Digital Comprehensive Summaries of Uppsala Dissertations
from the Faculty of Science and Technology 2623

Systemic insight into bacterial protein processing by live-cell single-molecule tracking

TORA HÄVERMARK

ACTA UNIVERSITATIS
UPSALIENSIS
2026

UPPSALA
UNIVERSITET

Dissertation presented at Uppsala University to be publicly examined in A1:107a Föreläsningssal, Uppsala biomedicinska centrum (BMC), 752 37 Uppsala, Uppsala, Friday, 6 February 2026 at 09:15 for the degree of Doctor of Philosophy. The examination will be conducted in English. Faculty examiner: Professor Patricia L. Clark (University of Notre Dame: Notre Dame, Indiana, US).

Abstract
Hävermark, T. 2026. Systemic insight into bacterial protein processing by live-cell single-molecule tracking. *Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology* 2623. 57 pp. Uppsala. ISBN 978-91-513-2697-9.

A single Escherichia coli cell contains thousands of different kinds of proteins in a total of millions of copies, all working together to ensure that the cell can grow and divide. New proteins are made by large cellular machines known as ribosomes, and during optimal growth, the ribosomes must produce roughly 20 new proteins every second. To ensure that these new proteins end up correctly folded at their intended cellular location with correct chemical modifications, the ribosome and the nascent protein require assistance from a number of processing factors. Importantly, different proteins require help from different factors. How these factors find their target proteins among the tens of thousands of ribosomes, each translating one of thousands of possible proteins, and engage with them in a coordinated fashion without impeding the other factors, remains mysterious. Traditional in vitro studies have provided high-resolution details about the function of individual factors. However, these fail to mimic the dynamic, crowded conditions of a living cell required to get the full picture of the system. In this work, we have employed a range of single-molecule super-resolution microscopy techniques to study different aspects of protein synthesis directly inside living E. coli cells. In Paper I and Paper II, we used two different microscopy methods to benchmark the cellular activity of Trigger factor (TF), a processing factor that guides the folding of new proteins. We discovered that TF binding to ribosomes is more dynamic than previously perceived. We also saw that TF competes for ribosome binding with another vital factor, namely, the Signal recognition particle. In Paper III, we took a closer look on the biogenesis of outer-membrane proteins, assessing the possibility of using in vivo single-molecule FRET as a reporter for their folding. In all, this work provides fundamental insight into how proteins are processed within live bacterial cells. Such knowledge is key to developing new treatments for bacterial infections and for understanding human diseases linked to protein misfolding.

Keywords: co-translational processing, post-translational processing, Escherichia coli, Trigger factor, outer-membrane protein folding, OmpA, single-molecule tracking, single-molecule FRET

Tora Hävermark, Department of Cell and Molecular Biology, Molecular Systems Biology, Box 596, Uppsala University, SE-751 24 Uppsala, Sweden.

ISSN 1651-6214
ISBN 978-91-513-2697-9
URN urn:nbn:se:uu:diva-572623 (http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-572623)

List of Papers

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

I. Amselem, E., Broadwater, B., **Hävermark, T.**, Johansson, M., Elf, J. (2023) Real-time single-molecule 3D tracking in E. coli based on cross-entropy minimization. *Nat Commun*. 14, 1336.

II. **Hävermark, T.**, Metelev, M., Lundin, E., Volkov, I.L., Johansson, M. (2025) Dynamic binding of the bacterial chaperone Trigger factor to translating ribosomes in Escherichia coli, *Proc. Natl. Acad. Sci. U.S.A.,* 122 (1)

III. **Hävermark, T.**, Volkov, I.L., Johansson, M. (2025) Outer-membrane protein folding assessed by live-cell single-molecule FRET. *Manuscript.*

Contents

Abbreviations

2D	two-dimensional
3D	three-dimensional
ATP	adenosine triphosphate
BAM	beta-barrel assembly machinery
Cryo-EM	Cryogenic electron microscopy
DNA	deoxyribonucleic acid
EMCCD	electron multiplying charge-coupled device
FOV	field of view
FP	fluorescent protein
FRET	Förster resonance energy transfer
GTP	guanosine triphosphate
HMM	hidden Markov model
IM	inner membrane
kDa	kilodalton
LPS	lipopolysaccharide
MAP	methionine aminopeptidase
mRNA	messenger ribonucleic acid
OM	outer membrane
OMP	outer-membrane protein
PDF	peptide deformylase
PSF	point spread function
Ribo-Seq	Ribosome sequencing (profiling)
RNC	ribosome-nascent chain complex
sCMOS	scientific complementary metal-oxide semiconductor
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
smFRET	single-molecule Förster resonance energy transfer
SMT	single-molecule tracking
SNR	signal-to-noise ratio
SRP	Signal recognition particle
TF	Trigger factor
tRNA	transfer ribonucleic acid
WT	wild-type

Introduction

A single *Escherichia coli* cell – one of Nature's simplest designs, contains thousands of different kinds of proteins in a total of millions of copies, all working in coalition to ensure that the cell can respond to its surroundings and proliferate[1,2]. The blueprint for each protein is stored as separate genes in the DNA, which is translated into functional proteins by ribosomes via mRNA transcripts. Under optimal conditions, an *E. coli* cell divides roughly every 20 minutes. To prepare for this division, it must duplicate its biological content, synthesizing an astounding number of approximately 20 new proteins every second.

In order to perform its intended biological function, a protein must fold into its native 3D structure, but the high translation speed in the crowded cellular environment poses substantial risk of protein misfolding. Producing erroneously folded proteins is in the best case merely a waste of energy as the synthesized proteins lack function, but often, misfolded proteins cause accumulation of insoluble aggregates that are fatal to the cell. Consequently, all organisms have developed several quality control systems in the protein synthesis machinery. Among these are the chaperones – they themselves proteins, whose sole function is to assist the folding of other proteins during and after synthesis.

Further, to ensure that proteins end up at their intended cellular compartment with proper chemical modifications, a number of accessory factors engage with the ribosome-nascent chain complex (RNC), either during synthesis (co-translational processing), or after the nascent peptide is released off the ribosome (post-translational processing). Many co-translational factors share docking site on the ribosome, and not all nascent proteins require engagement by all processing factors. We still do not fully understand how these factors find their targets among tens of thousands of ribosomes, each translating one of thousands of possible proteins, and engage with them in a coordinated fashion without impeding other processing factors.

In vitro studies using reconstituted systems have provided high-resolution structures and mechanistic details of individual processing factors and their target specificities, but no reconstituted system can accurately mimic how all factors work in coalition in the crowded dynamic environment of the living

cell. During my PhD, I have studied aspects of co- and post-translational processing, especially related to protein folding, directly inside living cells. In **Paper I** and **II**, we used two different single-molecule tracking (SMT) methods to study ribosome-binding kinetics of the chaperone Trigger factor (TF) and found that TF displays highly dynamic binding to translating ribosomes. Using genetic perturbations, we also established an antagonistic interplay between TF and the Signal recognition particle (SRP) with respect to ribosome binding. In **Paper III**, we took a closer look on outer-membrane protein biogenesis, assessing the possibility of using *in vivo* single-molecule FRET (smFRET) as a reporter for folding.

Protein biogenesis in *E. coli*

The fundamentals of bacterial protein synthesis follow the Central dogma (Figure 1a). Genetic DNA encodes the amino acid sequence of a protein. When the cell needs that particular protein, the gene is transcribed into an mRNA molecule. Ribosomes bind to the mRNA transcript and, with the help of tRNAs, decode the mRNA codons into the corresponding amino acid sequence. The ribosome adds amino acids one by one to the growing protein chain which extends from the N-terminus to the C-terminus, pushing the chain through the ribosomal exit tunnel. Elongation continues until the ribosome encounters a stop codon on the mRNA, signaling the release of the nascent protein chain from the ribosome (Figure 1b).

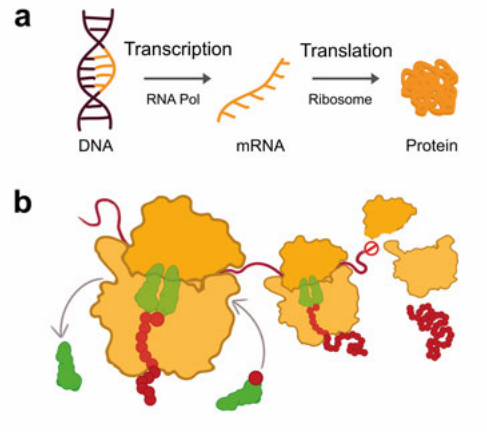

Figure 1: Principles of protein translation. a. The Central dogma. A gene (orange) is transcribed to mRNA by RNA polymerase. Ribosomes translate the mRNA to make proteins. **b.** A ribosome (orange) decodes mRNA by codon matching to amino-acylated tRNAs (green) and facilitates peptide bond formation between amino acids. As elongation proceeds, the nascent chain (red) emerges from the ribosome exit tunnel. At translation termination, the nascent chain is released off the ribosome.

Based on the nature and cellular location of the nascent protein, one or several accessory factors will bind to the ribosome-nascent chain complex (RNC) during synthesis to ensure correct enzymatic processing, folding and targeting (Figure 2a), described in closer detail in the sections below. In a comprehensive review, Koubek *et al.* (2021) constructed a model explaining the coordinated timing of bacterial co- and post-translational processing[3]. Their model is summarized in Figure 2b.

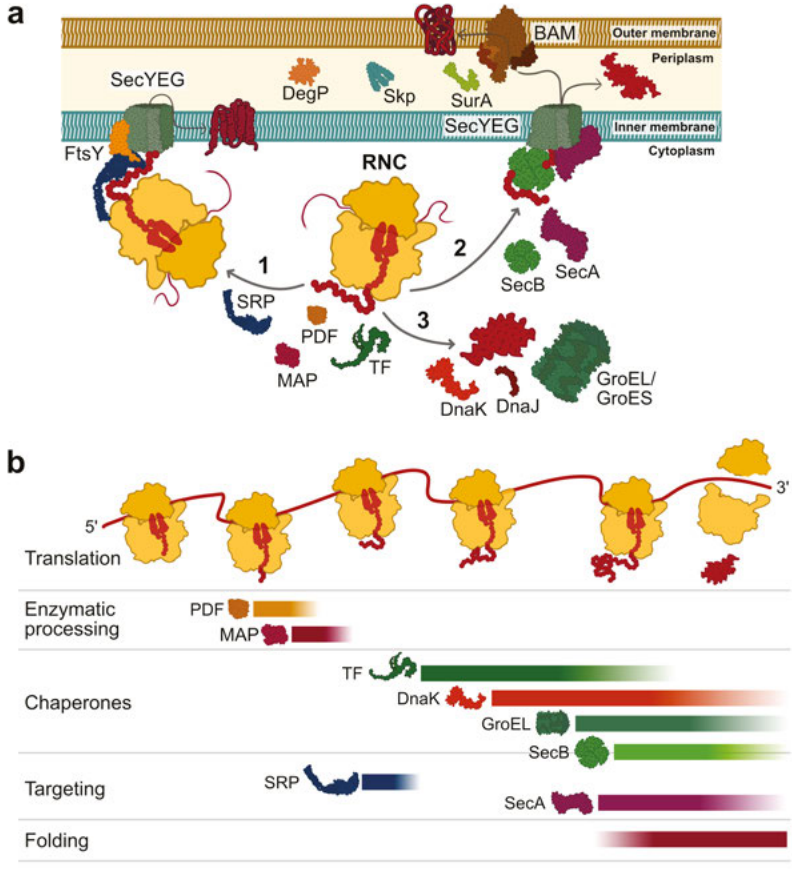

Figure 2: Co- and post-translational protein processing. a. Schematic of the main factors in co-translational processing and Sec-mediated translocation. Co-translational translocation of primarily inner membrane proteins is mediated by SRP and FtsY (1). Post-translational translocation of primarily periplasmic and outer-membrane proteins is mediated by SecB and SecA (2). Cytosolic proteins can fold independently, but often require assistance from chaperones (3). **b.** Timing of RNC engagement by different processing factors, grouped by their function. The model is adapted from Koubek *et al.* (2021).

Co-translational protein folding

Back in the 1960's, Christian B. Anfinsen proposed that the necessary information for a protein to fold into its correct structure is contained within its amino acid sequence. Given the right physiological environment, a protein will spontaneously fold into its most stable, thermodynamically favorable conformation[4]. While this holds true for small globular proteins in solution, the vectorial growth from the N- to the C terminus of the nascent protein during translation in the cell poses additional problems, as folding most often begins before the entire protein is released off the ribosome. During synthesis, the protein may obtain intermediate folds, presenting local energy minima. These can be stabilized or destabilized by co-translational chaperones, or through interactions with the ribosome surface[5]. Moreover, for a subset of proteins, folding must be delayed by chaperones due to downstream membrane targeting.

In *E. coli*, the major co-translational chaperone is the ATP-independent TF, which binds to the ribosome via its N-terminal signature GFRxGxxP motif[6]. The C-terminal domain forms the center of the structure and is responsible for most interactions with the nascent chain[7–10]. As the nascent protein is released off the ribosome, it may stay associated with TF[11,12], get delivered to other cytosolic chaperones[13,14], or fold independently[15]. TF has a broad substrate pool, preferentially interacting with sequences of positively charged amino acids flanked by aromatic residues or hydrophobic patches[16], and is particularly enriched on ribosomes translating outer-membrane proteins[17]. TF-RNC affinity is substrate-specific, with target recognition leading up to a 30-fold increase in binding strength. However, the duration of target binding varies greatly, from 1 s to 50 s[18–21]. Importantly, because these measurements are typically from reconstituted systems using stalled ribosomes, they may not accurately reflect how TF interacts with a growing nascent protein chain *in vivo*.

Membrane protein biogenesis

Gram-negative bacteria are diderm, with a symmetric phospholipid inner membrane (IM) and an asymmetric outer membrane (OM). The inner leaflet of the OM consists of phospholipids, whereas the outer leaflet is composed of lipopolysaccharides (LPS). The lipid contents of the inner leaflet and the lipid A-part of LPS in the outer leaflet provide a barrier against hydrophilic molecules. Moreover, the sugar chains of LPS provide a polar mesh around the cell, making it impermeable also to polar molecules. Pore-like proteins in the OM, known as outer-membrane proteins (OMPs) provide selective permeability to the cell, operating either as passive ion channels, or as specific pumps[22]. These

proteins also play vital roles in structural support, cellular movement, adhesion, and virulence. Because they serve as the primary interface between the cell and its environment, OMPs are an attractive target for developing new antibiotics aimed at gram-negative bacteria[23].

Between the OM and the IM is the periplasmic space; an aqueous and oxidative environment ranging between 10-70 nm in width. Substantial space in the periplasm is taken up by the peptidoglycan layer which provides rigidity to the cell. The periplasm is further populated by a large number of proteins, creating an exceptionally crowded and confined milieu[24].

Together, all membrane and periplasmic proteins comprise ca 30% of the *E. coli* proteome. As protein synthesis occurs within the cytoplasm, these proteins require targeting mechanisms to the IM for insertion or translocation across the membrane. This is predominantly facilitated by the identification of specific stretches of N-terminal amino acids, referred to as signal peptides[25–27]. While there are several pathways in place for protein translocation governed by different signal peptides and accessory factors, this work focuses on SecYEG-dependent translocation, described in closer detail below. The central component of this pathway is the SecYEG translocon; a trimeric protein complex forming dynamic pores in the IM, thus facilitating protein translocation.

SRP-mediated targeting and translocation

Guided by SRP, most inner membrane proteins are inserted into the IM co-translationally, i.e., during translation. SRP, which is composed of the GTPase Ffh and 4.5S RNA, recognizes a highly hydrophobic and SRP-specific signal peptide on the nascent chain[28]. Upon signal recognition, SRP binds the RNC through interactions with the nascent chain and to L23 on the ribosome[29,30]. The SRP-bound RNC has affinity for the membrane-associated SRP receptor FtsY, thus facilitating membrane targeting of the RNC (Figure 2a). The SRP-FtsY-RNC complex has affinity for the SecYEG transolcon[31]. Upon translocon docking, GTP hydrolysis triggers release of SRP and FtsY[32]. Via active translation, the nascent chain is pushed through the translocon, where lateral gate opening facilitates insertion into the inner membrane[33].

SecB- and SecA-mediated targeting and translocation

SecYEG-mediated translocation of periplasmic and OM proteins is primarily governed post-translationally by the chaperone SecB and the ATPase SecA. SecB predominantly binds to unfolded stretches of aromatic and basic residues in the nascent protein[34], either late in the synthesis, or after nascent chain release from the ribosome. SecB maintains the precursor protein in an unfolded state and has affinity for the membrane-associated ATPase SecA, thus

enabling membrane targeting[35]. SecA has affinity for SecYEG and a SecA-specific signal peptide on the precursor protein[36,37]. Upon SecA-SecYEG docking, a conformational change in SecB causes nascent chain handover to SecA[38]. SecA pushes the unfolded protein through the SecYEG pore by ATP hydrolysis. Upon translocation, the N-terminal signal peptide is cleaved off and the protein is further processed and matured into its functional state, either in the periplasm or in the OM[39].

Apart from SecA and SecB, TF has been implied as a significant chaperone in the processing of post-translationally secreted proteins[40,41]. SecB and TF share nascent chain binding specificities[34,42], and TF is overrepresented on RNCs translating SecB-specific secretory proteins[17]. In cells lacking TF, SecB will bind to its target proteins earlier during translation, suggesting that TF delays SecB-target interactions[43].

Although the description above suggests two distinct pathways for SecYEG-mediated translocation and insertion, these pathways are likely somewhat redundant and not strictly co- or post-translational in the cell[44]. For example, Cryo-EM and ribo-Seq profiling has implied co-translational roles of SecA and SecB, revealing a co-translational preference for RNCs translating inner membrane proteins[43,45,46].

Kinetics of membrane targeting and translocation

In vitro-based assays have reported that SecA translocates roughly 4-40 or 2.2 amino acids per second, using *E. coli* or *Thermus thermophilus* translocons, respectively[47–49]. Importantly, while such *in vitro* assays can measure translocation rates, they are blind to the membrane-targeting aspect of translocation. Instead, kinetic measurements on both the targeting and translocation requires *in situ* experiments at high temporal resolution. In our lab, we have measured SRP-mediated membrane targeting in live *E. coli* by *in vivo* SMT, showing that the majority of SRP-RNC targeting events start in the cytosol. The average SRP cycle is on the order of 2 s., and hence does not necessitate translation arrest during targeting in *E. coli*[50]. Some of our ongoing work focuses on *in vivo* SMT of SecA and SecB for similar measurements of post-translational targeting and translocation. Finally, we aspire to measure the complete cycle of targeting, translocation, and folding of an OMP by labeling and tracking the protein directly, as initiated in **Paper III**.

Outer-membrane protein biogenesis

Over the years, the biogenesis of OMPs has gained focus in my PhD studies. First, when it became evident to me that TF is overrepresented on RNCs translating OMPs[17], and may have a direct and antagonistic role to SecB in post-translational targeting[40]. Second, we expanded our work to SecA- and SecB-

tracking (not included in this thesis as it is still in progress). Third, this evolved into the project presented in **Paper III**, focusing directly on OMP biogenesis. Hence, the structure, translocation, and folding of OMPs deserve a more detailed description.

Practically all OMPs share the β-barrel as a structural feature (Figure 3). A β-barrel is composed of 8-36 antiparallel amphipathic β-strands forming a cylinder, with a hydrophilic interior and a hydrophobic exterior, enabling embedment in the hydrophobic OM[51,52]. Some OMPs contain additional water-soluble domains, and some OMPs oligomerize in the OM while others are monomeric[22].

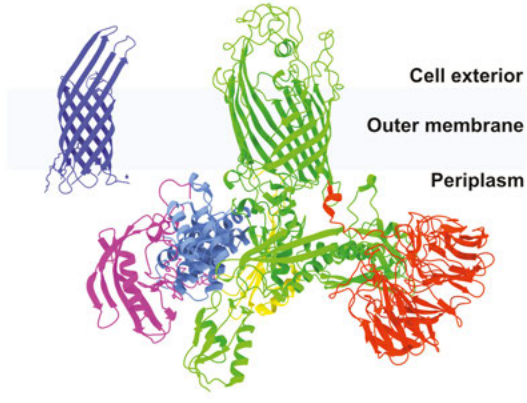

Figure 3: Examples of outer-membrane proteins. The 8-stranded β-barrel OmpX (left, PDB 1QJ8) is one of the smallest OMPs, while the BAM complex (right, PDB 5LJO) is more intricate. The BAM complex is composed of five proteins: BamA (green), BamB (red), BamC (purple), BamD (blue), and BamE (yellow). BamA has a 16-stranded β-barrel domain with a lateral gate embedded in the OM, and periplasmic domains (POTRA) which interact with BamB-E.

OMPs are post-translationally targeted and translocated over the IM in an unfolded state by SecB and the SecA-SecYEG translocon. The OMP then traverses the crowded periplasm toward the OM, bypassing the peptidoglycan layer[53]. Finally, the β-barrel assembly machinery (BAM) complex catalyzes folding and insertion of the β-barrel into the OM. As the periplasm is devoid of ATP, these processes are not driven by external energy. Periplasmic transport of the OMP to the BAM complex is guided by a plethora of periplasmic ATP-independent chaperones, for example SurA, Skp, FkpA, Spy, and the protease-chaperone DegP. These are partially functionally redundant, yet, they seem to all play distinctive roles which may vary depending on the OMP substrate at hand[53]. As such, there seems to be no universal mechanism for periplasmic OMP transport and processing, and the mechanism of BAM-mediated folding and OM insertion is also elusive and possibly substrate-dependent[22].

OMP biogenesis studies

OMP folding studies are generally employed by isolation and storage of the OMP in a high-urea buffer at which the OMP is chemically denatured, thus resembling the unfolded state of the precursor protein. Upon dilution in the presence of lipid vesicles, the OMP can fold into its presumed native structure[54]. By quenching the reaction at different timepoints, apparent folding kinetics can be obtained. For example, the ratio of folded and unfolded protein can be assessed by cold SDS-PAGE, which separates folded and unfolded OMPs into distinct bands[55]. Other techniques for assessing folding include bulk FRET[56], fluorescence or circular dichroism spectroscopy[57,58], *in vivo* pulse-chase experiments[59], proteinase K digestion[60,61], and phage resistance assays[62].

The folding kinetics observed in the *in vitro* lipid-based assays are highly dependent on factors like lipid composition, temperature, and pH[55], and are often unphysiological. Remarkably, these methods typically report slow folding rates, on the order of a few minutes or even hours[56–59,63]. Further, while *in vivo* pulse-chase experiments can capture folding kinetics under native conditions, they have limitations in temporal resolution, making them suitable only for studying inherently slow-folding OMPs[64].

The relatively slow folding rates of smaller OMPs measured in *in vitro* bulk experiments were recently challenged in an *in vitro* smFRET study on OmpG. Using a single-molecule approach, the authors concluded that individual folding events are very rare, but when they occur, they fold on a millisecond timescale[65]. This discrepancy in folding rates is likely due to the rarity of successful folding in these synthetic lipids, leading to slow apparent kinetics in ensemble measurements. Although insightful, it should be noted that this smFRET study also utilized synthetic planar lipids far from resembling the native *E. coli* OM, and accessory factors such as chaperones and the BAM complex, otherwise crucial for efficient insertion into native *E. coli* OM vesicles[63], were absent.

What is clearly lacking in OMP biogenesis research is an experimental setup which accounts for the complexity of the cellular environment, regarding the involved accessory factors as well as providing the physical constraints of the crowded periplasm. Moreover, a single-molecule approach is crucial to probe the kinetics of these rare events.

The present work

In this work, we have used a set of *in vivo* single-molecule microscopy tools to study aspects of protein biogenesis in *E. coli*. In the first sections, I will present the details of each methodology. Then, I will discuss how they were implemented in each study, and summarize the main biological findings from each application.

Methodology

Over the years, *in vitro* experiments on isolated reconstituted systems have been instrumental in deducing the mechanisms behind biochemical processes central to life. However, all cellular machineries cooperate in a finely tuned and highly complex manner that cannot be sufficiently portrayed *in vitro*. Instead, *in vivo* single-molecule methods have gained increasing power for studying biomolecules in their native environment. Single-molecule methods are powerful in detecting transient states that are otherwise hidden in the averaging effect of conventional biochemical kinetics experiments[66]. One such method is *in vivo* SMT by means of super-resolution fluorescence microscopy. The molecule of interest is fluorescently labeled, and the position of the fluorescence signal is recorded over time, giving information about where and how the labeled molecule is moving around inside the living cell.

Basics of fluorescence

All methods described here depend on the detection of a fluorescently labeled molecule. Fluorescence occurs when electrons in a molecule are excited to a higher energy state upon absorbing energy from an external light source of a suitable wavelength. The electron quickly returns to its ground state, typically within nanoseconds, releasing the excess energy in the form of photons (emission). In the process of excited state relaxation, preceding emission, some energy is dissipated. Consequently, the emitted light has a longer wavelength (lower energy) than the excitation light source. Fluorescence competes with other, non-radiative, excitation decays, so not all molecules fluoresce efficiently at all wavelengths. Molecules that fluoresce exceptionally well at a

specific wavelength are called fluorescent dyes or fluorophores. For fluorescence microscopy, dyes with a good separation between their maximal absorption and emission wavelengths (Stokes shift) allows for good excitation and emission signal filtering. Another important property is the quantum yield (QY), describing the relative occurrence of fluorescence to other non-radiative decays from the excited state. Further, the extinction coefficient (ε) describes the probability of light of a specific wavelength to be absorbed by the dye. The product of the QY and ε determines the brightness of the dye. Thus, a bright dye requires both a high quantum yield and extinction coefficient. A final important property is the photostability of the fluorophore, describing how many times it can be excited before being permanently damaged and no longer subject to fluorescence – a phenomenon known as photobleaching. A high photostability allows for higher excitation powers or longer exposure times. In summary, a good fluorophore for single-molecule applications should have a large Stokes shift, high brightness, and high photostability.

Fluorescence labeling for SMT

The advent of fluorescent proteins (FPs) revolutionized the field of *in vivo* fluorescence microscopy as it allowed for easy fluorescence labeling of a protein of interest by genetic FP fusions. However, in our SMT applications, we require very strong fluorescence from a single fluorophore and long enough trajectories to capture transitions between binding states within a trajectory. Such brightness and photostability is unattainable from any FP. Instead, we rely on using organic fluorophores, such as rhodamine-based (Janelia-Fluor family) or cyanine dyes. In this work, these organic dyes were fused to the molecule of interest by using two different approaches (Figure 4).

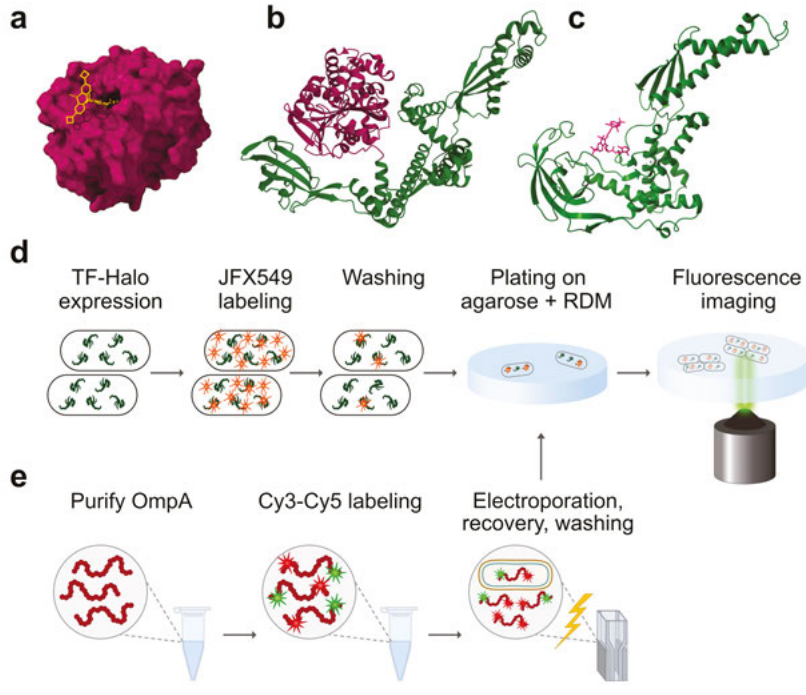

Figure 4: Labeling methods. a. Crystal structure of HaloTag (magenta, PDB (8SW8) bound to JF669-HaloTag ligand (orange). **b.** Alphafold3 prediction[67] of TF (green, 48 kDa) with a C-terminal HaloTag fusion (magenta, 33 kDa). **c.** Schematic of TF labeled with sulfoCy5-maleimide (pink, 0.8 kDa), based on a crystal structure of TF (PDB code 1T11). **d.** *In vivo* HaloTag labeling procedure. **e.** *In vitro* labeling and electroporation procedure.

In the first approach, we use HaloTag[68], a 33 kDa modified haloalkane dehalogenase that covalently binds chloroalkane-activated organic dyes with high specificity (Figure 4a). HaloTag is genetically fused to the protein of interest (Figure 4b) and expressed in *E. coli*. Cells are soaked in the fluorescent dye for conjugation to HaloTag. Excess dye is then washed off (Figure 4d). Hence, any dye that can be chloroalkane-activated and is cell-permeable can be used for HaloTag labeling *in vivo*. This approach was used in **Paper I and II.** Although extremely powerful, HaloTag labeling suffers from its bulkiness. In bacteria, many proteins are of equal size or even smaller than HaloTag, and consequently, labeling may disturb the function of the protein of interest.

The second approach circumvents the bulkiness by directly conjugating the relatively small organic fluorophore to the protein of interest (Figure 4c). This can be done for example by maleimide-labeling of cysteines[69]. The fluorophore of interest is functionalized with a maleimide group, which reacts with cysteine residues in the protein, enabling site-specific labeling. This

commonly requires prior deletion of any native cysteines followed by cysteine insertion at the site of interest in the protein sequence. Hence, while the fluorophore itself is relatively small and may not affect protein function, it is important to validate that the cysteine mutations do not damage the protein.

Maleimide-labeling is performed *in vitro* and hence requires prior protein purification (Figure 4e). After labeling, the protein can be delivered back into *E. coli* by electroporation[70,71]. In **Paper III**, we used this approach for labeled proteins, but more commonly, we have employed electroporation of labeled nucleic acids[50,72]. When implementing this method, it is important to ensure that electroporation does not destroy the labeled molecule.

Diffusion rate gauges the *in vivo* activity of a molecule

All molecules inside a bacterial cell are in motion. Generally, this motion is driven by passive diffusion, and the characteristics of the diffusion depend on factors like particle size, viscosity, confinement, temperature, and pressure. The characteristics of the diffusion is described by the diffusion rate constant, D ($\mu m^2 s^{-1}$). Changes in the diffusional behavior, e.g., due to a change in particle size, is reflected in D. For a small protein, the diffusion rate will be significantly different depending on whether it is free in the cytoplasm, or bound to macromolecular complexes such as ribosomes, or tethered to a cell membrane (Figure 5a). A simple model with two different diffusion states representing a bound and free state can be used to deduce the binding kinetics, as the average time spent in the slow state is the inverse of the dissociation rate constant, k_{off} (Figure 5b). Further, by mapping the coordinates of diffusion trajectories to the imaged cell, the cellular location of a molecule can be determined, providing insight into its biological function (Figure 5c). In this work, we have used different methods under the umbrella of super-resolution fluorescence microscopy and SMT to extract the diffusion, localization, binding kinetics, and folding of various proteins inside live *E. coli*. Each of these techniques and how they differ are described in the next sections.

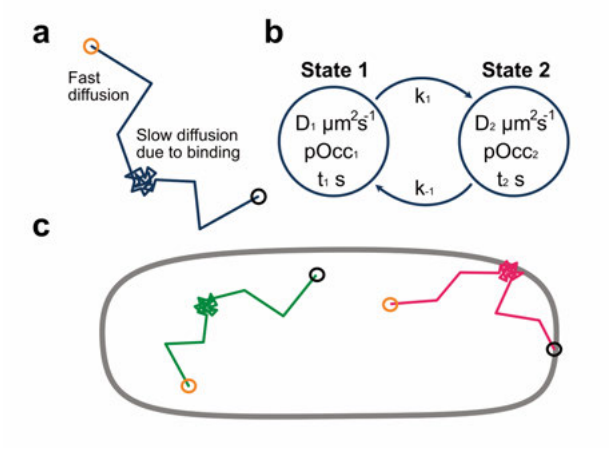

Figure 5: single-molecule diffusion analysis. a. Single-molecule trajectory, recording one molecule's motion over time. Orange and black circles mark the first and last frame. The displacement between consecutive frames (step length) correlates with the molecule's diffusion rate. **b.** Model describing discrete states of diffusion based on an ensemble of single trajectories. In our implementation, the model describes the state's average diffusion rate (D), occupancy (Occ), and dwell time (t) from which the rate constant (k) is derived. **c.** Mapping trajectories to the cell geometry provides insight into the biological function behind diffusion states. The slow diffusion in the pink trajectory corresponds to membrane binding, whereas the slow diffusion in the green trajectory is binding to some macromolecular complex (e.g., ribosomes) in the cytoplasm.

2D-SMT using camera-based super-resolution fluorescence microscopy

Camera-based 2D super-resolution microscopy, as employed for SMT in **Paper II** and **Paper III**, uses live-cell widefield epi-fluorescence microscopy, where the full volume of a field of view is evenly illuminated with light for fluorophore excitation. The excitation light travels through the objective lens, hitting the cell sample, and fluorophore emission travels back through the objective along with reflected excitation light. A dichroic mirror with an emission filter separates the longer-wavelength emission from the much more intense shorter-wavelength excitation light, allowing only the emitted photons to hit a camera detector (Figure 6a).

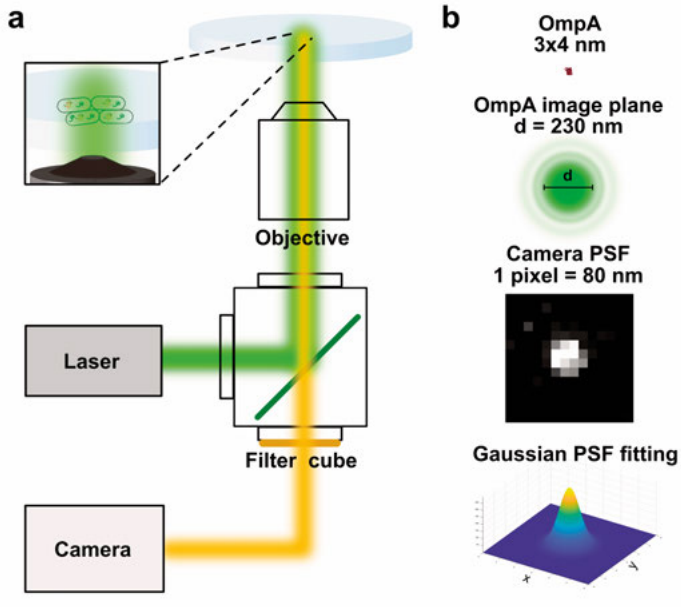

Figure 6: Camera-based 2D widefield epifluorescence microscopy. a. Microscopy setup. **b.** Super-resolved localization of a single protein. A fluorescently labeled protein is a few nm in size, but appears much larger in the image plane due to the diffraction limit of visible light. On the camera, the emitter is normally portrayed as a gaussian PSF. Sub-pixel localization of the protein is obtained by PSF fitting.

As proteins are smaller (a few nm) than the diffraction limit of visible light (in the range of a few hundred nm), a single fluorescently labeled protein appear on the camera as a diffraction pattern known as the point spread function (PSF). For regular lenses, the PSF usually has the shape of a disc with concentric rings (Airy disc). The main disc contains roughly 84% of all emitted photons and has a size approximately equal to the diffraction limit, i.e., much larger than the molecule of interest. Consequently, to obtain a spatial resolution on the tens of nm-scale, the PSF is fitted to a distribution, often a Gaussian, by least-squares fitting, to find the centroid of the fluorescence emission. The coordinates of the center correspond to the most likely localization of the single protein (Figure 6b). The x-y precision of the localization is in this case given by the variance of the Gaussian. Notable parameters that increase the precision, and thus, the spatial resolution, are a high signal-to-noise ratio (SNR), and a large number of collected photons[73].

Another important consideration for precise localization is motion blur, occurring if the displacement of the emitter is larger than the localization

precision within the image acquisition time, causing distortion of the PSF. As an example, a localization precision of 20 nm for a molecule displaying Brownian diffusion at 1 $\mu m^2 s^{-1}$ would entail that the acquisition time is only 0.1 ms. No current cameras can record a sufficient field of view (FOV) in such short time (i.e., high frame rate). To circumvent this, stroboscopic illumination is employed, where the sample is illuminated with only ms-long laser pulses while allowing longer camera exposures for recording the signal from a larger FOV. The localization is thus not significantly influenced by motion blur due to the short laser exposure[66].

Camera-based widefield epi-fluorescence microscopy produces a 2D image of the sample at the focal plane. However, as the entire sample is illuminated, emission from out-of-focus fluorescent molecules and sample autofluorescence yield a reduced SNR, thus negatively impacting the spatial resolution. Moreover, for SMT applications, the axial movement of the emitter is lost. This leads to shorter trajectories due to out-of-focus movement, and makes detection of axial diffusion confined to, e.g., the *E. coli* membrane, more difficult.

Finally, in addition to a high spatial resolution, the tracking of fast-diffusing molecules over time also requires a high temporal resolution. In camera-based SMT, the temporal resolution is largely determined by the camera detection speed, i.e., how long it takes for pixel readout of the entire sensor. The frame rate may be increased by specifying a smaller FOV. Still, current state-of-the art sCMOS and EMCCD cameras are limited to approximately 1 kHz frame rates (1 ms per frame) for a FOV of practical size. In order to reach sub-ms temporal resolution, other detection methods are required, as exemplified in the next section.

3D-SMT using modulation-enhanced localization microscopy with confocal detection

As described in the previous section, widefield epi-fluorescence microscopy with camera detection suffers from some drawbacks:

1. Illumination of the entire sample (cell) volume causes low SNR, which has a negative effect on the spatial resolution.
2. 2D images only give high resolution for molecules in the focal plane, which leads to shorter trajectories for fast-moving molecules and possibly underestimation of axial confined diffusion.
3. The temporal resolution is limited by the camera speed to a ms-timescale at best.

A plethora of different microscopy techniques have been developed to circumvent these issues. Here, I will briefly present some aspects that were applied to the microscopy development in **Paper I** for 3D-SMT of proteins displaying heterogeneous (fast and slow) diffusion.

First, this method is based on modulation-enhanced localization techniques[74]. Instead of applying uniform illumination to the sample, thus illuminating any single emitter directly, the sample is instead illuminated by a pattern of sparsely spread excitation PSFs. By recording how the emitter responds to each relative illumination point, more precise localization is obtained (Figure 7). Further, this structured illumination leads to less direct excitation of the fluorophore, generally enabling longer recordings before photobleaching. In **Paper I**, localization was inferred by cross-entropy minimization, where the difference between the expected photon counts from the pattern of Gaussian excitation PSFs and the actual number of detected photons (which is altered due to excitation PSF interactions with the emitter) were minimized.

A second improvement of the spatial resolution was obtained by confocal filtering, where a 200 nm pinhole is placed in front of the detector, filtering out any out-of-focus emission. Further, photon counting by single-photon avalanche diodes was used for detection, allowing for highly sensitive (single photon) detection at a ns-timescale.

Thirdly, in order to obtain 3D localization, the PSF excitation pattern was applied in three different focal planes surrounding the emitter, thus providing additional high-resolution z information about its position.

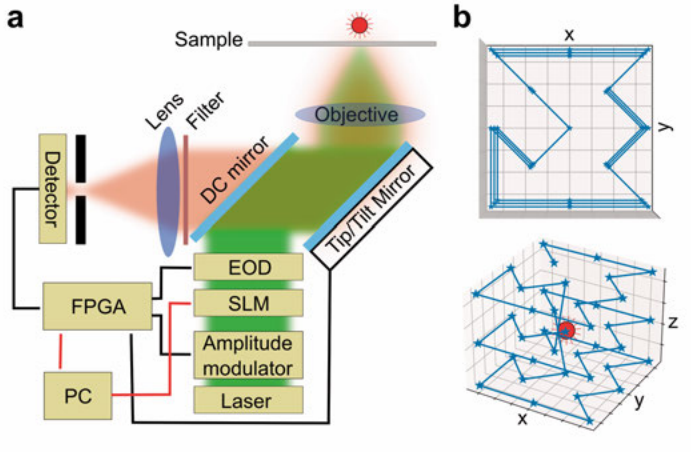

Figure 7: Microscopy setup employed in Paper I. a. Schematic of components of the microscope. **b.** Gaussian PSF shot-pattern viewed from the top and from the side. Stars indicate focal points of each Gaussian, where every point will provide relative information on the emitter's localization (red). Adapted from Figure 1 in Paper I.

Even though the detection of photons is on the ns-timescale, applying modulation-enhanced localization patterns in 3D traditionally requires time-consuming movement of the aperture or sample, with some real-time analysis for concurrent repositioning of the pattern to account for any emitter movement. This has rendered such methods, for example MINFLUX[75], to high-resolution imaging of immobile structures. Those methods are too slow for tracking fast-moving molecules, such as proteins diffusing freely in the bacterial cytosol. The purpose of **Paper I** was thus to develop a setup employing a larger structured illumination pattern, allowing for sampling of a larger volume of the cell compared to MINFLUX, enabling detection of fast-moving molecules. Further, fast online feedback of the emitter position was employed, enabling repositioning the PSF patterns during data acquisition. The setup was optimized for obtaining long trajectories of relatively fast molecules, thus yielding a temporal resolution of 0.84 ms and a precision of 67 nm and 109 nm in the lateral and axial directions, respectively.

Data analysis – dot detection and trajectory building

In this section, I will describe the methods used for analysis of the fluorescence microscopy data. As my contribution to **Paper I** was merely to provide biological samples for testing the microscope developed by Amselem and Broadwater, a thorough description of their analysis is beyond the scope of this Thesis summary. Instead, I will focus on the analysis tools used in **Papers II** and **III**. A comparison between the diffusion analyses employed in **Paper I** and **II** is given in the chapter *TF diffusion – a comparison between Paper I and II*.

The raw data from a SMT microscopy experiment as performed in **Paper II** and **III** is a phase contrast image of a small cell colony, and a corresponding time-lapse movie of fluorescence images acquired with different exposure times (typically 5 or 20 ms per frame and 150-500 frames per movie). The phase contrast image is used for cell segmentation, marking the outlines of each cell (Figure 8). Segmentation masks are created using the Per Object Ellipse fit (POE) method for adaptive thresholding, well-suited for images where cells are adjacent[76].

The segmentation masks are then aligned with the fluorescence images for fluorophore detection, made on a per-cell basis, using a radial symmetry-based algorithm[77]. In short, this algorithm calculates per-pixel intensity gradients in an image pointing to where the intensity is highest. The gradients are perpendicular to intensity contour lines, and for perfect radially symmetric dots, these lines intersect at the center of the PSF, i.e., at the sub-pixel localization of the emitter. In a real image with noise, the radial symmetry is not perfect. Instead, the algorithm finds the point that minimizes the distance between the lines.

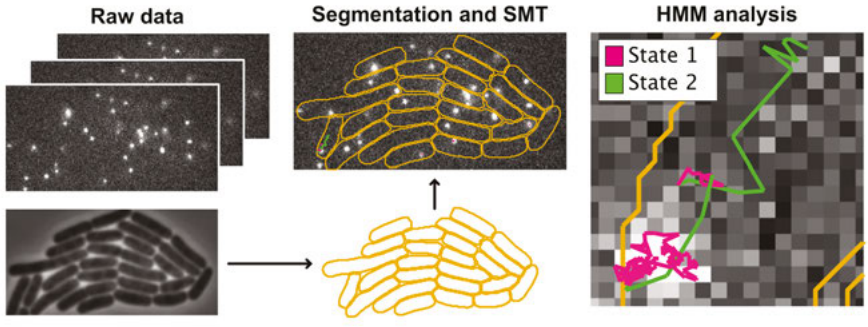

Figure 8: Analysis procedure. Cell segmentation masks are constructed from a phase contrast image (bottom panel) and aligned with a sequence of fluorescence images (top panel). SMT is performed, tracking one dot per cell. Trajectories are analyzed with HMM to extract diffusion rates. In the example trajectory (right panel), the trajectory steps are color-coded based on fitting to a 2-state diffusion model.

Next, for all detected dots and their corresponding localization, additional symmetric Gaussian PSF modelling and maximum a posteriori fitting is performed. This allows for localization refinement, and importantly, for estimation of dot localization uncertainty[78].

Each detected dot in a frame is assigned to a segmented cell. The next step is to connect dot localizations in consecutive frames, constructing a single-molecule trajectory, describing the molecular movement over time. This is not a trivial task, partly because of transient detection failures, e.g., due to out-of-focus movement and spontaneous blinking of the fluorophore. Additionally, samples commonly contain several fluorescently labeled particles. Together, these phenomena provide a risk of trajectory misconnections[79]. Many algorithms have been developed to account for these issues. We have used uTrack[80], where the user defines parameters for handling high particle densities. For example, these parameters define the appropriate search radii in dot connections, allowing gaps in the dot detection with user-defined penalties. The algorithm constructs the most likely set of trajectories based on a global optimization of the trajectories. Although the algorithm can handle relatively high signal densities, we build trajectories once there is only one fluorescent dot detected in a cell. This further reduces the risk of misconnections. Finally, while the fluorescence imaging is super-resolved, our cell segmentation is based on the lower-resolution phase contrast images. Consequently, it is occasionally dubious to which of two adjacent cells a fluorescent dot belongs, especially in the case of membrane protein SMT. The algorithm handles these cases by searching for dots within a cell, as well as within a user-defined distance from the cell boundaries. The dot is then localized to the cell which produced the longest trajectory.

Diffusion analysis using HMM

To extract the diffusional behavior and kinetic parameters from our tracked molecules, an ensemble of trajectories is analyzed using a diffusive hidden Markov model (HMM) algorithm[78]. HMM is a statistical model that assumes that measured data is generated by an underlying system transitioning between unobservable (hidden) states. It is assumed that transitions between hidden states are memory-less, i.e. the probability of transitioning depends only on the current state, not on any previous states in the sequence (i.e., a Markov chain). The goal of HMM is to infer the hidden state parameters and sequence from the observable outputs.

It should be noted that HMM can be implemented in various ways for analysis of diffusion data, as further discussed in the chapter *TF diffusion – a comparison between Paper I and II*. In the implementation in **Paper II** and **III**, diffusion states were treated as hidden states, with the observables being trajectory steps of varying lengths. An ensemble of trajectories (normally in the range of thousands for a certain experimental condition) were fitted to a user-defined number of hidden states. Using MLE, diffusion rates and transition probabilities for the given model size are estimated, thus providing each state's average diffusion rate, partial occupancy and average dwell time (Figure 5b). Unlike other traditional HMM-based algorithms such as vbSPT, this algorithm makes explicit use of the dot localization uncertainties for localization refinement and motion blur correction[81].

Importantly, since both the diffusion rate and the transition frequency are used as fitting parameters, the HMM is able to separate states not only based on differences in diffusion rates, but also based on differences in transition frequencies. In principle, the model should therefore be able to separate biological states with similar diffusion, but different binding kinetics. To our knowledge, **Paper II** is the first instance that this phenomenon has been observed in an actual biological context.

One drawback of our HMM method in contrast to, for example, vbSPT, is that it does not provide any measure for what model size best fits the data. While there are statistical criteria for finding the best fit, such as Akaike's information criterion (AIC) or the Bayesian information criterion (BIC), we find that these perform badly on our data. This is likely due to the heterogeneous nature of diffusion *in vivo*, where molecular motion likely displays a set of diffusion distributions rather than discrete diffusion states. As such, for our large datasets, multi-state models will fit the data better according to AIC (*SI Appendix*, Figure S2 in **Paper II**).

As our objective is to find a model describing a biological system, we aim to obtain the simplest possible model that can explain the data and makes

biological sense. As such, we routinely fit our data to models of 2-9 diffusion states, assessing if states accumulate in discrete clusters with increasing size (Figure 9a). If so, we coarse-grain larger models into fewer states, using diffusion rate thresholds. We base our diffusion thresholds and number of states in the coarse-graining on the following:

1. Clustering of diffusion rates in multi-state models.
2. A-priori knowledge about the biological system at hand, i.e., how many biological states do we expect to see?
3. Comparison to diffusion rates determined from other molecular species. For example, a small protein bound to a ribosome should display ribosome-like diffusion.

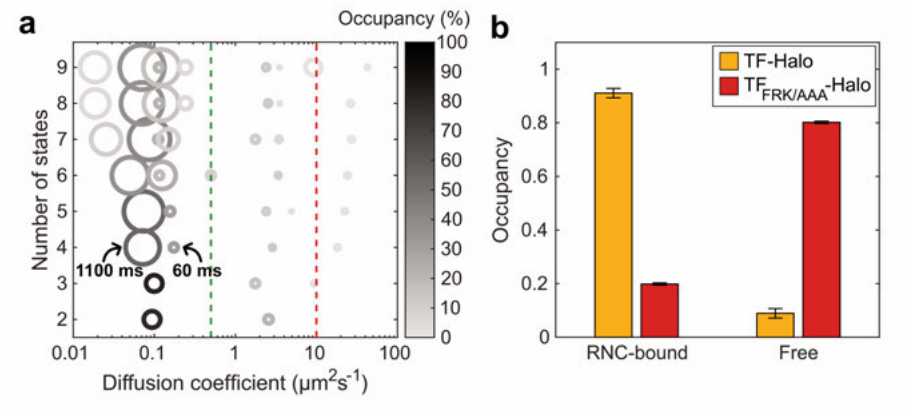

Figure 9: Assessment of TF-Halo HMM models. a. Empirical assessment of diffusion clustering in multi-state models. Data was fitted to models with 2–9 states. Each model is resolved along the y-axis with discrete diffusion states represented by circles, separated along the x-axis based on diffusion rate. In this example, circles are color-coded according to state occupancy and the area is proportional to the state dwell time. As reference points, two dwell times in the 4-state model are explicitly indicated. Green and red lines mark diffusion thresholds at 0.5 and 10 $\mu m^2 s^{-1}$, respectively, separating the three clusters of diffusion states (RNC binding, free TF-Halo, and tracking artifacts). **b.** Example of a biological perturbation control experiment. A TF-Halo mutant compromised in RNC binding displays less slow diffusion, confirming that slow diffusion reflects RNC binding. This figure is also presented in Paper II.

Coarse-graining further circumvents that HMM assumes a single-exponential distribution of the state dwell time. This assumption is not problematic for simple kinetics, such as a one-step binding and unbinding event, which indeed follows an exponential distribution. In our experiments, a diffusion state often corresponds to a more complex series of reactions, such as translation elongation when tracking ribosomes[82]. In such cases, multistate coarse-graining

allows for more accurate measurements of otherwise underestimated non-exponential dwell times.

With this empirical coarse-graining approach, our ambition is to obtain the simplest possible model that sufficiently describes the biology. The validity of the model must then be assessed by biological control experiments, as exemplified in Figure 9b.

Using *in vivo* smFRET to study molecular interactions

In **Paper III**, we employed a third *in-vivo* single-molecule microscopy technique – smFRET. A comprehensive view on the experimental and analytical aspects of smFRET is provided in a recent review by Ha *et al* (2024), and briefly summarized below.

FRET is the non-radiative transfer of energy from an excited molecule (donor) to another molecule (acceptor) through dipole-dipole interactions (Figure 10a). In the case of smFRET, the acceptor and donor molecules (FRET pair) are dyes. The donor is a fluorophore, while the acceptor is either a fluorophore or a non-fluorescent dye (quencher). A sample is illuminated with light for donor excitation, and upon FRET, there is a detectable change in fluorescence from both fluorophores. Depending on the dyes used, FRET either causes acceptor emission with anticorrelated donor intensity decrease, or quenching of the donor fluorescence if the acceptor is a quencher.

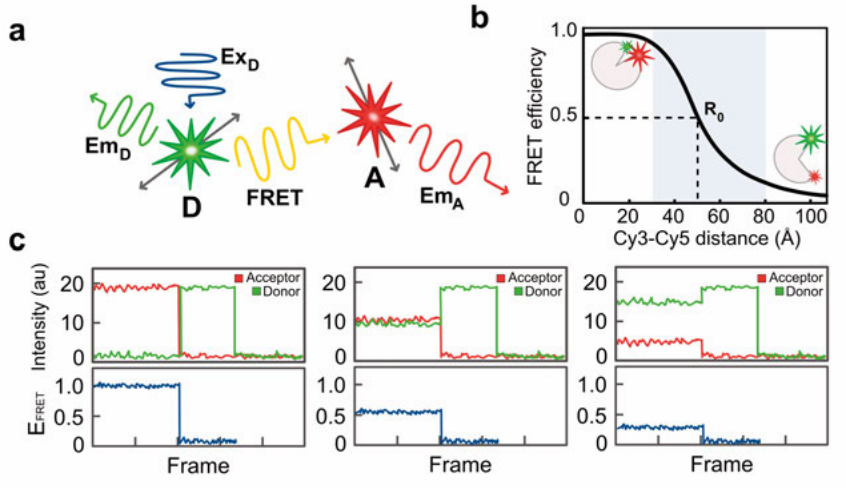

Figure 10: Basics of smFRET. a. A donor fluorophore (D) is excited with light (Ex_D). Some light is emitted as fluorescence (Em_D) and some is subject to FRET, causing fluorescence emission from a nearby acceptor (A). **b.** Schematic of distance dependence on FRET efficiency (E_{FRET}) between Cy3 and Cy5. **c.** Schematics of smFRET traces displaying high- (left), mid- (center), and low-FRET (right). The acceptor and donor emission is anti-correlated.

The efficiency of FRET (E_{FRET}) is highly sensitive to the distance, r, between the FRET pair, displaying a dependence of $1/r^6$. The distance at which a FRET pair displays 50% E_{FRET} is called the Förster Radius (R_0). One commonly used pair in smFRET is Cy3 (donor) and Cy5 (acceptor). These display large changes in E_{FRET} at distances between 30-80 Å, with R_0 around 50 Å (Figure 10b). Hence, if observed over time, Cy3-Cy5-based FRET is a great molecular ruler for measuring intermolecular interactions (e.g., binding kinetics) and intramolecular interactions (e.g., protein folding). Further, the single-molecule methodology allows for detection of transient intermediates and detection of rare events, otherwise hidden in classical ensemble FRET measurements[83].

A comprehensive mathematical description of FRET is beyond the scope of this Thesis summary. Here, I will highlight only the aspects that are relevant to understanding the analysis in **Paper III**, where we used smFRET to probe the folding state of Cy3Cy5-labeled OmpA in live *E. coli*. To this end, we used our camera-based widefield epifluorescence microscope with Cy3 (donor) excitation and simultaneous recording of any donor and acceptor (Cy5) emission. The signal output is thus intensities from detected dots in a fluorescence movie. In the case of FRET, there should be matching signal in the acceptor and donor channels. The FRET efficiency can be quantified accordingly:

$$E_{FRET} = \frac{I_A}{I_A + I_D} \quad (1)$$

Where I_A and I_D denote corrected (as described below) acceptor and donor intensity, respectively, and E_{FRET} takes on a value between 0 and 1. At maximum FRET ($E_{FRET} = 1$), it becomes evident that $I_A >> I_D$, which in practice means maximum acceptor emission and very little or no donor emission (Figure 10c). Further, for $E_{FRET} = 0.5$, acceptor and donor display equal emission. At $E_{FRET} < 0.5$, donor emission is higher than acceptor emission. Moreover, FRET is characterized by an anti-correlated nature between donor and acceptor emission.

In order for Equation 1 to accurately describe the FRET efficiency, background noise must be subtracted from the signal as the two different channels may suffer from different levels of background noise. The background subtraction also ensures zero-intensity signal after a photobleaching event. Further, for maximal robustness and reproducibility, three additional correction factors should be included in the analysis.

First, the α-factor accounts for donor emission bleed-through in the acceptor channel. Second, the δ-factor corrects for any direct excitation of the acceptor fluorophore from the donor excitation light. Third, the γ-factor corrects for differences in quantum yield and signal detection between the two fluorophores. These correction factors eliminate distortions in the data and are crucial when measuring absolute distances based on the E_{FRET} signal. However, in applications where E_{FRET} is used as binary on/off signal (high/low E_{FRET}), partial correction may suffice[83].

In addition to these sources of distortion, other physical phenomena affect E_{FRET}, especially when applied on a dynamic molecule in a complex molecular environment. These include the orientation factor accounting for the relative orientation between the fluorophores' transition dipoles, which is approximated to 2/3 when the dipoles display complete random, orientational freedom with extensive rotation during excitation. Fluorophores tethered to rigid or otherwise confined molecules, may experience rotational bias[84,85]. Next, differences in local chemical environments, such as presence of lipids or differences in pH and oxygen concentration, may locally alter the photophysical properties of a fluorophore. Such local chemical differences are prevalent in cells and may result in, for example, a high degree of sporadic signal quenching, blinking, or bursts of emission[58,86,87]. Lastly, when applying in-cell smFRET with widefield epi-fluorescence microscopy, signal heterogeneity can arise from out-of-focus signal and a relatively low SNR, possibly adding even more noise to the FRET efficiency measurements.

Results

TF diffusion – a comparison between Paper I and II

In the first part of this work, we focused on the RNC-binding dynamics of TF, a chaperone that co-translationally aids *de novo* folding of the nascent chain. We applied two different SMT methods to measure the binding kinetics of TF in cells with active translation. In **Paper I**, we used a 3D scanning-oriented modulation-enhanced localization approach, as described in *3D-SMT using modulation-enhanced localization microscopy with confocal detection*. In **Paper II**, we used camera-based 2D super-resolution fluorescence microscopy, as described in *2D-SMT using camera-based super-resolution fluorescence microscopy*. In this chapter, I will compare the results from the two methods.

TF labeling

For SMT of TF in *E. coli*, we fused the HaloTag protein to the C terminus of TF (TF-Halo), expressed either from the chromosome, from a plasmid in a WT background or in a TF knockout strain (Figure 4d). The fusion is functional *in vivo*, as shown in Supplementary Note 10 of **Paper I** and further discussed in subsection *RNC binding kinetics depend on the cellular level of TF-Halo* in **Paper II**.

Diffusion rate measurements in Paper I

In Paper I, TF-Halo was used as a proof-of-concept for 3D, sub-millisecond SMT. Tracking was applied to TF-Halo and a TF-Halo mutant, compromised in ribosome binding, with a temporal resolution of 0.84 ms. To extract pointwise diffusion rates from the trajectories, a covariance estimator (CVE) method[88–90] was extended to not only consider the displacement between consecutive steps, but rather, to any user-defined number of steps (ECVE). The rationale was that, at such high temporal resolution, a single step will yield a shorter displacement, thus making the diffusion rate estimation more sensitive to noise and localization error.

When applied on TF-Halo data, a lag-time of 8 steps with averaging over 16 points had to be employed in order to yield robust diffusion rate estimates. Thus, although the method itself allows for sub-millisecond tracking, using these pointwise diffusion rates for inferring ribosome binding limits our temporal resolution to at least 0.84 ms*8 frames = 6.7 ms. Consequently, transitions between slow and fast diffusion below this timescale will not be resolved.

Finally, to yield a model describing discrete diffusion states with partial occupancies, the ECVE-estimated diffusion rates were subsequently trained on a 3-state HMM, assuming that the diffusion rates in each state follow an inverse-gamma distribution. With this, we obtained a model with two distinct

RNC-bound diffusion states (states 1 and 2) and a third state with freely diffusing TF-Halo (Table 1).

Table 1: Comparison of TF-Halo diffusion models from Paper I and II

		State 1	State 2	State 3
Paper I	D ($\mu m^2 s^{-1}$)	0.32 ± 0.02	0.86 ± 0.12	2.67 ± 0.20
	Occupancy (%)	35 ± 6	44 ± 7	21 ± 3
Paper II*	D ($\mu m^2 s^{-1}$)	0.043 ± 0.001	0.135 ± 0.002	2.46 ± 0.04
	Occupancy (%)	51.5 ± 1.3	38.6 ± 1.1	8.4 ± 0.3
Paper II**	D ($\mu m^2 s^{-1}$)	0.072 ± 0.002	0.174 ± 0.002	2.87 ± 0.04
	Occupancy (%)	55.6 ± 1.2	31.4 ± 1.0	12.1 ± 0.4
Paper II***	D ($\mu m^2 s^{-1}$)	0.079 ± 0.002	0.214 ± 0.003	3.03 ± 0.03
	Occupancy (%)	47.8 ± 1.2	27.4 ± 0.7	22.8 ± 0.6

* Same expression system as in Paper I (Paper II, Dataset S1, Tab 4, model size 4).
** Chromosomal TF-Halo (Paper II, Dataset S1, Tab 4, model size 4)
*** Low TF-Halo WT background (Paper II, Dataset S1, Tab 13, model size 4)

Diffusion rate measurements in Paper II

In **Paper II**, we used camera-based super-resolution fluorescence microscopy for 2D SMT. In contrast to Paper I, this method employs wide-field, stroboscopic, epi-illumination of TF-Halo molecules, and the emitted photons are recorded using an EMCCD camera, with a frame rate of 200 Hz (5 ms per frame), thus yielding a temporal resolution similar to that in Paper I. After SMT, we applied HMM to build diffusion models directly from the trajectory step lengths, using diffusion rates and transition frequencies as fitting parameters[78]. Thus, although HMM is utilized to obtain diffusion models in both papers, it is employed differently. Another notable difference is that this direct illumination with longer laser exposures generally leads to shorter trajectories and fewer time steps compared to the modulation-enhanced localization approach. Moreover, tracking in 3D provides axial coordinates, whereas tracking in a 2D focal plane comes with the caveat of molecules moving out of focus, further leading to shorter trajectories.

Comparison of diffusion rates between Paper I and II

Despite the vast differences in the SMT and analysis methods, both approaches identify two RNC-bound and one free diffusion state (Table 1), as verified by tracking of the TF-Halo mutant unable to bind ribosomes. We note that the exact numbers obtained from the models vary to some extent. Most notable is the discrepancy in the RNC-bound diffusion rates, where the rates in Paper I are faster than expected from translating ribosomes[82]. As demonstrated by tracking of immobile fluorescent beads (Paper I, Figure 4), this is

an overestimation caused by an inherent tradeoff in building a system well-suited for tracking slow- and fast-moving dots at the same time. Obtaining a high temporal resolution for fast-moving dots requires a larger illumination pattern than what is necessary for slow dots. The larger illumination pattern consequently affects the spatial resolution negatively. Thus, this method becomes more sensitive to localization errors for slow-moving dots, with faster apparent diffusion as a consequence. In contrast, the diffusion rates obtained from the camera-based SMT in Paper II are very robust independent of expression system used, though with some variation in state occupancies and dwell times. Differences in state occupancies and dwell times are further discussed in the next section.

A high temporal resolution resolves different binding modes of TF with high on- and off-rates

As discussed in **Paper II**, the ratio between TF and RNCs will affect the occupancy and average RNC-bound dwell time in the model, explaining the observed variations between different expression systems (Paper II, Figures 3 and 4). Nevertheless, we repeatedly found that the two RNC-bound states differ in that they have distinct dwell times; one is on the second-long timescale and the other lasts for tens of milliseconds. We show that these represent distinct distributions of longer and shorter RNC bindings. The states are separated in the HMM not due to different diffusion rates, but rather, due to different underlying binding kinetics, and consequently, different transition frequencies (Paper II, Figure 5). By reducing the camera exposure to 2.5 ms, the HMM results were similar (Paper II, *SI Appendix* Figure S10), suggesting that what we observe in our 5 ms-tracking is the lower limit for the kinetics we can resolve in our setup. When increasing the exposure to 60 ms, all shorter bindings become invisible to the camera. This filters out the longer bindings alone, which on average last for 1-2 s. Importantly, identical 60-ms tracking and coarse-graining analysis on translating ribosomes yielded average dwell times of 20-50 s[82], attesting that the second-long average dwell time of our TF-RNC binding is not an underestimation caused by experimental artifacts. This average is shorter than what is commonly proposed from target-specific binding to stalled ribosomes[19], although binding times vary greatly between experiments, and shorter times have also been reported[18].

It is not inconceivable that ribosome stalling severely affects binding kinetics as the TF-target specificity will remain static on a stalled ribosome. During active translation, however, the chain is growing, with a new amino acid appearing from the exit tunnel on average every 50 ms. As such, the target specificity on an active RNC is continuously changing with respect to TF, which

could explain why TF most frequently stays bound for on average only 50 ms. Hypothetically, TF stays associated for longer only when extended patches of target sequences appear from the ribosome exit tunnel (1 long binding event per 16-25 shorter bindings, Paper II, *SI Appendix*, Table S9 and S11). It should be stressed that our model only reports on averages from two distributions of dwell times that are distinct enough to be separated. As such, we cannot exclude that bindings lasting for tens of seconds occur, but if they do, they must be too rare to appear separately in our analysis.

Another important aspect of our kinetic model is the short residence time in the free state – on average only 40 ms between binding events. At 60 ms tracking, as performed by us and by Yang *et al*. (2016), release events are thus most often missed. This causes poorly-resolved, "averaged" diffusion states with rates faster than those of RNCs, but slower than free TF (Paper II, *SI Appendix* Supplementary Note S2). A similar averaging effect was observed in our 5-ms tracking, where the average free diffusion rate is ca 3 $\mu m^2 s^{-1}$, and the short-lived RNC-bound state has a slightly faster apparent diffusion rate compared to the long-lived RNC state. In contrast, the TF mutant which cannot bind ribosomes, shows an average free diffusion of ca 6 $\mu m^2 s^{-1}$. Indeed, there thus seems to be an averaging effect caused by TF's ribosome-binding capability. Hence, it is likely that the binding and unbinding kinetics are sometimes even faster than our temporal resolution.

Studying the RNC-binding competition between co-translational processing factors

The co-translational processing factors SRP and TF share docking site on the ribosome[6,7,29]. TF has a large substrate pool of cytosolic and secretory protein-RNCs, and is particularly enriched on outer-membrane RNCs[17]. In contrast, TF is underrepresented on inner-membrane RNCs, which instead are targeted by SRP[91], suggesting that the two factors exert some kind of RNC-binding competition *in vivo*. However, this competition has been elusive, as there are also models supporting concurrent binding[18,92–96]. In the last part of **Paper II**, we therefore investigated the interplay between these factors *in vivo*, performing SMT on labeled SRP at gradually increasing concentrations of (unlabeled) TF-Halo.

We found that the RNC-bound occupancy of SRP decreases with increasing TF concentration, concluding that they indeed compete for binding *in vivo* (Paper II, Figure 8). Interestingly, the RNC-bound dwell time displayed the opposite trend; at increasing TF-Halo levels, SRP's average RNC-bound time increased. We pose that a high TF concentration increases the binding specificity of SRP, as reflected in a higher proportion of productive co-translational

targeting cycles compared to short-lived probing of non-targets, that are otherwise occupied by TF. We could show that this effect was TF-specific, as overexpression of HaloTag alone did not affect the RNC-bound occupancy, nor dwell time, of SRP.

These results mark our first attempt in studying the interplay between co-translational processing factors *in vivo*. Future work will focus on additional factors. Among these are the N-terminal processing factors Peptide deformylase (PDF), which removes the formyl group on fMet[97], and Methionine aminopeptidase (MAP), which subsequently hydrolyzes the deformylated methionine[98,99]. In *E. coli,* most proteins are deformylated during translation, whereas the methionine is removed in only about 50% of them[100]. PDF and MAP share ribosome binding site, however, in the presence of PDF, MAP may bind to a secondary site[101]. Further, PDF and SRP may bind simultaneously[18], but prior SRP binding prevents deformylation by PDF[102]. Lastly, PDF and TF may bind simultaneously[18,101]. These results are compiled from structural data and isolated *in vitro* binding kinetics experiments in reconstituted systems with little resemblance to the cellular environment. Our SMT method could address whether the proposed competitions are applicable to live cells.

Pitfalls in biological perturbation control experiments

As described in *Diffusion analysis using HMM*, we rely on verifying our models by biological perturbations. In **Paper II**, we could show that the effect on SRP-RNC binding was TF-Halo-specific, because overexpression of HaloTag alone did not cause similar effects. However, not all control experiments were as successful.

In initial interpretations of the kinetic model of TF, I implied that the long-lived RNC-bound state corresponds to target-specific bindings, and that the shorter-lived state is non-target binding, as expected from the literature. I tried to verify this by overexpressing different proteins, for example, the TF-target OmpA in hopes of enriching the long-lived RNC-bound state.

Upon plasmid-induced OmpA overexpression we observed a small increase in the coarse-grained RNC-bound dwell time, as expected (Figure 10a). However, closer inspection revealed that the increase was due to the appearance of an additional long-lived (1-4 s depending on model size), very slowly-diffusing state close to the cell membrane with significant occupancy (ca 20%, Figure 10b-c).

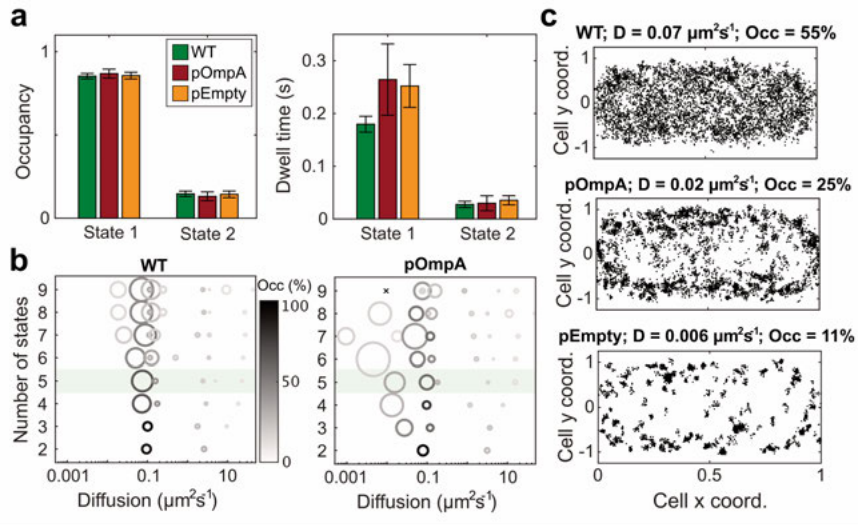

Figure 11: OmpA overexpression causes membrane dwelling by TF-Halo. a. Occupancy and dwell time in coarse-grained 2-state models of chromosomally expressed TF-Halo (green), with induced plasmid expression of OmpA (deep red), or the empty plasmid with inducer (IPTG) added (orange). **b.** Comparison of multi-state models with and without OmpA overexpression. Upon OmpA expression, long-lived very slow states appear that are not present in the WT. Circles are color-coded according to state occupancy and the area is proportional to the state dwell time. 5-state models are highlighted in green. **c.** Projections of TF-Halo coordinates assigned to state 1 in 5-state models in a normalized cell geometry. N = 99,352; 58,563; and 27,700 trajectory steps from 3, 3, and 2 experiments for WT, pOmpA, and pEmpty, respectively.

Since OmpA is post-translationally targeted to translocons, this new state most likely does not represent our hypothesized "target" RNC-bound state, but rather, TF-OmpA complexes or aggregates accumulating by the membranes. Our 2-state coarse-graining thus erroneously incorporates this in the RNC-bound state, resulting in the longer average dwell time. As such, the experiment did not help define the different RNC-bound states of TF, but instead caused adverse effects that made direct comparisons to the wild-type (4-state) model difficult. Further, when we inserted the empty plasmid as a negative control for the OmpA overexpression, we also observed a long-lived, membrane-associated state, albeit less occupied than for the OmpA overexpression (Figure 10c).

This exemplifies how one must be very cautious to the adverse systemic effects that a biological perturbation, such as protein overexpression, may pose. Further, the results from the perturbation experiment itself need verification by negative control experiments. These were successful in the SRP experiments, but not in our OmpA controls. Therefore, the OmpA control results were not included in Paper I, and we changed the assignment from "target" and "non-target" binding to "long-" and "short-lived" binding. This also

highlights that our methodology is dependent on prior biochemical work for explaining our models, and that we lack resolution to find detailed and groundbreaking mechanistic features. Instead, the strength of our method lies in complimenting seminal biochemical and structural data with a systemic and dynamic perspective from living cells.

In vivo measurements of outer-membrane protein biogenesis – the bigger picture behind Paper III

In the final year of my PhD, I dove deeper into OmpA and its biogenesis. This resulted in some preliminary results presented in the unpublished **Paper III**. In this section, I will first introduce the overall aim and experimental design of this project. Then, I will summarize the progress as presented in Paper III.

Experimental design

Our big-picture aim of this project is to follow the entire post-translational targeting, translocation and OM folding of OmpA in live *E. coli* using SMT combined with smFRET. The preliminary experimental design is based on *in vitro* labeling of OmpA and delivering it back in an unfolded conformation to *E. coli* by electroporation[50,71]. Details are given in **Paper III**, and summarized here in Figure 12a.

After electroporation, cell recovery and washing causes a delay of at least a few hours from the electroporation event until image acquisition. To prevent all incorporated OmpA molecules from translocating before image acquisition, we will add a blocking domain upstream of the N terminus, making the signal peptide inaccessible to the translocation machinery (Figure 12b). The blocker contains a TEV protease cleavage site. Upon TEV protease (TEVp) expression, the blocker domain is cleaved off, yielding the native, translocation-competent OmpA. The blocker domain is labeled with a FRET quencher (L1), placed such that it quenches a nearby second label (L2) downstream of the signal peptide on OmpA. The onset of fluorescence signal from L2 hence is due to cleavage of the L1-containing blocker domain by TEV protease. This marks the zero-timepoint for an OmpA targeting and translocation event. By regular SMT, we record the time from the appearance of the L2 signal in the cytosol, until it translocates to the periplasm. The photophysical properties of L2 should be sensitive to the chemical environment, yielding a detectable shift in emission upon entering the periplasm[103,104]. This could for example be due to the oxidative environment in the periplasm, or tunable differences in pH between the periplasm and cytoplasm[105,106]. Finally, L2 is a FRET donor to a third label (L3) placed on the C terminus of OmpA. In the unfolded conformation, L2 and L3 are too far apart for FRET, whereas they end up within a

high-FRET distance in the folded structure. Hence, the onset of a FRET signal marks the timepoint of OM insertion and folding. With this design, there are three detectable changes in fluorescence, marking checkpoints in the OMP biogenesis pathway; the zero-timepoint of a membrane-targeting event from the cytoplasm; translocation to the periplasm; and folding in the OM.

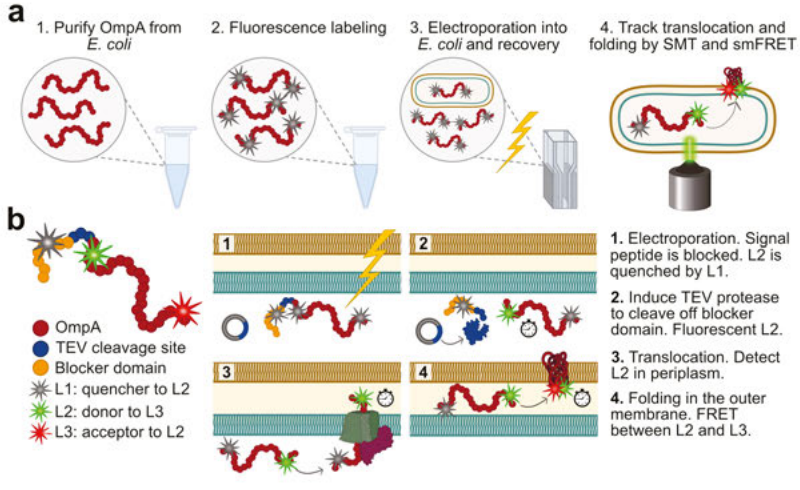

Figure 12: Experimental design. a. Overview of experimental protocol. **b.** Design of inducible translocation competence of blocker-OmpA, and SMT and smFRET to record its targeting, translocation and folding.

Control experiments to assess the practical limitations of our design

For the design described above, we require three different site-specific dye labels with high brightness, photostability and FRET-compatibility. They also need to be cell-permeable and small enough not to interfere with any other components of the biogenesis pathway, such as the SecYEG translocon or the BAM complex. Finding such fluorophores is not trivial and will require careful *in vivo* characterization. However, my present work has been focused on two other important aspects of the design that require verification:

1. Is it possible to delay and trigger membrane targeting by masking the signal peptide with a TEV-cleavable blocker domain?
2. Is it possible to electroporate unfolded OmpA and have it end up folded in the OM?

Assessing blocker-dependent translocation delay

To address the first question, we fused the signal peptide of OmpA (ss_{OmpA}) to the N terminus of the fluorescent protein mCherry, yielding a fluorescent periplasmic reporter protein (Figure 13). We designed a set of constructs with

additional blocker domains of different sizes upstream of ss_{OmpA}. The smallest domain was a 6x-Histidine tag (0.8 kDa). This domain was not sufficient in blocking translocation of ss_{OmpA}-mCherry. Next, we extended the construct with three different domains; the 3xFLAG tag (22 residues, 2.7 kDa), the native *E. coli* protein Thioredoxin (Trx, 113 amino acids, 12 kDa), and finally, the SNAP-tag (186 amino acids, 19 kDa). The larger the domain, the less peripheral fluorescence was observed. We next verified that the blocker domains could be cleaved off by induction of TEV protease, as evident by an increase in periplasmic fluorescence. Out of all blockers tested, Trx and 3xFLAG seem to employ the best trade-off between efficient blocking and cleavage. With this, we conclude that it is possible to delay and induce translocation by introducing a cleavable blocker domain on the N terminus of secretory proteins.

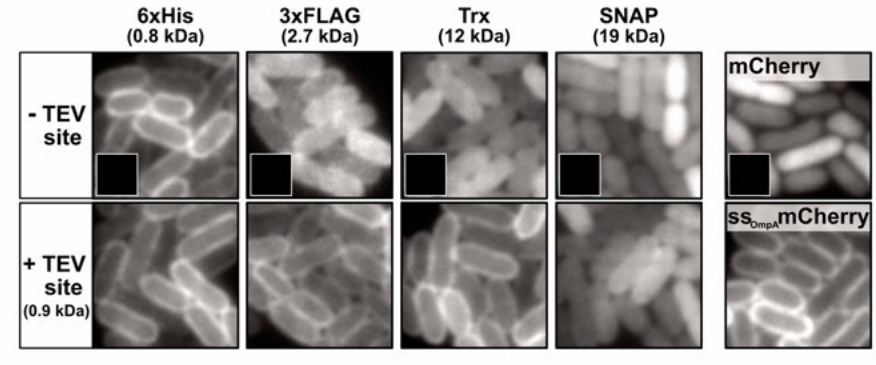

Figure 13: Blocker-dependent translocation delay. *E. coli* expressing TEV protease, with plasmids encoding ss_{OmpA}mCherry + different N-terminal blocker domains ± tev site. Translocation inhibition becomes more pronounced with increasing blocker size (top panel). Inhibition can be relieved by TEV protease when a tev cleavage site (ENLYFQG, 0.9 kDa) is inserted between the blocker and ss_{OmpA}mCherry. The rightmost panel is cytosolic mCherry (top) and periplasmic ss_{OmpA}mCherry (bottom). Brightness and contrast settings are the same for each blocker domain ± tev site, but may vary between blockers due to different protein expression levels. The small images are of WT *E. coli* for autofluorescence comparison. Images are representative examples from two independent experiments.

Assessing in-cell smFRET for detecting folded OmpA – Paper III

The second question listed above was addressed in **Paper III**, where we were able to purify unfolded OmpA, label it with Cy3 and Cy5 *in vitro*, and electroporate it back into *E. coli*. Regular SMT showed that a considerable fraction of the incorporated Cy3Cy5-OmpA ended up by the outer membrane (Paper III, Figure 2), as expected. Further, preliminary results show that we can detect very rare events of high FRET, which could correspond to folded OmpA (Figure 14a-b). However, there are also instances of molecules displaying lower

FRET. These could correspond to sporadic signal from misfolded proteins where Cy3 and Cy5 end up at varying distances between each other, or, it could be due to experimental and technical impurities arising from the very noisy *in vivo* environment.

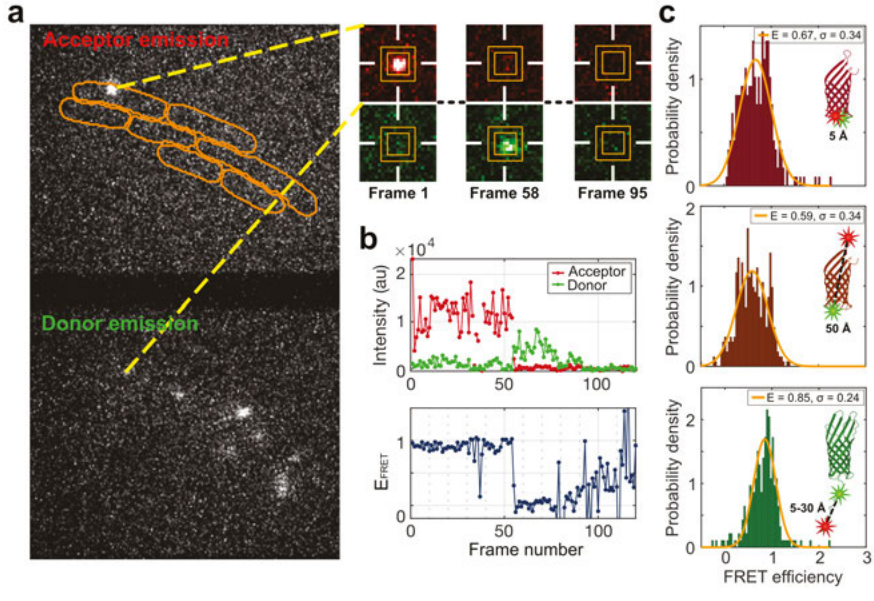

Figure 14: smFRET of Cy3Cy5-OmpA constructs. a. Example of anti-correlated high-FRET signal in fluorescence movie. **b.** Acceptor and donor intensity (top) and E_{FRET} (bottom) for the molecule displayed in panel a. **c.** E_{FRET} distributions for the folding reporter (top), a low-FRET control (middle), and a high-FRET control (bottom). The distributions are cumulated frame-wise E_{FRET} values from all traces passing the FRET selection criteria. Details are found in Paper III.

To assess if the FRET efficiency distribution contained a fraction of "true" signal from folded OmpA, we designed a low-FRET OmpA construct expected to display lower FRET in the folded conformation, and a high-FRET OmpA expected to always display high FRET, irrespective of its conformation. We argued that any misfolding or photophysical effects would be stochastic and equally prevalent in the low- and high-FRET OmpA controls as for the folding reporter. Thus, if there was a significant difference in E_{FRET} distributions between these reporters, this should be due to a fraction of correctly folded protein in the data. Our preliminary results point to a detectable difference (Figure 14c), but more data and additional control experiments are required for validation. For example, the correction factors described in *Using in vivo smFRET to study molecular interactions* have not been included in this preliminary analysis, likely causing distortion to the inherently noisy signals

from *in vivo* smFRET. I refer the reader to **Paper III** and Future Perspective for more details on further control experiments. Although preliminary, these results mark the first steps in developing a system for studying OMP biogenesis at high spatiotemporal resolution in live cells.

Future perspective

In this work, I first used SMT to study the RNC-binding kinetics of TF, and for investigating the binding competition between SRP and TF. Apart from TF and SRP, there are a number of other co- and post-translational factors to investigate, as summarized in Figure 2. Although these factors have been extensively studied in reconstituted systems over the years, our *in vivo* single-molecule methodology may provide additional insight into the binding competition, and into transient interactions that are hidden in ensemble measurements. The end-goal would be to build a quantitative model describing the cooperation and competition between factors, derived from SMT in knockout strains and during overexpression. Since such experiments rely on rather invasive biological perturbations, much caution and rigorous controls are required. Further, this model should describe the frequency of RNC binding by each factor. For such a quantitative comparison, we need a way to quantify the relative abundances of ribosomes and processing factors in our cells. This is not trivial, in part because the number of ribosomes and factors can vary greatly depending on growth conditions. Some insight might be gained from performing quantitative proteomics on our particular strains, grown under the same conditions as for our microscopy experiments. Alternatively, all SMT could be performed in a strain where the ribosome is labeled, for example with a FLAG-tag. Western blots using anti-FLAG and anti-Halo antibodies could then be performed routinely to obtain relative levels of ribosomes (translating and non-translating) and HaloTagged factors.

In the last part of my work, I started developing an experimental design to directly measure each step in the post-translational targeting, translocation and folding of an outer-membrane protein in live cells. Here, there is still a lot of work to be done. First, preliminary results presented in **Paper III** point to the possibility of using *in vivo* smFRET as a checkpoint for a folded OMP in the OM. However, the data is very noisy and even the low-FRET folding control display rare instances of high-FRET. As such, we cannot blindly trust that any high-FRET signal is due to OmpA folding. More data is required before we can determine the frequency of finding high-FRET traces in the low-FRET control. But once obtained, this may be used as a correction factor for the likelihood of a high-FRET trace corresponding to a correctly folded protein.

A comprehensive discussion about required control experiments for validation of our smFRET results is presented in **Paper III**. Here, I will reiterate the most critical aspects. First, although we showed that Cy3Cy5-OmpA could fold in the presence of synthetic lipids *in vitro*, it is crucial to assess the folding competence under more native-like conditions, as the dyes may disrupt OM embedment and interactions with the BAM complex. Hence, we should repeat the folding assay, for all three OmpA constructs, in reconstituted *E. coli* OM vesicles with BAM present. Second, we need to verify that the immobile membrane-associated dots actually represent correctly folded OmpA embedded in the OM. We are working on an immunofluorescence microscopy assay for detection of the folded conformation on the surface of electroporated cells.

Another concern is that the occurrence of FRET is extremely rare (less than 2% chance of detection), despite the prevalence of considerable donor signal in the majority of all imaged cells. It is not unimaginable that electroporation of unfolded Cy3Cy5-OmpA only produces successful folding and OM insertion on very rare occasion. However, the high-FRET Cy3Cy5-OmpA control should always display FRET, irrespective of its conformational state. Even so, FRET signal was only marginally more prevalent, suggesting that there is something wrong in the FRET detection. One possible cause could be that unfolded OmpA frequently gets truncated during electroporation and sample preparation, resulting in tracking of mostly single-labeled truncations. Another possibility is that the Cy5 signal is subject to extensive quenching in the OM, or that it easily enters a dark triplet state[107]. Preliminary attempts to add Trolox for reducing the triplet state, and oxygen scavenging systems (PCA/PCO) for increased photostability, did not seem to make a difference. However, further investigation and optimization of the Cy5 signal is required.

The purpose of Paper III was to assess if we can electroporate "native" Cy3Cy5-OmpA and detect the folded conformation using in-cell smFRET. This is in essence a positive control for the experimental design described in *In vivo measurements of outer-membrane protein biogenesis – the bigger picture behind Paper III*. The fact that FRET detection was so rare even in Paper III is concerning for the chances of capturing any kinetics in the final experimental design. There, we rely on combining these rare events with TEV-mediated cleavage of the blocker domain, which is likely quite inefficient *in vivo*. Further, the photostability of Cy3Cy5 under the current conditions leads to very short trajectories (on average only 7 frames). Hence, capturing the timing of all events (targeting, translocation and folding) in one single trajectory seems impossible at this stage. Instead, it may be better to focus on the analysis, developing tools that can combine the kinetics from the individual steps

captured in individual trajectories to form a comprehensive model of the entire OMP biogenesis pathway.

As discussed in Paper III, using smFRET may not be the easiest way to study OMP folding *in vivo* since it produces very noisy data and requires two dye labels. Instead, it might suffice to use SMT and only one fluorescent reporter that produces detectable changes in photophysical properties depending on the chemical environment (cytoplasm, periplasm, and OM). Further, advances in site-specific *in vivo* labeling of unnatural amino acids may eliminate the need for electroporation in the future[60]. As of today, our in-lab protocol for such labeling and SMT sample preparation takes approximately 10 hours. Granted that OmpA can evade degradation for such an extended time, *in vivo* labeling of a blocker-OmpA construct and concurrent blocker cleavage could be tested.

Finally, our *in vivo* smFRET setup is merely in its early stages of development. Regardless of its future with respect OMP folding utility, *in vivo* smFRET may be extended to study intermolecular interactions, which would be a crucial complement to our diffusion-based binding models. For example, we could directly measure interactions between a labeled secretory protein and labeled SecYEG translocons and BAM, and from that extract translocation and OM insertion rates.

Popular science summary

The principle for how life is maintained in a cell is more or less the same, regardless of whether it is a single-celled bacterium, or the approximately 30 trillion cells working together to build up an entire human. DNA acts as the recipe book, storing all the necessary information for cellular function. When a specific recipe is required by the cell, the information is copied into a temporary molecule called mRNA. Molecular machines known as ribosomes decode this mRNA to build **proteins**, which carry out most essential functions required for an organism to survive.

A single bacterial *Escherichia coli* cell, although a seemingly simple organism, contains thousands of different types of proteins, in millions of copies, all collaborating to ensure growth. Under optimal conditions, an *E. coli* cell divides roughly every 20 minutes. To prepare for this division, it must duplicate its entire biological content, synthesizing an astounding number of approximately 20 new proteins every second.

For a protein to perform its intended biological function, it must fold into its specific three-dimensional shape. The high speed of protein synthesis within the crowded cell causes significant risk of misfolding. These misfolded proteins can clump together into insoluble aggregates that are deadly to the cell. Nature has therefore evolved quality control systems within the protein synthesis machinery to manage this risk. Protein misfolding is linked to many severe diseases, for example, Alzheimer's, Parkinson, Huntington's, and Creutzfeldt-Jakob disease. Understanding the protein quality control systems is therefore of great importance in treatment development. Among prominent quality systems are the **chaperones**. These are special proteins, whose only purpose is to assist the correct folding of other proteins, both during and after their synthesis.

In addition to folding, each new protein must end up at their intended cellular location and receive proper chemical modifications in order to function correctly. This is ensured by a number of **processing factors** engaging with the ribosome and/or the nascent protein during synthesis. This engagement can happen while the protein is still being made (co-translational processing) or after it is released from the ribosome (post-translational processing).

Many of the co-translational factors bind at the same position on the ribosome, namely by the protein exit site. Not every new protein requires assistance from every factor, and often times, each factor has to engage at a critical timepoint during synthesis. How these factors find their specific targets among tens of thousands of ribosomes, each translating a different protein, engaging precisely when needed without hindering the activity of other processing factors, is quite poorly understood, despite decades of research.

Traditionally, such research is based on isolating the molecules involved in the process of interest from cells, and measuring their biochemical activity in a test tube. These methods have provided high-resolution mechanistic details of individual processing factors. However, the test tube cannot mimic how all these factors work together in the complex, crowded, dynamic environment of a living cell, and consequently, might not be representative for how the processes play out in cells. Just imagine trying to learn how to play a game of football by studying one player at a time, isolated from the playing field. Moreover, these methods often rely on measuring the activity of millions of molecules that are synchronized to react at the same time in the test tube. The measurement is hence an average of the total reaction, whereby intermediate reaction steps and rare events become undetectable. In contrast, single-molecule methodologies allow for measurements on individual molecules over time, thus enabling detection of transient and rare states.

In this work, we have used **single-molecule microscopy** techniques to study different aspects of co- and post-translational processing of proteins, in real time, inside living *E. coli* cells. First, we employed two different microscopy techniques to benchmark the activity of the chaperone Trigger factor (TF). This chaperone binds to the ribosome, ensuring that the protein does not get misfolded during synthesis. TF assists folding of most new proteins as they emerge from ribosomes, but binds stronger to certain target proteins. In our studies, we could distinguish long and short ribosome-binding kinetics, a technical advancement that has not been accomplished before. We saw that TF-ribosome binding is highly dynamic, and the average ribosome-bound time was shorter than commonly perceived from test-tube measurements. We believe that this discrepancy is because the ribosomes are not performing protein synthesis in the test tube, in contrast to our live-cell experiments. We could also observe that TF competes for ribosome binding with another processing factor, the Signal recognition particle (SRP). Importantly, the presence of TF makes SRP's engagement with ribosomes more efficient, as TF blocks SRP from engaging with ribosomes producing non-SRP target proteins. Taken together, our study provided an overall view on the activity of TF in live cells.

In the second half of this work, we focused on the synthesis of a specific class of proteins, namely, the outer-membrane proteins (OMPs). Many pathogenic bacteria are enclosed by an outer membrane (OM), containing a large number of OMPs which are crucial for survival. The OM and OMPs make these cells practically impermeable. Consequently, this class of bacteria is especially resilient against many antibiotics, as the antibiotics cannot enter the cell. Hence, attacking the OM and the OMPs is an attractive approach for killing these pathogens. OMP synthesis has proven particularly challenging to study in test-tube experiments. Thus, we started developing an experimental platform for seminal live-cell studies on OMP synthesis. Preliminary results show that we can record the activity and folding of OMPs in live *E. coli*, but further control experiments are required before our method can be applied to novel in-cell OMP studies.

Populärvetenskaplig sammanfattning

Den centrala dogmen beskriver den grundläggande principen för hur liv upprätthålls i en cell, vare sig det gäller en encellig bakterie, eller de cirka 30 biljoner celler som samarbetar för att bygga upp en hel människa. I cellen finns DNA som fungerar som en receptbok. När cellen behöver använda specifika recept kopieras informationen till en kortlivad mRNA-molekyl. Ribosomer avkodar sedan mRNA för att bygga **proteiner**, som i sin tur utför majoriteten av de funktioner som krävs för att en cell ska överleva.

En till synes enkel organism, den encelliga *Escherichia coli*-bakterien, innehåller tusentals olika typer av proteiner i miljontals kopior som alla samarbetar för att säkerställa att cellen kan växa och dela på sig. Under optimala förhållanden delar sig en *E. coli*-cell ungefär var 20:e minut. För att förbereda sig för delning måste bakterien kopiera hela sitt biologiska innehåll och därmed tillverka så många som 20 nya proteiner varje sekund.

För att ett protein ska ha sin rätta biologiska funktion måste det veckas ihop till sin korrekta tredimensionella struktur. Men cellens höga proteinsynteshastighet och trånga miljö leder till en stor risk att nya proteiner felveckas. Felveckade proteiner kan ansamlas i olösliga klumpar som blir dödliga för cellen. Naturen har därför utvecklat olika kvalitetskontrollsystem inom proteinsyntesmaskineriet för att kunna hantera detta. Proteinfelveckning är kopplat till många allvarliga sjukdomar, till exempel Alzheimers, Parkinson, Huntingtons och Creutzfeldt-Jakobs sjukdom. Därför kan en ökad förståelse för cellers proteinkvalitetskontrollsystemen vara viktigt för utveckling av nya behandlingar. Bland framträdande kvalitetssystem finns **chaperon-proteinerna**. Dessa är speciella proteiner vars syfte är att assistera i veckningen av andra proteiner, både under och efter syntes.

Utöver veckningen måste varje nytt protein hamna på sin avsedda plats i cellen och genomgå kemiska modifieringar för att fungera korrekt. Detta säkerställs av ett antal **hjälpfaktorer** som interagerar med ribosomen och det nybildade proteinet. Denna interaktion kan ske under pågående syntes (kotranslationell bearbetning) eller kort efter att det nya proteinet har lossnat från ribosomen (posttranslationell bearbetning).

Många av de kotranslationella hjälpfaktorerna binder till samma yta på ribosomen, och alla nya proteiner kräver inte bearbetning från alla faktorer. De

olika faktorerna måste dessutom ingripa vid specifika tidpunkter under proteinsyntesen. Hur de hittar sina målproteiner bland tiotusentals ribosomer, som var och en producerar olika proteiner, och interagerar exakt när det behövs utan att hindra andra hjälpfaktorer är fortfarande svårbegripligt, trots årtionden av forskning.

Traditionellt har sådan forskning bedrivits genom att isolera de molekyler som tros vara involverade i en viss process från celler, följt av aktivitetsmätningar i provrör. Dessa experiment har givit detaljerad kunskap om individuella hjälpfaktorers funktion. Däremot kan provröret inte efterlikna hur alla dessa faktorer samverkar i den oerhört trånga och dynamiska miljön i en levande cell. Det finns därför en risk att provrörsbaserade mätningar inte är representativa för hur processerna utspelar sig i celler. Det är ungefär som att försöka förstå hur en fotbollsmatch går till genom att studera en spelare i taget, isolerad från fotbollsplanen och alla andra spelare. Dessa metoder bygger dessutom ofta på att mäta aktiviteten hos miljontals molekyler som är synkroniserade att reagera samtidigt i provröret. Mätningen blir därmed ett medelvärde för den totala reaktionen där intermediära reaktionssteg och sällsynta händelser blir omöjliga att upptäcka. Enmolekylsmetoder tillåter istället mätningar av enskilda molekyler över tid, vilket möjliggör upptäckt av kortlivade och sällsynta tillstånd.

I det här arbetet har vi använt **enmolekylsmikroskopi** för att studera olika aspekter av ko- och posttranslationell bearbetning av proteiner inuti levande *E. coli*-celler. Först använde vi två olika mikroskopimetoder för att studera aktiviteten hos chaperonproteinet Trigger factor (TF). TF binder till ribosomer och säkerställer att det nytillverkade proteinet inte felveckas under pågående syntes. TF guidar veckningen av flertalet nya proteiner när de kommer ut från ribosomerna, men binder starkare till vissa målproteiner. Genom banbrytande teknologi och analys kunde vi särskilja längre och kortare bindningshändelser. Vi såg även att ribosombindningen är mycket dynamisk, och den genomsnittliga ribosombundna tiden var kortare än vad man vanligtvis ser i provrörsmätningar. Vi tror att denna skillnad beror på att ribosomerna inte utför proteinsyntes i provrörsmätningar, till skillnad från i våra cellbaserade experiment. Vi kunde också observera att TF konkurrerar om ribosombindning med en annan hjälpfaktor, Signal Recognition Particle (SRP). Närvaron av TF verkar dessutom göra SRP:s interaktioner med ribosomer mer effektiv genom att TF hindrar SRP från att binda till ribosomer som inte producerar SRP-specifika proteiner. Sammantaget gav det här arbetet en övergripande bild av hur TF används i levande celler.

I nästa del fokuserade vi på syntesen av en särskild proteinklass, nämligen yttermembranproteiner. Många patogena bakterier är omslutna av ett yttre

membran, som innehåller ett stort antal livsviktiga yttermembranproteiner. Den komplexa kemiska strukturen hos yttermembranet och dess proteiner gör att dessa celler praktiskt taget är ogenomträngliga och särskilt motståndskraftiga mot många antibiotika som måste kunna komma in i cellen för att ha en dödlig verkan. Därför är yttermembranet i sig ett attraktivt mål för att döda dessa baciller. Yttermembranproteinsyntes har visat sig vara mycket utmanande att studera i provrörsbaserade experiment. Därför ville vi utveckla en experimentell plattform för cellbaserade studier. Preliminära resultat visar att vi kan följa aktiviteten och veckningen av yttermembranproteiner i levande *E. coli*-celler, men mer arbete och verifiering krävs innan metoden är färdig att användas för att beskriva ännu okända fenomen kring yttermembranproteinsyntes.

Acknowledgements

I would like to thank you, **Magnus**, for having me as a PhD student. I am grateful for your support and encouragement, and for giving me a great level of academic freedom. **Suparna**, thank you for being my co-supervisor. I have always appreciated your feedback and your positive attitude. **Irmeli**, thank you for the invaluable work you do, for your social and administrative support. **Jimmy**, thank you for always being very helpful in the lab.

I would like to thank **everyone in the Johansson lab** for your feedback and friendliness. Regarding my academic work in particular, I thank you, **Misha,** for introducing me to the lab and for being a role model in what an efficient scientist looks like. **Ivan**, thank you for your crucial work building and maintaining our microscopes, and for helping me understand how to use them. **Erik**, thank you for sharing my figure-making struggles. I am still figuring it out. I would also like to thank all **past and current members of the Elf lab**, and **Javier** and **Laura,** for showing me kindness. I am thankful for all the social activities we have enjoyed together, and for receiving much valuable feedback during presentations, always delivered in a kind and respectful way. Thank you, **Elias** and **Bo**, for your hard work on Paper I.

My deepest gratitude goes to my friends and family for your support during these challenging years. **Saga, Ulf, Elli, Julia, Linda,** and **Ellinor**, thank you for recharging my batteries through playing music and laughing together. **Elin** and **Jenny**, I am so thankful for your emotional support, for our AWs, wine evenings, parties, and most importantly, for not giving up on me during difficult times. **Olle (Nerd)**, thank you for so many fun times and for the shared frustration over our PhD studies. **Magda, Martin,** and **Vendela,** thank you for amazing parties, celebrations, and many laughs, and for introducing me to the brilliant **Gabriella**. **Minna**, thank you for being such a sweet and thoughtful friend. **Emma, Saga** and **Yvonne**, thank you for the incredible voodoo rituals. **Åsa** and **Dado**, thank you for all late evenings at your home, for always being so generous, and for allowing me to cuddle with all your animals and children. **Ulf**, thank you for all music, laughs, parties, and for sharing countless hours of silly podcast humor with me. **Tony**, I am glad you and Saga produced baby **Dino**, otherwise I would not have had anything to thank you for. **Mamma, Pappa,** and **all siblings;** I love you. Thank you for all your

support and for shaping me into who I am today. Thank you also for all your **lovely little pets** who have given me the comfort and joy that only four-legged friends can provide. **Olle**, thank you for being so good-looking, and for filling my everyday with laughter, love, and comfort. Thank you also for always honoring Partymåndag. Finally, although you have been mentioned many times already, I thank you, **Saga,** for being the best friend, role model, and big sister anyone could ever ask for. I truly could not have done this without you.

Kram till er alla.

References

1. Lopez-Campistrous, A. *et al.* Localization, Annotation, and Comparison of the Escherichia coli K-12 Proteome under Two States of Growth. *Molecular & Cellular Proteomics* **4**, 1205–1209 (2005).
2. Dolgalev, G. V., Safonov, T. A., Arzumanian, V. A., Kiseleva, O. I. & Poverennaya, E. V. Estimating Total Quantitative Protein Content in Escherichia coli, Saccharomyces cerevisiae, and HeLa Cells. *Int J Mol Sci* **24**, 2081 (2023).
3. Koubek, J., Schmitt, J., Galmozzi, C. V. & Kramer, G. Mechanisms of Cotranslational Protein Maturation in Bacteria. *Front Mol Biosci* **8**, (2021).
4. Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **181**, 223–230 (1973).
5. Mecha, M. F., Hutchinson, R. B., Lee, J. H. & Cavagnero, S. Protein folding in vitro and in the cell: From a solitary journey to a team effort. *Biophys Chem* **287**, 106821 (2022).
6. Kramer, G. *et al.* L23 protein functions as a chaperone docking site on the ribosome. *Nature* **419**, 171–174 (2002).
7. Ferbitz, L. *et al.* Trigger factor in complex with the ribosome forms a molecular cradle for nascent proteins. *Nature* **431**, 590–596 (2004).
8. Kramer, G. *et al.* Functional Dissection of Escherichia coli Trigger Factor: Unraveling the Function of Individual Domains. *J Bacteriol* **186**, 3777 (2004).
9. Merz, F. *et al.* The C-terminal Domain of Escherichia coli Trigger Factor Represents the Central Module of Its Chaperone Activity. *Journal of Biological Chemistry* **281**, 31963–31971 (2006).
10. Merz, F. *et al.* Molecular mechanism and structure of Trigger Factor bound to the translating ribosome. *EMBO Journal* **27**, 1622–1632 (2008).
11. Kaiser, C. M. *et al.* Real-time observation of trigger factor function on translating ribosomes. *Nature* **444**, 455–460 (2006).
12. Agashe, V. R. *et al.* Function of Trigger Factor and DnaK in Multidomain Protein Folding: Increase in Yield at the Expense of Folding Speed. *Cell* **117**, 199–209 (2004).
13. Teter, S. A. *et al.* Polypeptide Flux through Bacterial Hsp70: DnaK Cooperates with Trigger Factor in Chaperoning Nascent Chains. *Cell* **97**, 755–765 (1999).
14. Schaffitzel, E., Rüdiger, S., Bukau, B. & Deuerling, E. Functional dissection of trigger factor and DnaK: Interactions with nascent polypeptides and thermally denatured proteins. *Biol Chem* **382**, 1235–1243 (2001).

15. Niwa, T. *et al.* Bimodal protein solubility distribution revealed by an aggregation analysis of the entire ensemble of Escherichia coli proteins. *Proc. Nat. Acad. Sci. U.S.A.* **106**, 4201–4206 (2009).
16. Lakshmipathy, S. K., Gupta, R., Pinkert, S., Etchells, S. A. & Hartl, F. U. Versatility of trigger factor interactions with ribosome-nascent chain complexes. *Journal of Biological Chemistry* **285**, 27911–27923 (2010).
17. Oh, E. *et al.* Selective Ribosome Profiling Reveals the Cotranslational Chaperone Action of Trigger Factor In Vivo. *Cell* **147**, 1295–1308 (2011).
18. Bornemann, T., Holtkamp, W. & Wintermeyer, W. Interplay between trigger factor and other protein biogenesis factors on the ribosome. *Nat Commun* **5**, (2014).
19. Raine, A., Lovmar, M., Wikberg, J. & Ehrenberg, M. Trigger Factor Binding to Ribosomes with Nascent Peptide Chains of Varying Lengths and Sequences. *Journal of Biological Chemistry* **281**, 28033–28038 (2006).
20. Rutkowska, A. *et al.* Dynamics of Trigger Factor Interaction with Translating Ribosomes. *Journal of Biological Chemistry* **283**, 4124–4132 (2008).
21. Niwa, T. *et al.* Application of fluorescence correlation spectroscopy to investigate the dynamics of a ribosome-associated trigger factor in Escherichia coli. *Front Mol Biosci* **9**, 891128 (2022).
22. Konovalova, A., Kahne, D. E. & Silhavy, T. J. Outer Membrane Biogenesis. *Annu Rev Microbiol* **71**, 539–556 (2017).
23. Saxena, D. *et al.* Tackling the outer membrane: facilitating compound entry into Gram-negative bacterial pathogens. *npj Antimicrobials and Resistance* **1**, 1–22 (2023).
24. Tran, B. M., Punter, C. M., Linnik, D., Iyer, A. & Poolman, B. Single-protein Diffusion in the Periplasm of Escherichia coli. *J Mol Biol* **436**, 168420 (2024).
25. Nielsen, H., Tsirigos, K. D., Brunak, S. & von Heijne, G. A Brief History of Protein Sorting Prediction. *The Protein Journal 2019 38:3* **38**, 200–216 (2019).
26. Schierle, C. F. *et al.* The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the Escherichia coli periplasm via the signal recognition particle pathway. *J Bacteriol* **185**, 5706–5713 (2003).
27. Bendtsen, J. D., Nielsen, H., Widdick, D., Palmer, T. & Brunak, S. Prediction of twin-arginine signal peptides. *BMC Bioinformatics 2005 6:1* **6**, 167- (2005).
28. Valent, Q. A. *et al.* Nascent membrane and presecretory proteins synthesized in Escherichia coli associate with signal recognition particle and trigger factor. *Mol Microbiol* **25**, 53–64 (1997).
29. Gu, S. Q., Peske, F., Wieden, H. J., Rodnina, M. V. & Wintermeyer, W. The signal recognition particle binds to protein L23 at the peptide exit of the Escherichia coli ribosome. *RNA* **9**, 566–573 (2003).
30. Zopf, D., Bernstein, H. D., Johnson, A. E. & Walter, P. The methionine-rich domain of the 54 kd protein subunit of the signal recognition particle contains an RNA binding site and can be crosslinked to a signal sequence. *EMBO J* **9**, 4511–4517 (1990).

31. Angelini, S., Deitermann, S. & Koch, H. G. FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon. *EMBO Rep* **6**, 476–481 (2005).
32. Voigts-Hoffmann, F. *et al.* The Structural Basis of FtsY Recruitment and GTPase Activation by SRP RNA. *Mol Cell* **52**, 643–654 (2013).
33. Egea, P. F. & Stroud, R. M. Lateral opening of a translocon upon entry of protein suggests the mechanism of insertion into membranes. *Proc. Nat. Acad. Sci. U.S.A.* **107**, 17182–17187 (2010).
34. Knoblauch, N. T. M. *et al.* Substrate Specificity of the SecB Chaperone. *Journal of Biological Chemistry* **274**, 34219–34225 (1999).
35. Komarudin, A. G. & Driessen, A. J. M. SecA-Mediated Protein Translocation through the SecYEG Channel. *Microbiol Spectr* **7**, (2019).
36. Zimmer, J., Nam, Y. & Rapoport, T. A. Structure of a complex of the ATPase SecA and the protein-translocation channel. *Nature* **455**, 936–943 (2008).
37. Wang, L., Miller, A. & Kendall, D. A. Signal peptide determinants of SecA binding and stimulation of ATPase activity. *Journal of Biological Chemistry* **275**, 10154–10159 (2000).
38. Fekkes, P., Van der Does, C. & Driessen, A. J. M. The molecular chaperone SecB is released from the carboxy-terminus of SecA during initiation of precursor protein translocation. *EMBO J* **16**, 6105–6113 (1997).
39. Auclair, S. M., Bhanu, M. K. & Kendall, D. A. Signal peptidase I: Cleaving the way to mature proteins. *Protein Science* **21**, 13–25 (2012).
40. De Geyter, J. *et al.* Trigger factor is a bona fide secretory pathway chaperone that interacts with SecB and the translocase. *EMBO Rep* **21**, (2020).
41. Ullers, R. S., Ang, D., Schwager, F., Georgopoulos, C. & Genevaux, P. Trigger Factor can antagonize both SecB and DnaK/DnaJ chaperone functions in Escherichia coli. *Proc. Nat. Acad. Sci. U.S.A.* **104**, 3101-3106 (2007).
42. Patzelt, H. *et al.* Binding specificity of Escherichia coli trigger factor. *Proc. Nat. Acad. Sci. U.S.A.* **98**, 14244–14249 (2001).
43. Eismann, L. *et al.* Selective ribosome profiling reveals a role for SecB in the co-translational inner membrane protein biogenesis. *Cell Rep* **41**, 111776 (2022).
44. Valent, Q. A. *et al.* The Escherichia coli SRP and SecB targeting pathways converge at the translocon. *EMBO J* **17**, 2504–2512 (1998).
45. Zhu, Z., Wang, S. & Shan, S. ou. Ribosome profiling reveals multiple roles of SecA in cotranslational protein export. *Nat Commun* **13**, (2022).
46. Singh, R. *et al.* Cryo-electron microscopic structure of SecA protein bound to the 70S ribosome. *Journal of Biological Chemistry* **289**, 7190–7199 (2014).
47. Tomkiewicz, D., Nouwen, N., Van Leeuwen, R., Tans, S. & Driessen, A. J. M. SecA Supports a Constant Rate of Preprotein Translocation. *Journal of Biological Chemistry* **281**, 15709–15713 (2006).
48. Fessl, T. *et al.* Dynamic action of the sec machinery during initiation, protein translocation and termination. *Elife* **7**, (2018).

49. Kanaoka, Y. *et al.* AFM observation of protein translocation mediated by one unit of SecYEG-SecA complex. *Nature Communications* **16**, 225 (2025).
50. Volkov, I. L. *et al.* Spatiotemporal kinetics of the SRP pathway in live E. coli cells. *Proc. Nat. Acad. Sci. U.S.A.* **119** (2022).
51. Vogt, J. & Schulz, G. E. The structure of the outer membrane protein OmpX from Escherichia coli reveals possible mechanisms of virulence. *Structure* **7**, 1301–1309 (1999).
52. Lauber, F., Deme, J. C., Lea, S. M. & Berks, B. C. Type 9 secretion system structures reveal a new protein transport mechanism. *Nature* **564**, 77–82 (2018).
53. Devlin, T. & Fleming, K. G. A team of chaperones play to win in the bacterial periplasm. *Trends Biochem Sci* **49**, 667–680 (2024).
54. Surrey, T. & Jahnig, F. Refolding and oriented insertion of a membrane protein into a lipid bilayer. *Biophysics (Oxf)* **89**, 7457–7461 (1992).
55. Burgess, N. K., Dao, T. P., Stanley, A. M. & Fleming, K. G. β-Barrel proteins that reside in the Escherichia coli outer membrane in vivo demonstrate varied folding behavior in vitro. *Journal of Biological Chemistry* **283**, 26748–26758 (2008).
56. Kang, G. *et al.* Förster resonance energy transfer as a probe of membrane protein folding. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1818**, 154–161 (2012).
57. Bulieris, P. V., Behrens, S., Holst, O. & Kleinschmidt, J. H. Folding and insertion of the outer membrane protein OmpA is assisted by the chaperone Skp and by lipopolysaccharide. *Journal of Biological Chemistry* **278**, 9092–9099 (2003).
58. Bergman, W. N. & Sousa, M. C. A fluorescent reporter and single-turnover kinetics reveal new insight into BAM complex function. *bioRxiv* doi 2025.05.30.657120 (2025).
59. Wang, X., Peterson, J. H. & Bernstein, H. D. Bacterial Outer Membrane Proteins Are Targeted to the Bam Complex by Two Parallel Mechanisms. *mBio* **12**, (2021).
60. Kipper, K. *et al.* Application of Noncanonical Amino Acids for Protein Labeling in a Genomically Recoded Escherichia coli. *ACS Synth Biol* **6**, 233–255 (2016).
61. Saurí, A. *et al.* Autotransporter β-Domains Have a Specific Function in Protein Secretion beyond Outer-Membrane Targeting. *J Mol Biol* **412**, 553–567 (2011).
62. Wang, X. & Bernstein, H. D. The Escherichia coli outer membrane protein OmpA acquires secondary structure prior to its integration into the membrane. *Journal of Biological Chemistry* **298**, 101802 (2022).
63. Hussain, S. & Bernstein, H. D. The Bam complex catalyzes efficient insertion of bacterial outer membrane proteins into membrane vesicles of variable lipid composition. *Journal of Biological Chemistry* **293**, 2959–2973 (2018).
64. Chng, S. S. *et al.* Disulfide Rearrangement Triggered by Translocon Assembly Controls Lipopolysaccharide Export. *Science* **337**, 1665 (2012).
65. Weatherill, E. E. *et al.* Fast slow folding of an outer membrane porin. *Proc. Nat. Acad. Sci. U.S.A.* **119** (2022).

66. Elf, J. & Barkefors, I. Single-Molecule Kinetics in Living Cells. *Annu Rev Biochem* **88**, 635–659 (2019).
67. Abramson, J. *et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* **630**, 493–500 (2024).
68. Los, G. V. *et al.* HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem Biol* **3**, 373–382 (2008).
69. Ernst, M., Mahoney-Kruszka, R., Zelt, N. B. & Robertson, J. L. A roadmap to cysteine specific labeling of membrane proteins for single-molecule photobleaching studies. *Methods* **234**, 21 (2024).
70. Crawford, R. *et al.* Long-lived intracellular single-molecule fluorescence using electroporated molecules. *Biophys J* **105**, 2439–2450 (2013).
71. Sustarsic, M. *et al.* Optimized delivery of fluorescently labeled proteins in live bacteria using electroporation. *Histochem Cell Biol* **142**, 113 (2014).
72. Volkov, I. L. *et al.* tRNA tracking for direct measurements of protein synthesis kinetics in live cells. *Nature Chemical Biology* **14**, 618–626 (2018).
73. Liu, Z., Lavis, L. D. & Betzig, E. Imaging Live-Cell Dynamics and Structure at the Single-Molecule Level. *Mol Cell* **58**, 644–659 (2015).
74. Reymond, L., Huser, T., Ruprecht, V. & Wieser, S. Modulation-enhanced localization microscopy. *Journal of Physics: Photonics* **2**, 041001 (2020).
75. Balzarotti, F. *et al.* Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science (1979)* **355**, 606–612 (2017).
76. Ranefall, P., Sadanandan, S. K. & Wahlby, C. Fast adaptive local thresholding based on ellipse fit. *Proceedings - International Symposium on Biomedical Imaging* **2016-June**, 205–208 (2016).
77. Loy, G. & Zelinsky, A. Fast radial symmetry for detecting points of interest. *IEEE Trans Pattern Anal Mach Intell* **25**, 959–973 (2003).
78. Lindén, M., Ćurić, V., Amselem, E. & Elf, J. Pointwise error estimates in localization microscopy. *Nature Communications 2017 8:1* **8**, 1–9 (2017).
79. Kuhn, T., Hettich, J., Davtyan, R. & Gebhardt, J. C. M. Single molecule tracking and analysis framework including theory-predicted parameter settings. *Scientific Reports* **11**, 9465- (2021).
80. Jaqaman, K. *et al.* Robust single-particle tracking in live-cell time-lapse sequences. *Nat Methods* **5**, 695–702 (2008).
81. Persson, F., Lindén, M., Unoson, C. & Elf, J. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nat Methods* **10**, 265–269 (2013).
82. Metelev, M. *et al.* Direct measurements of mRNA translation kinetics in living cells. *Nat Commun* **13**, 1852 (2022).
83. Ha, T. *et al.* Fluorescence resonance energy transfer at the single-molecule level. *Nature Reviews Methods Primers* **4**, 21 (2024).
84. Vogel, S. S., Van der Meer, B. W. & Blank, P. S. Estimating the distance separating fluorescent protein FRET pairs. *Methods* **66**, 131–138 (2014).
85. Iqbal, A. *et al.* Orientation dependence in fluorescent energy transfer between Cy3 and Cy5 terminally attached to double-stranded nucleic acids. *Proc. Nat. Acad. Sci. U.S.A.* **105**, 11176–11181 (2008).

86. Hei, Y. *et al.* Photostability of organic fluorophore influenced by adjacent amino acid residues. *Communications Chemistry* **8**, 1–8 (2025).
87. Vaiana, A. C. *et al.* Fluorescence Quenching of Dyes by Tryptophan: Interactions at Atomic Detail from Combination of Experiment and Computer Simulation. *J Am Chem Soc* **125**, 14564–14572 (2003).
88. Berglund, A. J. Statistics of camera-based single-particle tracking. *Phys Rev E Stat Nonlin Soft Matter Phys* **82**, 011917 (2010).
89. Michalet, X. & Berglund, A. J. Optimal diffusion coefficient estimation in single-particle tracking. *Phys Rev E Stat Nonlin Soft Matter Phys* **85**, 061916 (2012).
90. Shuang, B. *et al.* Improved analysis for determining diffusion coefficients from short, single-molecule trajectories with photoblinking. *Langmuir* **29**, 228–234 (2013).
91. Schibich, D., Gloge, F., Pöhner, I. *et al.* Global profiling of SRP interaction with nascent polypeptides. *Nature* **536**, 219–223 (2016).
92. Ullers, R. S. *et al.* Interplay of signal recognition particle and trigger factor at L23 near the nascent chain exit site on the Escherichia coli ribosome. *Journal of Cell Biology* **161**, 679–684 (2003).
93. Ullers, R. S. *et al.* Sequence-specific Interactions of Nascent Escherichia coli Polypeptides with Trigger Factor and Signal Recognition Particle. *Journal of Biological Chemistry* **281**, 13999–14005 (2006).
94. Raine, A., Ivanova, N., Wikberg, J. E. S. & Ehrenberg, M. Simultaneous binding of trigger factor and signal recognition particle to the E. coli ribosome. *Biochimie* **86**, 495–500 (2004).
95. Buskiewicz, I. *et al.* Trigger factor binds to ribosome-signal-recognition particle (SRP) complexes and is excluded by binding of the SRP receptor. *Proc. Nat. Acad. Sci. U.S.A.* **101**, 7902–7906 (2004).
96. Ariosa, A., Lee, J. H., Wang, S., Saraogi, I. & Shan, S. O. Regulation by a chaperone improves substrate selectivity during cotranslational protein targeting. *Proc. Nat. Acad. Sci. U.S.A.* **112**, 3169–3178 (2015).
97. Giglione, C., Fieulaine, S. & Meinnel, T. N-terminal protein modifications: Bringing back into play the ribosome. *Biochimie* **114**, 134–146 (2015).
98. Solbiati, J., Chapman-Smith, A., Miller, J. L., Miller, C. G. & Cronan, J. E. Processing of the N termini of nascent polypeptide chains requires deformylation prior to methionine removal. *J Mol Biol* **290**, 607–614 (1999).
99. Yang, C. I., Hsieh, H. H. & Shan, S. O. Timing and specificity of cotranslational nascent protein modification in bacteria. *Proc. Nat. Acad. Sci. U.S.A.* **116**, 23050–23060 (2019).
100. Bienvenut, W. V., Giglione, C. & Meinnel, T. Proteome-wide analysis of the amino terminal status of Escherichia coli proteins at the steady-state and upon deformylation inhibition. *Proteomics* **15**, 2503–2518 (2015).
101. Bhakta, S., Akbar, S. & Sengupta, J. Cryo-EM Structures Reveal Relocalization of MetAP in the Presence of Other Protein Biogenesis Factors at the Ribosomal Tunnel Exit. *J Mol Biol* **431**, 1426–1439 (2019).
102. Ranjan, A., Mercier, E., Bhatt, A. & Wintermeyer, W. Signal recognition particle prevents N-terminal processing of bacterial membrane proteins. *Nature Communications* **8**, 1–6 (2017).
103. Takakura, H. *et al.* Long Time-lapse Nanoscopy with Spontaneously Blinking Membrane Probes. *Nat Biotechnol* **35**, 773 (2017).

104. Uno, S. N. *et al.* A spontaneously blinking fluorophore based on intramolecular spirocyclization for live-cell super-resolution imaging. *Nature Chemistry* **6**, 681–689 (2014).
105. Knoke, L. R. *et al.* The role of glutathione in periplasmic redox homeostasis and oxidative protein folding in Escherichia coli. *Redox Biol* **64**, 102800 (2023).
106. Wilks, J. C. & Slonczewski, J. L. pH of the cytoplasm and periplasm of Escherichia coli: Rapid measurement by green fluorescent protein fluorimetry. *J Bacteriol* **189**, 5601–5607 (2007).
107. Byrdin, M. & Byrdina, S. Impact of triplet state population on GFP-type fluorescence and photobleaching. *Biol Cell* **117** (2025).

Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 2623

Editor: The Dean of the Faculty of Science and Technology

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

Distribution: publications.uu.se
urn:nbn:se:uu:diva-572623

ACTA UNIVERSITATIS
UPSALIENSIS
2026