From the Department of Microbiology, Tumor and Cell Biology

Karolinska Institutet, Stockholm, Sweden

DIVERSE CYCLIC DINUCLEOTIDE SIGNALS REGULATE *ESCHERICHIA COLI* LIFESTYLE TRANSITION

Fengyang Li

Stockholm 2020

Published by Karolinska Institutet.
Printed by Arkitektkopia AB

ISBN 978-91-7831-822-3

DIVERSE CYCLIC DINUCLEOTIDE SIGNALS REGULATE *ESCHERICHIA COLI* LIFESTYLE TRANSITION
THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Fengyang Li

Principal Supervisor:
Prof. Ute Römling
Karolinska Institutet
Department of Microbiology, Tumor, and Cell Biology

Co-supervisor(s):
Prof. Mark Gomelsky
University of Wyoming
Department of Molecular Biology

Dr. Annika Cimdins-Ahne
University of Münster
Institute of Hygiene

Opponent:
Prof. Karin Sauer
Binghamton University
Department of Biological Sciences

Examination Board:
Prof. Karl-Eric Magnusson
Linköping University
Department of Biomedical and Clinical Sciences

Prof. Jan-Willem de Gier
Stockholm University
Department of Biochemistry and Biophysics

Prof. Roland Möllby
Karolinska Institutet
Department of Microbiology, Tumor, and Cell Biology

咬定青山不放松，

立根原在破岩中。

千磨万击还坚劲，

任尔东西南北风。

-清 郑燮《竹石》

ABSTRACT

Bacteria have the ability to change their lifestyle to adapt to various environmental conditions. Cyclic dinucleotides (cDNs) are ubiquitous second messengers that can regulate fundamental lifestyle switches, such as motility versus sessility and acute versus chronic virulence, in bacteria. Investigation of the diverse established cDNs, cyclic di-GMP, cyclic di-AMP, and the recently identified hybrid molecule cyclic GAMP, expanded our knowledge of the complexity of regulation of bacterial physiology by nucleotide-based second messengers and unravel common and distinct regulatory patterns.

In this thesis, we provided the molecular basis to assess the regulatory mechanisms of semi-constitutive rdar biofilm formation by Illumina Miseq or PacBio sequencing of the genomes of eight rdar biofilm forming *E. coli* strains **(Paper I)**. By using phenotypic, genetic and biochemical approaches, we showed that animal commensal isolate *E. coli* ECOR31 expresses a semi-constitutive rdar biofilm morphotype on agar plates characterized by expression of the extracellular matrix components cellulose and curli fimbriae. This morphotype is conventionally dependent on the major biofilm regulator CsgD and positively regulated by cyclic di-GMP signaling **(Paper II)**. As expected, flagella-dependent motility is negatively regulated by cyclic di-GMP signaling. Bioinformatic analysis suggested the presence of a dinucleotide cyclase DncV homolog, hypothesized to possess cyclic GAMP synthase activity, encoded by the *E. coli* ECOR31 genome. DncV synthesized 3'3'-cGAMP *in vitro* and *in vivo* and, via its catalytic activity, negatively regulated *csgD* expression at the transcriptional level with subsequent suppression of rdar biofilm formation and cell aggregation. DncV also suppressed swimming and swarming motility post-transcriptional of class 1 flagellar regulon genes. In liquid culture, expression of *dncV* restricted cell aggregation, but showed a complex temporal pattern of biofilm formation at the abiotic surface.

The patatin-like phospholipase CapV is a receptor for 3'3'-cGAMP. In this thesis, we showed that expression of $CapV_{Q329R}$, a single amino acid variant of CapV, induced extensive cell filamentation in *E. coli* MG1655 independently of the 3'3'-cGAMP synthase DncV **(Paper III)**. Moreover, overexpression of $CapV_{Q329R}$ repressed swimming motility by inhibiting flagella biosynthesis and reduced rdar biofilm formation and CsgD expression, possibly through modulation of cyclic di-GMP levels. The observed phenotypes of $CapV_{Q329R}$ are not restricted to *E. coli* MG1655, but common to other *E. coli* strains and *S.* typhimurium UMR1 suggesting that conserved pathway(s) are required for their expression.

Based on our genome sequences, in the last study, we investigated the molecular basis of temperature-independent expression of the rdar biofilm morphotype and subsequently *csgD* expression in seven semi-constitutive rdar biofilm forming *E. coli* strains **(Paper IV)**. Based on the observation that amino acid variations in cyclic di-GMP turnover proteins correlated with the expression of a semi-constitutive rdar biofilm, in particular, we demonstrated that distinct amino acid changes outside of conserved signature motifs that potentially lead to an alteration in the trigger activity of the hybrid cyclic di-GMP phosphodiesterase/diguanylate

cyclase YciR contributed to regulate rdar biofilm formation and *csgD* expression in semi-constitutive rdar biofilm forming *E. coli* strains.

In conclusion, this thesis highlights that diverse cDN second messenger signals differentially regulate the bacterial sessile/motile lifestyle transition. Furthermore, the effect of $CapV_{Q329R}$ on bacterial phenotypes and physiology is an example of rapid evolution of protein functionality.

LIST OF SCIENTIFIC PAPERS

I. Cimdins, Annika, Petra Lüthje, **Fengyang Li**, Irfan Ahmad, Annelie Brauner, and Ute Römling. "Draft Genome Sequences of Semiconstitutive Red, Dry, and Rough Biofilm-forming Commensal and Uropathogenic *Escherichia coli* Isolates." Genome Announcement. 5, no. 4 (2017): e01249-16.

II. **Li, Fengyang**, Annika Cimdins, Manfred Rohde, Lothar Jänsch, Volkhard Kaever, Manfred Nimtz, and Ute Römling. "DncV Synthesizes Cyclic GMP-AMP and Regulates Biofilm Formation and Motility in *Escherichia coli* ECOR31." mBio 10, no. 2 (2019): e02492-18.

III. **Li, Fengyang**, Sulman Shafeeq, Heike Bähre, Manfred Rohde, and Ute Römling. "A Single Amino Acid Substitution in CapV Leads to Pronounced Cell Filamentation in *Escherichia coli*." manuscript.

IV. Cimdins, Annika, Roger Simm, **Fengyang Li**, Petra Lüthje, Kaisa Thorell, Åsa Sjöling, Annelie Brauner, and Ute Römling. "Alterations of C-di-GMP Turnover Proteins Modulate Semi-constitutive Rdar Biofilm Formation in Commensal and Uropathogenic *Escherichia coli*." MicrobiologyOpen 6, no. 5 (2017): e00508.

Publications by the author, which are not included in this thesis:

I. Sun, Lei, Peter Vella, Robert Schnell, Anna Polyakova, Gleb Bourenkov, **Fengyang Li**, Annika Cimdins et al. "Structural and Functional Characterization of the BcsG Subunit of the Cellulose Synthase in *Salmonella* typhimurium." Journal of Molecular Biology 430, no. 18 (2018): 3170-3189.

CONTENTS

LIST OF ABBREVIATIONS

ATP	Adenosine 5’-triphosphate
Bcs	Bacterial cellulose synthesis
Bdar	Brown, dry, and rough
C-di-AMP	Cyclic diadenosine monophosphate
C-di-GMP	Cyclic diguanosine monophosphate
CFU	Colony forming unit
cGAMP	Cyclic guanosine monophosphate-adenosine monophosphate
Csg	Curli subunit gene
CTP	Cytidine 5’-triphosphate
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
GTP	Guanosine 5’-triphosphate
HPLC	High performance liquid chromatography
LB	Luria-Bertani broth
MS/MS	Tandem mass spectrometry
OD	Optical density
Pdar	Pink, dry, and rough
PVDF	Polyvinylidene fluoride
PCR	Polymerase chain reaction
Rdar	Red, dry, and rough
RNA	Ribonucleic acid
Saw	Smooth and white
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEM	Scanning electron microscopy
S. typhimurium	*Salmonella enterica* serovar typhimurium
TEM	Transmission electron microscopy
TLC	Thin layer chromatography
UTP	Uridine 5’-triphosphate
EPS	Extracellular polymeric substance
DGC	Diguanylate cyclase

DAC	Diadenylate cyclase
PDE	Phosphodiesterase
GAC	Cyclic GMP AMP cyclase
DncV	di-nucleotide cyclase from *Vibrio cholerae*

1 INTRODUCTION

The capability of bacteria to sense and respond to environmental signals is vital for survival. Bacteria use nucleotide-based molecules as second messengers to regulate various modes of physiological processes in response to a first signal, which can be an environmental signal such as divers nutrients, various stress signals, temperature, osmolarity, pH, light, and quorum sensing (1, 2). Cyclic dinucleotides (cDNs) act as intracellular second messengers that amplify the signal, thereby distinctively modulating bacterial physiology to regulate, for example, the fundamental lifestyle transition between motility and sessility (biofilm formation) on the transcriptional to the post-translational level (3). Four major cDNs have been discovered up to now: cyclic dimeric guanosine monophosphate (c-di-GMP), cyclic dimeric adenosine monophosphate (c-di-AMP), and the hybrid cyclic guanosine monophosphate-adenosine monophosphate (cGAMP), which exists as a canonical and non-canonical structural analog. 3'3'-cGAMP occurs in bacteria, while 2'3'-cGAMP has been identified as a major innate immunoregulator in eukaryotes (1, 2, 4). Cyclic di-GMP has been recognized as a key regulator of the bacterial sessility/motility lifestyle switch in all Gram-negative and Gram-positive bacteria where it has been investigated (2, 5-7). On the other hand, c-di-AMP is involved in physiological functions such as the monitoring of DNA integrity, osmoprotection, cell-wall synthesis, potassium homeostasis and virulence mainly in Gram-positive bacteria and Mycobacteria (8). Cyclic GAMP was first identified in *Vibrio cholerae* O1 biovar El Tor 7th pandemic strains as a signaling molecule involved in chemotaxis and virulence (4). In eukaryotes, all of these cDNs, besides the intrinsic signaling molecule 2'3'-cGAMP, are recognized as important pathogen- or microbial-associated molecular patterns (PAMPs or MAMPs) that trigger a distinct innate immune response (9).

1.1 CYCLIC DI-GMP

The bacterial second messenger cyclic di-GMP (c-di-GMP) was firstly discovered by Moshe Benziman and his colleagues in 1987 as an allosteric activator of the bacterial cellulose synthase in the Gram-negative fruit-degrading bacterium *Komagataeibacter* (*Gluconacetobacter) xylinus*, previously known as *Acetobacter xylinum* (10). More than fifteen years after this seminal discovery, c-di-GMP has been established as a ubiquitous bacterial second messenger involved in the regulation of complex lifestyle changes, such as the transition between sessility and motility, cell cycle regulation, the transition between acute and chronic virulence and various additional associated phenotypes in bacteria (2, 5, 6) (Figure 1). Cyclic di-GMP has been primarily investigated in Gram-negative bacteria, including *Escherichia coli*, *Salmonella typhimurium, Pseudomonas aeruginosa*, *Caulobacter crescentus*, *Klebsiella pneumoniae, Legionella pneumophila, Yersinia pestis* and *Vibrio cholerae*, and a few Gram-positive bacteria, such as *Bacillus subtilis, Streptomyces* spp., *Staphylococcus aureus*, *Clostridium difficile* and *Listeria monocytogenes*, but not in archaea (5, 11).

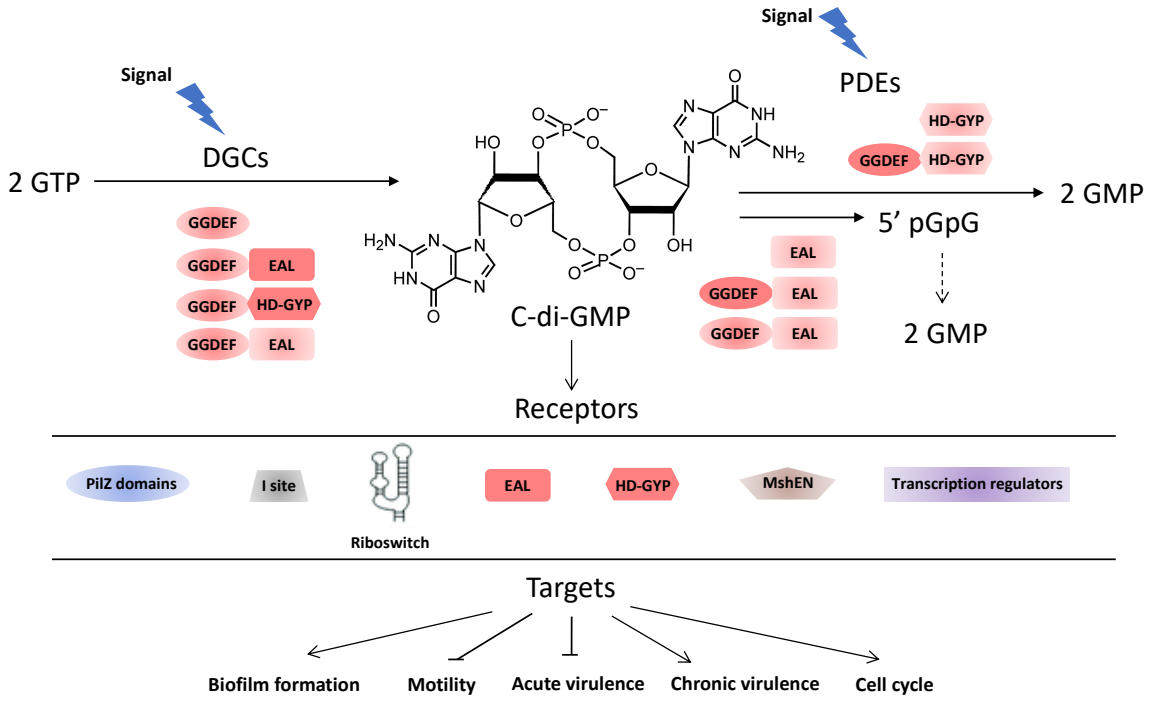

Figure 1 Overview of the c-di-GMP signaling network in bacteria. Cyclic di-GMP is synthesized by DGCs with a GGDEF domain and hydrolyzed by PDEs with an EAL or HD-GYP domain. Cyclic di-GMP is sensed by protein or RNA receptors, including PilZ domain receptors, I-site receptors, inactive EAL or HD-GYP domain receptors, MshEN domain receptors, divers transcriptional regulators e.g. FleQ and VpsT, and riboswitches e.g. c-di-GMP binding GEMM-I and GEMM-II riboswitches. The diagram shows the protein domains involved in c-di-GMP turnover and signaling. Enzymatically active and inactive GGDEF, EAL, and HD-GYP domains are indicated by light pink and pink, respectively.

1.1.1 C-di-GMP mediated physiological functions

In nature, most microorganisms exist primarily in larger communities rather than single planktonic cells whereby they can attach to a biotic or abiotic surface and subsequently build up single and multispecies cell aggregates to accumulate at interfaces (12). The ubiquituousness and predominance of biofilm formation in almost any environment with an adequate combination of moisture, carbon and energy provision covers plant and body tissues, natural materials, metals, plastics, medical implant materials and natural environments with even extreme conditions such as host springs and alkaline hydrothermal vents (13, 14). In clinical settings, biofilms are a major cause of persistent human infections. Such arrangement into multicellular communities (biofilm formation) is possible through the production of extracellular polymeric substances (EPS), which is similar to tissue formation in higher organisms (15) (Figure 2). A microbial biofilm is mainly comprised of EPS (approx. 85% dry biomass), which provides a physical and chemical protection from adverse growth conditions and environmental threats, such as desiccation, phages (bacteria-attacking viruses), bactericides and antibiotics for the residing bacteria.

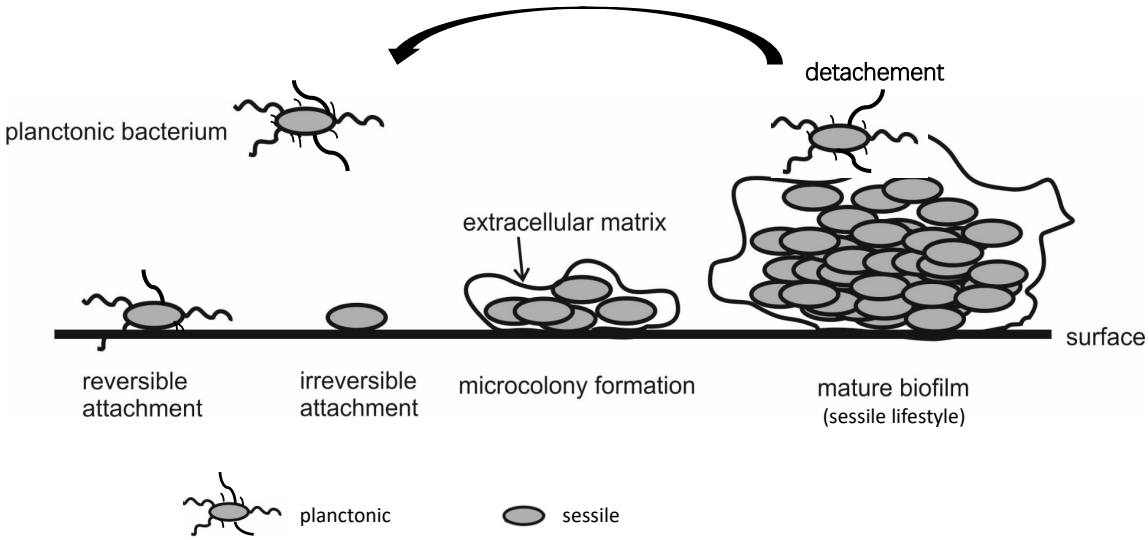

Figure 2 Schematic illustration of subsequent developmental stages leading to biofilm formation. Biofilm formation is initiated with the reversible attachment of a single planktonic cell followed by surface scanning using surface appendages such as type IV pili (*P. aeruginosa*) or type I pili (*E. coli*) and subsequent adhesion to the surface by extracellular matrix components such as cellulose (*S.* typhimurium) (16) (stage 1). The bacteria then form a monolayer and irreversibly attach by producing an extracellular matrix (stage 2). Next, a microcolony is formed where multilayers appear (stage 3). During later stages, a mature biofilm is developed, which can display as a characteristic mushroom or slime-like structure due to EPS production (stage 4). Eventually, some cells start to detach and the biofilm will disperse (stage 5). The figure is adapted from a figure originated from Annika Cimdins-Ahne.

Cyclic di-GMP has been shown to regulate multiple modes of biofilms formation. Model system biofilms demonstrated to be c-di-GMP regulated include pellicle formation at the air liquid interface, rdar (red, dry and rough) or rugose morphotype formation on agar plates, adhesion to abiotic surfaces, cell aggregates and biofilm formation in continuous flow systems (17). Consecutively, the distinct expression of a multitude of combinatorial extracellular matrix components including exopolysaccharides, adhesive pili, non-fimbrial adhesins as well as extracellular DNA is affected by c-di-GMP signaling (2, 18). In *Enterobacteriaceae*, c-di-GMP regulates a distinct multicellular behavior termed the rdar morphotype. Thereby, c-di-GMP promotes the expression of the transcriptional regulator CsgD, which subsequently activates production the extracellular biofilm components the exopolysaccharide cellulose and amyloid curli fimbriae in *S.* typhimurium and *E. coli* (2, 18). The rdar morphotype (Figure 3), a well-established validated biofilm model, is visualized on agar plates containing Congo Red and Coomassie Brilliant Blue, which differentially binds to the extracellular matrix components cellulose and curli fimbriae (18).

Figure 3 The rdar morphotype of *E. coli* ECOR31 after 3 days of growth on a Congo Red agar plate.

Transition from motility to sessility is an important universal switch regulated by c-di-GMP. Motility, which is defined as the ability to actively move, is a major survival mechanism of most microorganisms (19). For example, *P. aeruginosa*, *E. coli,* and other bacteria can move on surfaces by the extension and retraction of type IV pili in a process called twitching motility (20). A very common mode of motility which many bacteria exhibit is motility mediated by propelling of flagella; in peritrichously flagellated *E. coli* up to 10 µm long and 20 nm wide appendages. Two forms of flagella mediated motility have been described: swimming motility in liquid or semi solid medium and the swarming motility on the surface of semi-solid medium (19, 21). In contrast to biofilm formation, flagellum-based swimming and swarming motility in *E. coli* and *S.* typhimurium is inhibited by c-di-GMP on the post-transcriptional level (21). Cyclic di-GMP binds to the PilZ domain protein YcgR, which interacts with FliG and FliM, the subunits of the flagella switch complex to induce a counterclockwise (CCW) rotational bias and slow down the flagellar motor rotation. As a consequence, smooth swimming for a longer time with fewer frequencies of tumbling intervals is achieved (22-24). According to another proposed model, c-di-GMP bound YcgR directly interacts with MotA, the stator component of the flagella motor thereby reducing the rotation speed of the motor (25). Moreover, c-di-GMP can also bind FliI, the flagella export AAA+ ATPase, to repress swimming motility (26). In other bacteria, in addition to posttranslational regulation of motility, c-di-GMP regulates motility at the transcriptional level. For example, in *V. cholerae*, the transcription factor VpsT requires c-di-GMP to be active and to suppress flagellar gene expression (27).

Cyclic di-GMP also contributes to cell cycle progression and alterations in cell morphology in the aquatic bacterium *Caulobacter crescentus* (28, 29). *C. crescentus* undergoes asymmetric cell division, which generates a motile and replication-inert swarmer cell and a sessile and replication-competent stalked cell. The transition from a motile to a sessile state of *C. crescentus* is coordinated by c-di-GMP through its metabolizing proteins. Specifically, the diguanylate cyclase (DGC) PleD, which is inactive in swarmer cells, is activated upon phosphorylation and localizes to the stalked cell pole to control flagella ejection and subsequently biogenesis of stalk and holdfast (30, 31). Localization of phosphorylated PleD helps to recruit PopA, a c-di-GMP effector protein that binds c-di-GMP through its I-site, sequesters to the stalked cell pole and helps to recruit the replication initiation inhibitor CtrA via the mediator protein RcdA to be delivered and degraded by the polar protease ClpXP (32-

34). DgcB is another DGC involved in *C. crescentus* holdfast biogenesis and surface attachment. DgcB inhibits motility by counteracting the c-di-GMP phosphodiesterase (PDE) PdeA to lower c-di-GMP level in the swarmer cell (35). Also, in other bacteria, c-di-GMP contributes to alterations in cell morphology and regulation of the cell cycle (36, 37). In *Mycobacterium smegmatis*, c-di-GMP signaling is not only required for the maintenance of proper growth under stress conditions, but also affects the cell size and cell division (37). YfiN, a bifunctional DGC in *E. coli*, blocks cell division through preventing the initiation of septal peptidoglycan synthesis in response to envelope stress (36).

Besides, c-di-GMP also regulates other physiological processes in bacteria, such as survival and transmission of the obligate intracellular pathogen *Borrelia burgdorferi* in insect and mammalian hosts, photosynthesis and heterocyst formation in *Cyanobacteria*, multicellular development, sporulation and antibiotic production in *Streptomycetes*, long-term survival and lipid metabolism and transport in *Mycobacteria*, the type III secretion system in *S*. typhimurium, *P. aeruginosa* and other bacteria and antioxidant regulation in *M. smegmatis* (38-42).

1.1.2 C-di-GMP metabolism

Diguanylate cyclases (DGCs), the enzymes that perform c-di-GMP synthesis, are structurally similar to type III nucleotidyl transferases and adenylate cyclases, but contain a unique conserved GGDEF signature motif (2, 43). GGDEF domain proteins function as homodimers, whereby each monomer contributes one bound GTP molecule to form the phosphodiester bonds. In the majority of cases, one of the two divalent metal ions Mg^{2+} and Mn^{2+} is required for phosphodiester bond formation between two GTP molecules to form 5'-pppGpG, an intermediate product that is converted into c-di-GMP with the release of two inorganically phosphate molecules (1, 2). The first two glycine residues of the central GGDEF motif are involved in GTP binding whereas the fourth amino acid glutamate is involved in metal ion coordination. The third residue which can be aspartate or glutamate is required for catalysis and metal ion coordination (44). The RxxD motif is located five amino acids upstream of the GGDEF motif. Upon c-di-GMP binding to the RxxD motif, DGC activity of the GGFEF domain is inhibited. The RxxD motif, an essential part of this c-di-GMP binding site, is conserved in approximately half of the GGDEF domain proteins and also named "I site" (inhibitory site) due to its allosteric inhibitory function, is required to potentially restrict the signal (2, 44, 45). However, the function of the I-site extends beyond the allosteric regulation of enzymatic activity (46, 47).

The hydrolysis of c-di-GMP is achieved by two major groups of phosphodiesterases (PDEs): HD-GYP and EAL domain PDEs. Cyclic di-GMP PDEs with HD-GYP domain hydrolyze c-di-GMP to two GMP molecules, while PDEs with EAL domain cleave c-di-GMP into a linear 5'-pGpG that is further hydrolyzed into two GMP molecules by HD-GYP domain PDEs or oligoribonucleases (2, 48-51). Metal ions, Mg^{2+}, Fe^{3+}, or Mn^{2+} are coordinated for catalysis by both HD-GYP and EAL domain PDEs, with Ca^{2+} to strongly inhibit the PDE enzymatic activity of the EAL domain (2, 52, 53). Conserved signature motifs of EAL domains extend beyond the conserved EAL motif (51, 54). The first glutamate residue of the EGVE motif of the EAL

domain serves as a general base catalyst, which accepts the proton from a water molecule. The produced hydroxide ion subsequently conducts the nucleophilic attack, resulting in the weakening and subsequent hydrolysis of the phosphoester bond (55). In addition, the conserved DFG(T/A)GYSS motif of the loop 6 of the EAL domain which connects elements of secondary structure, a β-sheet and an α-helix is essential for catalytic activity (54, 56). Structural studies showed that c-di-GMP in the EAL domains is present in an extended (open) conformation which facilitates hydrolysis of one of the phosphodiester bonds in c-di-GMP. In contrast, c-di-GMP in HD-GYP domains is revealed as a bent, U-shaped (closed) conformation as observed in the I sites of DGCs and some PilZ domain c-di-GMP receptors (2, 57, 58). Besides HD-GYP domain PDEs and EAL domain PDEs, some proteins with DHH–DHHA1 motifs (signature motif for c-di-AMP hydrolysis) can also hydrolyze c-di-GMP, e.g. Rv2837c from *M. tuberculosis* specifically targets the 3'5' phosphodiester bond and degrades c-di-GMP to GMP (59, 60).

GGDEF/EAL/HD-GYP domains belong to abundant bacterial protein superfamilies. The number of GGDEF/EAL/HD-GYP domain proteins can, though, vary dramatically within a genus even between closely related bacterial species (11, 61). Genomic analyses showed that a GGDEF domain is often coupled to an EAL domain or a HD-GYP domain. For example, the *S.* typhimurium genome encodes twenty-one GGDEF/EAL domain proteins, of which five proteins contain a GGDEF, nine an EAL domain and seven contain both a GGDEF and an EAL domain (18, 62). The genome of the laboratory model organism *E. coli* K-12 strains contains twenty-nine GGDEF/EAL domain proteins, including twelve proteins with a GGDEF domain, ten proteins with an EAL domain and seven proteins with both a GGDEF and an EAL domain (63). Although less investigated than GGDEF-EAL domain proteins, GGDEF-HD-GYP domain proteins are also widespread in bacteria, particularly among the *Aquificae*, *Deinococci*, *Firmicutes* and *Planctomycetes* (2). Typically, in such protein tandems one of the two domains is enzymatically inactive, but bifunctional DGCs/PDEs exist. For example, the BphG1 protein from *Rhodobacter sphaeroides*, a bacteriophytochrome with a PAS-GAF-PHY photosensory module followed by a GGDEF-EAL output domain (64), the PAS-GGDEF-EAL domain protein PP2258 from *Pseudomonas putida* (65) and the STM1703/YciR from *S.* typhimurium (66), possess both, DGC and PDE, activities. However, for example, YciR, which has homologs in species outside of the gamma-proteobacteria in the phylogenetic tree, regulates rdar morphotype and *csgD* expression independently of the catalytic activities through protein-protein interactions (66-68).

1.1.3 C-di-GMP receptors

Many different types of c-di-GMP binding proteins (receptors) with different binding sites have been identified in bacteria so far: PilZ domain receptors, DGC I-site receptors, catalytically inactive EAL and HD-GYP domain receptors, MshEN domain receptors, and other c-di-GMP receptors which include transcriptional regulators such as the major flagella gene regulator, the AAA+ ATPase enhancer-binding protein FleQ in *P. aeruginosa*, and VpsT, a CsgD-family transcriptional factor of *V. cholerae* involved in regulating biofilm formation (27, 58, 69). In

mammals, STING (stimulator of interferon genes) is a direct innate immune sensor of c-di-GMP (70).

The first identified c-di-GMP protein receptor was the PilZ domain protein YcgR (71, 72). The PilZ domain is widespread in bacteria from different branches of the phylogenetic tree. The PilZ domain can be coupled to a variety of other domains and affect their catalytic or protein binding activity (73). The C terminus of the cellulose synthase BcsA of *G. xylinus*, the flagellar-break YcgR from *E. coli*, the DgrA protein from *C. crescentus*, and the PlzC and PlzD from *V. cholerae* are identified PilZ domain c-di-GMP receptors (71, 74-76). Upon c-di-GMP binding the PilZ domain protein YcgR undergoes a profound conformational change to inhibit the flagella associated motility through protein-protein interactions. The conformational change upon c-di-GMP binding has been utilized to construct an *in vivo* c-di-GMP sensor based on the alteration in Förster transfer between the fluorophores (77). Equally binding of c-di-GMP to the PilZ domain protein BcsA activates cellulose biosynthesis (10, 23, 78). The conserved sequence motifs RxxxR and [D/N]xSxxG of PilZ domains, which are responsible for c-di-GMP binding, were disclosed by bioinformatics analyses and subsequently determination of crystal structures and biochemical experiments confirmed their role in c-di-GMP binding (61, 71).

Cyclic di-GMP can also bind to the I-site of GGDEF domains, which allosterically inhibits c-di-GMP synthesis by active DGCs. However, even degenerated GGDEF domains deficient in DGC activity can serve as c-di-GMP receptors. As such, PopA from *C. crescentus* contains a catalytically non-functional GGDEF domain, but serves still as a c-di-GMP receptor as it binds c-di-GMP through its functional I-site (79). In EAL domain proteins, such as SgmT from *M. xanthus*, PelD and FimX from *P. aeruginosa*, the substrate binding site of a defunct EAL domain functions as the c-di-GMP binding site (32, 80, 81).

Another widespread type of receptor is the MshEN domain. In *V. cholerae*, two MshE domain ATPases involved in the bacterial type II secretion system (T2SS) and type IV pilus formation were found to specifically bind c-di-GMP (82-84). As one of the first receptors, the cAMP Receptor-Like Protein CLP was identified as a novel c-di-GMP receptor coordinating cell–cell signaling to virulence gene expression in *Xanthomonas campestris* (85). Moreover, in *E. coli* c-di-GMP also directly binds to polynucleotide phosphorylase (PNPase) which processes mRNA transcripts to regulate signal dependent RNA processing (86).

Besides binding to protein receptors, c-di-GMP binds to riboswitches such as GEMM domain riboswitches to transcriptionally and post-transcriptionally regulate target gene expression and translation (87, 88). Riboswitches are sequences in the 5' untranslated region of bacterial mRNAs that consist of an aptamer capable of binding small molecules whereby the structural change upon binding allows to regulate a downstream expression platform for gene transcription or protein translation (87, 89). So far, two main classes, the c-di-GMP I and II riboswitches, which can function as 'on' and 'off' riboswitches, with elevated c-di-GMP triggering and inhibition transcription or translation, respectively, have been identified.

1.2 CYCLIC DI-AMP

The bacterial second messenger cyclic di-AMP (c-di-AMP) was originally discovered upon elucidation of the crystal structure of the sporulation checkpoint protein DisA from *Thermotoga maritima* by Hopfner and colleagues in 2008 (90-92). Subsequently, a consecutive physiological role for c-di-AMP signaling has emerged from the analyses of this signaling pathway in several Gram-positive bacteria and mycobacteria including *S. aureus*, *M. tuberculosis*, *M. smegmatis*, *B. subtilis, B. thuringiensis*, *L. monocytogenes*, *S. pyogenes* and *S. pneumoniae* and *Lactococcus lactis*, but also in some Gram-negative bacteria such as *Chlamydia trachomatis* and *Borrelia burgdorferi*, as well as in archaea (93-98). Concluded from these studies, c-di-AMP is considered as a broadly conserved second messenger with a delicate homeostasis required for optimal microbial growth and physiology, with effects on, for example, DNA integrity, cell-wall synthesis, osmoprotection, potassium homeostasis and virulence (1, 99) (Figure 4). In mammals, c-di-AMP binds the cytosolic DNA sensor STING and induces the production of type I interferons (IFNs) in response to pathogen infections (97).

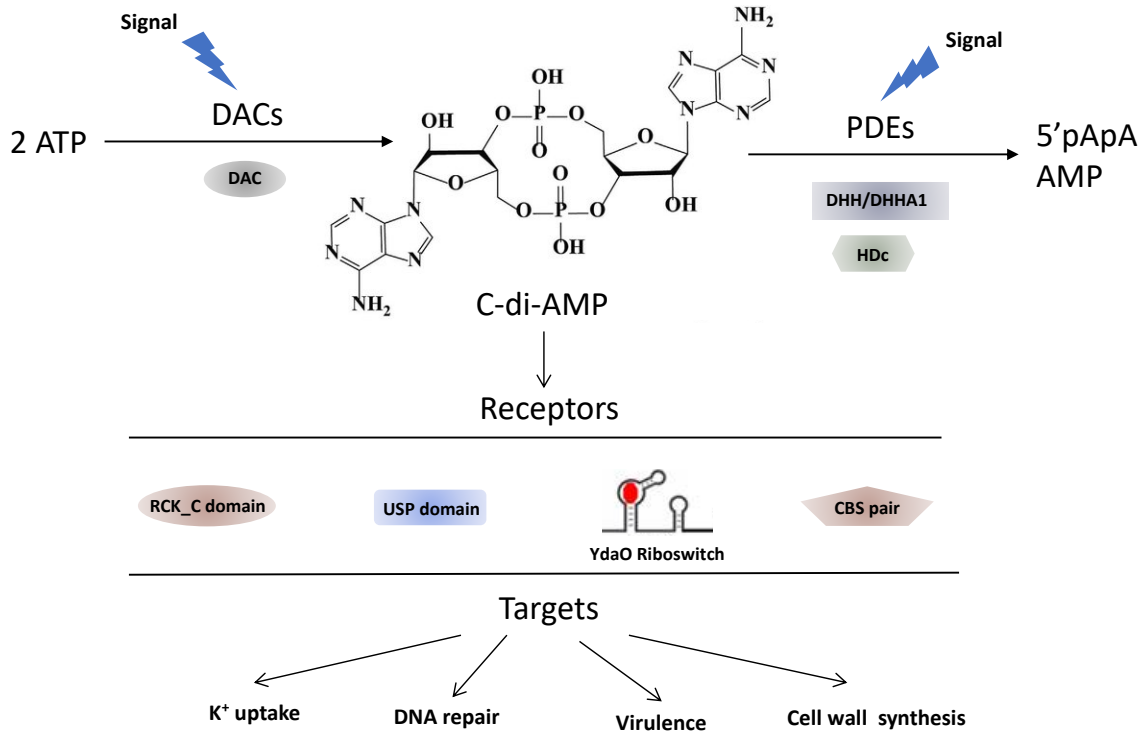

Figure 4 Overview of the c-di-AMP signaling network in bacteria. Cyclic di-AMP is produced by DAC domain diadenylate cyclases and degraded by PDEs with DHH-DHHA1 or HDc domains. Cyclic di-AMP is sensed by protein or RNA receptors, including RCK_C domains, USP domains, CBS pair domains, and riboswitches e.g. the YdaO riboswitch (100). The diagram shows the protein domains involved in c-di-AMP turnover and signaling.

1.2.1 C-di-AMP metabolism

As for each second messenger, the cellular concentration of c-di-AMP is tightly regulated and indeed, overexpression and depletion of c-di-AMP is deleterious (101). Cyclic di-AMP is synthesized from two ATP molecules by diadenylate cyclases (DACs) via its enzymatic

activity (91). Up to now, the DAC domain is the only domain that has been identified to perform c-di-AMP synthesis *in vivo* (102). The active pocket of DACs contains two metal ions which coordinate the phosphate group of the two ATP molecules. A condensation reaction of the two ATP molecules by active site residues produces c-di-AMP. The DAC domain is conserved among bacteria and archaea and is characterized by a catalytic site containing the conserved RHR and DGA motifs (90, 103, 104). DAC domain enzymes of different classes have been reported from multiple bacteria such as DisA, CdaA and CdaS from *B. subtilis*, CdaA from *S. pyogenes*, CdaA from *S. pneumoniae*, MtDisA from *M. tuberculosis*, CdaA from *C. trachomatis*, and CdaA from *L. monocytogenes* (90, 95-97, 105-107). Moreover, the DAC domain and c-di-AMP signaling seem to be as widespread as the GGDEF and EAL domains, however, usually only 1-3 DAC domain proteins are encoded per genome (91, 92).

Cyclic di-AMP is subsequently degraded into linear pApA or AMP by PDEs (106, 108). Up to now, two major classes of PDEs have been identified to degrade c-di-AMP. The first is the DHH-DHHA1 family PDEs, such as GdpP from *B. subtilis*, which is characterized by the DHH-DHHA1 catalytic domain coupled to two transmembrane helical domains, a signal-sensing PAS domain and a degenerate GGDEF domain (109). A related type of PDE contains the catalytic DHH-DHHA1 domain only, such as DhhP from *B. burgdorferi*, Pde1 and Pde2 from *S. pneumoniae*, and CnpB from *M. tuberculosis* (106, 110-112). The second type of PDE, which belongs to the 7TMR-HD family, possesses a catalytic histidine-aspartate (HD) domain, such as PgpH from *B. subtilis* (113) and *L. monocytogenes* (114).

1.2.2 C-di-AMP receptors

The identified bacterial c-di-AMP receptors encompass RNA riboswitches (115), and domains associated with transcriptional regulators (116, 117) and ion uptake (92). The first identified c-di-AMP receptor is DarR, a transcriptional regulator that belongs to the TetR family in *M. smegmatis* (117). Upon binding c-di-AMP, the DNA binding activity of DarR is enhanced and negatively regulates target gene expression and fatty acids synthesis (117). Another identified c-di-AMP receptor is KtrA, the potassium uptake transporter subunit which binds c-di-AMP within its RCK_C domain to contribute to *S. aureus* growth under low K^+ conditions (118). Moreover, the RCK_C domain protein CpaA and the histidine kinase protein KdpD from *S. aureus* and CabP from *S. pneumoniae* have also been identified as c-di-AMP receptors to regulate K^+ transport (118, 119). Other identified c-di-AMP receptors include DarA from *B. subtilis* and PstA from *S. aureus* (120, 121), both of which are PII-like signal transduction proteins, and LmPC from *L. monocytogenes* (116). Besides, c-di-AMP also binds to RNA riboswitches, e.g. the *ydaO* riboswitch, to regulate gene expression in response to intracellular ATP (115, 122).

1.3 CYCLIC GAMP AS A SECOND MESSENGER

In 2012, a hybrid cyclic GMP-AMP (3’3’-cGAMP) molecule was identified as a new type of bacterial cDNs in *V. cholerae* O1 biovar El Tor 7th pandemic strains by John J. Mekalanos and colleagues (4). Whereas the specific enzyme classes that produce, degrade, and respond to both

c-di-GMP and c-di-AMP second messenger signaling have been well described, all consecutive components of the 3'3'-cGAMP signaling pathway, synthase, phosphodiesterase, receptor and target output, have only been uncovered in *V. cholerae* (Figure 5).

1.3.1 3'3'-cGAMP synthesis

The novel bacterial second messenger 3'3'-cGAMP is synthesized by a novel class of dinucleotide cyclases termed DncV, which are sequentially and structurally distinct from classical c-di-GMP as well as c-di-AMP cyclases. In *V. cholerae* El Tor, DncV is required for efficient intestinal colonization and chemotaxis, a phenotype previously associated with hyper-infectivity (4). The presence of DncV homologs has been predicted for other bacterial species, suggesting a more global signaling role for 3'3'-cGAMP signaling (123-125). Genome analyses showed a homolog of DncV to be present, for example, in the *E. coli* commensal strain ECOR31 (ATCC 35350) from the *E. coli* reference (ECOR) collection (126-128). ECOR31 contains a high-pathogenicity island (HPI) encoding the siderophore yersiniabactin (Ybt), which was first identified in *Y. pestis* as a widely distributed virulence determinant among *E. coli* and other *Enterobacteriaceae* species that cause extra-intestinal infections (126, 129). ECOR31 HPI comprises an additional 35 kb fragment at the right border (RB) compared to the HPI of *E. coli* CFT073 and *Yersinia* species; the homolog of *dncV* is located on this fragment flanked up- and downstream by *V. cholerae* 7th pandemic island-1 (VSP-1) homologs of *capV*, *vc0180*, and *vc0181* (126, 130). Studies from our laboratory demonstrated DncV to synthesize 3'3'-cGAMP and to regulate biofilm formation and motility in ECOR31 (130). Others have shown that *dncV* and adjacent proteins are involved in the resistance against phage infections in *E. coli* (124). DncV homologs have also been detected in clinically relevant species such as *Enterobacter* and hypervirulent *Klebsiella pneumoniae* (130). In *V. cholerae*, CapV was identified as a cGAMP-activated phospholipase that leads to growth retardation upon overexpression (131). Moreover, based on structural and *in vitro* evidence, the catalytic activity of *V. cholerae* DncV is inhibited by folates (132).

Besides the di-nucleotide cyclase DncV, a sub-class of GGDEF domains, the hybrid promiscuous (Hypr) GGDEF enzymes, which are homologs to and structurally highly related to the GGDEF family of diguanylate cyclases (DGCs) associated with the c-di-GMP signaling pathway, were also identified to produce 3'3'-cGAMP. Those enzymes are, for example, involved in regulation of genes associated with extracellular electron uptake in the Gram-negative bacterium *Geobacter sulfurreducens* (133-135). The Hypr GGDEF homologs are also found in divers other bacterial species, including the predatory species *Bdellovibrio bacteriovorus* (now within the class *Oligoflexia*) and social myxobacterial species such as *Myxococcus xanthus* (133).

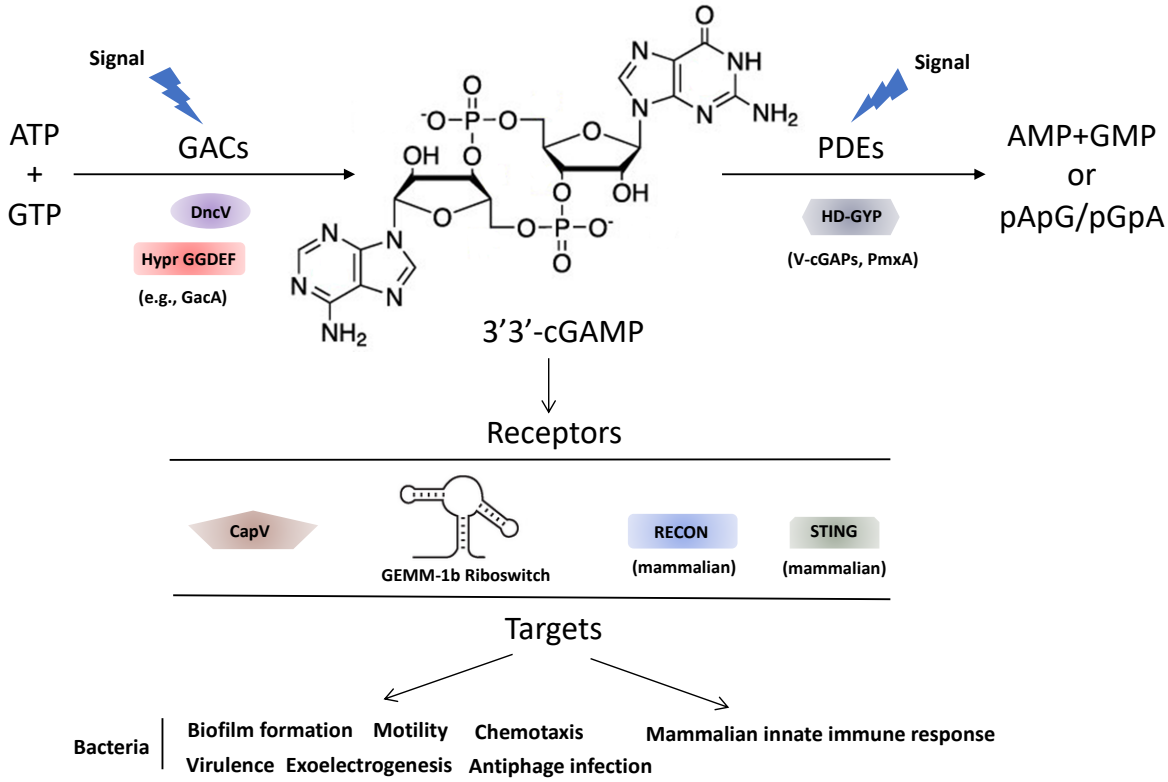

Figure 5 Overview of the 3'3'-cGAMP signaling network. 3'3'-cGAMP is produced by the GAC DncV and proteins with Hypr GGDEF domain and degraded by PDEs with HD-GYP domains. 3'3'-cGAMP is sensed by protein or RNA receptors, including CapV from *V. cholerae* (131), RECON (136) and STING (137) in mammals, and riboswitches e.g. GEMM-1b riboswitch (138). The diagram shows the protein domains involved in 3'3'-cGAMP metabolism and signaling as identified in *V. cholerae* and *Geobacter* spp.

Biochemical studies showed that the bacterial cGAMP contains two 3'-OH-5'-phosphate phosphodiester bonds (cyclic G(3',5')pA(3',5')p; 3'3'-cGAMP) characteristic for bacterial cyclic di-nucleotide signaling molecules (139). Besides producing 3'3'-cGAMP, two GTP or ATP molecules can also fit into the catalytic pocket of DncV, which generates the corresponding cyclic dimeric GMP (c-di-GMP) and c-di-AMP molecules *in vitro*, suggesting that one enzyme might also synthesize multiple cDNs in bacterial cells (4, 138). However, the predominant product of DncV is 3'3'-cGAMP with GTP and ATP predominantly used as substrates both *in vitro* and *in vivo* (4, 139, 140). These studies suggest DncV to preferably recognize ATP and GTP as acceptor and donor nucleotides. Structural studies determined that DncV comprises an elongated pocket for binding the nucleotide substrates in the middle of the enzyme, which is flanked by two distinct domains. The catalytic residues, Asp131, Asp133, and Asp193, characteristic for nucleotidyl transferases, are located on a β sheet at the bottom of the pocket, with two Mg^{2+} ions and two nucleotides bound to the active site (132, 141) (Figure 6). Despite the low sequence identity (<10%), the structure of bacterial DncV is highly similar to eukaryotic cGAS, the cyclic GMP-AMP synthase that produces 2'3'-cGAMP (142, 143).

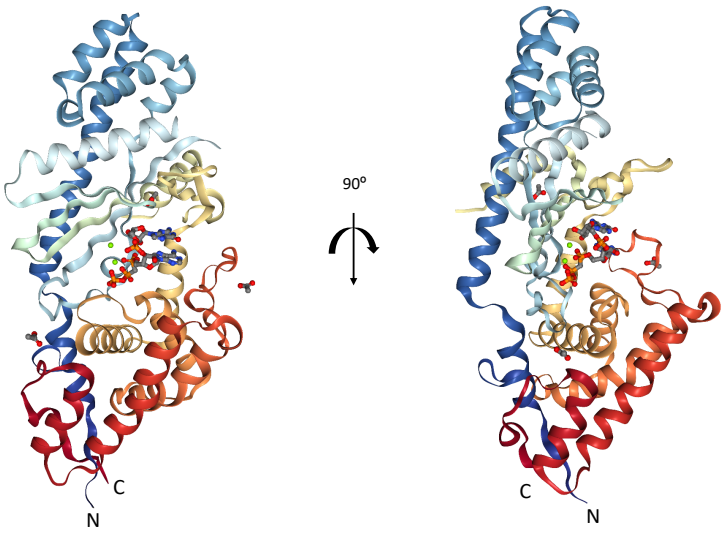

Figure 6 Crystal structure of DncV from *V. cholerae* containing a bound linear dinucleotide intermediate pp(c)pA[3'-5']pG shown as a ribbon representation (PDB 4TY0) (144). The bound Mg^{2+} ions and nucleotides are shown as green spheres and as ball-and-stick model, respectively.

In eukaryotes, the cGAS produced cGAMP subsequently stimulates the adaptor protein STING and thereby promotes a type I interferon innate immune response in mammalian cells. Biochemical studies have demonstrated that the mammalian cGAMP comprises a canonical 3'5' phosphodiester bond and a non-canonical 2'5' phosphodiester bond (cyclic G(2',5')pA(3',5')p; 2'3' cGAMP) (4, 139, 140). Subsequently structural studies revealed that cGAS sequentially catalyzes the reactions in one active site within the nucleotide acceptor and donor pockets (145, 146). The first reaction starts with the formation of the 2'5' phosphodiester linkage between GTP and ATP bound to the acceptor and donor pockets, respectively, to produce a linear pppG(2',5')pA intermediate molecule which subsequently rebinds to the catalytic pocket in the reverse direction. Secondly, the 3'5' phosphodiester linkage within the linear pppG(2',5')pA intermediate is formed leading to the production of 2'3'-cGAMP (Figure 7) (141, 144).

Distinct from eukaryotic cGAS, which produces 2'3'-cGAMP only upon binding of DNA, DncV is constitutively active to produce 3'3'-cGAMP *in vitro* (4, 139, 140, 143). However, the catalytic activity of DncV is inhibited by folic acid and its derivatives such as 5-methyltetrahydrofolate diglutamate, which was found to bind to a pocket opposite to the catalytic side upon determination of the crystal structure (132). DncV functions as a monomer and does not undergo major conformational changes upon substrate binding. Furthermore, DncV lacks regulatory dimerization contacts as observed in eukaryotic cGAS and conventional bacterial cyclic dinucleotide cyclases (90, 141, 144, 147-149). Similar to eukaryotic cGAS, DncV also produces cGAMP through sequential reactions in one active site. Crystallization of DncV with the linear dinucleotide intermediate pp(c)pA[3'-5']pG indicated that the first step in DncV nucleotide cyclization occurs through the attack of the 3' hydroxyl of the ribose of adenosine to the α-phosphate group of the guanosine, utilizing ATP as the nucleophile donor

and GTP as the electrophile acceptor; while in the second step, the 3’ hydroxyl of the ribose of guanosine attacks the α-phosphate group of the adenosine, to eventually form 3’3’-cGAMP (A[3’-5’]pG[3’-5’]p) (144) (Figure 7).

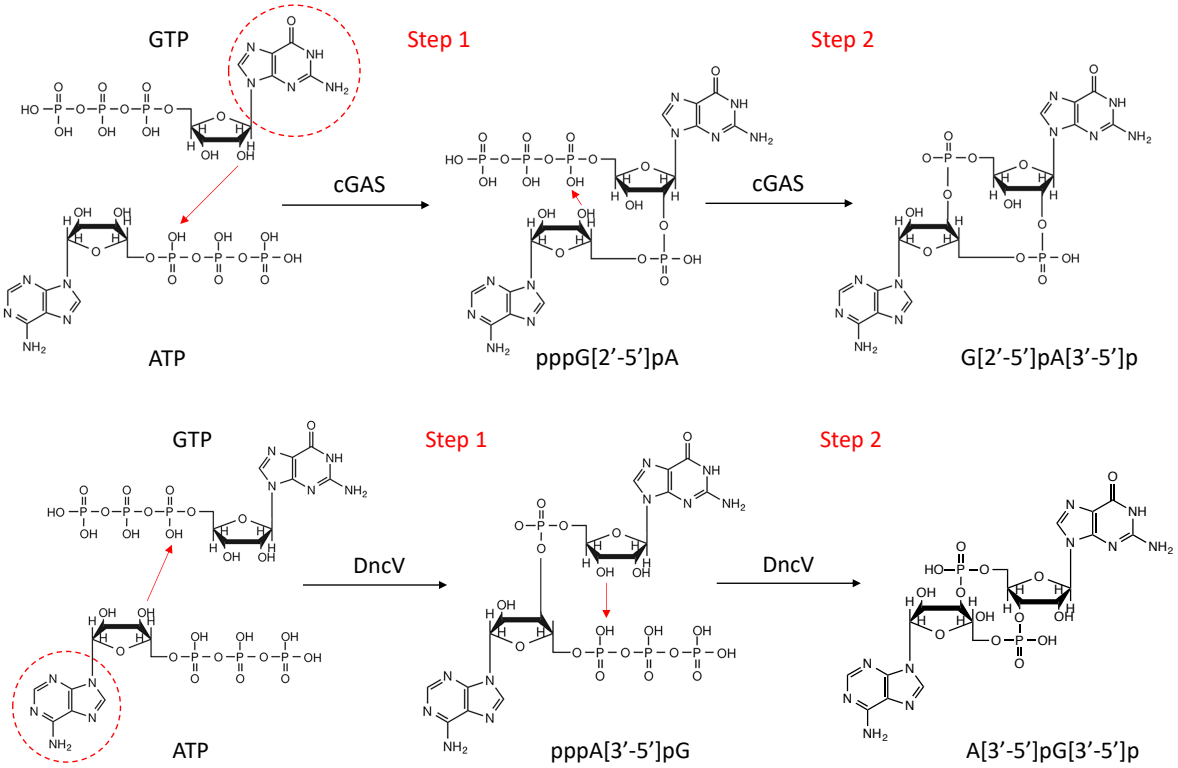

Figure 7 Schematic model of the two-step process of cGAS- (above) and DncV-catalyzed (below) cyclic dinucleotide product synthesis.

In contrast to cGAS and DncV, both of which produce 2’3’- and 3’3’-cGAMP, respectively,via a linear dinucleotide intermediate at one active site, the Hpyr GGDEF enzyme, e.g. GacA, a homodimeric enzyme from *G. sulfurreducens*, uses either of the nucleotides as a substrate in the first bond-forming step, which generates both types of linear intermediates (pp(c)pA[3’-5’]pG and pppG(2’,5’)pA), to produce 3’3’-cGAMP (135). Besides bacterial DncV and Hypr GGDEF domain proteins as 3’3’-cGAMP cyclases, 3’3’-cGAMP has been observed to be produced by the cGAMP synthase in the anemone *Nematostella vectensis* (150).

1.3.2 3’3’-cGAMP hydrolysis

Using an eukaryotic screening system, the first cGAMP-specific PDEs were identified as HD-GYP domain proteins, named V-cGAP1/2/3 in *V. cholerae* (151). All three PDEs were able to degrade 3’3’-cGAMP, but not other cGAMP analogues such as 2’3’, 2’2’, 3’2’-cGAMP. 3’3’-cGAMP is first hydrolyzed to linear 5’-pApG by all three V-cGAPs, which is subsequently further cleaved into 5’-ApG in a second step by V-cGAP1 via its 5’-nucleotidase activity (151). Structural studies showed that compared to other characterized HD-GYP domain protein structures which contain only one HD-GYP domain (53, 152, 153), V-cGAP3 contains a novel tandem HD-GYP domain within its N- and C-terminal domain, the structures of which are highly similar to each other despite low-sequence identity (154). The N-terminal domain of V-

cGAP3 is catalytically nonfunctional, but can enhance the catalytic activity of functional C-terminal domain. Upon binding with 3'3'-cGAMP, the C-terminal domain of V-cGAP3 undergoes a conformational change to form a bi-nuclear metal center for 3'3'-cGAMP and two Mg^{2+} (154).

The discovery of the *V. cholerae* V-cGAPs on 3'3'-cGAMP indicates the existence of specific PDEs of cGAMP in other bacteria, including the mammalian 2'3'-cGAMP. Recently, ectonucleotide pyrophosphatase/phosphodiesterase (ENPP1) was identified as the dominant 2'3'-cGAMP PDE using biochemical studies (155). In contrast to V-cGAPs that only degrades 3'3'-cGAMP, ENPP1 can hydrolyze not only 3'2'-cGAMP, but also 3'3'-cGAMP *in vitro*, but preferentially hydrolizes 2'3'-cGAMP (155, 156). Structural studies showed that 2'3'-cGAMP, but not 3'3'-cGAMP, binds to the active site of ENPP1 in a conformation suitable for the in-line attack by the catalytic residue Thr238, resulting in the specific hydrolysis of 2'3'-cGAMP by ENPP1 (157).

Besides V-cGAPs, the HD-GYP domain protein PmxA was recently identified as a 3'3'-cGAMP specific PDE that promotes resistance to osmotic stress in *M. xanthus* (158). In contrast to V-cGAPs, which can actually hydrolyze any guanine nucleobase containing molecules (151), PmxA, which contains a Qxx(K/R) motif instead of the canonical Rxx(K/R) motif in the HD-GYP domain, preferentially degrades 3'3'-cGAMP over c-di-GMP (158), suggesting the glutamine of the Qxx(K/R) motif may serve as a signature residue for the substrate specificity of HD-GYP domain PDEs.

1.3.3 3'3'-cGAMP receptors

The first protein receptor identified for 3'3'-cGAMP has been CapV from *V. cholerae* El Tor (131). Genomic analysis shows that *capV* is located immediately upstream of *dncV*, the gene which encodes the novel bacterial second messenger 3'3'-cGAMP, on the VSP-1 island of *V. cholerae* El Tor (4). CapV contains a canonical patatin-like phospholipase A (PNPLA) domain which belongs to the phospholipase A_2 (PLA_2) superfamily (159). PLA_2 and PLA_1 are two major subclasses of the PLA superfamily. PLA_2 catalytically cleaves the sn-2 position of acyl chain of phospholipids, while PLA_1 cleaves the sn-1 position, both release free fatty acid and lysophospholipid (160, 161). Patatin-like phospholipases (PLPs) have been originally reported in plants and are, for example, potato tuber storage proteins (162, 163), are a subfamily of PLA_2 enzymes. PNPLA domain PLPs are distributed both in plants and mammals. In plants, PLPs do not only act as enzymes to cleave fatty acids from membrane lipids, but also contribute to the defense against plant pathogens, while in mammals, PLPs are mostly involved in lipid metabolism and turnover (164). Besides plants and mammals, PLPs are also widespread in bacteria, particularly Gram-negative species. For example, ExoU, one well characterized PLP in *P. aeruginosa*, is activated by host ubiquitin and targets host cell membranes as a cytotoxic effector through the type III secretion system (T3SS) (165-168). But the ubiquitin is not the only activator for PLPs enzymatic activity. In *V. cholerae* El Tor, the phospholipase activity of CapV was activated upon binding to the novel bacterial second messenger 3'3'-cGAMP, leading to the degradation of specific bacterial cell membrane phospholipids including

phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) and the release of free fatty acids (FFAs) and lysophospholipids to eventually cause cell death (131).

The first 3'3'-cGAMP riboswitch was developed based on the c-di-GMP class I riboswitches (169, 170). Recent studies showed that similar to c-di-GMP, 3'3'-cGAMP in the Gram-negative environmental bacterium *G. sulfurreducens* utilizes a subclass of GEMM-I (GEMM-I b, Genes for the Environment, Membranes, and Motility) class riboswitch as specific receptor to regulate genes involved in exoelectrogenesis (134, 138). In addition, 3'3'-cGAMP signaling was found to be activated by the second messenger cyclic AMP (cAMP), but inhibited by c-di-GMP (158), suggesting a cross regulation of these nucleotide-based second messenger signaling pathways. However, no other receptors, effectors and activators of bacterial 3'3'-cGAMP signaling have been identified.

1.4 ADDITIONAL CYCLIC NUCLEOTIDE SECOND MESSENGERS

Furthermore, in recent years, there had been an explosion in the discovery of various bacterial cyclic di- and multiple nucleotide second messengers. CdnE, which is located on the horizonal transferred RB-HPI of *E. coli* ECOR31, uses both purine and pyrimidine nucleotides to synthesize cyclic UMP-AMP. Other homologs synthesize a diverse range of cDNs, including the cyclic trinucleotide AMP-AMP-GMP (125). Bioinformatic and biochemical analyses suggest CdnE to be a member of the cGAS/DncV-like nucleotidyltransferases (CD-NTases) family with homologs present in nearly every bacterial phylum. Similar to c-di-AMP and 3'3'-cGAMP, cyclic UMP-AMP and cyclic AMP-AMP-GMP can also bind to the RECON receptor in mammalian cells, but block its activity rather than stimulate it (125). The biological role of these molecules in bacteria, though, is still unknown. Besides cyclic di- and tri-nucleotides, cyclic oligonucleotides have recently been identified in bacteria embedded in CRISPR/Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated genes) antiviral defense systems (171, 172). For example, in *Streptococcus thermophilus*, Csm, an effector complex that is involved in bacterial type III CRISPR-Cas systems, synthesizes cyclic oligoadenylates (cAn; n = 2 to 6) via a GGDEF domain homolog, and then binds Csm6 ribonuclease to activate its nonspecific RNA degradation to prevent phage infection and propagation (171, 172).

2 AIMS OF THE THESIS

1. Providing the molecular basis to assess the regulatory mechanisms of semiconstitutive rdar biofilm formation by sequencing the genomes of eight semiconstitutive rdar biofilm forming commensal and uropathogenic *E. coli* strains by Illumina Miseq or PacBio sequencing (Paper I).

2. Definition of the cGAMP signaling network and investigation of its physiological impact and molecular mechanisms to regulate of rdar biofilm formation and motility in the animal commensal strain *E. coli* ECOR31 (Paper II).

3. Characterization of the role of CapV and $CapV_{Q329R}$ in the regulation of *E. coli* phenotypes such as cell morphology, flagella biosynthesis and activity and *csgD* dependent rdar biofilm formation (Paper III).

4. Analysis of the molecular basis of temperature-independent expression of the *csgD* dependent rdar morphotype by c-di-GMP turnover proteins in seven semiconstitutive rdar biofilm forming *E. coli* strains. In particular, to investigate the role of the c-di-GMP phosphodiesterase/diguanylate cyclase/trigger enzyme YciR to regulate rdar biofilm formation and *csgD* expression (Paper IV).

3 MATERIALS AND METHODS

Experimental methods frequently and uniquely used in this thesis are described below. A detailed description of experimental methods can be found in the *Materials and Methods* section of each paper.

3.1 DNA MANIPULATION OF *E. COLI*

Construction of gene deletions was performed by using the recombineering vector pSIM7, according to a protocol published previously with slight modifications (173). This plasmid is composed of an origin of replication and a segment of the bacteriophage λ genome comprising the Red recombinase genes (*exo*, *bet* and *gam*) under their native transcriptional control from the *pL* promoter and the temperature-sensitive repressor CI857, a variant of native CI repressor. At low temperature (30 - 34°C), the repressor is active and blocks the *pL* promoter, shutting off transcription of the Red genes. Following a brief temperature shift to 42°C results in a transient denaturation of the repressor, allowing the expression of *gam*, *bet* and *exo*. Upon transformation of PCR created dsDNA composed of an antibiotic marker flanked by FRT (Flippase recognition target) sites and nucleotide sequences homologous to the start and stop region of the target gene as short as 30 nucleotides, a gene can be replaced by the antibiotic resistance cassette. Upon shifting back to low temperature (30 - 34°C), the repressor again renatures and binds to *pL*, to turns off the Red system. The recombinant colonies can be selected on agar plates containing the respective antibiotic. The antibiotic marker can be excised using another temperature-sensitive helper plasmid, e.g. pCP20, which encodes the FLP recombinase that acts directly on the flanking FRT sites to excise the antibiotic gene marker.

3.2 PHENOTYPIC ANALYSES

3.2.1 Rdar morphotype

Predominantly sessile multicellular behavior, also known as biofilm formation, is a universal character of bacteria important for transmission, colonization and persistent infections. The rdar (red, dry and rough) morphotype commonly expressed by *S.* typhimurium and *E. coli* is a well-established biofilm model visualized by growth of bacteria on LB without salt agar plates as a distinct colony morphology. It is characterized by the expression of the adhesive biofilm extracellular matrix components cellulose and curli fimbriae, both of which can selectively bind to the dye Congo Red. Colonies were photographed at different time points to analyze the development of the morphology of the macrocolony and dye binding capacity indicative for differential expression of extracellular matrix components.

3.2.2 Motility assay

Apparent flagella-mediated swimming motility has been analyzed by measuring the swimming diameter in tryptone broth containing 0.25% agar, 0.5% NaCl, and 1% tryptone (23). Inoculation of bacteria into the agar of such a plate promotes movement of bacteria through the agar that can be detected as a ring or halo that enlarges over time depending on the ability of

the bacteria to swim and to conduct chemotaxis. Upon lack of apparent swimming motility, isolation of bacterial cell-associated flagellin can be a first step to extend and complement the flagella-mediated swimming motility assay in order to assess the molecular mechanisms of apparent swimming deficiency e.g. upon deletion of a gene or overexpression of a gene product.

3.2.3 Aggregation assay

Many bacterial cells including *E. coli* cells can bind to each other which is termed autoaggregation (174). Autoaggregation can be observed combining microscopic and macroscopic observation of the formation of bacterial multicellular aggregates that settle to the bottom of the culture medium when reaching a certain size. Bacterial aggregation was visually observed throughout the growth phase for up to 24 h in conjunction with detection of expression of the major biofilm regulator CsgD by Western blot analysis.

3.2.4 Biofilm formation on abiotic surfaces

Bacterial adhesion to abiotic surfaces under suitable growth conditions is an often used biofilm model (12, 175). Crystal violet staining of adherent cells and subsequent OD measurement of the amount of the bound dye after dissolution is a commonly used method to initially assess biofilm formation, for example, in a 96 well microtiter plate format. Crystal violet is a dye that interacts with negatively charged molecules (i.e., peptidoglycan) and the extracellular polysaccharide matrix (176), which makes it not only suitable for measuring the amount of biofilm, but also for the visualization of the pattern of biofilm distribution in the well.

3.3 ELECTRON MICROSCOPY

Electron microscopy (EM) uses the diffraction of electron to create an image of biological samples with high resolution. It is a commonly used technique for investigating the detailed structures of various biological samples (177). Transmission EM (TEM) and the scanning EM (SEM) are two major types of electron microscopy. In microbiology, TEM can be used, for example, to image the bacterial cell and the organization of bacterial filaments, e.g. the flagellum, in conjunction with the negative staining technique. SEM can be used to assess the details of bacterial cell morphology and atomic compositions in conjunction with secondary detectors.

3.4 REAL-TIME QUANTITATIVE REVERSE TRANSCRIPTION PCR

Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) is a nucleotide amplification-based method to detect and measure the mRNA levels for a target gene (178). To allow assessment of expression of the initial transcript, the mRNA is first transcribed into DNA by template-dependent reverse transcriptase. The double stranded DNA template is then exponentially amplified which allows relative quantification of the initial concentration according to the number of cycles required before a certain threshold concentration of the amplified DNA is reached. Two common methods have been applied for monitoring the amount of the PCR products generated by qRT-PCR: non-specific fluorescent dyes and sequence-specific DNA probes. The first approach uses a fluorescent dye, such as SYBR Green,

that intercalates non-specifically with any double-stranded DNA product, including PCR products; while the other approach consists of fluorescent reporter labelled oligonucleotides probe, which only detects DNA containing its complementary sequence. The two approaches require housekeeping gene(s) as internal standard for normalization, which are selected based on their almost constant level of expression. In this thesis, qRT-PCR was run with monitoring by SYBR green and the data were analyzed by the $2^{-\Delta\Delta CT}$ method (179). The *rpsV* gene was used as an endogenous control for internal normalization.

3.5 PROTEIN BIOCHEMICAL ASSAYS

3.5.1 Western blot analysis

Western blot is a widely applied technique used in cell and molecular biology to estimate the relative or absolute amount of expression of specific proteins within a cell extract after separation by protein gels (180). Proteins in a cell extract are separated by SDS-PAGE gel electrophoresis based on their negative charge upon binding of the sodium dodecyl sulfate detergent molecules which is correlated with the molecular weight, followed by an electrophoretic transfer onto a nitrocellulose or PVDF membrane. The membrane is then incubated with an antibody that recognizes specific epitope(s) of the protein of interest. The constant part of the first antibody is subsequently recognized by a secondary antibody to which e.g. a dye or an enzyme (frequently horseradish peroxidase) is covalently coupled. Oxidation of the substrate luminol by hydrogen peroxide results in the production of a chemo fluorescence signal which can be monitored.

3.5.2 Protein purification by affinity tag

Protein purification is a basic experimental approach that will subsequently allow characterization of functionality, structure and interactions of the protein of interest (181). Adding specific affinity tags, e.g. a 6*His-tag, to the protein of interest simplifies the purification of a recombinant fusion protein by employing affinity column chromatography for purification. During purification, the 6*His-tag of the protein of interest selectively binds divalent metal ions such as nickel and cobalt on affinized resin, while all untagged proteins pass through the column. The 6*His-tagged protein can be eluted with imidazole, which competes with the 6*His tag for binding to the column, or by a reduced pH (typically to 4.5 - 5.3), which decreases the affinity of the tag for the resin. The enzymatic activity of the purified protein can be subsequently assessed by the respective enzymatic assay.

3.5.3 Thin layer chromatography

Thin layer chromatography (TLC) is a straightforward method to analyze mixtures of small molecules by separating the compounds in the mixture by liquid/solid phase separation using e.g. an immobilized silica matrix coated on a support plate as the solid phase (182). TLC can be used to aid the determination of the complexity in a mixture, unravel the identity of a compound and assess the purity of a compound. As the solvent front subsequently moves along the stationary phase of the plate by capillary forces, for each molecule, an equilibrium is

established between adsorption onto the solid phase and dissolution in the solution. Due to their chemical structure and charge as well as the nature of the solid phase, the components differ in solubility and in the strength of their absorption to the absorbent, this equilibrium results in some components to be carried further up the plate than others. This distinct running behavior is quite reproducible for a given system and characterized by the R_f value. Thus, TLC was used to primarily analyze the enzymatic products of the respective proteins. Using a silica solid phase that contains a fluorescence indicator can detect molecules such as nucleotides that absorb light of the same wave length.

3.5.4 Tandem mass spectrometry

Tandem mass spectrometry (MS/MS) is a technique where two or more mass spectrometers are consecutively coupled to increase their resolution to analyze protein or peptides samples (183). The compounds are ionized and separated according to their mass-to-charge ratio (m/z) by one mass spectrometer and fragmented by a second mass spectrometer by their m/z ratio for identification. The enzymatic products of DncV were scraped out from the TLC plate to be identified by MS/MS using standard compounds.

3.6 BIOINFORMATIC ANALYSIS

3.6.1 Assembly and annotation of genomic DNA sequences

The genomic DNA from *E. coli* strains Tob1 and ECOR31 was sequenced with the PacBio RS II system (Pacific Biosciences; NGI Uppsala, Science For Life Laboratory [SciLifeLab], Uppsala, Sweden). The assembly was done on SMRT portal version 2.3, using HGAP3 with default settings. The genomic DNA from *E. coli* strains Fec67, Fec101, No.12, and B-11870 was sequenced using an Illumina MiSeq version 3 platform with read length up to 2 * 300 bp (NGI Stockholm, SciLifeLab, Solna, Sweden). *De novo* assembly was performed using SPAdes (http://bioinf.spbau.ru/spades), with contigs smaller than 500 bp omitted (184). The coverage was calculated according to Lander and Waterman (185). Initial analysis was performed by the Rapid Annotations using Subsystems Technology (RAST; version 2.0) server (http:// rast.nmpdr.org/rast.cgi) (186, 187). Genome sequences were submitted to DDBJ/ENA/GenBank and annotated with the NCBI Prokaryotic Genome Annotation Pipeline.

4 RESULTS AND DISCUSSION

4.1 PAPER I: DRAFT GENOME SEQUENCES OF SEMICONSTITUTIVE RED, DRY, AND ROUGH BIOFILM-FORMING COMMENSAL AND UROPATHOGENIC *ESCHERICHIA COLI* ISOLATES

E. coli is a well-studied model organism for various aspects of bacterial physiology and behavior. Biofilm formation is defined as multicellular microbial communities surrounded by a self-produced matrix of extracellular polymeric and non-polymeric substances adherent to each other, to interfaces and/or to biotic or abiotic surfaces (188). Strains of *E. coli* were proven to exhibit diverse biofilm phenotypes. A well-studied biofilm type characterized by the formation of an extracellular matrix consisting mainly of amyloid curli fimbriae and the exopolysaccharide cellulose is the red, dry, and rough (rdar) morphotype. Whereas *E. coli* K-12 model strains express at most a rudimentary rdar morphotype without production of cellulose exclusively at temperature below 30°C (18, 189), clinical isolates have been shown to frequently form the rdar morphotype semi-constitutively at 30°C and body temperature (190-192).

In this study, the genomes of eight semi-constitutive rdar biofilm forming *E. coli* strains including *E. coli* commensal strains Tob1, Fec67, Fec101, and ECOR31, *E. coli* UPEC strains No.12, B-11870, 80//6 and B8638 were sequenced by Illumina Miseq or PacBio sequencing. PacBio sequencing gave rise to 4 contigs for the human commensal strain *E. coli* Tob1(one plasmid) and 6 contigs for *E. coli* ECOR31 (three plasmids). The Illumina sequencing followed by SPAdes assembly resulted in assembly of the respective genomes into 90-140 scaffolds. In depth homology search by Blast (193) of characterized c-di-GMP turnover proteins showed that the ECOR31 genome comprises twelve proteins with a GGDEF domain, ten with an EAL domain and six with a GGDEF-EAL tandem. Besides, a putative 3’3’-cGAMP synthetase and putative c-di-AMP synthetase (DAC domain; data not shown) homolog have also been found encoded on the ECOR31 genome. All sequences have been submitted to the NCBI WGS database. In depth analysis of the sequenced genomes of semi-constitutive rdar morphotype expressing *E. coli* will shed light on the variability of the *E. coli* genomes, in particular, in this case, on the regulatory mechanisms of semi-constitutive versus temperature regulated rdar biofilm formation.

4.2 PAPER II: DNCV SYNTHESIZES CYCLIC GMP-AMP AND REGULATES BIOFILM FORMATION AND MOTILITY IN *ESCHERICHIA COLI* ECOR31

E. coli ECOR31 is an animal commensal strain isolated from leopard feces, which belongs to the *E. coli* reference strain collection (128). Similar to *E. coli* clinical isolates that usually shows semi-constitutive rdar morphotype expression, ECOR31 also displayed a semi-constitutive rdar morphotype on LB agar plates containing Congo Red at both 28°C and 37°C. Moreover, deletion of the *csgBA* genes encoding the major and minor subunits CsgA and CsgB of amyloid curli fimbriae, respectively, and *bcsA* encoding the cellulose synthase resulted in a pdar and bdar colony morphotype, respectively. The regulation of the rdar phenotype is controlled by the expression of the orphan response regulator CsgD protein, a transcriptional regulator. *CsgD* expression is regulated by a multitude of global regulatory pathways including the c-di-GMP signaling pathway in *E. coli* (18, 189). Cyclic di-GMP not only regulates the expression of cellulose and curli fimbriae, but also represses various modes of motility, such as flagellar-mediated swimming and swarming motility. Consistently, a *csgD* deletion mutant resulted in a saw colony morphotype, while swimming motility was not affected. Overexpression of c-d-GMP producing DGCs AdrA (7) and YdeH (7), upregulated the rdar morphotype and CsgD expression and downregulated swimming motility, while overexpression of c-d-GMP PDEs YhjH and YE2225 (194) downregulated the rdar morphotype and CsgD expression. Coexpression of AdrA or YdeH with increasing amounts of the PDE YhjH or YE2225 gradually relieved the inhibitory effect of the DGCs on swimming motility of ECOR31. Taken together, these results indicate that the rdar morphotype is dependent on the transcriptional activator CsgD and c-di-GMP signals in ECOR31 as it has been observed in other *E. coli* strains and other bacterial species (18, 62, 191, 195).

Bioinformatic analyses of the *E. coli* ECOR31 genome showed that it contains a DncV homolog (paper I), which has been demonstrated to synthesize 3'3'-cGAMP in *V. cholerae* biovar El Tor (4). ECOR31 contains a horizontally transferred high-pathogenicity island (HPI) encoding the siderophore yersiniabactin on the chromosome, which is a widely distributed virulence factor among *E. coli*, *Yersinia* spp. and other *Enterobacteria* that cause extra-intestinal infections (126, 129). The *dncV* locus is located on the 34.5 kbp ECOR31 right border HPI (RB-HPI) (Figure 6). Compared with DncV of *V. cholerae*, DncV of ECOR31 shows 61% identity and 74% in similarity. Blast search for DncV homologues with >40% identity indicates that DncV is widely distributed in various bacterial species and with DncV homologs to be classified into at least three subgroups based on their phylogenetic relationship.

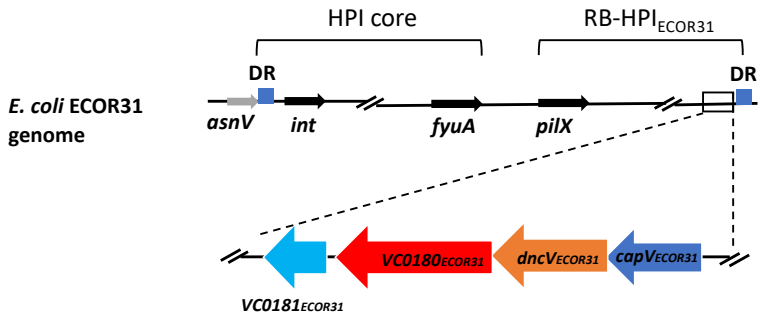

Figure 8 *E. coli* ECOR31 contains a horizonal transferred composite HPI homologs to the Yersiniabactin encoding HPI of *Yersinia*, *E. coli* and *Klebsiella* spp. and a RB-HPI, which contains part of the *Vibrio* 7th pandemic island-I (VSP-I) that encodes a 3'3'-cGAMP synthase *dncV* and its putative phospholipase receptor *capV* on a four gene operon. Grey arrow, HPI integration site. Blue boxes, direct repeat (DR) flanking the *E. coli* ECOR31 HPI.

In *V. cholerae*, DncV was found to regulate virulence and chemotaxis (4). Structural studies of DncV of *V. cholerae* demonstrated that amino acid Q110 is required for GTP binding, while amino acids D131 and D133 are critical catalytic residues that bind Mg^{2+} (132, 141, 144). Interestingly, though deletion of *dncV* showed no obvious effect on rdar biofilm formation, *dncV* overexpression downregulated the rdar biofilm morphotype and CsgD expression, as well as cellulose and curli fimbriae expressions in *E. coli* ECOR31. In contrast, the catalytic mutants of DncV, $DncV_{Q110A}$ and $DncV_{D129A/D131A}$, had no effect. Consistent with its effect on rdar morphotype expression, overexpression of *dncV* inhibited CsgD production leading to resolution of aggregate formation, while the two protein mutants had no effect compared to the ECOR31 vector control. qRT-PCR analysis showed that the mRNA levels of *csgD* were dramatically reduced upon *dncV* overexpression, but not upon overexpression of the mutant protein. On the other hand, overexpression of *dncV* enhanced biofilm formation on the abiotic surface after 24 h of growth in liquid culture, while it downregulated adherence after 48 h. Consistent with this finding, biofilm formation was down- and up-regulated, respectively, in the *dncV* deletion mutant. In conclusion, these data suggest the catalytic activity of *dncV* is required to negatively regulate the rdar morphotype and *csgD* mRNA steady state levels. The complex effect of *dncV* on biofilm formation indicated that *dncV* and/or 3'3'-cGAMP affects biofilm formation in multiple ways.

Besides rdar biofilm formation, DncV also inhibited swimming and swarming motility of *E. coli* ECOR31, while the DncV catalytic mutants had no effect. Consistently, a *dncV* chromosomal deletion mutant enhanced swimming and swarming motility compared to the ECOR31 wild type. Interestingly, in *V. cholerae*, expression of *dncV* affected chemotaxis, but had no effect on motility and biofilm formation (4). Moreover, coexpression of the motility-specific c-di-GMP PDE YhjH with DncV did not be relieved the swimming inhibition by DncV expression, indicating cGAMP may be responsible for the observed phenotype repression. Indeed, both the production of cell associated extracellular flagellin and the number of flagella per cell as visualized by TEM was dramatically reduced upon *dncV* expression. Flagella

biosynthesis is a complex process that involves flagellar regulon genes divided into three classes (196). qRT-PCR analysis showed that the level of mRNA for *flhD* encoding class 1 flagella regulator $FlhD_4C_2$ was not altered upon overexpression of *dncV*, while class 2 and class 3 gene expression as estimated by expression of the gene *fliA* encoding the flagella specific sigma factor and *fliC* encoding the flagellar subunit protein FliC, respectively, was diminished upon *dncV* overexpression. These findings indicate that in ECOR31, *dncV* inhibits motility by interacting with class 1 *flhDC* genes at the post-transcriptional level.

Results from the *in vitro* DncV enzymatic assay using purified DncV protein suggests that DncV can synthesize a major product which is different from c-di-AMP and c-di-GMP standards when incubated with its substrates ATP and GTP. Further analysis by MS/MS confirmed the product to be 3'3'-cyclic GAMP. Interestingly, using GTP and ATP as substrates, DncV can also synthesize c-di-GMP and c-di-AMP, respectively. In the presence of all four nucleotides, DncV only uses ATP or GTP, but not CTP or UTP as substrates, which indicates DncV to preferentially use purine nucleotide triphosphates as substrates. Interestingly, when incubated with CTP alone or GTP plus CTP, an unknown product was produced by DncV, suggesting that DncV can also use CTP as a substrate to some extent. This finding extends the substrate and possibly product spectrum of DncV homologs as a previous report documented that DncV from *V. cholerae* only uses ATP and/or GTP as a substrate(s) (4). By digestion with enzymes RNase T1, S1 nuclease, and RNase T2, the *in vitro* DncV synthesized products were confirmed to be 3'3'-cGAMP, 3'3' c-di-GMP, and 3'3' c-di-AMP, respectively. We also detected 3'3'-cGAMP upon overexpression of DncV in cell extracts, while in the vector control, upon overexpression of the catalytic mutant, no signal was detected. Moreover, the levels of *in vitro* bi-products c-di-GMP and c-di-AMP were under the detection limit in all assays.

We also found that, after mutation of the catalytic residues glutamine 110 to alanine, DncV almost completely lost its enzymatic activity; when both aspartate 129 and aspartate 131 were mutated to alanine, DncV lost its enzymatic activity completely, which indicates the enzymatic activity of DncV is partially dependent on glutamine 110, and absolutely dependent on aspartic acid 129 in combination with aspartic acid 131. Cumulatively, these results further demonstrate that DncV synthesizes 3'3'-cGAMP and regulates biofilm formation and motility in *E. coli* ECOR31.

4.3 PAPER III: A SINGLE AMINO ACID SUBSTITUTION IN CAPV LEADS TO PRONOUNCED CELL FILAMENTATION IN *ESCHERICHIA COLI*

In the *E. coli* ECOR31 genome, *dncV* is flanked by the *V. cholerae* VSP-1 homologs of *capV* upstream and *vc0180* and *vc0181* downstream (126, 130) (Figure 8, paper II). In this paper, we characterized $CapV_{Q329R}$, a variant of the patatin-like phospholipase CapV of strain ECOR31, demonstrating that its expression from a plasmid induces pronounced cell filamentation, repressed apparent swimming motility and flagella expression and rdar biofilm formation. This effect was not only observed in *E. coli* MG1655, but also in other *E. coli* commensal and clinical isolates as well as *S.* typhimurium UMR1, showing that induction of this morphological and physiological phenomenon is not only restricted to a single strain.

In *V. cholerae*, CapV, which encodes a patatin-like phospholipase A (PNPLA), was identified as a cGAMP receptor that results in growth retardation upon activation (131). Bioinformatic analysis showed that CapV_ECOR31, which shows 65% amino acid sequence identity compared to CapV_*V.cholerae,* also contains a canonical PNPLA domain with the three main characteristic conserved signature motifs (160), the phosphate or anion binding motif G-G-G-x-[K/R]-G, the esterase box G-x-S-x-G, and D-G-[A/G], a part of the catalytic dyad. Blast homology search of CapV_ECOR31 demonstrated that CapV_ECOR31 homologs with >60% identity is not only found in *E. coli* and *V. cholerae* strains, but also in many other bacterial species. Phylogenetic analysis of representative CapV_ECOR31 homologs supported classification into four different subgroups.

Surprisingly, overexpression of CapV did not affect cell growth in *E. coli* MG1655, while a variant of CapV_ECOR31, $CapV_{Q329R}$, induced a mild cell growth arrest after 7 h as estimated by the optical density (OD) of the cell suspension. Since the enzymatic activity of CapV_*V.cholerae* is activated by 3'3'-cGAMP which *E. coli* MG1655 does not possess, the results suggest that the $CapV_{Q329R}$ induced apparent arrest of cell growth may occur via a different mechanism, rather than by activation of CapV by the 3'3'-cGAMP signaling pathway and subsequent cell lysis.

For rod-shaped cells, filamentation is a dramatic morphology adaptation under various stressful growth conditions. After exposure to stress, cell filamentation often occurs upon activation of the SOS response system and expression of the cell division inhibitor SulA, resulting in a cell division block (197-199). Surprisingly, we found that $CapV_{Q329R}$ induced extensive *sulA*-independent cell filamentation upon overexpression in *E. coli* MG1655, while CapV wild type had only a minor effect. $CapV_{Q329R}$-induced cell filamentation was initiated 2 h after the initiation of $CapV_{Q329R}$ expression, and continuous filament elongation was observed for around 6 h, while cells returned to rod-shape after 22 h. Indeed, follow-up experiments showed that the observed cell filamentation phenotype is dependent on the L-arabinose concentration and subsequently the $CapV_{Q329R}$ expression level. Moreover, results from mutation analyses indicate that the G-G-G-x-[K/R]-G and D-G-[AG] motifs, but not the G-x-S-x-G motif in the canonical PNPLA domain are required for $CapV_{Q329R}$ functionality. We further showed that, under our experimental conditions, $CapV_{Q329R}$ did hardly affect the viability of MG1655 cells.

Since the *in vivo* 3'3'-cGAMP signal was under the detection limit (data not shown), we assumed that the $CapV_{Q329R}$-induced OD retardation suggesting an apparent cell growth arrest might be due to the pronounced cell filamentation. This hypothesis is in line with the viability counts. Alignment of the amino acid sequences of CapV homologues extracted by Blast homology search showed that Q329 is highly conserved among the CapV proteins. The position of Q329 is located outside of the putative PNPLA domain of CapV_ECOR31 and consequently not part of the characteristic motifs required for the catalytic activity of this phospholipase A_2 (PLA_2) superfamily member (159, 160). 3D model construction using the closest homolog of CapV_ECOR31 in the PDB database (FabD from *Solanum cardiophyllum,* PDB: 1oxwC) revealed that Q329_CapV is located in the context of the $RARGRR_{329}$ sequence within the third last α helix with the arginine side chain pointing outwards causing no change in the overall structure of CapV_ECOR31. However, the switch from a potential hydrogen-bond acceptor (Q) to a hydrogen-bond donor (R) with a longer side chain indicates that the CapV_ECOR31 mutant might shows altered enzymatic activity, ligand binding properties or protein-protein interactions. Furthermore, Arginine with its positively charged side chain is involved in a variety of different functionalities such as in binding of negative charged molecules such as phosphates present, for example, in DNA molecules or nucleotides and the RR twin arginine motif is part of the N-terminal signal sequence for the Twin-Arginine Translocation (TAt) pathway (200, 201). Furthermore, a RxxR motif constitutes a conserved peptidase cleavage site, while an RxxxR motif is part of the binding motif for c-di-GMP in PilZ domain proteins, whereby Arginine residues can bind O-6 and N-7 at the Hoogsteen edge of the guanine base (202).

Filamentation induced by $CapV_{Q329R}$ is regulated by environmental conditions. We found that the $CapV_{Q329R}$ overexpression cells are less filamentous in LB medium than in TB medium. Supplementation with pyridoxine (vitamin B6) at 5 mg/ml in TB dramatically restricted $CapV_{Q329R}$-induced cell filamentation, while supplementation with other B vitamins had no effect. $CapV_{Q329R}$ also inhibited swimming motility of MG1655 in the soft agar plate. Microscopic observation showed that the highly filamentous cells induced by $CapV_{Q329R}$ with more than 20 times the length did not show any movement, while shorter filaments were still motile, indicating that cells might gradually loose flagella and become non-motile upon $CapV_{Q329R}$ overexpression.

Interestingly, expression of $CapV_{Q329R}$ also modulated rdar biofilm formation in *E. coli* strains. In the UPEC strain *E. coli* No.12 and other *E. coli* strains, $CapV_{Q329R}$ expression is accompanied by downregulation of the expression of the transcriptional regulator CsgD and the levels of the ubiquitous second messenger c-di-GMP. Besides c-di-GMP, the level of cAMP and 3'3'-cGAMP, both of which were reported to participate in biofilm regulation (130, 203), was also altered. In UPEC, the cAMP-CRP complex regulates curli and cellulose production and the formation of rugose and pellicle biofilms through activation of *csgD* expression (203). In Paper II, we demonstrated that DncV synthesized 3'3'-cGAMP and participated in regulation of rdar biofilm formation in *E. coli* ECOR31 (130). Besides DncV, Hypr GGDEF domain containing proteins can also produce 3'3'-cGAMP (133, 135),

indicating Hypr GGDEF domain proteins might be directly or indirectly activated by $CapV_{Q329R}$ to produce 3'3'-cGAMP, and thus inhibit rdar biofilm formation of *E. coli* No.12. Moreover in *M. xanthus*, cAMP is an activator of the Hypr GGDEF enzyme GacB to produce 3'3'-cGAMP, whereas GacB is inhibited directly by c-di-GMP (158), providing evidence for cross regulation between different cDNs signaling pathways.

We also found that $CapV_{Q329R}$ induced cell filamentation without affecting cell arrangement of *E. coli* No.12 upon agar plate growth. The length of the filamentous cells on the agar plate is extended compared to TB medium suggesting that surface sensing or another environmental parameter associated with agar plate growth induced extended elongation. The $CapV_{Q329R}$ induced cell filamentation, swimming motility and rdar biofilm repression was not restricted to *E. coli* MG1655 and No.12, but also common to other *E. coli* strains and even *S.* typhimurium UMR1, suggesting a general role of $CapV_{Q329R}$ on various aspects of bacterial physiology.

4.4 PAPER IV: ALTERNATIONS OF C-DI-GMP TURNOVER PROTEINS MODULATE SEMI-CONSTITUTIVE RDAR BIOFILM FORMATION IN COMMENSAL AND UROPATHOGENIC *ESCHERICHIA COLI*

The expression of the rdar morphotype varies among different *E. coli* strains. For example, while certain pathogenic *E. coli* strains frequently express a typical rdar morphotype (expression at 28°C only), some commensal, but also UPEC *E. coli* strains express a semi-constitutive rdar morphotype (expression at both 28°C and 37°C at a higher level) (127, 191, 192). In this paper, we aimed to investigate the underlying molecular mechanism of semi-constitutive rdar biofilm formation in three commensal *E. coli* strains, Tob1, Fec67, and Fec101, and four UPEC isolates, No.12, B-11870, 80//6, and B-8638. As a reference, we used the human commensal strain *E. coli* Fec10, a close homolog of *E. coli* K-12, which expresses a temperature-dependent rdar morphotype.

The rdar morphotype is characterized by the expression of extracellular matrix components cellulose and curli fimbriae in *E. coli* (18). We found that all isolates expressed the rdar morphotype at both 28°C and 37°C, but showed variations in color intensity and roughness, suggesting variable ratio of cellulose versus curli fimbriae expression and variability in expression of the master transcriptional regulator CsgD. CsgD expression accompanied rdar morphotype expression for all the strains at 28°C and 37°C, however, limited CsgD expression at 37°C was observed for Fec67 and 80//6. This is consistent with the previous finding that the rdar morphotype is independent on CsgD in the probiotic strain *E. coli* Nissle 1917 (204). However, a deletion mutant for *csgD* has to be constructed to confirm this hypothesis. Besides, cell aggregation and swimming motility also varied among the strains. While strains Fec101, B-11870, 80//6 showed strong adherence to glass (ring formation) and clumping, other strains did not display this phenotype. Only Tob1 showed a pronounced swimming motility, other strains were either nonmotile or much restricted in swimming motility compared to Tob1. Collectively, these results demonstrated that the level of rdar biofilm morphotype expression on Congo Red agar plate and biofilm formation in liquid culture cannot be directly correlated as previously observed (191, 205). Alternative extracellular matrix components other than curli fimbriae and cellulose can contribute to biofilm formation in liquid culture.

In paper I, the genomes of commensal and pathogenic semi-constitutive rdar biofilm forming *E. coli* isolates were sequenced. *E. coli* has previously been phylogenetically classified into 7 phylogroups (206). Genomic analysis indicated that strain Fec101 belongs to phylogroup B1, UPEC strain B-8638 to group D, and the remaining strains to group B2. Rdar morphotype expression and regulation in *E. coli* is characterized by distinct gene modules. Bioinformatic analysis found that the semi-constitutive *E. coli* isolates differ in their genomic content leading to rdar biofilm formation. The genomes of commensal strains Tob1, Fec101, and UPEC B-8638 lack the *pga* operon, which encodes the biosynthesis gene products for poly-β-1,6-*N*-acetyl-D-glucosamine (PNAG), a linear exopolysaccharide homopolymer that serves as an adhesin for the maintenance of the structural stability of biofilms in diverse eubacteria (207). In contrast, the *pga* operon is present in the commensal strain Fec67 and UPEC strain B-11870, 80//6, and No.12, but in a different genomic context compared to *E. coli* K-12. The biosynthesis

operons for curli (*csgBAC* and *csgDEFG* operons (189, 208)) and cellulose (*bcsABZC* and *bcsEFG* operons (13, 209)) are present in all strains.

The ubiquitous second messenger c-di-GMP is the key regulator of rdar biofilm formation in *Enterobacteriaceae*, in particular in *Escherichia* and *Salmonella* spp. (2, 18). The complex c-di-GMP signaling network involves numerous c-di-GMP turnover proteins, e.g. DGCs and PDEs. We also observed a high variability in the number of c-di-GMP turnover proteins in the genomes of the semi-constitutive *E. coli* strains using *E. coli* K-12 strain MG1655 and *E. coli* Fec10 as a reference (63). Specifically, the DGCs YcdT, and YddV/DosC are not present in three (Tob1, Fec101, B-8638) and four (Tob1, Fec67, B-11870, No.12) strains, respectively, while other conserved core reference c-di-GMP turnover proteins including AdrA, YliF, Rtn, YhjH, YhjK, YlaB, YhdA, YeaI are present but show substantial amino acid variations in all strains. Moreover, stop codons and frameshift mutations are found within several open reading frames (ORFs), indicating potentially nonfunctional truncated proteins. Interestingly, several c-di-GMP turnover proteins not present in *E. coli* K-12 (63) are also found in several strains, including the GGDEF protein DgcX (UniProt B7LBD9_ECO55) in Fec101, the EAL protein PdeX (UniProt Q707K1_ECOLX) in Tob1, and PdeY (UniProt Q1RDG4_ECOUT) in No.12, B-11870 and Fec67. In addition, three novel EAL domain proteins were identified in UPEC strains, namely PdeU1 in B-11870 and 80//6, PdeU2 in B-11870, and PdeU3 in B-8638. Previous studies indicated that c-di-GMP turnover proteins can potentially contribute to the semi-constitutive rdar morphotype. For example, the *E. coli* probiotic strain Nissle 1917 and commensal strain 1094 show CsgD and DGC AdrA independent cellulose expression (204, 210).

In *E. coli*, YciR has been identified as a trigger enzyme that, besides possessing dual catalytic function, controls *csgD* transcription and thus affect rdar biofilm formation through protein-protein interactions (67, 68, 211). Bioinformatic analysis showed that YciR contains a tandem PAS-GGDEF-EAL domain, and distinct amino acid substitutions are found in all semi-constitutive rdar strains compared to YciR from *E. coli* K-12 and Fec10. In strain B-11870, a nonsense mutation resulted in the insertion of a stop codon into the EAL domain, which leads to the truncated PAS-GGDEF $YciR_{B\text{-}11870}$ protein. Overexpression of YciR from Fec10 and Fec101 dramatically downregulated the rdar morphotype of Tob1 at both temperatures concomitant with downregulation of CsgD expression, whereas YciR from Tob1 only decreased the rdar morphotype moderately. Overexpression of YciR from B-11870 predominantly upregulated the rdar morphotype and CsgD expression at 28°C, while overexpression of YciR ($YciR_{B\text{-}11870\_Full}$) with the reversion of the stop codon into a sense codon resulted in decreased rdar morphotype at both temperatures.

Of note, YciR from *E. coli* Tob1 contains eleven amino acid substitutions compared to YciR from Fec10 and exhibits substantially altered activity. Results from the site-directed mutants demonstrate that T37I, T371A, and combined T37I/T371A mutants of YciR from Fec10 resembled the protein activity of YciR from Tob1 on rdar morphotype of Tob1 at 28°C, while YciR variants T37I, A90G, A361S and T371A closely resembled YciR from Tob1 at 37°C.

Other mutants are either failed to assemble or showed inconsistent phenotype. Similarly, the same result was observed when mutant $YciR_{T37I}$, $YciR_{T371A}$, and combined $YciR_{T37I/T371A}$ derived from Fec101 was introduced into Tob1. In conclusion, T37I is the most determinative amino acid exchange in combination with T371A. Of note, T37 and T371 of YciR are located outside of the catalytic motifs and on GGEDF motif, suggesting that these amino acid substitutions probably affect protein–protein interactions. In addition, mutations in the consensus catalytic motifs in YciR from Fec101 and YciR from Tob1 consistently showed that the $E_{440}AL$ mutant of YciR from Fec101 and YciR from Tob1 downregulated the rdar morphotype of Tob1 more effectively than the respective wild type proteins, while the $GGD_{316}E_{317}F$ mutant of YciR from Fec101 and YciR from Tob1 resulted in a colony with increased roughness and increased fluorescence under UV light compared to the wild type protein, suggesting enhanced cellulose production.

5 CONCLUDING REMARKS AND PERSPECTIVES

In this thesis, we characterized the networks of metabolizing proteins for the second messengers c-di-GMP and 3'3'-cGAMP in *E. coli* with respect to the regulation of multicellular behavior such as semi-constitutive rdar biofilm formation and swimming motility (Paper I, Paper II, Paper III, Paper IV) as well as regulation of cell morphology (Paper III).

3'3'-cGAMP is novel cyclic dinucleotide second messenger that regulates many aspects of bacterial physiology. We demonstrated that DncV synthesized 3'3'-cGAMP and regulates biofilm formation and flagella biosynthesis in the animal commensal strain *E. coli* ECOR31 (Paper I). However, the 3'3'-cGAMP turnover proteins, PDEs and GACs as well as receptors/effectors involved in these phenotypes are still not identified. Whether 3'3'-cGAMP is involved in the regulation of alternative bacterial phenotypes also need to be further investigated.

We also characterized the role of the patatin-like phospholipase CapV and its $CapV_{Q329R}$ variant in the regulation of phenotypes such as cell morphology (filamentation), biosynthesis of flagella and swimming motility and rdar biofilm formation and CsgD expression (Paper III). *In vivo* evidence suggests that CapV and $CapV_{Q329R}$ might exert as a patatin-like phospholipase as we observed a dramatic change in the content of charged phospholipids (unpublished data), however, further biochemical and structural studies need to be conducted to confirm this hypothesis and give insights into the detailed molecular mechanisms of the enzymatic functionality such as substrate specificity and regulation of the catalytic activity. It will also be interesting to further unravel the molecular basis of extended cell filamentation, and repression of biofilm formation and flagella biosynthesis and the overall physiological changes that occur upon expression of $CapV_{Q329R}$.

We also analyzed the genomes of seven semi-constitutive rdar biofilm forming *E. coli* strains and investigated the role of the c-di-GMP turnover and trigger protein YciR to regulate rdar biofilm formation and *csgD* expression (Paper I, Paper IV). Besides YciR, other c-di-GMP turnover proteins show distinct amino acid substitutions correlated with semi-constitutive rdar morphotype expression (unpublished data), however, whether and how these gene products contribute to the regulation of the temperature-independent expression of the rdar morphotype and biofilm components needs to be further tested.

6 ACKNOWLEDGEMENTS

To be a PhD student and obtain a PhD degree is not easy. I still clearly remember the first day when I landed in the foreign country, the first day when I joined the group and the first experiment I did. Everything is so fresh in my head just like I did yesterday. But if there's time machine that takes me back to six years ago, I will definitely make the same decision, to be a PhD, not only because for the love for science, but you, who made this journey unique and irreplaceable.

First of all, I would like to express my sincere gratitude to my supervisor Prof. **Ute Römling** for the continuous support of my PhD study and related research, for your patience, motivation, and immense knowledge, and most importantly, for triggering my desires for the wonderful microbiology world. Your guidance helped me in all the time of research and writing of this thesis, especially help me walk through my desperate dark moments of research. All of these efforts you made on me in these years will benefit my future work. I could not have imagined having a better supervisor for my PhD study. It's so lucky to be a student of you, thanks for the opportunity!

Besides my main supervisor, I would like to thank my co-supervisor Prof. **Mark Gomelsky** and Dr. **Annika Cimdins-Ahne**. **Mark**, thank you for all of the inspiring scientific discussions and advices during your visit in the lab and joint group meetings after you went to back to US, I learned a lot from you. **Annika**, I might not be a good student, but you truly are a good supervisor and friend. You are always so patient with me and have so many scientific suggestions on the projects. You set a good example of how to be a qualified supervisor, thank you!

I would like to thank my mentor Prof. **Hans-Gustaf Ljunggren**. Thanks for being my mentor for my PhD study. I was always encouraged by your wisdom through our nice talks!

To my colleagues at UR lab, **Shady Kamal** and **Sulman Shafeeq**, you guys are the most awesome people in this world, I cannot imagine the days without you guys in the lab. Thanks for the stimulating discussions, for the days we were working together, and for all the fun we have had in the last years. **Shady**, Yo, my PhD buddy, you were here since I joined the lab, and we have been through so much together and I have to say that we witnessed the glorious period of our lab history. I was always encouraged by your enthusiasm and optimism when shit happens in life. I wish all my best for your PhD defense and future work! Mr. **Sulman**, the man with broad knowledge of science and daily life. Though sometimes I want to disrupt your endless talk when I ask you questions, I always find an answer from you. I wish I could have learned a lot more from you in these years, especially your kindness, generosity, and always being cool and keep positive under pressure.

To the students I had the honor to supervise, **Lydia**, **Filip**, **Yasemin**, **Vittoria**, **Aysha** and **Elsa**. You all are not only students of me, but also I learned from you about how to be a supervisor.

Thanks for the patience with me and I really appreciate your contributions to my projects. I wish you all good in your future studies. The future is yours, you youngsters!

To the people who have been in the lab, **Changhan** and **Hyunehee**, I was always affected by your funny jokes during group lunch and fika, haha, and so happy to hear that you had a baby, enjoy lives in US! **Lei**, thanks for the wonderful talks while working in the lab and your generous help as a good friend, and the delicious food by your Mom, bring my best wishes to her. **Ying**, thanks for your endless talking to me, I really wish I could be as talkative as you, stay safe in Gemerny. **Jonathan**, thanks for the joyful discussions and suggestions with your magic jokes in the last years, you are a good listener and friend. **Stefanie**, though you stayed not as long as others in the group, you are quite impressive and offered me lots of help, nothing is better than talking with a beauty.

I would like to acknowledge all the collaborators for their contributions in the publications in this thesis: **Manfred Rohde**, **Lothar Jänsch**, **Volkhard Kaever**, **Manfred Nimtz**, **Heike Bähre**, **Petra Lüthje**, **Irfan Ahmad**, **Roger Simm**, **Annelie Brauner**, **Kaisa Thorell** and **Åsa Sjöling**, thanks for the fruitful collaborations.

Special thanks to the Professors at KI and MTC administration and organization group who participated in my PhD. Prof. **Roland Möllby** for the wonderful biosafety class. Prof. **Ingemar Ernberg** for being my chairman. **Åsa**, **Eva**, **Gesan**, **Kristina** for your support and guidance.

To all my friends in Sweden during my PhD, **Xiaogang**, **Yu**, **Honglei**, **Qingyang**, **Xiaolei**, **Jing**, **Jiawei**, **Zilong**, **Lifeng**, **Leonie**, **Sharesta**, **Patricia**, **DaeHo**... Many thanks to you guys, I had a fantastic time with you!

I also would like to thank my friends in China, **Nan**, **Rui**, **Xuanduo**, **Mingshu**, for the constant friendship and generous support!

Last but not the least, I would like to thank my family: my parents and my sister for supporting me spiritually throughout my PhD and my life in general. I love you all!

7 REFERENCES

1. Opoku-Temeng C, Zhou J, Zheng Y, Su J, Sintim HO. 2016. Cyclic dinucleotide (c-di-GMP, c-di-AMP, and cGAMP) signalings have come of age to be inhibited by small molecules. Chem Commun (Camb) 52:9327-42.

2. Römling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev 77:1-52.

3. Römling U. 2012. Cyclic di-GMP, an established secondary messenger still speeding up. Environmental microbiology 14:1817-1829.

4. Davies BW, Bogard RW, Young TS, Mekalanos JJ. 2012. Coordinated regulation of accessory genetic elements produces cyclic di-nucleotides for *V. cholerae* virulence. Cell 149:358-370.

5. Kalia D, Merey G, Nakayama S, Zheng Y, Zhou J, Luo Y, Guo M, Roembke BT, Sintim HO. 2013. Nucleotide, c-di-GMP, c-di-AMP, cGMP, cAMP,(p) ppGpp signaling in bacteria and implications in pathogenesis. Chemical Society Reviews 42:305-341.

6. Tamayo R, Pratt JT, Camilli A. 2007. Roles of cyclic diguanylate in the regulation of bacterial pathogenesis. Annual review of microbiology 61:131.

7. Simm R, Morr M, Kader A, Nimtz M, Römling U. 2004. GGDEF and EAL domains inversely regulate cyclic di-GMP levels and transition from sessility to motility. Molecular microbiology 53:1123-1134.

8. Commichau FM, Dickmanns A, Gundlach J, Ficner R, Stülke J. 2015. A jack of all trades: the multiple roles of the unique essential second messenger cyclic di-AMP. Molecular microbiology 97:189-204.

9. Danilchanka O, Mekalanos JJ. 2013. Cyclic dinucleotides and the innate immune response. Cell 154:962-970.

10. Ross P, Weinhouse H, Aloni Y, Michaeli D, Weinberger-Ohana P, Mayer R, Braun S, De Vroom E, Van der Marel G, Van Boom J. 1987. Regulation of cellulose synthesis in *Acetobacter xylinum* by cyclic diguanylic acid. Nature 325:279-281.

11. Galperin MY. 2004. Bacterial signal transduction network in a genomic perspective. Environmental microbiology 6:552-567.

12. O'Toole G, Kaplan HB, Kolter R. 2000. Biofilm formation as microbial development. Annual Reviews in Microbiology 54:49-79.

13. Römling U, Galperin MY. 2015. Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. Trends in microbiology 23:545-557.

14. Römling U, Balsalobre C. 2012. Biofilm infections, their resilience to therapy and innovative treatment strategies. Journal of internal medicine 272:541-561.

15. Römling U, Sierralta WD, Eriksson K, Normark S. 1998. Multicellular and aggregative behaviour of *Salmonella* typhimurium strains is controlled by mutations in the *agfD* promoter. Molecular microbiology 28:249-264.

16. Grantcharova N, Peters V, Monteiro C, Zakikhany K, Römling U. 2010. Bistable expression of CsgD in biofilm development of *Salmonella enterica* serovar typhimurium. Journal of bacteriology 192:456-466.

17. Manuel CM, Nunes OC, Melo L. 2007. Dynamics of drinking water biofilm in flow/non-flow conditions. Water Research 41:551-562.

18. Römling U. 2005. Characterization of the rdar morphotype, a multicellular behaviour in *Enterobacteriaceae*. Cellular and Molecular Life Sciences CMLS 62:1234-1246.

19. Harshey RM. 2003. Bacterial motility on a surface: many ways to a common goal. Annual Reviews in Microbiology 57:249-273.

20. Merz AJ, So M, Sheetz MP. 2000. Pilus retraction powers bacterial twitching motility. Nature 407:98.

21. Wolfe AJ, Visick KL. 2008. Get the message out: cyclic-di-GMP regulates multiple levels of flagellum-based motility. Journal of bacteriology 190:463-475.

22. Girgis HS, Liu Y, Ryu WS, Tavazoie S. 2007. A comprehensive genetic characterization of bacterial motility. PLoS genetics 3:e154.

23. Fang X, Gomelsky M. 2010. A post-translational, c-di-GMP-dependent mechanism regulating flagellar motility. Molecular microbiology 76:1295-1305.

24. Paul K, Nieto V, Carlquist WC, Blair DF, Harshey RM. 2010. The c-di-GMP binding protein YcgR controls flagellar motor direction and speed to affect chemotaxis by a "backstop brake" mechanism. Molecular cell 38:128-139.

25. Boehm A, Kaiser M, Li H, Spangler C, Kasper CA, Ackermann M, Kaever V, Sourjik V, Roth V, Jenal U. 2010. Second messenger-mediated adjustment of bacterial swimming velocity. Cell 141:107-116.

26. Trampari E, Stevenson CE, Little RH, Wilhelm T, Lawson DM, Malone JG. 2015. Bacterial rotary export ATPases are allosterically regulated by the nucleotide second messenger cyclic-di-GMP. Journal of Biological Chemistry 290:24470-24483.

27. Krasteva PV, Fong JC, Shikuma NJ, Beyhan S, Navarro MV, Yildiz FH, Sondermann H. 2010. *Vibrio cholerae* VpsT regulates matrix production and motility by directly sensing cyclic di-GMP. Science 327:866-868.

28. Jenal U, Malone J. 2006. Mechanisms of cyclic-di-GMP signaling in bacteria. Annu Rev Genet 40:385-407.

29. Lori C, Ozaki S, Steiner S, Böhm R, Abel S, Dubey BN, Schirmer T, Hiller S, Jenal U. 2015. Cyclic di-GMP acts as a cell cycle oscillator to drive chromosome replication. Nature 523:236-239.

30. Paul R, Weiser S, Amiot NC, Chan C, Schirmer T, Giese B, Jenal U. 2004. Cell cycle-dependent dynamic localization of a bacterial response regulator with a novel di-guanylate cyclase output domain. Genes & development 18:715-727.

31. Aldridge P, Paul R, Goymer P, Rainey P, Jenal U. 2003. Role of the GGDEF regulator PleD in polar development of *Caulobacter crescentus*. Molecular microbiology 47:1695-1708.

32. Duerig A, Abel S, Folcher M, Nicollier M, Schwede T, Amiot N, Giese B, Jenal U. 2009. Second messenger-mediated spatiotemporal control of protein degradation regulates bacterial cell cycle progression. Genes & development 23:93-104.

33. McGrath PT, Iniesta AA, Ryan KR, Shapiro L, McAdams HH. 2006. A dynamically localized protease complex and a polar specificity factor control a cell cycle master regulator. Cell 124:535-547.

34. Jenal U, Fuchs T. 1998. An essential protease involved in bacterial cell-cycle control. The EMBO journal 17:5658-5669.

35. Abel S, Chien P, Wassmann P, Schirmer T, Kaever V, Laub MT, Baker TA, Jenal U. 2011. Regulatory cohesion of cell cycle and cell differentiation through interlinked phosphorylation and second messenger networks. Molecular cell 43:550-560.

36. Kim HK, Harshey RM. 2016. A diguanylate cyclase acts as a cell division inhibitor in a two-step response to reductive and envelope stresses. MBio 7:e00822-16.

37. Gupta KR, Baloni P, Indi SS, Chatterji D. 2016. Regulation of growth, cell shape, cell division, and gene expression by second messengers (p) ppGpp and cyclic Di-GMP in *Mycobacterium smegmatis*. Journal of bacteriology 198:1414-1422.

38. He M, Ouyang Z, Troxell B, Xu H, Moh A, Piesman J, Norgard MV, Gomelsky M, Yang XF. 2011. Cyclic di-GMP is essential for the survival of the Lyme disease spirochete in ticks. PLoS pathogens 7:e1002133.

39. Neunuebel MR, Golden JW. 2008. The *Anabaena sp*. strain PCC 7120 gene *all2874* encodes a diguanylate cyclase and is required for normal heterocyst development under high-light growth conditions. Journal of bacteriology 190:6829-6836.

40. Bharati BK, Sharma IM, Kasetty S, Kumar M, Mukherjee R, Chatterji D. 2012. A full-length bifunctional protein involved in c-di-GMP turnover is required for long-term survival under nutrient starvation in *Mycobacterium smegmatis*. Microbiology 158:1415-1427.

41. Li W, He Z-G. 2012. LtmA, a novel cyclic di-GMP-responsive activator, broadly regulates the expression of lipid transport and metabolism genes in *Mycobacterium smegmatis*. Nucleic acids research 40:11292-11307.

42. Li W, Li M, Hu L, Zhu J, Xie Z, Chen J, He Z-G. 2018. HpoR, a novel c-di-GMP effective transcription factor, links the second messenger's regulatory function to the mycobacterial antioxidant defense. Nucleic acids research 46:3595-3611.

43. Seshasayee AS, Fraser GM, Luscombe NM. 2010. Comparative genomics of cyclic-di-GMP signalling in bacteria: post-translational regulation and catalytic activity. Nucleic Acids Res 38:5970-81.

44. Chan C, Paul R, Samoray D, Amiot NC, Giese B, Jenal U, Schirmer T. 2004. Structural basis of activity and allosteric control of diguanylate cyclase. Proceedings of the National Academy of Sciences of the United States of America 101:17084-17089.

45. Christen B, Christen M, Paul R, Schmid F, Folcher M, Jenoe P, Meuwly M, Jenal U. 2006. Allosteric control of cyclic di-GMP signaling. Journal of Biological Chemistry 281:32015-32024.

46. Dahlstrom KM, Giglio KM, Collins AJ, Sondermann H, O'Toole GA. 2015. Contribution of physical interactions to signaling specificity between a diguanylate cyclase and its effector. MBio 6:e01978-15.

47. Dahlstrom KM, Giglio KM, Sondermann H, O'Toole GA. 2016. The inhibitory site of a diguanylate cyclase is a necessary element for interaction and signaling with an effector protein. Journal of bacteriology 198:1595-1603.

48. Orr MW, Donaldson GP, Severin GB, Wang J, Sintim HO, Waters CM, Lee VT. 2015. Oligoribonuclease is the primary degradative enzyme for pGpG in

Pseudomonas aeruginosa that is required for cyclic-di-GMP turnover. Proceedings of the National Academy of Sciences 112:E5048-E5057.

49. Cohen D, Mechold U, Nevenzal H, Yarmiyhu Y, Randall TE, Bay DC, Rich JD, Parsek MR, Kaever V, Harrison JJ. 2015. Oligoribonuclease is a central feature of cyclic diguanylate signaling in *Pseudomonas aeruginosa*. Proceedings of the National Academy of Sciences 112:11359-11364.

50. Petchiappan A, Naik SY, Chatterji D. 2020. Tracking the homeostasis of second messenger cyclic-di-GMP in bacteria. Biophysical Reviews:1-12.

51. Römling U, Liang Z-X, Dow JM. 2017. Progress in understanding the molecular basis underlying functional diversification of cyclic dinucleotide turnover proteins. Journal of bacteriology 199:e00790-16.

52. Schmidt AJ, Ryjenkov DA, Gomelsky M. 2005. The ubiquitous protein domain EAL is a cyclic diguanylate-specific phosphodiesterase: enzymatically active and inactive EAL domains. Journal of bacteriology 187:4774-4781.

53. Bellini D, Caly DL, McCarthy Y, Bumann M, An SQ, Dow JM, Ryan RP, Walsh MA. 2014. Crystal structure of an HD-GYP domain cyclic-di-GMP phosphodiesterase reveals an enzyme with a novel trinuclear catalytic iron centre. Molecular microbiology 91:26-38.

54. Rao F, Qi Y, Chong HS, Kotaka M, Li B, Li J, Lescar J, Tang K, Liang Z-X. 2009. The functional role of a conserved loop in EAL domain-based cyclic di-GMP-specific phosphodiesterase. Journal of bacteriology 191:4722-4731.

55. Rao F, Yang Y, Qi Y, Liang Z-X. 2008. Catalytic mechanism of cyclic di-GMP-specific phosphodiesterase: a study of the EAL domain-containing RocR from *Pseudomonas aeruginosa*. Journal of bacteriology 190:3622-3631.

56. Römling U. 2009. Rationalizing the evolution of EAL domain-based cyclic di-GMP-specific phosphodiesterases. Journal of bacteriology 191:4697-4700.

57. Wigren E, Liang ZX, Römling U. 2014. Finally! The structural secrets of a HD-GYP phosphodiesterase revealed. Molecular microbiology 91:1-5.

58. Cheang QW, Xin L, Chea RYF, Liang Z-X. 2019. Emerging paradigms for PilZ domain-mediated C-di-GMP signaling. Biochemical Society Transactions 47:381-388.

59. He Q, Wang F, Liu S, Zhu D, Cong H, Gao F, Li B, Wang H, Lin Z, Liao J. 2016. Structural and biochemical insight into the mechanism of Rv2837c from *Mycobacterium tuberculosis* as a c-di-NMP phosphodiesterase. Journal of Biological Chemistry 291:3668-3681.

60. Gu L, He Q. 2020. A Unified Catalytic Mechanism for Cyclic di-NMP Hydrolysis by DHH–DHHA1 Phosphodiesterases, p 79-92, Microbial Cyclic Di-Nucleotide Signaling. Springer.

61. Römling U, Galperin MY, Gomelsky M. 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiology and Molecular Biology Reviews 77:1-52.

62. Simm R, Ahmad I, Rhen M, Le Guyon S, Römling U. 2014. Regulation of biofilm formation in *Salmonella enterica* serovar Typhimurium. Future microbiology 9:1261-1282.

63. Hengge R, Galperin MY, Ghigo J-M, Gomelsky M, Green J, Hughes KT, Jenal U, Landini P. 2016. Systematic nomenclature for GGDEF and EAL domain-containing cyclic di-GMP turnover proteins of *Escherichia coli.* Journal of bacteriology 198:7-11.

64. Tarutina M, Ryjenkov DA, Gomelsky M. 2006. An unorthodox bacteriophytochrome from *Rhodobacter sphaeroides* involved in turnover of the second messenger c-di-GMP. Journal of Biological Chemistry 281:34751-34758.

65. Österberg S, Åberg A, Herrera Seitz MK, Wolf-Watz M, Shingler V. 2013. Genetic dissection of a motility-associated c-di-GMP signalling protein of *Pseudomonas putida*. Environmental microbiology reports 5:556-565.

66. Ahmad I, Cimdins A, Beske T, Römling U. 2017. Detailed analysis of c-di-GMP mediated regulation of *csgD* expression in *Salmonella* typhimurium. BMC microbiology 17:27.

67. Lindenberg S, Klauck G, Pesavento C, Klauck E, Hengge R. 2013. The EAL domain protein YciR acts as a trigger enzyme in ac-di-GMP signalling cascade in *E. coli* biofilm control. The EMBO journal 32:2001-2014.

68. Deng Y, Schmid N, Wang C, Wang J, Pessi G, Wu D, Lee J, Aguilar C, Ahrens CH, Chang C. 2012. Cis-2-dodecenoic acid receptor RpfR links quorum-sensing signal perception with regulation of virulence through cyclic dimeric guanosine monophosphate turnover. Proceedings of the National Academy of Sciences 109:15479-15484.

69. Hickman JW, Harwood CS. 2008. Identification of FleQ from *Pseudomonas aeruginosa* as ac-di-GMP-responsive transcription factor. Molecular microbiology 69:376-389.

70. Burdette DL, Monroe KM, Sotelo-Troha K, Iwig JS, Eckert B, Hyodo M, Hayakawa Y, Vance RE. 2011. STING is a direct innate immune sensor of cyclic di-GMP. Nature 478:515-518.

71. Ryjenkov DA, Simm R, Römling U, Gomelsky M. 2006. The PilZ domain is a receptor for the second messenger c-di-GMP THE PilZ DOMAIN PROTEIN YcgR CONTROLS MOTILITY IN *ENTEROBACTERIA*. Journal of Biological Chemistry 281:30310-30314.

72. Amikam D, Galperin MY. 2006. PilZ domain is part of the bacterial c-di-GMP binding protein. Bioinformatics 22:3-6.

73. Benach J, Swaminathan SS, Tamayo R, Handelman SK, Folta-Stogniew E, Ramos JE, Forouhar F, Neely H, Seetharaman J, Camilli A, Hunt JF. 2007. The structural basis of cyclic diguanylate signal transduction by PilZ domains. EMBO J 26:5153-66.

74. Christen M, Christen B, Allan MG, Folcher M, Jenö P, Grzesiek S, Jenal U. 2007. DgrA is a member of a new family of cyclic diguanosine monophosphate receptors and controls flagellar motor function in *Caulobacter crescentus*. Proceedings of the National Academy of Sciences 104:4112-4117.

75. Whitney JC, Whitfield GB, Marmont LS, Yip P, Neculai AM, Lobsanov YD, Robinson H, Ohman DE, Howell PL. 2015. Dimeric c-di-GMP is required for post-translational regulation of alginate production in *Pseudomonas aeruginosa*. Journal of Biological Chemistry 290:12451-12462.

76. Pratt JT, Tamayo R, Tischler AD, Camilli A. 2007. PilZ domain proteins bind cyclic diguanylate and regulate diverse processes in *Vibrio cholerae*. Journal of Biological Chemistry 282:12860-12870.

77. Christen M, Kulasekara HD, Christen B, Kulasekara BR, Hoffman LR, Miller SI. 2010. Asymmetrical distribution of the second messenger c-di-GMP upon bacterial cell division. Science 328:1295-1297.

78. Morgan JL, Strumillo J, Zimmer J. 2013. Crystallographic snapshot of cellulose synthesis and membrane translocation. Nature 493:181.

79. Whitney JC, Colvin KM, Marmont LS, Robinson H, Parsek MR, Howell PL. 2012. Structure of the cytoplasmic region of PelD, a degenerate diguanylate cyclase receptor that regulates exopolysaccharide production in *Pseudomonas aeruginosa*. Journal of Biological Chemistry 287:23582-23593.

80. Petters T, Zhang X, Nesper J, Treuner-Lange A, Gomez-Santos N, Hoppert M, Jenal U, Søgaard-Andersen L. 2012. The orphan histidine protein kinase SgmT is ac-di-GMP receptor and regulates composition of the extracellular matrix together with the orphan DNA binding response regulator DigR in *Myxococcus xanthus*. Molecular microbiology 84:147-165.

81. Kazmierczak BI, Lebron MB, Murray TS. 2006. Analysis of FimX, a phosphodiesterase that governs twitching motility in *Pseudomonas aeruginosa*. Molecular microbiology 60:1026-1043.

82. Wang Y-C, Chin K-H, Tu Z-L, He J, Jones CJ, Sanchez DZ, Yildiz FH, Galperin MY, Chou S-H. 2016. Nucleotide binding by the widespread high-affinity cyclic di-GMP receptor MshEN domain. Nature communications 7:1-12.

83. Jones CJ, Utada A, Davis KR, Thongsomboon W, Sanchez DZ, Banakar V, Cegelski L, Wong GC, Yildiz FH. 2015. C-di-GMP regulates motile to sessile transition by modulating MshA pili biogenesis and near-surface motility behavior in *Vibrio cholerae*. PLoS pathogens 11.

84. Roelofs KG, Jones CJ, Helman SR, Shang X, Orr MW, Goodson JR, Galperin MY, Yildiz FH, Lee VT. 2015. Systematic identification of cyclic-di-GMP binding proteins in *Vibrio cholerae* reveals a novel class of cyclic-di-GMP-binding ATPases associated with type II secretion systems. PLoS pathogens 11.

85. Chin K-H, Lee Y-C, Tu Z-L, Chen C-H, Tseng Y-H, Yang J-M, Ryan RP, McCarthy Y, Dow JM, Wang AH-J. 2010. The cAMP receptor-like protein CLP is a novel c-di-GMP receptor linking cell–cell signaling to virulence gene expression in *Xanthomonas campestris.* Journal of molecular biology 396:646-662.

86. Carzaniga T, Antoniani D, Dehò G, Briani F, Landini P. 2012. The RNA processing enzyme polynucleotide phosphorylase negatively controls biofilm formation by repressing poly-N-acetylglucosamine (PNAG) production in *Escherichia coli* C. BMC microbiology 12:270.

87. Sudarsan N, Lee E, Weinberg Z, Moy R, Kim J, Link K, Breaker R. 2008. Riboswitches in eubacteria sense the second messenger cyclic di-GMP. Science 321:411-413.

88. Tuckerman JR, Gonzalez G, Gilles-Gonzalez M-A. 2011. Cyclic di-GMP activation of polynucleotide phosphorylase signal-dependent RNA processing. Journal of molecular biology 407:633-639.

89. Lee ER, Baker JL, Weinberg Z, Sudarsan N, Breaker RR. 2010. An allosteric self-splicing ribozyme triggered by a bacterial second messenger. Science 329:845-848.

90. Witte G, Hartung S, Büttner K, Hopfner K-P. 2008. Structural biochemistry of a bacterial checkpoint protein reveals diadenylate cyclase activity regulated by DNA recombination intermediates. Molecular cell 30:167-178.

91. Römling U. 2008. Great times for small molecules: c-di-AMP, a second messenger candidate in Bacteria and Archaea. Sci Signal 1:pe39-pe39.

92. He J, Yin W, Galperin MY, Chou S-H. 2020. Cyclic di-AMP, a second messenger of primary importance: tertiary structures and binding mechanisms. Nucleic Acids Research.

93. Corrigan RM, Abbott JC, Burhenne H, Kaever V, Gründling A. 2011. c-di-AMP is a new second messenger in *Staphylococcus aureus* with a role in controlling cell size and envelope stress. PLoS pathogens 7:e1002217.

94. Woodward JJ, Iavarone AT, Portnoy DA. 2010. c-di-AMP secreted by intracellular *Listeria monocytogenes* activates a host type I interferon response. Science 328:1703-1705.

95. Bai Y, Yang J, Zhou X, Ding X, Eisele LE, Bai G. 2012. *Mycobacterium tuberculosis* Rv3586 (DacA) is a diadenylate cyclase that converts ATP or ADP into c-di-AMP. PLoS One 7:e35206.

96. Kamegaya T, Kuroda K, Hayakawa Y. 2011. Identification of a *Streptococcus pyogenes* SF370 gene involved in production of c-di-AMP.

97. Barker JR, Koestler BJ, Carpenter VK, Burdette DL, Waters CM, Vance RE, Valdivia RH. 2013. STING-dependent recognition of cyclic di-AMP mediates type I interferon responses during *Chlamydia trachomatis* infection. MBio 4:e00018-13.

98. Gomelsky M. 2011. cAMP, c-di-GMP, c-di-AMP and now cGMP: bacteria use them all! Molecular microbiology 79:562-565.

99. Commichau FM, Dickmanns A, Gundlach J, Ficner R, Stülke J. 2015. A jack of all trades: the multiple roles of the unique essential second messenger cyclic di-AMP. Molecular microbiology.

100. Gao A, Serganov A. 2014. Structural insights into recognition of c-di-AMP by the *ydaO* riboswitch. Nature chemical biology 10:787.

101. Mehne FM, Gunka K, Eilers H, Herzberg C, Kaever V, Stülke J. 2013. Cyclic di-AMP homeostasis in *Bacillus subtilis* both lack and high level accumulation of the nucleotide are detrimental for cell growth. Journal of Biological Chemistry 288:2004-2017.

102. Xayarath B, Freitag NE. 2015. Uncovering the nonessential nature of an essential second messenger. Cell host & microbe 17:731-732.

103. Müller M, Deimling T, Hopfner K-P, Witte G. 2015. Structural analysis of the diadenylate cyclase reaction of DNA-integrity scanning protein A (DisA) and its inhibition by 3′-dATP. Biochemical Journal 469:367-374.

104. Braun F, Thomalla L, van der Does C, Quax T, Allers T, Kaever V, Albers S. 2019. Cyclic nucleotides in archaea: cyclic di-AMP in the archaeon *Haloferax volcanii* and its putative role. Microbiologyopen 8 (9): e00829.

105. Mehne FM, Schröder-Tittmann K, Eijlander RT, Herzberg C, Hewitt L, Kaever V, Lewis RJ, Kuipers OP, Tittmann K, Stülke J. 2014. Control of the Diadenylate Cyclase CdaS in *Bacillus subtilis* AN AUTOINHIBITORY DOMAIN LIMITS CYCLIC DI-AMP PRODUCTION. Journal of Biological Chemistry 289:21098-21107.

106. Bai Y, Yang J, Eisele LE, Underwood AJ, Koestler BJ, Waters CM, Metzger DW, Bai G. 2013. Two DHH subfamily 1 proteins in *Streptococcus pneumoniae* possess cyclic di-AMP phosphodiesterase activity and affect bacterial growth and virulence. Journal of bacteriology 195:5123-5132.

107. Rosenberg J, Dickmanns A, Neumann P, Gunka K, Arens J, Kaever V, Stülke J, Ficner R, Commichau FM. 2015. Structural and biochemical analysis of the essential diadenylate cyclase CdaA from *Listeria monocytogenes*. Journal of Biological Chemistry 290:6596-6606.

108. Huynh TN, Woodward JJ. 2016. Too much of a good thing: regulated depletion of c-di-AMP in the bacterial cytoplasm. Current opinion in microbiology 30:22-29.

109. Rao F, See RY, Zhang D, Toh DC, Ji Q, Liang Z-X. 2010. YybT is a signaling protein that contains a cyclic dinucleotide phosphodiesterase domain and a GGDEF domain with ATPase activity. Journal of Biological Chemistry 285:473-482.

110. Ye M, Zhang J-J, Fang X, Lawlis GB, Troxell B, Zhou Y, Gomelsky M, Lou Y, Yang XF. 2014. DhhP, a cyclic di-AMP phosphodiesterase of *Borrelia burgdorferi,* is essential for cell growth and virulence. Infection and immunity 82:1840-1849.

111. Manikandan K, Sabareesh V, Singh N, Saigal K, Mechold U, Sinha KM. 2014. Two-step synthesis and hydrolysis of cyclic di-AMP in *Mycobacterium tuberculosis*. PLoS One 9:e86096.

112. Yang J, Bai Y, Zhang Y, Gabrielle VD, Jin L, Bai G. 2014. Deletion of the cyclic di-AMP phosphodiesterase gene (cnpB) in *Mycobacterium tuberculosis* leads to reduced virulence in a mouse model of infection. Molecular microbiology 93:65-79.

113. Gundlach J, Mehne FM, Herzberg C, Kampf J, Valerius O, Kaever V, Stülke J. 2015. An essential poison: synthesis and degradation of cyclic di-AMP in *Bacillus subtilis*. Journal of bacteriology 197:3265-3274.

114. Huynh TN, Luo S, Pensinger D, Sauer J-D, Tong L, Woodward JJ. 2015. An HD-domain phosphodiesterase mediates cooperative hydrolysis of c-di-AMP to affect bacterial growth and virulence. Proceedings of the National Academy of Sciences 112:E747-E756.

115. Nelson JW, Sudarsan N, Furukawa K, Weinberg Z, Wang JX, Breaker RR. 2013. Riboswitches in eubacteria sense the second messenger c-di-AMP. Nature chemical biology 9:834.

116. Sureka K, Choi PH, Precit M, Delince M, Pensinger DA, Huynh TN, Jurado AR, Goo YA, Sadilek M, Iavarone AT. 2014. The cyclic dinucleotide c-di-AMP is an allosteric regulator of metabolic enzyme function. Cell 158:1389-1401.

117. Zhang L, Li W, He Z-G. 2013. DarR, a TetR-like transcriptional factor, is a cyclic di-AMP-responsive repressor in *Mycobacterium smegmatis*. Journal of Biological Chemistry 288:3085-3096.

118. Corrigan RM, Campeotto I, Jeganathan T, Roelofs KG, Lee VT, Gründling A. 2013. Systematic identification of conserved bacterial c-di-AMP receptor proteins. Proceedings of the National Academy of Sciences 110:9084-9089.

119. Bai Y, Yang J, Zarrella TM, Zhang Y, Metzger DW, Bai G. 2014. Cyclic di-AMP impairs potassium uptake mediated by a cyclic di-AMP binding protein in *Streptococcus pneumoniae*. Journal of bacteriology 196:614-623.

120. Gundlach J, Dickmanns A, Schröder-Tittmann K, Neumann P, Kaesler J, Kampf J, Herzberg C, Hammer E, Schwede F, Kaever V. 2015. Identification, characterization, and structure analysis of the cyclic di-AMP-binding PII-like signal transduction protein DarA. Journal of Biological Chemistry 290:3069-3080.

121. Campeotto I, Zhang Y, Mladenov MG, Freemont PS, Gründling A. 2015. Complex structure and biochemical characterization of the *Staphylococcus aureus* cyclic diadenylate monophosphate (c-di-AMP)-binding protein PstA, the founding member of a new signal transduction protein family. Journal of Biological Chemistry 290:2888-2901.

122. Watson PY, Fedor MJ. 2012. The *ydaO* motif is an ATP-sensing riboswitch in *Bacillus subtilis*. Nature chemical biology 8:963-965.

123. Burroughs AM, Zhang D, Schäffer DE, Iyer LM, Aravind L. 2015. Comparative genomic analyses reveal a vast, novel network of nucleotide-centric systems in biological conflicts, immunity and signaling. Nucleic acids research 43:10633-10654.

124. Cohen D, Melamed S, Millman A, Shulman G, Oppenheimer-Shaanan Y, Kacen A, Doron S, Amitai G, Sorek R. 2019. Cyclic GMP–AMP signalling protects bacteria against viral infection. Nature 574:691-695.

125. Whiteley AT, Eaglesham JB, de Oliveira Mann CC, Morehouse BR, Lowey B, Nieminen EA, Danilchanka O, King DS, Lee AS, Mekalanos JJ. 2019. Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. Nature 567:194-199.

126. Schubert S, Dufke S, Sorsa J, Heesemann J. 2004. A novel integrative and conjugative element (ICE) of *Escherichia coli*: the putative progenitor of the *Yersinia* high-pathogenicity island. Molecular microbiology 51:837-848.

127. Cimdins A, Lüthje P, Li F, Ahmad I, Brauner A, Römling U. 2017. Draft genome sequences of semiconstitutive red, dry, and rough biofilm-forming commensal and uropathogenic *Escherichia coli* isolates. Genome Announc 5:e01249-16.

128. Ochman H, Selander RK. 1984. Standard reference strains of *Escherichia coli* from natural populations. Journal of bacteriology 157:690-693.

129. Schubert S, Picard B, Gouriou S, Heesemann J, Denamur E. 2002. *Yersinia* high-pathogenicity island contributes to virulence in *Escherichia coli* causing extraintestinal infections. Infection and immunity 70:5335-5337.

130. Li F, Cimdins A, Rohde M, Jänsch L, Kaever V, Nimtz M, Römling U. 2019. DncV Synthesizes Cyclic GMP-AMP and Regulates Biofilm Formation and Motility in *Escherichia coli* ECOR31. mBio 10:e02492-18.

131. Severin GB, Ramliden MS, Hawver LA, Wang K, Pell ME, Kieninger A-K, Khataokar A, O'Hara BJ, Behrmann LV, Neiditch MB. 2018. Direct activation of a phospholipase by cyclic GMP-AMP in El Tor *Vibrio cholerae*. Proceedings of the National Academy of Sciences 115:E6048-E6055.

132. Zhu D, Wang L, Shang G, Liu X, Zhu J, Lu D, Wang L, Kan B, Zhang J-r, Xiang Y. 2014. Structural biochemistry of a *Vibrio cholerae* dinucleotide cyclase reveals cyclase activity regulation by folates. Molecular cell 55:931-937.

133. Hallberg ZF, Wang XC, Wright TA, Nan B, Ad O, Yeo J, Hammond MC. 2016. Hybrid promiscuous (Hypr) GGDEF enzymes produce cyclic AMP-GMP (3′, 3′-cGAMP). Proceedings of the National Academy of Sciences 113:1790-1795.

134. Nelson JW, Sudarsan N, Phillips GE, Stav S, Lünse CE, McCown PJ, Breaker RR. 2015. Control of bacterial exoelectrogenesis by c-AMP-GMP. Proceedings of the National Academy of Sciences 112:5389-5394.

135. Hallberg ZF, Chan CH, Wright TA, Kranzusch PJ, Doxzen KW, Park JJ, Bond DR, Hammond MC. 2019. Structure and mechanism of a Hypr GGDEF enzyme that activates cGAMP signaling to control extracellular metal respiration. Elife 8:e43959.

136. McFarland AP, Luo S, Ahmed-Qadri F, Zuck M, Thayer EF, Goo YA, Hybiske K, Tong L, Woodward JJ. 2017. Sensing of bacterial cyclic dinucleotides by the oxidoreductase RECON promotes NF-κB activation and shapes a proinflammatory antibacterial state. Immunity 46:433-445.

137. Zhang X, Shi H, Wu J, Zhang X, Sun L, Chen C, Chen ZJ. 2013. Cyclic GMP-AMP containing mixed phosphodiester linkages is an endogenous high-affinity ligand for STING. Molecular cell 51:226-235.

138. Kellenberger CA, Wilson SC, Hickey SF, Gonzalez TL, Su Y, Hallberg ZF, Brewer TF, Iavarone AT, Carlson HK, Hsieh Y-F. 2015. GEMM-I riboswitches from *Geobacter* sense the bacterial second messenger cyclic AMP-GMP. Proceedings of the National Academy of Sciences 112:5383-5388.

139. Diner EJ, Burdette DL, Wilson SC, Monroe KM, Kellenberger CA, Hyodo M, Hayakawa Y, Hammond MC, Vance RE. 2013. The innate immune DNA sensor cGAS produces a noncanonical cyclic dinucleotide that activates human STING. Cell reports 3:1355-1361.

140. Ablasser A, Goldeck M, Cavlar T, Deimling T, Witte G, Röhl I, Hopfner K-P, Ludwig J, Hornung V. 2013. cGAS produces a 2′-5′-linked cyclic dinucleotide second messenger that activates STING. Nature 498:380-384.

141. Kato K, Ishii R, Hirano S, Ishitani R, Nureki O. 2015. Structural basis for the catalytic mechanism of DncV, bacterial homolog of cyclic GMP-AMP synthase. Structure 23:843-850.

142. Wu J, Sun L, Chen X, Du F, Shi H, Chen C, Chen ZJ. 2013. Cyclic GMP-AMP is an endogenous second messenger in innate immune signaling by cytosolic DNA. Science 339:826-830.

143. Sun L, Wu J, Du F, Chen X, Chen ZJ. 2013. Cyclic GMP-AMP synthase is a cytosolic DNA sensor that activates the type I interferon pathway. Science 339:786-791.

144. Kranzusch PJ, Lee AS, Wilson SC, Solovykh MS, Vance RE, Berger JM, Doudna JA. 2014. Structure-guided reprogramming of human cGAS dinucleotide linkage specificity. Cell 158:1011-1021.

145. Civril F, Deimling T, de Oliveira Mann CC, Ablasser A, Moldt M, Witte G, Hornung V, Hopfner K-P. 2013. Structural mechanism of cytosolic DNA sensing by cGAS. Nature 498:332.

146. Gao P, Ascano M, Wu Y, Barchet W, Gaffney BL, Zillinger T, Serganov AA, Liu Y, Jones RA, Hartmann G. 2013. Cyclic [G (2′, 5′) pA (3′, 5′) p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. Cell 153:1094-1107.

147. Li X, Shu C, Yi G, Chaton CT, Shelton CL, Diao J, Zuo X, Kao CC, Herr AB, Li P. 2013. Cyclic GMP-AMP synthase is activated by double-stranded DNA-induced oligomerization. Immunity 39:1019-1031.

148. Zhang X, Wu J, Du F, Xu H, Sun L, Chen Z, Brautigam CA, Zhang X, Chen ZJ. 2014. The cytosolic DNA sensor cGAS forms an oligomeric complex with DNA and undergoes switch-like conformational changes in the activation loop. Cell reports 6:421-430.

149. De N, Pirruccello M, Krasteva PV, Bae N, Raghavan RV, Sondermann H. 2008. Phosphorylation-independent regulation of the diguanylate cyclase WspR. PLoS biology 6.

150. Kranzusch PJ, Wilson SC, Lee AS, Berger JM, Doudna JA, Vance RE. 2015. Ancient origin of cGAS-STING reveals mechanism of universal 2′, 3′ cGAMP signaling. Molecular cell 59:891-903.

151. Gao J, Tao J, Liang W, Zhao M, Du X, Cui S, Duan H, Kan B, Su X, Jiang Z. 2015. Identification and characterization of phosphodiesterases that specifically degrade 3′ 3′-cyclic GMP-AMP. Cell research 25:539-550.

152. Lovering AL, Capeness MJ, Lambert C, Hobley L, Sockett RE. 2011. The structure of an unconventional HD-GYP protein from *Bdellovibrio* reveals the roles of conserved residues in this class of cyclic-di-GMP phosphodiesterases. MBio 2:e00163-11.

153. Rinaldo S, Paiardini A, Stelitano V, Brunotti P, Cervoni L, Fernicola S, Protano C, Vitali M, Cutruzzolà F, Giardina G. 2015. Structural basis of functional diversification of the HD-GYP domain revealed by the *Pseudomonas aeruginosa* PA4781 protein, which displays an unselective bimetallic binding site. Journal of bacteriology 197:1525-1535.

154. Deng M-j, Tao J, Chao E, Ye Z-y, Jiang Z, Yu J, Su X-d. 2018. Novel mechanism for cyclic dinucleotide degradation revealed by structural studies of *vibrio* phosphodiesterase V-cGAP3. Journal of molecular biology 430:5080-5093.

155. Li L, Yin Q, Kuss P, Maliga Z, Millán JL, Wu H, Mitchison TJ. 2014. Hydrolysis of 2′ 3′-cGAMP by ENPP1 and design of nonhydrolyzable analogs. Nature chemical biology 10:1043.

156. Gao J, Tao J, Liang W, Jiang Z. 2016. Cyclic (di) nucleotides: the common language shared by microbe and host. Current opinion in microbiology 30:79-87.

157. Kato K, Nishimasu H, Oikawa D, Hirano S, Hirano H, Kasuya G, Ishitani R, Tokunaga F, Nureki O. 2018. Structural insights into cGAMP degradation by Ecto-nucleotide pyrophosphatase phosphodiesterase 1. Nature communications 9:1-8.

158. Wright TA, Jiang L, Park JJ, Anderson WA, Chen G, Hallberg ZF, Nan B, Hammond MC. 2019. Second messengers and divergent HD-GYP phosphodiesterases regulate 3′, 3′-cGAMP signaling. Molecular microbiology.

159. Six DA, Dennis EA. 2000. The expanding superfamily of phospholipase A2 enzymes: classification and characterization. Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids 1488:1-19.

160. Scherer GF, Ryu SB, Wang X, Matos AR, Heitz T. 2010. Patatin-related phospholipase A: nomenclature, subfamilies and functions in plants. Trends in plant science 15:693-700.

161. Kudo I, Murakami M. 2002. Phospholipase A2 enzymes. Prostaglandins & other lipid mediators 68:3-58.

162. Rosahl S, Schmidt R, Schell J, Willmitzer L. 1986. Isolation and characterization of a gene from *Solanum tuberosum* encoding patatin, the major storage protein of potato tubers. Molecular and General Genetics MGG 203:214-220.

163. Andrews D, Beames B, Summers M, Park W. 1988. Characterization of the lipid acyl hydrolase activity of the major potato (Solanum tuberosum) tuber protein, patatin, by cloning and abundant expression in a baculovirus vector. Biochemical Journal 252:199-206.

164. Wilson SK, Knoll LJ. 2018. Patatin-like phospholipases in microbial infections with emerging roles in fatty acid metabolism and immune regulation by *Apicomplexa*. Molecular microbiology 107:34-46.

165. Sawa T, Hamaoka S, Kinoshita M, Kainuma A, Naito Y, Akiyama K, Kato H. 2016. *Pseudomonas aeruginosa* type III secretory toxin ExoU and its predicted homologs. Toxins 8:307.

166. Sato H, Frank DW, Hillard CJ, Feix JB, Pankhaniya RR, Moriyama K, Finck-Barbançon V, Buchaklian A, Lei M, Long RM. 2003. The mechanism of action of the *Pseudomonas aeruginosa*-encoded type III cytotoxin, ExoU. The EMBO journal 22:2959-2969.

167. Anderson DM, Schmalzer KM, Sato H, Casey M, Terhune SS, Haas AL, Feix JB, Frank DW. 2011. Ubiquitin and ubiquitin-modified proteins activate the *Pseudomonas aeruginosa* T3SS cytotoxin, ExoU. Molecular microbiology 82:1454-1467.

168. Anderson DM, Sato H, Dirck AT, Feix JB, Frank DW. 2015. Ubiquitin activates patatin-like phospholipases from multiple bacterial species. Journal of bacteriology 197:529-541.

169. Kellenberger CA, Wilson SC, Sales-Lee J, Hammond MC. 2013. RNA-based fluorescent biosensors for live cell imaging of second messengers cyclic di-GMP and cyclic AMP-GMP. Journal of the American Chemical Society 135:4906-4909.

170. Ren A, Wang XC, Kellenberger CA, Rajashankar KR, Jones RA, Hammond MC, Patel DJ. 2015. Structural basis for molecular discrimination by a 3′, 3′-cGAMP sensing riboswitch. Cell reports 11:1-12.

171. Niewoehner O, Garcia-Doval C, Rostøl JT, Berk C, Schwede F, Bigler L, Hall J, Marraffini LA, Jinek M. 2017. Type III CRISPR–Cas systems produce cyclic oligoadenylate second messengers. Nature 548:543-548.

172. Kazlauskiene M, Kostiuk G, Venclovas Č, Tamulaitis G, Siksnys V. 2017. A cyclic oligonucleotide signaling pathway in type III CRISPR-Cas systems. Science 357:605-609.

173. Datta S, Costantino N, Court DL. 2006. A set of recombineering plasmids for gram-negative bacteria. Gene 379:109-115.

174. Trunk T, Khalil HS, Leo JC. 2018. Bacterial autoaggregation. AIMS microbiology 4:140.

175. Flint S, Palmer J, Bremer P, Seale B, Brooks J, Lindsay D, Burgess S. 2011. Biofilm formation.

176. Petrachi T, Resca E, Piccinno MS, Biagi F, Strusi V, Dominici M, Veronesi E. 2017. An alternative approach to investigate biofilm in medical devices: A feasibility study. International journal of environmental research and public health 14:1587.

177. Ohi M, Li Y, Cheng Y, Walz T. 2004. Negative staining and image classification-powerful tools in modern electron microscopy. Biological procedures online 6:23-34.

178. Freeman WM, Walker SJ, Vrana KE. 1999. Quantitative RT-PCR: pitfalls and potential. Biotechniques 26:112-125.

179. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. methods 25:402-408.

180. Mahmood T, Yang P-C. 2012. Western blot: technique, theory, and trouble shooting. North American journal of medical sciences 4:429.

181. Scopes RK. 2013. Protein purification: principles and practice. Springer Science & Business Media.

182. Santiago M, Strobel S. 2013. Thin layer chromatography, p 303-324, Methods in enzymology, vol 533. Elsevier.

183. McLafferty FW. 1981. Tandem mass spectrometry. Science 214:280-287.

184. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology 19:455-477.

185. Lander ES, Waterman MS. 1988. Genomic mapping by fingerprinting random clones: a mathematical analysis. Genomics 2:231-239.

186. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M. 2008. The RAST Server: rapid annotations using subsystems technology. BMC genomics 9:75.

187. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic acids research 42:D206-D214.

188. Scott HM, Costerton J. 1995. Microbial biofilms. Annu Rev Microbiol 49:711745Cunningham.

189. Hammar Mr, Arnqvist A, Bian Z, Olsén A, Normark S. 1995. Expression of two *csg* operons is required for production of fibronectin-and congo red-binding curli polymers in *Escherichia coli* K-12. Molecular microbiology 18:661-670.

190. Bian Z, Brauner A, Li Y, Normark S. 2000. Expression of and Cytokine Activation by *Eschevichia coli* Curi Fibers in Human Sepsis. The Journal of infectious diseases 181:602-612.

191. Bokranz W, Wang X, Tschäpe H, Römling U. 2005. Expression of cellulose and curli fimbriae by *Escherichia coli* isolated from the gastrointestinal tract. Journal of Medical Microbiology 54:1171-1182.

192. Kai-Larsen Y, Lüthje P, Chromek M, Peters V, Wang X, Holm Å, Kádas L, Hedlund K-O, Johansson J, Chapman MR. 2010. Uropathogenic *Escherichia coli* modulates immune responses and its curli fimbriae interact with the antimicrobial peptide LL-37. PLoS pathogens 6.

193. Gish W, States DJ. 1993. Identification of protein coding regions by database similarity search. Nature genetics 3:266-272.

194. El Mouali Y, Kim H, Ahmad I, Brauner A, Liu Y, Skurnik M, Galperin MY, Römling U. 2017. Stand-alone EAL domain proteins form a distinct subclass of EAL proteins involved in regulation of cell motility and biofilm formation in enterobacteria. Journal of bacteriology 199:e00179-17.

195. Cimdins A, Simm R, Li F, Lüthje P, Thorell K, Sjöling Å, Brauner A, Römling U. 2017. Alterations of c-di-GMP turnover proteins modulate semi-constitutive rdar biofilm formation in commensal and uropathogenic *Escherichia coli*. MicrobiologyOpen 6:e00508.

196. Terashima H, Kojima S, Homma M. 2008. Flagellar motility in bacteria: structure and function of flagellar motor. International review of cell and molecular biology 270:39-85.

197. El-Hajj ZW, Newman EB. 2015. How much territory can a single *E. coli* cell control? Frontiers in microbiology 6:309.

198. Cordell SC, Robinson EJ, Löwe J. 2003. Crystal structure of the SOS cell division inhibitor SulA and in complex with FtsZ. Proceedings of the National Academy of Sciences 100:7889-7894.

199. Bi E, Lutkenhaus J. 1993. Cell division inhibitors SulA and MinCD prevent formation of the FtsZ ring. Journal of bacteriology 175:1118-1125.

200. Frain KM, Dijl JMV, Robinson C. 2019. The Twin-Arginine Pathway for Protein Secretion. Protein Secretion in Bacteria:53-66.

201. Lin B-L, Chen C-Y, Huang C-H, Ko T-P, Chiang C-H, Lin K-F, Chang Y-C, Lin P-Y, Tsai H-HG, Wang AH-J. 2017. The Arginine Pairs and C-Termini of the Sso7c4 from *Sulfolobus solfataricus* Participate in Binding and Bending DNA. PloS one 12.

202. Galperin MY, Chou S-H. 2020. Structural conservation and diversity of PilZ-related domains. Journal of Bacteriology 202.

203. Hufnagel DA, Evans ML, Greene SE, Pinkner JS, Hultgren SJ, Chapman MR. 2016. The catabolite repressor protein-cyclic AMP complex regulates csgD and biofilm formation in uropathogenic *Escherichia coli*. Journal of bacteriology 198:3329-3334.

204. Monteiro C, Saxena I, Wang X, Kader A, Bokranz W, Simm R, Nobles D, Chromek M, Brauner A, Brown Jr RM. 2009. Characterization of cellulose production in *Escherichia coli* Nissle 1917 and its biological consequences. Environmental microbiology 11:1105-1116.

205. Lasaro MA, Salinger N, Zhang J, Wang Y, Zhong Z, Goulian M, Zhu J. 2009. F1C fimbriae play an important role in biofilm formation and intestinal colonization by the *Escherichia coli* commensal strain Nissle 1917. Appl Environ Microbiol 75:246-251.

206. Clermont O, Christenson JK, Denamur E, Gordon DM. 2013. The Clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environmental microbiology reports 5:58-65.

207. Itoh Y, Rice JD, Goller C, Pannuri A, Taylor J, Meisner J, Beveridge TJ, Preston JF, Romeo T. 2008. Roles of *pgaABCD* genes in synthesis, modification, and export of the *Escherichia coli* biofilm adhesin poly-β-1, 6-N-acetyl-d-glucosamine. Journal of bacteriology 190:3670-3680.

208. Römling U, Bian Z, Hammar M, Sierralta WD, Normark S. 1998. Curli fibers are highly conserved between *Salmonella* typhimurium and *Escherichia coli* with respect to operon structure and regulation. Journal of bacteriology 180:722-731.

209. Ahmad I, Rouf SF, Sun L, Cimdins A, Shafeeq S, Le Guyon S, Schottkowski M, Rhen M, Römling U. 2016. BcsZ inhibits biofilm phenotypes and promotes virulence by blocking cellulose production in *Salmonella enterica* serovar Typhimurium. Microbial cell factories 15:177.

210. Da Re S, Ghigo J-M. 2006. A CsgD-independent pathway for cellulose production and biofilm formation in *Escherichia coli*. Journal of bacteriology 188:3073-3087.

211. Hengge R. 2016. Trigger phosphodiesterases as a novel class of c-di-GMP effector proteins. Philosophical Transactions of the Royal Society B: Biological Sciences 371:20150498.