Perspectives and Conclusion
6.1. Introduction

This chapter summarizes the work described in this thesis on TransLesion Synthesis DNA polymerases (TLS polymerases) in *Escherichia coli* using single-molecule fluorescence microscopy. Single-molecule imaging has great potential in revealing the molecular details of fundamental cellular processes by direct visualization of individual molecules in live cells. Even though it represents a powerful technique for quantitative and qualitative measurements, single-molecule imaging methods still require specific expertise and solutions tailored to the problem. This thesis reports on technological advances enabling the visualization of individual proteins in live bacteria and describes new mechanistic insights into the regulation of DNA polymerase activity.

On the methodological side, we have developed a software framework for data management, exploration, and archiving that greatly speeds up the handling and analysis of the large amounts of imaging data that single-molecule approaches generate. Further, we describe an optimized microfluidic flow-cell device for time-lapse imaging that allows for long-term visualization of single molecules in live *E. coli* cells. On the mechanistic side, we investigate the localization and regulation of DNA polymerase V (Pol V) and the bacterial replisome to test if current models for Pol V regulation and recruitment to the replisome hold true in live cells.

In this concluding chapter, I will briefly summarize and discuss the main findings and the technical contributions described in this thesis. I will discuss the relevance of our methodological work to the field of single-molecule microscopy and the implications of our results on the understanding on TLS polymerases regulation and polymerase switching mechanisms.

6.2. Technological contributions and perspectives

Single-molecule imaging data sets are often very large and thus place significant requirements on data storage and analysis. Faster cameras with larger pixel numbers, fully automated microscope stage scanners, and the ability to collect multiparameter images (such as wavelength, lifetime, polarization) are just a few factors that contribute to increasing sizes of data sets. Triggered by these developments and the need for generalizable approaches, we developed new methodology for data management and image analysis, ranging from an optimized microfluidic flow-cell design to a series of software solutions.

6.2.1. A need for flexible analysis platforms

Chapter 2 of this thesis reports on our efforts to integrate the myriad of algorithms for image analysis that are available in ImageJ [1] and Fiji [2] to work with the highly hierarchical datasets typical of live-cell fluorescence imaging studies. ISBatch – in Singulo-Biology Batch tools – streamlines multiple image processing steps. In a field where data exploration and analysis are intricately intertwined, automation has the potential to significantly lower the burden of writing customized code for
every single task.

Cellular segmentation is a fundamental problem in image analysis. It consists of detecting and annotating cellular outlines for further analysis. Single-molecule imaging data sets possess a highly hierarchical structure with multiple levels of data (experiments, fields of view, individual cells, individual molecules) that each place different requirements on the exact method of analysis. Just identifying a cell outline in the absence of additional contextual information is not sufficient.

Unfortunately, despite enormous efforts, segmentation algorithms are still not sufficiently mature to support general use. Existing algorithms are tailored for very specific situations related to, for example, a particular type of cell or imaging parameters. Further, those algorithms are often too complex for the average user and not easily expandable or integrated with other platforms.

A fundamental challenge is to combine what is available for cellular segmentation and localization with the vast selection of existing single-molecule detection and analysis algorithms. Spatial coordinates of each molecule have to be translated to cellular coordinates so we can obtain quantitative information that is relevant in a cellular context. For example, such analysis is critical when imaging replisome localization throughout the cell cycle or identifying the dynamics of proteins in the membrane and cytosol.

Considering the large dependence of data-analysis methods on fast and efficient algorithms and the increasing complexity of analysis pipelines, open-source development of scientific algorithms is playing an increasingly important role in science. To maximally streamline this process, all code used in scientific publications should be made publicly available upon publication. References in text to ‘in-house algorithms’ are frequently found but should be avoided at all costs, as they do not aim for reproducibility and may hide unintentional errors. When developing analysis algorithms, the only certainty is that the code changes frequently. It is unrealistic to demand perfectly designed software – in the sense of satisfactory user experience – for scientific projects, but transparency should be the goal. Sharing code and allowing others to optimize algorithms seems to be the most efficient way towards establishing software standards in the sciences that most optimally support reproducibility and progress.

6.2.2. Imaging live cells with fluorescence

In chapter 3, we reported on the artefactual localization of red-fluorescent proteins (RFPs) in *E. coli*. Our data indicate that RFPs tend to bind to the cellular membrane at the single-molecule level in the cell. Interestingly, this mislocalization only seemed to occur at very low intracellular concentrations. This intriguing behaviour is still not yet well understood, but reveals potential biases when using RFPs. The usual controls that monitor the pair protein-fluorescent tag at high intracellular concentrations (under conditions of over expression), would be inadequate in detecting such an artefact. Instead, controls of fluorescent proteins should be performed over a wide range of concentrations. Further, our work shows that control experiments should always be done on both the labelled protein of interest and the isolated flu-
oescent proteins. After all, depending on the affinities of each protein to a cellular structure, protein chimeras may play a "tug of war", where the fluorescent protein could enhance a cytosolic or membrane localization, or even bind to other cellular structures. The artefactual mislocalization then becomes a complicated net effect of the affinities of both protein of interest and fluorescent label to a cellular structure.

The discovery and development of improved fluorescent tags is a key factor in our ability to characterize protein dynamics and localization. The ideal fluorophore should be inert in that it has no influence on the behaviour of the protein of interest. However, it is unlikely that such Holy Grail exists. The intimate relationship between structure and function are at the core of molecular biology and any modification to a protein has the potential to affect the observed outcome. Because of the intricate and highly complex nature of the networks that proteins are involved in, "better" and "worse" are highly speculative adjectives for describing protein activity. Protein tags are likely to interfere with the "wild-type" behaviour and protein characterization will remain a challenge. In vitro experiments will not be able to reproduce with fidelity all the nuances of a complex living system and current in vivo experiments are biased by the tags used. Therefore, both approaches have to coexist wherever possible. The use of plasmid-based protein expression systems allows us to conveniently test for different concentrations as we are able to tune protein expression at will and monitor the phenotypes. On the other hand, modifications performed directly on the genome, such as presented in this thesis, allow us to directly measure the cellular response to a given stimuli. A combination of these various approaches and methods is essential to enable improved understanding of protein activity inside living systems.

Besides the technical challenges involved with imaging fluorescent proteins in bacteria, the results presented in chapters 4 and 5 critically rely on new chemical surface treatment strategies we developed for our microscope glass slides. Specifically in chapter 4, we describe the use of 3-aminopropyltriethoxysilane (APTES) as the chemical of choice for cellular attachment and growth on a glass support. Compared to the previously frequently used poly-L-lysine [3], APTES-treated surfaces have little effect on the cells, allowing hours-long time-lapse experiments with cells growing with division times as short as 30-40 minutes. Poly-L-lysine, on the other hand, is toxic and likely to kill the cells throughout the experiment.

The methodological improvements required to reach single-molecule sensitivity and resolution in live cells are far from straightforward. In the course of my PhD work, we overcame several barriers and established a robust set of designs and protocols that enabled us to work towards our main goal: study of the dynamics of DNA polymerases in live E. coli cells.

6.2.3. Contributions and perspectives on TLS polymerases

As is laid out in the introductory Chapter 1 and the experimental Chapters 4 and 5, our results affect current models of TLS polymerase regulation and activation in E. coli cells. Chapters 4 and 5 report new mechanistic perspectives on TLS polymerases accessing DNA, especially in relation to DNA polymerase V. By com-
bining λ-red recombination-mediated genetic engineering [4] and single-molecule microscopy, we are able to directly visualize individual molecules in their natural environment and under control of their native promoters. This thesis describes the first single-molecule studies of translesion DNA polymerases in live E. coli cells. The main results and conclusions of this work will be briefly recapitulated here, setting the stage for a broader discussion and conclusions.

6.2.4. Pol V associates with the cellular membrane

One of the most intriguing observations we made is the association of DNA polymerase V with the inner cellular membrane directly after its upregulation in response to DNA damage. What is a DNA polymerase doing at the cellular membrane?

Due to the significantly higher rate of insertion errors in Pol V (thus leading to increased mutagenesis), the cell delays enzyme activity until all the necessary components are fully available (e.g. RecA*, UmuD’ and damage sites). After the cells being UV-damaged, replication in E. coli cell stops for 15 minutes before restart [5]. During this delay, other mechanisms, such as NER, act on damage sites and several mechanisms act to restart replication, such as DnaC-dependent restart, lesion-skipping. Pol IV, which is present at reasonably high levels during the early stages of the DNA damage response, and is capable of carrying out TLS at stalled replication forks [6]. In vitro experiments have shown that in the context of replisomes stalled at CPD lesions, Pol IV is able to bypass CPD lesions in an error-free manner, whereas Pol II cannot [6].

Our experiments have shown that, independently of the cellular concentration, dissociation of Pol V from the membrane depends on RecA nucleoprotein filaments (RecA*) and on UmuD → UmuD’ conversion. Our experiments were not able to identify whether the UmuD is also bound to the membrane as part of Pol V. Such a binding-release mechanism would suggest that inactive Pol V is kept at the cellular membrane while active Pol V (Pol V Mut) is cytosolic. Because Pol V upregulation follows after 45 minutes of damage, taken together, this also suggest that Pol V is activated to Pol V Mut after the replisomes are resumed (15-20 min) and the substrate for Pol V Mut might lie the products many different restart pathways, including lesion-skipping mechanisms. We considered mass-action models for polymerase switching at the replisome by visualizing the behaviour of Pol V in UV-damaged cells in time-lapse experiments – and repeated these in in ΔpolB, ΔdinB and ΔpolB ΔdinB backgrounds. These experiments showed that Pol V localization is independent on the presence of Pol II and Pol IV. Earlier reports proposed an association between UmuD and Pol IV [7], which could act as an additional limiting step in Pol V activation. If true, Pol IV deletion would dramatically increase the availability of UmuD for conversion to UmuD’, which stimulates Pol V release from the membrane. These observations strongly suggest that Pol V localization is decoupled from the other TLS polymerases.
6.2.5. Pol V acts away from the stalled replisome

Visualizing Pol V membrane binding upon activation by RecA* and release after UmuD → UmuD', another intriguing phenomenon becomes apparent: Pol V does not colocalize with stalled replisomes. RecA* appears to act as a ‘traffic cop’ directing Pol V activity. If not at the replisome, where is Pol V acting and what are the implications to current models of polymerase switching?

Long-held models for polymerase switching and translesion DNA synthesis assume a highly concerted mechanism. Firstly, replisomes stall at a damage site and RecA* filaments are formed on the accumulated ssDNA, triggering the SOS response. Next, activated Pol V displaces Pol III via mass-action mechanisms and synthesises over the lesion. The replisome then restarts and resumes replication [6, 8–12]. Many recent experiments, including the ones reported in this thesis, suggest that it is now time to revise this view.

In vitro experiments show that replisomes are inherently damage tolerant and upon encountering CPD lesions stall briefly before continuing, leaving ssDNA gaps behind [13]. When stalling, DnaB helicase continues unwinding dsDNA and allows Pol III to reengage with a newly synthesized primer on the other side of the lesion [13]. This result supports the notion that leading-strand synthesis can jump over a lesion and leave ssDNA gaps that could be target for TLS Pols. These ssDNA gaps left behind by replication forks may be bound by RecA supporting a RecA*-dependent activation of Pol V far from the site of DNA replication, termed activation in trans [14–16]. Our data show that Pol V associates with the DNA in regions away from the replisome, indicating that TLS may occur in these DNA gaps.

Interestingly, other studies have indicated that TLS polymerases could act at stalled transcription bubbles, carrying out transcription-coupled TLS. In E. coli, it has been shown that the transcription elongation factor NusA may recruit TLS DNA polymerases (in particular Pol IV) to gaps encountered during translesion [17]. This result suggests that stalled RNAPs can actually recruit TLS polymerases to the transcribed strand opposed to the lesion [18]. Interestingly, DNA lesions caused by UV were insufficient to trigger this similar response, but the application N2-dG did initiate this response. N2-dG is described as a preferable lesion for Pol IV [18]. Considering the lack of phenotype when deleting Pols II and Pol IV, our data supports that Pol V may not participate in this mechanism and that the choice of polymerase could the highly related to the substrate created by the damaging agent.

6.2.6. Pol V does not compete with Pol II and Pol IV

We monitored colocalization of Pol V foci (Pol V Mut) with the bacterial replisome inside living cells in different genetic backgrounds. Our results suggest that mass action, and therefore the presence of other polymerases occupying the replisome, plays little or no role in the regulation of Pol V mut activity and spatial distribution. Strains containing constitutive RecA* activity (RecA(E38K)), showed Pol V localized at the cytosol independently of the presence of other TLS polymerases. The access of Pol V to the DNA was not affected by the presence of Pol II and Pol IV in the cell, either when switching with Pol III or when accessing DNA in regions distinct from...
the replisome. We were not able to observe mass-action effects in rapidly growing cells – irrespective of whether UV-induced DNA damage was present.

Our results indicate that any mass-action mechanisms for polymerase exchange are overwhelmed by RecA* acting as a ‘traffic cop’. Also, Pols II and IV have little or no influence on Pol V activity on the DNA in RecA(E38K) cells. This finding together with the tight control on Pol V activity suggests that mechanisms exist to keep Pol V away from the DNA. Because all the substrates to activate Pol V are frequently available and Pol V bypasses competition when active, several layers of regulation could work as a mechanism to avoid excessive mutagenicity, thus, favouring genomic stability. Pol V could then act as a last resort for the cell to rapidly generate genetic variability to overcome stress conditions.

6.2.7. The role of Pol V in mutagenesis and implications of TLS in antibiotic resistance

Many mechanistic insights on the role of TLS polymerase come from work on Pol IV. In relation to polymerase switching, previous work has described active exchange between Pol IV and Pol III [19, 20]. Recently, it was reported that high concentration of Pol IV could block replisomes catalysing leading-strand synthesis. Contrary to what we observed for Pol V, the exchange between Pol IV – Pol III is dependent on stalled Pol III [21]. TLS polymerases, in general, may also switch with non-stalled replisomes.

The SOS response is reported to control adaptive mutation in non-growing cells [17, 22, 23]. In non-growing cells, where DNA replication is slow or inactive, RNA polymerase molecules would frequently encounter lesions before DNA replication forks. A recent study in bacteria provided evidence for a transcription-driven adaptive mutation mechanism operating under replication-inhibited conditions, termed retromutagenesis [23]. Here, lesions encountered by RNA polymerase cause the enzyme to produce mutant mRNA transcripts and thus mutant proteins (transcriptional mutagenesis). Because DNA replication is inactive the presence of the lesion does not lead immediately to a mutation fixing on the chromosome. If however, a transcriptional mutation produces a variant protein that restores DNA replication, for example by circumventing the action of a replication-inhibiting drug, replication will be restored and proceed over the lesion and produce a mutation. In this context, the links drawn between Pol IV and RNA polymerase become extremely interesting. The ability to trial mutations though transcriptional mutagenesis, then subsequently only lock in those that restore DNA replication (possibly through TLS polymerase action) could provide a potent means for adaptive mutation.

It is well demonstrated that the SOS response influences the development of antibiotic resistance in bacteria. Removing recA or making LexA non-cleavable abolishes antibiotic-induced mutagenic effects, especially in sub-lethal doses of antibiotics [24, 25]. Ampicillin, an antibiotic that inhibits cell wall formation, also has been shown to induce SOS response via RecA activation [26]. Complementing those results, specialized TLS polymerases impacts cell fitness when in stationary phase [27] and, their removal lowers mutation rates and slows acquisition of an-
tibiotic resistance [24, 28]. There is clearly a link between TLS polymerases and the development of antibiotic resistance. However the question still remains: do TLS polymerases provide all cells in a population with higher tolerance to antibiotics and thus increase the chance that a sub-population will become resistant, or do they speed the development of resistance by participating in adaptive mutation pathways?

We are currently losing the race against bacterial infections. Bridging gaps in our understanding of bacterial evolution and how bacteria respond to environmental stresses will help in finding better solutions for the post-antibiotic era. Developing new drugs is paramount, especially since currently no drug exists that can inhibit TLS polymerases. New exploratory technologies such as metagenomics or *in silico* design could provide parts of the answer. Antibiotics targeting bacterial growth or essential metabolic functions are not enough. Instead, we need to target bacterial evolution in a combined approach if we want to avoid development of resistance to antibiotics.

### 6.3. Conclusion

Studying polymerases in live bacterial cells is laborious and challenging. The lack of clear phenotypes in some cases (e.g. when performing DNA-damage experiments) makes it barely possible to establish relationships between damaging agents and phenotypes.

The experiments reported in this thesis represent a foundation for future work on DNA polymerases in live bacteria, especially – but not limited to – *E. coli*. The protocols and devices presented in this thesis provide a toolbox for further exploration of metabolic pathways. Pushing the boundaries of this technology will soon allow us to reliably image polymerase binding properties *in vivo*, and expand these studies to tagged Pol II and Pol IV. In combination with computational approaches that will allow us to deal with the experimental data and that will enable the modelling of intracellular molecular behaviour, we will be well placed to test current models and gain great mechanistic insight in the processes that are responsible for genomic maintenance.
References


