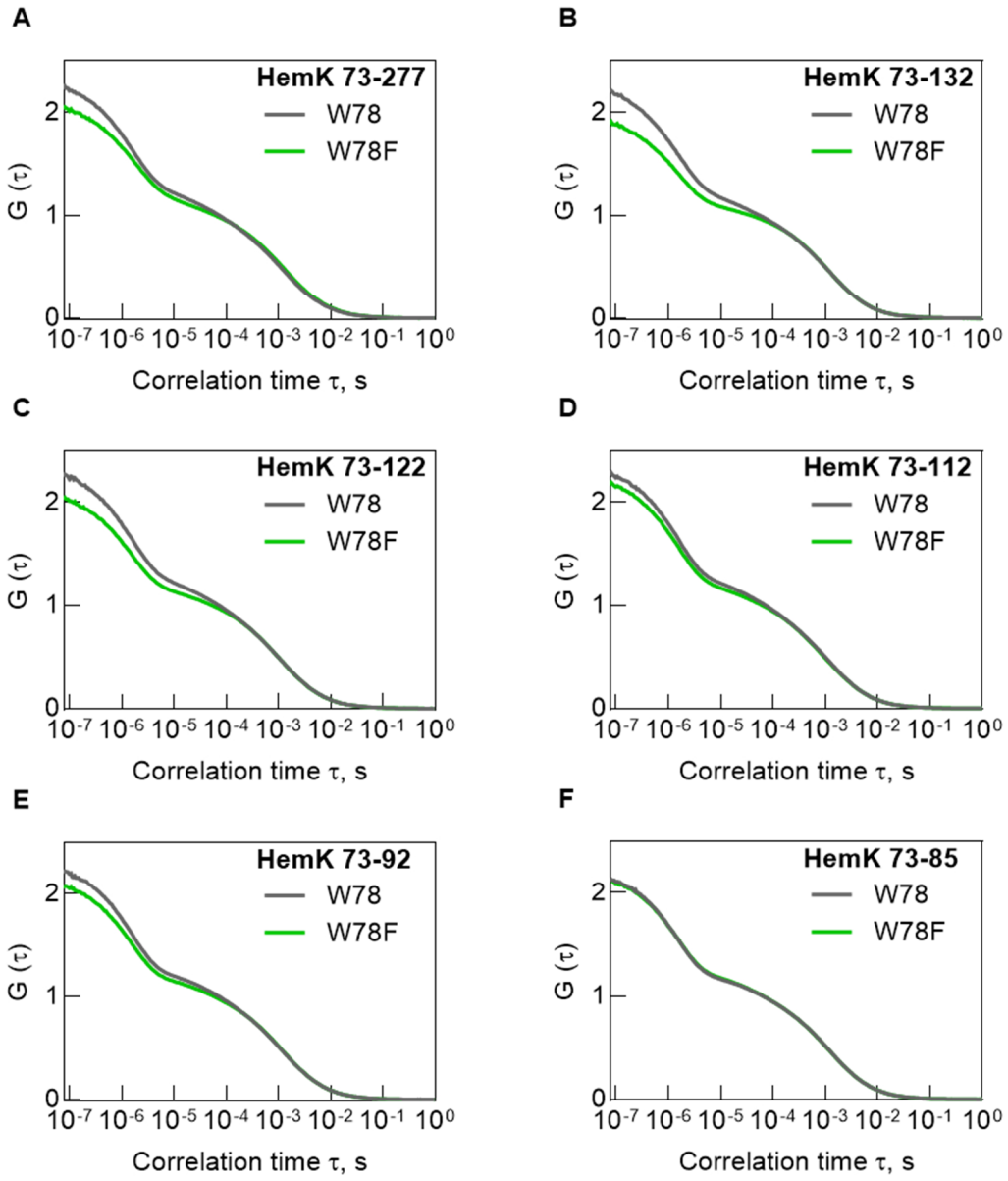



Figure 26. Fast dynamics of the nascent chain in the polypeptide exit tunnel measured by PET-FCS. Autocorrelation function was calculated from fluorescent traces with N-terminally Atto655-labeled HemK 73-277 wt or W78F RNCs and its truncations. Each autocorrelation function is averaged from 3 biological replicates with 4 technical replicates each and was normalized to N=1 according to equation 14 (chapter 5.12).

The best fit parameters of the normalized autocorrelation function (ACF; Figure 26) are listed in Table 2. The lifetime τ is reported as the reciprocal of the fitted rate k.

Table 2. Best fit parameters of all ACFs of all constructs. Values are reported ± SEM.

| Construct | Type | $c_1$ | $\tau_1$, µs | $c_2$ | $\tau_2$, µs | F | $\tau_{trip}$, µs | N | $k_d$, $s^{-1}$ |
|---|---|---|---|---|---|---|---|---|---|
| HemK 73-85 | wt (W78) | 0.15 ± 0.01 | 0.40 ± 0.03 | 0.67 ± 0.01 | 1.81 ± 0.03 | 0.17 ± 0.01 | 49 ± 2 | 1 ± 0.01 | 941 ± 6 |
| | W78F | 0.17 ± 0.01 | 0.49 ± 0.03 | 0.61 ± 0.01 | 1.89 ± 0.03 | 0.18 ± 0.01 | 50 ± 1 | 1 ± 0.01 | 948 ± 4 |
| HemK 73-92 | wt (W78) | 0.19 ± 0.02 | 0.54 ± 0.04 | 0.64 ± 0.02 | 2 ± 0.04 | 0.19 ± 0.01 | 56 ± 2 | 1 ± 0.01 | 940 ± 6 |
| | W78F | 0.14 ± 0.01 | 0.33 ± 0.02 | 0.66 ± 0.01 | 1.79 ± 0.02 | 0.16 ± 0.01 | 44 ± 1 | 1 ± 0.01 | 925 ± 4 |
| HemK 73-112 | wt (W78) | 0.21 ± 0.01 | 0.56 ± 0.03 | 0.64 ± 0.01 | 2.08 ± 0.04 | 0.2 ± 0.01 | 54 ± 1 | 1 ± 0.01 | 1022 ± 5 |
| | W78F | 0.2 ± 0.01 | 0.46 ± 0.02 | 0.66 ± 0.01 | 2 ± 0.03 | 0.17 ± 0.01 | 51 ± 1 | 1 ± 0.01 | 1052 ± 5 |
| HemK 73-122 | wt (W78) | 0.22 ± 0.01 | 0.49 ± 0.03 | 0.63 ± 0.01 | 2.17 ± 0.05 | 0.21 ± 0.01 | 52 ± 2 | 1 ± 0.01 | 1021 ± 7 |
| | W78F | 0.17 ± 0.01 | 0.41 ± 0.02 | 0.61 ± 0.01 | 1.98 ± 0.03 | 0.15 ± 0.01 | 43 ± 1 | 1 ± 0.01 | 1031 ± 5 |
| HemK 73-132 | wt (W78) | 0.20 ± 0.01 | 0.37 ± 0.02 | 0.65 ± 0.01 | 2.03 ± 0.03 | 0.19 ± 0.01 | 40 ± 1 | 1 ± 0.01 | 1009 ± 5 |
| | W78F | 0.16 ± 0.01 | 0.34 ± 0.02 | 0.57 ± 0.01 | 1.91 ± 0.03 | 0.12 ± 0.01 | 31 ± 1 | 1 ± 0.01 | 1001 ± 4 |
| HemK 73-277 | wt (W78) | 0.25 ± 0.02 | 0.62 ± 0.04 | 0.57 ± 0.02 | 2.37 ± 0.07 | 0.2 ± 0.01 | 55 ± 2 | 1 ± 0.01 | 907 ± 6 |
| | W78F | 0.21 ± 0.01 | 0.55 ± 0.04 | 0.54 ± 0.01 | 2.65 ± 0.08 | 0.16 ± 0.01 | 49 ± 2 | 1 ± 0.01 | 819 ± 6 |

In order to quantify the dynamics that originate from Trp-Atto655 interaction, the differences of the Y-intercept of the wt and W78F variant were calculated with the fitting parameter in Table 2. The Y-intercept represents the average number of fluorescent molecules, i.e., molecules that are not in the dark state due to quenching (reviewed in Krichevsky and Bonnet, 2002; Lakowicz, 2006), allowing to dissect the contributions of PET by the ribosome residues (measured with the W78F construct) from that of the intramolecular Trp-PET. The Trp-Atto655 dynamics are quantified by dG(0), which is calculated as $G(0)_{wt} - G(0)_{W78F}$ and is used as a proxy to highlight the contribution of the Trp78 to quench the N-terminal Atto655 by PET due to dynamic motion of the peptide (Table 3 ; Figure 27).

Table 3. Analysis of the autocorrelation functions from PET-FCS. The Y-intercept G(0) was calculated according to equation 15 (chapter 5.12) with the fitted parameters listed in Table 2. dG(0) was calculated with $G(0)_{wt}$ - $G(0)_{W78F}$.

| **Construct** | **Type** | **G(0) ± SEM** | **dG(0) ± SEM** |
|---|---|---|---|
| HemK 73-85 | wt (W78) | 2.2 ± 0.02 | 0.03 ± 0.03 |
| | W78F | 2.17 ± 0.02 | |
| HemK 73-92 | wt (W78) | 2.28 ± 0.03 | 0.13 ± 0.03 |
| | W78F | 2.14 ± 0.01 | |
| HemK 73-112 | wt (W78) | 2.32 ± 0.02 | 0.08 ± 0.03 |
| | W78F | 2.25 ± 0.02 | |
| HemK 73-122 | wt (W78) | 2.33 ± 0.03 | 0.24 ± 0.03 |
| | W78F | 2.09 ± 0.01 | |
| HemK 73-132 | wt (W78) | 2.28 ± 0.01 | 0.32 ± 0.02 |
| | W78F | 1.97 ± 0.01 | |
| HemK 73-277 | wt (W78) | 2.29 ± 0.03 | 0.2 ± 0.04 |
| | W78F | 2.08 ± 0.02 | |

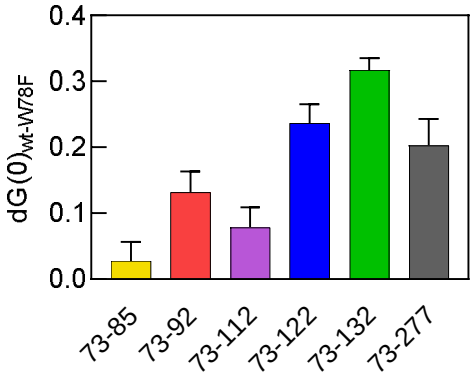



Figure 27. Intramolecular Trp-PET dynamics of HemK CTD with different length. The values are calculated from the fitted parameter and error reported as SEM (reported in Table 3).

In general, the differences of the wt and W78F Y-intercepts dG(0) of the three longer nascent peptides are larger than those of the three shorter peptides. A larger value can be considered as an indication for a higher relative intramolecular nascent chain dynamics. To check whether the trend for a more pronounced steady-state dynamics from PET-FCS correlates with that derived from the ensemble Trp-PET efficiencies, the two values from each construct were plotted against each other (Figure 28).

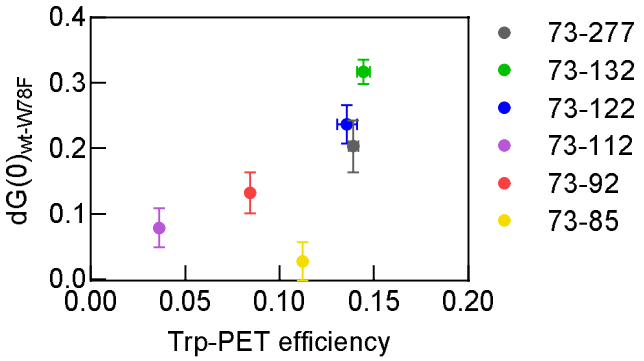



Figure 28. Correlation between PET-FCS dynamics and PET efficiencies from the stopped-flow experiment. dG(0) values are from Table 4, Trp-PET efficiencies from Figure 25.

In general, larger Trp-PET efficiencies from the stopped-flow data correlate well with larger Trp-PET dynamics from PET-FCS. This suggests that Trp-PET mainly arises from constant conformational transitions between a quenched and a fluorescent species. However, the construct HemK 73-85 (13 aa) seems to be an outlier with measurable Trp-PET efficiency and almost no Trp-PET dynamics (yellow in Figure 28). It is possible that there are two stable conformations of this very short peptide. One of these populations is constantly quenched by the formation of a stable Trp-Atto655 complex and is not visible in the ACF, but contributes to the Trp-PET efficiency measured in the stopped-flow, whereas the second population is fluorescent. The fluorescent population gives rise to dynamic interactions with the polypeptide exit tunnel, accounting for the same fast dynamic term in the ACF of both the wt and W78F variant (Figure 26F).

### 3.3.3. Folding pathway outside of the polypeptide exit tunnel

We next studied the folding pathway of the HemK CTD as it emerges from the narrow confinement of the polypeptide exit tunnel. We used the FRET system established here (chapter 3.2.3 and 5.9) to extract the information for the CTD folding pathway. FRET reports on the compaction of the part of the nascent chain that is located between the two fluorophores. By moving the second internal dye further along the CTD sequence, an increasing part of the nascent chain is stepwise included into the analysis. By following FRET in a time-resolved manner during active translation, we obtain the time of nascent chain compaction. If folding is concerted, i.e., the protein compacts rapidly after growing to a certain critical length, then the time taken for folding will remain constant, irrespective of the position of the second dye along the protein sequence. In contrast, if folding is sequential, i.e., it begins with the formation of

local interactions between neighboring protein regions followed by docking of structural elements appearing later in translation, the time of folding will increase steadily with later incorporation sites for the second dye.

Additionally, comparing FRET time courses of peptides with different internal labeling positions makes it possible to capture folding events potentially hidden in a folding pattern of a longer peptide. As all folding experiments are performed during ongoing translation, the folding times measured by FRET can be compared to translation time courses of C-terminally truncated versions of the CTD mRNAs. The translation time courses provide the information about the translation time of peptides of a given length. Comparison of the folding time with translation time indicates the timing of a given compaction event during translation.

Based on this experimental strategy, analyzing the sequential folding pathway of HemK CTD included 3 major steps. First, we obtained translation time courses of HemK CTD wt and its different C-terminally truncated versions. These time courses served as a reference to capture the length of the peptide that permitted a folding event. Second, we screened for the potential BOF-Lys incorporation positions throughout the HemK CTD sequence. The positions of the dye incorporation had to be tested by translation time courses in order to ensure that translation kinetics do not change upon dye incorporation. Third, we performed time-resolved FRET measurements with all the chosen BOF-Lys positions to analyze the timing of nascent peptide compaction.

#### 3.3.3.1. Translation time courses of C-terminally truncated HemK 73-277 wt

We produced mRNAs coding for HemK CTD with truncations from the C-terminus in 20 codon steps (25 codons in the first step) to have a detailed grid of translation time courses that is used as a reference to compare to the folding times. The translation products were visualized on SDS-PAGE by the N-terminal BOF (Appendix Figure 1) and processed as described in chapter 5.8.3.1 (Figure 29). The fitting parameters are listed in Table 4.

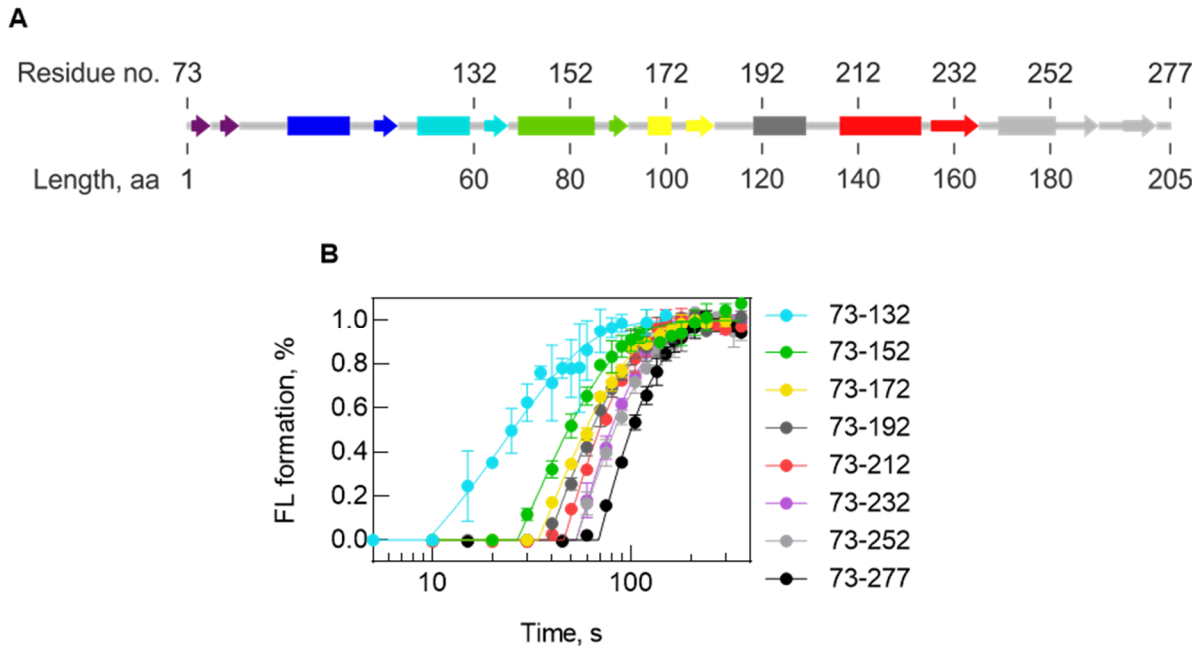



Figure 29. Translation time courses of HemK 73-277 wt and its truncations. (A) Scheme of HemK CTD secondary structure indicating the residue of the truncation and the length (aa) of the truncated peptide of the CTD. (B) Translation time courses of proteins with different length as indicated. The band intensities (SDS-PAGE; Appendix Figure 1) of the respective full-length product (FL) were quantified and normalized to 1 (chapter 5.8.3.1) and reported as mean ± SD (N=3). Smooth lines are fits to a delay-single exponential function (equation 4).

Table 4: Translation times and average translation rates of HemK CTD constructs. Translation time $\tau_{trans}$ ± SEM (N=3) of each mRNA construct and average translation rate in aa/s calculated according to equation 5 and equation 7 (chapter 5.8.3.1), respectively, are listed together with the corresponding length of the translated full-length product.

| **Construct** | **Length, aa** | **Translation time $\tau_{trans}$ ± SEM** | **Translation rate aa/s ± SEM** |
|---|---|---|---|
| HemK 73-132 | 60 | 32 ± 3 | 1.9 ± 0.2 |
| HemK 73-152 | 80 | 58 ± 2 | 1.4 ± 0.1 |
| HemK 73-172 | 100 | 70 ± 2 | 1.4 ± 0.1 |
| HemK 73-192 | 120 | 74 ± 2 | 1.6 ± 0.1 |
| HemK 73-212 | 140 | 80 ± 2 | 1.8 ± 0.1 |
| HemK 73-232 | 160 | 91 ± 2 | 1.8 ± 0.1 |
| HemK 73-252 | 180 | 96 ± 3 | 1.9 ± 0.1 |
| HemK 73-277 | 205 | 114 ± 2 | 1.8 ± 0.1 |

In general, with longer constructs, translation takes more time to complete, which is represented by a longer translation time. The information on the translation rate in aa/s provides insight into the translation pattern of HemK CTD. A slight decrease of the translation rate of the HemK 73-152 and 73-172 constructs indicates a non-uniform translation rate along the mRNA, such as the presence of pause sites or slowly translating codon stretches.

In the following chapter, the translation time course of HemK CTD (73-277 in Figure 29) was used as a reference to test the suggested incorporation sites of BOF-Lys to control the influence of cotranslational dye incorporation on the translation rate.

#### 3.3.3.2. Screening BOF-Lys incorporation sites throughout the CTD sequence

In order to monitor the cotranslational folding events, we wanted to incorporate BOF-Lys at different positions of the nascent peptide. Previously mentioned rules (chapter 3.2.1.2) about the positioning of the incorporation site downstream of a Gly or Ala could not be applied throughout the entire sequence of HemK CTD, as Gly is rare in HemK and Ala is mostly found inside secondary structures. Therefore, we did not follow this rule, but chose potential BOF-Lys incorporation sites in loop regions to avoid the disruption of secondary structures.

First, all selected positions for the BOF-Lys incorporation were screened by using translation time courses to determine the efficiency and rate of translation. To correlate the stopped-flow data with the translation time courses from chapter 3.3.3.1, identical translation conditions were used, including the presence of both N-terminal Atto655 and BOF-Lys. The time courses were analyzed by SDS-PAGE (Appendix Figure 2). Before considering translation rates with BOF-Lys at different positions, we tested whether the quantification of the time courses (chapter 5.8.3.1) using the fluorescence signal of the N-terminal Atto655 or internal BOF returns the same time course, which is exemplarily shown for BOF-Lys incorporation at the T120K, R131K and K159 residue (Figure 30).

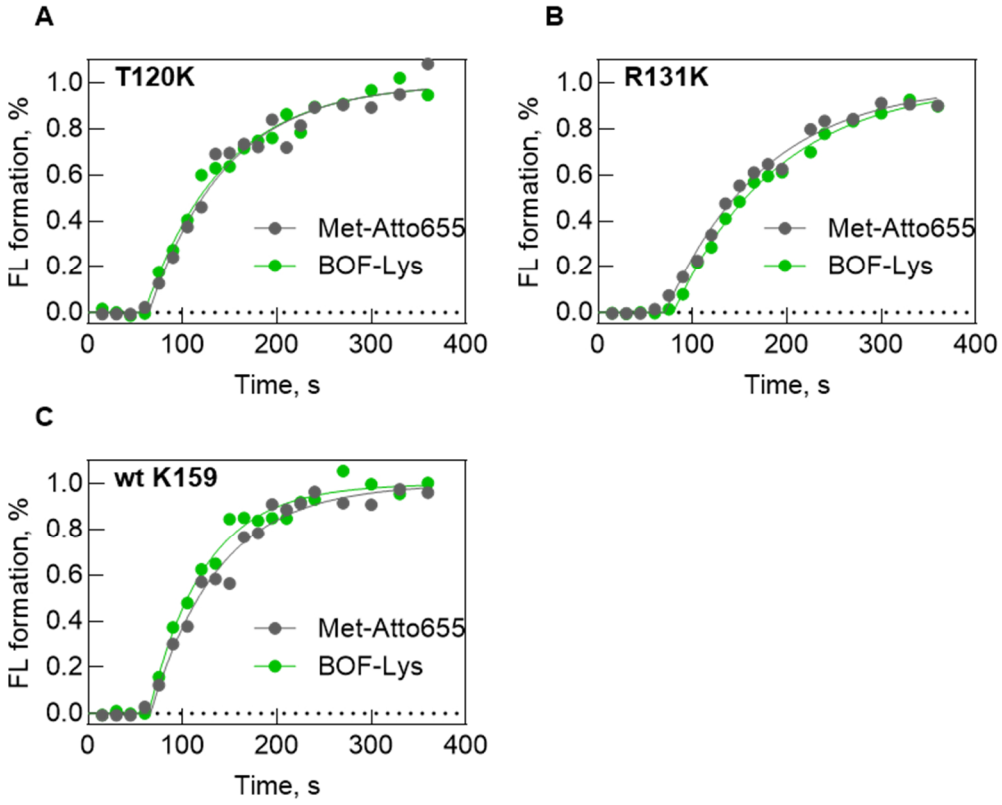



Figure 30. Comparison of the Atto655 and BOF signals in translation time courses. Example of the quantification and plotting of full-length product (FL) of both internal BOF signal and N-terminal Atto655 signal of translation time courses from Appendix Figure 2. Band intensities of the respective full-length product were quantified and normalized to 1 as described in chapter 5.8.3.1. Solid lines represent fits using a delay-single exponential function (equation 4).

The time courses for the two dyes are identical and are not split in two populations by considering the N-terminal or internal dye, indicating that the translation rate is the same independent of the signal origin (Figure 30). Additionally, taking into account that the identity of the N-terminal dye does not have an effect on the translation rate (chapter 3.2.2), all differences in translation rate arise from the incorporation of BOF-Lys at different positions. For the final analysis the internal BOF signal was chosen for quantification and was compared to the reference time course of HemK CTD with N-terminal BOF (73-277 in Figure 29).

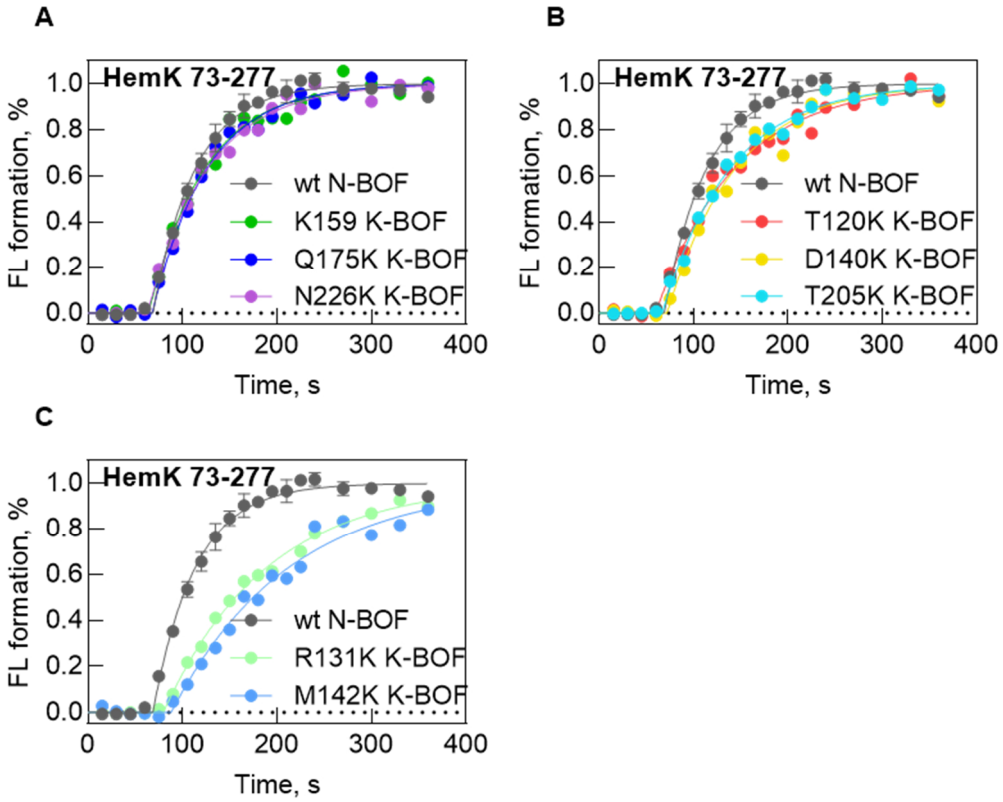



Figure 31. Effect of BOF-Lys incorporation at various positions on the translation rate. The band intensities of the respective full-length product (FL) from SDS-PAGE (Appendix Figure 2) were quantified and normalized to 1 according to chapter 5.8.3.1. Data points were fitted as a delay-single exponential function and compared to translation time course of HemK 73-277 (Figure 29).

Comparing the translation time courses of the CTD with different BOF-Lys incorporation sites (Figure 31) to HemK 73-277 wt with N-terminal BOF only (73-277 in Figure 29), allows to group the translation time courses into 3 categories. Category A (Figure 31A) shows no effect of the BOF-Lys incorporation on the translation rate. Category B (Figure 31B) has a slightly slower translation rate than the wt control. These constructs can be used further, but these results need to be discussed more carefully. In terms of sequential folding, they can be interpreted unambiguously if the folding rates are faster than the folding rate monitored by the label at the downstream incorporation site with an unchanged translation rate (category A). Category C (Figure 31C) has a clear decrease in translation rate with a pronounced delay in full-length product formation. These constructs were not used further, because connecting these kinetic data to the reference translation time courses would be unreliable.

### 3.3.3.3. Real-time folding of HemK CTD with various BOF-Lys positions

The constructs with BOF-Lys incorporation sites (Figure 32), chosen in the previous chapter (3.3.3.2), were used to measure time-resolved FRET in a cotranslational folding experiment as shown in chapter 3.2.3 according to chapter 5.9.

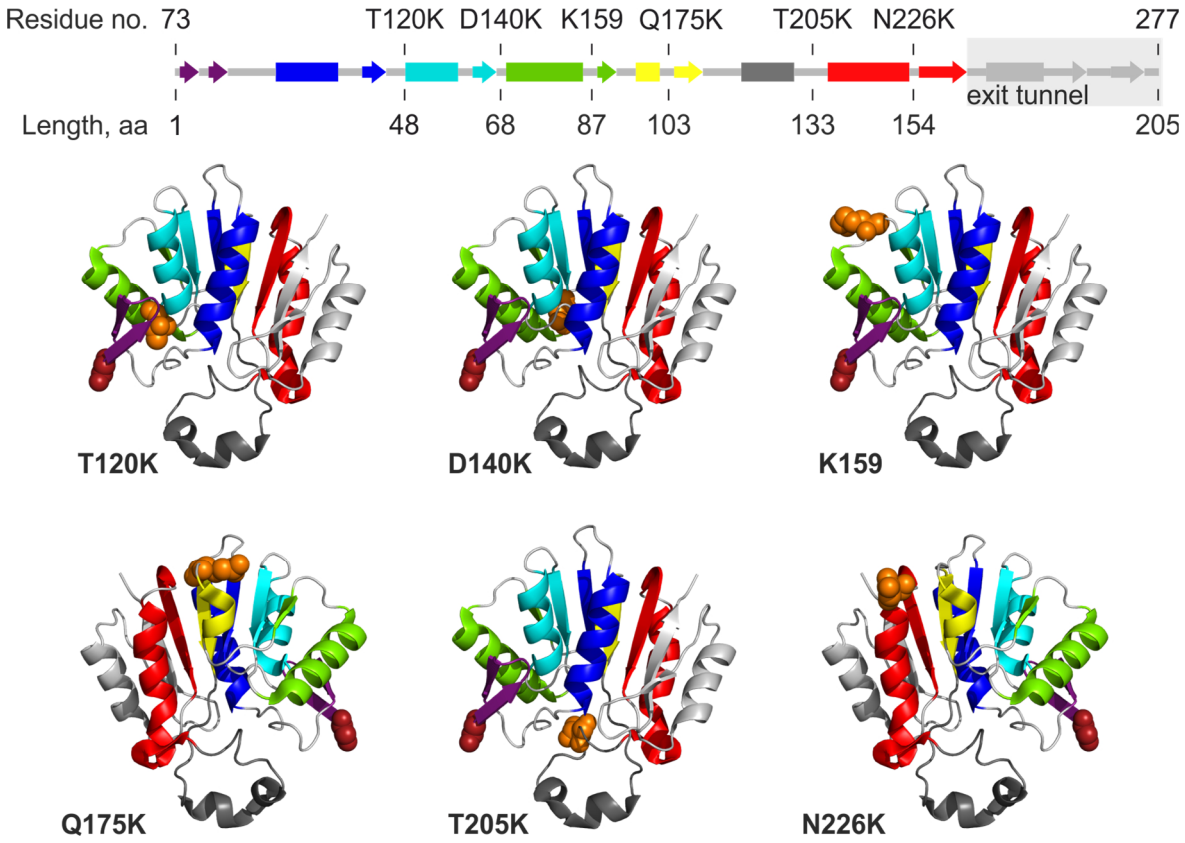



Figure 32. Schematic of HemK CTD secondary structure and the respective crystal structure (PDB: 2B3T). Color code is the same. Additionally, the positions of the BOF-Lys incorporation sites are indicated as the residue number and the length between the dyes (aa) in the HemK CTD. The polypeptide exit tunnel encompasses 40 aa. The position of the N-terminal and internal dye is indicated as red and orange spheres, respectively.

The FRET time courses (Figure 32) were evaluated by fitting to a delay-single exponential or delay-double exponential function (chapter 5.9).

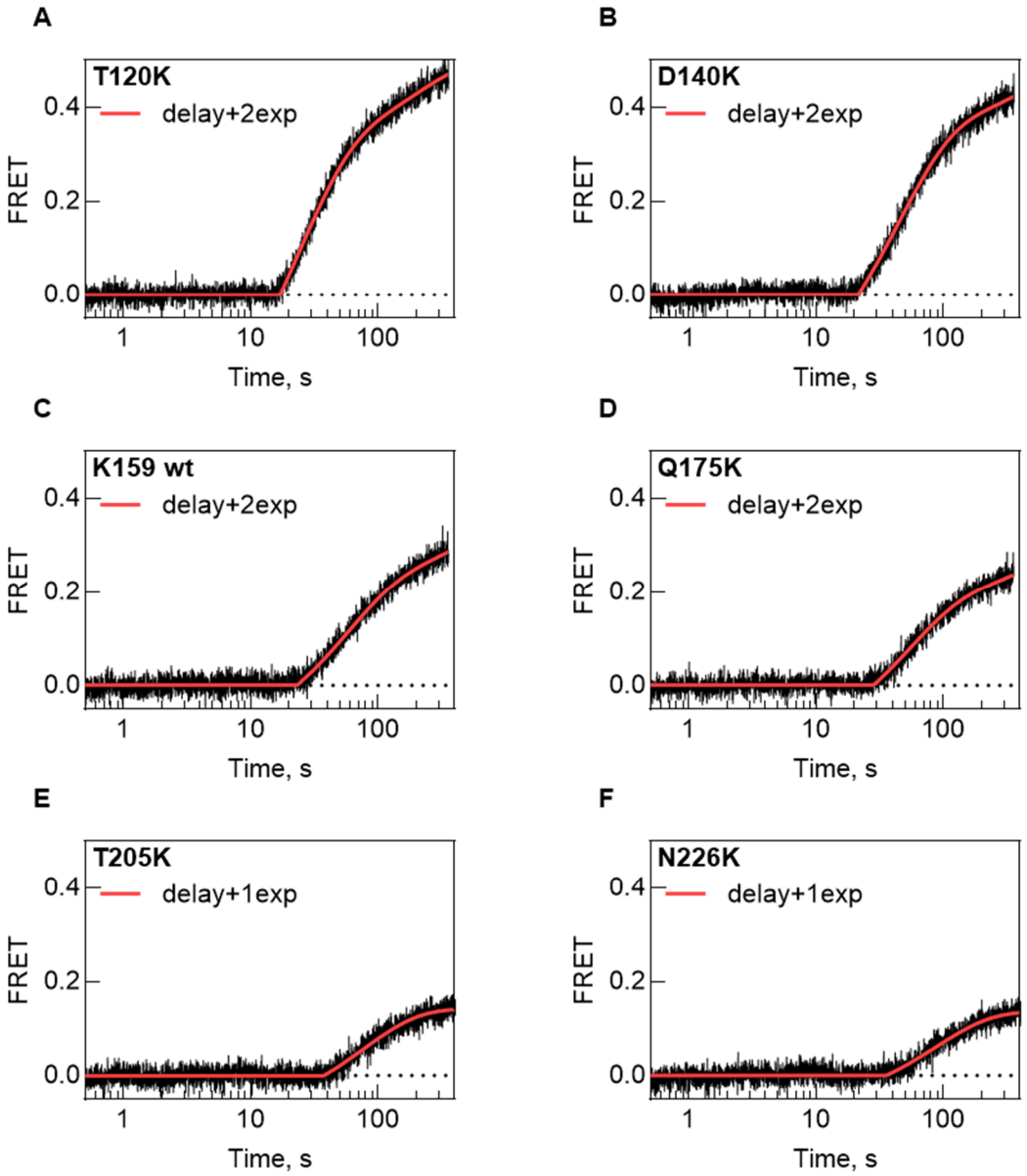



Figure 33. Cotranslational folding of HemK 73-277 with BOF-Lys incorporation at various residues in the polypeptide sequence. Real-time folding experiment observing FRET between the N-terminal Atto655-Met and internal BOF-Lys using HemK 73-277 wt or HemK 73-277 K159R variants as indicated. Experiments were performed according to chapter 5.9. FRET time courses were corrected for non-FRET fluorescence following equation 8. Red lines represent best fits using a delay-exponential function with one (equation 9) or two (equation 10) exponential terms as indicated.

Most of the FRET time courses, except the ones originating from HemK 73-277 K159R T205K and N226K constructs, required a delay-double exponential function (Table 5).

Table 5. Fitting of FRET time courses of HemK 73-277 with different Lys-BOF incorporation sites from Figure 33. Lifetimes $\tau_n$ were calculated according to equation 11 (chapter 5.9). n.d.: not determined, i.e., SEM is larger than fitted parameter.

| **Construct** | **$Amp_1$ ± SEM** | **$\tau_1$ , s ± SEM** | **$Amp_2$ ± SEM** | **$\tau_2$ , s ± SEM** |
|---|---|---|---|---|
| HemK 73-277 K159R T120K | 0.31 ± 0.01 | 39 ± 1 | 0.19 ± 0.01 | 215 ± 36 |
| HemK 73-277 K159R D140K | 0.36 ± 0.02 | 65 ± 2 | n.d. | n.d. |
| HemK 73-277 wt (K159) | 0.22 ± 0.03 | 77 ± 6 | n.d. | n.d. |
| HemK 73-277 K159R Q175K | 0.18 ± 0.02 | 80 ± 5 | n.d. | n.d. |
| HemK 73-277 K159R T205K | 0.14 ± 0.01 | 115 ± 2 | - | - |
| HemK 73-277 K159R N226K | 0.14 ± 0.01 | 124 ± 3 | - | - |

As the translation time of HemK 73-277 wt mRNA is 114 ± 2 s (chapter 3.3.3.1, Table 4), the lifetime of the second phase ($\tau_2$) of all FRET time courses is longer than the translation time and is thus posttranslational. Therefore, $\tau_2$ is not considered further with regard to cotranslational folding. Consequently, $\tau_1$ is the folding time $\tau_{fold}$. In order to find the nascent chain length that allows the folding of an element between two dyes in a certain construct, its folding time $\tau_{fold}$ is assigned to the closest translation times $\tau_{trans}$ of the HemK CTD truncations (Table 4). This sets the limits to the peptide length for a folding event.

Table 6. Comparison of folding times for the constructs with different BOF-Lys incorporation sites to translation times of HemK CTD constructs. Folding times $\tau_{fold}$ are taken from $\tau_1$ in Table 5. Translation times $\tau_{trans}$ are taken from Table 4 with the respective construct length.

| **BOF-Lys residue** | **Peptide length between dyes, aa** | **$\tau_{fold}$ ± SEM, s** | **closest $\tau_{trans}$ ± SEM, s** | **Bordering constructs by $\tau$ (number of aa)** |
|---|---|---|---|---|
| T120K | 48 | 39 ± 1 | 32 ± 3<br>58 ± 2 | HemK 73-132 (60 aa)<br>HemK 73-152 (80 aa) |
| D140K | 68 | 65 ± 2 | 58 ± 2<br>70 ± 2 | HemK 73-152 (80 aa)<br>HemK 73-172 (100 aa) |
| wt K159 | 87 | 77± 6 | 74 ± 2<br>80 ± 2 | HemK 73-192 (120 aa)<br>HemK 73-212 (140 aa) |
| Q175K | 103 | 80 ± 5 | 80 ± 2.<br>91 ± 2 | HemK 73-212 (140 aa)<br>HemK 73-232 (160 aa) |
| T205K | 133 | 115 ± 2 | 114 ± 2 | HemK 73-277 (205 aa) |
| N226K | 154 | 124 ± 3 | 114 ± 2 | HemK 73-277 (205 aa) |

For better visualization, the folding times were plotted together with the translation times against the corresponding peptide length in Figure 34. The peptide length of the particular folding event was set to be in the middle of the limits determined by the translation time courses (Table 6).

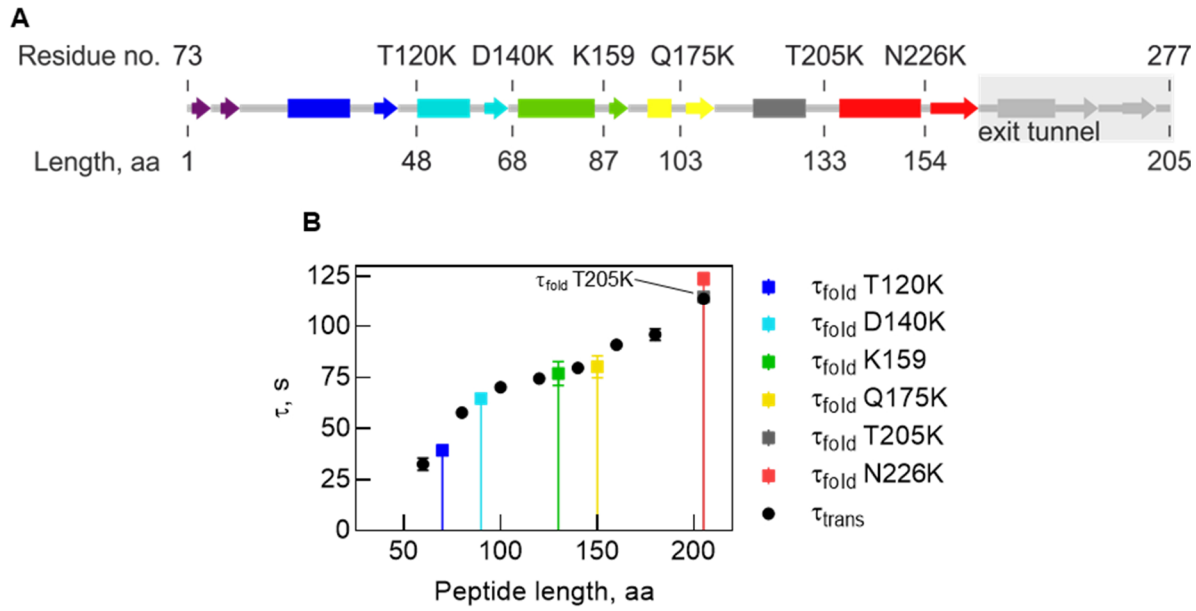



Figure 34. Comparison of folding times and translation times of HemK CTD constructs. (A) Schematic of HemK CTD secondary structure. The positions of the BOF-Lys incorporation sites are indicated at the residue number and the length between the dyes (aa) in the HemK CTD. The polypeptide exit tunnel encompasses 40 aa. (B) Translation times ($\tau_{trans}$; black circles) of HemK CTD constructs of different length. Folding time ($\tau_{fold}$; colored squares) of HemK 73-277 plotted against the average length of the bordering constructs (Table 6). The folding time of the T205K incorporation site is hidden behind the black dot at the very right.

The folding time increases with the distance between the N-terminal Atto655 and the BOF-Lys incorporation site. The folding of the T120K variant (48 aa between the dyes) is indicated to happen when the nascent chain reaches a total length between 60 and 80 aa and the following elements are added as they emerge from the polypeptide exit tunnel. The element until D140K (68 aa between the dyes) folds between 80 and 100 aa. The structural element containing K159 (87 aa between the dyes) folds as the chain length is between 120 and 140 aa and the structure containing Q175K (103 aa between the dyes) folds between 140 and 160 aa. The FRET time courses of the T205K (133 aa between the dyes) and N226K (154 aa between the dyes) construct coincides with the translation time course of the HemK CTD (73-277), suggesting that the structural elements following Q175 fold at the end of synthesis of the full-length CTD.

All the constructs that were earlier assigned to category B (T120K, D140K and T205K; small, but visible effects on the translation rate by BOF-Lys incorporation; chapter 3.3.3.2) can be interpreted with regard to sequential folding. As mentioned before, as long as the folding times increase with a later incorporation site, even when the BOF-Lys incorporation induces a slower translation rate, the folding occurs sequentially.

### 3.3.3.4. Determination of folding peptide length from FRET amplitudes

We further validated our conclusions by comparing the amplitude of the FRET change during the translation of mRNA that was truncated every 20 aa from the C-terminus (Figure 35A). By incorporating the BOF-Lys at the same position, it is possible to capture the minimal length of the folding event independent of the translation rate.

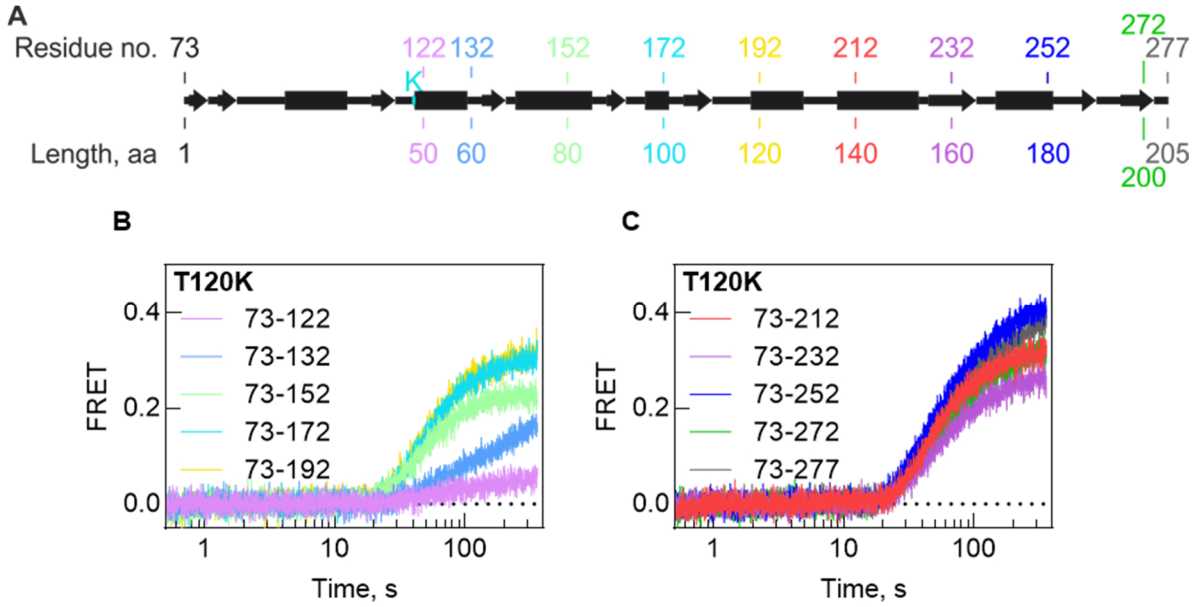



Figure 35. Cotranslational folding of HemK 73-277 T120K constructs of different length. (A) Secondary structure scheme of HemK CTD. Truncations are color-coded in the same way as FRET time courses shown in B and C. The BOF-Lys incorporation site at residue number 120 is indicated with 'K' in cyan. (B, C) Real-time folding experiment monitoring FRET between N-terminal Atto655-Met and internal BOF-Lys using truncated HemK constructs as indicated. Experiments were performed as described in chapter 5.9. FRET time courses were corrected for non-FRET fluorescence according to equation 8.

The amplitude of the FRET signal increases with increasing nascent peptide length (Figure 35B, C). In order to identify co- and posttranslational processes, the FRET traces were fitted with a delay-exponential function with one or two exponential terms as described for Figure 33 and Table 5 in chapter 3.3.3.3. The lifetime of the second amplitude $\tau_2$ is very large

and variable (not shown), with the shortest $\tau_2$ being 249 s. As the translation time of HemK CTD wt is 114 ± 2 s (Table 4), the processes reflected by the 2nd exponential phase must occur posttranslationally and were excluded from further analysis. Figure 35 reveals that the signal change onset is very similar for all truncations and can be fitted to a delay time between 19 and 24 s. FRET can only be observed after BOF-Lys is incorporated into the peptide chain. The similarity between the delay times shows that independent of the truncation, the incorporation of BOF-Lys occurs at the same point (in time and protein sequence) during mRNA translation. The folding time $\tau_{fold}$ (delay + exponential; Table 7) of the two shortest truncations 73-122 (105 s) and 73-132 (75 s) differ from and appear to be larger than the folding times of all longer constructs, which are all in the range of 59-67 s (Table 7).

Table 7. Comparison of FRET signal change of HemK 73-277 T120K truncations. Folding time $\tau_{fold}$ was calculated according to equation 11. P value of the amplitudes was calculated with the unpaired t test tool from GraphPad using HemK 73-277 as reference. N=2. n.s.: not significant. *: significant.

| **Construct** | **Length, aa** | **$\tau_{fold}$, s ± SEM** | **Amplitude ± SEM** | **Significance (P value)** |
|---|---|---|---|---|
| HemK 73-122 | 50 | 105 ± 6 | 0.05 ± 0.001 | * (0.01) |
| HemK 73-132 | 60 | 75 ± 15 | 0.07 ± 0.03 | * (0.03) |
| HemK 73-152 | 80 | 62 ± 1 | 0.22 ± 0.001 | n.s. (0.07) |
| HemK 73-172 | 100 | 62 ± 2 | 0.27 ± 0.01 | n.s. (0.47) |
| HemK 73-192 | 120 | 62 ± 2 | 0.28 ± 0.01 | n.s. (0.70) |
| HemK 73-212 | 140 | 60 ± 2 | 0.27 ± 0.01 | n.s. (0.47) |
| HemK 73-232 | 160 | 67 ± 3 | 0.23 ± 0.02 | n.s. (0.17) |
| HemK 73-252 | 180 | 59 ± 3 | 0.30 ± 0.03 | n.s. (0.81) |
| HemK 73-272 | 200 | 65 ± 4 | 0.24 ± 0.03 | n.s. (0.30) |
| HemK 73-277 | 205 | 63 ± 3 | 0.29 ± 0.02 | - |

This indicates that FRET during translation of the shorter mRNA originates from slow structural rearrangements that are probably off-pathway. In contrast, with the longer constructs, the compaction of the nascent chain occurs at the same time and hence, in a single step at a similar point during the translation. To validate this finding independently of the observed rate, the amplitudes of the FRET signal were analyzed and plotted as a bar-chart (Figure 36).

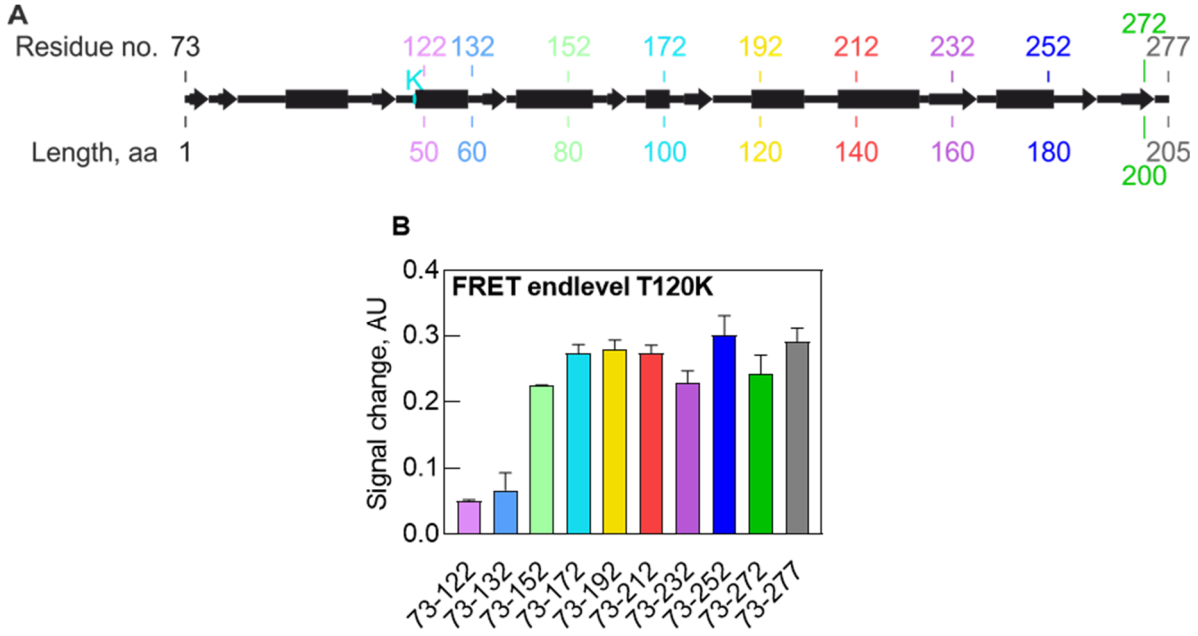



Figure 36. Cotranslational folding of HemK 73-277 T120K depending on the construct length. (A) Secondary structure scheme of HemK CTD labeled with the residue of the truncation position and the corresponding length (aa). The BOF-Lys incorporation site at residue 120 (48 aa between the dyes) is indicated with 'K' in cyan. (B) Plotted are the fitted parameters of the first amplitude ± SEM, extracted by delay-exponential fitting (Figure 35; Table 7).

The FRET change of the HemK 73-122 truncation (Figure 36) serves as an incorporation control for BOF-Lys. This construct codes for 50 aa, i.e., there are only 2 aa added after the dye incorporation. We can assume that in this case the nascent peptide resides in the polypeptide exit tunnel, where only limited folding events can occur, such as α-helix formation, and no long-range tertiary interactions are formed. The observation that the FRET amplitude is small when the tertiary structure formation is disallowed supports the notion that the following FRET changes arise from the 3D compaction of the nascent chain outside the polypeptide exit tunnel when more of the protein is translated. In order to interpret the amplitude with regard to the transition from the extended to a compacted peptide conformation, the values with their respective errors (Table 7) were submitted to the unpaired t test calculator from GraphPad in order to calculate the P value, using the null hypothesis that the amplitude

of the signal change of all the truncations compared to the HemK 73-277 is the same. The threshold to reject the null hypothesis is <0.05.

The FRET amplitudes of HemK 73-132 and HemK 73-122 are significantly smaller than the amplitude of HemK 73-277 (Table 7; P value of <0.05). The FRET amplitude for the longer constructs are the same, within the limits of the error. Therefore, the transition from the unfolded to the compacted state of the first structural element on the ribosome occurs between 60 and 80 aa (HemK 73-132 and HemK 73-152). This confirms the results obtained in the previous chapter (3.3.3.3).

### 3.3.4. Posttranslational compaction of HemK CTD

At the end of translation, the C-terminal part of the protein is still inside the polypeptide exit tunnel and cannot participate in folding. Additionally, the ribosome can partially destabilize the nascent chain (Cabrita et al., 2016; Liutkute et al., 2020a). Therefore, it is possible that before the protein is released from the ribosome it forms a native-like intermediate that is only partially compacted. To test whether the CTD adopts a more compact state in solution than on the ribosome, or in other words, whether there are posttranslational folding events or structural rearrangements in solution, the nascent chain was released from the RNC and FRET was monitored between the N-terminal Atto655 and BOF-Lys at residues 226 or 175.

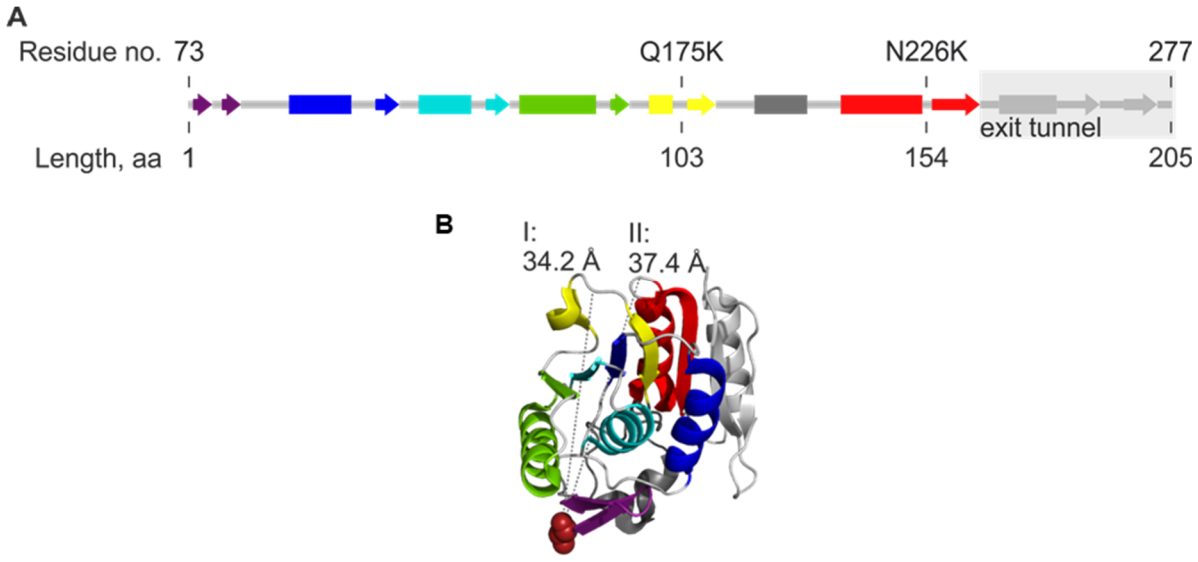



Figure 37. FRET change upon release of HemK CTD from the ribosome. (A) Scheme of the secondary structure of HemK CTD with BOF-Lys incorporation sites. (B) Crystal structure of HemK CTD (PDB: 2B3T). Distances are measured between the Cα atoms of the N-terminal Met (for approximate Atto655 position) and Lys-BOF incorporation site at residue (I) 175 and (II) 226.

In both cases, the distance between the amino acids that are labeled with the dyes is very similar in the crystal structure (Figure 37A and B). However, the residue 226 could still be inside the polypeptide exit tunnel at the end of HemK synthesis. To test whether posttranslational folding involves adjustments of the C-terminal part of the protein, FRET release experiments were performed in the fluorimeter as described in chapter 5.10. The FRET change upon release was calculated from the raw fluorescence time course (donor + acceptor; Figure 38A and C).

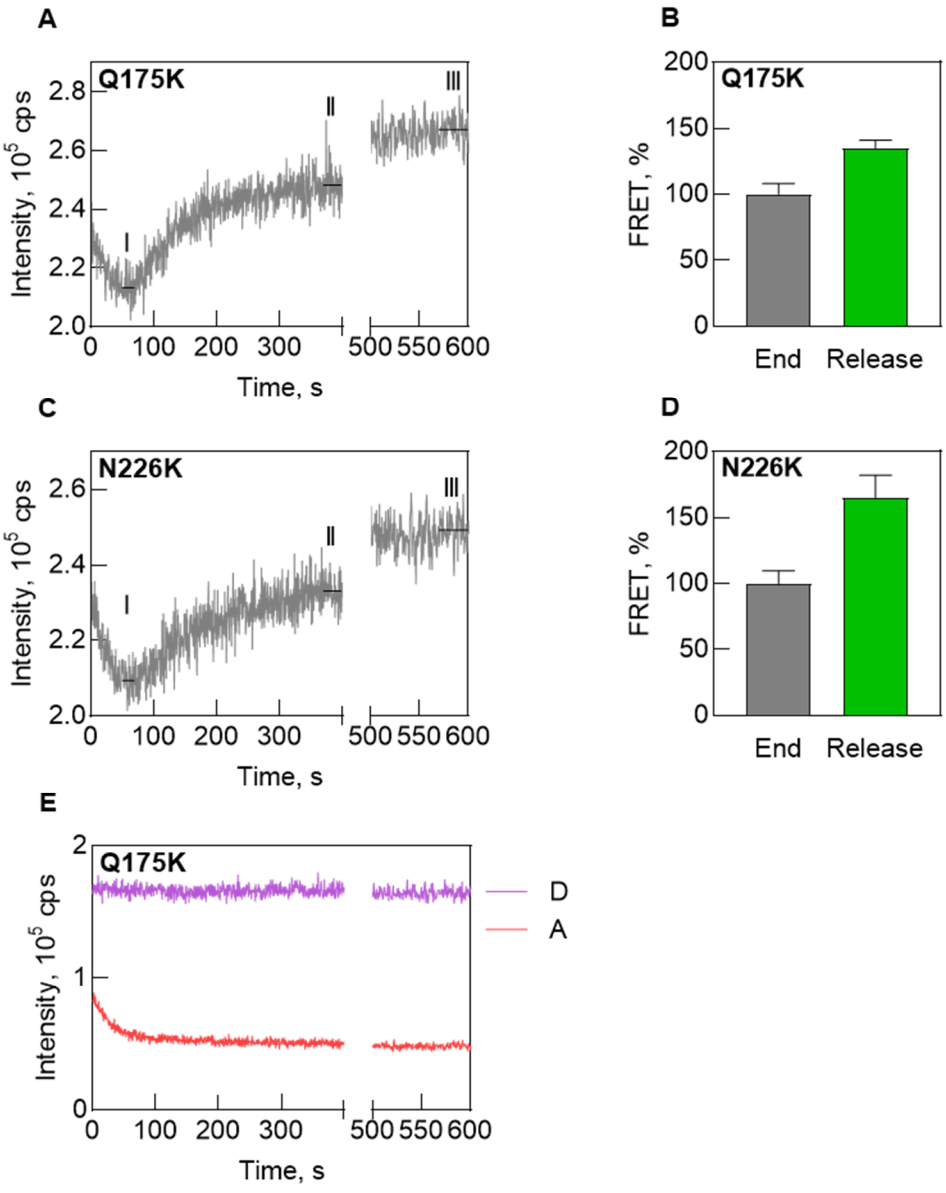



Figure 38. FRET change upon release of HemK CTD from the ribosome. (A) and (C) Example time courses with BOF-Lys at different residues in HemK CTD as indicated and N-terminal Atto655-Met. Time window from 0 to 400 shows fluorescence change upon peptide synthesis. For normalization, the lowest value at point (I) is set to 0% FRET and (II) to 100% FRET. These values are used to calculate the end level after peptide release from the ribosome (III) in (B) and (D). Reported values are mean ± SD (N=3). (E) Donor-only and acceptor-only control for the release of the Q175K variant.

The fluorescence of the donor-only and acceptor-only sample does not increase after peptide release (Figure 38E), indicating that the increase of the sample containing both donor and acceptor originate from FRET (Figure 38B and D). The FRET signal increases for both constructs (+35 ± 10% for Q175K and +65 ± 19% for N226K) after peptide release, suggesting additional compaction of the protein upon release from the ribosome. Furthermore, the C-terminal tail that is constrained by the ribosome appears to be contributing to the final compaction of the CTD after release, as the peptide with the incorporation site at residue 226 has a larger FRET increase than 175 after release. However, submitting the values of the additional FRET increase after release to an unpaired t test (GraphPad) returns a P value of 0.08 and therefore no significance for the difference. However, in general, the release of HemK CTD from the ribosome demonstrate that the peptide release leads to the final folding step of the HemK CTD and shows that the ultimate folding of the entire domain is finalized in solution.

# 4. Discussion

## 4.1. Folding mode of HemK CTD

Proteins can fold stepwise through the formation of intermediates or in a concerted fashion in a single step folding reaction. In order to understand the principles of protein folding and to understand the role of the ribosome, it is often instructive to compare protein folding on the ribosome with that in solution, obtained by refolding experiments with purified proteins. Until now, there is no information of how HemK CTD refolds in solution. However, there are several examples available of proteins with similar Rossmann fold-like or similar α/β topologies in solution. These proteins are CheY, flavodoxin, and both domains of phosphoglycerate kinase. Equilibrium unfolding / refolding studies indicate that they are able to fold in a three state mode, i.e., transition from unfolded to native state through the formation of an intermediate state, with relatively fast reaction kinetics in the ms to s time range (reported in Garbuzynskiy et al., 2013). However, further investigations employing pre-steady state techniques revealed that for CheY and flavodoxin additional transient off-pathway intermediates are formed, which is an indication that these proteins are prone to misfolding (Kathuria et al., 2008; Stagg et al., 2010). The formation of misfolded intermediates could be prevented during cotranslational folding on the ribosome, as the vectorial appearance of the peptide allows only a limited number of conformations at any given time. The FPA of HemK CTD shows that folding on the ribosome is more complicated than the simple three state mode folding mechanism inferred from refolding experiments in solution, as it folds cotranslationally through the formation of several intermediates on the ribosome (Figure 21). Therefore, it is possible that the differences between the folding pathways on and off the ribosome reduce the formation of misfolded intermediates that are observed in solution.

The FPA is a suitable approach to estimate the number and the peptide length required for a folding event. The FPA analysis is sensitive towards folding and structural rearrangements of the peptide both inside and outside of the polypeptide exit tunnel (Agirrezabala et al., 2022; Farias-Rico et al., 2018; Kemp et al., 2019; Liutkute et al., 2020a; Nilsson et al., 2015). As both the folding inside and outside the polypeptide exit tunnel can occur simultaneously with increasing length of the nascent chain, it becomes increasingly difficult to interpret single tension events that are observed with the FPA. One way to solve this problem is to employ a

differential analysis of the FPA by comparing two profiles in the absence or presence of point mutations at key positions, which play a role in secondary and tertiary structure formation (Liutkute et al., 2020a). Similarly, here we compare the FPA of the coding sequence of HemK (NTD + CTD) with HemK CTD. This allows the direct assignment of folding events that correspond to the deleted part of the NTD because tension peaks can be directly compared for their shape and intensity. Comparison of the HemK and HemK CTD FPA shows that the overall shape and tension distribution is the same, which suggests that the effect of the NTD on the folding of HemK CTD is small and thus the folding of the CTD seems to proceed through the same folding pathway independently of the NTD. There are only two sections (Figure 21), II (peaking at 109 aa) and V (with midpoint at 161 aa), which are affected by the removal of HemK NTD, with a pronounced reduction in tension in both areas for the isolated CTD. The additional tension in the HemK FPA could reflect interactions of the NTD with the emerging sequence of the CTD. Alternatively, it is possible that the transition of HemK NTD from a native-like structure to the fully natively folded domain causes the increase in tension.

At a linker length of 40 aa, the HemK NTD has fully emerged from the ribosome. However, it does not acquire a fully natively folded, but rather forms a native-like conformation, which is induced by ribosome contacts (Liutkute et al., 2020a). At this length, the NTD adopts its native structure only after release from the ribosome (Holtkamp et al., 2015; Liutkute et al., 2020a; Mercier and Rodnina, 2018). The additional synthesis of parts of the CTD could serve as a linker that places the NTD further away from the ribosome surface and abolishes the destabilizing effect on the protein structure as it was demonstrated for the FLN5 domain (Cabrita et al., 2016). FLN5 acquires its native structure on the ribosome, but only when 67 additional aa of FLN6 are produced, although only 30 aa are accommodated in the polypeptide exit tunnel, by lifting the destabilization that is induced by ribosome contacts with the ribosomal protein L24 (Ahn et al., 2022; Cabrita et al., 2016). The length of the linker that is needed to alleviate HemK NTD destabilization at the ribosome surface is difficult to predict, as the polypeptide exit tunnel can accommodate 30 unfolded and up to 70 aa of a highly structured peptide (Agirrezabala et al., 2022; Ahn et al., 2022; Cabrita et al., 2016; Holtkamp et al., 2015). Additionally, the relative orientation of the NTD and CTD on the ribosome or the conformation of the CTD may affect the length of the peptide linker. Despite the difficulties of interpreting the FPA in detail, the N-terminal peaks can be interpreted directly because the ‘folding background’ that arises from folding outside the exit tunnel is absent or low (Kemp et al., 2019; Liutkute et al., 2020a). As section I and II of the FPA (Figure 21) involve folding

event with up to 40 aa of HemK CTD that have passed the constriction site, they can be assigned to folding events of secondary structures inside the polypeptide exit tunnel, as no peptide has emerged from the ribosome at this point (Kemp et al., 2019; Liutkute et al., 2020a).

### 4.2. Folding inside the polypeptide exit tunnel

A more detailed insight at the conformational dynamics of the N-terminal part of HemK CTD, provided by a combination of ensemble PET and PET-FCS measurements (Figure 28; chapter 3.3.2) indicates that the peptide undergoes several compaction / relaxation events inside the polypeptide exit tunnel during the synthesis of the nascent chain. The earliest event that we capture in this study is the mobility restriction of the first 13 aa of HemK CTD. Although Trp-PET suggests that the N-terminal dye and Trp78 (5 aa downstream of N-terminal dye) are in close proximity, the peptide itself does not appear to be flexible (73-85 in Figure 28). Structurally, it is possible that with the length of 13 aa, the fluorophore is located close or even at the constriction site, which is 30 Å away from the PTC. About 20 aa in an α-helical and 9 aa in an extended conformation can be accommodated before the constriction site, but only unfolded peptide can pass through it (Bhushan et al., 2010; Lu and Deutsch, 2005).

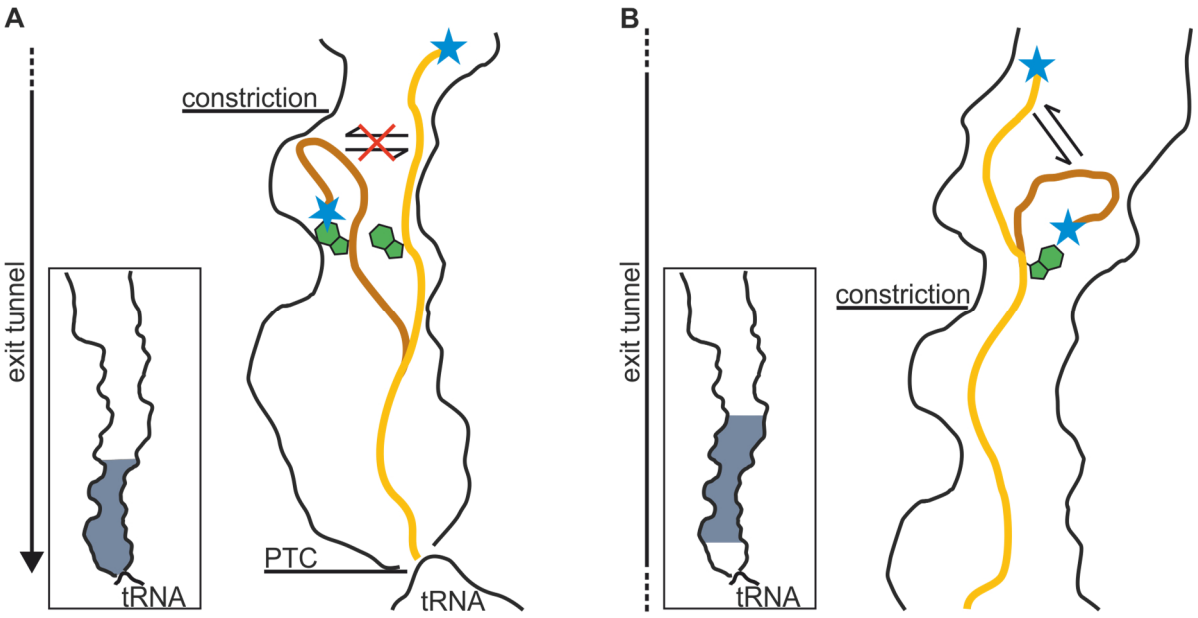



Figure 39. Continues on following page.

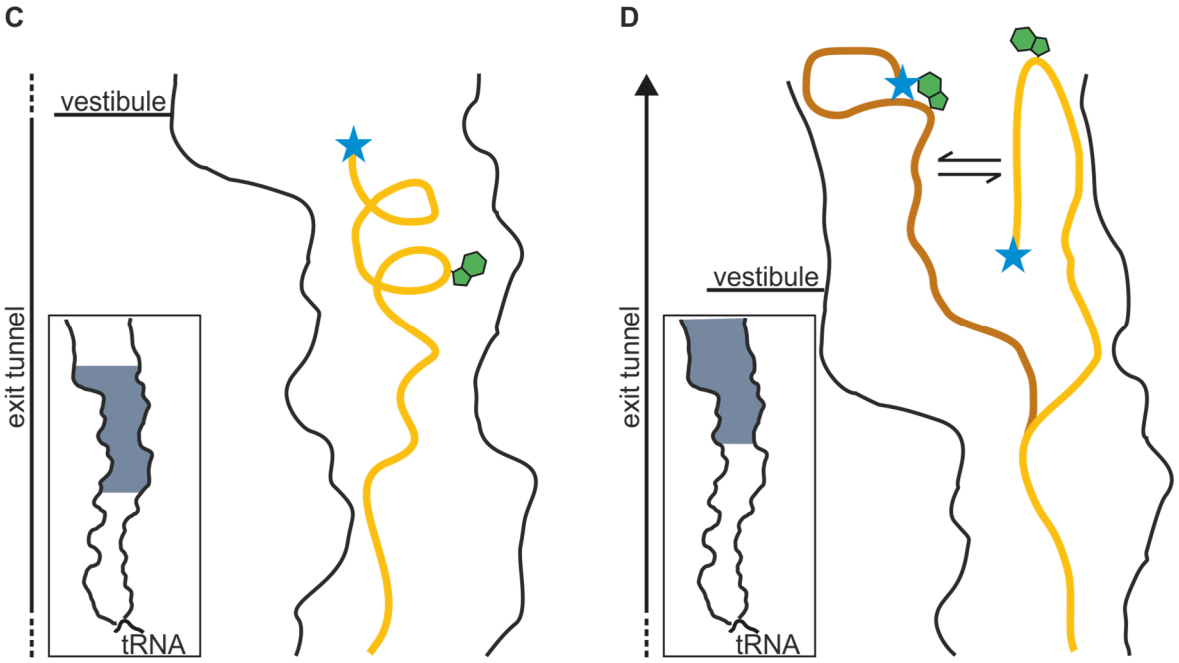



Figure 39. Folding model of HemK CTD inside the polypeptide exit tunnel. The nascent chain conformation that does or does not give rise to PET is shown in brown or yellow, respectively. PET originates from van der Waals contact of the N-terminal fluorophore (blue star) and Trp (green) that is 5 aa away. Arrows on the axis of each panel indicate the end or beginning of the tunnel, whereas a dashed line indicate that a segment in the middle of the exit tunnel is shown. The box in each panel shows the entire polypeptide exit tunnel with the segment that is emphasized in grey. (A) Model based on data of HemK 73-85. (B) Model based on data of HemK 73-92. (C) Model based on data of HemK 73-112. (D) Model based on data of HemK 73-122, 73-132 and 73-277.

Depending on the extension of the peptide inside the tunnel, it is possible that two populations of conformations are formed, one that has the dye constrained at the constriction site and therefore not interacting with the intramolecular Trp, and another that adopts a compact conformation with the dye coming into the proximity of Trp (Figure 39A). The high PET efficiency does not seem to be unusual for this area of the exit tunnel, as a 14 aa-long peptide of CspA also had a high PET value inside the exit tunnel (Agirrezabala et al., 2022).

At a length of 20 aa, the N-terminus of the HemK CTD peptide is probably located at a narrow section of the polypeptide exit tunnel (Figure 39B). This could explain the relatively low Trp-PET efficiency and Trp-PET dynamics (73-92 in Figure 28). With a peptide length of 40 aa, the N-terminal part of the nascent chain has moved beyond the constriction site and is located in a wider part of the tunnel that allows folding of α-helical structures, including simple α-helices, but also α-turns (Agirrezabala et al., 2022; Bhushan et al., 2010; Kosolapov and Deutsch, 2009). Structural data on CspA shows that the nascent protein folds into an α-helix

past the constriction site (Agirrezabala et al., 2022). Thus, at the nascent peptide length of 40 aa of HemK CTD, the dye and Trp should not have space restrictions imposed by the exit tunnel that limit the peptide mobility. And yet the overall PET efficiency and dynamics of the nascent chain as indicated by PET-FCS are relatively low (73-112 in Figure 28). Thus, it is reasonable to conclude that the mobility restriction are induced by conformations of the peptide itself. It is tempting to speculate that the N-terminus could acquire an α-helical conformation, as residue 1 and 6 (positions of dye and Trp) would face opposite directions in an α-helix and could not form a stable dye-Trp complex (Figure 39C), but assessing the exact conformation is difficult. Further translation of only 10 additional aa, i.e., to a total length of 50 aa, allows the N-terminal peptide to acquire the same dynamic folded state as upon the synthesis of all 205 aa of HemK CTD (73-122, 73-132 and 73-277 in Figure 28) with a highly dynamic structure and contacts between the dye and the Trp (Figure 39D). Structural rearrangements of the nascent peptide at this length are also supported by the FPA data of HemK CTD (Figure 21) as they coincide with the tension in area II.

In general, the method of PET-FCS turns out to be invaluable to get a detailed view into the dynamics of the nascent chain inside the polypeptide exit tunnel. In particular in combination with ensemble PET measurements (Trp-PET efficiency) it is possible to distinguish between conformations that are forced upon the nascent peptide by space restriction (13 aa, close or at the constriction site) and mobility restrictions of the peptide that originate from folding (40 aa). PET-FCS has been employed before to investigate protein or peptide dynamics in solution (Luitz et al., 2017; Neuweiler et al., 2007; Rajab et al., 2021). They all have in common that removing Trp by substitution reduces the amount of fitting parameters needed to represent the ACF, allowing the direct assignment of certain molecular motions to a relaxation time. On the ribosome this is not the case due to the interference by the PET-active guanines of the ribosome that surrounds the fluorophore at every position inside the polypeptide exit tunnel and to some extent even outside of the ribosome. In a previous study, extensive modelling had to be employed in order to extract elemental rates of fundamental peptide motions on the ribosome (Liutkute et al., 2020a). In this study, by considering the Y-intercept of the ACF, a more general approach was chosen to estimate the extent of quenching that originates from dynamic peptide motion. Of note, one potential limitation of this approach is the uncertainty of the ACF Y-intercept calculation due to a large number of fitting parameters, and in particular the presence of the triplet state. The triplet state reduces the amount of available dye molecules for fluorescence and therefore by extension the amount of detectable Trp-quenching (reviewed in

Krichevsky and Bonnet, 2002). Additionally, there seems to be an effect of the ribosome on the triplet state itself. FCS measurements on Atto655-labeled peptide yield a triplet state relaxation time of 2 µs for Atto655 (Luitz et al., 2017), but on the ribosome it is 40 µs (Liutkute et al., 2020a). While the estimations for the triplet state relaxation times in this study are in the same order of magnitude as in Liutkute et al. (2020a), the range of the exact values, ranging from 31 to 56 µs, is relatively large (Table 2). For future experiments it would be helpful to know the exact way the ribosome is modulating the triplet state of the fluorophore. This would reduce the uncertainty of PET-FCS experiments on the ribosomes and refine the technique further.

### 4.3. Folding outside of the polypeptide exit tunnel

After the N-terminus emerges from the polypeptide exit tunnel, the compaction of the peptide progresses outside of the ribosome. By incorporating a second fluorophore and measuring FRET, the folding of the peptide that is located between the two dyes was observed (Figure 40).

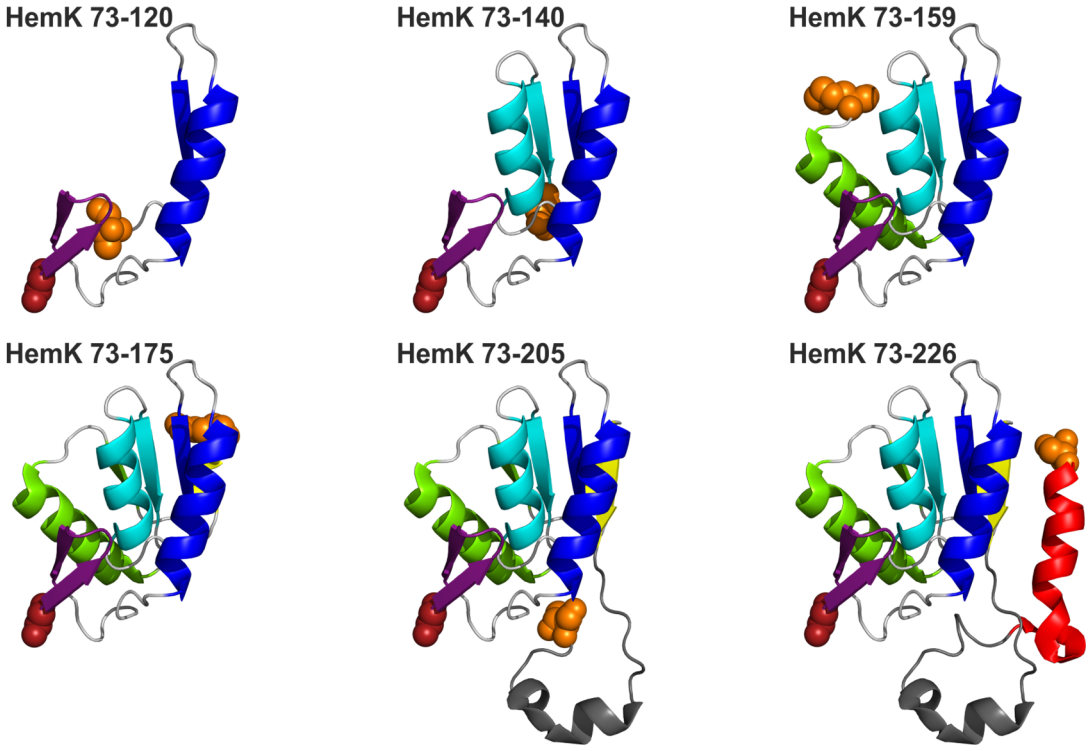



Figure 40. Solution structure of HemK CTD elements tested in cotranslational folding experiments monitoring FRET (PDB: 2B3T). Elements are defined by the position of the internal BOF-Lys; the downstream sequence is not shown, as only the distance between the N-terminal dye (red) and BOF-Lys (orange) is monitored.

Our FRET results suggest that compaction of the nascent peptide continues outside the ribosome tunnel step-by-step. The smallest fragment of HemK CTD that was examined is 48 aa long (Figure 40; HemK 73-120). As the folding of this part occurs between the translation of 60 and 80 aa, it is likely that 12-32 additional aa after the BOF-Lys incorporation have to be synthesized to allow folding of the 48 aa long peptide of HemK CTD (Table 8). Similar calculations for reporters at different positions of the protein show that the length of the linker that has to be synthesized for the folding to happen increases with the size of the folding unit (Table 8).

Table 8. Summary of sequential folding of HemK CTD. Given are the length of the observed folding segment that is determined by the BOF-Lys incorporation site, the range of peptide length that allows the folding and by subtracting both values the number of amino acids that have to be synthesized right after the BOF-Lys incorporation site to allow folding.

| **BOF-Lys position** | **Length of folding segment, aa** | **Peptide length for folding event, aa** | **Additional peptide after BOF-Lys incorporation for folding event, aa** |
|---|---|---|---|
| T120K | 48 | 60-80 | 12-32 |
| D140K | 68 | 80-100 | 12-32 |
| wt K159 | 87 | 120-140 | 33-53 |
| Q175K | 103 | 140-160 | 37-57 |
| T205K | 133 | 205 | 72 |
| N226K | 154 | 205 | 72 |

For shorter folding segments, the linker is relatively short and would support folding of the nascent peptide inside the polypeptide exit tunnel. Notably, at the time of BOF-Lys incorporation at residue 120, which is the shortest folding element studied by FRET, the N-terminal β-hairpin has attained a compact and dynamic conformation, as shown by PET experiments (chapter 3.3.2). Assuming that the downstream helical segment also adopts a dynamic compact conformation, as observed for HemK NTD and CspA (Agirrezabala et al., 2022; Holtkamp et al., 2015; Liutkute et al., 2020a; Mercier and Rodnina, 2018), folding of the 73-120 segment is likely to occur inside the polypeptide exit tunnel. In fact, smaller tertiary structures can be accommodated inside the exit tunnel, such as the 29 aa ADR1α domain (Nilsson et al., 2015; Wruck et al., 2021). The first structural element with 48 aa of the CTD is larger than ADR1α, but it would be possible to accommodate the structural element at the vestibule with its diameter of roughly 20 Å, in particular considering the highly dynamic nature of the folded structure and the fact that the exact conformation of this intermediate and the orientation of the secondary structures relative to each other inside the vestibule is unknown.

It is noteworthy that only the C-terminal part of the structural element (location of BOF-Lys) would need to be at the vestibule, whereas the N-terminal part can already have emerged from the exit tunnel. Similar conclusions were made for HemK NTD that folds through partially compacted high-FRET intermediates that form at the early stages of the translation of HemK NTD and reside inside the polypeptide exit tunnel (Holtkamp et al., 2015; Mercier and Rodnina, 2018). However, folding of HemK NTD seems to occur sequentially, with more compaction depending on the translated segment (Holtkamp et al., 2015; Liutkute et al., 2020a; Mercier and Rodnina, 2018), whereas folding of the first 48 aa of HemK CTD occurs in a single folding step at a specific length as shown by the FRET measurements with the HemK CTD truncations (Figure 35; Table 7). This is not a discrepancy, but simply highlights that proteins with different structures follow different folding pathways.

Next, folding of the first 68 aa of HemK CTD (Figure 40; HemK 73-140) occurs at a peptide length between 80 and 100 aa. This also requires a 12-32 aa linker to be present inside the tunnel. This structure is already too large to be accommodated inside the polypeptide exit tunnel and the folding has to occur after the whole segment of the peptide emerged from the ribosome. An unstructured linker of about 30 aa may be long enough to bring the structured domain outside the ribosome. However, a notion of an unstructured linker seems contradictory to the FPA (area I and II; Figure 21) and ensemble PET and PET-FCS measurements (Figure 28; chapter 3.3.2), which indicate that the N-terminus of HemK CTD folds inside the polypeptide exit tunnel. But it is yet unknown whether peptide during the synthesis of the entire domain folds inside the tunnel or if mainly the N-terminus is prone to fold inside the exit tunnel. It seems possible, considering the nature of FPA, that tension that originates from folding outside of the tunnel, can unfold previously folded peptide inside the tunnel. Such information is currently not available; by analogy, it would be interesting to know whether the unfolded peptide of FLN6 of the FLN5/FLN6 structures (PDB: 2N62, 7ZP8) is folded at some point during translation.

In HemK, intriguingly, area III of the FPA is a low tension area that would coincide with the folding of the first 48 aa, which would be consistent with the idea that the folding of the first structural element unfolds peptide inside the tunnel. This would provide additional length of the linker bringing the first 68 aa outside of the tunnel. Moreover, simultaneous folding outside of the tunnel and the unfolding inside the tunnel could remove any tension on the PTC and lead to the low tension area III. Interestingly, area IV of the FPA (Figure 21) does coincide with the

folding event of the 68 aa part of HemK CTD, but it is not obvious whether the tension of the FPA originates from the folding inside or outside the polypeptide exit tunnel.

As translation progresses, 87 aa and 103 aa of HemK CTD (Figure 40; HemK 73-159 and 73-175) fold successively after the additional synthesis of 33-53 and 37-57 aa, respectively (Table 8). This increase in length of the linker (approximately 20 aa longer than the linker for the 68 aa folding segment), which allows this folding event, can be explained by two scenarios. First, the nascent chain can partially fold inside the polypeptide exit tunnel, after being presumably in the extended conformation at an earlier point. Secondly, it is possible that the folded structure has to be further away from the ribosome in order to reach its compacted state on the ribosome, as larger structures require more space to fold. In principle, there is evidence that contacts with the ribosome prevent folding to the native structure and induce an intermediate state (Chan et al., 2022) and that in particular contacts with L24 can delay folding to the native state (Cabrita et al., 2016; Kudva et al., 2018). However, it is unclear to what extent the ribosome is able to exclude the formation of intermediate structures by occluding the docking of newly emerged peptide.

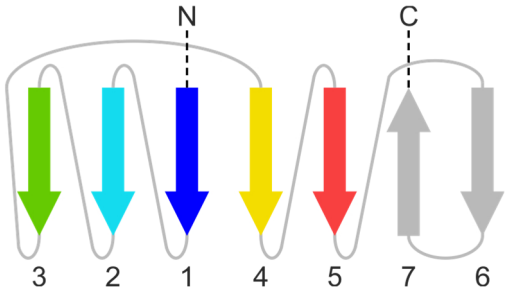



Figure 41. Topology of the central β-sheet of HemK CTD.

The two latest incorporation sites of BOF-Lys in HemK CTD report on the folding of the majority of the domain (Figure 40; HemK 73-205 and 73-226). In both cases, the folding occurs at the very end of translation and represents the formation of the final intermediate structure on the ribosome. Although there is a difference in folding times by roughly 9 s (Table 6), both reporters provided the information of the same folding event. This seems to be breaking with a sequential folding pathway of HemK CTD, but considering that both labels are located between β-strand 4 and 5 of the central β-sheet (Figure 41), it is possible that the α-helical loop at the end of HemK 73-205 (Figure 40) is not stabilized by the core structure until the C-terminal helix of HemK 73-226 (red α-helix in Figure 40) or more likely the following β-strand as part of the central β-sheet (red β-strand in Figure 32 and Figure 41) has emerged

from the ribosome. Therefore, in conclusion, on the ribosome the HemK CTD folds through the continuous compaction in a sequential fashion, fully dependent on the progress of translation and is therefore similar to the folding pathway of HemK NTD (Holtkamp et al., 2015; Liutkute et al., 2020a; Mercier and Rodnina, 2018), but in contrast to the reports that support domain-wise folding in a concerted fashion. The SH3 domain and EF-G are discussed to stay unfolded until they fully emerge from the ribosome (Chen et al., 2020; Eichmann et al., 2010; Guinn et al., 2018). At least in case of EF-G it is possible that similar intermediate structures form, as we observed for HemK CTD, as the experiments were conducted at a different stage of translation, with the domain already emerged from the ribosome (Liu et al., 2019; Liu et al., 2017). Only for FLN5 the formation of intermediate states are confirmed by NMR measurements (Chan et al., 2022). Structurally they could resemble what is published for the following domain FLN6, which is called compacted disordered state (Cabrita et al., 2016). Therefore, even in these cases the sequential folding pathway cannot be fully excluded.

### 4.4. Transition to the native state

Small proteins often attain their native structure in solution after release of the nascent chain from the ribosome. This is very well illustrated using the small fully β-stranded protein CspA as an example. The cryo-EM structure shows that only the first β-hairpin is natively folded on the ribosome. The rest of CspA is accommodated inside the polypeptide exit tunnel in a highly structured, α-helical conformation (Agirrezabala et al., 2022). For CspA and similar proteins, release of the peptide is an essential step in protein folding. For large proteins or domains, most of the peptide chain can emerge from the ribosome during ongoing translation. The emerged protein can acquire a native or native-like structure at the ribosome surface before the release from the ribosome (Cabrita et al., 2016; Chan et al., 2022; Chen et al., 2020; Eichmann et al., 2010; Guinn et al., 2018; Holtkamp et al., 2015; Liutkute et al., 2020a). However, at least 30 aa from the C-terminus are still accommodated inside the polypeptide exit tunnel and cannot participate in the cotranslational and sequential folding. Additionally, proteins possess highly dynamic structures outside of the ribosome as long as they are in close proximity to the ribosomal surface (Cabrita et al., 2016; Liutkute et al., 2020a). The release experiments that were performed in this study clearly show that the protein structure of HemK CTD compacts further after the release from the ribosome (chapter 3.3.4). It is possible that the compaction is induced by alleviating the ribosome-induced destabilization of the nascent chain, which is

manifested as peptide dynamics measured by PET-FCS (Figure 26) that are very pronounced even at the end of translation of the entire HemK CTD.

In summary, this work shows the detailed folding pathway of HemK CTD to an intermediate state on the ribosome and the transition to the native state in solution after release from the ribosome. This work contributes to the understanding of protein folding and could present a basic mechanism of cotranslational folding of α/β proteins.

### 4.5. Conclusion and perspective

This work presents a detailed folding pathway of the HemK CTD dissected by a combination of pre-steady state and equilibrium techniques in combination with FRET and PET measurements. The N-terminal part of the CTD starts folding inside the polypeptide exit tunnel and undergoes various conformational changes depending on the precise location of the peptide inside the polypeptide exit tunnel and the corresponding space restriction at these locations. After the N-terminus emerges from the ribosome, the CTD continues to compact sequentially to an intermediate state as the rest of the protein is being synthesized. This intermediate protein structure remains dynamic on the ribosome and compacts into its final native state only upon release from the ribosome at the end of translation. In order to assess the nature of the intermediate state on the ribosome, single molecule FRET measurement could help to understand the heterogeneity of the compact intermediate state on the ribosome as was done for protein–protein interactions on the ribosome (Marino et al., 2018). Additionally, the effect of the cofactor SAM on the folding pathway could be analyzed, as SAM is located close to strand 2, 1 and 4 of the central β-sheet (Figure 41) and could affect the cotranslational folding pathway. In general, as a next major step in understanding of how protein fold on the ribosome, it would be interesting to analyze whether structurally similar proteins follow a conserved folding pathway and whether there are general folding rules that can be applied to different types of proteins.

# 5. Material and Methods

## 5.1. Lists of buffers, plasmids and primers

Table 9: Buffers and their composition

| **Buffer** | **Composition** |
|---|---|
| $HAMK_7$ | 50 mM Hepes, 70 mM $NH_4Cl$, 30 mM KCl, 7 mM $MgCl_2$, pH 7.5 |
| TBE | 89 mM Tris, 89 mM borate, 2.5 mM EDTA |
| TAE | 40 mM Tris, 20 mM acetic acid, 1 mM EDTA |
| Cathode buffer | 100 mM Tris, 100 mM Tricine, 0.1% SDS |
| Anode buffer | 200 mM Tris-HCl, pH 8.9 |
| Gel buffer (3x) | 3 M Tris-HCl, pH 8.45, 0.3% SDS |
| HiFi | 50 mM Hepes, 70 mM $NH_4Cl$, 30 mM KCl, 3.5 mM $MgCl_2$, 8 mM putrescine, 0.5 mM spermidine, pH 7.5 |
| Compensation buffer | 50 mM Hepes, 70 mM $NH_4Cl$, 30 mM KCl, 16 mM Putrescine, 1 mM spermidine, pH 7.5 |
| tRNA buffer A | 50 mM NaOAc, 10 mM $MgCl_2$, pH 4.5 |
| tRNA buffer B | 50 mM NaOAc, 10 mM $MgCl_2$, 1.1 M KCl, pH 4.5 |
| SDS sample buffer | 100 mM Tris-HCl pH 6.8, 20% glycerol, 7% SDS, 2% BME |
| 5x HPLC buffer | 100 mM $NH_4OAc$, 50 mM $MgCl_2$, 2000 mM NaCl, pH 5 |
| HPLC buffer A | 20 mM $NH_4OAc$, 10 mM $MgCl_2$, 400 mM NaCl, 5% EtOH, pH 5 |
| HPLC buffer B | 20 mM $NH_4OAc$, 10 mM $MgCl_2$, 400 mM NaCl, 40% EtOH, pH 5 |
| EF-Tu His buffer A | 50 mM Hepes, 60 mM $NH_4Cl$, 300 mM KCl, 7 mM $MgCl_2$, 20 mM imidazol, 10% glycerol, 5 mM β-mercaptoethanol, 30 µM GDP, pH 8 |
| EF-Tu His buffer B | 50 mM Hepes, 60 mM $NH_4Cl$, 300 mM KCl, 7 mM $MgCl_2$, 500 mM imidazol, 10% glycerol, 5 mM β-mercaptoethanol, 30 µM GDP, pH 8 |
| EF-Tu Q buffer A | 25 mM Tris-HCl, 10% glycerol, pH 7.5 |
| EF-Tu Q buffer B | 25 mM Tris-HCl, 1.5 M KCl, 10% glycerol, pH 7.5 |
| EF-Tu gel filtration buffer | 50 mM Hepes, 70 mM $NH_4Cl$, 30 mM KCl, 3.5 mM $MgCl_2$, 30 µM GDP, pH 7.5 |
| *In vitro* transcription buffer | 50 mM Tris-HCl, 15 mM $MgCl_2$, 5 mM DTT, 2 mM spermidine, pH 7.5 |

Table 10. List of the original plasmid constructs used in this study. In black is the 5´UTR that includes the T7 promotor for *in vitro* transcription and the ribosome binding site for initiation. Both 5´UTR originate from the pet24a+ vector. In green is the corresponding HemK sequence. In blue is SecM arrest peptide, in red the 20 aa of CspA and in purple a 2 aa linker, a 6-histidine tag and the stop codon.

| **Original plasmids** | **DNA sequence** |
|---|---|
| HemK wt pet24a+ (Holtkamp et al., 2015) | TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAG AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGAATATCAACA CTGGTTACGTGAAGCAATAAGCCAACTTCAGGCGAGCGAAAGCCCGCGGC GTGATGCTGAAATCCTGCTGGAACATGTTACCGGCAGAGGGCGTACTTTTA TCCTCGCCTTTGGTGAAACGCAGCTGACTGACGAACAATGTCAGCAACTTG ATGCGCTACTGACACGTCGTCGCGATGGTGAACCCATTGCTCATTTAACCG GGGTGCGAGAATTCTGGTCGTTGCCGTTATTTGTTTCGCCAGCGACCTTAA TTCCGCGCCCGGATACGGAGTGTCTGGTGGAGCAGGCACTGGCGCGGTTG CCTGAACAACCTTGCCGTATTCTCGATCTCGGGACGGGTACCGGGGCGATT GCGCTTGCGCTGGCTAGCGAGCGCCCGGACTGCGAAATTACCGCTGTAGA TCGTATGCCCGATGCTGTCGCTCTGGCACAACGTAATGCCCAGCATCTGGC GATCAAAAATATCCGCATTCTGCAAAGCGACTGGTTTAGTGCGCTGGCCGG GCAGCAGTTTACGATGATTGTCAGCAATCCGCCGTATATTGACGAGCAGG ACCCACATCTTCAACAAGGCGATGTCCGCTTTGAGCCGCTCACTGCGCTGG TTGCGGCAGACAGTGGAATGGCAGACATCGTGCATATCATCGAACAGTCG CGTAACGCGCTGGTATCCGGCGGCTTTCTGCTTCTGGAACATGGCTGGCAG CAGGGCGAAGCGGTGCGACAAGCATTTATCCTCGCGGGGTATCATGACGT CGAAACCTGCCGTGACTATGGTGATAACGAGCGCGTAACGCTCGGCCGCT ATTATCAACTCGAGCACCACCACCACCACCACTGA |
| HemK NTD 1-103 SecM CspA20 pEX-K168 (Liutkute et al., 2020a) Produced by Eurofins | TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAG AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGAATATCAACA CTGGTTACGTGAAGCAATAAGCCAACTTCAGGCGAGCGAAAGCCCGCGGC GTGATGCTGAAATCCTGCTGGAACATGTTACCGGCAAAGGGCGTACTTTTA TCCTCGCCTTTGGTGAAACGCAGCTGACTGACGAACAATGTCAGCAACTTG ATGCGCTACTGACACGTCGTCGCGATGGTGAACCCATTGCTCATTTAACCG GGGTGCGAGAATTCTGGTCGTTGCCGTTATTTGTTTCGCCAGCGACCTTAA TTCCGCGCCCGGATACGGAGTGTCTGGTGGAGCAGGCACTGTTCAGCACG CCCGTCTGGATAAGCCAGGCGCAAGGCATCCGTGCTGGCCCTATGTCCGGT AAAATGACTGGTATCGTAAAATGGTTCAACGCTGACAAAGGCTTCGGCTTC ATCACTCCT |

Table 11: Primer list for site-directed mutagenesis. Listed are the templates, the primers and their sequence and the final products. Highlighted in yellow are the target codons or segments. The PCR templates originate either from previous work (Table 10), from preceding site-directed mutagenesis (listed in this table as cloning product) or from Gibson assembly (Table 18). Forward (fwd) primer are designed to bind the antisense strand and reverse (rev) primer are designed to bind the sense strand. Primers were ordered from Eurofins.

| **PCR template** | **Primer name** | **Primer sequence (5‘→ 3‘)** | **Cloning product** |
|---|---|---|---|
| HemK wt pet24a+ (Table 10) | HemK G73AUG fwd | ATGGTGCGAGAATTCTGGTCG | HemK 73-277 wt pet24a |
| | CTD truncation rev | ATGTATATCTCCTTCTTAAAGTTA | |
| HemK wt pet24a+ (Table 10) | HemK P93AUG fwd | ATGGATACGGAGTGTCT | HemK 93-277 wt pet24a |
| | CTD truncation rev | ATGTATATCTCCTTCTTAAAGTTA | |
| HemK 73-277 wt pet24a+/<br><br>HemK 93-277 wt pet24a+/ | HemK K159R fwd | GGCGATCCGTAATATCCGCA | HemK 73-277 K159R pet24a+/<br><br>HemK 93-277 K159R pet24a+ |
| | HemK K159R rev | AGATGCTGGGCATTACGTTGT | |
| HemK 73-277 K159R pet24a+/<br><br>HemK 93-277 K159R pet24a+ | HemK T120K fwd | CGGGACGGGTAAAGGGGCGATTGC | HemK 73-277 T120K K159R pet24a+/<br><br>HemK 93-277 T120K K159R pet24a+ |
| | HemK T120K rev | AGATCGAGAATACGGCAAGGTTGTTCAGG | |
| HemK 73-277 K159R pet24a+ | HemK R131K fwd | GCTTGCGCTGGCTAGCGAGAAACCGGACTGCGAAATTACC | HemK 73-277 R131K K159R pet24a+ |
| | HemK R131K rev | GCAATCGCCCCGGTACCCGTCCCG | |
| HemK 73-277 K159R pet24a+ | HemK D140K fwd | GAAATTACCGCTGTAAAACGTATGCCCGATG | HemK 73-277 D140K K159R pet24a+ |
| | HemK D140K rev | GCAGTCCGGGCGCTCGCTAGCCAGC | |
| HemK 73-277 K159R pet24a+ | HemK M142K fwd | GAAATTACCGCTGTAGATCGTAAACCCGATGCTGTCGCTCTGG | HemK 73-277 M142K K159R pet24a+ |
| | HemK M142K rev | GCAGTCCGGGCGCTCGCTAGCCAGCGCAAGC | |
| HemK 73-277 K159R pet24a+ | HemK Q175K fwd | GGTTTAGTGCGCTGGCCGGGAAACAGTTTACGATGATTG | HemK 73-277 Q175K K159R pet24a+ |
| | HemK Q175K rev | AGTCGCTTTGCAGAATGCGGATAT | |
| HemK 73-277 K159R pet24a+ | HemK T205K fwd | CCGCTTTGAGCCGCTCAAAGCGCTGGTTGCGGCAGAC | HemK 73-277 T205K K159R pet24a+ |
| | HemK T205K rev | ACATCGCCTTGTTGAAGATGTGGGTC | |

| | | | |
|---|---|---|---|
| HemK 73-277 K159R pet24a+ | HemK N226K fwd | CATCGAACAGTCGCGTAAAGCGCTGGTATCCGGCGG | HemK 73-277 N226K K159R pet24a+ |
| | HemK N226K rev | ATATGCACGATGTCTGCCATTCCAC | |
| HemK 73-277 wt pet24a+ | HemK W78F fwd | GTGCGAGAATTCTTTTCGTTGCCGTTATTTG | HemK 73-277 W78F pet24a+ |
| | HemK W78F rev | CATATGTATATCTCCTTCTTAAAGTTAAAC | |
| HemK FL wt SecM CspA20 pet24a+ (Table 18) | HemK G73AUG fwd | ATGGTGCGAGAATTCTGGTCG | HemK 73-277 wt SecM CspA20 pet24a+ |
| | CTD truncation rev | ATGTATATCTCCTTCTTAAAGTTA | |
| HemK FL wt SecM CspA20 pet24a+ (Table 18) | Insert CspA 21-50 fwd | CCAGAACGATGGTTACAAATCTCTGGACGAAGGTCAGAAACTCGAGCACCACCACCACCACCACTGAG | HemK FL wt SecM CspA50 pet24a+ |
| | Insert CspA 21-50 rev | ATAGCAGAGAAGTGTACGAACACATCTTTAGAGCCATCGTCAGGAGTGATGAAGCCGAAGCCTTTGTC | |
| HemK wt SecM CspA20 pet24a+ (Table 18)/<br><br>HemK wt SecM CspA50 pet24a+/<br><br>HemK 73-277 wt SecM CspA20 pet24a+ | SecM arrest fwd | TTCAGCACGCCCGTCTGGATAAGCCAGG | HemK 1-xxx wt SecM CspA20 pet24a+/<br><br>HemK 1-xxx wt SecM CspA50 pet24a+/<br><br>HemK 73-xxx wt SecM CspA20 pet24a+ |
| | HemK truncation xxx rev | 5´end of the primer binds to the last codon (XXX) that is still included in the new product. Sections that decode for HemK and are located between the binding sites of fwd and rev primer are deleted.<br>Several primers were designed to stepwise truncate the HemK sequence from the C-terminus | |

Table 12: Primers for PCR used for *in vitro* transcription. Fwd primer are designed to bind the antisense strand and rev primer are designed to bind the sense strand. Primers were ordered from Eurofins. FPA: force profile assay.

| **Primer name** | **Primer sequence (5‘→ 3‘)** |
|---|---|
| T7 fwd (pet24a+) (all mRNA) | TAATACGACTCACTATAGGGG |
| HemK85 mRNA rev | CGAAACAAATAACGGCAACGACCAG |
| HemK85 W78F mRNA rev | CGAAACAAATAACGGCAACGAAAAG |
| HemK92 mRNA rev | GCGCGGAATTAAGGTCGCTGGCG |
| HemK112 mRNA rev | ACGGCAAGGTTGTTCAGGCAACCG |
| HemK122 mRNA rev | CGCCCCGGTACCCGTCCCGAGATCG |
| HemK122 T120K mRNA rev | CGCCCCTTTACCCGTCCCGAGATCG |
| HemK132 mRNA rev | CGGGCGCTCGCTAGCCAGCGCAAG |
| HemK152 mRNA rev | ATTACGTTGTGCCAGAGCGACAGC |
| HemK172 mRNA rev | CAGCGCACTAAACCAGTCGCTTTGC |
| HemK192 mRNA rev | TGGGTCCTGCTCGTCAATATACGG |
| HemK212 mRNA rev | ACTGTCTGCCGCAACCAGCGCAGTG |
| HemK232 mRNA rev | GCCGCCGGATACCAGCGCGTTACG |
| HemK252 mRNA rev | GAGGATAAATGCTTGTCGCACCGC |
| HemK272 mRNA rev | GAGCGTTACGCGCTCGTTATCACC |
| HemK FL mRNA rev | TTGATAATAGCGGCCGAGCGTTACG |
| CspA20 mRNA rev (most mRNA used in FPA) | AGGAGTGATGAAGCCGAAGCC |
| CspA50 mRNA rev (long mRNA used in FPA) | TTTCTGACCTTCGTCCAGAGATTTGTAACC |

### 5.2. Site-directed mutagenesis

Point mutations, deletions or insertions into the plasmid template were introduced by site-directed mutagenesis. The PCR reaction was performed according to Table 13 and Table 14, using primers and templates from Table 11.

Table 13: PCR mix for site-directed mutagenesis.

| **Component** | **Final concentration** |
|---|---|
| Phusion GC buffer (Thermo Scientific) | 1x |
| dNTP (NEB) | 0.2 mM |
| Forward primer | 0.4 µM |
| Reverse primer | 0.4 µM |
| Plasmid template | 0.5 ng/µl |
| Phusion polymerase (Thermo Scientific) | 0.02 U/µl |

Table 14: PCR program for site-directed mutagenesis.

| **Step** | **Temperature, °C** | **Time, s** | **Number of cycles** |
|---|---|---|---|
| Initial denaturation | 98 | 120 | 1 |
| Denaturation | 98 | 30 | 30 |
| Primer annealing | 55 | 20 | |
| Extension | 72 | 150 | |
| Final extension | 72 | 300 | 1 |

The integrity of the linearized PCR product was checked on 1% agarose gel in TAE buffer and visualized with Stain G (Serva). A final concentration of 40% PCR reaction was phosphorylated in 1 x T4 DNA ligase buffer (NEB, compatible with phosphorylation), supplemented with 2 mM ATP and 0.2 U/µl T4 Polynucleotide Kinase (NEB) at 37°C for 30 min followed by a heat inactivation at 60°C for 20 min. Ligation of the PCR product was done with a final concentration of 50% phosphorylation mix in 1 x T4 DNA ligase buffer (NEB), 2.5 mM additional ATP and 20 U/µl T4 DNA ligase (NEB) at room temperature for 2 h, followed by heat inactivation at 65°C for 10 min. To degrade the original template,

1 U/µl DpnI (NEB) was added. Cleavage reaction was performed at 37°C overnight. Finally, 5 µl of the final cloning product was used for heat shock transformation of 50 µl *E. coli* NovaBlue cell suspension (Merck, in house preparation) according to the manufacturer's instructions. The success of the mutagenesis was controlled by sequencing performed by SeqLab (Mycrosynth).

### 5.3. Gibson assembly to generate initial plasmid for force profile assay

The Gibson assembly was used to introduce the SecM arrest peptide and the 20 aa CspA reporter for the force profile assay downstream of the HemK sequence in the pet24a+ vector.

Table 15. Templates and primers for the PCR needed for the Gibson assembly of the HemK SecM CspA pet24a+ construct.

| **PCR template** | **Primer name** | **Primer sequence (5'→ 3')** | **PCR product** |
|---|---|---|---|
| HemK wt pet24a+ (Table 10) | Gibson primer 1 | CTCGAGCACCACC | Linearized plasmid, opened right after HemK coding sequence (green/purple acc. to Table 10) → vector |
| | Gibson primer 2 | TTGATAATAGCGGCCG | |
| HemK NTD 1-103 SecM CspA20 pEX-K168 (Table 10) | Gibson primer 3 | GGCCGCTATTATCAATTCAGCACGCCC | SecM-CspA20 (blue and red acc. to Table 10) with overhang that pair partially with HemK and His-Tag in the pet24a+ construct → insert |
| | Gibson primer 4 | GGTGGTGCTCGAGAGGAGTGATGAAGCC | |

The PCR to produce the vector was performed in a total volume of 150 µl according to Table 13 and Table 16.

Table 16. PCR program for vector amplification for the Gibson assembly.

| **Step** | **Temperature, °C** | **Time, s** | **Number of cycles** |
|---|---|---|---|
| Initial denaturation | 98 | 120 | 1 |
| Denaturation | 98 | 30 | 25 |
| Primer annealing | 58 | 20 | |
| Extension | 72 | 240 | |
| Final extension | 72 | 300 | 1 |

The PCR to produce the insert was performed in a total volume of 150 µl according to Table 19 and Table 17.

Table 17. PCR program for insert amplification for the Gibson assembly.

| **Step** | **Temperature, °C** | **Time, s** | **Number of cycles** |
|---|---|---|---|
| Initial denaturation | 98 | 120 | 1 |
| Denaturation | 98 | 30 | 20 |
| Primer annealing | 58 | 20 | |
| Extension | 72 | 30 | |
| Final extension | 72 | 300 | 1 |

The total reaction mixture of the insert and the vector was separated on a 1% agarose gel in TAE buffer and visualized with Stain G (Serva). The PCR products were purified from the agarose gel, by gel extraction using the NucleoSpin Gel and PCR Clean-up, Mini kit (MACHEREY-NAGEL) according to the manufacturer's instructions. The Gibson assembly was performed with pre-mixed Gibson assembly mix (prepared by Tessa Hübner according to (Gibson et al., 2009)). 50 ng of vector and 180 ng insert was mixed with the Gibson assembly mix to have a total volume of 20 µl. Reaction was performed at 50°C for 60 min. Finally, 10 µl of the final cloning product was used for heat shock transformation of 50 µl *E. coli* NovaBlue cell suspension (Merck, in house preparation) according to the manufacturer's instructions. The success of the mutagenesis was controlled by sequencing performed by SeqLab (Mycrosynth). The sequence of the final product is listed in Table 18.

Table 18. Sequence of the final product of the Gibson assembly. In black is the 5´UTR that includes the T7 promotor for *in vitro* transcription and the ribosome binding site for initiation. Both 5´UTR originate from the pet24a+ vector. In green is the corresponding HemK sequence. In blue is SecM arrest peptide, in red the 20 aa of CspA and in purple a 2 aa linker, a 6-histidine tag and the stop codon.

| **Original plasmids** | **DNA sequence** |
|---|---|
| HemK FL wt SecM CspA20 pet24a+ | TAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAG AAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATGGAATATCAACA CTGGTTACGTGAAGCAATAAGCCAACTTCAGGCGAGCGAAAGCCCGCGGC GTGATGCTGAAATCCTGCTGGAACATGTTACCGGCAGAGGGCGTACTTTTA TCCTCGCCTTTGGTGAAACGCAGCTGACTGACGAACAATGTCAGCAACTTG ATGCGCTACTGACACGTCGTCGCGATGGTGAACCCATTGCTCATTTAACCG GGGTGCGAGAATTCTGGTCGTTGCCGTTATTTGTTTCGCCAGCGACCTTAA TTCCGCGCCCGGATACGGAGTGTCTGGTGGAGCAGGCACTGGCGCGGTTG CCTGAACAACCTTGCCGTATTCTCGATCTCGGGACGGGTACCGGGGCGATT GCGCTTGCGCTGGCTAGCGAGCGCCCGGACTGCGAAATTACCGCTGTAGA TCGTATGCCCGATGCTGTCGCTCTGGCACAACGTAATGCCCAGCATCTGGC GATCAAAAATATCCGCATTCTGCAAAGCGACTGGTTTAGTGCGCTGGCCGG GCAGCAGTTTACGATGATTGTCAGCAATCCGCCGTATATTGACGAGCAGG ACCCACATCTTCAACAAGGCGATGTCCGCTTTGAGCCGCTCACTGCGCTGG TTGCGGCAGACAGTGGAATGGCAGACATCGTGCATATCATCGAACAGTCG CGTAACGCGCTGGTATCCGGCGGCTTTCTGCTTCTGGAACATGGCTGGCAG CAGGGCGAAGCGGTGCGACAAGCATTTATCCTCGCGGGGTATCATGACGT CGAAACCTGCCGTGACTATGGTGATAACGAGCGCGTAACGCTCGGCCGCT ATTATCAATTCAGCACGCCCGTCTGGATAAGCCAGGCGCAAGGCATCCGT GCTGGCCCTATGTCCGGTAAAATGACTGGTATCGTAAAATGGTTCAACGCT GACAAAGGCTTCGGCTTCATCACTCCTCTCGAGCACCACCACCACCACCAC TGA |

### 5.4. EF-Tu purification

EF-Tu was overexpressed in *E. coli* BL21 (DE3)pLysS using a pet24a+ EF-Tu wt carrying a C-terminal His-Tag by Michael Zimmermann. Approximately 50 g cells were reconstituted in 2 ml/g cell pellet EF-Tu His buffer A, supplemented with 1 tablet cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail (Roche)/ 50 ml buffer, 200 µl Rnase-free DNase I (Jena Bioscience)/50 ml buffer and one small spatula lysozyme from chicken egg white (Sigma-Aldrich)/ 50 ml buffer. Cell lysis was performed by sonication with an amplitude of 40% for 20 min as 5 s pulses followed by 5 s pauses and 5 min pauses after 2.5 min accumulated pulses. Lysate was cleared by centrifugation at 20000 rpm using a Ti50.2 rotor (Beckman Coulter) for 40 min at 4°C. Supernatant was loaded on 3x5 ml HisTrap HP (Cytiva), equilibrated with EF-Tu His buffer A, at 1 ml/ml. To remove unspecifically bound proteins, the column was washed with EF-Tu His buffer A at 3 ml/min until the baseline was reached. EF-Tu was eluted applying a linear gradient from 0-100% EF-Tu His buffer B in 75 min at 3 ml/min (15 CV). Fractions were collected in 5 ml. Fractions to pool were determined by 12% Tris-glycine SDS-

PAGE. Pooled fractions were diluted with EF-Tu Q buffer A 1:4 and loaded on 2x5 ml HiTrap Q HP (Cytiva), equilibrated with EF-Tu Q buffer A at 2 ml/ml. Column was washed with 10 CV EF-Tu Q buffer A at 4 ml/ml. Elution was started by applying a linear gradient from 0-50% EF-Tu Q buffer B in 125 min (100% in 100 CV) at 4 ml/min. Fractions were collected in 5 ml. Fractions to pool were determined by 12% Tris-glycine SDS-PAGE. Pool was concentrated to 5 ml with Vivaspin 20, 10 kDa MWCO (Satorius). Sample was split in two and loaded separately on a HiLoad® 26/600 Superdex® 75 pg (Cytiva) equilibrated with EF-Tu gel filtration buffer. Separation was performed at a flow rate of 1.5 ml/min. Fractions were collected in 2 ml. Fractions to pool were determined by 12% Tris-glycine SDS-PAGE. Pool was concentrated to a final concentration of ~1000 µM with Vivaspin 20, 10 kDa MWCO (Satorius). The final concentration was determined by absorbance measurement at $\lambda$ = 280 nm and an extinction coefficient of 32900 $M^{-1}*cm^{-1}$.

### 5.5. Aminoacylation of total tRNA

Aminoacylation was performed in a total volume of 40 ml. 80 OD/ml total RNA (Roche) were aminoacylated with 0.3 mM of each amino acid (all 20 amino acids or without Lys for incorporation reaction (Lys-deficient aa-tRNA)) by using 3% of S100 *E. coli* extract (prepared by Olaf Geintzer). Reaction was carried out in $HAKM_7$ supplemented with 13 mM $MgCl_2$ (final $MgCl_2$ concentration of 20 mM), 3 mM ATP and 2 mM DTT for 30 min at 37°C. 1/10 volumes of 20% KOAc, pH 5 was added to stop the reaction. In order to remove the synthetases of the S100 extract the solution was mixed with equal volumes of ROTI®Aqua-Phenol (Roth). The aqueous phase was separated by centrifugation at 4000 rpm for 10 min at 16°C and transferred into a fresh tube, followed by RNA precipitation with 1/10 volumes of 20% KOAc, pH 5 and 2.5 volumes ice cold ethanol overnight at -20°C. Pellet was collected by centrifugation and dissolved in tRNA buffer A, split into two to be purified separately. Sample was loaded on 3x5 ml HiTrap Q columns (Cytiva), equilibrated with tRNA buffer A using a flow rate of 3 ml/min. Columns were washed at a flow rate of 3 ml/min until baseline was reached. aa-tRNA was eluted with a linear gradient from 0-100% tRNA buffer B in 10 CV at a flow rate of 3 ml/min. Two peaks elute from the column during the linear gradient. The first peak is discarded because it only contains free nucleotides and residual phenol. The second peak contains aa-tRNA. The main fractions were pooled and the tRNA precipitated with 1/10 volumes of 20% KOAc, pH 5 and 2.2 volumes ice cold ethanol overnight at -20°C. The pellet was collected by centrifugation and dissolved in 40 ml water. The precipitation was

repeated at least 3 times. The final pellet was dissolved in water to a final volume of 1/10 volumes of the initial reaction mixture. Final concentration was determined by absorbance measurement at $\lambda$ = 260 nm and an extinction coefficient of 0.575 $\mu M^{-1}*cm^{-1}$.

### 5.6. Fluorescent labeling of Lys-tRNA$^{Lys}$

[$^{14}$C]Lys-tRNA$^{Lys}$ was prepared by Olaf Geintzer as described previously (Mittelstaet et al., 2013). The pre-reaction mixture contained 40 µM [$^{14}$C]Lys-tRNA$^{Lys}$ , 4 mM BODIPY™ FL sulfonated succinimidyl ester dissolved in water (Invitrogen) and 30 mM potassium phosphate buffer, pH 7 ($pK_a$ = 6.86; mixing 1.59 and 1 part $K_2HPO_4$ and $KH_2PO_4$, respectively). The labeling of the ε-nitrogen of Lys was started by adding 48 mM KOH and incubated at room temperature for 20 s under constant agitation and stopped by adding 48 mM acetic acid. Free dye was removed by RNA precipitations with 1/10 volumes of 20% KOAc, pH 5 and 2.5 volumes ice cold ethanol overnight at -20°C. The pellet was collected by centrifugation and dissolved in 3 ml water. The RNA precipitation was repeated until the supernatant of the ethanol precipitation was completely colorless, but at least 3 times. The final pellet was dissolved in ~1 ml water. ¼ volumes 5x HPLC buffer was added prior to loading the sample on the reverse-phase chromatography column WP-300 RP-18 (Merck) equilibrated with HPLC buffer A using a flow rate of 1 ml/min. Elution was started right after loading by applying a linear gradient from 0-50% HPLC buffer B over 70 min at a flow rate of 3 ml/min. 3 ml fractions were collected and main fractions were determined by chromatograms measured at $\lambda$ = 260 nm for tRNA, $\lambda$ = 502 nm for BOF dye and liquid scintillation counting of the $^{14}$C label. The main fractions were pooled and precipitated overnight with 1/10 volumes of 20% KOAc, pH 5 and 2.2 volumes ice cold ethanol at -20°C. The pellet was collected by centrifugation and dissolved in 10 ml water and the RNA precipitation was repeated once. The final pellet was dissolved in 200 µl water. Final concentration was determined by liquid scintillation counting the $^{14}$C label and calculated according to

equation 1

$$\text{concentration} = \frac{\text{counts [dpm]}}{\text{specific activity } \left[\frac{\text{dpm}}{\text{pmol}}\right]} * \text{counted volume}^{-1} * \text{dilution factor}$$

### 5.7. *In vitro* transcription and purification of mRNA

The template for the *in vitro* transcription was amplified by PCR according to Table 19 and Table 20, using templates from Table 11 (both listed as templates and products) and primers from Table 12 in a total volume of 300 µl.

Table 19: PCR mix for *in vitro* transcription.

| **Component** | **Final concentration** |
|---|---|
| Phusion HF buffer (Thermo Scientific) | 1x |
| dNTP (NEB) | 0.2 mM |
| Forward primer | 0.4 µM |
| Reverse primer | 0.4 µM |
| Plasmid template | 0.5 ng/µl |
| Phusion polymerase (Thermo Scientific) | 0.02 U/µl |

Table 20: PCR program for *in vitro* transcription.

| **Step** | **Temperature, °C** | **Time, s** | **Number of cycles** |
|---|---|---|---|
| Initial denaturation | 98 | 120 | 1 |
| Denaturation | 98 | 30 | 30 |
| Primer annealing | 55 | 20 | |
| Extension | 72 | 15 | |
| Final extension | 72 | 90 | 1 |

The efficiency of the template amplification was checked on a 1% agarose gel in TAE buffer and visualized by Stain G (Serva). *In vitro* transcription was performed in *in vitro* transcription buffer supplemented with 10 mM DTT, 3 mM of each NTP, 5 mM GMP, 1/10 volumes PCR mix, 0.005 U/µl thermostable inorganic pyrophosphatase (NEB), 0.5% T7 polymerase, in the presence of 0.2 U/µl RNase inhibitor (Molox) at 37°C for 3 h. mRNA was purified using the RNeasy Midi Kit (Qiagen) according to the company's protocol for RNA clean-up, with the following exception. mRNA was eluted with 3x500 µl Milli-Q water. Elutions were pooled and precipitated with 1/10 volumes of 20% KOAc, pH 5 and 2.2 volumes ice cold ethanol

overnight at -20°C. The pellet was collected by centrifugation and dissolved in 200 µl Milli-Q water. The mRNA concentration was determined by measuring the absorbance at λ = 260 nm and calculated according to

equation 2

$$\text{concentration}\left[\frac{\mu g}{\mu l}\right] = 0.04\ \frac{\mu g}{\mu l} * OD_{260}$$

and

equation 3

$$MW(RNA) = \left(\#nt * 320.5\frac{g}{mol}\right) + 159.0\frac{g}{mol}$$

where #nt is the number of nucleotides in the mRNA and MW(RNA) is the approximated molecular weight of the RNA.

### 5.8. Translation

All the components for mRNA translation, i.e., ribosomes, initiation factors IF1, IF2 and IF3, initiator tRNA, EF-G, EF-Tu, and EF-Ts were prepared by Olaf Geintzer, Michael Zimmermann, Christina Kothe, Sandra Kappler, Anna Pfeifer and Vanessa Herold as described in 5.4 and in earlier reports (Cunha et al., 2013; Milon et al., 2007; Mittelstaet et al., 2013; Rodnina and Wintermeyer, 1995; Wieden et al., 2002).

#### 5.8.1. Assembly of initiation complex

Initiation complex (IC) was assembled using a final concentration of 1 µM 70S *E. coli* ribosomes, 1.5 µM IF1, 1.5 µM IF2, 1.5 µM IF3, 1.5 µM initiator tRNA (Atto655-[$^3$H]Met-tRNA$^{fMet}$, Bodipy FL-[$^3$H]Met-tRNA$^{fMet}$ or f[$^3$H]Met-tRNA$^{fMet}$) and 8 µM mRNA in $HAMK_7$ supplemented with 2 mM DTT and 2 mM GTP. All components were mixed and heated at 37°C for 45 min. If needed, IC was purified by sucrose cushion ultracentrifugation. 140 µl IC was loaded on 60 µl 1.1 M sucrose in $HAKM_7$ and collected in a Beckman Optima XP centrifuge using a TLA-100 rotor at 68 000 rpm and 4°C for 1 h. The pellet was dissolved in $HAKM_7$. Concentration of the purified IC was determined by liquid scintillation counting of $^3$H label and calculated according to equation 1.

### 5.8.2. Formation of ternary complex

Ternary complex (TC) was prepared by using a final concentration of 240 µM EF-Tu, 0.4 µM EF-Ts and 200 µM aa-tRNA in $HAMK_7$ supplemented with 2 mM DTT, 2 mM GTP, 10 mM phosphoenolpyruvate, 3.36% pyruvate kinase (Roche) and additionally 8 µM EF-G. First, all the components, except aa-tRNA, were mixed and incubated at 37°C for 15 min. aa-tRNA was added and again heated at 37°C for 1 min. In case that BOF-Lys was incorporated into the peptide nascent chain, 200 µM Lys-deficient aa-tRNA was used and 2 µM BOF-[$^{14}$C]Lys-$tRNA^{Lys}$ was added. In case that the TC was prepared for the translation of the mRNA used in force profile assays, 400 µM aa-tRNA was used in order to compensate for longer constructs. In case that TC was prepared for the mRNA translation used for the purification of stalled ribosome-nascent chain complexes, 40 µM EF-Tu was used to avoid excessive spontaneous release of the nascent chain.

### 5.8.3. Translation

To translate an mRNA, first, the initiation complex (chapter 5.8.1; purified or unpurified) was diluted to a concentration of 0.16 µM using $HAMK_7$ and additionally mixed 1:1 with compensation buffer to have a concentration of 0.08 µM IC in HiFi supplemented with a final concentration of 1 mM DTT and 1 mM GTP. The TC (chapter 5.8.2) was mixed 1:1 using compensation buffer to a concentration of 100 µM aa-tRNA, 120 µM EF-Tu and 4 µM EF-G in HiFi supplemented with a final concentration of 1 mM DTT and 1 mM GTP. Translation was started by mixing IC and TC 1:1 to have a final concentration of 0.04 µM IC, 50 µM aa-tRNA, 60 µM EF-Tu, 0.1 µM EF-Ts and 2 µM EF-G in HiFi supplemented with 1 mM DTT and 1 mM GTP. The translation reaction was performed at 37°C. End point translations were incubated for 10 min. Samples for translation time courses were taken at the indicated time points and quenched immediately for SDS-PAGE analysis. Samples analyzed by SDS-PAGE, were quenched by adding 1/10 volumes of 2 M NaOH and hydrolyzed at 37°C for 30 min. To avoid extensive bleaching of the fluorophores, the high pH was re-adjusted by adding the same volume of 2 M HEPES. To samples that were labeled only with the dye Bodipy FL, SDS sample buffer was added 1:1. Samples that were labeled with Atto655 (either single-labeled or doubled-labeled together with Bodipy FL) were dried in a SpeedVac for 35 min at 65°C. Pellets were dissolved in SDS sample buffer using the same volume as of the initial translation mixture. Samples were heated for 10 min at 70°C and separated on a three layer 16% Tris-Tricine SDS-PAGE (Agirrezabala et al., 2022; Schagger, 2006; Schägger and von Jagow,

1987). Translation products on SDS-PAGE were visualized by fluorescence using the Amersham™ Typhoon™ RGB (Cytiva). If needed, RNCs that are stalled at the end of translation were purified by sucrose cushion ultracentrifugation. 140 µl of unquenched translation mixture was loaded on 60 µl 1.1 M sucrose in HiFi and collected in a Beckman Optima XP centrifuge using a TLA-100 rotor at 68 000 rpm and 4°C for 1 h. The pellet was dissolved in HiFi. The concentration of the purified RNC was determined by liquid scintillation counting the $^{3}H$ label and calculated according to equation 1.

#### 5.8.3.1. Data analysis of translation kinetics

The SDS-PAGE analysis of translation time courses were used to extract the information of the translation kinetics. Band intensities of the full-length translation product were quantified with Image Studio Lite (Version 5.2; LI-COR). The band intensities were plotted against their respective time and analyzed in Graphpad Prism (Version 8.3.0) with

equation 4

$$Y(t) = IF(t > delay, Y_0, Y_0 + Amp * \left(1 - e^{-k*(t-delay)}\right)$$

The IF condition accounts for the time delay until the first product appears on the SDS-PAGE. $Y_0$ is the initial intensity of the time delay. Amp is the amplitude of the signal change and k is the apparent rate of the signal change. For better visualization and comparability the graphs of the translation time courses were normalized according to the fitting parameter to Y0 = 0 and Y0+Amp = 1.

Fitting parameters were processed according to

equation 5

$$\text{translation time } \tau_{trans} = delay + \frac{1}{k}$$

The error of the lifetime was propagated according to

equation 6

$$\Delta\tau_{trans} = \sqrt{\Delta delay^2 + \left(-\frac{\Delta k}{k^2} *\right)^2}$$

The standard error of the mean (SEM) was used for propagating the error (Δx) of the fitted parameters.

The average translation rate was calculated according to

equation 7

$$\text{translation rate} = \frac{\#\text{aa}}{\tau_{\text{trans}}}$$

where #aa is the length of translation product (aa)

The error (SEM) of the translation rate was propagated according to

$$\Delta\text{translation rate} = \#\text{aa} * \frac{\Delta\tau_{\text{trans}}}{\tau_{\text{trans}}^2}$$

### 5.9. Measurements of cotranslational protein folding monitoring FRET

The full set of experiments, including the controls, for each construct comprised of two biological replicates with both donor fluorophore (Bodipy FL-[$^{14}$C]Lys-tRNA$^{Lys}$ , TC) and acceptor fluorophore (Atto655-[$^{3}$H]Met-tRNA$^{fMet}$, IC) (donor/acceptor), one donor-only, one acceptor-only and one buffer-only reaction. IC and TC were prepared as described in 5.8.1, 5.8.2 and 5.8.3. All experiments were done with purified IC. FRET stopped-flow experiments were carried out in the stopped-flow device SX-20MV (Applied Photophysics) using a 470 nm LED at 2 mA for excitation, a RG665 emission filter (main channel; Schott) in combination with a R2228 red sensitive fluorescence photomultiplier (Applied Photophysics) and a KV550 emission filter (side channel; Schott) in combination with a R6095 fluorescence photomultiplier (Applied Photophysics) for the acceptor and donor signal, respectively. IC and TC were mixed rapidly in equal volumes at 37°C to start the translation reaction. Fluorescence change was traced with 4000 points over 6 min in logarithmic scale. Each biological replicate was measured with 6 technical replicates. All technical replicates were averaged. Additionally, the biological replicates of the donor/acceptor sample were averaged as well. The fluorescence signal of each time point of the donor/acceptor sample were corrected for non-FRET fluorescence according to

equation 8

$$\text{FRET} = (\overline{F_{DA}} - F_B) - (F_D - F_B) - (F_A - F_B) = \overline{F_{DA}} - F_D - F_A + F_B$$

, where FRET is the fluorescence due to FRET, $\overline{F_{DA}}$ is the average fluorescence of the two biological replicates with both donor and acceptor. $F_B$ is the fluorescence of the blank. $F_D$ the fluorescence of the donor-only sample and $F_A$ the fluorescence of the acceptor-only sample. Additionally, in order to have the same noise-to-signal ratio as after the correction, the controls were subtracted as simulations. The donor-only and blank do not have a signal change over time, therefore all points were averaged and all time points of the D+A sample were corrected with these averaged values. The signal of the acceptor-only does change over time. Thus, the fluorescence time course was submitted to TableCurve2D (Version 5.01.03) and fitted with all equations. The best fit was used as simulation for the correction.

The fluorescence signal over time was analyzed in GraphPad Prism (Version 8.3.0) with

equation 9

$$Y(t) = IF(t > delay, Y_0, Y_0 + Amp * (1 - e^{-k*(t-delay)})$$

and

equation 10

$$Y(t) = IF(t > delay, Y_0, Y_0 + Amp_1 * (1 - e^{-k_1*(t-delay)}) + Amp_2 * (1 - e^{-k_2*(t-delay)})$$

The IF condition accounts for the time delay until the signal changes. $Y_0$ is the initial intensity of the time delay. Amp is the amplitude of the signal change and k is the apparent rate of the signal change. For best fit analysis, all values of time points smaller than 0.5 s were excluded.

Fitting parameters were processed according to

equation 11

$$\text{lifetime } \tau_n = delay + \frac{1}{k_n}$$

The error of the lifetime was propagated according to

equation 12

$$\Delta\tau_n = \sqrt{\Delta delay^2 + \left(-\frac{\Delta k_n}{{k_n}^2} *\right)^2}$$

The standard error of the mean (SEM) was used for propagating the error (Δx) of the fitted parameters.

### 5.10. Measurements of posttranslational protein folding monitoring FRET

The release experiment comprised of three biological replicates with both donor fluorophore (Bodipy FL-[$^{14}$C]Lys-tRNA$^{Lys}$, TC) and acceptor fluorophore (Atto655-[$^{3}$H]Met-tRNA$^{fMet}$, IC) (donor/acceptor). IC and TC were prepared as described in 5.8.1, 5.8.2 and 5.8.3. All experiments were performed with purified IC. First, cotranslational folding was observed, followed by release with ArfB. FRET release experiments were carried out in the fluorimeter Fluoromax-4 (Horiba) using the monochromator at λ = 480 nm for excitation (slits set to bandpass of 7 nm) and at λ = 680 nm for emission (slits set to bandpass of 7 nm). IC and TC were mixed in equal volumes in a 5*5 mm QS cuvette (HellmaAnalystics) at 37°C to start the translation. Fluorescence change was traced in 0.5 s intervals. After the end of translation at around 400 s, 1 µM ArfB was added and the measurement was continued until a total measuring time of approximately 700 s. For analysis, the intensities between time points 50-70 s (baseline), 370-400 s (folded on the ribosome) and 570-600 s (folded off the ribosome) were averaged, representing the starting point of the FRET change (baseline), final FRET on the ribosome and final FRET off the ribosome, respectively. The baseline was subtracted from both FRET values, and the replicates averaged. The FRET value on the ribosome was normalized to 100% and the FRET value off the ribosome accordingly. The error was calculated as the standard deviation of the biological replicates.

### 5.11. Measurements of cotranslational protein folding monitoring PET

A full set of experiments for PET measurements included two biological replicates of both, a Trp (wt)- and a Phe-decoding version of a construct, labeled with N-terminal Bodipy FL (Bodipy FL-[$^{3}$H]Met-tRNA$^{fMet}$, IC). IC and TC were prepared as described in 5.8.1, 5.8.2 and 5.8.3. All experiments were done with purified IC. PET stopped-flow experiments were carried out in the stopped-flow device SX-20MV using the monochromator (Applied Photophysics) at λ = 470 nm for excitation (slits = 2 mm with a bandpass of 4.65 nm/mm) and a KV500 emission filter (Schott) in combination with a R6095 fluorescence photomultiplier (main channel). IC and TC were mixed rapidly in equal volumes at 37°C to start the translation reaction. Fluorescence change was traced with 4000 points over 6 min in logarithmic scale. Each biological replicate was measured with 6 technical replicates. The technical replicates were

averaged. After normalization of the fluorescence change by its initial fluorescence, i.e., the average fluorescence between 0.001-0.005 s, the biological replicates were averaged. PET efficiencies of each time point were calculated according to

equation 13

$$\text{Trp} - \text{PET efficiency} = 1 - \frac{F_W}{F_F}$$

where $F_W$ is the fluorescence originating from the wt sample containing the Trp and $F_F$ from the Phe control.

### 5.12. PET-FCS

PET-FCS measurements were done on a MicroTime 200 confocal microscope (PicoQuant). Sample was excited with a 636.5 nm laser using an intensity of 77% adjusting the filters to approximately 25000 counts (~ 40 µW). The excitation beam was reflected on a 485+640 dichroic mirror onto the sample. The emission passes through the same filter before it is focused by a 50 µm pinhole. The emission signal was split 50/50 on a beam splitter. Both beams were passed through H690/70 bandpass filters and detected on single photon avalanche diodes. The detection of the emission by two detectors allows the cross-correlation of the fluorescence trace and reduces the noise caused by afterpulsing of the detector. This in particular allows the accurate analysis of dynamics with fast relaxation times. Atto655-labeled purified RNC (prepared according to 5.8.3) diluted in HiFi was used. The concentration was kept low (1-10 nM) to have on average 1 molecule in the focal volume. All experiments were done with 3 biological replicates, measured for 4x10 min at room temperature. Fluorescence traces were analyzed with the software SymPhoTime 64 (PicoQuant). Each fluorescence trace was checked for peaks originating from aggregates. Aggregate containing sections were removed by considering the parts on the left and the right of the fluorescence trace separately. The ACF of every fluorescence trace was calculated with 314 points for correlation times between 0 and 1 s. All ACFs were initially normalized by setting the time point at 0.1 ms to 1. The ACF of the technical and biological replicates were averaged and analyzed in Graphpad Prism (Version 8.3.0) with a 2D model, according to

equation 14

$$G(\tau) = \left(1 + c_1 e^{-k_1 \tau} + c_2 e^{-k_2 \tau}\right)\left(\frac{1 - F + F e^{-k_f \tau}}{1 - F}\right)\left(\frac{1}{N}\right)\left(\frac{1}{1 + k_d \tau}\right)$$

The parameter $c_x$ describes the amplitude of the respective term for fast dynamic and $k_x$ the corresponding relaxation time. F reflects the fraction of the triplet state and $k_f$ the corresponding relaxation time. N is the average number of molecules in the focal volume and $k_d$ the diffusion rate.

The fitted average number of molecules N was used to normalize the autocorrelation function to N=1 by

$G(\tau) * N.$

Final ACFs were fitted again with equation 14. In order to quantify the fraction of fluorescent molecules, the Y-intercept G(0) was calculated according to

equation 15

$$G(0) = (1 + c_1 + c_2)\left(\frac{1}{1-F}\right)\left(\frac{1}{N}\right)$$

The error of the Y-intercept was propagated according to

equation 16

$$\Delta G(0) = \sqrt{\begin{array}{c}\left(\left(\frac{1}{1-F}\right)\left(\frac{1}{N}\right) * \Delta c_1\right)^2 + \left(\left(\frac{1}{1-F}\right)\left(\frac{1}{N}\right) * \Delta c_2\right)^2 + \\ \left((1 + c_1 + c_2)\left(\frac{1}{N}\right)\left(\frac{1}{(1-F)^2}\right) * \Delta F\right)^2 + \left((1 + c_1 + c_2)\left(\frac{1}{1-F}\right)\left(-\frac{1}{N^2}\right) * \Delta N\right)^2\end{array}}$$

The standard error of the mean (SEM) was used for propagating the error ($\Delta x$) of the fitted parameters.

# 6. References

Agirrezabala, X., Samatova, E., Klimova, M., Zamora, M., Gil-Carton, D., Rodnina, M.V., and Valle, M. (2017). Ribosome rearrangements at the onset of translational bypassing. Sci Adv *3*, e1700147.

Agirrezabala, X., Samatova, E., Macher, M., Liutkute, M., Maiti, M., Gil-Carton, D., Novacek, J., Valle, M., and Rodnina, M.V. (2022). A switch from alpha-helical to beta-strand conformation during co-translational protein folding. EMBO J *41*, e109175.

Agris, P.F. (2008). Bringing order to translation: the contributions of transfer RNA anticodon-domain modifications. Embo Rep *9*, 629-635.

Agris, P.F., Eruysal, E.R., Narendran, A., Vare, V.Y.P., Vangaveti, S., and Ranganathan, S.V. (2018). Celebrating wobble decoding: Half a century and still much is new. Rna Biology *15*, 537-553.

Agris, P.F., Vendeix, F.A., and Graham, W.D. (2007). tRNA's wobble decoding of the genome: 40 years of modification. J Mol Biol *366*, 1-13.

Ahn, M., Wlodarski, T., Mitropoulou, A., Chan, S.H.S., Sidhu, H., Plessa, E., Becker, T.A., Budisa, N., Waudby, C.A., Beckmann, R., *et al.* (2022). Modulating co-translational protein folding by rational design and ribosome engineering. Nat Commun *13*, 4243.

Anfinsen, C.B. (1973). Principles that govern the folding of protein chains. Science *181*, 223-230.

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 A resolution. Science *289*, 905-920.

Bhatia, S., and Udgaonkar, J.B. (2022). Heterogeneity in Protein Folding and Unfolding Reactions. Chem Rev *122*, 8911-8935.

Bhushan, S., Gartmann, M., Halic, M., Armache, J.P., Jarasch, A., Mielke, T., Berninghausen, O., Wilson, D.N., and Beckmann, R. (2010). alpha-Helical nascent polypeptide chains visualized within distinct regions of the ribosomal exit tunnel. Nat Struct Mol Biol *17*, 313-317.

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*, e1000581.

Boycheva, S., Chkodrov, G., and Ivanov, I. (2003). Codon pairs in the genome of Escherichia coli. Bioinformatics *19*, 987-998.

Braselmann, E., Chaney, J.L., and Clark, P.L. (2013). Folding the proteome. Trends Biochem Sci *38*, 337-344.

Broom, A., Gosavi, S., and Meiering, E.M. (2015). Protein unfolding rates correlate as strongly as folding rates with native structure. Protein Sci *24*, 580-587.

Buchan, J.R., Aucott, L.S., and Stansfield, I. (2006). tRNA properties help shape codon pair preferences in open reading frames. Nucleic Acids Res *34*, 1015-1027.

Buhr, F., Jha, S., Thommen, M., Mittelstaet, J., Kutz, F., Schwalbe, H., Rodnina, M.V., and Komar, A.A. (2016). Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations. Mol Cell *61*, 341-351.

Burkhardt, D.H., Rouskin, S., Zhang, Y., Li, G.W., Weissman, J.S., and Gross, C.A. (2017). Operon mRNAs are organized into ORF-centric structures that predict translation efficiency. Elife *6*.

Buskiewicz, I., Peske, F., Wieden, H.J., Gryczynski, I., Rodnina, M.V., and Wintermeyer, W. (2005). Conformations of the signal recognition particle protein Ffh from Escherichia coli as determined by FRET. J Mol Biol *351*, 417-430.

Buskirk, A.R., and Green, R. (2017). Ribosome pausing, arrest and rescue in bacteria and eukaryotes. Philos Trans R Soc Lond B Biol Sci *372*.

Cabrita, L.D., Cassaignau, A.M.E., Launay, H.M.M., Waudby, C.A., Wlodarski, T., Camilloni, C., Karyadi, M.E., Robertson, A.L., Wang, X., Wentink, A.S., *et al.* (2016). A structural ensemble of a ribosome-nascent chain complex during cotranslational protein folding. Nat Struct Mol Biol *23*, 278-285.

Cannarrozzi, 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.

Carbone, C.E., Demo, G., Madireddy, R., Svidritskiy, E., and Korostelev, A.A. (2020). ArfB can displace mRNA to rescue stalled ribosomes. Nat Commun *11*, 5552.

Chadani, Y., Niwa, T., Izumi, T., Sugata, N., Nagao, A., Suzuki, T., Chiba, S., Ito, K., and Taguchi, H. (2017). Intrinsic Ribosome Destabilization Underlies Translation and Provides an Organism with a Strategy of Environmental Sensing. Mol Cell *68*, 528-539 e525.

Chadani, Y., Sugata, N., Niwa, T., Ito, Y., Iwasaki, S., and Taguchi, H. (2021). Nascent polypeptide within the exit tunnel stabilizes the ribosome to counteract risky translation. EMBO J *40*, e108299.

Chan, K.H., Petrychenko, V., Mueller, C., Maracci, C., Holtkamp, W., Wilson, D.N., Fischer, N., and Rodnina, M.V. (2020). Mechanism of ribosome rescue by alternative ribosome-rescue factor B. Nat Commun *11*, 4106.

Chan, S.H.S., Wlodarski, T., Streit, J.O., Cassaignau, A.M.E., Woodburn, L.F., Ahn, M., Freiherr von Sass, G.J., Waudby, C.A., Budisa, N., Cabrita, L.D., *et al.* (2022). The ribosome stabilizes partially folded intermediates of a nascent multi-domain protein. Nat Chem.

Chandrasekaran, V., Juszkiewicz, S., Choi, J., Puglisi, J.D., Brown, A., Shao, S., Ramakrishnan, V., and Hegde, R.S. (2019). Mechanism of ribosome stalling during translation of a poly(A) tail. Nat Struct Mol Biol *26*, 1132-1140.

Charneski, C.A., and Hurst, L.D. (2013). Positively Charged Residues Are the Major Determinants of Ribosomal Velocity. Plos Biol *11*.

Chartier, M., Gaudreault, F., and Najmanovich, R. (2012). Large-scale analysis of conserved rare codon clusters suggests an involvement in co-translational molecular recognition events. Bioinformatics *28*, 1438-1445.

Chen, X., Rajasekaran, N., Liu, K., and Kaiser, C.M. (2020). Synthesis runs counter to directional folding of a nascent protein domain. Nat Commun *11*, 5096.

Chen, Y.L., and Wen, J.D. (2022). Translation initiation site of mRNA is selected through dynamic interaction with the ribosome. Proc Natl Acad Sci U S A *119*, e2118099119.

Clarke, T.F.t., and Clark, P.L. (2008). Rare codons cluster. Plos One *3*, e3412.

Cortazzo, P., Cerveñansky, C., Marín, M., Reiss, C., Ehrlich, R., and Deana, A. (2002). Silent mutations affect in vivo protein folding in Escherichia coli. Biochemical and Biophysical Research Communications *293*, 537-541.

Cunha, C.E., Belardinelli, R., Peske, F., Holtkamp, W., Wintermeyer, W., and Rodnina, M.V. (2013). Dual use of GTP hydrolysis by elongation factor G on the ribosome. Translation (Austin) *1*, e24315.

Dao Duc, K., Batra, S.S., Bhattacharya, N., Cate, J.H.D., and Song, Y.S. (2019). Differences in the path to exit the ribosome across the three domains of life. Nucleic Acids Res *47*, 4198-4210.

Del Campo, C., Bartholomaus, A., Fedyunin, I., and Ignatova, Z. (2015). Secondary Structure across the Bacterial Transcriptome Reveals Versatile Roles in mRNA Regulation and Function. PLoS Genet *11*, e1005613.

Dill, K.A., and MacCallum, J.L. (2012). The protein-folding problem, 50 years on. Science *338*, 1042-1046.

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., 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.

Doerfel, L.K., Wohlgemuth, I., Kubyshkin, V., Starosta, A.L., Wilson, D.N., Budisa, N., and Rodnina, M.V. (2015). Entropic Contribution of Elongation Factor P to Proline Positioning at the Catalytic Center of the Ribosome. J Am Chem Soc *137*, 12997-13006.

Doose, S., Neuweiler, H., and Sauer, M. (2009). Fluorescence quenching by photoinduced electron transfer: a reporter for conformational dynamics of macromolecules. Chemphyschem *10*, 1389-1398.

Duc, K.D., Batra, S.S., Bhattacharya, N., Cate, J.H.D., and Song, Y.S. (2019). Differences in the path to exit the ribosome across the three domains of life. Nucleic Acids Research *47*, 4198-4210.

Duc, K.D., and Song, Y.S. (2018). The impact of ribosomal interference, codon usage, and exit tunnel interactions on translation elongation rate variation. Plos Genetics *14*.

Dykeman, E.C. (2020). A stochastic model for simulating ribosome kinetics in vivo. Plos Comput Biol *16*.

Eichmann, C., Preissler, S., Riek, R., and Deuerling, E. (2010). Cotranslational structure acquisition of nascent polypeptides monitored by NMR spectroscopy. Proc Natl Acad Sci U S A *107*, 9111-9116.

Elf, J., Nilsson, D., Tenson, T., and Ehrenberg, M. (2003). Selective charging of tRNA isoacceptors explains patterns of codon usage. Science *300*, 1718-1722.

Englander, S.W., and Mayne, L. (2014). The nature of protein folding pathways. Proceedings of the National Academy of Sciences *111*, 15873-15880.

Farias-Rico, J.A., Ruud Selin, F., Myronidi, I., Fruhauf, M., and von Heijne, G. (2018). Effects of protein size, thermodynamic stability, and net charge on cotranslational folding on the ribosome. Proc Natl Acad Sci U S A *115*, E9280-E9287.

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 A resolution by Cs-corrected cryo-EM. Nature *520*, 567-570.

Gamble, C.E., Brule, C.E., Dean, K.M., Fields, S., and Grayhack, E.J. (2016). Adjacent Codons Act in Concert to Modulate Translation Efficiency in Yeast. Cell *166*, 679-690.

Garbuzynskiy, S.O., Ivankov, D.N., Bogatyreva, N.S., and Finkelstein, A.V. (2013). Golden triangle for folding rates of globular proteins. Proc Natl Acad Sci U S A *110*, 147-150.

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.

Goldman, D.H., Kaiser, C.M., Milin, A., Righini, M., Tinoco, I., Jr., and Bustamante, C. (2015). Ribosome. Mechanical force releases nascent chain-mediated ribosome arrest in vitro and in vivo. Science *348*, 457-460.

Goyal, A., Belardinelli, R., Maracci, C., Milon, P., and Rodnina, M.V. (2015). Directional transition from initiation to elongation in bacterial translation. Nucleic Acids Res *43*, 10700-10712.

Goyal, A., Belardinelli, R., and Rodnina, M.V. (2017). Non-canonical Binding Site for Bacterial Initiation Factor 3 on the Large Ribosomal Subunit. Cell Rep *20*, 3113-3122.

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.

Grosjean, H., and Westhof, E. (2016). An integrated, structure- and energy-based view of the genetic code. Nucleic Acids Res *44*, 8020-8040.

Gualerzi, C.O., and Pon, C.L. (2015). Initiation of mRNA translation in bacteria: structural and dynamic aspects. Cell Mol Life Sci *72*, 4341-4367.

Guinn, E.J., Tian, P., Shin, M., Best, R.B., and Marqusee, S. (2018). A small single-domain protein folds through the same pathway on and off the ribosome. Proc Natl Acad Sci U S A *115*, 12206-12211.

Guo, F.B., Ye, Y.N., Zhao, H.L., Lin, D., and Wei, W. (2012). Universal Pattern and Diverse Strengths of Successive Synonymous Codon Bias in Three Domains of Life, Particularly Among Prokaryotic Genomes. DNA Res *19*, 477-485.

Han, Y., David, A., Liu, B., Magadan, J.G., Bennink, J.R., Yewdell, J.W., and Qian, S.B. (2012). Monitoring cotranslational protein folding in mammalian cells at codon resolution. Proc Natl Acad Sci U S A *109*, 12467-12472.

Heurgue-Hamard, V., Champ, S., Engstrom, A., Ehrenberg, M., and Buckingham, R.H. (2002). The hemK gene in Escherichia coli encodes the N(5)-glutamine methyltransferase that modifies peptide release factors. EMBO J *21*, 769-778.

Holtkamp, W., Kokic, G., Jager, M., Mittelstaet, J., Komar, A.A., and Rodnina, M.V. (2015). Cotranslational protein folding on the ribosome monitored in real time. Science *350*, 1104-1107.

Hou, Y.M., Masuda, I., and Gamper, H. (2019). Codon-Specific Translation by m(1)G37 Methylation of tRNA. Front Genet *9*.

Hu, S., Wang, M., Cai, G., and He, M. (2013). Genetic code-guided protein synthesis and folding in Escherichia coli. J Biol Chem *288*, 30855-30861.

Huter, P., Arenz, S., Bock, L.V., Graf, M., Frister, J.O., Heuer, A., Peil, L., Starosta, A.L., Wohlgemuth, I., Peske, F., *et al.* (2017a). Structural Basis for Polyproline-Mediated Ribosome Stalling and Rescue by the Translation Elongation Factor EF-P. Mol Cell *68*, 515-527 e516.

Huter, P., Muller, C., Arenz, S., Beckert, B., and Wilson, D.N. (2017b). Structural Basis for Ribosome Rescue in Bacteria. Trends Biochem Sci *42*, 669-680.

Huter, P., Muller, C., Beckert, B., Arenz, S., Berninghausen, O., Beckmann, R., and Wilson, D.N. (2017c). Structural basis for ArfA-RF2-mediated translation termination on mRNAs lacking stop codons. Nature *541*, 546-549.

Ingolia, N.T., Ghaemmaghami, 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.

Ismail, N., Hedman, R., Schiller, N., and von Heijne, G. (2012). A biphasic pulling force acts on transmembrane helices during translocon-mediated membrane integration. Nat Struct Mol Biol *19*, 1018-1022.

Ito, K., and Chiba, S. (2013). Arrest peptides: cis-acting modulators of translation. Annu Rev Biochem *82*, 171-202.

Juhling, F., Morl, M., Hartmann, R.K., Sprinzl, M., Stadler, P.F., and Putz, J. (2009). tRNAdb 2009: compilation of tRNA sequences and tRNA genes. Nucleic Acids Res *37*, D159-162.

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

Kathuria, S.V., Day, I.J., Wallace, L.A., and Matthews, C.R. (2008). Kinetic traps in the folding of beta alpha-repeat proteins: CheY initially misfolds before accessing the native conformation. J Mol Biol *382*, 467-484.

Kemp, G., Kudva, R., de la Rosa, A., and von Heijne, G. (2019). Force-Profile Analysis of the Cotranslational Folding of HemK and Filamin Domains: Comparison of Biochemical and Biophysical Folding Assays. J Mol Biol *431*, 1308-1314.

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.

Kirchner, S., Cai, Z., Rauscher, R., Kastelic, N., Anding, M., Czech, A., Kleizen, B., Ostedgaard, L.S., Braakman, I., Sheppard, D.N., *et al.* (2017). Alteration of protein function by a silent polymorphism linked to tRNA abundance. PLoS Biol *15*, e2000779.

Korostelev, A.A. (2011). Structural aspects of translation termination on the ribosome. RNA *17*, 1409-1421.

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. (2007). Codon reading by tRNAAla with modified uridine in the wobble position. Mol Cell *25*, 167-174.

Koutmou, K.S., Schuller, A.P., Brunelle, J.L., Radhakrishnan, A., Djuranovic, S., and Green, R. (2015). Ribosomes slide on lysine-encoding homopolymeric A stretches. Elife *4*.

Kramer, G., Shiber, A., and Bukau, B. (2019). Mechanisms of Cotranslational Maturation of Newly Synthesized Proteins. Annu Rev Biochem *88*, 337-364.

Krichevsky, O., and Bonnet, G. (2002). Fluorescence correlation spectroscopy: the technique and its applications. Reports on Progress in Physics *65*, 251-297.

Kudva, R., Tian, P., Pardo-Avila, F., Carroni, M., Best, R.B., Bernstein, H.D., and von Heijne, G. (2018). The shape of the bacterial ribosome exit tunnel affects cotranslational protein folding. Elife *7*.

Kyte, J., and Doolittle, R.F. (1982). A simple method for displaying the hydropathic character of a protein. Journal of Molecular Biology *157*, 105-132.

Lakowicz, J.R. (2006). Principles of Fluorescence Spectroscopy, 3 edn (Springer New York, NY).

Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S., and Noller, H.F. (2008). Structural basis for translation termination on the 70S ribosome. Nature *454*, 852-857.

Ledoux, S., and Uhlenbeck, O.C. (2008). Different aa-tRNAs are selected uniformly on the ribosome. Mol Cell *31*, 114-123.

Leininger, S.E., Rodriguez, J., Vu, Q.V., Jiang, Y., Li, M.S., Deutsch, C., and O'Brien, E.P. (2021). Ribosome Elongation Kinetics of Consecutively Charged Residues Are Coupled to Electrostatic Force. Biochemistry *60*, 3223-3235.

Lim, V.I., and Curran, J.F. (2001). Analysis of codon : anticodon interactions within the ribosome provides new insights into codon reading and the genetic code structure. Rna *7*, 942-957.

Liu, K., Maciuba, K., and Kaiser, C.M. (2019). The Ribosome Cooperates with a Chaperone to Guide Multi-domain Protein Folding. Mol Cell *74*, 310-319 e317.

Liu, K., Rehfus, J.E., Mattson, E., and Kaiser, C.M. (2017). The ribosome destabilizes native and non-native structures in a nascent multidomain protein. Protein Sci *26*, 1439-1451.

Liutkute, M., Maiti, M., Samatova, E., Enderlein, J., and Rodnina, M.V. (2020a). Gradual compaction of the nascent peptide during cotranslational folding on the ribosome. Elife *9*.

Liutkute, M., Samatova, E., and Rodnina, M.V. (2020b). Cotranslational Folding of Proteins on the Ribosome. Biomolecules *10*.

Lu, J., and Deutsch, C. (2005). Folding zones inside the ribosomal exit tunnel. Nat Struct Mol Biol *12*, 1123-1129.

Lu, J.L., and Deutsch, C. (2008). Electrostatics in the Ribosomal Tunnel Modulate Chain Elongation Rates. Journal of Molecular Biology *384*, 73-86.

Luitz, M.P., Barth, A., Crevenna, A.H., Bomblies, R., Lamb, D.C., and Zacharias, M. (2017). Covalent dye attachment influences the dynamics and conformational properties of flexible peptides. PLoS One *12*, e0177139.

Makhoul, C.H., and Trifonov, E.N. (2002). Distribution of rare triplets along mRNA and their relation to protein folding. J Biomol Struct Dyn *20*, 413-420.

Maracci, C., and Rodnina, M.V. (2016). Review: Translational GTPases. Biopolymers *105*, 463-475.

Marino, J., Buholzer, K.J., Zosel, F., Nettels, D., and Schuler, B. (2018). Charge Interactions Can Dominate Coupled Folding and Binding on the Ribosome. Biophys J *115*, 996-1006.

Melnikov, S., Ben-Shem, A., Garreau de Loubresse, N., Jenner, L., Yusupova, G., and Yusupov, M. (2012). One core, two shells: bacterial and eukaryotic ribosomes. Nat Struct Mol Biol *19*, 560-567.

Mercier, E., and Rodnina, M.V. (2018). Co-Translational Folding Trajectory of the HemK Helical Domain. Biochemistry *57*, 3460-3464.

Mercier, E., Wintermeyer, W., and Rodnina, M.V. (2020). Co-translational insertion and topogenesis of bacterial membrane proteins monitored in real time. EMBO J *39*, e104054.

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. In Translation Initiation: Reconstituted Systems and Biophysical Methods, pp. 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.

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.

Mohammad, F., Green, R., and Buskirk, A.R. (2019). A systematically-revised ribosome profiling method for bacteria reveals pauses at single-codon resolution. Elife *8*.

Moura, G., Pinheiro, M., Silva, R., Miranda, I., Afreixo, V., Dias, G., Freitas, A., Oliveira, J.L., and Santos, M.A. (2005). Comparative context analysis of codon pairs on an ORFeome scale. Genome Biol *6*, R28.

Muller, C., Crowe-McAuliffe, C., and Wilson, D.N. (2021). Ribosome Rescue Pathways in Bacteria. Front Microbiol *12*, 652980.

Mustoe, A.M., Busan, S., Rice, G.M., Hajdin, C.E., Peterson, B.K., Ruda, V.M., Kubica, N., Nutiu, R., Baryza, J.L., and Weeks, K.M. (2018). Pervasive Regulatory Functions of mRNA Structure Revealed by High-Resolution SHAPE Probing. Cell *173*, 181-+.

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.

Navon, S.P., Kornberg, G., Chen, J., Schwartzman, T., Tsai, A., Puglisi, E.V., Puglisi, J.D., and Adir, N. (2016). Amino acid sequence repertoire of the bacterial proteome and the occurrence of untranslatable sequences. Proc Natl Acad Sci U S A *113*, 7166-7170.

Nedialkova, D.D., and Leidel, S.A. (2015). Optimization of Codon Translation Rates via tRNA Modifications Maintains Proteome Integrity. Cell *161*, 1606-1618.

Neuweiler, H., Lollmann, M., Doose, S., and Sauer, M. (2007). Dynamics of unfolded polypeptide chains in crowded environment studied by fluorescence correlation spectroscopy. J Mol Biol *365*, 856-869.

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.

Nick Pace, C., Scholtz, J.M., and Grimsley, G.R. (2014). Forces stabilizing proteins. FEBS Lett *588*, 2177-2184.

Nilsson, O.B., Hedman, R., Marino, J., Wickles, S., Bischoff, L., Johansson, M., Muller-Lucks, A., Trovato, F., Puglisi, J.D., O'Brien, E.P., *et al.* (2015). Cotranslational Protein Folding inside the Ribosome Exit Tunnel. Cell Rep *12*, 1533-1540.

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.

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.

Oh, E., Becker, A.H., Sandikci, A., Huber, D., Chaba, R., Gloge, F., Nichols, R.J., Typas, A., Gross, C.A., Kramer, G., *et al.* (2011). Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell *147*, 1295-1308.

Peil, L., Starosta, A.L., Lassak, J., Atkinson, G.C., Virumae, K., Spitzer, M., Tenson, T., Jung, K., Remme, J., and Wilson, D.N. (2013). Distinct XPPX sequence motifs induce ribosome stalling, which is rescued by the translation elongation factor EF-P. P Natl Acad Sci USA *110*, 15265-15270.

Petrychenko, V., Peng, B.Z., de, A.P.S.A.C., Peske, F., Rodnina, M.V., and Fischer, N. (2021). Structural mechanism of GTPase-powered ribosome-tRNA movement. Nat Commun *12*, 5933.

Pierson, W.E., Hoffer, E.D., Keedy, H.E., Simms, C.L., Dunham, C.M., and Zaher, H.S. (2016). Uniformity of Peptide Release Is Maintained by Methylation of Release Factors. Cell Rep *17*, 11-18.

Plaxco, K.W., Simons, K.T., Ruczinski, I., and Baker, D. (2000). Topology, stability, sequence, and length: defining the determinants of two-state protein folding kinetics. Biochemistry *39*, 11177-11183.

Pollo-Oliveira, L., Klassen, R., Davis, N., Ciftci, A., Bacusmo, J.M., Martinelli, M., DeMott, M.S., Begley, T.J., Dedon, P.C., Schaffrath, R., *et al.* (2020). Loss of Elongator- and KEOPS-Dependent tRNA Modifications Leads to Severe Growth Phenotypes and Protein Aggregation in Yeast. Biomolecules *10*.

Qu, X.H., Wen, J.D., Lancaster, L., Noller, H.F., Bustamante, C., and Tinoco, I. (2011). The ribosome uses two active mechanisms to unwind messenger RNA during translation. Nature *475*, 118-121.

Rajab, S., Bismin, L., Schwarze, S., Pinggera, A., Greger, I.H., and Neuweiler, H. (2021). Allosteric coupling of sub-millisecond clamshell motions in ionotropic glutamate receptor ligand-binding domains. Commun Biol *4*, 1056.

Ranjan, N., and Leidel, S.A. (2019). The epitranscriptome in translation regulation: mRNA and tRNA modifications as the two sides of the same coin? Febs Lett *593*, 1483-1493.

Ranjan, N., and Rodnina, M.V. (2016). tRNA wobble modifications and protein homeostasis. Translation *4*.

Rezgui, V.A., Tyagi, K., Ranjan, N., Konevega, A.L., Mittelstaet, J., Rodnina, M.V., Peter, M., and Pedrioli, P.G. (2013). tRNA tKUUU, tQUUG, and tEUUC wobble position modifications fine-tune protein translation by promoting ribosome A-site binding. Proc Natl Acad Sci U S A *110*, 12289-12294.

Rodnina, M.V. (2018). Translation in Prokaryotes. Cold Spring Harb Perspect Biol *10*.

Rodnina, M.V., Fischer, N., Maracci, C., and Stark, H. (2017). Ribosome dynamics during decoding. Philos Trans R Soc Lond B Biol Sci *372*.

Rodnina, M.V., Gromadski, K.B., Kothe, U., and Wieden, H.J. (2005). Recognition and selection of tRNA in translation. Febs Lett *579*, 938-942.

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.

Rudorf, S. (2019). Efficiency of protein synthesis inhibition depends on tRNA and codon compositions. Plos Comput Biol *15*.

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.

Sabi, R., and Tuller, T. (2015). A comparative genomics study on the effect of individual amino acids on ribosome stalling. BMC Genomics *16 Suppl 10*, S5.

Sander, I.M., Chaney, J.L., and Clark, P.L. (2014). Expanding Anfinsen's principle: contributions of synonymous codon selection to rational protein design. J Am Chem Soc *136*, 858-861.

Schagger, H. (2006). Tricine-SDS-PAGE. Nat Protoc *1*, 16-22.

Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Analytical Biochemistry *166*, 368-379.

Schmeing, T.M., and Ramakrishnan, V. (2009). What recent ribosome structures have revealed about the mechanism of translation. Nature *461*, 1234-1242.

Seely, S.M., and Gagnon, M.G. (2022). Mechanisms of ribosome recycling in bacteria and mitochondria: a structural perspective. RNA Biol *19*, 662-677.

Seip, B., Sacheau, G., Dupuy, D., and Innis, C.A. (2018). Ribosomal stalling landscapes revealed by high-throughput inverse toeprinting of mRNA libraries. Life Sci Alliance *1*.

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*.

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.

Sorokina, I., Mushegian, A.R., and Koonin, E.V. (2022). Is Protein Folding a Thermodynamically Unfavorable, Active, Energy-Dependent Process? Int J Mol Sci *23*.

Stagg, L., Samiotakis, A., Homouz, D., Cheung, M.S., and Wittung-Stafshede, P. (2010). Residue-specific analysis of frustration in the folding landscape of repeat beta/alpha protein apoflavodoxin. J Mol Biol *396*, 75-89.

Su, T., Cheng, J., Sohmen, D., Hedman, R., Berninghausen, O., von Heijne, G., Wilson, D.N., and Beckmann, R. (2017). The force-sensing peptide VemP employs extreme compaction and secondary structure formation to induce ribosomal stalling. Elife *6*.

Subramaniam, A.R., Zid, B.M., and O'Shea, E.K. (2014). An integrated approach reveals regulatory controls on bacterial translation elongation. Cell *159*, 1200-1211.

Takyar, S., Hickerson, R.P., and Noller, H.F. (2005). mRNA helicase activity of the ribosome. Cell *120*, 49-58.

Tats, A., Tenson, T., and Remm, M. (2008). Preferred and avoided codon pairs in three domains of life. BMC Genomics *9*, 463.

Tesina, P., Lessen, L.N., Buschauer, R., Cheng, J., Wu, C.C., Berninghausen, O., Buskirk, A.R., Becker, T., Beckmann, R., and Green, R. (2020). Molecular mechanism of translational stalling by inhibitory codon combinations and poly(A) tracts. EMBO J *39*, e103365.

Thiaville, P.C., El Yacoubi, B., Kohrer, C., Thiaville, J.J., Deutsch, C., Iwata-Reuyl, D., Bacusmo, J.M., Armengaud, J., Bessho, Y., Wetzel, C., *et al.* (2015). Essentiality of threonylcarbamoyladenosine (t(6)A), a universal tRNA modification, in bacteria. Mol Microbiol *98*, 1199-1221.

Thommen, M., Draycheva, A., and Rodnina, M.V. (2022). Ribosome selectivity and nascent chain context in modulating the incorporation of fluorescent non-canonical amino acid into proteins. Sci Rep *12*, 12848.

To, P., Whitehead, B., Tarbox, H.E., and Fried, S.D. (2021). Nonrefoldability is Pervasive Across the E. coli Proteome. J Am Chem Soc *143*, 11435-11448.

To, P., Xia, Y., Lee, S.O., Devlin, T., Fleming, K.G., and Fried, S.D. (2022). A proteome-wide map of chaperone-assisted protein refolding in a cytosol-like milieu. Proc Natl Acad Sci U S A *119*, e2210536119.

Tuller, T., Veksler-Lublinsky, I., Gazit, N., Kupiec, M., Ruppin, E., and Ziv-Ukelson, M. (2011). Composite effects of gene determinants on the translation speed and density of ribosomes. Genome Biology *12*.

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.

Vieira, J.P., Racle, J., and Hatzimanikatis, V. (2016). Analysis of Translation Elongation Dynamics in the Context of an Escherichia coli Cell. Biophys J *110*, 2120-2131.

von Heijne, G. (1989). Control of topology and mode of assembly of a polytopic membrane protein by positively charged residues. Nature *341*, 456-458.

Walsh, I.M., Bowman, M.A., Soto Santarriaga, I.F., Rodriguez, A., and Clark, P.L. (2020). Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness. Proc Natl Acad Sci U S A *117*, 3528-3534.

Wang, J., and Panagiotou, E. (2022). The protein folding rate and the geometry and topology of the native state. Sci Rep *12*, 6384.

Weinberg, D.E., Shah, P., Eichhorn, S.W., Hussmann, J.A., Plotkin, J.B., and Bartel, D.P. (2016). Improved Ribosome-Footprint and mRNA Measurements Provide Insights into Dynamics and Regulation of Yeast Translation. Cell Rep *14*, 1787-1799.

Weixlbaumer, A., Murphy, F.V., Dziergowska, A., Malkiewicz, A., Vendeix, F.A.P., Agris, P.F., and Ramakrishnan, V. (2007). Mechanism for expanding the decoding capacity of transfer RNAs by modification of uridines. Nat Struct Mol Biol *14*, 498-502.

Wieden, H.J., Gromadski, K., Rodnin, D., and Rodnina, M.V. (2002). Mechanism of elongation factor (EF)-Ts-catalyzed nucleotide exchange in EF-Tu. Contribution of contacts at the guanine base. J Biol Chem *277*, 6032-6036.

Wilson, D.N., Arenz, S., and Beckmann, R. (2016). Translation regulation via nascent polypeptide-mediated ribosome stalling. Curr Opin Struc Biol *37*, 123-133.

Wohlgemuth, I., Brenne, 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.

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. P Natl Acad Sci USA *110*, E878-E887.

Wruck, F., Tian, P., Kudva, R., Best, R.B., von Heijne, G., Tans, S.J., and Katranidis, A. (2021). The ribosome modulates folding inside the ribosomal exit tunnel. Commun Biol *4*, 523.

Wu, C.C.C., Zinshteyn, B., Wehner, K.A., and Green, R. (2019). High-Resolution Ribosome Profiling Defines Discrete Ribosome Elongation States and Translational Regulation during Cellular Stress. Mol Cell *73*, 959-+.

Wu, P., and Brand, L. (1994). Resonance energy transfer: methods and applications. Anal Biochem *218*, 1-13.

Xu, Y., Ma, P., Shah, P., Rokas, A., Liu, Y., and Johnson, C.H. (2013). Non-optimal codon usage is a mechanism to achieve circadian clock conditionality. Nature *495*, 116-120.

Yang, Z., Shipman, L., Zhang, M., Anton, B.P., Roberts, R.J., and Cheng, X. (2004). Structural characterization and comparative phylogenetic analysis of Escherichia coli HemK, a protein (N5)-glutamine methyltransferase. J Mol Biol *340*, 695-706.

Youngman, E.M., McDonald, M.E., and Green, R. (2008). Peptide release on the ribosome: mechanism and implications for translational control. Annu Rev Microbiol *62*, 353-373.

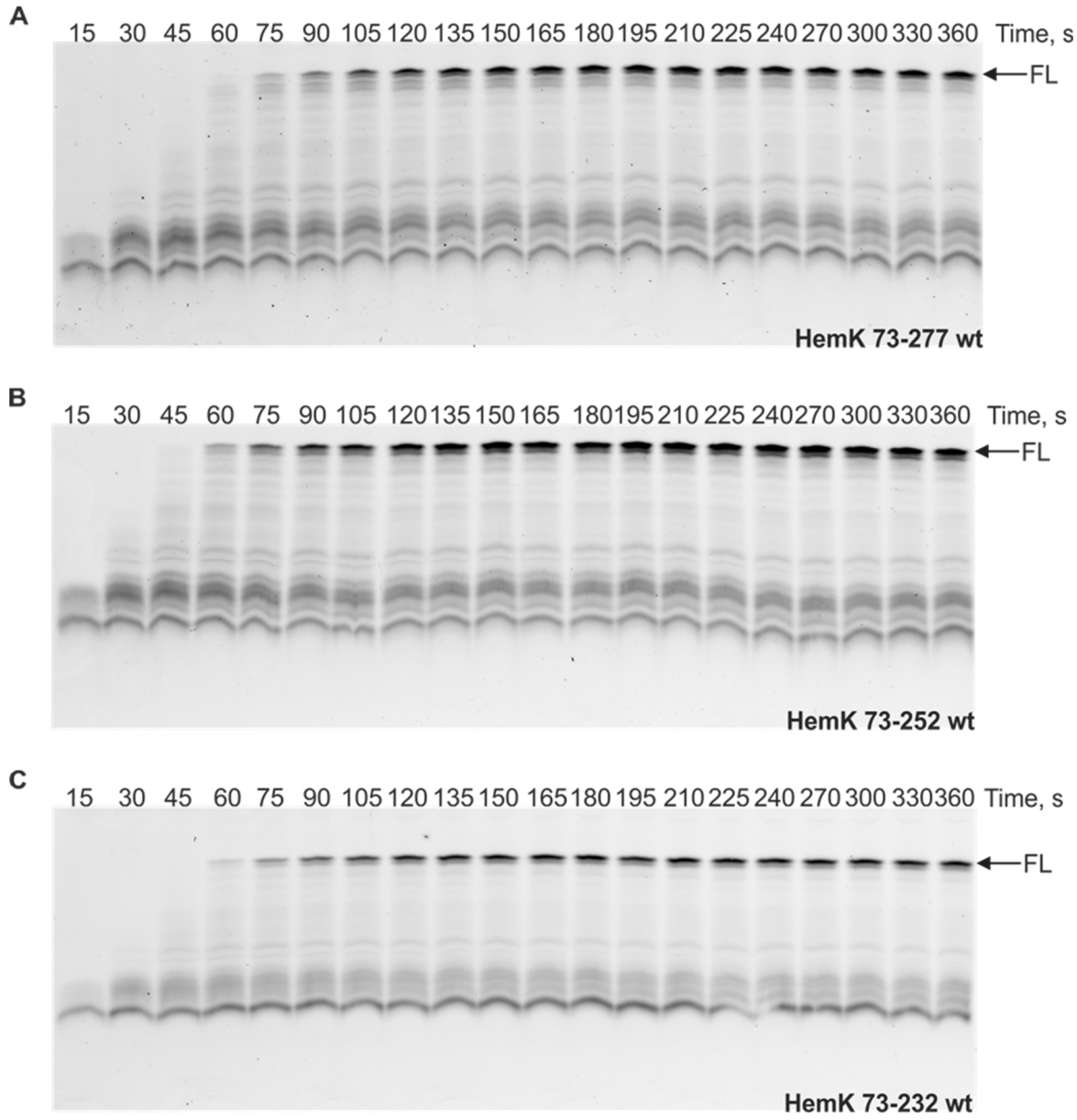
Zhang, G., Hubalewska, M., and Ignatova, Z. (2009). Transient ribosomal attenuation coordinates protein synthesis and co-translational folding. Nat Struct Mol Biol *16*, 274-280.

Zhang, J., Pan, X., Yan, K., Sun, S., Gao, N., and Sui, S.F. (2015). Mechanisms of ribosome stalling by SecM at multiple elongation steps. Elife *4*.

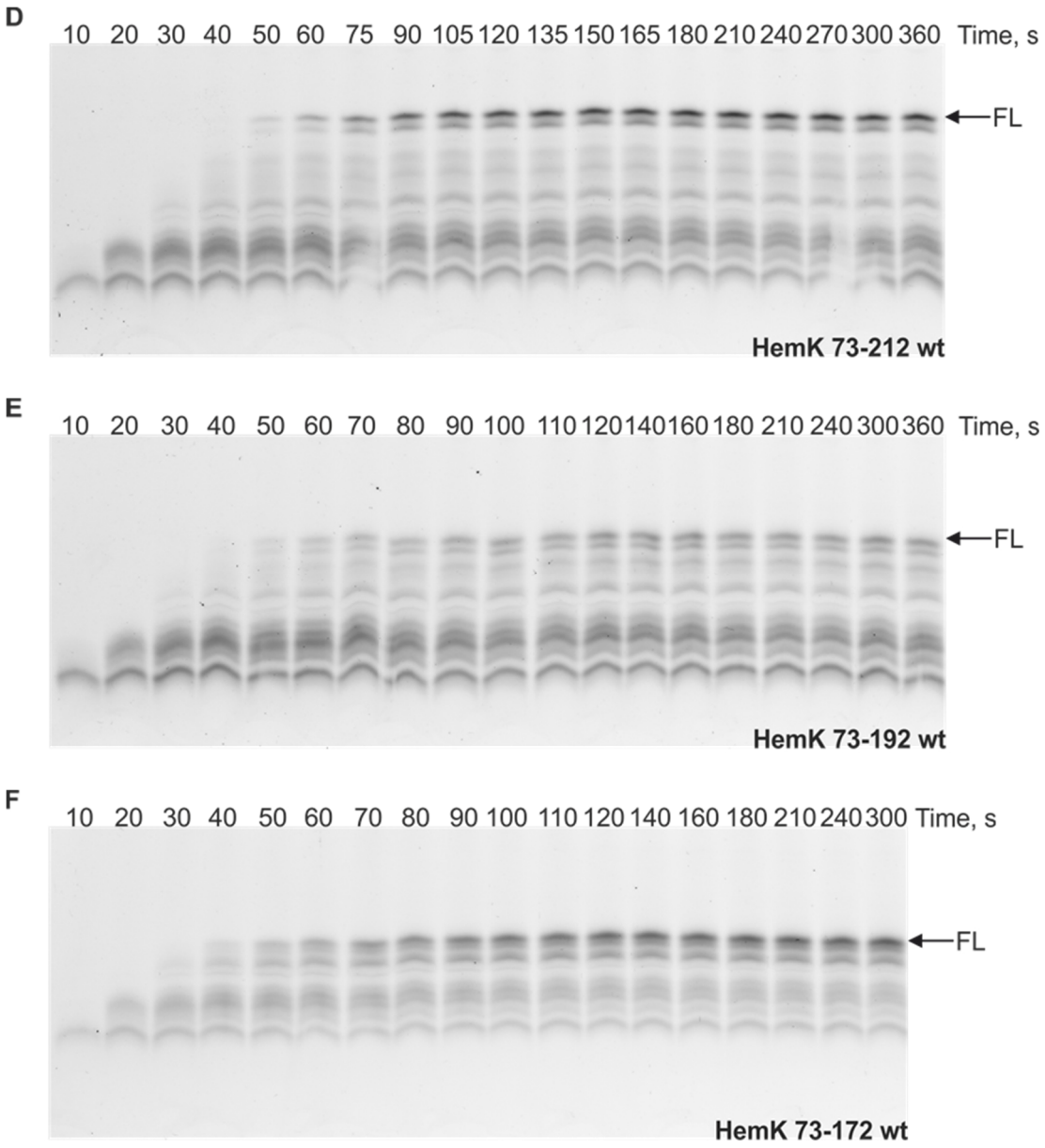
Zhou, M., Guo, J., Cha, J., Chae, M., Chen, S., Barral, J.M., Sachs, M.S., and Liu, Y. (2013). Non-optimal codon usage affects expression, structure and function of clock protein FRQ. Nature *495*, 111-115.

Zinshteyn, B., and Gilbert, W.V. (2013). Loss of a conserved tRNA anticodon modification perturbs cellular signaling. PLoS Genet *9*, e1003675.

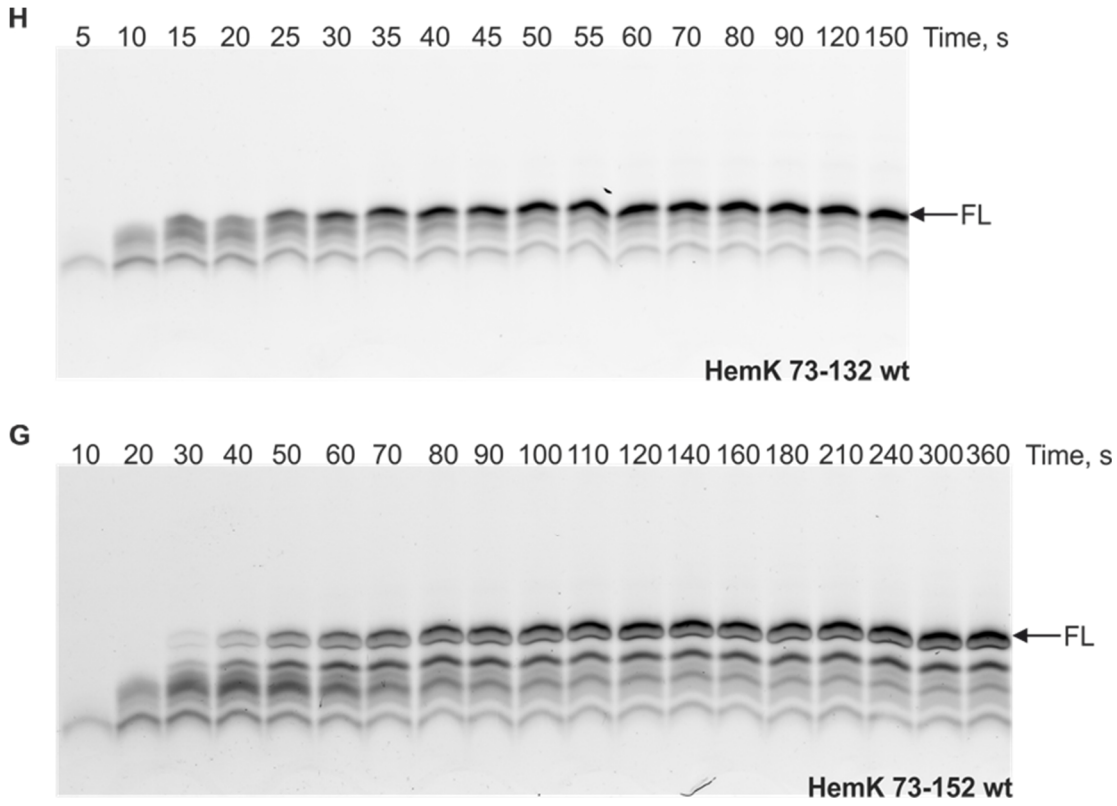
# 7. Appendix



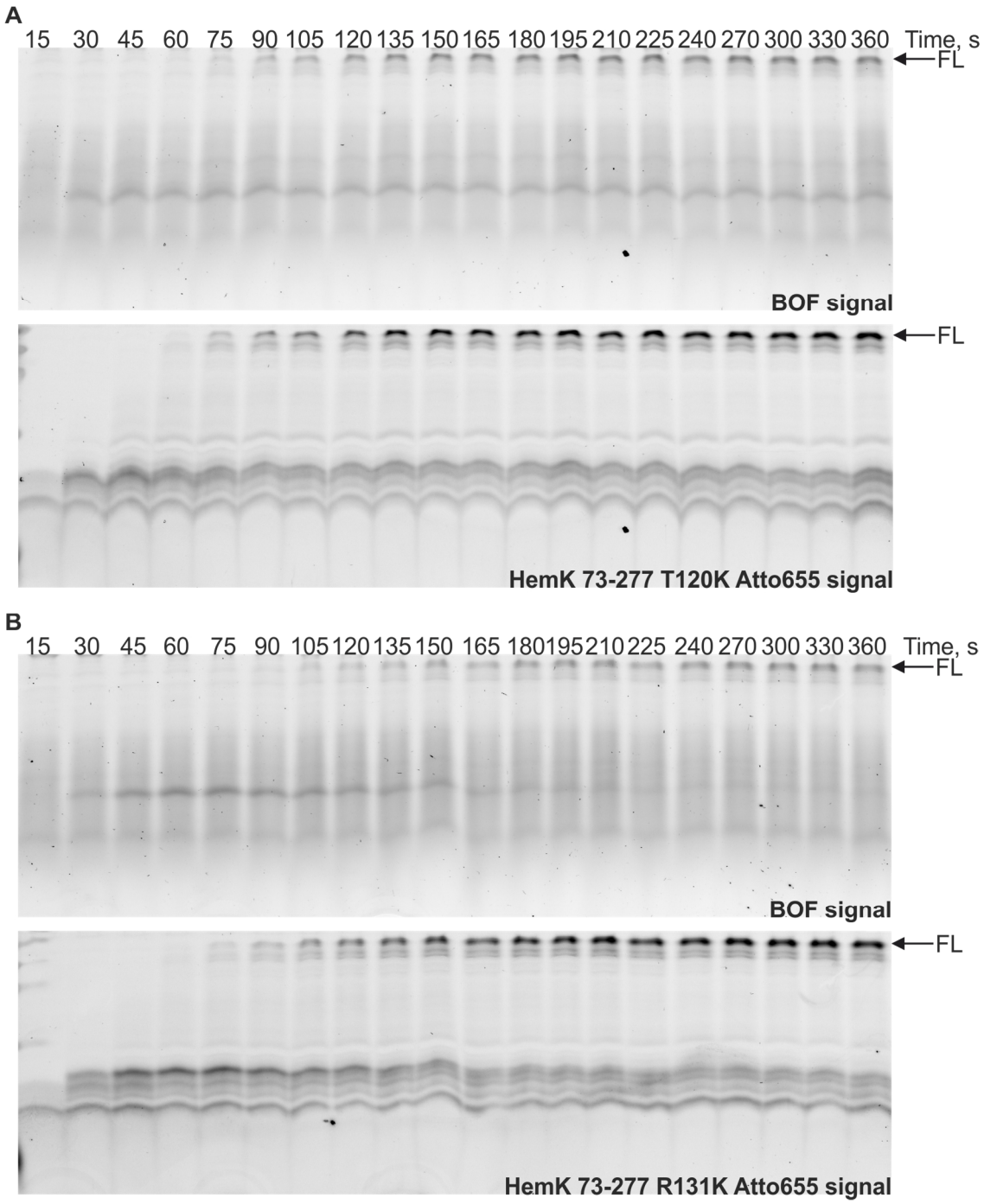

Appendix Figure 1. Continues on following page.



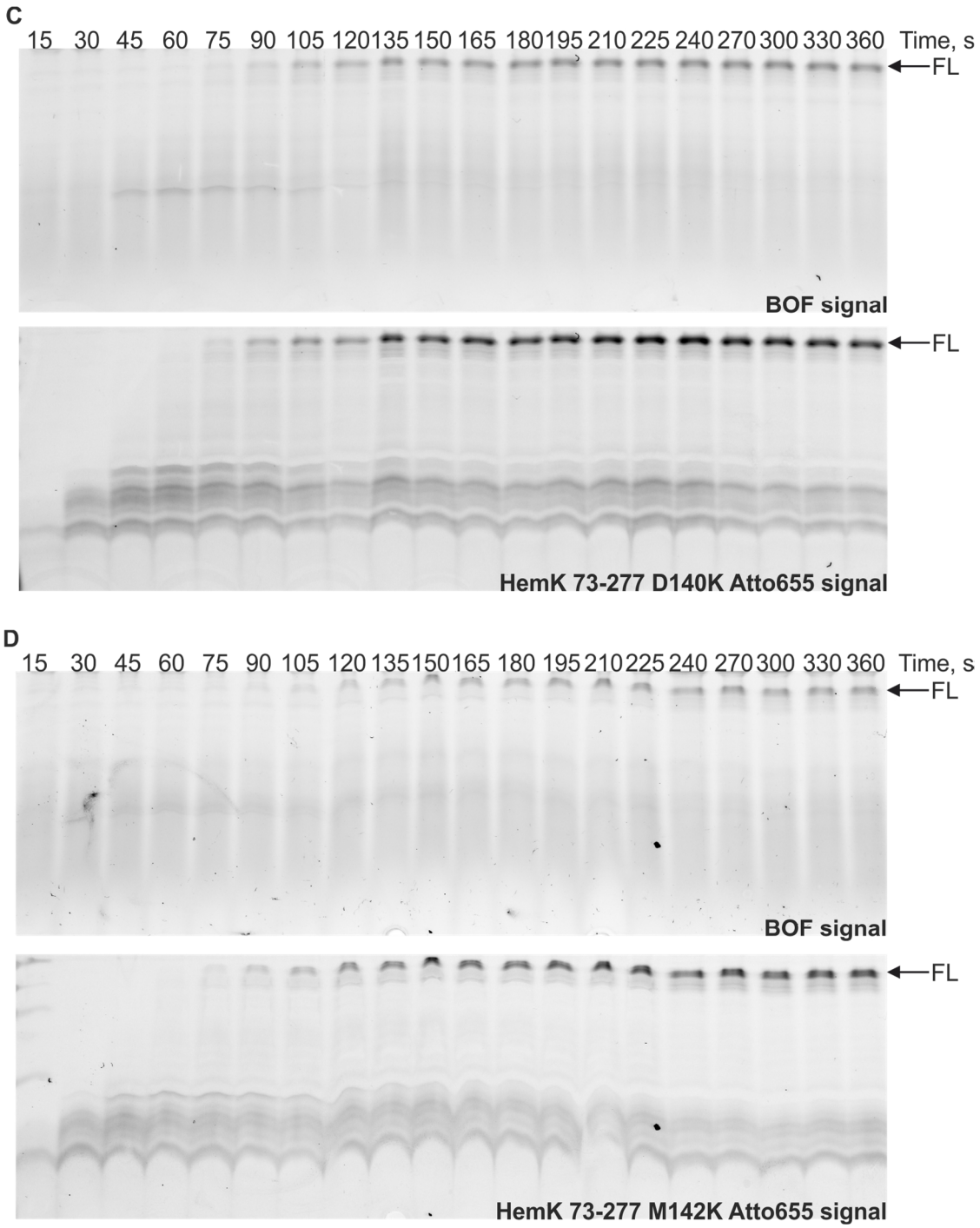

Appendix Figure 1. Continues on following page.



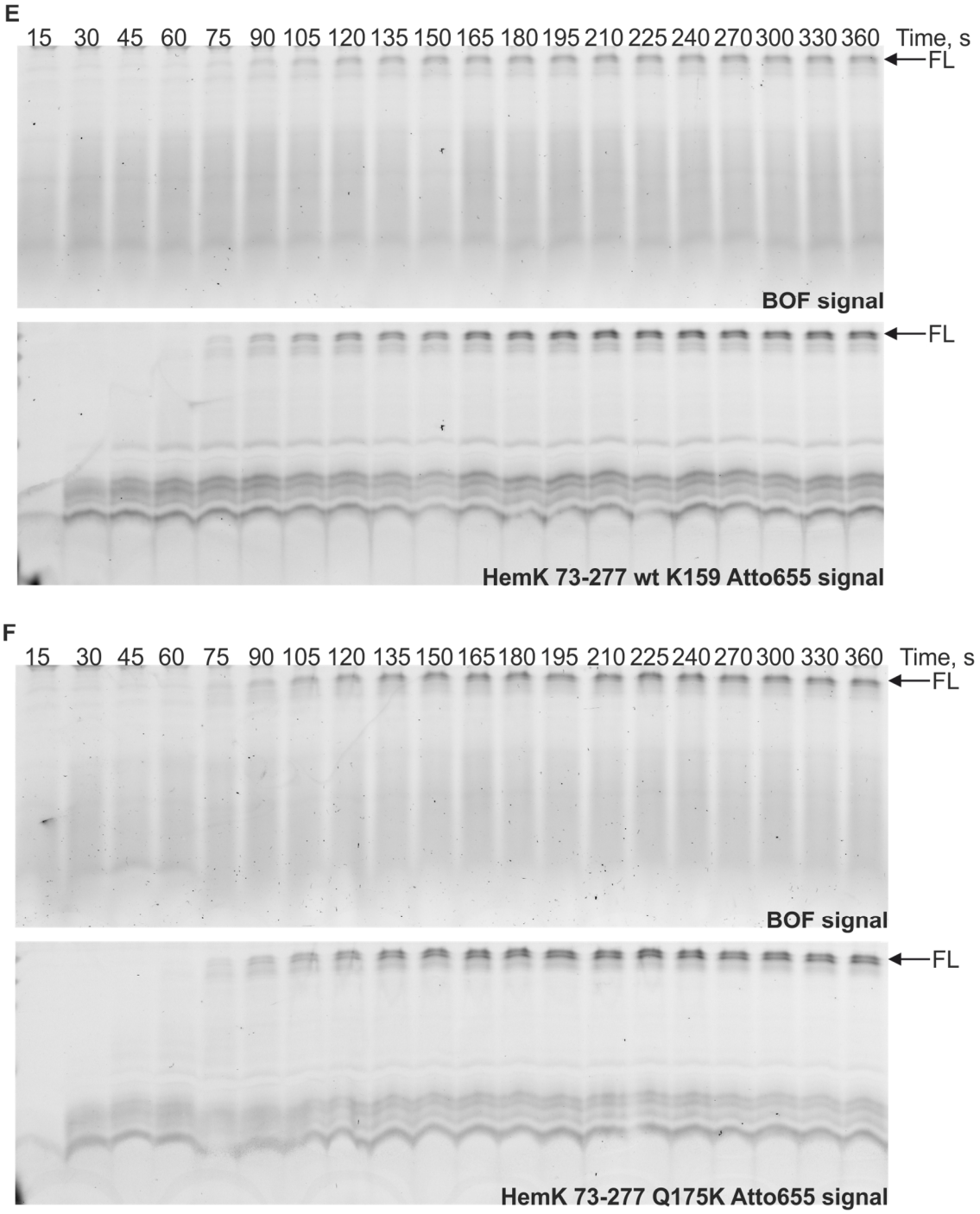

Appendix Figure 1. Translation time courses of HemK 73-277 wt and its truncations. Representative SDS-PAGE of time-resolved formation of (A) HemK 73-277 wt – (H) HemK 73-132 wt translation products. Translation products were visualized by N-terminal BOF. All experiments were carried out in triplicates (only one is shown). FL: full-length product.



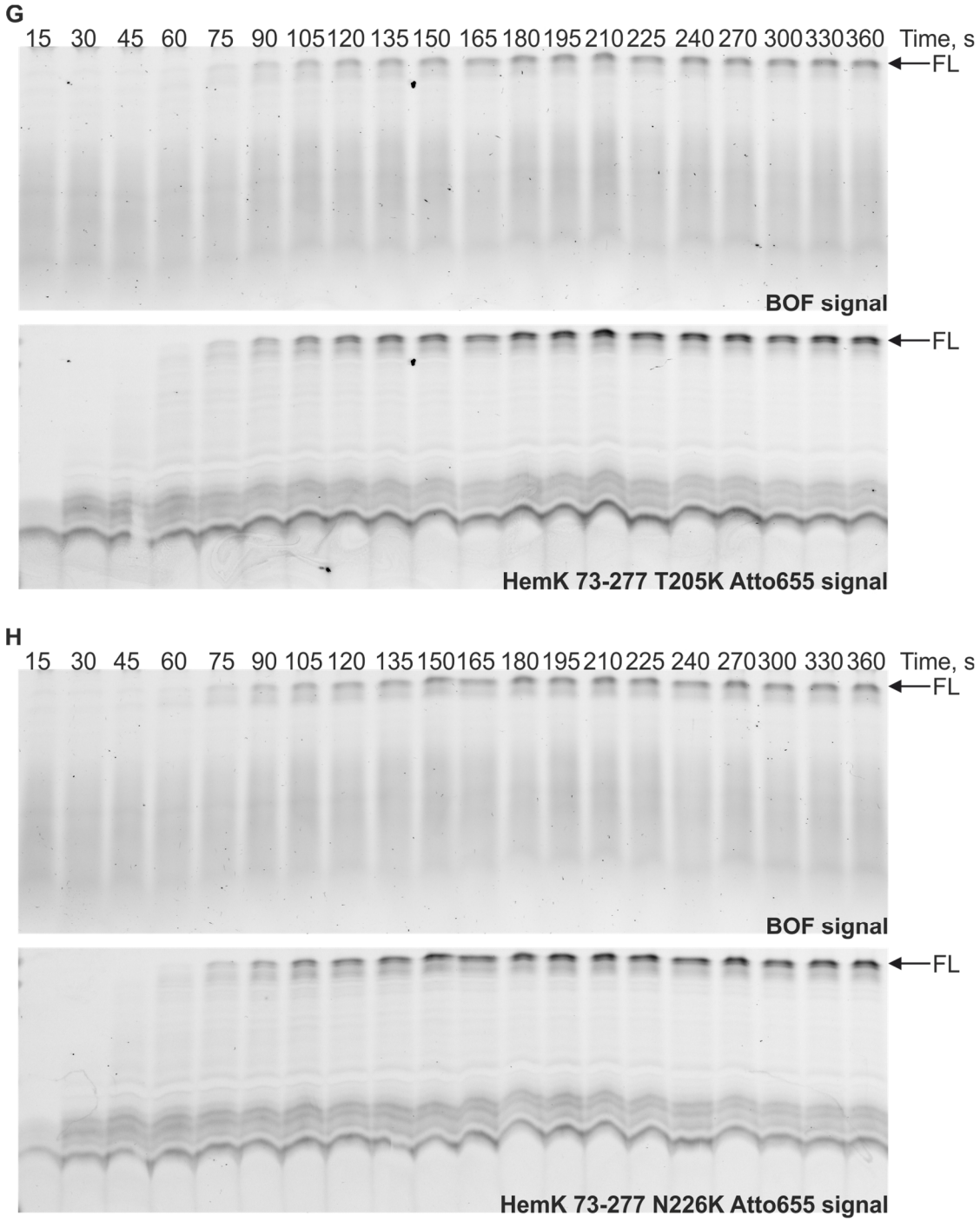

Appendix Figure 2. Continues on following page.




Appendix Figure 2. Continues on following page.




Appendix Figure 2. Continues on following page.




Appendix Figure 2. Effect of BOF-Lys incorporation on the translation rate at various residues. SDS-PAGE of time-resolved formation of HemK 73-277 translation products upon incorporation of BOF-Lys at Lys codon as indicated, visualized by N-terminal Atto655 and internal BOF fluorescence. FL: full-length product.

# 8. Acknowledgements

First and foremost I want to thank Marina Rodnina for this great opportunity to work on this project in her lab and all the helpful and insightful discussions we had over the past years. I also would like to thank the members of my thesis advisory committee Jörg Enderlein and Alexander Stein for the discussion and helpful suggestions we had during the meetings. Additionally, I want to thank Sonja Lorenz, Alex Faesen and Kai Tittmann for joining my examination board and the GGNB office and the Biomolecules program for the administrative organization.

In particular, I want to thank Katya Samatova for keeping me on track during the last 4 years. This would not have been possible without her. I am very grateful for all the time we spent on so many discussions, even though you wanted to do or even started your experiments. I also want to thank Evan Mercier who answered all my questions very, very patiently, Mani Maiti who taught me to operate the confocal microscope and Riccardo Belardinelli for teaching me how to use the stopped-flow machine and perhaps more importantly for teaching me how not to use the stopped-flow machine. Special thanks, and perhaps an apology goes to Vanessa Herold for all the EF-Tu I so recklessly shot in the stopped-flow machine and the hundreds of constructs she was cloning for me. Without the excellent technical support in our lab this project could not have been possible. Thanks to Vanessa, Olaf (thank you for accepting all the tRNA orders and pre-orders and the time you spent on the suppressor tRNA), Sandra, Anna, Christina, Michael, Tessa, Theo, Franzi and Manuela I never had to worry about the material I used and could focus on my experiments. I am also very grateful to Dimitra who was most helpful in every situation, in particular, with handling all the paper work and organizational matters, and Mario who always fixed my computer problems, although I probably caused them all myself. I also want to thank everyone in the department and in my lab who made this entire experience even more enjoyable. Thanks to Xiaolin, Lena and Justas, spending the days in the lab with the science and the evenings with food, drinks and games and all the crazy conversations (thanks to Xiaolin I know how rain works), these four years passed almost too quickly.