Dynamics of DNA repair enzymes
And competence proteins
In *Bacillus subtilis*

Inauguraldissertation zur Erlangung der Doktorenwürde

des Fachbereichs Chemie

der Philipps-Universität Marburg

Vorgelegt von Stephan Altenburger, Diplom-Biologe,

geboren in Freiburg im Breisgau

1. GutachterBetreuer: Prof. Dr. Peter L. Graumann
   2. 2. Gutachter: Dr. Gert bange

Einreichungsdatum: 04.02.2016
Prüfungsdatum: 22.03.2016

Marburg, Februar 2016
UMR = 1180

Lebenslauf:

■ Persönliche Daten

Name: Stephan Altenburger

Geburtsdatum/-ort: 02.05.1978, Freiburg im Breisgau

Familienstand: ledig

Staatsangehörigkeit: deutsch

■ Wissenschaftliche Tätigkeit

04/2013-03/2016	Akademischer wissenschaftlicher Angestellter (Doktorand) am der LOEWE-Zentrum für Synthetische Mikrobiologie SYNMIKRO (Uni Marburg)
04/2011-03/2013	Akademischer wissenschaftlicher Angestellter (Doktorand) am Institut für Biologie II (FB Mikrobiologie) der Uni Freiburg

■ Studium

03/2010-03/2011	Diplomarbeit in der Abteilung Mikrobiologie an der Albert-Ludwigs Universität Freiburg
10/2002-03/2011	Studium der Biologie an der Universität Freiburg Schwerpunkte: Genetik, Zellbiologie, Ökologie, biologische Anthropologie

- **Publikationen:**

  - Altenburger S, Pediaditakis M, Fries SJ, El Najjar N, Schmitt B, Dahlke S, Rösch T, Fritz G, Graumann PL.

    „*Single-molecule tracking of recombination proteins RecN, RecO and RecJ reveals a DNA scanning mode of movement and the establishment of DNA repair centres at many sites within the nucleoid.*" **in prep.**

  - Gillon L, Altenburger S, Graumann PL, Schalk IJ.

    "*Deciphering Protein Dynamics of the Siderophore Pyoverdine Pathway in Pseudomonas aeruginosa*" PLoS One., **8(10)**: e79111. Published online, 2013

  - Guillon L, El Mecherki M, Altenburger S, Graumann PL, Schalk IJ.

    „*High cellular organization of pyoverdine biosynthesis in Pseudomonas aeruginosa: clustering of PvdA at the old cell pole.*" Environ Microbiol, **14(8)**: 1982-94. doi: 10.11112012, 2012

Zusammenfassung

DNA Doppelstrangbrüche (DSBs) sind eine schwere Bedrohung der Integrität des Genoms, deshalb gibt es seine Vielzahl von Proteinen, die solche Schäden reparieren. Der in Bakterien wichtigste Signalweg ist die homologe Rekombination (HR), mit der ATPase RecA als zentrales Enzym. RecA bildet dabei Filamente, die die Distanz zwischen den beiden Schwesterchromatiden überbrückt. Aber bereits vor RecA gibt es eine Vielzahl von Proteinen (RecNJORFX), die aktiv bei HR werden. In *Bacillus subtilis* startet dieser Prozess damit, dass RecN Fokusse bildet, 15 Minuten nachdem DSBs induziert werden. Beendet wird der Vorgang 3 Stunden später durch den Abbau der RecA-Filamente und der Wiederaufnahme des Wachstums. Ich wollte einen detailreicheren Einblick in die Dynamik der beteiligten Proteine haben und habe deshalb die Methode der Einzelmolekülmikroskopie (*single molecule microscopy*) in lebenden Zellen angewandt. Dabei habe ich Videos mit 40 ms Belichtungszeit erstellt, die Bewegung der Rec-Proteine gemessen und die resultierenden Trajektorien mathematisch analysiert. In exponentiell wachenden Zellen konnte ich beobachten, dass RecN, RecO und viele RecJ-Moleküle kontinuierlich das Chromosom abrastern, was ein Model für die Einzelmoleküle der distributiven Suche unterstützt. Im Gegensatz zu RecN und RecO verbleibt ein Anteil der RecJ-Moleküle an der Replikationmaschine. Sobald DSBs induziert werden, verharren RecNOJ an mehreren Stellen auf dem Nukleoid. RecN bildet keine statischen Reparaturzentren, wie man in Eukaryoten beobachten konnte, sondern kurzlebige (~2,5 s) Cluster die als Rekrutierungsplattform für Reparaturenzym dienen. So wird die lokale Konzentration von Rec-Proteinen erhöht und das Einfangen von Interaktionspartnern aus einem diffusen Reservoir ermöglicht. Der Großteil der RecNJO-Moleküle sucht, selbst in Gegenwart von DSBs, weiterhin das Chromosom nach Schäden bzw. Interaktionspartnern ab. *In toto* zeigt meine Arbeit, dass die initiale Detektion von DSBs, das Prozessieren der freien DNA-Enden und das Beladen der hergestellten ssDNA mit RecA in sehr kurzen Zeiträumen abläuft und nur von einer Minderheit der Proteinpopulation bewerkstelligt wird.

Synopsis

DNA double strand breaks (DSBs) are a severe threat to genome integrity and thus a variety of proteins are dedicated to repair such threats. The major repair route in bacteria is that of homologous recombination (HR), with the ATPase RecA as a key player. In HR, a broken DNA strand is repaired using a second intact DNA copy present on a homologous chromosome. This process involves the exchange of DNA strands, mediated by RecA, which forms filamentous polymers on ssDNA and initiates strand exchange. Prior to RecA, a plethora of Rec-proteins (Rec NJORFX) act to initiate HR. In *Bacillus subtilis* this process starts with RecN forming foci 15 minutes after DSB induction and is finished when RecA filaments disassemble and cell growth resumes after three hours. I wished to obtain a more detailed view on the dynamics of these proteins, and therefore employed single molecule fluorescence microscopy in live cells. Using 40 ms stream acquisition, I detected the movement of single Rec proteins and analyzed these trajectories mathematically. In exponentially growing cells I observed that RecN, RecO and, partially, RecJ continuously scan the nucleoid, supporting a distributive search model of individual molecules. In contrast to RecN and RecO, a fraction of the exonuclease RecJ is retained at the replication machinery. Upon induction of DSBs, RecNJO arrest at several sites on the nucleoid. RecN does not form static repair centers as proposed for eukaryotes, but short-lived (~2.5 s) clusters that act as repair enzyme recruitment platforms. Thus the local concentration of Rec-proteins increases to trap interaction partners out of a pool of diffusive enzymes. A majority of the RecNJO molecules keep on scanning for lesions or interaction partners, even in the presence of DSBs. *In toto*, my work indicates that the initial detection of a DSB, processing of free DNA ends, and loading of RecA on the generated ssDNA site takes place in a very short time frame, performed by a minority of the protein population.

Table of contents

1. Introduction

1.1. *Bacillus subtilis*

B. subtilis is a Gram-positive, rod-shaped and flagellated soil bacterium. It is an important member of the community of microorganisms in the rhizosphere [1], mobilizing high-molecular compounds and forming symbiotic biofilms with *viridiplantae*, the terrestrial primary producers [2]. *B. subtilis* belongs to the phylum *firmicutes* that are part of *eubacteria* domain. It is a well understood Gram-positive model organism and serves investigations of molecular and cell biology, e.g. replication, DNA-repair or gene expression and regulation [3, 4, 5]. *B. subtilis* cells grow mostly filamentous in exponential growth phase and become unicellular and highly motile in stationary phase. At the transition from exponential to stationary phase, when resources become scarce, subpopulations in a *B. subtilis* culture can develop sets of different genetic programs or cell fates [4]. The cells can acquire the above mentioned status of peritrichously flagellated and mobile stationary cells. Another cell fate is sporulation. Here, as a result of an asymmetrical cell division the mother cell forms a heat and desiccation resistant endospore that can endure adverse environmental conditions [6]. An alternative cell fate that only a small subpopulation develops is the state of natural competence. This is the ability to take up and incorporate exogenous DNA (exoDNA), either to promote chromosomal integrity or to acquire new genetic traits, also known as horizontal gene transfer (HGT). These are three examples of differentiation a *B. subtilis* cell can undergo, but there are several more [4, 7, 8]. Each of these underlying genetic programs has a specific sigma factor, or even a set of specific sigma factors, e.g. σ^E, σ^F, σ^G and σ^K for Sporulation or ComK for competence. The differentiation of individuals of isogenic cultures is usually subject to bistability; meaning the regulation of these transcriptional regulators results in an either/or decision, both equally stable and the on-switch being of lower probability. Usually these genetic programs also receive input from other circuits

1

integrating environmental signals such as the nutrition level or the population density by quorum sensing [9, 8].

1.2.   Horizontal gene transfer (HGT)

Bacteria are not only able to achieve genetic information vertically that means from mother cell to daughter cells, but are able to transmit genetic information horizontally to contemporaries. The mechanisms involved are transduction, conjugation and natural competence [10, 11].
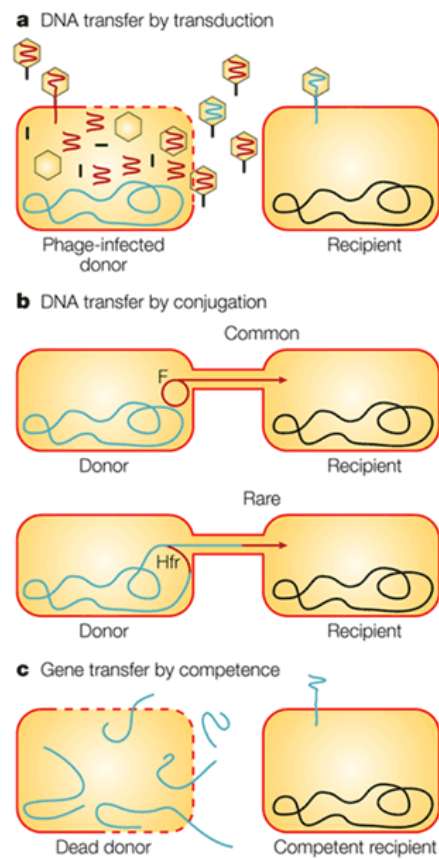
Figure 1 Horizontal Gene Transfer (HGT) [11]: Three paths for HGT; transduction (a); conjugation (b); natural competence (c)

In transduction the vector is a bacteriophage that introduces DNA from its former host (donor) to the subsequently infected cell (recipient) (see Figure 1 a). A bacteriophage is virus that propagates in bacteria; upon infection it integrates into the genome, where it can stay and be replicated within the host (lysogenic cycle) or it can directly switch the host metabolism towards phage reproduction resulting in cell lysis (lytic cycle). There is a low probability to incorporate parts of the host genome into a phage particle. These were loci neighboring the integration sites of the phage. These phages are called transducing phages. Transduction is a common tool for genetic engineering [10, 11].

The second mechanism for horizontal gene transfer is conjugation. Here, cell-to-cell contact is necessary for DNA transfer (see Figure 1 b). A molecular machine mobilizes the transforming DNA, initiates the cell contact and establishes a cytoplasmic bridge to transfer the genetic information. The ability for conjugation of a donor is encoded on specific conjugative plasmids, the fertility (F) or sex plasmids [10, 11, 12]. These plasmids could be stabile maintained or they are integrated in the genome and replicated with it [12, 13]. Conjugation is, as transduction, part of the genetic toolbox of many model organisms.

The third mechanism of HGT is natural competence. Here, the recipient expresses a molecular machine that facilitates the uptake of exoDNA (see Figure 1 c). The source of exoDNA can be diverse; usually it originates from congeners that lyse during stationary growth phase. Fragments from their chromosome, now exoDNA, are substrate to the competence machinery. These fragments are actively taken up and incorporated in the recipient's genome following homologues recombination [10, 11]. Competence is again part of the genetic toolbox and will be discussed in the following section.

1.3.   Natural competence

Natural competence is widely spread in prokaryotes; there are examples in Gram-positive as well as Gram-negative organisms [10, 14]. All organisms showing natural competence

express a molecular machinery for uptake of exoDNA. The key players are conserved; these proteins show homology even when compared in evolutionarily distinct species, such as Gram-negative and Gram-positive organisms. The main distinction is that Gram-negative bacteria have two membranes to span that are separated by the periplasmatic space and a thin layer of peptidoglycan. In contrast, Gram-positive organisms have one membrane surrounded by a thin periplasmatic cleft and a sturdy and thick cell wall [10, 15]. In any case the initial step of the uptake is to bind double stranded DNA (dsDNA) to the cell surface (see Figure 2). In Gram-positives, this is followed by the transport trough the cell wall and fragmentation of the DNA. In Gram-negatives, exoDNA is first imported through the outer membrane into the periplasm, and then conveyed through the peptidoglycan, followed by fragmentation. In both cases, there are now short fragments of dsDNA bound to the (inner) membrane. Then a single strand is transferred into the cytoplasm (See chapter 1.4.1). This single stranded DNA (ssDNA) is protected by cytoplasmic proteins and prepared for homologous recombination with the chromosome [10, 16].
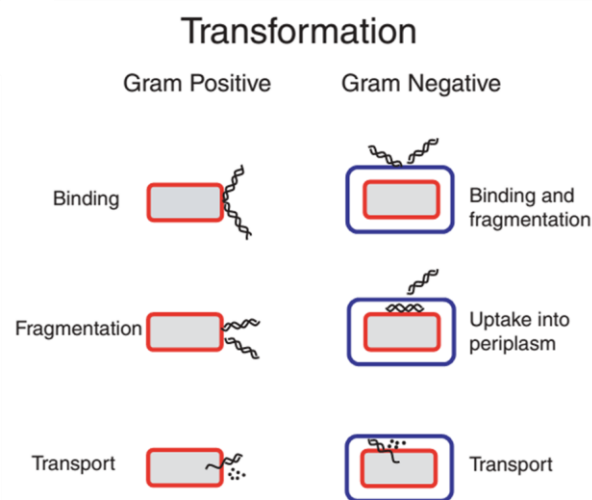

Figure 2 Natural competence in bacteria [10]: Comparison of Gram-negative and Gram-positive model of the DNA-uptake during natural competence.

1.4. Natural competence in *Bacillus subtilis*

Up to now the state of competence in *B. subtilis* is not fully understood. It is known that only a subpopulation of cells (up to 20 %) is able to develop the involved molecular machinery and they do so at the onset of stationary phase [10]. Development of competence is under control of the transcriptional regulator ComK [17, 9]. The ComK regulon consists of around 100 genes [18] (See chapter 1.4.4). The proteins that form the molecular machinery that imports exoDNA are the so-called competence (Com-) proteins. They are encoded in four operons, the so-called "late competence" operons: *comC*, *com*E, *comF* and *comG*.

1.4.1. Model of the competence machinery
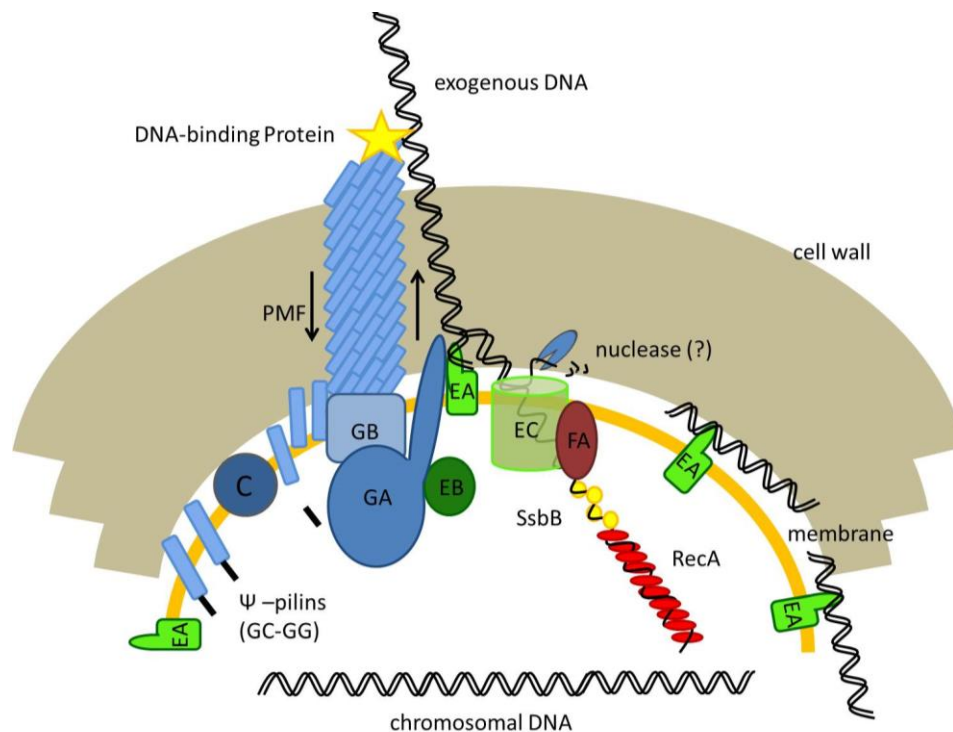
Figure 3 Cartoon model after Chen *et al*. and Kaufenstein *et al*. [10, 19]: The figure shows the current working model that describes the binding and transport trough cell wall and membrane of exogenous DNA.

Figure 3 presents a model of the competence machinery (after [10] modified after [19]), that will be explained in more detail in the following. ExoDNA is bound to a polymeric structure termed competence "pseudo" (Ψ)-pilus, due to its similarity to a type IV pilus. This structure is composed of the gene products of the "late competence" operon *comG* (See chapter 1.4.2). The Ψ-pilus is assumed to bring the exoDNA to the vicinity of the membrane where it is fragmented by the endonuclease NucA [20]. Then the dsDNA-acceptor protein ComEA binds to the double strand fragments. Then a single strand of DNA is transported through the membrane by a channel formed by the permease ComEC, while the other strand is degraded. The helicase-like protein ComFA is thought to actively support this process. Incoming ssDNA is protected from degradation by the single strand binding protein A (SsbA). The single strand binding is replaced by RecA and homologous recombination (HR) can occur [19, 10].

It has been shown that a single functional competence machinery localizes exclusively to the pole [21, 22, 23, 24]. Until now, it is not clear whether there is just individual functional machinery at the pole or whether an assembly of subcomplexes occurs and these subcomplexes form arrays of higher numbers. Additionally, there are a number of cytosolic, membrane-bound or membrane-associated Com-proteins that have, till now, no assigned function, e.g. ComEB, ComFB or ComFC. All of them localize to the cell pole, too [10, 19, 22, 23, 24].

1.4.2.   The competence proteins

The late competence gene *comC* encodes for the Ψ-prepilin peptidase ComC (248 amino acids (aa)). After translation and insertion into the membrane, all Ψ-prepilins have to be N-terminally processed by ComC before being assembled into the competence-pilus. The late competence operon *comE* contains four open reading frames (ORFs). That is *comEA*, *comEB*, *comEC* and *comER*. The former three are transcribed in forward direction. The latter is transcribed in reverse direction and overlapping with *comEA*. The transcription is

controlled by ComK, except for *comER* having a σ$_A$, the housekeeping sigma factor, promotor. It is not known whether ComER is essential for competence but the arrangement of the operon suggests a regulatory role of the ORF, repressing transcription of *comE* efficiently in exponential phase [25]. ComEA (205 aa) is essential for transformation, for dsDNA-binding and ssDNA-uptake. It consists of a single membrane-spanning helix at the N-terminus and a predicted DNA-binding domain at the C-terminus [26]. Kaufenstein *et al.* showed that ComEA-dsDNA complexes form a pool of DNA-fragments bound to the membrane awaiting further processing and internalization. ComEB (189 aa) the product of the second ORF of the operon, has been reported as unessential for transformation and its function is unknown [18, 27]. It shows a predicted homology to dCMP deaminases. The fact that it is present in the operon suggests a connection to the nucleotide metabolism (See chapter 1.5) or that parts of the competence machinery might have evolved out of the pool of enzymes from nucleotide metabolism. The deletion of *comEB* results in the delocalization of ComGA, so it might also have a role in the proper positioning of individual proteins or subcomplexes of the competence machinery at the pole [27]. The third ORF of the *comE*-operon encodes for the permease ComEC. It has a size of 776 aa. It is currently assumed that ComEC forms a membrane channel as a homodimer [28]. Every subunit contains seven membrane-spanning helices, a C-terminal loop, an N-terminal loop and an amphipathic helix that is inserted in the membrane. It is stabilized by intramolecular disulfide bonds [28].

The *comF*-operon encodes three ORFs, *comFA*, *comFB* and *comFC*. ComFA is the largest product with 462 aa. It shows a significant similarity to the DEAD box family of ATP-dependent DNA/RNA helicases and is essential for transformation [29, 10]; it is currently assumed that ComFA is involved in ssDNA import through the ComEC-channel. The second ORF encodes for ComFB (98 aa) which is of unknown function. The last ORF encodes for ComFB (229 aa) which is also of unknown function and is assumed to be dispensable for transformation. It shows sequence homology to an essential Com-protein in *Haemophilus influenzae* [30].

The late competence operon *comG* encodes proteins that are thought to form a pilus-like structure with homology to the type IV pilus [10, 16] (See chapter 1.4.1). The conserved proteins include a cytoplasmic ATPase of the AAA$^+$-ATPase superfamily, ComGA (356 aa). Followed by a polytopic membrane protein, ComGB (323 aa). There a several Ψ-pilins and Ψ-pilin-like proteins, the major Ψ-pilin being ComGC (98 aa) and the three minor Ψ-pilin-like proteins being ComGD (143 aa), ComGE (115 aa), Com GF (127 aa) and ComGG (124 aa) [10]. All of these Ψ-pilins have to be processed before incorporation into the pilus (See above). For stabilization of the pilus-structure intramolecular disulfide bonds are introduced by heterodimer BdbCD (138 aa and 222 aa, respectively) forming a thiol-disulfide oxidoreductase [10]. It has been shown that the *comG* gene products are dispensable for transformation when the cell walls have been removed. A possible explanation is that the Ψ-pilus transports exoDNA across the cell wall bringing it to the dsDNA acceptor protein ComEA. A proposed model is the extension/retraction by assembly/disassembly; this hypothesis is tempting due to the homology to type IV-pili where two different ATPases energize the dynamic system, PilT for assembly and PilF for disassembly. Until now we know only of a single ATPase, ComGA, in *B. subtilis*. So how this dynamic process is energized is unknown. Also, the transport of DNA across the membrane needs the proton motive force (PMF), a proton gradient established actively by the cell; it is possible that energizes pilus assembly as well [31, 32].

### 1.4.3. Factors processing incoming ssDNA

Other than in some organisms (e.g. *Vibrio cholerae*), there is no DNA-sequence-based restriction of exoDNA to be taken up by *B. subtilis*. A competent *B. subtilis* cell can import any kind of naturally occurring circular or linear DNA, such as chromosomal, viral or plasmid DNA. On the cytosolic side, proteins differentiate between the different origins of DNA [22]. This process can be divided into the protection of the incoming ssDNA and the promotion of HR or other recombinational events. Due to the fact that the integration into the chromosome follows the HR path, there is a strong interplay with enzymes for the

repair of DNA double strand breaks (DSBs). The following proteins are known to be part of the cytosolic pathway of HR dependent transformation: RecA, RecU, RecO, RecR, RecN, SsbA, DprA and CoiA (YjbF). Even without exoDNA RecA, RecU, and to some extent RecN are positioned at the pole in vicinity of the uptake machinery In contrast to this, RecO and the majority of RecN are dispersed over the nucleoid or localize in foci upon the nucleoid when DNA damage is present. Three pathways have been proposed for the differentiation depending on the origin of the exoDNA [5, 24].

First, if the ssDNA has sufficient homology to the recipient chromosome the process involves RecA, the major recombinase of *B. subtilis*; here RecA is loaded onto ssDNA forming dynamic heteroduplex-filaments that are able to screen the nucleoid for homologous regions where HR is performed. This pathway and the involved proteins are similar to the preferred repair route in DSB repair and will be described in more detail elsewhere (See chapter 1.6.2). The second pathway occurs in the absence of significant homology to the chromosome and when the ssDNA derives from circular dsDNA (plasmid transformation). Here, RecA is not needed. RecU is essential for transformation with plasmid DNA, possibly due to the down-regulation of RecA. The process itself seems to be RecO depended. RecO is recruited to the pole upon addition of plasmid DNA to competent cells. Its suggested role is the annealing from ssDNA to dsDNA. With sufficient internal homology the dsDNA will be assembled in a circular plasmid by intramolecular homologous recombination and if the origin of replication (*ori*) is present in the sequence, the plasmid will be replicated [24, 5, 22]. Another mechanism is proposed for transformation with viral DNA. The model proposes a mixed mechanism of the former two pathways. Initially, the incoming viral DNA forms dsDNA in a replicative manner, as observed with plasmid DNA. Next the full length linear sequence is assembled, by recombining overlapping regions, in a RecA dependent manner. Finally, the linear sequence is circularized by intramolecular recombination to a replicative circular phage molecule [22, 24].

1.4.4. Regulation of competence

The development of the state of competence is subject to bistability. This means that in a monoclonal isogenic culture the expression of the competence genes is heterogeneously distributed. Only a small (< 20 %) subpopulation develops competence [8]. This is achieved by the control of the master regulator of competence ComK. This occurs on several levels including a positive feedback loop by ComK activating its own transcription. ComK is a 192 aa protein that is active as a tetramer. There is a stringent transcriptional control and ComK is constantly degraded by delivery of ComK through the adaptor protein MecA to the protease complex ClpC/X (see Figure 4) [17, 9].
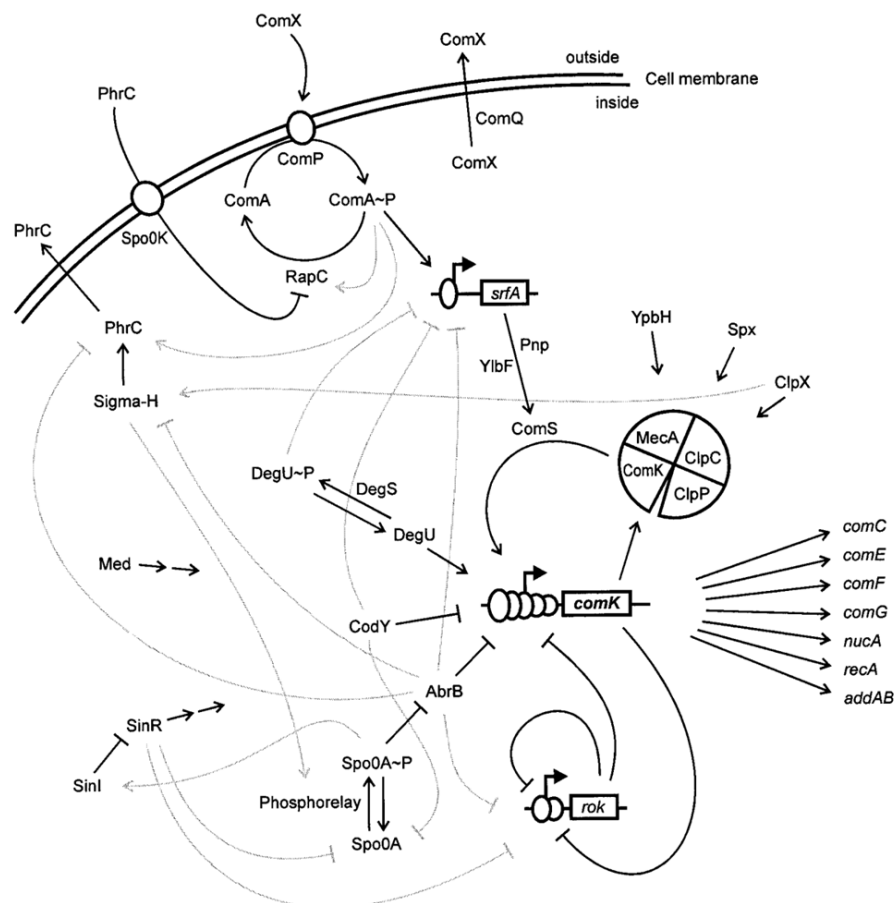
Figure 4 Regulation of competence development [9]: Various cues from cell density to nutrition state influence the upregulation of competence genes.

The regulon controlled by ComK consists not only of the late competence genes but of around 100 genes that are either activated or repressed, e.g. *nucA*, *recA*, *addBD* and *comK* itself [9, 18, 33, 34]. There are several external and internal signals starting the transcriptional activation of *comK*. Firstly, the expression of *comK* is tightly controlled, mainly by the <u>r</u>epressor <u>o</u>f Com<u>K</u> *RoK*. *RoK* binds to specific sequences in the promotor region of a ComK-dependent promotor, repressing transcription. By activation through DegU, ComK can compete with RoK for its own promotor [17]. There is also an influence of Spo0A in its phosphorylated form on the system. Spo0A is the phosphorelay response regulator that is thought to form a cellular clock, accumulating phosphate residues over time, comparable to cytokines in eukaryotes, and initiates sporulation late in stationary growth phase [34]. Dubnau and colleagues proposed that low to intermediate levels of Spo0A-P open a window of opportunity for cells to develop competence by competing with RoK for repressor boxes and enhancing binding of ComK to the promotor region [34]. Secondly, quorum sensing pathways have a strong influence on the delivery of ComK to the protease complex; there are two pathways involved, the ComX/ComP and the PhrC/Spo0K pathway. The peptide ComX (9-10 aa) is a competence pheromone. It is cleaved by the peptidase ComQ and activates its receptor: the membrane bound histidine kinase ComP. After a phosphorelay, starting with the phosphorylation of ComA, the transcription of *comS* gets upregulated. ComS is essential for competence development since it relieves ComK from the delivery to ClpC/. PhrC, a 40 aa long peptide, also acts as a pheromone. It is detected by a separate sensor kinase, Spo0K. Upon phosphorylated it inhibits RapC, a phosphatase that would otherwise inhibit ComA (see above) [9, 17].

There are several other connections to housekeeping pathways, e.g. the cell cycle via AbrB, a transcriptional regulator that regulates gene expression during the transition from growth to stationary phase, and nutrition state via CodY, a transcriptional pleiotropic repressor [17].

1.5. "Why does natural competence exist?"

Natural competence is a wide spread trait. There are only few known cases in which species secrete DNA actively in the environment [4]. Nonetheless, DNA is a common and abundant contamination of subterranean water and soil. Concentrations constitute up to 10 µg of free DNA per gram of soil, the majority of which is of bacterial origin. The phosphor esters of pentose sugars are very stable compounds, not only in a cell or aqueous solution, but especially in water-free complexes with clay or other soil minerals [35]. Another habitat with comparable concentrations of free DNA is the respiratory tract of mammals with concentrations of approximately 300 mg per ml mucus [36]. In this environment competent bacterial species are also commonly present, e.g. *H. influenzae*, *Streptococcus pneumoniae* and *Neisseria meningitides*. So why do bacteria take up DNA from the environment? There are several reasons discussed.

The first, and for laboratory routine the most important one, is HGT. A newly established genetic trait can be acquired faster via HGT than by mutation and selection cycles. However, transformation could be detrimental, since disadvantageous mutations or toxins, without the respective antitoxin, could be taken up along a new trait. When performing a gedankenexperiment with the first organism to have evolved an early variant of the uptake machinery, this species would risk their newly acquired fitness gain. The cells could minimize this risk by only taking up short stretches of exoDNA or my minimizing the subpopulation of competent cells [37].

Another reason could be the utilization of DNA as a nutrient source. In every known competent organism only ssDNA is internalized. Meaning 50 % of a high-molecular energy-rich compound would remain unused. In many organism development of competence is growth phase dependent and sets on at the transition to stationary phase. During this phase resources are becoming scarce, cell lysis starts and DNA gets released into the medium. So the uptake of DNA is possible and can increase the fitness of the competent cells significantly [37, 4, 1].

A last theory describes competence as a method for chromosomal integrity of a culture; meaning that the exoDNA originating from congeners serves as template for DNA repair. This hypothesis has been discussed controversially in literature. Some organisms such as *Vibrio cholerae* that are naturally competent are known to take up solely their own or closely related DNA identified by specific recognition sites [38, 39]. This would hint to DNA repair. Additionally it has been shown that *B. subtilis* may survive DNA damage to a higher extent in presence of exoDNA, compared to cells damaged in the absence. In contrast to this early work, Redfield [37, 36] reports for *H. influenzae* that repair of the chromosome via transformation could only minimally increase the survival rate of damaged cells. Also he states that there is no induction or upregulation of competence during DNA repair, in *B. subtilis* as well as in *H. influenzae* [37]. But since different pathways of DNA repair are not always induced at the same time or upon the same signal [40] and even a subpopulation in the single digit percentage could be sufficient for a strain to survive adverse environments, DNA repair could potentially constitute pressure towards evolution of competence, but it would probably not be the only factor [11].

1.6.   DNA repair in *Bacillus subtilis*

It is essential for all organisms to repair damages in their genetic information quickly and faithfully. In general, DNA damage can have several reasons. These can be divided in endogenous factors, e.g. mistakes in replication or reactive oxygen species originating in sugar metabolism, and exogenous reasons, e.g. toxic substances or ionizing radiation. *B. subtilis* has several pathways for the repair of different DNA damages.

The base excision repair (BER) system is recruited to repair post-replicationaly modified pyrimidine or purine bases. These bases get excised to create an apurinic or apyrimidinic site. In a second step, this site is recognized, nicked and later on corrected by polymerases using the complementary strand as template [41]. The nucleotide excision repair (NER) system is employed to correct damages resulting in helix-distortion such as thymine dimers. This UvrABC pathway is highly conserved among kingdoms and best studied in

Escherichia coli. The respective homologs in *B. subtilis* are named identically [42]. Here, the bulk in the DNA-helix is recognized, excised by exonucleases in 3' and 5' direction, and re-synthesized by a polymerase [42, 43]. Mismatch repair (MMR) is a system employed to repair falsely incorporated bases that are not directly recognized by the proofreading function of the DNA polymerases. Key players are the highly conserved proteins MutS and MutL. In regard to this system *B. subtilis* differs from the much better understood organism *E. coli* since it does not possess a Dam-methylase, allowing differentiation between old and newly synthesized DNA strands. The differentiation between old and newly synthesized strand thus seems to be dependent on the orientation of the beta clamp (DnaN) of the DNA-dependent DNA-polymerase complex [43, 44]. MutS is recruited by DnaN, and then recruits MutL; its latent endonuclease activity is stimulated followed by strand removal, resynthesis and ligation [44, 43]. A different quality of DNA damage is the break of the DNA double strand. A strand break results in the collapse of the replisome, arrest in cell cycle, and ultimately cell death. One way to repair DSBs is the non-homologous end joining (NHEJ). NHEJ is a low-fidelity DSBs repair pathway. Here the highly conserved enzymes Ku and LigD are the key players [45, 43]. Ku tags the loose dsDNA ends and recruits LigD that ligates the ends. But there is the danger of genetic information being lost; there might be several DSBs or the loose ends might me processed by exonucleases before end joining. NHEJ seems to be of special importance in the endospore development or during extended periods of stationary phase [45]. In both of these cases *B. subtilis* has predominantly one chromosome per cell [46]. In contrast to textbooks, *B. subtilis* has during exponential growth more than one chromosome per cell. But when cells are polyploid the danger of falsely connected DSBs by NHEJ is much higher. But with an extra copy there is a template for faithful repair by homologous recombination.

1.6.1.   Overview on homologous recombination
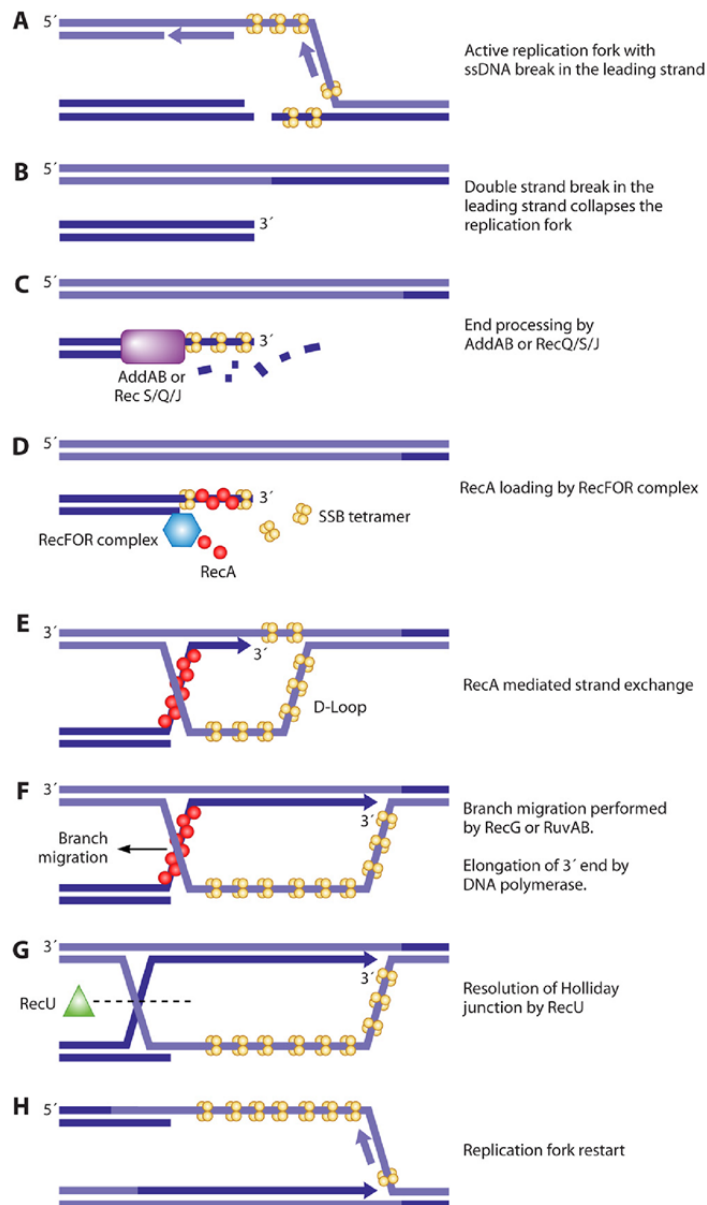
Figure 5 Repair via HR of a single DSB [43]: The figure shows a scheme of the current model of DSB repair in *B. subtilis*

Homologous recombination (HR) is a universal mechanism to mediate exchange of homologous DNA stretches and is therefore employed for the faithful repair of DBSs. In

every organism the genome is, at least temporarily, present as a duplicate. This second copy is used as a blueprint for repair. The model of HR in *B. subtilis* lists five steps (see Figure 5): (1) recognition of the DSBs by RecN; (2) resection/processing of the dsDNA ends; (3) first protection of ssDNA by SsbA and the subsequent loading of RecA onto the strand; (4) search for homology, strand invasion, formation of a D-loop and finally a Holliday-junction (HJ); (5) resolving of the HJ and separation of the two intact chromosomes. The steps (1) to (3) are termed presynapsis. First part of step (4) is called synapsis, when the two chromatids are connected. The end of (4) and (5) are called postsynapsis [40, 47]. In general we do have a working model of the HR in *B. subtilis* but the role of individual enzymes in this concerted reaction cascade still has to be investigated. The alternative helicase nuclease complex ReJQS, or the recombinase RecA itself are examples that need further elucidation.
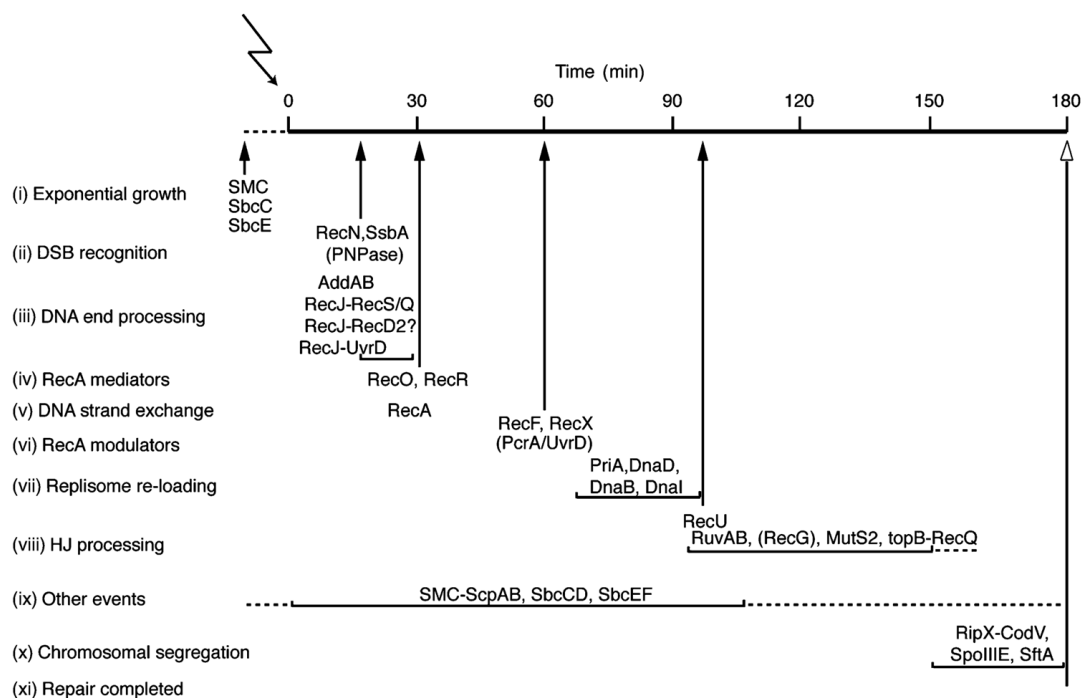
Figure 6 Well timed orchestra [47]: The key players do perform DSB repair in a tightly orchestrate interplay; only after three hours cytokinesis goes on.

16

In an initial step, the sensor for DSBs RecN binds to the dsDNA ends and starts the HR repair cascade. In eukaryotes the formation of repair centers (RCs) was observed upon the induction of DSBs [48]. Epifluorescence studies propose such RCs for *B. subtilis* as well [49], composed of multimerizing RecN proteins tethering loose DNA ends. In the ongoing process RecN recruits the downstream acting enzymes, first the PNPase, a polynucleotide phosphorylase that processes the 3'-ends removing the first few, possibly aberrant nucleotides. Next, the double strand is relaxed by a helicase and the 5'-end is resected by an exonuclease. In case there is no overhang, this task is performed by the AddAB complex [47]. If there is a 3'-overhang RecJ, a 5'-exonuclease, together with the helicase RecQ are performing the reaction. The resulting overhang of ssDNA is protected by the single strand binding protein A (SsbA). Later in the process SsbA is replaced by RecA. Since RecA has a lower affinity to ssDNA than SsbA, this is mediated by the RecA loading complex consisting of RecO and RecR [50, 47]. The resulting RecA-nucleofilament performs the search for the homology region in the sister chromosome [49, 47]. The growth of the RecA filament is regulated by accessory factors [51]. At the site of homology, the RecA-nucleofilament can invade the double strand, forming an intermediate of three strands, the D-loop. There, the new synthesis takes place. This strand invasion is promoted to a HJ by translocases such as RecG or RuvAB. Upon finished synthesis the nick is ligated and the HJ is resolved by RecU resulting in two intact double strands [47, 52].
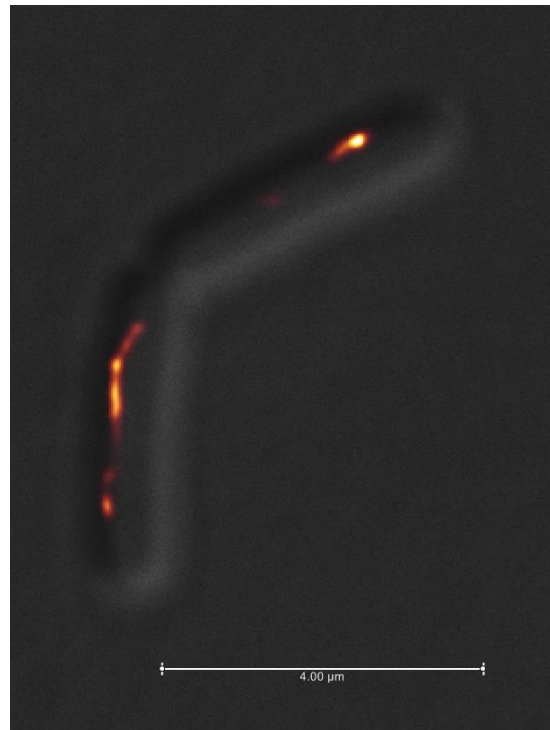
Figure 7 GFP-RecA: The figure shows a superresolution micrograph (Leica Sp8 gSTED) of a merodiploid strain showing RecA filaments 90 min after the induction of DSB. White represents 4 μm

1.6.2.   Proteins involved in HR

RecN is the first protein to localize in clusters upon induction DSBs. This recombination (Rec) protein is a member of the SMC-like protein family and has a length of 576 aa. SMC, short for <u>s</u>tructural <u>m</u>aintenance of the <u>c</u>hromosome is part of the *Bacillus* condensing complex condensing the chromosome. Proteins of the SMC-like family consist of a central coiled-coil domain and, formed by C- and N-terminus, an ATP-binding cassette. The current model is that RecN forms, comparable to eukaryotes, a repair center (RCs) tethering several dsDNA ends and recruiting enzymes that act downstream [48, 53]. Under exponential growth conditions RecN is homogenously distributed over the nucleoid and starts to form foci upon the induction of DSBs. There is also a threefold upregulation of expression levels upon induction of DNA damage [53]. Other than SMC there is no hinge-

18

domain in RecN present that would serve as a self-interaction surface for dimerization. Nonetheless a 3D structure derived from crystallography data from *Deinococcus radiodurans* shows that RecN does forms dimers via the distal end of the coiled-coil-domain. Although sequence similarity to RecN$_{B. subtilis}$ is poor, there are identical key residues suggesting a conserved structure [54]. There is data that proposes multimerization through interaction of the ATP-binding domains in the presence of Mg^{2+} [55].
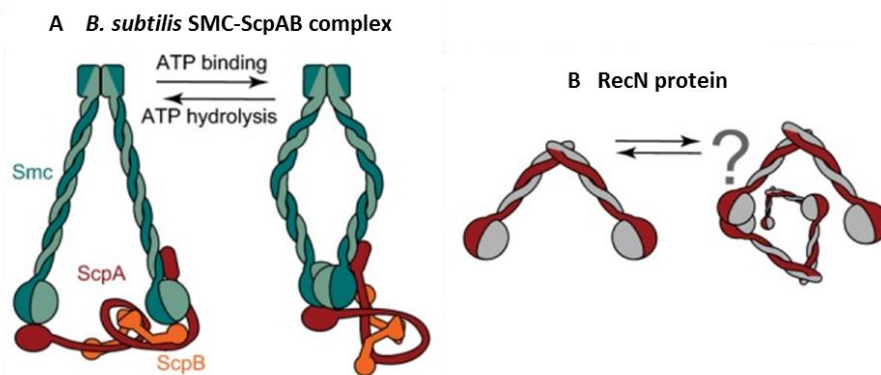
Figure 8: Cartoon model of RecN [55]: A member of the SMC-like protein family

The polynucleotide phosphorylase PNPase is one of the first enzymes recruited to the RCs. It is a 705 aa large protein. Its primary function is that of a 3'-5 exoribonuclease and it is also part of the RNA degradosome. In the context of DSB repair it functions as 3'-5' exonuclease, with ssDNA as substrate, to resect aberrant nucleotides at the end of dsDNA that could arise through adduct reactions that initially caused the DSB [47, 40].

Afterwards a helicase/nuclease complex, which would be RecBCD in *E. coli*, co-localizes with RecN foci, presumably processing dsDNA, generating stretches of ssDNA. *B. subtilis* lacks the RecBCD-complex; the functional analogue is the AddAB complex, the ATP-dependent deoxyribonuclease [43]. The complex is a heterodimer consisting of AddA (1232 aa) and AddB (1166 aa). The subunits share some homology and carry an UvrD-like helicase domain at the C-terminus. It is a rapid DNA helicase with a 3'-5'-polarity motor in

the AddA subunit. The exact nuclease domain is not known. There is an additional ATP-binding domain in the AddB subunit that seems to be involved in recombination hotspot detection. The *chi* recombination hotspots are DNA sequences that slow down the AddAB nuclease activity to promote downstream recombination to occur [43, 56].

RecJ, an ssDNA specific exonuclease, is processing the dsDNA if there is a 3'-overhang [57]. RecJ has a size of 786 aa [47, 58]. It has been shown that RecJ interacts with the C-terminus of single strand binding protein SsbA. SsbA is sequestered to the replisome during exponential growth, so the exonuclease RecJ is kept in proximity to an endogenous cause of DSBs: the roadblock induced collapse of the replication fork [59, 43, 57].

RecQ (496 aa) and the RecQ-paralogue RecS (352 aa) are ATP-dependent DNA helicases. They do interact with SsbA similarly to RecJ [59]. There is evidence that RecJQS act together as DNA helicase nuclease complex comparable to AddAB [56]. Otherwise the role of RecQ and RecS has to be further elucidated.

RecO does have a role in plasmid transformation (See chapter 1.4.3). But the best understood role is its part in the RecA-loading complex. RecO has a size of 255 aa and there is evidence from crystallography that it forms a heterocomplex with RecR (198 aa) in the ratio of 1:2 (RecO: RecR). As a complex they provide RecA access to ssDNA in DSBs repair and in transformation with chromosomal DNA [50].

The key player in HR is the recombinase RecA. It is one of the 100 most abundant proteins in *B. subtilis* [60]. A RecA monomer has a size of 347 aa. RecA has a lower affinity to ssDNA then SsbA and has to be loaded on ssDNA (see above) [50, 47]. In its ATP-bound form it does facilitate the strand exchange in concert with cofactors, but also in the absence of any accessory protein. In the exponential growth phase it is dispersed over the nucleoid. In presence of DNA damage it localizes to the replisome, later on forms foci on the nucleoid, and then starts filamentation. Filaments are stable over one to one and a half hours [49]. There is evidence that is has a regulatory effect on proteins that are thought to

act upstream of RecA in the current model (see Figure 5) [47]. There are regulatory factors that either stimulated the filament growth or negatively regulate filament length (see below). The RecA-ssDNA nucleoprotein filament spans the distance between the two sister chromatids. In *B. subtilis* the segregation of the newly synthesized chromosome happens in parallel to the replication and the homologous regions are spatially separated [49, 47]. The RecA-filament does perform the search for the exact site and the invasion in the intact double strand, forming a three-strand intermediate (D-loop) (see Figure 5). In Figure 9 the molecular structure of *E. coli* RecA is shown; RecA$_{E\,coli}$ has 62% identity and 86% similarity to the protein of *B. subtilis* [61].
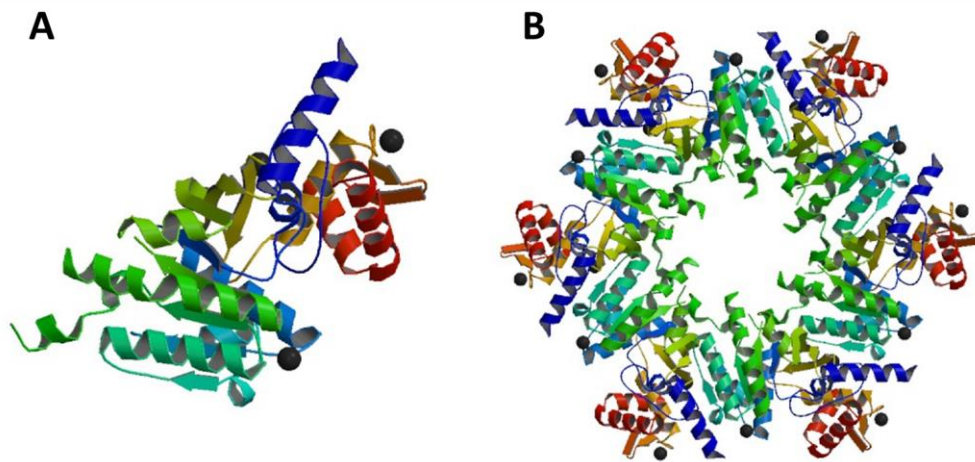
Figure 9: Structure of *E. coli* RecA [61]. The figure shows the structure of the major recombinase RecA, a homohexamer.

RecF and RecX are known facilitators of RecA. RecF (370 aa) stimulates the RecA DNA repair center assembly. It is a positive regulator of RecA-ssDNA-nucleoprotein filaments, stimulating filament growth [51]. RecX (246 aa) modulates the length or packing of RecA filaments. It stimulates the start of recombination and negatively regulates filament length in later steps of HR [51]. Overproduction of RecX increases recombination. It localizes in foci on the nucleoid in case of DNA damage and forms distinct polar foci in competent cells [51, 43, 47].

2. Material & Methods

2.1. Chemicals

Standard chemicals were purchased from Roth (Karlsruhe, Germany) or Applichem (Darmstadt, Germany). Fine chemicals or chemical dyes were purchased from Life Technologies (Carlsbad, USA). DNA polymerases, restriction endonucleases and other DNA modifying enzymes and markers were purchased from New England Biolabs (Ipswich, USA). DNA purification kits were manufactured by Qiagen (Hilden, Germany)

2.2. Plasmids & deoxyribo-oligonucleotides

To monitor specific protein localization in *Bacillus subtilis*, coding regions were cloned to vectors of the pSG-series (Table 3) designed to allow C- or N-terminal fusion with different variants of fluorescent proteins and subsequent selection for antibiotic resistance in *Escherichia coli* and *B. subtilis* [62]. The fusions were integrated into the chromosome by homologous recombination, either via a single Campbell-type integration at the original locus, or via double crossover integration at the *amyE*-locus. *amyE* encodes for the exoenzyme amylase. If this locus was targeted, clones were tested on loss of amylase activity to confirm integration. Depending on the vector, the fusion proteins were under the control of either their native promoter or the artificially introduced xylose promotor (P_{xyl}). Expression was then induced by adding xylose to varying final concentration ranging from 0.01 % to 0.5 % (m/w) (Table 4). The p*CM::tet* plasmid was used to exchange a chloramphenicol (cm) resistance by a tetracycline (tet) resistance in *B. subtilis* [63]. The respective strain was transformed with this plasmid, disrupting the cm-cassette and establishing tet-resistance. Protein overexpression in *E. coli* was accomplished by the expression vector pET24-d (Novagen, Nottingham, UK) that provides IPTG inducible expression based on the T7 phage system.

An overview of the deoxyribo-oligonucleotides used to generate plasmids for strains used in this study is provided in Table 3.

2.3.   Preparation and transformation of competent *E. coli* cells

Competent *E. coli* cells, either DH5α for cloning or BL21 for overexpression (see 2.6), were prepared following a slow growth protocol as described in [64]. For transformation, plasmid DNA was added to the cells with different concentrations: in the case of a preceding ligation the total ligation volume was added, when closed plasmid was used < 1000 µg were added. Cells and DNA were incubated for 5 min on ice, then heat shocked for 2 min at 42 °C, then cooled for 10 min. In a final step, 900 µl of super optimal broth with catabolite repression (SOC) (see Table 4) was added and the cells were incubated at 37 °C for one hour and then plated on selective plates (see Table 5).

2.4.   Preparation and transformation of competent *B. subtilis* cells

B. subtilis is a naturally competent organism (see 1.4). The subpopulation of competent cells can be enriched by growth in media that suppresses sporulation and can be stored until transformation. 10 ml SpC medium (see Table 4) were inoculated by collecting cells from a LB-agar plate grown O/N at 30 °C, and incubated at 37 °C until optical density at λ = 600 nm (OD_{600}) detected constant results. The culture was diluted 1:5 in 50 ml SpII medium and further incubated for 90 min at 37 °C. Cells were harvested by centrifugation (4000 rpm, RT) and resuspended in 10 ml supernatant with 5% (v/v) glycerol. The suspension was aliquoted and stored at -80 °C. For transformation, 125 µl of the suspension were incubated with approximately 0.5 µg chromosomal DNA or 5 µg plasmid DNA for 30 min (37 °C, 200 rpm), and then plated on selective plates (see Table 5).

2.5.   Growth medium and supplements

E. coli cells were usually grown in Luria-Bertani (LB) medium (see Table 4) at 37 °C. To prepare LB-agar plates 1.5 % (w/v) agar was added. To prepare competent *E. coli* cells SOB and SOC was used. *B. subtilis* was grown in LB at 30 °C, or in the minimal medium S7$_{50}$ at 30 °C. To prepare competent *B. subtilis* cells SpC and SpII was used (see Table 4). The media were either sterilized in an autoclave (121 °C, 2 bar) or by filtration (0.2 µm pore size). To quantify cell densities the OD$_{600}$ was measured. The concentrations of selective antibiotics are listed in Table 5.

2.6.   Bacterial strains

E. coli strain XL-1 Blue (Agilent Technologies) (see table 2) was used for the propagation of constructed plasmids. *E. coli* strain BL21 Star DE3 (Life technologies) (see table 2) was used for heterologous overexpression of proteins. All *B. subtilis* strains constructed and used in this work were generated in the background of the prototrophic wild type strain PY79. The strains in used in this work are listed in Table 6.

2.7.   Molecular biology

All procedures related to the construction of vectors (extraction of plasmid and chromosomal DNA, PCR reactions, agarose gel electrophoresis, purification, digestion and ligation of DNA) were performed following standard protocols as described in Molecular Cloning [65]. Enzymes were used according to the manufacturers' recommendation.

2.8.   Heterologous protein overexpression and indirect immunodetection

In order to observe localization behavior in the absence of assumed interactors in *B. subtilis,* proteins of interest have been cloned as full length constructs in expression vectors (see Table 3) and *E. coli* BL21 was transformed with the resulting plasmid. To

observe localization the expression was induced by the addition of IPTG (see Table 4). To verify expression of fusion proteins, to check expression levels or to adjust the level of expression of inducible promotors, I performed indirect immunodetection. Appropriate amounts of *B. subtilis cells* were lysed. Cell lysates were separated by SDS-PAGE [66]. Proteins were transferred to nitrocellulose membranes. Detection followed the protocol described by Dempwolff *et al*. in [67].

2.9. Microscopy

Epifluorescence microscopy

Specimens were mounted on top of a 24x50 mm high precision (d = 170±5 µm) coverslip and covered with an agarose pad (1% (w/v) agarose poured in S7$_{50}$). Images were acquired with an Observer.A1 (Zeiss) equipped with a Plan Fluar objective (NA: 1.45; Zeiss; Jena, Germany) and a Cascade II 512 EMCCD camera (Photometrics; Tuscon, USA). Image data were acquired with VisiView 1.7.4 (Visitron Systems GmbH; Pucheim, Germany). Fluorophores (e.g. CFP and YFP) were excited by exposing the specimen with a laser of 445 nm or 514 nm wavelength respectively coupled in by a Visitron VisiTIRF system. Fluorescence signals were acquired using the appropriate filter cubes.

Stimulated emission depletion (STED) microscopy

Specimen were mounted on top of a 24x50 mm high precision (d = 170±5 µm) coverslip and covered with a pad of 1% (w/v) agarose in S7$_{50}$. STED microscopy was performed on a Leica TCS SP8 gSTED with a 100x objective (NA 1.40). The image data were acquired and treated with the LAS AF software (Leica Microsystems; Wetzlar, Germany).

Single molecule microscopy (SMM)

Specimen were mounted on top of an r = 24 mm high precision (d = 170±5 µm) sapphire glass coverslip and covered with a pad of 1% (w/v) agarose in S7$_{50}$. SMM was performed on an Olympus IX71 equipped with an ApoN (100x, NA 1.70 HOil) objective (Olympus;

Hamburg, Germany) and an iXON Ultra EMCCD (Andor; Belfast, North Ireland). Streams were recorded using the Andor Solis software (Andor; Belfast, North Ireland)) with 24.4 Hz and a kinetic cycle time of 41 ms. Specimens were illuminated with an argon ion laser (Laser Drive Inc.; Gibsonia, USA) where the 514 nm band was extracted using appropriate filter cubes.

2.10. Data treatment/ Single Molecule Tracking (SMT)

All image data acquired were prepared for analysis or presentation in Fiji ImageJ [68]. Streams acquired during single molecule microscopy (SMM) were formatted in Fiji as well. Pretreated streams were then further analyzed with MATLAB (The MathWorks Inc.; Natick, Massachusetts), using the implements u-track and MicrobeTracker [69, 70]. Statistical tests, *e.g.* χ^2-test were performed using the statistic toolbox of MATLAB. The cell borders were established in the MicrobeTracker suite from the Jacobs-Wagner lab to ensure that all acquired trajectories would be inside a bacterial cell. U-track from the Danuser lab has been established for SMT and has been evaluated in a contest [69]. We established the parameters with our experimental data (see chapters 3.5.1 and 3.5.2). This resulted in x- and y-coordinates as a function of time, and experimental Gaussian distribution of the fluorophores and cell meshes for each micrograph.

The values were analyzed in a custom-written MATLAB software package [71], kindly established and provided by Dr. Thomas Rösch. In this MATLAB implement, the cell outlines and coordinates are rotated and normalized, which results in an average cell length of 3 µm that spreads along the X-axis and an average cell width of 1 µm that spreads along the y-axis. The trajectories were rotated and normalized accordingly.

A central quantity used in the analysis is the mean squared displacement (MSD) that describes the deviation or the area that a "random walk"-trajectory covers as a function of time.
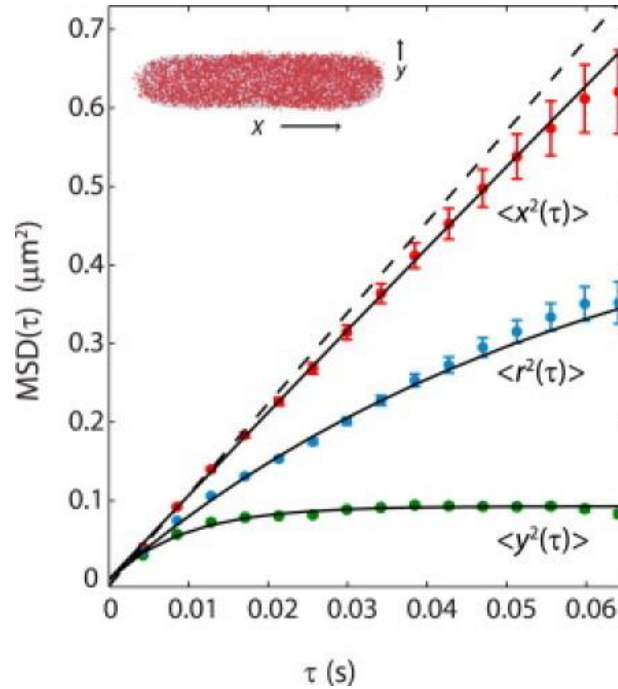
Figure 10 MSD of the fluorophore Kaede [72]: The inlet shows the coordinates of a bacterial cell. The red dots show the MSD of the squared X-displacement of the trajectories, the green dots show the MSD of the squared Y-displacement of the trajectories and the blue dots shows the MSD of the squared and summed displacements in the x- and y-axis. The apparent confinement for movement along the Y-axis is demonstrated in an asymptotical behavior instead of a linear behavior (compared to $x^2(\tau)$).

A given trajectory is fragmented in the distances, Δx, it spanned in single time increments and multiple increments up to the total timespan of the trajectory; for example, a trajectory with five time steps would give rise to four different Δx-values for $\Delta t = 1$, three different Δx-values for $\Delta t = 2$, two different Δx-values for $\Delta t = 3$ and one Δx-value for $\Delta t = 4$. As an equation the MSD would be described as:

$$\text{MSD} = \frac{1}{T}\sum_{t=1}^{T}(x(t) - x_0)^2$$

With T = total time, t = time increment, $x(t)$ = distance at a given time point, x_0 = starting position.

27

To derive the diffusion coefficient from the MSD curves, the MSD was only calculated for the first four time points, which were linearly fitted. The relation of the MSD and the diffusion constant is given in the Einstein-Smoluchowski-equation:

$$\delta^2\,(\tau) = 2N\,D\,\tau$$

With δ^2 (τ) ≙ MSD, N = dimensionality of the system (usually 2), D = diffusion constant, τ = time increment (i.e. acquisition time). Since the derived values could still be subjected to confinement and are therefore underestimated, we have to address this as the apparent diffusion constant D_{app}. Also the systematic error of the microscope setup, which corresponds to the offset of the linear fit at the y-axis were included in the calculation of the Dapp in the following formula:

$$D_{app} = \frac{MSD - 4\sigma_{loc}{}^2}{4\Delta t}$$

With $4\sigma_{loc}{}^2$ = the offset, corresponding to the Y-axis intercept of the fitted MSD-curves, σ_{loc} the estimated localization error and Δt = the frame rate of the acquired streams. We calculated D_{app} for each MSD curve at Δt = 1 and plotted the distribution as a probability density function which describes the relative likelihood that D_{app} has a given value and which therefore has the SI-unit [per μm^2 s^{-1}].

3. Results

3.1.   ComEB is necessary for the recruitment of ComGA to the pole

Deletion of comEB results in delocalized ComGA

The competence machinery does localize to one, or both, cell poles and sometimes the septum of dividing cells that is the "future" cell pole [19, 27, 5]. This localization pattern is persistent even in protoplasts and has been shown to not be disturbed by deletions of individual late competence genes [27]. A *comEB* mutant strain was constructed and combined with the fusion of ComGA-CFP. ComGA is an ATPase and a key player in the transport of ssDNA across the membrane (see chapter 1.4.2). The strain was grown to competence and prepared for microscopy as described in chapter 2.4. The micrographs are shown in Figure 11.
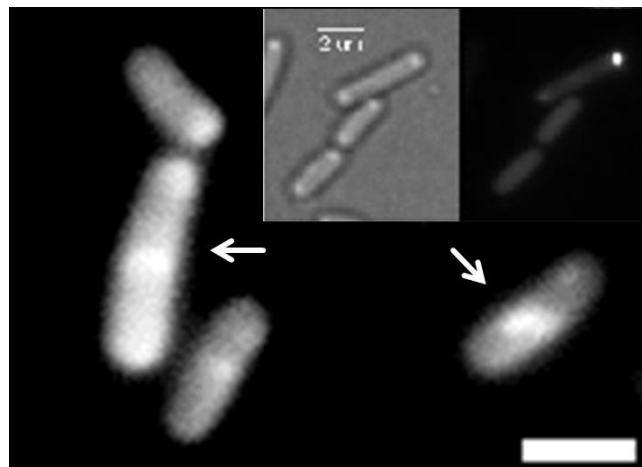
Figure 11 PY79 *comEB* ComGA-CFP: In the absence of ComEB, ComGA fails to form defined polar foci; Inlet shows the localization pattern in presence of ComEB. White arrows point at diffuse fluorescing cells. White bar represents 2 μm

There was no influence on the localization of the other Com-proteins (Data not shown). Since the *comEB* mutant was not competent a strain carrying a copy of the permease

ComEC under transcriptional control was created to exclude polar effects of the knock-out (see Table 6). However, the non-transformable phenotype persisted: while the PY79 wildtype strain showed a relative transformation efficiency of 1.0, the strain with *comEC* at the *thrC*-site showed a relative efficiency of 1.66. The strain *thrC::comEC* in the mutant background showed a relative transformation efficiency of 0. This observation was confirmed in two biological replicates (see Figure 12).
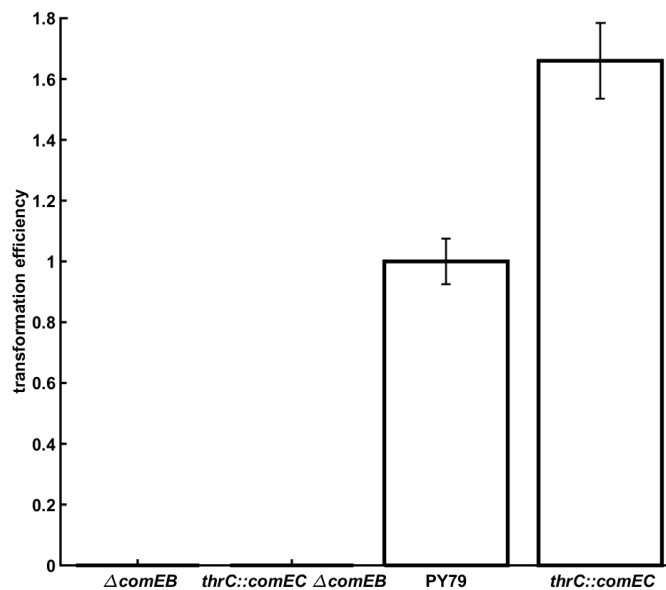

**Figure 12 Relative Transformation Efficiency of *comEB* mutant*:* The figure shows from right to left: overexpression strain of the permease ComEC (*thrC::comEC*), *wt*, *thrC::comEC ΔcomEB* and *ΔcomEB* alone. Values are normalized against the *wt;* error bar gives SE

ComEB-YFP and ComGA-CFP co-localize at the cell pole and the septum

Only the polar localization of the competence machinery results in a functional machinery and it does so only at a single cell pole [23, 24]. To further investigate whether (1) *comEB* is translated into a functional protein, (2) ComEB is a part of the polar localized machinery as well as whether (3) ComEB and ComGA do in fact co-localize, a strain carrying fusions of

a fluorophore to each of the two proteins was constructed. The fusions were integrated at the original loci and under the control of the native promotor (see Table 6).
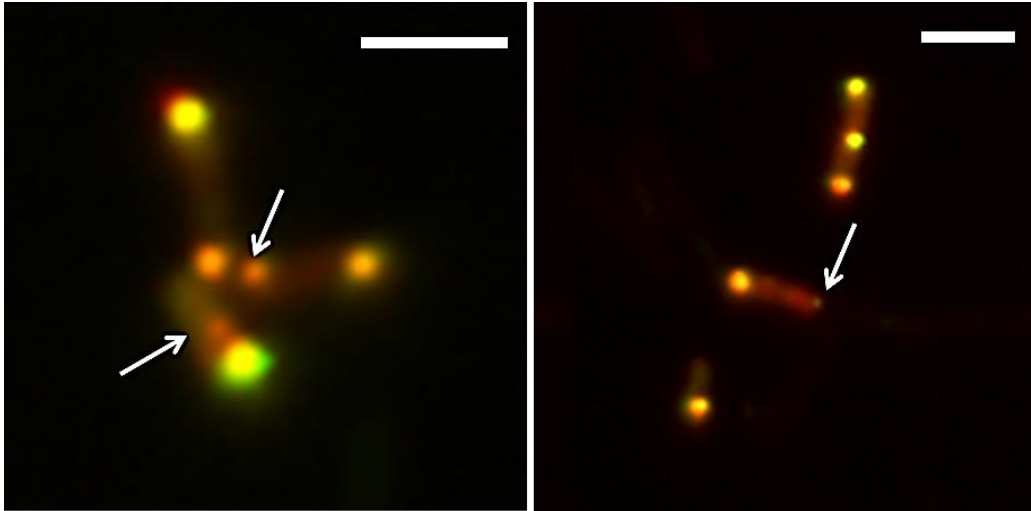

Figure 13 Overlay of ComEB-YFP (green) and ComGA-CFP (red): The figure shows the merge of the two channels; a co-localization event results in a yellow signal. The intensity varies for each foci and the colocalization is not exclusive. In the left panel arrows indicate ComGA-CFP foci without ComEB-YFP, the right panel vice versa. White bar represents 2 μm

The epifluorescence micrograph Figure 13 shows a false color overlay of two channels: ComEB-YFP in green and ComGA-CFP in red. The fluorophores were excited consecutively with the respective wavelengths. A co-localization would show, as a result of additive color mixing, in yellow.

From the total amount of cells (N = 53) 13 % showed a fluorescence signal. Of these cells, a 28 % showed a single focus and 72 % showed two or more foci. In cells with more than one focus, 20 % of the ComEB-YFP signals were not co-localizing, whereas 40 % of the ComGA-CFP signals were not co-localizing with its assumed counterpart.

3.2. ComEB expressed during exponential phase localizes to the cell pole

As mentioned in chapter 1.4.2 the localization of the competence machinery to the cell pole is very persistent. To investigate whether the localization of ComEB-GFP was independent of any known interaction partner, be it under control of ComK or any other stationary phase specific interaction, I constructed a strain carrying an additional copy of *comEB* tagged with *gfp* at the integration locus *amyE* (see Table 6) [62]. The vector had a xylose-promoter to control the expression of *comEB-gfp*. Cells were grown to early exponential phase before expression was induced by addition of xylose to a final concentration of 0.005 % (w/v). After continued incubation (45 min), the cells were prepared for microscopy (see chapter 2.9). The micrograph is shown in Figure 14. The cell borders are marked with white dashes, although the lack of membrane stain results in some uncertainty. Cells that did express ComEB-GFP showed localization at the pole or the midcell (septum).
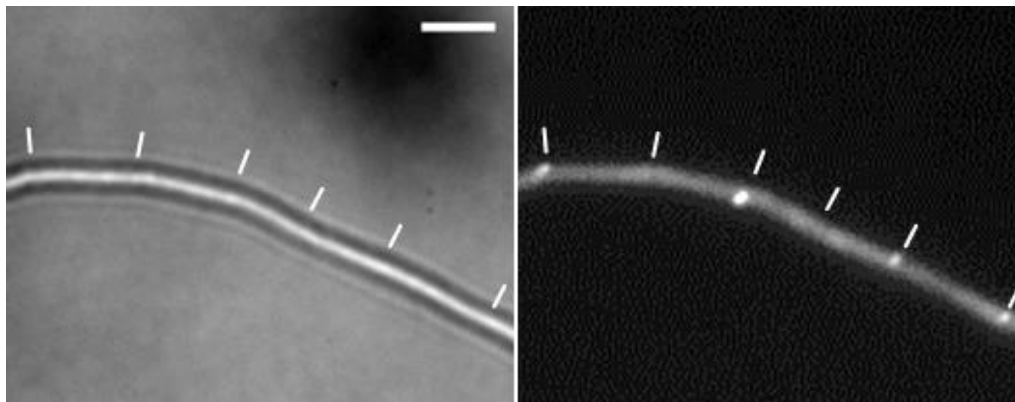
Figure 14 ComEB-GFP: The figure shows a merodiploid strain carrying a *comEB-gfp* under control of a Xylose-promoter to start the expression in absence of other Com- or stationary phase proteins. ComEB-GFP localizes to the cell pole in log-phase under low-induction conditions. White dashes mark cell border. White bar represents 4 μm

3.3.   ComEB-YFP from *Bacillus subtilis* localizes at the cell pole in *Escherichia coli*

To investigate whether ComEB from *B. subtilis* (ComEB$_{Bsub}$) displays specific affinity to the bacterial cell pole even without species-specific polar markers, such as DivIVA [73], a vector was constructed (Table 3 and Table 6) with the full length gene under the control of a T7 based expression system (Novagen). *E. coli* BL21 cells were transformed with the resulting plasmid and grown to exponential phase (OD$_{600}$ = 0.6). Expression was induced with 0.1 mM IPTG for 45 min and cells were prepared for microscopy. The resulting micrograph of ComEB$_{Bsub}$ is shown in Figure 15. In the left panel the brightfield channel is shown; in the middle panel the YFP-signal. The right panel shows an overlay of both channels with the fluorescence signal yellow false colored. The foci only appeared at the pole or in the septal region. There were a number of small and faint foci localizing to the septum as well.
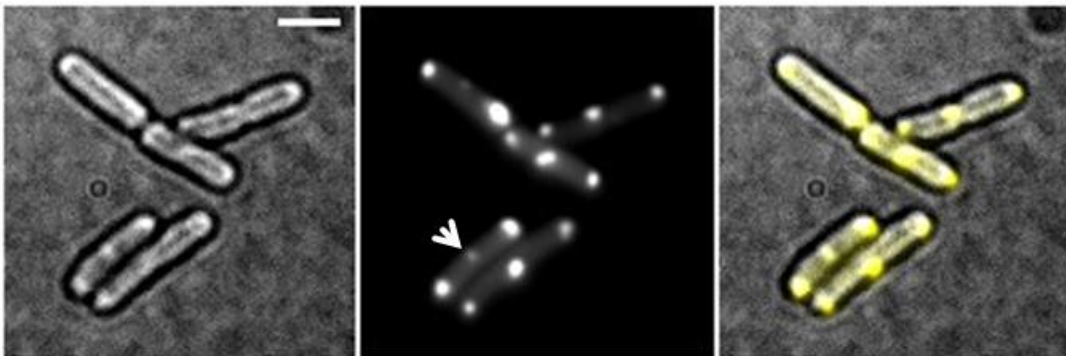

Figure 15 ComEB$_{Bsub}$-YFP: The figure shows micrographs where the protein was expressed from plasmid in *E. coli* BL21. Left panel bright field; in the middle YFP-channel; right panel overlay. (White arrowhead, see text above) White bar represents 2 µm.

## 3.4.   RecN foci persist in presence of DSBs in the seconds time-scale

RecN is one of the first proteins that form foci upon the induction of DSBs in exponentially growing *B. subtilis* cells. It has been shown that 15 min after drug treatment one, rarely

two or more foci are formed. To investigate the dynamics of this focal localization with higher temporal resolution, I performed epifluorescence imaging with low intensity of a 515 nm LED-laser, 100 ms exposure and stream acquisition.
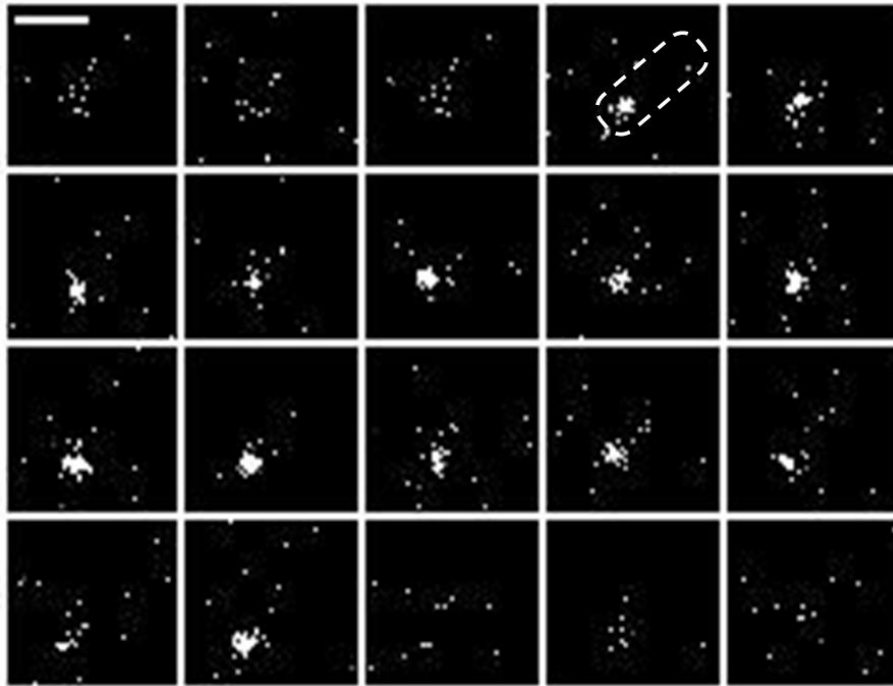
Figure 16 RecN-YFP: The figure shows a montage of a series of images showing a RecN-YFP focus moving in a single *B. subtilis* cell and finally disappearing in the last frames. Chronological order from upper left to lower right, a frame represents 100 ms; White bar represent 2 μm

In 12 streams 18 RecN foci were counted which formed and disassembled during the observation. The average time of focus persistence was 2.46 s±0.6 s (raw data videos are on the data DVD). Most foci were static, persisting at a given position, then disappearing (see Figure 16). A single focus could be observed showing movement (see Figure 17). Some foci disassembled and reestablished shortly after disassembly on a different locus on the chromosome. A data DVD was attached, please refer to chapter 7.7 for video data.
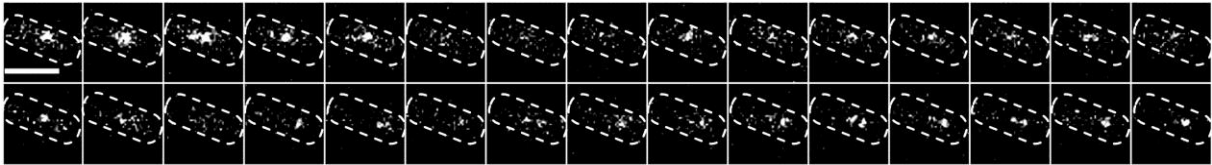
Figure 17 RecN-YFP: The figure shows a montage of a series of images showing a RecN focus resting several frames at a specific position, then moving to another and finally disassembling. A frame represents 100 ms; White bar represent 2 µm

3.5. Single Molecule Microscopy and Tracking (SMT): dynamics of Rec-proteins on the single molecule level

As described above, the proteins of the REC-pathway, or Rec-proteins, are employed by the bacterial cell to faithfully repair any occurring DNA-damage. In this study, two experimental setups were compared, that is (1) an unharmed, fast growing exponential *B. subtilis* culture, and (2) the same fast growing exponential culture treated with the DNA damaging agent Mitomycin C (MMC) with a concentration of 50 ng/µl. At this dosage around 50 % of the cells survived [49]. The duration of the treatment was dependent on the times known for the Rec-proteins to form foci in epifluorescence as stated before [47]. In all strains used in these experiments the fusion protein was expressed under the control of the native promotor. The terms induced and uninduced used in the following sections describe the chemical induction of DSBs. In the initial examination of the movie data three distinct types could be observed by bare eye: (1) single fluorescent molecules stopped in their movement and after some time moved on (see Figure 18 A), (2) arresting implied a minimal amount of displacement (see Figure 18 B), (3) most trajectories moved the whole time they were observed (see fFigure 18 C). This was in contrast to the dynamic behavior of SMC [74], for example. Literature data on the *lac*-operon inhibitor LacI support behavior like this for a DNA-binding protein, as the observed Rec-proteins RecN, RecJ and RecO are [72, 75, 74]. The data treatment that was established during this study evolved around these modes of movement and is described in chapter 2.10. A data DVD was attached, please refer to chapter 7.7 for video data.
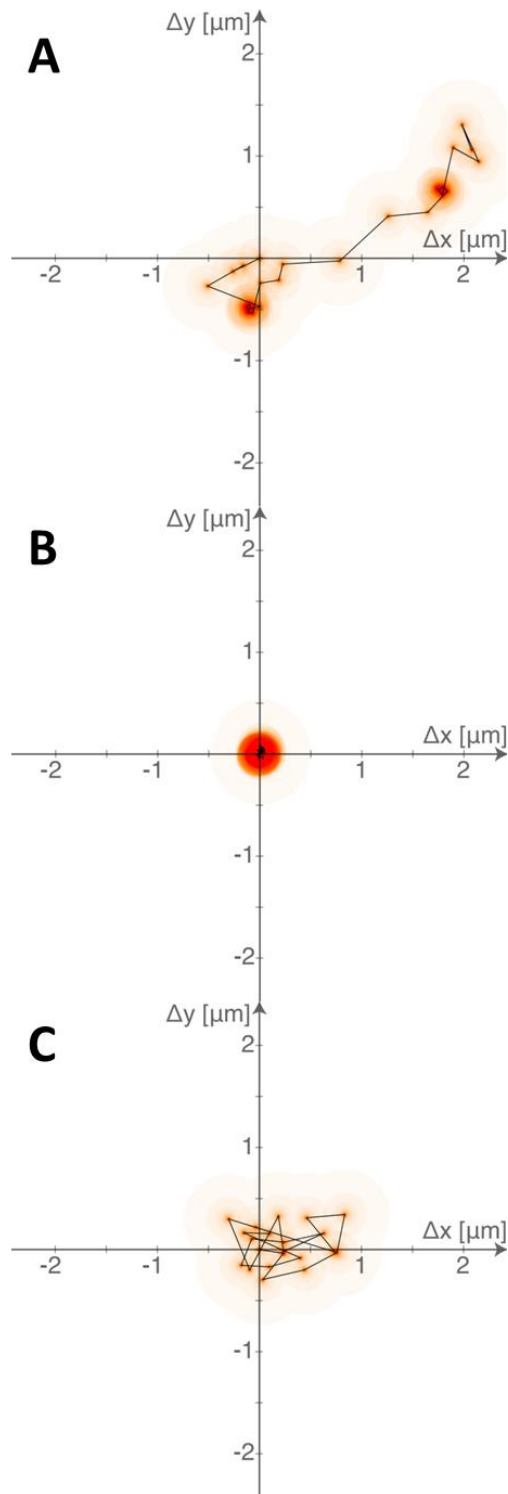
Figure 18 Distinct types of movement: Panel (A) shows a trajectory of a RecJ molecule in the absence of DSBs. The localization is emphasized with a red radius, the shades of red increase during stopping events. This trajectory moved over a distance of more than 2 μm. Panel (B) shows the trajectory of a RecN molecule in presence of DSBs, that did not move

over the detected time. Panel (C) shows the diffusion of a RecO molecule in absence of DSBs without a stopping event. All trajectories cover around 20 frames (~800 ms). This figure was edited using SMMtrack established by Prof. Dr. Schmitt **[71]**.

3.5.1.   Evaluation of tracking parameters: upper boundary of the allowed displacement and temporal tracking window

The documentation to the Matlab software u-Track (see chapter 2.10) [69] emphasizes the need to estimate the upper value of the allowed displacement, which corresponds to the frame to frame distance travelled by a molecule. This is a parameter that could be varied and should be evaluated with the video data. First, the threshold should be set at a point where there is no fragmentation of long tracks. Next, the maximal upper bound should be in a range that depicts physiological diffusion constants of biological macromolecules. Lastly, there should be no or few trajectories that are falsely connected by allowing a maximal displacement that is too large. I varied the upper bound for allowed displacement from one pixel to 15 pixels using the same data set, viz. "RecO-YFP uninduced". Accordingly, I estimated the minimal length of the trajectories using the RecO-YFP uninduced data set.
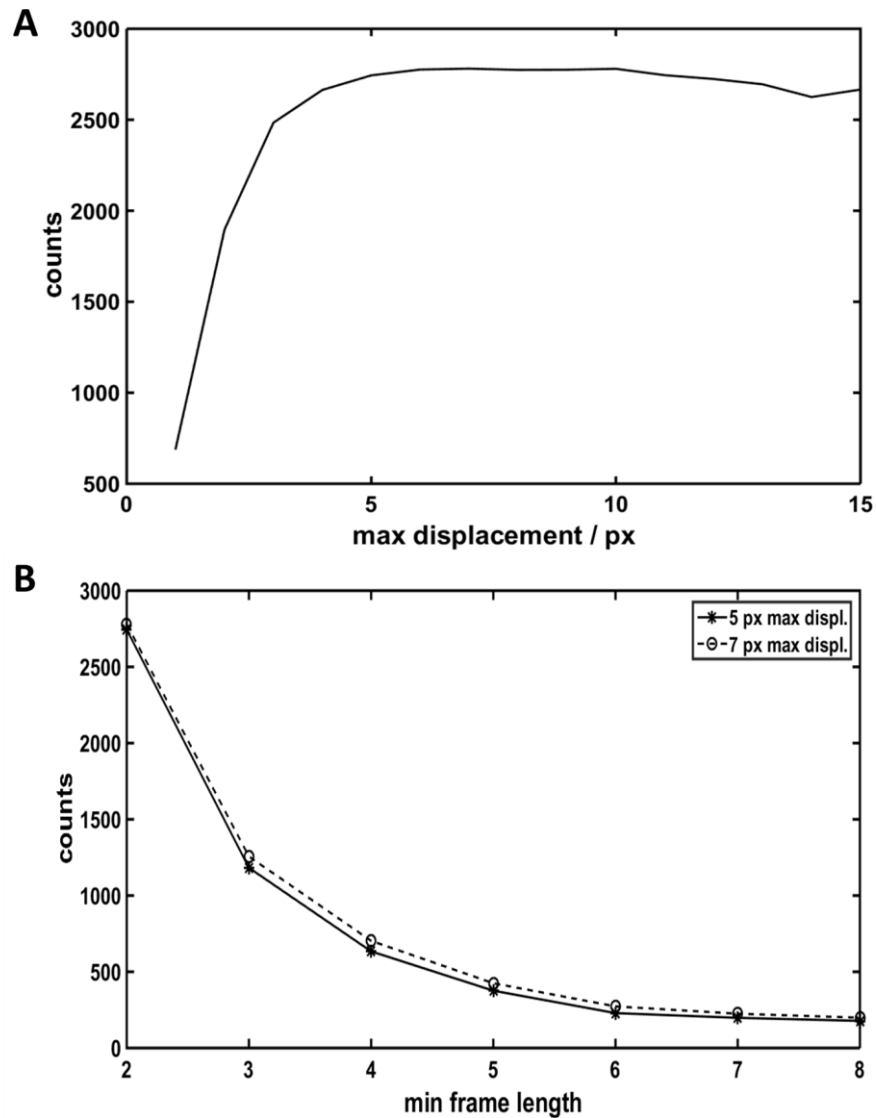
Figure 19 Number of trajectories: The figure shows the number of tracks as a function of the upper boundary that was allowed in the tracking with u-track in pixels (A) and the number of trajectories as a function of the minimal frame length. The curve in (A) reaches a plateau after five to six pixels; the curve in (B) after six to seven frames;

In Figure 19 (A) the absolute number of tracks resulting from the variation of the upper bound is plotted. There is a saturation effect reaching its maximum around six pixels. In Figure 19 (B) the absolute number of tracks is given as a function of the minimal length of a trajectory, starting with two frames ending with eight frames. This graph includes the

values for two upper bound values, 5 pixels and 7 pixels. This graph was used to estimate the temporal tracking window.

Figure 20 shows that the diffusion constant also increases when the upper boundary of the displacement increases from 1 pixel, which represents a displacement of around 100 nm, to the maximal displacement of 15 pixels (1500 nm). As explained in chapter 2.10, I applied a two population fit on the PDF of D_{app} that resulted in two values for D_{app}, D1 and D2 (one population was "freely" diffusing, the other was "thwarted by unspecific DNA interaction). Accordingly we calculated a fraction corresponding to the number of molecules with the different diffusion constant, A1 and A2. Again, we can see a plateau reached for D1 and D2 at around five to six pixels. For the weighted diffusion constant D_{weigh} a discrete increase at around five to six pixels is observed. The value of D at 5 pixels upper bound likely constitutes an outlier, with the value of six pixels being more reliable. (see Figure 20 (B)). In correlation to the diffusion constants, we can see a plateau for A1 and A2, being reached between five to six pixels. Again, the value at five pixels constitutes an outlier. A1 represents the population with the higher diffusion constant, with an asymptotic approximation to 60 % and A2 represents the slower population around with an asymptotic approximation to 40 %.

In the downstream data analysis the upper bound of allowed displacement for tracking was set to seven pixels. This value is located indisputably in the saturated regions of the above graphs and ensures that the whole population recruited for the analysis is represented in the data sets. The temporal window for the analysis was set to a minimum of five frames corresponding to 200 ms.
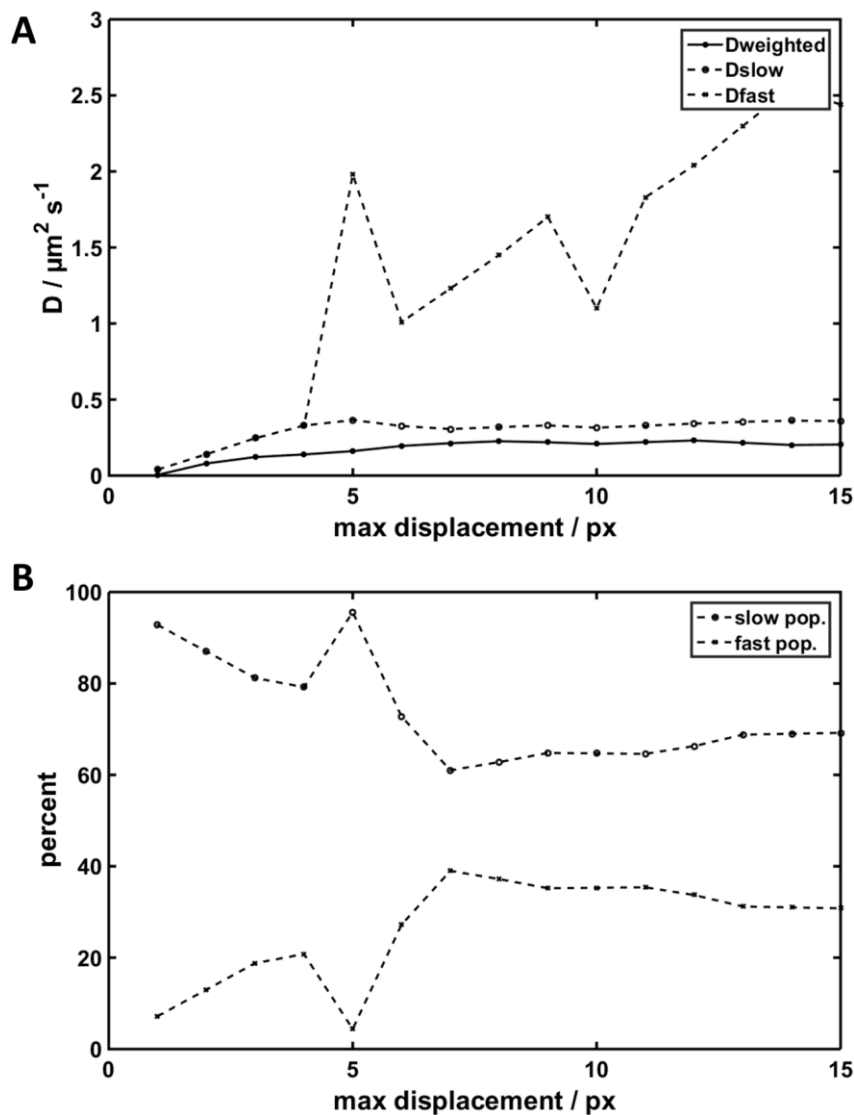
Figure 20 Estimated diffusion constants: The figure shows diffusion constants as a function of the upper boundary that was allowed in the tracking with u-track (A). D_{weigh} is the weighted diffusion constant, derived from the MSD of the whole population of trajectories. D1 is the diffusion constant derived from the fast population of molecules, D2 is the diffusion constant derived from the slow population of molecules (see chapter 2.10). (B) Shows the corresponding fraction size (A1 and A2)

3.5.2. Estimating the tracking reliability of u-track for the used data sets

The SMT software u-track, a Matlab implement, was bench-marked in a software contest [69]. Nonetheless I measured and confirmed the reliability of u-track with our microscopy data. Therefore, I imaged the wildtype strain PY79 not expressing any fusion protein or fluorescence marker under identical conditions as the strains containing fluorescent fusion proteins. In total, the software identified 4 tracks in the absence of DSBs and 6 tracks in the presence of DSBs with a minimum length of five frames in a comparable amount of streams (see Figure 21). Data sets from strains carrying a Rec-protein fusion or strains carrying a fusion proteins used as control resulted in a minimum of 500 trajectories (as stated below). This results in a maximum of 1% false positive tracks, probably even less than that.
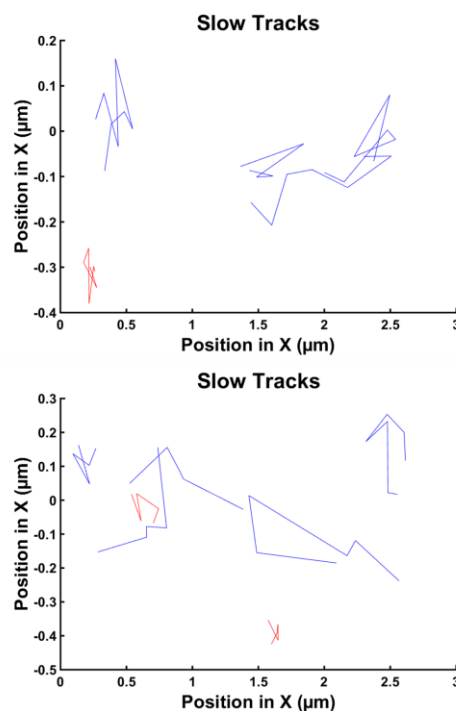

Figure 21 PY79: The figure shows the wildtype strain PY79 used to analyze the rate of false positive trajectories detected by the software used for the analysis. Upper panel shows tracks in absence of DSBs. Lower panel shows tracks in presence of DSBs. There were a one-digit number of tracks in streams treated similar to the data sets carrying

41

fluorophores. This would result in less than one percent false positive trajectories (see text below).

### 3.5.3. Evaluation of the instrumental and analytical localization error

To investigate the systematic error of our microscopy system applied for the experiments in this study, I imaged a strain expressing *comEB-gfp* from an ectopic locus (see chapter 3.1. and Table 6). ComEB-GFP has been described as a polar marker (see chapter 1.4.2. and [27]). It is a membrane associated protein with no predicted transmembrane helices [76]. A total of 593 tracks with a minimal frame length of five frames was acquired and used for downstream analysis. From these 593 tracks, the MSD curves over the first 4 frames were plotted and linear fits to these curves were applied. From these straights the apparent diffusion constant was derived (as described in chapter2.10.) and the probability density function (PDF) of these diffusion constants was plotted as a histogram (see Figure 22). We assumed three distinct populations on the single molecule level, namely a "free" diffusing cytosolic fraction; a membrane-associated diffusive fraction; and a static membrane associated fraction that is locked at the poles. In Figure 22 (B) a three population fit was therefore applied to the occurring distribution, resulting in three mean values for the apparent diffusion constant (D1-D2) and the three fraction values (A1-A3) derived from the integral of these curves: $D1 = 0.0089 \ \mu m^2 \ s^{-1} \pm SD$ with 23.4%; $D2 = 0.236 \ \mu m^2 \ s^{-1} \pm SD$ with $D2 = 53.8\%$; $D3 = 0.8 \ \mu m^2 \ s^{-1} \pm SD$ with $A3 = 22.8\%$. We compared this fit with the distribution of the PDF for GFP-MreB, where the expectation for the above mentioned three populations is discussed [77, 78] [and Dr. Christian Reimold personal communication] (see chapter 7.4). The slow fraction in this experiment refers to the static polar foci in the epifluorescence micrographs above (see Figure 15).
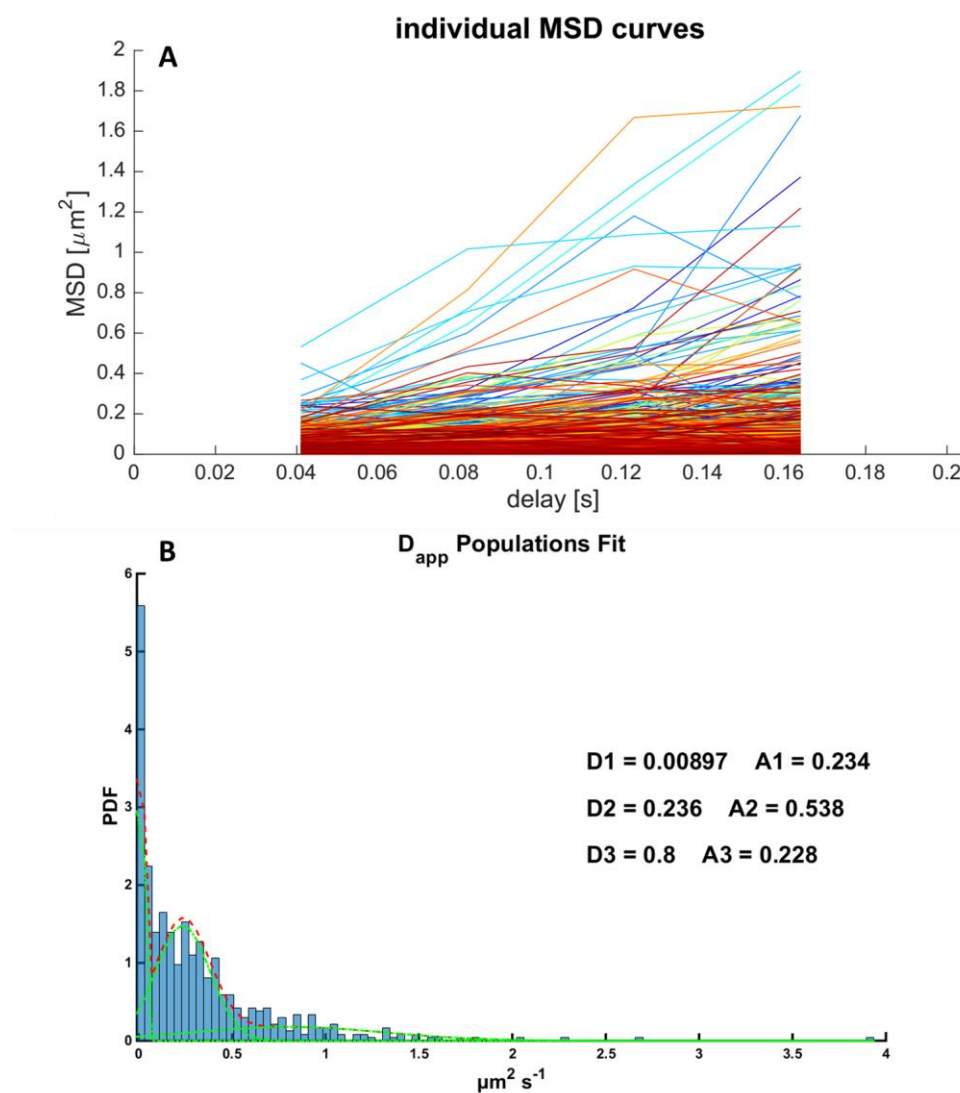
Figure 22 Single molecule tracking of ComEB-GFP: The figure shows in (A) all MSD curves over the first four time frames that were used for linear fits. Apparent diffusion constants derived from these fits are shown in the histogram in panel (B). Three populations were assumed to constitute the data distribution of this membrane-associated protein (see text below). D1, D2 and D3 refer to the individual apparent diffusion constants and A1-A3 refers to the fraction of trajectories in the three populations.

No movement was observed in these micrographs and we assumed the slow movement observed on the single molecule level, with D1 = 0.0089 $\mu m^2 s^{-1}$, refers to the offset and localization error of the applied setup, i.e. the microscope and camera. So we selected the MSD curves that refer to D1 (see Figure 23), cured for positive y-intercepts and a minimal

fit quality of R = X. From these MSD curves we determined the y-intercepts that in turn correspond to the localization error $4\sigma_{loc}{}^2$ (see Figure 23 (B) and the third equation in chapter 2.10). Finally, we used the mean localization error to estimate the localization precision of our setup σ_{loc} = 20 nm ± 8 nm. A data DVD was attached, please refer to chapter 7.7 for video data.
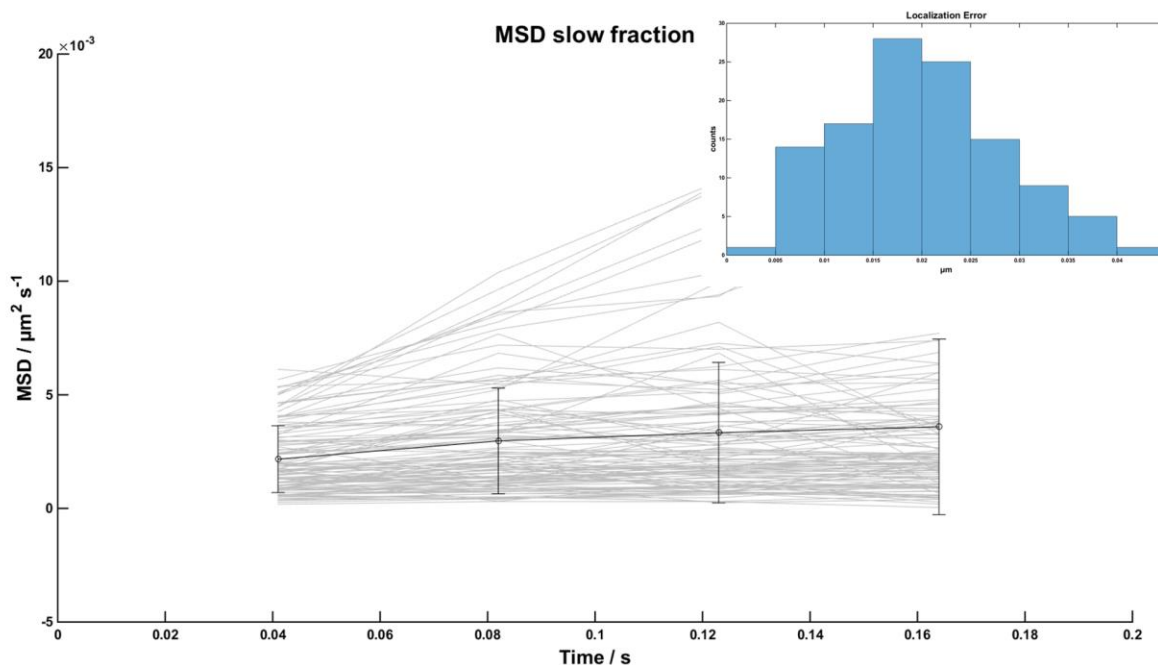

Figure 23 Estimation of the localization precision: The figure shows the MSD curves of slowly diffusing ComEB-GFP molecules that were used to calculate the localization error σ_{loc}. The inlet shows the distribution of σ_{loc}. The mean value of the localization error was 20 nm ± 8 nm.

3.5.4. Estimation of static DNA-bound single molecule trajectories

As described above (see chapter 3.5), the Rec-proteins showed a distinct behavior by pausing for some time during an acquired track or even for the whole track. These pausing events implied a minimal movement or wiggling likely resulting from the movement of DNA. To distinguish these nearly static tracks from the whole population, I performed SMT

experiments (streams with 40 ms per frame) using a strain carrying a *tetO*-array in proximity to the origin of replication of the chromosome, *oriC*, and a TetR-YFP fusion under a constitutive promotor. This strain was kindly provided by Dr. Katrin Schenk from the Graumann Lab. The TetR-repressor has a high binding affinity to its operator with 10^{-9} M [79]. We assumed that the TetR-YFP fusion must exist as two distinctly behaving populations, one bound to the operator displaying a static DNA-bound state corresponding to this high affinity, and in a cytosolic "free" diffusive state. I acquired 1336 trajectories with a minimal length of five frames. The MSD curves over the first four frames were plotted and linear fitted (see Figure 24(A)). Prior to calculating the D_{app} (as described in chapter 2.10.); the localization error determined in chapter 3.5.3 was subtracted from the MSD values. The PDF of D_{app} is given in the histogram shown in Figure 24 (B). Here, a two population fit was applied, which showed a static DNA-bound population with $D1 = 0.0408\ \mu m^2\ s^{-1} \pm SD$ with a fraction size of 49 % and a population of freely diffusing molecules with $D2 = 0.406\ \mu m^2\ s^{-1} \pm SD$ and a fraction of 51 %.

Thereby, we concluded that the threshold for a molecule that binds with high affinity to a specific DNA stretch to fulfill its biological function likely has a comparable diffusion coefficient as the slowly diffusing TetR molecules. Finally, we included the standard deviation of this diffusion constant to allow for a tolerance leading to a $D_{threshold}$ of 0.07 $\mu m^2\ s^{-1}$. A data DVD was attached, please refer to chapter 7.7 for video data.
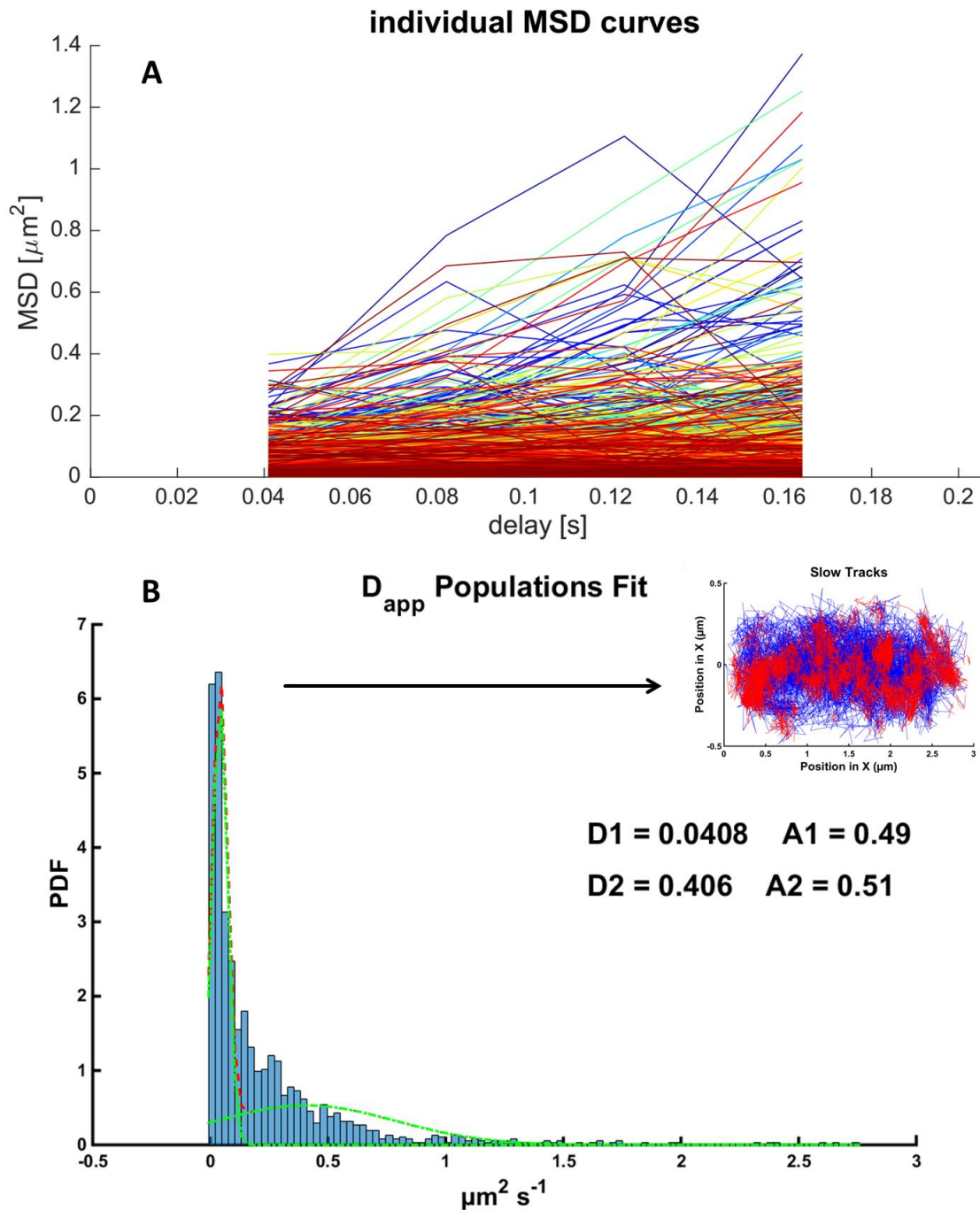
Figure 24 SMT of *oriC-tetO*-TeTR-YFP: This figure shows (A) All MSD curves over the first four time frames that were used for linear fits. (B) Histogram of the apparent diffusion constants derived from these straights. The data were fitted under the assumption that there are two populations of TetR-YFP, a fraction of bound repressor and a fraction of diffusive molecules. D_{app} are stated as D1 and D2. Tracks with the slow diffusion constant are shown in red in the inlet.

3.5.5. SMT of RecN, both in absence and presence of chemically induced DSBs

RecN is a member of the group of SMC-like proteins and one of the first proteins recruited to the site of a DSB. In the absence of chromosomal lesions it is dispersed over the chromosome [49]. Still tracks with a D_{app} below the threshold of static DNA-bound occur rarely, but in this case they showed slow movement (i. e. below $D_{threshold}$ (see Figure 26 B).

The localization of RecN alters when DSBs are introduced in cells. In the working model, RecN has the role of a sensor for DSBs [53]. When the static DNA-bound tracks are visualized a localization pattern shows hotspots of trajectories (see Figure 25 D). This is similar to the change in the localization pattern that we know from epifluorescence micrographs [49].
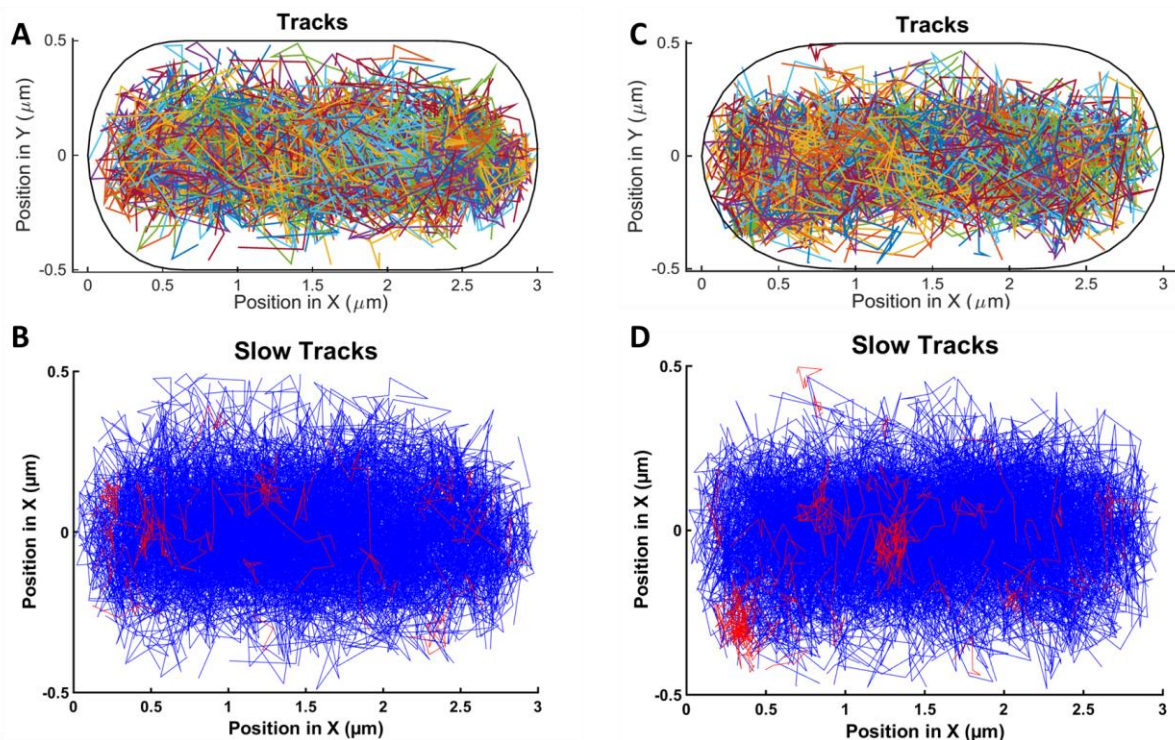

Figure 25 SMT of RecN-YFP: This figure shows all acquired track in a normalized cell in absence (A) and presence (C) of DSBs. In the lower panels the tracks are sorted, blue tracks

exceed the threshold for DNA-binding, red tracks are below the threshold. (B) Shows the uninduced state and (D) the state after the induction of DSBs.

From these (see Figure 26 A inlet), the MSD over the first four time frames was plotted and the apparent diffusion constant was calculated. The PDF of the D_{app} was plotted and a two population fit was performed (see Figure 26 A). The broad distribution of diffusion constants could be described with two populations superpositioned in the distribution given by the experimental data, as there were a high peak with a mean in the lower D_{app}-range and a higher mean with broad shoulder in the higher D_{app}-range; logically these would refer to a slower population "thwarted" by unspecific DNA interaction and a faster "free" diffusing population. These fits resulted in two mean values for the D_{app} (D1 and D2) as well as the respective fraction by calculating the integral of the curves (A1 and A2). In the uninduced condition we used for the calculation of D_{app} 1215 trajectories with the minimal length of five frames. For the induced condition we acquired 1238 trajectories (see Figure 26 B).

In the absence of DSBs this resulted in: $D1 = 0.303\ \mu m^2\,s^{-1} \pm SD$ with $A1 = 75.2\,\%$; $D2 = 0.962\ \mu m^2\,s^{-1} \pm SD$ with $A2 = 24.8\,\%$ (see Figure 26 A). There were $3.5\,\%$ of the observed RecN molecules that showed a D_{app} below the threshold derived from the DNA-bound TetR repressor. In the presence of DSBs this resulted in: $D1 = 0.283\ \mu m^2\,s^{-1} \pm SD$ with $A1 = 76.4\,\%$; $D2 = 0.854\ \mu m^2\,s^{-1} \pm SD$ with $A2 = 23.6\,\%$ (see Figure 26 B). There were $5.3\,\%$ of the observed RecN molecules that showed a D_{app} below the threshold derived from the DNA-bound TetR repressor. This constituted an increase in static DNA-bound molecules of 51%. A χ^2-test between proportions was performed to determine whether there was a significant difference in relative amount of static tracks. The χ^2-statistics rejected the nil-hypothesis significantly with $p = 0.0149$. The increase of static tracks was significant. A data DVD was attached, please refer to chapter 7.7 for video data.
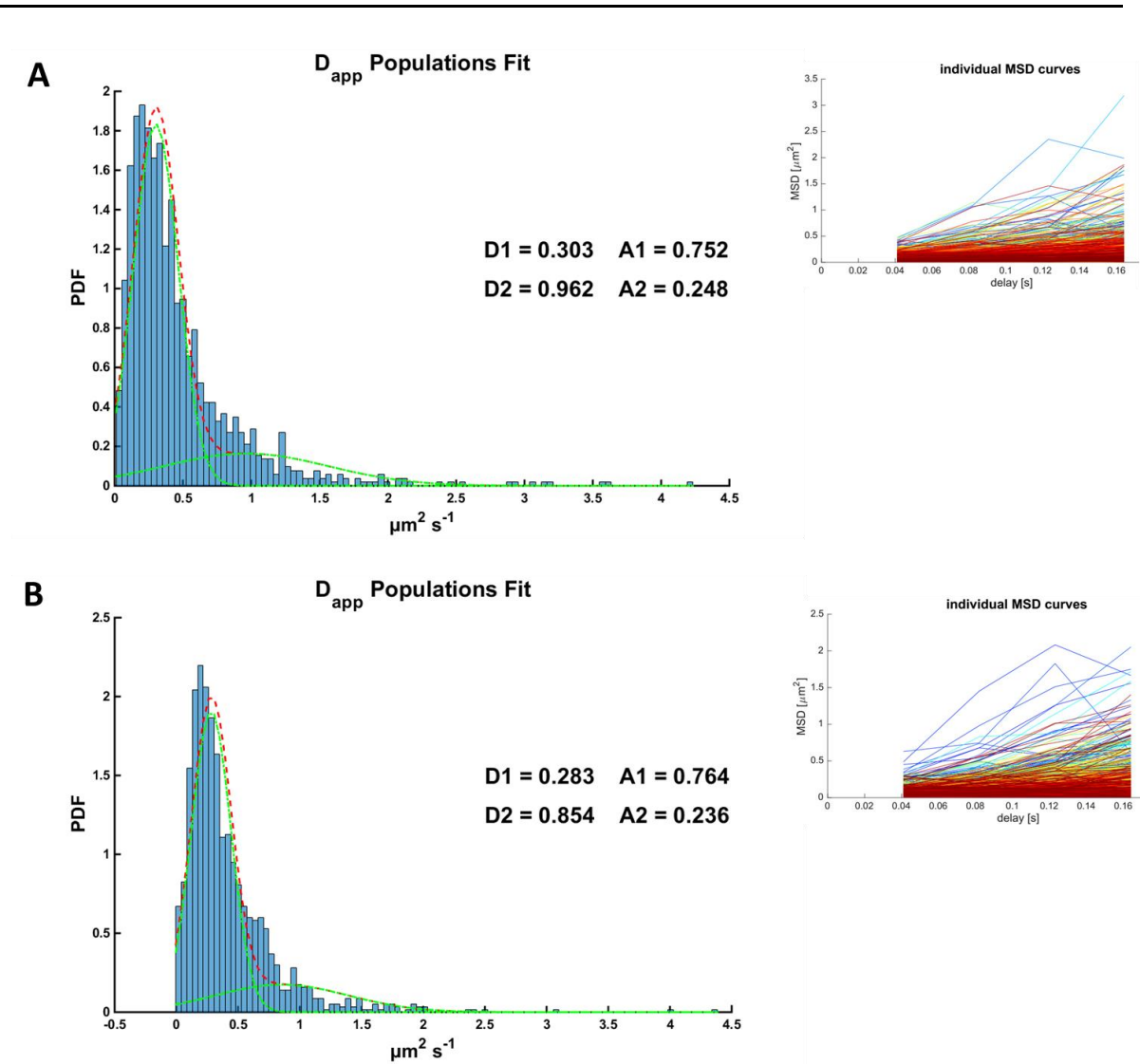
Figure 26 SMT RecN-YFP: The figure shows the PDF of the Dapp for the uninduced condition (A) and the induced condition (B). The inlet shows the MDS curves from that the diffusion coefficient was derived. D1 and A1 are the diffusion constant and fraction of the slower population; D2 and A2 are the diffusion constant and fraction of the faster population. The standard deviation is shown in tTable 1; X-axis in $\mu m^2\, s^{-1}$; Y-axis shows the PDF.

3.5.6.   SMT of RecJ in the absence and presence of chemically induced DSBs

The exonuclease RecJ, which processes dsDNA resulting in stretches of ssDNA, shows a high affinity to the C-terminus of SsbA. It has been shown that RecJ is bound to the replication machinery via SsbA (see chapter 1.6.2.) [59, 57]. When the static DNA-bound tracks are visualized a localization pattern shows hotspots of trajectories that are either at midcell or at the ¼ or ¾ positions. The replication fork would be either at midcell or in the case of a replication start prior to cell division at the ¼ or ¾ positions (see Figure 27 B).

In the presence of DSBs the exonuclease is presumably released from the replisome. The substrate is 3'-ends of dsDNA [57]. In the current working model it would be recruited to DSB sites via the sensor protein RecN. In epifluorescence micrographs this results in localization foci away from the replication fork [71]. Again SMT does show comparable localization patterns. The "hotspots" of static DNA-bound trajectories increased in number and were dispersed over the nucleoid (see Figure 27 D).

In the uninduced condition we used for the estimation of D_{app} 604 trajectories with the minimal length of five frames and in the induced condition we used 523 tracks. From these the MSD over the first four time frames was plotted and the apparent diffusion constant was calculated. The PDF of the D_{app} was plotted and a two population fit was performed, we argued that two distinct behaving subpopulations constitute the given distribution: One population being bound (either at the replication fork or on DNA) and a free population. These fits resulted in two mean values for the diffusion constants (D1 and D2) as well as the respective fractions by calculating the integral of the curves (A1 and A2). (Please refer to Figure 30 in the appendix)

In the absence of DSBs we calculated: $D1 = 0.247\ \mu m^2\ s^{-1} \pm SD$ with $A1 = 38.7\%$ and $D2 = 1.02\ \mu m^2\ s^1 \pm SD$ with $A2 = 46.2\%$. There were 17.4% of the observed RecJ molecules below the threshold of the DNA-bound state. In the presence of DSBs we calculated:

D1 = 0.258 µm^2 s^{-1} ±SD with A1 = 94.1% and D2 = 1.5 µm^2 s^{-1} ±SD with A2 = 5.9%. There were 21.8% of the observed RecJ molecules below the threshold of the DNA-bound state. A χ^2-test between proportions was performed to determine whether there was a significant difference in relative amount of static tracks. The χ^2-statistics rejected the nil-hypothesis significantly with p = 0.037. The increase of static tracks was significant.
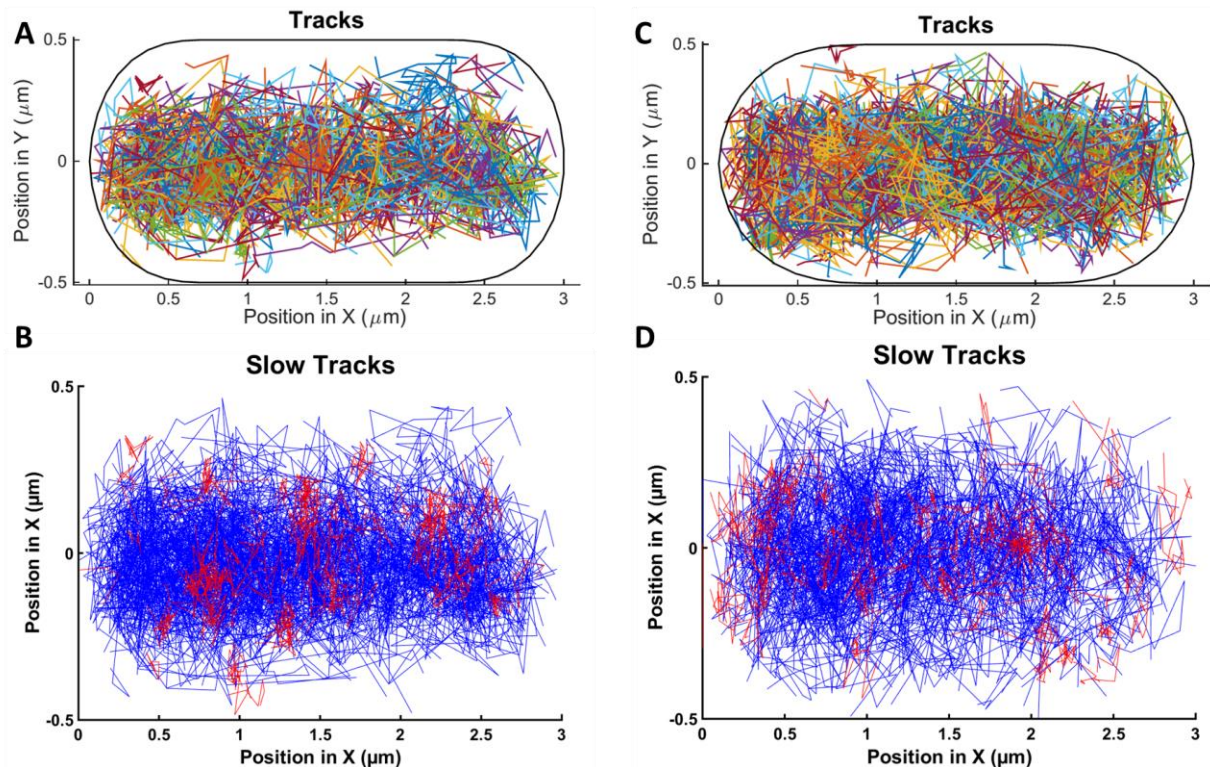
Figure 27 SMT of RecJ-YFP: This figure shows all acquired track in a normalized cell in absence (A) and presence (C) of DSBs. In the lower panels the tracks are sorted, blue tracks exceed the threshold for DNA-binding, red tracks are below the threshold. (B) Shows the uninduced state and (D) the state after the induction of DSB.

3.5.7. SMT for RecO in the absence and presence of chemically induced DSBs

The Rec-protein RecO is part of the RecA loading complex (see 1.6.2) [50]. The fluorescence signal in the absence of DSBs is too weak to show any localization in

epifluorescence micrographs. The distribution in Figure 28 (A) suggested a nucleoid-confined localization. Trajectories with a D_{app} below the threshold of static DNA-bound occur, but in this case they did not persist over time at a given position, but moved slowly (see Figure 28B).

In the presence of DSBs RecO forms foci (see chapter 1.6.2). Those foci, predominantly one focus and rarely two foci, localize on the nucleoid. Again SMT gives a rough resemblance of the localization pattern seen in epifluorescence pictures, when we visualize the tracks below the threshold for D_{app} of statically DNA-bound molecules (see Figure 28 D).

In the uninduced condition we used 425 trajectories with the minimal length of five frames for the calculation of D_{app}. From these the MSD over the first four time frames was plotted and the apparent diffusion constant was calculated. The PDF of the D_{app} was plotted and a two population fit was performed. For the induced condition we acquired 523 trajectories. The broad distribution of diffusion constants could be described with two populations superpositioned in the distribution given by the experimental data, as there was a high peak with a mean in the lower D_{app}-range and a higher mean with broad shoulder in the higher D_{app}-range; logically these would refer to a slower population "thwarted" by unspecific DNA interaction and a faster "free" diffusing population. These fits resulted in two mean values for the D_{app} (D1 and D2) as well as the respective fraction by calculating the integral of the curves (A1 and A2). (Please refer to Figure 31 in the appendix)

In the absence of DSBs we calculated: $D1 = 0.349\ \mu m^2\,s^{-1} \pm SD$ with $A1 = 71.4\%$, $D2 = 0.946\ \mu m^2\,s^{-1} \pm SD$ with $A2 = 28.6\%$. There were to 4% of the observed RecO molecules with a D_{app} below the threshold of the DNA-bound state. In the presence of DSBs we calculated: $D1 = 0.54\ \mu m^2\,s^{-1} \pm SD$ with $A1 = 91.9\%$, $D2 = 1.78\ \mu m^2\,s^{-1} \pm SD$ with $A2 = 8.1\%$. There were to 6.8% of the observed RecO molecules with a D_{app} below the threshold of the DNA-bound state. This constitutes an increase of 70%. A χ^2-test between

proportions was performed to determine whether there was a significant difference in relative amount of static tracks. The χ^2-statistics rejected the nil-hypothesis significantly with p = 0.018. The increase of static tracks was significant. A data DVD was attached, please refer to chapter 7.7 for video data.
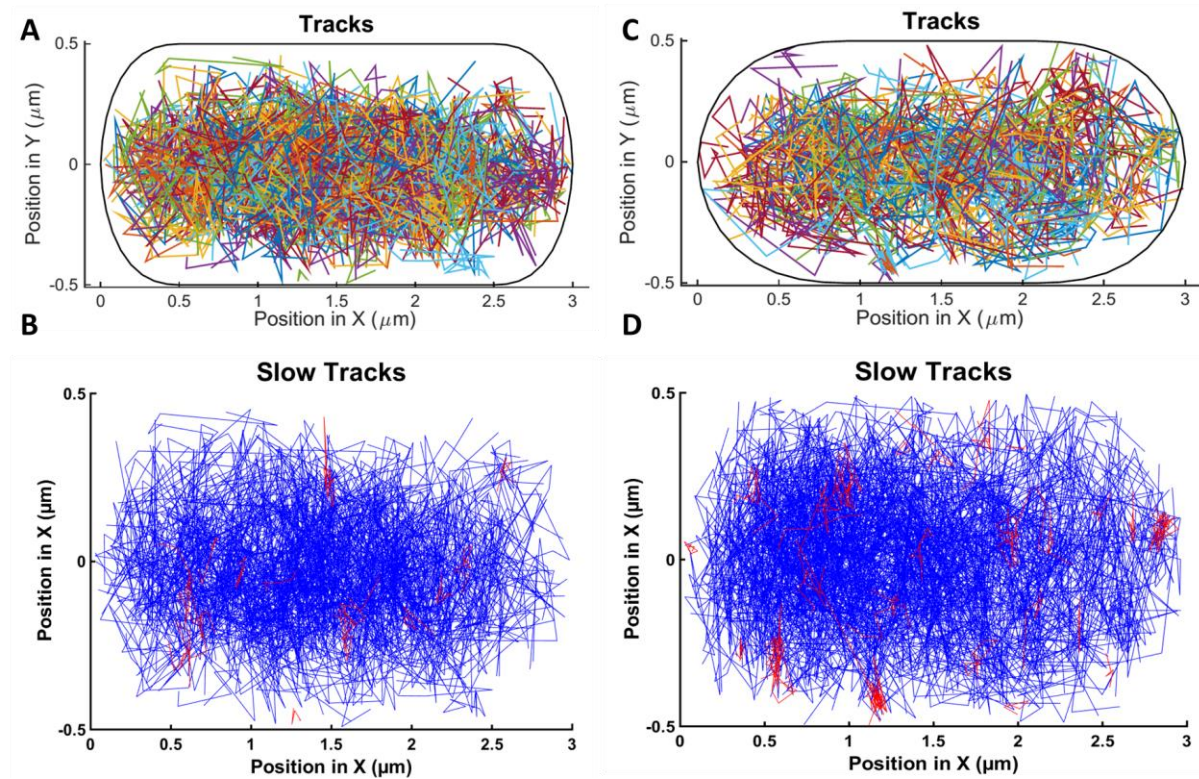
Figure 28 SMT of RecO-YFP: This figure shows all acquired track in a normalized cell in absence (A) and presence (C) of DSBs. In the lower panels the tracks are sorted, blue tracks exceed the threshold for DNA-binding, red tracks are below the threshold. (B) Shows the uninduced state and (D) the state after the induction of DSBs.

3.5.8.   SMT of PfkA-GFP in the absence and presence of DSBs

PfkA is part of a hetero-tetramer, the phosphofructokinase complex. Its molecular weight is comparable to the Rec-proteins investigated before, and literature data suggest the complex is cytosolic [80]. Nicking the chromosome could have had an influence on the viscosity and composure of the cytosol. To investigate this influence on D_{app}, a strain

carrying the PfkA-GFP fusion was utilized. The fusion was under control of the original promotor. The cells were grown to exponential phase, stream acquisition and data treatment was the same as for the strains carrying the Rec-protein fusions.

The overlay of all tracks in a standardized cell outline suggested less confinement to central regions of the cell (see Figure 32 A and C in the appendix). Just a few trajectories diffused with the rate below DNA-binding and they were not confined to hotspots but moved over the area with a slow D_{app} (red tracks in Figure 32 B and D). There was no difference observable between the uninduced and induced experimental conditions.

The MSD curves over the first four time frames were plotted and the apparent diffusion constant was calculated. The PDF of the D_{app} was plotted and a two population fit was performed (see Figure 33 A in the appendix); we applied a two population fit for two reasons. First, to achieve comparable values and second, the distribution suggest here as well that there is superposition of two populations, that might be monomers vs. homotetramers. These fits resulted in two mean values for the D_{app} (D1 and D2) as well as the respective fraction by calculating the integral of the curves (A1 and A2) (Please refer to Figure 33 B in the appendix). In the uninduced condition we used for the calculation 656 trajectories with the minimal length of five frames. For the induced condition we acquired 420 trajectories.

In the absence of DSBs we calculated: $D1 = 0.512\,\mu m^2\,s^{-1} \pm SD$ with $A1 = 84.3\%$, $D2 = 1.21\,\mu m^2\,s^{-1} \pm SD$ with $A2 = 15.7\%$. There were 1.1% of the observed PfkA molecules with D_{app} below the threshold of the DNA-bound state. In the presence of DSBs we calculated $D1 = 0.584\,\mu m^2\,s^{-1} \pm SD$ with $A1 = 77\%$, $D2 = 1.29\,\mu m^2\,s^{-1} \pm SD$ with $A2 = 23\%$. There were 1.43% of the observed PfkA molecules with D_{app} below the threshold of the DNA-bound state. This constitutes an increase of 14%. A χ^2-test between proportions was performed to determine whether there was a significant difference in relative amount of static tracks. The χ^2-statistics did not reject the nil-hypothesis with $p = 0.6$. The increase of

static tracks was not significant. A data DVD was attached, please refer to chapter 7.7 for video data.

Table 1 D_{app}, fractions size and relative amount of static tracks from Rec-Proteins

Protein	Condition	$D_{app}[\mu m^2 s^{-1}]$ / Fraction [%]		Static Tracks	Increase /p-value*
RecN	-MMC	D1=0.3±0.16 A1=75.2	D2=0.96±0.6 A2=24.8	3.5%	**51%**
	+MMC	D1=0.28±0.15 A1=76.4	D2=0.85±0.5 A2=23.6	5.3%	**0.015**
RecJ	-MMC	D1=0.25±0.19 A1=74.1	D2=1.02±0.69 A2=25.9	17.4%	**25%**
	+MMC	D1=0.29±0.24 A1=94	D2=1.05±0.69 A2=5.9	21.8%	**0.037**
RecO	-MMC	D1=0.35±0.18 A1=71.4	D2=0.95±0.48 A2=28.6	4%	**71%**
	+MMC	D1=0.54±0.35 A1=91.9	D2=1.78±1.05 A2=8.1	6.8%	**0.018**

* Critical p-value for a χ^2-test: p = 0.05 to reject the null-hypothesis that the increase was not significant.

Table 2 D_{app}, fractions size and relative amount of static tracks of controls

Protein	Condition	$D_{app}[\mu m^2 s^{-1}]$ / Fraction [%]			Static Tracks	Increase/ (p-value)*
TetR-YFP	-MMC	D1=0.048±0.03 A1=49	D2=0.4±0.38 A2=51%		49%	N/A
ComEB-YFP	-MMC	D1=0.009 ±0.015 A1=23.4	D2=0.236 ±0.14 A2=53.8	D3=0.81± 0.51 A3=22.8		N/A
GFP-MreB	-MMC	D1=0.067 ±0.05 A1=46.1	D2=0.37± 0.21 A2=46.7	D3=1.02± 0.61 A3=7.2	46.1	N/A
PfkA-GFP	-MMC	D1=0.51±0.26 A1=84.3	D2=1.21±0.59 A2= 15.7		1.1%	25%
PfkA-GFP	+MMC	D1=0.584±0.29 A1=77	D2=1.29±0.56 A2=23		1.4%	0.6

* Critical p-value for a χ^2-test: p = 0.05 to reject the null-hypothesis that the increase was not significant.

4. Discussion

4.1. ComEB might be a recruiting factor for polar localization of ComGA

As shown in chapter 3.1 ComGA fails to localize to the cell pole in distinct foci in a *comEB* deletion strain. ComGA is a putative traffic ATPase and a key player in the competence machinery in *Bacillus subtilis* [81, 5, 23]. Up to now, it has been subject to discussion whether the second ORF of the late competence operon *comE, ComEB,* is even expressed when the ComG-regulon is turned on [25, 26, 18]. This work shows that the *comEB* mutant displays a non-transformable phenotype. The ORF *comEB* is expressed and the fusion protein ComEB-YFP under the control of the native promotor also localizes to the cell pole. In the presence of ComGA-CFP, they both localize to the pole and the overlaid channels merge in color, meaning the two foci are in a diffraction limited proximity (d ≤ λ/2 = 250 nm). However, this co-localization is not exclusive: there are substantial numbers of foci of both ComEB and ComGA that do not localize with its assumed partner (see Figure 14 and chapter 3.1). But in every analyzed cell there was at least one focus, if not several, that showed co-localization. This is well in the line with the fact that a functional competence machinery only localizes at a single cell pole and that the other signals are excess clusters [23, 24].

The question to be clarified remained whether ComEB possesses an intrinsic property to localize to any physical or biochemical feature provided by the pole. There are numerous examples of polar localizations, DivIVA being the most prominent one [73]. There are also several reasons why these proteins might act in such a way: DivIVA is attracted to the negative curvature of the membrane [73], others are potentially attracted to the specific lipid composition at the cell pole [82]. I wanted to ensure that ComEB recognizes the pole by itself and that it is not captured by a protein-protein interaction. To this end, the gene was put under transcriptional control of the xylose promotor and expressed in exponential

growth, when the expression of ComK is tightly repressed. I still observed polar localization of ComEB-YFP (see chapter 3.1 and Figure 14).

Furthermore, ComEB finds the pole even in the absence of any *B. subtilis*-specific interaction partner, namely when expressed in *Escherichia coli*. I used a pET-plasmid with *comEB* under the control of the T7-expression system. These systems are usually high copy plasmids with strongly transcribed expression systems, so one might argue the localization pattern observed in Figure 15 is an overexpression artefact. However, the expression of ComEB was induced with low concentrations of IPTG, (the T7-system has more a neither/nor response then a linear dose-dependent induction increase) and only for a short time span before image data were acquired. Secondly, precipitation due to overexpression would result in localization at the poles and not at the septum. The fluorescence signals in Figure 15 occur on every pole and every septum. The signals at the septa are especially weak suggesting low local concentrations of ComEB. That indicates to me that we in fact observe the intrinsic affinity of ComEB to localize to bacterial cell pole and no overproduction artefact.

The mechanics for this behavior are unknown so far as the amino acid sequence of ComEB does allow the prediction of domains that suggest membrane association. In fact there are no predicted domains [83], but a zinc binding domain that contains a putative active site for a hydrolase activity, or more specific dCMP deaminase [76, 83]. Consequently, a mechanism comparable to the activity of DivIVA could be excluded [73]. The second ORF of the *comE*-operon is conserved among firmicutes [76]. The *Bacillus cereus*-group harbors the homologue with high identity and even phylogenetically further distinct *firmicutes* species such as the thermophile *Geobacillus denitrificans* carry the gene [84]. The apparent conservation due to selective pressure on this gene indicates that *comEB* might indeed have a relevant function likely linking DNA uptake to the nucleotide metabolism.

4.1. Single Molecule Microscopy and Tracking (SMT) of Rec-proteins reveal a highly dynamic behavior

Evaluation of the tracking software

We established a new technique in our lab. Imaging is core skill in our group, so establishing SMM was feasible and needs no further discussion. We used u-track as detection and tracking tool [69]. Before, I tested the parameters set for tracking (see chapter 3.5.1), so the testing was performed with the same data set to ensure comparability.

We determined an upper bound of displacement that seems high with 7 pixels (equivalent to 700 µm), but we observed an asymptotical saturation effect on several levels: The mean apparent diffusion constant D_{weigh}, D_{app} for the two subpopulations D1 and D2, their corresponding fractions A1 and A2, and the total number of tracks showed an asymptotic approximation that was saturated at seven pixels. The data treatment methods used in SMT is based on total populations, so we carefully made sure to detect this total population. There is and will always be the discussion about false positive detection and tracking, but Jaqaman *et al.* stated that improving largely the quantity of detected tracks, a given lack of quality in some of these tracks is negligible [69]. To this end, I tested u-track with the wild type not carrying a fluorophore, since the background fluorescence in the Gram-positive bacterium *B. subtilis* is in many microscopy experiments an issue. The result with 0.5 % to 1 % false positive detected trajectories is insignificant. But there might be need to evaluate these settings for different protein classes and microscope settings. Also, it should be thought about more conservative settings when the molecules are tracked with fluorophores having a higher photon yields.

Dynamics of Rec-proteins on the single molecule level in real-time

To analyze the dynamics of Rec-proteins, I started with RecN that is thought to be the sensor of double strand breaks (DSBs), because of its high affinity to 3' ssDNA ends *in vitro*

[47, 53]. I initially employed epifluorescence stream acquisition with low intensity continuous laser illumination and 100 ms exposure times. After 30 min incubation with the DNA damaging agent MMC, the formation of DSBs occurred. We then observed the formation of focal RecN-YFP assemblies. These assemblies persisted for a few second, with a mean of 2.5±0.6 s, before they disassembled. Rarely, the assembly of a new focus in proximity could be observed. In a single case the movement of one of these foci was observed (see chapter 3.4). In eukaryotes, the formation of repair centers (RCs) was described [48]. In these RCs multiple strands get coordinated, meaning the dsDNA-ends are brought to proximity [48]. So the formation of RCs is discussed for bacteria as well [53], because this would be a simple way to spatially organize the loose DNA strands. These preliminary epifluorescence results questioned whether these RCs exist at all or might be very short lived clusters. Single molecule microscopy and tracking (SMT), followed by a data treatment and analysis was established in the Graumann lab (thanks to Dr. Thomas Rösch) and was employed for a more detailed insight of the behavior of the Rec-proteins. At the beginning we needed to consider two features:

On one hand, due to the fact that we excite and detect single molecules we achieved subpixel resolution, as in every pointillism microscopy technique [72, 85, 86]. When reaching resolutions notably below 100 nm (i.e. one pixel) we had to determine the localization error. Therefore, we used a data set that I acquired from the membrane-associated protein ComEB-GFP (see detailed in chapter 3.5.3). By plotting the probability density function of D_{app} (for the calculation of D_{app} see chapter 2.10.), we could describe three different populations, one that is freely diffusing, a second one that is membrane attached and a third one that is bound to static polar clusters (see chapters 1.4.2 and 3.3). In this static fraction there was so little movement that we concluded to calculate the localization offset ($4\sigma^2$) from this and estimated the localization error with $\sigma_{loc} = 20$ nm±8nm (refer to chapter 3.5.3). This is a range that is well comparable to the values stated in other publications [86, 85, 87]. This localization error was used to cure all other values for D_{app} that we further calculated.

On the other hand, we observed binding events of the Rec-proteins: they stopped in their movement and after some time moved on (see Figure 18). We had to define exactly what a binding event really was bearing in mind that the population of which we derived the localization error has a D_{app} of 0.009 $\mu m^2 s^{-1}$. To this end, we employed a data set acquired from the *tet*-repressor TetR fused to YFP. The binding affinity of TetR to its operator with 1 nM is very high [88]. From this, we assumed that there must be two populations in a strain expressing this fusion, a static one bound to the operator DNA showing the movement of the DNA and a "freely" diffusing one. We calculated $D_{app} = 0.04$ $\mu m^2 s^{-1}$ for the slowest population (see chapter 3.5.4), and set this diffusion coefficient as threshold value in which we also included the standard deviation leading to $D_{threshold} = 0.07$ $\mu m^2 s^{-1}$, which we finally used to analyze all other Rec-proteins. We visualized trajectories of individual molecules below this threshold and then observed the static DNA-bound subpopulation in the total amount of tracks (the figures are shown in the result section and the appendix).

We now have two levels of results: first, we have subpopulations of proteins that diffuse with different apparent diffusion constants, probably due to the fact that the slower fraction is "thwarted" by unspecific interactions, comparable to LacI [75], while the other is diffusing "freely". And second, we can directly observe and quantify the number of molecules that are DNA bound. When we now compare the two experimental conditions we can see a much clearer image of the presynaptic DSB repair in *B. subtilis* (see Table 1 as overview). We show that RecN, RecO and RecJ proteins scan through the entire *B. subtilis* genome during exponential growth, searching for DNA lesions. Upon the induction of DSBs we observe a change in this behavior: RecN movement slows down when we compared the D_{app} experimental conditions (see Table 1).

Furthermore we observed an increase of static DNA-bound molecules from 3.5 % (no DSBs) to 5.3 % (with DSBs). This constitutes an increase of 51 %, while the majority of the molecules still keep on scanning the nucleoid. This can be compared to Liao *et al.* [89]. Their work on MutS, the mismatch recognition protein in MMR (see chapter 1.6), suggests

that the local density of MutS is increased by interaction with the replisome, where mismatches might be incorporated in newly synthesized DNA. This would circumvent the fact that 3D diffusion alone is not efficient enough to recruit MutS quickly to DNA-bulks. Compared to my work, this would suggest that the reaction (of binding dsDNA ends) is highly efficient, and the high numbers of molecules in motion circumvent the low probability to detect a lesion. Then there is RecJ, the only investigated Rec-protein that forms spots of static tracks even in the absence of DSBs (see Figure 27). This was described being caused by the interaction with the C-terminus of SsbA [59]. After induction of DSBs, the exonuclease RecJ is released from theses clusters and is recruited by RecN to sites of DNA lesions. This was visualized in Figure 27 where we see a multitude of static "hotspots". For RecJ the static tracks increase from 17.4 % (no DSBs) to 21.8 % (with DSBs), which represents an increase of 25 %. This can again be compared to Liao *et al.* as well [89]. The interaction with SsbA and therefore the retention of RecJ to the replication machinery increases its local concentration to account for DSBs that might occur in exponential growing cells, which would result in the stalling and collapse of the replication fork [90]. RecO changes its behavior as well. There are little tracks with a D_{app} below the threshold in the absence of DSBs. After the induction of DSBs, we can again observe "hotspots" of statically bound RecO-molecules dispersed over the nucleoid (see Figure 28 the appendix). The relative increase of static tracks is 70 % (4 % without DSBs and 6.8 % with DSBs). Again the majority of molecules are in a diffusive state, which ensures that the density of RecO-molecules on the nucleoid is high enough to react on any cue.

At first glance, the absolute number of molecules showing static behavior seems rather low, but this is well comparable with other enzymes having the DNA as a substrate, e.g. DNA polymerase I (Pol) and DNA ligase (Lig) in *E. coli* [91]. Uphoff *et al.* described that although the absolute number of molecules per cell is much higher, but relatively the same amount is bound to the DNA: 2.7 % of Pol and 3.8 % of Lig. When they treated the cells with methyl methanesulfonate (MMS), a compound that adds methyl groups to DNA

and stalls the replisome, the bound fraction increased to 13 % for Pol and 17 % for Lig [91]. These are higher values than the values we observed for the Rec-proteins, but when compared with the dose-dependent damage curves, there is a quite dynamic range [91]. Interestingly, they do observe a dispersed localization pattern of the repair enzymes [91] similar to what we observed (compare Figure 27 and Figure 28). This dispersed pattern supports a distributive search model of individual molecules of RecN, RecJ and RecO as well as described for MutS and PolI in *E. coli* [89, 91]. So the model of RCs, as observed in eukaryotes, must indeed be discarded.

My work indicates that RCs in *B subtilis* only exist as short-lived (~2.5 s) repair enzyme recruitment platforms that, upon DNA damage, increase the local density of protein-protein interaction partners to trap single molecules out of a pool of diffusive enzymes.

The data presented here can be summarized as shown by the cartoon model in Figure 29: (A) In exponential growth many RecJ molecules are bound to the replication fork, while others diffuse throughout the nucleoid. All of them act at spontaneously occurring sites of DNA damage (bear in mind that a common cause of DSBs is the replication fork collapse). In this context, the interaction of RecJ with SsbA increases the local concentration of the exonuclease RecJ at the replication machinery. RecN and RecO are free to diffuse throughout the nucleoid, where both proteins stochastically bind to DNA in an unspecific manner. (B) Then DNA damage occurs: RecN, having high affinity to dsDNA ends, binds to the DSBs (when several are present) and oligomerizes in clusters. The other players are recruited by protein-protein interactions leading to diffusion capture. Free RecJ gets recruited and initiates the strand resection, thereby providing ssDNA. RecO is recruited to the site of lesion and starts to form the RecA-loading complex (see chapter 1.6.2). (C) While the RecA loading starts, the RecN focus is disassembled and moves to the next site of DSB, where the process occurs in a similar manner. Based on our observation that presynaptic steps can take place at many sites on the nucleoids, we propose that RecA is loaded at distinct break sites, in case of a large number of breaks occurring, and forms filamentous structures that bundle together from any of these sites. Similarly, it was

shown that in *E. coli* break sites and the homologous site in the other cell half can be moved together during the formation of crossovers [92].
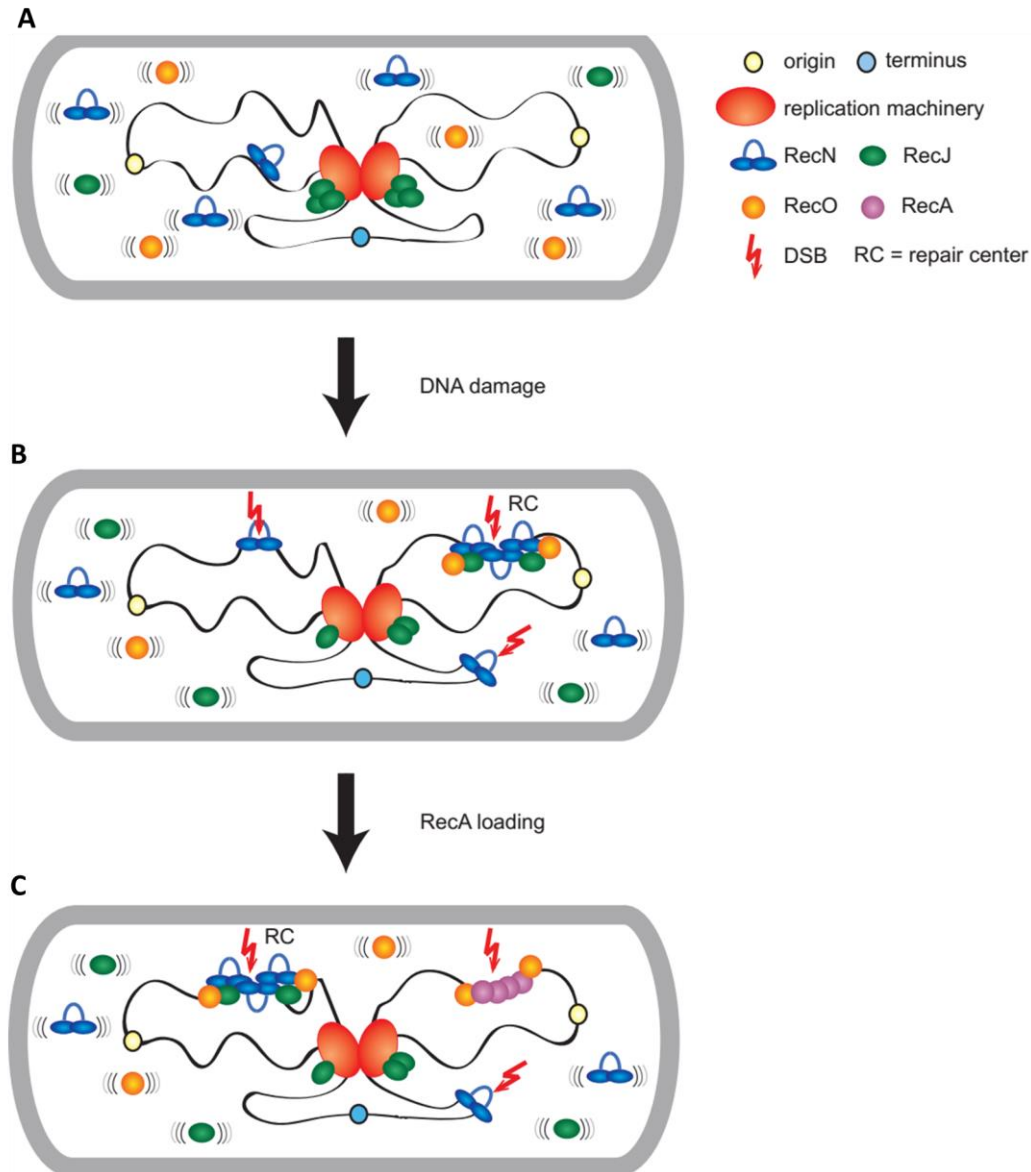
Figure 29 Cartoon Model of Rec proteins acting in DSB repair prior to RecA: (A) shows a growing and replicating cell. RecJ is stalled at the replication and RecN and RecO are diffusing over the chromosome, rarely binding to it for longer periods. (B) Upon DSB induction RecN forms clusters and tethers loose dsDNA ends. Then it recruits other Rec-proteins. RecJ is at least partially released and localized to the repair centers (RC) to produce ssDNA. RecO is recruited as well. (C) The RC is moving on to another site of DSB, while RecO loads RecA to the ssDNA.

5. References

[1] A. Earl, R. Losick and R. Kolter, "Ecology and genomics of Bacillus subtilis," *Trends Microbiol,* **16(6)**, 269-275, 2008.

[2] F. Gagné-Bourque , B. Mayer , J. Charron , H. Vali , A. Bertrand und S. Jabaji , „Accelerated Growth Rate and Increased Drought Stress Resilience of the Model Grass Brachypodium distachyon Colonized by Bacillus subtilis B26.," *PLoS One,* **10(6):**e0130456, 2015.

[3] D. Lopez, H. Vlamakis und R. Kolter, „Generation of multiple cell types in Bacillus subtilis.," *FEMS Microbiol Rev,* **33(1)**, 152-163, 2009.

[4] J. González-Pastor , E. Hobbs und R. Losick, „Cannibalism by sporulating bacteria.," *Science,* **301(5632)**, 510-3, 2003.

[5] D. Kidane, S. Ayora , J. Sweasy, P. Graumann und J. Alonso, „The cell pole: the site of cross talk between the DNA uptake and genetic recombination machinery.," *Crit Rev Biochem Mol Biol.,* **(6)**, 531-55, 2012.

[6] M. Fujita und R. Losick, „Evidence that entry into sporulation in Bacillus subtilis is governed by a gradual increase in the level and activity of the master regulator Spo0A.," *Genes Dev.,* **19(18)**, 2236-44, 2005.

[7] C. Höfler , J. Heckmann , A. Fritsch , P. Popp, S. Gebhard , G. Fritz und T. Mascher, „Cannibalism Stress Response in Bacillus subtilis.," *Microbiology,* doi:

10.1099/mic.0.000176 [Epub ahead of print], 2015.

[8]   J. Veening, W. Smits und O. Kuipers, „Bistability, epigenetics, and bet-hedging in bacteria.," *Annu Rev Microbiol.,* **62**, 193-210, 2008.

[9]   L. Hamoen, G. Venema und O. Kuipers, „Controlling competence in Bacillus subtilis: shared use of regulators.," *Micobiology,* **149(Pt 1)**, 9-17, 2003.

[10]   I. Chen, P. Christie und D. Dubnau, „The ins and outs of DNA transfer in bacteria.," *Science,* **310(5753)**, 1456-60, 2005.

[11]   R. Redfield , „Do bacteria have sex?," *Nat Rev Genet. ,* **2(8)**, 634-9, 2001.

[12]   J. Auchtung , C. Lee C, R. Monson, A. Lehman und A. Grossman, „Regulation of a Bacillus subtilis mobile genetic element by intercellular signaling and the global DNA damage response.," *Proc Natl Acad Sci U S A. ,* **102(35)**, 12554-9.

[13]   T. Rösch, W. Golman, L. Huckelsby, J. Gonzalez-Pastor und P. Graumann, „The presence of conjugative plasmid pLS20 affects global transcription of Its Bacillus subtilis host and confers beneficial stress resistance to cells.," *Appl Environ Microbiol. ,* **80(4)**, 1349-58, 2014.

[14]   D. Hofreuter, S. Odenbreit und R. Haas, „Natural transformation competence in Helicobacter pylori is mediated by the basic components of a type IV secretion system.," *Mol Microbiol.,* **41(2)**, 379-91, 2014.

[15]   J. Errington und R. Daniel, „Control of cell morphogenesis in bacteria: two distinct ways to make a rod-shaped cell.," *Cell,* **113(6)**, 767-76, 2003.

[16]   B. Burton und D. Dubnau, „Membrane-associated DNA transport machines." *Cold Spring Harb Perspect Biol.,* **2(7)**, p. a000406. doi: 10.1101/cshperspect.a000406, 2010.

[17]   W. Smits , C. Eschevins, K. Susanna, S. Bron, O. Kuipers und L. Hamopen, „Stripping Bacillus: ComK auto-stimulation is responsible for the bistable response in competence development.," *Mol Microbiol.,* **56(3)**, 604-14, 2005.

[18]   M. Ogura, H. Yamaguchi, K. Kobayashi, N. Ogasawara, Y. Fujita und T. Tanaka, „Whole-genome analysis of genes regulated by the Bacillus subtilis competence transcription factor ComK.," *J Bacteriol.,* **184(9)**, 2344-51, 2002.

[19]   M. Kaufenstein , M. van der Laan und P. Graumann, „The three-layered DNA uptake machinery at the cell pole in competent Bacillus subtilis cells is a stable complex.," *J Bacteriol.,* **193(7)**, 1633-42, 2011.

[20]   R. Provvedi, I. Chen und D. Dubnau, „NucA is required for DNA cleavage during transformation of Bacillus subtilis," *Mol Microbiol,* **40(3)**, 634-644, 2001.

[21]   D. Kidane, S. Ayora, J. Sweasy, P. Graumann und J. Alonso, „The cell pole: the site of cross talk between the DNA uptake and genetic recombination machinery.," *Crit Rev Biochem Mol Biol.,* **47(6)**, 531-55, 2012.

[22]   D. Kidane, B. Carrasco, C. Manfredi, K. Rothmaier, S. Ayora, S. Tadesse, J. Alonso
         und P. Graumann, „Evidence for different pathways during horizontal gene
         transfer in competent Bacillus subtilis cells.," *PLoS Genet. ,* **5(9)**, e1000630,
         2009.

[23]   J. Hahn, M. Maier, B. Haijema, M. Sheetz und D. Dubnau, „Transformation
         proteins and DNA uptake localize to the cell poles in Bacillus subtilis.," *Cell,*
         122(1), 59-71, 2005.

[24]   D. Kidane und P. Graumann, „Intracellular protein and DNA dynamics in
         competent Bacillus subtilis cells.," *Cell,* Nr. **122(1)**, pp. 73-84, 2005.

[25]   J. Hahn, G. Inamine, Y. Kozlov und D. Dubnau, „Characterization of comE, a late
         competence operon of Bacillus subtilis required for the binding and uptake of
         transforming DNA.," *Mol Microbiol. ,* **10(1)**, 99-111, 1993.

[26]   G. Inamine und D. Dubnau, „ComEA, a Bacillus subtilis integral membrane
         protein required for genetic transformation, is needed for both DNA binding
         and transport.," *J Bacteriol.,* **177(11)**, 3045-51, 1995.

[27]   M. Kaufenstein, Assembly of the Competence Machinery in Bacillus subtilis
         [Dissertation], Freiburg im Breisgau: Universität Freiburg, 2011.

[28]   I. Draskovic und D. Dubnau, „Biogenesis of a putative channel protein, ComEC,
         required for DNA uptake: membrane topology, oligomerization and formation
         of disulphide bonds.," **55(3)**, 881-96, 2005.

[29]   M. Takeno, H. Taguchi und T. Akamatsu, „Role of ComFA in controlling the DNA uptake rate during transformation of competent Bacillus subtilis." *J Biosci Bioeng.,* **111(6)**, 618-23, 2011.

[30]   T. Sysoeva , L. Bane , D. Xiao , B. Bose , S. Chilton , R. Gaudet und B. Burton, „Structural characterization of the late competence protein ComFB from Bacillus subtilis." *Biosci Rep.,* **35(2)**, e00183, 2015.

[31]   B. Maier , I. Chen, D. Dubnau und M. Sheetz , „DNA transport into Bacillus subtilis requires proton motive force to generate large molecular forces.," *Nat Struct Mol Biol.,* **11(7)**, 643-9, 2004.

[32]   B. Maier, M. Koomey und M. Sheetz, „A force-dependent switch reverses type IV pilus retraction," *Proc Natl Acad Sci U S A,* **101(30)**, 10961-6, 2004.

[33]   R. Berka, J. Hahn, M. Albano, I. Draskovic, M. Persuh, X. Cui, A. Sloma, W. Widner und D. Dubnau, „Microarray analysis of the Bacillus subtilis K-state: genome-wide expression changes dependent on ComK.," *Mol Microbiol.,* **43(5)**, 1331-45, 2002.

[34]   N. Mirouze, Y. Desai, A. Raj und D. Dubnau, „Spo0A~P imposes a temporal gate for the bimodal expression of competence in Bacillus subtilis.," *PLoS Genet.,* **8(3)**, e1002586., 2012.

[35]   M. Brigulla und W. Wackernagel, „Molecular aspects of gene transfer and foreign DNA acquisition in prokaryotes with regard to safety issues." *Appl*

Microbiol Biotechnol., **86(4)**, 1027-41, 2010.

[36]   A. Gagne, K. Stevens, M. Cassone, A. Pujari, O. Abiola, D. Chang und M. Sebert, „Competence in Streptococcus pneumoniae is a response to an increasing mutational burden.," *PLoS One,* **8(8)**, e72613, 2013.

[37]   R. Redfield , „Genes for breakfast: the have-your-cake-and-eat-it-too of bacterial transformation.," *J Hered.,* **84(5)**, 400-4, 1993.

[38]   M. Blokesch und P. Seitz, „DNA-uptake machinery of naturally competent Vibrio cholerae.," *Proc Natl Acad Sci U S A.,* **110(44)**, 17987-92, 2013.

[39]   M. Blokesch, „A quorum sensing-mediated switch contributes to natural transformation of Vibrio cholerae," *Mob Genet Elements.,* **2(5)**, 224-227, 2012.

[40]   J. Alonso, P. Cardenas, H. Sanchez, J. Hejna, Y. Suzuki und K. Takeyasu, „Early steps of double-strand break repair in Bacillus subtilis.," *DNA Repair (Amst).,* **12(3)**, 162-76, 2013.

[41]   H. Krokan und M. Bjørås , „Base excision repair.," *Cold Spring Harb Perspect Biol.,* **5(4)**, a012583, 2013.

[42]   C. Kisker , J. Kuper und B. Van Houten, „Prokaryotic nucleotide excision repair.," *Cold Spring Harb Perspect Biol.,* **5(3)**, a012591, 2013.

[43]   J. Lenhart, J. Schroeder, B. Walsh und L. Simmons, „DNA repair and genome

maintenance in Bacillus subtilis.," *Microbiol Mol Biol Rev.,* **76(3)**, 530-64, 2012

[44]   R. Fishel, „Mismatch Repair.," *J Biol Chem.,* **290(44)**, 26395-403, 2015.

[45]   S. Shuman und M. Glickman, „Bacterial DNA repair by non-homologous end joining.," *Nat Rev Microbiol.,* **5(11)**, 852-61, 2007.

[46]   I. Tan und K. Ramamurthi, „Spore formation in Bacillus subtilis.," *Environ Microbiol Rep. ,* **6(3)**, 212-25, 2014.

[47]   S. Ayora , B. Carrasco , P. Cárdenas, C. César, C. Cañas, T. Yadav, C. Marchisone und J. Alonso, „Double-strand break repair in bacteria: a view from Bacillus subtilis.," *FEMS Microbiol Rev.,* **35(6)**, 1055-81, 2011.

[48]   P. Thorpe , D. Alvaro, M. Lisby und R. Rothstein, „Bringing Rad52 foci into focus.," *J Cell Biol. ,* **194(5)**, 665-7, 2011.

[49]   D. Kidane und P. Graumann, „Dynamic formation of RecA filaments at DNA double strand break repair centers in live cells.," *J Cell Biol. ,* **170(3)**, 357-66, 2005.

[50]   J. Lenhart , E. Brandes, J. Schroeder , R. Sorenson, H. Showalter und L. Simmons, „RecO and RecR are necessary for RecA loading in response to DNA damage and replication fork stress," *J Bacteriol.,* **196(15)**, 2851-60, 2014.

[51]   P. Cárdenas , B. Carrasco, C. Defeu Soufo, C. César, K. Herr, M. Kaufenstein, P. Graumann und J. Alonso, „RecX facilitates homologous recombination by

modulating RecA activities.,“ *PLoS Genet.,* **8(12)**, e1003126, 2012.

[52]   C. Cañas , B. Carrasco, S. Ayora und J. Alonso, „The RecU Holliday junction
       resolvase acts at early stages of homologous recombination.,“ *Nucleic Acids
       Res.,* **36(16)**, 5242-9, 2008.

[53]   P. Cardenas, C. Gándara und J. Alonso, „DNA double strand break end-
       processing and RecA induce RecN expression levels in Bacillus subtilis.,“ *DNA
       Repair (Amst).,* **14**, 1-8, 2014.

[54]   S. Pellegrino, J. Radzimanowski, D. de Sanctis, E. Boeri Ebra, S. McSweeney und
       J. Timmins, „Structural and functional characterization of an SMC-like protein
       RecN: new insights into double-strand break repair.,“ *Structure.,* **20(12)**, 2076-
       89, 2012.

[55]   S. Novolios und D. Sheratt, „The bacterial chromosome: architecture and action
       of bacterial SMC and SMC-like complexes.,“ *FEMS Microbiol Rev.,* **38(3)**, 380-92,
       2014.

[56]   M. Wilkinson und D. Wigley, „Structural features of Chi recognition in AddAB
       with implications for RecBCD.,“ *Cell Cycle. ,* **13(18)**, 2812-20, 2014.

[57]   S. Humberto , D. Kidane, M. G. Castillo Cozar, P. Graumann und J. Alonso,
       „Recruitment of Bacillus subtilis RecN to DNA Double-Strand Breaks in the
       Absence of DNA End Processing,“ *J Bacteriol.,* **188(2)**, 353–360, 2006.

[58]   H. Sanchez, D. Kidane, M. Castillo Cozar, P. Graumann und J. Alonso, „Recruitment of Bacillus subtilis RecN to DNA double-strand breaks in the absence of DNA end processing.,“ *J Bacteriol.,* **188(2)**, 353-60, 2006.

[59]   A. Costes , F. Lecointe, S. McGovern, S. Quevillon-Cheruel und F. Polard, „The C-terminal domain of the bacterial SSB protein acts as a DNA maintenance hub at active chromosome replication forks.,“ *PLoS Genet.,* **6(12)**, e1001238, 2010.

[60]   C. Eymann, A. Dreisbach, D. Albrecht, J. Bernhardt, D. Becher, S. Gentner, L. Thi Tam, K. Buettner, G. Buurman, C. Scharf, S. Venz, U. Völker und M. Hecker, „A comprehensive proteome map of growing Bacillus subtilis cells.,“ *Proteomics.,* **4(10)**, 2849-7, 2004.

[61]   X. Xing und C. Bell, „Crystal structures of Escherichia coli RecA in a compressed helical filament.,“ *J Mol Biol. ,* **342(5)**, 1471-85, 2004.

[62]   P. Lewis und A. Marston, „GFP vectors for controlled expression and dual labelling of protein fusions in Bacillus subtilis.,“ *Gene.,* **227(1)**, 101-10, 1999.

[63]   M. Steinmetz und R. Richter, „Plasmids designed to alter the antibiotic resistance expressed by insertion mutations in Bacillus subtilis, through in vivo recombination,“ *Gene,* **142(1)**, 79-83, 1994.

[64]   H. Inoue, H. Nijima und H. Okayama, „High efficiency transformation of Escherichia coli with plasmids.,“ *Gene.,* **96(1)**, 23-8, 1990.

[65]   J. Sambrook und D. e. Russell, Molecular Cloning: A laboratory Manual, Cold Spring Harbor: Cold Spring Harbor Laboratory Press, 2001.

[66]   U. Laemmli, „Cleavage of structural proteins during the assembly of the head of bacteriophage T4.," *Nature.,* **227(5259**), 680-5, 1970.

[67]   F. Dempwolff, H. Wischhusen, M. Specht und P. Graumann, „The deletion of bacterial dynamin and flotillin genes results in pleiotrophic effects on cell division, cell growth and in cell shape maintenance.," *BMC Microbiol.,* **12**, 298, 2012.

[68]   J. Schindelin, I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, S. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Tinevez, D. White, V. Hartenstein, K. Eliceiri, P. Tomanacak und A. Carcona, „Fiji: an open-source platform for biological-image analysis.," *Nat Methods. ,* **9(7)**, 676-82, 2012.

[69]   K. Jaqaman, D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S. L. Schmid und G. Danuser, Robust single-particle tracking in live-cell time-lapse sequences, *Nat. methods* **5(8)**, 695 - 702, 2008.

[70]   O. Sliusarenko, J. Heinritz, T. Emonet und C. Jacobs-Wagner, „High- throughput, subpixel-precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics," *Mol Micro.,* **80(3)**:, 612:627, 2011.

[71]   Altenburger S, Pediaditakis M, Fries SJ, El Najjar N, Schmitt B, Dahlke S, Rösch T, Fritz G, Graumann PL., „*Single-molecule tracking of recombination proteins RecN, RecO and RecJ reveals a DNA scanning mode of movement and the*

establishment of DNA repair centres at many sites within the nucleoid." **in prep**.

[72]   S. Bakshi, B. Bratton und J. Weishaar, „Subdiffraction-limit study of Kaede diffusion and spatial distribution in live Escherichia coli.," *Biophys J.,* **101(10)**, 2535-44, 2011.

[73]   R. Lenarcic , S. Halbedel, L. Visser, M. Shaw, L. Wu, J. Errington, D. Marenduzzo und L. Hamoen, „Localisation of DivIVA by targeting to negatively curved membranes.," *EMBO J.,* **28(15)**, 2272-82.

[74]   L. Kleine Borgmann , J. Ries, H. Ewers und P. Graumann , „The bacterial SMC complex displays two distinct modes of interaction with the chromosome.," *Cell Rep.,* **3(5)**:, 483-92, 2012.

[75]   J. Elf , G. Li und X. Xie, „Probing transcription factor dynamics at the single-molecule level in a living cell.," *Science.,* **316(5828)**, 1191-4, 2007.

[76]   „Basic logical alignment tool," NCBI, [Online]. Available: http://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE_TYPE=BlastHome. [Zugriff am 2015].

[77]   C. Reimold, H. Defeu Soufo, F. Dempwolff und P. Graumann, „Motion of variable-length MreB filaments at the bacterial cell membrane influences cell morphology.," *Mol Biol Cell.,* **24(15)**, 2340-9, 2013.

[78]   P. Olshausen, H. Defeu Soufo, K. Wicker, R. Heintzmann, P. Graumann und A. Rohrbach, „Superresolution imaging of dynamic MreB filaments in B. subtilis--a

multiple-motor-driven transport?," *Biophys J.,* **105(5)**, 1171-81, 2013.

[79]   J. Ramos, M. Martines-Bueno, A. Molina-Henares, W. Teran, K. Watanabe, Z. Xiadong, M. Gallegos, R. Brennan und R. Tobes, „The TetR Family of Transcriptional Repressors," *Microbiol Mol Biol Rev.,* **69(2)**, 326–356, 2005.

[80]   J. Meile, L. Wu , S. Ehrlich, J. Errington und P. Noirot, „Systematic localisation of proteins fused to the green fluorescent protein in Bacillus subtilis: identification of new proteins at the DNA replication factory.," *Proteomics. ,* **6(7)**, 2135-46, 2006.

[81]   K. J. Briley, P. Prepiak, M. Dias, J. Hahn und D. Dubnau, „The secretion ATPase ComGA is required for the binding and transport of transforming DNA.," *Mol Microbiol.,* **81(3)**, 818-30, 2011.

[82]   T. Romantsov , A. Battle , J. Hendel , B. Martinac und J. Wood, „Protein localization in Escherichia coli cells: comparison of the cytoplasmic membrane proteins ProP, LacY, ProW, AqpZ, MscS, and MscL.," *J Bacteriol.,* **192(4)**, 912-24, 2010.

[83]   EMBL-EBI, SIB, PIR und NIH, „uniprot," UniProt consortium, 2016. [Online]. Available: http://www.uniprot.org/uniprot/P32393.

[84]   L. Feng, W. Wang, J. Cheng, Y. Ren, G. Zhao, C. Gao, Y. Tang, X. Liu, W. Han, X. Peng, R. Liu und L. Wang, „Genome and proteome of long-chain alkane degrading Geobacillus thermodenitrificans NG80-2 isolated from a deep-subsurface oil reservoir.," *Proc Natl Acad Sci U S A,* **104(13)**, 5602-7, 2007.

[85]   S. Bakshi, H. Choi, J. Mondal und J. Weishaar, „Time-dependent effects of transcription- and translation-halting drugs on the spatial distributions of the Escherichia coli chromosome and ribosomes.," *Mol Microbiol. ,* **94(4)**, 871-87, 2014.

[86]   S. Upphoff, D. Sheratt und N. Achillefs, „Visualizing Protein-DNA Interactions in Live Bacterial Cells Using Photoactivated Single-molecule Tracking," *J Vis Exp,* **(85)**, 51177.

[87]   M. Stracy , C. Lesterlin, F. Garza de Leon, S. Uphoff, P. Zawadski und A. Kapanidis, „Live-cell superresolution microscopy reveals the organization of RNA polymerase in the bacterial nucleoid.," *Proc Natl Acad Sci U S A,* **112(32)**, E4390-9, 2015.

[88]   M. Gossen und H. Bujard, „Tight control of gene expression in mammalian cells by tetracycline-responsive promoters.," *Proc Natl Acad Sci U S A.,* **89(12)**, 5547–5551, 1992.

[89]   Y. Liao, J. Schroeder, B. Gao, L. Simmons und J. Biteen, „Single-molecule motions and interactions in live cells reveal target search dynamics in mismatch repair.," *Proc Natl Acad Sci U S A.,* **112(50)**, E6898-906, 2015.

[90]   H. Masai, T. Tanaka und D. Kohda, „Stalled replication forks: making ends meet for recognition and stabilization.," *Bioessays.,* **32(8)**, 687-97, 2010.

[91]   S. Uphoff, R. Reyes-Lamothe, F. Garza de Leon, D. Sheratt und A. Kapanidis, „Single molecule DNA repair in live bacteria," *Proc Natl Acad Sci U S A.,* **110(20)**,

8063-8, 2013.

[92]   C. Lesterlin, G. Ball, L. Schermelleh und D. Sherratt, „RecA bundles mediate homology pairing between distant sisters during DNA break repair.," *Nature, 506(7487)*, 357-366, 2005.

[93]   R. Taylor, D. Walker und R. McInnes, „E. coli host strains significantly affect the quality of small scale plasmid DNA preparations used for sequencing.," *Nucleic Acids Res., 21(7)*, 1677-8, 1993.

[94]   F. Studier und B. Moffat, „se of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.," *J Mol Biol., 189(1)*, 113-30, 1986.

[95]   P. Youngman, J. Perkins und R. Losick, „Construction of a cloning site near one end of Tn917 into which foreign DNA may be inserted without affecting transposition in Bacillus subtilis or expression of the transposon-borne erm gene," *Plasmid, 12(1)*, 1-9, 1984.

[96]   H. Defeu Soufo und P. Graumann, „Dynamic movement of actin-like proteins in Bacillus subtilis," *EMBO Rep., 5(8)*, 789-794, 2004.

[97]   C. Rui, S. Guttenplan, K. Blair und D. Kearns, „Role of the σD-Dependent Autolysins in Bacillus subtilis Population Heterogeneity," *J Bacteriol.,* Nr. *191(18)*, 5775–5784, 2009.

6. Acknowledgements

7. Appendix

7.1. SMT of RecJ: Two Population Fit

We argued that three distinct behaving subpopulations constitute the given distribution. The slowest being DNA bound, an intermediate being at the replication fork and a free population. These fits resulted in three mean values for the diffusion constants (D1 and A2) as well as the respective fractions by calculating the integral of the curves (A1 and A2).
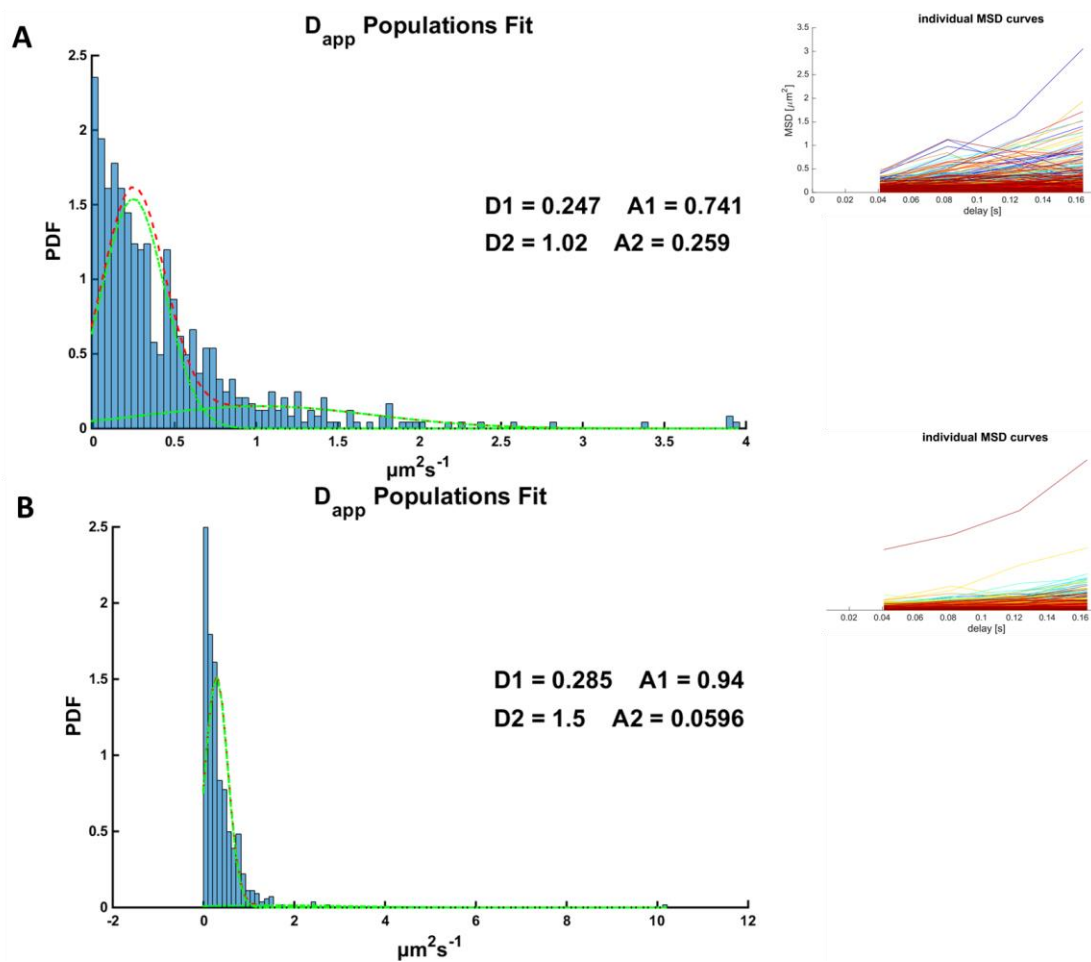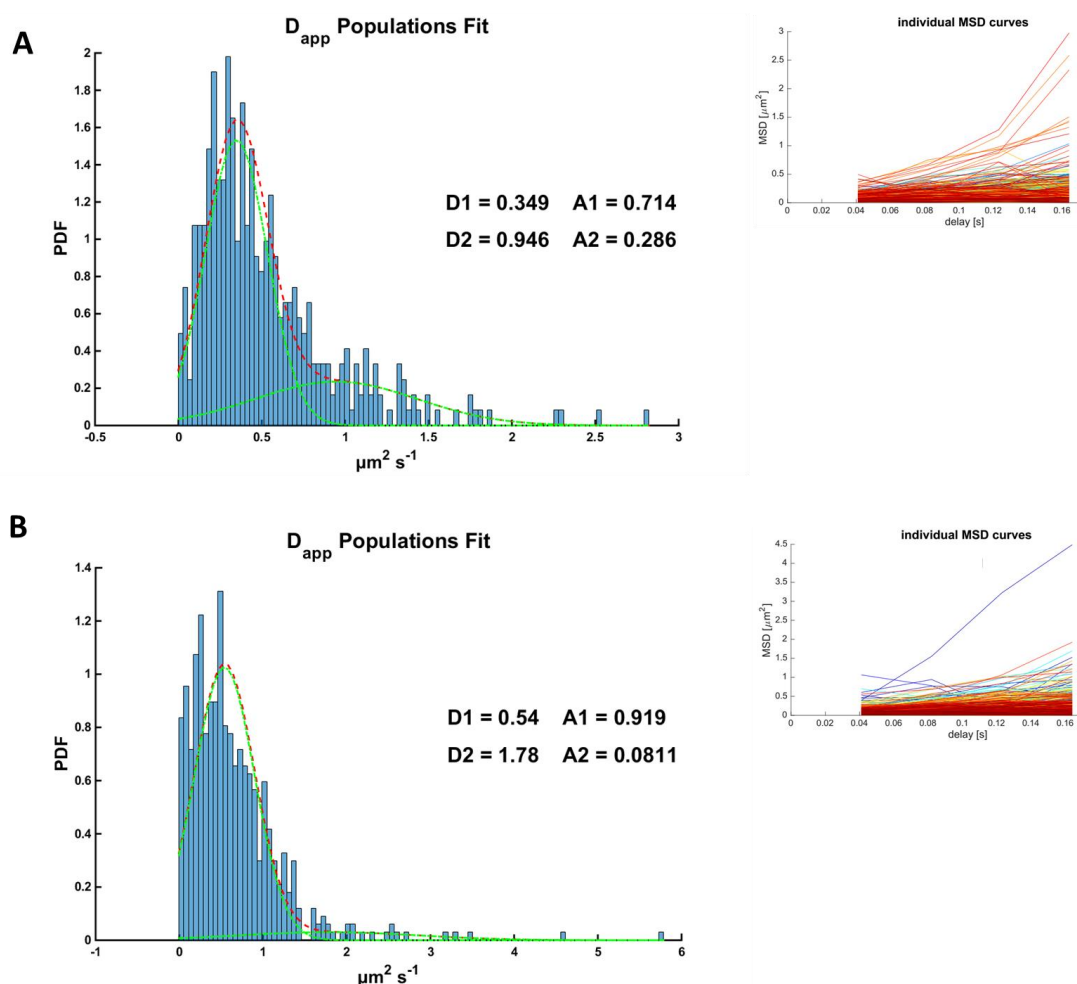

Figure 30 SMT of RecJ-YFP: The figure shows the PDF distribution of D_{app}. Three populations fit over the D_{app}-distribution. (A) Without DSBs and (B) with DSBs. D1 and D2 are the diffusion constants of the three populations and A1 and A2 are their fractions. Inlet

in the upper right shows all MSD curves that were used to derive D_{app}. The standard deviation is give in tTable 1; X-axis in $\mu m^2\ s^{-1}$; Y-axis shows the PDF

7.2. SMT of RecO: Two Population Fits

The broad distribution of diffusion constants could be described with two populations superpositioned in the distribution given by the experimental data, as there were a high peak with a mean in the lower D_{app}-range and a higher mean with broad shoulder in the higher D_{app}-range; logically these would refer to a slower population "thwarted" by unspecific DNA interaction and a faster "free" diffusing population.


Figure 31 SMT of RecO: The figure shows the PDF distribution of D_{app}. Two populations fit over the D_{app}-distribution. (A) Without DSBs and (B) with DSBs. D1 and D2 are the diffusion

constants of the two populations and A1 and A2 are their fractions. Inlet in the upper right shows all MSD curves that were used to derive D_{app}. The standard deviation is given in table 1; X-axis in $\mu m^2\, s^{-1}$; Y-axis shows the PDF
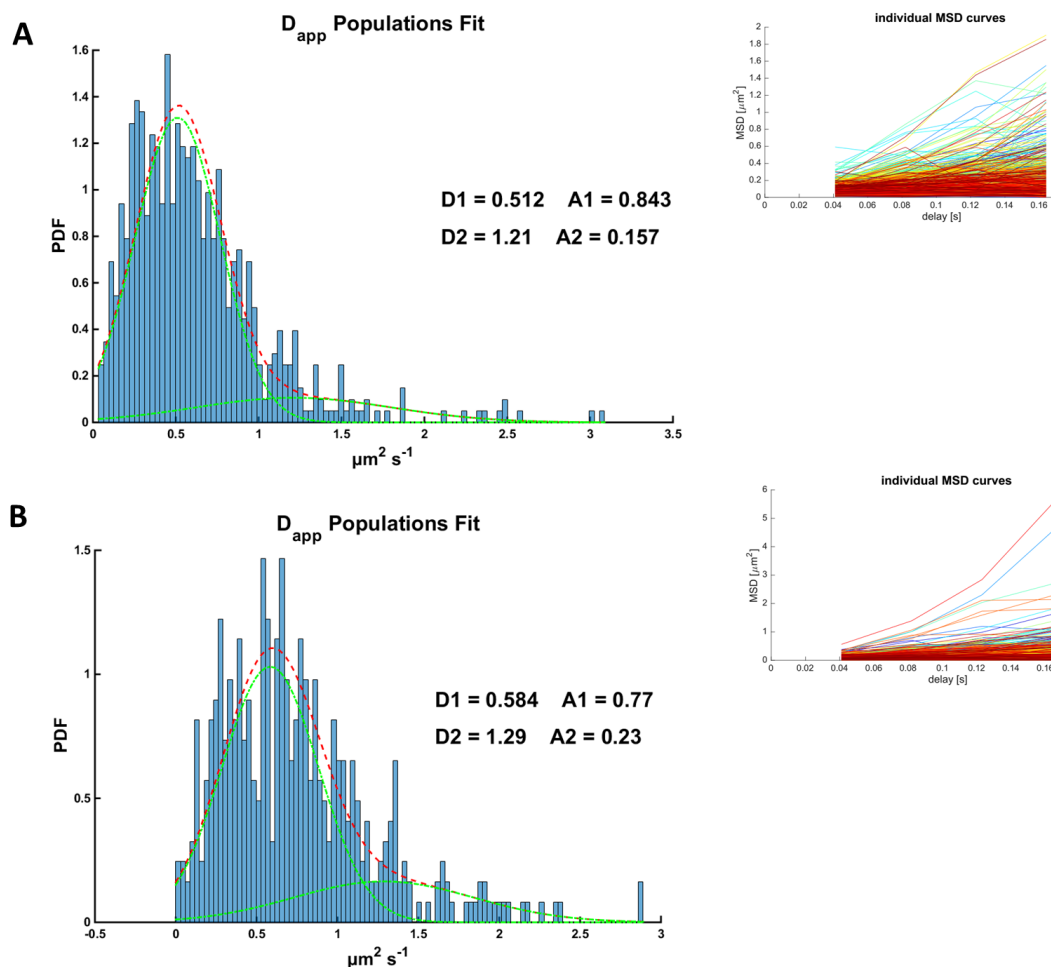
7.3.   SMT of PfkA

The overlay of all tracks in a standardized cell outline suggested less confinement to central regions of the cell. Just a few trajectories diffused with the rate below DNA-binding and they were not confined to hotspots but moved over the area with a slow D_{app} (red tracks in Figure 32 B and D). There was no difference observable between the uninduced and induced experimental condition. We applied a two population fit for two reasons. First, to achieve comparable values and second, the distribution suggest here as well that there is superposition of two populations, that might be monomers vs. homotetramers. These fits resulted in two mean values for the D_{app} (D1 and D2) as well as the respective fraction by calculating the integral of the curves (A1 and A2) (Please refer to Figure 33 B in the appendix).
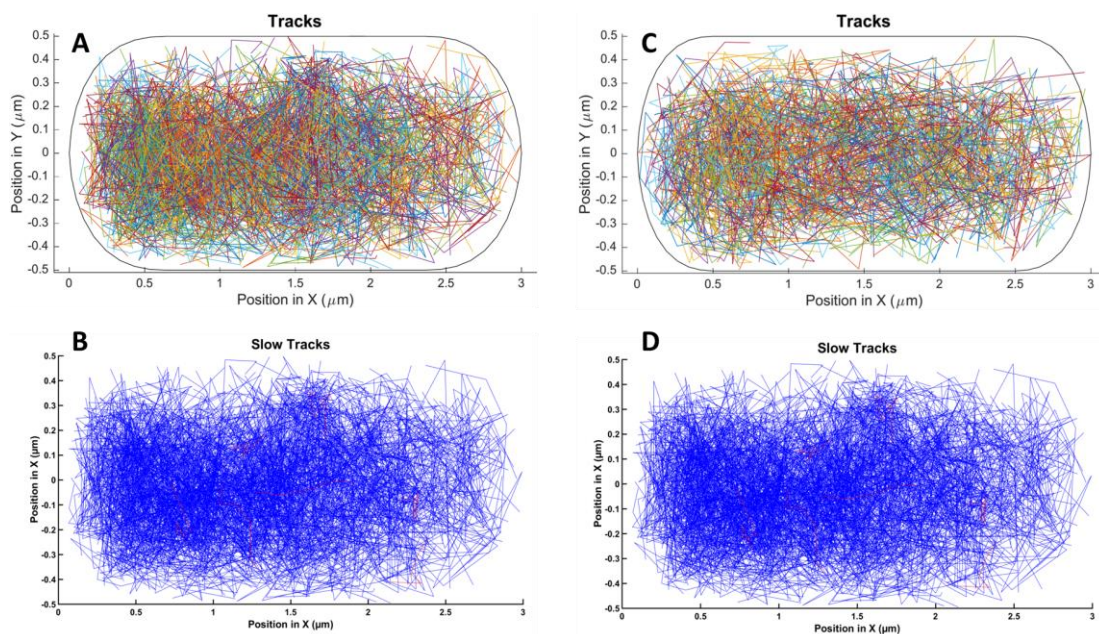
Figure 32 SMT of PfkA-GFP: This figure shows all acquired track in a normalized cell in absence (A) and presence (C) of DSBs. In the lower panels the tracks are sorted, blue tracks

exceed the threshold for DNA-binding, red tracks are below the threshold. (B) Shows the uninduced state and (D) the state after the induction of DSBs.

Figure 33 SMT of PfkA-GFP: The figure shows the PDF of the Dapp for the uninduced condition (A) and the induced condition (B). The inlet shows the MDS curves from that the diffusion coefficient was derived. D1 and A1 are the diffusion constant and fraction of the slower population; D2 and A2 are the diffusion constant and fraction of the faster population. The standard deviation is given in Table 2; X-axis in $\mu m^2\,s^{-1}$; Y-axis shows the PDF.

7.4.   SMT results of GFP-MreB

The actin-like protein MreB is a well-studied protein; still there is a heated discussion on the dynamic behavior. In the course of my experiments I used GFP-MreB to further investigate the behavior of a membrane associated protein that does not have any

transmembrane helices. Also our lab had estimated the movement speed for GFP-MreB in epifluorescence and super resolution experiments, a filament has the average speed of 60 nm s^{-1} [78].

There is the assumption that MreB exist in three populations: First, a free diffusive that is not yet bound to the membrane with the relative highest D_{app}. Second, a membrane attached population that shows an intermediate D_{app}. Third, a population that is membrane attached and bound in the filamentous structures GFP-MreB forms in exponentially growing *B subtilis* cell, with the slowest D_{app}. I performed SMT with a merodiploid strain having *gfp-mreB* und the control of the xylose promotor. I prepared the cell for microscopy 45 min after induction. Figure 34 depicts the all acquired tracks (Top) in the normalized cell and in the lower panel the tracks that are below the threshold for DNA-binding. This threshold might or might not be inaccurate in this case, the slow tracks might as well refer to the filament bound MreB-molecules; the overlay pattern has some similarity the localization pattern in epifluorescence [78, 77]. In Figure 35 we applied a three population fit to the occurring distribution, resulting in three mean values for the apparent diffusion constant (D1-D3) and the three fraction values (A1-A3) derived from the integral of these curves: D1 = 0.0674 µm^2 s^{-1} with 46.1%; D2 = 0.374 µm^2 s^{-1} with D2 = 46.7%; D3 = 1.02 µm^2 s^{-1} with A3 = 7.2%. This seemed well in the range of other D_{app} we estimated and D1 is well in the range of the observed average speed of an MreB-filament (60 nm s^{-1} [78]).
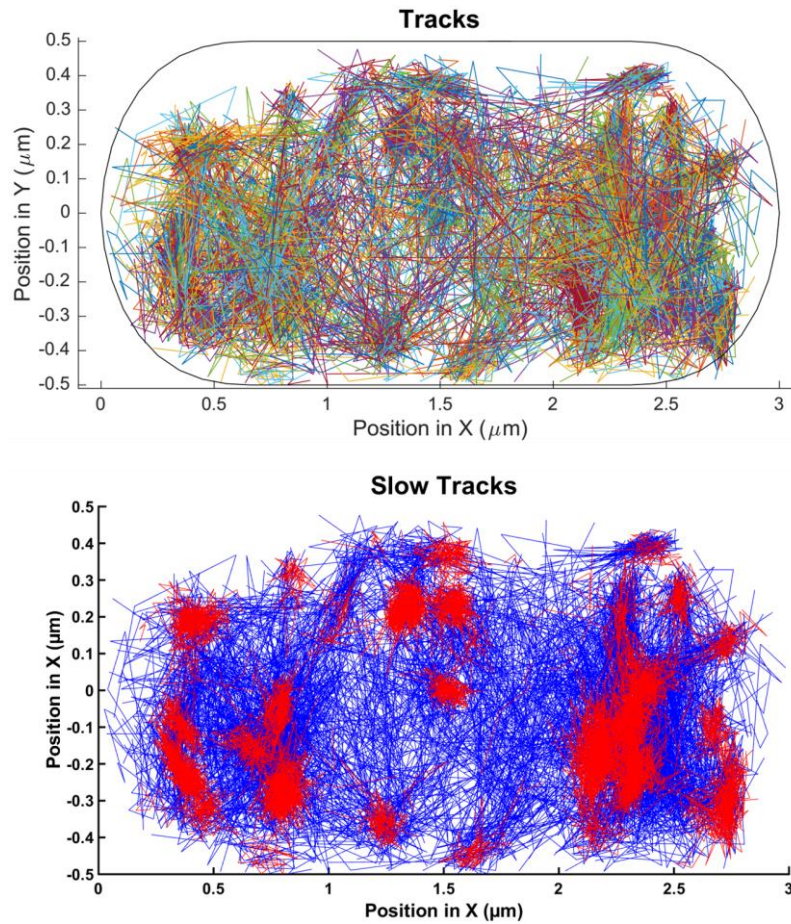
Figure 34 SMT of GFP-MreB: Top panel shows all occurring tracks in the outline of a normalized cell. In the lower panels the tracks are sorted, blue tracks exceed the threshold for DNA-binding, red tracks are below the threshold.
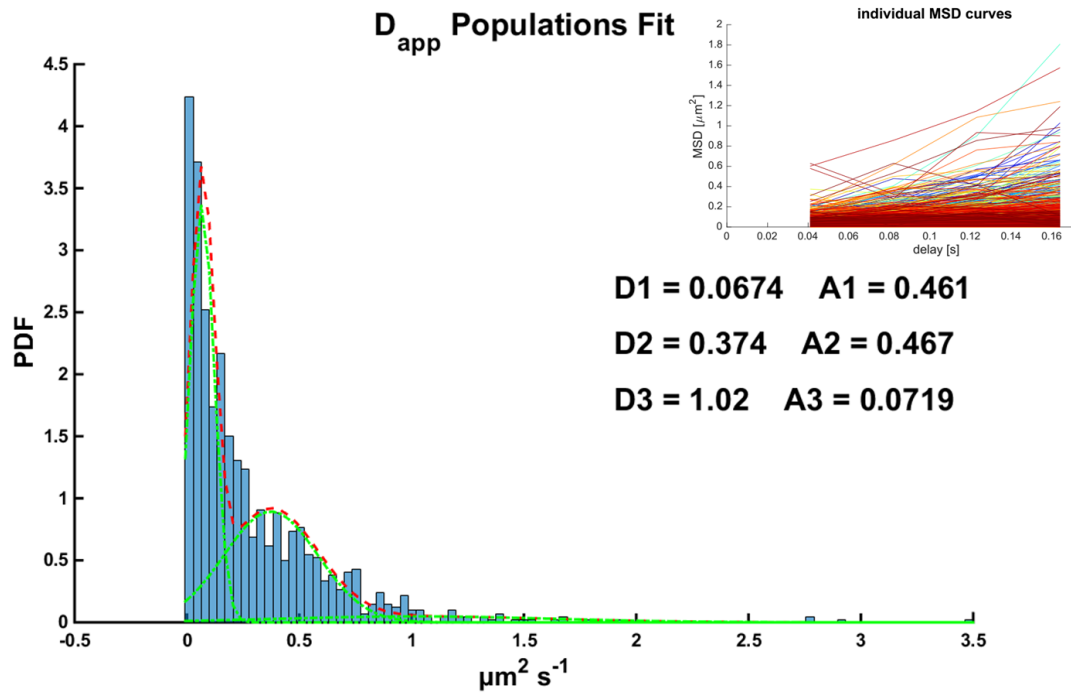
Figure 35 SMT of GFP-MreB: The figure shows the PDF distribution of D_{app}. Three populations fit over the D_{app}-distribution. (A) Without DSBs and (B) with DSBs. D1-D3 are the diffusion constants of the three populations and A1-A3 are their fractions. Inlet in the upper right shows all MSD curves that were used to derive $D_{app.}$ The standard deviation is given in tTable 1; X-axis in $\mu m^2\ s^{-1}$; Y-axis shows the PDF

7.5.   Tables

Table 3: deoxyribo-oligonucleotides/ restriction endonuclease sites

No.	Sequence (5´→3´)	Construct (vector)
1476	AAGGATATATATCAAGCGGC	*ΔcomEB*
1477	<u>GAACAACCTGCACCATTGCAAGA</u>GTTGTTCCCTCAAATGTTG	*ΔcomEB*
1478	<u>TTGATCCTTTTTTTATAACAGGAATTC</u>TGATGAATGCGTAATTCGCG	*ΔcomEB*
1479	CGATAGCTGGAAAACCCGGC	*ΔcomEB*
2272	CAT<u>GCTAGC</u>ATGCGTAATTCGCGCTTATT	*thrC::comEC (pdP150)*
2273	CAT<u>CGTACG</u>GTTCGTCTCTGTTATATCTG	*thrC::comEC (pdP150)*
1154	ACT<u>GGGCCC</u>GTCCGAGATTATCTGTCG	*comEB-YFP (psg1164)*
1155	ACT<u>GAATTC</u>CACGTAGCTCGTGAAAAG	*comEB-YFP*

2593	CATATCGTCACGTAGCTCGTGAAAAGTG	*comEB* full length (psg1193)
5397	CAT<u>GGATCC</u>TCATGTTCCCCTCCGCC-	*comEB*_{*Geo*} *(pET-24d)*
5398	CAT<u>CCATGG</u>CCCACCATCACCATCACCATATGGAACGAATGACATGGGAC	*comEB*_{*Geo*} *(pET-24d)*

Table 4: growth media and supplements

Name	Component	concentration
LB growth medium	tryptone	1 % (w/v)
	yeast extract	0.5 % (w/v)
	NaCl	1 % (w/v)
SOB	tryptone	2 % (w/v)
	yeast extract	0.5 % (w/v)
	NaCl	10 mM

	KCl	2.5 mM
SOC	Identical to SOB except for the addition of D-glucose	 20 mM
S7$_{50}$ minimal medium sterilize by filtration	1x S7$_{50}$ salts 1x S7$_{50}$ metals D-glucose glutamate casamino acids	See below See below 1 % (w/v) 0.1 % (w/v) 0.004 % (w/v)
10x S7$_{50}$ salts	$(NH_4)2SO_4$ KH_2PO_4 Adjust to pH 7.0	0.5 M 100 mM 50 mM
100x S7$_{50}$ metals sterilize by filtration	$MgCl_2$ $CaCl_2$ $ZnCl_2$	0.2 M 70 mM 0.1 mM

	$MnCl_2$	5 mM
	$FeCl_3$	0.5 mM
	thiamine-HCl	0.01% (w/v)
SpC medium	1x T-Base	See below
	D-glucose *	0.5%
sterilize by filtration**	$MgSO_4$	0.018%
	yeast extract	0.2%
	casamino acids	0.025%
SpII medium	1x T-Base	See below
	D-glucose *	0.5%
sterilize by filtration	$MgSO_4$	0.084%
	yeast extract	0.1%
	casamino acids	0.01%
	$CaCl_2$	0.5 mM
10x T-Base	$(NH4)_2SO_4$	150 mM

	K_2PO_4	1 M
	KH_2PO_4	440 mM
	trisodium citrate $\cdot$ 2 H_2O	39 mM
TB buffer	Pipes (or Hepes)	10 mM
	$CaCl_2$	15 mM
sterilize by filtration	KCl	250 mM
	Adjust pH to 6.7 with KOH or HCl	
	Add $MnCl_2$	55 mM
Mg^{2+} solution for SOB	$MgCl_2$	1 M
	$MgSO_4$	1 M

* If P_{xyl} was used to control downstream gene expression, glucose was substituted by fructose and xylose (0.5 % (w/v) each); **if not stated otherwise solutions were sterilized by autoclaving

Table 5: selective antibiotics

Name	Final concentration
Ampicillin	100 µg/ml
Chloramphenicol	5 µg/ml
Spectinomycin	100 µg/ml
Tetracycline	20 µg/ml

Table 6: bacterial strains used in this work

Strain	Genotype	References
Escherichia coli **DH5α**	*fhuA2 lac(del)U169 phoA glnV44 Φ80' lacZ(del)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*	[93]
Escherichia coli **BL21 (D3E)**	*F⁻ ompT gal dcm lon hsdS$_B$(r$_B^-$m$_B^-$) [malB⁺]$_{K-12}$(λ^S)*	[94]
Bacillus subtilis **(PY79)**	*Wildtype*	[95]

DK01	*recN-yfp*	[24]
DK02	*recO-yfp*	[24]
MH04	*recJ-yfp*	[71]
JS17	*amyE::P_{xyl} gfp-mreB*	[96]
DS3520	*amyE::Physpank-lytF*	[97]
Mik26	*comGA-cercfp, Δrok*	[19]
SA01	*comEB::tet*	This work
SA02	*thrC::comEC*	This work
SA03	*thrC::comEC, ΔcomEB*	This work
SA04	*comGA-cfp, rok-, comEB::tet*	This work
SA05	*comEB-yfp*	This work
SA06	*comEB-yfp, comGA-cfp*	This work
SA07	*amyE::comEB-GFP*	This work

SJF01	*amyE::P_{hyspank}-lytF, recN-yfp*	This work

7.6.   Abbreviations

aa	amino acids
ATP	adenosine tri-phosphate
bp	base pairs
CM	chloramphenicol
DNA	deoxyribonucleic acid
dsDNA	double stranded DNA
DSB/DSBs	double strand break/double strand breaks
GFP/YFP/CFP	green/yellow/cyan fluorescent protein
H_2O_d / H_2O_{dd}	deionized water / double deionized water
HR	homologous recombination
IPTG	isopropanol-β-D-thiogalactopyranoside
Kan	kanamycin
Kb	kilo base pairs
kDa	kilo Dalton
Mb	mega base pairs

Mg   magnesium

MMC   mitomycin C

MMS   methylmethanosulphonate

MW   molecular weight

nt   nucleotide(s)

OD_x optical density at x nm

ORF   open reading frame

PAA   poly-acrylamide

PAGE   polyacrylamide gel electrophorese

PCR   polymerase chain reaction

px   pixel

SDS   sodium dodecylsulfate

ssDNA   single stranded DNA

tet   tetracycline

w/v   weight over volume

vol   volume

Xyl   xylose

7.7. Attached Data DVD