STRUCTURAL AND FUNCTIONAL STUDIES ON THE TRANSCRIPTIONAL REGULATION OF FLAGELLAR MOTILITY AND BIOFILM FORMATION

DISSERTATION

ZUR ERLANGUNG DES AKADEMISCHEN GRADES

DES DOKTORS DER NATURWISSENSCHAFTEN

(DR. RER. NAT.)

DEM FACHBEREICH CHEMIE DER PHILIPPS-UNIVERSITÄT MARBURG VORGELEGT

(HOCHSCHULKENNZIFFER 1180)

VON

DEVID MRUSEK

MASTER OF SCIENCE

AUS FRANKFURT AM MAIN

MARBURG (LAHN), DEZEMBER 2018

Die Untersuchungen zur vorliegenden Arbeit wurden von Oktober 2014 bis Dezember 2018 am LOEWE Zentrum für synthetische Mikrobiologie (SYNMIKRO) der Philipps-Universität Marburg unter der Leitung von Prof. Dr. Gert Bange durchgeführt.

Vom Fachbereich Chemie
der Philipps-Universität Marburg (Hochschulkennziffer 1180)
als Dissertation am 21.12.2018 angenommen
Erstgutachter: Prof. Dr. Gert Bange
(Fachbereich Chemie, Philipps-Universität Marburg)
Zweitgutachter: Prof. Dr. Peter Graumann
(Fachbereich Chemie, Philipps-Universität Marburg)
Weitere Mitglieder der Prüfungskommission:
Prof. Dr. Erhard Bremer
(Fachbereich Biologie, Philipps-Universität Marburg)
Jun.-Prof. Dr. Olalla Vázquez
(Fachbereich Chemie, Philipps-Universität Marburg)

Tag der Disputation: 27.03.2019

Declaration of authorship:

I hereby declare that this submission is entirely of my own and to the best of my knowledge has not been submitted, either in part or whole, for a degree at this or any other educational institution except where due acknowledgement is made in this work. Quotes and paraphrased material are clearly acknowledged and all sources are referenced.

Devid Mrusek Marburg, December 19th, 2018

Jesus sagte zu ihm: Weil du mich gesehen hast, glaubst du. Selig sind, die nicht sehen und doch glauben.

Joh 20,29.

Surface is an illusion, but so is depth.

-David Hockney

Citizen bacillus.

-John Brunner

Publications

The majority of the work presented herein has been published in the following articles or is part of the following manuscripts:

Karniel A*, **Mrusek D***, Steinchen W, Dym O, Bange G, Bibi E (2018) Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J Mol Biol 430(11):1607-1620.
* Both authors contributed equally to this work

Mrusek D, Hook JC, Schwan M, Blagotinsek V, Schier L, Thormann KM, Bange G FlhG couples C-Ring assembly to flagellar gene expression, *manuscript in preparation*.

Mrusek D, Hoffman T, Kearns DB, Bremer E, Bange G Bacterial biofilm activator RemA features a Histone-like mode of DNA interaction, *manuscript in preparation*.

Publications not included in this work are listed below:
Brenzinger S, Pecina A, **Mrusek D**, Mann P, Völse K, Wimmi S, Ruppert U, Becker A, Ringgaard S, Bange G, Thormann KM (2018) ZomB is essential for flagellar motor reversals in *Shewanella putrefaciens* and *Vibrio parahaemolyticus*, Mol Micro 109(5):694-709.

Rossmann FM, Rick T, **Mrusek D**, Sprankel L, Dörrich AK, Leonhard T, Bubendorfer S, Kaever V, Bange G, Thormann KM (2019) The GGDEF Domain of the Phosphodiesterase PdeB in Shewanella putrefaciens Mediates Recruitment by the Polar Landmark Protein HubP, J Bacteriol, 201:e00534-18.

Table of Contents

Zusammenfassung 6

Abstract 9

1 FlhG couples C-Ring assembly to flagellar gene expression 11

1.1 Introduction 11
1.1.1 Flagellar bacterial motility 11
1.1.2 MinD-like ATPase FlhG numerically regulates flagella biosynthesis in monotrichous bacteria 12
1.2 FlrA shows a domain architecture reminiscent to response regulators from two-component signalling systems 13
1.3 Aim of the work 15
1.4 Results 16
1.4.1 FlhG traverses with FliM to the nascent flagellar C-ring. 16
1.4.2 FliM forms a stable complex with FlhG and stimulates the ATPase activity of FlhG 18
1.4.3 The flagellar C-Ring can be reconstituted *in vitro* 20
1.4.4 Loss of FlhG does not affect C-ring assembly 22
1.4.5 FlrA interacts with FlhG in an ATP-dependent manner via its C-terminal domain 24
1.4.6 Towards the co-crystallization of FlhG and FlrA-HTH 26
1.4.7 The interaction with FlhG does not abrogate DNA-binding by FlrA-HTH 28
1.4.8 FlhG inhibits transcriptional activity of FlrA, but is itself not regulated by FlrA 29
1.5 Discussion 31
1.6 Summary and future perspectives 36

2 RemA features a histone-like mode of DNA interaction 39

2.1 Introduction 39
2.1.1 A small subclass of transcriptional regulators binds DNA with a fold different from HTH 39
2.1.2 RemA is the master regulator for biofilm formation and salt-stress response 40
2.2 Aim of this work 43
2.3 Results 43
2.3.1 The crystal structure of RemA shows a donut-shaped octamer 43
2.3.2 RemA-octamers further associate to hexadecamers 47
2.3.3 Homologues of RemA show a high degree of conservation over the whole sequence 48
2.3.4 RemA hexadecamers and octamers are in dynamic equilibrium with one another 49
2.3.5 Multiple binding sites form a continuous DNA interaction interface on RemA 51
2.3.6 Towards the visualization of RemA-DNA interaction by Atomic force microscopy 52
2.3.7 RemA represents the minimal motif of the LytTR fold 54
2.3.8 Mutants of RemA maintain a ring-shaped oligomeric structure 56
2.3.9 Mutations of *Gt*RemA affect the ratio between hexadecamer and octamer (or: tetradecamer and heptamer for R18W) 57
2.3.10 Other mutants which were analysed by analytical SEC but did not yield diffracting crystals: 59
2.3.11 Mutants of RemA are impaired in their ability to bind DNA and act as transcriptional activators 59
2.3.12 Summary of structural and functional results 61
2.4 Discussion 62
2.4.1 The structure of RemA shows a single-domain protein with a compact fold that oligomerizes into higher-order arrangements 62
2.4.2 Single-amino acid mutations do not abrogate oligomerization of RemA 63

2.4.3 RemA can exist in two states of unknown biological function 63
2.4.4 RemA binds DNA via its lateral convex sides 64
2.4.5 Sequence-specificity of RemA could be an intrinsic feature of the DNA 68
2.4.6 Sequence-specificity of RemA could be promoted by interaction partners of RemA 69
2.4.7 Identification of similar proteins to RemA that bind DNA 70
2.4.8 Future perspectives 71

3 Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor 73

4 Material and Methods 88

4.1 Materials 88
4.1.1 Chemicals, enzymes, combustibles 88
4.1.2 Enzymes and cloning equipment 88
4.1.3 Oligonucleotides for cloning 88
4.1.4 Protein biochemistry 88
4.1.5 Protein Crystallization 89
4.1.6 Data collection and structure determinantion 89
4.1.7 Plasmids 89
4.1.8 *E. coli* strains 90
4.1.9 Buffers and growth media 90
4.2 Methods 90
4.2.1 Molecular cloning 90
4.2.2 Plasmid preparation 91
4.2.3 Agarose gel electrophoresis 91
4.2.4 Protein production and purification 91
4.2.5 SDS-PAGE 92
4.2.6 Glutathione-S-transferase (GST) binding assays 92
4.2.7 Ni-NTA affinity binding assays 92
4.2.8 Hydrolysis assays 93
4.2.9 Analytical Size-exclusion chromatography 93
4.2.10 Atomic Force Microscopy 93
4.2.11 Microscale Thermophoresis 93

5 Bibliography 95

6 Appendix 106

6.1 Crystallization conditions and details of data acquisition for RemA variants 106

7 Scientific curriculum 110

8 Acknowledgements 110

Zusammenfassung

Teil 1: Numerische Regulierung im monotrichen Bakterium *Shewanella putrefaciens*

Die Fähigkeit von Mikroorganismen, sich an wechselnde Umgebungen anzupassen, hat dazu geführt, dass Bakterien beinahe jede Nische des Planeten Erde besiedelt haben. Eine Schlüsselfähigkeit für das Überleben von Bakterien stellt die Motilität dar. Diese erlaubt Bakterien, bevorzugte Lebensumgebungen anzusteuern und solche Umgebungen, die für das Überleben nicht förderlich sind, zu verlassen. In Verbindung mit einem Sensorium, durch das Bakterien Nährstoffe und andere Parameter wahrnehmen können, ermöglicht die Motilität den Bakterien die Bewegung in Richtung von Nährstoffgradienten. Bakterielle Motilität wird großmehrheitlich durch Flagellen ermöglicht. Die Biogenese eines Flagellums ist für die Zelle ein sehr kostspieliger Prozess. Dementsprechend ist die Flagellen-Biogenese hochgradig reguliert. In dem monotrich flagellierten Bakterium *Shewanella putrefaciens* sind die Proteine FlhF und FlhG verantwortlich für die Etablierung des flagellaren Musters. Im Rahmen dieser Arbeit konnte gezeigt werden, dass **FlhG** die Anzahl der Flagellen auf eins reguliert, indem es direkt mit dem Master-Regulator der flagellaren Biogenese, **FlrA**, interagiert. FlhG ist darüberhinaus am Zusammenbau des cytoplasmatischen Teils des Flagellums, dem **C-Ring**, beteiligt. Transkriptionelle Kontrolle via FlrA sowie die C-Ring-Assemblierung via **FliM** werden durch **dieselbe Bindestelle** an FlhG gesteuert. Dies macht deutlich, dass FlhG eine bisher unerkannte Schlüsselrolle spielt und den Prozess der Flagellen-Biogenese mit der transkriptionellen Regulierung integriert. Zusammen genommen bilden diese Erkenntnisse einen wichtigen Schritt in Richtung einer vollständigen Beschreibung der Flagellenbiogenese und der numerischen Regulierung derselben. Damit bilden die Einblicke auch die Basis für weitere Untersuchungen.

Teil 2: Transkriptionelle Regulierung von Biofilmen geschieht durch RemA, welches histon-artig mit DNA interagiert

Anstelle der motilen Lebensart sind viele Bakterien in der Lage, in einem gesellschaftlichen Lebensstil zu existieren in Form von Biofilmen. Biofilme sind mehrzellige Gruppierungen von bakteriellen Zellen, in welchen Aufgabenteilung stattfindet und die eine erhöhte Resistenz gegenüber Antibiotika und Umwelteinflüssen bieten. Dies wird massgeblich bewerkstelligt durch die **Sekretion von extrazellulären Proteinen und anderen Biomolekülen**. RemA ist ein zentrales Protein während dieses Prozesses, welches die Sekretion von diesen extrazellulären Bausteinen transkriptionell aktiviert. Darüber hinaus aktiviert RemA **Schutzprozesse der Zelle**, um hohen Salinitäten entgegenzusteuern, die sich bei der Biofilm-Bildung zwangsläufig ergeben. Im Rahmen dieser Arbeit konnte die Struktur von **RemA** aus *Geobacillus thermodenitrificans* aufgeklärt werden. Die Struktur von RemA zeigt eine gänzlich neue Form der DNA-Interaktion, die an das **DNA-Looping** von Histon-Komplexen erinnert. Mittels biochemischer Methoden konnte die Art und Weise eruiert werden, wie RemA DNA bindet, und es konnten die strukturellen Auswirkungen der Mutation von funktionell wichtige Aminosäuren identifiziert werden. Damit bilden die hier gewonnenen Erkenntnisse eine wichtige Grundlage, um die Funktion von RemA im zellulären Kontext zu verstehen. Gleichzeitig ermöglicht diese Arbeit weiterreichende Untersuchungen, um RemA gemeinschaftlich mit DNA strukturell aufzuklären und weitere Proteine zu identifizieren, die RemA ähnlich sind.

Teil 3: Membranprotein-Biogenese wird durch einen co-translationellen Zustand von FtsY gesteuert.

Membranproteine werden durch **Ribosomen** translatiert und großmehrheitlich co-translational durch das **SecYEG-Translocon** in die Plasmamembran insertiert. Ein wichtiger Faktor beim co-translationalen Einbau von Membranproteinen in die Membran spielt der SRP-Rezeptor **FtsY**, welcher zusammen mit dem SRP-Partikel **FFH** und **SRP-RNA** die Zielführung von Ribosomen zum SecYEG-Translocon ermöglicht. Im Rahmen dieser Arbeit konnte gezeigt werden, dass ein co-translationaler Zustand von FtsY, die **helikale**

Domäne N2-4, massgeblich zur Zielsteuerung zur Membran beiträgt. Durch kristallographische Studien und Untersuchungen in Lösung konnte gezeigt werden, dass die Subdomäne N2-4 isoliert eine andere Faltung zeigt als im Kontext der G-Domäne von FtsY. Diese Beobachtung stellt ein Novum dar, denn die strukturelle Bi-stabilität von N2-4 geht offenbar einher mit dedizierten Funktionen. Diese Erkenntnisse stellen somit einen wichtigen Baustein dar im Feld der Membranprotein-Biogenese. Die Arbeiten dienen auch als Basis für weitergehende Untersuchungen, ob solch ein strukturell bimodaler Zustand bei Homologen von FtsY (etwa **FlhF**) oder anderen Proteinen ebenfalls auftritt.

Abstract

Part 1: Numerical regulation in the monotrichous bacterium *Shewanella putrefaciens*

Microorganisms have the ability to adapt to changing environmental conditinos. This has enabled them to colonize virtually nearly every niche on the planet Earth. Key to this ability is bacterial motility, which allows bacteria to move away from unfavourable conditions and to move towards favourable conditions. In connection with a sensory system, which detects chemical cues and other stimuli, bacteria can move towards nutrients. Bacterial motility is largely enabled by flagella. The biogenesis of a flagellum is a very costly process, which is for this reason highly regulated. In the monotrichous bacterium *Shewanella putrefaciens*, **FlhF** and **FlhG** are responsible for maintaining number and location of the single polar flagellum. In the course of this work, it could be shown that FlhG limits the number of flagella to one by directly interacting with the master transcriptional regulator of the flagellum, **FlrA**. Furthermore, FlhG is implicated in assembly of the cytosolic face of the flagellum, the **C-Ring**. The transcriptional control via FlrA as well as the C-Ring assembly via **FliM** occur through the **same binding site on FlhG**. This highlights the central role of FlhG and shows that FlhG integrates the two processes to regulate flagellar number. Taken together, these observations represent an important step towards a complete conceptual description of flagellar biogenesis. Thereby, these results also form the basis for further research.

Part 2: Transcriptional regulation of biofilms is mediated by RemA, which interacts with DNA in a histone-like manner

Instead of a motile lifestyle, bacteria can also establish a multicellular, sessile lifestyle in the form of biofilms. In biofilms, bacterial cells establish a division of labour and establish an increased resistance against antibiotics and environmental hazardous conditions. This is mediated by the secretion of extracellular proteins and other biological molecules. The protein **RemA** is central to this process, as it activates the **secretion of these extracellular components**. Furthermore, RemA is implicated in processes which enable a cellular **protection against osmotic pressure**, which occurs during biofilm formation. In the context

of this work, the structure of **RemA** from *Geobacillus thermodenitrificans* could be elucidated. RemA interacts with DNA in a novel and unique way, which is reminiscent of **DNA-looping** by histone-complexes. By means of biochemical methods, crucial residues of RemA responsible for DNA interaction could be functionally investigated. Furthermore, the structural fate of amino acid mutations, which impair the functionality of RemA, could be investigated. Taken together, this work represents an important step towards the understanding of the transcriptional processes that govern biofilm-formation and osmoprotection in *Bacillus subtilis*. This work also provides the basis to further investigate the function of RemA in the cellular context. In the future, the structural investigation of RemA-DNA-interaction is facilitated by the insights obtained in the context of this work.

Part 3: Membrane protein biogenesis is regulated by a structurally unique, co-translational state of FtsY.

Membrane proteins are translated by **ribosomes** and predominantly inserted into the membrane by the **SecYEG-translocon**. A factor critical for this process is the **SRP-receptor FtsY**, which enables co-translational targeting to the translocon in coopration with the **SRP-particle FFH and SRP-RNA**. In the context of this work it could be shown that a co-translational state of FtsY, the helical domain **N2-4**, critically mediates membrane targeting of the receptor. By means of crystallographic analyses and studies in solution, it could be shown that the subdomain of N2-4 possesses a different fold when isolated than in the context of the G-domain of FtsY. This observation represents a unique paradigm, which indicates that nascent N2-4 executes a different function during its own translation than when N2-4 is part of the mature FtsY-receptor. These results are an important step towards the conceptual understanding of membrane protein biogenesis and –targeting. Further work could elucidate, whether this concept also applies to homologs of FtsY such as **FlhF**.

1 FlhG couples C-Ring assembly to flagellar gene expression

1.1 Introduction

1.1.1 Flagellar bacterial motility

The first microscopic observation of bacteria by Leuwenhook in 1673 revealed that these microscopic organisms are highly motile. Bacterial motility is a major characteristic of microorganisms and enables them to move towards nutrients or colonize host organisms. Motility can be mediated by different mechanisms, e.g. flagella, pili or other systems.[1] Flagellar motility is vastly distributed and conserved among bacterial species. Flagella enable motility both in liquids as well as on surfaces.[2] The flagellum is, broadly speaking, a membrane-embedded structure with a long extracellular filament.[3] In the following, a short overview of flagellar assembly, architecture and flagellation pattern is given:

The flagellum is assembled in a hierarchical manner.[4] Membrane-embedded parts are assembled first. Thereafter, cytosolic components (C-Ring and ATPase complex) assemble. In a third step, extracellular components (rod-, hook-proteins) and finally the highly abundant filament-protein flagellin are produced, exported and assembled outside the cell.[5] The exact mechanism of flagellar assembly is not understood in detail: For instance, it is not clear whether the C-Ring is assembled stepwise or »en-bloc«. Hierarchical production of flagellar building blocks is regulated, among other mechanisms, by transcriptional tiers[6] as well as by translational control[7]. The ordered export of substrates is regulated by membrane proteins in the type-3 secretion system such as FlhA.[8]

The flagellar structure can be divided into four segments, namely the cytosolic part (C-Ring and ATPase-Complex), the membrane-embedded basal-body (Type-3 secretion system and MS-Ring), the rod/hook and the filament (**Fig. 1A**). The C-ring of the flagellum is composed of multiple copies of the proteins FliN, FliM and FliG. FliG is connected to the basal body via the MS-Ring protein FliF

(**Fig. 1B**). The stoichiometry of the proteins FliG:FliM:FliN is approximately 26:34:100.[9,10] The C-Ring transmits torque from the stator complexes to the rod. This is achieved by FliG, which interacts with MotA[11] and FliF.[12,13] The C-Ring is also named »Switch-complex«, because it controls flagellar rotation and direction in response to environmental stimuli. Nutrients are sensed by the chemotaxis system and transmitted by CheY to FliM/FliN, which induces conformational changes in the C-Ring and a reversal of rotational direction from counter-clockwise to clockwise.[14]

The number and arrangement of flagella on the surface of the cell, the so-called flagellation pattern, is a species-specific characteristic and used for the taxonomic characterization of bacterial species.[15] The flagellation pattern has undergone evolutionary development[16] and the flagellation pattern of a given species can be informative for the living environment of that species. Generally, flagellation at the poles and on the lateral sides are differentiated. Some species display a mixed type of flagellation.[17–19] The numerical and spatial regulation of flagella, which is the prerequisite for establishing and maintaining flagellation patterns, is mediated in some species by the regulatory proteins (FlhF, FlhG, among others).[20–23] To date, the spatial regulation is not understood and is thought to respond to landmark proteins, cell envelope curvature or lipid composition.

1.1.2 MinD-like ATPase FlhG numerically regulates flagella biosynthesis in monotrichous bacteria

In polar flagellates, such as *Vibrio*,[20] *Pseudomonas*,[24] and *Shewanella*,[25] FlhG (also: FleN) restricts the number to one flagellum per cell pole. Deletion of *flhG* in these species leads to hyperflagellation and impaired motility. FlhG consists of an N-terminal activator helix, a central ATPase domain and a C-terminal amphipatic helix (also: membrane targeting sequence MTS, **Fig. 1C**). FlhG acts in concert with the signal recognition particle (SRP)-GTPase FlhF[21,26,27] that has been suggested to recruit the flagellar protein FliF to the cell pole in the polar-flagellated species, such as *V. cholera*.[28] FlhG shares significant homology to the MinD/ParA-type ATPases at the structural and functional level. Like the cell division site

determining MinD protein,[29] FlhG forms ATP-dependent homodimers that interact with the inner membrane through a C-terminal MTS.[30] In species that lack MinD, FlhG cooperates with FtsZ to initiate cell division.[31] Moreover, FlhG interacts with the C-ring protein FliM via the N-terminus of FliM in polar flagellated bacteria (**Fig. 1C**). However, the functional consequences of these interactions are far from being understood.

1.2 FlrA shows a domain architecture reminiscent to response regulators from two-component signalling systems

Two-component signalling (TCS) is the most prevalent prokaryotic mechanism by which bacteria regulate their transcriptional response in response to signal transduction.[32] Two-component mechanisms are composed of two conserved proteins, a histidine protein kinase and a response regulatory protein. Environmental stimuli lead to autophosphorylation of the histidine kinase and subsequent transfer of the PO_4-group to a conserved aspartate (or: Ser/Thr/Tyr) in the receiver domain of the response regulator. The phosphorylated response regulator then undergoes a conformational change, which enables the response. Two-component systems occur also in the context of the flagellum (i.e. CheA, CheB/CheY).[33] Additionally, in monotrichous flagellated species such as *Pseudomonas aeruginosa* and *Shewanella putrefaciens*, a TCS (composed of the proteins FleS and FleR) is involved in the hierarchical biosynthesis of the single polar flagellum.[34] FleS and FleR are under the control of the sigma^{54}-factor RpoN, which requires the enhancer binding protein (eBP) FlrA (also: FleQ) to effect transcription.[35] FlrA is composed of three domains separated by linkers, namely a N-terminal receiver domain, a central AAA+-ATPase domain and a C-terminal domain comprising a helix-turn-helix-motif (HTH), which enables DNA-binding (**Fig. 1C**). While FlrA shows a domain architecture similar to TCS-response regulators, no cognate histidine kinase is known. The close homolog from *Pseudomonas aeruginosa* also lacks the conserved residue for phosphotransfer in the Rec-domain. However, this is not the case for *Sp*FlrA and therefore, the capacity of *Sp*FlrA to undergo phosphorylation is not known. FlrA from *P. aeruginosa* responds to cyclic-di-GMP, but no signalling partner for the Rec-

domain is known.[36] FlrA belongs to the family of NtrC-like eBPs, of which several examples have been crystallized.[37] Structurally, the receiver-domain and the central AAA+-ATPase-domain from *P. aeruginosa* FleQ are known. The structure of the Receiver-domain of FlrA[38] confirms the absence of a phosphorylation relay and shows a novel dimerization mode of *Pa*FlrA. FlrA, like homologs from the NtrC-family, is dimeric in solution, but forms ATP- or cdG-dependent hexamers.[36] The second messenger cyclic di-GMP (cdG), a signalling molecule which mediates the transition between planktonic and sessile lifestyle, binds to FlrA. Structures of the ATPase-domain of FlrA with cdG as ligand have revealed that cdG binds at a site distinct from the ATP binding pocket. Binding of cdG obstructs ATP binding and thereby prevents FlrA function in the context of flagellar gene transcription.[39] Interestingly, FlrA in *P. aeruginosa* also inhibits transcription of genes involved in exopolysaccharide production when cdG levels are low.[39] In *Shewanella putrefaciens*, FlrA also suppresses the transcription of the adhesion BpfA, which is important for biofilm formation. Cyclic di-GMP abolishes binding of FlrA to the BpfA-promoter.[40]

ATPase activity is critical for FleQ function. ATP binding is required for oligomerization of FleQ and its homologs. In the ATP-bound hexameric state, a cooperative mechanism involving a conserved arginine leads to the intermolecular stimulation and hydrolysis of ATP. The cascade-like hydrolysis of six ATP molecules induces to a power-stroke, which is transmitted to the RNAP in the holo-enzyme and promotes DNA open-complex formation and transcription.[41] Hence, bound cdG inhibits ATP binding and hydrolysis, and FleQ cannot fulfil its role as an transcriptional activator.

In *P. aeruginosa*, FlrA interacts with FlhG.[39,42,43] FlhG, as mentioned above, is crucial for maintaining the correct number of flagella in monotrichous bacteria. FlhG inhibits the enzymatic activity of FlrA and thus negatively impacts flagellar gene expression.[43] The nature of this interaction and the mechanism by which FlhG inhibits FlrA are not understood to date. It is also not clear how the interaction

of these proteins in monotrichous bacteria enforce a strict numerical flagellar control.

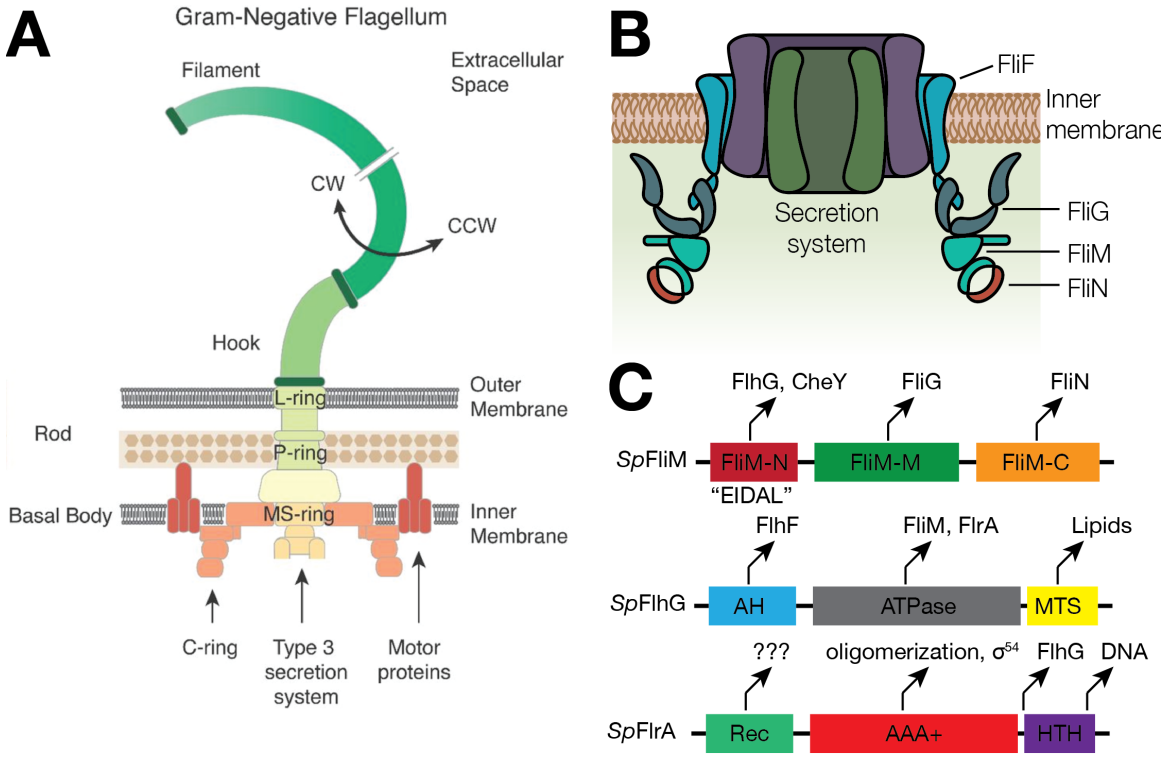

Figure 1. A: Architecture of the flagellum in gram-negative organisms (used with permission from [15]). **B:** Detailed architecture of the flagellar C-Ring. **C:** Domain architecture of FliM, FlhG and FlrA from *Shewanella putrefaciens*. Proteins that interact with the respective domains are indicated.

1.3 Aim of the work

The ability of bacterial cells to numerically regulate cell organelles represents a crucial advantage and important factor of virulence. This is especially so as the numerical regulation is reproducibly established during each cell division cycle. Flagellar building block production is a costly process which involves in the last steps the expression and export of circa 20'000 copies of flagellin. Therefore, bacteria have established intricate mechanisms of regulation and checkpoints. Conceptually, numerical regulation of cell organelles represents a valuable feature for synthetically assembled cells. To understand how bacterial cells spatially and numerically assemble organelles of locomotion can help to implement such signalling networks for other purposes.

To date, the exact mechanism how flagellar number is controlled by FlhG remains elusive. While the structure and interaction network of FlhG has been described, the dynamic impact that FlhG has on flagellar number is not understood. FlhG has been shown to interact with lipids, the SRP-like GTPase FlhF and the C-Ring protein FliM. However, the integration of these binding events is still lacking.
In the context of this work, I set out to determine additional binding partners of FlhG and to assess the functional consequence of their interaction with FlhG. I planned to determine a structural model of these interactions. Furthermore, I aimed at a quantitative description of the interaction network of FlhG. The goal of this work is a full description of the FlhG-interactome, so that these information can be integrated in a mathematical model which can explain how FlhG in *S. putrefaciens* limits the flagellar number to one.

1.4 Results

1.4.1 FlhG traverses with FliM to the nascent flagellar C-ring.

Previous studies showed that in *S. putrefaciens* CN-32 the C-Ring protein FliM localizes to the flagellated cell pole and that this localization pattern is reduced in the absence of FlhG.[25] Therefore, FlhG could allocate FliM and other C-ring proteins to the nascent polar flagellum. However, it could also be that FlhG passively traverses with FliM to the nascent flagellar structure. To better understand the role of FlhG in the biological context of C-ring assembly, the cellular localization of FlhG was analyzed in dependence of FliM. To do so, FliM was expressed from the genome as a mCherry-fusion protein. FlhG was deleted in the genome and ectopically overexpressed as a GFP-fusion protein. Then, the localization of FliM and FlhG were analyzed in a wild-type strain (»FliM-mCherry«) and in a strain in which FliM was truncated at the N-terminus (»FliM-mCherry ΔEIDAL«). The »EIDAL«-motif at the N-terminus of FliM is responsible for the interaction with FlhG.[25] Dr. Florian Roßmann performed the experiments and kindly provided the images. Truncation of the N-terminus of FliM does not affect FliM-localization. However, polar localization of FlhG is drastically reduced in the »FliM-mCherry ΔEIDAL«-strain (**Fig. 2A**). These findings clearly show that FlhG passively traverses with the FliM/FliN complex to the nascent flagellar structure.

Off-polar localization of FlhG also leads to hyperflagellation, which phenocopies the *dflhG*-strain. Therefore, the correct numerical regulation critically depends on the correct localization of FlhG.

To verify that the EIDAL motif mediates interaction with FlhG, a GST-interaction assay was performed. Briefly, the 44 N-terminal residues of FliM (which contain the »EIDAL«-motif) where fused to GST (GST-FliM-N). The interaction assays show that the N-terminus of FliM is necessary and sufficient to interact with FlhG independent of ATP. A mutant of FlhG, which cannot hydrolyse ATP (FlhG D58A) also binds GST-FliM-N (**Fig. 2B**, upper panel). These findings show that FliM-N does not recognize a specific functional state of FlhG.

To gain a better quantitative understanding of the interaction between FlhG and FliM, the dissociation constant (K_D) of FlhG and GST-FliM-N was determined by microscale thermophoresis (MST): 2.12 ± 0.373 µM. To further characterize the interaction between FlhG and FliM-N, mutants of FlhG were prepared based on structural information from the *B. subtilis* homologue. In all mutants tested, the interaction with GST-FliM-N was markedly reduced (**Fig. 2B**, lower panel). These residues are located on helices α6 and α7, which are solvent exposed and distal from the ATP-binding site (**Fig. 2C**). Notably, amino acids K175, K205 and F213 (K177, R207 and F213 in *G. thermodenitrificans*) are conserved between homologs of FlhG (**Fig. 2D**), except for FlhG from *Campylobacter jejuni*. The EIDAL-motif at the N-terminus of FliM is strictly conserved among homologs of FliM (**Fig. 2D**).

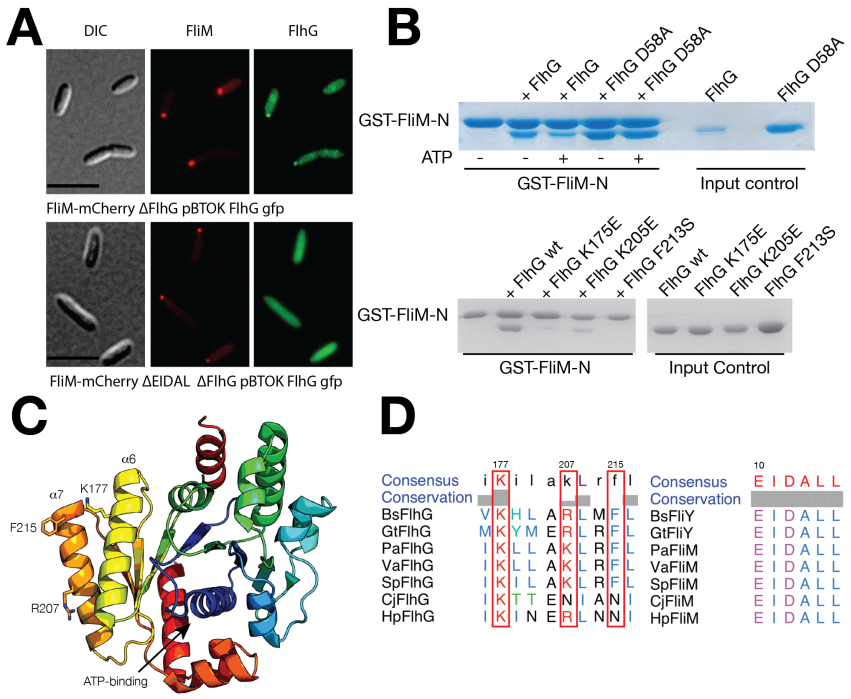

Figure 2. A. Cellular localization of GFP-tagged FlhG and mCherry-tagged FliM. Dr. Florian Roßmann performed the experiments and kindly provided the images. **B**: GST-Pulldown of FliM-N versus FlhG and FlhG D58A in the presence and absence of ATP (upper panel). GST-Pulldown of FliM-N versus FlhG and mutants of FlhG (lower panel). **C**: Structural representation of FliM-interacting residues in *Gt*FlhG (PDB: 4RZ2). The numbering is according to the *Gt*FlhG-sequence and differs from *S. putrefaciens* FlhG. **D**: Left: Sequence alignment of FlhG-homologs (Bs, *Bacillus subtilis*; Gt, *Geobacillus thermodenitrificans*; Pa, *Pseudomonas aeruginosa*; Va, *Vibrio alginolyticus*; Sp, *Shewanella putrefaciens*; Cj, *Campylobacter jejuni*; Hp, *Helicobacter pylori*). The FliM-interacting residues are boxed in red. Right: Sequence alignment of FliM-homologs. In *B. subtilis* and *G. thermodenitrificans*, both FliM and FliY are present and contain an EIDAL-motif, but only FliY interacts with FlhG.[25]

1.4.2 FliM forms a stable complex with FlhG and stimulates the ATPase activity of FlhG

To further verify the interaction between FliM and FlhG, size exclusion chromatography was employed. The interaction of FliM with FlhG was probed both with the isolated EIDAL-segment fused to Protein G-B1-domain (GB1)[44] or

Glutathione-S-Transferase (GST)[45] as well as with FliM in context of FliN (FliM/FliN-complex). On analytical size-exclusion chromatography (SEC), GB1-FliM-N (15 kDa) eluted at a volume corresponding to 31 kDa. GB1 and FliM-N are monomeric proteins. The difference between calculated and observed molecular weight could arise from a non-globular protein shape. FlhG eluted at a volume corresponding to its calculated molecular weight (31 kDa). When mixed and incubated, both proteins formed a complex which eluted at a volume corresponding to 59 kDa, indicating a heterodimeric stoichiometry (**Fig. 3A**). Similarly, FlhG also forms a SEC-stable complex with FliM/FliN (**Fig. 3B**). As FlhG is an ATPase, the effect of FliM-N on the enzymatic activity was assessed. To do so, FlhG was incubated with ATP in the presence or absence of GST-FliM-N. In the presence of GST-FliM-N, FlhG showed a 5-fold increase of enzymatic activity (**Fig. 3C**) Taken together, these findings clearly show that the N-terminus of FliM enables FlhG to interact with the FliM/FliN-complex. The interaction between FliM and FlhG stimulates the ATPase-activity of FlhG by a sofar unknown mechanism. Functionally, the binding of FliM to FlhG stimulates the ATPase activity of FlhG, resulting in dissociation of the ATP-bound FlhG-homodimer. During this process, FliM likely stays attached to FlhG (**Fig. 3D**).

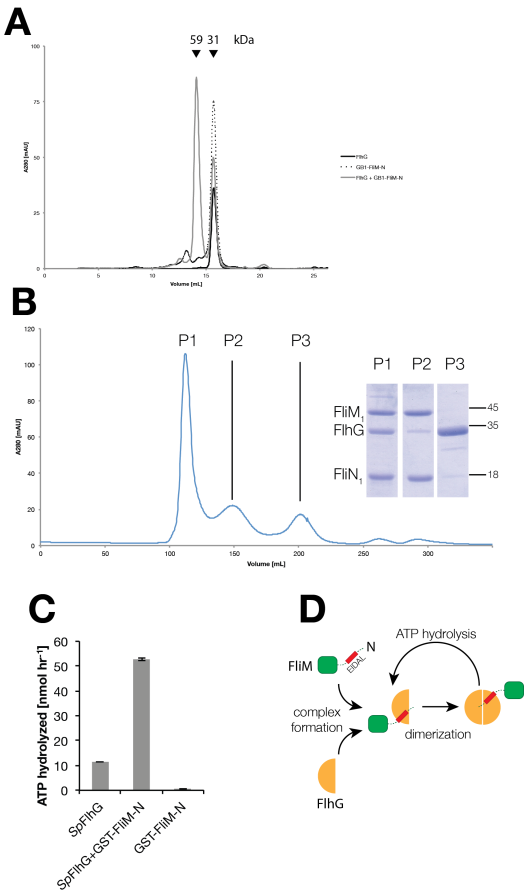

Figure 3. A: Analytical size-exclusion chromatography (SEC, Superdex200 10/300 GL) of GB1-FliM-N, FlhG and a mixture of both. When GB1-FliM-N and FlhG are incubated together and applied to analytical SEC, an additional peak is observed eluting at a volume which corresponds to a molecular weight of 59 kDa. Observed molecular weights were calculated by employing a molecular-weight calibration kit (GE Healthcare). **B**: Preparative SEC (Superdex200 26/600) of a Ni-purified co-expression of FliM/FliN and FlhG. FlhG forms a ternary complex with the FliM/FliN-complex, which elutes in the void fraction of the column due to its large molecular weight. **C**: FliM-N stimulates the ATPase activity of FlhG. **D**: Model for the interaction between FliM and FlhG.

1.4.3 The flagellar C-Ring can be reconstituted *in vitro*

To further challenge the ability of creating flagellar protein-complexes *in vitro*, a series of GST-pulldown-experiments was conducted with the aim of reconstituting the flagellar C-Ring. In a first step, GST-FliG was immobilized on GSH-beads and the SEC-purified FliM/FliN-complex was added to the reaction.

As a results, a stoichiometric complex between GST-FliG, FliM and FliN was observed (**Fig. 4A**, left). To consolidate this finding, di-domains of FliG (FliG-NM, FliG-MC, **Fig. 4A**, center) and single domains FliG-M and FliG-C were employed. The results show that GST-FliG-M is necessary and sufficient to bind FliM/FliN, while GST-FliG-C is not sufficient for the interaction with FliM/N (**Fig. 4A**, right).

Next, the situation at FliF was probed. FliF is an integral membrane protein with a small cytosolic domain (FliF-C) to which FliG binds via its N-terminus. To do so, GST-FliF-C was immobilized and full length FliG and truncated variants of FliG (FliG-NM, FliG-MC) were added. The experiment shows that FliG and the di-domain FliG-NM can stably interact with GST-FliF-C. In contrast, FliG-MC is not capable of interacting with GST-FliF-C (**Fig. 4B**, left). Finally, a complete C-Ring was reconstituted by immobilization of GST-FliF-C and subsequent addition of FliG and FliM/FliN (**Fig 4B**, center). Likewise, a »minimal« C-Ring was reconstituted by employing FliG-NM (**Fig. 4B**, right).

These findings show that the proteins $FliF_C$, FliG and FliM/FliN can be assembled *in vitro* to mimic the flagellar C-Ring, demonstrating that the assembly does not require co-factors or chaperons.

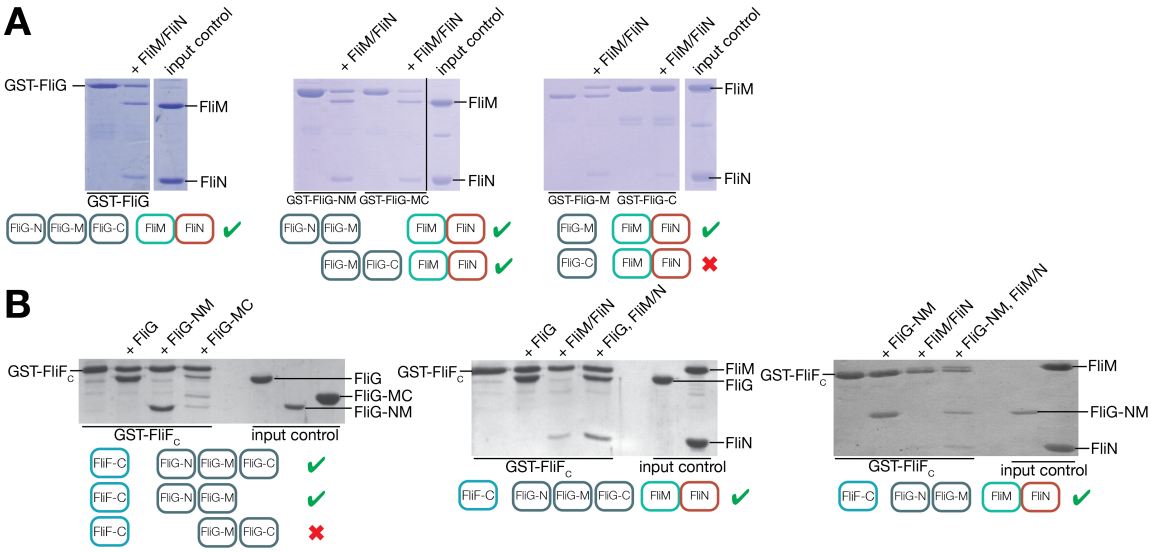

Figure 4. A: GST-Pulldown of full-length FliG (left panel), di-domains of FliG (middle panel) and single domains of FliG (right panel) versus a SEC-purified complex of FliM/FliN. The GST-pulldowns show that GST-FliG-M is necessary and sufficient to bind FliM/FliN. The employed domains of FliG and the result are drawn

schematically below the SDS-PAGE gels. **B**: GST-Pulldown of the cytosolic domain of FliF ($FliF_C$) versus full-length FliG and di-domains (FliG-NM, FliG-MC, left panel). Only full-length FliG and FliG-NM can interact with GST-$FliF_C$. Middle panel: GST-pulldown of GST-$FliF_C$ versus FliG and FliM/FliN. In lane 4, both FliG and FliM/FliN are added subsequently with a washing step in between. The results is a quaternary complex between GST-$FliF_C$, FliG and FliM/FliN. Right panel: The same as middle panel, but FliG-NM is employed. The employed domains of FliG and the result are drawn schematically below the SDS-PAGE gels.

1.4.4 Loss of FlhG does not affect C-ring assembly

Previously, it has been shown that the presence of FlhG (more specifically: presence of FlhG at the pole) is critical for correct numerical regulation in the monotrichous bacterium *Shewanella putrefaciens*. In the absence of FlhG (or mislocalization from the pole), hyperflagellation was observed. This indicates that flagella were assembled, but the integrity of C-Ring structures is not known. To determine the integrity of C-Ring structures in cells grown without FlhG, the basal body of *S. putrefaciens* polar flagella were imaged by cryo-electron tomography. Dr. Florian Roßmann performed the experiments and kindly provided the images. Image analysis shows that the motors display an 11-fold symmetry, and that the C-ring structures appear unaltered in strains lacking FlhG (**Fig. 5A**). This result indicates that FlhG traverses to the cell pole via FliM but is not part of the mature flagellum. To gain a deeper understanding of the fate of FlhG upon C-Ring assembly, the situation was reconstituted *in vitro*. Briefly, GST-FliG was immobilized on GSH-beads, and then the SEC-purified FliM/FliN- or FlhG-FliM/FliN-complexes (see above) were added, respectively. The complex of FliM/FliN bound stoichiometrically to GST-FliG, as observed before. However, in the case of the FlhG-FliM/FliN, FlhG was released during the process (**Fig. 5B**, left panel). To consolidate this finding, the experiment was repeated in the context of GST-$FliF_C$. To do so, a preformed complex of GST-$FliF_C$ and FliG was immobilized and probed with FliM/FliN and FlhG-FliM/FliN. A stoichiometric complex between GST-$FliF_C$, FliG and FliM/FliN was observed, but again FlhG was not part of this complex (**Fig. 5B**, right panel). These findings indicate that FlhG is released upon

binding of FliM/FliN to FliG, irrespective whether this happens at isolated FliG or in the context of $FliF_C$.

To gain furter insight into the mechanism underlying FlhG dissociation, pulldown-experiments were repeated with truncated variants of FliG. In pulldowns emplyoing di-domains of FliG, the addition of FlhG-FliM/FliN led to stoichiometric binding of FliM/FliN only in the case of FliG-MC, but not FliG-NM (**Fig. 5C**, left panel). In the next step, FliG-M was employed. Addition of FliM/FliN to GST-FliG-M led to stoichiometric binding of FliM/FliN (**Fig. 5C**, middle panel, as above). However, addition of FlhG-FliM/FliN to GST-FliG-M did not lead to an interaction between FliM/FliN and GST-FliG-M (**Fig. 5C**, right panel).

Taken together, these findings indicate that truncation of FliG abrogates the ability of FliM/FliN in the context of FlhG to interact with FliG.

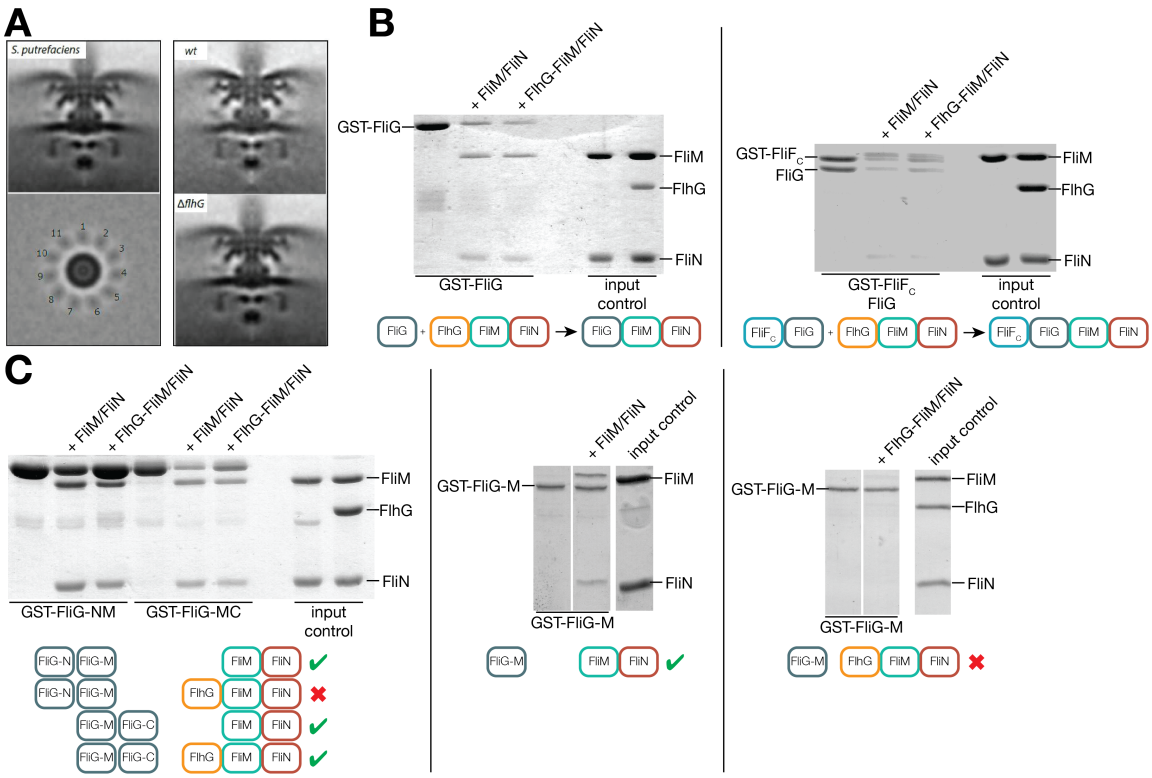

Figure 5. A: Electron-cryo-tomographic images of flagellar C-Rings from S. putrefaciens. Dr. Florian Rossmann performed the experiments and kindly provided the images. **B**: GST-Pulldown of full-length FliG versus a SEC-purified complexes of FliM/FliN and FlhG-FliM/FliN, respectively (left panel). Right panel: The preformed complex of GST-$FliF_C$/FliG is probed with SEC-purified complexes of FliM/FliN and FlhG-FliM/FliN, respectively. In all cases, dissociation of FlhG is

observed. The employed proteins and the result are drawn schematically below the SDS-PAGE gels. **C**: GST-pulldowns emplyoing di-domains (left panel) and single domain of FliG (middle and right panel) versus SEC-purified complexes of FliM/FliN and FlhG-FliM/FliN, respectively. The binding-efficiency of the FlhG-FliM/FliN-complex to FliG-NM is reduced, and no binding is observed to FliG-M. The employed domains of FliG and the result are drawn schematically below the SDS-PAGE gels.

1.4.5 FlrA interacts with FlhG in an ATP-dependent manner via its C-terminal domain

Previously, it was shown that FlhG interacts with FliM and that abrogation of this interaction leads to hyperflagellation. As FlhG is not part of the assembled C-Ring and dissociates upon interaction of FliM and FliG by a sofar unknown mechanism, FlhG must control flagellar number in another way. In order to identify further binding partners of FlhG, a yeast-two-hybrid (Y2H) interaction assay was performed.[46] Constructs were cloned by D. Mrusek and the experiments were performed by Dr. Dieter Kressler, who kindly provided the images. It was found that FlhG interacts with FlhF, FliM and FlrA, the master regulator of the flagellum (**Fig. 6A**). In GST-pulldown assays it was found that the C-terminal, DNA-binding domain of FlrA (FlrA-HTH, residues 373-478) interacts with FlhG in the presence of ATP. A mutant of FlhG, which is able to bind ATP, but unable to hydrolyze it, interacted with HTH irrespective of the nucleotides added (**Fig. 6B**). Further truncation of the HTH-domain revealed that residues 389-408 of FlrA are critical for the interaction with FlhG (**Fig. 6C**, upper panel). These residues correspond to helix »A« in the HTH-nomenclature and the linker region which connects the HTH to the ATPase-domain (**Fig. 6C**, lower left). The C-terminal part of FlrA is similar to that of other enhancer binding proteins, such as *Aquifex aeolicus* NtrC4, which has been structurally characterized as an isolated protein as well as in complex with DNA.[47] The C-terminal region of NtrC-type proteins is composed of 4 helices A-D, of which only helix D recognizes DNA. Helices A and B mediate dimerization of the HTH-domain (**Fig. 6C**, lower right).

The primary sequence of residues 389-408 shows a highly acidic (pI 3.7) segment rich in glutamates and serines, similar to the EIDAL-motif of FliM-N (pI 3.7, rich in aspartates and lysines).

Identification of the binding interface between FlhG and HTH was difficult, because the two proteins dissociated after size-exclusion chromatography. Preliminary competitive GST-pulldown experiments with FlhG, FliM-N and HTH indicated that HTH and FliM-N could bind to the same region of FlhG. To challenge the hypothesis that FliM and FlrA share a binding site on FlhG, GST-pulldown experiments were performed with mutants of FlhG. It was found that mutations K175E, K205E and F213S abrogated binding of FlhG to FlrA-HTH (**Fig. 6D**). Notably, these mutations also led to a loss of binding of FliM-N to FlhG. Thus, FliM and FlrA share the same binding site on FlhG.

To gain a better quantitative understanding of the interaction between FlhG and FlrA, the dissociation constant (K_D) of FlhG and FlrA-HTH was determined by microscale thermophoresis (MST) in the presence of AMPPNP: 1.027 ± 0.089 µM. The dissociation constant (K_D) of FlhG and AMPPNP is: 14.891 ± 4.1 mM.

Next, the functional consequence of the interaction of FlrA with FlhG was investigated. It was found that HTH strongly stimulates the ATPase activity of FlhG (**Fig. 6E**).

Taken together, these findings clearly show that helix A of the DNA-binding domain of FlrA enables interaction with the ATP-bound state of FlhG, which is dimeric. The interaction between FlrA-HTH and FlhG stimulates the ATPase-activity of FlhG by a sofar unknown mechanism. Functionally, the binding of FlrA-HTH to FlhG stimulates the ATPase activity of FlhG, resulting in dissociation of the ATP-bound FlhG-homodimer and FlrA-HTH (**Fig. 6F**).

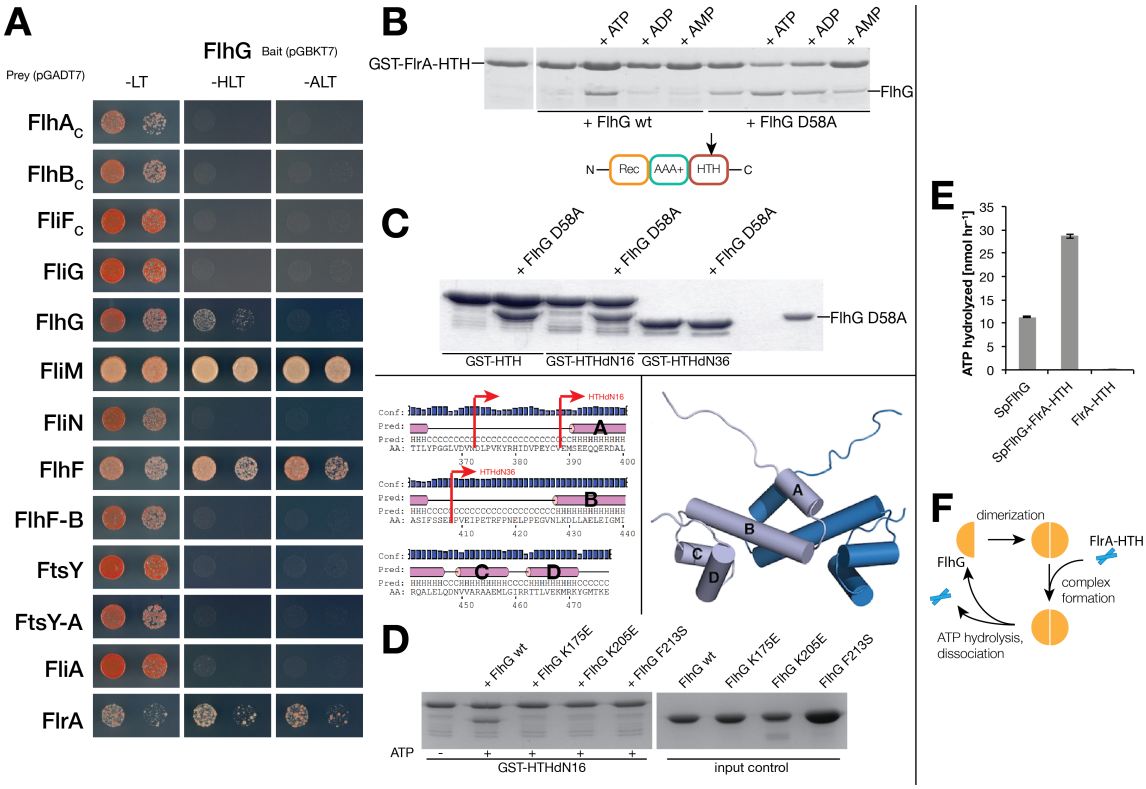

Figure 6. A: Yeast-two-hybrid assay emplyoing SpFlhG as bait versus various flagellar proteins. Dr. Dieter Kressler performed the experiments and kindly provided the images. FlhG interacts strongly with FliM and FlhF and also with FlrA. FlhG shows weak self-interaction. **B**: GST-Pulldown of FlrA-HTH versus FlhG and FlhG D58A in the absence and presence of nucleotides. The domain architecture of FlrA is drawn schematically below, indicating the binding site for FlhG. **C**: GST-Pulldown of FlrA-HTH and truncated variants versus FlhG D58A (upper panel). Lower panel, left: Predicted secondary structure of FlrA-HTH. Employed variants of FlrA-HTH are indicated by red arrows. Helices of HTH are indicated with letters A-D. Lower panel, right: Cylinderview of the HTH-domain of *A. aeolicus* NtrC4 (PDB: 3E7L). Reproduced with permission.[37] **D**: GST-Pulldown of FlrA-HTHdN16 versus FlhG and mutants in the presence of ATP. **E**: FlrA-HTH stimulates the ATPase-activity of FlhG. **F**: Model of the interaction between FlhG and FlrA-HTH and the fate during ATP hydrolysis.

1.4.6 Towards the co-crystallization of FlhG and FlrA-HTH

Coexpression of hexahistidine-tagged FlrA-HTH and untagged FlhG D58A led to co-elution of both proteins from Ni-affinity chromatography. However, this

complex dissociated on size-exclusion chromatography. When employing hexahistidine-tagged FlhG D58A and untagged FlrA-HTH, a complex of FlhG-HTH was observed (**Fig. 7A**). The combined fractions were concentrated and subjected to crystallization experiments. Crystals appeared after 24 hours in several conditions but did not diffract. In one condition (0.2 M tri-sodium citrate, 20% PEG3350 (v/v)), crystals appeared after five weeks (**Fig. 7B**). Two datasets were recorded (27.02.2018, ESRF Grenoble, beamline ID29, puck-code AA001A, entries x6 and x8). Both crystals belong to space group I222 (dimensions a=109.0 b=126.4 c=190.0 α=89.8 β=90.0 γ=90.0) and diffracted to 3.1-3.2 Å. Based on the solvent fraction, the unit cell contains 4 copies of *Sp*FlhG. It was attempted to solve the phase problem by molecular replacement employing the model of FlhG from *P. aeruginosa* (PDB: 5J1J, 64% identity), but no definite solution was found. The crystals could not be reproduced in fine-screens. Analytical size-exclusion chromatography of the concentrated peak fractions (peak 1, **Fig. 7A**) showed that the complex between FlhG D58A/FlrA-HTH dissociated over time (**Fig. 7C**). Cocrystallization was also attempted emplyoing a synthesized peptide (residues 389-409 of FlrA; H_2N-EMSEEQQERDALASIFSSEEP-COOH). Peptide and protein had to be added separately to crystallization drops, because the mixture of protein+peptide precipitated. No crystals have been oberved so far.

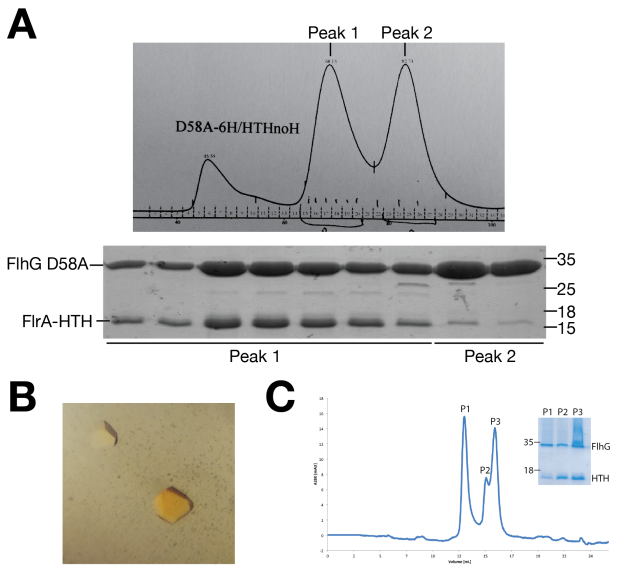

Figure 7. A: SEC-chromatogram of a co-expression of hexahistidine-tagged FlhG D58A and untagged FlrA-HTH (upper panel). SDS-PAGE of selected fractions

(lower panel). **B**: Crystals grown in a condition which consisted of 0.2 M tri-sodium citrate, 20% PEG3350 (v/v). **C**: Analytical SEC of combined and concentrated fractions (Peak 1, **Fig. 7A**) shows that the complex between FlhG D58A and FlrA-HTH dissociates.

1.4.7 The interaction with FlhG does not abrogate DNA-binding by FlrA-HTH

Enhancer binding proteins such as FlrA bind to upstream activating sequences (UAS), 80 to 150 bp upstream of the transcription start site.[48] To understand the impact of the interaction between FlhG and FlrA, the DNA-binding capacity of FlrA alone and in context with FlhG was investigated.

First, the ability of FlrA and FlrA-HTH to interact with target DNA (*flrBC*-promoter fragment)[49] was probed with electromobility shift assays (EMSA). To do so, a DNA fragment spanning 1000 bp around the *flrBC*-promoters was amplified by polymerase chain reaction (PCR), incubated with FlrA or FlrA-HTH, and the reaction loaded onto a 5% polyacrylamide-gel. The experiment showed that full-length FlrA, but also the single domain FlrA-HTH, are capable of interacting with DNA (**Fig. 8A**). In the case of FlrA-HTH, significant DNA-shifting was observed with concentrations above 25 µM. Full-length FlrA led to a similar shift at higher concentrations. The Walker-B mutant FlrA D233A, which can bind but not hydrolyze ATP,[50] shifted the DNA more drastically. FlrA-HTH shifted the DNA to a defined state also at higher concentrations (**Fig. 8B**). This finding indicates that FlrA-HTH and the DNA form a complex with a defined stoichiometry.

Previously, the FlhG-homologe MinD was implicated in DNA-binding.[51] Therefore, it was tested whether FlhG shows any interaction with DNA. The experiment shows that wild-type FlhG does not shift DNA irrespective of AMPPNP. However, FlhG D58A shows a dramatic shift of DNA both in the absence and presence of AMPPNP. When FlrA-HTH is employed, the DNA-shift is further enhanced in the presence of FlhG. The »super-shift« of FlrA-HTH and FlhG is not dependent on AMPPNP, which is contrary to previous observations. The combination of FlrA-HTH and FlhG D58A shows the same dramatic DNA-shift as in the case of FlhG D58A alone (**Fig. 8C**). In a model of a structural homolog of FlrA-HTH bound to

DNA (*A. aeolicus* NtrC4-DBD), it can be seen that the FlhG-interacting part of the DBD is not engaged in DNA-binding (**Fig. 8D**, market in red).

Taken together, these findings clearly show that the HTH-domain of FlrA is necessary and sufficient for DNA-interaction and that FlhG does not abrogate DNA-binding of FlrA-HTH, which is in agreement with structural information for a FlrA-HTH-homolog.

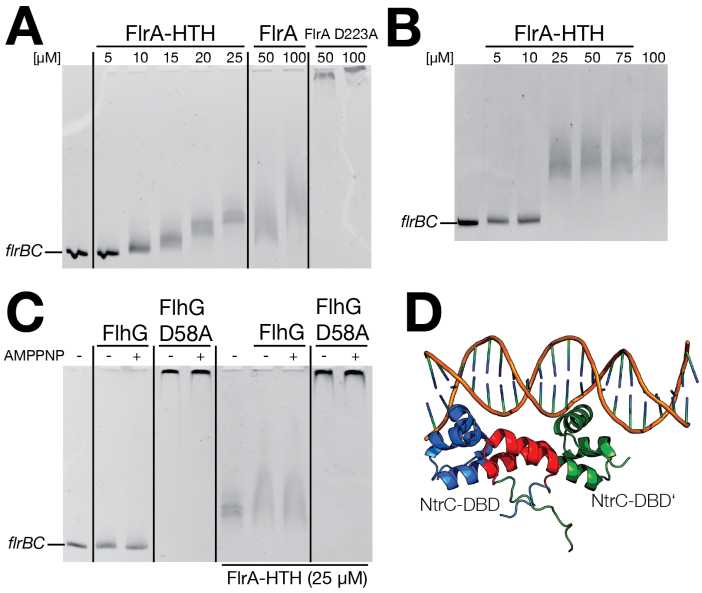

Figure 8. A: Electromobility shift assay employing FlrA, FlrA D233A and FlrA-HTH. **B**: Significant DNA-shifting of FlrA-HTH is observed with concentrations of 25 µM and greater. **C**: FlhG does not abrogate DNA-binding of FlrA-HTH. **D**: Crystal structure of the NtrC4-DBD-dimer bound to DNA (PDB: 4FTH). The two monomers of NtrC4-DBD are coloured in blue and green, respectively. The segment corresponding to the FlhG-binding site is coloured in red.

1.4.8 FlhG inhibits transcriptional activity of FlrA, but is itself not regulated by FlrA

In the sections above, it was established that FlhG binds FlrA and that this interaction does not abrogate DNA-binding of FlrA. FlhG acts as a negative regulator of flagellar biosynthesis, and thus the impact of FlhG on the transcriptional activity of FlrA was assessed by quantitative PCR (qPCR). Meike Schwan performed the experiments and kindly provided the data. The transcript levels of *flrA* and *fliF* were assessed in wild-type background, *flhG*-knockout and in strains transformed with an empty vector control (pBTOK EVC) or with a vector

encoding FlhG (pBTOK flhG). As shown in **Figure 9A**, the *fliF* promoter activity, determined by qPCR, was significantly enhanced in the Δ*flhG* strain. In contrast, overabundant FlhG decreased transcription of this gene below wild-type levels. In all conditions tested, *flrA*-promoter activity was not affected. In conclusion, FlhG negatively influences transcriptional activity from the FlrA-dependent *fliF*-promoter.

In the literature, FlrA is described as the master regulator of the flagellar gene hierarchy and is responsible for transcription of class II genes (e.g. *fliEFGHIJ, flhFG*, among others).[6] Following this reasoning and keeping the above findings in mind, transcription of *flhG* by FlrA represents a negative feedback loop. To challenge this hypothesis, the impact of FlrA on the presence of FlhG was determined *in vivo*. To do so, FLAG-tagged FlhG was detected by western blotting in strains which lacked FlrA, RpoN (alternative sigma^{54}-factor) or FliA (alternative sigma^{28}-factor). Meike Schwan performed the experiments and kindly provided the data. As is shown in **Figure 9B**, FLAG-FlhG was detected even in the absence of FlrA, RpoN and FliA. Preliminary quantifications by western blotting showed that FlhG is overabundant (ca 3-fold) when compared to FlrA or FliM (**Fig. 9C**, Dr. Florian Rossmann performed experiments and kindly provided images).

Taken together, these findings clearly show that while FlhG downregulates transcriptional activity of FlrA, FlhG itself is not under the control of FlrA. FlhG also does not impact the level of FlrA, which shows that FlrA is produced by a factor different from FlrA itself.

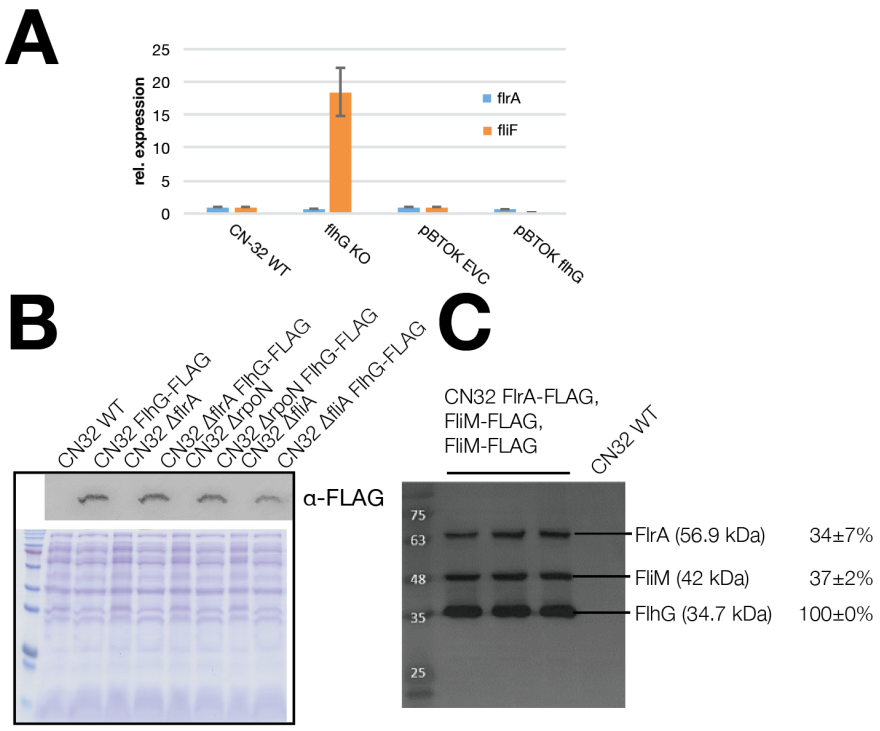

Figure 9. A: Quantitative PCR results for *flrA* and *fliF* determined in wild-type and *flhG*-knockout strain. Overabundance of FlhG was achieved by ectopic expression from a vector (pBTOK flhG) with the empty vector as control (EVC). **B**: FlhG is produced in the absence of FlrA, RpoN and FliA. Western blotting with α-FLAG antibody (upper) and coomassie-strained SDS-PAGE (lower). Meike Schwan performed the experiments and kindly provided the images. **C**: Quantitation of FLAG-tagged proteins in *S. putrefaciens*, grown at exponential phase. Dr. Florian Roßmann performed the experiments and kindly provided the image.

1.5 Discussion

In the sections above, a series of experiments was performed to shed light on the regulatory mechanism by which *Sp*FlhG restricts the number of flagella to one. In the following, three hypotheses are formulated and discussed regarding the role of FlhG in the context of flagellar assembly and numerical control:

1) FlhG assists in flagellar assembly, thereby limiting the amount of flagellar building blocks engaged in a single C-Ring to 24 (FliG) and 36 (FliM).

2) FlhG is produced alongside the other flagellar components and limits their production via a negative feedback loop and the inhibition of transcriptional activity of FlrA.
3) FlhG, which is not regulated by FlrA, continuously inhibits FlrA‘s transcriptional activity. Polar localization of FlhG is the main parameter which limits flagellar number to 1.

Hypothesis 1: FlhG actively assists in flagellar assembly

The experimental data presented above clearly show that FlhG does not play an active role in flagellar assembly itself. In *S. putrefaciens*, FlhG does not recruit flagellar building blocks to the nascent flagellar structure (**Fig. 2A**), nor is FlhG part of the mature C-Ring (**Fig. 5A**). Given the over-stoichiometric excess of FlhG over FliM (**Fig. 9C**), it is questionable whether FlhG could detect changes in free or assembled FliM-molecules. On the contrary: C-Ring building blocks can assemble without assistance *in vitro* (**Fig. 4B**) and *in vivo* (**Fig. 5A**). FlhG is recruited to the nacent flagellar structure by FliM and can only then excert its regulatory action. Interaction assays in vitro showed that FlhG is released during C-Ring assembly (i.e. interaction of FliG and FliM, **Fig. 5B**), therefore a discrete localization of FlhG in proximity to the C-Ring appears unlikely. FlhG dimerizes in an ATP-dependent manner. This dimerization could facilitate flagellar building block assembly. While this mechanism cannot be ruled out, it is unlikely given that FliM stimulates FlhG‘s ATPase activity (**Fig. 3C**). FlhG can thus be expected to be monomeric in complex with FliM.

In the amphitrichous bacterium *Campylobacter jejuni*, a Δ*flhG*-strain shows a hyperflagellation phenotype. However, no interaction between FliM and FlhG could be detected (see appendix). Although *Cj*FliM contains the N-terminal EIDAL-motif (**Fig. 2D**), FliM-interacting residues at FlhG are not conserved. Therefore, there is no evidence that flagellar number is regulated via the interaction of FliM and FlhG in *C. jejuni*.

Truncation of the EIDAL-motif in *Sp*FliM led to a drastic loss of polar localization of FlhG. This strain phenocopies the hyperflagellated Δ*flhG*-strain. Therefore, the polar localization of FlhG is required for numerical regulation. In order to provide

further evidence that the interaction between FliM and FlhG is not solely responsible for numerical flagellar control, FlhG could be targeted to the pole via a different mechanism (i.e. construction of a FlhG-variant which possesses a HubP-interacting segment).

Hypothesis 2: FlhG is transcriptionally synchronized with other flagellar building blocks

Proteins of the flagellum are produced in a hierarchical manner by four transcriptional tiers.[6] Following this model, proteins of the C-Ring as well as FlhG are under the control of FlrA. Therefore, it might be possible that expression of FlhG is synchronized with the C-Ring proteins FliG, FliM and FliN. Thereby, FlhG is produced alongside flagellar building blocks FliM/FliN and possibly even forms a co-translational complex with the nascent N-terminus of FliM, which contains the EIDAL-motif at its N-terminus. Thus, nearly all of FlhG is engaged in complex with FliM. Progressive C-Ring assembly (and thus: interaction of FliG and FliM) leads to the release of FlhG, which is then free (given the overlappig binding site) to interact with FlrA (**Fig. 10A**). FlhG inactivates transcriptional activity of FlrA (**Fig. 9A**). Previously, it has been shown that FlhG inhibits the ATPase activity of FlrA.[43] Enhancer binding proteins such as FlrA form ATP-dependent hexamers. Cooperative ATP-hydrolysis of the hexamer-state provides mechanical force to drive transcriptional activation.[41] Therefore, the mechanism by which FlhG inactivates FlrA likely is by inhibiting ATP-hydrolysis of FlrA. FlhG binds to a linker region of FlrA which connects the DNA-binding HTH-motif and the ATPase-domain. In FlrA-homologs of the NtrC-family, this linker is described to be variable in sequence in mostly unstructured. The linker is thought to act as a flexible tether, which provides translational freedom for the hexamerization of the ATP-domain.[47] Therefore, binding of FlhG to this linker might inhibit FlrA-hexamerization. In summary, FlhG might regulate building block number by progressively inactivating FlrA upon successive C-Ring assembly.

However, experimental evidence provided above shows that this hypothesis is flawed: Primarily, FlhG is not transcriptionally synchronized with other flagellar proteins (**Fig. 9C**). Importantly, FlhG is not regulated on the transcriptional level

by FlrA (**Fig. 9B**). Furthermore, FlhG is overabundant compared to both FliM and FlrA (**Fig. 9C**). Therefore, a progressive inactivation of FlrA is not synchronized to C-Ring assembly. Furthermore, while keeping in mind that amount of incorporated C-Ring building blocks does not represent total amount produced,[52–54] the stoichiometry of flagellar building blocks in the flagellum do not provide strong evidence for transcriptional synchronization.

Hypothesis 3: Polar localization of FlhG continuously inhibits transcriptional activity of FlrA.

As mentioned above, the presence of FlhG in the cell alone is not sufficient for FlhG to act as the numerical regulator. Additionally, polar localization of FlhG is required. As demonstrated above, FlhG is not present in the the mature C-Ring (**Fig. 5A**) and its interaction with FliM is abrogated once FliM binds to FliG (**Fig. 5B**). Therefore, FlhG can only detect free levels of FliM, but not the stoichiometry of FliM attached to the nascent C-Ring.

The question is: What is the molecular mechanism which leads to the release of FlhG from FliM? Two scenarios are possible:

i) The N-terminus of FliM, which mediates interaction with FlhG, undergoes a conformational change upon binding of FliM to FliG. Subsequently, FlhG dissociates.

ii) Binding of FliM to FliG leads to a steric obstruction of helices a6 and a7 of FlhG, which mediate contact to FliM-N. Subsequently, FlhG dissociates.

Given the fact that CheY interacts with the EIDAL-motif of FliM[14] and can presumably do so while FliM is incorporated into the flagellar C-Ring, scenario i) appears unlikely.

Additionally, two indications provide evidence that numerical control is independent of FliM: 1) FlhG is present also in the absence of FlrA. Thus, FlhG inhibits FlrA even if no FliM is produced. (Here, additional experiments need to be performed to show that FliM indeed is produced by FlrA) 2) FlhG is overabundant compared to FliM and shows a localization which is not restricted to the pole (**Fig. 2A**). Therefore, FlhG might fulfill other functions which do not require targeting by FliM.

The overabundance of FlhG and the fact that FlhG is not produced by FlrA lead to an important question: How is transcriptional activation of FlrA possible at all when assuming that for every FlrA molecule, there are 2-3 molecules of FlhG? Firstly, FlhG only interacts with FlrA when bound to ATP. Secondly, a major difference between the FliM-FlhG and FlrA-FlhG interaction is the forced dissociation in the case of FlrA: Upon binding a FlhG-dimer, ATPase hydrolysis occurs and leads to a dissociation of the FlhG-FlrA-complex (**Fig. 6F**). This forced dissociation might enable FlrA to achieve hexamerization and transcriptional activation, which leads to the production of flagellar building blocks. Studies about the dynamics of C-Ring proteins FliM and FliN have shown that these proteins exist in two populations: one that is attached to the flagellar rotor; and a second which is freely diffusing.[52–54] Fluorescence anisotropy measurements also indicate that the amount of FliM/FliN bound to the flagellar rotor varies with the rotational direction.[55] These findings clearly show that an excess of C-Ring building blocks is produced. In the context of numerical regulation by FlhG, these results can be interpreted as follows: The continuous inhibition of FlrA by FlhG leads to a slow accumulation of flagellar building blocks. Once a treshold concentration is reached, the completion of flagellar assembly is achieved. Excess C-Ring proteins are produced, but the total level of freely diffusing flagellar components is maintained, perhaps due to increased degradation of freely diffusing flagellar building blocks (**Fig. 10B**).

The exact details of this proposed model require further attention. However, given the experimental evidence, hypothesis 3 represents the most likely case.

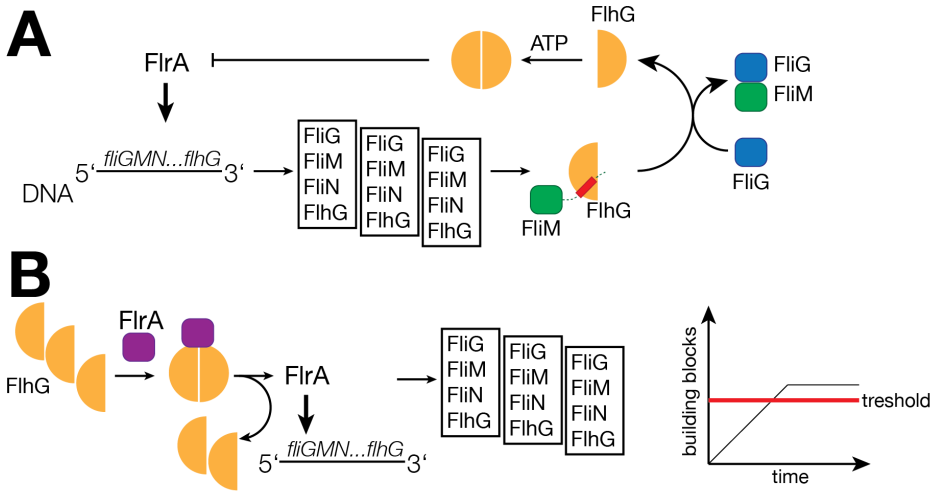

Figure 10. A: Schematic representation of hypothesis 2. FlrA interacts with DNA, which leads to the synchronized production of flagellar components. Subsequently, flagellar building blocks form (FlhG-FliM) and flagellar C-Ring assembly occurs. This in turn leads to the liberation of FlhG, which inactivates FlrA. **B**: Schematic representation of hypothesis 3. Overabundant FlhG continuously inactivates FlrA. Interaction between FlrA and FlhG and stimulation of the FlhG-dimer enable FlrA to bind DNA and effect flagellar component production. The concentration of building blocks increases until a critical treshold concnetration is reached and flagellar biogenesis is accomplished.

1.6 Summary and future perspectives

Summary points:

1. FlhG regulates flagellar number by inhibiting the transcriptional activator FlrA. FlhG binds FlrA in the linker region between ATPase-domain and DNA-binding domain, thereby likely preventing FlrA from forming the transcriptionally-active hexamer. FlhG does not inhibit DNA-binding of FlrA.

2. FlhG is overabundant in *S. putrefaciens* cells and is not under the control of FlrA.

3. FlhG localizes to the nascent flagellar structure in response to FliM, but is released upon FliM-FliG interaction.

4. FliM and FlrA share a binding site on FlhG. Both proteins stimulate the ATPase activity of FlhG. FlrA recognizes specifically the ATP-bound dimer of FlhG. The

interaction of FlrA with dimeric FlhG and the enzymatic stimulation thus lead to complex dissociation of FlrA-FlhG.

5. C-Ring building blocks can assemble without co-factors or chaperons. This is different from the situation of exported substrates, which have dedicated chaperones that prevent premature polymerization.[56]

Future issues:

1. The polar localization of FlhG is achieved by FliM. To gain a better understanding of this process, a control experiment should be performed, in which FlhG localizes at the pole, but is unable to interact with FliM.

2. The mechanism by which FlhG is released upon FliM-FliG interaction remains unclear. To rule out EIDAL-obstruction, the EIDAL motif could be placed somewhere else (C-terminus of FliM, N-terminus of FliN, etc). To rule out FliG as the releasing agent, the effect of truncations of FliG should be assessed by cryo-tomograpy. Transient contacts between FlhG and FliG could be detected by crosslinking.

3. The release of FlhG during C-Ring assembly should be assessed with regard to the membrane association of FlhG. To mimic the situation in vitro, a variant of FliG with an N-terminally fused MTS from FlhG is constructed (MTS-FliG). This construct should be tested with regard to its lipid association in flotation assays. Then, in flotation assays with liposomes the fate of FlhG should be assessed in the context of FliM/FliN (FlhG-FliM/FliN) and MTS-FliG with/without ADP/AMPPNP.

4. The MTS of FlhG can localize GFP to the membrane.(Schuhmacher 2015) In the context of the ATPase FlhG, the MTS is not sufficient to enable membrane binding of FlhG (**Fig. 2A**). The regulation of the MTS by the ATPase domain should therefore be investigated. To do so, the ATPase activity of FlhG should be assessed in the absence and presence of liposomes. Additionally, flotation assays with liposomes should be performed with wild-type FlhG and a variant of FlhG in which the MTS is inactivated (FlhG F275A F276A). The flotations should be performed with and without liposomes and with and without AMPPNP.

5. The fact that FlhG is present irrespective of FlrA, RpoN or FliA leads to the question: what other roles does FlhG play in *S. putrefaciens*? Additional interaction partners of FlhG could be identified by co-immunoprecipitation or UV-crosslinking.

6. The transcriptional activation of FlrA leads to the production of flagellar building blocks (i.e. FliF, FliG, FliM). The experiments above indicate that the inhibitory action of FlhG permit the production of just enough building blocks to assemble one flagellum, but not enough for two flagella. To challenge this theory, the flagellar building blocks FliF, FliG or FliM could be knocked out in the genome and be placed under the control of a tuneable promoter.

7. In TCS, the response regulator is phosphorylated by its cognate sensor kinase at a conserved aspartate-residue in its N-terminal domain. In *Sp*FlrA, this residue is conserved. However, no cognate sensor kinase is known. It should be investigated whether FlrA can be phosphorylated, whether phosphorylation impacts the interaction network of the N-terminus of FlrA (no interaction partner known so far, **Fig. 1C**). In case FlrA can be phosphorylated, the cognate sensor kinase should be identified.

8. The aim of this work was a full quantitative description of the FlhG interactome. This data should now be used for the construction of a mathematical model of flagellar assembly. Such a model could help to investigate missing factors or identify parameters which are crucial for the numerical regulation of flagellar assembly (This work is currently being undertaken in the group of Prof. Frey, Munich).

2 RemA features a histone-like mode of DNA interaction

2.1 Introduction

2.1.1 A small subclass of transcriptional regulators binds DNA with a fold different from HTH

In prokaryotes, a large majority of DNA-binding proteins interacts with DNA via a helical motif termed helix-turn-helix (HTH).[57] A well known exception to this pattern is the *E. coli* methionine repressor,[58] which inserts β-strands into the major groove of DNA and belongs to the superfamily of ribbon-helix-helix DNA-binding proteins.[59] Bacterial histone-like proteins (e.g. HU, N-HS, IHF) also bind DNA via β-strands.[60] Apart from transcriptional response regulators that bind DNA via a helix-turn-helix motif (see above), there is a smaller subgroup which interacts with DNA via a conserved domain named LytTR. A recent database-search estimated that of 5589 DNA-binding response regulators, only 4.7% are predicted to bind DNA via a LytTR motif.[61] An analysis about the distribution of LytTR-regulatory systems (LRS) showed that the majority of LRS are encoded within the Firmicutes phylum.[62] Only recently the first structure of a LytTR domain was solved.[63] The C-terminal DNA-binding domain of the *S. aureus* response regulator AgrA, which regulates quorum sensing and is the global regulator of virulence,[64] shows a novel topology, where 10 β-strands are arranged into 3 antiparallel β-sheets which form a β- β- β sandwich. DNA binding by AgrA-C occurs via loops which insert into successive major grooves. AgrA-C interacts with DNA both nonspecifically (10 contacts) and base-specifically (2 contacts). DNA-footprinting with full-length AgrA showed that the protein recognizes two 9 bp direct repeats separated by a 12 bp spacer on the promoter region P2 in the RNAIII-*agr* intergenic region.[65] A second interaction pattern, albeit with lower affinity, is observed ca 40 bp upstream in the P3 promoter region. The co-crystal structure of AgrA-C complexed to a 15 bp minimal DNA recognition sequence shows AgrA-C as a monomer. However, given the close proximity of the direct recognition repeats, AgrA could form dimers via its N-terminal domain or engange in higher oligomeric states, given that a total number of 4 DNA recognition motifs have been observed (perhaps like in the case of SinR, see below). The binding of AgrA-C to DNA

induces a DNA-bending of 38°, which could further hint at an oligomeric assembly of AgrA. The molecular mechanism by which AgrA induces transcription is not yet understood. Another transcriptional activator with the LytTR-fold, AlgR, binds well upstream of the promoter region and also bends DNA.[66,67] AgrA alone can promote transcription from the *agr* P2 promoter, and its effect is attenuated by two accessory proteins, SarA and SarR, which bind DNA close to AgrA.[68] The DNA-binding ability of proteins like AgrA might be further influenced by small molecules.[69]

The protein AgrA-C is structurally similar to archaeal Sac7d, which non-specifically binds DNA by inserting hydrophobic sidechains into its minor grooves and thereby increasing the DNA‘s melting temperature. Sac7d (PDB: 1AZP) and a related protein (Sso7d) are chromosomal proteins found in the hyperthermophilic archaeal species *Sulfolobus acidocaldarius* (or *S. solfactaricus*). Sac7d, being a very small protein (7 kDa) and possessing the OB-fold (oligonucleotide-binding), itself is extremely stable to heat, acid and chemical agents and has recently found application as highly stable artificial proteins that bind antigens with high-affinity (»Affitin«).[70] The relevance of AgrA-C and other proteins of the LytTR-family lies in the fact that their mode of binding DNA occurs via a completely new mechanism (not helices, not b-strands, but loops). Undiscovered DNA-interacting proteins might be found by applying secondary-structure-guided search algorithms.

2.1.2 RemA is the master regulator for biofilm formation and salt-stress response

B. subtilis lifestyle can change between planktonic and sessile in response to environmental stimuli (sporulation is not considered here, as it represents a largely metabolically dormant state, **Fig. 11A**). When cells enter the sessile (biofilm) lifestyle, flagellar motility is inhibited both at the functional level (flagellar rotation is prevented by EpsE, which decouples motor components from the C-Ring of the flagellum)[71,72] and at the transcriptional level.[73] Cells that enter the sessile state express extracellular matrix components, exopolysaccharide (EPS), TasA amyloid proteins and the film-protein BslA,[74] among others. Under planktonic conditions, their expression is repressed by the protein SinR.

SinR, commonly known as the master regulator of biofilm formation in *B. subtilis*, is a pleiotropic DNA binding protein which was first described as the repressor of the sporulation controller Spo0A.[75] More recently, Spo0A has been identified as an activator of SinR.[76,77] SinR is a 14 kDa-protein comprised of two domains; a N-terminal DNA binding domain of the HTH-type and a C-terminal domain which mediates oligomerization. SinR represses transcription from the *tasA*-operon by looping the DNA upstream of the promoter.[78] The tetrameric arrangement of SinR on the DNA is broken by its antagonist, SinI. SinI binds to the C-terminus of SinR, which mediates oligomerization, and enforces hetero-dimerization, causing deprepression.[79]

While the inhibitory effect of SinR is relieved by SinI, SinR also occludes binding of the **r**egulator of **e**xtracellular **m**atrix **A**, RemA. RemA is a 89-amino acid protein and highly conserved in the bacterial domain. RemA has been found to compete with SinR for DNA binding upstream of the *epsA* operon (**Fig. 11B**). In contrast to SinR, which loops DNA and inhibits transcription, RemA is thought to also bend DNA but acts as an activator of biofilm formation.[80] Recently, RemA has been identified as a DNA-binding protein that binds to multiple sites upstream of the *tasA* and *epsA*-promoters. Notably, RemA-binding sites overlap with those of SinR.[81] It has also been found that SinR-derepression alone is not sufficient for biofilm formation, since *sinR-remA*-double mutants showed defects in pellicle formation.[80] Apart from its role as a global activator of biofilm formation, RemA is also implicated in osmostress response. Here, RemA positively regulates the transcription of *opuA*, *opuB* and *opuC*, which encode importers for compatible solutes such as choline, proline betaine and glycin betaine.[82] Notably, RemA shows a similar DNA-binding pattern upstream of the *opuA* promoter as in the case of *epsA* (Hoffmann and Bremer, unpublished). Biofilm-formation and osmostress-protection are interlinked, as the biogenesis of the extracellular matrix leads to a high osmotic pressure. The »osmostress«-importers protect the cell from high osmotic pressure, as the import of compatible solutes (*de-novo* synthesis of compatible solutes represents an alternative pathway) prevents water efflux, maintains turgor and stabilizes macromolecules. For this reason, promoters of *opuA*, *opuB* and *opuC* are also osmotically inducible.[82]

RemA is transcribed from its own promoter and also co-transcribed from the promoter of the gene *yloC* immediately upstream.[83] The genetic cluster of *remA* contains three genes: the aforementioned *yloC*, which encodes a protein of unknown function. RemA is co-transcribed from the *yloC*-promoter.[80] Immediately downstream of *remA* is the gene encoding guanylate kinase (GMK), which is an essential enzyme and responsible for purine metabolism.[84] GMK is implicated in the stringent response, because the »alarmone« (p)ppGpp binds GMK and inhibits GDP production, which helps bacteria to adapt to amino acid starvation.[85] Downstream of *gmk* lies the gene *rpoZ*, which encodes the omega-subunit of the RNA polymerase (**Fig. 11C**).[86] The »sister«-protein, named RemB, which also positively regulates biofilm formation, is located at another position in the genome.[80]

The mode by which RemA interacts with DNA is still unclear, and its functional relevance remains to be investigated. RemA could be the factor that transcriptionally interlinks osmotic stress response and biofilm formation, as biofilm-formation leads to osmotic pressure gradients inside the biofilm.[87,88]

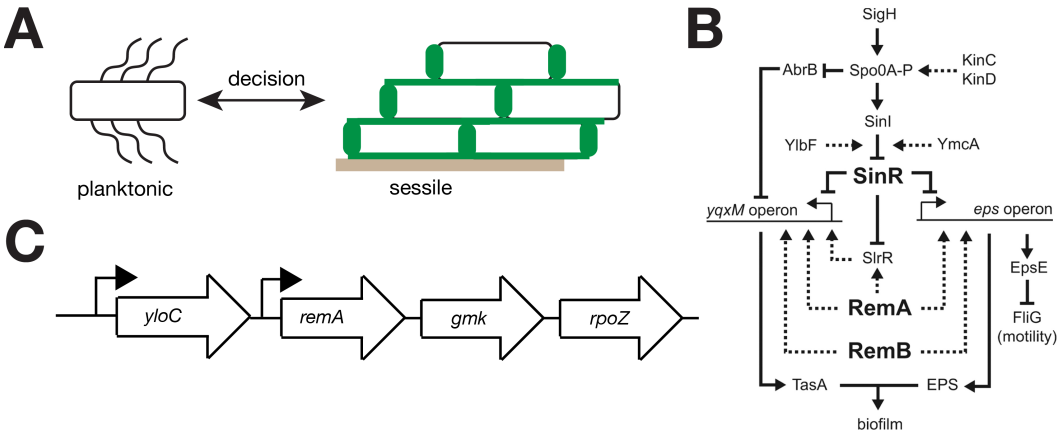

Figure 11. A: Cells from *B. subtilis* and other species can change their lifestyle in response to certain signals from planktonic (motile) to sessile (biofilm, secreted components e.g. TasA and EPS are indicated in green). **B**: Regulatory network of SinR, RemA, RemB. Reproduced with permission.[80] **C**: Operon architecture of RemA.

2.2 Aim of this work

The processes of biofilm-formation and osmoprotection are regulated at the transcriptional level by the protein RemA. RemA acts as a transcriptional activator and binds individual repeats on the DNA. However, RemA does not possess a classical DNA-interacting motif. No binding partners or structural information is available for RemA.

In this work I set out to elucidate the structure of RemA. The structural information should help to understand the DNA-binding mode of RemA. Furthermore, the interaction of RemA with other proteins, such as SinR and proteins encoded in the operon of RemA, should be attempted.

2.3 Results

2.3.1 The crystal structure of RemA shows a donut-shaped octamer

To gain a deeper understanding into the regulatory activity of RemA, the crystal structure was determined. As proteins from thermophilic organisms are often better suited for structural studies, RemA from the moderate thermophilic organisms *Geobacillus thermodenitrificans* (Gt) was chosen (**Fig. 12A**). *Gt*RemA was amplified with a N-terminal hexahistidine-tag, cloned into pET24d and was produced in *E. coli* BL21(DE3) and purified by Ni-ion affinity and size exclusion chromatography (SEC). A high NaCl-concentration of 1 M was required in all buffers to maintain solubility of *Gt*RemA (pI=6.72).

The so-prepared RemA crystallized within 24 hours and its structure was determined at 2.3 Å resolution by single-wavelength anomalous diffraction (SAD). The crystals belong to space group P23 with four RemA molecules per asymmetric unit and a solvent content of 51%. Residues 2 to 83 of *Gt*RemA could be unambiguously assigned within the electron density map, while the remaining 4 residues at its C-terminus were only poorly visible likely due to flexibility (**Fig. 12A**). The overall structure of RemA shows a wedge-shaped domain with a β-β-α fold, consisting of six β-strands and two α-helices (**Fig. 12C**). The central core of six beta-strands has a topology of 1↑2↓3↑5↓6↑4↑. The N- and C-termini of the monomer are in close proximity with a distance of only 5.6 Å (in a molecule comprising residues 2-78), this conformation is stabilized by a backbone

interaction (M3-L77, 2.6 Å). Major intramolecular interactions of *Gt*RemA include the two β-sheets formed by β-strands β1 and β2, as well as β3, β4, β5 and β6. Residues from these β-sheets are mostly hydrophobic and pack inside the molecule. The two β-sheets are flanked on the top by helix α1, which projects hydrophobic residues towards the beta-sheet formed by β3, β4, β5 and β6. Helix α1 is oriented by K31, which forms a salt-bridge to backbone carbonyl of D26 (2.3 Å), located between β3 and α1. On the lower left side of the protein, helix α2 packs against the core formed by both β-sheets. Two residues of α2 orient the helix; Q69 interacts with R53 (2.5 Å) and R76 contacts the backbone-carbonyl of N7 (2.8 Å). The long loop between β4 and β5, which is fully resolved in the electron density, contains three arginine residues (R50, R51 and R53). These complex SO_4-anions, which have been placed into blobs of electron density and likely origin from the crystallization condition. Sulphate-anions from the crystallization conditions are not present at every monomer due to crystal packing. Data collection and refinement statistics are listed in **table 1**.

Table 1: Data collection and refinement statistics of *Gt*RemA. Values in parenthesis refer to the highest resolution shell. For R_{free} calculation, 5% of the total reflections from the working set were used.

	Gt*RemA*
Wavelength	0.966
Resolution range	35.64 - 2.301 (2.384 - 2.301)
Space group	P 2 3
Unit cell	106.91 106.91 106.91 90 90 90
Total reflections	680824 (44272)
Unique reflections	18378 (1809)
Multiplicity	37.0 (24.5)
Completeness (%)	99.86 (100.00)
Mean I/sigma(I)	19.73 (1.35)
Wilson B-factor	41.15
R-merge	0.1939 (3.204)
R-meas	0.1966 (3.27)
R-pim	0.03207 (0.6506)
CC1/2	0.999 (0.617)
CC*	1 (0.874)
Reflections used in refinement	18361 (1809)
Reflections used for R-free	984 (106)
R-work	0.1935 (0.2697)
R-free	0.2305 (0.2940)
CC(work)	0.957 (0.794)
CC(free)	0.922 (0.710)
Number of non-hydrogen atoms	2530
macromolecules	2425
ligands	15
solvent	90
Protein residues	314
RMS(bonds)	0.003
RMS(angles)	0.65
Ramachandran favored (%)	98.04
Ramachandran allowed (%)	1.96
Ramachandran outliers (%)	0.00
Rotamer outliers (%)	1.49
Clashscore	2.82
Average B-factor	47.37
macromolecules	47.18
ligands	70.96
solvent	48.75

Analysis of the asymmetric unit shows that four RemA monomers are arranged into a tetramer reminiscent of a semicircle. Following the crystallographic 2-fold rotation axis, the other half of the semicircle is found in the neighboring unit cell (**Fig. 12C**). The four copies of RemA in the asymmetric unit superimpose with RMSD of 0.125-0.188 Å (470 or 473 atoms, respectively). In the octamer, the total solvent area buried upon oligomerization is 1868 $Å^2$ (37%) per monomer. The RemA-octamer is ring-shaped with dimension of 71.8 Å and 34.2 Å and with a central hole of 20.1 Å in diameter at the smallest constriction (**Fig. 12C**). Each of the RemA monomers primarily interacts with its two neighbors through β-sheet argumentation involving β2 (M13 to 15) and β3‘ (I20 to I22). Furthermore, polar contacts on top of the ring (R32-E39‘, 2.9 Å) and inside the ring (R18-A17‘, 2.6 Å and R18-D59‘, 2.5 Å) mediate oligomerization (**Fig. 12D**).

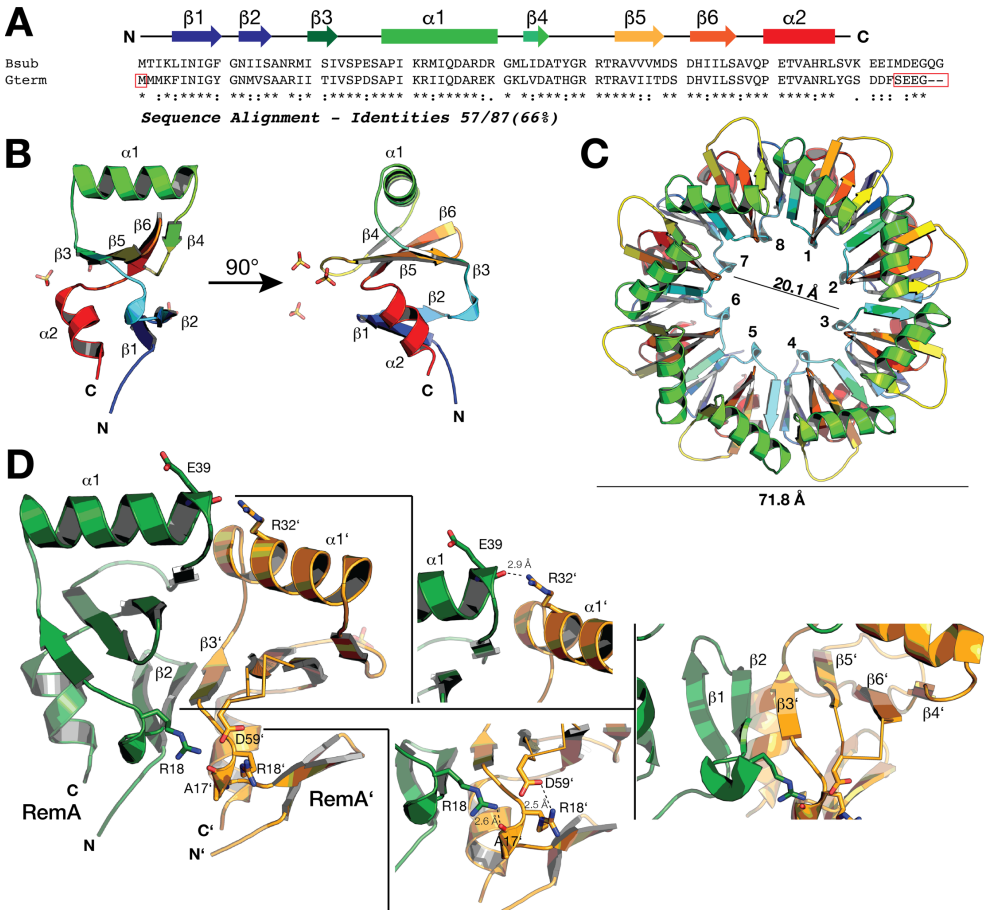

Figure 12. A: Sequence alignment of RemA from *B. subtilis* and *G. thermodenitrificans* NG-80 with secondary structural elements aligned to the sequences. Parts of the sequence which are not resolved in the model are boxed

in red. **B**: Crystal structure of RemA. Sulphate-ions from the crystallization condition have been placed in the electron density. These are complexed by arginine residues located in the loop between β4 and β5. **C**: Oligomeric state of RemA. 8 copies of RemA from two neighboring unit cells form a ring-like octamer. **D**: Overview of major contacts between adjacent monomers of RemA (left panel), polar contacts R32-E39‘ (middle panel, top), R18-A17‘ and R18-D59‘ (middle panel, bottom) and β-sheet complementation β2-β3‘ (right panel).

2.3.2 RemA-octamers further associate to hexadecamers

In the crystal lattice, RemA-octamers are arranged in a defined pattern (**Fig. 13A**) and two octamers always stack together via their helical face to form hexadecamers (**Fig. 13B**). The interaction between the octamers is mediated by polar residues of opposite charge which are located in helix α1 (**Fig. 13C**). Residues R32 and E39, which also link neighboring monomers (see above), stack on top of each other. The helical face of the octameric ring shows a corresponding, alternating distribution of positive and negative charges (**Fig. 13D**). Electron microscopic images of GtRemA show a heterogenic distribution of particles with an average size of 7.63 ± 1.02 nm (23 particles considered) (**Fig. 13E**, insets).

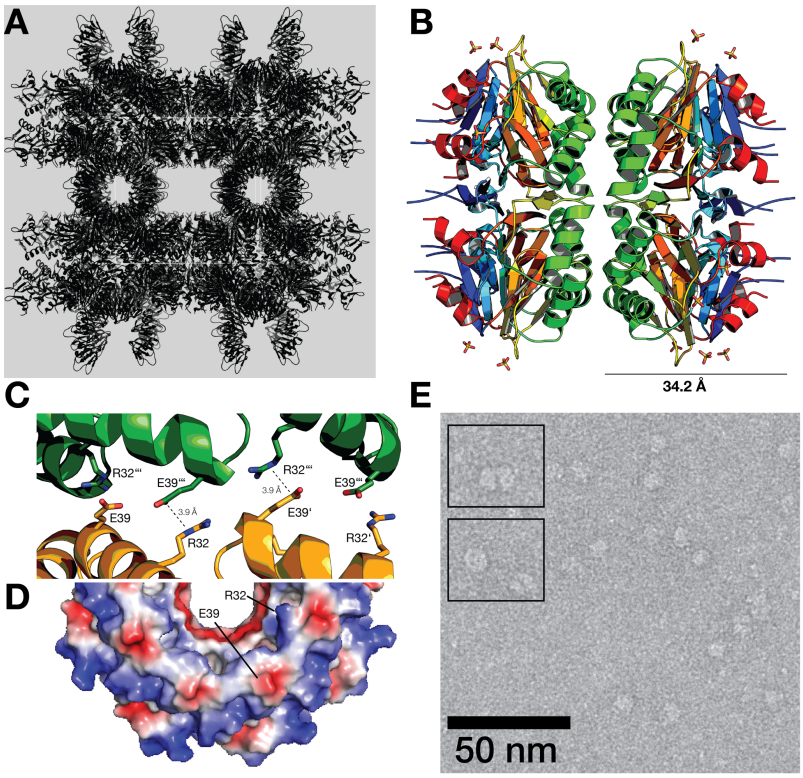

Figure 13. A: Crystal lattice of wildtype RemA (image prepared with Chimera). **B**: The RemA-hexadecamer shown from the side. The octamers stack together with their helical faces, the N- and C-termini point outward and the flanks of the rings are fully solvent accessible. **C**: Interface between the two octamers with residues R32 and E59 shown as sticks. **D**: Electrostatic surface of the helical face. **E**: Negative-stain electron microscopic image of GtRemA. Insets show examples of particles that are reminiscent of a hexadecamer (top) and octamer (bottom). The electron microscopic image was prepared by Dr. Thomas Heimerl, Marburg.

2.3.3 Homologues of RemA show a high degree of conservation over the whole sequence

Alignment of 187 reference proteomes against the primary sequence of RemA from *G. thermodenitrificans* shows that the sequences share a high degree of conservation. The projection of the sequence alignment onto the structure of RemA shows that the outward pointing face (putative DNA-interface) as well as the helical face (hexadecamer-interface) are nearly 100% conserved. The N- and C-termini as well as the β-sheet core and the inner opening show more variability (**Fig. 14A**).

2.3.4 RemA hexadecamers and octamers are in dynamic equilibrium with one another

Analysis of *Gt*RemA by size-exclusion chromatography (SEC) coupled to multi-angle light scattering (MALS) suggests that the protein forms two species in solution with determined molecular weights of 123 and 75 kDa (**Fig. 14B**). Based on the crystallographic analysis, the two species likely are hexadecamer and octamer. Notably, the higher-order arrangement of RemA is observed in a buffer containing 1000 mM NaCl. The observed molecular weight of 75 kDa is close to the calculated molecular weight of a RemA-octamer (84 kDa). The observed molecular weight of 123 kDa differs from the calculated molecular weight of a RemA-hexadecamer (168 kDa). The deviation (11% for the octamer, 27% for the hexadecamer) could arise from the extremely compact shape of the RemA-ring or double-ring. Furthermore, the observed molecular weight was calculated by integrating the UV-absorption signal, which might yield non-optimal results given the small extinction coefficient of RemA. The ratio between the two species (16mer:8mer) is 0.58 ± 0.03 or 0.49 ± 0.03 in SEC-buffers containing 1 M NaCl or 1 M KCl (over a range of concentrations 1-25 mg/mL), respectively (**Fig. 14C**). These measurements indicate that the two states of *Gt*RemA exist in a dynamic equilibrium with one another. As RemA is a global activator for osmostress-response, the impact of small molecules on the 16mer:8mer-ratio was also investigated. To do so, a fixed concentration of *Gt*RemA (2.3 mg/mL) was incubated for 60 minutes at room temperature either without or with a defined concentration of sodium-glutamate (GluNa, f.c. 100 mM), proline (Pro, f.c. 500 mM), or glycin betaine (GB, f.c. 500 mM). Of the three molecules tested, proline showed the strongest effect and lowered the 16mer:8mer-ratio to 0.39 ± 0.02 (1 M NaCl-buffer) or 0.34 ± 0.01 (1 M KCl-buffer), respectively. The effects of glutamate and glycine betaine on the 16mer:8mer-ratio were not as pronounced (**Fig. 14C**, right panel).

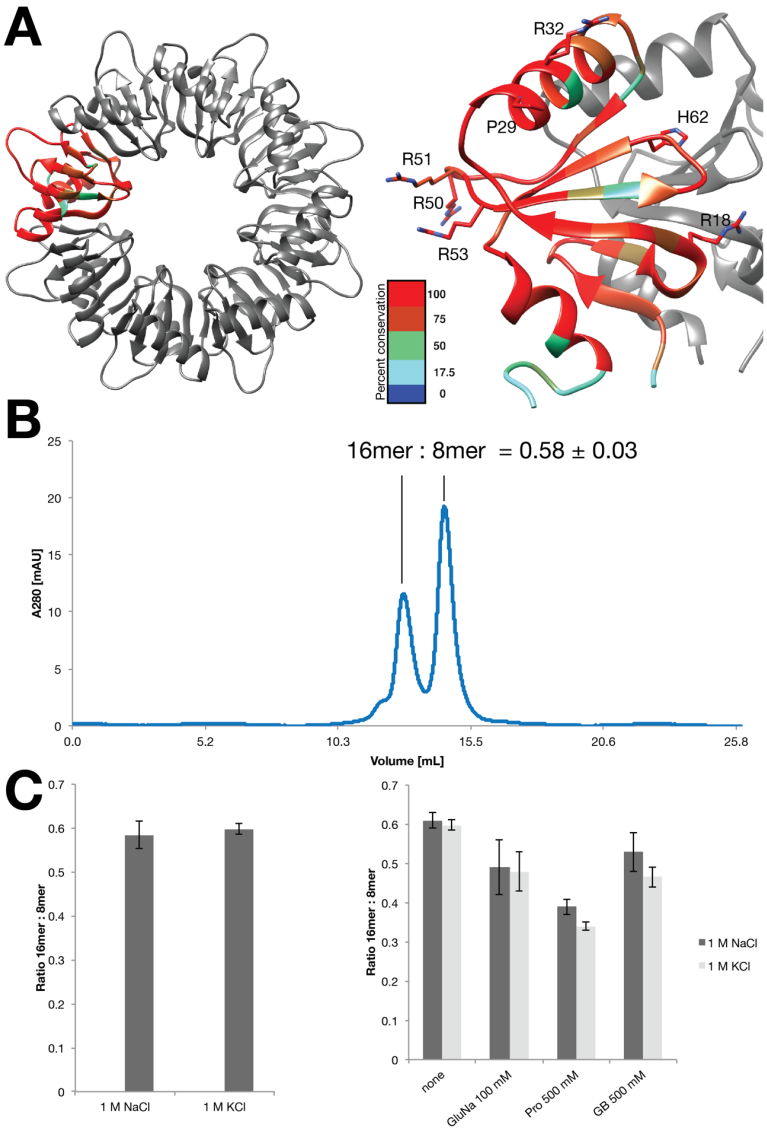

Figure 14. A: Sequence alignment of RemA against 187 proteomes projected onto the structural model. Amino acids on the outer rim and top of the molecule show a high degree of conservation, while amino acids at the structural core and at the N- and C-termini are less conserved (right panel). **B**: Size-exclusion chromatogram (SEC, Superdex200 10/300 GL with SEC-buffer including 1 M NaCl) of wildtype GtRemA reveals that octameric RemA further associates into hexadecamers in solution. Peak 1 was assigned to the RemA 16mer (MW_{obs}=123 kDa, MW_{calc}=168 kDa), Peak 2 was assigned to the RemA octamer (MW_{obs}=75 kDa, MW_{calc}=84 kDa). **C**: Diagram of 16mer:8mer ratios obtained from integration of peak areas observed for wild-type GtRemA in buffers containing 1 M NaCl or 1 M KCl (left panel). Different concentrations of GtRemA were employed (1-25

mg/mL) and all runs were performed in duplicates. Right panel: Diagram of 16mer:8mer ratios obtained from integration of peak areas observed for GtRemA wt (2.3 mg/mL) in buffers containing 1 M NaCl or 1 M KCl in the absence or presence of sodium-glutamate (GluNa, f.c. 100 mM), proline (Pro, f.c. 500 mM) or glycin betaine (GB, f.c. 500 mM). All runs were performed in duplicates.

2.3.5 Multiple binding sites form a continuous DNA interaction interface on RemA

RemA has been shown to bind successive repeats 50 bp upstream of the *epsA*-promoter. In footprinting assays, a MBP-*Bs*RemA-construct protected six DNA-regions of 7 bp length with 3 bp-spacers in between. The spacer regions show DNA hyperdigestions, indicating DNA-bending.[89,90] At high concentrations, additional binding sites closer to the promoter were observed. The six DNA-sites do not contain a consensus sequence.[81]

To assess the ability of purified *Gt*RemA to interact with DNA, electromobility shift assay (EMSA) was employed. *Gt*RemA showed a continuous, concentration-dependent shift of target DNA (*opuC*-promoter region, 224 bp) in the EMSA (**Fig. 15A**, left panel). In a control experiment, *Gt*RemA also shifted a non-related PCR-product (**Fig. 15A**, right panel), indicating that *Gt*RemA interacts nonspecifically with DNA. To rationalize the DNA-binding capacity of *Gt*RemA, its electrostatic surface properties were analyzed. The RemA octamer shows an extended positively charged surface area on its inner surface and outer rim. The following residues contribute to the positively charged surface: R18, which links monomers together; K4; and the N-terminus (**Fig. 15B**, left panel). The outside of the ring is also predominantly positive due to 3 arginine residues located in the β4-β5-loop (i.e. R50, R51, R53, **Fig. 15B**, right panel). This surface is covered by several sulfate ions originating from the crystallization buffer (see above). The arginine residues on the outer rim are 24-30 Å apart (**Fig. 15C**, left panel). Another octameric protein-DNA-complex is the nucleosome core particle of chromatin (PDB: 1AOI).[91] The curvature of the chromatin-DNA fits the RemA-octamer and the distance between major or minor clefts is 26-30 Å, which mirrors the arginine-distances of RemA (**Fig. 15C**, right panel). Taken together, the biochemical

assays verify that RemA nonspecifically binds to DNA. This interaction likely is mediated by positively charged residues on the outside of the ring. The shape of RemA could support a highly bent DNA structure, because putative DNA-binding residues in the octamer are spaced so that they would match with a double-stranded, bent DNA.

2.3.6 Towards the visualization of RemA-DNA interaction by Atomic force microscopy

In order to visualize the DNA-interaction of RemA and to verify whether RemA bends the DNA, atomic force microscopy (AFM) was employed. AFM of DNA-protein complexes is a sensitive and robust method and allows for the determination of protein-induced DNA-bending.[67,92] Duplex DNA of different lengths (*opuA*-promoter fragment 968 bp, 186 bp) was readily visualized by AFM. In the case of short DNA fragments, extensive linking was observed (**Fig. 15D**, middle panel). This might be due to the high temperatures employed during electro-elution from the agarose slice and circumvented by purifying the PCR-reaction via spin-columns. When incubating RemA together with long or short DNA-fragments in varying ratios, no co-localization of RemA with the *opuA*-promoter fragment was observed. The deposition efficiency was decreased and the mica surface had to be scanned extensively in order to detect DNA or protein. DNA doublestrands (average height 0.8 nm) were visible as well as blobs that presumably were RemA complexes (average height 3.5 nm) (**Fig. 15E**, right panel).

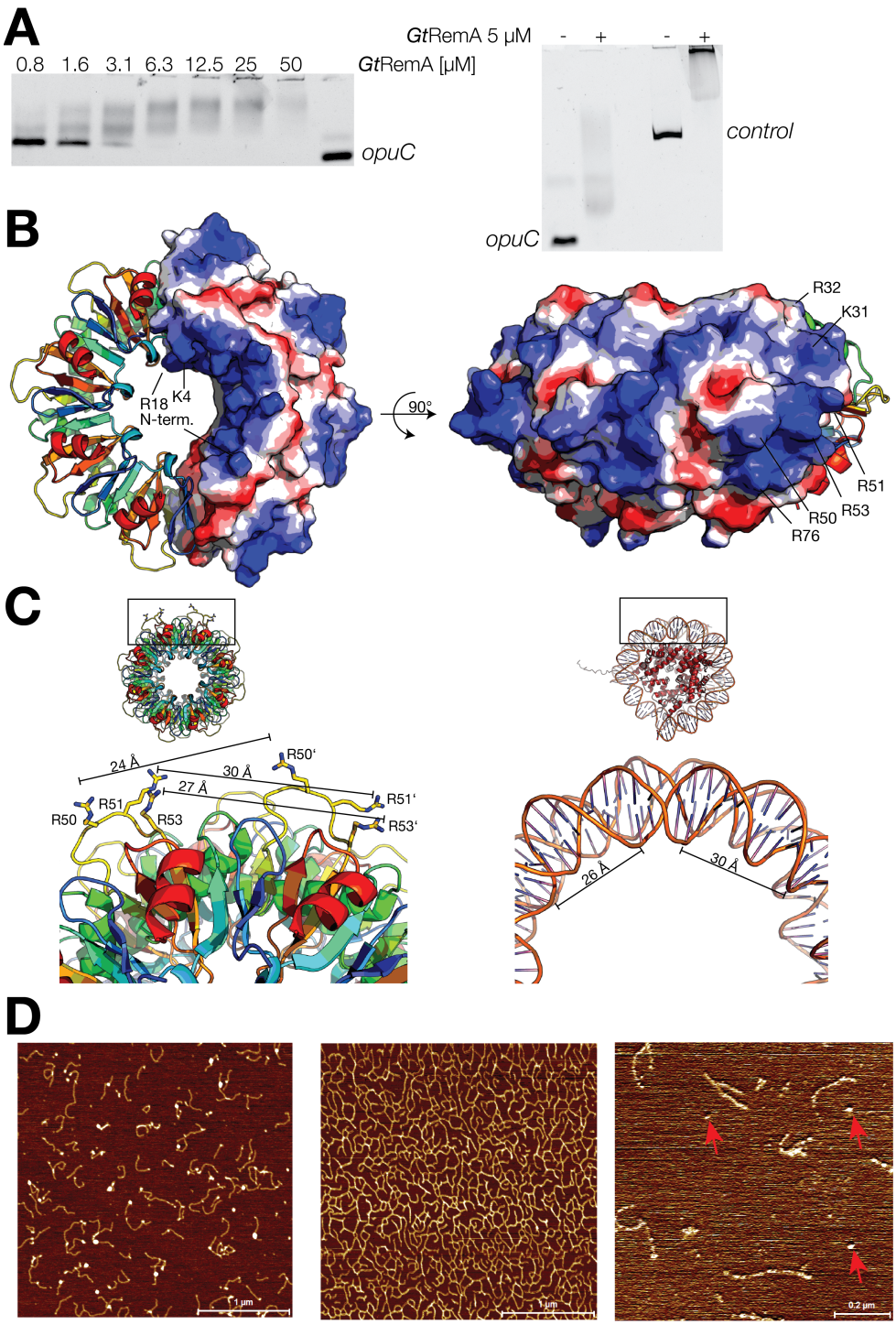

Figure 15. A: Electromobility shift assay (EMSA) of wildtype *Gt*RemA. *Gt*RemA shifts the *opuC*-promoter region (224 bp) with increasing concentration (left panel). *Gt*RemA shifts also control DNA (PCR-product of *fliM*-gene from *S. putrefaciens*, right panel). **B**: Electrostatic surface of *Gt*RemA seen from the bottom (non-helical face, left panel) and side (right panel). **C**: The distance

between outward-facing arginine residues of neighboring RemA-octamer-subunits (left panel) corresponds to the spacing of large clefts in looped DNA (PDB: 1AOI, right panel). **D**: AFM-images of long (968 bp, left panel) and short (186 bp, middle panel) *opuA*-promoter region fragments deposited on mica. RemA incubated with long *opuA*-DNA and deposited on mica did not lead to co-localization of protein and DNA (right panel, putative RemA-particles are indicated with arrows). The DNA was prepared and purified by Dr. Tamara Hoffmann. The AFM-images were prepared by Darius Günder (Prof. Witte, Department of Physics, University Marburg).

2.3.7 RemA represents the minimal motif of the LytTR fold

A primary sequence search of RemA against the protein data bank (PDB) did not yield any results. However, a structural homology search using the DALI server[93] revealed that RemA is similar to the C-terminal, DNA-binding domain of *S. aureus* AgrA, the global regulator of virulence.[63] The authors of this study already noticed that AgrA-C (PDB: 3BS1) exhibits a two-fold symmetry among its 10 β-strands, and argued that the LytTR-fold, of which AgrA-C is the first structurally solved example, may be derived from duplication of a minimal unit (**Fig. 16B**, right panel). RemA represents this minimal unit, as a superimposition of two RemA monomers and AgrA-C shows (**Fig. 16A**). A dimer of RemA superimposes with AgrA-C with a RMS of 1.379 Å (175 atoms). AgrA-C contains all secondary structural elements of RemA except for α1 and β4 (numbering according to *Gt*RemA, **Fig. 16B**). In stark contrast to AgrA-C, β4 and α1 form a head for the RemA-ring, and two important residues that mediate multimerization are located in α1 (see above). The curvature of two RemA-monomers is matched by AgrA-C, however AgrA-C is monomeric and thus bends the DNA to a lesser degree than RemA putatively does. AgrA is proposed to assemble into dimers via its N-terminal domain.[63,65] While AgrA-C possesses a duplicated RemA-fold, its primary sequence does not show a duplication of a certain motif or sequence. Residues of AgrA-C which interact specifically with DNA are H169, N201 and R233, all of which are located in loops on the DNA-binding face of the protein. The putative DNA-interacting residues of RemA (R50, R51, R53) are located in a similar region of the protein

(**Fig. 16C**). The linkage between RemA-monomers on the inside of the ring is achieved by arginine R18, which is interlinked with aspartate D59 (**Fig. 16D**, left panel). In AgrA-C, several residues are involved in this linkage: D157-H208 (2.7 Å), D157-N206 (3.0 Å), D157-N209 (2.9 Å), I159-N209 (2.9 Å) (**Fig. 16D**, right panel). Taken together, RemA represents the minimal motif of the LytTR-fold.

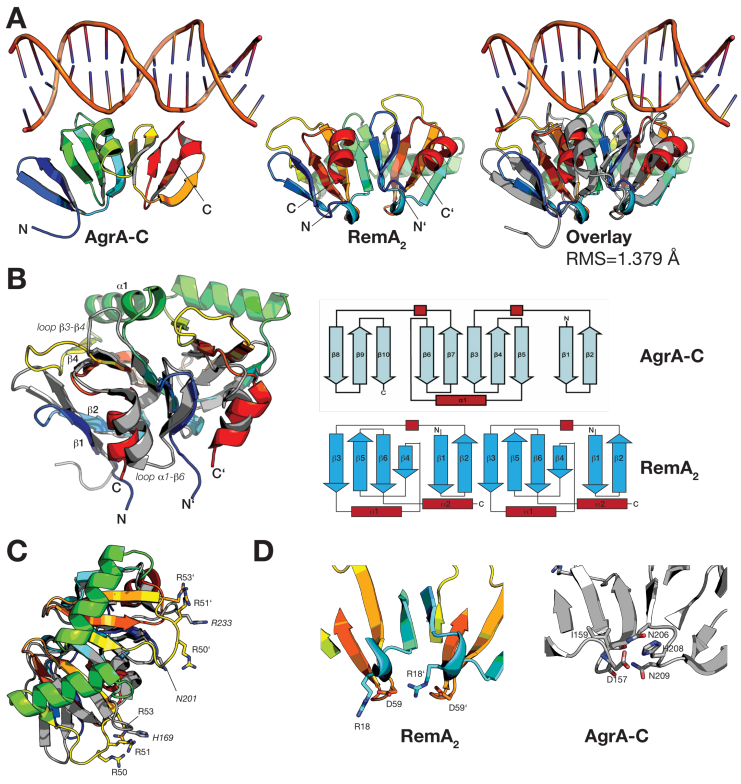

Figure 16. A: Structure of *Sa*AgrA-C complexed to target DNA (PDB: 3BS1) (left panel); Structure of the *Gt*RemA-dimer (middle); superimposition of AgrA-C with the RemA-dimer (right panel). For clarity, AgrA-C is coloured in grey. **B**: Structural differences between AgrA-C and the RemA-dimer (left panel): AgrA-C contains a duplicated RemA-fold. For this reason, the loop between α1 and β6 (labeled in italics) in AgrA-C is not present in RemA. Correspondingly, the C- and N-termini of neighboring RemA-monomers are in close proximity (10 Å). Furthermore, AgrA-C does not possess helix α1 and β-strand β4 of RemA, which are replaced by loop β3-β4 (labeled in italics). The β-β-β-fold of AgrA-C is mirrored by $RemA_2$ (right panel). **C**: Superimposition of AgrA-C and the RemA-dimer. DNA-binding residues of AgrA-C are shown as sticks and labeled in italics. These residues are located

in similar regions where putative DNA-binding arginines of RemA are found. **D**: Polar contacts between RemA-dimer on the inside of the ring (left panel) and corresponding polar contacts of AgrA-C (right panel).

2.3.8 Mutants of RemA maintain a ring-shaped oligomeric structure

In order to challenge the biochemical understanding and to evaluate the structural basis for defects of RemA observed *in vivo*, several mutants of *Gt*RemA were constructed (mutated residues are conserved between *B. subtilis* and *G. thermodenitrificans*, see **Fig. 12A**). The mutants were cloned and purified as the wild-type protein (see above). A list of all mutants and their crystallographic data can be found in the appendix. All mutants of *Gt*RemA that yielded diffracting crystals showed an octameric arrangement of monomers reminiscent of the wild type protein (**Fig. 17A**). *Gt*RemA R18W is an exception: In this case, the octameric ring collapsed into a heptameric arrangement of monomers. The change in higher order arrangement for the mutant R18W leads to a decrease of ring diameter (63.2 vs 71.8 Å in the wildype) and a corresponding increase in spacing between arginine residues (R50, R51, R53: 28-35 Å versus 24-30 Å in the wildtype). In all cases, the octameric (or heptameric for R18W) assemblies further associate into hexadecameric (or tetradecameric for R18W) as evident from the crystal lattice. The mutations induce global changes in the RemA-ring. To quantify these, a reference system was established by choosing the C-alphas of R50 as well as the center of mass of the protein-complex. Then, diameters and angles were determined. Contact surface area between monomers was calculated by the jsPISA-server (http://www.ccp4.ac.uk/pisa/). The mutations affect not only diameters, angles and contact interfaces (**Fig. 17A**). Superimposition of mutant monomers with the wild-type show that the overall-RMSD is small. However, when superimposing defined monomers of the octameric arrangements, the deviations accumulate across the ring. To this end, mutant octamer rings were superimposed on the wild-type residues 46-54. Then, the deviations between pairs of R50-C_α were measured (**Fig. 17A**). Immediate structural differences between the mutants and wild type RemA are primarily the mutated amino acids and their bonding network (**Fig. 17B**).

2.3.9 Mutations of *Gt*RemA affect the ratio between hexadecamer and octamer (or: tetradecamer and heptamer for R18W)

Crystallographic analysis showed that all mutants retain oligomeric arrangements of monomers in the crystal lattice and form octamers which further stack together to form hexadecamers (heptamers and tetradecamers for R18W). To validate the ability of *Gt*RemA mutants to oligomerize in solution, analytical SEC was performed. The experiments showed that all mutants elute as two species. This indicates that the two states of the protein observed in the crystal also exist in solution. Molecular weights of the two species were determined with MALS/UV (**Table 2**). Deviations of these values from the calculated molecular weights likely origin from the compact shape of the particles as well as the low extinction coefficient of RemA (see above). Notably, mutants of RemA show differences in the ratio between the two states (**Table 2**). Based on the results obtained from analytical SEC, selected mutants prefer the octameric state (P29S, R50A, H62Q) while others prefer the hexadecameric state (Y10I, R53A).

Table 2: Analytical SEC-results of *Gt*RemA mutants.

Variant	MW observed [kDa]	peak ratio P1:P2	Average angle	Average offset (R50)
*Gt*RemA wt	123/75	0.58 ± 0.03	44.95±0.472	-
RemA Y10I	99/60	0.8 (single run)	-	-
RemA R18W	142/82	0.45 ± 0.09	51.37±1.964	6.36±5.062
RemA P29S	120/100	0.13 ± 0.01	45.01±0.613	1.575±0.9
RemA R32A	tbd	0.54 ± 0.05	-	-
RemA R50A	132/78	0.12 (single run)	-	-
RemA R51A	115/70	0.48 (single run)	44.95±0.35	1.238±0.789
RemA R53A	126/75	0.68 (single run)	-	-
RemA H62Q	100/73	0.27 ± 0.01	45.03±1.453	1.8875±0.819

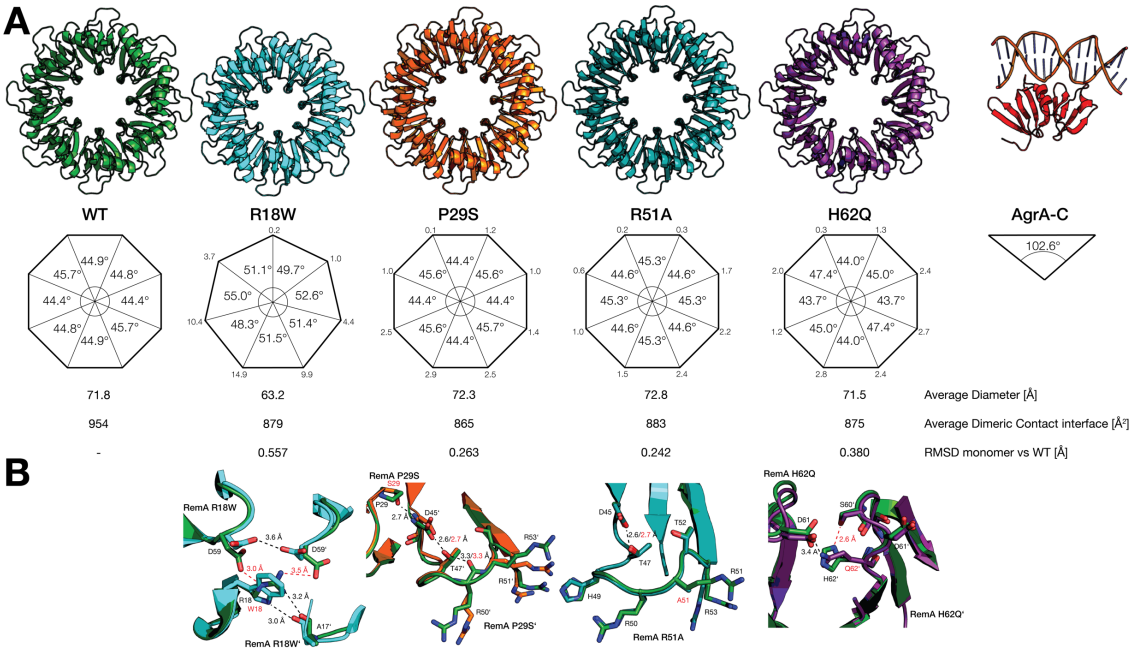

Figure 17. A: Overview of the oligomeric structure of *Gt*RemA wild-type protein and mutants (top row). Middle row: C_α's of residue R50 and the center of mass of the protein-complex were used to determine angles between the monomers (angular values). The values outside the octagon (heptagon for R18W) indicate the offset between corresponding C_α's of R50 (in Å). Measured values for average diameter, average dimer contact interface and the RMSD between mutant monomers and wild-type are given below. **B**: Immediate structural consequences of the mutations. First panel: The mutation R18W impacts the interaction network on the inside of the ring. The contacts D59-R18-D59' are lost. In the mutant, W18 maintains the H-bond to A17; and D59 and D59' interact directly. Second panel: The mutation P29S leads to formation of a H-bond S29-D45'. This increases the distance between the backbones of S29 and D45' (0.4 Å vs wild-type) and affects the bonding network of D45' and the position of the adjacent arginines R50', R51' and R53'. Third panel: The mutation R51A does not lead to a major rearrangement of the loop. Fourth panel: In the mutant H62Q, the H-bond between H62' and S60' is lost. Q62' instead binds to D61 directly.

2.3.10 Other mutants which were analysed by analytical SEC but did not yield diffracting crystals:

Further mutants of *Gt*RemA were constructed: Y10I, R32A, R50A, R53A. *Gt*RemA Y10I, R50A and R53A did not yield diffracting crystals. *Gt*RemA R32A crystallized as lens-shaped crystals in a condition composed of 0.1 M NaOAc pH5.0, 10-20% MPD. These crystals could be reproduced in fine screens, but only diffracted to 4 Å and lower. The data obtained by analytical SEC are summarized in **Table 2** (see above).

2.3.11 Mutants of RemA are impaired in their ability to bind DNA and act as transcriptional activators

Dr. Tamara Hoffmann performed the experiments detailed in this paragraph and kindly provided the images. GtRemA variants for the EMSA-experiments were cloned and purified by D.M.

*Gt*RemA can complement a *B. subtilis*-RemA-K.O.-strain (*Gt*RemA is expressed from an IPTG-inducible promoter at the *ytoI*-locus), showing that the two homologs are functionally equivalent. In reporter-assays, mutants of *Bs*RemA show a loss of transcriptional functionality. The ability of *Bs*RemA to effect transcription from its target promoters was determined in LacZ/TreA-assays. Briefly, the promoters and the upstream region were fused at the 5'-end of the *lacZ/treA*-gene and the transcriptional activity quantified by chromogenic analysis of LacZ or TreA.[81,94] In a first step, transcriptional activity of *Bs*RemA and mutants was determined in a *sinR*-K.O.-background for the *epsA*-promoter. The experiment shows that while *Bs*RemA is functionally active, all mutants except D36S and D39A are inactive. *Bs*RemA H62Q retains a basal activity 4% of that of the wildtype protein (**Fig. 18A**, left panel). Amino acids D36 and D39 (E39 in *Gt*RemA) are located in helix α1 and likely mediate hexadecamer-formation. In summary, all tested mutants except D36S and D39A showed complete loss of transcriptional activitiy on the *epsA*-promoter.

In a next step, *Bs*RemA-transcriptional activity was assessed on the *opuA*-promoter in the absence and presence of 1.2 M NaCl. *OpuA* is an osmotically

inducible promoter and hence shows a strong increase in transcriptional activity in the presence of high salinities.

In the absence of 1.2 M NaCl, wildtype *Bs*RemA as well as D36S and D39A show basal activies, while all other mutants show no activity (**Fig. 18A**, middle panel). This is similar to the situation on the *epsA*-promoter. In the presence of 1.2 M NaCl, the activity of wildtype *Bs*RemA and the mutants D36S and D39S increases approximately five-fold. Additionally, R18W and H62Q show an activity similar to the wildtype protein (**Fig. 18A**, right panel).

Taken together, all tested mutants except D36S and D39A are inactive in the *epsA*-background. The same holds true for the *opuA*-background. However, under high salinities, R18W and H62Q variants of RemA regain activity.

The DNA-binding capacity of *Gt*RemA and its mutants was assessed in EMSAs employing the promoter regions of *epsA* and *opuA* (**Fig. 18B**, upper and lower panel, respectively). The EMSAs show that the DNA-binding capacity of *Gt*RemA is the same for *PepsA* and *PopuA*, confirming the observation that *Gt*RemA shifts DNA nonspecifically (see above). Based on the EMSA-experiments, three functional groups can be identified:

1) Variants of *Gt*RemA that shift DNA like the wildtype (Y10I, H62Q).

2) Variants of *Gt*RemA that shift the DNA to a lesser degree than the wild type protein (R18W, R32A).

3) Variants of *Gt*RemA that have lost the ability to shift DNA (P29S, R50A, R51A, R53A).

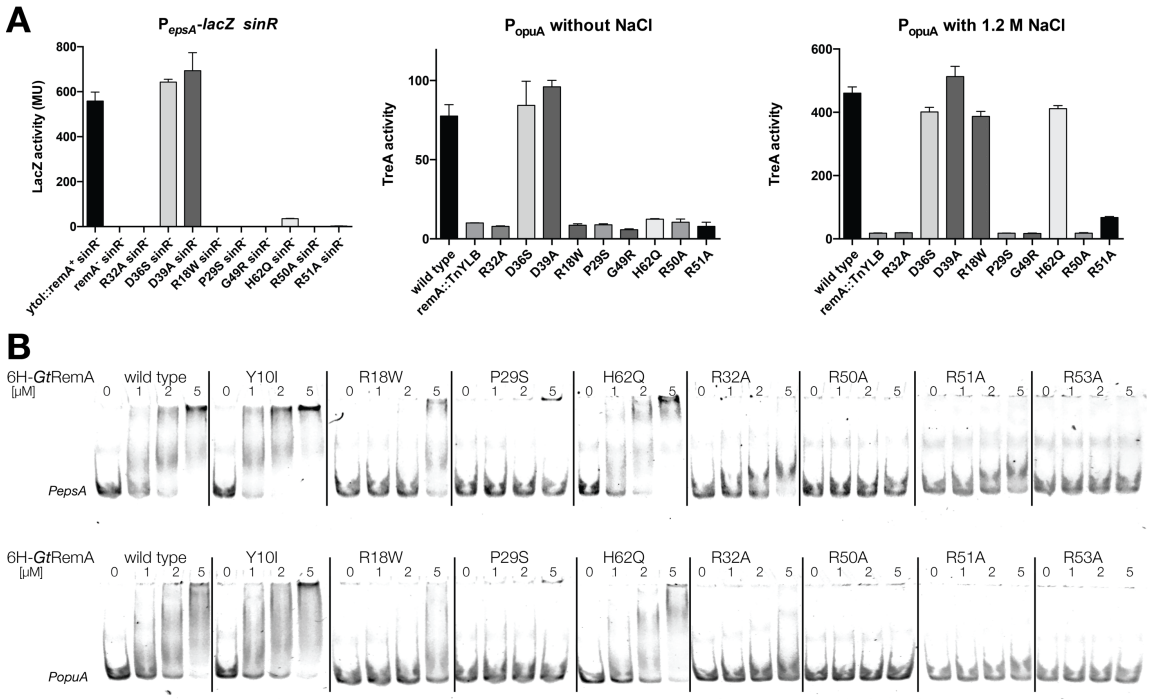

Figure 18. Functional experimental results of RemA. **A**: Reporter-activity of *Bs*RemA and mutants on *epsA* (left), *opuA* witout NaCl (center) and *opuA* with 1.2 M NaCl (right). *Bs*RemA R18W and H62Q regain transcriptional activity only in the case of the *opuA*-promoter and in the presence of 1.2 M NaCl. **B**: EMSAs of *Gt*RemA and mutants on the *epsA*-promoter (top row) and the *opuA*-promoter (bottom row). The preparation of strains and DNA-fragments were conducted by Dr. Tamara Hoffmann, the *Gt*RemA-variants used for EMSAs were purified by Devid Mrusek. The reporter experiments as well as the EMSAs were conducted by Dr. Tamara Hoffmann. The diagrams and images were prepared by Dr. Tamara Hoffmann.

2.3.12 Summary of structural and functional results

Taken together, RemA from *Geobacillus thermodenitrificans* was structurally and functionally characterized. The structure of *Gt*RemA shows a novel oligomeric architecture and represents the minimal motif of the previously identified LytTR-fold. *Gt*RemA is the first example for an octameric arrangement of a single-domain, DNA-binding protein. The relevance of this novel structure lies in its DNA-recognition, which is in all likelihood mediated by loops, and and the highly stable quarternary structure. DNA-binding of *Gt*RemA occurs nonspecifically and is mediated by arginine-residues which are located on the outside rim of the

octamer. The function of these arginines appears to be cooperative, as single mutants lose their ability to bind DNA. So far, only one mutation (R18W) has been identified which disrupts the octamer and collapses the quarternary fold to a heptamer. While this mutant shows dramatic defects in geometry, it can still bind DNA and is transcriptionally active in high salinities. Other mutants induce minute changes in protein geometry, which accumulate in the octamer and lead to offsets of functionally critical residues on the order of 2-3 Å. So far, no interaction partner of RemA has been identified. The exact way in which RemA binds DNA remains to be investigated.

2.4 Discussion

2.4.1 The structure of RemA shows a single-domain protein with a compact fold that oligomerizes into higher-order arrangements

Based on the crystallographic data, a model for RemA could be built which shows a compact tertiary structure with a β-β-α fold. Monomeric RemA (which was never observed in solution) provides three means for homo-oligomerization:

a) β-complementation originating from β2
b) β-complementation originating from β3
c) Dimerization via residues located in helix α1.

The wedge-shaped fold of RemA (curvature ca 45°, see **Fig. 17A**) leads to the exclusive formation of octameric rings. The interface between two monomers accounts for 18.6% (=surface area 5076 $Å^2$; interface area 942 $Å^2$) of the total solvent-accessible surface (37.1% when considering that each RemA monomer has two neighbors). Search of the protein data bank (www.rcsb.org) with parameters [protein-stoichiometry=A8, polymer type=protein, protein symmetry=cyclic] returned 18 hits within the bacterial domain of life. Of these 18, 10 are membrane proteins which form pores for transport (e.g. curli transport lipoprotein, PDB: 4UV2). Of the remaining 8, only one is a small-molecular weight, single-domain α+β protein: MPN314 from *Mycoplasma pneumoniae* belongs to the UPF0040-family of proteins implicated in bacteria cell division.[95] In contrast to

RemA, MPN314 shows two subdomains which are related by a pseudo two-fold axis (which is similar to AgrA-C). Additionally, MPN314 oligomerizes by a »hand-in-hand« arrangment and forms an octamer with a star-shaped appearance (**Fig. 19A**)

Two octamers of RemA can further stack together with their helical faces to form a hexadecamer (mode c), interface between monomers (180 $Å^2$) only 3.5%). No structural model is found in the protein data bank which shows a similar stoichiometry with a cyclic 8-fold symmetry.

2.4.2 Single-amino acid mutations do not abrogate oligomerization of RemA

For all eight RemA-mutants considered in this study, analyical SEC/MALS indicated that they form homo-oligomeric species in solution. Four of these mutants could also be crystallized and confirmed this observation. One mutant (R18W) led to a collapse of quaternary structure and assembled into heptamers instead of octamers. Interestingly, R18W is still partially functional. This shows that the fold and oligomeric state of RemA is highly stable. Furthermore, it opens the possibility to use the RemA-protein as a tool for synthetic biology to study the impact of changes in primary sequence and the consequences in oligomeric state.

2.4.3 RemA can exist in two states of unknown biological function

*Gt*RemA has been characterized structurally and the impact of single amino-acid mutations has been elucidated for selected examples. It has been found that wild-type RemA oligomerizes into octamers, which further associate into hexadecamers. These higher-order oligomeric structures are not an artifact from crystal packing, because they have been observed also for mutants of RemA, which crystallized in entirely different space groups. Furthermore, RemA also exists as hexadecamer and octamer in solution. It has been found that the two species exist in a dynamic equilibrium and preliminary results suggest that the equilibrium can be influenced by small molecules such as proline. Furthermore, certain mutants of RemA show a different ratio between hexadecamer and octamer, indicating that structural changes within the molecules (even such that

are distant from the interface) influence the affinity between octamers. So far, it is not known which amino acids mediate the interaction between octamers. There is a characteristic sequence of alternating charges (R32, D36, E39) on the helical face of the octamer. However, preliminary results of single amino-acid mutants (R32A) show that hexadecamerization is not impaired and the 16mer:8mer-ratio of *Gt*RemA R32A was not significantly different from the wild-type protein. The molecular factors which enable the hexadecamer remain to be investigated. The impact of salinity, nucleotides and DNA as well as small molecules on the 16mer:8mer-ratio should also be investigated in more depth.

It can only be assumed that the two states of RemA also occur under physiological conditions (**Fig. 19B**). If this were the case, then the biological relevance of the two states of RemA need to be investigated. It is not clear which of the states, or if both, are functionally active. To elucidate the functional relevance of the two states of RemA, the selective preparation or purification of one state exclusively is an important prerequisite (see above). The two states of RemA could show different specificities, either towards binding partners or towards different DNA regions (see also below). If the two states of RemA impact downstream signalling, then factors need to be identified which mediate the equilibrium between the two states.

Given the fact that RemA acts as a transcriptional activator both for biofilm-formation and osmo-stress protection, the two states of RemA could mirror this functional duality.

2.4.4 RemA binds DNA via its lateral convex sides

The key function of RemA is transcriptional activation. To do so, RemA must be able to bind DNA. It has been found that RemA binds DNA via three conserved arginine residues which are located on an exposed loop. In the crystal, the electron density map of the loop is of excellent quality and allows for unambigious construction of amino acid backbone and sidechains, indicating that this loop is not flexible. Therefore, the arginine-rich loop might act as a DNA-recognizing motif. Indeed, arginines are the most frequent residues to interact with minor grooves of the DNA.[96] DNA binding is abolished if one of the arginines is mutated

to alanine, indicating that the DNA-binding occurs cooperatively. A similar observation has been made for other DNA-binding proteins which have conserved arginines in their DNA-binding helices.[97] DNA binding is also abolished in mutants P29S and R32A. *Gt*RemA P29S has lost its ability to interact with DNA (**Fig. 18B**). The structural basis for this is the altered bonding network of D45‘, which leads to a repositioning of the adjacent DNA-binding arginines R50‘, R51‘ and R53‘. As *Gt*RemA R32A only yielded poorly diffracting crystals, a structural basis for the loss of DNA-binding is not clear to date. In conclusion, RemA presents eight identical DNA-binding interfaces in a cyclic arrangement (16 if the hexadecamer is considered, **Fig. 19C**, left panel).

Dimeric RemA (which was never observed in solution) shows structural similarity to *S. aureus* AgrA-C, which is the first characterized member of the novel family of DNA-binding proteins with the LytTR-fold. The identification of RemA reveals that LytTR is not the minimal motif which can bind DNA. However, RemA was never observed in any stoichiometry other than octameric/hexadecameric. Therefore, it remains to be investigated what the evolutionary relationship between AgrA-C and RemA is. While homologs of AgrA are highly abundant in the phylum Firmicutae, *B. subtilis* does not encode a homolog of AgrA. Homologs of RemA are also not present in *S. aureus*. Phylogenetic analyses need to be undertaken in order to elucidate whether proteins of the LytTR-fold exist in organisms that encode *remA*.

AgrA-C and RemA display different curvatures (100.2° vs 2x45°, **Fig. 17A**) and the superimposition of 4 copies of AgrA-C onto octameric RemA shows that the DNA bound by AgrA-C indeed attains a curvature different from that of the RemA octamer (**Fig. 19C**, right panel). The circular decoration of RemA with DNA in **Fig. 19B** is reminiscent of eukaryotic histone-DNA complexes (e.g. chromatin, PDB: 1AOI). Indeed, octameric RemA fits nearly perfectly into the looped DNA of the chromatin-DNA complex (**Fig. 19D**, left panel). In the nucleosome, the eight histone proteins lead to a two-turn looping of DNA. Importantly, the two-turns of the histone-binding DNA-fragment (146 bp) also match the back-to-back-arrangement of hexadecameric RemA (**Fig. 19D**, right panel). Notably, eukaryotic

nucleosomes are composed of eight proteins, two copies each of histones H2A, H2B, H3 and H4, respectively.[91] Histone proteins require chaperons to prevent misguided DNA-interactions, which would lead to the aggregation of DNA-protein-complexes.[98] The high basicity of histone proteins could also lead to unwanted interactions with other acidic partners in the cell. So far, no chaperone or binding partner has been identified for RemA and it remains mysterious how RemA is prevented from associating with random acidic components, or how sequence specificity of RemA is achieved (see below).

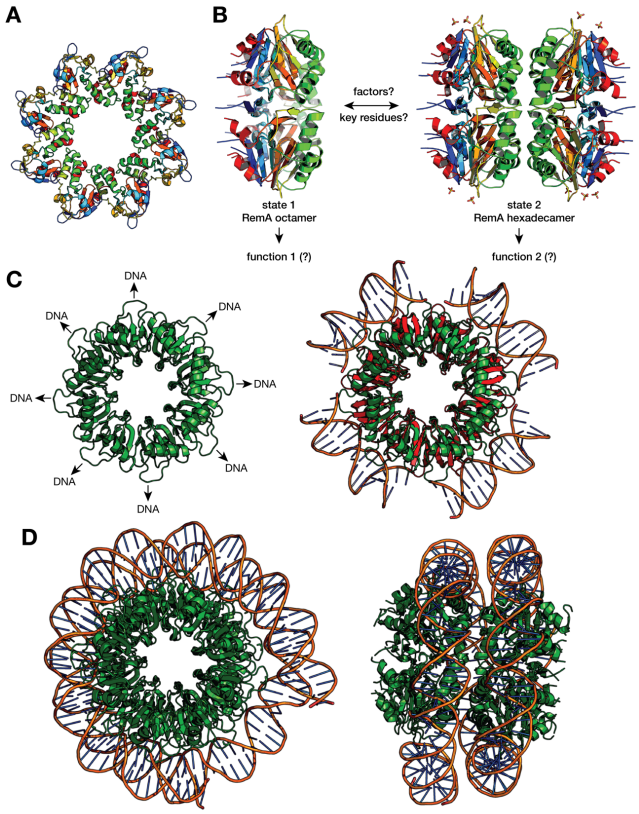

Figure 19. A: Structure of MPN314 (PDB: 1N0G), a protein from *Mycoplasma pneumoniae* which forms octameric complexes with cyclic symmetry. In contrast to RemA, MPN314 contains two sub-domains that are equivalent to each other. **B**: The octameric and hexadecameric states of RemA are in a dynamic equilibrium with each other, and unknown factors might promote the interconversion. Furthermore, each state might serve a distinct function. **C**: Octameric RemA provides eight identical DNA binding interfaces (left panel). Right panel: Superimposition of four copies of AgrA-C (PDB: 3BS1) onto an octamer of RemA.

D: Model of RemA fitted into the looped DNA of the chromatin nucleosome complex (PDB: 1AOI) front view (left panel) and side view (right panel).

In preliminary EMSAs, RemA was found to bind DNA non-specifically, likely due to exposed arginine residues, which coordinate sulphate-ions in the crystal structure. These arginines might provide a strong, yet nonspecific DNA-interaction, as their mutation completely abrogates DNA-binding. Therefore, primary DNA-binding of RemA is achieved by three exposed arginine residues which likely bind to phosphates from the DNA-backbone.

Footprinting assays employing RemA show a characteristic protection pattern on target-DNA. This finding seems to contradict the observation that RemA binds DNA irrespective of sequence. RemA shows a high specificity towards certain promoters and even binds some recognition sequences stronger than others.(Winkelman 2013) Residues in exposed loops of RemA might also be involved in DNA-binding (e.g. K31, H48, Q69, E71) and provide a low-energy, sequence-specific interaction.

However, it is not clear how the same DNA-binding interface can specifically interact with varying DNA-recognition sequences: While octamers of RemA provide eight copies of the same interface to bind DNA in a circular array (**Fig. 19C**), the 7 bp DNA-sequences that RemA recognizes do not show a consensus motif but are entirely different from each other (**Fig. 20A**). The same holds true if RemA were to bind DNA as a hexadecamer. Therefore, the higher-order structure of RemA suggests a sequence-independent binding mode to DNA. This mirrors the situation of eukaryotic histone-proteins, which mainly contact DNA via backbone-phosphates.[91] Additionally, the interaction between DNA and histone-complexes is known to be very dynamic: histone-complexes can »slide« on the DNA, which further highlights the sequence-nonspecificity of the histone DNA-binding moieties.[99] Other proteins also bind DNA as as oligomers and show a high sequence specificity. As for RemA, the exact mode of how these proteins recognize DNA is not clear to date.[100]

The question therefore is: how can a repeating binding interface specifically recognize a variable strech of DNA (**Fig. 19B**)? In the following, the explanation of this phenomenon shall be attempted. As a closing remark, experimental approaches will be sketched out with the aim to elucidate the DNA-binding mode of RemA.

2.4.5 Sequence-specificity of RemA could be an intrinsic feature of the DNA

DNA-recognition by proteins can either occur via the formation of base-specificic hydrogen bonds or by the recognition of DNA-deformation. For example, the histone-like protein from *E. coli* recognizes DNA by »indirect readout«, for which sequence-dependent DNA-shaping likely forms the basis.[60] DNA shape can be influenced, among other factors, by AT-rich regions, which tend to narrow minor grooves.[96] RemA utilizes arginines to bind DNA: this amino acid most frequently binds to narrow minor grooves.[96] Notably, the majority of RemA target-DNAs identified to date have AT-contents of up to 75% (**Fig. 20A**). Mathematical simulations have shown that especially during the binding stage between protein and DNA, intrinsic conformational features of the DNA play a significant role.[101] A more recent conceptual framework describing DNA-protein interactions defines DNA-sequence and DNA-structure-guided protein recognition as two extremes on a DNA-binding continuum. As purely sequence-guided recognition would be extremely prone to DNA-mutations, and DNA-structure is not completely independent of sequence, the authors conclude that both extremes are not likely to exist.[102] An overview of DNA-structure and DNA-protein-recognition is given elsewhere.[103] On a side note, in-cell salt concentration also impacts DNA-supercoiling, which could enhance the recognition of curved DNA by RemA.[104] Taken together, RemA could recognize a distinct three-dimensional structure of its target DNA. The role of RemA as an osmostress-regulator and the fact that DNA-struture is affected by osmostress further highlight this possibility.

2.4.6 Sequence-specificity of RemA could be promoted by interaction partners of RemA

In prokaryotes, sigma^{54}-dependent transcriptional activation requires the presence of »enhancer binding proteins« (see also section 1.1.3), which bind upstream of the promoter sequence and facilitate open-complex formation of the RNAP-sigma^{54}-complex.[48] However, genes under the control of RemA are not sigma^{54}-dependent, but are controlled by the sigma^{70}-factor sigmaA. Sigma^{70}-dependent genes do not require the action of enhancer binding proteins. Thus, RemA likely functions similar to cAMP receptor protein (CRP, or catabolite activator protein, CAP) and enhances the ability of RNAP-holoenzyme to bind and initiate transcription.[105,106] In contrast to CRP, the stoichiometry and putative DNA-bending of RemA is more reminiscent of enhancer binding proteins. However, RemA lacks the canonical ATPase-domain which provides mechanical work for open-complex formation. Furthermore, RemA binds DNA differently from CRP (which uses a classical HTH motif). Two facts hint at the possibility that the transcriptional regulation of RemA occurs in concert with interaction partners:

1) RemA relieves transcriptional repression by SinR. SinR itself binds DNA segments close to or overlapping with RemA-recognition motifs (**Fig. 20A**). Furthermore, SinR binds DNA as a dimer and has been proposed to further associate into tetramers, thereby looping DNA (**Fig. 20C**). Given the considerations above about DNA-recognition based on DNA-structure, SinR could directly or indirectly interact with RemA: The RemA-binding sites upstream of the *epsA*-promoter are located in a region of DNA which is bent by SinR. Thereby, SinR indices a structural change in target DNA that RemA recognizes. Effectively, SinR recruits RemA. SinR and RemA could possibly also directly interact with each other. In this scenario, interaction of SinR with the two states of RemA should be differentiated.

2) RemA is encoded in an operon immediately upstream of *rpoZ*, the omega-subunit of the bacterial RNAP. As proteins encoded in one operon often interact with each other, a direct interaction between RemA and RNAP could be a

possibility. In this scenario, RemA would be guided to its recognition sequence via the interaction with RNAP. As transcriptional regulators of two component systems often are phosphorylated, which affects their DNA-binding capacity, phosphorylation of RemA should also be considered.[107] In gram-negative bacteria, the omega-subunit of RNAP (rpoZ) is required for correct RNAP-folding and sigma-factor binding. In gram-positive bacteria, fewer research results are available for *rpoZ*. However, in *M. smegmatis* depletion of *rpoZ* leads to diminished biofilm formation and a change in biofilm matrix composition.[108] A similar observation has recently been made in *S. aureus*.[109]

2.4.7 Identification of similar proteins to RemA that bind DNA

Identification of RemA as the minimal fold of the DNA-binding LytTR-family poses the question how proteins similar to RemA can be identified. One example will be given here in the context of this work:

Since RemA and the structurally and functionally similar protein AgrA-C do not share significant primary-sequence identity, other bioinformatic methods should be used to identify proteins which are functionally similar to RemA. The search algorithm could utilize the pattern of secondary structure elements of RemA (**Fig. 12A**). The protein RemB was identified together with RemA and also activates transcription from the *epsA*-operon.[80] RemA and RemB do not share sequence similarity. However, they share the same pattern of secondary structural elements (**Fig. 20D**). Therefore, RemB could possess a similar ternary fold like RemA. Structural investigation of RemB is currently in process and it will be interesting to assess the DNA-binding capacity of RemB, as the protein contains fewer arginines than RemA.

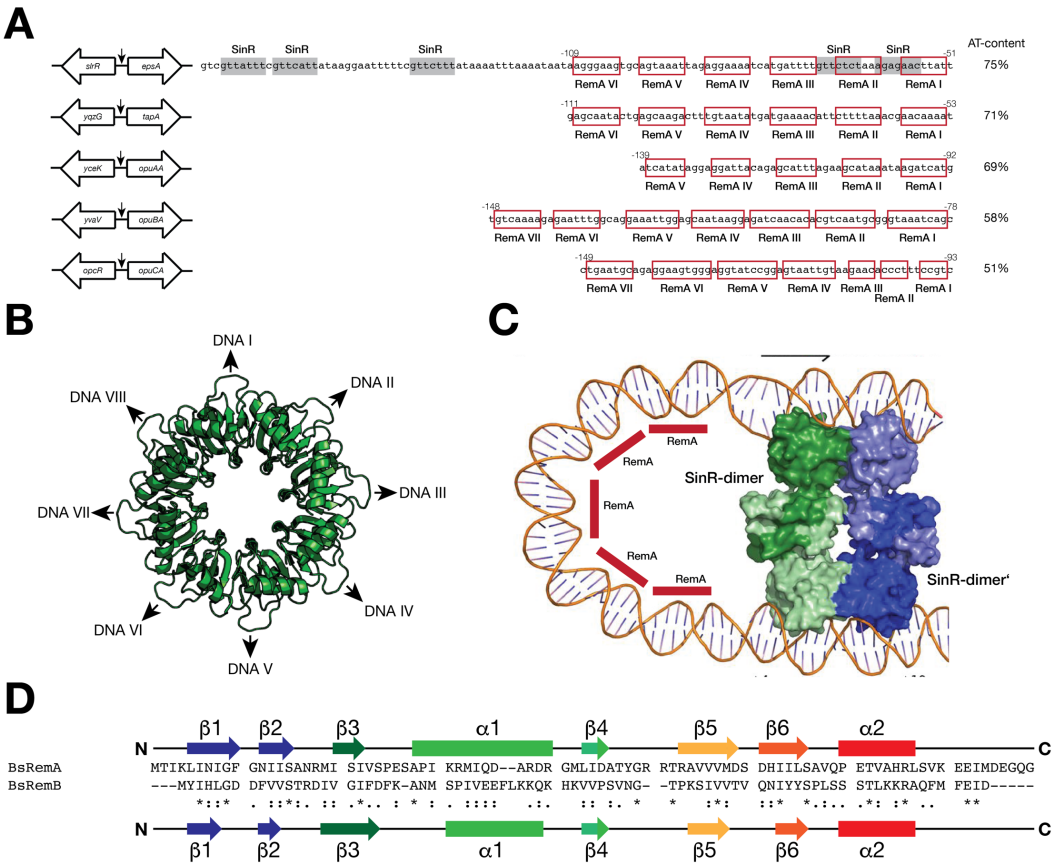

Figure 20. A: Experimentally determined binding sites of RemA and SinR on target DNA (adapted from [81], Hoffman, unpublished) **B**: Conceptual depiction of octameric RemA and its putative circular mode of DNA-interaction. **C**: Model of DNA-interaction on the *epsA*-promoter region proposed for two SinR-dimers. Selected RemA-binding-sites are indicated as red bars. Adapted with permission from [78].

2.4.8 Future perspectives

To elucidate the two putative states of RemA and to investigate the RemA-DNA-complex, the following systematic approach should be undertaken:

1) Identification of an optimal DNA-construct for complex-reconstitution. To this end, binding energies of RemA to long DNA-fragments of the various RemA-binding regions should be quantified (either with fluorescently labeled DNA by microscale thermophoresis, or by label-free bio-layer interferometry.

2) Elucidation, which state of RemA preferably interacts with DNA: RemA can exist as an octamer or hexadecamer in solution. Therefore, employing corresponding

DNA-segments (ca 75 bp or 150 bp) and quantifying the resulting binding strength to RemA might shed light on the question whether both states can interact with DNA. The preparation of RemA-variants which are exclusively hexadecameric or octameric could be beneficial in this regard.

3) The DNA-protein complex should be assembled by mixing DNA and protein together at a high salt concentration (1-2 M NaCl) and then reducing the salt concentration in a step-wise process. Detailed method descriptions can be found elsewhere.[110,111]

3 Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor

In the following, the published manuscript with the title “Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor” is attached. For this manuscript, I designed the construct and performed protein purification, crystallization and structure determination of E. coli FtsY NGdN1. Likewise, I designed the constructs and performed protein expression and purification of GB1-N2-4 for Hydrogen-Deuterium Exchange. Together with Prof. Dr. Gert Bange and Prof. Dr. Eitan Bibi, I contributed to writing the manuscript and creating the figures.

Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor

Amihai Karniel [1,†], Devid Mrusek [2,†], Wieland Steinchen [2], Orly Dym [3], Gert Bange [2] and Eitan Bibi [1]

*1 - **Department of Biomolecular Sciences,** Weizmann Institute of Science, Rehovot 7610001, Israel*
*2 - **Center for Synthetic Microbiology (SYNMIKRO) and Department of Chemistry,** Philipps University Marburg, 35043 Marburg, Germany*
*3 - **Structural Proteomics Unit (SPU),** Weizmann Institute of Science, Rehovot 7610001, Israel*

***Correspondence to Eitan Bibi:** e.bibi@weizmann.ac.il*
https://doi.org/10.1016/j.jmb.2018.04.017
Edited by Bert Poolman

Abstract

Much of our knowledge on the function of proteins is deduced from their mature, folded states. However, it is unknown whether partially synthesized nascent protein segments can execute biological functions during translation and whether their premature folding states matter. A recent observation that a nascent chain performs a distinct function, co-translational targeting *in vivo*, has been made with the *Escherichia coli* signal recognition particle receptor FtsY, a major player in the conserved pathway of membrane protein biogenesis. FtsY functions as a membrane-associated entity, but very little is known about the mode of its targeting to the membrane. Here we investigated the underlying structural mechanism of the co-translational FtsY targeting to the membrane. Our results show that helices $N_{2–4}$, which mediate membrane targeting, form a stable folding intermediate co-translationally that greatly differs from its fold in the mature FtsY. These results thus resolve a long-standing mystery of how the receptor targets the membrane even when deleted of its alleged membrane targeting sequence. The structurally distinct targeting determinant of FtsY exists only co-translationally. Our studies will facilitate further efforts to seek cellular factors required for proper targeting and association of FtsY with the membrane. Moreover, the results offer a hallmark example for how co-translational nascent intermediates may dictate biological functions.

Introduction

Folding is usually required to achieve the three-dimensional structure of proteins and therefore shapes their function [1,2]. It is well-established that individual domains of multi-domain proteins can already reach their “final” folding state during protein synthesis at the ribosome [3–5]. However, co-translational folding of a sub-domains might differ from how they finally fold in the mature state. Such bistable folding intermediates could execute distinct functions during translation and other functions or none after translation when the proteins mature. If this was true, as we propose here for the signal recognition particle (SRP) receptor, it offers an additional layer of biological regulation.

The SRP receptor is a major player in the pathway of membrane protein biogenesis in all living cells [6–9]. Its functional interactions with the SRP [10,11] and the translocon [12–14] and its association with membrane-bound ribosomes [15,16] underscore its central role in ribosome targeting and biogenesis of membrane proteins. The *Escherichia coli* receptor FtsY contains three domains, termed A, N and G [17,18] (Fig. 1a). The A domain is not essential for bacterial growth [19], although it was proposed that this long acidic domain participates in regulatory aspects of the pathway [14,20]. The G domain is responsible for GTP binding and it interacts with SRP [21,22] through the homologous domain of the SRP protein, Ffh [24]. The structure of a functional NG + 1 [19] core of FtsY (Fig. 1B) [23] shows that the N-domain is composed of four well-defined α-helices (termed $N_{1–4}$), oriented in an anti-parallel manner. A short helical segment that precedes helix N_1 mediates the essential regulation of FtsY by anionic

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

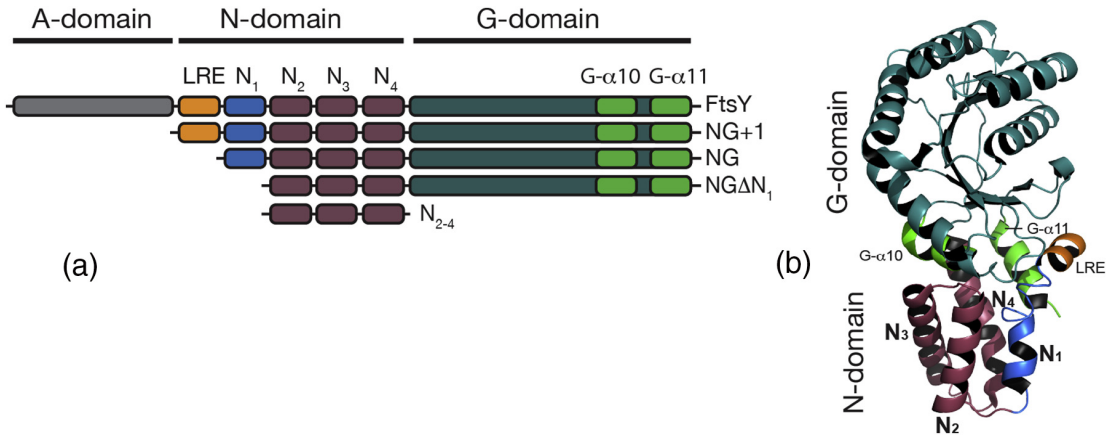

Fig. 1. Composition and structure of the *E. coli* SRP receptor. (a) Schematic representation of FtsY. Full-length FtsY comprises domains A, N and G. LRE is a short helical LRE. NG + 1 is a functional core receptor [19] comprising domains N and G without the A-domain. NG is a non-functional receptor variant containing domains N and G without the LRE and the A-domain. $NG\Delta N_1$ is NG + 1 deleted of N_1 and LRE (contains only helices $N_{2–4}$ of the N-domain). (b) Crystal structure of NG + 1 [23]. The LRE is shown in orange, helix N_1 is shown in blue, helices $N_{2–4}$ are shown in brown and the G-domain is shown in dark green except the C-terminal helices G-α10 and G-α11, which are shown in light green.

phospholipids at the inner membrane [23,25–28]. Proper formation of this amphipathic element is essential for the receptor function *in vivo* [19], through its lipid-responsive effect on the GTPase activity of the SRP/SRP receptor complex [25]. Its ability to interact with lipids and its similarity to the most C-terminal membrane targeting segment of MinD [29] have led to the suggestion that this short helix is the primary membrane targeting sequence (MTS) of FtsY [30]. However, we have shown that NG, a construct defective in forming this short helical structure, is also able to target the membrane *in vivo* [25]. Moreover, our recent studies with translation intermediates (TIs) lacking this helix, further emphasized the question what mediates targeting of FtsY to the membrane [31]. Based on these and other studies, we proposed that the short helix is not required for membrane targeting but rather serves as a lipid-responsive element (LRE). The work described here offers an interesting structural view on how FtsY targeting to the membrane occurs in the absence of the LRE and helix N_1.

Clearly, as outlined above, FtsY executes its essential function at the membrane, but only little is known about the underlying mechanisms that ensure its productive membrane localization. Previously, we hypothesized that FtsY might target the membrane co-translationally [6,32,33] and recent studies with FtsY TIs demonstrated feasibility [31]. Moreover, this work also showed that exposing helices $N_{2–4}$ of the N-domain at the ribosome during translation ($TI\text{-}N_{2–4}$) is necessary and sufficient for membrane targeting. However, it remains a mystery how these helices ($N_{2–4}$) can target the stalled ribosomes to membranes *in vivo*, especially since N_1 and the LRE, which were previously assigned to mediate functional membrane lipid interactions, are absent in $TI\text{-}N_{2–4}$. In addition, helices $N_{2–4}$ do not contain a bona fide transmembrane segment [31], suggesting that they interact with the membrane peripherally, maybe through membrane-associated proteins. All together, these observations suggested that helices $N_{2–4}$ (i) are not involved in the interaction with lipids, but possibly with non-lipid components at the membrane [34]; (ii) exhibit a specific conformational state in the context of $TI\text{-}N_{2–4}$; and (iii) undergo a conformational rearrangement in the context of the fully translated G-domain of FtsY. If true, such a co-translational folding intermediate could carry the information for targeting of the SRP receptor to its destination(s) at the membrane.

To challenge these ideas, we performed a series of experiments that strongly support the notion that $N_{2–4}$ of FtsY adopts two distinct conformational states: an elongated helix and a three-helical bundle. The elongated helical state represents an intermediate that exists only throughout the translation of FtsY. Upon completion of FtsY synthesis, residues at the C-terminus of the G-domain facilitate the formation of a stable, three-helical bundle state of $N_{2–4}$. The conformational switch of $N_{2–4}$ might separate between its role in co-translational targeting of FtsY and ribosomes to the cytoplasmic membrane and its role in the assembly and downstream functions of the mature receptor. We propose that the underlying mechanistic basis for this behavior of $N_{2–4}$ is “conformational bistability.” We speculate that FtsY presents a hallmark example for other, yet unknown proteins with bistable folding states that execute distinct biological functions during and after their translation.

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

Results

$N_{2–4}$ localizes with the membrane fraction and does not co-migrate with cytoplasmic ribosomes

We previously showed that exposing helices $N_{2–4}$ of the N-domain at the ribosome during translation (TI-$N_{2–4}$) is necessary and sufficient for membrane targeting *in vivo* [31]. These results suggested to us that $N_{2–4}$ performs an individual function, independently of the other FtsY domains, co-translationally. To investigate whether $N_{2–4}$ can target the membrane also in the absence of the ribosome, we analyzed the subcellular distribution of $N_{2–4}$ as an independently expressed protein. Cell fractionation results show that $N_{2–4}$ appears in the cytosol and in the membrane fraction (Fig. 2a, left panel). In FtsY, $N_{2–4}$ is "sandwiched" between N_1 and the G-domain, and thus, we asked what would be the effect on membrane association of fusing an unrelated protein at the N- or C-terminus of $N_{2–4}$ (N_4-C). To this end, we utilized protein GB1 domain from streptococci [35]. The results show that both hybrids (GB1-$N_{2–4}$ and $N_{2–4}$-GB1) migrate with the membrane fraction to a lesser extent compared with $N_{2–4}$ (Fig. 2a, right panels), suggesting that the unrelated flanking GB1 peptides interfere. We also asked whether the fully expressed $N_{2–4}$ co-migrates with cytosolic ribosomes in density gradient, which would indicate interaction between $N_{2–4}$ and the ribosome. The results show that $N_{2–4}$ does not co-migrate with the cytosolic ribosomes (Fig. 2b, c); it mostly sediments in the pellet, together with the membranes. Collectively, these and previous experiments [31] suggest the $N_{2–4}$ serves as an autonomous membrane targeting domain, which is responsible for membrane targeting of FtsY in the context of the translational intermediate TI-$N_{2–4}$, but probably less so when fused to unrelated proteins. The results also suggest that as an independently expressed protein, $N_{2–4}$ can target the membrane post-translationally. This is in contrast to the

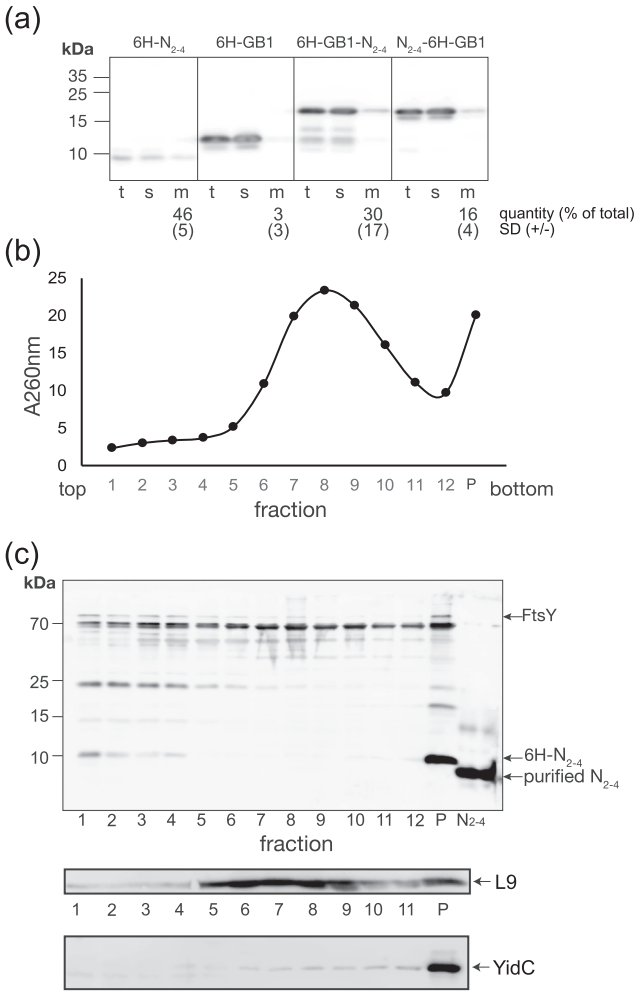

Fig. 2. Cellular distribution of independently expressed $N_{2–4}$. (a) Cellular distribution of independently expressed $N_{2–4}$, GB1 and two versions of their hybrids. Samples from total extracts (t), supernatant after untracentrifugation of the total extract (s) and membranes purified by flotation (m) (10 μg protein in all cases) were subjected to SDS-PAGE, electroblotted into a nitrocellulose membrane and detected by HisProbe™. For densitometry and statistical calculations, the gels were loaded with both 10 and 20 μg samples. The experiment was repeated three times and the blots were analyzed by densitometry. The ratio m/t is shown as a %. (b) Cells expressing 6H-$N_{2–4}$ were sonicated and ultracentrifuges, and the pellet fraction was suspended and analyzed by sucrose density gradient centrifugation and RNA content measurements. (c) Equal volumes taken from each fraction of the sucrose gradient (b) were assayed by Western blotting with anti-FtsY antibodies. Right lane, purified $N_{2–4}$.

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

$N_{2–4}$-mediated obligatory co-translational targeting mode of NG [25,31], which does not interact efficiently with lipids or with *E. coli* membranes post-translationally [30,36].

Crystal structure of $N_{2–4}$ shows an elongated helix

To better understand the function of $N_{2–4}$, we determined its crystal structure. $N_{2–4}$ was overexpressed with a hexahistidine tag (6His) followed by a thrombin cleavage site (TCS), purified and crystallized. Two crystal forms of $N_{2–4}$ were obtained, monoclinic and orthorhombic, that led to structure determination at 2.65- and 1.95-Å resolution, respectively (Table 1). Both crystal structures unambiguously showed $N_{2–4}$ as a 74.8-Å-long helix with 13.8 turns (Fig. 3a). Surprisingly, this structure is entirely different from the structure of $N_{2–4}$ in the context of the active core receptor, NG + 1 [23] (Fig. 1b), where $N_{2–4}$ forms a three-helical bundle. The elongated $N_{2–4}$ helix is in all likelihood not a crystal packing artifact because it was observed in both crystal forms, with completely different crystal contacts (Fig. S1). However, the notion that N_1 of the N-domain might influence the structure of $N_{2–4}$ in the context of NG + 1 has not escaped our attention. Therefore, we also determined the crystal structure of $NG\Delta N_1$ at a 2.1-Å resolution (Fig. 3b). $NG\Delta N_1$ was crystallized in a space group *I*4 with two copies per asymmetric unit (Fig. S2A). Structural comparison of $NG\Delta N_1$ (Fig. S2B, left panel) and NG + 1 (Fig. S2B, right panel) shows no difference in the structure of the helical bundle formed by $N_{2–4}$, strongly suggesting that N_1 is not required for formation of the three-helical bundle of $N_{2–4}$.

$N_{2–4}$ exists in two distinct conformational states depending on the presence of the G-domain

Importantly, all the above considerations are based on crystal structures and we therefore asked what happens in solution. To test whether $N_{2–4}$ retains its extended helical structure in solution and forms a three-helical bundle only in the context of $NG\Delta N_1$, we initially utilized cysteine cross-linking. Notably, FtsY has no native cysteines and cysteines were introduced at sites 227 (L227C) and 264 (A264C). As seen in Fig. 4, these cysteines may form a disulfide bond only in the helical bundle-state of $N_{2–4}$ (Fig. 4b), but not in the extended conformation (Fig. 4a). We show that inserting these two cysteines did not compromise expression and subcellular distribution (Fig. 4c), and NG + 1(L227C/A264C) was functional *in vivo* in FtsY-complementation assay (Fig. 4d).

Table 1. Data collection and refinement statistics for the FtsY structures

Data collection			
PDB code	6FPK	6FPR	6FQD
Space group	$C222_1$	*C*2	*I*4
Cell dimensions			
a, *b*, *c* (Å)	56.10, 82.55, 29.81	164.57, 35.49, 31.70	110.69,110.96, 110.63
α, *β*, *Υ* (≡)	90, 90, 90	90, 96.48, 90	90, 90, 90
No. of copies in a.u.	1	2	2
Resolution (Å)	46.40–1.95	40.88–2.65	45.28–2.10
Upper resolution shell (Å)	2.06–1.95	2.79–2.65	2.175–2.10
Unique reflections	5333 (749)[a]	5450 (781)[a]	38,419 (3856)[a]
Completeness (%)	99.9 (99.9)	100.0 (100.0)	98.28 (98.32)
Multiplicity	7.1 (7.2)	7.4 (7.6)	3.9 (3.7)
Mean $I/\sigma(I)$	18.6 (4.1)	15.1 (3.9)	10.34 (2.64)
R_{sym} (*I*) (%)	6.7 (49.5)	9.9 (58.6)	8.1 (43)
Refinement			
Resolution range (Å)	46.40–1.95	40.88–2.65	45.28–2.10
No. of reflections ($I/\sigma(I) > 0$)	4767	5178	38,414
No. of reflections in test set	551	257	2051
R-working (%)/*R*-free (%)	21.2/23.4	23.8/27.7	17.75/20.82
No. of protein atoms	407	849	4098
No. of ions/ligands atoms	3 MPD		2 GDP, 2 K
No. of water molecules	3	5	327
Overall average *B* factor (Å^2)	34.0	52.7	45.04
Root mean square deviations			
Bond length (Å)	0.021	0.014	0.005
Bond angle (°)	2.35	1.6	0.81
Ramachandran plot			
Most favored (%)	100.0	96.0	98.87
Additionally allowed (%)	0.0	2.0	1.13
Disallowed (%)	0.0	1.0	0.00

[a] Values in parentheses refer to the data of the corresponding upper resolution shell.

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

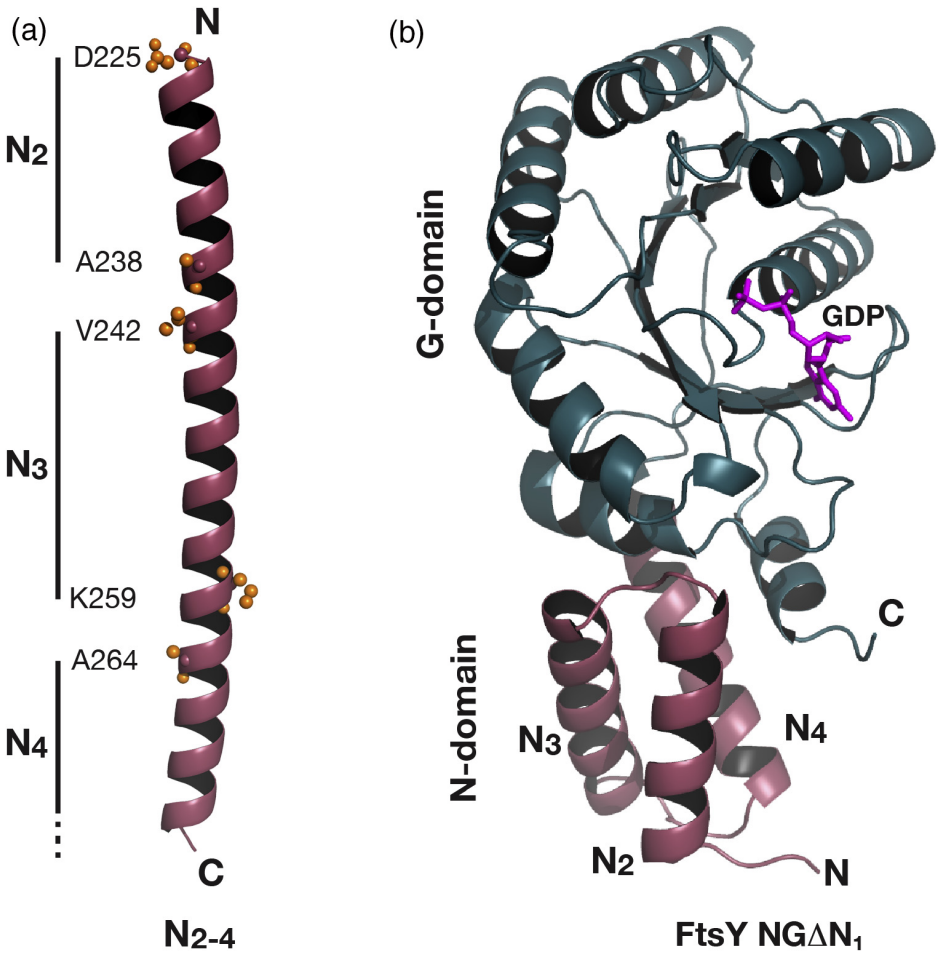

Fig. 3. Structural analysis of $N_{2–4}$ alone and in the presence of the G-domain. (a) Crystal structure of $N_{2–4}$, space group $C222_1$, determined at 2 Å. The boundaries of $N_{2–4}$ as seen in the three-helical bundle of the NG-domain are indicated as orange spheres (compare also to Fig. 1b). (b) Crystal structure of NGΔN_1. A GDP nucleotide in the active site of the GTPase is shown as magenta sticks.

Next, the mutants were expressed and cell extracts were incubated in the absence or presence of the oxidizing complex Cu-phenanthroline (Cu/phe), which catalyzes disulfide bond formation between adjacent cysteines. Disulfide bonds can sometimes be detected by SDS-PAGE separation because they increase the compactness of proteins. The more compact (cross-linked) forms usually migrate faster on the gel. The results show that the cysteine pairs form a disulfide bond in the context of NGΔN_1(L227C/A264C) (Fig. 4e). Even in the absence of harsh oxidation by Cu/phe, disulfide bonds form, although less efficiently. Next, we tested whether $N_{2–4}$(L227C/A264C) forms an intramolecular disulfide bond. The results show that this construct forms mainly intermolecular disulfide bonds (Fig. 4f). In the absence of Cu/phe, $N_{2–4}$(L227C/ A264C) seems to form dimers and other heavier adducts, whereas with Cu/phe, it mostly forms various high molecular adducts. These results suggest that $N_{2–4}$(L227C/A264C) does not form a three-helix bundle in solution. In all cases, treatment with β-ME reduced the disulfide bonds.

We next asked what is the conformation of $N_{2–4}$ during translation at the stalled ribosomes [in the context of the TI TI-$N_{2–4}$(L227C/A264C)]. The results show that $N_{2–4}$(L227C/A264C) forms mainly inter-molecular disulfide bonds observed as high-molecular-weight adducts (Fig. 4g), especially in the presence of Cu/phe. This suggests that $N_{2–4}$ does not form a helical bundle also in the context of translation.

Taken together, the cross-linking experiments clearly support the notion that $N_{2–4}$ can exist in two distinct and G-domain-dependent conformational states (i.e., an elongated structure and a three-helical bundle). The results suggest that $N_{2–4}$ does not form a three-helix bundle, neither when fully expressed nor when it is still tethered to the ribosome as a TI.

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

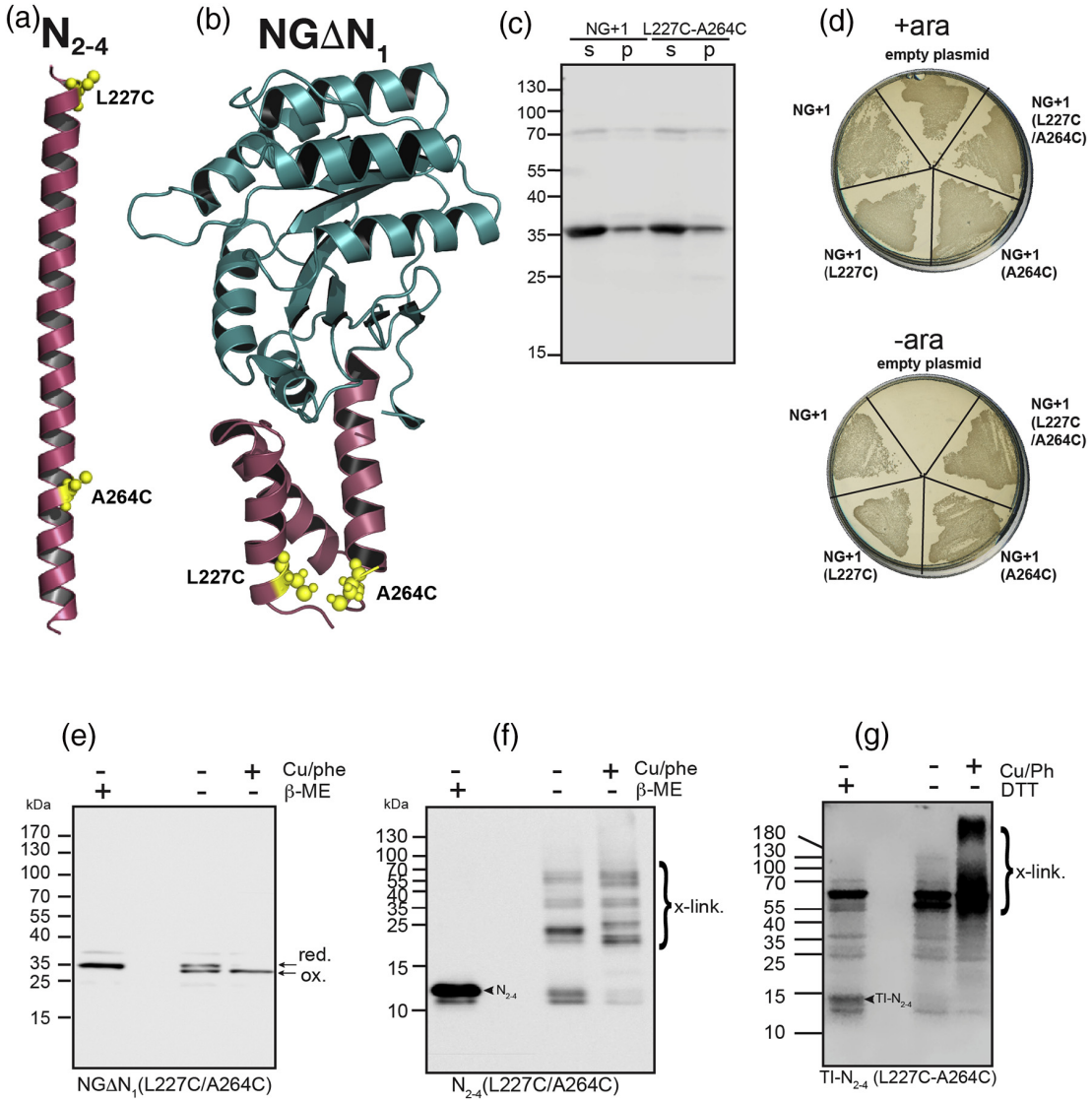

Fig. 4. Biochemical analyses of disulfide bond formation in the double cysteine mutants of $N_{2–4}$, $NG\Delta N_1$ and TI-$N_{2–4}$. (a and b) Inserted cysteines in $N_{2–4}$ and $NG\Delta N_1$ are shown as yellow spheres. (c) Expression and distribution of NG + 1 and its double mutant NG + 1(L227C/A264C). Each lane contains 10 μg protein. (d) Complementation of FtsY depletion by the double cysteine mutant of the active core receptor NG + 1. (e) Disulfide bond formation in extracts prepared from *E. coli* expressing 6His-tagged $NG\Delta N_1$(L227C/A264C). (f) Disulfide bond formation in extracts prepared from *E. coli* expressing 6His-tagged $N_{2–4}$(L227C/A264C). (g) Disulfide bond formation in extracts prepared from *E. coli* expressing TI-$N_{2–4}$(L227C/A264C). In panels E–G, each lane contains 20 μg protein.

Conformational dynamics of $N_{2–4}$ in relation to the G-domain in $NG\Delta N_1$

The cross-linking studies enabled assessment of the distance between only two residues at steady state. We therefore employed *h*ydrogen–*d*euterium e*x*change (HDX) *m*ass *s*pectrometry (HDX-MS), a method for the dynamic analysis of protein conformation and structure in solution, which allows for analyses of much wider areas of interactions. Briefly, purified $N_{2–4}$ was incubated in D_2O-containing buffer and the reaction was quenched after 15, 30, 60, 120 and 600 s. After digesting with pepsin or fungal protease XIII, the resulting peptides were separated by liquid chromatography and analyzed by mass spectrometry. The results show that $N_{2–4}$ readily exchanges protons over the whole sequence already after 15 s and no further changes were observed over

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

the time course of the experiment (Fig. 5a). This fast HDX reaction of $N_{2–4}$ means that it does not contain any ternary structure, in agreement with our crystallographic and biochemical analyses.

In the next step, we studied the conformational dynamics of $NG\Delta N_1$ in which $N_{2–4}$ is arranged in a three-helical bundle (Fig. 3b). In the context of $NG\Delta N_1$, $N_{2–4}$ shows a high degree of protection (Fig. 5b), indicating a stable ternary structure, as was described by Stjepanovic et al [30]. Taken together, our HDX results clearly support the notion that $N_{2–4}$ can exist in two distinct conformational states, that is, an elongated helix in the absence of the G-domain and a three-helical bundle in the presence of the G-domain.

So far, our data strongly support the notion that formation of the three-helical bundle requires the presence of the G-domain. We therefore asked whether specific determinants within the G-domain C-terminal part would provide a driving force for the conformational rearrangement, as proposed previously [31]. Close inspection of the $NG\Delta N_1$ crystal structure together with our HDX-MS data revealed that the N_4-C (i.e., residues 271–279) is highly protected in the context of the G-domain. The crystal structure shows that N_4-C is clamped between helices G-α10 and G-α11 at the C-terminus of the G-domain (Fig. S2B). These, mainly hydrophobic interactions, are stabilized by hydrogen bond pairs formed between $N_{2–4}$ and the C-terminal part of the

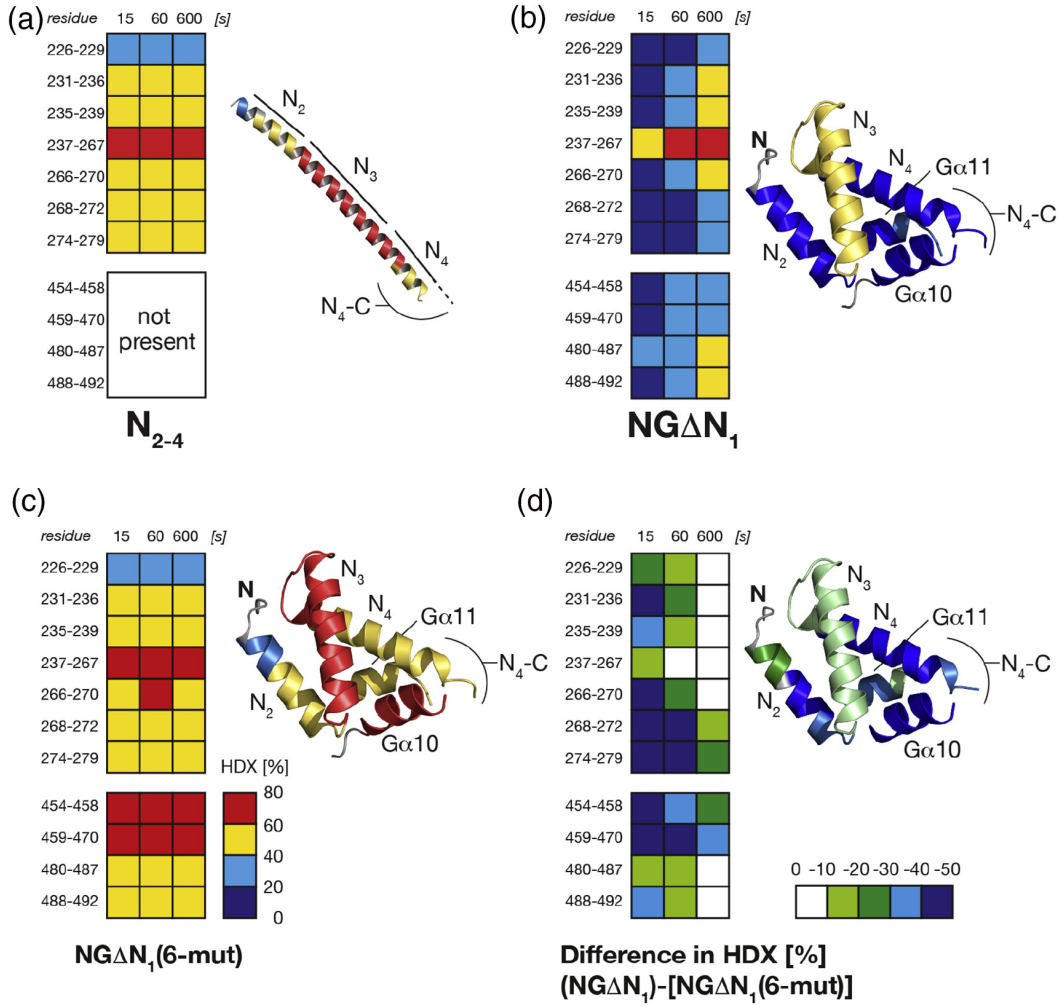

Fig. 5. Conformational dynamics of $N_{2–4}$ and the effect of the G-domain. (a) HDX-MS analyses of helices $N_{2–4}$ as independently expressed protein. (b) HDX-MS analyses of helices $N_{2–4}$ in construct $NG\Delta N_1$. (c) HDX analysis of $N_{2–4}$ in construct $NG\Delta N_1$(6-mut). (d) Difference in relative deuterium incorporation between $NG\Delta N_1$ and $NG\Delta N_1$(6mut). In all panels, the relative deuterium incorporation of $N_{2–4}$ or $NG\Delta N_1$ is mapped onto their crystal structures (right panels). Most of the G-domain of $NG\Delta N_1$ was omitted for clarity.

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

G-domain (K247/D462 and D239/K453). All these contacts might provide the driving force for the transition of $N_{2–4}$ from an elongated helical state to the three-helical bundle once the complete G-domain is synthesized and released from the ribosome. To test this idea, we created a NGΔN1 variant in which residues critical for the stabilizing interactions between N_4-C and the G-domain were exchanged [$NG\Delta N_1$(6mut); Fig. S3]. Strikingly, $N_{2–4}$ of $NG\Delta N_1$(6mut) shows an up to 50% higher deuterium incorporation compared to $NG\Delta N_1$ in HDX-MS, which could be explained by $N_{2–4}$ adopting an elongated helical state (Fig. 5c). Notably, the N_4-C segment showed the highest relative change in HDX, from fully protected in $NG\Delta N_1$ to almost fully labeled in $NG\Delta N_1$(6mut) (Fig. 5d). In line with these observations, G-α10 (−60%) and G-α11 (−30%) displayed marked differences in HDX between $NG\Delta N_1$ and $NG\Delta N_1$(6mut) after 15 s of deuteration. Therefore, we conclude that clamping of N_4-C between G-α10 and G-α11 is critical for the conformational rearrangement of $N_{2–4}$ from the elongated helical state into the three-helical bundle.

Discussion

The *E. coli* SRP receptor FtsY performs its biological role at the membrane together with SRP and the translocon(s) [37,38]. A poorly answered question is how FtsY functionally targets and binds the membrane? Previous work identified a short amphipathic helix at the N-terminus of N_1 in the N-domain, capable of interacting with anionic phospholipids. This amphipathic element is essential for FtsY function because it is necessary for the functional response of the receptor to anionic lipids [23,25,28]. Consequently, this short helix has been proposed to be the primary element for membrane localization of FtsY and termed MTS for membrane targeting sequence [30]. However, work in our laboratory has shown that NG, a construct missing this helix, can target the membrane *in vivo* [25,28]. Thus, we concluded that this acidic lipid-interacting helix may not be required for targeting the receptor to the membrane [31] but rather to sense acidic lipids and regulate the GTPase activity of FtsY. This led to the suggestion that this short helix is a LRE (Fig. 1a). The notion that this short helix is not required for targeting FtsY to the membrane was later substantiated by the observation that a subdomain of FtsY, $N_{2–4}$ targets the membrane during the receptor translation even in the absence of the LRE and N_1 [31]. However, these results do not explain why NG variants that contain $N_{2–4}$ did not efficiently interact with acidic lipids post-translationally [30]. To resolve this mystery, we speculated that FtsY might contain a signal for membrane localization that is only present during synthesis of FtsY at the ribosome, but absent in the mature FtsY and is independent of the LRE.

Our previous experiments showed that exposing helices $N_{2–4}$ of the N-domain at the nascent chain exit tunnel of a SecM-stalled ribosome (TI-$N_{2–4}$) is necessary and sufficient for membrane targeting [31]. Therefore, $N_{2–4}$ must contain the signal, which targets nascent FtsY to the membrane independently of the LRE. Moreover, these studies have shown that a TI exposing the entire NG + 1 (TI-NG + 1) is unable to target the membrane, unless the interactions between the N-domain and the C-terminal part of the G-domain were disrupted [31]. These results thus support the notion that the signal is hidden in the mature receptor and only exposed as a nascent chain during translation. Our current structural and mechanistic analyses suggest that indeed, $N_{2–4}$ can exist in two distinct conformational states: (i) the targeting determinant, which is an elongated helix (Fig. 6a) and (ii) a three-helical bundle (Fig. 6b). We show that if the G-domain is not present, $N_{2–4}$ favors the elongated state, while once the G-domain is present the three-helical bundle forms. Interestingly, all residues critical for promoting the conformational transition of $N_{2–4}$ into the three-helical bundle localize at the very C-terminus of the G-domain. This is in agreement with our hypothesis that only full translation and release from the ribosome of the G-domain provides the driving force for folding $N_{2–4}$ into the three-helical bundle seen in many FtsY structures. In conclusion, we suggest that the elongated helical state of $N_{2–4}$ persists only during translation, until the entire receptor has been synthesized and released from the ribosome (Fig. 6). The conformational transition of $N_{2–4}$ from an elongated helical conformation to a three-helical bundle could be imagined as a molecular switch that separates between the role of $N_{2–4}$ in co-translational membrane targeting of FtsY and ribosomes and its role in the assembly and downstream functions of the mature receptor.

Our analysis leaves an important question open: What are the down-stream events and interactions of the elongated $N_{2–4}$ when it emerges from an FtsY-translating ribosome? In all likelihood, helices $N_{2–4}$ do not contain a *bona fide* transmembrane segment [31] or any further lipid-interacting elements [23]. This idea is supported by the fact that the previously identified LRE domain of FtsY [23,25,26] does not exist in $N_{2–4}$ and is not required for RNC targeting [31]. Therefore, we favor the idea that $N_{2–4}$ targets the membrane during translation of FtsY by interacting with membrane-associated proteins. Interestingly, previous *in vitro* studies on the membrane targeting of the mammalian homolog of the SRP receptor, SRα demonstrated co-translational targeting to the membrane integral SRβ subunit of the receptor [39]. However, bacteria do not have an SRβ homolog, and the primary docking site for the FtsY–ribosome

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

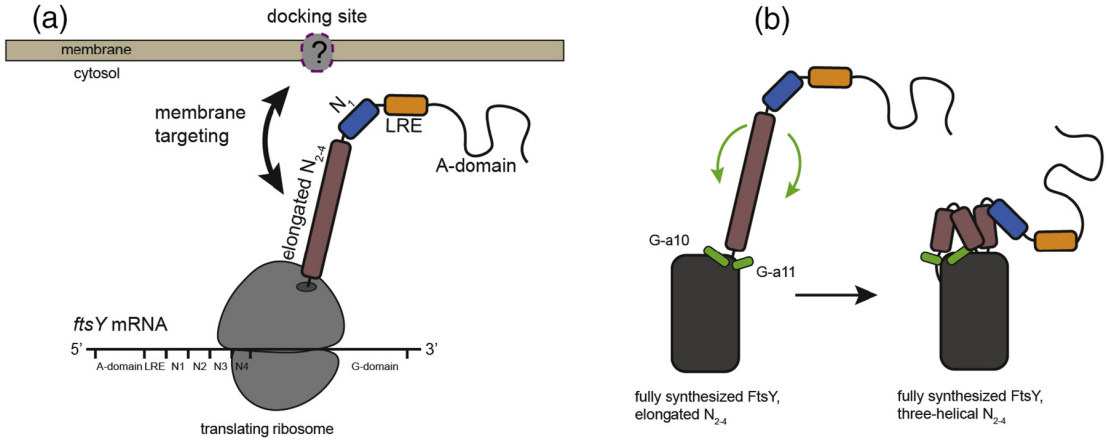

Fig. 6. A model for co-translational targeting of FtsY. (a) Proposed order of folding events during and after translation of FtsY. $N_{2–4}$ exists in an elongated helical state during translation and is targeted to the membrane by unknown partners. (b) Upon translation termination of FtsY and its release from the ribosome (left panel), refolding of the elongated $N_{2–4}$ into a three-helical bundle (right panel) occurs rapidly and is mediated by the two C-terminal helices of the G-domain shown in green.

RNC complex has to be identified. $N_{2–4}$ may bind to several promising membrane protein candidates, including SecY, YidC, and maybe other membrane-associated proteins [12,28,38,40]. These possibilities are currently being investigated.

Finally, our study demonstrates that the co-translational folding state of a nascent domain can dramatically differ from its folding in the mature protein. The case of FtsY shows that such a discrete folding state, which is buried in the mature protein, is temporally exposed only during translation and executes a function, which is absent in the mature protein. We speculate that other such cases might exist, but they are hidden and will only be identified if investigated co-translationally. Therefore, we propose FtsY as a hallmark example for an additional layer of spatiotemporal regulation that is present only during protein translation.

Materials and Methods

E. coli strains, plasmids and growth conditions

E. coli TOP10 was used for propagation and preparation of various plasmid constructs. *E. coli* BL21(DE-03) (Novagen) served in cell fractionation studies and protein purifications. *E. coli* IY28 [28], which contains a chromosomal ftsY gene under the arabinose promoter, was used for FtsY-complementation experiments. *E. coli* UT5600 ΔssrAsmpB was used for expression of TIs. Typically, cells were cultured in LB medium supplemented with the appropriate antibiotics (kanamycin was used at 30 μg/mL and ampicillin at 200 μg/mL). The gene encoding $N_{2–4}$ was amplified from pT7–5($N_{1–4}$) [31] and cloned into pET28a (after the 6His-encoding sequence), by restriction-free (RF) cloning [41] utilizing the forward primer that encodes also a TCS, 5′-GTGCCGCGCGGCAGCAAAAAAA-TCGACGATGATCTG-3′ and the reverse primer 5′-GACGGAGCTCGAATTCGGATCCTTACGCCA-GAATCTCGC-3′. The gene encoding NGΔN_1 was amplified from pT7–5-tacP(NGΔN_1) [31] with the forward primer 5′-GTGCCGCGCGGCAGCAAAAAA-ATCGACGATGATCTGTTTGAGG-3′ and reverse primer 5′-ACGGAGCTCGAATTCGGATCCT-TAGTCCTCTCGGGCATCAAG-3′, and cloned into pET28a with a 6His tag and a TCS to generate pET28a(6His-TCS-NGΔN_1). Mutants L227C, A264C, or L227C/A264C were constructed in pT7–5-tacP (NG + 1), pET28a(6His-TCS-$N_{2–4}$) and pET28a(6His-TCS-NGΔN_1) by PCR, using the following primers: for L227C, forward 5′-CGACGATGATTGTTTT-GAGGAGCTGGAAGAGCAG-3′ and reverse 5′-ATTTTTTTGCTGCCGCGCGGCAC-3′; for A264C, forward 5′-TGACTGTGAGGCGCTCTATGGCCTG-3′ and reverse 5-CGAAGCTGCTTGCGGGATG-3′. The double mutants L227C/A264C were constructed by transferring L227C into the single A264C mutants utilizing PCR. For testing disulfide formation in $N_{2–4}$ in the context of translation, we used PCR and RF-cloning to create plasmid that encodes a TI, termed briefly TI-$N_{2–4}$(L227C/A264C) in the text. In addition to N2–4 (L227C?A264C) the plasmid encoding this TI (pBAD24/araP/$N_{2–4}$(L227C-A264C)/TSS/UP12) harbors the following elements: araP, arabinose promoter, a 33-residue long linker from protein SecM [31]; TSS, a translation stalling sequence (WWWPPIRGPPGS) adopted from the work of Cymer *et al.* [42]; and uspG, a gene encoding UP12 [43], used as a reporter for arrest-bypass translation.

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

The 6His-NGΔN1(6mut) protein contains the following mutations in the G domain: K453A, F458A, D462A, F464A, I489D and F493D. The DNA section, which contains the mutations, was synthesized by Gen9 and cloned into pET28a(6His-TCS-NGΔN_1) using RF cloning with the forward primer 5′-GCATGAAGTTATGCTGACTATTGATGCCAGCAC-CGGGCAGAACGCGG-3′ and reverse primer 5′-CGGAGCTCGAATTCGGATCCTTAGTCCTC-TCGGGCATCAAG-3′. The final plasmid was verified by DNA sequencing. The gene encoding 6His-NGΔN_1(6mut) was then amplified (without the TCS) and subcloned into plasmid pET24d (Novagen) by utilizing enzymes NcoI and XhoI.

For HDX studies, the gene encoding $N_{2–4}$ was amplified from the genome and cloned into pEM (GB1). This plasmid is a derivative of plasmid pCFX3 [44]. The gene encoding $N_{2–4}$ was cloned into this plasmid using the restriction enzymes NcoI and XhoI to generate the expression plasmid pEM(6His-GB1-TEV-$N_{2–4}$). For $N_{2–4}$ GB1 fusion cell fractionation tests, $N_{2–4}$ was cloned before GB1 in pEM(6His-GB1) (vector) by RF cloning using the forward primer 5′-CATCGAAGGCCGCGGCCGCAAAAAAATCGAC-GATGATCTGTTTGAG-3′ and the reverse primer 5′-GTTTTACCGTTCAGGATCAGTTTGTACTGCGC-CAGAATCTCGCCCATC-3′, then this construct was added a stop codon after GB1 by PCR (the pEM system takes its stop from the fusion protein which is downstream to GB1) using the forward primer 5′-TAAAGTCGACTCGAGCGAGCTC-3′ and the reverse primer 5′-CATGGTCATATGGCTGCCGCTC-3′. This primer pair was also utilized to generate a stop codon in pEM(6His-GB1)(vector) to form pEM(6His-GB1).

Expression and purification of $N_{2–4}$ and NGΔN_1

6His-TCS-$N_{2–4}$ was expressed in *E. coli* BL21(DE-3) from plasmid pET28a(6His-TCS-$N_{2–4}$) under regulation of the T7 promoter. Cultures were grown overnight at 37 °C in LB, diluted to an A_{600} of 0.01–0.05 and induced at an A_{600} of 0.4–0.7 with 0.5 mM IPTG for 1–2 h at 37 °C. For purification, 6His-TCS-$N_{2–4}$ in 50 mM NaPi (pH 7.5), 300 mM NaCl, 5 mM imidazole and 1 mM PMSF was purified by cobalt affinity chromatography (Talon, Clonetech) and eluted by thrombin cleavage using the thrombin cleavage capture kit and company instructions [buffer: 20 mM Tris–HCl (pH 8.4), 150 mM NaCl and 2.5 mM $CaCl_2$; 69022, Novagen]. Removal of the thrombin was accomplished by a second cycle of purification on streptavidin agarose beads (Novagen). Buffer exchange and concentration were done using 3 K MWCO NanoSep centrifugal device (Pall). The purified protein was analyzed and confirmed by mass spectrometry. The protein (~ 10 mg/mL) in 20 mM Tris–HCl (pH 8) was then subjected to crystallization screens.

6His-NGΔN1 and 6His-NGΔN_1(6mut) were produced in BL21(DE-3). Briefly, cultures of LB medium were inoculated from an LB overnight culture and grown at 37 °C, up to A_{600} ~ 0.5–0.7. Then the temperature was reduced to 18 °C for 60 min, before IPTG induction (0.1 mM), and growth continued for 18 h. After cell lysis by a Microfluidizer (M110-L, Microfluidics), cell debris was removed by high-speed centrifugation and proteins were purified by Ni-NTA and size exclusion chromatography (SEC) as described recently [45].

For HDX studies, 6His-GB1-TEV-$N_{2–4}$ was produced in BL21(DE-3). Briefly, overnight culture was diluted 1:200 into fresh LB with ampicillin 100 μg/mL and grown in the presence of inducer [1.75% (w/v) D (+)-lactose-monohydrate] for 16 h at 30 °C. Cell harvest, lysis and Ni-NTA purification were performed as described above. $N_{2–4}$ was then released from the Ni-NTA beads by cleavage with TEV (0.1 mg/1 mg hybrid) in buffer B (20 mM Hepes, 20 mM $MgCl_2$, 20 mM KCl, 200 mM NaCl, 500 mM imidazole) at 4 °C for 16 h in a dialysis tube against buffer A (as buffer B, but 40 mM imidazole). The cleaved mixture was passed over a 1 mL HisTrap HP column (GE Healthcare) for removal of the 6His-GB1 fragment and 6His-TEV, and $N_{2–4}$ was further purified by SEC (see above). For production of 6His-TEV protease, we used plasmid pET(6His-TEV), and *E. coli* Rosetta (Novagen). Transformed cells were grown in LB media supplemented by lactose at 30 °C for 16 h and processed as described above. 6His-TEV was purified by Ni-NTA chromatography, eluted in buffer B [20 mM Hepes–Na (pH 8.0), 200 mM NaCl, 20 mM KCl, 20 mM $MgCl_2$, 500 mM imidazole] and concentrated to 0.5 mg/mL. The SEC buffer consisted of 20 mM Hepes–Na (pH 7.5), 200 mM NaCl, 20 mM KCl and 20 mM $MgCl_2$.

For subcellular distribution studies, 6His-TCS-$N_{2–4}$, 6His-GB1-TEV-$N_{2–4}$, 6His-GB1 and 6His-$N_{2–4}$-GB1 were produced in BL21(DE-3). Briefly, Cultures were grown overnight at 37 °C in LB, diluted to an A_{600} of 0.01–0.05 and induced at an A_{600} of 0.4–0.7 with 0.5 mM IPTG for 1–2 h at 37 °C. Induced cells were fractionated by membrane flotation as described below.

FtsY complementation

Complementation experiments were conducted by plating transformed *E. coli* IY28 [28], harboring the indicated pT7–5-tacP constructs on LB agar plates with ampicillin (200 μg/mL) and with or without 0.02% arabinose (for induction of the chromosomal *ftsY* gene) and with IPTG, the inducer of the complementing construct. Plates were scanned after 17 h at 37 °C.

Isolation of membranes by flotation

Harvested cultures were washed in a buffer containing 25 mM Hepes (pH 7.5), 10 mM $MgAc_2$,

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

20 mM KAc, 100 mM NH_4Cl and 5% sucrose. After centrifugation, the cell pellets were suspended in the same buffer with 1 mM Pefabloc (Merck). Cell suspensions were then sonicated, incubated on ice for 20 min and subjected to low-speed centrifugation for cell debris removal. For separation by floatation [46], membranes were collected by ultracentrifugation (Beckman, TLA120.2, 0.5 h, 53,000 rpm, 4 °C) and re-suspended in 60 μL of the same buffer. The membranes were then purified by flotation in a three-layer sucrose buffer solution (400 μL of 61% sucrose mixed with the resuspended sample, 680 μL of 53% sucrose and 270 μL of buffer with no sucrose) utilizing ultracentrifugation (Beckman, TLS55, 5–17 h, 54,000 rpm, 4 °C). Membranes were collected and associated proteins were resolved by SDS-PAGE and Western blot analysis.

Density gradient centrifugation

Harvested cultures were washed in buffer MNH [20 mM Hepes (pH 7.5), 20 mM $MgCl_2$, 20 mM KCl, 5% sucrose, 60 mM NH_4Cl, 0.1 mM DTT] and re-suspended in the same buffer. Extracts were prepared by three cycles of brief sonication (10 s) at 1-min intervals on ice, followed by a low-speed centrifugation (16,000*g* for 15 min) for removal of cell debris. Ribosomes and membranes were collected by ultracentrifugation (Beckman, TLA-120.2, 90 min, 75,000 rpm, 4 °C). Pellets were re-suspended in ice-cold 5% sucrose solution in buffer MNH. Samples were loaded on top of a preformed sucrose density gradient (1.35 mL containing 0.27-mL layers of 22.5%, 20%, 15%, 10% and 7.5% sucrose). Following ultracentrifugation (Beckman, TLS-55, 48 min, 54,000 rpm, 4 °C), fractions were collected from the top. The sucrose gradient pellet was resuspended in 10 mM Tris (pH 8) 1% SDS. A_{260} was measured for each fraction using a NanoDrop spectrophotometer and the values were plotted *versus* the fraction number. Fraction samples were subjected to Western blotting with αFtsY and αL9 antibodies or with αYidC antibodies, kindly provided by Hans-Georg Koch (Freiburg University), as described previously [15].

Disulfide bond formation

E. coli BL21 (DE-3) expressing the 6His-TCS-$N_{2–4}$ (L227C-A264C) or 6His-TCS-NGΔN_1(L227C-A264C) were grown as described in the previous section and induced with 0.5 mM IPTG for 1 h at 37 °C or overnight 16 °C, respectively. Total extracts were prepared as previously described. Oxidative cross-linking with Cu/Phe was performed essentially as previously described [47]. Briefly, Cu/Ph was prepared by mixing 40 μL of 1,10-phenanthroline (1.25 M in 50% ethanol) with 60 μL of $CuSO_4$ (250 mM). Then the extracts were incubated with Cu/Ph (0.5 mM) for 5 min at 37 °C, and the reaction was terminated by a non-reducing protein sample buffer. Samples (10 μg protein) were then subjected to SDS/PAGE separation and Western blot analysis using αFtsY antibodies and/or HisProbe™ for detection.

In experiments with the TI TI-$N_{2–4}$(L227C-A264C), cells were disrupted by three 5-s cycles of sonication in 50 mM Hepes (pH 7), 300 mM NaCl, 1 mM pefabloc and 4 U of RQ1 DNase, followed by removal of cell debris. Cross-linking was performed on total extracts at 1-mg/mL protein concentration, as described above, followed by ultracentrifugation (53,000 rpm, 1 h, TLA 120.2 rotor, Beckman-Coulter), for concentrating the TIs. Pellet was then suspended in 50 mM Na–Hepes (pH = 7), 300 mM NaCl and 12 mM EDTA, and treated with RNaseA (50 μg/mL, 1 h) at room temperature, before the addition of sample buffer (+/–) DTT. Samples were separated on a 10%–17% step gradient SDS PAGE and Western blots were probed by αFtsY antibodies.

Crystallization and structure determination of FtsY $N_{2–4}$

Crystals of $N_{2–4}$ were obtained using the hanging drop vapor diffusion method with a Mosquito robot (TTP LabTech). The crystals were obtained in a screen buffer containing 42% MPD and 0.1 M NaCacodylate (pH = 6.2). The crystals formed in the orthorhombic space group $C222_1$, with two monomers per asymmetric unit. A complete data set to 1.95-Å resolution was collected at 100 K on a single crystal on beamline BM14 of the European Synchrotron Radiation Facility (ESRF), at Grenoble, France. Diffraction images were indexed and integrated using the Mosflm program [48], and the integrated reflections were scaled using the SCALA program [49]. Structure factor amplitudes were calculated using TRUNCATE [50] from the CCP4 program suite. Initially molecular replacement (MR) was conducted with the program PHASER [51] using the relevant segment (70 residues S212–A280) of the NG + 1 construct of the *E. coli* SRP receptor FtsY solved to 1.9 Å (PDB code 2QY9). This segment shares 92% sequence identity with the FtsY and contains three consecutive α helices (Fig. 1A). Initially, all potential MR models had a high R and R_{free} values. In case the relative orientations of the three helices in the model are different from those in the crystal, we broke the model into separate three PDB files containing rigid-body unit-residues. Specifically, the segments K221–D239, V240–L261 and R262–K281 of PDB 2QY9 were used as separate ensembles in the MR search. To our surprise, a continuous electron density of a long α-helix containing all the tree α helices arranged consecutively was observed. All steps of atomic refinements were carried out with the CCP4/REFMAC5 program [52]. The model was built into $2mF_{obs} - DF_{calc}$, and $mF_{obs} - DF_{calc}$ maps by using the COOT program [53]. The R_{free} value is

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

23.4% (for the 5% of reflections not used in the refinement), and the R_{work} value is 21.1% for all data to 1.95 Å. The FtsY model was evaluated with the MOLPROBIDITY program [54]. Details of the refinement statistics of the FtsY structure are described in Table 1. Figures were prepared with Pymol (www.pymol.org). The coordinates of $N_{2–4}$ were deposited in the RCSB Protein Data Bank with accession code 6FPK. Another crystal form was obtained and solved as described to 2.65 Å. Also, these coordinates were deposited with accession code 6FPR.

Crystallization and structure determination of NGΔN$_1$

Crystallization was performed by the sitting-drop method at 20 °C in 500-nL drops consisting of protein and precipitation solutions in ratios of 1:1 and 1:2. NGΔN$_1$ crystallized at 30-mg/mL concentration with a 1.5-fold excess of GDP within 1 to 5 days in 0.1 M Na_3citrate (pH 5.6), 20% (v/v) ^{i}PrOH and 20% (w/v) polyethylene glycol 4000. Crystals were flash-frozen in liquid nitrogen employing a cryo-solution that consisted of mother-liquor supplemented with 20% (v/v) glycerol. Data were collected under cryogenic conditions at the ESRF, Grenoble, France, at beamlines ID30b and ID29. Data were processed with XDS [55] and ccp4-implemented AIMLESS [56]. Structures were determined by MR with PHASER [57], manually built in COOT [53] and refined with PHENIX [58]. Search models were the structures of the *E. coli* FtsY-NG (PDB ID: 1FTS). Figures were prepared with Pymol (www.pymol.org). The coordinates of NGΔN$_1$ were deposited in the RCSB Protein Data Bank with accession code 6FQD.

HDX-MS

Preparation of samples for HDX analysis was aided by a two-arm robotic autosampler (LEAP Technologies). 7.5 μL of 50 μM FtsY 6His-NGΔN$_1$, 6His-NGΔN$_1$(6mut) or $N_{2–4}$ was mixed with 67.5 μL of D_2O-containing SEC buffer and incubated at 25 °C. After 15/30/60/120/600 s, 55 μL of the HDX reaction was added to 55 μL of ice-cold quench buffer [400 mM KH_2PO_4/H_3PO_4, 2 M guanidine–HCl (pH 2.2)] and 95 μL of the mixture injected into an ACQUITY UPLC M-class system with HDX technology (Waters) [59]. Online digestion was carried out with immobilized porcine pepsin or protease type XIII from *Aspergillus saitoi* [60] at 12 °C at 100 μL/min flow rate of water +0.1% formic acid and the resulting peptides were trapped on a C18 column kept at 0.5 °C. After 3 min, the trap column was placed in line with an ACQUITY UPLC BEH C18 1.7-μm 1.0 × 100-mm column (Waters) and the peptides were eluted at 0.5 °C using a gradient of water +0.1% formic acid (A) and acetonitrile +0.1% formic acid (B) at 30 μL/min flow rate: 0–7 min/95%–65% A, 7–8 min/65%–15% A, 8–10 min/15% A, 10–11 min/5% A, and 11–16 min/95% A. Mass spectra of deuterated FtsY were acquired in High-Definition MS (HDMS)-positive ion mode by a G2-Si HDMS mass spectrometer equipped with an ESI source (Waters). Continuous lock mass correction was performed using [Glu1]-Fibrinopeptide B standard (m/z = 785.8427; Waters). Undeuterated FtsY was prepared similar by dilution of the sample in H_2O-containing SEC buffer and mass spectra were recorded in Enhanced HDMS (HDMS^{E}) mode [61,62]. All measurements were performed in triplicates. The pepsin column was washed three times with 80 μL of 4% (v/v) acetonitrile and 0.5 M guanidine hydrochloride during each run and blank runs performed between each sample to avoid peptide carry-over. Peptide identification and determination of deuterium uptake was carried out as described previously [63,64] aided by the PLGS and DynamX 3.0 softwares (Waters). Data from pepsin and protease type XIII digestion were analyzed separately.

Acknowledgments

We thank members of the E.B. and G.B. laboratories for helpful discussions. We thank E. Michel and G. Stier for plasmids pEM(GB1) and pET(6His-TEV), respectively. This work was supported by a grant from the Israel Science foundation (ISF 600/11) and the Erica Drake Fund to E.B., and the LOEWE excellence initiative of the state of Hessen, Germany to G.B. D.M. was supported by the Peter and Traudl Engelhorn foundation. We acknowledge great support provided by the ESRF, Grenoble, France. G.B. thanks the Deutsche Forschungsgemeinschaft (DFG) for the installation grant INST160/621-1FUGG and the DFG core facility for interactions, dynamics and macromolecular assembly structure (BA 5311/6-1). G.B. and D.M. further acknowledge CRC-TRR174 of the DFG.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2018.04.017.

Received 18 February 2018;
Received in revised form 17 April 2018;
Accepted 18 April 2018
Available online xxxx

Keywords:
SRP receptor FtsY;
co-translational targeting;
co-translational folding;
ribosome targeting;

†A.K. and D.M. contributed equally to this work.

Abbreviations used:
SRP, signal recognition particle; TIs, translation intermediates; LRE, lipid-responsive element; TCS, thrombin cleavage site; Cu/phe, Cu-phenanthroline; HDX-MS, hydrogen–deuterium exchange mass spectrometry; N_4-C, C-terminus of $N_{2–4}$; RF, restriction-free; SEC, size exclusion chromatography; MR, molecular replacement.

References

[1] J. Frydman, Folding of newly translated proteins in vivo: the role of molecular chaperones, Annu. Rev. Biochem. 70 (2001) 603–647.

[2] F.U. Hartl, A. Bracher, M. Hayer-Hartl, Molecular chaperones in protein folding and proteostasis, Nature 475 (2011) 324–332.

[3] A.A. Komar, A pause for thought along the co-translational folding pathway, Trends Biochem. Sci. 34 (2009) 16–24.

[4] O.B. Nilsson, A.A. Nickson, J.J. Hollins, S. Wickles, A. Steward, R. Beckmann, et al., Cotranslational folding of spectrin domains via partially structured states, Nat. Struct. Mol. Biol. 24 (2017) 221–225.

[5] F. Trovato, E.P. O'Brien, Insights into cotranslational nascent protein behavior from computer simulations, Annu. Rev. Biophys. 45 (2016) 345–369.

[6] E. Bibi, Early targeting events during membrane protein biogenesis in *Escherichia coli*, Biochim. Biophys. Acta 1808 (2011) 841–850.

[7] P. Walter, A.E. Johnson, Signal sequence recognition and protein targeting to the endoplasmic reticulum membrane, Annu. Rev. Cell Biol. 10 (1994) 87–119.

[8] P. Grudnik, G. Bange, I. Sinning, Protein targeting by the signal recognition particle, Biol. Chem. 390 (2009) 775–782.

[9] E.A. Costa, K. Subramanian, J. Nunnari, J.S. Weissman, Defining the physiological role of SRP in protein-targeting efficiency and specificity, Science 359 (2018) 689–692.

[10] D. Braig, M. Mircheva, I. Sachelaru, E.O. van der Sluis, L. Sturm, R. Beckmann, et al., Signal-sequence-independent SRP-SR complex formation at the membrane suggests an alternative targeting pathway within the SRP cycle, Mol. Biol. Cell 22 (2011) 2309–2323.

[11] J.D. Miller, H.D. Bernstein, P. Walter, Interaction of *E. coli* Ffh/4.5S ribonucleoprotein and FtsY mimics that of mammalian signal recognition particle and its receptor, Nature 367 (1994) 657–659.

[12] S. Angelini, S. Deitermann, H.G. Koch, FtsY, the bacterial signal-recognition particle receptor, interacts functionally and physically with the SecYEG translocon, EMBO Rep. 6 (2005) 476–481.

[13] A. Draycheva, T. Bornemann, S. Ryazanov, N.A. Lakomek, W. Wintermeyer, The bacterial SRP receptor, FtsY, is activated on binding to the translocon, Mol. Microbiol. 102 (2016) 152–167.

[14] P. Kuhn, A. Draycheva, A. Vogt, N.A. Petriman, L. Sturm, F. Drepper, et al., Ribosome binding induces repositioning of the signal recognition particle receptor on the translocon, J. Cell Biol. 211 (2015) 91–104.

[15] A.A. Herskovits, E. Shimoni, A. Minsky, E. Bibi, Accumulation of endoplasmic membranes and novel membrane-bound ribosome-signal recognition particle receptor complexes in *Escherichia coli*, J. Cell Biol. 159 (2002) 403–410.

[16] E.C. Mandon, Y. Jiang, R. Gilmore, Dual recognition of the ribosome and the signal recognition particle by the SRP receptor during protein targeting to the endoplasmic reticulum, J. Cell Biol. 162 (2003) 575–585.

[17] H.D. Bernstein, M.A. Poritz, K. Strub, P.J. Hoben, S. Brenner, P. Walter, Model for signal sequence recognition from amino-acid sequence of 54K subunit of signal recognition particle, Nature 340 (1989) 482–486.

[18] K. Romisch, J. Webb, J. Herz, S. Prehn, R. Frank, M. Vingron, et al., Homology of 54K protein of signal-recognition particle, docking protein and two *E. coli* proteins with putative GTP-binding domains, Nature 340 (1989) 478–482.

[19] A. Eitan, E. Bibi, The core *Escherichia coli* signal recognition particle receptor contains only the N and G domains of FtsY, J. Bacteriol. 186 (2004) 2492–2494.

[20] P. Kuhn, B. Weiche, L. Sturm, E. Sommer, F. Drepper, B. Warscheid, et al., The bacterial SRP receptor, SecA and the ribosome use overlapping binding sites on the SecY translocon, Traffic 12 (2011) 563–578.

[21] P.F. Egea, S.O. Shan, J. Napetschnig, D.F. Savage, P. Walter, R.M. Stroud, Substrate twinning activates the signal recognition particle and its receptor, Nature 427 (2004) 215–221.

[22] S.F. Ataide, N. Schmitz, K. Shen, A. Ke, S.O. Shan, J.A. Doudna, et al., The crystal structure of the signal recognition particle in complex with its receptor, Science 331 (2011) 881–886.

[23] R. Parlitz, G. Stjepanovic, L. Bahari, G. Bange, E. Bibi, I. Sinning, *Escherichia coli* signal recognition particle receptor FtsY contains an essential and autonomous membrane-binding amphipathic helix, J. Biol. Chem. 282 (2007) 32176–32184.

[24] K. Wild, G. Bange, D. Motiejunas, J. Kribelbauer, A. Hendricks, B. Segnitz, et al., Structural basis for conserved regulation and adaptation of the signal recognition particle targeting complex, J. Mol. Biol. 428 (2016) 2880–2897.

[25] L. Bahari, A. Eitan, G. Stjepanovic, E.S. Bochkareva, I. Sinning, E. Bibi, Membrane targeting of ribosomes and their release require distinct and separable functions of FtsY, J. Biol. Chem. 282 (2007) 32168–32175.

[26] D. Braig, C. Bar, J.O. Thumfart, H.G. Koch, Two cooperating helices constitute the lipid-binding domain of the bacterial SRP receptor, J. Mol. Biol. 390 (2009) 401–413.

[27] E. de Leeuw, K. te Kaat, C. Moser, G. Menestrina, R. Demel, B. de Kruijff, et al., Anionic phospholipids are involved in membrane association of FtsY and stimulate its GTPase activity, EMBO J. 19 (2000) 531–541.

[28] E. Erez, G. Stjepanovic, A.M. Zelazny, B. Brugger, I. Sinning, E. Bibi, Genetic evidence for functional interaction of the *Escherichia coli* signal recognition particle receptor with acidic lipids in vivo, J. Biol. Chem. 285 (2010) 40508–40514.

[29] T.H. Szeto, S.L. Rowland, C.L. Habrukowich, G.F. King, The MinD membrane targeting sequence is a transplantable lipid-binding helix, J. Biol. Chem. 278 (2003) 40050–40056.

[30] G. Stjepanovic, K. Kapp, G. Bange, C. Graf, R. Parlitz, K. Wild, et al., Lipids trigger a conformational switch that regulates signal recognition particle (SRP)-mediated protein targeting, J. Biol. Chem. 286 (2011) 23489–23497.

[31] A. Bercovich-Kinori, E. Bibi, Co-translational membrane association of the *Escherichia coli* SRP receptor, J. Cell Sci. 128 (2015) 1444–1452.

[32] E. Bibi, Is there a twist in the *Escherichia coli* signal recognition particle pathway? Trends Biochem. Sci. 37 (2012) 1–6.

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

[33] A.A. Herskovits, E.S. Bochkareva, E. Bibi, New prospects in studying the bacterial signal recognition particle pathway, Mol. Microbiol. 38 (2000) 927–939.
[34] J.S. Millman, F. Vulcu, H.D. Bernstein, D.W. Andrews, FtsY binds to the *Escherichia coli* inner membrane via interactions with phosphatidylethanolamine and membrane proteins, J. Biol. Chem. 276 (2001) 25982–25989.
[35] J.R. Huth, C.A. Bewley, B.M. Jackson, A.G. Hinnebusch, G.M. Clore, A.M. Gronenborn, Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR, Protein Sci. 6 (1997) 2359–2364.
[36] J.S. Millman, D.W. Andrews, A site-specific, membrane-dependent cleavage event defines the membrane binding domain of FtsY, J. Biol. Chem. 274 (1999) 33227–33234.
[37] A. Kuhn, H.G. Koch, R.E. Dalbey, Targeting and insertion of membrane proteins, EcoSal Plus 7 (2017).
[38] S. Cristobal, P. Scotti, J. Luirink, G. von Heijne, J.W. de Gier, The signal recognition particle-targeting pathway does not necessarily deliver proteins to the sec-translocase in *Escherichia coli*, J. Biol. Chem. 274 (1999) 20068–20070.
[39] J.C. Young, D.W. Andrews, The signal recognition particle receptor alpha subunit assembles co-translationally on the endoplasmic reticulum membrane during an mRNA-encoded translation pause in vitro, EMBO J. 15 (1996) 172–181.
[40] T. Welte, R. Kudva, P. Kuhn, L. Sturm, D. Braig, M. Muller, et al., Promiscuous targeting of polytopic membrane proteins to SecYEG or YidC by the *Escherichia coli* signal recognition particle, Mol. Biol. Cell 23 (2012) 464–479.
[41] T. Unger, Y. Jacobovitch, A. Dantes, R. Bernheim, Y. Peleg, Applications of the restriction free (RF) cloning procedure for molecular manipulations and protein expression, J. Struct. Biol. 172 (2010) 34–44.
[42] F. Cymer, R. Hedman, N. Ismail, G. von Heijne, Exploration of the arrest peptide sequence space reveals arrest-enhanced variants, J. Biol. Chem. 290 (2015) 10208–10215.
[43] E.S. Bochkareva, A.S. Girshovich, E. Bibi, Identification and characterization of the *Escherichia coli* stress protein UP12, a putative in vivo substrate of GroEL, Eur. J. Biochem. 269 (2002) 3032–3040.
[44] E. Michel, K. Wuthrich, High-yield *Escherichia coli*-based cell-free expression of human proteins, J. Biomol. NMR 53 (2012) 43–51.
[45] J.S. Schuhmacher, F. Rossmann, F. Dempwolff, C. Knauer, F. Altegoer, W. Steinchen, et al., MinD-like ATPase FlhG effects location and number of bacterial flagella during C-ring assembly, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 3092–3097.
[46] A.A. Herskovits, E. Bibi, Association of *Escherichia coli* ribosomes with the inner membrane requires the signal recognition particle receptor but is independent of the signal recognition particle, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 4621–4626.
[47] O. Tirosh, N. Sigal, A. Gelman, N. Sahar, N. Fluman, S. Siemion, et al., Manipulating the drug/proton antiport stoichiometry of the secondary multidrug transporter MdfA, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 12473–12478.
[48] T.G. Battye, L. Kontogiannis, O. Johnson, H.R. Powell, A.G. Leslie, iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM, Acta Crystallogr. D Biol. Crystallogr. 67 (2011) 271–281.
[49] P. Evans, Scaling and assessment of data quality, Acta Crystallogr. D Biol. Crystallogr. 62 (2006) 72–82.
[50] S. French, K. Wilson, Treatment of Negative Intensity Observations, Acta Crystallogr. A 34 (1978) 517–525.
[51] A.J. McCoy, Solving structures of protein complexes by molecular replacement with Phaser, Acta Crystallogr. D Biol. Crystallogr. 63 (2007) 32–41.
[52] G.N. Murshudov, A.A. Vagin, E.J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, Acta Crystallogr. D Biol. Crystallogr. 53 (1997) 240–255.
[53] P. Emsley, K. Cowtan, Coot: model-building tools for molecular graphics, Acta Crystallogr. D Biol. Crystallogr. 60 (2004) 2126–2132.
[54] V.B. Chen, W.B. Arendall III, J.J. Headd, D.A. Keedy, R.M. Immormino, G.J. Kapral, et al., MolProbity: all-atom structure validation for macromolecular crystallography, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 12–21.
[55] W. Kabsch, Xds, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 125–132.
[56] P.R. Evans, G.N. Murshudov, How good are my data and what is the resolution? Acta Crystallogr. D Biol. Crystallogr. 69 (2013) 1204–1214.
[57] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, Phaser crystallographic software, J. Appl. Crystallogr. 40 (2007) 658–674.
[58] P.D. Adams, P.V. Afonine, G. Bunkoczi, V.B. Chen, I.W. Davis, N. Echols, et al., PHENIX: a comprehensive Python-based system for macromolecular structure solution, Acta Crystallogr. D Biol. Crystallogr. 66 (2010) 213–221.
[59] T.E. Wales, K.E. Fadgen, G.C. Gerhardt, J.R. Engen, High-speed and high-resolution UPLC separation at zero degrees Celsius, Anal. Chem. 80 (2008) 6815–6820.
[60] L. Cravello, D. Lascoux, E. Forest, Use of different proteases working in acidic conditions to improve sequence coverage and resolution in hydrogen/deuterium exchange of large proteins, Rapid Commun. Mass Spectrom. 17 (2003) 2387–2393.
[61] S.J. Geromanos, J.P. Vissers, J.C. Silva, C.A. Dorschel, G.Z. Li, M.V. Gorenstein, et al., The detection, correlation, and comparison of peptide precursor and product ions from data independent LC-MS with data dependant LC-MS/MS, Proteomics 9 (2009) 1683–1695.
[62] G.Z. Li, J.P. Vissers, J.C. Silva, D. Golick, M.V. Gorenstein, S.J. Geromanos, Database searching and accounting of multiplexed precursor and product ion spectra from the data independent analysis of simple and complex peptide mixtures, Proteomics 9 (2009) 1696–1719.
[63] S. Schaper, W. Steinchen, E. Krol, F. Altegoer, D. Skotnicka, L. Sogaard-Andersen, et al., AraC-like transcriptional activator CuxR binds c-di-GMP by a PilZ-like mechanism to regulate extracellular polysaccharide production, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) E4822-E31.
[64] W. Steinchen, J.S. Schuhmacher, F. Altegoer, C.D. Fage, V. Srinivasan, U. Linne, et al., Catalytic mechanism and allosteric regulation of an oligomeric (p)ppGpp synthetase by an alarmone, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 13348–13353.

Please cite this article as: A. Karniel, et al., Co-translational Folding Intermediate Dictates Membrane Targeting of the Signal Recognition Particle Receptor, J. Mol. Biol. (2018), https://doi.org/10.1016/j.jmb.2018.04.017

4 Material and Methods

4.1 Materials

4.1.1 Chemicals, enzymes, combustibles

Chemicals were purchased from Sigma Aldrich, Roth and AppliChem in highest purity available. Chemicals were used as received without further purification except where stated.

Combustible laboratory equipment (1.5/2.0 ml reaction tubes, 15/50 ml Falcon tubes, pipette tips as well as syringes) was supplied by Sarstedt and Braun. Other equipment (pipettes, heating block, vortexers and power supplies) was purchased from Neolab, Eppendorf, Biozym. Electronic pipettes were purchased from Eppendorf.

4.1.2 Enzymes and cloning equipment

Restriction enzymes and further reagents (dNTPs, reaction buffers) for molecular cloning and genetic manipulations were supplied by New England Biolabs. Plasmid preparation and gel extraction of amplified or plasmid DNA were performed using kits from ThermoFisher (GeneJet Miniprep Kit, GeneJet PCR-purification Kit) accord ing to the manual provided by the manufacturer. As size standard for agarose gels, GeneRuler 1 kB Plus DNA Ladder (ThermoFisher) was employed. Sequencing of all plasmids was performed by Eurofins-MWG AG.

4.1.3 Oligonucleotides for cloning

Oligonucleotides were purchased from Sigma-Aldrich and diluted with deionized water to a final concentration of 10 mM.

4.1.4 Protein biochemistry

Purified proteins were concentrated using Amicon Ultra-15 centrifugal filter units (10 k, 30 k and 50 k size exclusion) purchased from Merck Millipore. Unstained protein marker (ThermoFisher) was used as size standard for SDS-PAGEs. Ni-NTA agarose and glutathione sepharose 4B were purchased from ThermoFisher and GE Healthcare, respectively. Spin columns and other equipment for pull down experiments were supplied by MoBiTec GmbH, Göttingen.

4.1.5 Protein Crystallization

Crystallization experiments were performed at room temperature using the JCSG Core Suites I – IV (QIAGEN). Proteins were crystallized by the sitting-drop method using SWISSCI MRC 2 Well and 3 Well (Jena Bioscience) crystallization plates. 250 nL of protein solution was mixed with mother liquor in a 1:1 or 1:2 ratio and a reservoir volume of 50 (2 well) or 30 µl (3 well) was used. The experiments were set up with a Crystal Gryphon (Art Robbins) according to the manufacturers manual. Individual fine screens and additive screens were prepared in hanging or sitting drop plates (Molecular dimensions).

4.1.6 Data collection and structure determinantion

Prior data collection, crystals were flash-frozen in liquid nitrogen after incubation in a cryo- protecting solution containing either 20 % (v/v) glycerol or MPD. Crystals were harvested with Adjustable Mounted CryoLoopsTM (Hampton Research) of different diameters and data collected at the European Synchrotron Radiation Facility (ESRF, Grenoble).

The data were processed, integrated and scaled with XDS[112] and merged with the program AIMLESS from the CCP4 suite.[113] The resolution cutoff was determined with the program AIMLESS and according to several values from the CORRECT-file obtained by the XDS-program. Cross validation of refinement was performed with 5% of the total reflections (Rfree).

The merged-MTZ file was either used for molecular replacement (MR) using Phaser from Phenix[114] or experimental phasing.[115] Refinement was performed with PHENIX.refine[116] and models manually built and corrected with COOT.[117] Figures containing crystal structures or superimpositions of crystal structures were designed with PyMol (www.pymol.org).

4.1.7 Plasmids

Various plasmids were used in the scope of this work for different purposes. Firstly, pET24d (Kanamycin resistance) and pET16b (Ampicillin resistance) served as vectors for protein production of hexahistidine-tagged proteins, which also allowed co-production of different proteins due to different resistance markers. N-

terminal GST-fusions were generated using pGAT3 (Ampicillin resistance). N-terminal GB1-fusions were generated using pEM GB1 (Ampicillin resistance).

4.1.8 *E. coli* strains

Large-scale protein production for crystallography and biochemical assays was carried out in phage resistant, chemically competent *E. coli* BL21 (DE3) (Life technologies). For plasmid amplification chemically competent *E. coli* DH5α (Life technologies) were employed.

4.1.9 Buffers and growth media

E. coli was cultured in LB broth (Roth). The media was sterilized before usage.
list of buffers

For Ni-affinity chromatography, buffer A was used: 20 mM HEPES, pH 8.0 250 mM NaCl, 20 mM KCl, 20 mM MgCl2, 40 mM imidazole.
For elution of proteins from the resin, buffer B was used (as buffer A, but 500 mM imidazole).

For size-exclusion chromatography, SEC-buffer was used:
20 mM HEPES, pH 7.5, 200 mM NaCl, 20 mM KCl, 20 mM MgCl2

For GST-interaction assays, SEC-buffer was employed.

GST-tagged proteins were eluted from glutathione sepharose or from a GST-trap column employing GSH elution buffer:
50 mM TRIS, pH 8.0, 20 mM glutathione

4.2 Methods

4.2.1 Molecular cloning

PCR amplification of each gene was performed using Phusion® High-Fidelity DNA Polymerase (NEB) according to the manufacturer's protocol. 100 µM dNTP's (NEB), 200 µM of each Oligonucleotide and 0.01 U/µl polymerase was used to set up a PCR reaction. An estimation of the optimal annealing temperature for

each oligonucleotide was calculated by the webpage OligoCalc (http://biotools.nubic.northwestern.edu/OligoCalc.html). PCR-reactions were carried out in PCR-cyclers (Biorad, Techne). The reaction program was used according to the manufacturers manual.

4.2.2 Plasmid preparation

Transformation of chemically competent *E. coli* strains was performed with 50 – 200 ng of plasmid and 100 µl of cells. The transformation was as follows: Cells and DNA incubated at 4°C (10 min), heatshock 42°C (1 min), regeneration with 150 µL LB-media at 37°C and shaking (30 min). The whole reaction was transferred to selective media and incubated at 37 °C overnight.

To extract plasmid DNA, 5 mL cultures of *E. coli* were inoculated with a single colony and incubated overnight at 37 °C and 200 rpm. The cells were harvested by centrifugation at 5000 g (5 min) and further processed according to the manufacturers manual supplied with the GeneJET Plasmid Miniprep Kit (Thermo Scientific). The plasmid DNA was finally eluted in 50 µl deionized water.

4.2.3 Agarose gel electrophoresis

DNA restriction fragments and PCR-amplificons were analyzed by agarose gel electrophoresis. Depending on the DNA fragment size, gels were prepared with an agarose concentration ranging from 0.8 % to 2 % (w/v). The agarose was dissolved in running buffer (90 mM Tris, 30 mM Taurine, 0.5 mM EDTA) by microwave heating and poured into gel casts. 5 µL of a 0.025 % ethidium bromide solution (Roth) was added to 100 ml of gel solution. Gel Loading Dye, Purple (NEB) was added in appropriate amounts to the samples prior to loading. A voltage of 100 V was applied and DNA visualized with a UV-imager (INTAS).

4.2.4 Protein production and purification

For gene expression, *E. coli* BL21(DE3) were grown in LB medium under autoinduction conditions [D(+)-lactose-monohydrate, 1.75% (w/v)] supplemented with the respective antibiotics (50 µg/mL kanamycin or 100 µg/mL ampicillin) at 30 °C for 16 h under constant shaking (180 rpm). After cell lysis by a Microfluidizer

(M110-L, Microfluidics), cell debris was removed by high-speed centrifugation and proteins were purified by Ni-ion affinity- and size exclusion chromatography (SEC).

4.2.5 SDS-PAGE

Sodium-dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) was performed to separate proteins according to their size. 15% polyacrylamide gels were prepared using a Mini-PROTEAN 3 Multi-Casting Chamber (Biorad) and stored at 4 °C. A 5x loading buffer (100 mM Tris-HCl (pH 6.3), 10% (v/v) glycerol, 2 mg/ml SDS, 3% (v/v) β-mercaptoethanol, 1 mg/ml brom-phenol blue) was added to samples for SDS-PAGE analysis. Samples containing whole cell extracts or ribosomes were denatured prior to electrophoresis at 98 °C (10 min). Gels were run at 270 V (40 min, Mini-PROTEAN Tetra Cell, Biorad). Coomassie-staining was performed with Coomassie Brilliant Blue R-250 (Roth). For destaining, a solution containing 60% H2O, 30% ethanol and 10% AcOH was used.

4.2.6 Glutathione-S-transferase (GST) binding assays

GST pulldown assays were performed in HEPES buffer (20 mM HEPES, 150 mM NaCl, 20 mM $MgCl_2$, 20 mM KCl, pH 7.5) at 4°C. GST-protein (1 nmol) was applied to 30 µL glutathione Sepharose 4B (GE Healthcare) in small filter columns (MoBiTec GmbH) by incubation on a wheel for 15 min. Thereafter, the column was washed once with HEPES-buffer. Subsequently, 10 nmol of binding partners were added and incubated for 15 min at 4°C in the wheel. After centrifugation (4000 rpm, 1 min in a table-top centrifuge) the column was washed 3 times with HEPES-buffer. Proteins were eluted with 50 µL of GSH-buffer (50 mM Tris, 20 mM GSH, pH 8) and analyzed by coomassie-stained SDS-PAGE.

4.2.7 Ni-NTA affinity binding assays

For Ni-pulldowns, hexahistine-tagged protein was co-expressed with untagged putative interaction partners. Cell lysate was incubated with Ni-sepharose at 4 °C for 20 min, then the resin was washed with buffer A (20 mM HEPES, 150 mM NaCl, 20 mM $MgCl_2$, 20 mM KCl, 40 mM imidazole, pH 8.0) and bound proteins were eluted with buffer B (like buffer A, but 500 mM imidazole) and analyzed by Coomassie-stained SDS-PAGE.

4.2.8 Hydrolysis assays

ATP hydrolysis was investigated using an HPLC-based assay. FlhG (50 µM) was supplemented with ATP (1 mM) and interaction partner (50 µM) in SEC-buffer, incubated at 30°C (30 min). Addition of chloroform, vortexing and subsequent flash freezing in liquid nitrogen stopped the hydrolysis reaction. For HPLC measurements, the aqueos phase was analyzedwith an Agilent 1100 Series HPLC system and a C18 column (Macherey- Nagel). The samples were injected onto HPLC and run for 30 min with a buffer containing 50 mM KH2PO4, 50 mM K2HPO4, 10 mM tetrapentylammonium bromide (TPAB) and 15 % (v/v) acetonitrile at 0.8 ml/min flow rate. ADP and ATP were detected by UV (260.8 nm) and quantified (by peak area) using ChemStation (version: B.04.03; Agilent technologies).

4.2.9 Analytical Size-exclusion chromatography

For analytical SEC, proteins were prepared in a defined concentration (100 µL), mixed with 100 µL SEC-buffer and injected (500 µL Loop) on a Äkta Purifier or Pure system, connected to a Superdex200 10/300 GL column. Flowrate: 0.5 mL/min.

4.2.10 Atomic Force Microscopy

Briefly, long and short DNA segments of the opuA-promoter region (986, 186 bp) were amplified by PCR and purified by gel extraction (Dr. Tamara Hoffmann). The so obtained DNA was diluted in deposition buffer (20 mM Hepes pH7.4, 5 mM $MgCl_2$) to 10 nM. 5 µL of this dilution were pipetted in the center of freshly cleaved mica (ca 10x10mm) and allowed to deposit for 10 minutes. The mica-surface was then rinsed with 200 µL ddH_2O and dried under a nitrogen-stream. The film morphology was characterized by atomic force microscopy (AFM, Agilent SPM 5500) operated in tapping mode at ambient conditions (Darius Günder in the group of Prof. Witte, Fachbereich Physik Universität Marburg).

4.2.11 Microscale Thermophoresis

MST was performed on a Monolith NT.115 (NanoTemper Technologies GmbH, Munich, Germany) at 21°C (red LED power was set to 70% and infrared laser

power to 25%).[118] Target protein (50 µM) was labeled with the dye NT 647 according to the supplier's protocol (NanoTemper Technologies). 200 nM of the labeled target protein was titrated with putative interaction partner starting from a concentration of 0.5 mM in Buffer C (20 mM HEPES-Na, pH 7.5, 200 mM NaCl). To each measurement, Tween20 (Sigma) was added to a final concentration of 0.05 mM. At least nine independent MST experiments were recorded at 680 nm and processed by NanoTemper Analysis 1.2.009. For fitting of the experimental data and K_d determination Origin8G was used.

5 Bibliography

1. Bardy, S. L. Prokaryotic motility structures. *Microbiology* **149,** 295–304 (2003).
2. Harshey, R. M. Bacterial Motility on a Surface: Many Ways to a Common Goal. *Annu. Rev. Microbiol.* **57,** 249–273 (2003).
3. Chevance, F. F. V & Hughes, K. T. Coordinating assembly of a bacterial macromolecular machine. *Nat. Rev. Microbiol.* **6,** 455–465 (2008).
4. Macnab, R. M. How Bacteria Assemble Flagella. *Annu. Rev. Microbiol.* **57,** 77–100 (2003).
5. Altegoer, F. & Bange, G. Undiscovered regions on the molecular landscape of flagellar assembly. *Curr. Opin. Microbiol.* **28,** 98–105 (2015).
6. Dasgupta, N. *et al.* A four-tiered transcriptional regulatory circuit controls flagellar biogenesis in Pseudomonas aeruginosa. *Mol. Microbiol.* **50,** 809–824 (2003).
7. Altegoer, F., Rensing, S. A. & Bange, G. Structural basis for the CsrA-dependent modulation of translation initiation by an ancient regulatory protein. *Proc. Natl. Acad. Sci.* **113,** 10168–10173 (2016).
8. Bange, G. *et al.* FlhA provides the adaptor for coordinated delivery of late flagella building blocks to the type III secretion system. *Proc. Natl. Acad. Sci.* **107,** 11295–11300 (2010).
9. Brown, P. N., Mathews, M. A. A., Joss, L. A., Hill, C. P. & Blair, D. F. Crystal Structure of the Flagellar Rotor Protein FliN from Thermotoga maritima. *J. Bacteriol.* **187,** 2890–2902. (2005).
10. Thomas, D. R., Morgan, D. G. & DeRosier, D. J. Rotational symmetry of the C ring and a mechanism for the flagellar rotary motor. *Proc. Natl. Acad. Sci.* **96,** 10134–10139 (1999).
11. Yakushi, T., Yang, J., Fukuoka, H., Homma, M. & Blair, D. F. Roles of charged residues of rotor and stator in flagellar rotation: comparative study using H+-driven and Na+-driven motors in Escherichia coli. *J. Bacteriol.* **188,** 1466–72 (2006).
12. Lynch, M. J. *et al.* Co-Folding of a FliF-FliG Split Domain Forms the Basis of the MS:C Ring Interface within the Bacterial Flagellar Motor. *Structure*

25, 317–328 (2017).

13. Xue, C. *et al.* Crystal structure of the FliF-FliG complex from Helicobacter pylori yields insight into the assembly of the motor MS-C ring in the bacterial flagellum. *J. Biol. Chem.* **293,** 2066–2078 (2018).
14. Sarkar, M. K., Paul, K. & Blair, D. F. Chemotaxis signaling protein CheY binds to the rotor protein FliN to control the direction of flagellar rotation in Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 9370–9375 (2010).
15. Schuhmacher, J. S., Thormann, K. M. & Bange, G. How bacteria maintain location and number of flagella? *FEMS Microbiol. Rev.* (2015). doi:10.1093/femsre/fuv034
16. Pallen, M. J. & Matzke, N. J. From The Origin of Species to the origin of bacterial flagella. *Nat. Rev. Microbiol.* **4,** 784–790 (2006).
17. Liu, R. & Ochman, H. Origins of Flagellar Gene Operons and Secondary Flagellar Systems. *J. Bacteriol.* **189,** 7098–7104 (2007).
18. McCarter, L. L. Multiple modes of motility: a second flagellar system in Escherichia coli. *J. Bacteriol.* **187,** 1207–9 (2005).
19. Merino, S. *et al.* Bacterial lateral flagella: an inducible flagella system. *FEMS Microbiol. Lett.* **263,** 127–135 (2006).
20. Kusumoto, A. *et al.* Regulation of polar flagellar number by the flhF and flhG genes in Vibrio alginolyticus. *J. Biochem.* **139,** 113–121 (2006).
21. Kusumoto, A. *et al.* Collaboration of FlhF and FlhG to regulate polarflagella number and localization in Vibrio alginolyticus. *Microbiology* **154,** 1390–1399 (2008).
22. Gulbronson, C. J. *et al.* FlhG employs diverse intrinsic domains and influences FlhF GTPase activity to numerically regulate polar flagellar biogenesis in Campylobacter jejuni. *Mol. Microbiol.* **99,** 291–306 (2016).
23. Kazmierczak, B. I. & Hendrixson, D. R. Spatial and numerical regulation of flagellar biosynthesis in polarly flagellated bacteria. *Mol. Microbiol.* **88,** 655–663 (2013).
24. Dasgupta, N., Arora, S. K. & Ramphal, R. fleN, a gene that regulates flagellar number in Pseudomonas aeruginosa. *J. Bacteriol.* **182,** 357–64 (2000).

25. Schuhmacher, J. S. *et al.* MinD-like ATPase FlhG effects location and number of bacterial flagella during C-ring assembly. *Proc. Natl. Acad. Sci.* **112,** 201419388 (2015).
26. Bange, G., Petzold, G., Wild, K., Parlitz, R. O. & Sinning, I. The crystal structure of the third signal-recognition particle GTPase FlhF reveals a homodimer with bound GTP. *Proc. Natl. Acad. Sci.* **104,** 13621–13625 (2007).
27. Rossmann, F. *et al.* The role of FlhF and HubP as polar landmark proteins in Shewanella putrefaciens CN-32. **98,** 727–742 (2015).
28. Green, J. C. D. *et al.* Recruitment of the Earliest Component of the Bacterial Flagellum to the Old Cell Division Pole by a Membrane-Associated Signal Recognition Particle Family GTP-Binding Protein. *J. Mol. Biol.* **391,** 679–690 (2009).
29. Raskin, D. M. & de Boer, P. A. Rapid pole-to-pole oscillation of a protein required for directing division to the middle of Escherichia coli. *Proc. Natl. Acad. Sci. U. S. A.* **96,** 4971–6 (1999).
30. Szeto, T. H., Rowland, S. L., Habrukowich, C. L. & King, G. F. The MinD membrane targeting sequence is a transplantable lipid-binding helix. *J. Biol. Chem.* **278,** 40050–40056 (2003).
31. Balaban, M. & Hendrixson, D. R. Polar flagellar biosynthesis and a regulator of flagellar number influence spatial parameters of cell division in campylobacter jejuni. *PLoS Pathog.* **7,** 18–25 (2011).
32. Stock, A. M., Robinson, V. L. & Goudreau, P. N. Two-Component Signal Transduction. *Annu. Rev. Biochem.* **69,** 183–215 (2000).
33. Prüß, B. M. Involvement of Two-Component Signaling on Bacterial Motility and Biofilm Development. (2017). doi:10.1128/JB.00259-17
34. Ritchings, B. W., Almira, E. C., Lory, S. & Ramphal, R. Cloning and phenotypic characterization of fleS and fleR, new response regulators of Pseudomonas aeruginosa which regulate motility and adhesion to mucin. *Infect. Immun.* **63,** 4868–76 (1995).
35. Arora, S. K., Ritchings, B. W., Almira, E. C., Lory, S. & Ramphal, R. A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene

expression in Pseudomonas aeruginosa in a cascade manner. *J. Bacteriol.* **179,** 5574–5581 (1997).

36. Matsuyama, B. Y. *et al.* Mechanistic insights into c-di-GMP–dependent control of the biofilm regulator FleQ from *Pseudomonas aeruginosa*. *Proc. Natl. Acad. Sci.* **113,** E209–E218 (2016).
37. Vidangos, N. K. *et al.* DNA Recognition by a σ54Transcriptional Activator from Aquifex aeolicus. *J. Mol. Biol.* **426,** 3553–3568 (2014).
38. Su, T. *et al.* The REC domain mediated dimerization is critical for FleQ from Pseudomonas aeruginosa to function as a c-di-GMP receptor and flagella gene regulator. *J. Struct. Biol.* **192,** 1–13 (2015).
39. Hickman, J. W. & Harwood, C. S. Identification of FleQ from Pseudomonas aeruginosa as a c-di-GMP-responsive transcription factor. *Mol. Microbiol.* **69,** 376–389 (2008).
40. Cheng, Y.-Y. *et al.* FlrA Represses Transcription of the Biofilm-Associated bpfA Operon in Shewanella putrefaciens. *Appl. Environ. Microbiol.* **83,** e02410-16 (2017).
41. Chen, B. *et al.* Engagement of Arginine Finger to ATP Triggers Large Conformational Changes in NtrC1 AAA+ ATPase for Remodeling Bacterial RNA Polymerase. *Structure* **18,** 1420–1430 (2010).
42. Dasgupta, N. & Ramphal, R. Interaction of the antiactivator FleN with the transcriptional activator FleQ regulates flagellar number in pseudomonas aeruginosa. *J. Bacteriol.* **183,** 6636–6644 (2001).
43. Chanchal, Banerjee, P. & Jain, D. ATP-Induced Structural Remodeling in the Antiactivator FleN Enables Formation of the Functional Dimeric Form. *Structure* **25,** 243–252 (2017).
44. Huth, J. R. *et al.* Design of an expression system for detecting folded protein domains and mapping macromolecular interactions by NMR. *Protein Sci.* **6,** 2359–64 (1997).
45. Harper, S. & Speicher, D. W. Purification of Proteins Fused to Glutathione S-Transferase. in 259–280 (2011). doi:10.1007/978-1-60761-913-0_14
46. Mehla, J., Caufield, J. H., Sakhawalkar, N. & Uetz, P. A comparison of two hybrid approaches for detecting protein-protein interactions. *Methods*

Enzymol. **586,** 333 (2017).

47. Vidangos, N. *et al.* Structure, function, and tethering of DNA-binding domains in ??54 transcriptional activators. *Biopolymers* **99,** 1082–1096 (2013).
48. Bush, M. & Dixon, R. The Role of Bacterial Enhancer Binding Proteins as Specialized Activators of 54-Dependent Transcription. *Microbiol. Mol. Biol. Rev.* **76,** 497–529 (2012).
49. Shi, M., Gao, T., Ju, L., Yao, Y. & Gao, H. Effects of FlrBC on flagellar biosynthesis of Shewanella oneidensis. *Mol. Microbiol.* **93,** 1269–1283 (2014).
50. Baraquet, C., Murakami, K., Parsek, M. R. & Harwood, C. S. The FleQ protein from Pseudomonas aeruginosa functions as both a repressor and an activator to control gene expression from the Pel operon promoter in response to c-di-GMP. *Nucleic Acids Res.* **40,** 7207–7218 (2012).
51. Di Ventura, B. *et al.* Chromosome segregation by the Escherichia coli Min system. *Mol. Syst. Biol.* **9,** 1–12 (2013).
52. Delalez, N. J., Berry, R. M. & Armitage, J. P. Stoichiometry and turnover of the bacterial flagellar switch protein FliN. *MBio* **5,** e01216-14 (2014).
53. Delalez, N. J. *et al.* Signal-dependent turnover of the bacterial flagellar switch protein FliM. *Proc. Natl. Acad. Sci. U. S. A.* **107,** 11347–51 (2010).
54. Diepold, A., Kudryashev, M., Delalez, N. J., Berry, R. M. & Armitage, J. P. Composition, Formation, and Regulation of the Cytosolic C-ring, a Dynamic Component of the Type III Secretion Injectisome. *PLOS Biol.* **13,** e1002039 (2015).
55. Hosu, B. G. & Berg, H. C. CW and CCW Conformations of the E. coli Flagellar Motor C-Ring Evaluated by Fluorescence Anisotropy. *Biophys. J.* **114,** 641–649 (2018).
56. Altegoer, F. *et al.* FliS/flagellin/FliW heterotrimer couples type III secretion and flagellin homeostasis. *Sci. Rep.* **8,** 11552 (2018).
57. Aravind, L., Anantharaman, V., Balaji, S., Babu, M. M. & Iyer, L. M. The many faces of the helix-turn-helix domain: Transcription regulation and beyond. *FEMS Microbiol. Rev.* **29,** 231–262 (2005).

58. Somers, W. S. & Phillips, S. E. V. Crystal structure of the met repressor-operator complex at 2.8 Å resolution reveals DNA recognition by β-strands. *Nature* **359,** 387–393 (1992).
59. Schreiter, E. R. & Drennan, C. L. Ribbon-helix-helix transcription factors: Variations on a theme. *Nat. Rev. Microbiol.* **5,** 710–720 (2007).
60. Swinger, K. K. & Rice, P. A. IHF and HU: Flexible architects of bent DNA. *Curr. Opin. Struct. Biol.* **14,** 28–35 (2004).
61. Galperin, M. Y. Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *J. Bacteriol.* **188,** 4169–82 (2006).
62. Zou, Z. *et al.* LytTR Regulatory Systems: A potential new class of prokaryotic sensory system. (2018). doi:10.1371/journal.pgen.1007709
63. Sidote, D. J., Barbieri, C. M., Wu, T. & Stock, A. M. Structure of the Staphylococcus aureus AgrA LytTR Domain Bound to DNA Reveals a Beta Fold with an Unusual Mode of Binding. *Structure* **16,** 727–735 (2008).
64. Novick, R. P. Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol. Microbiol.* **48,** 1429–1449 (2003).
65. Koenig, R. L. *et al.* Staphylococcus aureus AgrA Binding to the RNAIII-agr Regulatory Region. *J. Bacteriol.* **186,** 7549–7555 (2004).
66. Okkotsu, Y., Little, A. S. & Schurr, M. J. The Pseudomonas aeruginosa AlgZR two-component system coordinates multiple phenotypes. *Front. Cell. Infect. Microbiol.* **4,** 82 (2014).
67. Crespo, A., Pedraz, L., Van Der Hofstadt, M., Gomila, G. & Torrents, E. Regulation of ribonucleotide synthesis by the Pseudomonas aeruginosa two-component system AlgR in response to oxidative stress. *Sci. Rep.* **7,** 1–15 (2017).
68. Reyes, D. *et al.* Coordinated Regulation by AgrA, SarA, and SarR To Control agr Expression in Staphylococcus aureus. *J. Bacteriol.* **193,** 6020–6031 (2011).
69. Leonard, P. G., Bezar, I. F., Sidote, D. J. & Stock, A. M. Identification of a hydrophobic cleft in the LytTR domain of AgrA as a locus for small molecule interactions that inhibit DNA binding. *Biochemistry* **51,** 10035–10043

(2012).

70. Mouratou, B. *et al. Remodeling a DNA-binding protein as a specific in vivo inhibitor of bacterial secretin PulD*. (2007).
71. Guttenplan, S. B., Blair, K. M. & Kearns, D. B. The EpsE Flagellar Clutch Is Bifunctional and Synergizes with EPS Biosynthesis to Promote Bacillus subtilis Biofilm Formation. *PLoS Genet.* **6,** e1001243 (2010).
72. Blair, K. M., Turner, L., Winkelman, J. T., Berg, H. C. & Kearns, D. B. A Molecular Clutch Disables Flagella in the Bacillus subtilis Biofilm. *Science (80-.).* **1636,** 1636–1639 (2009).
73. Guttenplan, S. B. & Kearns, D. B. Regulation of flagellar motility during biofilm formation. *FEMS Microbiol. Rev.* **37,** 849–71 (2013).
74. Hobley, L. *et al.* BslA is a self-assembling bacterial hydrophobin that coats the Bacillus subtilis biofilm. *Proc. Natl. Acad. Sci.* **110,** 13600–13605 (2013).
75. Mandic-Mulec, I., Doukhan, L. & Smith, I. *The Bacillus subtilis SinR Protein Is a Repressor of the Key Sporulation Gene spo0A*. *JOURNAL OF BACTERIOLOGY* **177,** (1995).
76. Mielich-Süss, B. & Lopez, D. Molecular mechanisms involved in Bacillus subtilis biofilm formation. *Environ. Microbiol.* **17,** 555–565 (2015).
77. Cairns, L. S., Hobley, L. & Stanley-Wall, N. R. Biofilm formation by Bacillus subtilis: New insights into regulatory strategies and assembly mechanisms. *Mol. Microbiol.* **93,** 587–598 (2014).
78. Newman, J. A., Rodrigues, C. & Lewis, R. J. Molecular Basis of the Activity of SinR Protein, the Master Regulator of Biofilm Formation in Bacillus subtilis *. (2013). doi:10.1074/jbc.M113.455592
79. Colledge, V. L. *et al.* Structure and Organisation of SinR, the Master Regulator of Biofilm Formation in Bacillus subtilis. *J. Mol. Biol.* **411,** 597–613 (2011).
80. Winkelman, J. T., Blair, K. M. & Kearns, D. B. RemA (YlzA) and RemB (YaaB) regulate extracellular matrix operon expression and biofilm formation in Bacillus subtilis. *J. Bacteriol.* **191,** 3981–3991 (2009).
81. Winkelman, J. T. *et al.* RemA is a DNA-binding protein that activates biofilm

matrix gene expression in Bacillus subtilis. *Mol. Microbiol.* **88,** 984–997 (2013).

82. Hoffmann, T. *et al.* Osmotic Control of opuA Expression in Bacillus subtilis and Its Modulation in Response to Intracellular Glycine Betaine and Proline Pools. *J. Bacteriol.* **195,** 510–522 (2013).
83. Buescher, J. M. *et al.* Global Network Reorganization During Dynamic Adaptations of Bacillus subtilis Metabolism. *Science (80-.).* **335,** 1099–1103 (2012).
84. Gentry, D., Bengraj, C., Ikeharall, K. & Cashel, M. *Guanylate Kinase of Escherichia coli K-12" THE JOURNAL OF BIOLOGICAL CHEMISTRY* **268,** (1993).
85. Liu, K. *et al.* Molecular mechanism and evolution of guanylate kinase regulation by (p)ppGpp. (2014). doi:10.1016/j.molcel.2014.12.037
86. Gentry, D. R. & Burgess, R. R. rpoZ, encoding the omega subunit of Escherichia coli RNA polymerase, is in the same operon as spoT. *J. Bacteriol.* **171,** 1271–7 (1989).
87. Mah, T.-F. C. & O'Toole, G. A. Mechanisms of biofilm resistance to antimicrobial agents. *Trends Microbiol.* **9,** 34–39 (2001).
88. Rubinstein, S. M. *et al.* Osmotic pressure can regulate matrix gene expression in Bacillus subtilis. (2012). doi:10.1111/j.1365-2958.2012.08201.x
89. Krämer, H. *et al.* lac repressor forms loops with linear DNA carrying two suitably spaced lac operators. *EMBO J.* **6,** 1481–1491 (1987).
90. Plumbridge, J. & Kolb, A. CAP and Nag repressor binding to the regulatory regions of the nagE-B and manX genes of Escherichia coli. *J. Mol. Biol.* **217,** 661–679 (1991).
91. Luger, K., Mäder, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389,** 251–260 (1997).
92. Amidani, D. *et al.* Study of DNA binding and bending by Bacillus subtilis GabR, a PLP-dependent transcription factor. *Biochim. Biophys. Acta - Gen. Subj.* **1861,** 3474–3489 (2017).

93. Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20,** 478–480 (1995).
94. Brill, J., Hoffmann, T., Putzer, H. & Bremer, E. T-box-mediated control of the anabolic proline biosynthetic genes of Bacillus subtilis. *Microbiology* **157,** 977–987 (2011).
95. Chen, S., Jancrick, J., Yokota, H., Kim, R. & Kim, S.-H. Crystal structure of a protein associated with cell division from Mycoplasma pneumoniae (GI: 13508053): A novel fold with a conserved sequence motif. *Proteins Struct. Funct. Bioinforma.* **55,** 785–791 (2004).
96. Rohs, R. *et al.* The role of DNA shape in protein-DNA recognition. *Nature* **461,** 1248–1253 (2009).
97. Ouhammouch, M. & Geiduschek, E. P. A thermostable platform for transcriptional regulation: the DNA-binding properties of two Lrp homologs from the hyperthermophilic archaeon Methanococcus jannaschii. *EMBO J.* **20,** 146–56 (2001).
98. Andrews, A. J., Chen, X., Zevin, A., Stargell, L. A. & Luger, K. The Histone Chaperone Nap1 Promotes Nucleosome Assembly by Eliminating Nonnucleosomal Histone DNA Interactions. *Mol. Cell* **37,** 834–842 (2010).
99. Lyubchenko, Y. L. Nanoscale Nucleosome Dynamics Assessed with Time-lapse AFM. *Biophys. Rev.* **6,** 181–190 (2014).
100. Peeters, E. & Charlier, D. The Lrp family of transcription regulators in archaea. *Archaea* **2010,** (2010).
101. Chen, C. & Pettitt, B. M. DNA Shape versus Sequence Variations in the Protein Binding Process. *Biophys. J.* **110,** 534–544 (2016).
102. Duzdevich, D., Redding, S. & Greene, E. C. DNA Dynamics and Single-Molecule Biology. *Chem. Rev.* **114,** 3072–3086 (2014).
103. Rohs, R. *et al.* Origins of Specificity in Protein-DNA Recognition. *Annu. Rev. Biochem.* **79,** 233–269 (2010).
104. Hsieh, L. S., Rouviere-Yaniv, J. & Drlica, K. Bacterial DNA supercoiling and [ATP]/[ADP] ratio: Changes associated with salt shock. *J. Bacteriol.* **173,** 3914–3917 (1991).
105. Lawson, C. L. *et al.* Catabolite activator protein: DNA binding and

transcription activation. *Curr. Opin. Struct. Biol.* **14,** 10–20 (2004).

106. Liu, B., Hong, C., Huang, R. K., Yu, Z. & Steitz, T. A. Structural basis of bacterial transcription activation. *Science (80-.).* **358,** 947–951 (2017).
107. Reynolds, J. & Wigneshweraraj, S. Molecular Insights into the Control of Transcription Initiation at the Staphylococcus aureus agr operon. *J. Mol. Biol.* **412,** 862–881 (2011).
108. Weiss, A. & Shaw, L. N. Small things considered: The small accessory subunits of RNA polymerase in Gram-positive bacteria. *FEMS Microbiol. Rev.* **39,** 541–554 (2015).
109. Weiss, A. *et al.* The ω Subunit Governs RNA Polymerase Stability and Transcriptional Specificity in Staphylococcus aureus. *J. Bacteriol.* **199,** 1–16 (2017).
110. Gurard-Levin, Z. A., Quivy, J.-P. & Almouzni, G. Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics. *Annu. Rev. Biochem.* **83,** 487–517 (2014).
111. Dyer, P. N. *et al.* Reconstitution of Nucleosome Core Particles from Recombinant Histones and DNA. *Methods Enzymol.* **375,** 23–44 (2003).
112. Kabsch, W. XDS. *Acta Crystallogr D Biol Crystallogr* **66,** 125–132 (2010).
113. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. D. Biol. Crystallogr.* **69,** 1204–14 (2013).
114. McCoy, A. J. *et al.* Phaser crystallographic software. *J Appl Crystallogr* **40,** 658–674 (2007).
115. McCoy, A. J., Storoni, L. C. & Read, R. J. Simple algorithm for a maximum-likelihood SAD function. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **60,** 1220–1228 (2004).
116. Afonine, P. V. *et al.* Towards automated crystallographic structure refinement with *phenix.refine*. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **68,** 352–367 (2012).
117. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2126–32 (2004).
118. Jerabek-Willemsen, M., Wienken, C. J., Braun, D., Baaske, P. & Duhr, S. Molecular interaction studies using microscale thermophoresis. *Assay Drug*

Dev. Technol. **9,** 342–53 (2011).

6 Appendix

6.1 Crystallization conditions and details of data acquisition for RemA variants

Table 3 Crystallization conditions and data mesurements.

Variant	Crystallization condition	Space group	Unit Cell	Resolution	Oligomeric State	Dataset recorded
GtRemA wt	1.6 M NH4SO4, 0.1 M citric acid pH 3.5, final pH 4	P23	106.5 106.5 106.5 90 90 90	2.3 Å	Octamer/16mer	ID30-A1 (10.11.2017 22:14:41), CD190A x8
GtRemA Y10I	No crystals obtained	-	-	-	-	-
GtRemA R18W	2.5 M NaCl, 0.2 M Li2SO4, 0.1 M NaOAc pH4.5	C2221	116.2 102.4 113.0 90.1 90.0 90.3	2.6 Å	Heptamer/14mer	ID23-2 (22.10.2018 00:12:03), NA1208 x8
GtRemA P29S	0.07 M NaOAc pH4.6, 5.6% (v/v) PEG4000, 30% (v/v) Glycerol	I422	78.4 78.4 171.0 90 90 90	2.2 Å	Octamer/16mer	ID23-1 (27.09.2018 06:14:45), NA1209 x5
GtRemA R32A	0.1 M NaOAc pH5.0, 10-20% MPD	-	-	>4 Å	no dataset available	-
GtRemA R50A	0.1 M Tris pH8.5, 2.0 NH4H2PO4	-	-	>4 A	no dataset available	-
GtRemA R51A	0.1 M Tris pH8.5, 2.0 NH4H2PO4	I422	89.3 89.3 109.3 90.0 90.0 90.1	1.8 Å	Octamer/16mer	ID23-1 (27.09.2018 05:43:21), NA1214 x3
GtRemA R53A	-	-	-	-	-	-
GtRemA H62Q	0.8 M N2HPO4, 0.8 M NaH2PO4, 0.1 M Hepes pH7.5	P23	104.6 104.6 104.6 90 90 90	3.2 Å	Octamer/16mer	ID29 (27.02.2018 02:25:37), CD190A x5

Table 4 Data collection and refinement statistics. Statistics for the highest-resolution shell are shown in parentheses.

	GtRemA R18W	GtRemA P29S	GtRemA R51A	GtRemA H62Q
Wavelength				
Resolution range	47.28 - 2.6 (2.693 - 2.6)	40.95 - 2.2 (2.279 - 2.2)	37.52 - 1.8 (1.864 - 1.8)	42.64 - 3.2 (3.315 - 3.2)
Space group	C 2 2 21	I 4 2 2	I 4 2 2	P 2 3
Unit cell	103.89 116.97 114.135 90 90 90	89.061 89.061 107.825 90 90 90	89.323 89.323 109.236 90 90 90	104.436 104.436 104.436 90 90 90
Total reflections	157247 (16164)	164372 (16755)	224975 (19824)	183877 (18313)
Unique reflections	21727 (2148)	11357 (1101)	20833 (2048)	6498 (629)
Multiplicity	7.2 (7.5)	14.5 (15.2)	10.8 (9.7)	28.3 (29.1)
Completeness (%)	99.58 (99.77)	99.85 (99.82)	99.88 (99.76)	99.77 (100.00)
Mean I/sigma(I)	9.10 (0.79)	14.11 (0.77)	14.71 (0.75)	21.38 (0.91)
Wilson B-factor	76.73	62.23	33.87	144.67
R-merge	0.1129 (1.786)	0.09565 (3.629)	0.093 (2.007)	0.1014 (3.918)
R-meas	0.1221 (1.918)	0.09928 (3.755)	0.09771 (2.12)	0.1033 (3.987)
R-pim	0.04578 (0.6932)	0.02616 (0.9567)	0.02957 (0.6742)	0.01953 (0.7339)
CC1/2	0.996 (0.404)	0.999 (0.409)	1 (0.43)	0.999 (0.445)
CC*	0.999 (0.758)	1 (0.762)	1 (0.776)	1 (0.785)
Reflections used in refinement	21649 (2143)	11345 (1099)	20817 (2043)	6489 (629)
Reflections used for R-free	1089 (101)	525 (49)	1080 (88)	294 (24)
R-work	0.2553 (0.3880)	0.2293 (0.3926)	0.2298 (0.3350)	0.2293 (0.3506)
R-free	0.3030 (0.4588)	0.2866 (0.3927)	0.2374 (0.4074)	0.2449 (0.4083)
CC(work)	0.955 (0.573)	0.955 (0.612)	0.929 (0.647)	0.966 (0.613)
CC(free)	0.927 (0.373)	0.915 (0.561)	0.954 (0.333)	0.816 (0.674)
Number of non-hydrogen atoms	4145	1178	1255	2267
macromolecules	4145	1174	1180	2257
ligands		4		10
solvent			75	
Protein residues	533	152	154	295
RMS(bonds)	0.014	0.011	0.010	0.006
RMS(angles)	2.06	1.10	1.08	0.87
Ramachandran favored (%)	93.45	96.62	98.00	93.73
Ramachandran allowed (%)	5.59	3.38	2.00	5.92
Ramachandran outliers (%)	0.96	0.00	0.00	0.35
Rotamer outliers (%)	0.00	0.00	0.77	1.20
Clashscore	22.01	7.47	9.93	12.69
Average B-factor	78.42	72.20	31.64	134.49
macromolecules	78.42	72.12	31.61	134.41
ligands		93.96		151.75
solvent			32.16	

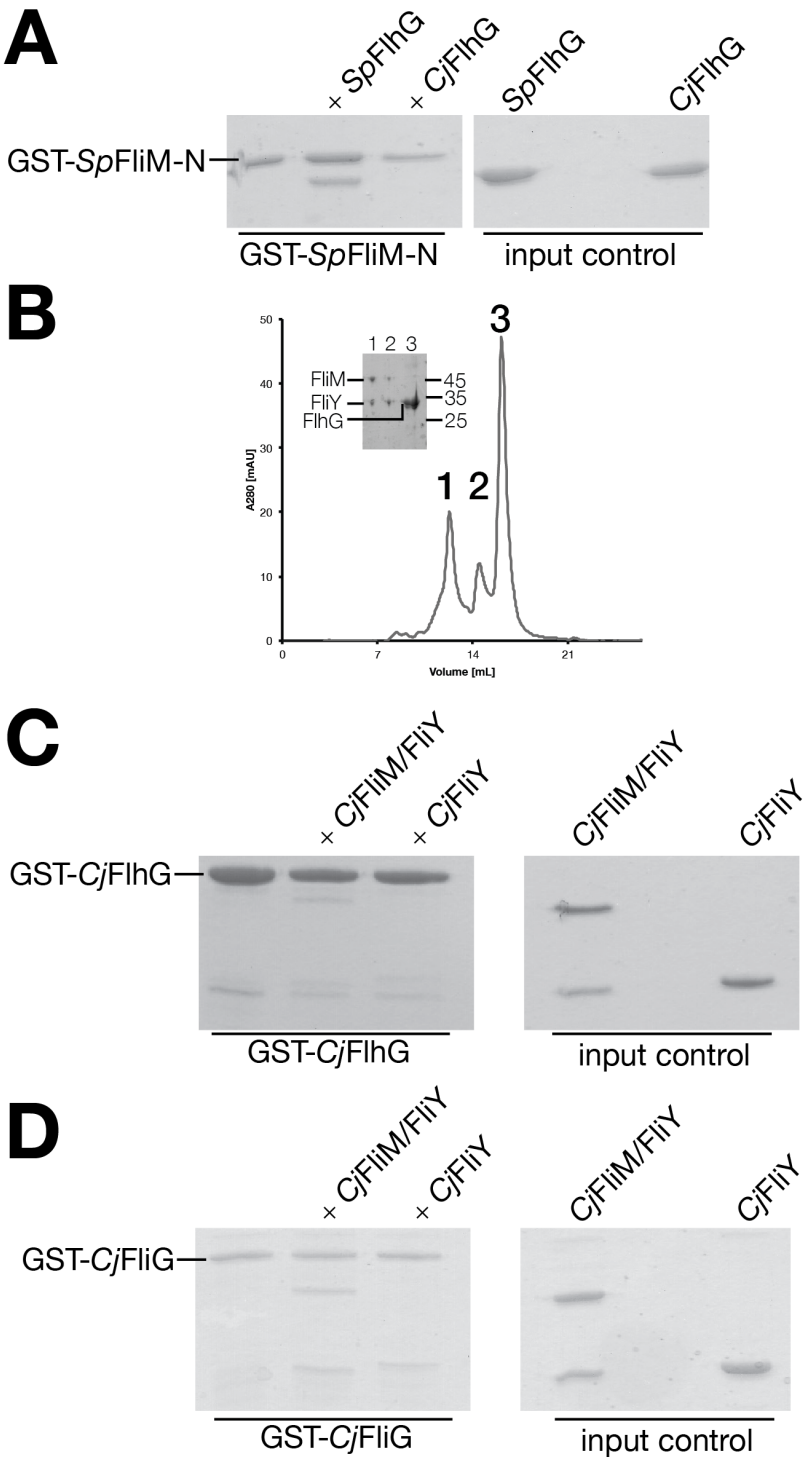

Figure A1. A: GST-interaction assay employing GST-*Sp*FliM-N (EIDAL) versus FlhG from *S. putrefaciens* and *C. jejuni*. Only *Sp*FlhG binds to the EIDAL motif. **B**: Analytical size-exclusion chromatography of *Cj*FliM/FliY and *Cj*FlhG. The proteins were mixed stoichiometrically and incubated for 30 min at room temperature. Analytical SEC shows that they do not form a ternary complex. **C**: GST-interaction

assay employing GST-*Cj*FlhG versus *Cj*FliY and the SEC-purified *Cj*FliM/FliY-complex. No interaction is detected. **D**: GST-interaction assay employing GST-*Cj*FliG versus *Cj*FliY and the SEC-purified *Cj*FliM/FliY-complex. Stoichiometrix interaction is deteced in the case of *Cj*FliM/FliY. *Cj*FliY also binds weakly to GST-*Cj*FliG.

7 Scientific curriculum

8 Acknowledgements

First and foremost, I am grateful to Gert Bange for his invitation to Marburg, for his trust and his continuous support during my doctoral studies.

I also want to thank Peter Graumann and Erhard Bremer for their support and their willingness to act in my thesis advisory committee. I thank Olalla Vázquez for co-examining my dissertation.

I thank Prof. Thormann, Prof. Frey, Prof. Thanbichler, Prof. Witte and Prof. Bibi for their support. I thank Dr. Sven Freibert, Dr. Thomas Heimerl, Dr. Tamara Hoffmann and Dr. Wieland Steinchen and Darius Günder for their helpful experimental work.

I am grateful for the support and the friendly atmosphere in the AG Bange, especially to Florian Altegoer.

And I am especially grateful to Andreza Maraschin Woolf de Oliveira, who made it possible, with love.