Study on late competence proteins involved in natural transformation of *Bacillus subtilis*

Dissertation

zur Erlangung des Grades eines

Doktor der Naturwissenschaften

(Dr. rer. nat.)

des Fachbereichs Chemie der Philipps-Universität Marburg

vorgelegt von

Marie Thérèse Burghard-Schrod

geboren in Offenbach am Main

Marburg, März 2021

Die vorliegende Dissertation wurde von Juli 2016 bis Juni 2020 am Fachbereich Chemie unter Leitung von Prof. Dr. Peter L. Graumann angefertigt.

Vom Fachbereich Chemie der Philipps-Universität Marburg (Hochschulkennziffer 1180) als Dissertation angenommen am _____

Erstgutachter(in): Prof. Dr. Peter L. Graumann

Zweitgutachter(in): Prof. Dr. Lars Oliver Essen

Tag der Disputation: _____

Erklärung

Ich erkläre, dass eine Promotion noch an keiner anderen Hochschule als der Philipps-Universität Marburg, Fachbereich Chemie, versucht wurde.

Hiermit versichere ich, dass ich die vorliegende Dissertation

"Study on late competence proteins involved in natural transformation of Bacillus subtilis"

selbstständig, ohne unerlaubte Hilfe Dritter angefertigt und andere als die in der Dissertation angegebenen Hilfsmittel nicht benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen sind, habe ich als solche kenntlich gemacht. Dritte waren an der inhaltlich-materiellen Erstellung der Dissertation nicht beteiligt; insbesondere habe ich hierfür nicht die Hilfe eines Promotionsberaters in Anspruch genommen. Kein Teil dieser Arbeit ist in einem anderen Promotions- oder Habilitationsverfahren verwendet worden. Mit dem Einsatz von Software zur Erkennung von Plagiaten bin ich einverstanden.

_____ _____

Ort/Datum Unterschrift (Vor- und Nachname)

Danksagung

Ich möchte mich bei meiner Familie, Eltern und Schwiegereltern und meinem Bruder Patrick für die stetige Unterstützung bedanken. Besonderer Dank geht an meinen Mann Christian, der mir in allen Lebenslagen immer bedingungslos zur Seite steht, sowie an meine Kätzchen Godzilla und Pumpkin, die immer mit mir am Schreibtisch saßen.

Großer Dank geht an alle Graumänner und die AG Bange für die gute Zusammenarbeit. Besonders bedanken möchte ich mich jedoch bei Nina El Najjar, Patricia Bedrunka, Hector Romero, Wieland Steinchen und Barbara Waidner für die fachliche und moralische Unterstützung in den vergangenen 4 Jahren.

Des Weiteren danke ich der wunderbarsten TA Antje Schäfer und meinen Masteranden Kai Krämer und Alexandra Kilb, mit denen mir die Arbeit im Labor, aber auch der wissenschaftliche Austausch immer unheimlich viel Spaß gemacht hat.

Zuletzt geht mein Dank an Peter, der mir die Möglichkeit gegeben hat als wissenschaftliche Mitarbeiterin in seiner AG zu arbeiten, und mich dadurch persönlich und fachlich weiter zu entwickeln.

Zusammenfassung

In der hier vorgelegten Studie wurden sogenannte Kompetenzproteine, die an der natürlichen Transformation von *Bacillus subtilis* beteiligt sind, untersucht. Das grampositive Bakterium *B. subtilis* gehört zu einer Vielzahl von Bakterien, die in der Lage sind, DNA aus ihrer Umgebung aufzunehmen und die fremde DNA durch homologe Rekombination in ihr eigenes Chromosom zu integrieren; eine Eigenschaft, die als Kompetenz bezeichnet wird. Diese faszinierende Fähigkeit wird nur von einem Teil der Bakterienkultur ausgeführt, was als Heterogenität bezeichnet wird. Eine bestimmte Anzahl von Zellen exprimiert spezifische Proteine, welche von den späten Kompetenz- Operons kodiert werden. Wenn exogene doppelsträngige DNA aus der Umgebung aufgenommen werden soll, muss diese zunächst die dicke Zellwand von *B. subtilis* mit einem Durchmesser von ~40 nm passieren. Im Falle von *B. subtilis* wird diese DNA über eine Pilus-ähnliche Struktur, den sogenannten Pseudopilus in das Periplasma der Zelle transferiert. Die Energie für diesen speziellen Prozess wird wahrscheinlich von der Assemblierungs-/ Disassemblierungs- ATPase ComGA bereitgestellt. Liegt die aufgenommene DNA schließlich im Periplasma vor, wird diese durch den Kompetenzkomplex weiter in das Cytosol der Zelle transportiert. Dabei wird der Komplex aus spezifischen Kompetenzproteinen gebildet, welche sich an der Membran assemblieren. Hierzu gehören ein DNA-bindendes Transmembranprotein (ComEA) und ein Kanalprotein, welches eine hydrophile Pore für die DNA ausbildet (ComEC).

In der folgenden Arbeit wurde die unbekannte Rolle des Proteins ComEB im Zusammenhang mit der Kompetenz weiter aufgeklärt und seine enzymatische Funktion *in vitro* analysiert. Es wurde festgestellt, dass das Protein eine Deaminase-Aktivität besitzt, welche jedoch für die Transformation nicht essentiell ist. Des Weiteren wurde eine Aminosäure von ComEC (D573) als essentiell für die Transformation von *B. subtilis* identifiziert. Verkürzungen des Proteins ComEC, welches vermutlich eine Exonuklease- Funktion ausübt, wurden heterolog exprimiert und als GST-Tag-Fusionen aufgereinigt. Eine enzymatische Aktivität konnte wahrscheinlich aufgrund von Aggregationen der Proteine nicht nachgewiesen werden. Die intrazelluläre Diffusion von Fluorophore-Fusionen verschiedener Kompetenzproteine, ComEB-mV, ComGA-mV, ComEC-mV und mV-ComEA,

wurde mittels Einzelmolekül-Tracking in An- und Abwesenheit von exogener DNA analysiert. Im Falle von ComGA wurde eine C-terminale Fusion mit mVenus, (ComGA-mV) analysiert und es wurde festgestellt, dass das Protein in Anwesenheit von DNA dynamischer wird. Weiterhin wurden erstmals die Lokalisation und Diffusion eines fluoreszenzmarkierten PCR-Produktes innerhalb von kompetenten Bacillus-Zellen analysiert. Hierbei wurde herausgefunden, dass die Diffusion der fluoreszierenden DNA der Diffusion des DNA-bindenden Transmembranproteins mV-ComEA sehr ähnelt. Dies führte zu der Hypothese, dass ComEA als Reservoir für aufgenommene DNA dient, ähnlich wie es bereits für ComEA-Orthologe anderer, natürlich kompetenter Bakterien bekannt ist.

Synopsis

The following study comprises *in vivo* and *in vitro* data on several of the so-called late competence proteins, which are involved in natural transformation of *Bacillus subtilis*. The gram-positive bacterium *B. subtilis* belongs to those bacteria, who are able to take up DNA from their environment and incorporate the foreign DNA by homologous recombination into their own chromosome; a feature named competence. This fascinating ability is carried out by only a portion of the bacterial culture, expressing specific proteins, encoded by the late competence operons. If exogenous double-stranded DNA is about to be taken up from the environment, it needs to first cross the thick cell wall of *B. subtilis*, with a width of ~40 nm. In case of *B. subtilis*, this first border is crossed by a putative pseudopilus who transfers the DNA inside of the cell. The energy for this particular process is probably provided by the assembly/disassembly ATPase ComGA. The taken-up DNA is then further transferred into the cytosol by the so-called competence complex or competence machinery. The complex consists out of specific competence proteins, which assemble at the membrane, including a DNA-binding transmembrane protein (ComEA) and an aqueous channel protein (ComEC).

In the following thesis, the unknown role of the protein ComEB has been further elucidated in the context of competence, and its enzymatic function was analysed *in vitro*. It was found that the protein carries out deaminase activity, which is not essential for transformation. In case of ComEC, an amino acid, D573, has been identified as essential for transformation. Truncations of the protein, supposed to carry out an exonuclease function, were heterologously expressed and purified as GST-tag fusions, but, probably due to aggregations of the proteins, no enzymatic activity was detected. The intracellular diffusion of fluorophore fusions of several competence proteins, namely ComEB-mV, ComGA-mV, ComEC-mV and mV-ComEA was analysed via single-molecule tracking, in the presence and absence of exogeneous DNA. In case of ComGA, a C-terminal fusion to mVenus was analysed and it was found that the protein becomes more dynamic in the presence of DNA. Further, the localization and diffusion of a fluorescently labeled PCR product inside of competent *Bacillus* cells was analysed for the first time. The diffusive behaviour and localization of the stained DNA resembles the diffusion of mV-ComEA. This led to the hypothesis that ComEA serves

as a reservoir for taken-up DNA, similar to what is already known for orthologues of ComEA from other naturally competent bacteria.

Table of Contents

List of Figures XII
List of Tables XV
Abbreviations XVI

1. Introduction 1

1.1 Natural transformation in bacteria 1
1.2 Natural transformation of *Bacillus subtilis* 8
1.2.1 The putative pseudopilus and DNA-binding of competent cells 8
1.2.2 The secretion ATPase ComGA 11
1.2.3 Transfer of exogenous DNA into the cytosol 13
1.2.4 Homologous recombination 18
1.2.5 Expression of late competence genes is controlled by ComK 20

2. Results 23

2.1 Manuscript I: The *Bacillus subtilis* dCMP- deaminase ComEB acts as a dynamic polar localization factor for ComGA within the competence machinery, *Molecular Microbiology (2020)* 23
2.1.1 Abstract 24
2.1.2 Abbreviated summary 24
2.1.3 Introduction 25
2.1.4 Results 28
2.1.5 Discussion 47
2.1.6 Methods 51
2.1.7 References 58
2.1.8 Manuscript I- supplementary material 63
2.2 Manuscript II: Single molecule dynamics of DNA receptor ComEA, membrane permease ComEC and taken up DNA in competent *Bacillus subtilis* cells, *www.bioRxiv.org (2020)* 77
2.2.1 Abstract 77

2.2.2 Importance 78
2.2.3 Introduction 79
2.2.4 Methods 82
2.2.5 Results 89
2.2.6 Discussion 109
2.2.7 References 116
2.2.8 Manuscript II- supplementary material 121
2.3 Unpublished results 133
2.3.1 Expression and purification of *B. subtilis* ComEC periplasmic domains 133
2.3.2 Localization of stained DNA in competent *B. subtilis* cells 141

3. Methods (unpublished results) 143

3.1 Growth conditions 143
3.2 Strain construction 143
3.3 Preparation of chemical competent cells 144
3.4 Transformation of *E. coli* 145
3.5 Protein expression and lysate preparation 145
3.6 Affinity chromatography 146
3.7 Size exclusion chromatography 147
3.8 Mass photometry 148
3.9 Electrophoretic mobility shift assay (EMSA) & Assays of nuclease activity 149
3.10 Single molecule tracking of fluorescently labeled DNA 150

4. Discussion 151

4.1 Analysing the function of the *B. subtilis* late competence gene, *comEB* 151
4.2 ComGA - the putative assembly ATPase of the *B. subtilis* competence apparatus 154
4.3 ComEC - the putative exonuclease of the *B. subtilis* competence complex .. 156
4.4 The *B. subtilis* competence complex might be a mobile system, assembling by a diffusion/capture mechanism 158

5. Appendix 162

5.1 Calibration– mass photometry 162

5.2 Literature 169

6. *Curriculum vitae* 183

7. Inlay - Supplementary movies of manuscript I & manuscript II 183

List of Figures

1. Introduction

Figure 1: Scheme of horizontal gene transfer and natural transformation 4
Figure 2: Competence pili and DNA-uptake in different bacteria..... 7
Figure 3: Natural transformation and homologous recombination in *B. subtilis*..... 19
Figure 4: Regulation of expression of late competence genes in *B. subtilis* 22

2. Results

2.1 Manuscript I

Figure 1: Localization of different competence proteins in PY79 cells grown to competence 30
Figure 2: Relative transformation frequency of PY79 strains modified to investigate the function of ComEB in *B. subtilis*. 33
Figure 3: Artificial expression of ComEB-GFP. 35
Figure 4: Ni-NTA-purification of $ComEB_{Geo}$ and $ComEB_{Geo}$ C98A purified from *E. coli* BL21 (DE3). 37
Figure 5: Enzyme kinetics of $ComEB_{Geo}$ expressed from *E. coli* BL21 (DE3)..... 38
Figure 6: Apparent Diffusion- incubation with or without DNA. 41
Figure 7: Bleaching curves of the polar complexes formed within the cell by our ComGA-mV fusion (PG3788). 45
Figure 8: Heat maps and probability of all analysed tracks (see table 1) plotted against the x-axis of a normalized cell. 46

Figure S1: Selection of the alignment of B. subtilis ComEB (1), Geobacillus thermodenitrificans ComEB (2) and Streptococcus mutans dCMP deaminase (3)... 64
Figure S2: In-gel fluorescence of samples applied in single-molecule tracking..... 65
Figure S3: Transformation frequencies of strains PG3781 (empty-vector control), PG3782, PG3783 and PG3784 66

2.2 Manuscript II

Figure 1: Scheme of horizontal gene transfer and natural transformation 90
Figure 2: Competence pili and DNA-uptake in different bacteria..... 91
Figure 3: Natural transformation and homologous recombination in *B. subtilis*..... 93
Figure 4: Regulation of expression of late competence genes in *B. subtilis* 95h
Figure 5: Overview of *B. subtilis* ComEC C-terminal truncations..... 134
Figure 6: Strep-tag affinity chromatography, heparin affinity chromatgraphy and mass photometry of construct (4) 136
Figure 7: GST-tag affinity chromatography and mass photometry of construct (5) 138
Figure 8: GST-tag affinity chromatography, size exclusion and mass photometry of construct (7) 139
Figure 9: Localization of fluorescently labeled DNA in competent *B. subtilis* cells 142
Figure 10: Model of a competent B. subtilis cell; competence proteins comEA and ComEC diffuse along the membrane..... 161
Figure 11: Calibration of mass photometry measurements 162

Figure S1: Overlay of tracks on one cell for ComECΔ128-mV and ComECΔ301-mV. *122*
Figure S2: t-stacks of movies of ComECΔ128-mV. *123*
Figure S3: Agarose-gel of labelled PCR-product. *123*
Figure S4: Negative control for Single-molecule tracking. *124*
Figure S5: Transformation frequency of mV-ComEA. *125*
Figure S6: Single-molecule tracking of mV-ComEA induced with 0.5% Xylose. ... *126*

2.3 Unpublished results

Figure 5: Overview of *B. subtilis* ComEC C-terminal truncations 134
Figure 6: Strep-tag affinity chromatography, heparin affinity chromatgraphy and mass photometry of construct (4) 136
Figure 7: GST-tag affinity chromatography and mass photometry of construct (5) 138
Figure 8: GST-tag affinity chromatography, size exclusion and mass photometry of construct (7) 139
Figure 9: Localization of fluorescently labeled DNA in competent *B. subtilis* cells 142

4. Discussion

Figure 10: Model of a competent *B. subtilis* cell; competence proteins comEA and ComEC diffuse along the membrane……..…161

5. Appendix

Figure 11: Calibration of mass photometry measurements 162

List of Tables

2. Results

2.1 Manuscript I

Table 1: Apparent Diffusion of ComGA-mV, ComEB-V and ComEB-mV C98A .. 42

Table S1: List of strains *67*
Table S2: List of primer *71*

2.2 Manuscript II

Table 1: Localization of ComEC-mV and C-terminal truncations 92
Table 2: Transformation frequencies of strains expressing ComEC-mV and truncations 94
Table 3: Transformation frequencies of mutants of ComEC 96
Table 4: Single-molecule tracking of ComEC-mV, ComECΔ128-mV and ComECΔ301-mV 101
Table 5: Single-molecule tracking of fluorescently labeled DNA compared to mVenus-ComEA 107

Table S1: List of strains, B. subtilis *127*
Table S2: List of strains, E. coli DH5α *129*
Table S3: List of primer, named by E. coli strains or purpose. Restriction sites and mutations are indicated in italic, SD-sequence in bold. *130*

5. Appendix

Table 1: List of Primer 163
Table 2: List of strains 164
Table 3: Polymerase chain reaction 166
Table 4: List of chemicals 167

Abbreviations

%	percentage
°C	degree Celsius
µg	microgram
µl	microliter
µm	micrometer
aa	amino acid
Amp	Ampicillin
APS	Ammonium persulfate
ATP	adenosine triphosphate
bp	base pair
CHES	2-(cyclohexylamino) ethane sulfonic acid
cm	centimeter
Da	dalton
dCMP	deoxycytidine monophosphate
dCTP	deoxycytidine triphosphate
DNA	deoxyribonucleotide acid
dNTP	deoxynucleotide triphosphate
dsDNA	double- stranded DNA
DTT	1,4-dithiothreitol
dTTP	deoxythymidine triphosphate
dUMP	deoxyuridine monophosphate
EMSA	electrophoretic mobility shift assay

fw	forward/upstream
g	gram
GST	glutathione
GST-tag	glutathione-S-transferase-tag
h	hour
HEPES	n-(2-hydroxyethyl) piperazine-n′-(2-ethanesulfonic acid)
IPTG	Isopropyl-β-D-thiogalactopyranosid
Kan	Kanamycin
kDa	kilodalton
kHrz	kiloHerz
l	liter
M	molar
MBP	Maltose Binding Protein
mg	milligram
min	minute
ml	milliliter
mM	millimolar
mRNA	messenger RNA
ms	millisecond
mV	millivolt
ng	nanogram
nM	nanomolar
nm	nanometer

OB-fold	oligonucleotide/oligosaccharide binding fold
OD	optical density
PCR	polymerase chain reaction
pg	picogram
rcf	relative centrifugal force
rev	reverse/downstream
RFU	relative fluorescent unit
RNA	ribonucleotide acid
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
sec	second
ssDNA	single- stranded DNA
strep	streptavidin
strep-tag	streptavidin-tag
TEMED	N, N, N′, N′-Tetramethylethylenediamine
Tris	tris(hydroxymethyl)aminomethane
U	unit

1. Introduction

1.1 Natural transformation in bacteria

Horizontal gene transfer (HGT) in bacteria has been known since the early 20th century. Back then, *Streptococcus pneumoniae* was first "transformed" by Griffith in 1928, who studied the development of new phenotypes in bacteria and discovered the exchange of pathogenic factors (Griffith, 1928). DNA has been isolated later on and identified as "transforming principle" (Avery *et al.*, 1944).

HGT can be divided in three different types: Transduction, conjugation and natural transformation (Dubnau & Blokesch, 2019, Seitz & Blokesch, 2013). While transduction was found to involve bacteriophages infecting and thereby fortuitously transferring foreign DNA into the bacterial host organism (Chiang *et al.*, 2019), conjugation requires the formation of a conjugating pilus, creating cell-to-cell contact, and sharing genetic information by an active and directed mechanism (Holmes & Jobling, 1996). In addition to transduction and conjugation, natural transformation describes the uptake of exogenous double-strand DNA (dsDNA) by some bacterial species, depending *only* on the recipient cell (see Figure 1 a)). It does not require any artificial treatment of the cells compared to e.g. well-established methods for chemical transformation and electroporation in case of *Escherichia coli*, and has been widely used by various laboratories, in order to genetically modify bacteria, namely *Lactobacillus, Bacillus* and *Streptococcus* (Mandel & Higa, 1970, Hanahan, 1983, Chassy, 1987, Quinn & Dancer, 1990, Powell *et al.*, 1988).

Growth conditions or substances leading to the formation of competent cells within a bacterial culture differ strongly between bacterial species. The mechanism, how dsDNA is taken up by naturally competent bacteria, has been intensively studied in gram- positive and gram- negative model organisms, such as *Bacillus subtilis, Streptococcus pneumoniae, Neisseria gonorrhoeae, Haemophilus influenza*, *Vibrio cholerae* and *Helicobacter pylori* (Dubnau, 1991, Muschiol *et al.*, 2015, Israel *et al.*, 2000, Alexander & Redman, 1953). *H. pylori* and *N. gonorrhoeae* are constitutively competent through every growth phase (Israel *et al.*, 2000, Sparling, 1966). *H. influenzae,* like *B. subtilis*, requires a downshift of nutrition; *H. influenzae* develops

competence from the late exponential phase on, and frequency of natural transformation is even increased when cells are transferred to a nutrition depleted medium or subjected to transient, limited aeration (Herriott *et al.*, 1970, Gromkova & Goodgal, 1979). *B. subtilis* is an obligate aerobe soil bacterium, which can also be sampled from aquatic habitats and plants (Nicholson *et al.*, 2000, Molina-Santiago *et al.*, 2019). It is able to develop competence in the beginning of stationary growth phase in a subpopulation of cells (10-20%), a feature termed heterogenous gene expression (Davis & Isberg, 2016, Berka *et al.*, 2002). Special nutrient-deficient media are required to reach this state in the laboratory (Spizizen, 1958). In case of *Vibrio cholerae,* competence is only induced when cells are grown on chitinous surfaces, which resembles its natural habitat, namely the exoskeletons of aquatic crustaceans (Meibom *et al.*, 2005).

H. pylori colonizes the human stomach in half of the world´s population, where it can cause gastritis and peptic ulcer disease (Suerbaum & Michetti, 2009). In addition, an infection increases the risk of developing gastric cancer (Parsonnet *et al.*, 1991, Wang *et al.*, 2007). *H. pylori* is thought to be one of the most genetically diverse bacterial species, which is directly associated with its high ability to take up foreign DNA, thereby adapting to its environment, the human gastric epithelia (Suerbaum & Achtman, 1999). Considering its pathogenicity, its constitutive competence may be required due to high selective pressure of the host's immune system. A similar observation has been made for *N. gonorrhoeae,* where the genetical exchange by natural transformation leads to high interspecies diversity and a variety of surface antigens, preventing an adaptation of the hosts immune system to gonococcal infections (Hobbs *et al.*, 1994). It has been shown that DNA-damage causes an increase of transformation frequency in *H. pylori* (Dorer *et al.*, 2010). Interestingly, competence is also induced by DNA damage in *S. pneumoniae,* which was demonstrated by incubating cells with the DNA-modifying agent Mitomycin C (Prudhomme *et al.*, 2006). In addition, transformation frequency increases under stress conditions such as antibiotic treatment (Domenech *et al.*, 2018). Regarding the effect of DNA damage on induction of natural competence, it is important to note that these events are probably linked by the lack of an SOS-response system in *H. pylori* as well as in *S. pneumoniae* (Dorer *et al.*, 2010). In bacteria, special genes are transcribed in

case of DNA-damage, a process named SOS-response. The expression of these genes is mainly regulated by two proteins: LexA and RecA. The regulon is directly controlled by LexA (or a LexA-like protein, in *Bacillus*: DinR) binding to a specific site on the DNA called ´SOS box`, where it acts as a repressor (Winterling *et al.*, 1997). If the replication fork is stalled by DNA mutation, RecA will bind to the free single-stranded DNA at the replication fork, and further induce self-cleavage of LexA. In case the concentration of LexA is decreasing, the SOS regulon, consisting of over 50 genes, will be expressed. Repair proteins are then produced by the cell in order to fix the problem. These repair proteins include *ruv* genes for excision of wrong nucleotides, *recN* and *ruvAB* genes for homologous recombination and low-fidelity DNA-polymerases (Maslowska *et al.*, 2019). It was found that in some *Streptococcus* species the described SOS-response is missing, e.g. in *S. pneumoniae*. In these bacteria, the presence of ssDNA inside of the cell induces natural transformation instead, indicating that natural transformation might serve as a repair mechanism in some organisms (Prudhomme *et al.*, 2006).

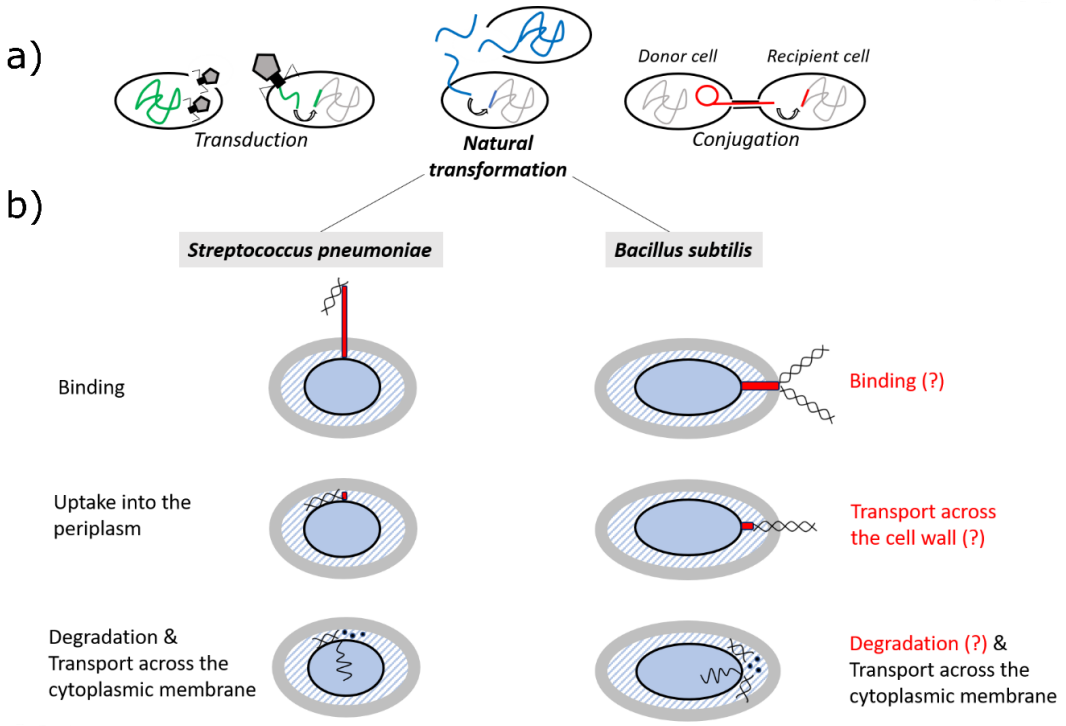


Figure 1: Scheme of horizontal gene transfer and natural transformation

In a) the three different types of DNA-uptake in bacteria are shown. From left to right: environmental DNA can enter the cells via a bacteriophage infection (transduction), taken up by competent cells (natural transformation) or provided by a donor cell, and cells induce the formation of a conjugating pilus (Conjugation) in order to transfer DNA. b) compares two gram-positive model organisms for natural transformation, *S. pneumoniae* and *B. subtilis*. Light grey= cell wall, light blue (dashed)= periplasm, black= cytosolic membrane, light blue= cytosol, red= type IV pilus of *S. pneumoniae* or putative pseudopilus of *B. subtilis*, black helix= double-stranded DNA. While for *S. pneumoniae* the most critical steps (indicated on the left) during DNA-uptake have been elucidated in the past decade, in case of *B. subtilis*, some open questions remain (highlighted on the right, in red and with a question mark).

In case of *N. gonorrhoeae*, transforming DNA needs to encode specific sequences of 10 base pairs (DUS; DNA-uptake sequences, 5´-GCCGTCTGAA-3´), which are present quite frequently in *Neisseria spp.* chromosomes. These sequences enable the identification and integration of DNA deriving only from the same species (Hamilton & Dillard, 2006). For *H. influenzae*, DNA-uptake requires a 11 base pair DNA-uptake sequence (5'-AAGTGCGGTCA-3') to be encoded in the transforming DNA (Danner et al., 1980). Such sequences are not known for *Bacillus*, but it was found by experiments carried out using bacteriophage transduction, that a minimum size of approximately 50-70 base pairs of general homology to the *Bacillus* chromosome were needed for integration of plasmid DNA (Khasanov *et al.*, 1992, Alonso *et al.*, 1986). Maximum efficiency of transduction has been reached with a size of 500 homologous base pairs, or larger (Khasanov *et al.*, 1992, Alonso *et al.*, 1986, Deichelbohrer *et al.*, 1985). Therefore, the source of transforming DNA is considered to be of interspecies origin, deriving from dead cells of a bacterial culture from the same species (Veening & Blokesch, 2017). This theory is supported by the fact that if DNA is taken up in case of *S. pneumoniae*, which colonizes the human nasopharynx, non-competent neighbouring cells are lysed at the same time, for which special bacteriocins are secreted (Wholey *et al.*, 2016, Kjos *et al.*, 2016), a process termed fratricide (Wei & Håvarstein, 2012).

Formation of antibiotic resistances and the distribution of pathogenic islands, as well as DNA-repair, are possible reasons why natural competence could have evolved among many bacteria, leading to genome diversity as an evolutionary benefit for the population (Brito *et al.*, 2018a). According to the current state of knowledge, about 80 bacterial species are naturally transformable (Johnston *et al.*, 2014). Special proteins are expressed in order to facilitate the uptake of foreign DNA, and are, except for *H. pylori*, well conserved among gram- positive and gram- negative species. These proteins will be described in more detail later on. DNA-uptake across the cell wall or membrane is carried out by pilins of the type IV secretion system, or so-called pseudopilins, similar to proteins which are involved in type II secretion systems of gram-positive and gram-negative bacteria, as the DNA needs to cross the outer cell membrane or the cell wall, in order to be taken up by the cell (Muschiol *et al.*, 2015, Giltner *et al.*, 2012). A pilus can be described as a macromolecular, filamentous

structure at the cell surface of bacteria. These filaments are involved in twitching motility (Mattick, 2002), biofilm formation (O'Toole & Kolter, 1998), surface adhesion (Rudel *et al.*, 1995), but also DNA-uptake (Dubnau & Blokesch, 2019). For *B. subtilis*, such a structure has not been visualized yet, but it is considered to be 40 to 100 nm long, just long enough to cross the cell wall (Chen *et al.*, 2006). It is not clear whether the putative pseudopilus of *B. subtilis* is able to directly bind DNA, or how large this filamentous structure is, as some revealing experiments are missing. In both organisms, only ssDNA enters the cytosol, meaning one strand has to be degraded ((Draskovic & Dubnau, 2005, Bergé *et al.*, 2002), see Figure 1b)).

For *S. pneumoniae*, these surface exposed filaments can be quite larger than *B. subtilis* putative pseudopilus; they have been detected by electron micrographs with a length of several micrometres and a width of 6 to 7 nm, binding to DNA throughout the filament ((Muschiol *et al.*, 2017, Laurenceau *et al.*, 2013) see Figure 2). In case of *V. cholerae*, the uptake of labeled DNA by competence pili has been reported, which assemble and retract in order to internalize DNA (Ellison et al., 2018). It has been shown that DNA- uptake takes place at the cell pole of *V. cholerae, H. pylori* and *B. subtilis*, while for *S. pneumoniae* and *N. gonorrhoeae*, no specific localization was observed (Laurenceau *et al.*, 2013, Boonstra *et al.*, 2018, Corbinais *et al.*, 2016, Gangel *et al.*, 2014). Lately, sequence- alignments revealed high abundance of a core competence domain of the protein ComEC, which is known to be essential for transformation, as it is thought to serve as an aqueous channel transferring the DNA from the periplasm into the cytosol of the cells (Draskovic & Dubnau, 2005). The domain was found to be present in 89% of the proteomes and 96% of the analysed genomes (Pimentel & Zhang, 2018). The identified sequence homology provides evidence of a common origin of the protein among bacteria, and gives rise to the idea of a very broad distribution of natural competence; more than it has yet been shown under laboratory conditions.

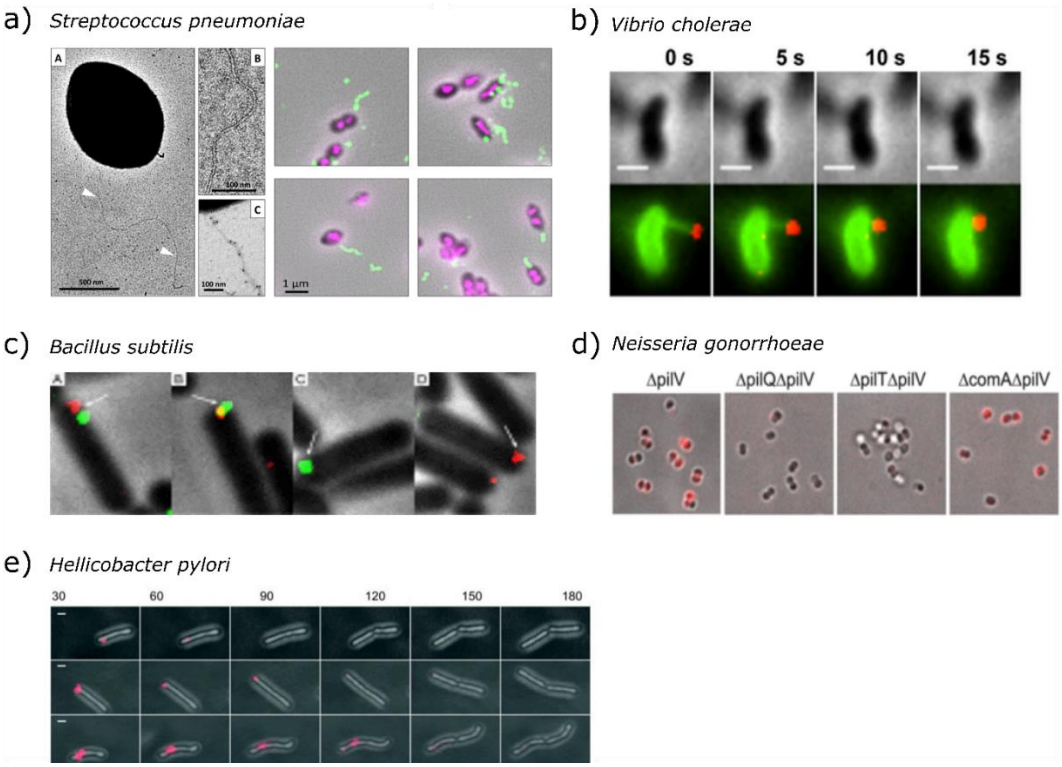


Figure 2: Competence pili and DNA-uptake in different bacteria

a) shows an electron-micrograph, where a competence pilus of *S. pneumoniae* has been detected. White arrows indicate pilus-structure; smaller images show zoomed pilus. On the right, structures were detected by immunofluorescence, using an anti-FLAG antibody (green). Nucleoids were stained with DAPI (magenta), and an overlay of the fluorescence signals with the corresponding bright field image was generated. (Laurenceau *et al.*, 2013). In b) a time lapse of a *V. cholerae* cell is shown. A PilA-Cys strain was created, where the major subunit of the type IV pilus was labelled with AF488-malemeide (green), which is taking up fluorescently labelled DNA (Cy3-DNA, red). Scale bar indicates 1 µm (Ellison *et al.*, 2018). c) shows *B. subtilis* cells taking up differently labelled fluorescent DNA at the cell pole, indicated by white arrows (Boonstra *et al.*, 2018). In d) fluorescently labelled DNA was added to *N. gonorrhoeae* (stained with Cy3), where it was found that the signal was most of the time homogenously distributed in the periplasm. Different deletion strains (indicated on the top of each image) were analysed (Gangel *et al.*, 2014). e) shows a time-lapse of H. pylori cells taking up labelled DNA (stained with 550-ATTO), which localizes to the cell pole first (time scale indicated on top of each image). Scale bar indicates 1 µm (Corbinais *et al.*, 2016). All images were assembled from the indicated publications.

1.2 Natural transformation of *Bacillus subtilis*

1.2.1 The putative pseudopilus and DNA-binding of competent cells

In general, natural competence is associated with the formation of type IV pili or structurally related fibres, which are similar to the type II secretion system of bacteria. Most of the proteins involved in the formation of these DNA-uptake machineries are considered as homologs, and the basic mechanism by which DNA is transferred across the membrane or cell wall is assumed to be similar, even though some proteins remain unique for the individual systems (Piepenbrink, 2019, Korotkov & Sandkvist, 2019). One polytopic membrane protein forms the base of the filament, while typically, one or two traffic ATPases are involved in the system, responsible for assembly and/or retraction of the pilus (Misic *et al.*, 2010, Satyshur *et al.*, 2007). Pseudopilins or Pilins are small, hydrophobic proteins, with a size of ~7-20 kDa, possessing a conserved, hydrophobic, N-terminal membrane helix. In every system, one of these proteins, called the major pilin/pseudopilin, reversibly polymerizes by addition of single subunits at the base (or platform) of the pilus into a helical fibre. The resulting filament is tightly packed and stabilised by hydrophobic and electrostatic subunit-subunit interactions of the N-terminal helices of the corresponding major pilin (Giltner *et al.*, 2012, Craig *et al.*, 2003). Typically, a prepilin-peptidase is involved in pilin maturation. At least one minor pilin participates in the formation or function of the pilus, while the purpose of minor pilins seems to vary among the different systems (Giltner *et al.*, 2012).

In the gram-positive bacteria, *Staphylococcus*, *Streptococcus* and *Bacillus*, fibers involved in DNA-uptake and are also called competence pili or pseudopili. Initially, pilins of the type II secretion system were not known to form fibres and therefore named "pseudopilins" (Pugsley, 1993). In the aforementioned organisms, components of the pseudopilus and other involved competence proteins are encoded by the *com*-operons (Muschiol *et al.*, 2015, Kovács *et al.*, 2009, Fagerlund *et al.*, 2014). In case of *B. subtilis*, the putative assembly/retraction ATPase (ComGA), a polytopic membrane protein (ComGB) and the pseudopilins (ComGC-ComGG) are encoded by the *comG*-operon (Chung *et al.*, 1998, Chung & Dubnau, 1998, Chen *et al.*, 2006, Briley *et al.*, 2011a). In order to facilitate DNA-uptake, it is generally assumed that *B. subtilis*

expresses a short pseudopilus, which only spans the cell envelope, similar to the type II secretion systems of gram-negative bacteria (Green & Mecsas, 2016, Chen *et al.*, 2006). On the other hand, when overexpressed, some major pseudopilins polymerize into large, surface-exposed fibres, as demonstrated for XcpT of *Pseudomonas aeruginosa* (Durand *et al.*, 2003, Sauvonnet *et al.*, 2000). The type II secretion systems are well characterized, and are known to be involved in the secretion of toxins, proteases and biofilm associated proteins, as well as enzymes facilitating the break-down of complex carbohydrates, RNA and DNA, in order to increase availability of carbon and phosphate sources (Kirn *et al.*, 2005, Rossier *et al.*, 2009, Wilton *et al.*, 2018). Modelling of a pseudopilus of *V. cholerae´s* type II secretion system, using structural subunit information, resulted in a right-handed one-start helix with a width of 50-60 Å (Giltner *et al.*, 2012, Laurenceau *et al.*, 2013). The pseudopilus of *B. subtilis* probably consists of 40-100 monomers; indicating a variable size of the filaments. These multimers of ComGC were isolated from *B. subtilis* cell walls treated with lysozyme (Chen *et al.*, 2006). For comparison, a type IV pilus is assembled out of 500 – 1000 subunits or more (Sauvonnet *et al.*, 2000).

It has been demonstrated that all seven genes of the *comG* operon are required for transformation in *B. subtilis* (Chung & Dubnau, 1998, Briley *et al.*, 2011a), but the function of the minor-pilins, ComGD-ComGG, which are low abundant compared to ComGC, remains to be elucidated (Mann *et al.*, 2013). The major pseudopilin ComGC and some of the minor pseudopilins, ComGD, ComGE and ComGG, need to undergo a posttranslational cleavage event by the prepilin peptidase ComC, which removes a short, N-terminal leader sequence, in order to release pre-pilins from the cytoplasmic membrane and to facilitate their secretion into the periplasmic space (Chung & Dubnau, 1995, Strom & Lory, 1991). It is possible that one minor pilin is required to bind the DNA at the tip of the pseudopilus, similar to what has been postulated for the minor pilin ComP of *N. gonorrhoea,* which binds double-stranded DNA *in vitro* (Cehovin *et al.*, 2013, Wolfgang *et al.*, 1999). In addition, it has been postulated that the minor pilins of *V. cholerae* facilitate DNA-binding at the tip of the pilus, but so far, no DNA-binding properties of *B. subtilis* minor pilins have been detected (Ellison *et al.*, 2018). The formation of an initiation complex by at least some of the pseudopilins has been assumed, which primes the assembly of the pilus. Minor pilins

would then be transferred to the outside of the cell during pilus assembly at the tip of the nascent filament (Giltner *et al.*, 2012). This particular theory is in agreement with the finding that the minor pilins ComGD, ComGE and ComGG of *B. subtilis* interact directly, as demonstrated by *in vitro* studies of affinity-tagged proteins (Mann *et al.*, 2013). On the other hand, in *S. pneumoniae*, DNA binding was detected to occur throughout the whole filament, leading to the hypothesis that the interspace of the major pseudopilin ComGC creates a DNA-binding motif, assuming that DNA-binding could only be detected in fully assembled pili. Another idea would be the integration of a minor pilin along the pilus to accomplish DNA-binding (Laurenceau *et al.*, 2013). Interestingly, recent studies showed an influence of the composition of cell-wall teichoic acids on DNA- binding in competent *B. subtilis* cells. Transformation efficiency decreased in the presence of Tunicamycin, an antibiotic which acts on the synthesis of peptidoglycan and glycopolymers such as teichoic acids, the latter decorating the bacterial cell wall. In addition, binding of fluorescently labeled DNA of competent *B. subtilis* cells, which were treated with Tunicamycin, decreased (Mirouze *et al.*, 2018). The authors demonstrated that cell-wall teichoic acids were accumulating at the cell pole of competent cells when incubated with labelled Concavalin A, a compound which binds glucose residues at the cell surface. These studies also indicate an importance of the putative glycosyl transferase TuaH for transformation, as it is specifically induced in competent cells, and disruption of its encoding gene leads to an 80fold decrease of transformation efficiency (Mirouze *et al.*, 2018, Berka *et al.*, 2002). Thus, binding of DNA in competent cells might occur directly or indirectly through specially glycosylated teichoic acids accumulating at the cell pole, guiding DNA in close proximity to the pseudopilus.

1.2.2 The secretion ATPase ComGA

In competent *B. subtilis cells*, fluorophore-fusions of ComGA localise to the septum and the pole, with one or two foci per cell. In some cases, up to 10 foci per cell were reported, which displayed a punctuate pattern throughout the membrane, in addition to its polar localization (Hahn *et al.*, 2005). Performing epifluorescence microscopy, recent studies colocalized the protein with fluorescently labelled DNA in *B. subtilis* (Boonstra *et al.*, 2018). *ComGA*, the first gene of the *comG* operon, is the only gene of this operon which is absolutely required for DNA-binding of competent *B. subtilis* cells. In addition, it is needed for internalization of transforming DNA (Briley *et al.*, 2011a). ComGA belongs to the large superfamily of AAA^+ ATPases (**A**TPases **A**ssociated with diverse cellular **A**ctivities) which are widely distributed in all organisms (Hanson & Whiteheart, 2005). These proteins participate in multiple cellular processes by inducing conformational changes of their specific substrate; AAA^+ ATPases are involved in the formation of the proteasome, in ribosome biogenesis of eukaryotic cells, DNA replication and regulation of gene expression (Yedidi *et al.*, 2017, Prattes *et al.*, 2019). The proteins are known to self-oligomerize into hexamers, leading to the formation of their typical ring-shape, which can be detected by electron microscopy (Hospenthal *et al.*, 2017, Hanson & Whiteheart, 2005). AAA^+ proteins possess a characteristic 200-250 amino-acid ATP-binding domain, located at its N-terminus, which contains a Walker A and Walker B motif required to bind and hydrolyse ATP (Hanson & Whiteheart, 2005). Orthologs of ComGA can be found in every competence-associated bacterial system, type II or type IV secretion system, which encodes an ATPase providing energy for pilus assembly. About 30% sequence similarity was detected for N-terminal and C-terminal domains of PilB, PilT, PilU, (type IV pilus systems), GspE (type II secretion system) and ComGA (*com* system) (Piepenbrink, 2019).

Briley *et al.* (2011) showed by mutating the Walker A and Walker B motifs of ComGA in *B. subtilis*, that the activity of ComGA is required for transformation of *B. subtilis*, as transformability of these strains was abolished completely. But, in case of a Walker B motif mutation, which would allow binding but no hydrolysis of ATP, competent cells were at least still able to bind DNA (Briley *et al.*, 2011a). These studies indicate

that hydrolysis of ATP might not be inevitably required for binding of transforming DNA, while binding of ATP (by the presence of an intact Walker A motif) would be sufficient. Data were obtained by a filter-based binding-assay, under usage of radiolabelled DNA. Briley *et al.* analysed several mutant strains, where they detected radiolabelled DNA which becomes DNAse resistant when taken up by the cells. The authors demonstrated that the major pilin ComGC only assembles in the periplasm in case ComGA mutants were able to take up DNA. These results support the idea that ComGA is the secretion ATPase of the *B. subtilis* system, providing energy for pilin assembly facilitating DNA-uptake (Briley *et al.*, 2011a); In general, it is assumed that a conformational change of the hexameric protein might occur through ATP-hydrolysis, converting chemical into mechanical energy leading to a rotation of the multimer (Piepenbrink, 2019). Structural analysis of PilB and PilT are each thought to form a pore, rotating in opposing directions, clockwise and counterclockwise, in order to facilitate assembly and retraction of the type IV pilus systems (McCallum *et al.*, 2017, Solanki *et al.*, 2018). In *B. subtilis*, it is not clear whether ComGA might take over both functions, or if pilus retraction/degradation might be accomplished independently of ComGA (Giltner *et al.*, 2012). Beside its crucial role in competence, ComGA serves as a checkpoint during competence inhibiting cell division, until the protein is degraded and cells start to divide again, thereby escaping from the competent state (Haijema *et al.*, 2001). In competent *B. subtilis* cells, FtsZ-ring formation is prevented by ComGA (Briley *et al.*, 2011b). Further, Hahn *et al.* (2015) demonstrated that ComGA contributes to the inhibition of rRNA transcription. Replisome assembly was bypassed in cells where ComGA was inactivated. It was found that fluorescent fusions of ComGA and DnaX (DNA-polymerase III) colocalized in competent cells, and less DnaX-foci where detected in *comGA* mutant strains (Hahn *et al.*, 2015). ComGA inhibits replisome assembly and it was found to directly interact with RelA, a GTP pyrophosphokinase, which synthesises and degrades the alarmone (p)ppGpp, provoking *B. subtilis* stringent response. As increased (p)ppGpp levels inhibit DNA replication and rRNA synthesis, the authors propose that an interaction of ComGA and RelA prevents hydrolysis of (p)ppGpp synthesis (Hahn *et al.*, 2015). These data reveal multiple cellular functions for ComGA during competence, beside its role as putative assembly and retraction ATPase.

1.2.3 Transfer of exogenous DNA into the cytosol

DNA is bound by ComEA in the periplasm

B. subtilis takes up double-stranded DNA into the periplasm with a rate of 80 bp/s (Maier *et al.*, 2004). But before the transforming DNA enters the *B. subtilis* cell, it has to be degraded into smaller DNA-fragments, which can then be isolated from the competent culture already 30 s after addition of the double-stranded DNA (Davidoff-Abelson & Dubnau, 1973). Former studies showed that circular plasmid DNA, which is bound at the cell-surface of competent cells, can only become linearized in the presence of *nucA* (Provvedi et al., 2001). The gene encodes a membrane-associated endonuclease (NucA) which is required to create cell-surface-associated DNA-fragments, and carries out nuclease activity *in vitro* (Provvedi *et al.*, 2001, Vosman *et al.*, 1988, Mulder & Venema, 1982). DNA seems to be unspecifically hydrolysed to an average size of 25.000 bp in the periplasm, providing smaller fragments for further internalisation into the cytosol (Ayora *et al.*, 2011). The double-stranded DNA (dsDNA) is then bound by the integral membrane protein ComEA, encoded by the first gene of the *comE* operon (Provvedi & Dubnau, 1999, Inamine & Dubnau, 1995, Hahn *et al.*, 1993), (see Figure 3a)). Inamine & Dubnau showed that ComEA is absolutely relevant for transport, but not for binding of DNA (Inamine & Dubnau, 1995). Its C-terminus is located at the outside of the cytosolic membrane, with which it binds the incoming dsDNA by a helix-loop-helix motif (Inamine & Dubnau, 1995, Provvedi & Dubnau, 1999). The topology was experimentally proven, as the protein is protease-accessible when expressed in protoplasts, and an alkaline phosphatase assay of two C-terminal PhoA-fusions resulted in expression of active PhoA, indicating an extracellular location of the C-terminus (Inamine & Dubnau, 1995, Takeno *et al.*, 2012). Orthologues of ComEA are found in many naturally competent bacteria, which are all double-strand-DNA binding proteins. In *N. gonorrhoea*, the protein ComE localises in the periplasm and exhibits DNA-binding function during competence. It is considered to form a reservoir for incoming DNA in the periplasm, as a ComE-mCherry fusion homogenously distributes in the periplasm until DNA is added to the cells, leading to the formation of discrete foci (Gangel *et al.*, 2014). ComEA-mCherry accumulates also in case DNA is added to cells of *V. cholerae*, and is required for

uptake of transforming DNA into the periplasm (Seitz *et al.*, 2014). In addition, *B. subtilis*, ComEA has been found to localise throughout the membrane when analysed by epifluorescence microscopy. This is in contrast to the localization of the fluorophore fusions of other competence proteins of *B. subtilis*, which exhibit a more polar pattern, such as the second and the third gene product of the *comE* operon, ComEB and ComEC (Kaufenstein *et al.*, 2011). Localization of a N-terminal fluorophore-fusion of ComEA should reflect the proper cellular localization, allowing proper folding of the protein in the cytosol.

Uptake of ssDNA into the cytosol by the membrane permease ComEC

In *H. pylori*, the orthologue of ComEA, ComH, hands over the bound dsDNA to the N-terminus of ComEC, a polytopic membrane protein, which transfers the incoming DNA into the cytosol, serving as a hydrophilic aqueous channel for DNA internalisation (Damke *et al.*, 2019b). Even though systems tend to vary among different bacteria, at the current state of knowledge, interaction of ComEA and ComEC is assumed to be similar among the corresponding homologues or orthologues (Muschiol *et al.*, 2015, Dubnau & Blokesch, 2019). Furthermore, only single-strand DNA is transferred into the cytosol, while the homologous strand is degraded, and nucleotides are released into the medium (Lacks *et al.*, 1974, Dubnau & Cirigliano, 1972b, Dubnau & Cirigliano, 1972a, Davidoff-Abelson & Dubnau, 1973). Provvedi *et al.* showed that the non-transforming strand cannot be degraded in absence of *comEC*, encoding the putative channel protein of the system, indicating a direct or indirect involvement of the protein in strand degradation (Provvedi *et al.*, 2001). In contrast, a competence-specific nuclease, EndA, is expressed in *S. pneumoniae* and facilitates degradation of the complementary, non-transforming strand (Puyet *et al.*, 1990). In case *comEC* is deleted in this organism, free nucleotides can still be detecting in the periplasm, proving that degradation is independent of ComEC (Bergé *et al.*, 2002). As such a nuclease has not been identified in *B. subtilis* so far, a recent theory favours the idea, that ComEC carries out the required nuclease function (Dubnau & Blokesch, 2019, Baker *et al.*, 2016).

ComEC contains three important domains: an oligo-saccharide binding domain (OB-fold), predicted to be present at the N-terminus of the protein, which is located within a domain of unknown function (DUF4131), a transmembrane competence domain including multiple transmembrane helices (number varies from 7 to 11 helices, depending on two different publications), and a Lactamase_B-domain, localized at the C-terminus of the protein (Baker *et al.*, 2016, Draskovic & Dubnau, 2005, Pimentel & Zhang, 2018). The C-terminal domain belongs to a broad family of zinc-dependent nucleases (Metallo-β-lactamases), and could facilitate the degradation of the incoming DNA in *B. subtilis* (González, 2020, Baker *et al.*, 2016). The presence of conserved sequences of ComEC led to the identification of many homologues in multiple bacterial phyla and species by computational analysis (Kantor *et al.*, 2013, Mell & Redfield, 2014). Recently, Pimentel & Zhang found that 89% of 5.574 analysed bacterial proteomes contained the characteristic competence domain of ComEC (Pimentel & Zhang, 2018).

In former studies, the topology of the protein was analysed by creating PhoA-and LacZ fusions at different positions of the protein, in order to detect the enzymatic activity after isolation of the corresponding fusions (Draskovic & Dubnau, 2005). While PhoA would only be properly folded in the periplasm, LacZ only remains functional when expressed in the cytosol, revealing localization of the distinct parts which were fused to the enzymes. In addition, ComEC was isolated from membrane fractions of competent *Bacillus* cells and it was found that the protein dimerises through formation of a disulphide bond (Draskovic & Dubnau, 2005). It contains two soluble loops which are localized at the N- and C-terminus, and both are probably present in the periplasm of the cell (Draskovic & Dubnau, 2005). Dimerization possibly occurs by interaction of these domains, as indicated by co-purification of both soluble parts when co-expressed in *E. coli* (Dubnau & Blokesch, 2019).

The uptake of ssDNA into the cytosol through the ComEC channel requires energy, which is assumed to be provided by ComFA, a membrane associated, DEAD-box helicase-like ATPase encoded by the *comF* operon (Londoño-Vallejo & Dubnau, 1993, Chilton *et al.*, 2017). The protein would then be acting as the DNA-translocase of the system, similar to a cyclic molecular motor driving DNA import through conformational change (Chilton *et al.*, 2017). ComFA is, in addition to ComEC,

needed for degradation of the non-transforming strand, and the proteins have been found to be in close proximity to each other, measured by fluorescence resonance energy transfer (FRET) (Provvedi *et al.*, 2001, Kramer *et al.*, 2007). Importantly, mutations in the Walker A, Walker B motif and the DEAD box motif of *comFA* decreased transformation efficiency 100-fold, thereby revealing a need for the ATPase and helicase activity of ComFA during transformation (Chilton *et al.*, 2017). The effect of proton motive force on DNA translocation has been investigated, too, using uncoupling agents (protonophores), which were added to competent *Bacillus* cells. The rate of DNA transport, measured by using a bead-based assay, immediately decreased in these experiments, indicating a requirement of transmembrane motive force for DNA uptake into the cell, or proton-symport. Therefore, it seems that multiple energy sources empower the translocation of ssDNA into the cell. In addition, transport into the cell was inhibited when ComEC was deleted, highlighting again its role as a channel (Maier *et al.*, 2004).

The unknown role of comEB, the second gene of the comE operon

It has been reported that *comEB*, the second gene of the *comE* operon, is not relevant for transformation. This statement was based on an in-frame truncation of ComEB, where 84 N-terminal residues of the protein were removed and transformation efficiency was not significantly affected. However, the authors state that since 56% of the protein sequence would still be expressed in their specific strain, a functional domain could be left carrying out its activity during competence (Inamine & Dubnau, 1995). Further, the authors assume an indirect role in competence and postulated the presence of an internal ribosomal binding site within the *comEB* sequence, which might regulate the expression of the downstream gene *comEC,* as only a tandem pair of UGA codons is located between these genes (Inamine & Dubnau, 1995, Hahn *et al.*, 1993). ComEB is annotated as a putative 2´deoxy-cytosidine-(dCMP) deaminase (Inamine & Dubnau, 1995). These proteins deaminate the nucleotide dCMP to form dUMP, which is commonly found as intermediate in nucleotide metabolism, serving as the precursor for dTTP, the final product of the pathway (Munch-Petersen & Mygind, 1976). Similar deaminases, which are found in *S. mutans,* T4 bacteriophages

and viruses, are allosterically regulated proteins which form hexamers in solution (Rami Almog *et al.*, 2004, Hou *et al.*, 2008, Zhang *et al.*, 2007). The addition of the positive effector dCTP increases enzyme activity by binding the allosteric site, leading to a conformational change which increases affinity for the substrate at the active site, while the addition of dTTP decreases activity due to product inhibition (Li *et al.*, 2016, Keefe *et al.*, 2000). dCMP-deaminase activity of ComEB was confirmed by activity-assays, when the corresponding sequence of *B. halodurans* was overexpressed in *E. coli* (Oehlenschlaeger *et al.*, 2015). The authors purified the protein with the aim to solve the crystal structure, which was not successful, but postulated an alternate route to form dTTP via ComEB, as *Bacillus* species encode a second, essential, tetrameric dCMP deaminase (encoded be the *cdd* gene) (Song & Neuhard, 1989, Johansson *et al.*, 2004). Finally, to the current state of knowledge and at the beginning of the following studies, no investigations were done whether competence and the enzyme activity of ComEB might be correlating. Interestingly, in fluorescence microscopy, the protein localises to the cell pole in competent *B. subtilis* cells, where DNA-uptake takes place, similar to other late competence genes, indicating a specific role during uptake (Boonstra *et al.*, 2018, Kidane *et al.*, 2012).

1.2.4 Homologous recombination

SsDNA, which enter the cytosol, is directly bound by single-strand DNA binding proteins. SsbB (YwpH), SsbA and DprA (Smf) cover the DNA and prevent it from unspecific degradation by cellular nucleases (Yadav et al., 2012, Yadav et al., 2014, Berge *et al.*, 2003). SsbB and DprA localize to the cell pole in competent cells, close to the competence machinery, and both are essential for transformation (Hahn *et al.*, 2005, Tadesse & Graumann, 2007), (see Figure 3b)). Further, the protected DNA is bound by several recombination (Rec) proteins, which then mediate homologous recombination of the single strand and the nucleoid, in order to maintain the integrity of the chromosome or increase genomic diversity (Brito *et al.,* 2018b). Interestingly, in *B. subtilis*, two different pathways exist depending of the type of transforming DNA, whether it is of chromosomal or plasmid origin (Kidane *et al.,* 2009). The key protein of chromosomal integration is RecA. The binding of RecA to the incoming DNA is mediated by RecO and DprA (Tadesse & Graumann, 2007). The protein localises at the cell pole during competence, followed by formation of helical filaments or threads during DNA-uptake. These threads probably transport ssDNA towards the bacterial chromosome, in order to carry out strand exchange (Kidane & Graumann, 2005). Homologous regions are then detected by the RecA-ssDNA filament which slides along the bacterial chromosome. In *E. coli,* the "sliding model" was investigated by single molecule tools, and it was found that the protein-DNA complex diffuses at 8000 bp/s (Ragunathan *et al.,* 2012). RecA is an ATP-dependent nuclease, carrying out strand exchange upon hydrolysis of ATP. It detects homologous regions by weak, transient random interactions, while it is able to ignore shorter sequences and identifies highly homologous sequences (Bell & Kowalczykowski, 2016). In case plasmid DNA is added to competent *B. subtilis* cells, fluorophore fusions of RecO and RecU localise to the cell pole close to the DNA uptake complex. This pattern has not been observed when chromosomal DNA was added. It seems that incoming plasmid DNA needs to be annealed into double strand DNA, dependent on RecO, but independent of RecA (Kidane *et al.,* 2009, Yadav *et al.,* 2012). If ssDNA is coated with SsbB and SsbA, RecO is able to anneal complementary strands in vitro to form a DNA duplex, which would then be able to replicate intracellularly (Yadav *et al.,* 2012, Yadav *et al.,* 2013). RecU localises to the cell pole dependent on RecA, where it modulates the activity of

RecA by preventing RecA-ssDNA disassembly, thereby leading to an inhibition of strand migration. (Yadav *et al.*, 2012, Carrasco *et al.*, 2005).

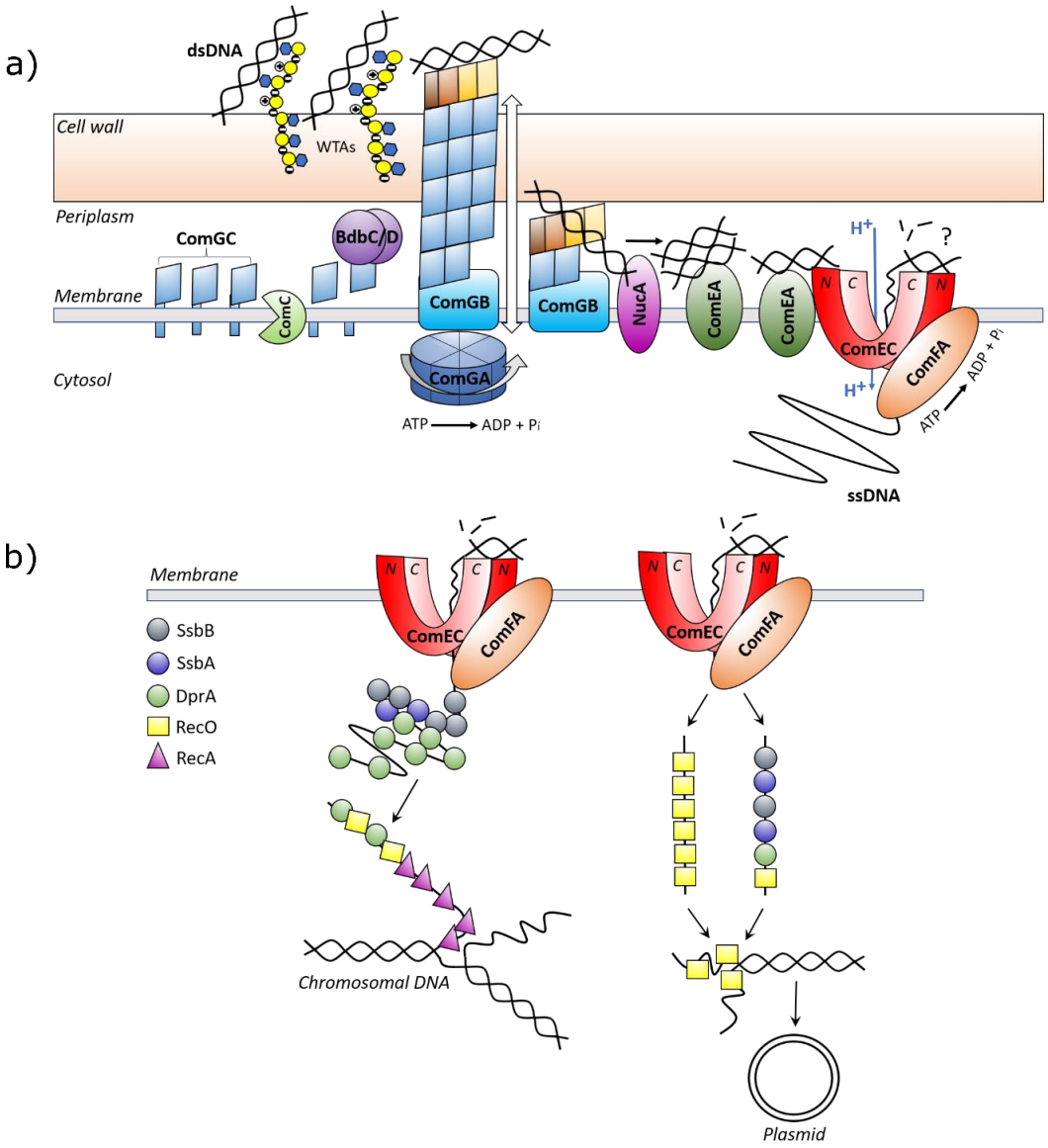


Figure 3: Natural transformation and homologous recombination in B. subtilis

In a), the DNA-uptake machinery at the cell pole of *B. subtilis* is shown. Wall teichoic acids (WTAs) with specific glycosylation may modulate the binding of double-strand DNA (dsDNA), accumulating specifically next to the uptake apparatus. DsDNA may also be directly bound by minor pseudopilins (ComGD-GG, indicated in brown tones). ComGA forms a hexamer providing the energy for pilus assembly by ATP hydrolysis. The N-terminal peptide sequence of the major pilin ComGC is hydrolysed by the

peptidase ComC. ComGC forms the pilus by which DNA is taken up into the periplasm by assembly and disassembly of subunits (white arrow). Intramolecular disulphide bonds are generated by the oxidoreductase pair BdbC/D. Probably, NucA degrades the incoming dsDNA, thereby increasing the amount of accessible free ends for the DNA-receptor ComEA. The dsDNA is then handed over to ComEC, where it is bound by its N-terminally located OB-fold. ComEC´s putative exonuclease function at the C-terminus is indicated by a question mark. Entry of single-strand DNA (ssDNA), may require proton symport (blue arrow), and translocation by ComFA (Dubnau & Blokesch, 2019, Giltner *et al.*, 2012). B) shows homologous recombination, which can follow two different pathways. Left: Import and homologous recombination of chromosomal DNA, right: Import and Annealing of plasmids (Burton & Dubnau, 2010).

1.2.5 Expression of late competence genes is controlled by ComK

In *B. subtilis*, competence is controlled by the master regulator ComK, which is expressed in the beginning of the stationary growth phase (van Sinderen *et al.*, 1995). The transcription factor binds the promotor at its specific recognition-site (AAAA-$[N]_5$-TTTT), also named K-box, where it stabilises the interaction of the RNA-polymerase, thereby increasing the transcription of the corresponding gene (Hamoen *et al.*, 1998, Susanna *et al.*, 2004). Binding of the protein to the DNA occurs by formation of a tetramer composed of two dimers (Maamar & Dubnau, 2005). ComK regulates up to 100 genes, which are, amongst other things, involved in the expression of extracellular and surface proteins, and controls its own promotor (Albano *et al.*, 2005, 2002, Berka *et al.*, 2002). In case ComK reaches high intracellular concentrations, cells become competent and enter the so-called K-state. Only a portion, meaning 5-50% of the *B. subtilis* culture becomes competent, a feature which is named heterogeneity (Berka *et al.*, 2002, Mirouze *et al.*, 2018, van Sinderen *et al.*, 1995, Hamoen *et al.*, 1998). It is assumed that the typical bimodal expression of the *B. subtilis* K-state is caused by stochastic fluctuations (noise) of expression of ComK, leading to the development of two different phenotypes within one culture (Gamba *et al.*, 2015). Further, competence development is regulated on a transcriptional, post-transcriptional and post-translational level. Expression of *comK* is repressed by Rok (**r**epressor **o**f **C**omk) and CodY (Smits & Grossman, 2010, Hoa *et al.*, 2002). CodY

encodes a GTP binding protein, a nucleotide which serves as a nutritional signal for competence development, and might therefore be involved in GTP-sensing (Serror & Sonenshein, 1996). In addition, CodY senses the amount of branched chain amino acids, namely leucine, isoleucine and valine (Shivers & Sonenshein, 2004). The Kre protein destabilizes the amount of ComK´s mRNA, which decreases the concentration of the protein, therefore inhibiting an entry into the K-state. ComK down-regulates the expression of *kre*, leading to a double negative feedback, which fine-tunes bi-modal gene expression, additionally (Gamba *et al.*, 2015). In case ComK will be degraded, cells escape the k-state and will start dividing again. In exponential phase, ComK is bound by MecA, which again is in a complex with the protease ClpCP, leading to hydrolyses of ComK. In case stationary-phase stress is appearing, the protein ComS is produced, competing for the MecA binding site. As a result, ComK is released and induces the K-state by positive autoregulation (Liu & Zuber, 1998, Ogura *et al.*, 1999, Turgay *et al.*, 1998). Expression of ComS is induced based on *quorum sensing.* A specific protein, ComA, becomes phosphorylated at high cell densities, when the concentration of the extracellular pheromones, ComX and CSF (competence stimulating factor) is increasing. It then binds the *srf*-operon which encodes ComS and induces its expression (Magnuson *et al.*, 1994, Solomon *et al.*, 1995, Lazazzera *et al.*, 1997). In addition to these regulatory pathways and the control of competence development on a proteolytic level, recent studies of Boonstra *et al.* showed an importance of non-coding antisense RNAs (ncRNAs), which seem to down regulate specific genes in competent cells, such as *degU* and *shigA*. The authors identified 17 ncRNAs, and the majority was preceded by potential K-boxes, indicating ComK-dependent expression (Boonstra *et al.*, 2020).

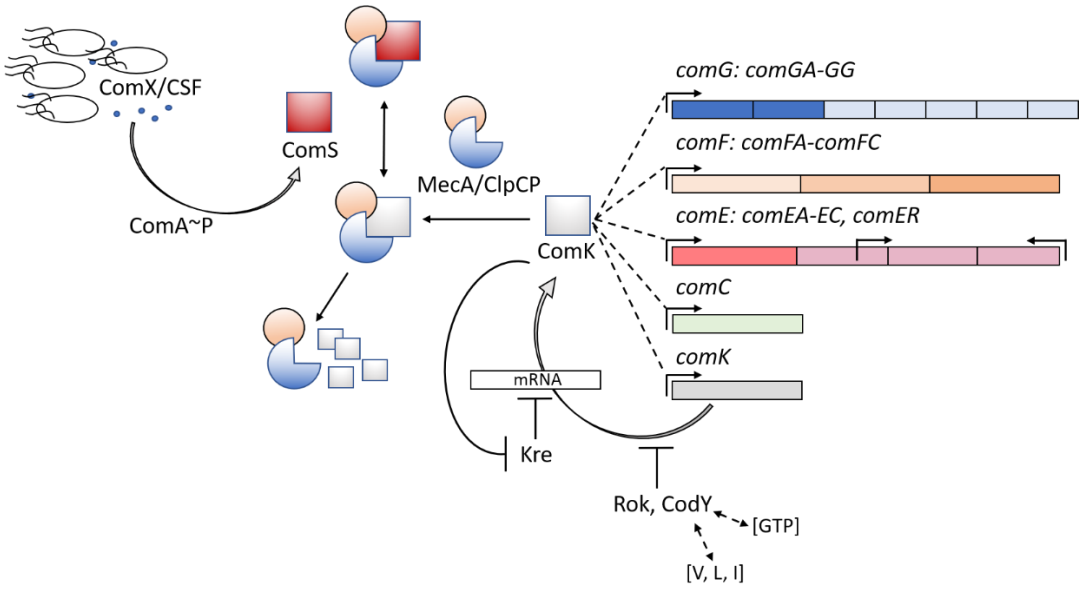


Figure 4: Regulation of expression of late competence genes in B. subtilis

Competence of *B. subtilis* is induced by quorum sensing. ComX and CSF (competence stimulating factors) are secreted in stationary phase. As a result, ComA is phosphorylated, which again induces the expression of ComS. The master regulator ComK and ComS are antagonists, competing for the binding to the MecA/ClpCP complex. ComK itself is targeted and further hydrolysed by the MecA/ClpCP complex, unless higher concentrations of ComS are produced. ComK leads to the expression of the late competence genes by inducing the operons *comC, comE, comF, comG.* The protein is autoinducing, but expression is also controlled by Rok and CodY. In addition, mRNA is destabilised by the Kre protein, whose expression is regulated by a double negative feedback loop through ComK.

2. Results

2.1 Manuscript I: The *Bacillus subtilis* dCMP- deaminase ComEB acts as a dynamic polar localization factor for ComGA within the competence machinery, *Molecular Microbiology (2020)*

Marie Burghard, Stephan Altenburger, and Peter L. Graumann[#]

SYNMIKRO, LOEWE Center for Synthetic Microbiology, Marburg, Germany, and Department of Chemistry, Philipps Universität Marburg, Marburg, Germany

KEY WORDS: Competence, DNA uptake, Single molecule microscopy, DNA recombination, *Bacillus subtilis*, moonlighting

RUNNING TITLE: ComEB plays a dual role in competence and nucleotide metabolism

ADDRESS FOR CORRESPONDENCE:

[#]Prof. Dr. Peter L. Graumann

SYNMIKRO, LOEWE Center for Synthetic Microbiology and Department of Chemistry, Philipps Universität Marburg, Marburg, Germany, Hans-Meerwein-Straße 4, 35032 Marburg, Germany. Phone number: +49-6421 28 22210

peter.graumann@synmikro.uni-marburg.de

AUTHORS CONTRIBUTION

The following manuscript was composed and written by Marie Burghard-Schrod and Peter L. Graumann. All experiments were carried out by Marie Burghard-Schrod, except for cloning and epifluorescence microscopy of *B. subtilis* fluorophore fusions of competence proteins and of corresponding deletion strains in order to study interaction of competence proteins (strains shown in Figure 1), which was carried out by Stephan Altenburger. In addition, cloning and heterologous overexpression of ComEB-YFP in *E. coli* BL21 was carried out by Stephan Altenburger (Figure 2).

2.1.1 Abstract

Bacillus subtilis can import DNA from the environment by an uptake machinery that localizes to a single cell pole, at the onset of stationary phase. We investigated the roles of ComEB and of ATPase ComGA during the state of competence by single-molecule tracking and *in vitro* assays. We show that ComEB plays an important role during competence, possibly because it is necessary for the recruitment of GomGA to the cell pole. ComEB localizes to the cell poles even upon expression during exponential phase, indicating that it can serve as polar marker. ComEB is also a deoxycytidylate monophosphate (dCMP) deaminase, for the function of which a conserved cysteine residue is important. However, cysteine-mutant ComEB is still capable of natural transformation, while a *comEB* deletion strain is highly impaired in competence, indicating that ComEB confers two independent functions. Single molecule tracking reveals a different behavior of ComEB and ComGA. While both proteins exchange at the cell poles between bound and unbound in a time scale of few milliseconds, turnover of ComGA increases during DNA-uptake, whereas the mobility of ComEB is not affected. Our data reveal a highly dynamic role of ComGA during DNA uptake and an unusual role for ComEB as a mediator of polar localization, localizing by diffusion-capture on an extremely rapid time scale, and functioning as a moonlighting enzyme.

2.1.2 Abbreviated summary

ComEB is an enzyme involved in nucleotide metabolism, which confers a second, enzyme-independent function. The ATPase ComGA, a component of the DNA uptake machinery in *Bacillus subtilis*, is localized to the cell pole by ComEB. Both, ComEB and ComGA, exchange rapidly between pole-bound and freely diffusive modes. ComGA changes its residence time during DNA uptake, revealing highly dynamic kinetics during formation of a polar multiprotein complex.

2.1.3 Introduction

Among many bacterial species, the uptake of extracellular DNA is described to serve as a source of phosphate, nitrogen, carbon and deoxyribonucleotides, to be either metabolised or directly incorporated by the cell during de novo synthesis of DNA (McDonough *et al.*, 2016, Munch-Petersen & Mygind, 1976, Kaufenstein *et al.*, 2011b). On the other hand, exogenous DNA encoding homologous regions can be integrated into the chromosome via homologous recombination, termed natural competence, as described e.g. for *Bacillus subtilis, Streptococcus pneumoniae* and the Gram-negative species *Acinetobacter* and *Helicobacter pylori* (Everitt et al., 2014, de Vries & Wackernagel, 2002, Yahara et al., 2012, Dubnau, 1991). Under nutrient starvation, a complex is formed at the cell pole, facilitating the uptake of DNA (Hahn *et al.*, 2005). The proteins involved in the formation of this complex are products of the so-called "late competence genes", encoded by the *comC*, *comE*, *comF* and *comG* operons. Expression of these genes is controlled by the transcriptional regulator ComK (van Sinderen *et al.*, 1995). Within one culture, the expression of these proteins is heterogeneous, meaning only 1-20% of cells move into the "K-state" (Mirouze *et al.*, 2012).

In *B. subtilis*, the uptake of dsDNA across the cell wall is carried out by a membrane-anchored structure called the competence pilus or pseudopilus, which is similar to the pseudopilus of the type II secretion system of Gram-negative bacteria. DNA is translocated through the cell wall by a so far unknown mechanism, possibly by pilus assembly and disassembly (Campos *et al.*, 2010, Giltner *et al.*, 2012). DNA entering through the cell wall is bound by the integral membrane receptor ComEA, encoded in the *comE* operon, and is transferred through the membrane by the permease-like channel protein ComEC (Inamine & Dubnau, 1995, Draskovic & Dubnau, 2005). Only ssDNA is translocated across the membrane (Draskovic & Dubnau, 2005), and can be integrated into the chromosome by a RecA-depended pathway (Yadav *et al.*, 2012, Ayora *et al.*, 2011), or a RecO-dependent pathway, in case a replicative plasmid has been taken up (Kidane *et al.*, 2009).

The energy needed for the assembly and disassembly of the pseudopilus may be provided by the secretion ATPase ComGA, as it is the only ATPase encoded in the

comG operon. ComGA may convert chemical energy into mechanical work by dynamic structural changes (Patrick *et al.*, 2011). For the type II secretion and the type IV pili system, a hexameric secretion ATPase is equally essential, e.g. GspE (PilB) isolated from *Vibrio cholerae* (Lu *et al.*, 2013) or EpsE in *Klebsiella* (Patrick *et al.*, 2011). All systems encode an ATPase belonging to the superfamily of AAA^+ ATPases, containing a P-loop NTPase domain and are located at the base of the pilus (Snider *et al.*, 2008). ComGA belongs to the same family of ATPases, carrying a Walker A and Walker B motif for ATP binding. A point mutation in the Walker B motif of ComGA inhibiting ATP hydrolysis was shown to be essential for DNA binding to the cell surface of *B. subtilis* (Briley *et al.*, 2011a). In addition, ComGA was found to localize in association with the competence complex at the cell pole of *B. subtilis* cells (Haijema *et al.*, 2001, Kaufenstein *et al.*, 2011a).

The function of ComEB, encoded in the middle of the *comE* operon, is not known, while the product of the homologous gene of *B. halodurans* carried out 2´-deoxycytidylate (dCMP) deaminase (dcd) activity *in vitro* (Oehlenschlaeger *et al.*, 2015). To our knowledge, no cellular function in the context of *B. subtilis* competence was annotated for this protein so far. An in-frame deletion removing 44% of the protein did not yield any phenotype, indicating that the protein is dispensable for transformation, but the authors could not exclude that a domain relevant for transformation was still encoded (Inamine & Dubnau, 1995). Several *Bacillus* species, including *Bacillus halodurans, Bacillus cereus* and *Bacillus megaterium,* encode a bifunctional dCTPase DCD:DUT, in addition to the *comEB* gene, to generate dUMP directly from dCMP (Oehlenschlaeger *et al.*, 2015). The enzyme is able to deaminate dCTP and hydrolyze dUTP to form dUMP, which then serves as a general precursor for dTTP in nucleotide metabolism (Oehlenschlaeger *et al.*, 2015, Jensen *et al.*, 2008). In *B. subtilis*, dUMP can be produced by a pathway involving ribonucleotide reductase (Reichard, 2002) and two dUTPases (YosS, YncF), in which dUTP is finally hydrolyzed to dUMP (Johansson *et al.*, 2004, Garcia-Nafria *et al.*, 2013). In addition, the cytidine deaminase (CAD) encoded by the *cdd* gene participates in the formation of dUMP (Song & Neuhard, 1989), in conjunction with a nucleotide kinase (NDK, EC:2.7.4.6). Given this redundancy of pathways for thymidine synthesis, it has been unclear why *comEB* should be expressed specifically during competence. ComEB

could play an indirect role in competence, as was shown for some SOS response genes that are induced during competence (Love *et al.*, 1985). It has also been speculated that *comEB*, encoding an internal promoter, may provide an additional ribosomal binding site to enhance transcription of the downstream gene, *comEC* (Inamine & Dubnau, 1995). Moreover, the gene product of *comEB* could be also part of a mechanism regulating the amount of ComEC in combination with ComER on the transcriptional level. *ComER* is located downstream of *comEC* and is encoded in reverse orientation to the *comE* operon, so that its expression might inhibit basal induction of the *comE* operon in the absence of ComK during early growth (Inamine & Dubnau, 1995, Draskovic & Dubnau, 2005). On the other hand, ComEB localises very distinctly at the cell pole, similar to ComGA, ComFA and ComEC, which indicates a direct interaction with proteins of the competence machinery (Kaufenstein *et al.*, 2011a, Hahn *et al.*, 2005).

In our work, we addressed the roles of ComEB and of ComGA during competence by epifluorescence microscopy, single-molecule tracking, and by *in vitro* and genetic approaches. Our aim was to see whether we could find an interaction of ComEB with other competence proteins, and we identified a link to ComGA. We show that the loss of enzymatic activity of ComEB has no influence on *B. subtilis* transformation, while the absence of the gene decreases transformation efficiency. We also studied the intracellular dynamics of ComEB and ComGA, showing that they behave differently on the single-molecule-level, although we find a function for ComEB in guiding ComGA to the cell pole.

2.1.4 Results

ComEB is necessary for the recruitment of ComGA to the pole

The competence machinery localizes to one cell pole, or rarely, to both poles and to the septum of dividing cells, which is the future cell pole [19, 27, 5] (Fig. 1A). This localization pattern is persistent even in protoplasts. While functions for several Com proteins have been described, for some of the proteins, the role during the state of competence is unknown. We set out to test if some of the competence genes without known function play a role in the recruitment of the entire machinery or of parts thereof, and generated or used (kind gifts from D. Dubnau, Rutgers, New Jersey) several deletion strains, *comEB, comEC, comFA, comFB* and *comFC*. Gene deletions were combined with strains expressing Com protein-FP fusions from the original gene locus, except the gene to be deleted and the sequence of the protein-fusion were identical, with the addition of xylose driving expression of downstream genes. We detected a failure for localization of GomGA-CFP in the absence of ComEB (Fig. 1B), but did not detect a loss of recruitment of ComFA-YFP, ComEB-YFP or ComFC-YFP proteins in the absence of all possible combinations, deleting the before mentioned genes (Fig. 1C-F, Fig. 1I, and data not shown). ComEB-YFP, ComEC-YFP, ComFA-YFP, ComFB-YFP and ComFC-YFP continued to localize at single or both poles in cells grown to competence, indicating that the machinery can cope with the lack of these proteins. Note that the *comFB* deletion may have a polar effect on *comFC*, which did not matter because of the absence of a visible phenotype for localization. We further investigated the localization of ComGA-CFP in deletions of different *com* genes. While ComGA-CFP retained clear polar localization on the absence of *comFA, comFB, comFC* and *comEC* (Fig. 1G, H, J and data not shown), ComGA-CFP showed highly aberrant localization in the absence of ComEB, with strongly increased staining throughout the cells, and much reduced fluorescence at cell poles (Fig. 1B). In contrast to wild type cells, which showed ComGA-CFP foci in 18.5% of cells grown to competence (n = 380 cells), only 3% of *comEB* mutant cells contained clear ComGA-CFP foci (n = 350 cells from 3 biological replicates). Instead, about 4% of cells contained patchy or arc-like localization of ComGA at various places in the cell (Fig.

1B, right panel). These findings reveal that ComEB plays an important role in the recruitment of ComGA to the cell poles, but not of other Com proteins investigated.

ComEB was shown to also localize to the cell pole(s) in cells grown to competence (Kaufenstein *et al.*, 2011a). Notably, cells expressing ComEB-YFP as sole source of the protein show reduced transformation activity, albeit still much more than the *comEB* mutant cells (Kaufenstein *et al.*, 2011a). Therefore, the fusion compromises the activity of ComEB (underlining its importance during competence), which must be taken into account when interpreting fluorescence data. However, because the fusion retains activity, its localization will to a large extent reflect that of the wild type protein. We wished to analyse if ComEB and ComGA colocalize, therefore we generated a strain expressing ComEB-YFP and ComGA-CFP. When grown to competence, 13% of all cells showed fluorescence signals; of these, 72% contained a single focus (containing both proteins), while 28% contained several foci (n = 530 cells) (Fig. 1 M, N, O). In cells with several foci, we found 4 cases (20% of the cells with multiple foci) in which ComGA-CFP was not accompanied by a ComEB-YFP signal (Fig. 1M-O), showing that ComEB and ComGA almost exclusively colocalize (as expected), but ComGA can also form foci in the absence of ComEB.

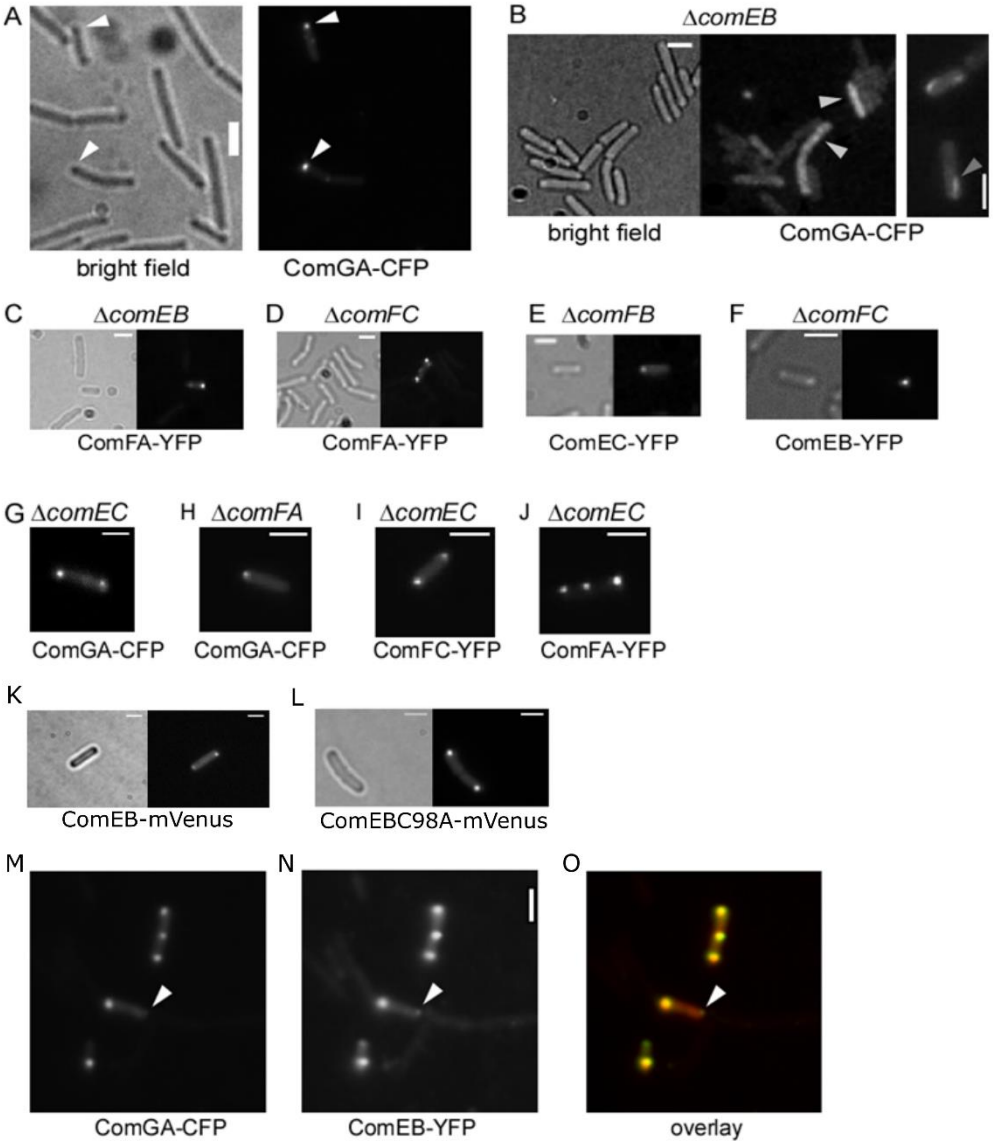


Figure 1: Localization of different competence proteins in PY79 cells grown to competence.

A) wild type cells, expressing ComGA-CFP (PG3792), B)-J) mutant cells as indicated above the images (PG3793-PG3798, PG3800, PG3814, PG3815). Fusion proteins indicated below the images. K) wild type cells expressing ComEB-mVenus (PG3786) L) wild type cells expressing mutant ComEB-mVenus (PG3787). M-O) Cells expressing ComGA-CFP and ComEB-YFP (PG3248), overlay of ComEB-YFP (green) and ComGA-CFP (red); a co-localization event results in a yellow signal. Arrowhead indicates a ComEB-YFP focus without a ComGA-CFP signal. White bars 2 µm.

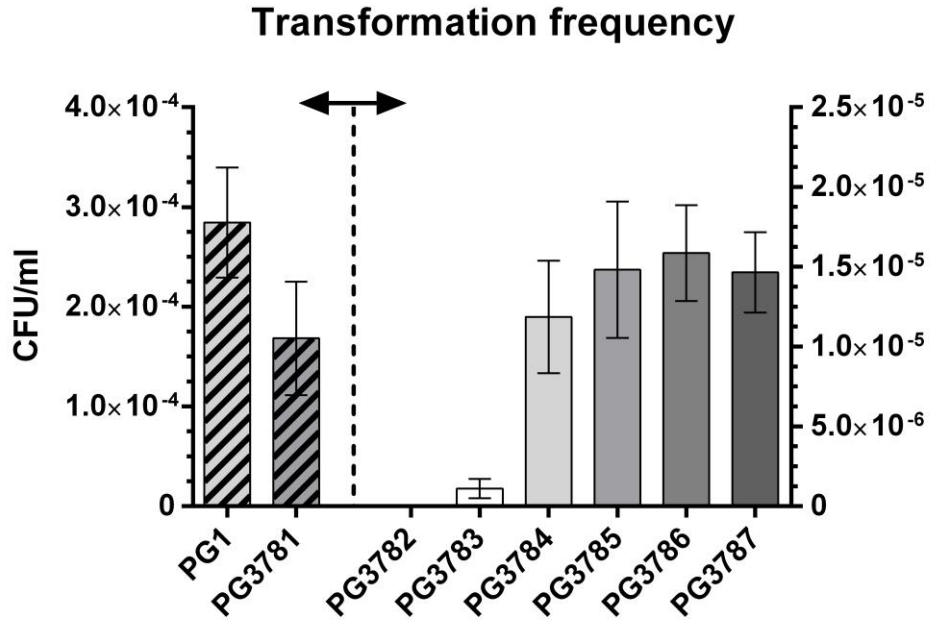
Ectopic expression of ComEB increases transformation frequency

In order to verify that ComEB plays an important role during transformation, we deleted *comEB* by insertion of a *tet* resistance cassette (PG3782). While the PY79 wild type strain (PG1) showed high transformation efficiency (2.85 x 10^{-4} CFU (colony forming units)/ml), the strain *comEB::tet* (PG3782) showed no detectable transformation activity (Fig. 2, A). We then tested a *comEB* mutant strain carrying a copy of the permease gene *comEC* under ectopic, inducible control, to exclude polar effects of the deletion (PG3783). In case *only* ComEC was expressed ectopically, transformation frequency was very low, resulting in 0.34% normalized to the wild type (1.11 x 10^{-6} CFU/ml). We also complemented the function of *comEB* by expressing *comEB* and *comEC* genes ectopically from the *amyE*-site (PG3784). Expressing both genes in the deletion background of *comEB* increased transformation frequency to 4.18% (3.5 x 10^{-6} CFU/ml), about six times higher compared to ectopic expression of only ComEC (Fig. 2 B, Fig. S3). In addition, we measured the frequency of transformation of cells expressing a catalytically inactive mutant of ComEB (PG3785), by expressing *comEBC98AEC* (cysteine 98 changed to alanine) from the *amyE*-site (see Fig. S1 and Fig. 4 for *in vitro* measurements). Cells were still transformable, reaching 5.19% normalized to the wild type strain (Fig. 2A and B).

To study the impact of the enzymatic activity of ComEB on the transformation frequency of cells and on the localization of the competence complex *in vivo* by single-molecule tracking (SMT), we created a C-terminal mVenus (mV)-fusion of the protein, and a version of it carrying a mutation at position 98 (see Fig 4, Fig. S1). We then measured the frequency of transformation in order to verify the functionality of our fusions. In case of the ComEB-mVenus fusion (PG3786) and the corresponding, inactive mutant ComEB-mV C98A (PG3787, the integration site was sequenced to insure that the mutant version was expressed as sole source of the protein), we found that transformation frequency remained in a range of 5%. Performing epifluorescence microscopy, the fusion ComEB-mV C98A still localized at one or two poles of the cells during the competence phase (Fig. 1L), comparable to the wild type ComEB-mV (Fig.1 K).

Overall, these data suggest that ComEB is not essential for transformation but plays an important role, while its enzymatic activity has no effect on transformation efficiency. Maximum level of transformants could not be achieved by ectopic expression for complementation of mutations in the *comE* operon. We detected a lower amount of colony forming units in the empty-vector-control (PG3781, about 59% of the wild type strain), indicating that the use of Spectinomycin appeared to lower the overall number of transformants, which was not the case of mVenus-Fusions (PG3786-PG3787), which were grown on Chloramphenicol as a resistance marker.

A



B

Strain	PG1	PG3781	PG3782	PG3783	PG3784	PG3785	PG3786	PG3787
Genotype	PY79 (wild type)	pDR111 x PY79	Δ *comEB*	Δ *comEB amyE:: comEC*	Δ *comEB amyE:: comEBEC*	Δ *comEB amyE:: comEBC98AEC*	*comEB-mV*	*comEB C98A-mV*
CFU/ml Mean	2.85×10^{-04}	1.68×10^{-04}	0	1.11×10^{-06}	1.19×10^{-05}	1.48×10^{-05}	1.59×10^{-05}	1.47×10^{-05}
+/- SD	5.52×10^{-05}	5.68×10^{-05}	0	6.08×10^{-07}	3.52×10^{-06}	4.28×10^{-06}	3.01×10^{-06}	2.53×10^{-06}
% Wild type	100	58.95	0	0.34	4.18	5.19	5.58	5.16
% +/- SD	19.3	19.9	0	0.21	1.24	1.5	1.06	0.89

Figure 2: Relative transformation frequency of PY79 strains modified to investigate the function of ComEB in B. subtilis.

In A), transformation frequencies are shown as bar plots. The dotted line and the black arrow on top of the line separate the columns left and right, indicating that the data were plotted on two y- axes. Three independent experiments were done (N=3), performed in technical triplicates, except for PG7383 (N=4) and PG3785 (N=5). Error bars represent the standard deviation of the mean. Plotted values are shown in B), additionally normalized to wild type level.

Artificial expression of ComEB-GFP results in polar localization in B. subtilis and in E. coli cells

We wished to analyse if ComEB requires the competence machinery to localize to the cell pole. Therefore, we constructed a strain carrying an additional copy of *comEB* tagged with *gfp* (PG3232) as a translational fusion at the integration locus *amyE* (Lewis & Marston, 1999). Cells were grown to early exponential phase, long before the state of competence is induced, and expression was induced by addition of xylose to a final concentration of 0.005 % (w/v). After continued incubation (45 min), cells expressed ComEB-GFP localized at the pole or occasionally, at midcell (septum) (Fig. 3A, note that the right panels shows several short chains of cells where cell poles are more easy to be seen), showing that ComEB can find the cell pole independent of other Com proteins, which are not expressed during this growth phase. In addition, we analysed its localization in a *comK* deletion background (PG3801), eliminating the possibility of any interaction with other competence proteins. ComEB-GFP still localized at the cell pole and the septum in this mutant background (Fig. 3B). Contrarily, ComEC-GFP or ComGA-GFP expression during exponential growth led to a homogeneous localization throughout the cells (PG3231, PG3233, Fig. 3 C and D), indicating specificity in polar localization for ComEB, but not for competence proteins in general.

To investigate whether ComEB from *B. subtilis* displays affinity to the bacterial cell pole even without species-specific polar markers, such as DivIVA (Lenarcic *et al.*, 2009), a vector was constructed with the full-length *comEB-yfp* gene under the control of a T7 based expression system (Novagen). *E. coli* BL21(DE3) 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 30 min, to reduce the likelihood of formation of inclusion bodies, and cells were prepared for microscopy. Fig. 3 E) shows that ComEB-YFP foci only appeared at the poles (always at both poles) or in the septal region (where inclusion bodies usually do not form). These experiments indicate that ComEB is a further example for a bacterial protein recognizing polar regions.

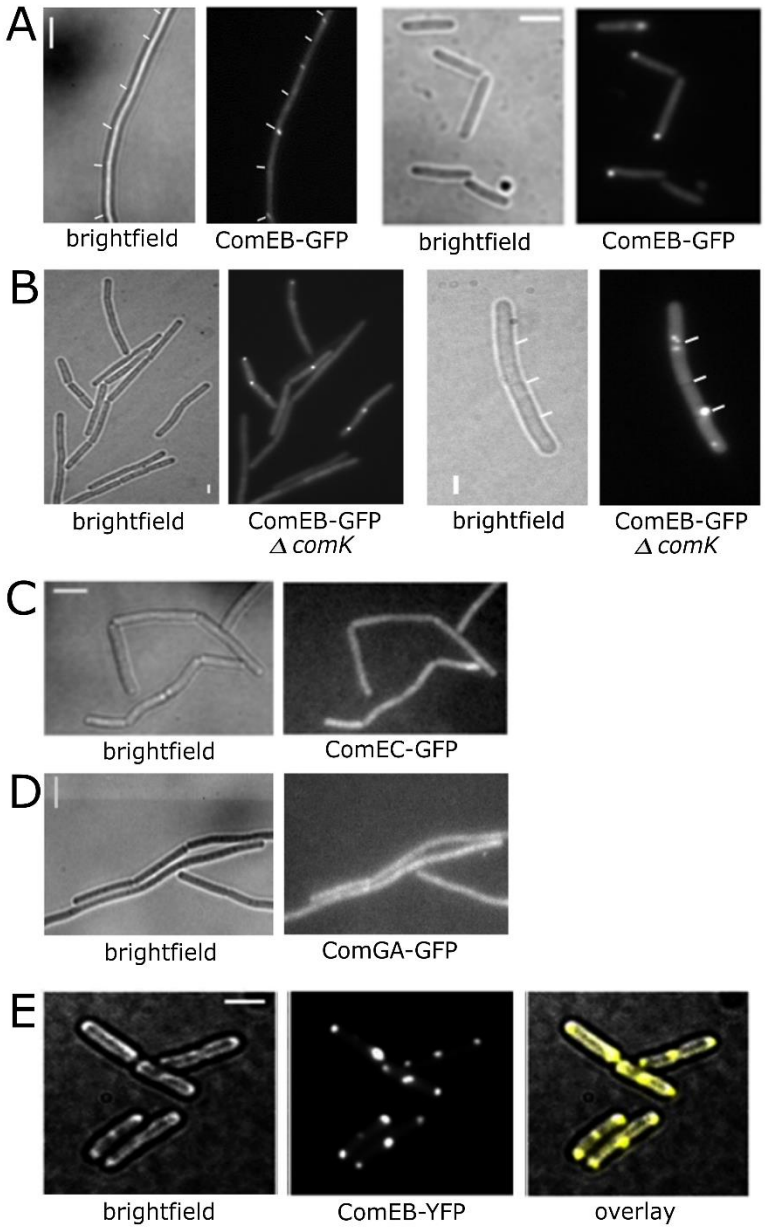


Figure 3: Artificial expression of ComEB-GFP.

A) merodiploid strain carrying comEB-gfp under the control of a xylose-inducible promoter to start expression during exponential growth, in the absence of other Com proteins in B. subtilis (PG3232). White bar 4 µm. B) ComEB-GFP localizes to the cell pole in log-phase under low-induction conditions, even in the absence of all other competence genes, when ΔcomK was deleted (PG3801). White bar 1 µm. In contrast, ComEC-GFP (PG3231) and ComGA-GFP (PG3233) localize diffusively C), D). White dashes mark cell border. White bar 4 µm. E) ComEB-YFP expressed from plasmid in E. coli BL21 (PG3806). Left panel bright

field; in the middle YFP-channel; right panel overlay. (White arrowhead, see text above) White bar represents 2 µm.

ComEB has cysteine-deaminase activity in vitro that can be inactivated at position C98

Further, we characterized the enzymatic function of ComEB in the context of the formation of *B. subtilis* competence complex *in vitro*. So far, it has not been known whether the enzymatic activity of ComEB plays a role in the formation of the competence complex or the uptake of DNA during natural transformation. In *Bacillus halodurans* strain 125, the homologue of *B. subtilis* ComEB carries out a 2´-deoxycytidylate (dCMP) deaminase activity (Oehlenschlaeger *et al.*, 2015), catalyzing the deamination of dCMP to 2´- deoxyuridylic acid (dUMP). We found that the *B. subtilis* protein remained insoluble during purification, but were able to purify ComEB of *Geobacillus thermodenitrificans*, heterologously expressed in *E. coli* BL21(DE3) cells via an N-terminal 6 x Histidin tag. Both enzymes contain the conserved nucleotide binding motifs PCXXC and HXE, as well as the dCTP binding motif GXNG, exhibiting an overall amino acid identity of 78%. An alignment of the *Geobacillus* ComEB (further referred to as $ComEB_{Geo}$) based on published crystal structures of the PDB revealed the highest similarity to the structure of *Streptococcus mutans* 2´deoxycytidylate deaminase, displaying the formation of a hexamer by three dimers, which is caused by the effector dCTP (Hou *et al.*, 2008) (Fig. S1). $ComEB_{Geo}$ was purified by affinity chromatography followed by dialysis (Fig. 4A shows 3 elution fractions), and likewise C98A cysteine mutant ComEB (Fig. 4B), and their kinetics were analysed by measuring product formation via HPLC. $ComEB_{Geo}$ showed high specificity for dCMP, while CMP was not converted (Fig. 4C). Upon addition of the effector dCTP the activity of $ComEB_{Geo}$ highly increased whereas a change of the incubation temperature of up to 42°C did not have any remarkable impact on the enzymatic activity in absence or presence of dCTP (Fig 4C). We therefore performed kinetic studies of the enzymatic activity at 30°C. Velocities were fitted to an allosteric sigmoidal curve (Fig. 5) and were found to be in a similar range than described for the *B. halodurans* ComEB, resulting in an K_m value of 0.137 +/- 0.033 mM and a V_{max}

value of 0.195 +/- 0.02 nmol/min in the presence of dCTP, and 1.745 +/- 0.283 mM and 0.195 +/- 0.03 nmol/min in the absence of dCTP. Hill coefficients were calculated to 2.73 +/- 0.7 without addition of effector and 1.7 +/- 0.56 in the presence of dCTP. To describe the conversion of the substrate without dCTP, the substrate-range of our assays had to be 3x increased. The enzyme concentration had to be 10x increased (10 µM instead of 1 µM enzyme) in order to saturate the reaction, resulting in k_{cat} values of 162.5/s in presence of dCTP and 16.25/s in the absence of the effector.

We created an inactive mutant of ComEB$_{Geo}$ by exchanging the conserved cysteine residue of the P**C**XXC motif to P**A**XXC at position 98, which is known to impair the activity but not the structural integrity of zinc- binding deaminases. The mutant ComEB$_{Geo}$C98A did not carry out any catalytic activity, even in the presence of its effector dCTP (Fig. 4D).

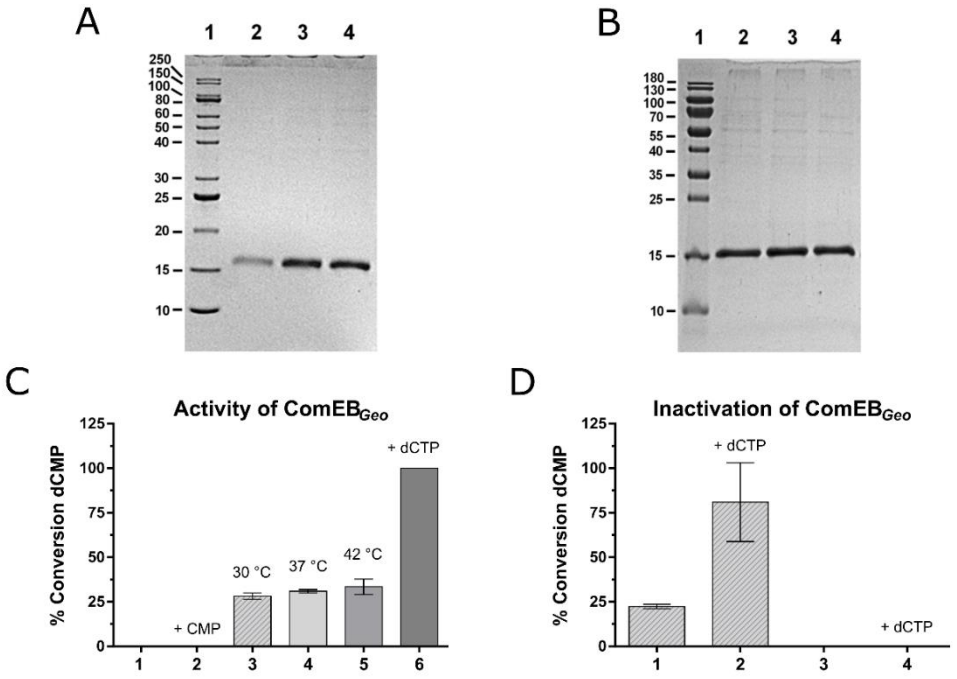


Figure 4: Ni-NTA-purification of ComEB$_{Geo}$ and ComEB$_{Geo}$ C98A purified from E. coli BL21 (DE3).

A): 1= Unstained Protein Ladder (NEB), 2-4= ComEB$_{Geo}$ elution fractions; B) 1= PageRuler prestained (Thermo Scientific), 2-4= ComEB$_{Geo}$ C98A elution fractions. Proteins were separated applying a 15% acrylamide gel and were detected at the same size of about 17.68

kDa. Activity assays were incubated in C) for 1 h, or D) 30 min and further analysed by RP-HPLC. The y- axis gives the ratio of the peak area of the product (dUMP) to the area of the substrate peak (dCMP) in percentage. Error bars indicate the standard deviation of the mean. C): 1= negative control; incubation of 1 mM dCMP without enzyme at 30°C, 37°C and 42°C. 2= 1 mM CMP, 1 µM ComEB at 30°C, 37°C at 42°C. No conversion of CMP as a putative substrate was detected. 3- 5= 1 mM dCMP, 1 µM ComEB$_{Geo}$ at the corresponding temperatures, 6 = 1 mM dCMP, 1 µM of ComEB$_{Geo}$, 0.1 mM of dCTP at 30°C, 37°C and 42°C. Upon addition of dCTP, 100% of substrate was converted in the given time at all temperatures. ComEBGeo was inactivated by mutation at position C98, shown in D): 1= 1 mM dCMP, 1 µM ComEB$_{Geo}$, 2= 1 mM dCMP, 1 µM ComEB$_{Geo}$, 0.1 mM dCTP, 3= 1 mM dCMP, 1 µM ComEB$_{Geo}$ C98A, 4= 1 mM dCMP, 0.1 mM dCTP, 1 µM ComEB$_{Geo}$ C98A.

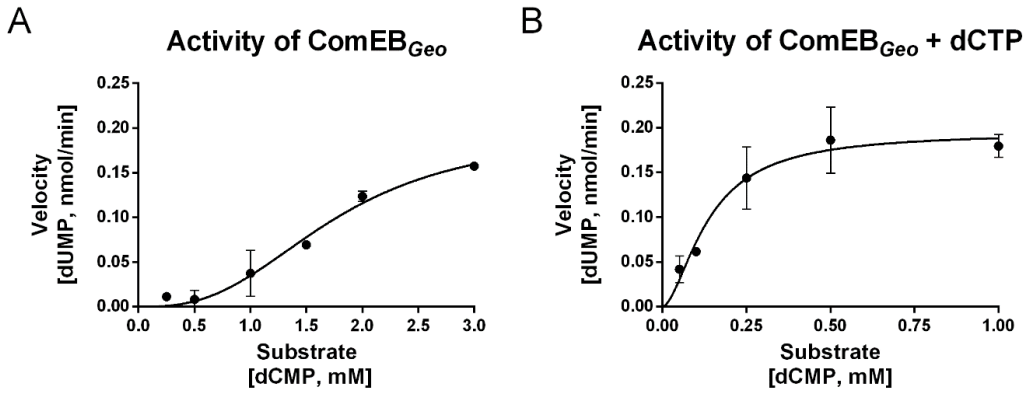


Figure 5: Enzyme kinetics of ComEB$_{Geo}$ expressed from E. coli BL21 (DE3).

In A) the activity was measured without addition of the effector dCTP, where a much higher substrate concentration was needed to finally saturate the reaction, shown on the x- axis. Activity increased dramatically upon addition of dCTP B). Note that we applied 10 µM A) compared to 1 µM B) of enzyme in the assays to be able to measure the activity without addition of effector. Therefore, the catalytic activity kcat increases 10 x within a solution of 100 x excess of dCTP (100 µM).

ComEB and ComGA form different static and mobile fractions in cells grown to competence

We employed single molecule tracking (SMT) microscopy (Schenk *et al.*, 2017, Rösch *et al.*, 2018b) to investigate the mobility of ComGA and of ComEB. To this end, we generated fusions of the proteins to mVenus (mV), which in our hands provides best results when tracking via YFP-bleaching is employed. Of note, we are not using photoactivation tracking, because *B. subtilis* cells are strongly affected in growth by blue light excitation, but not by green light (unpublished results).

ComGA-mV and ComEB-mV were tracked using 30 ms stream acquisition (movie S1), cell meshes were determined using the software MicrobeTracker (Sliusarenko *et al.*, 2011), and trajectories were determined by utrack (Jaqaman *et al.*, 2008) (movie S2). Final analysis was performed under application of SMTracker, a custom made graphical user interphase program (Rösch *et al.*, 2018a). We used apparent diffusion analyses to study the dynamics of proteins. Fig. 6A shows that the apparent diffusion data of ComEB could be best fitted using two gamma functions for two different populations, while data of ComGA could be best fitted by three gamma curves, rather than by one or two (Fig. 6B), indicating the presence of three different populations of ComGA. A diffusion constant of 0.01 +/- 0.01 $\mu m^2/s$ indicates that about 9% of ComGA-mV were statically positioned (table 1), while freely diffusive ComGA-mV had a constant of 0.61 +/- 0.02 $\mu m^2/s$. As a comparison, statically positioned chromosome condensation protein Smc has a diffusion constant of 0.012 ± 0.003 $\mu m^2/s$, measured with an almost identical exposure time of 29 ms (Kleine Borgmann *et al.*, 2013), and diffusive DnaA 0.68 ± 0.061 $\mu m^2/s$ (Schenk *et al.*, 2017), when 41 ms acquisition times were used. The slow mobile fraction (Fig. 6B, green curve) of ComGA could potentially consist of membrane-associated ComGA, because a diffusion constant of 0.17 +/- 0.01 $\mu m^2/s$ corresponds to that of a membrane-associated protein (Lucena *et al.*, 2018). However, visual inspection of acquired movies did not show a considerable membrane-associated fraction of ComGA having an intermediate diffusion constant (3.6 times slower than freely diffusive molecules). Interestingly, polar ComGA-mV foci showed strong oscillation in their intensity (movie S3): in spite of laser-induced bleaching, events of strong regain of fluorescence within few frames were apparent (Fig. 7A). After reaching of a single-molecule fluorescence level, events

of an up to three-fold increase in fluorescence were measured, detectable in 70% of 10 fluorescent complexes from three independent biological replicates (Fig. 7C), Taking into account specific recovery events, which were double in size compared to the base line of the corresponding bleaching curve, a mean of 9.3 +/- 5.5 events per complex was measured. These data reveal that ComGA molecules exchange within a rate well below 100 ms at the uptake machinery. The intermediate diffusion rate thus represents the fraction that changes between mobile and static state within the acquisition time of 30 ms. These data suggest that only a small proportion of ComGA molecules is associated with the polar DNA uptake machinery, featuring extremely rapid turnover, while a majority is diffusing through the cell.

Keeping in mind the caveat that ComEB-mV is not fully functional, we found that 51% of molecules are positioned at the cell poles, while 49% diffuse freely, with constants of 0.19 +/- 0.02 $\mu m^2/s$ or 0.604 +/-0.06 $\mu m^2/s$, respectively (Fig. 6A, table 1). In contrast to ComGA, apparent diffusion steps could be well explained by the presence of two populations, indicating that ComEB changes between free cytosolic diffusion and polar attachment, whereas it overall diffuses faster than ComGA.

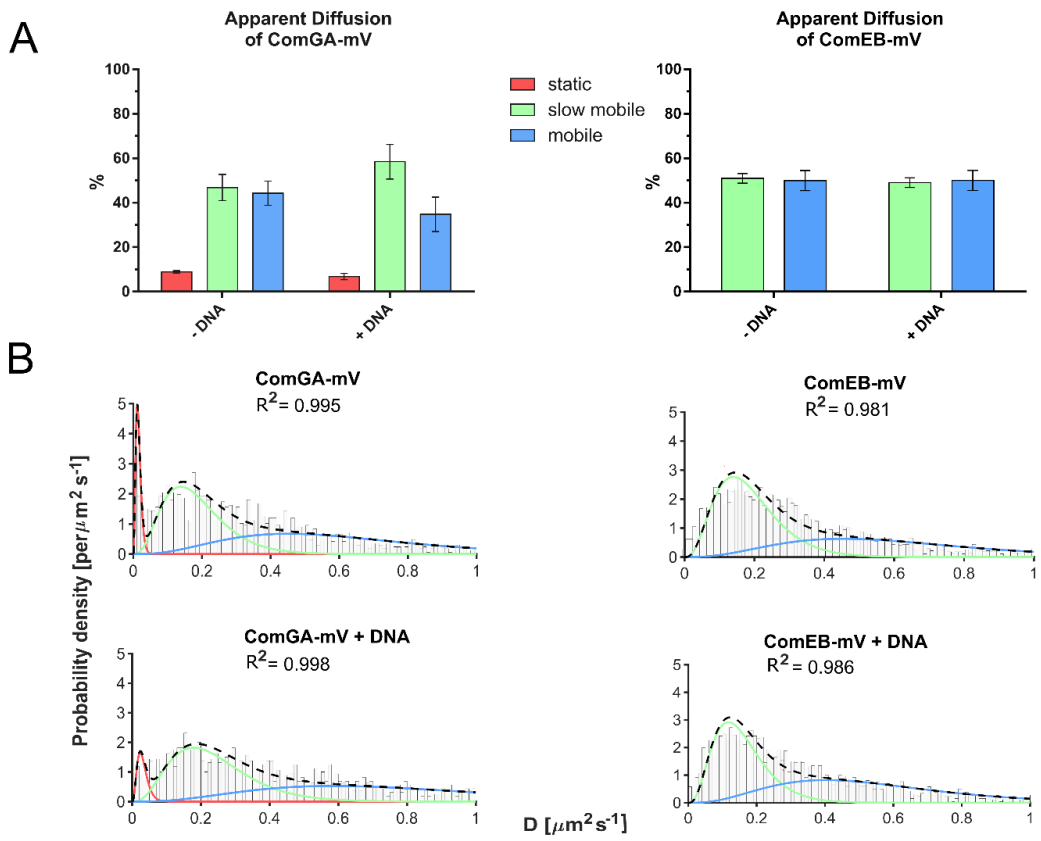


Figure 6: Apparent Diffusion- incubation with or without DNA.

The percentage of the fitted diffusive populations of ComGA-mV and ComEB-mV is depicted in A) as bar charts of a triplicate measurement (PG3788, PG3786). Bars show the percentage of the individual populations after incubation with chromosomal DNA of *B. subtilis* PY79. According to their Diffusion coefficient, populations were colored as shown in the legend (static, slow mobile, mobile). During incubation with DNA, the mobile population of ComGA-mV is decreasing, while the slow mobile fraction is increasing. For ComEB-mV, we did not observe any significant changes after incubation with DNA. Error bars indicate the standard deviation of the mean. B) shows the way values were fitted to calculate the diffusion coefficients for individual diffusive populations (ComGA-mV: N (number of trajectories) = 1559, ComGA-mV + DNA: N= 1596, ComEB-mV: N= 2745, ComEB-mV + DNA: N= 2606). The frequency of the Diffusion constants (probability density) was plotted against the specific Diffusion coefficient of each track in a Histogram. Values were fitted to Gamma-functions, in order to calculate the number of diffusive populations. The sum of the functions is depicted as black dashed line and its R^2 value is annotated as subtitle for each histogram. Diffusion of ComGA-mV can be described best as a three-populations fit with a static (red line) and one

slow mobile (green) and one mobile fraction (blue line). The diffusion of ComEB-mV is described by two functions, a slow mobile (green line) and mobile (blue line) population.

Table 1: Apparent Diffusion of ComGA-mV, ComEB-V and ComEB-mV C98A

	ComGA-mV	**ComGA-mV + DNA**
D1$_{(static)}$ µm²/s	0.01 +/- 0.01	0.04 +/- 0.01
D2 $_{(slow)}$ µm²/s	0.17 +/- 0.01	0.25 +/- 0.01
D3 $_{(fast)}$ µm²/s	0.61 +/- 0.02	0.76 +/- 0.01
static %	8.93 +/- 0.51	6.76 +/- 1.40
slow-mobile %	46.83 +/- 5.83	58.43 +/- 7.81
fast-mobile %	44.26 +/- 5.42	34.76 +/- 7.74
R²	0.994 +/- 0.001	0.998 +/- 0.001
number of trajectories	8622	7257
number of cells	106	129
	ComEB-mV	**ComEB-mV + DNA**
D1$_{(static)}$ µm²/s	0.19 +/- 0.02	0.17 +/- 0.01
D2 $_{(mobile)}$ µm²/s	0.604 +/- 0.06	0.58 +/- 0.04
static %	50.93 +/- 2.15	52.35 +/- 4.52
mobile %	49.07 +/- 2.16	47.65 +/- 4.52
R²	O.977 +/- 0.003	0.978 +/- 0.006
number of trajectories	8702	9286
number of cells	140	157
	ComEB-mV C98A	**ComEB-mV C98A + DNA**
D1$_{(static)}$ µm²/s	0.19 +/- 0.04	0.18 +/- 0.02
D2 $_{(mobile)}$ µm²/s	0.58 +/- 0.07	0.56 +/- 0.06
static %	55.05 +/- 4.8	56.93 +/- 4.2
mobile %	44.96 +/- 4.8	43.06 +/- 4.18
R²	0.996 +/- 0.002	0.998 +/- 0.001

number of trajectories	5974	6591
number of cells	62	110

Addition of DNA to competent cells induces changes in the dynamics of ComGA, but not of ComEB

Addition of DNA to competent cells has been shown to lead to a change in the localization pattern of HR proteins, in that for example, RecO accumulates at the poles upon addition of plasmid, but not of chromosomal DNA (Kidane *et al.*, 2009). As SMT is a highly sensitive method for analysing *in vivo* diffusion/binding modes, we added chromosomal DNA to cells, in order to investigate if this leads to changes within the DNA uptake machinery. Interestingly, this was the case for ComGA: in the absence of DNA, about 9% of the molecules were statically positioned, addition of DNA reduced this number to 6.8% (Fig. 8A, table 1), while the slow-mobile fraction increased about 12%, and the fast-mobile fraction decreased from 44% to 35%. These data show that a majority of the mobile molecules became slower, increasing the putative membrane- or protein-bound fraction, and a considerable number of proteins delocalized from the complex at the cell pole. This can be verified by the analyses of the likelihood of localization of molecules from an addition of all frames from movies, projected into a 1 x 1 µm large standardized cell (heat map): Fig. 8A shows that the cell poles are a major place of localization besides diffuse localization in the cytosol, which changes to a clearly more diffuse pattern at the expense of polar signal after addition of DNA (Fig. 8B). Note that a freely diffusive enzyme such as Phosphofructokinase (PfkA) homogeneously occupies the central part of the cell (El Najjar *et al.*, 2018, Rosch *et al.*, 2018a). We plotted the values of the probability for all tracks along the x-axis of the standardized cell (long axis), underneath each heat map, in order to quantify the distribution of tracks. Cells were separated into three parts, defining the polar regions of the cell as 0.3 µm from cell pole to midcell, and the septum as the middle of the normalized cell at 0.5 µm. While ComGA localized at the pole and the septum with a probability of 58.53% or 14.35%, respectively (see Fig. 8A and B marked red regions), the pattern changed upon addition of DNA to a more cytosolic localization (compare heat maps Fig. 8A and B), and the highest probability for localization changed from

8.8% at the septum, 8.5% and 10% at both poles, to a maximum value of 10% at only one cell pole. These data suggest that ComGA reacts to the addition of DNA (likely translocation of it) by moving and assembling at one cell pole, where DNA is taken up.

In case of ComEB, addition of DNA did not lead to significant changes in the static and dynamic fractions, or their diffusion constants (Fig. 6, table 1). Despite the fact that the size of the diffusive population did not change in our measurements, the probability of localization of ComEB changed after addition of DNA to a more polar pattern, with a highest probability of 7.4% at only one pole. The overall probability of localization at the pole changed between cells grown to competence with and without DNA from 49.88% to 59.15% (Fig. 8, red marked regions), indicating a clearer localization at both poles besides its diffuse localization in the cytosol (Fig. 8A, B).

Addition of DNA increases turnover of ComGA at the uptake machinery

To distinguish between the two possibilities (loss of binding sites or increased exchange), we studied the bleaching curves obtained from quantifying the intensity of polar ComGA-mV foci in cells grown to competence. Fig. 7A shows that bleaching of ComGA-mV involves large fluctuations, indicative of rapid turnover, which continues after addition of DNA (Fig. 7C). An overlay of 10 bleaching curves reveals that a) average signal intensity of ComGA-mV is lower after addition of DNA, and b) exchange is even more rapid (Fig. 7B). By fitting our data to a one phase exponential decay function, we found a decrease of the relative fluorescence intensity at the pole of 20% upon addition of DNA, based on measured Y_0 values of 19502 +/- 548.4 RFU (relative fluorescence intensity) and 15602 +/- 264.3 RFU. The rate of the bleaching curves changes from 0.076 +/- 0.012/s to 0.131 +/- 0.007/s when cells are incubated with DNA, which corresponds to an increase of the reaction rate of 72%. These experiments reveal that external DNA triggers an increase in ComGA turnover at the cell pole, possibly due to retraction/extension activity of the pseudopilus.

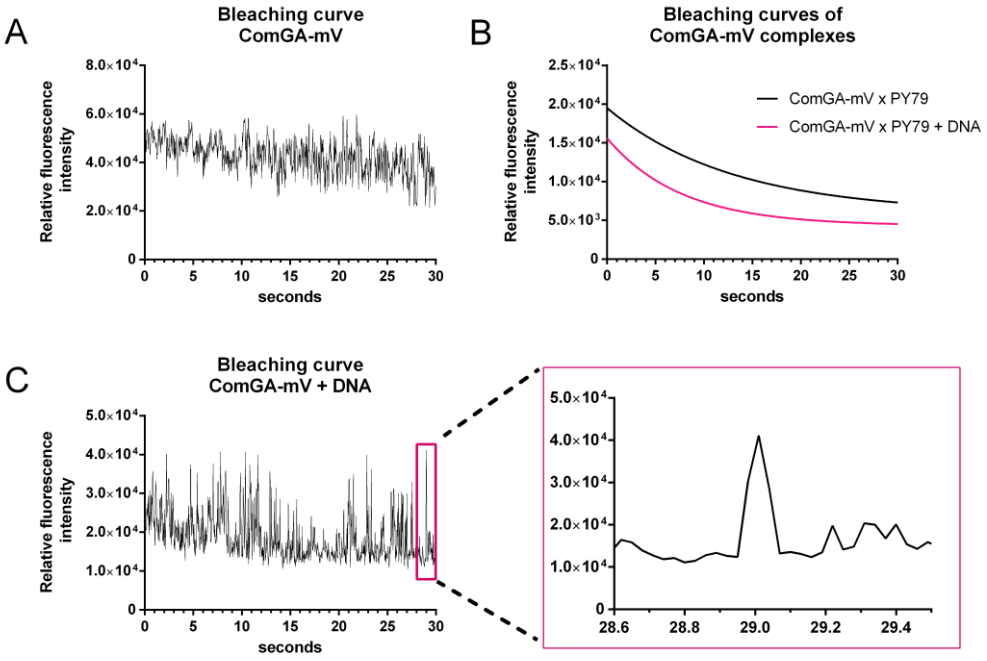


Figure 7: Bleaching curves of the polar complexes formed within the cell by our ComGA-mV fusion (PG3788).

An exemplary curve is shown in A), while addition of DNA leads to a lower total relative fluorescence intensity shown in B), where 10 complexes were fitted to a one phase exponential decay function, and a faster bleaching of the complexes. Tough complexes seemed to bleach faster upon addition of DNA, we detected distinct increased fluorescent signals, indicating enzyme activity of ComGA-mV, moving away from the competence complex and localize again, thereby recovering the signal at the pole (C), pink box). These events lasted 3-4 frames (0.09 - 0.12 sec).

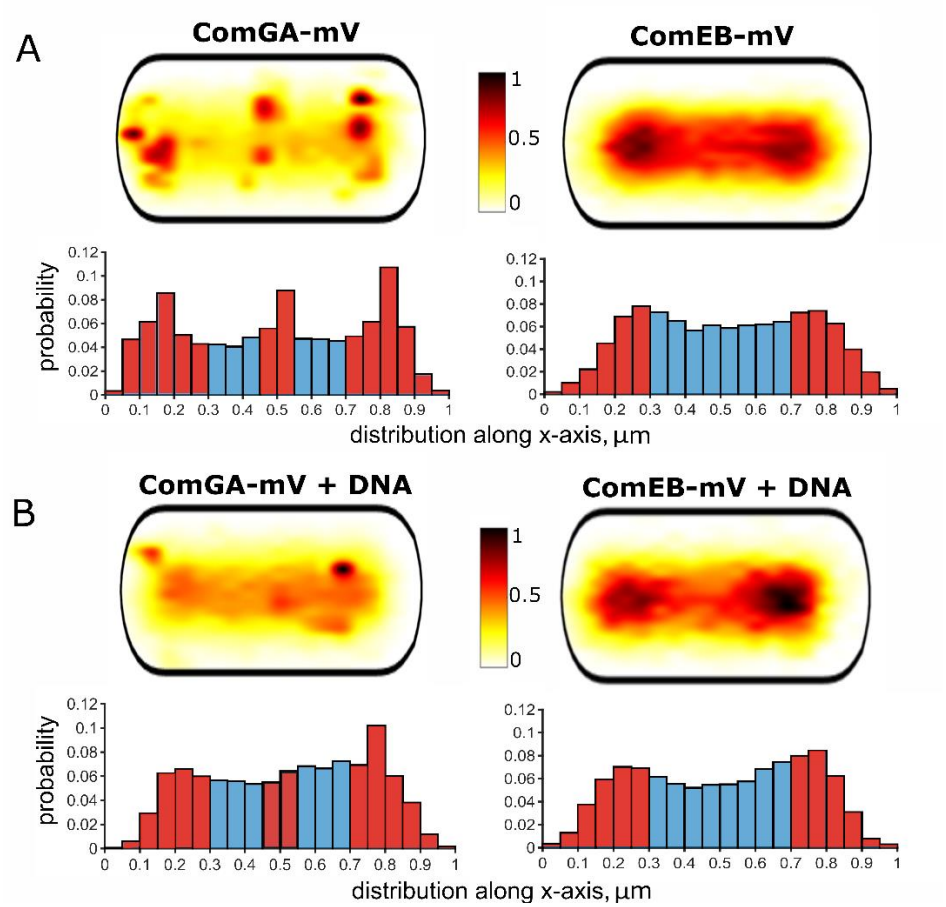


Figure 8: Heat maps and probability of all analysed tracks (see table 1) plotted against the x-axis of a normalized cell.

The localization of the tracks within the *B. subtilis* cells was analysed by Utrack and is shown here as heat maps according to the cell meshes set for tracking. Probability is indicated by a color shift from yellow to red and black (0, 0.5 and 1), depicting the highest probability for the location of the molecules in black. In order to calculate the cellular distribution of tracks, the probability of appearance was plotted in a histogram along the x-axis (meaning the long axis of the cell) according to a normalized cell of 1 µm. Polar regions were defined as 0.3 µm in size (one third of the size of a normalized cell) and septal regions (in case of ComGA-mV) were considered to be located in the middle of the normalized cell. Selected bars belonging to polar or septal region are marked in red. A) ComGA-mV fusions localize at the cell pole and the septum. ComEB-mV localizes only at the cell pole, thereby showing a higher probability for localization all over the cell compared to ComGA-mV. B) Upon addition of DNA the signal of ComGA-mV became more diffusive all over the cell, and the highest probability was only detected at one cell pole. In case of ComEB-mV, tracks were detected with a higher probability at the cell pole compared to A).

2.1.5 Discussion

Our work identifies a new player in DNA uptake during the state of competence in *Bacillus subtilis*, ComEB, which appears to serve as a polar recruitment factor for ATPase ComGA. The DNA uptake machinery plays a central role during the state of competence (Chen *et al.*, 2005), and localizes to a single cell pole during the so-called K-state. We show that the polar localization of the machinery, which consists of at least 13 proteins, is robust against the loss of ComFB, ComFC, ComEB (except for ComGA) and of ComEC, but that ComGA is delocalized in the absence of ComEB, which in turn is able to localize to the cell pole in exponentially growing cells, in cells lacking the master regulator for competence, ComK, and in *E. coli* cells. These data suggest that ComEB serves as a polar anchor for ComGA, or facilitates recruitment of ComGA to the pole. However, single molecule tracking experiments reveal that ComEB moves to and from the cell pole with exchange rates in the milliseconds range, indicating that it acts in an unusual manner, possibly, by mediating binding of ComGA to the cell pole, rather than being stably anchored at the pole by itself.

ComEB of *Bacillus subtilis* has been reported to be non-essential for transformation (Inamine & Dubnau, 1995). We confirm these data, but show that the absence of the gene significantly decreases transformation frequency. Our experiments exclude that the presence of an internal promoter is the only purpose of the gene, as suggested in previous studies (Inamine & Dubnau, 1995). In addition, we show that it has an enzymatic function for nucleotide metabolism that is not required for transformation. Therefore, we propose that the protein acquired a second function during evolution, in addition to its enzymatic function, participating in the localization of the *B. subtilis* competence complex.

The dual function of proteins (mostly enzymes) in two non-related physiological processes of the cell is termed moonlighting. Famous examples are aconitase, which acts as a citric acid cycle enzyme and also as transcriptional regulator of iron-regulated genes, or protein chaperones using ATP to refold proteins within the cytosol, acting as e.g. signaling molecules outside of the cell, after being secreted (Jeffery, 2014, Henderson & Martin, 2014). Moonlighting proteins have been found in all kinds of

organisms, and can be highly conserved in both functions, or arise in specific niches (Copley, 2012).

We show that ComEB$_{Geo}$ (from the related organism *Geobacillus*) has dCMP deaminase (DCD) activity, important for nucleotide metabolism, also shown for ComEB of *B. halodurans (Oehlenschlaeger et al., 2015)*. Our findings indicate an oligomeric structure for ComEB$_{Geo}$, based on the sigmoidal curve of enzyme kinetics, as was found for ComEB of *B. halodurans (Oehlenschlaeger et al., 2015)*. The structure of ComEB$_{Geo}$ is probably very similar to the 2´deoxycytidylate deaminase of *Streptococcus mutans*, which is the best hit in our structure-based alignments (Fig. S1) and which forms a homotrimeric structure in solution (Hou *et al.*, 2008). The presence of a hexameric structure in solution was also shown for deoxycytidylate deaminase of *E. coli* T4 phage R115E (Rami Almog *et al.*, 2004). We demonstrated that the activity of ComEB$_{Geo}$ is positively regulated by the addition of the allosteric effector dCTP, probably stabilizing substrate-binding of the subunits as described for *Streptococcus mutans* (Li *et al.*, 2016). Under 1 mM of substrate concentration and without addition of effector, enzyme velocity was almost not detectable for ComEB$_{Geo}$, which has been described 40 years ago for a DCD protein isolated from cell extracts of *B. subtilis* (Mollgaard & Neuhard, 1978). In spite of this general activity, ComEB appears to be specific for the recruitment of ComGA to the cell poles, because several other Com proteins tested still localized properly in the absence of ComEB. Additionally, several Com proteins continued to retain polar localization in the absence of ComEC, or ComFC or of ComFA, indicating relatively robust assembly of the machinery in the absence of individual components.

We employed single molecule tracking (SMT) to gain access to the question how ComEB or ComGA molecules may obtain their polar localization. We found that only 9% of ComGA molecules are statically positioned at the cell poles, while two other fractions exist that have distinct diffusion constants: a fast fraction likely representing freely diffusing molecules, and an intermediate fraction that likely consists of molecules that change their diffusive behaviour, from static to mobile, or from mobile to statically bound. This agrees with measurements of bleaching kinetics of polar ComGA-mV molecules, where general YFP-bleaching is continuously interrupted by rapid increases in intensity. This behaviour can only be explained by new molecules

becoming bound to the cell pole and others unbinding, with exchange occurring within a frame of 30 ms exposures, and even an increase from a single (fluorescent) molecule to four molecules could be observed within single frames. Thus, polar localization of ComGA is clearly achieved by a diffusion-capture mechanism.

We have previously employed FRAP to study the turnover of molecules at the cell pole, and found that only 20% of ComGA molecules show turnover, indicative of very stable binding to and slow dissociation from the uptake machinery (Kaufenstein *et al.*, 2011a). This is in apparent contradiction to the findings of this work. However, during FRAP, parts of cells are bleached by a strong laser for usually about one second or more. During this time, polar ComGA molecules have had a high turnover at the cell pole (30 molecules or more per second), such that even molecules that were diffusing through the cell will become bleached during FRAP, as they will have exchanged with polar ComGA. Therefore, for very rapid exchange reactions, FRAP is too slow to accurately determine subcellular dwell times.

Intriguingly, addition of DNA led to a considerable increase of mobility in case of ComGA, and as a result, the amount of polarly localized molecules decreased. This finding did not apply to ComEB, where the polar, slow-mobile fraction remained relatively constant, regardless of whether DNA was added or not. Addition of external DNA also increased the turnover of ComGA at the cell pole, revealing a switch-like behavior in ComGA dynamics. We speculate that increased ComGA exchange may be due to increased putative retraction/extension cycles of the pseudopilus, suggesting that it may indeed be a dynamic structure that pulls DNA through the cell wall, in analogy to the activity of *Vibrio cholerae* pili during competence (Ellison *et al.*, 2018).

It is interesting to note that ComGA plays several additional roles during the K-state: a) ComGA affects cell division via interaction with the highly conserved protein Maf, which inhibits Z-ring formation leading to slightly elongated cells in the later growth phase (Briley *et al.*, 2011b), and b) ComGA interacts with RelA, probably inhibiting hydrolysis of the alarmone (p)ppGpp by a direct interaction, which leads to rRNA transcription inhibition within the K-state (Hahn *et al.*, 2015). These functions contribute to the growth arrest of competent cells until ComGA is degraded. Thus, ComGA is a second example of a competence protein that plays multiple roles in the

cell. Release of ComGA from the poles due to higher turnover leads to a significant increase of the slow mobile fraction from 46.8 to 58.4%, which means the number of ComGA released from the pole upon addition of DNA, to either interact with other proteins, or to be loaded again with ATP, or both, is considerably different. If this turnover of ComGA is associated with changes in its ATPase cycle (which is likely), its additional roles may be connected to the DNA uptake, too. Interestingly, protein RecO assembles at the DNA uptake machinery in response to the uptake of plasmid DNA, but not of chromosomal DNA, due to an unknown trigger (Kidane *et al.*, 2012). Thus, ComGA is an example of a protein present at the pole that becomes more mobile when cells are incubated with chromosomal DNA, and ComEB a polar localization factor that also finds this subcellular site by diffusion/capture, in a highly dynamic manner that to our knowledge, has not been reported before. It will be interesting to understand these properties at a molecular detail in future studies.

2.1.6 Methods

Growth conditions

LB (Lysogenic Broth) medium was used to cultivate *E. coli* strains for protein production, for overnight cultures of *Bacillus subtilis* and for solid Agar plates. To induce competence, *B. subtilis* strain PY79 was grown at 37°C and 200 rpm in a modified Spizizen glucose minimal salt medium (10 x MC- medium competence, 100 ml: 14.01 g K_2HPO_4 x 3 H_20, 5.24 g KH_2PO_4, 20 g Glucose, 10 ml trisodium citrate (300 mM), 1 ml ferric ammonium citrate (22 mg/ml), 1 g casein hydrolysate, 2 g potassium glutamate), which was sterile filtrated and stored at -20°C in 1.5 ml aliquots (Spizizen, 1958). For transformation frequency, MC-medium (10x) was not used longer than 2 weeks. 1 ml of medium was 10 x diluted with autoclaved ddH_20 (8.7 ml) and 0.333 ml of 1 M $MgSO_4$ was added to a volume of 10 ml shortly before use. If necessary, Kanamycin was added to a final concentration of 50 µg/ml and Ampicillin to a concentration of 100 µg/ml. Erythromycin was applied in a concentration of 1 µg/ml and Chloramphenicol in a concentration of 5 µg/ml. Lincomycin was added to a final concentration of 25 µg/ml. For induction of gene expression in corresponding *B. subtilis* strains, 0.1% Xylose was added from a 50%, sterile filtrated stock solution in ddH_20. IPTG was applied at a final concentration of 0.5 mM (see fig. S3).

Strain construction

All strains and plasmids are listed and referenced in table S1. Primer are listed in table S2, named after the corresponding strain. For heterologous expression of *comEB* the nucleotide sequence of *Geobacillus thermodenitrificans (*strain *NG80-2,* GenBank: CP000557.1, gene: ABO67797.97) was amplified from genomic DNA and cloned into pET24d(+) (Novagen). An N-terminal 6 x His-tag was introduced to the sequence of the fw-Primer. For expression of ComEB-YFP in *E. coli* BL21(DE3) cells, the sequence of *B. subtilis comEB-yfp* (PG3245) was amplified and cloned into pET24d(+) to analyze its species-unspecific localization (PG3806). To create mVenus-fusions of *B. subtilis comEB* (PG3786) and *comGA* (PG3788) we used plasmid pDL-mVenus, which encodes the corresponding fluorophore *mVenus,* creating a single- crossover of

the desired gene with the introduced plasmid sequence, thereby generating a C-terminal fusion of the sequence to the fluorophore at the original locus of the gene (see Fig. S2 for expression control). To insert a homologous region needed for plasmid integration, 500 bp of the C-terminus of the desired gene were cloned into the vector next to the *linker*- and *mVenus* sequence. In case of CFP-fusions integrated at the original locus, vector pSG1164-CFP was used, following the aforementioned procedure for c-terminal fusion by homologous recombination, giving rise to strain PG3792 and PG3248. C-terminal GFP-fusions of *comEB*, *comEC* and *comGA* were ectopically integrated at the *amy*-site of *B. subtilis* (PG3232, PG3233, PG3231). For this purpose, vector pSG1193 was used. As a control strain for expression of protein-fusions, the sequence of *mVenus* was cloned into the *E. coli* vector pBbS2c by Gibson Assembly (Gibson *et al.*, 2009) and expressed in *E. coli* K12 cells (fig. S2).

In order to create strains suitable for co-localization, vector pCm::tc was used to exchange resistance *in vivo* by homologous recombination from Cm to Tet (PG3245, PG3248, PG3797-PG3800). The obtained clones were checked for Cm sensitivity and Tet resistance simultaneously. LFH (long flanking homology)-PCR was performed to delete *comEB* in *B. subtilis* and create PG3782 (Springer, 2000). By this method, the desired resistance cassette (Tet) was amplified, containing homologous overlaps to the surrounding regions of *comEB*, encoded in the primer. In a second PCR, overlaps surrounding comEB (800bp in size) were amplified and used as large Primer (see table S2). B. subtilis cells were transformed with the resulting PCR-product, leading to homologous recombination at the chromosomal region of the *comEB* gene and exchange from *comEB* to *tet.*

To accomplish the introduction of a sequence at the amy-site without fusion of the gene to any fluorophore, sequences *comEC* and *comEBEC* were cloned into vector pDR111 to complement deletion of *comEB* (PG3783, PG3784). Inactivation of the *amyE* gene was verified by selecting clones on plates containing 1% starch, followed by sequencing of the *amyE*-site. In case of a deletion of *comK*, strain PG3267 was transformed with chromosomal DNA isolated from the deletion strain, giving rise to strain PG3801. Deletion strains PG3793-PG3796 were created by transformation of chromosomal DNA of the corresponding strains.

In order to inactivate ComEB (PG3785, PG3787 and PG3790) a mutation of the gene was introduced at position 98 (Fig. S1) by vector PCR. The whole vector constructs were amplified, introducing the mutation by a modified forward primer containing the desired 2 substituted base pairs (TGC (Cys)→GCC (Ala)). The reverse primer was designed to be complementary to the downwards sequence of the vector, so that primer were designed back-to-back. The PCR product (50-100 ng) was purified, DpnI digested for 2 h followed by phosphorylation of PNK4 (polynucleotide kinase of T4 enterobacteria phage, NEB) for 1 h. Further, the linear ends of the vector were joined by blunt-end ligation (T4 ligase, NEB) over night at 25°C. All constructed vectors, and, in case of chromosomal integrated vectors or resistance cassettes, PCR-products of the desired fusion or gene, were sent for sequencing (Eurofins).

Assays of transformation frequency

B. subtilis strains (PG1, PG3781-PG3787) were grown in LB at 37°C overnight. Subsequently 10 ml of each strain were inoculated to an OD_{600} of 0.08 in MC-medium. Cells were grown to stationary phase (OD_{600} = 1.5), at 37°C and 200 rpm. 500 µl of liquid culture were distributed in test- tubes followed by addition of 1 µg DNA/ml. In all assays, chromosomal DNA encoding an Erythromycin resistance cassette was applied (PG3717). After an incubation time of 2 h at 37 °C (T2) and 200 rpm, a dilution series of cells was performed. 100 µl of culture were diluted in 900 µl LB, which was repeated, until a countable number of colonies were obtained (10^{-1}-10^{-10}). The dilutions were plated on LB-Agar containing 1 µg/ml Erytrhomycin and 25 µg/ml Lincomycin (MLS) as well as the corresponding antibiotic resistance of the strain. We found that the viable number of cells was obtained best by plating a dilution of 10^{-6} on LB Agar. Agar-plates were incubated for 2 days at 30°C. Transformation frequency was calculated by dividing the number of colony forming units (CFU= colonies/µgDNA/ml) by the number of viable cells per ml. For each strain, a technical triplicate was performed, which was repeated three times on different days.

Protein production

Genes of interest were heterologously expressed from the T7 based vector pET-24d (+) (Novagen) and proteins were produced in *E. coli* BL21 (DE3). 50 µg/ml Kanamycin were added to 4 x 500 ml of LB medium, which were inoculated from an overnight culture to an OD_{600} of 0.05. Cells were cultivated in 2 l Erlenmeyer flasks at 200 rpm and 37°C. Induction was carried out by addition of 0.4 mM IPTG at an OD_{600} of 0.5. After 4 h of induction, cells were harvested (Sorvall LYNX 6000 centrifuge, Thermo Scientific) at 4°C. Obtained pellets were frozen in liquid nitrogen and stored at -20°C until further analysis. Additional samples of 1 ml were taken before and after induction to verify protein expression. Cells were pelleted at 13000 rpm for 2 min and resuspended in 75 µl TE-buffer, pH8. 25 µl of 4x SDS buffer were added and the samples were heated for 15 min at 95°C and analysed on a 15% or 12% polyacrylamide gel (Lämmli).The OD_{600} of the samples was adjusted to an OD_{600} of 0.6 by loading the corresponding volume on the polyacrylamide gel.

Enzymatic activity of $ComEB_{Geo}$ and $ComEB_{Geo}$ C98A

Assays probing the dCMP deaminase activity of $ComEB_{\mathbf{Geo}}$ were carried out in buffer A (see protein purification) in a reaction volume of 50 µl under application of RP-HPLC as previously described (Gratani *et al.*, 2018, Steinchen *et al.*, 2018). 10 µM $ComEB_{Geo}$ (in absence of dCTP) or 1 µM $ComEB_{Geo}$ (in presence of dCTP) were incubated together with 1 mM dCMP for 60 minutes at 30/37/42°C. For a comparison with the mutant $ComEB_{Geo}$ C98A, samples were incubated for 30 min at 30°C. The effector dCTP was added in a concentration of 100 µM where indicated. Samples for enzyme kinetic analysis of $ComEB_{Geo}$ were obtained similarly. In absence of dCTP, 10 µM $ComEB_{Geo}$ was incubated together with varying concentrations of dCMP (i.e. 250/500/1000/2000/2500 or 3000 µM) at 30 °C and samples withdrawn after 12/24/36/48/60 minutes. In presence of dCTP (100 µM), 1 µM $ComEB_{Geo}$ was incubated together with 50/100/250/500 or 1000 µM dCMP at 30°C and samples withdrawn after 2/4/6/8/10 minutes. The reactions were stopped by addition of 120 µl chloroform, followed by thorough mixing for 15 seconds, heating to 95 °C for 15

seconds and flash-freezing in liquid nitrogen. While thawing, the samples were centrifuged (17300 *x g*, 15 min, 4 °C) and the aqueous phase analysed by high-performance liquid chromatography on an Agilent 1260 Series system (Agilent Technologies) equipped with a C18 column (EC 250/4.6 Nucleodur HTec 3 µM; Macherey-Nagel). Analytes were eluted at 0.8 ml/min flow rate with a buffer containing 25 mM KH_2PO_4, 25 mM K_2HPO_4, 5 mM TPAB (tetrapentylammonium bromide) and 20% (v/v) acetonitrile and detected at 260 nm wavelength in agreement with standards. The velocity of dCMP deamination was obtained by linear regression of the amount of dUMP quantified after different incubation times. The kinetic parameters (K_m, V_{max} and the Hill coefficient (h) ± standard deviation) of $ComEB_{Geo}$ were obtained from the fit of the v/S characteristic according to equation (2) using the software GraphPad Prism 6 (GraphPad Software, San Diego, California, USA). k_{cat} values were determined by dividing V_{max} by the amount of enzyme.

$$(2)\ y = \frac{V_{max} \cdot x^h}{(K_{half}^h + x^h)}$$

$V_{max} = maximum\ enzyme\ velocity,$

$k_{half} = concentration\ of\ substrate\ producing\ a\ half - maximal\ enzyme\ velocity,$

$h = Hill\ slope$

Fluorescence microscopy

A Zeiss Observer A1 microscope (Carl Zeiss) with an oil immersion objective (× 100 magnification, 1.45 numerical aperture, alpha Plan-FLUAR; Carl Zeiss) was used for wide-field microscopy. Images were acquired by a charge-coupled-device (CCD) camera (CoolSNAP EZ; Photometrics) and an HXP 120 metal halide fluorescence illumination with intensity control. Samples were prepared by growing *B. subtilis* cells under the appropriate conditions, spotting 3 µl of culture on a round coverslip (25 mm, Marienfeld). Cells were fixated with 1% agarose pads made from MC-medium by sandwiching 100 µl of the melted agarose between two smaller coverslips (12 mm, Menzel). Strains harboring a fusion to YFP or CFP were illuminated for 0.5-1 s. ComEB-YFP was excited at a wavelength of 514 nm and detected at 727 nm, ComGA-CFP at 436 nm and 480 nm. Images were processed via ImageJ (Schindelin, 2018).

Single-Molecule Tracking (SMT)

Tracking of single molecules was performed by slim-field microscopy under application of an inverted fluorescence microscope (Olympus IX71, Carl Zeiss Microscopy). A 514 nm laser diode (100 mW, Omicron Laser) was used with up to 50% of intensity to excite samples fused to mVenus in a slim-field, as well as an EMCCD camera (iXON Ultra EMCCD, Andor) to ensure high-resolution detection of the emission signal, leading to a calculated resolution of the position of the molecule of up to 20 nm. Samples were prepared by growing *B. subtilis* cells under the appropriate conditions (see method section, *growth conditions*), spotting 3 µl of culture on a round coverslip (25 mm, Marienfeld), and immobilized as described for fluorescence microscopy. Movies of 2500 frames were recorded under application of the camera´s program (AndorSolis) in an exposure time of 30 ms. For detection of signals and generation of tracks the software Utrack (Jaqaman *et al.*, 2008) was applied, where a minimal track length of 5 steps was selected. Strong fluorescence signal of the sample would lead to misconnected tracks, and to be sure of analysing movement only on a single-molecule-level, a bleaching curve of each movie was generated in ImageJ. Bleaching curves of static complexes at the pole during movies

were generated based on a manual set area in ImageJ, and further analysed under application of GraphPad Prism fitting a one phase exponential decay function (Fig. 9). At least 1000 frames were cut from the beginning of every movie. All movies were cut to a size of 1000 frames in total. For detection of signals the software Utrack was applied. Cell-meshes were set with MicrobeTracker (Sliusarenko *et al.*, 2011). Further analysis was carried out in a custom-made data analysis program, SMTracker.

Dynamics of competence proteins being fused to mVenus were analysed by growing *B. subtilis* strains to stationary phase (OD_{600}= 1.5) for 4 h in MC-medium. A volume of 100 µl of liquid culture was aliquoted in test tubes. 20 µg/ml of chromosomal DNA of *B. subtilis* PY79 was added to the sample and incubated for 1 h at 37°C and 200 rpm. Samples were analysed by SMT. As a negative control, EB buffer was added to the cells in an equal volume as the corresponding volume of DNA.

Acknowledgments

We would like to thank Wieland Steinchen for supporting our *in vitro* experiments and Andrea Becker for performing experiments on deletion strains. We thank David Dubnau of Rudgers, New Jersey, for the kind gift of *comFA and comEC* mutant cells. This work was supported by the Deutsche Forschungsgemeinschaft, DFG.

2.1.7 References

Ayora, S., B. Carrasco, P.P. Cardenas, C.E. Cesar, C. Canas, T. Yadav, C. Marchisone & J.C. Alonso, (2011) Double-strand break repair in bacteria: a view from Bacillus subtilis. *FEMS Microbiol Rev* **35**: 1055-1081.

Berge, M., I. Mortier-Barriere, B. Martin & J.P. Claverys, (2003) Transformation of Streptococcus pneumoniae relies on DprA- and RecA-dependent protection of incoming DNA single strands. *Mol Microbiol* **50**: 527-536.

Briley, K., A. Dorsey-Oresto, P. Prepiak, M.J. Dias, J.M. Mann & D. Dubnau, (2011a) The secretion ATPase ComGA is required for the binding and transport of transforming DNA. *Mol Microbiol* **81**: 818-830.

Briley, K., Jr., P. Prepiak, M.J. Dias, J. Hahn & D. Dubnau, (2011b) Maf acts downstream of ComGA to arrest cell division in competent cells of B. subtilis. *Mol Microbiol* **81**: 23-39.

Campos, M., M. Nilges, D.A. Cisneros & O. Francetic, (2010) Detailed structural and assembly model of the type II secretion pilus from sparse data. In: Proc Natl Acad Sci U S A. pp. 13081-13086.

Cardenas, P.P., T. Carzaniga, S. Zangrossi, F. Briani, E. Garcia-Tirado, G. Deho & J.C. Alonso, (2011) Polynucleotide phosphorylase exonuclease and polymerase activities on single-stranded DNA ends are modulated by RecN, SsbA and RecA proteins. *Nucleic Acids Res* **39**: 9250-9261.

Chen, I., P.J. Christie & D. Dubnau, (2005) The ins and outs of DNA transfer in bacteria. *Science* **310**: 1456-1460.

Copley, S.D., (2012) Moonlighting is mainstream: paradigm adjustment required. *BioEssays : news and reviews in molecular, cellular and developmental biology* **34**: 578-588.

de Vries, J. & W. Wackernagel, (2002) Integration of foreign DNA during natural transformation of Acinetobacter sp. by homology-facilitated illegitimate recombination. In: Proc Natl Acad Sci U S A. pp. 2094-2099.

Draskovic, I. & D. Dubnau, (2005) Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. *Mol Microbiol* **55**: 881-896.

Dubnau, D., (1991) Genetic competence in Bacillus subtilis. *Microbiol Rev* **55**: 395-424.

El Najjar, N., J. El Andari, C. Kaimer, G. Fritz, T.C. Rosch & P.L. Graumann, (2018) Single-Molecule Tracking of DNA Translocases in Bacillus subtilis Reveals Strikingly Different Dynamics of SftA, SpoIIIE, and FtsA. *Appl. Environ. Microbiol.* **84**: e02610-02617.

Ellison, C.K., T.N. Dalia, A.V. Ceballos, J.C.-Y. Wang, N. Biais, Y.V. Brun & A.B. Dalia, (2018) Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in Vibrio cholerae. *Nature Microbiology* **3**: 773.

Everitt, R.G., X. Didelot, E.M. Batty, R.R. Miller, K. Knox, B.C. Young, R. Bowden, A. Auton, A. Votintseva, H. Larner-Svensson, J. Charlesworth, T. Golubchik, C.L. Ip, H. Godwin, R. Fung, T.E. Peto, A.S. Walker, D.W. Crook & D.J. Wilson, (2014) Mobile elements drive recombination hotspots in the core genome of Staphylococcus aureus. *Nat Commun* **5**: 3956.

Garcia-Nafria, J., J. Timm, C. Harrison, J.P. Turkenburg & K.S. Wilson, (2013) Tying down the arm in Bacillus dUTPase: structure and mechanism. *Acta Crystallogr D Biol Crystallogr* **69**: 1367-1380.

Gibson, D.G., L. Young, R.Y. Chuang, J.C. Venter, C.A. Hutchison, 3rd & H.O. Smith, (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Methods* **6**: 343-345.

Giltner, C.L., Y. Nguyen & L.L. Burrows, (2012) Type IV Pilin Proteins: Versatile Molecular Modules. In: Microbiol Mol Biol Rev. pp. 740-772.

Gratani, F.L., P. Horvatek, T. Geiger, M. Borisova, C. Mayer, I. Grin, S. Wagner, W. Steinchen, G. Bange, A. Velic, B. Macek & C. Wolz, (2018) Regulation of the opposing (p)ppGpp synthetase and hydrolase activities in a bifunctional RelA/SpoT homologue from Staphylococcus aureus. *PLoS Genet* **14**: e1007514.

Hahn, J., B. Maier, B.J. Haijema, M. Sheetz & D. Dubnau, (2005) Transformation proteins and DNA uptake localize to the cell poles in *Bacillus subtilis*. *Cell* **122**: 59-71.

Hahn, J., A.W. Tanner, V.J. Carabetta, I.M. Cristea & D. Dubnau, (2015) ComGA-RelA interaction and persistence in the Bacillus subtilis K-state. *Mol Microbiol* **97**: 454-471.

Haijema, B.J., J. Hahn, J. Haynes & D. Dubnau, (2001) A ComGA-dependent checkpoint limits growth during the escape from competence. *Mol. Microbiol.* **40**: 52-64.

Henderson, B. & A.C. Martin, (2014) Protein moonlighting: a new factor in biology and medicine. *Biochem. Soc. Trans.* **42**: 1671-1678.

Hou, H.F., Y.H. Liang, L.F. Li, X.D. Su & Y.H. Dong, (2008) Crystal structures of Streptococcus mutans 2'-deoxycytidylate deaminase and its complex with substrate analog and allosteric regulator dCTP x Mg2+. *J Mol Biol* **377**: 220-231.

Inamine, G.S. & D. Dubnau, (1995) ComEA, a Bacillus subtilis integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. *J Bacteriol* **177**: 3045-3051.

Jaqaman, K., D. Loerke, M. Mettlen, H. Kuwata, S. Grinstein, S.L. Schmid & G. Danuser, (2008) Robust single-particle tracking in live-cell time-lapse sequences. *Nat. Methods* **5**: 695-702.

Jeffery, C.J., (2014) An introduction to protein moonlighting. *Biochem. Soc. Trans.* **42**: 1679-1683.

Jensen, K.F., G. Dandanell, B. Hove-Jensen & M. WillemoËs, (2008) Nucleotides, Nucleosides, and Nucleobases.

Johansson, E., J. Neuhard, M. Willemoes & S. Larsen, (2004) Structural, kinetic, and mutational studies of the zinc ion environment in tetrameric cytidine deaminase. *Biochemistry* **43**: 6020-6029.

Kaufenstein, M., M. van der Laan & P.L. Graumann, (2011a) The three-layered DNA uptake machinery at the cell pole in competent *Bacillus subtilis* cells is a stable complex. *Journal of bacteriology* **193**: 1633-1642.

Kaufenstein, M., M. van der Laan & P.L. Graumann, (2011b) The three-layered DNA uptake machinery at the cell pole in competent Bacillus subtilis cells is a stable complex. *J Bacteriol* **193**: 1633-1642.

Kidane, D., S. Ayora, J.B. Sweasy, P.L. Graumann & J.C. Alonso, (2012) The cell pole: the site of cross talk between the DNA uptake and genetic recombination machinery. *Crit Rev Biochem Mol Biol* **47**: 531-555.

Kidane, D., B. Carrasco, C. Manfredi, K. Rothmaier, S. Ayora, S. Tadesse, J.C. Alonso & P.L. Graumann, (2009) Evidence for different pathways during horizontal gene transfer in competent *Bacillus subtilis* cells. *PLoS genetics* **5**: e1000630.

Kidane, D. & P.L. Graumann, (2005) Intracellular protein and DNA dynamics in competent Bacillus subtilis cells. *Cell* **122**: 73-84.

Kleine Borgmann, L.A., J. Ries, H. Ewers, M.H. Ulbrich & P.L. Graumann, (2013) The bacterial SMC complex displays two distinct modes of interaction with the chromosome. *Cell Rep.* **3**: 1483-1492.

Lenarcic, R., S. Halbedel, L. Visser, M. Shaw, L.J. Wu, J. Errington, D. Marenduzzo & L.W. Hamoen, (2009) Localization of DivIVA by targeting to negatively curved membranes. *EMBO J.* **28**: 2272-2282.

Lewis , P. & A. Marston, (1999) GFP vectors for controlled expression and dual labelling of protein fusions in Bacillus subtilis. *Gene.*: 101-110.

Li, Y., Z. Guo, L. Jin, D. Wang, Z. Gao, X. Su, H. Hou, Y. Dong & IUCr, (2016) Mechanism of the allosteric regulation of Streptococcus mutans 2′-deoxycytidylate deaminase. *Acta Crystallographica Section D: Structural Biology* **72**: 883-891.

Love, P.E., M.J. Lyle & R.E. Yasbin, (1985) DNA-damage-inducible (din) loci are transcriptionally activated in competent Bacillus subtilis. *Proc Natl Acad Sci U S A* **82**: 6201-6205.

Lu, C., S. Turley, S.T. Marionni, Y.J. Park, K.K. Lee, M. Patrick, R. Shah, M. Sandkvist, M.F. Bush & W.G. Hol, (2013) Hexamers of the type II secretion ATPase GspE from Vibrio cholerae with increased ATPase activity. *Structure* **21**: 1707-1717.

Lucena, D., M. Mauri, F. Schmidt, B. Eckhardt & P.L. Graumann, (2018) Microdomain formation is a general property of bacterial membrane proteins and induces heterogeneity of diffusion patterns. *BMC Biol.* **16**: 97.

McDonough, E., H. Kamp & A. Camilli, (2016) Vibrio cholerae phosphatases required for the utilization of nucleotides and extracellular DNA as phosphate sources. *Mol Microbiol* **99**: 453-469.

Mirouze, N., Y. Desai, A. Raj & D. Dubnau, (2012) Spo0A∼P Imposes a Temporal Gate for the Bimodal Expression of Competence in Bacillus subtilis. In: PLoS Genet. pp.

Mollgaard, H. & J. Neuhard, (1978) Deoxycytidylate deaminase from Bacillus subtilis. Purification, characterization, and physiological function. *J Biol Chem* **253**: 3536-3542.

Munch-Petersen, A. & B. Mygind, (1976) Nucleoside transport systems in Escherichia coli K12: specificity and regulation. *J Cell Physiol* **89**: 551-559.

Oehlenschlaeger, C.B., M.N. Lovgreen, E. Reinauer, E. Lehtinen, M.L. Pind, P. Harris, J. Martinussen & M. Willemoes, (2015) Bacillus halodurans Strain C125 Encodes and Synthesizes Enzymes from Both Known Pathways To Form dUMP Directly from Cytosine Deoxyribonucleotides. *Appl Environ Microbiol* **81**: 3395-3404.

Patrick, M., K.V. Korotkov, W.G.J. Hol & M. Sandkvist, (2011) Oligomerization of EpsE Coordinates Residues from Multiple Subunits to Facilitate ATPase Activity*. In: J Biol Chem. pp. 10378-10386.

Rami Almog, Frank Maley, Gladys F. Maley, a. Robert MacColl & P.V. Roey*, (2004) Three-Dimensional Structure of the R115E Mutant of T4-Bacteriophage 2‘-Deoxycytidylate Deaminase†,‡.

Reichard, P., (2002) Ribonucleotide reductases: the evolution of allosteric regulation. *Arch Biochem Biophys* **397**: 149-155.

Rosch, T.C., S. Altenburger, L. Oviedo-Bocanegra, M. Pediaditakis, N.E. Najjar, G. Fritz & P.L. Graumann, (2018a) Single molecule tracking reveals spatio-temporal dynamics of bacterial DNA repair centres. *Sci. Rep.* **8**: 16450.

Rosch, T.C., L.M. Oviedo-Bocanegra, G. Fritz & P.L. Graumann, (2018b) SMTracker: a tool for quantitative analysis, exploration and visualization of single-molecule tracking data reveals highly dynamic binding of *B. subtilis* global repressor AbrB throughout the genome. *Sci. Rep.* **8**: 15747.

Schenk, K., A.B. Hervas, T.C. Rosch, M. Eisemann, B.A. Schmitt, S. Dahlke, L. Kleine-Borgmann, S.M. Murray & P.L. Graumann, (2017) Rapid turnover of DnaA at replication origin regions contributes to initiation control of DNA replication. *PLoS genetics* **13**: e1006561.

Schindelin, J., (2018) The ImageJ ecosystem: An open platform for biomedical image analysis - Schindelin - 2015 - Molecular Reproduction and Development - Wiley Online Library.

Serrano, E., B. Carrasco, J.L. Gilmore, K. Takeyasu & J.C. Alonso, (2018) RecA Regulation by RecU and DprA During Bacillus subtilis Natural Plasmid Transformation. *Front Microbiol* **9**: 1514.

Sliusarenko, O., J. Heinritz, T. Emonet & C. Jacobs-Wagner, (2011) High-throughput, subpixel precision analysis of bacterial morphogenesis and intracellular spatio-temporal dynamics. *Mol. Microbiol.* **80**: 612-627.

Snider, J., G. Thibault & W.A. Houry, (2008) The AAA+ superfamily of functionally diverse proteins. In: Genome Biol. pp. 216.

Song, B.H. & J. Neuhard, (1989) Chromosomal location, cloning and nucleotide sequence of the Bacillus subtilis cdd gene encoding cytidine/deoxycytidine deaminase. *Mol Gen Genet* **216**: 462-468.

Spizizen, J., (1958) TRANSFORMATION OF BIOCHEMICALLY DEFICIENT STRAINS OF BACILLUS SUBTILIS BY DEOXYRIBONUCLEATE. *Proc Natl Acad Sci U S A* **44**: 1072-1078.

Springer, Application of LFH-PCR for the disruption of SpoIIIE and SpoIIIG of B. subtilis | SpringerLink. file:///C:/Users/mt-bu/AppData/Local/Temp/Kim-Kim2000_Article_ApplicationOfLFH-PCRForTheDisr.pdf (2000)

Steinchen, W., M.S. Vogt, F. Altegoer, P.I. Giammarinaro, P. Horvatek, C. Wolz & G. Bange, (2018) Structural and mechanistic divergence of the small (p)ppGpp synthetases RelP and RelQ. *Scientific Reports* **8**: 2195.

van Sinderen, D., A. Luttinger, L. Kong, D. Dubnau, G. Venema & L. Hamoen, (1995) comK encodes the competence transcription factor, the key regulatory protein for competence development in Bacillus subtilis. *Mol Microbiol* **15**: 455-462.

Yadav, T., B. Carrasco, A.R. Myers, N.P. George, J.L. Keck & J.C. Alonso, (2012) Genetic recombination in Bacillus subtilis: a division of labor between two single-strand DNA-binding proteins. *Nucleic Acids Res* **40**: 5546-5559.

Yahara, K., M. Kawai, Y. Furuta, N. Takahashi, N. Handa, T. Tsuru, K. Oshima, M. Yoshida, T. Azuma, M. Hattori, I. Uchiyama & I. Kobayashi, (2012) Genome-Wide Survey of Mutual Homologous Recombination in a Highly Sexual Bacterial Species. *Genome Biol Evol* **4**: 628-640.

2.1.8 Manuscript I- supplementary material

The *Bacillus subtilis* dCMP deaminase ComEB acts as a polar anchor for ComGA within the competence machinery

Marie Burghard, Stephan Altenburger, and Peter L. Graumann

SYNMIKRO, LOEWE Center for Synthetic Microbiology, Marburg, Germany, and Department of Chemistry, Philipps Universität Marburg, Marburg, Germany

Supplementary movies

Movie S1: *Bacillus subtilis* cell expressing ComGA-mV grown to competence, 30 ms stream acquisition, 33 fps (real time).

Movie S2: same cell as in movie 1, tracks from selected molecules are shown in red, 5 fps (slow motion). Scale bar 3 µm.

Movie S3: *Bacillus subtilis* cell expressing ComGA-mV grown to competence, 30 ms stream acquisition, 33 fps (real time). Note that the cells contains a polar DNA uptake machinery and one localized at the division site, both of which show oscillation of signal intensity.

Supplementary figures

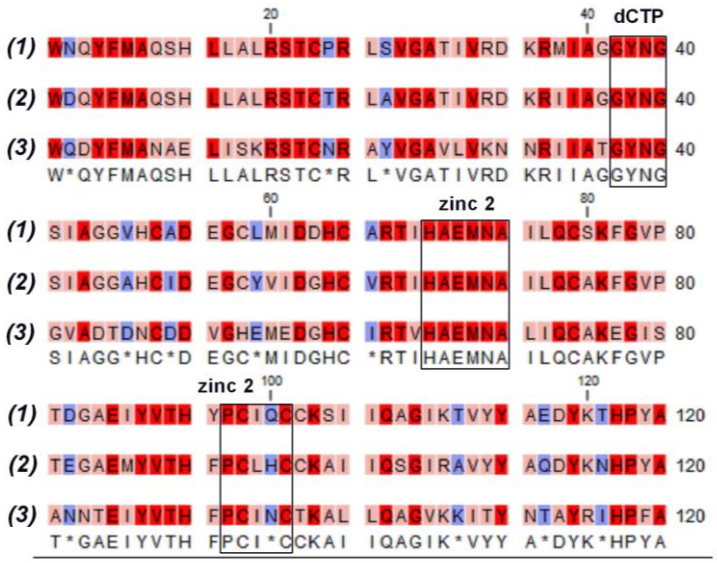


Figure S1: Selection of the alignment of B. subtilis ComEB (1), Geobacillus thermodenitrificans ComEB (2) and Streptococcus mutans dCMP deaminase (3).

Numbers on top indicate the position of the *B. subtilis* aminoacids. Conservation is indicated by dark red (100% conserved) and light red background color, while blue indicates unconserved residues. Consensus sequence is displayed underneath the alignment. The boxed areas indicate the allosterical dCTP-binding motif of the protein, and the putative binding sites of zinc 2 (catalytic site = HAEMNA or HXE, PCXXC). Note Cystein at position 98, which was exchanged for an Alanin in the mutants ComEB C98A$_{Geo}$ (PG3790) and ComEB-mV C98A (PG3787, B. subtilis).

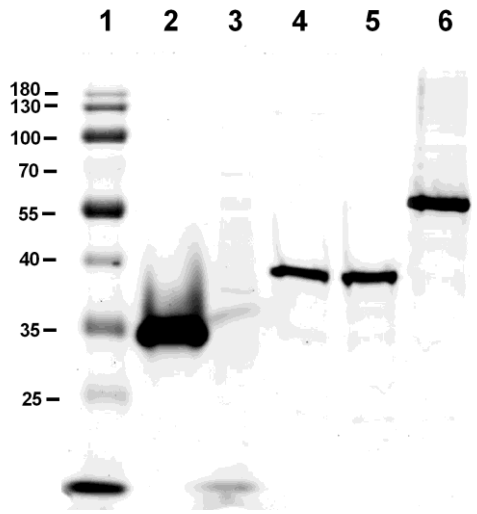


Figure S2: In-gel fluorescence of samples applied in single-molecule tracking.

Samples were run on 12% SDS-PAGE. 1= PageRuler Prestained protein ladder (NEB), 2= positive control: MVenus (PG3810), ~ 27.91 kDa, 3= negative control: *B. subtilis* PY79 (PG1), 4= ComEB-mV (PG3786), ~ 48.81 kDa, 5= ComEB-mV C98A (PG3787), ~ 48.81 kDa, 6= ComGA-mV (PG3788), ~ 68.29 kDa.

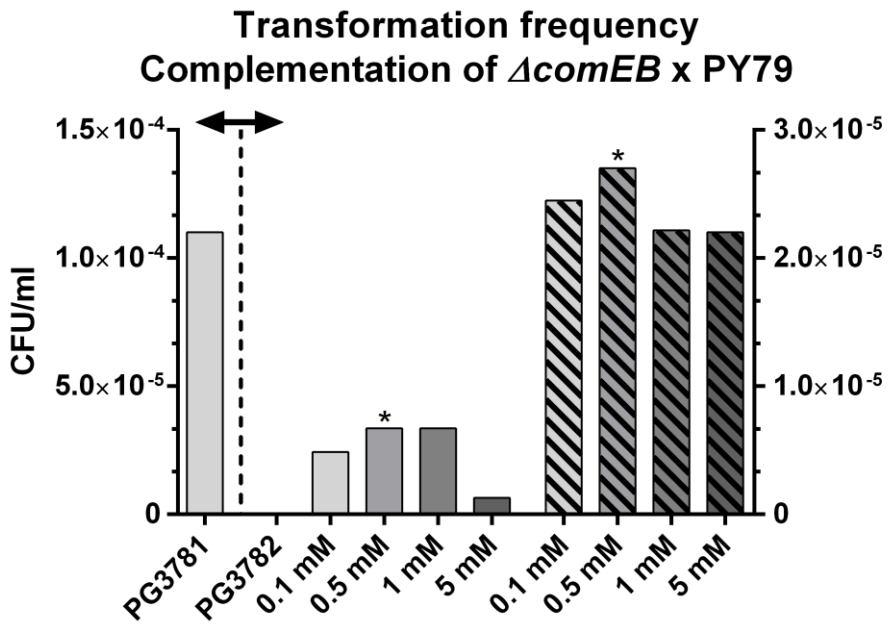


Figure S3: Transformation frequencies of strains PG3781 (empty-vector control), PG3782, PG3783 and PG3784.

Different concentrations of IPTG, depicted on the x-axis, were applied in order to complement the deletion of *comEB* (PG3782) by ectopical expression from the *amyE*-site, either by induction of *comEC* (PG3783, blank bars) or *comEBEC* (dashed bars). In case of PG3784, concentration of 0.5 mM was found to be optimal, while higher concentrations of IPTG lead to a decrease in transformation frequency. In case of PG3783, no difference was found between induction by 0.5 mM or 1 mM IPTG. Therefore, assays of transformation frequency were carried out under application of 0.5 mM IPTG (marked with asterisks).

Table S1: List of strains

Name of strain	**Relevant genotype**	**Integrated plasmid/ resistance casette**	**Resistance/ inducing agent**	**Reference**
PG1	PY79, laboratory strain	none	none	wild type
PG1305	*comFC-yfp*	pSG1164-YFP	Cm/ Xyl	Kaufenstein *et al.*, (2011)
PG3231	*amyE::*P_{xyl} *comEC-gfp*	pSG1193	Spec/ Xyl	this study
PG3232	*amyE::*P_{xyl} *-comEB-gfp*	pSG1193	Spec/ Xyl	this study
PG3233	*amyE::*P_{xyl}*-comGA-gfp*	pSG1193	Spec/ Xyl	this study
PG3235	*comFA-yfp*	pSG1164-YFP	Cm/ Xyl	Kaufenstein *et al.*, (2011)
PG3245	*comEB-yfp*	pSG1164-YFP	Cm/Xyl	Kaufenstein *et al.*, (2011)
PG3248	*comGA-cfp*, *comEB-yfp*	pSG1164-CFP cm::tet pSG1164-YFP	Tet, Cm/ Xyl	this study
PG3717	Δ*comGA*	resistance cassette(*)	MLS	Koo *et al.*, (2017)
PG3781	*amyE::*$P_{hyperspank}$*, empty vector control*	pDR111	Spec/IPTG	Ben Yehuda *et al.*, (2003)
PG3782	Δ*comEB*	resistance cassette	Tet	this study
PG3783	Δ*comEB, amyE::*$P_{hyperspank}$*-comEC*	pDR111	Spec, Tet/ IPTG	this study
PG3784	Δ*comEB, amyE::*$P_{hyperspank}$ *-comEBEC*	pDR111	Spec, Tet/ IPTG	this study

PG3785	Δ*comEB, amyE::P*$_{hyperspank}$ *- comEB_C98AEC*	pDR111	Spec,Tet/ IPTG	this study
PG3786	*comEB-mV*	pDL-mVenus	Cm/ Xyl	this study
PG3787	*comEB-mV C98A*	pDL-mVenus	Cm/ Xyl	this study
PG3788	*comGA-mV*	pDL-mVenus	Cm/ Xyl	this study
PG3792	*comGA-cfp*	pSG1164-CFP	Cm/ Xyl	this study
PG3793	*comFA-yfp,* Δ*comEB (*)*	pSG1164-YFP	Cm, Ery/ Xyl	this study
PG3794	*comFC-yfp,* Δ*comFA (**)*	pSG1164-YFP	Cm, Ery/ Xyl	this study
PG3795	*comEC-yfp,* Δ*comFB (*)*	pSG1164-YFP	Cm, Ery,/ Xyl	this study
PG3796	*comEB-yfp,* Δ*comFC (**)*	pSG1164-YFP	Cm, Ery/ Xyl	this study
PG3797	*comGA-cfp,* Δ*comEB*	pSG1164-CFP	Cm, Tet/ Xyl	this study
PG3798	*comGA-cfp,* Δ*comEC (**)*	pSG1164	Cm, Ery/ Xyl	this study
PG3799	*comGA-cfp,* Δ*comFC (**)*	pSG1164-CFP,	Cm, Ery/ Xyl	this study
PG3800	*comGA-cfp,* Δ*comFA (**)*	pSG1164-CFP	Cm, Ery/ Xyl	this study
PG3801	*comGA-yfp*	pSG1164-YFP	Cm/ Xyl	Kaufenstein *et al.*, (2011)
PG3801	*amyE::P*$_{xyl-}$ *comEB-gfp,* Δ*comk (**)*	psG1193	Spec, Kan/ Xyl	this study
PG3814	*comFC-yfp,* Δ*comEC*	pSG1164-YFP	Cm, Ery/ Xyl	this study
PG3815	*ComFA-yfp,* Δ*comEC*	pSG1164-YFP	Cm, Ery/Xyl	this study
Name	***E. coli* strain**	**Plasmid**	**Resistance**	**Reference**

PG128	DH5α	pCm::tc (BGSCID: ECE75)	Amp, Tet	Steinmetz M, Richter R (1994)
PG175	DH5α	pSG1164 (BGSCID: ECE155)	Amp, Cm	Lewis & Marsten, (1999)
PG332	DH5α	pSG1193 (BGSCID: ECE162)	Amp, Spec	Feucht A & Lewis PJ (2001)
PG429	DH5α	pSG1164-CFP	Amp, Cm	Dempwolff *et al.*; (2012)
PG430	DH5α	pSG1164-YFP	Amp, Cm	Feucht A & Lewis PJ (2001)
PG2056	DH5α	pDL-mVenus	Amp, Cm	Lucena *et al.*, (2018)
PG3099	BL21(DE3)	pET24d(+) x $comEB_{Geo}$	Kan	this study
PG3789	DH5α	pDR111 (BGSCID: ECE312)	Amp, Spec	Ben Yehuda *et al.*, (2003)
PG3790	BL21(DE3)	pET24d(+) x $comEB_{Geo}$ C98A	Kan	this study
PG3803	DH5α	none	none	ThermoFisher Scientific
PG3804	BL21(DE3)	none	none	New England Biolabs (NEB)
PG3805	DH5α	pET24d(+)	Kan	Novagen

PG3806	BL21(DE3)	pET24d(+) x *comEB-yfp*	Kan	this study
PG3807	DH5α	pBbSc2	Cm	Lee *et al.*,(2011)
PG3808	DH5α	pBbSc2 x *mVenus*	Cm	this study
PG3809	K12	none	none	Bachmann, BJ (1972)
PG3810	K12	pBbSc2 x *mVenus*	Cm	this study

(*) To create or work with these strains, chromosomal DNA was isolated from strains that were provided by (Koo *et al.*, 2017). (**) Original deletion strain from the Dubnau laboratory.

Table S2: List of primer

Primer	**Sequence 5´→ 3´**
PG3099 rev	CATGGATCCTCATGTTCCCCTCCGCC
PG3099 fw (contains His-tag)	CATCCATGGCCCACCATCACCATCACCATATGGA ACGAAT GACATGGGAC
PG3250- 1	AAGGATATATATCAAGCGGC
PG3250- 2 (overhang Tet cassette)	GAACAACCTGCACCATTGCAAGAGTTGTTCCCTC AAATGTT G
PG3250- 3 (overhang Tet cassette)	TTGATCCTTTTTTTATAACAGGAATTCTGATGAAT GCGTAATTCGCG
PG3250- 4	CGATAGCTGGAAAACCCGGC
PG3783 fw	GAGTCTGCTAGCATGCGTAATTCGCGCTTAT
PG3783 rev	GAGTCTGCATGCTTAGTTCGTCTCTGTTATATCTG AT
PG3784 fw	CTCAGAGCTAGCGTGGAACGAATTTCATGG
PG3784 rev	GAGTCTGCATGCTTAGTTCGTCTCTGTTATATCTG AT
PG3786 fw	GTAGTGGAATTCCAGAGACAAACGCATGATAG
PG3786 rev	GTGATAGGGCCCCACGTAGCTCGTGAAAAGTG
PG3787 fw	GCCATTCAATGCTGCAAATC
PG3787 rev	CGGATAATGCGTCACATAAA
PG3788 fw	GTGGTAGGTACCGAAATATTGTCACATTAGAG
PG3788 rev	GTGATAGGGCCCATCTTTTTCATGATAAACC
PG3790 fw (mutation C98A underlined)	CCTGCCTTGCATTGCTGTA
PG3790 rev	AAAATGGGTGACATACATCTC
PG3806 fw	ACTGGGCCCGTCCGAGATTATCTGTCG
PG3806 rev	ACTGAATTCCACGTAGCTCGTGAAAAG

PG3808- 1	ATGTATATCTCCTTCTTAAAAGATC
PG3808- 2 (*mVenus* underlined)	TTTAAGAAGGAGATATACATATGGTGAGCAAGG GCGAG
PG3808- 3 (*mVenus* underlined)	CCTTACTCGAGTTTGGATCCTTACTTGTACAGCTC GTCCATG
PG3808- 4	GGATCCAAACTCGAGTAAGGATCTCC
Sequencing primer	**Sequence 5´→ 3´**
***amyE*-site Sequencing fw**	ATGTTTGCAAAACGATTC
***amyE*-site sequencing rev**	ATGGGGAAGAGAACCG
mVenus/yfp rev	GTGGCCGTTTACGTCGCCG
comK fw	AATCTATCGACATATCCTGCAAAATG
comk rev	CCCGAAATTCGGACAAGG
tet^R fw	TGAGAAAATAAAGTCTTCCA
tet^R rev	GGCGCGGGTCTTGTAGTTGC
kan^R fw	ATGGCTAAAATGAGAATATCAC
kan^R rev	CTAAAACAATTCATCCAGTAAAATA

Supplemental methods

Isolation of plasmid and chromosomal DNA

Plasmid DNA was isolated from 3 ml of an overnight culture, applying a plasmid miniprep kit according to the manufactors protocol (GenElute Plasmid Miniprep Kit™, Sigma Aldrich). Chromosomal DNA of *Bacillus subtilis* was isolated from liquid culture by isopropanol precipitation. Cells were grown over night. 5 ml of culture were harvested by centrifugation. Obtained cell pellets were resuspended in 500 µl lysis buffer (100 mM Tris, 50 mM EDTA, 1% (w/v) SDS, pH 8). Subsequently, acid-washed glass beads were added to a final volume of 1.25 ml. The resuspended cells were vortexed for 2 min and the supernatant was transferred into a new vial. 7 M ammonium acetate was added (275 µl, pH 7) and the solution was incubated at 65 °C for 5 min, then further on ice 5 min. 500 µl of chloroform was added and the mixture was vortexed for 2 min. Samples were centrifuged for 5 min at 13.000 rpm and the supernatant was added to 800 µl of ice-cold isopropanol. Samples were incubated for 5 min on ice, followed by precipitation of chromosomal DNA by centrifugation (5 min., 13.000 rpm). Pellets were washed with 70 % ethanol, dried at 37 °C and resuspended in 50 µl of EB buffer (10 mM Tris, pH 8.5).

Transformation of chemically competent E. coli

Chemically competent *E. coli* cells were thawed on ice for 5 min. 1-5 ng of plasmid DNA were added and incubated further for 30 min, until cells were heat-shocked at 42 °C for 1:30 min. Aliquots were chilled on ice for 10 min, before addition of 900 ml LB and incubation at 350 rpm and at 37 °C for 1 h. Cells were plated on LB-Agar containing the corresponding antibiotic.

Transformation of B. subtilis strains

Solid and liquid medium used for transformation always contained the appropriate antibiotics and, in case the strain encoded an integrated vector, Xylose (0.1%). The application of water autoclaved glass ware seemed to be obligatory. Strains of *B. subtilis* PY79 to be transformed were streaked on a LB-broth Agar plate, and incubated for 18 h at 37°C. An overnight culture in LB was inoculated from one clone. 2 ml of a special competence medium (MC- medium competence; see growth conditions), were inoculated in test tubes, to an OD_{600} of 0.08. Cells were incubated for 3.5- 4 h at 200 rpm, until the culture reached an OD_{600} of 1.3- 1.5. 250 µl of the culture were separated and 0.5- 10 µg plasmid DNA or chromosomal DNA were added, followed by incubation of 1-2 h at 37°C and 200 rpm. Samples were plated on LB-Agar and incubated at 30°C for 2 days.

Ni-NTA affinity chromatography

Cell pellets were thawed on ice and resuspended in 30 ml of ice-cold buffer A (1 mM DTT, 20 mM Hepes, 250 mM NaCl, 50 mM KOH, 50 mM $MgS0_4$, pH 8), followed by a second incubation step on ice applying 1 protease inhibitor tablet (cOmplete ULTRA Tablets, Mini, EDTA-free, Roche Applied Science) for 1 h. DTT was always added freshly to all solutions shortly before purification. Under a pressure of 1200 psi, cells were disrupted applying a French pressure cell press (SLM AMINCO) by passing the suspension 3 times through the press. The obtained lysate was separated into the insoluble and soluble protein fraction by centrifugation at 17.000 rpm (RC 6+ centrifuge, Thermo Scientific). The soluble fraction was retained and sterile filtrated (0.45 µm Filtropur, Sarstedt). Desired 6his-tagged proteins were purified by affinity chromatography under application of the half-automated NGC (Biorad) system. After application of the soluble protein fraction to the system, a column filled with 1.5 ml Ni-NTA (Ni-NTA Superflow, Qiagen) was used to bind his-tagged proteins followed by a gradient of Imidazol in order to elute the desired proteins. Buffer A was mixed with buffer B (500 mM Imidazol, 1 mM DTT, 20 mM Hepes, 250 mM NaCl, 50 mM KOH, 50 mM $MgS0_4$, pH 8) and on the column in the following steps: 100% buffer A (wash), 15 ml; 3% buffer B, 15 ml (wash 2); 6% buffer B, 10 ml (wash3); 6-100%,

30 ml (gradient flow); 100 % buffer B, 20 ml (isocratic flow); 100% buffer A, 20 ml (isocratic flow). Elution fractions of 2 ml of the whole previous described protocol were collected. To identify the protein-containing fraction, 12 µl of each fraction showing a higher absorbance at 280 nm or 215 nm were analysed after addition of 4 x SDS buffer (4 µl) on a 15% SDS-polyacrylamide-gel. Fractions were combined and stored at 4 °C.

Preparation of concentrated protein

Proteins of interest were concentrated applying Amicon Spin columns (5MWCO, Sigma Aldrich). Columns were washed with 15 ml of ddH_2O and equilibrated with 15 ml of buffer A. Columns were filled with the obtained elution fractions from the purification (10- 12 ml) and concentrated at 3000 x *g* at 4 °C in a centrifuge (Mega Star 1.6R, VWR) to a volume of 2 ml. Further washing steps (4 x) ensured removal of residual Imidazol from the protein solution. Protein concentration was determined in a spectrophotometer (Ultrospec 3100 pro, Amersham Biosciences) at 280 nm and calculated accordingly to the specific extinction coefficient of the proteins (ε_{ComEB}= 17920 M^{-1} cm^{-1}, $\varepsilon_{ComEB\ C98A}$= 17795 M^{-1} cm^{-1}, ProtParam).

Immunoblotting and In-gel fluorescence

Proteins were transferred on a nitrocellulose membrane (transfer buffer: 2.98 g Glycin, 5.81 g Tris, 0.35 g SDS, 200 ml Ethanol, ad 1 l ddH_2O) in a semi-dry blotting device (Peqlab) under a voltage of 10 V. The membrane was blocked with 15 ml 5% fresh BSA in 1 x PBST (10 x stock solution: 11,5 g Na_2HPO_4 ,30 g NaH_2PO_4, 58.4 g NaCl, 10 ml Tween 20 ad 1 l ddH_2O, pH 7.5) buffer over night, followed by incubation for 1 h with a HRP-conjugated His-antibody (1:10.000) diluted in 5% BSA solution. After a 3 x washing step with PBST for 10 min each, the blot was developed by addition of solution 1 (2.5 mM Luminol, 0.9 mM Cumaric acid, 0.1 M Tris, pH 8.5) and 2 (6 µl of 30% H_2O_2, 0.1 M Tris, pH 8.5), 10 ml each, added simultaneously to the membrane. Bound protein was then detected by chemiluminescence reaction of the luminol (FusionSL, Analis). In case of proteins deriving from *Bacillus subtilis* cells fused to

mVenus, samples were collected from stationary phase and prepared as described previously (Bedrunka & Graumann, 2017) under mild denaturating conditions. Cell pellets were incubated at 37 °C for 20 min in lysis buffer and further on at 37 °C for 1 h in 1 x SDS-buffer. Protein fusions were detected via in-gel fluorescence (Typhoon TRIO, Amersham Biosciences) at 488 nm. The protein ladder was detected at a wavelength of 633 nm (Fig. S2).

Single molecule tracking

The mean-squared displacement (MSD, equation 1) of each track was calculated to determine the apparent diffusion coefficients (equation 2) of the movements. The probability density of each diffusion coefficient was plotted in a histogram (Fig. 6).

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

$MSD = mean\ squared\ displacement, T = total\ time, t = time\ increment,$

$x(t) = distance\ at\ a\ given\ timepoint, x_0 = starting\ position$

(2) $\delta^2(\tau) = 2nD\tau$

$\delta^2 \widehat{=} MSD, n = dimensionality\ of\ the\ system,$

$D = diffusion\ coefficient, \tau = time\ increment\ (interval)$

2.2 Manuscript II: Single molecule dynamics of DNA receptor ComEA, membrane permease ComEC and taken up DNA in competent *Bacillus subtilis* cells, *www.bioRxiv.org (2020)*

doi: https://doi.org/10.1101/2020.09.29.319830

Marie Burghard-Schrod, Alexandra Kilb, Kai Krämer and Peter L. Graumann*

Fachbereich Chemie und Zentrum für Synthetische Mikrobiologie, SYNMIKRO, Philipps-Universität Marburg, Hans-Meerwein Strasse 4, 35043 Marburg, Germany

*corresponding author: graumanp@uni-marburg.de

AUTHORS CONTRIBUTION

The following manuscript was composed and written by Marie Burghard-Schrod and Peter L. Graumann. All experiments were carried out by Marie Burghard-Schrod, except for cloning of strains ComECΔ128-mV (PG3818) and ComECΔ301-mV (PG3819) and single-molecule tracking of comEC-mV and of corresponding truncations (PG3817, PG3818, PG3819, upper cell level, Figure 6), which was carried out by Kai Krämer. Single-molecule tracking of mV-ComEA (PG3909, Figure 7 c), d), Figure 8 c), d), Figure S6) and analysis of functionality of mV-comEA (Figure S5) was carried out by Alexandra Kilb.

2.2.1 Abstract

In competent gram-negative and gram-positive bacteria, double stranded DNA is taken up through the outer cell membrane and/or the cell wall, and is bound by ComEA, which in *Bacillus subtilis* is a membrane protein. DNA is converted to single stranded DNA, and transported through the cell membrane via ComEC. We show that in *Bacillus subtilis*, the C-terminus of ComEC, thought to act as a nuclease, is not only important for DNA uptake, as judged from a loss of transformability, but also for the

localization of ComEC to the cell pole and its mobility within the cell membrane. Using single molecule tracking, we show that only 13% of ComEC molecules are statically localized at the pole, while 87% move throughout the cell membrane. These experiments suggest that recruitment of ComEC to the cell pole is mediated by a diffusion/capture mechanism. Mutation of a conserved aspartate residue in the C-terminus, likely affecting metal binding, strongly impairs transformation efficiency, suggesting that this periplasmic domain of ComEC could indeed serve a catalytic function as nuclease. By tracking fluorescently labeled DNA, we show that taken up DNA has a similar mobility within the periplasm as ComEA, suggesting that most taken up molecules are bound to ComEA. We show that DNA can be highly mobile within the periplasm, indicating that this subcellular space can act as reservoir for taken up DNA, before its entry into the cytosol.

2.2.2 Importance

Bacteria can take up DNA from the environment and incorporate it into their chromosome in case similarity to the genome exists. This process of "natural competence" can result in the uptake of novel genetic information leading to horizontal gene transfer. We show that fluorescently labelled DNA moves within the periplasm of competent *Bacillus subtilis* cells with similar dynamics as DNA receptor ComEA, and thus takes a detour to get stored before uptake across the cell membrane into the cytosol by DNA permease ComEC. The latter assembles at a single cell pole, likely by a diffusion-capture mechanism, and requires its large C-terminus, including a conserved residue thought to confer nuclease function, for proper localization, function and mobility within the membrane.

2.2.3 Introduction

B. subtilis is able to take up double-stranded DNA (dsDNA) from the environment at the onset of the stationary growth phase, e.g. under nutrient starvation conditions, and to integrate taken-up DNA into its chromosome. A multiprotein- complex is formed in 1-20% of the cells, dependent on the strain, and thus by heterogeneous expression within one culture, mediating the uptake of foreign DNA across the bacterial cell envelope (1-3). The competence-complex consists of the so-called late competence proteins, encoded by the *com*-operons *comE*, *comG*, *comF*, and *comC* (4-7). Expression of these operons depends on the transcription factor ComK (8), which is responsible for *B. subtilis* cells entering the so-called K-state, and is itself repressed by Rok (9). Rok is known to be a nucleoid associated protein, binding to A-T-rich regions, regulating a subset of genes associated with competence, but also other cell-surface and extracellular functions (10, 11). The expression of ComK increases in cells lacking *rok* from 10% to 60%, as demonstrated by GFP-fusions of ComK, leading to a much higher amount of cells entering the K-state (9).

Studies using epifluorescence-microscopy have shown that most of the major proteins essential for DNA-uptake localise to the cell poles in competent *B. subtilis* cells, such as ComGA, ComGB, ComEC, ComEB (12-14). It has been also shown that fluorescently stained dsDNA localises to the cell poles when incubated with competent *B. subtilis* cells. The labeled DNA can be integrated into the chromosome, leading to the formation of colonies displaying the particular resistance encoded by the transforming DNA (tDNA, (15)). These findings support the idea of a DNA-uptake complex located at the cell pole. Comparing data from other naturally competent bacteria, like the gram-negative *Vibrio cholerae* and the gram-positive *Streptococcus pneumoniae*, the concept of an uptake apparatus consisting of the competence complex and a type-II-/ type IV-secretion like pseudopilus has evolved in the field (16-18). In case of *B. subtilis*, its major component is thought to be the pseudopilin ComGC, facilitating the uptake of the tDNA through the cell wall (19).

The dsDNA-binding protein ComEA is an integral membrane protein that localizes throughout the membrane in a punctate manner (12, 14). It has been shown to be essential for binding of tDNA, carrying out the function of a DNA receptor during

transformation (20, 21). An orthologue of ComEA, ComH of *Helicobacter pylori*, has recently been found to directly interact with the N-terminus of ComEC, the aqueous channel protein, which transfers the tDNA through the membrane into the cytosol (22, 23). ComH hands over the bound dsDNA to ComEC, while in case of *B. subtilis*, this role is taken over by ComEA. After being bound to ComEC, tDNA passes the channel as single-stranded DNA (ssDNA)(23), probably driven by a proton symport and the DEAD box helicase ComFA, providing energy by ATP hydrolysis for the process of transfer (6, 24, 25). Entering the cytosol, the exogenous DNA is coated by the ssDNA binding proteins DprA, SsbB and SsbA (26, 27), followed by integration into the chromosome through recombination proteins, RecA or RecN (13, 28, 29). How the ssDNA is generated inside the *B. subtilis* periplasm before entering the channel is still a matter of current investigations (18). In case of *Streptococcus pneumoniae*, the incoming dsDNA is hydrolysed by the membrane associated nuclease EndA (30), while for *Bacillus*, such a nuclease has not been discovered yet. One theory assumes that the late competence-protein ComEC takes over the aforementioned function (18).

ComEC is an integral membrane protein, possessing either 1 putative amphipathic helix, laterally associated with the membrane and 7 transmembrane helices according to Draskovic & Dubnau, (2005) (23), or possibly 11 transmembrane helices according to recent modelling studies of Baker *et al.* (2016). Draskovic & Dubnau (23) have investigated the membrane topology of ComEC by fusing different parts of the protein, either to LacZ or PhoA, determining the cytosolic or periplasmic localization of protein domains based on the resulting enzyme-activity of the truncated LacZ-or PhoA- fusions. As a result, two soluble, periplasmic parts were identified. The N-Loop, located at the N-terminus of the protein, and the C-Loop, located at the C-terminus of ComEC. In addition, crosslinking-experiments of membrane protein fractions of competent *B. subtilis* cells were performed, identifying the formation of a dimer, probably by interaction of the two periplasmic parts (23). Recent *in silico* studies revealed the presence of a putative OB-fold, located at the N-terminus of the protein, which might be capable of ssDNA-binding, and a second domain at the C-terminus, putatively exhibiting exonuclease-function (31). These findings support the idea of ComEC providing the putative exonuclease among the competence-proteins, degrading the incoming DNA in order to pass the cell membrane. The ß-lactamase-

nuclease- like domain of ComEC contains 2 conserved zinc-binding motifs, which are located within the periplasmic part of the protein, already described earlier by Draskovic *et al.* (2005). The model of Baker *et al.* (2016) (31) supports the existence of a large, soluble C-terminal part.

Earlier *in vivo* studies of ComEC report that in epifluorescence microscopy, the overall signal of a C-terminal YFP-fusion of ComEC is very weak or undetectable, indicating low abundance of the protein (12, 26, 32). Fluorescence-microscopy of two other competence proteins essential for DNA-uptake, ComGA and ComFA, revealed a similar localization-pattern at the cell pole in competent *Bacillus* cells, combined with a non-uniform signal all over the cell. This prompted us to investigate the localization of ComEC again by Epifluorescence-microscopy, but in addition by Single-molecule-tracking (SMT), in order to understand how its characteristic, polar localization pattern is achieved. To find out how the DNA uptake into the cell takes place, we truncated the C-terminal soluble part of the protein. In the absence of its soluble part, ComEC localises throughout the membrane, and the percentage of a population expressing ComEC increased. Removing its hydrophilic, C-terminal part slowed down the diffusion of the protein, indicating that ComEC assembles during competence of *B. subtilis* by diffusion/capture. In addition, analysing mutations of ComEC *in vivo* by determining transformation frequencies of several mutants revealed that the presence of a conserved aspartate residue, located within a putative zinc-binding motif of ComEC, is required for transformation. We have been able to determine the dynamics of taken up DNA at a single molecule level, and compared it to the mobility of ComEA, finding a similar diffusive behaviour. Our data suggest that exogenous DNA is directly bound by ComEA and can move within the periplasm before it is being transported across the membrane.

2.2.4 Methods

Growth conditions

Escherichia coli (E. coli) cells were grown in lysogeny broth (LB) medium and on LB-Agar-plates (1.5% Agar), supplemented with 100 µg/ml Ampicillin, at 37°C and 200 rpm. *B. subtilis* cells were grown on LB-Agar plates and in LB-medium at 30°C, supplemented with the required antibiotics, applying the following final concentrations: 5 µg/ml Chloramphenicol (Cm), 100 µg/ml Spectinomycin (Spec), 25 µg/ml Lincomycin and 1 µg/ml Erythromcyin (MLS), 25 µg/ml Kanamycin (Kan). In case an Amylase-assay had to be performed, LB-Agar was supplemented with 1% starch. If necessary, gene expression was induced by addition of either 0.05% Xylose (50% stock solution in ddH_20, sterile filtrated, used for C-terminal and N-terminal *mVenus*-fusions at the original locus) or 0.5 M IPTG (1 M stock solution in ddH_20, sterile filtrated, used for ectopic expression from the *amyE*-site), which was already included in the Agar-plates. For transformation of *B. subtilis*, a modified competence medium (MC-medium) was used after Spizizen, and cells were grown at 37°C, 200 rpm (33). A volume of 100 ml 10x MC-medium was composed as follows: 14.01 g K_2HPO_4 x 3 H_20, 5.24 g KH_2PO_4, 20 g Glucose, 10 ml trisodium citrate (300 mM), 1 ml ferric ammonium citrate (22 mg/ml), 1 g casein hydrolysate, 2 g potassium glutamate. Medium was sterile filtrated and aliquots of 1.5 ml were frozen at -20°C until use. Shortly before use, 1 ml of 10x MC-medium was added to a volume of 10 ml with 8.7 ml of sterile ddH_20 and 1 ml of 1 M $MgCl_2$ (autoclaved) and supplemented with the corresponding antibiotics (or Xylose, if required). Frozen aliquots were not used longer than a week, as we found a loss of activity (transformation frequency) after long-term storage.

Strain construction

All *B. subtilis* strains are listed and referenced in table S1. Strains of *E. coli* are listed in table S2 and corresponding Primer are listed in table S3. The C-terminal mVenus-fusion of ComEC (PG3817) was created by restriction-based cloning of the last 500 bp of the gene *comEC* into the vector pDL-mVenus via *Bgl*II and *Apa*I. pDL-mVenus

encodes a MCS, followed by a 14 aa linker (GLSGLGGGGGSL) and *mVenus* (34). Inserting a *Bacillus* gene at the MCS of at least 500 bp, homology to the host genome is created, leading to a single-crossover of the vector with the *B. subtilis* chromosome. The gene of interest is thereby fused to the fluorophore at its C-terminus, still being located at its original locus. For C-terminal truncations of ComEC (PG3818, PG3819), the final protein sequence was chosen to be either 128 aa or 301 aa shorter than the full-length protein, removing either 1 putative helix and 1 zinc-binding-motif at position 667-672, aa: GDLEKE (PG3818), or two putative transmembrane helices and the C-terminal, soluble part of the protein, which is located in between the helices, and includes *both* putative conserved zinc-binding motifs at position 667-672 and 571-576, aa: GDLEKE and HADQDH (PG3819, (23, 31). Therefore, regions of 384 bp and 903 bp, starting from the C-terminus of *comEC*, were truncated from the template sequence for cloning. 500 bp of homology of the C-terminus were chosen, amplified and ligated to pDL_mVenus via *Eco*RI and *Apa*I restriction. After transformation of *B. subtilis* (PG001, see methods, transformation of *B. subtilis*), chromosomal integration was verified by PCR (see table S3, Primer EC fw and mV rev).

In order to investigate mutations of ComEC, the sequence of the gene was amplified and cloned via *Nhe*I and *Sph*I into pDR111 for ectopic expression at the *amyE*-site. In order to mutate amino acids H571, D573, D575, D668, and K671, plasmid DNA was isolated (GenElute Plasmid Miniprep Kit^{TM}, Sigma Aldrich) and further amplified by vector-PCR. Primer-pairs were designed back-to back with a size of 20 bp each, including the mutation at the beginning of the forward primer. through this technique, either one or two bp were exchanged, mutating the original amino acid to an alanine (used codons: *gcc, gca, gct*). Linearized plasmid DNA was phosphorylated and ligated (for detailed protocol see (35)), followed by transformation of *E. coli* DH5α. Plasmids were purified and sent for sequencing (Eurofins Genomics) to verify the presence of the desired mutations. In a final step, 24 bp encoding an RBS were added n-terminally to all constructs by a second vector-PCR (sequence, including **SD**: GATTAACTAATA**AGGAGG**ACAAAC (36, 37). Presence of the RBS was then verified by sequencing again. *B. subtilis* cells (PG001) were transformed by the resulting constructs. Ectopic integration at the *amyE*-site was proven by a 3x amylase-assay on starch-Agar. Therefore, plates were covered with 5 ml of Iodine-potassium

iodide solution (Carl Roth) and incubated for 10 min at room temperature. Each time the assay was performed, a positive and a negative control were included. Finally, strains were transformed with chromosomal DNA of a PG3722, genotype: *ΔcomEC*. This resulted in strains encoding either the original gene or a mutated version of *comEC* at the *amyE*-site, and an erythromycin resistance cassette at the original locus instead of *comEC* (PG3917-PG3919, PG3927-PG3929). Presence of the resistance cassette was verified by PCR (see Primer Ery fw and Ery rev, table S3).

To create an N-terminal mVenus-fusion of ComEA (PG3909), 500 bp of the 5' sequence of *comEA* were amplified and cloned into pHJDS-mV via *Apa*I and *Eco*RI, followed by transformation of *B. subtilis* PY79 (PG001). The original plasmid pHJDS1 (38) allows Xylose- inducible expression of N-terminal fusions of GFP at the original locus. In this case, *gfp* was exchanged for *mVenus* by restrictions-based cloning (PG3692, lab strain, unpublished, modified by Lisa Stuckenschneider via *Kpn*I and *Bam*HI). To facilitate SMT tracking of labeled DNA, we used a strain encoding a deletion of *rok* (PG876). In this background, *comEC* was deleted by transformation with chromosomal DNA of PG3722 to generate strain PG3816. Integration of the erythromycin resistance cassette was verified by PCR.

Transformation of E. coli

To generate chemical competent *E. coli* DH5α cells for cloning procedures, a 200 ml culture was grown in LB medium at 37°C and 200 rpm, centrifuged at an OD_{600} of 0.6 and 3000 rpm, followed by resuspension in 15 ml of sterile filtrated, ice-cold 0.1 M $CaCl_2$ solution, supplemented with 15% Glycerol. Aliquots of 150 µl were frozen in liquid nitrogen and stored at -80°C until use. For transformation, an aliquot was chilled on ice for approximately 5 min. 1 ng of plasmid DNA was added to cells, and incubated for 30 min on ice. After a heat shock for 90 sec at 42°C, aliquots were chilled on ice for another 10 min, followed by addition of 800 ml LB- medium and incubation at 37°C for 1 h (350 rpm). Cells were then plated on LB-Agar plates supplemented with the corresponding antibiotic and incubated at 37°C for 16 h.

Transformation of B. subtilis & assays of transformation frequency

For transformation of *B. subtilis*, the application of glassware was obligate. Best results were obtained when strains were freshly streaked on LB-Agar, supplemented with inducing agents (0.05% Xylose, 0.5 M IPTG) and antibiotics, depended on the inserted plasmids or resistance cassettes, followed by incubation at 37°C over-night *before* transformation. From these LB-Agar plates, a 3 ml LB-over-night culture was inoculated and grown at 37°C and 200 rpm. 0.5 ml MC-medium, supplemented with antibiotics and required inducing agents (see methods, growth conditions) was inoculated to an OD_{600} of 0.08, and grown at 200 rpm until the culture reached an OD_{600} of 1.5. Subsequently, 0.5-1 µg of chromosomal DNA, purified by isopropanol precipitation (35), was added to the cells and incubated further for 1-2 h at 37°C and 200 rpm. Cells were then plated on LB-Agar, containing the required antibiotics and incubated at 30°C for 48 h.

In order to perform assays of transformation frequency, 0.5 µg of chromosomal DNA were added to 0.5 ml of competent *B. subtilis* cells at an OD_{600} of 1.5 in a final concentration of 1 µg/ml. Further, cells were incubated for exactly 1 h at 37°C and 200 rpm. 100 µl of culture were diluted in 900 µl of LB-medium, followed by creation of a dilution-series in LB-medium. For calculating the number of viable cells/ml, a 10^{-6} dilution was plated on LB-Agar, while, depending on the strain, a 10^{-1}, 10^{-2} or 10^{-3} dilution was plated in order to count colony forming units/ml (CFU). Finally, transformation frequency was calculated by dividing the CFU/ml by the number of viable cells/ml/µg DNA. For each strain, a technical triplicate and a biological duplicate was performed. Data were visualized via GraphPad Prism6 (GraphPad Software, San Diego, California, USA).

Fluorescence staining of DNA

Preparation of labeled DNA was carried out following the exact protocol of Boonstra *et al.,* (2018). The modified nucleotide 5-[3-aminoallyl]-2'-deoxyuridine-5'-triphosphate (aminoallyl-dUTP, Thermo Scientific™) was incorporated by PCR using DreamTaq™ DNA Polymerase (Thermo Scientific™). A volume of 100 µl of dNTP-stock solution containing the modified nucleotide was set up as follows: 10 µl dATP (100 mM), 10 µl dGTP (100 mM), 10 µl aminoallyl- dUTP (50 mM), 10 µl dTTP (50 mM), 10µl dCTP (100 mM) 50 µl ddH_20. A 2300 bp fragment was amplified from the plasmid pDG1664 (PG311), using Primer prMB013 and prMB014 (15). The resulting PCR-product encoded for the erythromycin-resistance cassette of the vector including its flanking regions of the *thr*C-site, creating homology to the *B. subtilis* chromosome. PCR-reaction was carried out in a final volume of 400 µl. Subsequently residual template was removed from the PCR-reaction by a *Dpn*I digest for 1.5 h at 37°C. PCR product was purified (Qiagen PCR purification Kit) and eluted in 50 µl of 0.1 M $NaHCO_3$ solution, pH 9. For staining of DNA, the fluorescent dye DyLight488 NHS Ester was used (Thermo Scientific™). In order to allow the aminoallyl-group of the incorporated nucleotide to react with the stain, facilitating the formation of an aminoallyl-ester, a 10x excess of dye was added to the PCR-product and incubated for 3 h in the absence of light at 25°C. Amount of dye was calculated as follows:

(1) $\frac{\mathrm{m}\,[DNA]}{\mathrm{M}\,[DNA]} \times 10 \times \mathrm{M}\,[fluorophor] = \mathrm{m}\,[fluorophor]$

m= mass, M= molecular mass

Further, labeled DNA was purified from excess stain by PCR purification including an additional washing step (80% ethanol) and eluted in 50 µl ddH_20. In order verify staining and removal of residual dye, labeled PCR- product was detected in a 1% Agarose-gel by in-gel fluorescence at 488 nm (Typhoon TRIO, Amersham Biosciences, see figure S3). Aliquots were stored at -20°C under exclusion of light.

Epifluorescence microscopy

B. subtilis cells were grown in MC medium to an OD_{600} of 1.5 at 37°C and 200 rpm (see growth conditions). 1% agarose pads were made from MC-medium. 100 µl of the melted agarose were sandwiched between two smaller coverslips (12 mm, Menzel) and let rest for 2 min. 3 µl of culture were dropped on a round coverslip (25 mm, Marienfeld), and covered with an agarose pad to fix the cells. A Zeiss Observer A1 microscope (Carl Zeiss) with an oil immersion objective (100 x magnification, 1.45 numerical aperture, alpha Plan-FLUAR; Carl Zeiss) was used for wide-field microscopy. Images were acquired by a charge-coupled-device (CCD) camera (CoolSNAP EZ; Photometrics) and an HXP 120 metal halide fluorescence illumination with intensity control. Fusions of ComEC and ComEA to mVenus were excited at a wavelength of 514 nm and detected at 727 nm. Cells were illuminated for 0.5-1.5 s at the mid-cell plane. Images were processed via ImageJ (39).

Single-molecule tracking (SMT)

For single-molecule tracking, strains were grown under the appropriate conditions in order to obtain competent cells (see growth conditions). To study movement of mV-ComEA in the presence of DNA, 20 µg/ml chromosomal DNA (PY79) were added to cells entering stationary phase (t_1) and incubated for 1 h, as described in Burghard-Schrod *et al.* (2020). In case labeled DNA was added for SMT, all steps were carried out in the dark. To analyse the diffusion of labeled DNA, 50 µl of competent cells were incubated with 20 µg/ml labeled DNA for 30 min on a shaking platform. Further, cells were centrifuged at low speed (500 rpm) and resuspended in 50 µl MC-medium (supernatant of the negative control, PG876). DNAse was added to the sample in a final concentration of 200 mg/ml and incubated for additional 40 min (15) at 37°C on a shaking platform. Cells were centrifuged and resuspended again in 30 µl of used competence medium. Slides for microscopy were prepared as described (see methods, Epifluorescence microscopy).

For SMT, an inverted fluorescence microscope (Olypmus IX71, Carl Zeiss Microscopy) was used. A 514 nm laser diode (100 mW, Omicron Laser) was used with

50% intensity, which enables slim-field microscopy of proteins fused to mVenus. To facilitate high-resolution detection of the signal of the fluorophore-fusions, an EMCCD camera (iXON Ultra EMCCD, Andor) was applied. Movies of 2500 frames were recorded by use of AndorSolis software in an exposure time of either 50 ms (PG3817-PG3819) or 30 ms (PG876, PG3816, PG3909). In case ComEC-mV was analysed at mid-cell plane, SMT was performed using a Nikon microscope equipped with an $A = 1.49$ objective and an Image-EMCCD camera (Hamamatsu) using an integrated autofocus of the microscope.

Bleaching curves of each movie were generated via ImageJ (39). To obtain signal only at the single-molecule level, intense signal of the fluorophore-fusions was removed by cutting the movies according to the bleaching curves, as strong signal could lead to a wrong connection of tracks, hindering further processing of the data. Typically, the first 500 frames of each movie had to be removed, to analyse only the frames covered by the horizontal part of the bleaching curve (indicating no constant bleaching of the signal was occurring anymore). Cell meshes were set with Oufti (40). Prepared movies were analysed with UTrack (41), while for each track a minimal step length of 5 was selected. Finally, tracks were analysed by SMTracker software (42). Diffusion coefficients and diffusive populations were calculated under use of the Gaussian-mixture model (GMM) and Squared displacement (SQD, for detailed description see supplemental methods).

2.2.5 Results

A C-loop deletion of ComEC localizes at the cell- membrane in epifluorescence microscopy

In order to study the influence of the C-terminal part on the function of ComEC, we created a C-terminal mVenus fusion of the full-length protein (PG3817) and to two versions truncated of the terminal 128 amino acids (AS), N648-N776, and secondly, of the whole C-terminal part predicted to be soluble, by removing the final 301 AS, L475 – N776. We named the fused truncations after the size of the truncated parts ComECΔ128-mV (PG3818) or ComECΔ301-mV (PG3819). Fusions were created at the original locus of the gene and were analysed by epifluorescence - microscopy of *B. subtilis* strain PY79 grown to competence. Initially, the location of these truncations was chosen by the presence of the transmembrane-helices described by Draskovic *et al.* (2005), depicted in figure 1, a). We expected a more soluble or fast-diffusing protein in case transmembrane-helices were removed, and/or no polar localization pattern to be present anymore. The truncation ComECΔ128-mV displayed a polar signal similar to the wild-type fusion, but with fewer cells showing a signal at the cell pole, in 1.6 +/- 1.7% of the cells, compared to 5.7 +/- 2.05% (Fig. 1, b), Fig. 2) of the wild-type. Truncating a larger part of the C-terminus of ComEC in case of ComECΔ301-mV led to a punctate localization pattern all along the membrane in 68 +/- 5.2% of the cells (Fig. 1, c), Fig. 2). The fact that the C-terminal truncation is restricted to the cell membrane strongly argues against a defect in membrane insertion caused by proteolysis. In case of ComEC-mV and ComECΔ128-mV, a large number of cells, 35.12 +/-1.15% and 47.67 +/- 1.2%, displayed a diffusive signal throughout the cells, which seemed to be reduced in the truncation ComECΔ301-mV, where the signal was either present at the membrane (and little in the whole cell), or not present at all (Fig. 2 a)). Still, in all three analysed strains, cells without signal were detected, in agreement with only a subpopulation of cells entering the state of competence. This can be seen in Fig. 2, where about 30% of the cells express ComK (as seen by CFP expressed from a ComK-regulated promoter), while 70% show no expression. Raw data and number of replicates are presented in table 1. Curiously, more cells than expected show a signal for localized or diffusive ComEC-mVenus (about 40%), which

may be in a range expected for genetic noise, but almost 50% of the cells expressing the small ComEC truncation or more than 70% expressing the large truncation showed a fluorescence signal. While we have no explanation for this phenomenon, we note that a loss of the C-terminus of ComEC has an effect on the number of cells showing expression of the gene, or the entire *comE* operon.

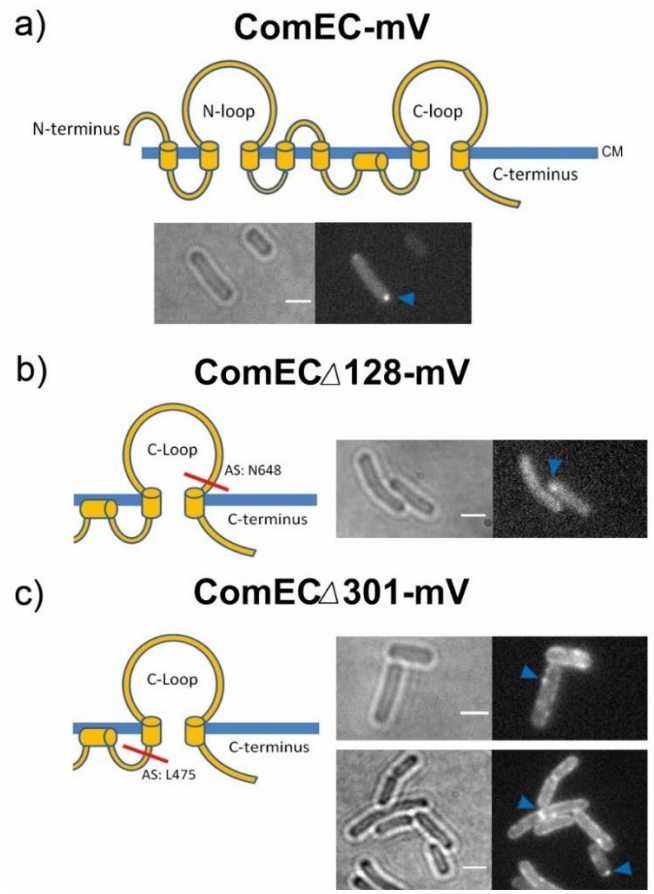


Figure 1: Epifluorescence-microscopy of ComEC-mV, ComECΔ128-mV, ComECΔ301-mV.

Panel a) shows epifluorescence-microscopy of ComEC-mV (PG3817). The topology of ComEC is shown after Drascovic et al. (2005). CM indicates cell-membrane, the loops are thought to be present in the periplasm of the cell. In panel b) and c) the truncated versions ComECΔ128-mV (PG3818) and ComECΔ301-mV (PG3819) are shown. Cartoon indicate the position of each truncation at the C-terminus of ComEC (red line, b), c)). While the wild-type protein localizes at the pole (a)), a truncation of 128 AS leads to a more diffusive signal of the fusion and a rare signal at the pole (b)). A truncation of the whole C-loop of 301 amino acids localizes throughout the cell membrane (c)). Blue arrows indicate positions of high fluorescence. White bars 2 µm.

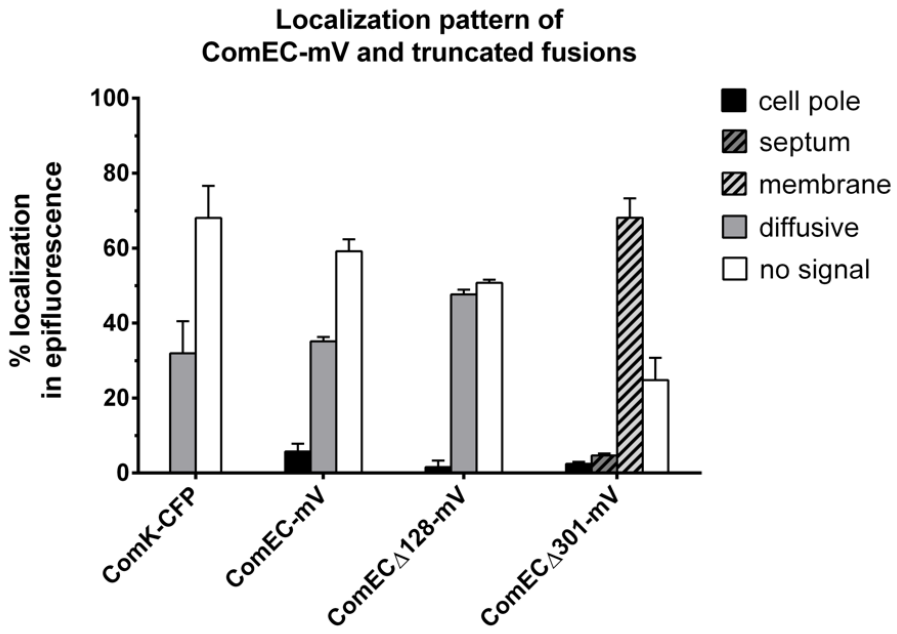


Figure 2: Localization pattern of ComEC-mV and C-terminal truncations.

Different strains are indicated at the x-axis by their corresponding fluorophore-fusion (from left to right: PG681, PG3817, PG3818, PG3819). Bars represent the different cellular localization of mV- fusions as indicated by the y axis legend in %. In case of ComECΔ301-mV, strong septal or polar localization were counted separately, but always included a localization of the fusions at the cell membrane, and little diffusive signal throughout the cell. Error bars represent standard deviation of the mean. Number of replicates and raw data are listed in table 1.

Table 1: Localization of ComEC-mV and C-terminal truncations

Strain genotype	**PG681 comK-CFP**	**PG3817 comEC-mV**	**PG3818 comECΔ128-mV**	**PG3819 comECΔ301-mV**
number of cells (n)	709 (224, 253, 232)	307 (167, 140)	475 (140, 151, 184)	487 (158, 189, 140)
number of replicates (N)	3	2	3	3
% cell pole (+/-SD)	-	5.74 +/-2.05	1.59 +/- 1.72	2.45 +/- 0.5
% septum (+/-SD)	-	-	-	4.67 +/- 0.54
% diffusive (+/-SD)	31.95 +/- 8.5	35.12 +/- 1.5	47.67 +/- 1.24	-
% membrane (+/-SD)	-	-	-	68.09 +/- 5.21
% no signal (+/-SD)	68.05 +/- 8.55	59.15 +/- 3.2	50.73 +/- 0.84	24.77 +/- 5.96

C-terminal truncations of ComEC lead to loss of transformability

Fusions were analysed with respect to their ability to take up DNA. Therefore, strains were transformed (see methods, transformation of *Bacillus subtilis* & transformation frequency) with chromosomal DNA of *B. subtilis*, encoding an Erythromycin resistance cassette (isolated from a Δ*comGA* strain, PG3717). Obtained colonies were selected for MLS- resistance, respectively. In case of a deletion of the whole C-loop of ComEC (PG3819, *comECΔ301-mV*) cells were hardly transformable (0.023% of the wild-type), while the mV-fusion of the full-length protein remained transformable, proving its functionality, as transformation frequency reached about 30% of the wild-type (see Fig. 3, table 2). Interestingly, activity was not completely abolished in the strain where we deleted only one helix and one putative zinc-binding site (PG3818,

comECΔ128-mV), but reduced it about 1000 fold compared to the wild- type. Comparing the transformation frequency of ComECΔ128-mV (PG3818) to the full-length protein fused to mVenus, transformation frequency reached about 3% (see table 2).

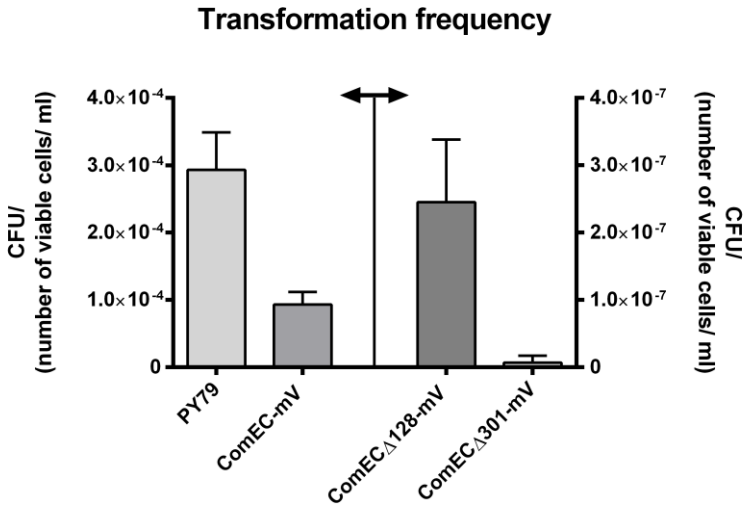


Figure 3: Transformation frequency of strains encoding comEC-mV and C-terminal truncations.

Bar plot shows CFU divided by the number of viable cells/ml per µg chromosomal DNA/ml. Assays of transformation frequency were carried out for every strain (from left to right: PG001, PG3817, PG3818, PG3819) in biological duplicates (on two different days). Each measurement was performed in technical triplicates. Arrow indicates separation of the graph into two y-axes. Error bars represent the standard deviation of the mean.

Table 2: Transformation frequencies of strains expressing ComEC-mV and truncations

Strain/ genotype	**Transformation frequency +/- SD**	**% wild-type +/- SD**	**% ComEC-mV +/- SD**
PG001 **PY79**	2.93 x 10^{-4} +/- 5.57^{-5}	100	-
PG3817 ***comEC-mV***	9.32 x 10^{-5} +/- 1.7 x 10^{-5}	31.78 +/-8.8	100
PG3818 ***comEC*Δ*128-mV***	2.45 x 10^{-7} +/- 9.3 x 10^{-8}	0.084 +/-0.031	2.63 +/- 9.9 x 10^{-3}
PG3819 ***comEC*Δ*301-mV***	6.77 x 10^{-9} + /- 1.05^{-8}	0.023 +/- 3.6 x 10^{-3}	0.072 +/- 0.01

A conserved aspartate residue in the C-terminus of ComEC is essential for transformation

In order to further study the function of the C-terminal part of ComEC, we mutated conserved amino acids, which were postulated to be relevant for its putative exonuclease activity (Baker et al, 2016). Corresponding strains were analysed *in vivo* by a transformation frequency assay. Two conserved motifs identified by Baker *et al.* (2016) were (aa) **H**A**D**Q**DH** (motif 1, aa 571-576, conserved aa indicated in bold) and G**D**LEKE (motif 2, aa 667-671), which potentially coordinate two zinc ions and therefore could be directly relevant for functionality. We chose amino acids located in these motifs, and exchanged them for alanine (figure 4 a)). Genes were expressed ectopically from the *amyE* site, while the gene at the original locus was deleted. Mutating a histidine of the first motif (H571) lead to a very low transformability (12.73%), and mutation of a conserved aspartate of motif 1 (D573) abolished transformability completely (figure 4 b)). Exchanging the second aspartate of motif 1 (D575) resulted in a higher (but compared to the positive control, still low) transformation frequency of 30.5%. Mutating the aspartate located in motif 2 (D668), lead to a transformation frequency of 40% relative to the wild-type (PG3929), the

highest value of all strains that were investigated, and where a putative zinc-binding amino acid has been exchanged. Interestingly, we found that transformation frequency decreased significantly in all mutations, except for one control, where we decided to mutate K671, which was postulated to be part of the conserved motif2 but *not* involved in zinc binding. In this case, transformation frequency reached 86% of the positive control (*ΔcomEC, amyE:comEC,* PG3929, see table 5 for transformation frequencies and percentages).

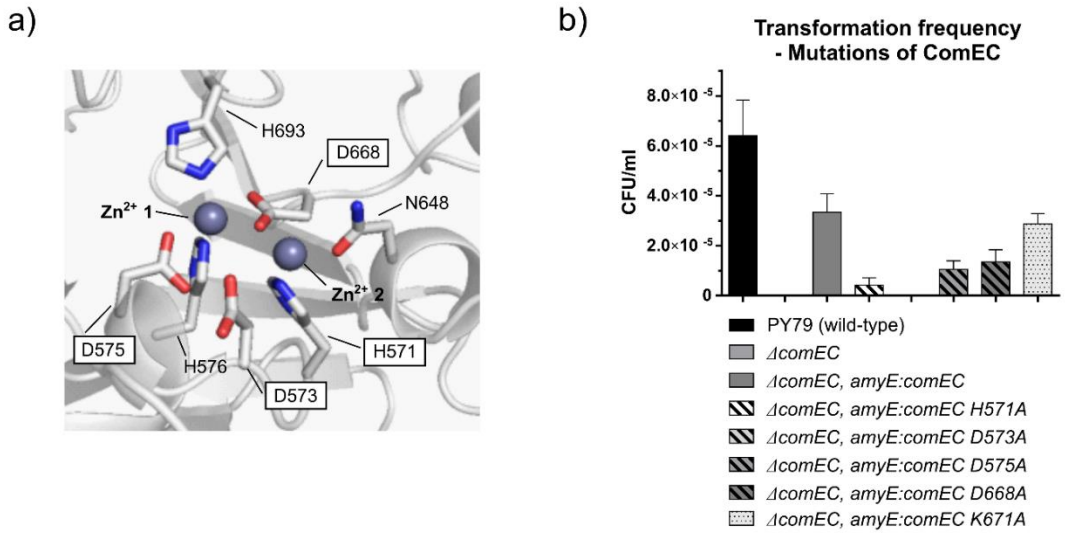


Figure 4: Transformation frequency of mutations of ComEC.

In a) the active site of the β-lactamase domain of ComEC is shown, based on the model of Baker *et. al.,* (2016). According to the model, two zinc ions (Zn^{2+} 1, Zn^{2+} 2) are coordinated by two conserved aspartate residues (D575, D668), one arginine (N648) and three histidine residues (H571, H576, H693). The mutated amino acids are indicated by white boxes. Respective residues were exchanged by an alanine. Mutated genes were integrated and induced from the *amyE*-site, while the original gene had been deleted, in order to study the effect of the mutations on transformation. b) shows transformation frequency of the corresponding strains, while genotype is indicated by the legend. As additional control, amino acid K671, which is located close to the conserved residue D668, has been exchanged for an alanine. We found that the conserved residue D573, which appears to be close to the active site and present in the conserved motif **H**A**D**Q**DH**, but originally not depicted in the model of Baker *et al.* (2016), is essential for transformation. Assays were performed in technical triplicates and in biological duplicates. Error bars indicate standard deviation of the mean.

Table 3: Transformation frequencies of mutants of ComEC

Strain	**genotype**	**Transformation frequency +/- SD**	**% wild-type +/- SD**	**% ΔcomEC amyE:comEC +/- SD**
PG001	Wild type PY79	6.407^{-005} +/- 1.425^{-005}	100	-
PG3722	*ΔcomEC*	0	0	0
PG3929	*ΔcomEC, amyE::comEC*	3.345^{-005} +/- 7.329^{-006}	52.2 +/-11.43	100
PG3917	*ΔcomEC, amyE::comEC H571A*	4.259^{-006} +/- 2.909^{-006}	6.65 +/- 4.5	12.73 +/- 8.7
PG3918	*ΔcomEC, amyE::comEC D573A*	0	0	0
PG3919	*ΔcomEC, amyE::comEC D575A*	1.055^{-005} +/- 3.413^{-006}	16.5 +/- 5.3	30.57 +/- 9.89
PG3927	*ΔcomEC, amyE::comEC D668A*	1.351^{-005} +/- 4.884^{-006}	21.08 +/- 7.62	40.38 +/- 14.6
PG3928	*ΔcomEC, amyE::comEC K671A*	2.880^{-005} +/- 4.105^{-006}	44.95 +/- 6.4	86.09 +/- 12.27

B. subtilis ComEC moves within the membrane during the state of competence

In order to investigate the localization of ComEC, we decided to analyse the fusion ComEC-mV by single-molecule tracking (SMT) in *B. subtilis* cells, grown to competence. Therefore, we applied an exposure time of 50 ms, where we made different observations by focusing on different cell planes during microscopy. By focusing to the central plane of the cell, we found a clear localization pattern of ComEC-mV along the membrane (Fig. 5a), lower panel, movie S1), where we found mobile molecule tracks (shown in blue) or molecules moving within a confined area (red, most pronounced at the cell poles) at the cell periphery (green tracks show transitions between diffusive and constrained motion). However, by focusing on the upper level of the cell, we gathered more tracks, because molecules moved all along a bent surface (Fig. 5b). In both cases, clear large foci were visible, which for the upper level focus were found at all cellular places, since we were looking onto the surface of the cell. Accumulation of ComEC at the cell poles was not seen using the upper cell focus (figure 5b)), but diffusive dynamics of the protein in x and y direction was much better captured in this mode, rather than following movement in largely only x direction along the membrane, as in the central focussing mode. We therefore decided to determine velocity of our fusions at the upper-cell level. Applying Gaussian-mixture-model (GMM) analyses, we fitted 2 diffusive populations to the overall movement of the full-length fusion, which resulted in an overall well representation of the obtained step lengths. The two population fits showed that ComEC molecules fall into two modes of movement, one static and one mobile population, having diffusion-coefficients of 0.017 +/- 0.00013 $\mu m^2/s$ and 0.31 +/- 0.0011 $\mu m^2/s$. The latter mobility corresponds to that of a membrane protein with several transmembrane spans (34). We found a large diffusive fraction in case of the full-length protein fusion, ComEC-mV, comprising 87.1% of all detected movements (see table 3). Comparing the sizes of the two diffusive fractions of the C-terminal truncations of ComEC, we found a decrease of the mobile fraction and an increase of the static fraction when the protein was truncated at its C-terminus in general. We found that ComECΔ128-mV showed a smaller mobile population of about 63% (24% more static than the wild-type), while

for ComEC301-mV, we detected even fewer mobile tracks of a population of 49.8%. As already indicated by its appearance in epifluorescence (figure 1), we found a large difference comparing the full-length protein fused to mVenus and ComECΔ301-mV, a 37.3% decrease of the mobile fraction when the whole C-terminus was deleted. We infer that our fusions are stable proteins, because free mVenus (or any fluorescent protein) has not been reported to localize at the membrane on its own, and all our fusions show clear membrane localization in SMT or epifluorescence. Unfortunately, we were unable to visualize ComEC-mVenus or mutant proteins by Western Blotting. Thus, we cannot exclude that we have been visualizing proteins lacking a part of their N-terminus, although this is unlikely because ComEC-mVenus is a functional protein fusion. We did not detect tracks in all cells as fusions were only expressed in cells being in the K-state, and therefore calculated the number of tracks per cell on the percentage of cells where tracks were present. Here we found no difference in case of ComEC-mV compared to ComECΔ128-mV (9.84 tracks, see table 3), but a moderate increase of tracks per cell in case of ComECΔ301-mV (11.8 tracks). These results indicate an increase of signal at the membrane in general when a part of the C-terminus was deleted (see overlay of tracks on cell, suppl. Fig. 1, in agreement with the epifluorescence analyses).

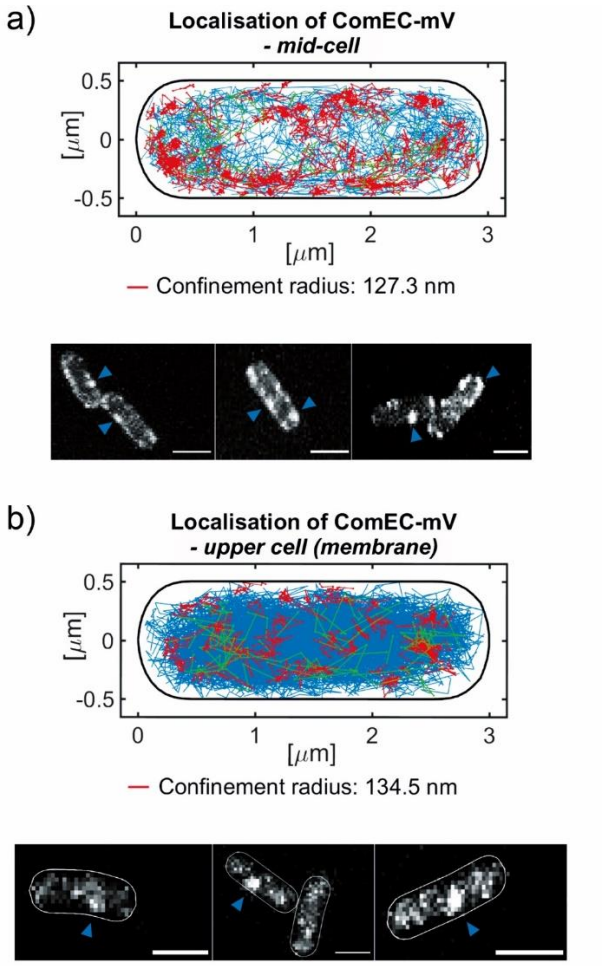


Figure 5: Single-molecule tracking of ComEC-mVenus.

In a) the localization of tracks is shown as an overlay in one normalised cell, molecules were tracked mid-cell (number of tracks: 548). Lower panels, exemplary t-stacks of an analysed movie with 2000 frames are shown. In b) the same experiment is shown except for the changed imaging plane (upper cell), where 1280 tracks were monitored. In red, the confined tracks are depicted, moving within a radius of either 127.3 nm (a)) or 134.5 nm (b)). The confinement radius corresponds to 3 times the standard error of the MSD (calculated from 5 intervals). Green colour indicates a change of the tracks from confined to a more mobile movement or *vice versa*; tracks are therefore considered as segmented. Steps of blue coloured tracks always exceed the confined radius; these tracks are considered as mobile. Blue arrows indicate localization of a strong intracellular signal. Measurements were carried out in biological triplicates. White bars 2 µm.

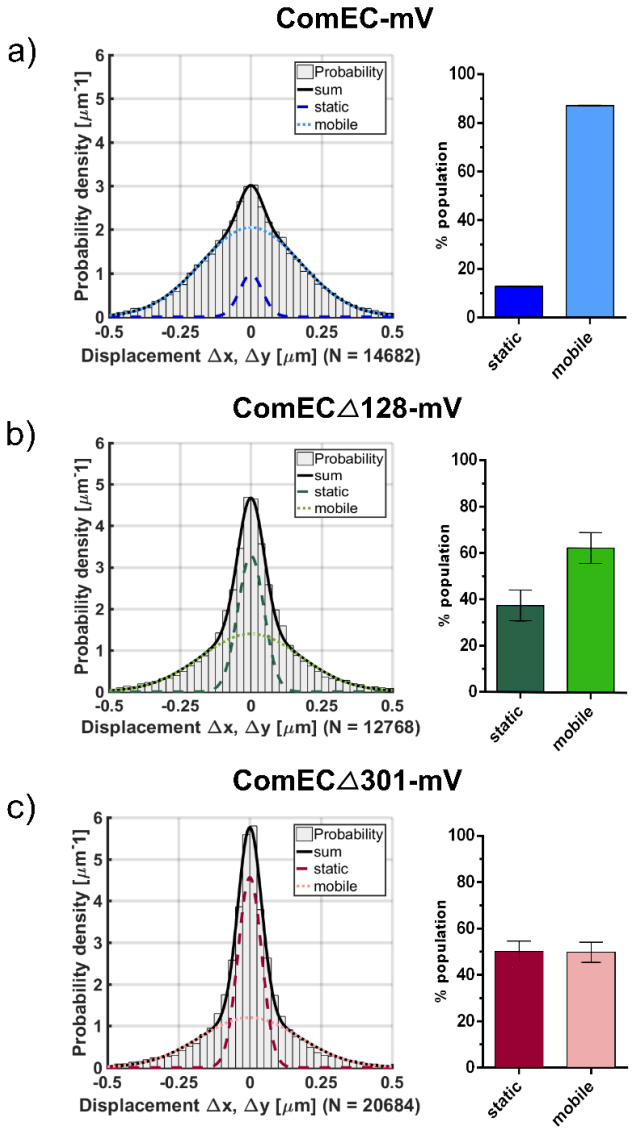


Figure 6: Analysis of populations of ComEC-mV and of C-terminal truncations with different mobility analysed by GMM.

In a) The step-sizes of all analysed tracks for ComEC-mV (PG3817) are plotted in a histogram, according to their probability density and fitted by two Gaussian functions, whose area is then summed up to one fitting (R^2= 1, black line). The % of static and mobile populations are depicted in a bar plot on the right. In b) the same analysis was done for ComECΔ128-mV (PG3818, R^2= 1) and c) for ComECΔ301-mV (PG3819, R^2= 1). Error bars of bar plots indicate standard deviation of the mean. N= number of steps.

Table 4: Single-molecule tracking of ComEC-mV, ComECΔ128-mV and ComECΔ301-mV

Strain/ genotype	**PG3817** ***comEC-mV***	**PG3818** ***comECΔ128-mV***	**PG3819** ***comECΔ301-mV***
D_1 +/- SD [µm²/s] (static population)	0.017 +/- 0.00013		
D_2 +/- SD [µm²/s] (mobile population)	0.31 +/- 0.0011		
Pop_1 +/- SD [%] (static population)	12.9 +/- 0.14	37.5 +/- 6.64	50.2 +/- 4.38
Pop_2 +/- SD [%] (mobile population)	87.1 +/- 0.14	62.5 +/- 6.64	49.8 +/- 4.38
Analysed cells	165	106	153
Tracks (total)	1280	915	1617
Cells with tracks	130	93	137
Cells with tracks (%)	78.78	87.73	89.54
Tracks/cell	9.84	9.84	11.8

Diffusion of fluorescently labeled DNA can be monitored in competent B. subtilis cells via single-molecule tracking inside of the periplasm

We fluorescently labeled DNA according to the protocol of Boonstra *et al.* (2018) (15), and incubated *B. subtilis* cells grown to competence with a fluorescently labeled PCR-product, which displayed homology to the *B. subtilis* chromosome, flanking an erythromycin resistance gene (2300 bp, partial sequence of vector pDG1664). After an incubation time of 30 min and an additional incubation time of 40 min with DNAse, we investigated our samples by single-molecule tracking at mid-cell level. We chose strain PG876 in order to increase the number of cells entering the K-state. Initial measurements with an exposure time of 50 ms, assuming slow movement of the up to 1.5 mDa large dsDNA, resulted in mostly blurry signals but no defined point spread

functions (data not shown). As a result, we recorded the signal with an exposure time of 30 ms, and to our surprise found dynamics that resemble those obtained for proteins (movie S2, movie S3). Data could be fitted best by assuming three diffusive populations, two diffusive population did not result in a sufficient quality of fitting (meaning not all measured data could be covered by a double fit). We calculated diffusion coefficients of 0.014 +/- 0.0 $\mu m^2/s$, 0.071 +/- 0.001 $\mu m^2/s$ and 0.621 +/- 0.004 $\mu m^2/s$ for the labeled DNA (PG876, figure 8a), table 4). These data correspond to static motion seen for ComEC (table 3), to slow diffusion in case of the medium-fast population, and free diffusion of cytosolic proteins (43-45).

After incubating the *rok* deletion strain with labeled DNA, we found a polar and mostly peripheral pattern of confined tracks (indicated in red), in addition to a fast-diffusive signal indicated in blue (PG876, *Δrok*, figure 7a)). This suggests that taken up DNA is largely present at the periphery of the cell, i.e. in the periplasm. As a control, we analysed wild-type cells (PG876) without addition of DNA, where we found an average background signal of 2 tracks/cell (see Fig. S4). Because only a low signal of 5.4 tracks/cell was detected when DNA was added, we note that many signals obtained could be due to background noise. We reasoned that an increased amount of signal might be obtained when *comEC* is deleted (in addition to *rok*), which was indeed the case (PG3816, figure 7b)). Signals changed from 445 to 1005 tracks in total (12.3 tracks/cell, see table 4). Diffusion constants for DNA in the *Δrok ΔcomEC* background were 0.015 +/- 0.0 $\mu m^2/s$, 0.077 +/- 0.001 $\mu m^2/s$ and 0.589 +/-0.002 $\mu m^2/s$ (PG3816, figure 8b), table 4), very closely resembling those in wild type (*Δrok*) cells. Because of the much higher signal to background ratio now allows us to deduce that taken up DNA indeed shows motion within the periplasm that can be best described by three distinct mobilities.

Performing a temporal overlay of frames acquired in SMT (t-stacks of movies), we discovered a punctate localization pattern of the labeled PCR-product along the membrane of the cells (figure 7 b), lower panel), in contrast to the fewer peripheral foci observed in wild type cells (Fig. 7a, lower panel). This observation was supported by SMT measurements, as confined tracks localized at the cell periphery (figure 7 b), upper panel, red tracks), in addition to mobile tracks, indicating a high mobility of the DNA (figure 7 b), upper panel, visualised in blue). Overall, we found a more intense

polar localization in case *comEC* was deleted (compare figure 7 a) to 7 b), lower panel), probably due to higher number of fluorescence signal in general. These findings show that in the absence of transport of DNA across the cell membrane, DNA accumulates within the periplasm, but also does so under wild type conditions, where it shows protein-like dynamics of very low to rapid diffusion.

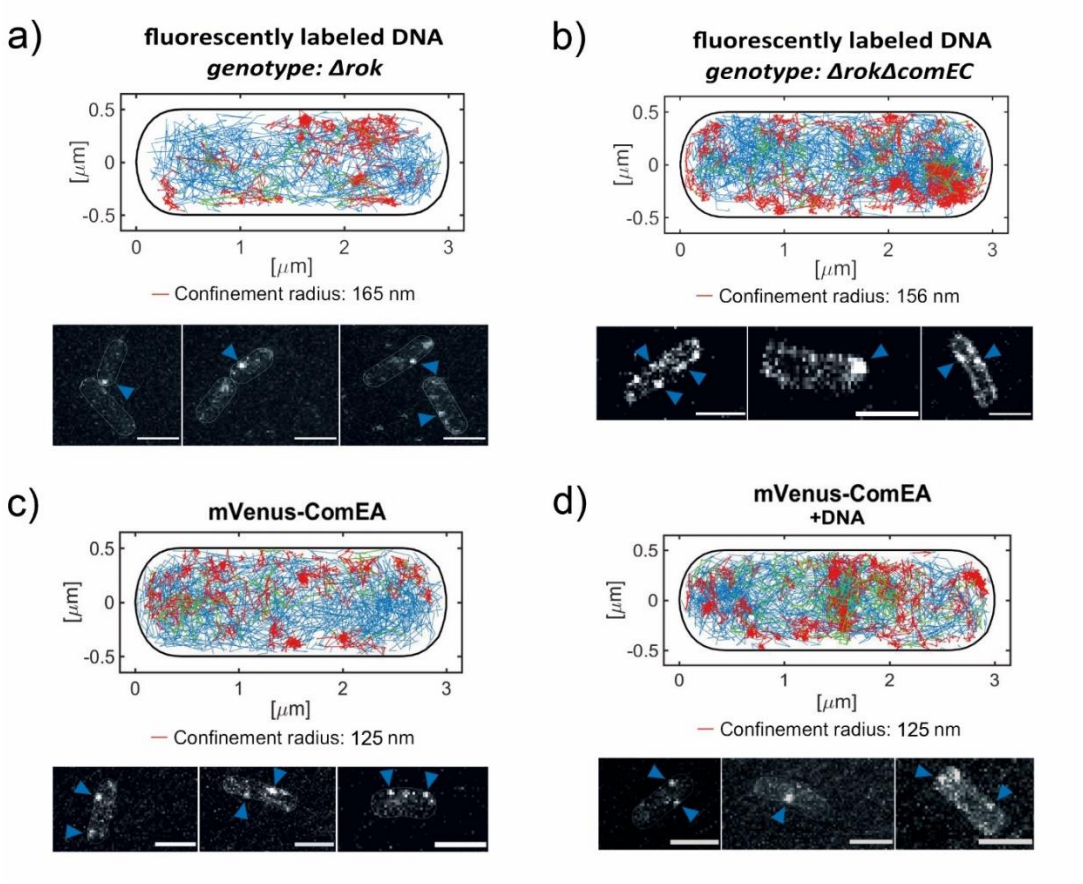


Figure 7: Single-molecule tracking of fluorescently labeled DNA and mVenus-ComEA.

Single-molecule tracking was carried out with 30 ms exposure time. In a), the signal of *B. subtilis* PY79 cells (PG876, *Δrok*) incubated with stained PCR-product, labeled with DyLight488, is shown by plotting tracks on a standardized cell (number of tracks: 251). Lower panel exemplary t-stack of a movie, signals indicated by blue arrows. Cell meshes show the borders of the cells. b), *comEC rok* deletion strain (PG3816, n = 555) showing more distinct localization of the signal along the membrane. c) shows localization of mVenus-ComEA, expressed under control of the Xylose-promotor using 0.05% Xylose (strain PG3909, number of tracks: 418). In d) chromosomal DNA (20 µg/ml) was added to the mVenus-ComEA

expressing cells and incubated for 1 h (n = 554). The confinement radius corresponds to 3 times the localization error. Green colour indicates a change of the tracks from confined to a more mobile movement or *vice versa*, while blue tracks consist of non-confined movements (steps) are thus of mobile molecules. Confined tracks are shown in red, moving within a radius as indicated. Cells were analysed close to mid-cell level. Scale bars 2 µm.

Movement of DNA is similar to that observed for ComEA

We wished to compare SMT data obtained for labeled DNA with the diffusion of membrane protein ComEA, which is thought to bind to DNA entering the periplasm during natural transformation of *B. subtilis*. Therefore, we created an N-terminal mVenus-fusion of the protein at the original locus (the N-terminus resides in the cytosol), which we induced with a low concentration of xylose (PG3909). Of note, even under low levels of induction, the fusion supported transformation (see figure S5). Fig. 7 shows that mVenus-ComEA was found to be mobile throughout the cell membrane. The protein accumulated at the cell-pole, where we detected a strong signal, in addition to the general localization pattern at the membrane (see figure 7 c), movie S4). The pattern became even more obvious when mVenus-ComEA was induced with 0.5% Xylose, which we did not use for analysis as such a strong signal would distort single-molecule tracking (figure S6). Further, we incubated strain PG3909 with chromosomal DNA for 1 h, where we detected an overall stronger signal at the membrane, septum and cell pole as depicted in figure 7d). Similar to the fluorescent DNA, the diffusion of mVenus-ComEA could be best fitted using three populations. We determined diffusion coefficients of 0.014 +/- 0.0001 $\mu m^2/s$, 0.075 +/- 0.001 $\mu m^2/s$ and 0.599 +/. 0.003 $\mu m^2/s$ (see figure 8 c, table 4), which are very similar to those determined for taken up DNA. Percentages of populations varied only little between the two conditions in which labeled DNA was investigated. About 40% of labeled DNA showed high and medium mobility, about 20% was statically positioned (see bubble plots, figure 8). In the *comEC* mutant background, more DNA molecules were static, close to 30%, while about 35% showed medium and high mobility (Fig. 8, see table 4 for exact values and standard deviations). For mVenus-ComEA, we monitored a static fraction of 16.1%, which increased (and almost

doubled) to 29.2% (figure 8c and d), while diffusion coefficients did not change significantly (see table 5). The slow-mobile fraction of ComEA contained 44.2% of molecules, and only slightly decreased to 41.1% upon addition of DNA, while the fast-mobile fraction decreased from about 40 to 31% after addition of DNA. While it is not straight forward to explain the different mobilities, our results clearly show that taken up DNA can rapidly move through the *B. subtilis* periplasm, which appears to be able to act as a reservoir for DNA that can not be pumped into the cell via ComEC. Of note, we closely followed the protocol devised by the Kuipers groups to obtain comparable conditions. We cannot state if there may be differences in dynamics of DNA or of ComEA soon after addition of DNA, different from our conditions.

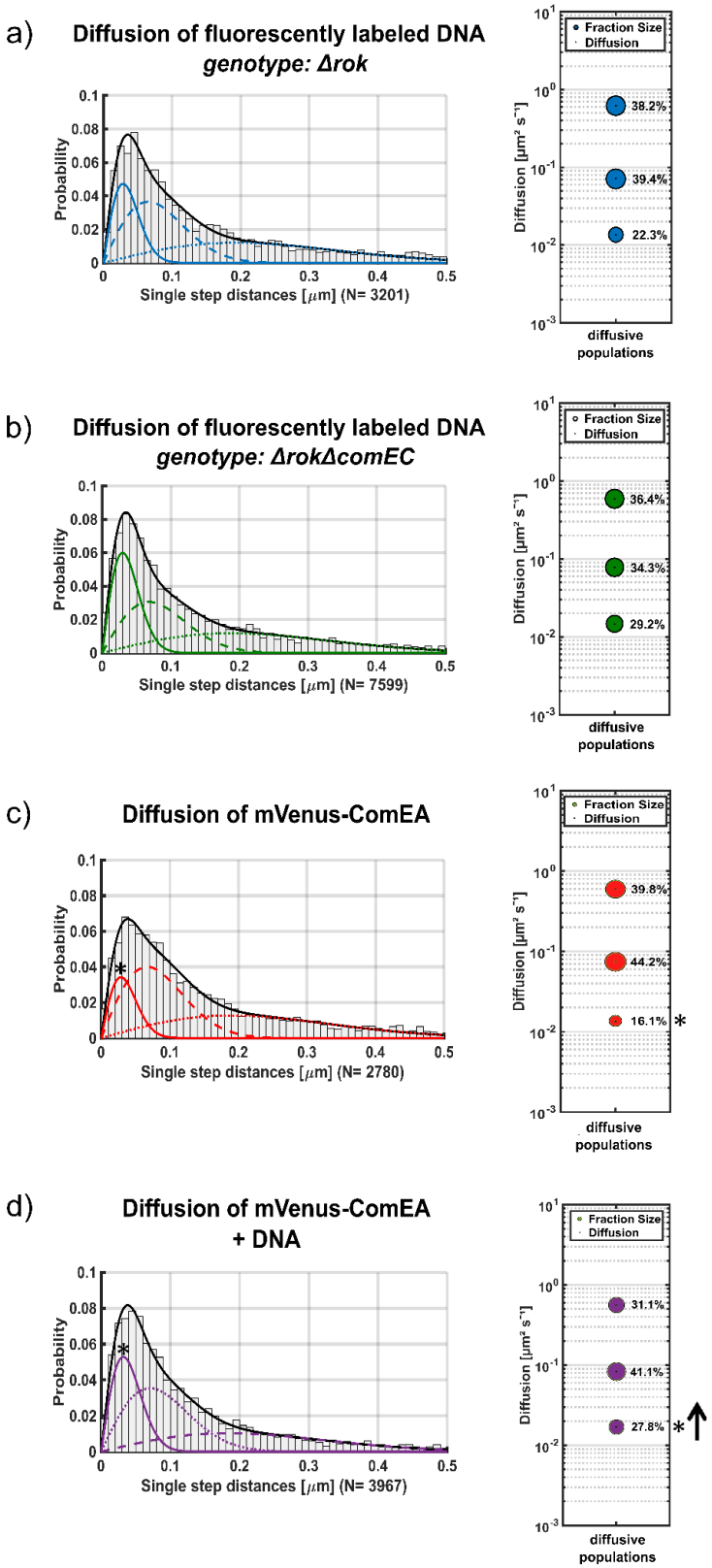


Figure 8: Diffusion of labeled PCR-products and of mVenus-ComEA inside competent B. subtilis cells.

Diffusion coefficients were calculated by the squared displacement, fitting a cumulative density function to the data. Histograms in a), b), c) and d) show the probability of single step distances (Jump distance) of all acquired data, and the percentage of the diffusive populations as a bubble-plot, next to the corresponding histogram. a) and b) show single molecule tracking of labeled DNA, c) results for mVenus-ComEA. In d) chromosomal DNA was added to competent cells expressing mVenus-ComEA. Comparing c) and d), we detected an increase of the static fraction of 11.7% (fraction indicated by asterisks and black arrow in d)). For all conditions, single-molecule tracking was carried out at mid-cell level, 30 ms stream acquisition). Measurements were carried out in biological triplicates. N= number of steps.

Table 5: Single-molecule tracking of fluorescently labeled DNA compared to mVenus-ComEA

Strain	**PG876 + labeled DNA genotype: *Δrok***	**PG3816 + labeled DNA genotype: *ΔrokΔcomEC***	**PG3909 genotype: *mV-comEA***	**PG3909 + DNA genotype: *mV-comEA***
D_1 +/- SD [$\mu m^2/s$] (static population)	0.014 +/- 0.0001	0.015 +/- 0.0001	0.014 +/- 0.0001	0.017 +/- 0.0001
D_2 +/- SD [$\mu m^2/s$] (slow mobile population)	0.071 +/- 0.001	0.077 +/- 0.001	0.075 +/- 0.001	0.083 +/- 0.001
D_3 +/- SD [$\mu m^2/s$] (mobile population)	0.621 +/- 0.004	0.589 +/- 0.002	0.599 +/- 0.003	0.558 +/- 0.003
Pop_1 +/- SD [%]	22.3 +/- 0.004	29.2 +/- 0.001	16.1 +/- 0.003	27.8 +/- 0.003

(static population)				
Pop_2 +/- SD [%] (slow mobile population)	39.4 +/- 0.003	34.3 +/- 0.001	44.2 +/- 0.002	41.1 +/- 0.002
Pop_3 +/- SD [%] (mobile population)	38.2 +/- 0.003	36.4 +/- 0.001	39.8 +/- 0.002	31.1 +/- 0.002
Analysed cells	83	83	100	99
Tracks (total)	443	1005	418	554
Tracks/cell	5.4	12.3	4.2	5.6

2.2.6 Discussion

Uptake of DNA by competent bacteria has been described to occur at single cell poles, for *B. subtilis*, *Vibrio cholerae* and for *Helicobacter pylori* (14, 46-48). It has been speculated that a single entry-point into the cytoplasm may facilitate search for homology to incoming DNA on the chromosome, and for incorporation of ssDNA at corresponding loci. It has remained an intriguing question how the cell can place a multiprotein complex guiding DNA through the cell envelope into the cell at a single cell pole. In recent work, we have found evidence suggesting that the polar DNA uptake machinery in *B. subtilis* may be assembled through a diffusion/capture mechanism, from highly dynamic molecules. In our present work, we have investigated the dynamics of membrane permease ComEC and of DNA receptor ComEA, and have investigated the role of the large C-terminus of ComEC. We have also been able to follow the fate of incoming DNA into the cell via single molecule tracking.

For ComEC-mV, we found a surprisingly low number of static molecules, and 87% molecules freely diffusing within the cell membrane. Of note, we found confined motion of ComEC at many places within the cell membrane, indicating that it does not only stop for some time at the cell pole, but here, it clearly stops most often. Hahn *et al.* described the movement of fluorescent foci towards the cell pole, studying a ComGA-CFP fusion by time-lapse microscopy, postulating a putative diffusion/capture mechanism for assembly of competence proteins at the cell pole, but stated that only high-speed microscopy could reveal the true nature of the assembly (49). Using SMT, our findings support the idea of a diffusion capture model for polar localization of ComEC, in order to internalize exogenous DNA that has entered the cell via the putative pseudopilus, similar to what we have found for ATPase ComGA. (35). Intriguingly, truncating the C-terminus of ComEC strongly reduced mobility of ComEC in the membrane. It has been determined that diffusion coefficients of membrane proteins decrease with increasing numbers of TMs, but not with the size of soluble parts of membrane proteins (34). Expecting an increase in mobility of ComEC, we removed one or two membrane helices from the C-terminus (ComECΔ128-mV and ComECΔ301-mV), but found an increase of static molecules by SMT. These findings

indicate that the putatively periplasmic C-terminus of ComEC is required for non-constrained diffusion in the membrane.

ComEC has only been seen to localize at the membrane in the end of exponential phase, when *B. subtilis* cells are developing competence (49) and at the cell pole in competent *Bacillus* cells, described by Kaufenstein *et al.*, (2011). Comparing the full-length fusion of ComEC-mV (PG3817) to ComECΔ128-mV (PG3818), where 128 amino acids were deleted from the C-terminus of ComEC, polar localization decreased from about 6% to 1.6% seen by epifluorescence microscopy. Both fluorescence fusions showed a high amount of uniform (or rather diffusive) fluorescence signal in 35% and 47% of the cells (see figure 2, table 1). In contrast, using SMT, we were able to detect a clear membrane-associated localization pattern of ComEC-mV and ComECΔ128-mV (figure 5a, figure S2). In case of ComECΔ301-mV, a large part of 301 amino acids was truncated, leading to a signal similar to a membrane-stain (figure 1), showing a distinct punctate localization pattern along the membrane. However, we detected a large, about threefold increase of the slow mobile/static fraction, showing that is spite of a loss of polar clustering, the truncation greatly lost diffusive mobility within the entire cell membrane.

To obtain further knowledge on the function of the C-terminus, we mutated several conserved amino acids of ComEC, identified by Baker *et al.* (2016), in order to study the putative exonuclease function possibly carried out by its C-terminus (figure 4). We found that an aspartate residue, **D**573, part of the putative zinc-binding motif HA**D**QDH, is essential for transformation. This finding supports the idea of ComEC being the putative exonuclease of the system. Unfortunately, we have been unable to provide *in vitro* assays to finally prove the theory, hampered by DNA contamination and insolubility when we expressed putative soluble parts of ComEC in *E. coli* (unpublished data).

Most importantly in this work, we have been able to follow the dynamics of fluorescently labeled DNA in real time, for which we have not found a precedence in the literature. We determined that a 2300 bp fragment of DNA (MW= 1.5 mDa) shows three distinct patterns of movement: we observed a slow mobile/static fraction, which may represent DNA that is taken up by the pseudopilus, possibly a slow event. This

fraction mildly increased in cells lacking ComEC, indicating that ComEC affects the dynamics of static DNA molecules to only a small degree. We found two additional populations, one having an intermediate (low) diffusion constant, and one having a diffusion constant that is comparable to that of freely diffusing cytosolic as well as membrane proteins (50, 51). In order to better understand the nature of the three populations, we followed the motion of DNA receptor ComEA by single molecule tracking. We found that also ComEA moves as three distinct populations and curiously, all three having very similar diffusion coefficients as those determined for DNA (see figure 8, table 4). ComEA showed confined motion at many places within the cell membrane, in addition to polar accumulation like ComEC or ComGA. The three populations of ComEA could be described as mobile molecules diffusing freely without bound DNA, molecules that move together with DNA, and molecules that bind to DNA that is moved through the cell wall by the pseudopilus. This idea is supported by the finding that the fast mobile fraction of ComEA moves with an average diffusion constant similar to that of single membrane proteins (34), while that of the medium fast and slow mobile fractions is much lower. Likewise, 1.5 mDa linear DNA might move like a large cytosolic protein through the periplasm by itself. On the other hand, it is striking that diffusion constants between the three DNA fractions and those of ComEA have such similar values. Of note, when large DNA strands are added to competent *B. subtilis* cells, the DNA uptake machinery generates on average 25.000 bp DNA fragments (13, 52, 53) through the activity of NucA endonuclease (54). Thus, low mobility DNA could be ComEA bound to 2300 bp DNA, and the fast population shorter DNA fragments bound to ComEA. However, addition of DNA to competent *B. subtilis* cells did not lead to strong changes in ComEA dynamics, except for an increase in the static fraction of ComEA. Note that a portion of *B. subtilis* cells entering stationary phase actively secrete chromosomal DNA, and fragments thereof (55), such that ComEA is likely also engaged in some DNA binding in the absence of externally added DNA. Thus, alternatively, the two mobile fractions could be due to few or many ComEA molecules bound to a single DNA fragment, because ComEA can bind cooperatively to incoming DNA (48).

We interpret these data in a way that taken up DNA is directly bound by ComEA (Fig. 9). Uptake of DNA through the pseudopilus may be a slow event, leading to statically

positioned ComEA and DNA. Only a small fraction of DNA is directly taken up further through ComEC, but DNA can be buffered and stored by ComEA (Fig. 9). The protein is very mobile during competence, similar to ComEA of *Vibrio cholerae*, despite the fact that *B. subtilis* ComEA is likely an integral membrane protein, and not a periplasmic receptor as it has been found for *V. cholerae* (21, 47). Because of similar diffusion constants, we favour the view that all taken up DNA is bound to ComEA, some freely diffusing, some being in contact with the uptake machinery, e.g. ComEC, or getting targeted to the pole by other factors.

Our findings show that unlike *V. cholerae*, ComEA is not accumulating only at the competence pole upon addition of DNA to receive incoming DNA, but remains diffusing through the membrane. The fact that diffusion constants and fraction sizes of DNA do not change markedly between wild type and *comEC* mutant cells suggests that ComEA can bind to excess incoming DNA that cannot be directly transported into the cytosol, and acts as a buffer or storage system, allowing DNA to enter at later time points. A similar function has been reported for ComE of the gram-negative *Neisseria gonorrhoeae*, where the homologue of ComEA (ComE) is considered to bind DNA in the periplasm and act as a reservoir for taken-up DNA during natural transformation (56, 57). Clearly, most of the tDNA accumulates in the periplasm, and only few molecules are transported into the cytosol, or otherwise the difference between wild type and *comEC* mutant cells would have been much bigger. We assume, that the signal we detected must have been present in the periplasm, as cells were treated with DNAse and washed, removing unspecific DNA from the cell surface. It is not clear whether the signal we detected in the wild type (PG867), was localized in the cytosol in addition, as bleed-through from the lower cell level could have also caused the detection of tracks localized at the cell center (see figure 7a)).

Our findings further improve our understanding of the *B. subtilis* competence machinery, a highly dynamic assembly of proteins, likely set up predominantly by a rapid diffusion/capture mechanism. Even large DNA molecules show high mobility within the bacterial periplasm, ensuring that it can also efficiently find the polar ComEC channel for eventual uptake into the cytosol. The idea of ComEA buffering DNA uptake agrees with findings from V. cholerae where DNA transport through the

outer membrane and through the cell membrane has been shown to be spatially but not temporally coupled (48).

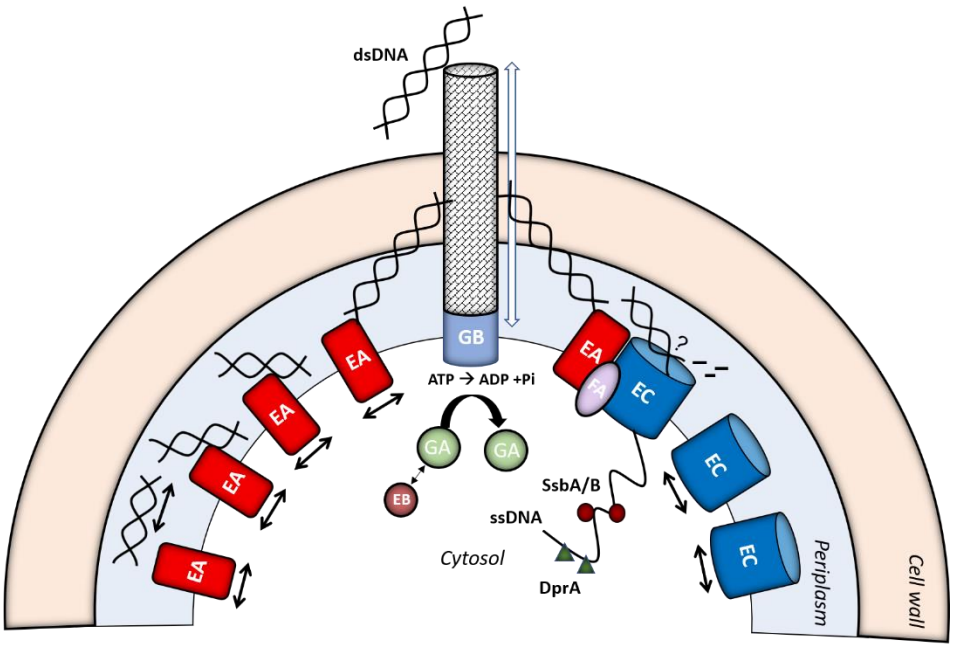


Figure 9: Model for DNA-uptake of B. subtilis.

Double-stranded DNA (dsDNA) is transferred into the periplasm via rapid assembly and disassembly of the putative pseudopilus at the cell pole (indicated by weight arrow), whose anchor might be the membrane protein ComGB. ComGA is guided to the cell pole by ComEB, where it provides the energy for pilus assembly/disassembly. ComEA binds the dsDNA in the periplasm, where it can either hand over the DNA to ComEC, and ssDNA is transferred to the Cytosol with energy provided by ComFA (complex on the right), or diffuses along the membrane to act as a kind of reservoir for transforming DNA (loaded ComEA on the left). ComEC and ComEA are diffusing freely until a complex is formed at the cell pole when DNA is present (diffusion indicated by black arrows). Probably, free DNA is present in the periplasm in addition, in case it has not been bound by ComEA yet. Once ssDNA enters the Cytosol, it is coated by single-strand DNA binding proteins SsbA, SsbB and DprA, followed by integration into the chromosome.

Acknowledgements

We would like to thank Lisa Stuckenschneider for providing vector pHJDS-mV, Luis M. Oviedo Bocanegra for supporting our Single-molecule tracking analysis, and Daniel J. Rigden for providing the model of the C-terminus of *B. subtilis* ComEC. In addition, we thank David Dubnau for providing deletion strains of competence proteins. This work was supported through the state of Hessen (central funds for University of Marburg) and the Deutsche Forschungsgemeinschaft (DFG).

2.2.7 References

1. Hadden C, Nester EW. 1968. Purification of Competent Cells in the Bacillus subtilis Transformation System. J Bacteriol 95:876-85.

2. Maamar H, Dubnau D. 2005. Bistability in the Bacillus subtilis K-state (competence) system requires a positive feedback loop. Mol Microbiol 56.

3. Mirouze N, Desai Y, Raj A, Dubnau D. 2012. Spo0A∼P Imposes a Temporal Gate for the Bimodal Expression of Competence in Bacillus subtilis, PLoS Genet, vol 8.

4. Chung YS, Dubnau D. 1998. All seven comG open reading frames are required for DNA binding during transformation of competent Bacillus subtilis. J Bacteriol 180:41-5.

5. Hahn J, Inamine G, Y K, Dubnau D. 1993. Characterization of comE, a late competence operon of Bacillus subtilis required for the binding and uptake of transforming DNA. Molecular microbiology 10.

6. Londoño-Vallejo J, Dubnau D. 1993. comF, a Bacillus subtilis late competence locus, encodes a protein similar to ATP-dependent RNA/DNA helicases. Molecular microbiology 9.

7. Mohan S, Aghion J, Guillen N, Dubnau D. 1989. Molecular cloning and characterization of comC, a late competence gene of Bacillus subtilis. Journal of bacteriology 171.

8. van Sinderen D, Luttinger A, Kong L, Dubnau D, Venema G, Hamoen L. 1995. comK encodes the competence transcription factor, the key regulatory protein for competence development in Bacillus subtilis. Mol Microbiol 15:455-62.

9. Hoa T, Tortosa P, Albano M, Dubnau D. 2002. Rok (YkuW) regulates genetic competence in Bacillus subtilis by directly repressing comK. Molecular microbiology 43.

10. Smits W, Grossman A. 2010. The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in Bacillus subtilis. PLoS genetics 6.

11. Albano M, Smits W, Ho LT, Kraigher B, Mandic-Mulec I, Kuipers OP, Dubnau D. 2005. The Rok protein of Bacillus subtilis represses genes for cell surface and extracellular functions. Journal of bacteriology 187.

12. Kaufenstein M, van der Laan M, Graumann PL. 2011. The three-layered DNA uptake machinery at the cell pole in competent Bacillus subtilis cells is a stable complex. J Bacteriol 193:1633-42.

13. Kidane D, Ayora S, Sweasy JB, Graumann PL, Alonso JC. 2012. The cell pole: the site of cross talk between the DNA uptake and genetic recombination machinery. Crit Rev Biochem Mol Biol 47:531-55.

14. Hahn J, Maier B, Haijema BJ, Sheetz M, Dubnau D. 2005. Transformation proteins and DNA uptake localize to the cell poles in Bacillus subtilis. Cell 122:59-71.

15. Boonstra M, Vesel N, Kuipers OP. 2018. Fluorescently Labeled DNA Interacts with Competence and Recombination Proteins and Is Integrated and Expressed Following Natural Transformation of Bacillus subtilis. mBio 9.

16. Ellison CK, Dalia TN, Ceballos AV, Wang JC-Y, Biais N, Brun YV, Dalia AB. 2018. Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in Vibrio cholerae. Nature Microbiology 3:773.

17. Muschiol S, Balaban M, Normark S, Henriques-Normark B. 2015. Uptake of extracellular DNA: Competence induced pili in natural transformation of Streptococcus pneumoniae. Bioessays 37:426-35.

18. Dubnau D, Blokesch M. 2019. Mechanisms of DNA Uptake by Naturally Competent Bacteria. https://doiorg/101146/annurev-genet-112618-043641 doi:10.1146/annurev-genet-112618-043641.

19. Chen I, Provvedi R, Dubnau D. 2006. A Macromolecular Complex Formed by a Pilin-like Protein in Competent Bacillus subtilis*S. J Biol Chem 281.

20. Provvedi R, Dubnau D. 1999. ComEA is a DNA receptor for transformation of competent Bacillus subtilis. Molecular microbiology 31.

21. Inamine GS, Dubnau D. 1995. ComEA, a Bacillus subtilis integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. J Bacteriol 177:3045-51.

22. Damke PP, Di Guilmi AM, Varela PF, Velours C, Marsin S, Veaute X, Machouri M, Gunjal GV, Rao DN, Charbonnier JB, Radicella JP. 2019. Identification of the periplasmic DNA receptor for natural transformation of Helicobacter pylori, Nat Commun, vol 10.

23. Draskovic I, Dubnau D. 2005. Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. Mol Microbiol 55:881-96.

24. Maier B, Chen I, Dubnau D, Sheetz M. 2004. DNA transport into Bacillus subtilis requires proton motive force to generate large molecular forces. Nature structural & molecular biology 11.

25. Chilton SS, Falbel TG, Hromada S, Burton BM. 2017. A Conserved Metal Binding Motif in the Bacillus subtilis Competence Protein ComFA Enhances Transformation, J Bacteriol, vol 199.

26. Kramer N, Hahn J, Dubnau D. 2007. Multiple interactions among the competence proteins of Bacillus subtilis. Mol Microbiol 65:454-64.

27. Yadav T, Carrasco B, Hejna J, Suzuki Y, Takeyasu K, Alonso JC. 2013. Bacillus subtilis DprA recruits RecA onto single-stranded DNA and mediates annealing of

complementary strands coated by SsbB and SsbA. The Journal of biological chemistry 288.

28. Ayora S, Carrasco B, Cardenas PP, Cesar CE, Canas C, Yadav T, Marchisone C, Alonso JC. 2011. Double-strand break repair in bacteria: a view from Bacillus subtilis. FEMS Microbiol Rev 35:1055-81.

29. Yadav T, Carrasco B, Myers AR, George NP, Keck JL, Alonso JC. 2012. Genetic recombination in Bacillus subtilis: a division of labor between two single-strand DNA-binding proteins. Nucleic Acids Res 40:5546-59.

30. Bergé M, Moscoso M, Prudhomme M, Martin B, Claverys JP. 2002. Uptake of transforming DNA in Gram-positive bacteria: a view from Streptococcus pneumoniae. Molecular microbiology 45.

31. Baker JA, Simkovic F, Taylor HM, Rigden DJ. 2016. Potential DNA binding and nuclease functions of ComEC domains characterized in silico. Proteins 84:1431-42.

32. dos Santos VT, Bisson-Filho AW, Gueiros-Filho FJ. 2012. DivIVA-Mediated Polar Localization of ComN, a Posttranscriptional Regulator of Bacillus subtilis, p 3661-9, J Bacteriol, vol 194.

33. Spizizen J. 1958. TRANSFORMATION OF BIOCHEMICALLY DEFICIENT STRAINS OF BACILLUS SUBTILIS BY DEOXYRIBONUCLEATE. Proc Natl Acad Sci U S A 44:1072-8.

34. Lucena D, Mauri M, Schmidt F, Eckhardt B, Graumann PL. 2018. Microdomain formation is a general property of bacterial membrane proteins and induces heterogeneity of diffusion patterns. BMC Biol 16:97.

35. Burghard-Schrod M, Altenburger S, Graumann PL. 2020. The Bacillus subtilis dCMP deaminase ComEB acts as a dynamic polar localization factor for ComGA within the competence machinery. Molecular microbiology 113.

36. Guiziou S, Sauveplane V, Chang HJ, Clerté C, Declerck N, Jules M, Bonnet J. 2016. A part toolbox to tune genetic expression in Bacillus subtilis. Nucleic Acids Res 44:7495-508.

37. Overkamp W, Beilharz K, Detert Oude Weme R, Solopova A, Karsens H, Kovács Á T, Kok J, Kuipers OP, Veening JW. 2013. Benchmarking Various Green Fluorescent Protein Variants in Bacillus subtilis, Streptococcus pneumoniae, and Lactococcus lactis for Live Cell Imaging, p 6481-90, Appl Environ Microbiol, vol 79.

38. Defeu Soufo HJ, Graumann PL. 2004. Dynamic movement of actin-like proteins within bacterial cells, p 789-94, EMBO Rep, vol 5.

39. Schindelin J. 2018. The ImageJ ecosystem: An open platform for biomedical image analysis - Schindelin - 2015 - Molecular Reproduction and Development - Wiley Online Library. doi:10.1002/mrd.22489.

40. Paintdakhi A, Parry B, Campos M, Irnov I, Elf J, Surovtsev I, Jacobs-Wagner C. 2016. Oufti: An integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. Mol Microbiol 99:767-77.

41. Jaqaman K, Loerke D, Mettlen M, Kuwata H, Grinstein S, Schmid SL, Danuser G. 2008. Robust single-particle tracking in live-cell time-lapse sequences. Nat Methods 5:695-702.

42. Rösch TC, Oviedo-Bocanegra LM, Fritz G, Graumann PL. 2018. SMTracker: a tool for quantitative analysis, exploration and visualization of single-molecule tracking data reveals highly dynamic binding of B . subtilis global repressor AbrB throughout the genome. Scientific Reports 8:1-12.

43. Rosch TC, Altenburger S, Oviedo-Bocanegra L, Pediaditakis M, Najjar NE, Fritz G, Graumann PL. 2018. Single molecule tracking reveals spatio-temporal dynamics of bacterial DNA repair centres. Sci Rep 8:16450.

44. Schenk K, Hervas AB, Rosch TC, Eisemann M, Schmitt BA, Dahlke S, Kleine-Borgmann L, Murray SM, Graumann PL. 2017. Rapid turnover of DnaA at replication origin regions contributes to initiation control of DNA replication. PLoS Genet 13:e1006561.

45. Schibany S, Kleine Borgmann LAK, Rosch TC, Knust T, Ulbrich MH, Graumann PL. 2018. Single molecule tracking reveals that the bacterial SMC complex moves slowly relative to the diffusion of the chromosome. Nucleic Acids Res 46:7805-7819.

46. Corbinais C, Mathieu A, Damke PP, Kortulewski T, Busso D, Prado-Acosta M, Radicella JP, Marsin S. 2017. ComB proteins expression levels determine Helicobacter pylori competence capacity. Scientific reports 7.

47. Seitz P, Pezeshgi Modarres H, Borgeaud S, Bulushev RD, Steinbock LJ, Radenovic A, Dal Peraro M, Blokesch M. 2014. ComEA Is Essential for the Transfer of External DNA into the Periplasm in Naturally Transformable Vibrio cholerae Cells, PLoS Genet, vol 10.

48. Seitz P, Blokesch M. 2014. DNA Transport across the Outer and Inner Membranes of Naturally Transformable Vibrio cholerae Is Spatially but Not Temporally Coupled, mBio, vol 5.

49. Hahn J, Kramer N, Briley K, Jr., Dubnau D. 2009. McsA and B mediate the delocalization of competence proteins from the cell poles of Bacillus subtilis. Mol Microbiol 72:202-15.

50. Schavemaker PE, Boersma AJ, Poolman B. 2018. How Important Is Protein Diffusion in Prokaryotes? Front Mol Biosci 5.

51. Najjar NE, Andari JE, Kaimer C, Fritz G, Rösch TC, Graumann PL, Vieille C. 2018. Single-Molecule Tracking of DNA Translocases in Bacillus subtilis Reveals Strikingly Different Dynamics of SftA, SpoIIIE, and FtsA. doi:10.1128/AEM.02610-17.

52. Chen I, Christie PJ, Dubnau D. 2005. The Ins and Outs of DNA Transfer in Bacteria. Science 310:1456-60.

53. Johnston C, Campo N, Bergé MJ, Polard P, Claverys JP. 2014. Streptococcus pneumoniae, le transformiste. Trends in microbiology 22.

54. Provvedi R, Chen I, Dubnau D. 2001. NucA is required for DNA cleavage during transformation of Bacillus subtilis. Mol Microbiol 40:634-44.

55. Zafra O, Lamprecht-Grandío M, de Figueras CG, González-Pastor JE. 2012. Extracellular DNA release by undomesticated Bacillus subtilis is regulated by early competence. PloS one 7.

56. Gangel H, Hepp C, Müller S, Oldewurtel ER, Aas FE, Koomey M, Maier B. 2014. Concerted Spatio-Temporal Dynamics of Imported DNA and ComE DNA Uptake Protein during Gonococcal Transformation, PLoS Pathog, vol 10.

57. Chen I, Gotschlich EC. 2001. ComE, a Competence Protein from Neisseria gonorrhoeae with DNA-Binding Activity, p 3160-8, J Bacteriol, vol 183.

2.2.8 Manuscript II- supplementary material

Single molecule dynamics of DNA receptor ComEA, membrane permease ComEC and taken up DNA in competent *Bacillus subtilis* cells

Marie Burghard-Schrod, Alexandra Kilb, Kai Krämer and Peter L. Graumann

Supplementary movies

All movies are played with a frame rate of 4 frames per second (fps) and not with its original frame rate (50 ms or 30 ms), in order to improve visualisation. Cell borders of *Bacillus subtilis* cells are indicated in white. Detected sub-resolution signals appearing with a minimum length of 5 frames, which defines the movements as tracks of a molecule, are shown in red.

Movie S1: *B. subtilis* cell expressing ComEC-mV, 50 ms stream acquisition. Fluorescence signal appears at both cell poles and at the septum. An overlay of the tracks is shown at the end of the movie.

Movie S2: *B. subtilis* cells incubated with labeled DNA (genotype*: Δrok*), 30 ms acquisition. A clear signal at the cell pole was detected (left cell), even though only little signal was detected in general (see figure 7, table 4)- The signal is bleached (frame 10) but reappears during DNA-uptake (frame 20).

Movie S3: competent *B. subtilis* cell (genotype: *ΔrokΔcomEC*) incubated with labeled DNA, 30 ms acquisition. Signal was most prominent at the cell border. See mobile track, starting at frame 45.

Movie S4: *B. subtilis* cell expressing mVenus-ComEA, 30 ms acquisition. Movie shows one exemplary mobile track (starting frame 4), moving from the cell pole towards the membrane, and one mobile track at the cell pole (frame 33.)

Supplementary figures

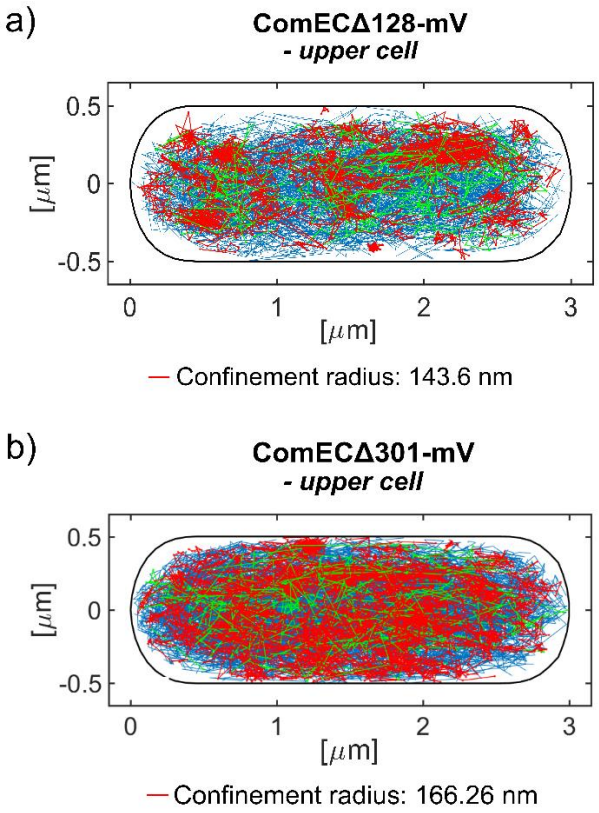


Figure S1: Overlay of tracks on one cell for ComECΔ128-mV and ComECΔ301-mV.

Additional data to single-molecule tracking of truncated versions of ComEC-mV. The figure shows overlays of all tracks (biological triplicate) on one cell, for a) ComECΔ128-mV (PG3818, 915 tracks) and b) ComECΔ301-mV (PG3819, 1617 tracks). We detected an increased static population (see figure 5) and a higher number of confined tracks (red, compare signal a) to signal b)), diffusing in a more static manner. Confinement radius is depicted underneath the normalised cells.

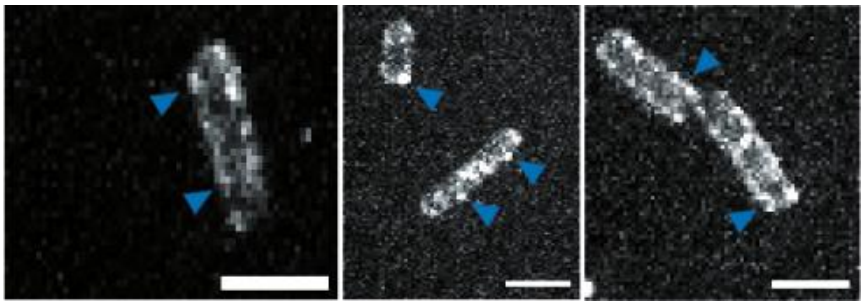


Figure S2: t-stacks of movies of ComECΔ128-mV.

We detected the truncated fusion ComECΔ128-mV (PG3818) at the membrane, when analysed by SMT (mid-cell), and t-stacks of movies were performed, similar to the wild-type full-length protein ComEC-mV (PG3817). Blue arrows indicate signal at the membrane (left and right), and a distinct signal at the cell pole (upper cell, centred image). Scale bar indicates 2 µm.

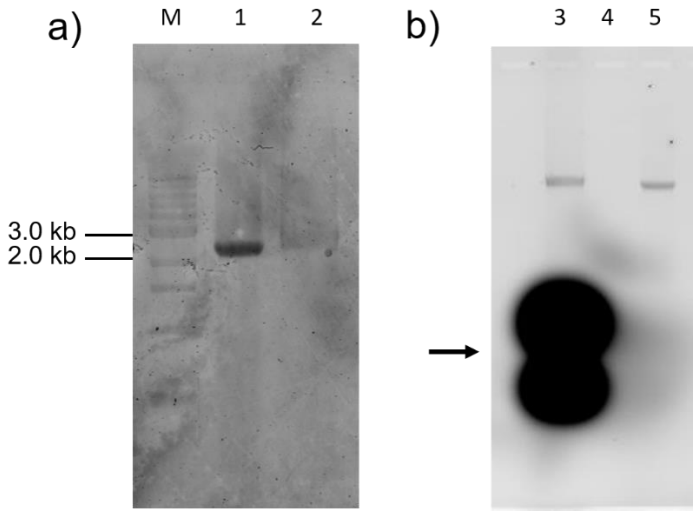


Figure S3: Agarose-gel of labelled PCR-product.

PCR-product was stained with DyLight488 (Thermo Scientific™), followed by detection of in-gel fluorescence in a 1% Agarose-gel (Typhoon TRIO, Amersham Biosciences) in order verify staining and removal of residual dye. a) Samples were excited at 488 nm and detected with an emission filter of 532 nm. For an overlay with the DNA-ladder, Marker was detected at 633 nm, leading to a slightly unspecific signal of the marker and the unstained PCR-product. M= Marker (1kb DNA ladder, NEB), 1= labelled PCR-product (2300 bp), 2= unlabelled PCR-product. b) Without washing, residual dye was detected (indicated by a black arrow). 3= PCR-

product without washing, 4= PCR-product, unstained, 5= washed, stained PCR-product. Gel was excited by 488 nm laser.

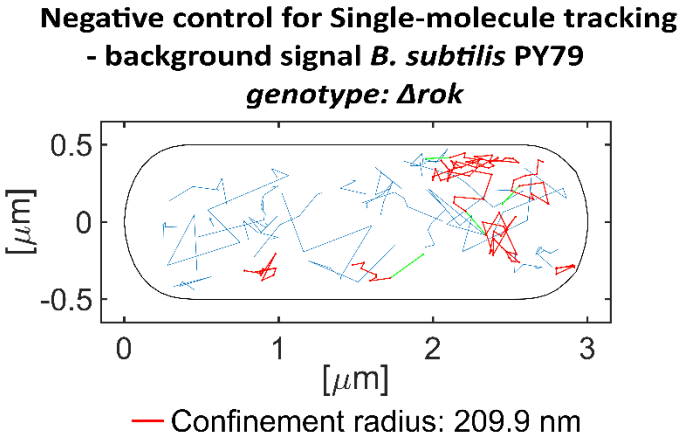


Figure S4: Negative control for Single-molecule tracking.

As a negative control for SMT, we analysed the background signal of *B. subtilis* cells, harbouring no fusion and without incubation of fluorescent DNA (PG876) at mid-cell (50% laser, 514 nm excitation). Figure shows an exemplary measurement of one day, where 29 tracks were detected (14 cells analysed). A large error was detected (69.96 nm), leading to a confinement radius of 209.9 nm. In total, a biological triplicate was performed leading to 135 tracks (55 cells analysed) and a mean of 2.45 tracks/cell.Typically, tracks did not show any defined pattern.

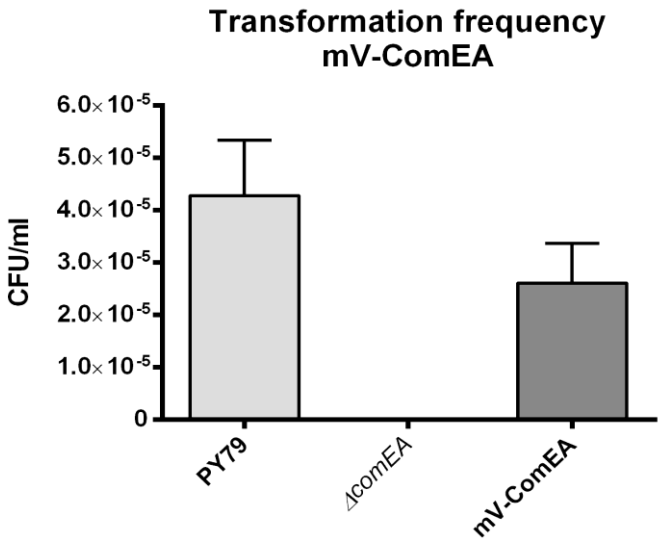


Figure S5: Transformation frequency of mV-ComEA.

The N-terminal mVenus-fusion of ComEA was transformable (PG3909, 2.60^{-5} +/- 7.59^{-6} CFU/ml), and reached 60.7% of the wild-type control PY79 (PG001, 4.27^{-5} +/- 1.06^{-5} CFU/ml). As a negative control, a strain encoding an erythromycin resistance cassette of *comEA* was used (PG3721, indicated by genotype, *ΔcomEA*). Measurements were carried out by use of chromosomal DNA encoding an Erythromycin or Kanamycin resistance cassette (PG3717, PG876) as technical triplicates on two different days. Error bar indicates standard deviation of the mean.

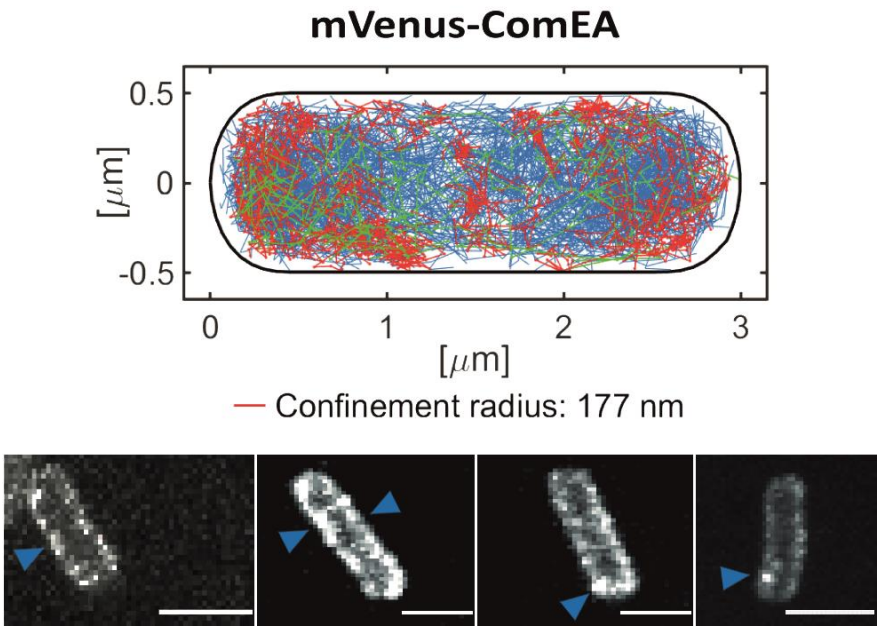


Figure S6: Single-molecule tracking of mV-ComEA induced with 0.5% Xylose.

We found an even more distinct localization pattern of mVenus-ComEA in case the protein was induced with 10 times excess of Xylose compared to our common work concentration of 0.05%. The protein appeared to localise at the membrane and the septum (see upper panel, overlay of 1228 tracks, exemplary measurement of one day), as well as at the cell pole (see lower panel, exemplary t-stacks of movies, blue arrows). Scale bar indicates 2 µm. In a biological triplicate, we detected 2886 tracks (41.8 tracks/cell), which we considered a non-physiological large amount of competence protein, and decided to use 0.05% Xylose for further measurements. In addition, if a concentration of fluorescent protein-fusion is too high, we could not exclude the formation of artefacts.

Table S1: List of strains, B. subtilis

strain	genotype	plasmid	resistance	reference
PG001	PY79	No plasmid	No resistance	wild-type strain
PG681	*comk-CFP*	original locus	Cm	gift from D. Dubnau
PG876	*Δrok*	resistance cassette, original locus	Kan	gift from D. Dubnau, name in publication: BD3196
PG3717	*ΔcomGA*	resistance cassette, original locus	MLS	BGSC-ID: BKE24730
PG3721	*ΔcomEA*	resistance cassette, original locus	MLS	BGSC-ID: BKE25590
PG3722	*ΔcomEC*	resistance cassette, original locus	MLS	BGSC-ID: BKE25570
PG3816	*ΔrokΔcomEC*	resistance cassette, original locus	Kan/MLS	this study
PG3817	*comEC-mVenus*	pDL_NLMV	Cm, Xyl	this study
PG3818	*comECΔ128-mVenus*	pDL_NLMV	Cm, Xyl	this study
PG3819	*comECΔ301-mVenus*	pDL_NLMV	Cm, Xyl	this study
PG3909	*mVenus-comEA*	pHJDS_NLMV	Cm, Xyl	this study

PG3917	*ΔcomEC, amyE:comEC H571A*	pDR111	Spec/MLS, IPTG	this study
PG3918	*ΔcomEC, amyE:comEC D573A*	pDR111	Spec/MLS, IPTG	this study
PG3919	*ΔcomEC, amyE:comEC D575A*	pDR111	Spec/MLS, IPTG	this study
PG3927	*ΔcomEC, amyE:comEC D668A*	pDR111	Spec/MLS, IPTG	this study
PG3928	*ΔcomEC, amyE:comEC K671A*	pDR111	Spec/MLS, IPTG	this study
PG3929	*ΔcomEC, amyE:comEC*	pDR111	Spec/MLS, IPTG	this study

Table S2: List of strains, E. coli DH5α

strain	plasmid	resistance	reference
PG2056	pDL-mVenus	Amp	Lucena *et al.*, (2018)
PG3692	pHJDS-mV	Amp	Defeu & Graumann (2004), modified via *Kpn* I and *BamH* I (Defeu Soufo & Graumann, 2004)
PG3789	pDR111 (BGSCID: ECE312)	Amp	Ben Yehuda *et al.*, (2003)
PG3803 DH5α wild-type	none	none	ThermoScientific™
PG3820	pDL-mV x *comEC*	Amp	this study
PG3821	pDL-mV x *comECΔ128*	Amp	this study
PG3822	pDL-mV x *comECΔ301*	Amp	this study
PG3877	pHJDS-mV x *comEA*	Amp	this study
PG3901	pDR111 x *comEC* H571A	Amp	this study
PG3902	pDR111 x *comEC* D573A	Amp	this study
PG3903	pDR111 x *comEC* D575A	Amp	this study
PG3904	pDR111 x *comEC* 668A	Amp	this study
PG3905	pDR111 x *comEC* K671A	Amp	this study
PG3906	pDR111 x *comEC*	Amp	this study

Table S3: List of primer, named by E. coli strains or purpose. Restriction sites and mutations are indicated in italic, SD-sequence in bold.

Primer	**Sequence 5´→ 3´**	**Usage/ created plasmid**
Ery fw	CACAGTCAAAACTTTATTACTTC	proves chromosomal integration (PG3917-19, PG23927-29)
Ery rev	CTTATTAAATAATTTATAGCTATTG AAAAG	
EC fw	GGGTCTAATAGGGAAATTGG	proves chromosomal integration (PG3817-19)
mV rev	GCTATTACGCCAGCTGGCG	
RBS fw	gattaactaata**AGGAGG**acaaacATGCGTA ATTCGCGCTTATTATTG	integration of RBS by PCR pf pDR111 x *comEC*
RBS rev	GCTAGCTGTCGACTAAGCTTAATT	
Seq pDR111 fw	GTTGCTCGAGGGTAAATG	Sequencing of genes clones into pDR111
Seq pDR111 rev	GCTAGCTGTCGACTAAGCTTAATT	
Seq pHJDS-mV fw	CCAACGAGAAGCGCGATCAC	Sequencing plasmid pHJDS-mV (PG3877)
Seq pDL-mV rev	GGCGCGGGTCTTCTAGTTGC	Sequencing plasmid pDL-mV (PG3820-22)
PG3820 fw	GTAATA*agatct*CCAGGGAAGAGGGA GTG	pDL-mV x comEC
PG3820 rev	GTGATA*gggccc*GTTCGTCTCTGTTAT ATCTGATG	
PG3821 fw	GAGTAT*gaattc*TTTGCTGCACGGTGA TGTTT	pDL-mV x comECΔ128
PG3821 rev	ATACTC*gggccc*TTTGCTTGCCGGATC AGGTG	

PG3822 fw	GAGTAT*gaattc*TTTGCCTTTCGTACA TCGTC	pDL-mV x comECΔ301
PG3822 rev	GTACGT*gggccc*CTGCGACAAGGAGC GTTTTT	
PG3877 fw	GTGATA*gggccc*AATTGGTTGAATCA GCA	pHJDS-mV x comEA
PG3877 rev	ATAGTG*gaattc*CTTCAGCTTTGGATG GC	
PG3901 fw	*gcc*GCTGACCAAGATCATATC	pDR111 x *comEC* H571A
PG3901 rev	CGTCAGAATTAAAGCGTCAAG	
PG3902 fw	CACGCT*gcc*CAAGATCATATC	pDR111 x *comEC* D573A
PG3902 rev	CGTCAGAATTAAAGCGTCAAG	
PG3903 fw	*gct*CATATCGGAGAGGCG	pDR111 x *comEC* D575A
PG3903 rev	TTGGTCAGCGTGCGTCAG	
PG3904 fw	*gct*CTGGAGAAAGAAGGGG	pDR111 x *comEC* 668A
PG3904 rev	ACCCGTCAAGATCCAGCTC	
PG3905 fw	*gca*GAAGGGGAACAAGAGG	pDR111 x *comEC* K671A
PG3905 rev	CTCCAGATCACCCGTCAAG	
PG3906 fw	GAGTCT*gctagc*ATGCGTAATTCGCG CTTAT	pDR111 x comEC
PG3906 rev	GAGTCT*gcatgc*TTAGTTCGTCTCTGT TATATCTGAT	

Supplemental methods

The GMM method analyses the diffusion of single molecules using their localization and their single-step displacements distribution (in each x- and y-direction). Assuming Brownian motion of single particles, a Gaussian Mixture Model can be fitted to the empirical probability density function of those single displacements using a multi-state diffusion model (see figure 6). Each component of the mixture model is considered to be a Gaussian distribution with zero mean and a standard deviation $\sigma=2D\Delta t$, where D is the respective diffusion constant and Δt the time interval between image frames. As a comparative methodology of one molecule species, diffusion constants are fixed among the different conditions, and up to three different diffusive populations can be determined. Using this method, focus lies on the change of population sizes, which is represented by the weight assigned to each single component of the model.

The Squared displacement (SQD) analysis uses the empirical cumulative distribution of squared displacements, which can be fitted to a lineal combination of up to 3 cumulative distribution functions of the probability that a particle stays in a circle of radius r within a time *t*. In the same manner as in the GMM methodology, diffusion constants and diffusive fractions of up to three different diffusive states can inferred from the fitting, indicating distinct biological functions. For reasons of clarity and better visual presentation of the three determined populations, Jump distance histograms are presented instead, along with the individual fitted distributions. Jump distance means the Euclidean distance between two localizations, equivalent to the square root of the square displacements. Check (Rösch *et al.*, 2018) for further details of both methods.

2.3 Unpublished results

2.3.1 Expression and purification of B. subtilis ComEC periplasmic domains

Apparently, the putative enzyme activity of ComEC as a DNA-binding nuclease is linked to hydrophilic domains that are located in the periplasmic space of *B. subtilis* (Draskovic & Dubnau, 2005). Baker *et al.* (2016) carried out a detailed modelling of the protein´s N-terminus, where an OB-fold (oligosaccharide-binding-fold) was discovered, possibly able to bind transforming DNA inside of the *B. subtilis* periplasm. In addition, an exonuclease function for the C-terminus was proposed (see introduction, 1.2.3). At the current state of knowledge, no *in vitro* data on ComEC´s activity were available. Former studies described a toxic effect of the full-length protein when expressed in *E. coli* cells (Draskovic & Dubnau, 2005), thereby hampering a purification and analysis of the proteins function. As a consequence of this, an experiment was designed to truncate the protein, thereby removing its transmembrane helices (the competence domain, located at the center of the protein sequence, please see unpublished methods for details and Figure 5), while its predicted, soluble N-terminal and C-terminal parts should be heterologously expressed in *E. coli*. It was planned to measure activity of the periplasmic protein domains *in vitro* by an EMSA (electro-mobility-shift assay) and an exonuclease-activity assay. Further, it has been expected that removal of hydrophobic transmembrane helices would ease protein purification by increasing the general solubility. Based on studies of Draskovic *et al.* (2005) and Baker *et al.* (2016), the N- and C-terminal sequences of *B. subtilis comEC* were cloned into the pGAT2 and pET16b expression systems and expressed in *E. coli* BL21(DE3) cells. The N-terminal and a C-terminal domain (named C-loop) were purified during supervision of the master thesis of Kai Krämer. It has been found that the truncations were only soluble in case a GST-tag was fused to the sequences (pGAT2 constructs). An EMSA of ComEC´s N-terminus (also named N-loop) was hampered by very high DNA- or RNA-contamination. The problem has been addressed by Mr. Krämer using an anionic-exchange column or higher salt concentrations of the buffer during protein purification, but purity of the N-loop remained very low. This was indicated by a ratio of 260/280 nm $\geq$ 1. In addition,

obvious DNA or RNA contaminations were detected when the protein was loaded on an agarose gel stained with midori green (see (Krämer, 2019)).

In order to analyse the nuclease activity of ComEC, several different C-terminal sequences were fused to GST based on the presence of a putative soluble domain named C-loop (Draskovic & Dubnau, 2005) and the conserved, zinc-binding motif, which might be required for its putative exonuclease activity (Baker *et al.*, 2016). It was found that the β-lactamase and the larger β-lactamase-like nuclease domain were insoluble under the applied conditions (see Methods, 3.5-3.7) when expressed in *E. coli* BL21 (DE3). Since the C-loop-region (Draskovic & Dubnau, 2005) was soluble when fused to GST, the protein sequence was step-by step enlarged until it contained the relevant Zinc-binding motifs identified by Baker *et al.* (2016). An overview of the protein sequences of all generated truncations is shown in Figure 5.

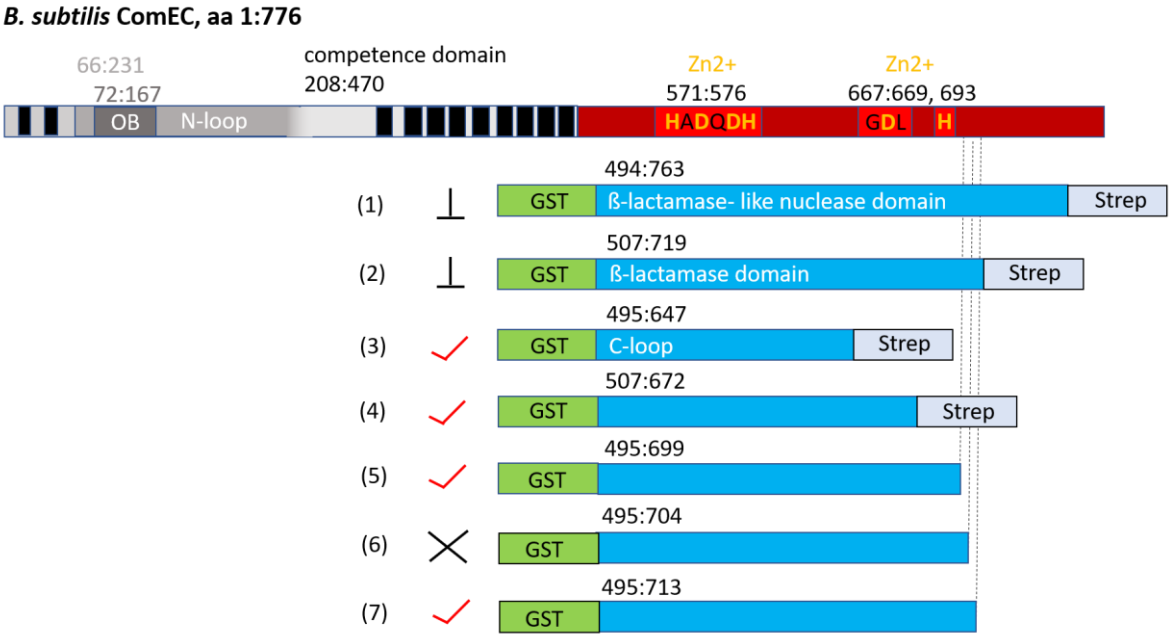


Figure 5: Overview of B. subtilis ComEC C-terminal truncations

The protein sequence of ComEC is shown at the top of the figure. The N-loop is indicated in grey and the C-loop in red. Putative Zinc-binding sites located at the C- terminus are indicated in yellow, as well as the predicted, conserved amino acids of the zinc-binding motifs. Transmembrane helices are indicated as black boxes at the very N-terminus of the protein sequence and within the competence domain. Expressed sequences are shown in blue with the

particular amino acid position indicated above. Sequences of constructs (1) and (2) were chosen based on former work of Baker *et al.* (2016). GST-tag fusions (green) of the sequences were found to be insoluble, indicated by the black inhibition symbol. Sequence (3) was predicted as soluble by Draskovic *et al.* (2005). The prediction was proven by the experiments described herein, at least as N-terminal GST-fusion. Solubility is indicated as red check mark. Furthermore, sequences were extended to include the conserved motifs required for putative exonuclease activity. C-terminal Strep-tag, which was used in the beginning for affinity purification, was removed in construct (5)-(7). *E. coli* BL21 (DE3) cells transformed with construct (6) did not grow twice (indicated by black X), so further experiments were carried out with construct (4), (5) and (7).

Protein sequences were purified either by strep-tag (Figure 6) or GST-tag affinity purification (Figure 7, 8). A high DNA/RNA contamination was detected, which was also visible on agarose and polyacrylamide gels. This would, again, hamper any further analysis. To overcome this issue, cells were lysed via sonification. Furthermore, proteins were subjected to a second purification via a heparin column, applying an increasing gradient of sodium chloride (Figure 6, construct 4), or an incubation of the lysate with a large amount DNAse followed by long washing step (Figure 7, construct 5), which was, in case of construct (7), followed by gel filtration (Figure 8). Applying these different settings (see 3. Methods, 3.5-3.7 protein purification), the ratio 260/280 nm decreased to 0.55 -0.7. However, contamination was still visible on agarose gels (Figure 6). A pure dimer of construct (7), which contained all conserved zinc-binding motifs required for activity, was finally isolated, but experiments under the conditions tested herein (see method 3. 9, EMSA & assays of nuclease activity) did not yield a hint for any exonuclease activity. A high tendency of the tagged proteins to aggregate/multimerise was found by gel filtration (Figure 6-8), and mass photometry. As the multimerization, which was most likely caused by the N-terminal GST-tag (Tudyka & Skerra, 1997, Vinckier *et al.*, 2011), could have hampered former assays of activity (e.g. EMSA, protein aggregates would remain in the slots of gel), experiments were carried out to remove the tag. Even though different thrombin concentrations and temperatures were used, constructs 4,5 and 7 precipitated shortly after addition of the protease (see method 3.6, affinity chromatography). The tag seemed to highly affect the oligomerization state of the protein truncations. Finally,

considering the obvious contamination by nucleotides, an exonuclease-activity of ComEC´s C-terminus, which commonly also requires DNA- binding, cannot be excluded.

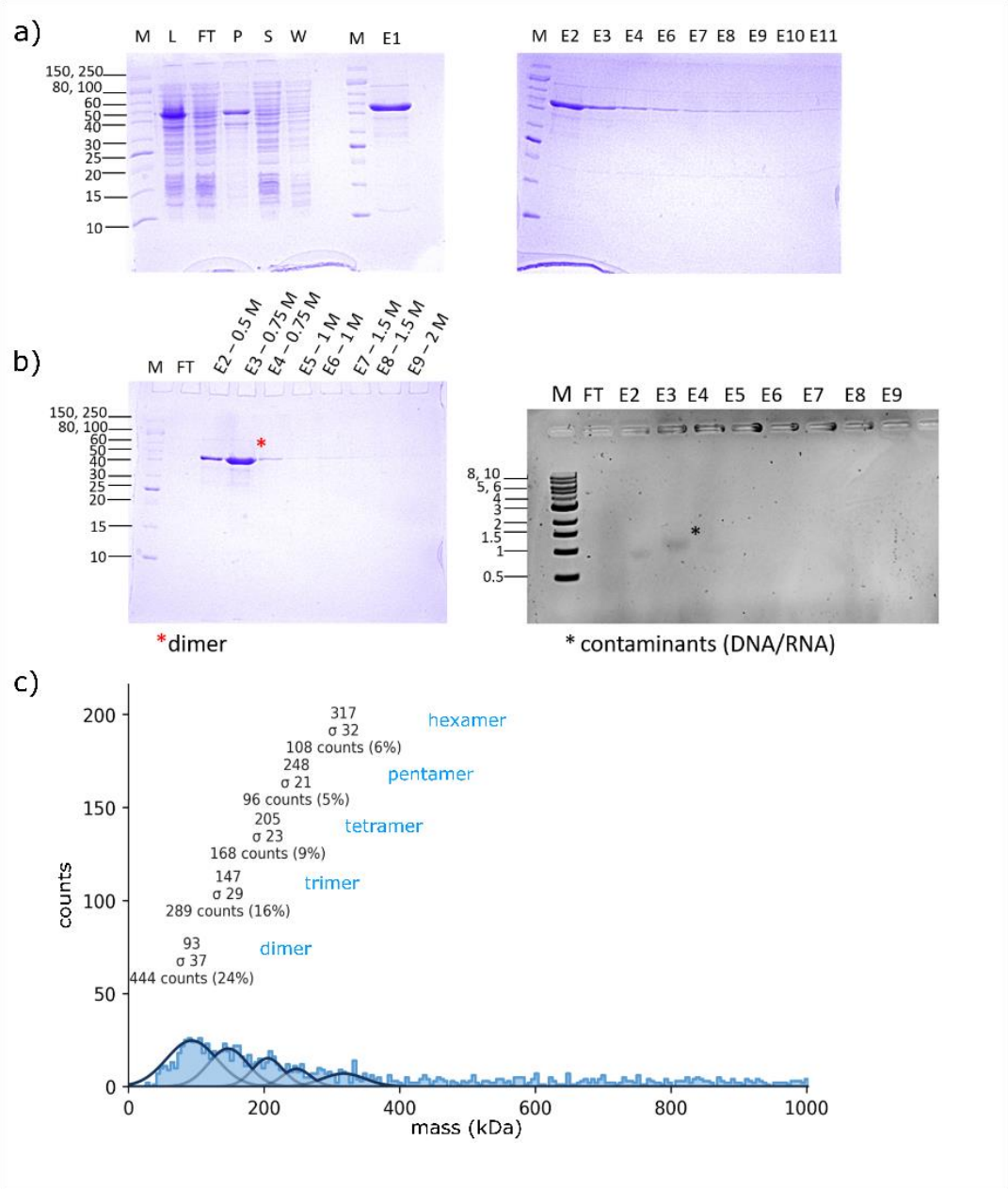


Figure 6: Strep-tag affinity chromatography, heparin affinity chromatgraphy and mass photometry of construct (4)

a) shows a 15% SDS- polyacrylamide gel stained with Coomassie. Samples were taken from a Strep-tag affinity chromatography of construct (4). The truncation had an expected size of 46.33 kDa. Left image: M= unstained protein standard broad range (NEB). L= lysate, FT= flow through, P= pellet/membrane fraction, S= supernatant, W= washing fraction, right image: M= unstained protein standard, broad range (NEB), E1-E11= elution fraction.

In b) the elution fractions from a), containing the desired protein, were pooled and subjected to a heparin affinity chromatography in order to decontaminate the sample from nucleic acids using a gradient of sodium chloride. Left image: Different concentrations are indicated on top of a 15%, Coomassie stained SDS- polyacrylamide gel by which the proteins on the gel were detected. Protein of interest is indicated by red star, while later on a dimeric main conformation was detected. M= unstained protein standard broad range (NEB), FT= flow through, E2-E9= elution fractions. Right image: 1% Agarose gel, stained with Midori Green. Contaminants are indicated with black star. M= 1kb DNA Ladder (NEB), FT= flow through, E2-E9= elution fractions.

c) shows the corresponding mass photometry histogram, carried out with 50 nM of pooled elution fractions from b). Several multimers were detected, as indicated in blue next to the molecular mass of the proteins. The percentage of the measured multimers and the number of counts are indicated and describe the abundance of the distinct protein species.

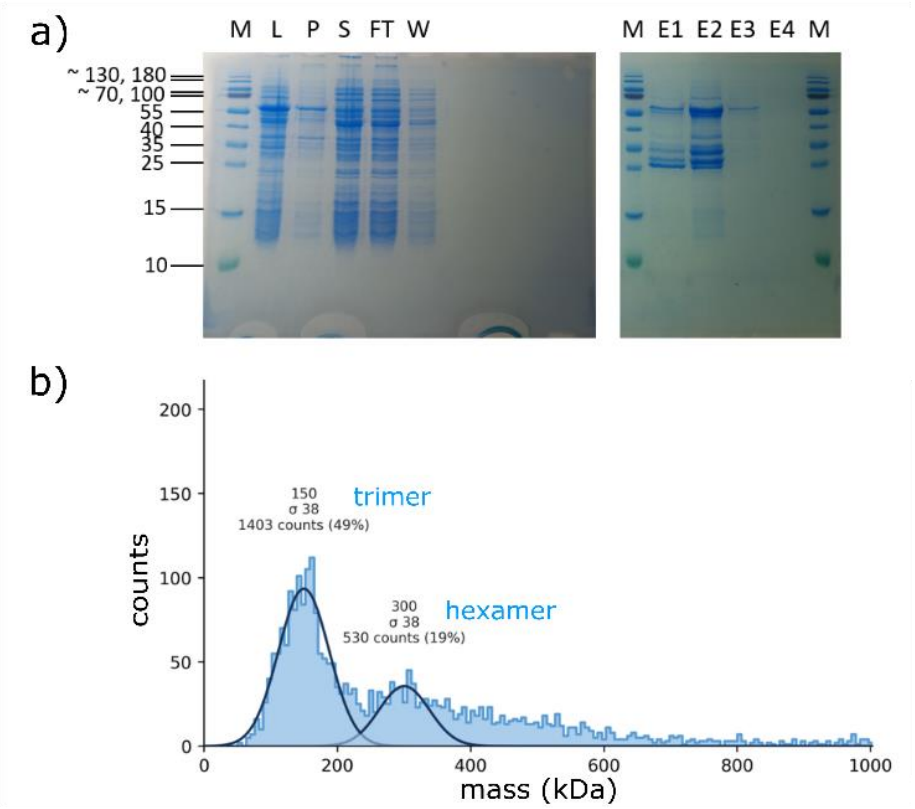


Figure 7: GST-tag affinity chromatography and mass photometry of construct (5)

In a) a 15% SDS-polyacrylamide gel is shown stained with Coomassie. Samples were taken from expression and purification of construct (5). The protein exhibited an expected size of 50.89 kDa. Left image: M= Marker/prestained protein standard (NEB). L= lysate, P= pellet/membrane fraction, S= supernatant, W= washing fraction, FT= flow through, right image: M= prestained protein standard (NEB), E1-E4= elution fractions.

b) shows mass photometry histogram (10 nmol protein) of fraction E3. A trimer and a hexamer were detected as predominant species.

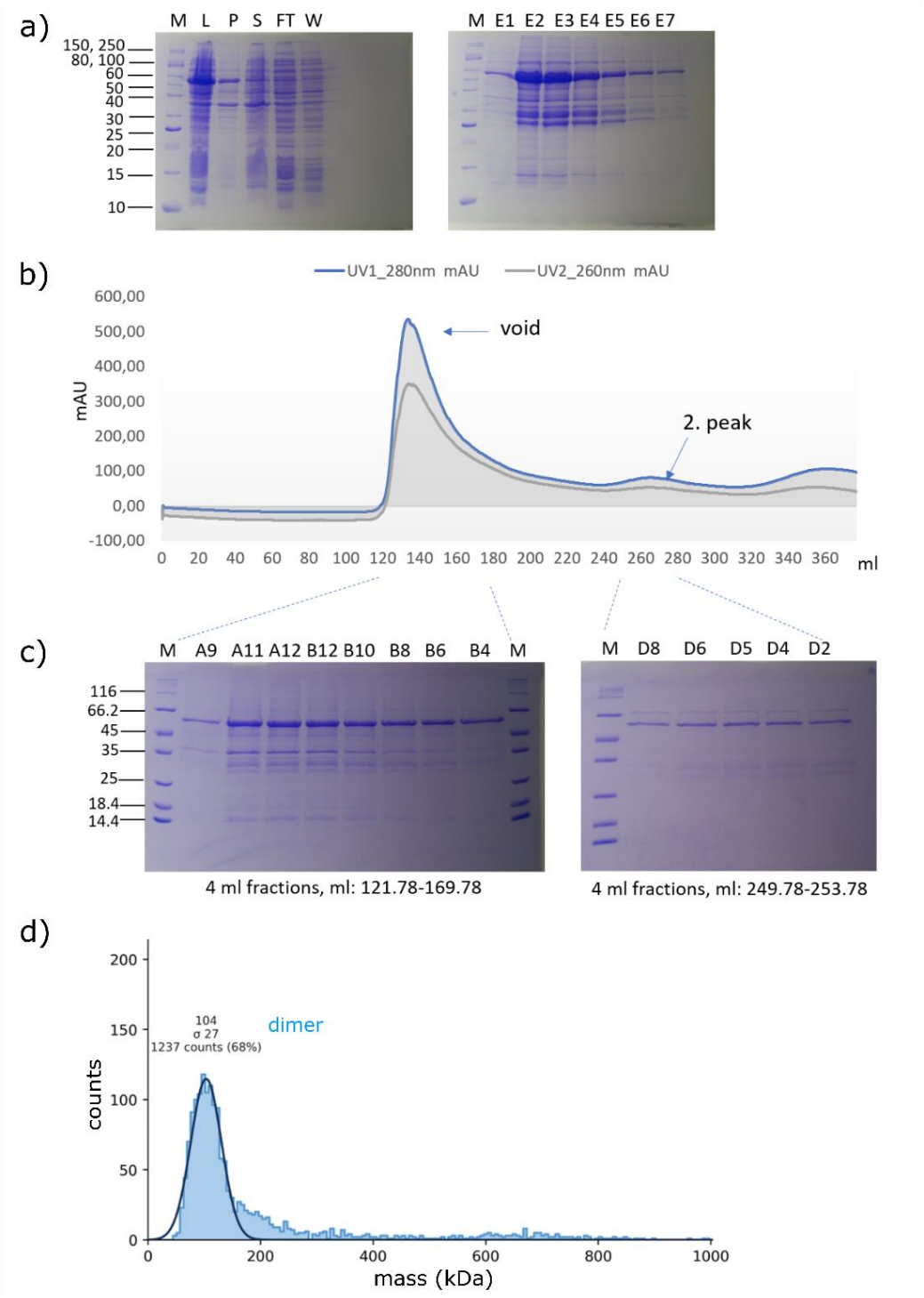


Figure 8: GST-tag affinity chromatography, size exclusion and mass photometry of construct (7)

a) shows two 15% SDS-polyacrylamide gels loaded with samples from GST-tag purification of construct (7). Expected size: 52.04 kDa. Left image: M= unstained protein standard broad range (NEB,) L= lysate, P= pellet/membrane fraction, S= supernatant, FT= flow through, W=

washing fraction. Right image: M= unstained protein standard broad range (NEB), E1-E7= elution fractions.

In b) a chromatogram of preparative size exclusion chromatography is shown, where elution fractions from a) were pooled and proteins eluted in the void volume and one distinct peak.

In c) the SDS-polyacrylamide gels of the taken samples from the above indicated fractions which corresponding to the detected peaks of the chromatogram are shown. Left image: M=unstained protein standard (gift from AG Bange), A9-B4= 4 ml elution fractions. Right image: M=unstained protein standard (gift from AG Bange), D8-D2= 4 ml fractions.

d) shows analysis of the 2. peak (fractions D8-D2) by mass photometry (50 nM protein). Only one species, a dimer was detected, with the doubled size of a monomer (104 kDa).

2.3.2 Localization of stained DNA in competent B. subtilis cells

These data are additional information to manuscript II and might serve as a control, showing that experiments involving single-molecule tracking of fluorescently labeled DNA, were carried out under proper conditions; meaning in particular the exposure time which has been chosen to monitor the single fluorescent particles inside of the cells.

Competent *B. subtilis* cells were incubated with labeled DNA in order to analyse the localization and behaviour of intracellular DNA during natural transformation via Single-molecule tracking (please see 3. Methods, 3.10 Single molecule tracking of fluorescently labeled DNA). Therefore, a modified protocol of Boonstra *et al.* (2018) was used, as described in manuscript II (Fluorescence staining of DNA, page 86). Cells were incubated with labeled DNA, followed by an incubation step with DNAse in order to remove DNA, which was unspecifically bound to the cell surface, and had not been taken up yet. Finally, samples were washed with medium to remove excess DNA. In a first experiment, an exposure time of 50 ms was used (Figure 9). It was found that the number of tracks was very low, resulting in 1.52 and 2.38 tracks/cell in case of *B. subtilis* PY79 (PG001) and *B. subtilis* PY79 *Δrok* (PG876). Static tracks, defined by a confined radius of 120 nm, were found to localise to the cell pole with a probability of 27.15% and 26.2% (number of cells =21, 25; number of tracks analysed =50, 38; Figure 9 a) and b)). The latter describes the sum of the probability of the polar localization in case the repressor of ComK, namely Rok, was deleted. In addition, in both strains, the longest tracks observed localized at the cell pole, one of them lasting even 6.5 s (Figure 9 c)). *B. subtilis* PY79*Δrok* was used in order to increase the number of tracks/cell. According to former studies, up to 60% of the cells of one culture can become competent in this particular genetic background (Hoa *et al.*, 2002). Surprisingly, the effect was not as intense as expected. Therefore, the experimental set-up was reconsidered. Using a shorter exposure time of 30 ms, it was found that the number of tracks increased up to 5.4 tracks/cell (see 2.2 Manuscript II, 2.2.4 Results, page 104). The usage of a higher exposure time of 50 ms might have caused a loss of signal, either by faster bleaching of the signal or by missing to detect movements inside of the cells.

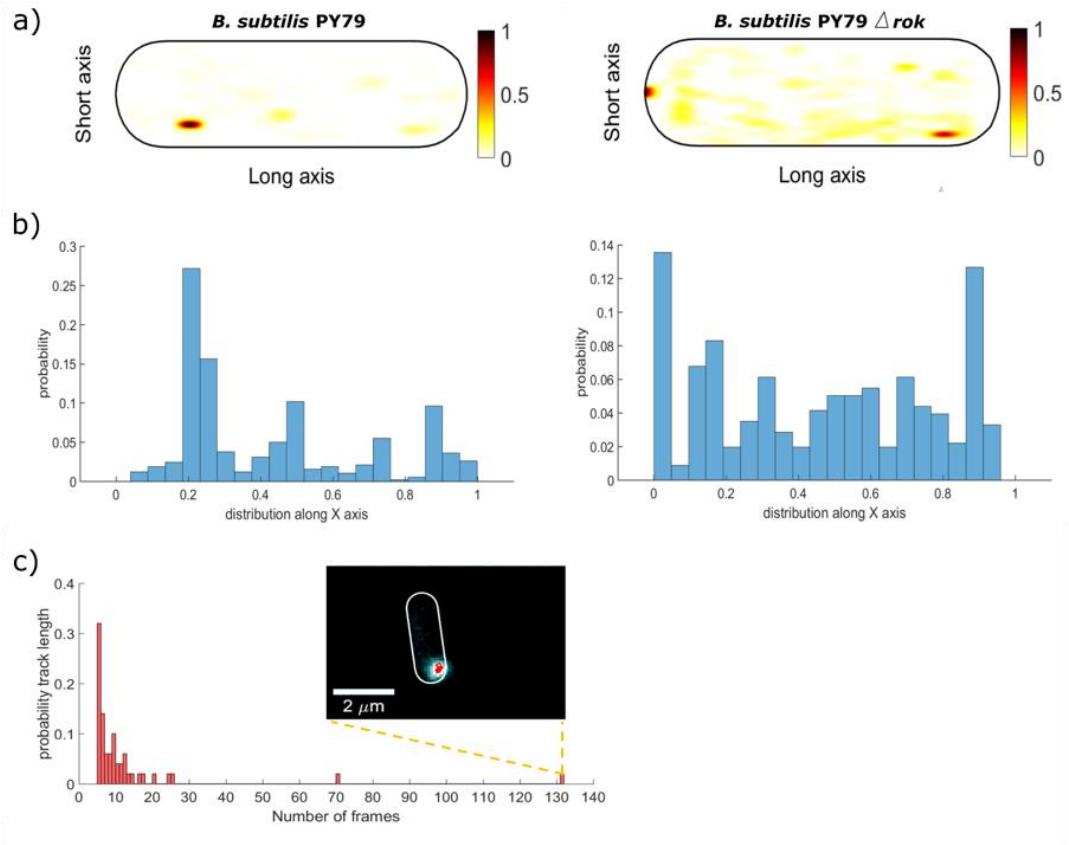


Figure 9: Localization of fluorescently labeled DNA in competent B. subtilis cells

In a) heat maps show the distribution of the fluorescent signal of labeled DNA which has been taken up by competent *B. subtilis* cells. Probability of localization is indicated by the color bar to the right of the images. Two different strains were used, PY79 (PG001, left image) and PY79 *Δrok* (PG876, right image). In both cases, a 514 nm laser was used to excite the DNA which was stained with DyLilght488 (see 2.2 Manuscript II, page 86). Signal was then detected with an exposure time of 50 ms. b) shows x-axis distribution of the tracks underneath the corresponding heat maps. in both cases, highest probability of the location of the signal is present at the cell pole. In c), the probability of tracks is plotted against the number of frames for strain *B. subtilis* PY79. Most tracks do have 5 frames, which was also set as a minimum for track detection. A long track at the cell pole (6.5 s) is highlighted exemplary.

3. Methods (unpublished results)

3.1 Growth conditions

E. coli cells were grown in lysogeny broth medium (LB- medium, composition: 10 g/l trypton, 10 g/l yeast extract, 5 g/l NaCl), where the pH was adjusted to 7.5 using a 10% NaOH solution. LB-medium was supplemented with 100 µg/ml Ampicillin or 25 µg/ml Kanamycin, if required. LB- Agar plates were prepared by usage of LB-medium supplemented with 1.5% Agar. Liquid cultures of *E. coli* and *B. subtilis* cells were always grown at 37°C and 200 rpm. *B. subtilis* cells were grown to develop competence in a special, sterile filtrated competence medium (MC- medium) after Spizizen (Anagnostopoulos & Spizizen, 1961, Spizizen, 1958), which was prepared as a 10x stock solution in a 100 ml volume and stored in 1.5 ml aliquots at -20°C (composition: 14.01 g K_2HPO_4 x 3 H_20, 5.24 g KH_2PO_4, 20 g Glucose, 10 ml trisodium citrate (300 mM), 1 ml ammonium iron(III) citrate (22 mg/ml), 1 g casein hydrolysate, 2 g potassium glutamate). The medium was diluted shortly before use to a volume of 10 ml by addition of 8.7 ml sterile ddH_2O and 1 ml $MgCl_2$ (1 M) to 1 ml of 10x MC-medium. Frozen aliquots were used for a week, until a loss of activity was detected. In case of *B. subtilis cells*, the usage of glass ware was obligate.

3.2 Strain construction

Additional Primer used in this study are listed in Table 1. The resulting strains are listed in Table 2. For protein production in *E. coli*, truncations of the *B. subtilis* gene *comEC* (construct (1)-(7), see Results, Figure 5) were amplified and cloned into vector pGAT2 (gift from AG Bange, modified pGAT2: his-tag was removed from vector backbone) under usage of restriction sites *NcoI* and *XhoI*, which were included in the primer and added to the truncated gene via polymerase chain reaction (PCR). The vector encodes a GST-tag which is fused n-terminally to a gene of interest which is cloned into the MCS. For construct (1)-(4), a *strep*-tag sequence (*strep* tag II: SAWSHPQFEK* or 5´- agc gcc tgg agc cac ccc cag ttc gag aag **tga** -3´) was included

at the c-terminus. In case a *NcoI*-site was present in the gene of interest, a *PciI*-site was added to the truncated genes via the downstream Primer instead of *NcoI*, and *PciI* was ligated to the *NcoI*-site of the digested vector. PCR was carried out following a standard protocol (see 5. Appendix, Table 3). PCR-products were purified via QIAquick PCR purification Kit (Qiagen). Preparative restriction digest was carried out at 37°C for 2 h in a 30 µl volume, according to the protocol of the company providing the enzymes (NEB). Further, PCR-products and linearized vectors were purified after restriction digest via agarose gel extraction kit (Qiagen). Ligation reactions were set up using T4 ligase (NEB) in a molar ratio PCR-product/vector of 1:3 and 1:5 for 1 h at 37°C. Reactions were stopped at 65°C for 15 min and cooled down on ice. Chemical competent *E. coli* DH5α cells were transformed via heat-shock (see 3.3, 3.4). 3 ml of LB medium supplemented with 100 µg/ml Ampicillin were inoculated with single colonies of *E. coli* DH5α. Cultures were grown over-night at 37°C and 200 rpm (KS 4000 I control, IKA). Resulting plasmid DNA was isolated via GenElute plasmid DNA Miniprep Kit (Sigma-Aldrich). Integration of the gene was verified by analytical restriction digest and sequencing (GATC) of the isolated plasmids. All constructs were generated using the software SerialCloner (v 6.2, Molecular Devices).

3.3 Preparation of chemical competent cells

Chemical competent cells of *E. coli* BL21 (DE3) and *E. coli* DH5α were prepared by inoculating 200 ml of LB-medium from a 3 ml overnight culture to an OD_{600} of 0.05. Further, cells were grown at 37°C and 200 rpm to an OD_{600} of 0.6. Cultures were cooled on ice for 30 min and subsequently harvested by centrifugation for 20 min at 3500 rpm and 4°C (Mega Star 1.6R, VWR). Supernatant was removed and cell pellets were carefully resuspended in 15 ml of sterile filtrated, ice-cold $CaCl_2$-buffer (0.1 M $CaCl_2$, 15% Glycerol), using a blue tip; the very tip was cut off to avoid shear forces. Aliquots of 120 µl were instantly frozen in liquid nitrogen and stored at -80°C.

3.4 Transformation of *E. coli*

For heat shock transformation, chemical competent cells were chilled on ice for 2 min. Either 10 µl if ligation reaction or 1-10 ng of purified vector were added to the cells and incubated for 30 min on ice. Samples were incubated at 42°C for 90 sec, quickly removed and incubated on ice for 10 min. Further, cells were resuspended in 800 µl LB medium and incubated at 37°C, 350 rpm (ThermoMixer C, Eppendorf). Finally, 100 µl were plated on agar plates containing the corresponding antibiotic. In order to plate all cells, residual sample was centrifuged at 500 rcf (Centrifuge 5415 D, Eppendorf) for 5 min. The resulting pellet was resuspended in 200 µl medium and plated. Finally, agar-plates were incubated for 16 h at 37°C.

3.5 Protein expression and lysate preparation

E. coli expression strain BL21 (DE3) was always transformed freshly with the plasmid encoding the gene of interest (construct (1)-(7)). For protein purification, 4x 500 ml of medium were distributed on 2 l Erlenmeyer flasks and supplemented with the corresponding antibiotic. Medium was inoculated with a LB- overnight-culture of the corresponding expression strain of BL21 (DE3) to an OD_{600} of 0.05. Protein expression was induced via 0.4 mM IPTG (1 M stock solution) and shifted to 30 °C. Cells were harvested at an OD_{600} 2.5-3 after 2 h of induction via centrifugation at 4 °C and 3500 rpm (Sorvall Lynx 6000 Centrifuge; Thermo Scientific). Supernatant was removed, cell pellets were frozen in liquid nitrogen and stored at -20°C until further use. In order to lyse the cells, pellets were re-suspended in 25 ml binding buffer (Strep-tag binding buffer: 100 mM Tris, 20 mM $MgCl_2$, 1M NaCl, pH 8; GST-tag binding buffer: 20 mM HEPES-Na, 1 M NaCl, 20 mM $MgCl_2$, 20 mM KCl, pH 7.5). One protease inhibitor tablet (EDTA-free, Roche) and 5-20 µl DNAse (100 mg/ml), 30 µl RNAse (100 mg/ml), 20 µl (100mg/ml) Lysozyme were added, mixed, and further incubated for 1.5 h on ice. Sonification was then carried out for 3 min with a pulse of 0.5 sec and an amplitude of 30-50% (SONOPULS HD 200, BANDELIN). The sample was then chilled on ice for 3 min. The sonification- cycle was repeated 3 times. Further, Lysate was centrifuged at 185.000 rpm for 45 min (Sorvall RC 6+, Thermo Scientific). During

the whole process of lysate preparation, 75 µl mixed with 25 µl of 4x loading dye (500 mM Tris/HCl pH 6.8, 8% (v/w) SDS, 40% (v/w) glycerine, 20% (v/w) β-mercaptoethanol, 10 mg/l bromophenol blue) and analysed via SDS-PAGE (15% polyacrylamide gel, running buffer: 25 mM Tris, 192 mM glycine, 0.1% (v/w) glycerine). In order to prepare a 15% polyacrylamide separating gel (20 ml 1.5 M Tris pH 8.8, 40 ml acrylamide, 18.8 ml ddH20, 800 µl 10% SDS, 400 µl 10% APS, 40 µl TEMED), solution was poured in a prepared gel caster (BIO-RAD) and covered with 99.5% isopropanol. After polymerisation for 30 min, the isopropanol was removed and a 5% collecting gel was poured on top (8.8 ml 0.5 M Tris pH 6.8, 5.8 ml acrylamide, 20 ml ddH_20, 350 µl 10% SDS, 135 µl 10% APS, 35 µl TEMED), and a gel comb was added (10 or 12 wells, BIO-RAD). After polymerisation, gels were stored at 4°C.

3.6 Affinity chromatography

Affinity chromatography of proteins was carried out at 4 °C using a peristaltic pump (Pump-P1, Pharmacia Biotech) and pre-cooled buffer, either via Streptavidin- affinity chromatography (Strep-tactin, IBA Life sciences, 1ml prepacked column) or via fusion of the protein of interest to Glutathion-S-transferase (GSTrap, Sigma-Aldrich, 5 ml prepacked column). The flow rate was set according to the information of the manufacturer. Columns were pre-equilibrated with 5x column volume of binding buffer and loaded with filtrated lysate (sterile filter Minisart NML, Sartorius, pore size: 0.45 µm). After loading, column was washed with 5x column volume. Elution was carried out with elution buffer containing either GST or desthiobiotin (20 mM GST or 2.5 mM desthiobiotin), which was added to the corresponding binding buffer. In case GST was added, the pH had to be adjusted again to 7.5. 10-12 fractions of 2 ml were collected manually. In case TEV protease (Sigma-Aldrich) was used to remove the GST-tag, elution fractions were combined and protease was added in a mass ratio of 1:50, protease: protein of interest. Samples were incubated either at 4°C or 25°C for 2 h or 16 h.

In case of construct (4), a Heparin column (1 ml, HiTrap Heparin column, GE Healthcare) was used in order to remove nucleotide contamination. The elution fractions collected by Strep-tag affinity chromatography were pooled, re-buffered via

centrifugation at 4 °C (Amicon, Millipore, 30 kDa; Mega Star 1.6R, VWR) in phosphate buffer (10mM Sodium phosphate, pH7) and loaded on the column via a peristaltic pump (Pump-P1, Pharmacia Biotech). In order to elute the protein of interest from the column several NaCl solutions (0.5, 0.75, 1.0, 1.5, 2.0 M) in phosphate buffer were prepared, and the protein was eluted from the column with an increased gradient of NaCl. Therefore, 2 x 2 ml of each NaCl solution was loaded on the column. 1-10 fractions of 2 ml were collected and analysed via SDS-PAGE. During affinity chromatography, 12 µl of a collected fraction were heated at 95°C for 15 min after addition of 4 µl of a 4x SDS loading dye and analysed via SDS-PAGE (for composition of buffers for SDS-PAGE see 3.5). Protein concentrations were determined by a spectrophotometer (NanoDrop Lite, Thermo Scientific) measured at 280 nm, while the contamination of nucleotides was determined detecting the ratio 260 nm/ 280 nm.

3.7 Size exclusion chromatography

Following affinity chromatography, proteins were separated according to their size via size-exclusion chromatography (SEC) on a HiLoad 26/600 Superdex 200 pg column (GE Healthcare, 317 ml column volume) equilibrated with SEC buffer (20 mM of HEPES-Na, pH 7.5, 20 mM KCl, and 200 mM NaCl) using an automated ÄKTA-FPLC system (GE Healthcare/ Amersham Biosciences). Flow rate was set to 2 ml/min according to the manufacturers protocol. Elution fractions gained from affinity chromatography were pooled, concentrated (Amicon, Millipore, 30 kDa) and typically a volume of 2ml was injected. Size exclusion was carried out at 4°C. Fractions of 4 ml were collected (frac- 900, Amersham Pharmacia Biotech) and analysed via SDS-PAGE (please see 3. Methods, section 3.6). Proteins of interest were pooled and re-buffered and/or concentrated via centrifugation (Mega Star 1.6R, VWR) at 4°C and 3500 rpm using a 50 ml centrifugal filter (Amicon, Millipore, 30 kDa). Protein concentration and nucleotide contamination was determined by a spectrophotometer (NanoDrop Lite, Thermo Fisher Scientific).

3.8 Mass photometry

Mass photometry (MP), originally introduced as interferometric scattering mass spectrometry (iSCAMS), was used to detect the oligomeric state of the proteins of interest (construct (4), (5), (7)). MP accurately measures molecular mass by quantifying light scattering from individual biomolecules in solution, and has the advantage that it is a very rapid, label-free approach requiring a very low amount of protein (50-100 nM), which resembles intracellular concentrations (Sonn-Segev *et al.*, 2020, Young *et al.*, 2018). The technique, which was first described by Young *et al.* (2018), is based on the fact that the amount of scattered light when a molecule touches the surface of a glass slide directly correlates in a linear manner with the molecular weight of the molecule (Young *et al.*, 2018). Data were acquired using a One^{MP} mass photometer (Refeyn, Oxford, UK) using the software AcquireMP (Refeyn, v 2.3.0) for detection. Analysis was then carried out using the software DiscoverMP (Refeyn, v 2.2.0), following the manufacturers protocol.

In order to calculate the molar concentration of protein solutions, the Lambert-Beer-law was used: $c = \frac{A}{\varepsilon \cdot l}$, while A= absorption, ε= extinction coefficient, l= length of the light path. The extinction coefficients of proteins were determined based on their amino acid sequence using the ProtParam tool of the platform ExPASy (the Expert Protein Analysis System, http://www.expasy.org/protparam/)(Gasteiger *et al.*, 2003). After affinity chromatography, samples were re-buffered using a centrifugal filter (Amicon, Millipore, 30 kDa) and the corresponding binding buffer, which was used during affinity chromatography, but containing a different, low salt concentration of 150 mM NaCl (see 3.6). In case a size exclusion was performed, the HEPES buffer (see 3. Methods, section 3.7 Size exclusion chromatography) was used to dilute the sample directly. Glass cover slides (No. 1.5H, 24x60 mm, Marienfeld) were cleaned three times by sonification alternately with ddH_2O and 100% isopropanol, 5 min each, and finally dried with compressed air. Gaskets (GBL103250 Grace Bio-Labs reusable CultureWell™ gaskets wells, 50, diam. × depth 3 mm × 1 mm, well capacity (3-10 µL), Merck) which were cleaned by sonification following the same protocol, were placed on the slides. The measurement was carried out by placing 18 µl of buffer in a gasket, in order to adjust the focus before measurement. Further, 2 µl were added and

the drop was mixed thoroughly. It was found that a final concentration of 50 nM protein was sufficient for the measurement. Movies were recorded with a frame rate of 10 kHrz. For each protein, measurement was at least repeated 3 times.

3.9 Electrophoretic mobility shift assay (EMSA) & Assays of nuclease activity

In order to analyse DNA-binding activity of the GST-tagged truncations of ComEC (construct (4), (6), (7)) an EMSA was performed according to the protocol of Ream *et al.*, where protein interactions with small RNA molecules were analysed by EMSA (Ream *et al.*, 2016). Custom-made oligonucleotides with a size of 50 bp (please see 5. Appendix, Table 5.2) were annealed before, creating double-strand DNA with 3´-overhang or 5´ overhang dsDNA and blunt-end dsDNA. In case ssDNA was tested, a single oligo was used as a template. Therefore, oligonucleotides were mixed in an equal molar ratio and heated at 95 °C for 5 min followed by slow cooling at room temperature for 3 h. In order to detect DNA- binding activity, 0.1 µg of protein were incubated with 1 µM of Primer in a total reaction volume of 10 µl, which was adjusted by addition of different buffers. A CHES buffer was used (50 mM CHES, 50 mM KCl, 5mM MgAc, 5% Glycerol, pH 7), optimized for the detection of exonuclease activity (Hernández-Tamayo *et al.*, 2019), which was modified by the addition of different divalent cations (1mM $MnCl_2$, 1mM $ZnCl_2$). The different CHES-buffers were used with and without addition of 1 mM DTT, which was added shortly before use. Samples were incubated for 1 h at 25 °C. In case nuclease activity was investigated, samples were incubated for 10, 20, 40 and 60 min at 37°C. Reactions were stopped by heating for 5 min at 95 °C in order to inactivate the protein and loaded on an agarose gel (2.5%), after addition of 2 µl 30% Glycerol. A negative control was carried out for every type of primer by incubating the reaction without addition of protein, and a control was carried out incubating only the protein with buffer. According to the method, an Agarose-gel electrophoresis was carried out in a pre-warmed 0.5 x TB – buffer system (45 mM Tris, 45 mM boric acid) at 25 °C, applying a voltage of 25 mV/cm, with a total duration time of 5-10 min. Gels were stained with 10 µl midori

green (NIPPON genetics)/ 50 ml 0.5 TB- buffer for 30 min and de-stained for 15 min in 0.5 x TB- buffer.

3.10 Single molecule tracking of fluorescently labeled DNA

Single-molecule tracking was carried out exactly as described in 2.2. Manuscript II (see 2.2.4, method "Single-molecule tracking (SMT)", page 87), except for the exposure time used to detect the fluorescently labeled DNA. DNA was added to competent *B. subtilis* PY79 (PG001) cells and *B. subtilis* PY79 *Δrok* (PG876). In order to detect the signal, an exposure time of 50 ms was used. Further analysis of diffusion coefficients was not carried out due to low number of tracks (number of tracks ≤ 500).

4. Discussion

4.1 Analysing the function of the *B. subtilis* late competence gene, *comEB*

The gene of the late competence protein ComEB is located between two highly important proteins involved in DNA- uptake during transformation, ComEA and ComEC (1. Introduction, Figure 4, page 22). Three genes form the *comE* operon of *B. subtilis. ComEA* encodes a DNA-binding transmembrane protein, *comEB* a putative dCMP-deaminase and *comEC* an aqueous channel protein with putative exonuclease function. Studies have been performed in order to elucidate the function of *comEB*, and it was found that a n-terminal truncation of 44% of the gene did not remarkably affect transformability (Inamine & Dubnau, 1995). An enzyme activity was, so far, only confirmed for ComEB of *Bacillus halodurans*, but the authors did not investigate its enzymatic function in the context of natural transformation of *Bacillus* (Oehlenschlaeger et al., 2015). In the study presented here (2.1 Manuscript I, page 23), it was investigated whether the enzyme activity of ComEB was relevant for natural transformation of *B. subtilis.* Therefore, the gene was expressed ectopically and inactivated by introducing a point mutation at position 98. A cysteine, important for the localization of a zinc ion at the active site, was exchanged for an alanine. Inactiviation of enzymatic activity was confirmed by *in vitro* assays of ComEB of *Geobacillus thermophilus* ($ComEB_{Geo}$, see 2.1 Manuscript I, 2.1.4 Results, Figure 4, page 37). Assays of transformation efficiency were carried out (see 2.1 Manuscript I, 2.1.4 Results, Figure 2, page 33), where it was found that inhibiting ComEB´s activity had no effect on transformation efficiency, indicating a different role of the gene.

However, ComEB localizes to the cell pole in competent *B. subtilis* cells when expressed as a c-terminal fluorophore fusion, which has been demonstrated by independent studies (Hahn *et al.*, 2005, Kaufenstein *et al.*, 2011) and the epifluorescence and SMT-studies presented in manuscript I. The fact that *comEB* encodes a dCMP deaminase suggests a role in connecting nucleotide metabolism with DNA-uptake during natural transformation. It might lead to a direct degradation of the incoming DNA, producing dUMP as a precursor of dTTP, which is needed synthesis

of DNA. Another hypothesis would be that its product, dUMP serves as a second messenger, maybe supporting other required cellular functions during DNA uptake or homologous integration, which would explain its polar localization. So far, several bacterial nucleotides - based second messengers, like cAMP, c-di-GMP and (p)ppGpp(guanosine-3´,5´-(pentaphosphate)bisdiphosphate) are well known. (p)ppGpp binds to RNA-Polymerase leading to a repressed transcription rate of RNA in order to switch from vegetative growth into survival mode (Pesavento & Hengge, 2009). Cells do sense c-di-GMP levels as a signal of phosphate starvation, and in *B. subtilis* a low level leads to the inhibition of biofilm formation (Kalia *et al.*, 2013). In case DNA is taken up by the cell, a high amount of dUMP, eventually generated by ComEB, could act as a second messenger. It could then be required to induce production of specific proteins for competence or homologous recombination. One would expect that the addition of chromosomal DNA to competent *B. subtilis* cells would affect ComEB´s localization, but the protein continuously localized to the cell pole (2.1 Manuscript I, 2.1.4 Results, Figure 6-8, pages 41, 45 and 46). If the product of comEB was a second messenger, the protein would indeed stay at the cell pole in order to produce its signal.

Some other circumstances might have affected the results presented herein. Maybe a different experimental set-up would be required to detect a change of its diffusion. On the other hand, dUMP could be simply used to synthesize RNA, promoting transcription of essential genes during competence at the cell pole, e.g. the downstream gene *comEC*. Therefore, its role would fit to a supporting, but rather unessential function in the context of natural transformation.

However, the enzymatic activity of ComEB seems not to be essential for transformation, but obviously the presence of the gene strongly increases transformation efficiency (see Manuscript I, 2.1.4 Results, Figure 2, page 33). It has been proposed earlier that the gene might encode a ribosomal binding site (RBS), which could be required for transcription of the downstream gene *comEC*, which is essential for transformation (Hahn *et al.*, 1993, Inamine & Dubnau, 1995). Recent studies of Dubnau & Hahn show that comEB and ComEC are translationally coupled. Apparently, *comEB* encodes a putative suboptimal RBS within its last 21 bp (ACGAGCT), which allows translation of the essential downstream gene *comEC* (De

Santis *et al.*, 2021*)*. Transformation efficiency was only affected in case the described motif was truncated from the gene, while a mutant encoding a stop codon did not influence transformation, meaning regulation appears on a translational level. The authors speculate that the enzymatic activity of the late competence proteins ComEB and ComFC (a putative type I phosphoribosyl transferase) is required to reuse nucleotides which are released during natural transformation by degradation of the homologous DNA strand. Now, the fact that a *comEB* deletion strain is deficient in transformation, which cannot be complemented by ectopic expression of *comEC*, appears in a different light, and actually supports the idea of De Santis *et al.* (2.1 Manuscript I, 2.1.4 Results, Figure 2, page 33).

Despite the apparent, unessential enzyme function of comEB during natural transformation, the question remains why the gene is (out of all competence operons) particularly located in the *comE* operon upstream of *comEC*. De Santis *et al.* answer this question by suggesting a benefit in regulating *comEC* stoichiometrically independent from *comEA*. The DNA- binding receptor ComEA might be required in a much higher amount to facilitate DNA-uptake than *comEC.* On the other hand, an RBS located in the beginning of comEC gene would have also led to an independent translation. It appears that *Lactobacillus planterium*, *Bacillus anthracis* and *Streptococcus pneumoniae* have a similar *comE*- operon structure, indicating a benefit of the uncoupled translation (De Santis *et al.*, 2021). Competence genes were probably acquired by HGT e.g. transduction or conjugation, being passed on and on, thereby evolving continuously (Chiang *et al.*, 2019). A change or shift of the nucleotide sequence of the open reading frame might have occurred and caused the described translational coupling, turning out to be beneficial for the organism. It might be a matter of evolutionary biology to answer the question why *comEB* evolved to be a part of the competence machinery, and was finally located at the *comE*.

4.2 ComGA - the putative assembly ATPase of the *B. subtilis* competence apparatus

ATPases are involved in every competence-associated machinery as the driving force of assembly and disassembly of the pilus or pseudopilus structure. It appears that either one or two ATPases are required in order to assemble and disassemble a pilus structure. In case of *B. subtilis*, two ATPases are encoded within the late competence operons, *comGA* and *comFA*, while *comFA* rather belongs to the family of ATP-binding helicases. It is probably involved in facilitating uptake of the ssDNA, which enters the cytosol through the aqueous membrane channel ComEC (Dubnau & Blokesch, 2019). For *S. pneumoniae*, it has been demonstrated that ComFA exhibits ATPase-dependent ssDNA-binding activity, as well as an interaction with ssDNA-binding protein DprA, which resembles the idea of a helicase (Diallo *et al.*, 2017). ComGA is orthologous to other hexameric assembly and disassembly ATPases of bacterial DNA-uptake machineries and the only ATPase left among the competence proteins of *B. subtilis*, which could provide the energy for pilus assembly (Piepenbrink, 2019). Mutations of ComGA´s ATP-binding/hydrolysing sites inhibit transformation. DNA is only bound by *B. subtilis* cells in case ATP is bound by the protein and the major pseudopilin ComGC only assembles in the periplasm of ComGA mutants, which allow intercellular uptake of DNA (Briley *et al.*, 2011a). Therefore, ComGA of *B. subtilis* is thought to be the assembly/disassembly ATPase of the system.

We analysed its function via single-molecule tracking, fusing the protein to mVenus at its C-terminus (2.1 Manuscript I, 2.1.4 Results, Figure 6, page 41). Chromosomal DNA was added to competent *B. subtilis* cells in order to induce a change of single-molecule dynamics. While we did not detect a significant change of the diffusion coefficients, the mobile population of the protein increased about 12% in case DNA was added. In addition, bleaching curves of the fluorescent signal at the cell pole were generated. In case DNA was added to the competent *B. subtilis* cells, we detected a lower average fluorescent signal at the cell pole, and a rapid decrease of the fluorescent signal at the cell pole (measured in RFU, see 2.1 Manuscript I, 2.1.4 Results section, Figure 7, page 45). Data were fitted to a one phase exponential decay function, resulting in an increased reaction rate of 72%. Both measurements indicate that

ComGA molecules (or subunits of ComGA) are exchanged at the cell pole, probably induced by a high enzyme activity, which is intensified by addition of DNA. Our studies support the current theory that ComGA acts as the assembly ATPase of the putative pseudopilus of *B. subtilis*.

As a suggestion for further studies and in order to analyse whether the enzymatic activity of ComGA had caused the described fluctuation at cell pole, the Single-molecule tracking experiment could be repeated using an inactivated, mutated version of *comGA*. If the detected fluctuations were dependent on enzymatic activity, these would be inhibited by the mutation. The protein could be inactivated via a point mutation (Briley *et al.*, 2011a). Furthermore, if ComGA really carries out the assembly and disassembly of the putative pseudopilus of *Bacillus*, its cellular role could be analysed by deleting or mutating the gene followed by an additional staining of the major pseudopilin. In other bacteria, these types of staining are well known, and such a staining is currently about to be established for *Bacillus subtilis* (Kilb, Alexandra, AG Graumann, unpublished data). In case the experiments will work out for *Bacillus*, the assembly and the disassembly could be analysed by visualizing the pilus in a genetic background where *comGA* had been mutated. According to the current state of knowledge, assembly should only occur in case the Walker A motif of ComGA, responsible for ATP and DNA-binding of competent cells, was intact (Briley *et al.*, 2011a), and a mutation would abolish formation of pili. In addition, if ComGA was responsible for disassembly in addition, a mutation of the Walker B motif, required for hydrolyzation of ATP, would lead to detectable pili, which were not able to disassemble anymore. Possibly, fluorescent signal could enhance, or an increased number of pseudopili would be detected as they were not able to retract anymore. An additional, interesting role to investigate would be ComGA´s ability to serve as a checkpoint during transformation, and its putative interaction with RelA, being involved in stress-response via (p)ppGpp. The interaction of both proteins seems to induce stringent response by preventing the hydrolyzation of (p)ppGpp. In case (p)ppGpp is not degraded by RelA due to its interaction with ComGA, the level of the alarmone would increase, thereby inhibiting DNA replication and rRNA synthesis. (Hahn *et al.*, 2015). According to this theory, it would be interesting to analyse the

diffusive behaviour of ComGA and RelA by single-molecule tracking, thereby investigating *B. subtilis* cells under different or changing stress/alarmone levels.

4.3 ComEC - the putative exonuclease of the *B. subtilis* competence complex

One of the basic questions concerning DNA-uptake of the *B. subtilis* competence machinery still remains unanswered. Until now, it is not known how the complement strand of the dsDNA, which enters the cell, is degraded. In case of *S. pneumoniae* the exonuclease EndA hydrolyses the complement strand, but at the current state of knowledge, such a nuclease is not known for *B. subtilis* (Puyet *et al.*, 1990). Recent studies support the idea that the channel protein ComEC is not only responsible for internalization of DNA, meaning the transport across the cytosolic membrane, but also for degradation of the incoming dsDNA, by an exonuclease function of its C-terminus (Baker *et al.*, 2016).

We tried to answer this question by heterologous gene expression followed by protein purification of different truncations of ComEC, which were fused N-terminally to a GST- tag (please see 2.3 Unpublished results, 2.3.1 Expression and purification of *B. subtilis* ComEC periplasmic domains, Figure 5, page 134). The strategy was chosen because a toxic effect had been reported, in case the full-length protein was expressed in *E. coli* (Draskovic & Dubnau, 2005). The presented data (please see 2.3 Unpublished results, 2.3.2 localization of stained DNA in competent *B. subtilis* cells, page 141) support the presence of soluble, periplasmic domain of the protein at the C-terminus, as described by Draskovic & Dubnau (2005) and Baker *et al*. (2016). Unfortunately, the purified C-terminus of the protein did not show any activity. Measurements carried out with purified truncation, using a mass photometer, revealed a strong tendency of the proteins to oligomerize (see 2.3 Unpublished results, figures 6-8, pages 136-139). This could have hampered the exonuclease assays which were carried out under the herein applied conditions (please see 3. Methods, 3.9 Electrophoretic mobility shift assay (EMSA) & Assays of nuclease activity, page 149). It cannot be assumed that oligomerization was caused by the truncated part of ComEC,

as the GST-tag, which was still fused to the overexpressed truncations, is known to form dimers and higher multimers (Tudyka & Skerra, 1997, Vinckier *et al.*, 2011). Therefore, oligomerization was probably independent of any fusion; even though the multimerization of ComEC would make sense. But it is still not known how many ComEC proteins interact in order to form a transmembrane pore. Attempts to remove the GST-tag led to precipitation of the protein, hindering further experiments concerning its DNA-binding oder exonuclease activity.

The only evidence for any enzyme activity could be the continuously examined DNA or RNA contamination of the purified proteins (2.3 Unpublished results, 2.3.1 Expression and purification of B. subtilis ComEC periplasmic domains, Figure 5, page 134). In addition, contamination with nucleic acids was, to a very high extend, detected in case the N-terminal, soluble part of the protein was purified (see (Krämer, 2019)). These observations are in agreement with data of an *in vitro* modelling of ComEC, indicating that the DNA-binding site of ComEC is located at the N-terminus (OB-fold), while the C-terminus would contain the β-lactamase domain, carrying out the putative exonuclease function (Baker *et al.*, 2016). It has been reported that both domains of ComEC interact when co-expressed in *E. coli* (Dubnau & Blokesch, 2019), and therefore it might be very likely, that both domains need to interact in order to generate a detectable enzymatic activity *in vitro*. Addressing this theory, a chimeric protein was created, cloning both parts (N-loop & C-loop) into an expression vector positioning a poly-Glycine- Serine linker in between. But unfortunately, the protein was not expressed (Krämer, 2019).

Switching the genetic background by choosing the *comEC* gene of *Geobacillus thermophilus* as a template for cloning and protein expression did not result in any differences concerning the solubility or nucleic-acids contamination of the truncated parts (Krämer, 2019). The usage of a different tag, e.g. MBP (Maltose binding protein), or of different *E. coli* strains for protein expression could increase solubility and could prevent oligomerization. In addition, it might be still beneficial to follow up the creation of a chimeric protein, thereby removing the insoluble competence domain, as it contains up to 9 putative transmembrane helices, which render the protein highly hydrophobic (Baker *et al.*, 2016).

Recent studies of ComEC of *H. pylori* show that the OB-fold domain is interacting with ComH, an orthologue of *B. subtilis* ComEA. The OB-fold was heterologously expressed and purified from an *E. coli* culture as a His-tag fusion. Interaction was then detected by a pulldown assay with MBP-ComH (Damke *et al.*, 2019a). In our case, purification of the OB-fold of *B. subtilis* via a C-terminal Strep-tag did not result in a stable protein (Krämer, 2019). However, considering the data of Damke *et al.*, the function of the *B. subtilis* comEC could be similar. The N-terminal domain of ComEC might be required during transformation to bind dsDNA in the periplasm, which is handed over by ComEA through a direct interaction of the two proteins.

4.4 The *B. subtilis* competence complex might be a mobile system, assembling by a diffusion/capture mechanism

In Manuscript II, we analysed the location and furthermore calculated the diffusion coefficients of ComEC-mVenus (PG3817) by singe-molecule tracking. The fusion was expressed under the control of the original promotor. We found a localization pattern of confined, single tracks at the membrane of competent *B. subtilis* cells within a radius of 127.3 nm (see 2.2 Manuscript II, 2.2.5 Results, Figure 5a), page 99). Data were fitted to two populations with diffusion coefficients of 0.017 $\mu m^2/s$ and 0.31 $\mu m^2/s$, resembling a static and a more mobile fraction. Diffusion coefficients of both populations are in the range expected for a membrane protein (Lucena *et al.*, 2018). In addition, the signal detected by single-molecule tracking was analysed by an overlay of the single frames, and fluorescence of ComEC-mV was always located at the membrane (in case cells were analysed mid-cell). A similar pattern was found for two different C-terminal truncations of ComEC, ComECΔ128-mV and ComECΔ301-mV (see 2.2 Manuscript II, 2.2.5 Results, Figure 1c), Figure 5a), and supplementary material, Figure S2, pages 90, 99 and 123). Interestingly, truncation ComECΔ301-mV, where the whole C-loop region of the protein was removed, located at the cell membrane in a very distinct manner even in epifluorescence (2.2 Manuscript II, 2.2.5 Results, Figure 1c), page 90). We did not detect any changes of the diffusion, if competent cells expressing ComEC-mV were incubated with chromosomal DNA (Krämer, 2019). Surprisingly, a large fraction of the protein of 87.1% was mobile

(analysed at the upper-cell or membrane level), while the mobility decreased in the truncations from 62.5% for ComECΔ128-mV to even 49.8% for ComECΔ301-mV. In parallel, transformation efficiency decreased remarkably when the C-terminus of the protein was truncated (see 2.2 Manuscript II, 2.2.5 Results, Figure 3, page 93). These data could indicate that the diffusive behaviour of ComEC is required for transformation. Furthermore, it was found, that a conserved amino acid, located within a conserved zinc-binding motif of the C-terminus, was required for transformation of *B. subtilis*. If the mutation of aa D573 inhibited transformation because the whole protein was destabilized, or if the described phenotype occurred because enzyme activity was abolished, cannot be stated. Stability of the protein could be further investigated by western-blotting of the corresponding strains using an anti-ComEC antibody. Enzyme activity could be verified by *in-vitro* assays. In case these would be successful (see 2.3 Unpublished results, 2.3.1 Expression and purification of ComEC periplasmic domains, page 133), the relevant, identified mutation D573A should be investigated, too.

Diffusion of mV-ComEA was analysed at mid-cell level, and again, the confined tracks of the protein localized to the membrane, similar to ComEC-mV and the fluorescent dsDNA, which was stained with DyLight488 and then added to competent *B. subtilis* cells (see Manuscript II, Figure 7 a), b), page 103). mV-ComEA and the labeled dsDNA/PCR-product showed similar diffusive behaviour and diffusion coefficients. The data were fitted to three populations with almost identical diffusion coefficients. It has been reported for other bacteria, like *Neisseria*, that ComE, the orthologue of ComEA, “slides” along the membrane, searching for dsDNA to bind during natural transformation (Chen & Gotschlich, 2001, Gangel *et al.*, 2014). Therefore, the authors hypothesized that ComEA might act as a reservoir for incoming DNA during transformation, until it can be taken up by the cell. Similar mobility has been reported for ComEA of *V. cholerae*, which accumulates at the membrane depending on the concentration of incoming DNA (Seitz & Blokesch, 2014). Therefore, the data on the diffusion of ComEA could be interpreted the same way, especially as mV-ComEA is becoming more static/immobile in case chromosomal DNA is added to competent *B. subtilis* cells. In addition, a more septal and polar localization was detected (see Manuscript II, Figure 7 c), d), page 103). The detected

mobility of ComEA was quite surprising, as ComEA is an integral membrane protein, while ComE of *Neisseria* and ComEA of *V. cholerae* are periplasmic DNA- receptors.

According to the observations of the single-molecule tracking in the thesis and the manuscripts presented here (meaning the analysed diffusion of ComGA-mV, ComEA-mV, ComEB-mV, ComEC-mV, and labeled fluorescent DNA), it is quite likely that in case of *B. subtilis*, ComEA and ComEC, which are absolutely essential for natural transformation, might assemble at the membrane through a diffusion/capture mechanism. In addition, this theory would explain the appearance of several polar and sometimes punctual signals of fluorescence fusions of competence proteins observed in epifluorescence (Hahn *et al.*, 2005, Kaufenstein *et al.*, 2011). The competence complex, or at least the proteins ComEC and ComEA, would assemble at the membrane, depending only on the concentration of taken-up DNA within the periplasm, and independent e.g. of the presence of the putative pseudopilus. A model of the described process is depicted in Figure 10.

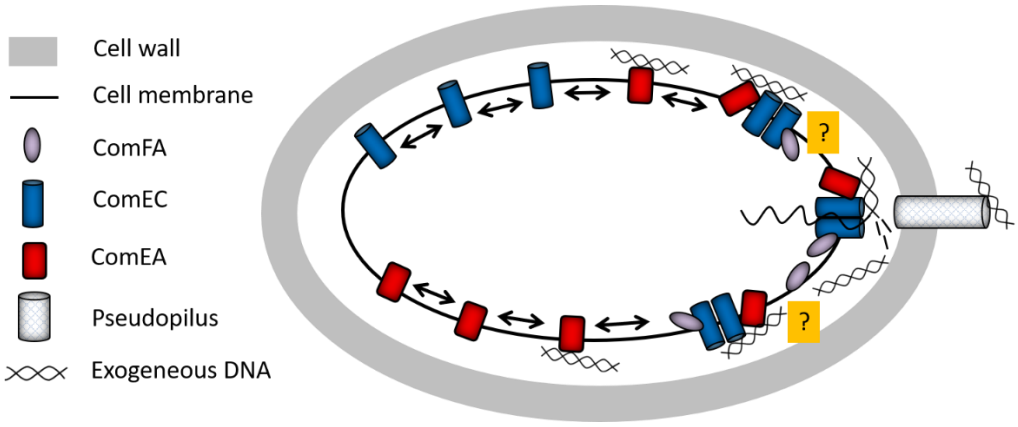


Figure 10: Model of a competent B. subtilis cell; competence proteins comEA and ComEC diffuse along the membrane

A competent *B. subtilis* cell internalizes exogeneous DNA via its putative pseudopilus. The proteins ComEA and ComEC diffuse along the membrane until they assemble at the cell pole (indicated by black arrows, see legend on the left for signs and colour code) together with the helicase ComFA. Possibly, the competence complex assembles also in close proximity of the pole (indicated by yellow boxes with question marks), in case excess DNA is taken up by the cell. It is still not clear how many complexes or ComEC proteins assemble during DNA-uptake, but it is known that ComEC forms a homodimer (Draskovic & Dubnau, 2005). ComEA acts as a reservoir for incoming DNA and further interacts with ComEC in order to hand over the transforming DNA. ComFA might provide the energy for the process, thereby interacting with ComEC.

5. Appendix

5.1 Calibration– mass photometry

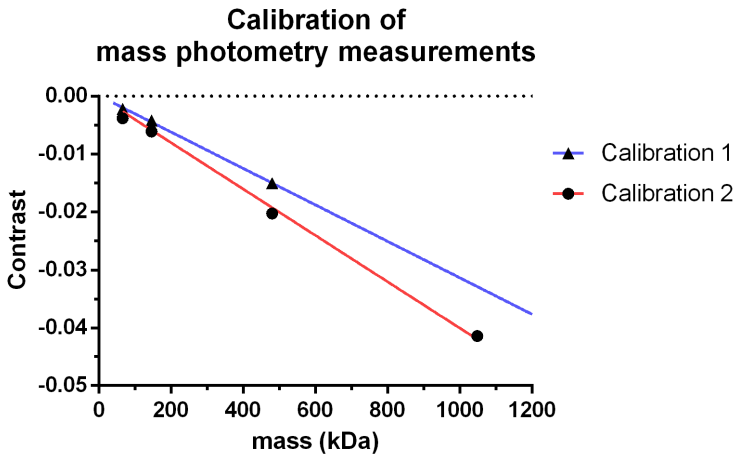


Figure 11: Calibration of mass photometry measurements

Calibration for mass photometry was carried out using standard solutions of proteins (NativeMark™ Unstained Protein Standard, Invitrogen) at a concentration of 50 nM: 66 kDa, 146 kDa, 480 kDa, 1048 kDa. Linear correlation of the measured contrast and the mass of the standard proteins prior to the measurement of construct (4) is shown by Calibration 1 (blue linear, Figure 6). Calibration 2 was carried out prior to the measurements of construct (5) and construct (7) (Figure 7, Figure 8).

Table 1: List of Primer

name of constructs, plasmid x *gene*	primer fw, sequence: 5´→ 3´	primer rev, sequence: 5´→ 3´
(1) pGAT2 x *comEC, bp 1480- 2303 **	GTGATA*ACATGTTG*TGTCTTAGTTCCG AAGGAGAAG	*tatcacCTCGAG*TCACTTCTCGAACTGGGG GTGGCTCCAGGCGCTGACAGAAAAGGTT CCAACTCTG
(2) pGAT2 x *comEC, bp 1519- 2157 **	*gcgata**ACATGTTG***GGACAGGGTGACA GCATGTT	*gcgctcCTCGAG*TCACTTCTCGAACTGGGG GTGGCTCCAGGCGCTTTTCCCGGCTGAG ATAATGG
(3) pGAT2 x *comEC, bp 1483- 1941 **	*gtgata**CCATGGGC***CTTAGTTCCGAAGG AGAAGT	*tatcacCTCGAG*TCACTTCTCGAACTGGGG GTGGCTCCAGGCGCTTTGCTTGCCGGAT C
(4) pGAT2 x *comEC, bp 1519- 2016 **	*gcgata**ACATGTTG***GGACAGGGTGACA GCATGTT	*gcgctcCTCGAG*TCACTTCTCGAACTGGGG GTGGCTCCAGGCGCTTTCTTTCTCCAGAT CACCCG
(5) pGAT2 x *comEC, bp 1483- 2098 **	*gtgata**ACATGTTG***CTTAGTTCCGAAGG AGAAGTGG	*tatcacCTCGAG**TTA***AGAGCCTTTGCTCCC ATG
(6) pGAT2 x *comEC, bp 1483- 2109 **	*gtgata**ACATGTTG***CTTAGTTCCGAAGG AGAAGTGG	*tatcacCTCGAG**TTA***TTCTTCACCGGTAGA GCCTTTG
(7) pGAT2 x *comEC, bp 1483- 2139 **	*gtgata**ACATGTTG***CTTAGTTCCGAAGG AGAAGTGG	*tatcacCTCGAG**TTA***GGCCGTTTTCGGCTG AAG
pGAT2, sequencing primer	TTTGGTGGTGGCGACC	TTGCTCAGCGGTGGCA
EMSA primer (ssDNA)	GAAGTGAATGAGTCCGGGGAAACGGT GTATTTGCGCTTATCCACGCTGCC	-
EMSA primer 2 (dsDNA, blunt)	GAAGTGAATGAGTCCGGGGAAACGGT GTATTTGCGCTTATCCACGCTGCC	GGCAGCGTGGATAAGCGCAAATACACCG TTTCCCCGGACTCATTCACTTC

EMSA primer 3 (dsDNA, *3´overhang*) *	GAAGTGAATGAGTCCGGGGAAACGGT GTATTTGCG*cttatccacgctgcc*	CGCAAATACACCGTTTCCCCGGACTCATT CACTTC
EMSA primer 4 (dsDNA, *5´overhang*) *	*gaagtgaatgagtcc*GGGGAAACGGTGTA TTTGCGCTTATCCACGCTGCC	GGCAGCGTGGATAAGCGCAAATACACCG TTTCCCC

* overhang, NcoI, PciI, XhoI, **Stop,** Spacer (ATG in frame), Strep-tag II

Table 2: List of strains

strain	**genotype**	**plasmid**	**Resistance/ inducing agent**	**reference**
PG001 ***B. subtilis*** **PY79**	PY79	no plasmid	no resistance	wild type strain
PG876 ***B. subtilis*** **PY79**	*Δrok*	resistance cassette, original locus	Kan	gift from D. Dubnau, name in publication: BD3196
PG3803 ***E. coli*** **DH5α**	*F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupGpurB20φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK–mK+), λ–*	no plasmid	no resistance	Thermo Fisher Scientific
PG3882 ***E. coli*** **DH5α**	PG3803 transformed with pGAT2, expression vector for N-terminal GST-tag fusions of proteins	pGAT2-modified: *his*-tag removed	Amp, IPTG	gift from AG Bange, original backbone: (Peränen *et al.*, 1996)
PG3890 ***E. coli*** **DH5α**	*construct (1): gst-comec-strep, aa: 494-763*	pGAT2-modified: *his*-tag removed	Amp, IPTG	this study
PG3889 ***E. coli*** **DH5α**	*construct (2) gst- comec-strep,*	pGAT2-modified:	Amp, IPTG	this study

	aa: 507-719	*his*-tag removed		
PG3883 ***E. coli* DH5α**	*construct (3) gst-comec-strep, aa: 495-647*	pGAT2-modified: *his*-tag removed	Amp, IPTG	this study
PG3884 ***E. coli* DH5α**	*construct (4) gst- comec-strep, aa: 507-672*	pGAT2-modified: *his*-tag removed	Amp, IPTG	this study
PG3887 ***E. coli* DH5α**	*construct (5) gst- comec, aa: 495-699*	pGAT2-modified: *his*-tag removed	Amp, IPTG	this study
PG3886 ***E. coli* DH5α**	*construct (6) gst- comec, aa: 495-704*	pGAT2-modified: *his*-tag removed	Amp, IPTG	this study
PG3886 ***E. coli* DH5α**	*construct (7) gst-comec, aa: 495-713*	pGAT2-modified: *his*-tag removed	Amp, IPTG	this study
PG3804 ***E. coli* BL21 (DE3)**	*B F– ompT gal dcm lon hsdSB(rB–mB–) λ(DE3 [lacI lacUV5-T7p07 ind1 sam7 nin5]) [malB+]K-12(λS)*	no plasmid	No resistance	Thermo Fisher Scientific

Table 3: Polymerase chain reaction

Substance	**Volume (concentration)**
Phusion HF buffer (5 x)	20 µl (1 x)
dNTP-mix	4 µl (10 mM)
Primer fw	4 µl (10 µM)
Primer rev	4 µl (10 µM)
$MgCl_2$	2 µl (50 mM)
Template (plasmid DNA)	1 µl (10 ng/ml)
Phusion Polymerase	1 µl (2,000 U/ml)
ddH_20	ad 64 µl
volume (total):	**100 µl**

Initial denaturation	5:00 min, 98 °C	
denaturation	0:30 min, 98°C	35 cycles
annealing	1:00 min, 60 °C *	
elongation	0:30 min/ 500 bp, 72 °C	
Final elongation	7:00 min, 72 °C	

* standard annealing temperature, which was only adjusted if necessary (+/- 5 °C). For calculation of annealing temperatures and design Primer, the online tool from New England Biolabs (NEB) was used: https://tmcalculator.neb.com.

Table 4: List of chemicals

agar- agar	Carl Roth
agarose	Carl Roth
agarose ultrapure	Thermo Fisher Scientific
ammonium iron(III) citrate	Sigma-Aldrich
Ampicillin	Carl Roth
APS	Carl Roth
boric acid	Chem-Solutions CmbH
Bromophenol blue	Carl Roth
$CaCl_2$	Merck KGaA
casein hydrolysate	Sigma-Aldrich
CHES	Sigma-Aldrich
Coomassie brilliant blue R-250	Thermo Fisher Scientific
desthiobiotin	Sigma-Aldrich
DNA-Ladder, 1 kb	New England Biolabs (NEB)
DNAse I	PanReac AppliChem
dNTPs	New England Biolabs (NEB)
DTT	PanReac AppliChem
DyLight™ 488 NHS Ester	Thermo Fisher Scientific
glucose	Carl Roth
glutathione	Carl Roth
glycerol	Carl Roth
glycine	Carl Roth
HCl	VWR Chemicals
HEPES-Na	Caarl Roth
IPTG	Carl Roth
Isopropanol, 99.5%	Thermo Fisher Scientific
K_2HPO_4	Carl Roth
Kanamycin	PanReac AppliChem
KCl	Chem-Solutions GmbH
KH_2PO_4	Carl Roth
Lysozyme	PanReac AppliChem

MgAc	Sigma-Aldrich
$MgCl_2$	Carl Roth
Midori green	NIPPON genetics
$MnCl_2$	Merck KGaA
NaCl	Thermo Fisher Scientific
NaOH	Carl Roth
NativeMark™ Unstained Protein Standard	Invitrogen
NcoI	New England Biolabs (NEB)
PageRuler Prestained Protein Ladder	Thermo Fisher Scientific
PciI	New England Biolabs (NEB)
Phusion High Fidelity DNA Polymerase	New England Biolabs (NEB)
polyacrylamide	Carl Roth
potassium glutamate	Carl Roth
Protease-Inhibitor cocktail (EDTA free)	F. Hoffmanns-La Roche AG
RNAse	PanReac AppliChem
SDS	PanReac AppliChem
T4 DNA Ligase	New England Biolabs (NEB)
TEMED	PanReac AppliChem
TEV-Protease	Sigma-Aldrich
Tris	Carl Roth
trisodium citrate	Honeywell Specialty Chemicals Seelze GmbH
trypton	PanReac AppliChem
XhoI	New England Biolabs (NEB)
yeast extract	Carl Roth
$ZnCl_2$	Carl Roth

5.2 Literature

Albano, M., Smits, W., Ho, L.T., Kraigher, B., Mandic-Mulec, I., Kuipers, O.P., and Dubnau, D. (2005) The Rok protein of Bacillus subtilis represses genes for cell surface and extracellular functions. *Journal of bacteriology* **187**.

Alexander, H.E., and Redman, W. (1953) Transformation of type specificity of meningococci; change in heritable type induced by type-specific extracts containing desoxyribonucleic acid. *J Exp Med* **97**: 797-806.

Alonso, J.C., Lüder, G., and Trautner, T.A. (1986) Requirements for the formation of plasmid-transducing particles of Bacillus subtilis bacteriophage SPP1. *The EMBO journal* **5**.

Anagnostopoulos, C., and Spizizen, J. (1961) REQUIREMENTS FOR TRANSFORMATION IN BACILLUS SUBTILIS1. *J Bacteriol* **81**: 741-746.

Avery, O.T., Macleod, C.M., and McCarty, M. (1944) STUDIES ON THE CHEMICAL NATURE OF THE SUBSTANCE INDUCING TRANSFORMATION OF PNEUMOCOCCAL TYPES : INDUCTION OF TRANSFORMATION BY A DESOXYRIBONUCLEIC ACID FRACTION ISOLATED FROM PNEUMOCOCCUS TYPE III. *J Exp Med* **79**: 137-158.

Ayora, S., Carrasco, B., Cardenas, P.P., Cesar, C.E., Canas, C., Yadav, T., Marchisone, C., and Alonso, J.C. (2011) Double-strand break repair in bacteria: a view from Bacillus subtilis. *FEMS Microbiol Rev* **35**: 1055-1081.

Baker, J.A., Simkovic, F., Taylor, H.M., and Rigden, D.J. (2016) Potential DNA binding and nuclease functions of ComEC domains characterized in silico. *Proteins* **84**: 1431-1442.

Bell, J.C., and Kowalczykowski, S.C. (2016) RecA: Regulation and Mechanism of a Molecular Search Engine. *Trends in biochemical sciences* **41**.

Berge, M., Mortier-Barriere, I., Martin, B., and Claverys, J.P. (2003) Transformation of Streptococcus pneumoniae relies on DprA- and RecA-dependent protection of incoming DNA single strands. *Mol Microbiol* **50**: 527-536.

Bergé, M., Moscoso, M., Prudhomme, M., Martin, B., and Claverys, J.P. (2002) Uptake of transforming DNA in Gram-positive bacteria: a view from Streptococcus pneumoniae. *Molecular microbiology* **45**.

Berka, R.M., Hahn, J., Albano, M., Draskovic, I., Persuh, M., Cui, X., Sloma, A., Widner, W., and Dubnau, D. (2002) Microarray analysis of the Bacillus subtilis K-state: genome-wide expression changes dependent on ComK. *Mol Microbiol* **43**: 1331-1345.

Boonstra, M., Schaffer, M., Sousa, J., Morawska, L., Holsappel, S., Hildebrandt, P., Sappa, P.K., Rath, H., de Jong, A., Lalk, M., Mäder, U., Völker, U., and Kuipers, O.P. (2020) Analyses of competent and non-competent subpopulations of Bacillus subtilis

reveal yhfW, yhxC and ncRNAs as novel players in competence. *Environmental microbiology* **22**.

Boonstra, M., Vesel, N., and Kuipers, O.P. (2018) Fluorescently Labeled DNA Interacts with Competence and Recombination Proteins and Is Integrated and Expressed Following Natural Transformation of Bacillus subtilis. *mBio* **9**.

Briley, K., Dorsey-Oresto, A., Prepiak, P., Dias, M.J., Mann, J.M., and Dubnau, D. (2011a) The secretion ATPase ComGA is required for the binding and transport of transforming DNA. *Mol Microbiol* **81**: 818-830.

Briley, K., Jr., Prepiak, P., Dias, M.J., Hahn, J., and Dubnau, D. (2011b) Maf acts downstream of ComGA to arrest cell division in competent cells of B. subtilis. *Mol Microbiol* **81**: 23-39.

Brito, P.H., Chevreux, B., Serra, C.R., Schyns, G., Henriques, A.O., and Pereira-Leal, J.B. (2018a) Genetic Competence Drives Genome Diversity in Bacillus subtilis. *Genome biology and evolution* **10**.

Brito, P.H., Instituto Gulbenkian de Ciência, O., Portugal, Nova Medical School, F.d.C.M., Universidade Nova de Lisboa, Portugal, Chevreux, B., DSM Nutritional Products, L., 60 Westview street, Lexington MA, USA, Serra, C.R., Instituto de Tecnologia Química e Biológica, O., Portugal, Schyns, G., DSM Nutritional Products, L., 60 Westview street, Lexington MA, USA, Henriques, A.O., Instituto de Tecnologia Química e Biológica, O., Portugal, Pereira-Leal, J.B., Instituto Gulbenkian de Ciência, O., Portugal, and Ophiomics—Precision Medicine, L., Portugal (2018b) Genetic Competence Drives Genome Diversity in Bacillus subtilis. *Genome Biology and Evolution* **10**: 108-124.

Burton, B., and Dubnau, D., (2010) Membrane-associated DNA Transport Machines. In: Cold Spring Harb Perspect Biol. pp.

Carrasco, B., Ayora, S., Lurz, R., and Alonso, J.C. (2005) Bacillus subtilis RecU Holliday-junction resolvase modulates RecA activities. *Nucleic acids research* **33**.

Cehovin, A., Simpson, P.J., McDowell, M.A., Brown, D.R., Noschese, R., Pallett, M., Brady, J., Baldwin, G.S., Lea, S.M., Matthews, S.J., and Pelicic, V. (2013) Specific DNA recognition mediated by a type IV pilin. *Proceedings of the National Academy of Sciences of the United States of America* **110**.

Chassy, B. (1987) Transformation of Lactobacillus casei by electroporation - Chassy - 1987 - FEMS Microbiology Letters - Wiley Online Library. *FEMS Microbiology Letters* **44**: 173-177.

Chen, I., and Gotschlich, E.C., (2001) ComE, a Competence Protein from Neisseria gonorrhoeae with DNA-Binding Activity. In: J Bacteriol. pp. 3160-3168.

Chen, I., Provvedi, R., and Dubnau, D. (2006) A Macromolecular Complex Formed by a Pilin-like Protein in Competent Bacillus subtilis*S. *J Biol Chem* **281**.

Chiang, Y.N., Penadés, J.R., and Chen, J., (2019) Genetic transduction by phages and chromosomal islands: The new and noncanonical. In: PLoS Pathog. pp.

Chilton, S.S., Falbel, T.G., Hromada, S., and Burton, B.M., (2017) A Conserved Metal Binding Motif in the Bacillus subtilis Competence Protein ComFA Enhances Transformation. In: J Bacteriol. pp.

Chung, Y.S., Breidt, F., and Dubnau, D. (1998) Cell surface localization and processing of the ComG proteins, required for DNA binding during transformation of Bacillus subtilis. *Mol Microbiol* **29**: 905-913.

Chung, Y.S., and Dubnau, D. (1995) ComC is required for the processing and translocation of comGC, a pilin-like competence protein of Bacillus subtilis. *Molecular microbiology* **15**.

Chung, Y.S., and Dubnau, D. (1998) All seven comG open reading frames are required for DNA binding during transformation of competent Bacillus subtilis. *J Bacteriol* **180**: 41-45.

Corbinais, C., Mathieu, A., Kortulewski, T., Radicella, J.P., and Marsin, S. (2016) Following transforming DNA in Helicobacter pylori from uptake to expression. *Molecular microbiology* **101**.

Craig, L., Taylor, R.K., Pique, M.E., Adair, B.D., Arvai, A.S., Singh, M., Lloyd, S.J., Shin, D.S., Getzoff, E.D., Yeager, M., Forest, K.T., and Tainer, J.A. (2003) Type IV pilin structure and assembly: X-ray and EM analyses of Vibrio cholerae toxin-coregulated pilus and Pseudomonas aeruginosa PAK pilin. *Molecular cell* **11**.

Damke, P.P., Di Guilmi, A.M., Varela, P.F., Velours, C., Marsin, S., Veaute, X., Machouri, M., Gunjal, G.V., Rao, D.N., Charbonnier, J.B., and Radicella, J.P., (2019a) Identification of the periplasmic DNA receptor for natural transformation of Helicobacter pylori. In: Nat Commun. pp.

Damke, P.P., Guilmi, A.M.D., Varela, P.F., Velours, C., Marsin, S., Veaute, X., Machouri, M., Gunjal, G.V., Rao, D.N., Charbonnier, J.-B., and Radicella, J.P. (2019b) Identification of the periplasmic DNA receptor for natural transformation of Helicobacter pylori. *Nature Communications* **10**: 1-11.

Danner, D.B., Deich, R.A., Sisco, K.L., and Smith, H.O. (1980) An eleven-base-pair sequence determines the specificity of DNA uptake in Haemophilus transformation. *Gene* **11**.

Davidoff-Abelson, R., and Dubnau, D. (1973) Kinetic analysis of the products of donor deoxyribonucleate in transformed cells of Bacillus subtilis. *Journal of bacteriology* **116**.

Davis, K.M., and Isberg, R.R. (2016) Defining heterogeneity within bacterial populations via single cell approaches. *BioEssays : news and reviews in molecular, cellular and developmental biology* **38**.

De Santis, M., Hahn, J., and Dubnau, D. (2021) ComEB protein is dispensable for the transformation but must be translated for the optimal synthesis of comEC. *Molecular microbiology*.

Deichelbohrer, I., Alonso, J.C., Lüder, G., and Trautner, T.A. (1985) Plasmid transduction by Bacillus subtilis bacteriophage SPP1: effects of DNA homology between plasmid and bacteriophage. *Journal of bacteriology* **162**.

Diallo, A., Foster, H.R., Gromek, K.A., Perry, T.N., Dujeancourt, A., Krasteva, P.V., Gubellini, F., Falbel, T.G., Burton, B.M., and Fronzes, R. (2017) Bacterial transformation: ComFA is a DNA-dependent ATPase that forms complexes with ComFC and DprA. *Molecular microbiology* **105**.

Domenech, A., Slager, J., and Veening, J.W. (2018) Antibiotic-Induced Cell Chaining Triggers Pneumococcal Competence by Reshaping Quorum Sensing to Autocrine-Like Signaling. *Cell reports* **25**.

Dorer, M.S., Fero, J., and Salama, N.R. (2010) DNA damage triggers genetic exchange in Helicobacter pylori. *PLoS Pathog* **6**: e1001026.

Draskovic, I., and Dubnau, D. (2005) Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulphide bonds. *Mol Microbiol* **55**: 881-896.

Dubnau, D. (1991) Genetic competence in Bacillus subtilis. *Microbiol Rev* **55**: 395-424.

Dubnau, D., and Blokesch, M. (2019) Mechanisms of DNA Uptake by Naturally Competent Bacteria. *https://doi.org/10.1146/annurev-genet-112618-043641*.

Dubnau, D., and Cirigliano, C. (1972a) Fate of transforming deoxyribonucleic acid after uptake by competent Bacillus subtilis: size and distribution of the integrated donor segments. *Journal of bacteriology* **111**.

Dubnau, D., and Cirigliano, C. (1972b) Fate of transforming DNA following uptake by competent Bacillus subtilis. Formation and properties of products isolated from transformed cells which are derived entirely from donor DNA. *Journal of molecular biology* **64**.

Durand, E., Bernadac, A., Ball, G., Lazdunski, A., Sturgis, J.N., and Filloux, A. (2003) Type II protein secretion in Pseudomonas aeruginosa: the pseudopilus is a multifibrillar and adhesive structure. *Journal of bacteriology* **185**.

Ellison, C.K., Dalia, T.N., Ceballos, A.V., Wang, J.C.-Y., Biais, N., Brun, Y.V., and Dalia, A.B. (2018) Retraction of DNA-bound type IV competence pili initiates DNA uptake during natural transformation in Vibrio cholerae. *Nature Microbiology* **3**: 773.

Fagerlund, A., Granum, P.E., and Håvarstein, L.S. (2014) Staphylococcus aureus competence genes: mapping of the SigH, ComK1 and ComK2 regulons by transcriptome sequencing. *Molecular microbiology* **94**.

Gamba, P., Jonker, M.J., and Hamoen, L.W. (2015) A Novel Feedback Loop That Controls Bimodal Expression of Genetic Competence. *PLoS genetics* **11**.

Gangel, H., Hepp, C., Müller, S., Oldewurtel, E.R., Aas, F.E., Koomey, M., and Maier, B., (2014) Concerted Spatio-Temporal Dynamics of Imported DNA and ComE DNA Uptake Protein during Gonococcal Transformation. In: PLoS Pathog. pp.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. (2003) ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic acids research* **31**.

Giltner, C.L., Nguyen, Y., and Burrows, L.L., (2012) Type IV Pilin Proteins: Versatile Molecular Modules. In: Microbiol Mol Biol Rev. pp. 740-772.

González, J.M. (2020) Visualizing the superfamily of metallo-β-lactamases through sequence similarity network neighborhood connectivity analysis.

Green, E.R., and Mecsas, J. (2016) Bacterial Secretion Systems: An Overview. *Microbiology spectrum* **4**.

Griffith, F. (1928) The Significance of Pneumococcal Types. *J Hyg (Lond)* **27**: 113-159.

Gromkova, R., and Goodgal, S. (1979) Transformation by plasmid and chromosomal DNAs in Haemophilus parainfluenzae. *Biochemical and biophysical research communications* **88**.

Hahn, J., Inamine, G., Y, K., and Dubnau, D. (1993) Characterization of comE, a late competence operon of Bacillus subtilis required for the binding and uptake of transforming DNA. *Molecular microbiology* **10**.

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

Hahn, J., Tanner, A.W., Carabetta, V.J., Cristea, I.M., and Dubnau, D. (2015) ComGA-RelA interaction and persistence in the Bacillus subtilis K-state. *Mol Microbiol* **97**: 454-471.

Haijema, B.J., Hahn, J., Haynes, J., and Dubnau, D. (2001) A ComGA-dependent checkpoint limits growth during the escape from competence. *Mol Microbiol* **40**: 52-64.

Hamilton, H.L., and Dillard, J.P. (2006) Natural transformation of Neisseria gonorrhoeae: from DNA donation to homologous recombination. *Molecular microbiology* **59**.

Hamoen, L.W., Van Werkhoven, A.F., Bijlsma, J.J., Dubnau, D., and Venema, G. (1998) The competence transcription factor of Bacillus subtilis recognizes short A/T-rich sequences arranged in a unique, flexible pattern along the DNA helix. *Genes & development* **12**.

Hanahan, D. (1983) Studies on transformation of Escherichia coli with plasmids. *J Mol Biol* **166**: 557-580.

Hanson, P.I., and Whiteheart, S.W. (2005) AAA+ proteins: have engine, will work. *Nature reviews. Molecular cell biology* **6**.

Hernández-Tamayo, R., Oviedo-Bocanegra, L.M., Fritz, G., and Graumann, P.L. (2019) Symmetric activity of DNA polymerases at and recruitment of exonuclease ExoR and of PolA to the Bacillus subtilis replication forks. *Nucleic acids research* **47**.

Herriott, R.M., Meyer, E.M., and Vogt, M. (1970) Defined nongrowth media for stage II development of competence in Haemophilus influenzae. *Journal of bacteriology* **101**.

Hoa, T., Tortosa, P., Albano, M., and Dubnau, D. (2002) Rok (YkuW) regulates genetic competence in Bacillus subtilis by directly repressing comK. *Molecular microbiology* **43**.

Hobbs, M.M., Seiler, A., Achtman, M., and Cannon, J.G. (1994) Microevolution within a clonal population of pathogenic bacteria: recombination, gene duplication and horizontal genetic exchange in the opa gene family of Neisseria meningitidis. *Molecular microbiology* **12**.

Holmes, R.K., and Jobling, M.G., (1996) *Genetics*. University of Texas Medical Branch at Galveston.

Hospenthal, M.K., Costa, T.R.D., and Waksman, G. (2017) A comprehensive guide to pilus biogenesis in Gram-negative bacteria. *Nature reviews. Microbiology* **15**.

Hou, H.F., Liang, Y.H., Li, L.F., Su, X.D., and Dong, Y.H. (2008) Crystal structures of Streptococcus mutans 2'-deoxycytidylate deaminase and its complex with substrate analog and allosteric regulator dCTP x Mg2+. *J Mol Biol* **377**: 220-231.

Inamine, G.S., and Dubnau, D. (1995) ComEA, a Bacillus subtilis integral membrane protein required for genetic transformation, is needed for both DNA binding and transport. *J Bacteriol* **177**: 3045-3051.

Israel, D.A., Lou, A.S., and Blaser, M.J. (2000) Characteristics of Helicobacter pylori natural transformation. *FEMS Microbiol Lett* **186**: 275-280.

Johansson, E., Neuhard, J., Willemoes, M., and Larsen, S. (2004) Structural, kinetic, and mutational studies of the zinc ion environment in tetrameric cytidine deaminase. *Biochemistry* **43**: 6020-6029.

Johnston, C., Martin, B., Fichant, G., Polard, P., and Claverys, J.-P. (2014) Bacterial transformation: distribution, shared mechanisms and divergent control. *Nature Reviews Microbiology* **12**: 181-196.

Kalia, D., Merey, G., Nakayama, S., Zheng, Y., Zhou, J., Luo, Y., Guo, M., Roembke, B.T., and Sintim, H.O. (2013) Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP, (p)ppGpp signaling in bacteria and implications in pathogenesis. *Chemical Society reviews* **42**.

Kantor, R.S., Wrighton, K.C., Handley, K.M., Sharon, I., Hug, L.A., Castelle, C.J., Thomas, B.C., and Banfield, J.F. (2013) Small genomes and sparse metabolisms of sediment-associated bacteria from four candidate phyla. *mBio* **4**.

Kaufenstein, M., van der Laan, M., and Graumann, P.L. (2011) The three-layered DNA uptake machinery at the cell pole in competent Bacillus subtilis cells is a stable complex. *J Bacteriol* **193**: 1633-1642.

Keefe, R.G., Maley, G.F., Saxl, R.L., and Maley, F. (2000) A T4-phage deoxycytidylate deaminase mutant that no longer requires deoxycytidine 5'-triphosphate for activation. *J Biol Chem* **275**: 12598-12602.

Khasanov, F.K., Zvingila, D.J., Zainullin, A.A., Prozorov, A.A., and Bashkirov, V.I. (1992) Homologous recombination between plasmid and chromosomal DNA in Bacillus subtilis requires approximately 70 bp of homology. *Molecular & general genetics : MGG* **234**.

Kidane, D., Ayora, S., Sweasy, J.B., Graumann, P.L., and Alonso, J.C. (2012) The cell pole: the site of cross talk between the DNA uptake and genetic recombination machinery. *Crit Rev Biochem Mol Biol* **47**: 531-555.

Kidane, D., Carrasco, B., Manfredi, C., Rothmaier, K., Ayora, S., Tadesse, S., Alonso, J.C., and Graumann, P.L. (2009) Evidence for different pathways during horizontal gene transfer in competent Bacillus subtilis cells. *PLoS genetics* **5**.

Kidane, D., and Graumann, P.L. (2005) Intracellular protein and DNA dynamics in competent Bacillus subtilis cells. *Cell* **122**: 73-84.

Kirn, T.J., Jude, B.A., and Taylor, R.K. (2005) A colonization factor links Vibrio cholerae environmental survival and human infection. *Nature* **438**.

Kjos, M., Miller, E., Slager, J., Lake, F.B., Gericke, O., Roberts, I.S., Rozen, D.E., and Veening, J.W. (2016) Expression of Streptococcus pneumoniae Bacteriocins Is Induced by Antibiotics via Regulatory Interplay with the Competence System. *PLoS pathogens* **12**.

Korotkov, K.V., and Sandkvist, M. (2019) Architecture, Function, and Substrates of the Type II Secretion System. *EcoSal Plus* **8**.

Kovács, A.T., Smits, W.K., Mirończuk, A.M., and Kuipers, O.P. (2009) Ubiquitous late competence genes in Bacillus species indicate the presence of functional DNA uptake machineries. *Environmental microbiology* **11**.

Kramer, N., Hahn, J., and Dubnau, D. (2007) Multiple interactions among the competence proteins of Bacillus subtilis. *Mol Microbiol* **65**: 454-464.

Krämer, K., (2019) In-vitro und in-vivo Analyse des Kompetenz-Proteins ComEC von *B. subtilis*. Master thesis, AG Graumann, FB 15, Philipps-University Marburg, pp. 1-81.

Lacks, S., Greenberg, B., and Neuberger, M. (1974) Role of a Deoxyribonuclease in the Genetic Transformation of Diplococcus pneumoniae.

Laurenceau, R., Péhau-Arnaudet, G., Baconnais, S., Gault, J., Malosse, C., Dujeancourt, A., Campo, N., Chamot-Rooke, J., Le Cam, E., Claverys, J.P., and Fronzes,

R. (2013) A type IV pilus mediates DNA binding during natural transformation in Streptococcus pneumoniae. *PLoS pathogens* **9**.

Lazazzera, B.A., Solomon, J.M., and Grossman, A.D. (1997) An exported peptide functions intracellularly to contribute to cell density signaling in B. subtilis. *Cell* **89**.

Li, Y., Guo, Z., Jin, L., Wang, D., Gao, Z., Su, X., Hou, H., Dong, Y., and IUCr (2016) Mechanism of the allosteric regulation of Streptococcus mutans 2'-deoxycytidylate deaminase. *Acta Crystallographica Section D: Structural Biology* **72**: 883-891.

Liu, J., and Zuber, P. (1998) A molecular switch controlling competence and motility: competence regulatory factors ComS, MecA, and ComK control sigmaD-dependent gene expression in Bacillus subtilis. *Journal of bacteriology* **180**.

Londoño-Vallejo, J., and Dubnau, D. (1993) comF, a Bacillus subtilis late competence locus, encodes a protein similar to ATP-dependent RNA/DNA helicases. *Molecular microbiology* **9**.

Lucena, D., Mauri, M., Schmidt, F., Eckhardt, B., and Graumann, P.L. (2018) Microdomain formation is a general property of bacterial membrane proteins and induces heterogeneity of diffusion patterns. *BMC Biol* **16**: 97.

Maamar, H., and Dubnau, D. (2005) Bistability in the Bacillus subtilis K-state (competence) system requires a positive feedback loop. *Mol Microbiol* **56**.

Magnuson, R., Solomon, J., and Grossman, A.D. (1994) Biochemical and genetic characterization of a competence pheromone from B. subtilis. *Cell* **77**.

Maier, B., Chen, I., Dubnau, D., and Sheetz, M.P. (2004) DNA transport into Bacillus subtilis requires proton motive force to generate large molecular forces. *Nature structural & molecular biology* **11**.

Mandel, M., and Higa, A. (1970) Calcium-dependent bacteriophage DNA infection. 1970. *Biotechnology* **24**: 198-201.

Mann, J.M., Carabetta, V.J., Cristea, I.M., and Dubnau, D. (2013) Complex formation and processing of the minor transformation pilins of Bacillus subtilis. *Mol Microbiol* **90**: 1201-1215.

Maslowska, K.H., Makiela-Dzbenska, K., and Fijalkowska, I.J. (2019) The SOS system: A complex and tightly regulated response to DNA damage. *Environmental and molecular mutagenesis* **60**.

Mattick, J.S. (2002) Type IV pili and twitching motility. *Annual review of microbiology* **56**.

McCallum, M., Tammam, S., Khan, A., Burrows, L.L., and Howell, P.L. (2017) The molecular mechanism of the type IVa pilus motors. *Nature communications* **8**.

Meibom, K.L., Blokesch, M., Dolganov, N.A., Wu, C.Y., and Schoolnik, G.K. (2005) Chitin induces natural competence in Vibrio cholerae. *Science (New York, N.Y.)* **310**.

Mell, J.C., and Redfield, R.J., (2014) Natural Competence and the Evolution of DNA Uptake Specificity. In: J Bacteriol. pp. 1471-1483.

Mirouze, N., Ferret, C., Cornilleau, C., and Carballido-López, R. (2018) Antibiotic sensitivity reveals that wall teichoic acids mediate DNA binding during competence in Bacillus subtilis. *Nature communications* **9**.

Misic, A.M., Satyshur, K.A., and Forest, K.T. (2010) P. aeruginosa PilT structures with and without nucleotide reveal a dynamic type IV pilus retraction motor. *Journal of molecular biology* **400**.

Molina-Santiago, C., Pearson, J.R., Navarro, Y., Berlanga-Clavero, M.V., Caraballo-Rodriguez, A.M., Petras, D., García-Martín, M.L., Lamon, G., Haberstein, B., Cazorla, F.M., de Vicente, A., Loquet, A., Dorrestein, P.C., and Romero, D. (2019) The extracellular matrix protects Bacillus subtilis colonies from Pseudomonas invasion and modulates plant co-colonization. *Nature communications* **10**.

Mulder, J.A., and Venema, G. (1982) Transformation-deficient mutants of Bacillus subtilis impaired in competence-specific nuclease activities. *Journal of bacteriology* **152**.

Munch-Petersen, A., and Mygind, B. (1976) Nucleoside transport systems in Escherichia coli K12: specificity and regulation. *J Cell Physiol* **89**: 551-559.

Muschiol, S., Balaban, M., Normark, S., and Henriques-Normark, B. (2015) Uptake of extracellular DNA: Competence induced pili in natural transformation of Streptococcus pneumoniae. *Bioessays* **37**: 426-435.

Muschiol, S., Erlendsson, S., Aschtgen, M.S., Oliveira, V., Schmieder, P., de Lichtenberg, C., Teilum, K., Boesen, T., Akbey, U., and Henriques-Normark, B. (2017) Structure of the competence pilus major pilin ComGC in Streptococcus pneumoniae. *The Journal of biological chemistry* **292**.

Nicholson, W.L., Munakata, N., Horneck, G., Melosh, H.J., and Setlow, P. (2000) Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. *Microbiology and molecular biology reviews : MMBR* **64**.

O'Toole, G.A., and Kolter, R. (1998) Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development. *Molecular microbiology* **30**.

Oehlenschlaeger, C.B., Lovgreen, M.N., Reinauer, E., Lehtinen, E., Pind, M.L., Harris, P., Martinussen, J., and Willemoes, M. (2015) Bacillus halodurans Strain C125 Encodes and Synthesizes Enzymes from Both Known Pathways To Form dUMP Directly from Cytosine Deoxyribonucleotides. *Appl Environ Microbiol* **81**: 3395-3404.

Ogura, M., Liu, L., Lacelle, M., Nakano, M.M., and Zuber, P. (1999) Mutational analysis of ComS: evidence for the interaction of ComS and MecA in the regulation of competence development in Bacillus subtilis. *Molecular microbiology* **32**.

Parsonnet, J., Friedman, G.D., Vandersteen, D.P., Chang, Y., Vogelman, J.H., Orentreich, N., and Sibley, R.K. (1991) Helicobacter pylori infection and the risk of gastric carcinoma. *The New England journal of medicine* **325**.

Peränen, J., Rikkonen, M., Hyvönen, M., and Kääriäinen, L. (1996) T7 vectors with modified T7lac promoter for expression of proteins in Escherichia coli. *Analytical biochemistry* **236**.

Pesavento, C., and Hengge, R. (2009) Bacterial nucleotide-based second messengers. *Current opinion in microbiology* **12**.

Piepenbrink, K.H. (2019) DNA Uptake by Type IV Filaments. *Front Mol Biosci* **6**.

Pimentel, Z.T., and Zhang, Y. (2018) Evolution of the Natural Transformation Protein, ComEC, in Bacteria. *Front Microbiol* **9**: 2980.

Powell, I.B., Achen, M.G., Hillier, A.J., and Davidson, B.E. (1988) A Simple and Rapid Method for Genetic Transformation of Lactic Streptococci by Electroporation. *Appl Environ Microbiol* **54**: 655-660.

Prattes, M., Lo, Y.H., Bergler, H., and Stanley, R.E. (2019) Shaping the Nascent Ribosome: AAA-ATPases in Eukaryotic Ribosome Biogenesis. *Biomolecules* **9**.

Provvedi, R., Chen, I., and Dubnau, D. (2001) NucA is required for DNA cleavage during transformation of Bacillus subtilis. *Mol Microbiol* **40**: 634-644.

Provvedi, R., and Dubnau, D. (1999) ComEA is a DNA receptor for transformation of competent Bacillus subtilis. *Molecular microbiology* **31**.

Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B., and Claverys, J.P. (2006) Antibiotic stress induces genetic transformability in the human pathogen Streptococcus pneumoniae. *Science (New York, N.Y.)* **313**.

Pugsley, A.P. (1993) The complete general secretory pathway in gram-negative bacteria.

Puyet, A., Greenberg, B., and Lacks, S.A. (1990) Genetic and structural characterization of endA. A membrane-bound nuclease required for transformation of Streptococcus pneumoniae. *Journal of molecular biology* **213**.

Quinn, C.P., and Dancer, B.N. (1990) Transformation of vegetative cells of Bacillus anthracis with plasmid DNA. *J Gen Microbiol* **136**: 1211-1215.

Ragunathan, K., Liu, C., and Ha, T. (2012) RecA filament sliding on DNA facilitates homology search. *eLife* **1**.

Rami Almog, Frank Maley, Gladys F. Maley, Robert MacColl, a., and Roey*, P.V. (2004) Three-Dimensional Structure of the R115E Mutant of T4-Bacteriophage 2'-Deoxycytidylate Deaminase†,‡.

Ream, J.A., Lewis, L.K., and Lewis, K.A. (2016) Rapid agarose gel electrophoretic mobility shift assay for quantitating protein: RNA interactions. *Analytical biochemistry* **511**.

Rossier, O., Dao, J., and Cianciotto, N.P. (2009) A type II secreted RNase of Legionella pneumophila facilitates optimal intracellular infection of Hartmannella vermiformis. *Microbiology (Reading, England)* **155**.

Rudel, T., Scheurerpflug, I., and Meyer, T.F. (1995) Neisseria PilC protein identified as type-4 pilus tip-located adhesin. *Nature* **373**.

Satyshur, K.A., Worzalla, G.A., Meyer, L.S., Heiniger, E.K., Aukema, K.G., Misic, A.M., and Forest, K.T. (2007) Crystal structures of the pilus retraction motor PilT suggest large domain movements and subunit cooperation drive motility. *Structure (London, England : 1993)* **15**.

Sauvonnet, N., Vignon, G., Pugsley, A.P., and Gounon, P. (2000) Pilus formation and protein secretion by the same machinery in Escherichia coli. *The EMBO journal* **19**.

Seitz, P., and Blokesch, M. (2013) Cues and regulatory pathways involved in natural competence and transformation in pathogenic and environmental Gram-negative bacteria. *FEMS Microbiol Rev* **37**: 336-363.

Seitz, P., and Blokesch, M., (2014) DNA Transport across the Outer and Inner Membranes of Naturally Transformable Vibrio cholerae Is Spatially but Not Temporally Coupled. In: mBio. pp.

Seitz, P., Pezeshgi Modarres, H., Borgeaud, S., Bulushev, R.D., Steinbock, L.J., Radenovic, A., Dal Peraro, M., and Blokesch, M., (2014) ComEA Is Essential for the Transfer of External DNA into the Periplasm in Naturally Transformable Vibrio cholerae Cells. In: PLoS Genet. pp.

Serror, P., and Sonenshein, A.L. (1996) CodY is required for nutritional repression of Bacillus subtilis genetic competence. *Journal of bacteriology* **178**.

Shivers, R.P., and Sonenshein, A.L. (2004) Activation of the Bacillus subtilis global regulator CodY by direct interaction with branched-chain amino acids. *Molecular microbiology* **53**.

Smits, W., and Grossman, A. (2010) The transcriptional regulator Rok binds A+T-rich DNA and is involved in repression of a mobile genetic element in Bacillus subtilis. *PLoS genetics* **6**.

Solanki, V., Kapoor, S., and Thakur, K.G. (2018) Structural insights into the mechanism of Type IVa pilus extension and retraction ATPase motors. *The FEBS journal* **285**.

Solomon, J.M., Magnuson, R., Srivastava, A., and Grossman, A.D. (1995) Convergent sensing pathways mediate response to two extracellular competence factors in Bacillus subtilis. *Genes & development* **9**.

Song, B.H., and Neuhard, J. (1989) Chromosomal location, cloning and nucleotide sequence of the Bacillus subtilis cdd gene encoding cytidine/deoxycytidine deaminase. *Mol Gen Genet* **216**: 462-468.

Sonn-Segev, A., Belacic, K., Bodrug, T., Young, G., VanderLinden, R.T., Schulman, B.A., Schimpf, J., Friedrich, T., Dip, P.V., Schwartz, T.U., Bauer, B., Peters, J.M., Struwe,

W.B., Benesch, J.L.P., Brown, N.G., Haselbach, D., and Kukura, P. (2020) Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nature communications* **11**.

Sparling, P.F. (1966) Genetic transformation of Neisseria gonorrhoeae to streptomycin resistance. *J Bacteriol* **92**: 1364-1371.

Spizizen, J. (1958) TRANSFORMATION OF BIOCHEMICALLY DEFICIENT STRAINS OF BACILLUS SUBTILIS BY DEOXYRIBONUCLEATE. *Proc Natl Acad Sci U S A* **44**: 1072-1078.

Strom, M.S., and Lory, S. (1991) Amino acid substitutions in pilin of Pseudomonas aeruginosa. Effect on leader peptide cleavage, amino-terminal methylation, and pilus assembly. *The Journal of biological chemistry* **266**.

Suerbaum, S., and Achtman, M. (1999) Evolution of Helicobacter pylori: the role of recombination. *Trends Microbiol* **7**: 182-184.

Suerbaum, S., and Michetti, P. (2009) Helicobacter pylori Infection. *http://dx.doi.org/10.1056/NEJMra020542*.

Susanna, K.A., van der Werff, A.F., den Hengst, C.D., Calles, B., Salas, M., Venema, G., Hamoen, L.W., and Kuipers, O.P. (2004) Mechanism of transcription activation at the comG promoter by the competence transcription factor ComK of Bacillus subtilis. *Journal of bacteriology* **186**.

Tadesse, S., and Graumann, P.L. (2007) DprA/Smf protein localizes at the DNA uptake machinery in competent Bacillus subtilis cells. *BMC Microbiol* **7**: 105.

Takeno, M., Taguchi, H., and Akamatsu, T. (2012) Role of ComEA in DNA uptake during transformation of competent Bacillus subtilis. *Journal of bioscience and bioengineering* **113**.

Tudyka, T., and Skerra, A. (1997) Glutathione S-transferase can be used as a C-terminal, enzymatically active dimerization module for a recombinant protease inhibitor, and functionally secreted into the periplasm of Escherichia coli. *Protein science : a publication of the Protein Society* **6**.

Turgay, K., Hahn, J., Burghoorn, J., and Dubnau, D. (1998) Competence in Bacillus subtilis is controlled by regulated proteolysis of a transcription factor. *EMBO J* **17**: 6730-6738.

van Sinderen, D., Luttinger, A., Kong, L., Dubnau, D., Venema, G., and Hamoen, L. (1995) comK encodes the competence transcription factor, the key regulatory protein for competence development in Bacillus subtilis. *Mol Microbiol* **15**: 455-462.

Veening, J.W., and Blokesch, M. (2017) Interbacterial predation as a strategy for DNA acquisition in naturally competent bacteria. *Nature reviews. Microbiology* **15**.

Vinckier, N.K., Chworos, A., and Parsons, S.M. (2011) Improved isolation of proteins tagged with glutathione S-transferase. *Protein expression and purification* **75**.

Vosman, B., Kuiken, G., Kooistra, J., and Venema, G. (1988) Transformation in Bacillus subtilis: involvement of the 17-kilodalton DNA-entry nuclease and the competence-specific 18-kilodalton protein. *Journal of bacteriology* **170**.

Wang, C., Yuan, Y., and Hunt, R.H. (2007) The association between Helicobacter pylori infection and early gastric cancer: a meta-analysis. *The American journal of gastroenterology* **102**.

Wei, H., and Håvarstein, L.S. (2012) Fratricide is essential for efficient gene transfer between pneumococci in biofilms. *Applied and environmental microbiology* **78**.

Wholey, W.Y., Kochan, T.J., Storck, D.N., and Dawid, S. (2016) Coordinated Bacteriocin Expression and Competence in Streptococcus pneumoniae Contributes to Genetic Adaptation through Neighbor Predation. *PLoS pathogens* **12**.

Wilton, M., Halverson, T.W.R., Charron-Mazenod, L., Parkins, M.D., and Lewenza, S. (2018) Secreted Phosphatase and Deoxyribonuclease Are Required by Pseudomonas aeruginosa To Defend against Neutrophil Extracellular Traps. *Infection and immunity* **86**.

Winterling, K.W., Levine, A.S., Yasbin, R.E., and Woodgate, R. (1997) Characterization of DinR, the Bacillus subtilis SOS repressor. *Journal of bacteriology* **179**.

Wolfgang, M., van Putten, J.P., Hayes, S.F., and Koomey, M. (1999) The comP locus of Neisseria gonorrhoeae encodes a type IV prepilin that is dispensable for pilus biogenesis but essential for natural transformation. *Molecular microbiology* **31**.

Yadav, T., Carrasco, B., Hejna, J., Suzuki, Y., Takeyasu, K., and Alonso, J.C. (2013) Bacillus subtilis DprA recruits RecA onto single-stranded DNA and mediates annealing of complementary strands coated by SsbB and SsbA. *The Journal of biological chemistry* **288**.

Yadav, T., Carrasco, B., Myers, A.R., George, N.P., Keck, J.L., and Alonso, J.C. (2012) Genetic recombination in Bacillus subtilis: a division of labor between two single-strand DNA-binding proteins. *Nucleic Acids Res* **40**: 5546-5559.

Yadav, T., Carrasco, B., Serrano, E., and Alonso, J.C. (2014) Roles of Bacillus subtilis DprA and SsbA in RecA-mediated genetic recombination. *J Biol Chem* **289**: 27640-27652.

Yedidi, R.S., Wendler, P., and Enenkel, C. (2017) AAA-ATPases in Protein Degradation. *Frontiers in molecular biosciences* **4**.

Young, G., Hundt, N., Cole, D., Fineberg, A., Andrecka, J., Tyler, A., Olerinyova, A., Ansari, A., Marklund, E.G., Collier, M.P., Chandler, S.A., Tkachenko, O., Allen, J., Crispin, M., Billington, N., Takagi, Y., Sellers, J.R., Eichmann, C., Selenko, P., Frey, L., Riek, R., Galpin, M.R., Struwe, W.B., Benesch, J.L.P., and Kukura, P. (2018) Quantitative mass imaging of single biological macromolecules. *Science (New York, N.Y.)* **360**.

Zhang, Y., Maley, F., Maley, G.F., Duncan, G., Dunigan, D.D., and Van Etten, J.L. (2007) Chloroviruses encode a bifunctional dCMP-dCTP deaminase that produces two key intermediates in dTTP formation. *J Virol* **81**: 7662-7671.

6. *Curriculum vitae*

Removed for elecronical upload.

7. Inlay - Supplementary movies of manuscript I & manuscript II