Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____   _____

Benjamin M. Fontaine   Date

Chemical Approaches to Elucidate Novel Facets of Nucleic Acid Metabolism and Signaling

By

Benjamin M. Fontaine
Doctor of Philosophy

Chemistry

_____
Emily E. Weinert, Ph.D.
Advisor

_____
Stefan Lutz, Ph.D.
Committee Member

_____
David G. Lynn, Ph.D.
Committee Member

Accepted:

_____
Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____
Date

Chemical Approaches to Elucidate Novel Facets of Nucleic Acid Metabolism and Signaling

By

Benjamin M. Fontaine
B.A. Wake Forest University, 2012

Advisor: Emily E. Weinert, Ph.D.

An abstract of
A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2018

Abstract

Chemical Approaches to Elucidate Novel Facets of Nucleic Acid Metabolism and Signaling

By Benjamin M. Fontaine

Regulation of nucleotide and nucleoside concentrations is critical for faithful DNA replication, transcription, and translation in all organisms. Additionally, certain nucleotide signalling molecules, such as adenosine 3',5'-cyclic monophosphate (3',5'-cAMP) and bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP), regulate organismal adaptations to environmental stimuli. Recently, functions have emerged for certain non-canonical nucleotides, including cytidine 3',5'-cyclic monophosphate (3',5'-cCMP) and the regioisomeric nucleoside 2',3'-cyclic monophosphates (2',3'-cNMPs), but the underlying mechanisms remain largely elusive. Prior studies have suggested the existence of unique enzymes involved in 3',5'-cCMP metabolism and binding; this work seeks to identify these potential enzymes through the development of affinity-based purification procedures and activity assays developed herein. In addition, LC-MS/MS methodology has been optimized to quantify 3',5'- and 2',3'-cNMP levels in various biological samples, resulting in the first quantification of 2',3'-cNMP pools in *Escherichia coli*. These studies have revealed that 2',3'-cNMP levels in *E. coli* are generated specifically from RNase I-catalyzed RNA degradation, presumably as part of a previously unidentified nucleotide salvage pathway. Furthermore, the present work reports the dissection of cellular processes modulated by RNase I and 2',3'-cNMPs through the development of cell-permeable 2',3'-cNMP analogs and a 2',3'-cyclic nucleotide phosphodiesterase. Utilization of these (bio)chemical tools, in conjunction with phenotypic and transcriptomic investigations, has identified different pathways regulated by 2',3'-cNMPs and RNase I, including chemotaxis, biofilm production, antibiotic tolerance, and nucleotide pool homeostasis. Along with chemical approaches to probe cryptic nucleotide-mediated signalling, novel chemical modulators of biofilm formation have been synthesized, providing additional tools to investigate microbial transduction. Through the development of these methods, this dissertation has elucidated enigmatic features of signal transduction, with a focus on understudied aspects of nucleotide metabolism.

Chemical Approaches to Elucidate Novel Facets of Nucleic Acid Metabolism and Signaling

By

Benjamin M. Fontaine
B.A. Wake Forest University, 2012

Advisor: Emily E. Weinert, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Chemistry
2018

Acknowledgements

Words cannot express my gratititude to my parents and grandfather for providing me with access to education and for giving me the opportunity to pursue science. None of this would have been possible without their immense support. In particular, I have to thank my mom for holding our family together through some very rough times, and for giving me every possible opportunity despite the hardships. I also am grateful to my brothers, Sam and Alec; we've had great times hanging out in Asheville (particularly in High Vista and at UNCA), Wild Dunes, Santa Barbara, and Wilmington. In addition, the time spent with my closest friends, Brad Blair and So Banerjee, will always remain a defining feature of my life. The nostalgia of these memories really gets to me, but I'm looking forward to the future.

Regarding my time at Emory, I am tremendously grateful to Emily Weinert for her exemplary mentorship. Her optimism, patience, and constructive guidance have been indispensable in conquering the challenges of this project; I cannot express how much I will miss working in her group. I greatly appreciate all of her support, both on an academic and a personal level, and I cannot imagine a more rewarding work environment. In this regard, I also must acknowledge Xin Jia for helping to cultivate the welcoming and productive atmosphere from the lab's inception, and for teaching me when I knew essentially nothing. Furthermore, the guidance I have received from Justin Burns and Johnnie Walker in the disciplines of microbiology and molecular biology has been invaluable in progressing these projects. I also am indebted to Kevin Martin, Parth Jariwala, Laura Briggs, Jennifer García-Rodriguez, and Matt Jenkins for their devoted efforts in overcoming the early obstacles associated with this work. In addition, I have been extremely fortunate to work with Shannon Rivera, Dayna Patterson, Yasha Duggal, Paul Young, and Adam Young. I have enjoyed working with all of you immensely, and the future of the lab is in good hands. I also am grateful to Stefan Lutz and Dave Lynn for helpful comments and constructive criticism pertaining to both this work and my independent proposal. Additionally, I appreciate Fred Strobel's guidance with LC-MS experiments, and I am indebted to Shaoxiong Wu and Bing Wang for sharing their expertise in NMR spectroscopy. Finally, I must thank my past colleagues from Uli Bierbach's lab at Wake Forest, namely Amanda Pickard, Jimmy Suryadi, and Leigh Ann Graham, who were pivotal in my early scientific training.

Contents

Chapter 4: Work toward elucidating the roles of cytidine 3',5'-cyclic monophosphate in mammals

Chapter 5: Development of phyto-inspired phenolic glycosides as inhibitors of
***Staphylococcus aureus* biofilm formation**

List of Figures and Tables

Chapter 1: Introduction to nucleotide signaling and metabolism

Chapter 2: RNase I regulates *Escherichia coli* nucleoside 2',3'-cyclic monophosphate levels and biofilm formation

Chapter 3: Chemical tools reveal diverse functions modulated by RNase I and nucleoside 2',3'-cyclic monophosphate pools in *Escherichia coli*

Chapter 4: Work toward elucidating the roles of cytidine 3',5'-cyclic monophosphate in mammals

Chapter 5: Development of phyto-inspired phenolic glycosides as inhibitors of *Staphylococcus aureus* biofilm formation

Chapter 6: Conclusions, ongoing work, and future directions

List of abbreviations

adenosine 3',5'-cyclic monophosphate (3',5'-cAMP)

G-protein-coupled receptor (GPCR)

guanosine triphosphate (GTP)

guanosine diphosphate (GDP)

adenylate cyclase (AC)

phosphodiesterase (PDE)

guanosine 3',5'-cyclic monophosphate (3',5'-cGMP)

hyperpolarization-activated cyclic nucleotide-gated ion channel (HCN)

 exchange protein directly activated by 3',5'-cAMP (EPAC)

3',5'-cAMP receptor protein (Crp)

glucose transport protein IIA (IIA^{glc})

guanosine 3'-diphosphate, 5'-(tri)diphosphate; ([p]ppGpp)

bis-(3'-5')-cyclic dimeric guanosine monophosphate (c-di-GMP)

diguanylate cyclase (DGC)

small regulatory RNA (sRNA)

glucose-6-phosphate (G6P)

quorum sensing (QS)

acyl-homoserine lactone (AHL)

4,5-dihydroxy-2,3-pentanedione; autoinducer-2 (AI-2)

S-adenosylmethionine (SAM)

nucleoside monophosphate (NMP)

5-phospho-α-D-ribose 1-diphosphate (PRPP)

inorganic diphosphate (PP$_i$)

2'-deoxy nucleoside 5'-triphosphate (dNTP)

methyl-accepting chemotaxis protein (MCP)

nucleoside 5'-monophosphate kinase (Nmk)

nucleoside 2',3'-cyclic monophosphate (2',3'-cNMP)

dipeptide permease (DppA)

P1,P4-di-adenosine 5'-tetraphosphate (Ap4A)

2',3'-cyclic nucleotide phosphodiesterase (CNPase)

2',3'-cyclic nicotinamide adenine dinucleotide phosphate (2',3'-cNADP)

oligoribonucleotides (oligoRNAs)

polynucleotide phosphorylase (PNPase)

5'-untranslated region (5'-UTR)

poly-A polymerase (PAPase)

liquid chromatography-tandem mass spectrometry (LC-MS/MS)

limit of detection (LOD)

casamino acids (CA)

anhydrotetracycline (AHT)

poly-*N*-acetyl-*β*-1,6-D-glucosamine (PNAG)

quantitative reverse transcription PCR (RT-qPCR)

phenol-soluble modulin (PSM)

ellagic acid (EA)

ellagic acid glycoside (EAG)

trimethylsilyl iodide (TMSI)

tert-butyldimethylsilyl (TBS)

tetrabutylammonium iodide (TBAI)

tris(dimethylamino)sulfonium difluorotrimethylsilicate (TASF)

confocal laser scanning microscopy (CLSM)

Chapter 1: Introduction to nucleotide signaling and metabolism

Chapter 1: Introduction to nucleotide signaling and metabolism

1.1 Nucleotide second messenger signaling

The survival of all organisms depends on implementation of appropriate phenotypic responses upon perception of relevant environmental stimuli. Sensory inputs are propagated *via* interconnected biochemical and/or electrical cascades mediated by diverse signaling molecules, including gases, metal cations, lipids, peptides, and nucleotides. These networks often comprise second messenger signaling systems in which a ligand (the primary messenger) binds an extracellular receptor, thereby altering the intracellular concentration of a second messenger molecule which ultimately modulates gene expression through interaction with various effectors. The discovery of adenosine 3',5'-cyclic monophosphate (3',5'-cAMP; **Figure 1.1**) in 1958 and the subsequent identification of its role in glycogenesis constitutes the first example of second messenger-mediated signaling,[1] eventually culminating in the 1971 Nobel Prize in Physiology or Medicine. 3',5'-cAMP functions as a second messenger in tandem with peptide hormones and neurotransmitters as primary messengers to coordinate numerous endocrine, immunological, and neurological processes in eukaryotes. Pathways mediated by 3',5'-cAMP often involve initial binding of the primary messenger to a transmembrane G-protein-coupled receptor (GPCR), which alters the receptor conformation and promotes interaction with associated intracellular heterotrimeric G-proteins consisting of G_α, G_β, and G_γ subunits[2] (**Figure 1.2**). Concomitant structural perturbation of the G_α subunit induces replacement of bound GDP for GTP, resulting in liberation of G_α from the herotrimeric G-protein complex and subsequent modulation of the intracellular 3',5'-cAMP level *via* G_α-induced activation or inhibition of an adenylate cyclase (AC),[2] which converts ATP to 3',5'-cAMP.[3] Nine AC families exist in mammals as
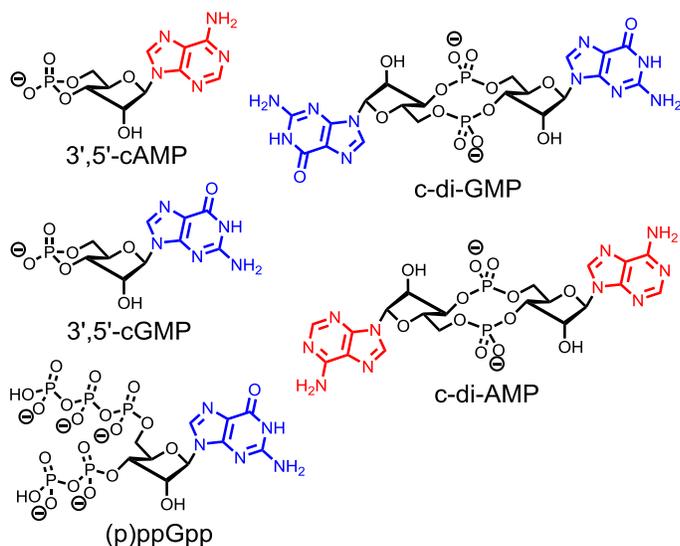
Figure 1.1. Chemical structures of representative nucleotide signaling molecules.

transmembrane proteins, along with one family.[3] The 3',5'-cAMP concentration is subject to additional regulation by cyclic nucleotide phosphodiesterases (PDEs), of which there are eleven known families in mammals.[4] Perturbation of the 3',5'-cAMP concentration directly alters the activity of various protein effectors to elicit transcriptional and physiological changes through phosphorylation cascades and interaction networks. Known 3',5'-cAMP-responsive effectors include protein kinase A (PKA), hyperpolarization-activated cyclic nucleotide-gated (HCN) ion channels, and exchange proteins directly activated by 3',5'-cAMP (EPACs).[3] The protein components of different 3',5'-cAMP signaling networks are spatially confined and often intersect with other eukaryotic second messenger pathways mediated by Ca^{2+} and inositol triphosphate,[5] illustrating the complexity of 3',5'-cAMP-dependent regulation. In addition to the expansive functions of 3',5'-cAMP, the purine 3',5'-cGMP controls an array of eukaryotic processes such as vasodilation,[6] visual transduction,[7] and other neurological functions (such as synaptic transmission and circadian regulation).[8]

Prokaryotic cellular processes also are governed by these two cyclic purine nucleotide second messengers. In *Escherichia coli*, 3',5'-cAMP primarily regulates carbon catabolism *via* the 3',5'-cAMP receptor protein (Crp), a transcription factor that interacts with RNA polymerase to induce expression of over 100 genes.[9] The activity of adenylate cyclase (CyaA) in *E. coli* is

modulated by carbon-source availability; an abundance of glucose indirectly inhibits CyaA-catalyzed 3',5'-cAMP synthesis, whereas glucose limitation activates cyclase activity and increases intracellular 3',5'-cAMP[9a] (**Figure 1.3**). Stimulation of CyaA is mediated by phosphorylated glucose transport protein IIA (IIA^{glc}) which accumulates during glucose starvation, as phosphoryl-IIA^{glc} participates in phosphate transfer to glucose during uptake.[9a] This cyclic nucleotide also mediates expression of virulence- and biofilm-associated genes in other bacterial taxa,[10] and regulates cell division in the archaeon *Halobacterium salinarum*.[11] In contrast to the wide conservation of 3',5'-cAMP signaling cascades in bacteria, the functions of 3',5'-cGMP appear more limited,[10] but this cyclic purine nucleotide promotes encystment in the α-proteobacterium *Rhodospirillum centenum*.[12] Other fundamental bacterial signaling molecules include the acyclic guanosine 3'-diphosphate, 5'-(tri)diphosphate ([p]ppGpp) which remodel the transcriptional and translational landscape to promote virulence gene expression and drive the bacterial stringent response in the wake of nutrient deprivation. (p)ppGpp engages in direct interactions with RNA polymerase, various sigma factors, and translation factors to facilitate these processes, while also modulating the levels of proteins which stabilize or degrade certain sigma factors to indirectly promote transcription of target genes.[10]

Among the most essential signaling molecules in bacteria is the second messenger bis-(3'-5')-cyclic dimeric guanosine
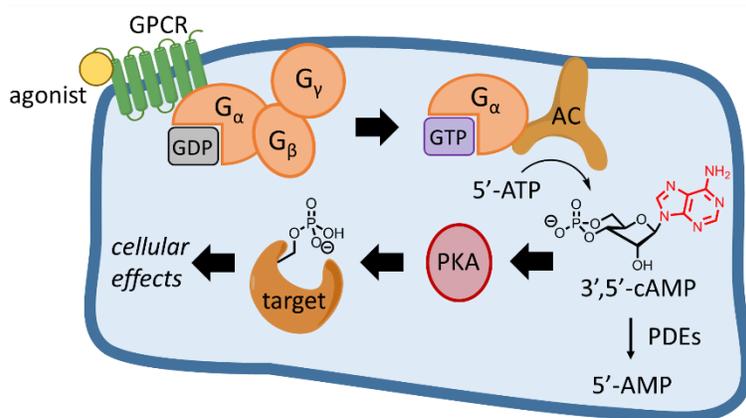
Figure 1.2. Schematic of a general 3',5'-cAMP-mediated pathway in mammals. Binding of the agonist (primary messenger) to the GPCR induces dissociation of the heterotrimeric G-protein. The G-*alpha* subunit stimulates a membrane-bound AC and increases the 3',5'-cAMP concentration, resulting in direct activation of PKA which phosphorylates target proteins to modulate gene expression. The 3',5'-cAMP signal is degraded by PDEs.

monophosphate (c-di-GMP).[13] Initially linked to the allosteric control of cellulose biogenesis in *Gluconacetobacter xylinum* 31 years ago,[14] c-di-GMP governs an intricate regulatory network orchestrating the transition from the motile planktonic state to sessile biofilm communities.[15] Cyclic-di-GMP is biosynthesized from two GTP molecules by diguanylate cyclases (DGCs), which contain a conserved GGDEF amino acid motif in the active site. The activity of some DGCs also is subject to allosteric inhibition *via* c-di-GMP binding to a remote site on the enzyme.[13] Hydrolysis of c-di-GMP to the dinucleotide 5'-phosphoguanylyl-(3'-5')-guanosine (pGpG) is catalyzed by specific phosphodiesterases (PDEs), which are characterized by conserved EAL or HD-GYP motifs in the active site.[15] The aforementioned active site residues are essential for catalytic activity, and genes encoding GGDEF-, EAL-, and HD-GYP-containing enzymes are conserved across bacteria, with many species expressing multiple functional diguanylate cyclases and c-di-GMP-specific phosphodiesterases.[13] For example, the *E. coli* K-12 genome encodes 12 DGCs and 13 PDEs. Among these 25 proteins, six contain domains homologous to both the GGDEF- and EAL-containing families, but lack one of the two associated enzymatic activities, while one such dual-domain-containing protein exhibits both DGC and PDE activity. In addition, four proteins in *E. coli* K-12 are homologous to either the GGDEF- and/or EAL-containing families, but are devoid of DGC and/or PDE activity.[16] These enzymatically inactive proteins nonetheless mediate biofilm production and motility in *E. coli* through interactions with other proteins and small regulatory RNAs (sRNAs). In certain cases, these macromolecular interactions are controlled allosterically by c-di-GMP or GTP binding to the degenerate GGDEF or EAL site.[13] Moreover, catalytically competent GGDEF- and EAL-containing enzymes also engage in this complex interaction network to generate spatially-resolved pools of c-di-GMP which vary over the course of the growth curve, facilitating activation of specific c-di-GMP-binding effectors with

exquisite spatial and temporal control to mediate distinct aspects of motility and biofilm formation[13,17] (**Figure 1.4**). These effectors include both proteins and structured mRNA elements known as riboswitches which undergo c-di-GMP-specific conformational changes to regulate translation of certain transcripts.[13] Families of cyclic-di-GMP-responsive proteins include PilZ, PelD, and FleQ effectors, along with certain enzymatically inactive GGDEF- and EAL-family proteins, as previously discussed herein.[13] Signaling cascades mediated by c-di-GMP are responsive to environmental factors such as oxygen availability,[18] light,[19] aminoglycosides,[20] and nutrient deprivation.[21] Perception of these diverse inputs relies on various GGDEF- and EAL-family proteins, as many contain a sensory domain which enables integration of the signal *via* modulation of the associated DGC or PDE activity.[13] Additional cyclic dinucleotides recently have emerged as second messengers, including c-di-AMP, which regulates sporulation and the response to membrane stress in gram-positive bacteria.[10]
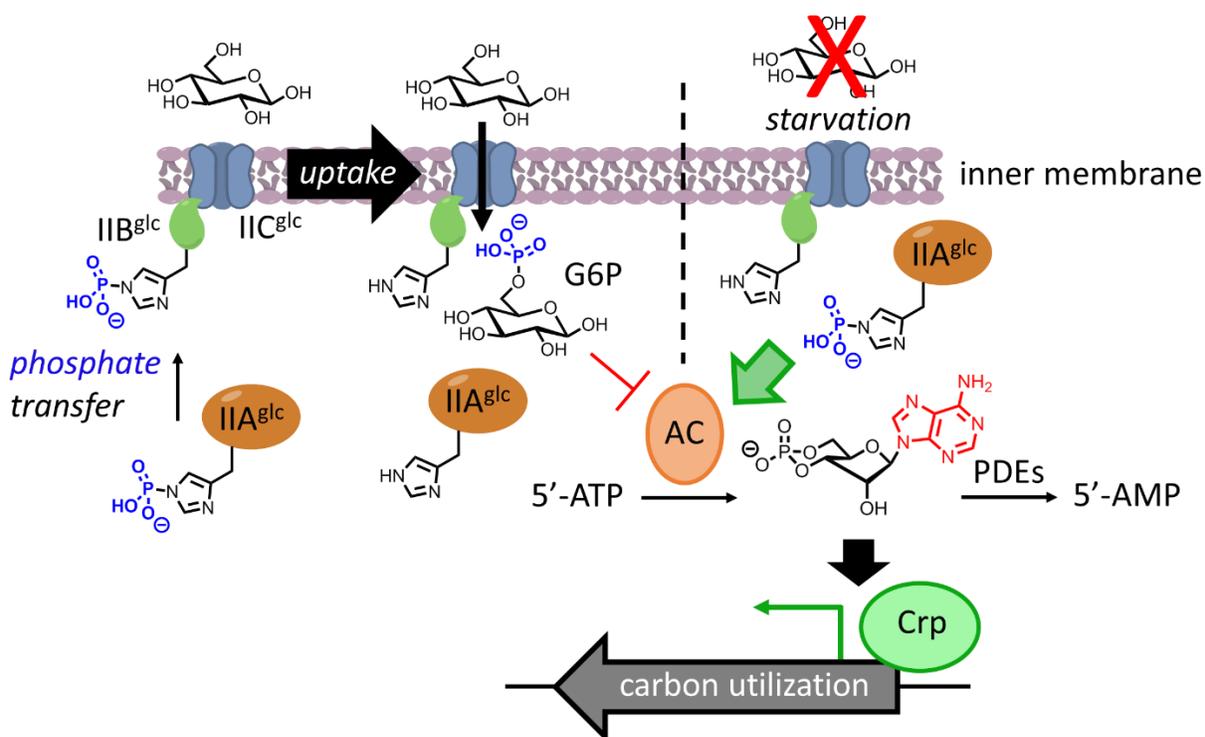
Figure 1.3. Schematic of 3',5'-cAMP-dependent carbon catabolism in *E. coli. Left*: Glucose, when abundant, is transported across the inner membrane by the glucose ABC transport system consisting of

transmembrane IIC^{glc}, membrane-bound IIB^{glc}, and soluble IIA^{glc}. During uptake, glucose is phosphorylated *via* phosphorelay from IIA^{glc} to IIB^{glc}, increasing the concentration of glucose-6-phosphate (G6P) which indirectly inhibits AC activity. *Right*: Under glucose-limiting conditions, accumulation of phosphoryl-IIA^{glc} increases AC activity, and the resulting 3',5'-cAMP binds CRP to upregulate transcription of alternative carbon utilization genes.

These diverse nucleotide-driven prokaryotic signaling cascades also intersect with bacterial quorum sensing (QS) networks, which involve inter- and intraspecies gene regulation *via* secreted autoinducer molecules in a cell density-dependent fashion.[22] QS in gram-negative bacteria relies primarily on diffusible acyl-homoserine lactones (AHL), while gram-positive taxa utilize cyclic peptide autoinducers. In addition, 4,5-dihydroxy-2,3-pentanedione (autoinducer-2 [AI-2]) exists as a mixture of hydrated diastereomeric tetrahydrofurans which regulate quorum interactions both within and across bacterial species.[22b] AHL-mediated QS and c-di-GMP signaling function antagonistically in *Vibrio cholerae*, as accumulation of AHL at high cell density inhibits transcription of c-di-GMP effector genes which promote biofilm production.[23] Another antagonistic link exists between 3',5'-cAMP and AI-2 in *E. coli*, as this cyclic nucleotide binds Crp to directly modulate transcription of AI-2-related genes and indirectly decrease the AI-2 concentration.[24] *Agrobacterium tumefasciens*, which induces crown gall tumor formation in plants, synthesizes (p)ppGpp in response to nutrient limitation, and this nucleotide signal activates transcription of the lactonase AttM, resulting in hydrolysis of AHL and downregulation of virulence gene expression.[25] These examples demonstrate the intermingled nature of nucleotide and quorum sensing signal transduction in bacteria. Moreover, AHLs and AI-2 are biosynthesized from *S*-adenosylmethionine (SAM), providing a link between QS, nucleotide signaling, and primary metabolism.[22]
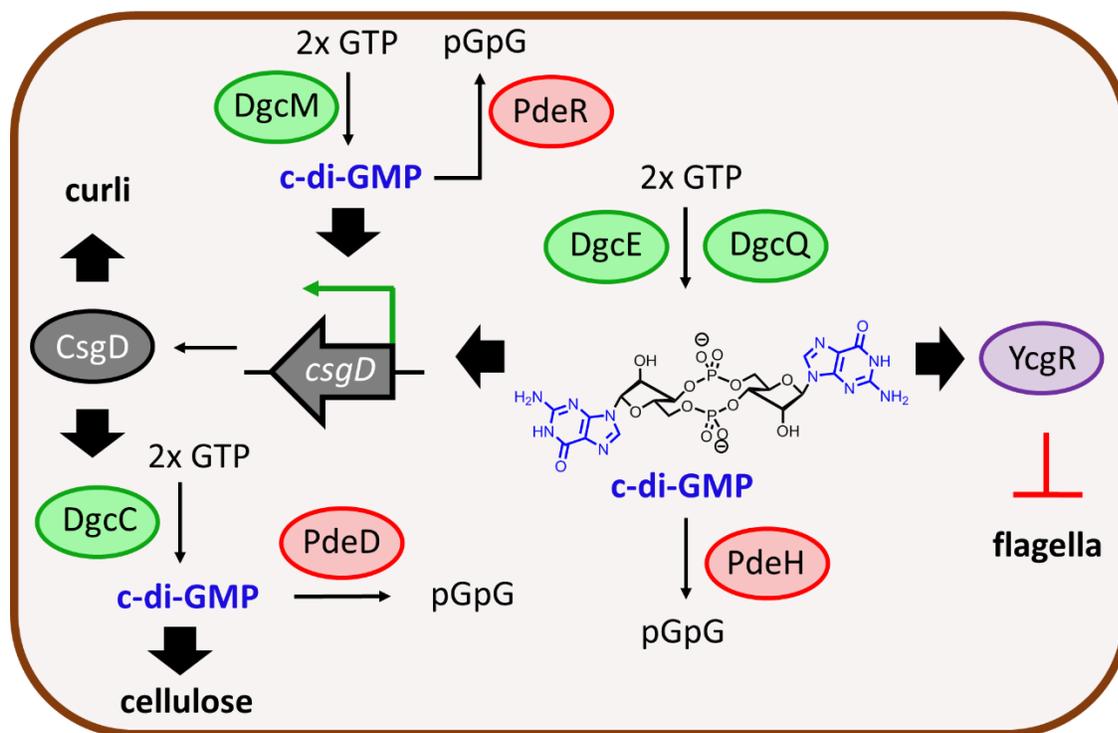
Figure 1.4. Simplified schematic depicting the general logic of c-di-GMP in the regulation of biofilm formation and motility. Cyclases DgcE and DgcQ synthesize c-di-GMP which functions to attenuate flagellar motility via direct interaction with the PilZ-family effector YcgR. In addition, c-di-GMP induces expression of the CsgD transcription factor to stimulate production of biofilm matrix components. Furthermore, induction of CsgD also requires cyclase DgcM, illustrating the multi-faceted nature of the c-di-GMP regulon. CsgD directly activates curli synthesis and also promotes expression of cyclase DgcC, thereby generating another c-di-GMP pool to increase cellulose production. These cellular processes are further regulated by specific phosphodiesterases PdeH, PdeR, and PdeD which engage in protein-protein interactions with certain DGCs and effectors to spatially resolve c-di-GMP signaling networks.

1.2 Primary nucleotide metabolism in regulation of prokaryotic pathways

In addition to the importance of nucleotide signaling in regulating gene expression, nucleotide metabolism is a vital aspect of organismal function. The *de novo* purine and pyrimidine biosynthetic pathways in all organisms are regulated by nucleotide availability *via* multiple levels of allosteric and orthosteric control at both the transcriptional and post-translational level (**Figure 1.5**). In *E. coli*, adenosine- and guanosine 5'-monophosphate (5'-AMP, 5'-GMP) are biosynthesized in 14 steps from 5-phospho-α-D-ribose 1-diphosphate (PRPP), with inosine 5'-monophosphate (5'-IMP) as a key intermediate. 5'-AMP and -GMP allosterically inhibit the

activity of PurF, which transfers ammonia from L-glutamine to PRPP, generating 5-phospho-β-D-ribose 1-amine with loss of diphosphate (PPi) in the first committed step of *de novo* purine biogenesis.[26] Moreover, the expression of *purF* and other *de novo* purine biosynthesis genes is repressed by the PurR transcription factor when hypoxanthine is abundant.[27] Similarly, multiple enzymes in the *de novo* biosynthesis of uridine- and cytidine 5'-monophosphate (5'-UMP, 5'-CMP) are inhibited either allosterically or competitively by 5'-UMP,[28] -UTP,[29] or -CTP.[30] Interestingly, 5'-UTP also attenuates transcription of the *de novo* biosynthesis genes *pyrBI* and *pyrE* via direct stabilization of intergenic structural termination motifs preceding these genes,[31] and hypoxanthine-bound PurR inhibits transcription of several pyrimidine biosynthesis genes.[27a] In addition to these aspects of negative regulation, certain enzymes in *de novo* pyrimidine biosynthesis are activated by the binding of certain nucleotides to allosteric sites, such as carbamoyl phosphate synthetase (encoded by *carBA*) which is stimulated by 5'-IMP binding, resulting in upregulation of the 5'-ATP-dependent conversion of carbonate to carbamoyl phosphate *en route* to 5'-UMP.[28] Ribonucleotide reductases, which catalyze the single-electron reduction of 5'-NDPs to generate 2'-deoxy-NDPs, also are regulated allosterically by nucleotide binding both in terms of substrate specificity and overall catalytic activity.[32]

Due to the energy-intensive process of *de novo* nucleotide biosynthesis, purine and pyrimidine salvage pathways are another valuable source of nucleotides. Most (2'-deoxy)nucleosides also can supply *E. coli* with a sole source of carbon and nitrogen, demonstrating the importance of nucleoside degradation.[33] (2'-deoxy)nucleoside catabolism relies on active transport into the cytosol by nucleoside:H$^+$ symporters and subsequent phosphorolysis by nucleoside phosphorylase enzymes to liberate the nucleobase and (2-deoxy)-α-D-ribose 1-

phosphate, which ultimately enters glycolysis or the citric acid cycle.[33] In accord with the transcriptional regulatory logic that governs induction of *de novo* synthesis genes, pyrimidine and
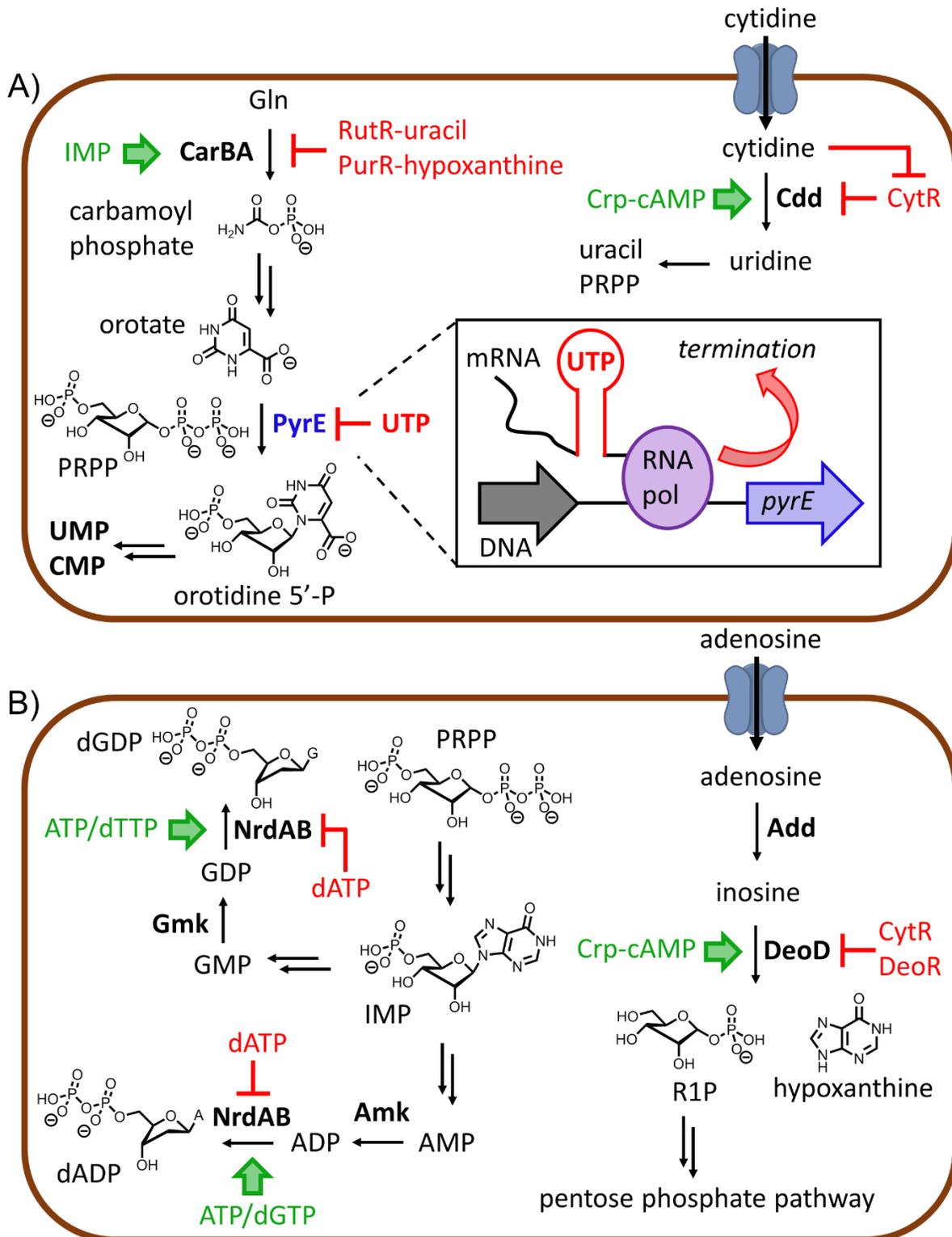
Figure 1.5. General schematic depicting regulatory features of *de novo* pyrimidine **(A)** and purine **(B)** biosynthesis and salvage, which is subject to allosteric and orthosteric regulation at both the transcriptional and post-translational level. Green denotes positive regulation; red denotes negative regulation. See text for details.

purine catabolism is controlled by similar mechanisms, as transcription of salvage genes is inhibited by the CytR and RutR transcription factors when intracellular pyrimidine concentrations are low, with de-repression occurring upon binding of cytidine to CytR and uracil/thymine to RutR (**Figure 1.5**).[34] DeoR is another transcriptional repressor that, in the absence of 2-deoxy-α-D-ribose 1-phosphate, occludes the promoters of genes involved in 2'-deoxynucleoside catabolism to prevent expression, demonstrating another negative feedback mechanism that represses catabolic gene expression in response to nucleotide concentrations.[35]

Malfunction in the aforementioned aspects of nucleotide metabolism impair essential cellular processes. For example, the frequency of genomic mutations increases in *Saccharomyces cerevisiae* expressing a mutated variant of ribonucleotide reductase due to elevated dTTP and dCTP pools relative to wild-type concentrations.[36] Similarly, *E. coli* lacking *ndk* (encoding nucleoside 5'-diphosphate kinase) exhibit approximately 16-fold and 5-fold higher concentrations of dCTP and dGTP, respectively, along with nearly 2-fold lower levels of dATP, compared to concentrations in the isogenic wild-type strain.[37] These aberrant dNTP pools reduced the fidelity of transcription *in vivo*, upregulating erroneous incorporation of dNTPs into RNA by RNA polymerase.[37] Moreover, *in vitro* ribosome dissociation experiments demonstrate that mis-incorporation of a single dNTP into the tRNA anti-codon or mRNA codon induces pre-mature termination of translation.[38] Faithful translation also depends on maintenance of guanine nucleotide pools due to the requirement of 5'-GTP in translation initiation and elongation.[39] Indeed, polysome formation is altered in *S. cerevisiae* deficient in 5'-GMP synthetase, indicating

an altered rate of translation initiation, and this impairment in protein synthesis was linked to an elevated 5'-GTP/-GDP ratio in the absence of functional 5'-GMP synthetase.[40]

In addition to perturbing the efficacy of DNA replication, transcription, and translation, imbalanced nucleotide pools alter other cellular processes, such as bacterial biofilm formation. The production of curli amyloid fibers, an abundant protein in the biofilm matrix, is downregulated in *E. coli* strains auxotrophic for pyrimidines or purines.[41] While the mechanisms linking pyrimidine and purine auxotrophy to impaired curli generation in *E. coli* are unclear, the cytidine-responsive CytR transcription factor (discussed above) appears to play an indirect role.[41] Moreover, the transcriptional repressor AriR which downregulates expression of several biofilm-related genes is induced by the pyrimidine anti-metabolite 5-fluorouracil, suggesting AriR as another possible link between pyrimidine metabolism and biofilm regulation.[42] Regulation of cellulose biosynthesis, another important process in biofilm formation, seemingly responds to the ratios between 5'-UMP and different intermediates in the *de novo* pyrimidine biosynthetic pathway, further demonstrating the role of nucleotide pool regulation in biofilm production.[41] The different cellulose levels in various *E. coli* pyrimidine auxotrophs is mediated by the DGC YedQ, suggesting that c-di-GMP signaling is modulated by pyrimidine biosynthesis.[41] Similarly, pharmacological inhibition of PurH, an essential enzyme in *de novo* purine biosynthesis, impairs biofilm production in *E. coli*, likely via indirect perturbation of c-di-GMP levels.[43]

Cyclic-di-GMP signaling and primary nucleotide metabolism also influence flagellar motility. Chemotaxis to pyrimidine nucleobases is mediated by the Tap methyl-accepting chemotaxis receptor protein (MCP), demonstrating a role for nucleotide metabolism in *E. coli* motility.[44] Curiously, the Tap MCP also induces a positive chemotactic response to dipeptides via direct interaction with dipeptide permease (DppA) in *E. coli*,[45] suggesting another unknown

protein binding partner potentially regulates Tap-dependent pyrimidine perception. The carbohydrate component of nucleosides, D-ribose, also functions as a chemoattractant recognized by the Trg MCP.[46] In contrast to pyrimidine nucleobases and D-ribose, 5'-AMP, 5'-ATP, 3',5'-cAMP, and the four canonical ribonucleosides do not modulate motility in *E. coli*,[46] suggesting a unique connection between pyrimidine salvage and chemotaxis, as pyrimidine nucleobases are byproducts of phosphorolytic nucleoside catabolism.[33]. Intriguingly, the related γ-proteobacterium *Vibrio fischerii* preferentially migrates to nucleosides and nucleotides in favor of nucleobases and D-ribose,[47] indicating different mechanistic links between nucleotide regulation and chemotaxis, even in closely related bacterial taxa. Pathogens rely on chemotaxis and biofilm formation for host-colonization, demonstrating the therapeutic potential of disrupting nucleotide pool regulation in attenuating these processes. Therefore, the present work has developed (bio)chemical tools to identify unknown facets of nucleotide metabolism and signaling in the regulation of bacterial phenotypes.

1.3 Objectives and scope of this dissertation

Nucleotides regulate diverse biological functions across the kingdoms of life, with the paradigm of nucleotide signaling emerging 60 years ago following the discovery of adenosine 3',5'-cyclic monophosphate (3',5'-cAMP) in mammalian liver homogenate.[1a] The ensuing investigations not only elucidated the myriad roles of 3',5'-cAMP in both eukaryotic and prokaryotic biology, but also resulted in the identification of guanosine 3',5'-cyclic monophosphate (3',5'-cGMP) as another fundamental nucleotide signaling molecule in eukaryotes.[48] More recently, additional (a)cyclic nucleotides, including cyclic dimeric (3',5') guanosine monophosphate (c-di-GMP), cyclic dimeric (3',5') adenosine monophosphate (c-di-AMP), guanosine 3'-diphosphate, 5'-(tri)diphosphate ([p]ppGpp), and P1,P4-di-adenosine 5'-

tetraphosphate (Ap4A), have been discovered to modulate numerous phenotypes in bacteria, including biofilm production and virulence.[10] Notably, the development of chemical tools, such as cell-permeable cyclic nucleotide analogs, have been critical in unraveling the roles of these 3',5'-cNMP and c-di-NMP second messengers in regulating organismal adaptation to varied environmental stimuli. These studies have resulted in a resurgence in the identification of nucleotide second messengers and suggested the possibility of additional nucleotide-based signaling molecules.

Recent studies have expanded the frontier of nucleotide signaling to include 3',5'-cyclic pyrimidines in mammals,[49] and the atypical regioisomeric nucleoside 2',3'-cyclic monophosphates (2',3'-cNMPs) in metazoans and plants.[50] Specifically, adenosine 2',3'-cyclic monophosphate (2',3'-cAMP) has been linked to mitochondrial-stimulated apoptosis in mammals and stress-induced translational remodeling in *Arabidopsis thaliana*,[50b,51] alluding to potential functions for other 2',3'-cNMPs in eukaryotes. While the physiological significance of 2',3'-cAMP in eukaryotes is beginning to emerge, the relevance of prokaryotic 2',3'-cNMPs was unknown at the onset of this work.

The present dissertation reports the development and deployment of novel cell-permeable 2',3'-cNMP analogs and a recombinant 2',3'-cNMP phosphodiesterase (CNPase) to manipulate 2',3'-cNMP concentrations *in vivo*. These complementary chemical biology-based approaches have enabled interrogation of transcriptional and phenotypic changes associated with aberrant intracellular 2',3'-cNMP levels, linking 2',3'-cNMPs to the regulation of diverse cellular functions in *Escherichia coli*, such as biofilm production, motility, *β*-lactam tolerance, ribosome homeostasis, and primary nucleotide metabolism.[52] This work also identified RNase I, a member of the widely conserved RNase T2 superfamily, as the enzyme responsible for producing 2',3'-

cNMP pools in *E. coli* from rRNA and mRNA degradation.[53] Importantly, the utilization of CNPase and cell-permeable 2',3'-cNMP derivatives enables modulation of intracellular 2',3'-cNMP levels independently of RNase I expression, facilitating differentiation of RNase I- and 2',3'-cNMP-dependent functions. Consequently, these tools will expedite investigations into cellular functions governed by 2',3'-cNMP pools and T2 family RNases in diverse organisms, illuminating a novel facet of nucleotide signaling which can be modulated to regulate biological (dys)function.

In addition to the optimization and utilization of (bio)chemical approaches to probe intracellular nucleotide signaling and metabolism, biofilm inhibitors inspired by phyto-phenolic glycosides have been synthesized as additional tools to interrogate and control bacterial behavior.[54] Collectively, this work reports the development of deployment of novel chemical tools to augment our understanding of organismal signal transduction.

1.4 References

(1) (a) Sutherland, E.; Rall, T. *J. Biol. Chem.* **1958**, *232*, 1077-1091; (b) Hardman, J.; Robison, G.; Sutherland, E. *Annu. Rev. Physiol.* **1971**, *33*, 311-336.

(2) Hanlon, C.; Andrew, D. *J. Cell Sci.* **2015**, *128*, 3533-3542.

(3) Kamenetsky, M.; Middelhaufe, S.; Bank, E.; Levin, L.; Buck, J.; Steegborn, C. *J. Mol. Biol.* **2006**, *362*, 623-639.

(4) Francis, S.; Blount, M.; Corbin, J. *Physiol. Rev.* **2011**, *91*, 651-690.

(5) Ahuja, M.; Jha, A.; Maleth, J.; Park, S.; Muallem, S. *Cell Calcium* **2014**, *55*, 385-393.

(6) Denninger, J.; Marletta, M. *Biochim. Biophys. Acta* **1999**, *1411*, 334-350.

(7) Mannu, G. *Neurosciences* **2014**, *19*, 275-280.

(8) Domek-Lopacinska, K.; Stroksznajder, J. *Journal of Physiology and Pharmacology* **2005**, *56*, 15-34.

(9) (a) Gorke, B.; Stulke, J. *Nat. Rev. Microbiol.* **2008**, *6*, 613-624; (b) Zheng, D.; Constantinidou, C.; Hobman, J.; Minchin, S. *Nucleic Acids Res.* **2004**, *32*, 5874-5893.

(10) Kalia, D.; Merey, G.; Nakayama, S.; Zheng, Y.; Zhou, J.; Luo, Y.; Guo, M.; Roembke, B.; Sintim, H. *Chem. Soc. Rev.* **2013**, *42*, 305-341.

(11) Baumann, A.; Lange, C.; Soppa, J. *BMC Cell Biol.* **2007**, *8*.

(12) Marden, J.; Dong, Q.; Roychowdhury, S.; Berleman, J.; Bauer, C. *Mol. Microbiol.* **2011**, *79*, 600-615.

(13) (a) Hengge, R. *Nat. Rev. Microbiol.* **2009**, *7*, 263-273; (b) Romling, U.; Galperin, M.; Gomelsky, M. *Microbiol. Mol. Biol. Rev.* **2013**, *77*, 1-52.

(14) Ross, P.; Weinhouse, H.; Aloni, Y.; Michaeli, D.; Weinbergerohana, P.; Mayer, R.; Braun, S.; Devroom, E.; Vandermarel, G.; Vanboom, J.; Benziman, M. *Nature* **1987**, *325*, 279-281.

(15) Simm, R.; Morr, M.; Kader, A.; Nimtz, M.; Romling, U. *Mol. Microbiol.* **2004**, *53*, 1123-1134.

(16) Hengge, R.; Galperin, M.; Ghigo, J.; Gomelsky, M.; Green, J.; Hughes, K.; Jenal, U.; Landini, P. *J. Bacteriol.* **2016**, *198*, 7-11.

(17) Sarenko, O.; Klauck, G.; Wilke, F.; Pfiffer, V.; Richter, A.; Herbst, S.; Kaever, V.; Hengge, R. *MBio* **2017**, *8*.

(18) Chang, A.; Tuckerman, J.; Gonzalez, G.; Mayer, R.; Weinhouse, H.; Volman, G.; Amikam, D.; Benziman, M.; Gilles-Gonzalez, M. *Biochemistry* **2001**, *40*, 3420-3426.

(19) Rajagopal, S.; Key, J.; Purcell, E.; Boerema, D.; Moffat, K. *Photochem. Photobiol.* **2004**, *80*, 542-547.

(20) Hoffman, L.; D'Argenio, D.; MacCoss, M.; Zhang, Z.; Jones, R.; Miller, S. *Nature* **2005**, *436*, 1171-1175.

(21) Gjermansen, M.; Ragas, P.; Sternberg, C.; Molin, S.; Tolker-Nielsen, T. *Environ. Microbiol.* **2005**, *7*, 894-906.

(22) (a) Fuqua, W.; Winans, S.; Greenberg, E. *J. Bacteriol.* **1994**, *176*, 269-275; (b) Waters, C.; Bassler, B. *Annu. Rev. Cell Dev. Biol.* **2005**, *21*, 319-346.

(23) Krasteva, P.; Fong, J.; Shikuma, N.; Beyhan, S.; Navarro, M.; Yildiz, F.; Sondermann, H. *Science* **2010**, *327*, 866-868.

(24) Wang, L.; Hashimoto, Y.; Tsao, C.; Valdes, J.; Bentley, W. *J. Bacteriol.* **2005**, *187*, 2066-2076.

(25) Zhang, H.; Wang, C.; Zhang, L. *Mol. Microbiol.* **2004**, *52*, 1389-1401.

(26) Zhou, G.; Smith, J.; Zalkin, H. *J. Biol. Chem.* **1994**, *269*, 6784-6789.

(27) (a) Cho, B.; Federowicz, S.; Embree, M.; Park, Y.; Kim, D.; Palsson, B. *Nucleic Acids Res.* **2011**, *39*, 6456-6464; (b) Meng, L.; Kilstrup, M.; Nygaard, P. *Eur. J. Biochem.* **1990**, *187*, 373-379.

(28) Trotta, P.; Pinkus, L.; Haschemeyer, R.; Meister, A. *J. Biol. Chem.* **1974**, *249*, 492-499.

(29) Serina, L.; Blondin, C.; Krin, E.; Sismeiro, O.; Danchin, A.; Sakamoto, H.; Gilles, A.; Barzu, O. *Biochemistry* **1995**, *34*, 5066-5074.

(30) Long, C.; Pardee, A. *J. Biol. Chem.* **1967**, *242*, 4715-4721.

(31) (a) Bonekamp, F.; Clemmesen, K.; Karlstrom, O.; Jensen, K. *EMBO J.* **1984**, *3*, 2857-2861; (b) Donahue, J.; Turnbough, C. *J. Biol. Chem.* **1994**, *269*, 18185-18191; (c) Turnbough, C.; Hicks, K.; Donahue, J. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1983**, *80*, 368-372.

(32) Torrents, E. *Front. Cell. Infect. Microbiol.* **2014**, *4*, doi: 10.3389/fcimb.2014.00052.

(33) Vogels, G.; van der Drift, C. *Bacteriological Reviews* **1976**, *40*, 403-468.

(34) (a) Shimada, T.; Hirao, K.; Kori, A.; Yamamoto, K.; Ishihama, A. *Mol. Microbiol.* **2007**, *66*, 744-757; (b) ValentinHansen, P.; SogaardAndersen, L.; Pedersen, H. *Mol. Microbiol.* **1996**, *20*, 461-466.

(35) Skerlov, J.; Fabry, M.; Hubalek, M.; Otwinowski, Z.; Rez, P. *The FEBS Journal* **2014**, *281*, 4280-4292.

(36) (a) Kumar, D.; Abdulovic, A.; Viberg, J.; Nilsson, A.; Kunkel, T.; Chabes, A. *Nucleic Acids Res.* **2011**, *39*, 1360-1371; (b) Watt, D. L.; Buckland, R. J.; Lujan, S. A.; Kunkel, T. A.; Chabes, A. *Nucleic Acids Res.* **2016**, *44*, 1669-1680.

(37) Wang, J.; Dong, H.; Chionh, Y.; McBee, M.; Sirirungruang, S.; Cunningham, R.; Shi, P.; Dedon, P. *Nucleic Acids Res.* **2016**, *44*, 8962-8975.

(38) Fahlman, R.; Olejniczak, M.; Uhlenbeck, O. *J. Mol. Biol.* **2006**, *355*, 887-892.

(39) Asano, K.; Clayton, J.; Shalev, A.; Hinnebusch, A. *Genes Dev.* **2000**, *14*, 2534-2546.

(40) Iglesias-Gato, D.; Martin-Marcos, P.; Santos, M.; Hinnebusch, A.; Tamame, M. *Genetics* **2011**, *187*, 105-122.

(41) Garavaglia, M.; Rossi, E.; Landini, P. *PLOS ONE* **2012**, *7*, doi:10.1371/journal.pone.0031252.

(42) Attila, C.; Ueda, A.; Wood, T. *Appl. Microbiol. Biotechnol.* **2009**, *82*, 525-533.

(43) Antoniani, D.; Rossi, E.; Rinaldo, S.; Bocci, P.; Lolicato, M.; Paiardini, A.; Raffaelli, N.; Cutruzzola, F.; Landini, P. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7325-7336.

(44) Liu, X.; Parales, R. *J. Bacteriol.* **2008**, *190*, 972-979.

(45) Manson, M.; Blank, V.; Brade, G.; Higgins, C. *Nature* **1986**, *321*, 253-256.

(46) Kondoh, H.; Ball, C.; Adler, J. *Proc. Natl. Acad. Sci. USA* **1979**, *76*, 260-264.

(47) DeLoney-Marino, C.; Wolfe, A.; Visick, K. *Appl. Environ. Microbiol.* **2003**, *69*, 7527-7530.

(48) Ashman, D.; Price, T.; Melicow, M.; Lipton, R. *Biochem. Biophys. Res. Commun.* **1963**, *11*, 330-&.

(49) Seifert, R. *Trends Biochem. Sci.* **2015**, *40*, 8-15.

(50) (a) Jackson, E. K. *Am. J. Physiol. Renal Physiol.* **2011**, *301*, F1160-F1167; (b) Kosmacz, M.; Luzarowski, M.; Kerber, O.; Leniak, E.; Gutierrez-Beltran, E.; Beltran, J. C. M.; Gorka, M.; Szlachetko, J.; Veyel, D.; Graf, A.; Skirycz, A. *Plant Physiol.* **2018**, *177*, 411-421; (c) Van Damme, T.; Blancquaert, D.; Couturon, P.; Van der Straeten, D.; Sandra, P.; Lynen, F. *Phytochemistry* **2014**, *103*, 59-66.

(51) Azarashvili, T.; Krestinina, O.; Galvita, A.; Grachev, D.; Baburina, Y.; Stricker, R.; Evtodienko, Y.; Reiser, G. *Am. J. Physiol. Cell Physiol.* **2009**, *296*, C1428-C1439.

(52) Fontaine, B. M.; Duggal, Y.; Weinert, E. E. *Submitted to Nucl. Acids Res.* **2018**.

(53) Fontaine, B. M.; Martin, K. S.; Garcia-Rodriguez, J. M.; Jung, C.; Southwell, J. E.; Jia, X.; Weinert, E. E. *Biochem. J* **2018**, *478*, 1491-1506.

(54) Fontaine, B. M.; Nelson, K.; Lyles, J. T.; Jariwala, P. B.; Garcia-Rodriguez, J. M.; Quave, C. L.; Weinert, E. E. *Front. Microbiol.* **2017**, *8*, DOI: 10.3389/fmicb.2017.00496.

Chapter 2: RNase I regulates *Escherichia coli* nucleoside 2',3'-cyclic monophosphate levels and biofilm formation

Chapter 2: RNase I regulates *Escherichia coli* nucleoside 2',3'-cyclic monophosphate levels and biofilm formation

2.1 Introduction

2.1.1 Discovery of nucleoside 2',3'-cyclic monophosphates and the early characterization of RNase I

In addition to the paradigmatic nucleotide signaling molecules discussed above, recent evidence demonstrates that the regioisomeric nucleoside 2',3'-cyclic monophosphates (2',3'-cNMPs) regulate certain processes in eukaryotes. 2',3'-cAMP and -cGMP were quantified in *Arabidopsis thaliana*, with increased concentrations observed in damaged leaves relative to control.[1] The four 2',3'-cNMPs derived from the canonical RNA nucleobases, along with 2',3'-cIMP, also exist in mammalian organs and cells.[2] Intriguingly, administration of 2',3'-cAMP to rodents *via* the renal artery dramatically increases concentrations of 2'-AMP, 3'-AMP, and adenosine in the urine, alluding to a role for 2',3'-cAMP in physiological processes.[3] Further *ex vivo* experiments with isolated mouse kidneys demonstrated that metabolic stress induces production of extracellular 2',3'-cAMP, 2'-AMP, 3'-AMP, and adenosine, likely by stimulating mRNA degradation.[2c] The extracellular adenosine produced from the dephosphorylation of 2',3'-cAMP in the wake of metabolic insult likely elicits a subsequent anti-inflammatory effect on the system through activation of purinogenic G-protein coupled receptors (GPCRs), further demonstrating the potential significance of 2',3'-cAMP[4] (**Figure 2.1**). Notably, recent work suggests that metabolism of 2',3'-cAMP to 2'-AMP, 3'-AMP, and adenosine occurs in diverse mammalian cell types, and a similar degradative pathway exists which generates extracellular adenosine from the canonical regioisomeric second messenger 3',5'-cAMP, suggesting the conservation of 2',3'-cAMP-mediated signaling in various mammalian tissues.[4] Exogenous 2',3'-

cAMP also stimulates Ca^{2+} efflux in rat kidney cells and oligodendrocytes, resulting in depolarization of the mitochondrial membrane and concomitant apoptosis, thus confirming a physiological role for this cyclic nucleotide in mammals[5] (**Figure 2.1**). The same effect was observed for exogenous 2',3'-cyclic nicotinamide adenine dinucleotide phosphate (2',3'-cNADP) as well,[5] but the endogenous occurrence of 2',3'-cNADP seemingly is unknown.
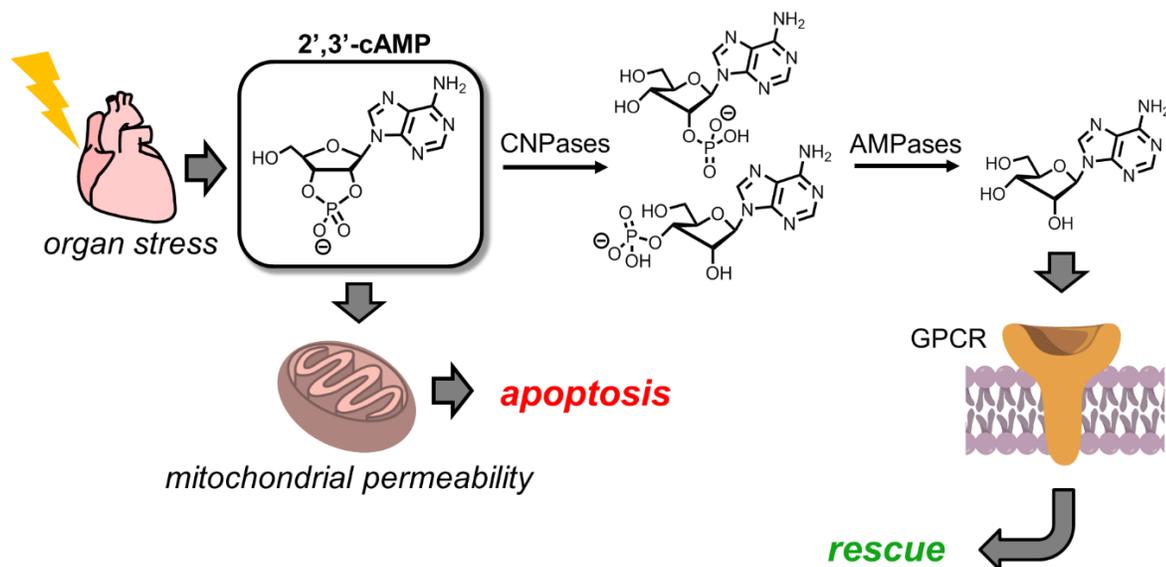
Figure 2.1. Schematic of 2',3'-cAMP metabolism and signaling in mammals. Ischemic organ insult stimulates 2',3'-cAMP production in the extracellular matrix, ultimately increasing the extracellular levels of 2'- and 3'-AMP, as well as adenosine. The balance between 2',3'-cAMP and adenosine pools likely elicits certain opposing cellular effects.

Previous *in vitro* studies identified a metal-independent 2',3'-cyclic nucleotide 3'-phosphodiesterase (CNPase) as a component of myelin in the mammalian nervous system.[6] Subsequent experiments using oligodendrocytes and renal cells either lacking or overexpressing CNPase confirmed that this enzyme hydrolyzes 2',3'-cAMP to 2'-AMP *in vivo*, and, importantly, 2',3'-cAMP is not metabolized by 3',5'-cAMP phosphodiesterases.[7] As mentioned above, metabolism of 2',3'-cAMP also generates 3'-AMP *in vivo*, but the eukaryotic enzymes that catalyze this conversion are unknown.[4] However, several prokaryotic 2',3'-cyclic nucleotide 2'-

phosphodiesterses have been characterized *in vitro*,[8] including EAL-family c-di-GMP phosphodiesterases.[9] Similarly, the mammalian nucleotidases that dephosphorylate 2'- and 3'-AMP to adenosine remain elusive, but the known 5'-AMPase CD73 does not hydrolyze these regioisomeric adenylates.[10] Moreover, the ribonuclease(s) (RNase[s]) presumably generating 2',3'-cAMP in eukaryotes have not been identified, but candidate enzymes include members of the RNase A and RNase T2 families, which cleave the phosphodiester backbone via transphosphorylation to produce a 2',3'-cyclic phosphate.[11] The potential relevance of other 2',3'-cNMPs in eukaryotic physiology also awaits further investigation.

In contrast to the expanding roles of 2',3'-cAMP in mammalian biology, the functions of 2',3'-cNMPs in prokaryotic processes remain entirely unknown, despite the fact that 2',3'-cNMPs were detected over five decades ago in *E. coli* lysate via thin-layer chromatography.[12] Recent work has quantified intracellular and extracellular 2'3,'-cCMP and -cUMP in *Pseudomonas fluorescens*,[13] as well as intracellular 2',3'-cAMP in *Staphylococcus aureus*.[14] 2',3'-cAMP, -cGMP, -cCMP, and -cUMP also recently were observed in *E. coli*, but the physiological concentrations were not determined.[15] 2',3'-cNMPs presumably arise from RNA degradation in prokaryotes, but the specific RNases involved remain unknown. Based on *in vitro* studies, RNase T2 family endoribonucleases appear particularly promising, as these enzymes activate the ribosyl 2'-hydroxyl moiety as a nucleophile, producing a 2',3'-cyclic phosphate and the cleaved RNA product.[11] ^{31}P NMR studies elucidated that these metal-independent nucleases also catalyze a second, slower step in which the liberated 2',3'-cyclic phosphodiester re-binds the active site and undergoes hydrolysis to the acyclic 3'-monophosphate.[16] Importantly, members of the RNase T2 family are non-specific enzymes, resulting in complete digestion of RNA to 2',3'-cNMP monomers *in vitro*. The lack of sequence specificity distinguishes the RNase T2 family from other

transferase-type endoRNases which typically cleave after particular nucleotides (e.g. RNase T1 and RNase A families)[11] or within specific sequence contexts (e.g. interferase family toxins of prokaryotic toxin-antitoxin systems)[17] to produce a 2',3'-cyclic phosphate at the 3'-terminus of the RNA, as opposed to free 2',3'-cNMP monomers. RNase T2 family enzymes are distributed throughout the kingdoms of life and function in diverse processes such as phosphate salvage, regulation of host immunity, and eukaryotic development.[11] Within the bacterial domain, members of the RNase T2 family are most abundant in gram-negative phyla, particularly Proteobacteria, but they exist in gram-positive taxa as well. However, investigations of these bacterial nucleases, both *in vitro* and *in vivo*, have been limited.

Figure 2.2. Schematic depicting the proposed functions of *E. coli* RNase I in RNA catabolism. The role of cytoplasmic RNase I and 2',3'-cNMPs was unknown at the start of this dissertation.

E. coli RNase I, the most well-studied bacterial member of the RNase T2 family, was identified over fifty years ago,[18] but its biological function remains enigmatic. RNase I initially was isolated from the *E. coli* periplasm,[19] and was shown to degrade RNA non-specifically to

produce 2',3'-cNMPs.[20] Subsequently, a cytoplasmic variant encoded by the same *rna* gene was purified from *E. coli*.[20-21] Notably, cytoplasmic RNase I is less stable than the periplasmic version and only digests short oligoribonucleotides (oligoRNAs) devoid of secondary structure *in vitro*, in contrast to periplasmic RNase I which lacks substrate specificity.[20] The periplasmic 2',3'-cNMP phosphodiesterase CpdB likely functions in catabolism of extracellular RNA along with periplasmic RNase I. Despite thorough *in vitro* characterization, the significance of RNase I in *E. coli* is unclear. Early reports identified a role for the enzyme in ribosome decay under certain starvation conditions,[22] but these studies were performed prior to the identification of cytoplasmic RNase I which complicates interpretation of the results. More recently, the 16S rRNA of the 30S ribosomal subunit was shown to inhibit RNase I activity, and *E. coli* mutants expressing rRNA chimeras exhibited altered ribosome decay profiles compared to wild-type *E. coli*. However, the potential physiological factors that stimulate ribosome decay by RNase I in wild-type *E. coli* are unclear. The study also failed to address whether cytoplasmic and/or periplasmic RNase I bind 16S rRNA.[23] Consequently, the physiological role of periplasmic RNase I remains ambiguous, and even less is understood about the function of the cytoplasmic variant. Periplasmic RNase I perhaps functions in catabolism of extracellular RNA, as it is co-localized with the 2',3'-cNMP PDE CpdB which non-specifically hydrolyzes 2',3'-cNMPs to 3'-NMPs *en route* to nucleosides.[8a,8b,24] An analogous combination of a T2 RNase and a 2',3'-cNMP PDE is expressed in the extracellular space of the tomato plant (*Solanum lycopersicum*) in response to phosphate deprivation,[25] further suggesting the importance of the RNase T2 family in nucleotide salvage. Prior work with cytoplasmic RNase I has suggested a role for the enzyme in one of the final steps of mRNA catabolism due to the specificity of RNase I for short oligoRNAs *in vitro*, but such a function has not been investigated experimentally.[20] The present work demonstrates that *E. coli* RNase I

degrades mRNA and rRNA to produce 2',3'-cNMPs, identifying a physiological role for RNase I and providing the first insight into the origin of 2',3'-cNMPs in bacteria.

2.1.2 RNA decay in bacteria

Degradation of messenger RNA enables regulation of protein synthesis and functions as an important source of nucleotides, as basal mRNA decay can proceed at half the rate of transcription in bacteria.[19b] In gram-negative bacteria, mRNA decay typically begins with endonucleolytic cleavage with A/U-rich regions by RNase E of the RNA degradasome,[26] a protein complex consisting primarily of the phosphorolytic 3'-5' exoRNase polynucleotide phosphorylase (PNPase), enolase, and the RNA helicase RhlB; minor components include chaperones DnaK and GroEL, as well as polyphosphate kinase (PPK),[27] but the biological relevance of the chaperones and kinase in the context of the degradosome is unclear. The multiple degradosome components assemble on the C-terminal scaffold domain of RNase E to facilitate transcript inactivation, with ATP-dependent RNA unwinding by RhlB promoting processive exonucleolytic digestion of certain mRNAs by PNPase.[28] The role of the associated enolase in RNA decay is incompletely understood, as this enzyme shows dehydratase activity, converting 2-phospho-D-glycerate to phosphoenolpyruvate. Interestingly, *E. coli* lacking enolase accumulate transcripts encoding genes relevant to carbon transport and metabolism, demonstrating that enolase regulates the half-life of specific transcripts.[29] Importantly, immunoblot experiments and transcriptomic analyses of *E. coli* mutants deficient in different degradosome proteins demonstrated that these proteins also operate independently of degradosome assembly, with the majority of RNase E, enolase, and PNPase existing free in the cell.[29-30] The activity of the degradosome is modulated by proteins RraA and RraB which bind the scaffolding region of RNase E, attenuating its nuclease activity and blocking the RhlB binding interface, thereby disrupting degradosome assembly.[31] The cellular abundance

of the main degradosome proteins also fluctuates over the *E. coli* growth curve, with stationary phase cultures displaying reduced levels of PNPase and increased levels of enolase relative to exponentially growing cultures.[30] In addition to mRNA degradation, RNase E cleaves polycistronic tRNA transcripts *en route* to mature tRNAs.[32] The enzyme also generate the 5S rRNA precursor from the 9S rRNA subsrate and catalyzes ribosome degradation during nutrient limitation.[33] In addition, antisense mRNA silencing in *E. coli* requires RNase E; short regulatory RNAs (sRNAs) hybridize to the 5'-untranslated region (5'-UTR) of target transcripts and sRNA-binding protein Hfq subsequently recruits RNase E for transcript inactivation.[34]
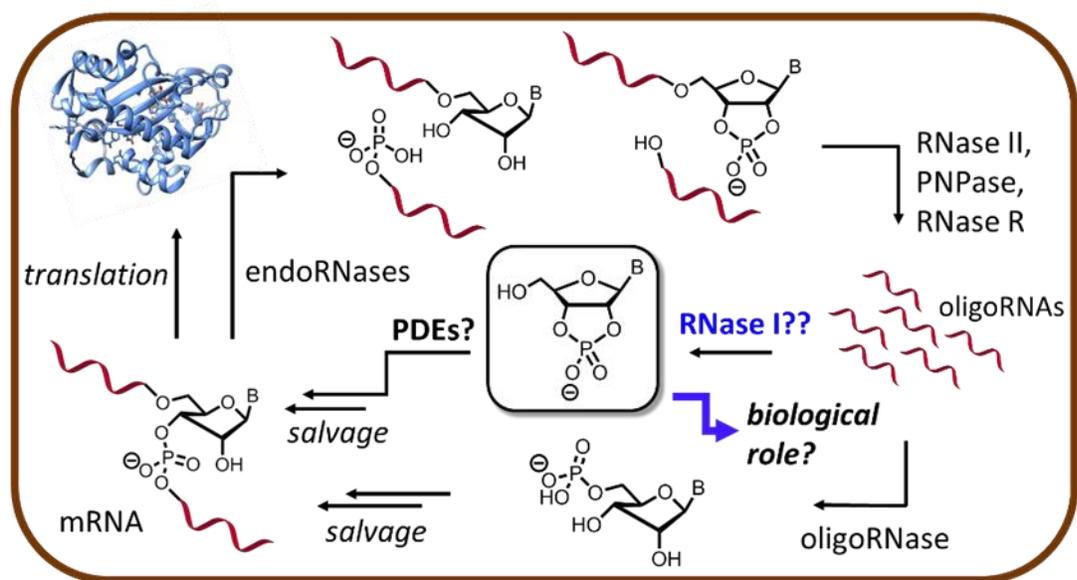
Figure 2.3. General schematic of mRNA degradation in *E. coli*. Transcript inactivation typically occurs with endonucleolytic cleavage by RNase E of the RNA degradasome to generate 3'-hydoxyl and 5'-phosphate termini. Alternatively, other endoRNases hydrolyze the transcript to generate 2',3'-cyclic phosphate and 5'-hydroxyl termini. Subsequently, the transcript fragments are digested by exoRNases to release 5'-NMPs and short (~10-nucleotide) oligoRNA fragments. Catabolism is completed by oligoRNase, and RNase I has been suggested to function in this final catabolic step as well.

Following RNase E-mediated inactivation of the transcript, the processive 3'-5' exonucleases RNase II, RNase R, and PNPase further digest the mRNA to 5'-phosphorylated nucleotides[35] (**Figure 2.3**). The function of these exoRNases is not entirely redundant, as global

transcript levels are differentially perturbed in *E. coli* mutants lacking each of these individual nucleases.[36] Importantly, PNPase and RNase II fail to completely digest substrates to monomeric nucleotides,[37] and this role is fulfilled by oligoRNase, which produces 5'-NMPs[38] (**Figure 2.3**) Moreover, the *in vitro* affinity of cytoplasmic RNase I for the degradation of short oligoRNAs suggests a similar function for this endonuclease in the terminal step of mRNA catabolism.[20] Intriguingly, PNPase not only phosphorolytically degrades mRNA to generate 5'-NDPs, but also catalyzes the reverse reaction in which 5'-NDPs are polymerized onto the 3'-terminus of an RNA in a template-independent fashion *in vivo*.[39] This polymerase activity likely facilitates 3'-5' exonucleolytic degradation by other exoRNases, similarly to 3'-polyadenylation by poly-A polymerase (PAPase).[40] PNPase and RNase R also facilitate destruction of defective tRNA and rRNA.[41]

Another important factor modulating transcript stability is secondary structure in the 5'- and 3'-UTR of the mRNA. For example, RNase E-mediated cleavage proceeds most efficiently upon binding to 5'-monophosphorylated transcripts (as opposed to 5'-triphosphate or 5'-hydroxyl termini),[42] and sequestration of the 5'-phosphate moiety in a stem-loop structure impairs endonucleolytic processing by RNase E, thereby increasing mRNA half-life.[43] Certain transcripts in *E. coli* also exhibit secondary structure in the 3'-UTR which terminate transcription independently of the Rho termination factor, a protein that binds 3'-single-stranded regions of Rho-dependent mRNAs to halt transcription.[44] The presence of structured Rho-independent terminator motifs inhibits exonucleolytic decay, as exoRNases require unstructured 3'-termini for binding.[45] Moreover, the coupling of transcription and translation in bacteria shields mRNA from endonucleolytic decay due to the presence of RNA polymerase and ribosomes along the transcript.[46] In addition to these aspects of mRNA structure in the regulation of transcript stability,

the cellular localization of RNA degrading enzymes, as well as the genomic coordinates of the chromosomal locus undergoing transcription, likely influence mRNA half-life.[35] Indeed, immunogold cryo-electron microscopy experiments demonstrated that the *E. coli* degradosome, RNase II, and PAPase occupy the periphery of the cytoplasm,[30,47] as does the sRNA-binding protein Hfq.[48] Furthermore, confocal microscopy studies identified conformational changes in the *E. coli* nucleoid upon induction of loci encoding membrane proteins, suggesting spatially resolved synthesis of the nascent mRNA and protein near the inner membrane.[49] These intriguing aspects of subcellular localization constitute an emerging feature of mRNA half-life regulation.

Environmental stress, such as nutrient limitation and the ensuing transition into stationary growth phase, alters several facets of RNA degradation. Starved *E. coli* induce the stringent response by upregulating synthesis of the nucleotide second messenger (p)ppGpp, which subsequently dampens transcription of rRNA genes and competitively inhibits GTP binding to translation initiation factor 2.[50] These combined phenomena repress translation and likely render mRNA more vulnerable to endonucleases like RNase E due to decreased shielding by translating ribosomes.[51] Moreover, RNase E-initialized ribosome destruction potentially contributes to this effect as well.[33b] Starvation also upregulates production of certain sRNAs which modify the transcriptome to counteract the stress.[35] Recent work also has demonstrated that RNases modulate biofilm formation, another important bacterial response to adverse stimuli. *E. coli* mutants deficient in either the RNase R, PNPase, or RNase II exonucleases exhibit aberrant biofilm production.[36] Similarly, transposon mutagenesis revealed a role for a T2 family RNase in promoting *Acinetobacter baumanii* biofilm formation,[52] demonstrating the expansive and increasing roles of RNases in controlling bacterial physiology. These discoveries allude to the medicinal value of elucidating novel features of RNase-mediated cellular processes. Experiments

reported herein have identified a role for RNase I and 2',3'-cNMPs in modulating biofilm formation and motility in *E. coli*.

2.2 Results

2',3'-cNMP levels fluctuate during E. coli growth

Physiological concentrations of 2',3'-cAMP, -cCMP, -cGMP, and -cUMP were quantified in wild-type (WT) *E. coli* BW25113[53] during growth in M9 minimal medium using a sensitive LC-MS/MS-based protocol.[2b] In exponentially growing *E. coli* BW25113 cultures, the various 2',3'-cNMPs exist at intracellular concentrations of approximately 10-30 μM (**Figure 2.4**). After 16 h of growth, the four 2',3'-cNMP concentrations fall to undetectable levels (limit of detection [LOD] of the LC-MS/MS assay is approximately 150-500 fmol for the different 2',3'-cNMPs[2b]). Intriguingly, the
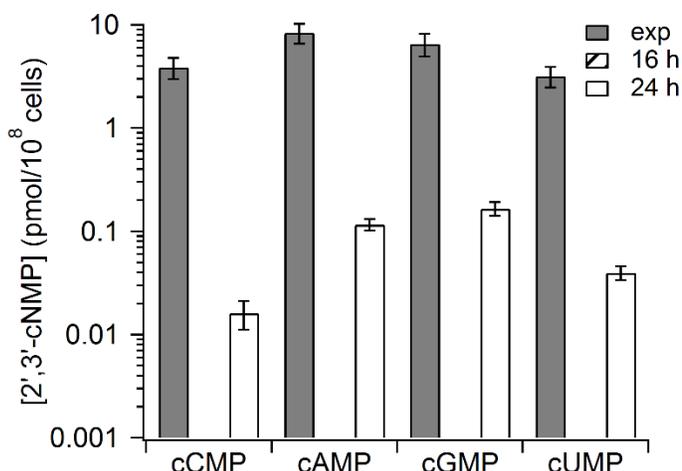
Figure 2.4. Growth phase-dependent 2',3'-cNMP concentrations in *E. coli*. Concentrations of 2',3'-cCMP, -cAMP, -cGMP, and -cUMP in *E. coli* BW25113 WT quantified by LC-MS/MS during mid-exponential (exp) and stationary phase growth (16 h and 24 h post-inoculation) in M9 minimal medium (0.4% glucose, 0.2% casamino acids). 2',3'-cNMPs were below the limit of detection (LOD) in the 16 h cultures. Unless otherwise specified, data are the mean of at least three biological replicates, and error bars denote standard deviation. All differences between exponential and stationary phase measurements and between purine and pyrimdine concentrations are statistically significant ($P < 0.05$). All subsequent 2',3'-cNMP quantifications were performed in exponentially growing cultures in M9 minimal media (supplemented with 0.4% glucose and 0.2% casamino acids), unless otherwise noted.

2',3'-cNMP concentrations then increase in 24 h-old cultures to nanomolar levels, approximately 40-240 fold lower than the exponential-phase concentrations, depending on the particular nucleotide (**Figure 2.4**). *E. coli* also exhibit distinct relative concentrations of the different 2',3'-cNMPs, maintaining 2-fold and 5-fold higher concentrations of the 2',3'-cyclic purines relative to

the pyridimines in exponential phase and stationary phase (24 h-old) cultures, respectively (**Figure 2.4**). These data demonstrate that relative and absolute 2',3'-cNMP concentrations are regulated during *E. coli* growth (**Figure 2.4**).

RNase I generates 2',3'-cNMPs in vivo from RNA degradation

To identify the enzyme(s) involved in 2',3'-cNMP production, cyclic nucleotide levels were quantified in an *E. coli* strain deficient in RNase I (BW25113 Δ*rna*)[54], as this enzyme generates 2',3'-cNMP monomers *in vitro* from short (~2-12-nt), unstructured oligoribonucleotides[19d]. The results demonstrate that seemingly all 2',3'-cNMPs produced during exponential and stationary phase growth arise from RNase I activity (**Figure 2.5A, 2.5C,** and **Supplementary Figure S2.10**), as 2',3'-cNMP levels in the Δ*rna* strain were below the limit of detection. To solidify the role of RNase I in 2',3'-cNMP production, the Δ*rna* strain was transformed with a plasmid encoding the *rna* gene, which restored 2',3'-cNMP production to approximately wild-type levels (**Figure 2.5C**).

Although the *rna* gene encoding RNase I contains a periplasmic localization sequence,[55] previous reports of a cytoplasmic RNase I variant encoded by the same gene[20,56] necessitated separate quantification of periplasmic and cytoplasmic 2',3'-cNMP concentrations to determine the cellular localization of 2',3'-cNMPs. Importantly, 2',3'-cNMPs exist in both the periplasm and the cytoplasm, demonstrating that cytoplasmic RNase I degrades cytosolic RNA (**Figure 2.5B**). To probe the possibility that cytoplasmic 2',3'-cNMPs accumulate due to transport across the inner membrane following RNase I-catalyzed degradation of periplasmic RNA, 0.1 mM exogenous 2',3'-cAMP was added to exponentially-growing cultures of the wild-type BW25113 strain in a separate experiment. This addition amounts to approximately 20,000 pmol exogenous 2',3'-cAMP added per 1×10^8 cells – over 4000-fold greater than the typical physiological 2',3'-

cAMP concentration of *ca*. 5 pmol per 1×10^8 cells in exponential phase (**Figure 2.4**). 2',3'-cAMP levels were unaffected by exogenous nucleotide addition, showing no change compared to cultures treated with exogenous 0.1 mM 3'-AMP as a control (**Figure 2.6A**), further supporting that 2',3'-cNMPs are generated from RNA degradation in the cytoplasm.
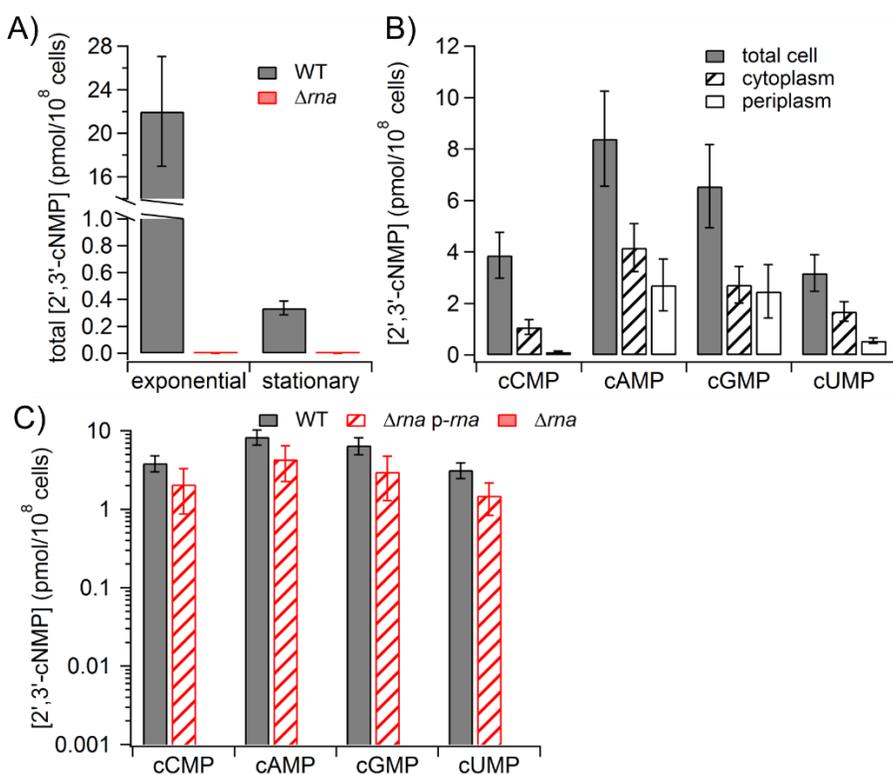
Figure 2.5. 2',3'-cNMPs are produced by RNase I and accumulate in both the cytoplasm and periplasm. (A) Sum of 2',3'-cAMP, -cCMP, cGMP, and cUMP concentrations in the total cell extract of WT and Δ*rna* during mid-exponential and late stationary phase growth (24 h post-inoculation). The *rna* mutant did not contain detectable 2',3'-cNMP levels (LOD *ca*. 150-500 fmol for the different 2',3'-cNMPs). **(B)** 2',3'-cNMP concentrations quantified in WT *E. coli* from total cell extract, cytoplasmic extract, and periplasmic extract. **(C)** Complementation of Δ*rna* with plasmid pCA24N-*rna* generates 2',3'-cNMPs in RNase I-deficient *E. coli*, demonstrating that RNase I produces 2',3'-cNMPs in *E. coli*.

Currently, the biological role of RNase I is unknown, but the nuclease is not essential for growth in *E. coli*.[56] Based on *in vitro* studies, cytoplasmic RNase I has been proposed to complete the catabolism of short oligoribonucleotides generated from mRNA degradation *in vivo*, but this function remains speculative.[20] In accord with the inability to degrade structured RNA substrates

in vitro, RNase I is not involved in the initial inactivation of mRNA, as previously determined by quantifying the half-life of functional β-galactosidase transcript.[57] However, the function of the enzyme in the downstream degradation of short mRNA fragments resulting from transcript inactivation has not been investigated experimentally. To this end, mRNA degradation was perturbed by amino acid starvation and by overexpression of a non-translatable mRNA to probe the effect on 2',3'-cNMP levels (**Figure 2.6**). The role of RNase I in the degradation of rRNA also was investigated by chloramphenicol-induced ribosome turnover (**Figure 2.6**). Amino acid starvation has been demonstrated to induce expression of a number of *Escherichia coli* toxin-antitoxin systems, including RelE and MazF, that cleave mRNA.[58] Moreover, *E. coli* lacking endoribonuclease toxin RelA display dysregulated activation of amino acid biosynthetic genes in the wake of nutrient deprivation, demonstrating the importance of toxin-antitoxin systems in responding to amino acid limitation [59]. Therefore, if 2',3'-cNMPs are formed during mRNA degradation, 2',3'-cNMP levels should be dependent on the presence of amino acids in the growth media. Indeed, *E. coli* BW25113 grown in minimal media with 1.2% casamino acids exhibit markedly lower concentrations of 2',3'-cNMPs than the same strain grown in minimal media with 0.2% casamino acids (**Figure 2.7B**).
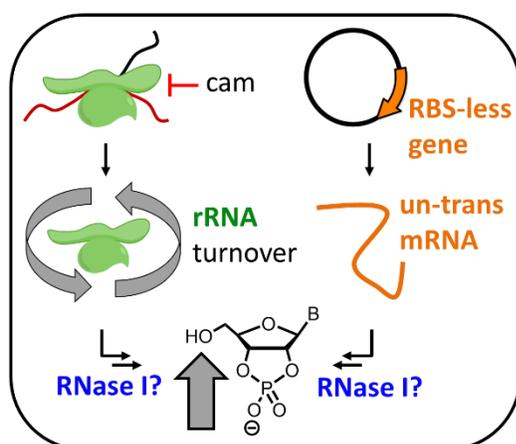
Figure 2.6. Chemical approaches to perturb RNA metabolism in *E. coli*. *Left:* Treatment with chloramphenicol (cam) up-regulates rRNA turnover; thus 2',3'-cNMP levels will increase in cam-treated cultures if RNase I degrades rRNA. *Right:* Over-expression of a gene lacking a ribosome-binding site (RBS) will flood the cell with an un-translatable mRNA substrate, resulting in elevated 2',3'-cNMP pools if RNase I degrades mRNA.

To further solidify a function for RNase I in mRNA degradation, a plasmid-borne gene lacking a ribosome-binding site (pACYC-*noRBS-mRNA*) was overexpressed to increase the intracellular mRNA

concentration, with the expectation that 2',3'-cNMP levels would increase upon expression of the mRNA substrate. In accord with the hypothesis, *E. coli* cultures expressing the non-translatable mRNA displayed ~2-2.7-fold higher 2',3'-cNMP levels compared to control cultures lacking the plasmid, providing additional evidence that RNase I degrades mRNA to produce 2',3'-cNMPs (**Figure 2.7C**). Collectively, these experiments identify a novel role for RNase I in mRNA catabolism.
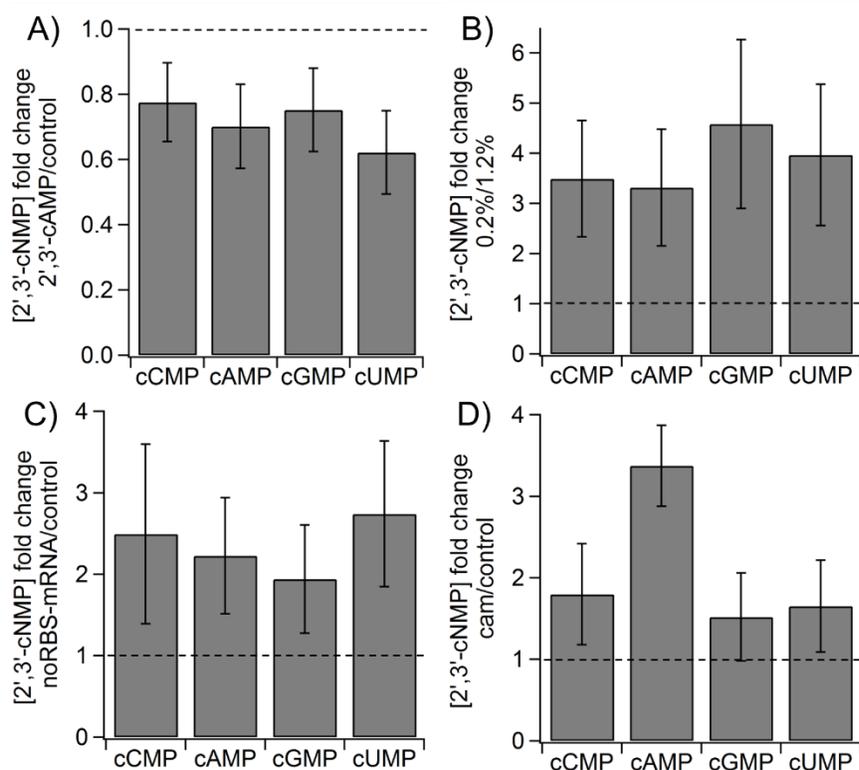
Figure 2.7. 2',3'-cNMPs are generated through RNA degradation in the cytoplasm. **(A)** 2',3'-cNMP levels in BW25113 WT cultures following addition of 0.1 mM exogenous 2',3'-cAMP or 0.1 mM exogenous 3'-AMP control quantified 20 min after nucleotide addition. **(B)** Effect of amino acid starvation on 2',3'-cNMP concentrations in BW25113 WT. BW25113 WT were cultured either in M9 minimal (0.4% glucose, 0.2% casamino acids [CA]) or in the same medium containing 1.2% casamino acids and the 2',3'-cNMPs were quantified. **(C)** 2',3'-cNMP levels increase upon expression of a non-translatable mRNA. 2',3'-cNMPs were quantified in BW25113 (DE3) overexpressing a non-translatable transcript lacking a ribosome binding site (pACYC-*noRBS-mRNA*) and compared to levels in BW25113 (DE3) lacking the construct. **(D)** Addition of chloramphenicol induces rRNA cycling and increases 2',3'-cNMP levels. 2',3'-cNMPs were quantified in BW25113 following treatment with either chloramphenicol (cam) or ethanol as a control. All fold changes reported (B, C, D) are statistically significant ($P < 0.05$). Graphs of all quantified 2',3'-cNMP levels can be found in **Supplementary Figure S2.12**.

The role of RNase I in ribosome decay was interrogated by treating WT *E. coli* cultures with chloramphenicol to stall translation and concomitantly increase rRNA turnover [60]. Cultures treated with chloramphenicol displayed higher 2',3'-cNMP levels relative to concentrations in the ethanol-treated control cultures (**Figure 2.7D**), demonstrating that the increased 2',3'-cNMP levels arise from rRNA degradation. Moreover, a control experiment confirmed that chloramphenicol treatment altered the total RNA concentration, as expected (**Supplementary Figure S2.6**). These studies suggest that cytoplasmic RNase I is involved in degradation of mRNA and rRNA to yield 2',3'-cNMPs. Additional work is underway to determine whether RNase I also degrades tRNA.

RNase I and 2',3'-cNMPs modulate biofilm formation

Several reports have established intriguing links between nucleoside/nucleotide pools and bacterial biofilms,[61] which are microbial communities of aggregated cells growing in an extracellular matrix of polysaccharides, nucleic acids, and other biopolymers.[62] Therefore, the roles of RNase I and 2',3'-cNMPs in biofilm formation were interrogated. Levels of 2',3'-cNMPs for *E. coli* BW25113 WT cells grown in shaking *versus* static culture first were investigated. Significant differences in 2',3'-cNMP levels were observed; quantification yielded approximately 15-fold lower levels of all 2',3'-cNMPs for cells grown in static, biofilm-forming cultures, as compared to shaking cultures (**Figure 2.7A**). To test the possibility that the decreased 2',3'-cNMP levels observed in sessile culture were simply a result of reduced metabolism compared to planktonic cells, the metabolic state of the cultures was assessed using a colorimetric tetrazolium-based assay. Although the static/20°C cultures exhibited approximately 1.3-fold decreased metabolism relative to the shaking/37°C cultures (**Supplementary Figure S2.3**), the metabolic

difference is not sufficient to explain the 15-fold lower 2',3'-cNMP concentrations in the static cultures (**Figure 2.8A**). Furthermore, compared to WT BW25113, which is known to form a relatively poor biofilm,[63] biofilm formation increased over 10-fold in the Δ*rna* strain (**Figure 2.8B** and **Supplementary Figure S2.5**), which does not have observable levels of 2',3'-cNMPs (**Figures 2.5A, 2.5C,** and **Supplementary Figure S2.10**). These results demonstrate that 2',3'-cNMP concentrations are correlated with biofilm formation, with sessile cells having low levels of 2',3'-cNMPs.
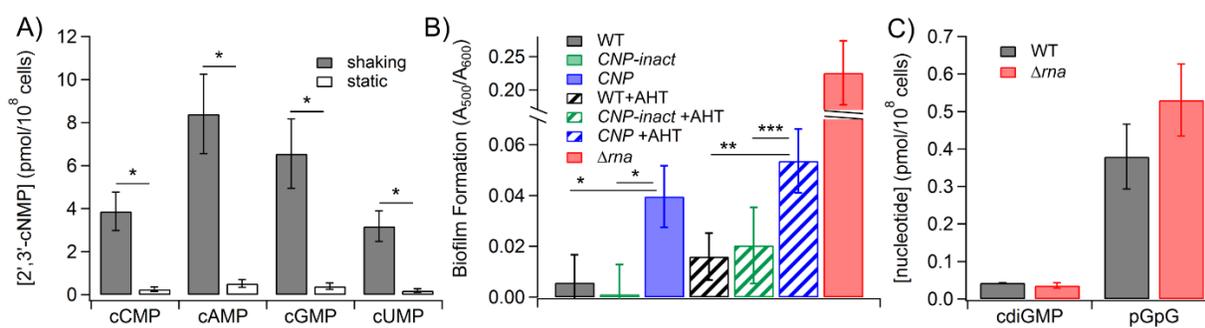
Figure 2.8. RNase I and 2',3'-cNMPs regulate biofilm formation. **(A)** 2',3'-cNMP levels are decreased in static *E. coli* BW25113 cultures relative to shaking cultures. BW25113 WT were cultured either at 37°C with shaking to exponential phase or at 20°C without shaking for 24 h prior to 2',3'-cNMP quantification (*$P < 0.05$). **(B)** Elevated biofilm production in the *rna* mutant and in WT cultures overexpressing a 2',3'-cyclic nucleotide phosphodiesterase (CNPase) relative to control. The biofilm formation of WT cultures harboring plasmid pKT-*CNP* or inactive variant pKT-*CNP-inact* as a control was quantified by Congo red staining after 48 h of static growth at 20°C in the presence or absence of 25 ng mL^{-1} anhydrotetracycline (AHT) and compared to biofilm production in the WT and Δ*rna* strains. (*$P < 0.05$, **$P < 0.01$, ***$P < 0.001$). **(C)** Cyclic-di-GMP and pGpG levels quantified in *E. coli* BW25113 WT and Δ*rna* strains grown in static cultures at 20°C for 18 h in M9 minimal (0.4% glucose, 0.2% casamino acids).

One possible explanation for the dysregulated biofilm formation in the *rna* mutant is aberrant c-di-GMP signaling. Previous studies with *Pseudomonas aeruginosa* demonstrated that deletion of oligoRNase, which degrades 2-5-nucleotide RNAs, increases the concentration of 5'-phosphoguanylyl-3':5'-guanosine (pGpG), the immediate degradation product of the important biofilm regulator c-di-GMP.[64] Accumulation of pGpG inhibits c-di-GMP-specific phosphodiesterases, thereby increasing the c-di-GMP concentration in *P. aeruginosa* lacking

oligoRNase, resulting in upregulated biofilm production.[64] However, *E. coli* contains both oligoRNase and RNase I,[65] while *P. aeruginosa* lacks a known homolog of RNase I. Due to the similar capability of oligoRNase and RNase I to hydrolyze short oligoRNAs *in vitro*,[20,38] the effect of RNase I deletion on intracellular pGpG and c-di-GMP levels in *E. coli* was investigated. Intriguingly, neither c-di-GMP nor pGpG levels were altered in Δ*rna* relative to WT *E. coli* (**Figure 2.8C**), suggesting alternative c-di-GMP-independent mechanisms for RNase I and/or 2',3'-cNMPs in modulating biofilm formation in this bacterium.

To independently investigate the role of RNase I *versus* the role of 2',3'-cNMPs in biofilm formation, the catalytic domain of a mammalian 2',3'-cyclic nucleotide phosphodiesterase (CNPase, UniProtKB-P1323)[66] was developed as an inducible tool to hydrolyze 2',3'-cNMPs in WT *E. coli* expressing RNase I. WT cells harboring plasmid pKT-*CNP* or inactive variant pKT-*CNP-inact* as a control were assayed for biofilm formation *via* Congo red assay. Biofilm formation increased in WT cultures expressing CNPase relative to control cultures expressing the inactive CNPase variant, both in the presence of the inducer (anhydrotetracycline, AHT) and under basal expression conditions in the absence of AHT (**Figure 2.8B**), thus demonstrating a functional link between 2',3'-cNMPs and biofilm formation. Importantly, expression of active CNPase in sessile cultures decreased levels of the 2',3'-cyclic purine nucleotides below the quantification limit of the LC-MS/MS assay, while reducing concentrations of 2',3'-cCMP and -cUMP 25-fold and 14-fold, respectively, compared to levels in cells expressing the inactive CNPase control (**Supplementary Figure S2.4**). These results further indicate that decreasing 2',3'-cNMP levels upregulates biofilm formation.

Curli production is up-regulated in the *rna* mutant

To provide mechanistic insight into the hyper-biofilm phenotype of the RNase I-deficient strain, which lacks 2',3'-cNMPs, the effect of *rna* deletion on transcript levels of biofilm-associated genes was investigated. Analysis of the transcriptome indicated 1.5-fold higher expression of curli structural gene *csgB* and 1.8-fold increased expression of curli accessory gene *csgC* in Δ*rna* compared to WT *E. coli* (**Figure 2.9**). Thus, the upregulated biofilm production in the mutant strain is perhaps due, at least in part, to increased curli synthesis. Curiously, Δ*rna* displays decreased expression of the divergently transcribed *csgDEFG* operon, which activates *csgBAC* transcription (CsgD) and regulates curli transport and assembly (CsgEFG) (**Figure 2.8**).[67]

Upregulated expression of curli genes is consistent with the increased Congo red staining observed in the *rna* mutant and in WT cells expressing active CNPase (**Figure 2.7B**), as Congo red primarily binds to amyloid curli fibers and cellulose.[68] Notably, decreased expression of the *pgaABCD* locus in Δ*rna* indicates that elevated poly-*N*-acetyl-*β*-1,6-D-glucosamine (PNAG) production is not contributing to the hyper-biofilm phenotype.[69] To validate the surprising downregulated expression of the PNAG biosynthetic operon, quantitative reverse transcription PCR (RT-qPCR) was performed to quantify abundance of the *pgaA* transcript. The *rna* mutant displayed reduced *pgaA* expression, further confirming that increased PNAG
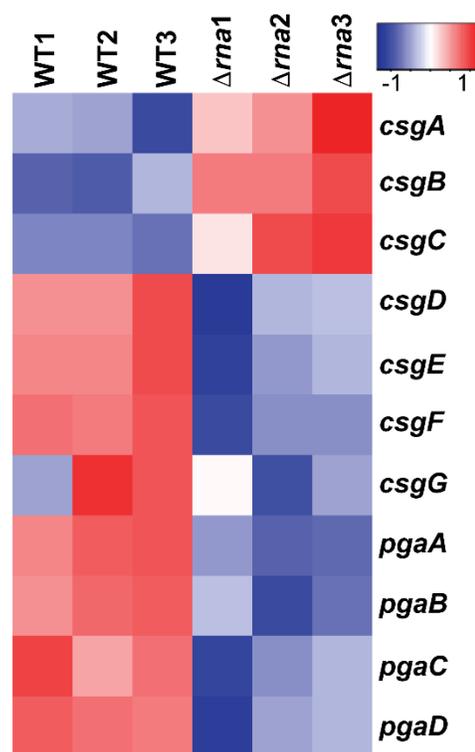

Figure 2.9. Deletion of *rna* alters expression of biofilm-related genes. Heatmap of expression levels of biofilm-associated genes for *E. coli* BW25113 WT and Δ*rna* strains. Heatmap was created using Heatmapper. Raw gene expression data can be found in **Supplementary Table S2.4**.

synthesis is not responsible for hyper-biofilm production in RNase I-deficient *E. coli* (**Supplementary Table S2.5**).

2.3 Discussion

Absolute and relative nucleotide concentrations are maintained by elaborate regulation of *de novo* synthesis and salvage pathways. These processes are vital not only in primary metabolism, but also in the coordination of specialized signal transduction cascades which rely on nucleotide second messengers. The present work demonstrates that 2',3'-cNMP concentrations are regulated over the *Escherichia coli* growth curve and are generated by RNase I-catalyzed degradation of mRNA and rRNA (**Figures 2.4, 2.5,** and **2.7**), presumably, based on the inability of RNase I to digest structured RNA substrates,[20] as one of the final steps in RNA catabolism. RNase I homologs exist in several classes within Proteobacteria, indicating that 2',3'-cNMPs likely govern certain biological processes in other bacterial taxa. In addition, genes encoding other RNase T2 superfamily enzymes are distributed across bacteria, eukaryotes, and viruses,[11] alluding to possible 2',3'-cNMP-dependent pathways in diverse kingdoms of life. The present results suggest that 2',3'-cNMP pools constitute a previously unknown facet of primary nucleotide metabolism and/or a novel nucleotide second messenger signaling system. 2',3'-cNMPs and the corresponding 3'-NMPs resulting from enzymatic hydrolysis possibly function as intermediates in a novel salvage pathway, as the nonspecific nucleotidase SurE in the cytoplasm accepts 3'-NMPs as substrates.[70] Analysis of previously published NTP, NDP, NMP, and nucleoside concentrations in *E. coli* fails to suggest many obvious parallels between the 2',3'-cNMP ratio and other nucleotide/nucleoside pools.[71] However, the finding that 2',3'-cNMP levels decrease in stationary phase *E. coli* cultures relative to exponential phase cultures mirrors the previously observed growth-dependent fluctuation in dNTP concentrations[71b] (**Figure 2.4**). The present study also reveals an increased

concentration of 2',3'-cAMP and 2',3'-cGMP compared to 2',3'-cCMP, and -cUMP in exponential and stationary phase cultures (**Figure 2.4**). The elevated 2',3'-cAMP level could be due to poly-A polymerase (PAPase) activity, as 3'-polyadenylation of mRNA facilitates exonucleolytic degradation in bacteria.[72] The different 2',3'-cNMP concentration ratios observed in exponential and stationary phase *E. coli* cultures (**Figure 2.4**) cannot be explained by preferential activity of RNase I because the enzyme does not display strong sequence or nucleobase specificity *in vitro*.[20] These results allude to a complex regulation of 2',3'-cNMP metabolism, which likely intersects with processes governing other nucleotide levels. Understanding the regulation of 2',3'-cNMP concentrations will require further investigation of growth-dependent relationships between 2',3'-cNMP, 3'-NMP, and other nucleotide/nucleoside concentrations.

Recent studies indicate that changes in nucleoside and nucleotide metabolism can alter biofilm formation,[61] demonstrating the importance of exploring 2',3'-cNMP pools in the context of overall nucleotide metabolism. This work demonstrates that decreasing 2',3'-cNMP levels increases biofilm formation in *E. coli* (**Figure 2.8B** and **Supplementary Figure S2.5**), likely due to upregulated production of curli (**Figure 2.9** and **Supplementary Table S2.4**), the major protein constituent of biofilms. Intriguingly, pyrimidine auxotrophy impairs synthesis of curli fibers, and conditions favoring UMP synthesis *via* pyrimidine salvage, as opposed to *de novo* UMP biosynthesis, also modify the biofilm matrix by increasing cellulose production.[61a] The altered cellulose synthesis in different *E. coli* pyrimidine auxotrophs is due, at least in part, to upregulated c-di-GMP synthesis by the DGC YedQ, suggesting that abnormal pyrimidine concentrations impact c-di-GMP signaling.[61a] Furthermore, inhibition of the essential purine biosynthesis enzyme PurH attenuates biofilm formation in *E. coli*, perhaps due to modulation of local c-di-GMP concentrations.[61b] In addition, the pyrimidine antimetabolite cancer drug 5-fluorouracil inhibits *E.*

coli biofilm formation by up-regulating expression of AriR, a transcriptional repressor of biofilm-related genes.[61c] Though some of the effectors involved in these processes have been identified in certain bacterial species, such as the CytR transcription factor in *Vibrio cholera* that de-represses pyrimidine metabolic genes in response to cytidine [61d], additional mechanistic details of the pathways connecting nucleoside/nucleotide levels to biofilm formation remain elusive. Notably, published data have shown that CytR is not directly involved in modulating pyrimidine-dependent biofilm phenotypes in *E. coli*,[61a] suggesting that additional unknown factors mediate this process in certain species. These findings suggest that disrupting normal 2',3'-cNMP regulation may alter biofilm formation by perturbing primary nucleotide/nucleoside metabolism, perhaps ultimately impacting c-di-GMP signaling. Although the total c-di-GMP concentration does not differ between WT and Δ*rna E. coli* (**Figure 2.8C**), it remains possible that cells lacking RNase I and 2',3'-cNMPs exhibit dysregulated levels of spatially isolated c-di-GMP pools, as local concentrations of this second messenger mediate biofilm formation.[73] Alternatively, the aberrant biofilm phenotype observed in Δ*rna* and in WT *E. coli* expressing CNPase (**Figure 2.8B**) potentially could be elicited by a novel second messenger signaling pathway mediated directly by 2',3'-cNMPs, as the micromolar 2',3'-cNMP concentrations in exponential phase *E. coli* cultures are similar to the basal level of 3',5'-cAMP,[71a,74] a canonical second messenger.

Ongoing work seeks to investigate potential 2',3'-cNMP-mediated signal transduction and elucidate the roles of the different 2',3'-cNMPs in modulating bacterial phenotypes. Gene expression data reported herein demonstrate that *E. coli* lacking RNase I exhibit aberrant expression of several transcripts relevant to biofilm production (**Figure 2.9** and **Supplementary Table S2.4**). Published phenotypic and deep sequencing investigations using *E. coli* deficient in RNase II, PNPase, or RNase R have linked these processive exoribonucleases to biofilm formation

via perturbation of biofilm-associated transcripts.[36,75] Although the mechanistic intricacies of the uniquely altered transcriptome in these different RNase mutants remain ambiguous, RNase II, PNPase, and RNase R directly impact transcript half-life,[76] suggesting that altered mRNA decay is directly influencing the biofilm phenotype. Conversely, previous work has shown that *rna* deletion does not directly affect transcript stability,[57] and the present study demonstrates that RNase I deletion does not perturb global c-di-GMP or pGpG levels (**Figure 2.8C**). These data allude to a more complex regulatory mechanism involving 2',3'-cNMPs, which is further demonstrated by the finding that inducing hydrolysis of 2',3'-cNMPs upregulates biofilm formation in WT cells expressing RNase I (**Figure 2.8B**). Another possible explanation for the altered transcriptome in the *rna* mutant is accumulation of short oligoRNAs, which can mis-prime RNA polymerase during transcription. In fact, this phenomenon is precedented in *P. aeruginosa* lacking oligoRNase,[77] an enzyme which exhibits a similar capacity for the degradation of short oligoribonucleotides. However, this pseudomonad lacks RNase I, whereas *E. coli* encodes both RNase I and oligoRNase. To further complicate this comparison, oligoRNase is dispensable in *P. aeruginosa*, yet it is essential in *E. coli*,[78] suggesting that *E. coli* oligoRNase performs unidentified functions that cannot be fulfilled by RNase I, despite the similar capacity of the two enzymes to degrade short oligoribonucleotides.[20,38] Future studies will investigate the possible buildup of oligoRNAs in *E. coli* deficient in RNase I.

The present study quantifies 2',3'-cNMPs in *E. coli*, demonstrates that RNase I generates 2',3'-cNMPs from mRNA and rRNA degradation, and identifies a role for 2',3'-cNMPs in regulating biofilm formation. The identification of RNase I as the enzyme responsible for generating 2',3'-cNMP pools provides the first insights into the phenotypic consequences of aberrant 2',3'-cNMP concentrations and RNase I levels on biofilm formation in bacteria.

Additional experiments are in progress to elucidate the mechanisms that control the relative concentration ratios of 2',3'-cNMP pools and the link to biofilm formation in *E. coli* and in other bacterial taxa. Given the importance of biofilm formation in the pathogenesis of numerous bacterial species,[79] elucidation of the mechanism(s) by which RNase I and 2',3'-cNMPs alter biofilm formation will provide insight into new methods to alter bacterial phenotypes.

2.4 Materials and Methods

Bacterial strains, plasmids, general culture conditions, chemicals, and statistical analyses

The *E. coli* strain BW25113 (wild-type, WT) (*lacI*^{q} *rrnB*$_{T14}$ Δ*lacZ*$_{WJ16}$ *hsdR514* Δ*araBAD*$_{AH33}$ Δ*rhaBAD*$_{LD78}$)[53] and Keio deletion mutant *rna*::kan^R (deficient in RNase I, Δ*rna*) in the BW25113 strain background[54] were used for all studies, unless specified otherwise. The *rna*::kan^R genotype of Δ*rna* was validated using locus-specific PCR amplifications (**Supplementary Figure S2.13, Supplementary Protocol S2.1,** and **Supplementary Table S2.1**). The pKT-*CNP* plasmid was generated by subcloning the catalytic domain (corresponding to the final 221 amino acid residues) of the *Rattus norvegicus CNP* gene[80] (UniProtKB-P13233; codon-optimized for *E. coli*; synthesized by GenScript) into the pKT vector[81] *via* double digest with restriction enzymes *Nde*I and *Spe*I, placing the gene under control of the TetA promoter (inducible with anhydrotetracycline). A catalytically-inactive variant of CNPase (H73L/H152L, numbering based on catalytic domain)[80] was generated *via* QuikChange mutagenesis (for primer sequences, see **Supplementary Table S2.2**). To construct plasmid pACYC-*noRBS-mRNA*, polymerase incomplete primer extension (PIPE) cloning was utilized.[82] To this end, the pACYCDuet-1 vector (EMD Millipore) was amplified by polymerase chain reaction (PCR) to delete both multiple cloning sites (including the ribosome binding site, T7 promoters, and T7 terminator). The 162-bp *noRBS-mRNA* insert containing its own T7 promoter and T7 terminator (54% GC, purchased as a

gBlock fragment from Integrated DNA Technologies) was PCR amplified to install 5'- and 3'-terminal regions complementary to the pACYC vector PCR product for PIPE cloning into the vector (for detailed PIPE cloning procedure and insert sequence, see **Supplementary Protocol S2.2** and **Supplementary Table S2.3**). T7-mediated expression was required because genes lacking a ribosome binding site are poorly transcribed by *E. coli* RNA polymerase.[83] Plasmid pCA24N-*rna* was obtained from the ASKA collection.[84] For bacterial growth, isolated colonies from Luria Broth (LB)-agar plates were cultured overnight at 37°C with 225-240 rpm shaking in 3 mL of M9 minimal medium (supplemented with 0.4% glucose and 0.2% casamino acids), unless otherwise noted. The resulting starter culture then was inoculated 1:100 into 10 mL of the same medium in 50-mL Celltreat® conical tubes (sterile, polypropylene; lids left loose for gas exchange) and incubated under the aforementioned conditions, unless specified otherwise. Kanamycin, chloramphenicol, and carbenicillin were used at working concentrations of 25, 30, and 100 μg mL^{-1}, respectively. Prior to 2',3'-cNMP extraction, cells were harvested by centrifugation at 3000g at 20°C for 10 min, frozen in liquid N_2, and stored at -80°C, unless otherwise noted. Analytical standards of adenosine 2',3'-cyclic monophosphate (2',3'-cAMP) and cytidine 2',3'-cyclic monophosphate (2',3'-cCMP) (monosodium salts) were purchased from Carbosynth (Berkshire, UK); standards of guanosine 2',3'-cyclic monophosphate monosodium salt (2',3'-cGMP), uridine 2',3'-cyclic monophosphate monosodium salt (2',3'-cUMP), cyclic dimeric-3':5'-guanosine monophosphate sodium salt (c-di-GMP), and 5'-phosphoguanylyl-3':5'-guanosine sodium salt (pGpG) were purchased from BioLog (Bremen, Germany). The sodium salt of 8-bromo adenosine 3',5'-cyclic monophosphate (8-Br 3',5'-cAMP) was obtained from Sigma-Aldrich. Adenosine 3'-monophosphate (3'-AMP) was purchased from Sigma-Aldrich as the free acid. All data represent at least three biological replicates. Statistical significance was evaluated using a two-sample *t*-test,

where equal or unequal variance was assessed using an *F*-test. Data were considered statistically-significant for $P < 0.05$.

Extraction of 2',3'-cNMPs

Aliquots (10-mL) were harvested from exponential-phase WT cultures (OD_{600} ~0.4-0.6) and stationary-phase cultures (16 or 24 h post-inoculation) by centrifugation. For 2',3'-cNMP extraction, frozen cell pellets were suspended in 500 µL of ice-cold acetonitrile/methanol/water (2/2/1, v/v/v), as previously described.[2b] The cells were lysed by sonication on ice and subsequently centrifuged at 4°C at 3000g for 10 min. The lysate was concentrated to dryness using a vacuum centrifuge and resuspended in 250 µL of sodium phosphate buffer (50 mM, pH 7.4) containing 0.5 µM 8-Br 3',5'-cAMP as internal standard. The extracts were centrifuged at 12000g for 30 min at 4°C and transferred to an LC-MS autosampler vial.

Quantification of 2',3'-cNMPs

Quantification of 2',3'-cNMPs was performed *via* an internal standard (IS) method, using 8-Br 3',5'-cAMP as the IS. Calibration curves for 2',3'-cAMP, -cCMP, -cGMP, and -cUMP analytes were constructed by plotting the peak area ratio of 2',3'-cNMP/IS against the concentration ratio of cNMP/IS, as described previously.[2b] 2',3'-cNMP concentrations were adjusted based on the recovery efficiency of each analyte (**Supplementary Figure S1**) and normalized to cell density. The concentration of IS was 0.5 µM in all samples for calibration. The concentrations of authentic 2',3'-cNMP analytes ranged from 0.02 – 20 µM (a range over which the analytical response remained linear). A linear regression model was used to generate the calibration curves. All nucleotide concentrations in stock solutions were determined by UV-Vis spectrophotometry (Cary Series, Agilent Technology, Santa Clara, CA, USA).

Extraction of c-di-GMP and pGpG

WT *E. coli* were cultured overnight (18 h) at room temperature without shaking to late-exponential/early stationary phase (OD_{600} ~0.7-1 A). The nucleotides were extracted essentially as described previously.[85] The protocol was performed analogously to the 2',3'-cNMP extraction described above, except that cell pellets were suspended in 0.5 mL of ice-cold sodium phosphate buffer (50 mM, pH 7.4) with 1 mM EDTA (0.05 mL of extraction buffer added per 1 mL of bacterial culture harvested).

Quantification of c-di-GMP and pGpG

Quantification of c-di-GMP and pGpG was performed using an IS method, in analogy to that detailed above for 2',3'-cNMP quantification. The concentration of IS was 0.1 µM in all samples for calibration. The concentrations of authentic c-di-GMP and pGpG analytes ranged from 0.0125 – 0.2 µM (a range over which the analytical response remained linear).

LC-MS/MS parameters

The LC-MS/MS methodology was performed as previously described,[2b] with minor modifications. A Thermo Electron LTQ-FTMS was employed for sample analysis. Chromatographic analysis was performed using a Shimadzu autosampler and a Dionex Ultimate 3000 dual gradient pump. LC-MS instrumentation was controlled by Xcalibur and DCMSlink software (Thermo Scientific). Samples were separated using a reversed-phase Leapsil C18 column (2.7 µm, 150 x 2.1 mm) (Dikma Technologies, Inc; Lake Forest, CA, USA). The mobile phase consisted of water with 0.1% formic acid (A) and methanol with 0.1% formic acid (B). The flow rate was 0.3 mL/min and the following chromatography program was employed: 0% B from 0 to 4 min, then a gradient from 0 to 1.5% B from 4 to 15 min, followed by a gradient from 1.5 to 8% B over 15 to 20 min, followed

by holding at 8% B from 20 to 25 min, then a gradient from 8 to 15% B from 25 to 28 min, followed by holding at 15% B from 28 to 35 min, and finally a gradient back to 0% B from 35 to 35.1 min. The column was re-equilibrated by holding at 0% B from 35.1 to 45 min. This chromatography method separates 2',3'-cNMPs from the 3',5'-cNMP regioisomers (**Supplementary Figures S2.7-S2.11**).[2b] The column was washed after analysis of every 2-4 extracts using the following chromatographic method: a gradient from 0 to 100% B from 0 to 2 min, followed by holding at 100% from 2 to 10 min, then a gradient from 0% to 100% C (acetonitrile) from 10 to 12 min, followed by holding at 100% C from 12 to 20 min, followed by a final gradient from 0% to 100% A over 20 to 25 min. The column was re-equilibrated to 100% A from 25 to 40 min. 2',3'-cNMPs were quantified with 10 to 30 uL injections; pGpG and c-di-GMP were quantified *via* a 45 uL injection. Electrospray ionization was performed in positive-ion mode in the LTQ-FTMS using a capillary voltage of 35 V, a 5 kV needle voltage, a capillary temperature of 275°C, and a 110 V tube lens voltage. Samples were detected in the ion trap using a 1 amu isolation window, and a normalized collision energy of 35 eV. An activation Q of 0.250 was used, with an activation time of 30 ms. Nucleotides were detected based on the protonated parent ions and quantified using the protonated nucleobase fragment ions (**Supplementary Figure S2.11**). Peaks were integrated using Xcalibur software (Thermo Fisher).

Quantification of 2',3'-cNMPs in Δ*rna* expressing pCA24N-*rna*

Cultures of Δ*rna* harboring plasmid pCA24N-*rna* were cultured to OD_{600} ~0.1-0.2 and subsequently induced by addition of 10 μM IPTG. Incubation was continued to OD_{600} ~0.5-0.6; the cells were harvested and the 2',3'-cNMPs were extracted and quantified, as described above.

Cytoplasm/periplasm fractionation

Separation of cytoplasmic and periplasmic fractions was performed according to a published procedure, and efficiency of the fractionation procedure was evaluated using SDS-PAGE analysis as described previously[86] (**Supplementary Figure S2.2**). Samples collected during exponential growth were resuspended in 100 μL TSE buffer (200 mM Tris-HCl pH 7.8, 500 mM sucrose, 1 mM EDTA). After incubation on ice for 30 min, the suspension was centrifuged at 14000g at 4°C for 40 min. The supernatant (final periplasmic extract) was stored at -80°C until LC-MS/MS and the pellet (spheroplast) was stored at -80°C until 2',3'-cNMP extraction. Spheroplasts were extracted in the same way as outlined above for cell pellets.

Addition of exogenous 2',3'-cAMP

WT cultures (60-mL) were grown to OD$_{600}$ ~0.4-0.5 A in 250-mL glass Erlenmeyer flasks. Each culture then was split into two equal portions (one for 0.1 mM 2',3'-cAMP addition and one for 0.1 mM 3'-AMP addition). 10-mL samples were collected 20 min after addition of the nucleotides for 2',3'-cNMP extraction.

Quantification of 2',3'-cNMP levels following growth +/- casamino acids

WT *E. coli* were cultured in 10 mL of either M9 minimal (0.4% glucose, 0.2% casamino acids) or M9 minimal (0.4% glucose, 1.2% casamino acids) in 50-mL Celltreat® conical tubes (sterile, polypropylene). Upon reaching OD$_{600}$ ~0.4-0.6, 10-mL samples were harvested for 2',3'-cNMP extraction.

Chloramphenicol-mediated induction of RNA degradation

WT cultures (50-mL) were grown to early exponential phase in 250-mL glass Erlenmeyer flasks and split into two equal portions. One portion was treated with 200 μg mL^{-1} chloramphenicol,[60a]

and the other portion was treated with an equal volume of ethanol as a control. After incubation for 30 min, 10 mL were harvested from all cultures by centrifugation.

Total RNA quantification +/- chloramphenicol treatment

BW25113 were cultured as described above for **Chloramphenicol-mediated inhibition of mRNA degradation**. From these cultures, 1-mL samples were collected pre-chloramphenicol treatment and 30 min post-treatment by centrifugation at 12000g at 24°C for 5 min to quantify total intracellular RNA *via* the RNA*snap*™ procedure.[87] The cell pellets were suspended in 300 µL of RNA*snap*™ extraction solution (95% formamide, 18 mM EDTA, 0.025% SDS, 1% β-mercaptoethanol) and incubated for 7 min in a 95°C sand bath. The samples were centrifuged at 14000g at 24°C for 10 min and the 260 nm absorbance (A260) of the supernatant was quantified using a NanoDrop™ 1000. The total RNA concentration was calculated from the A260 using an extinction coefficient of 0.025 µg mL^{-1} cm^{-1} and normalized to the OD$_{600}$- and volume-dependent cell density of each sample.[88]

λDE3 lysogenization of BW25113 WT

The WT BW25113 strain was lysogenized using the λDE3 Lysogenization Kit (Novagen) according to the manufacturer's instructions.

Overexpression of non-translatable mRNA pACYC-*noRBS-mRNA*

BW25113 (DE3) and BW25113 (DE3) harboring plasmid pACYC-*noRBS-mRNA* were cultured to early exponential phase (OD$_{600}$ ~0.2-0.3). All cultures then were treated with 0.4 mM isopropyl-*β*-D-thiogalactopyranoside (IPTG) to induce expression of the non-translatable mRNA. Upon reaching an OD$_{600}$ of 0.5-0.6, the cultures were harvested by centrifugation.

2',3'-cNMP quantification in shaking and static cultures

WT cultures (10-mL) were grown in 50-mL Celltreat® conical tubes (sterile, polypropylene) at either 37°C with 225 rpm shaking to mid-logarithmic phase (OD_{600} ~0.4-0.6) or at room temperature without shaking overnight to allow biofilm formation. Cells were harvested from 9 mL of culture and lysed for 2',3'-cNMP quantification as detailed above. Biofilm formation was qualitatively confirmed in the static cultures by crystal violet staining, in analogy to a published procedure.[89]

Assessment of metabolic state in shaking and static cultures

WT *E. coli* were grown in 100 μL cultures in a 96-well microplate (Corning Costar, sterile, non-treated, polystyrene). One set of cultures was incubated at 37°C with shaking to exponential phase, while the other set was incubated at room temperature without shaking for 24 h. The metabolic state then was assessed using the XTT Cell Proliferation Kit II (Roche) according to the manufacturer's protocol with minor modification. Upon reaching the desired cell density, the XTT labeling mixture (50 μL) was added to each culture and the 490 nm absorbance was immediately recorded using a microplate reader. The A_{490} was normalized to cell density using the OD_{600} of each culture.

Biofilm quantification by Congo Red staining

Congo Red assays were conducted as previously described.[90] Individual colonies of BW25113 WT and Δ*rna* from LB-agar plates were inoculated into 5 mL LB and cultured overnight in 15-mL plastic culture tubes. In addition, WT *E. coli* harboring plasmid pKT-*CNP* or inactive control pKT-*CNP-inact* were cultured in the same way. The overnights were inoculated 1:50 into 7 mL of YESCA (1% casamino acids, 0.12% yeast extract) containing 0.0025% Congo red in 50-mL

Celltreat® conical tubes (sterile, polypropylene) (lids left loose for gas exchange). After reaching an OD_{600} ~0.3-0.4, 1 mL of each culture was transferred to a 1.6-mL Eppendorf tube and either treated with vehicle or with 25 ng mL^{-1} anhydrotetracycline (AHT) to induce expression. The cultures were incubated for 48 h at room temperature without shaking (lids left open and tubes were loosely covered in plastic wrap and foil). For biofilm quantification, samples were centrifuged at 12000g for 15 min and 200 μL of supernatant were transferred to a 96-well microplate (Corning Costar; sterile, non-treated, polystyrene). The absorbance at 500 nm was recorded using a microplate reader. For normalization, each culture was disturbed by pipetting and 200 μL were transferred to a 96-well microplate prior to recording the OD_{600} using a microplate reader.

Biofilm quantification by crystal violet staining

Cultures of WT and Δ*rna* (2-mL) were incubated in 24-well Corning Costar microplates (sterile, non-treated, polystyrene) for 24 h at room temperature without shaking. Biofilm formation was quantified by crystal violet staining according to a published procedure with minor modification.[89] Non-adherent cells were poured out and the microplate was gently submerged twice in a beaker of water. A 0.1% aqueous solution of crystal violet (2.5 mL) was added to each well and the microplate was incubated at room temperature for 15 min. The crystal violet solution was poured out and the microplate was gently submerged three times in a beaker of water to remove residual crystal violet (blotting the plate on a stack of paper towels after each wash). The plate was allowed to dry overnight at room temperature. The crystal violet in each well was dissolved by addition of 3 mL of 30% aqueous acetic acid, and the 570 nm absorbance was measured using a microplate reader and normalized to colony-forming units (CFU; quantified by drop plating, according to published procedure[91]).

Quantification of biofilm-related gene transcript levels

Analysis of mRNA transcript levels for genes related to *E. coli* biofilm formation was performed by the Emory Integrated Genomics Core and analyzed by the Emory Integrated Computational Core. Six *E. coli* pellets (three biological replicates of WT and three of Δ*rna*) were submitted for extraction and expression profiling on the Affymetrix *E. coli* Genome 2.0 Array. RNA was extracted using the Qiagen miRNEasy Kit with the optional on-column DNase treatment. Cells were lysed using 700 μL Qiazol and 100 mg acid-washed beads (150-600 μm) on the Qiagen tissue lyser at 30 Hz for 5 min. RNA was eluted in 30 μL of nuclease free water. The nucleic acid concentration then was determined using a Nanodrop 1000, and sample profiles were assessed on the Agilent 2100 using the RNA 6000 Nano assay.

Whole-Transcript Expression Analysis (Gene ST Arrays) was performed as follows. RNA (10 ng) was processed according to the GeneChip® WT Pico Reagent Kit protocol. Labeled complementary DNA (cDNA) was hybridized to the *E. coli* Genome 2.0 microarray for 16-18 h at 45°C. Hybridized microarrays were washed and stained on an Affymetrix GeneChip 450 fluidics station using the appropriate chip-dependent fluidics script. Intensity data was extracted using an Affymetrix 7G scanner and the Command Console software suite.

The obtained expression data from the microarray experiment were analyzed using 'limma' package in R/Bioconductor (http://www.r-project.org). The raw data were $\log_2$ transformed and intensity was normalized between samples using Robust Multi-array Average (RMA). The differentially expressed genes were identified on the basis of Benjamini-Hochberg (BH) multiple test adjusted *P* values (i.e. FDR) and fold changes (the increase in number of gene copies). Genes with an FDR value <0.05 and $\log_2$ fold change ≥1.0 were considered significantly differentially expressed.

Gene expression data obtained from the microarray experiment have been submitted to ArrayExpress at EMBL-EBI (http://www.ebi.ac.uk/arrayexpress/) under accession number E-MTAB-6095.

Quantitative reverse transcription PCR (RT-qPCR) analysis

Total RNA was extracted using the Direct-zol™ RNA MiniPrep Kit (Zymo Research, Irvine, CA) according to the manufacturer's instructions with the optional on-column DNase treatment. Subsequently, total RNA (1 mg) was used as template to synthesize cDNA with the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Foster City, CA). Primers for all assays were designed using Primer 3[92] (also see http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi). For primer sequences, see **Supplementary Table S2.6**. Melting curve analysis was performed to insure single-product amplification for all primer pairs. Real time PCR was performed on the BioRad CFX384 Real Time System (BioRad, Hercules, CA) using assays specific to the genes of interest. Each reaction well contained 5 mL of PowerUp™ SYBR Green Master Mix (Applied Biosystems), cDNA equivalent to 20 ng of total RNA and 250 nM each of forward and reverse amplification primers in a final reaction volume of 10 mL. Cycling conditions were as follows: 95°C for 10 minutes for polymerase activation, followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. Data analysis was performed using CFX Manager software (BioRad, version 3.1). The experimental Cq (cycle quantification) was calibrated against the endogenous control products DNA-directed RNA polymerase subunit beta (*rpoC*). Samples were analyzed for relative gene expression by the ΔΔCt method.[93]

2.5 References

(1) Van Damme, T.; Blancquaert, D.; Couturon, P.; Van der Straeten, D.; Sandra, P.; Lynen, F. *Phytochemistry* **2014**, *103*, 59-66.

(2) (a) Bahre, H.; Kaever, V. *J. Chromatogr. B* **2014**, *964*, 208-211; (b) Jia, X.; Fontaine, B. M.; Strobel, F.; Weinert, E. E. *Biomolecules* **2014**, *4*, 1070-1092; (c) Ren, J.; Mi, Z.; Stewart, N.; Jackson, E. *J. Pharmacol. Exp. Ther.* **2009**, *328*, 855-865.

(3) Jackson, E.; Ren, J.; Mi, Z. *J. Biol. Chem.* **2009**, *284*, 33097-33106.

(4) Jackson, E. *Am. J. Physiol. Renal Physiol.* **2011**, *301*, F1160-F1167.

(5) Azarashvili, T.; Krestinina, O.; Galvita, A.; Grachev, D.; Baburina, Y.; Stricker, R.; Evtodienko, Y.; Reiser, G. *Am. J. Physiol. Cell Physiol.* **2009**, *296*, C1428-C1439.

(6) (a) Thompson, R. *Biochem. Soc. Trans.* **1992**, *20*, 621-626; (b) Vogel, U.; Thompson, R. *J. Neurochem.* **1988**, *50*, 1667-1677.

(7) (a) Jackson, E. K.; Gillespie, D. G.; Mi, Z. C.; Cheng, D. M.; Bansal, R.; Janesko-Feldman, K.; Kochanek, P. M. *Am. J. Physiol. Renal Physiol.* **2014**, *307*, F14-F24; (b) Verrier, J.; Jackson, T.; Gillespie, D.; Janesko-Feldman, K.; Bansal, R.; Goebbels, S.; Nave, K.; Kochanek, P.; Jackson, E. *Glia* **2013**, *61*, 1595-1606.

(8) (a) Anraku, Y. *J. Biol. Chem.* **1964**, *239*, 3412-3419; (b) Anraku, Y. *J. Biol. Chem.* **1964**, *239*, 3420-3424; (c) Nagata, M.; Kaito, C.; Sekimizu, K. *J. Biol. Chem.* **2008**, *283*, 2176-2184; (d) Podzelinska, K.; He, S.; Wathier, M.; Yakunin, A.; Proudfoot, M.; Hove-Jensen, B.; Zechel, D.; Jia, Z. *J. Biol. Chem.* **2009**, *284*, 17216-17226; (e) Shin, D.; Proudfoot, M.; Lim, H.; Choi, I.; Yokota, H.; Yakunin, A.; Kim, R.; Kim, S. *Proteins-Structure Function and Bioinformatics* **2008**, *70*, 1000-1009.

(9) Rao, F.; Qi, Y.; Murugan, E.; Pasunooti, S.; Ji, Q. *Biochem. Biophys. Res. Commun.* **2010**, *398*, 500-505.

(10) Jackson, E.; Ren, J.; Cheng, D.; Mi, Z. *Am. J. Physiol. Renal Physiol.* **2011**, *301*, F565-F573.

(11) Luhtala, N.; Parker, R. *Trends Biochem. Sci.* **2010**, *35*, 253-259.

(12) Wade, H. *Biochem. J* **1961**, *78*, 457-472.

(13) Bordeleau, E.; Oberc, C.; Ameen, E.; da Silva, A.; Yan, H. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 4520-4522.

(14) Zhang, Y.; Agrebi, R.; Bellows, L.; Collet, J.; Kaever, V.; Grundling, A. *J. Biol. Chem.* **2017**, *292*, 313-327.

(15) Liu, A.; Yu, Y.; Sheng, Q.; Zheng, X.; Yang, J.; Li, P.; Shi, M.; Zhou, B.; Zhang, Y.; Chen, X. *ACS Chem. Biol.* **2016**, *11*, 2414-2419.

(16) Thompson, J.; Venegas, F.; Raines, R. *Biochemistry* **1994**, *33*, 7408-7414.

(17) Yamaguchi, Y.; Park, J. H.; Inouye, M. *Annual Review Genetics, Vol 45* **2011**, *45*, 61-79.

(18) Elson, D. *Biochim. Biophys. Acta* **1959**, *36*, 372-386.

(19) (a) Abrell, J. *Arch. Biochem. Biophys.* **1971**, *142*, 693-700; (b) Meador, J.; Cannon, B.; Cannistraro, V.; Kennell, D. *Eur. J. Biochem.* **1990**, *187*, 549-553; (c) Neu, H.; Heppel, L. *J. Biol. Chem.* **1964**, *239*, 3893-3900; (d) Spahr, P.; Hollingworth, B. *J. Biol. Chem.* **1961**, *236*, 823-831.

(20) Cannistraro, V.; Kennell, D. *J. Bacteriol.* **1991**, *173*, 4653-4659.

(21) Meador, J.; Kennell, D. *Gene* **1990**, *95*, 1-7.

(22) (a) Beppu, T.; Arima, K. *J. Bacteriol.* **1969**, *98*, 888-897; (b) Kaplan, R.; Apirion, D. *J. Biol. Chem.* **1974**, *249*, 149-151; (c) Maruyama, H.; Mizuno, D. *Biochim. Biophys. Acta* **1965**, *108*, 593-&; (d) Wade, H.; Robinson, H.; Lovett, S. *Biochem. J* **1964**, *93*, 121-128.

(23) Kitahara, K.; Miyazaki, K. *Nature Communications* **2011**, *2*, DOI: 10.1038/ncomms1553.

(24) Anraku, Y.; Mizuno, D. *J. Biochem.* **1967**, *61*, 81-88.

(25) (a) Abel, S.; Nurnberger, T.; Ahnert, V.; Krauss, G.; Glund, K. *Plant Physiol.* **2000**, *122*, 543-552; (b) Nurnberger, T.; Abel, S.; Jost, W.; Glund, K. *Plant Physiol.* **1990**, *92*, 970-976.

(26) Mudd, E.; Krisch, H.; Higgins, C. *Mol. Microbiol.* **1990**, *4*, 2127-2135.

(27) (a) Blum, E.; Py, B.; Carpousis, A.; Higgins, C. *Mol. Microbiol.* **1997**, *26*, 387-398; (b) Miczak, A.; Kaberdin, V.; Wei, C.; LinChao, S. *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 3865-3869; (c) PY, B.; CAUSTON, H.; MUDD, E.; HIGGINS, C. *Mol. Microbiol.* **1994**, *14*, 717-729.

(28) Py, B.; Higgins, C.; Krisch, H.; Carpousis, A. *Nature* **1996**, *381*, 169-172.

(29) Bernstein, J.; Lin, P.; Cohen, S.; Lin-Chao, S. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 2758-2763.

(30) Liou, G.; Jane, W.; Cohen, S.; Lin, N.; Lin-Chao, S. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 63-68.

(31) (a) Gao, J.; Lee, K.; Zhao, M.; Qiu, J.; Zhan, X.; Saxena, A.; Moore, C.; Cohen, S.; Georgiou, G. *Mol. Microbiol.* **2006**, *61*, 394-406; (b) Lee, K.; Zhan, X.; Gao, J.; Ji, Q.; Feng, Y.; Meganathan, R.; Cohen, S.; Georgiou, G. *Cell* **2003**, *114*, 623-634.

(32) Ow, M.; Kushner, S. *Genes Dev.* **2002**, *16*, 1102-1115.

(33) (a) Misra, T. *Fed. Proc.* **1979**, *38*, 322-322; (b) Sulthana, S.; Basturea, G.; Deutscher, M. *RNA* **2016**, *22*, 1163-1171.

(34) Ikeda, Y.; Yagi, M.; Morita, T.; Aiba, H. *Mol. Microbiol.* **2011**, *79*, 419-432.

(35) Mohanty, B.; Kushner, S.; Gottesman, S. *Annual Review of Microbiology, Vol 70* **2016**, *70*, 25-44.

(36) Pobre, V.; Arraiano, C. *BMC Genomics* **2015**, *16*.

(37) (a) Klee, C.; Singer, M. *J. Biol. Chem.* **1968**, *243*, 923-927; (b) Nossal, N.; Singer, M. *J. Biol. Chem.* **1968**, *243*, 913-922.

(38) Niyogi, S.; Datta, A. *J. Biol. Chem.* **1975**, *250*, 7307-7312.

(39) Mohanty, B.; Kushner, S. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 11966-11971.

(40) Mohanty, B.; Kushner, S. *Mol. Microbiol.* **1999**, *34*, 1094-1108.

(41) Cheng, Z.; Deutscher, M. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6388-6393.

(42) Jiang, X.; Belasco, J. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9211-9216.

(43) Baker, K.; Mackie, G. *Mol. Microbiol.* **2003**, *47*, 75-88.

(44) Nudler, E.; Gottesman, M. *Genes Cells* **2002**, *7*, 755-768.

(45) Spickler, C.; Mackie, G. *J. Bacteriol.* **2000**, *182*, 2422-2427.

(46) Braun, F.; Le Derout, J.; Regnier, P. *EMBO J.* **1998**, *17*, 4790-4797.

(47) (a) Khemici, V.; Poljak, L.; Luisi, B.; Carpousis, A. *Mol. Microbiol.* **2008**, *70*, 799-813; (b) Lu, F.; Taghbalout, A. *J. Biol. Chem.* **2013**, *288*, 7241-7251; (c) Carabetta, V.; Silhavy, T.; Cristea, I. *J. Bacteriol.* **2010**, *192*, 3713-3721.

(48) Diestra, E.; Cayrol, B.; Arluison, V.; Risco, C. *Plos One* **2009**, *4*.

(49) Libby, E.; Roggiani, M.; Goulian, M. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 7445-7450.

(50) (a) Milon, P.; Tischenko, E.; Tomsic, J.; Caserta, E.; Folkers, G.; La Teana, A.; Rodnina, M.; Pon, C.; Boelens, R.; Gualerzi, C. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 13962-13967; (b) Ryals, J.; Little, R.; Bremer, H. *J. Bacteriol.* **1982**, *151*, 1261-1268.

(51) Kalia, D.; Merey, G.; Nakayama, S.; Zheng, Y.; Zhou, J.; Luo, Y.; Guo, M.; Roembke, B.; Sintim, H. *Chem. Soc. Rev.* **2013**, *42*, 305-341.

(52) Jacobs, A.; Blanchard, C.; Catherman, S.; Dunman, P.; Murata, Y. *Plos One* **2014**, *9*.

(53) Datsenko, K.; Wanner, B. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6640-6645.

(54) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K.; Tomita, M.; Wanner, B.; Mori, H. *Mol. Syst. Biol.* **2006**, *2*, doi: 10.1038/msb4100050.

(55) Beacham, I. *Int. J. Biochem.* **1979**, *10*, 877-883.

(56) Zhu, L. Q.; Gangopadhyay, T.; Padmanabha, K. P.; Deutscher, M. P. *J. Bacteriol.* **1990**, *172*, 3146-3151.

(57) Kivity-Vogel, T. E., D. *Biochim. Biophys. Acta* **1967**, *138*, 66-75.

(58) (a) Christensen, S.; Mikkelsen, M.; Pedersen, K.; Gerdes, K. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 14328-143233; (b) Christensen, S.; Pedersen, K.; Hansen, F.; Gerdes, K. *J. Mol. Biol.* **2003**, *332*, 809-819.

(59) Durfee, T.; Hansen, A.-M.; Zhi, H.; Blattner, F.; Jin, D. *J. Bacteriol.* **2008**, *190*, 1084-1096.

(60) (a) Lopez, P.; Marchand, I.; Yarchuk, O.; Dreyfus, M. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6067-6072; (b) Pato, M.; Bennett, P.; von Meyenburg, K. *J. Bacteriol.* **1973**, *116*, 710-718.

(61) (a) Garavaglia, M.; Rossi, E.; Landini, P. *PLOS ONE* **2012**, *7*, doi:10.1371/journal.pone.0031252; (b) Antoniani, D.; Rossi, E.; Rinaldo, S.; Bocci, P.; Lolicato, M.; Paiardini, A.; Raffaelli, N.; Cutruzzola, F.; Landini, P. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7325-7336; (c) Attila, C.; Ueda, A.; Wood, T. *Appl. Microbiol. Biotechnol.* **2009**, *82*, 525-533; (d) Haugo, A.; Watnick, P. *Mol. Microbiol.* **2002**, *45*, 471-483; (e) Ueda, A.; Attila, C.; Whiteley, M.; Wood, T. *Microb. Biotechnol.* **2009**, *2*, 62-74.

(62) Flemming, H.; Wingender, J. *Nat. Rev. Microbiol.* **2010**, *8*, 623-633.

(63) Wood, T.; Barrios, A.; Herzberg, M.; Lee, J. *Appl. Microbiol. Biotechnol.* **2006**, *72*, 361-367.

(64) (a) Cohen, D.; Mechold, U.; Nevenzal, H.; Yarmiyhu, Y.; Randall, T.; Bay, D.; Rich, J.; Parsek, M.; Kaever, V.; Harrison, J.; Banin, E. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11359-

11364; (b) Orr, M.; Donaldson, G.; Severin, G.; Wang, J.; Sintim, H.; Waters, C.; Lee, V. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E5048-E5057.

(65) Nicholson, A. W. *FEMS Microbiol. Rev.* **1999**, *23*, 371-390.

(66) Myllykoski, M.; Kursula, P. *BMC Research Notes* **2010**, *3*, DOI: 10.1186/1756-0500-1183-1112.

(67) Barnhart, M.; Chapman, M. *Annu. Rev. Microbiol.* **2006**, *60*, 131-147.

(68) Reichhardt, C.; Jacobson, A.; Maher, M.; Uang, J.; McCrate, O.; Eckart, M.; Cegelski, L. *PLOS ONE* **2015**, *10*.

(69) (a) Itoh, Y.; Rice, J.; Goller, C.; Pannuri, A.; Taylor, J.; Meisner, J.; Beveridge, T.; Preston, J.; Romeo, T. *J. Bacteriol.* **2008**, *190*, 3670-3680; (b) Wang, X.; Preston, J.; Romeo, T. *J. Bacteriol.* **2004**, *186*, 2724-2734.

(70) Proudfoot, M.; Kuznetsova, E.; Brown, G.; Rao, N.; Kitagawa, M.; Mori, H.; Savchenko, A.; Yakunin, A. *J. Biol. Chem.* **2004**, *279*, 54687-54694.

(71) (a) Bennett, B.; Kimball, E.; Gao, M.; Osterhout, R.; Van Dien, S.; Rabinowitz, J. *Nat. Chem. Biol.* **2009**, *5*, 593-599; (b) Buckstein, M.; He, J.; Rubin, H. *J. Bacteriol.* **2008**, *190*, 718-726.

(72) Carpousis, A. *Annu. Rev. Microbiol.* **2007**, *61*, 71-87.

(73) Sarenko, O.; Klauck, G.; Wilke, F.; Pfiffer, V.; Richter, A.; Herbst, S.; Kaever, V.; Hengge, R. *MBio* **2017**, *8*.

(74) Notley-McRobb, L.; Death, A.; Ferenci, T. *Microbiology* **1997**, *143*, 1909-1918.

(75) Carzaniga, T.; Antoniani, D.; Deho, G.; Briani, F.; Landini, P. *BMC Microbiol.* **2012**, *12*.

(76) (a) Cheng, Z.; Deutscher, M. *Mol. Cell* **2005**, *17*, 313-318; (b) Donovan, W.; Kushner, S. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 120-124.

(77) Goldman, S.; Sharp, J.; Vvedenskaya, I.; Livny, J.; Dove, S.; Nickels, B. *Mol. Cell* **2011**, *42*, 817-825.

(78) Ghosh, S.; Deutscher, M. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4372-4377.

(79) (a) Romling, U.; Galperin, M.; Gomelsky, M. *Microbiol. Mol. Biol. Rev.* **2013**, *77*, 1-52; (b) Hengge, R. *Nat. Rev. Microbiol.* **2009**, *7*, 263-273.

(80) Gravel, M.; Robert, F.; Kottis, V.; Gallouzi, I.; Pelletier, J.; Braun, P. *J. Neurosci. Res.* **2009**, *87*, 1069-1079.

(81) Neuenschwander, M.; Butz, M.; Heintz, C.; Kast, P.; Hilvert, D. *Nat. Biotechnol.* **2007**, *25*, 1145-1147.

(82) Klock, H.; Koesema, E.; Knuth, M.; Lesley, S. *Proteins-Structure Function and Bioinformatics* **2008**, *71*, 982-994.

(83) Adhya, S.; Gottesman, M. *Annu. Rev. Biochem* **1978**, *47*, 967-996.

(84) Kitagawa, M.; Ara, T.; Arifuzzaman, M.; Ioka-Nakamichi, T.; Inamoto, E.; Toyonaga, H.; Mori, H. *DNA Res.* **2005**, *12*, 291-299.

(85) Massie, J.; Reynolds, E.; Koestler, B.; Cong, J.; Agostoni, M.; Waters, C. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 12746-12751.

(86) Quan, S.; Hiniker, A.; Collet, J.; Bardwell, J. C. A. In *Bacterial Cell Surfaces: Methods and Protocols*; Delcour, A. H., Ed.; Springer Science+Business Media: New York, 2013; Vol. 966, p 359-366.

(87) Stead, M.; Agrawal, A.; Bowden, K.; Nasir, R.; Mohanty, B.; Meagher, R.; Kushner, S. *Nucleic Acids Res.* **2012**, *40*, doi:10.1093/nar/gks1680.

(88) Volkmer, B.; Heinemann, M. *PLOS ONE* **2011**, *6*, doi: 10.1371/journal.pone.0023126.

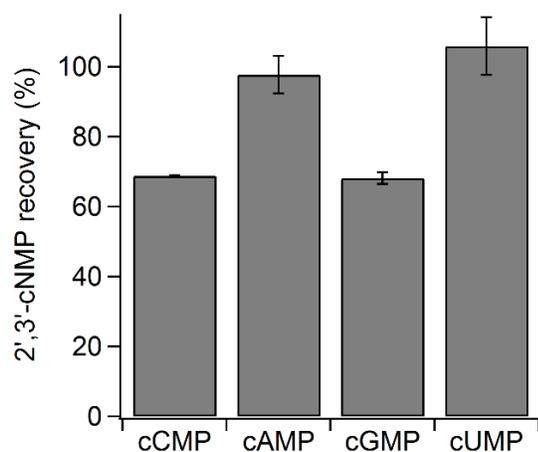(89) O'Toole, G. *J. Vis. Exp.* **2011**, doi: 10.3791/2437, doi: 10.3791/2437.

(90) Spiers, A.; Bohannon, J.; Gehrig, S.; Rainey, P. *Mol. Microbiol.* **2003**, *50*, 15-27.

(91) Chen, C.; Nace, G.; Irwin, P. *J. Microbiol. Methods* **2003**, *55*, 475-479.

(92) Ye, J.; Coulouris, G.; Zaretskaya, I.; Cutcutache, I.; Rozen, S.; Madden, T. *BMC Bioinformatics* **2012**, *13*.

(93) Pfaffl, M. *Nucleic Acids Res.* **2001**, *29*.

2.6 Supplementary Material

Supplementary Figure S2.1. Validation of 2',3'-cNMP extraction protocol. To evaluate the recovery of 2',3'-cNMP analytes, three aliquots from an exponential phase culture of *E. coli* BW25113 Δ*rna* were spiked with 10 μM of 2',3'-cAMP, -cCMP, -cGMP, and -cUMP. The samples then were subjected to the 2',3'-cNMP extraction protocol (see **Materials and Methods**; **section 2.4**). Unless otherwise specified, bars represent the mean of at least three replicates, and error bars denote the standard deviation.
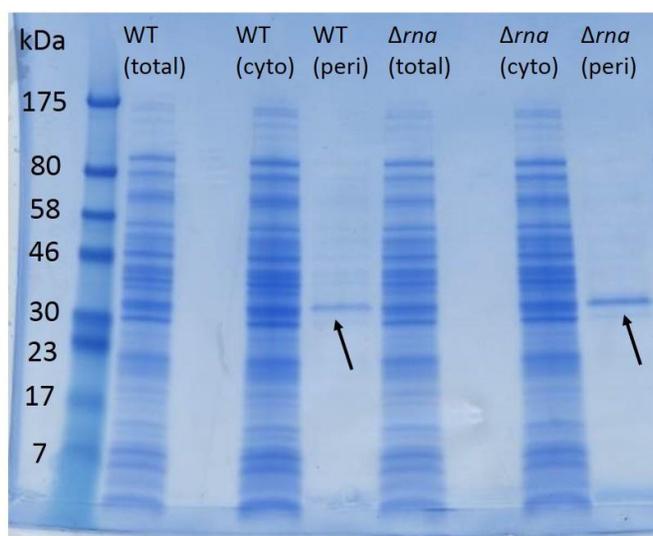

Supplementary Figure S2.2. SDS-PAGE validation of cytoplasmic/periplasmic fractionation. Exponential phase cultures of WT and Δ*rna* were subjected to tris/sucrose/EDTA osmotic shock, according to published procedure [1] (for details, see **Materials and Methods**; **section 2.4**). The arrow indicates enrichment of OmpF and OmpC in the periplasmic fraction. The multiple shared bands between the total lysate and cytoplasmic fraction, along with the absence of shared bands in the periplasmic fraction, indicates a clean osmotic shock fractionation.[1]
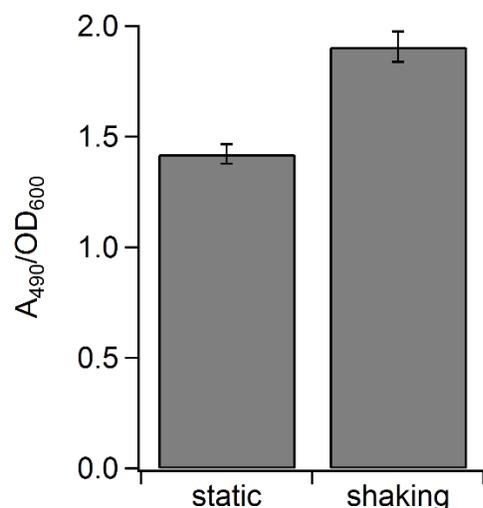
Supplementary Figure S2.3. Different 2',3'-cNMP levels in static *versus* shaking cultures are not a result of altered metabolism. BW25113 WT *E. coli* were cultured either at 37°C with shaking to exponential phase or at 20°C without shaking for 20 h, and the metabolic state was assessed spectrophotometrically by monitoring the *in vivo* reduction of the tetrazolium XTT (Roche) (for details, see **Materials and Methods**; **section 2.4**). The 1.3-fold reduction in metabolism of the static cultures relative to the shaking cultures is not sufficient to rationalize the 15-fold lower 2',3'-cNMP concentrations in the static cultures (see **Figure 2.7A** in main text).
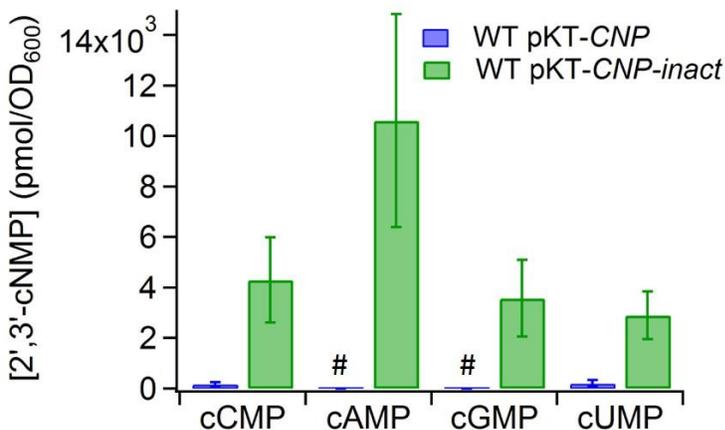
Supplementary Figure S2.4. Inducing *CNP* expression decreases 2',3'-cNMP concentrations *in vivo*. WT *E. coli* expressing either pKT-*CNP* or pKT-*CNP-inact* as a control were grown at room temperature without shaking and 2',3'-cNMP concentrations were quantified (for details, see **Materials and Methods**; **section 2.4**). Functional CNPase dramatically reduces 2',3'-cNMP levels, as compared to the catalytically-inactive control phosphodiesterase (# denotes value < LOQ).
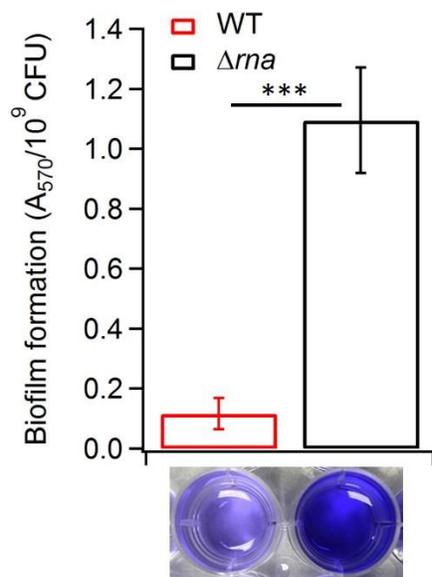
**Supplementary Figure S2.5.** Biofilm formation increases in *E. coli* lacking RNase I. Biofilm formation was quantified in static cultures of WT and Δ*rna E. coli* as previously described[2] (for details, see **Materials and Methods**; **section 2.4**; *** $P < 0.001$). Unless otherwise specified, statistical significance was evaluated using a two-sample *t*-test assuming unequal variances (variance assessed via an *F*-test). The image of the wells depicts crystal violet staining for each strain from a representative assay.
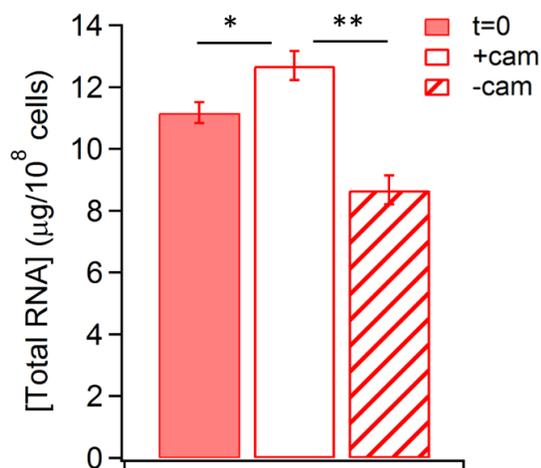
Supplementary Figure S2.6. Total RNA increases in *E. coli* treated with chloramphenicol. Total RNA was quantified via the RNA*snap*[TM] procedure[3] in WT cultures prior to treatment with chloramphenicol (cam) and 30 min post-treatment with either chloramphenicol (+cam) or ethanol as a control (-cam)[4] (for details, see **Materials and Methods**; **section 2.4**; * $P < 0.05$; ** $P < 0.01$).
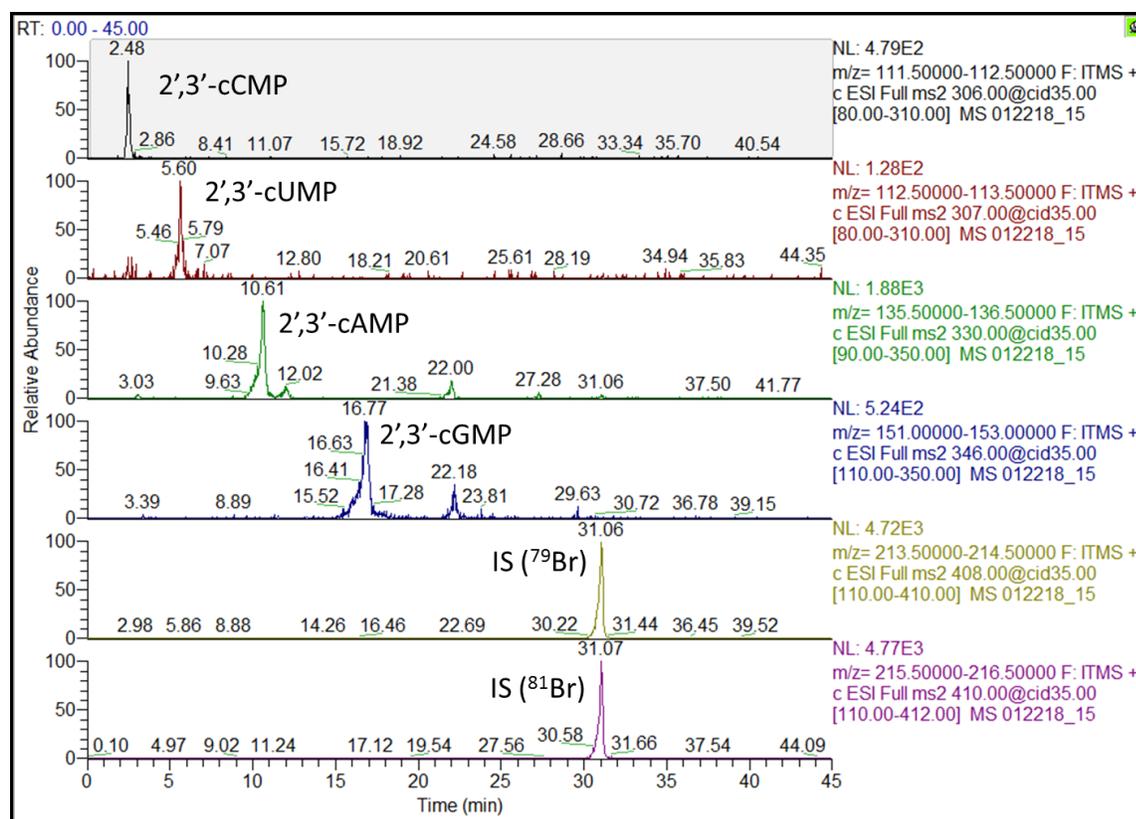
Supplementary Figure S2.7. Representative LC-MS/MS chromatograms. The chromatograms depict the retention time of each 2',3'-cNMP analyte and the internal standard (IS; 8-Br 3',5'-cAMP) from a representative WT sample (for details, see **Materials and Methods**; **section 2.4**). Nucleotides were detected *via* the protonated parent ion $[M+H]^+$ and quantified based on the protonated nucleobase cation.
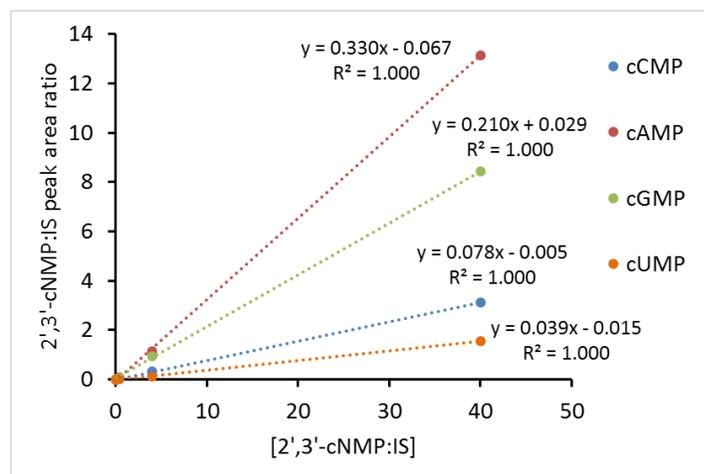

Supplementary Figure S2.8. Representative calibration curves used for LC-MS/MS-based quantification. The calibration curve for each of the 2',3'-cNMP analytes is plotted as the ratio of the 2',3'-cNMP:IS peak area against the ratio of the 2',3'-cNMP:IS concentration. The calibration curves were generated using samples containing 0.02, 0.2, 2, and 20 μM of each

2',3'-cNMP along with 0.5 µM 8-Br 3',5'-cAMP as internal standard (for details, see **Materials and Methods**; **section 2.4**).
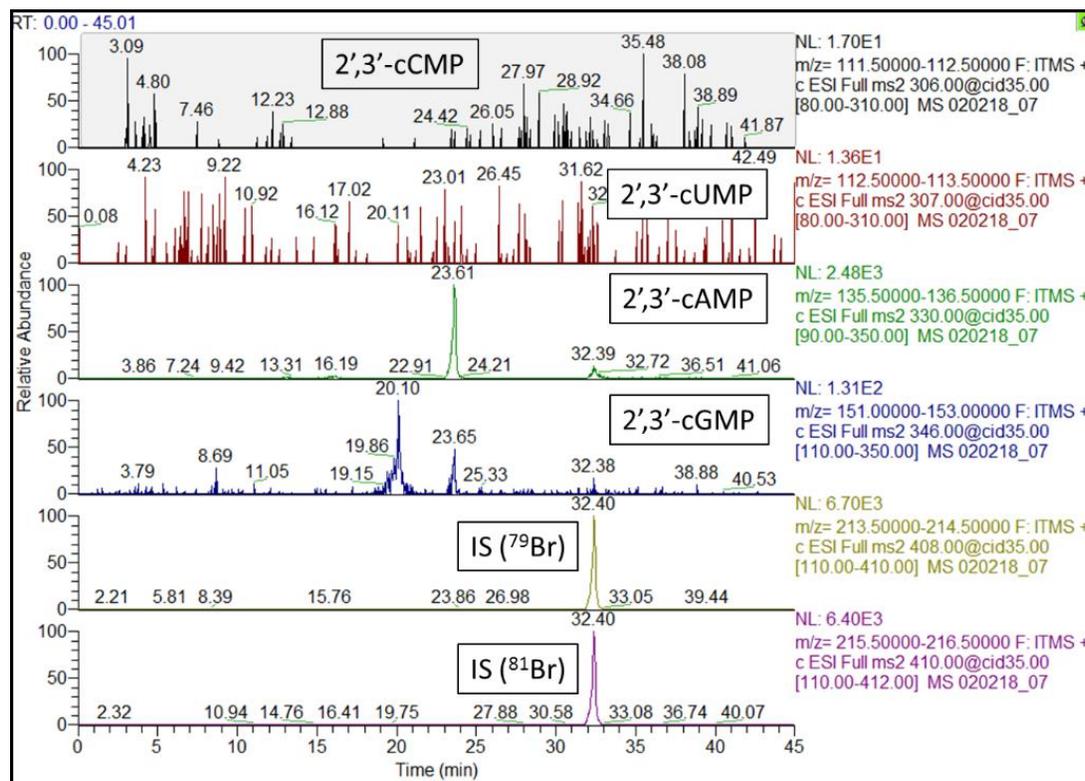
Supplementary Figure S2.9. 2',3'-cNMPs are undetectable in stationary phase WT cultures analyzed 16 h post-inoculation. Cultures of BW25113 WT were cultured for 16 h and subjected to 2',3'-cNMP extraction (for details, see **Materials and Methods**; **section 2.4**). 2',3'-cNMP concentrations were below the limit of detection (compare to control chromatograms in **Supplementary Figure S2.7**).
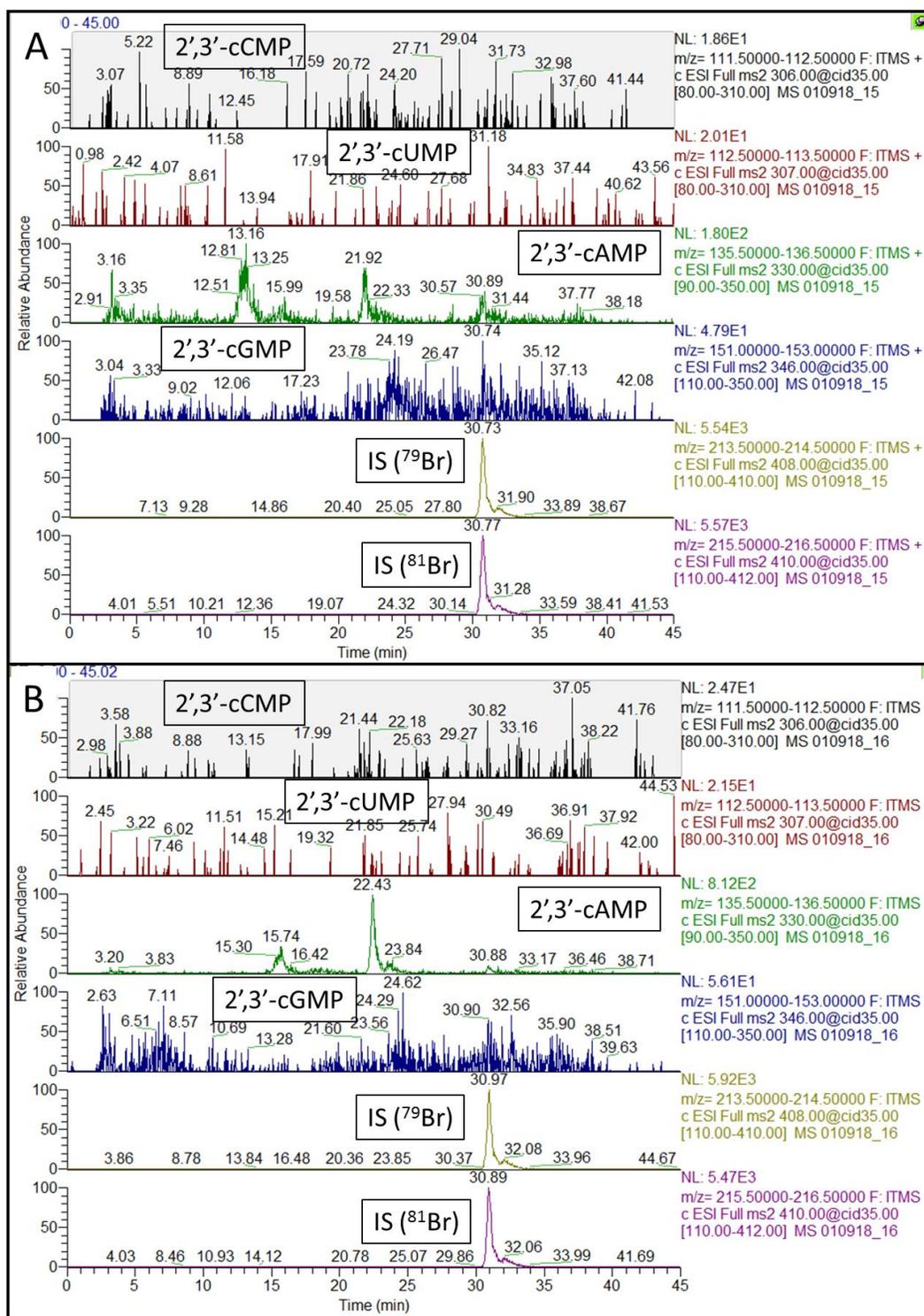
Supplementary Figure S2.10. 2',3'-cNMPs are undetectable in Δ*rna*. The *rna* mutant was subjected to 2',3'-cNMP extraction (for details, see **Materials and Methods**; **section 2.4**). The chromatograms demonstrate that 2',3'-cNMP levels are below the limit of detection during both exponential **(A)** and stationary **(B)** phase growth (compare to control chromatograms in **Supplementary Figure S2.7**).
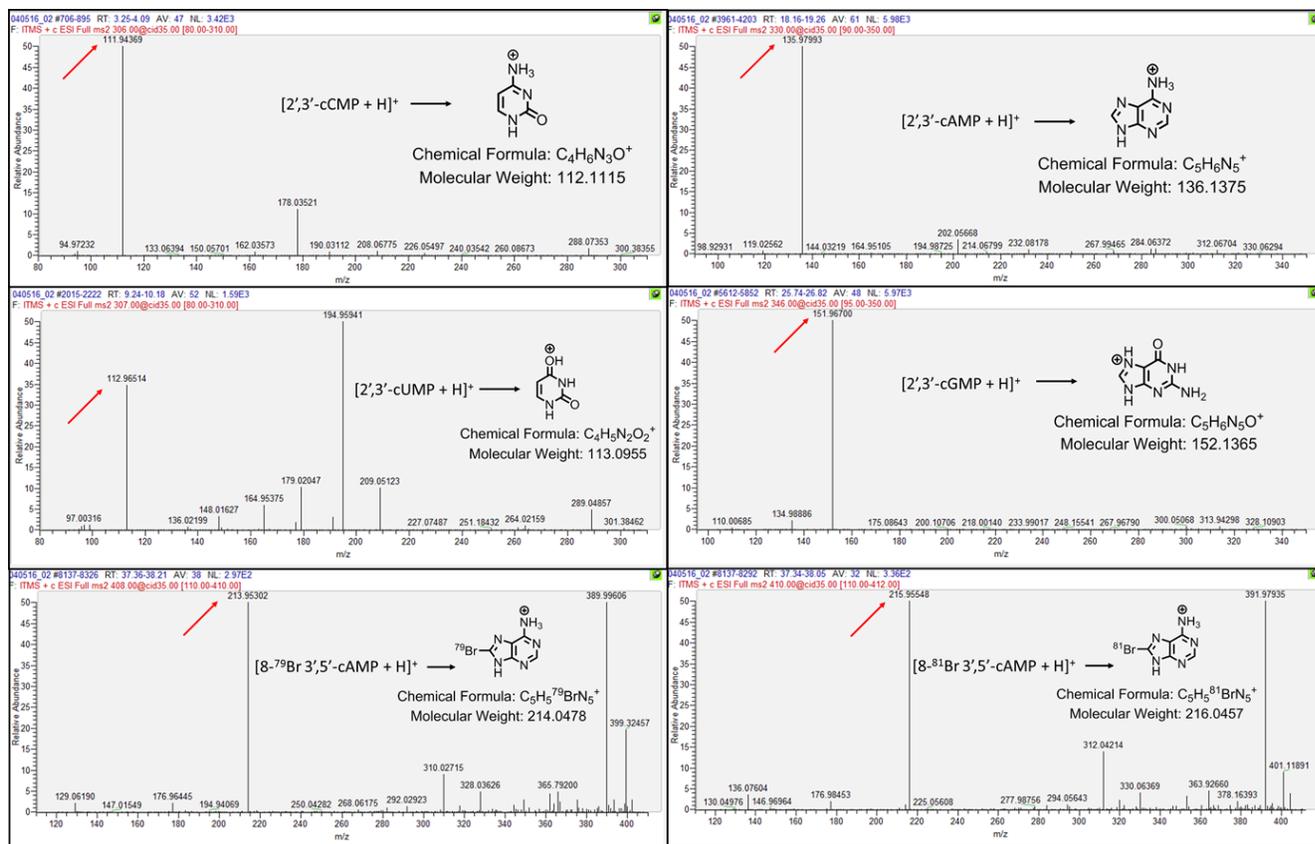
Supplementary Figure S2.11. Representative nucleotide MS/MS spectra. The spectra show the *m/z* of the protonated nucleobase cation (red arrow) derived from each protonated nucleotide parent ion [M+H]+, which was used for nucleotide quantification.
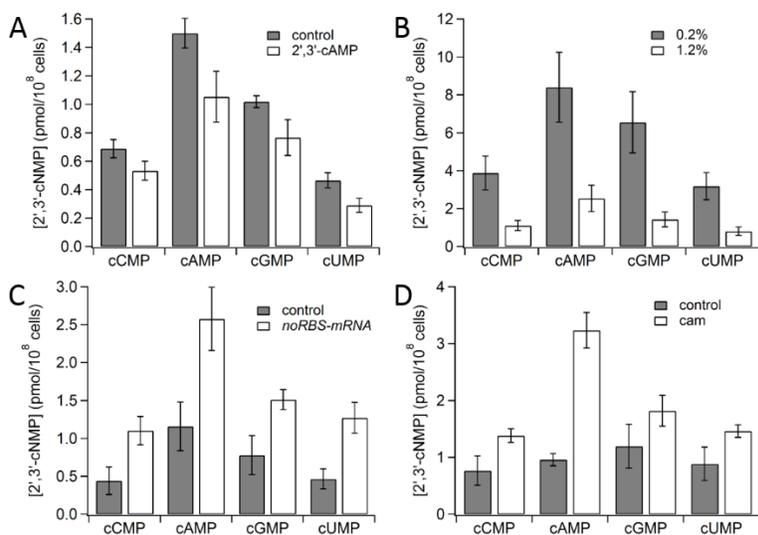
Supplementary Figure S2.12. Full normalized 2',3'-cNMP quantification data. 2',3'-cNMPs were quantified following exogenous 2',3'-cAMP addition **(A)**, in the presence of different casamino acid concentrations (0.2 or 1.2%) **(B)**, upon expression of non-translatable mRNA

noRBS-mRNA **(C)**, and in the presence of chloramphenicol **(D)**. The lower levels of 2',3'-cNMPs in **(A)**, **(C)** and **(D)** relative to **(B)** are a result of growth in 250-mL glass Erlenmeyer flasks *versus* 50-mL plastic conical tubes. For details of each experiment, see main text (**Figure 2.7**).
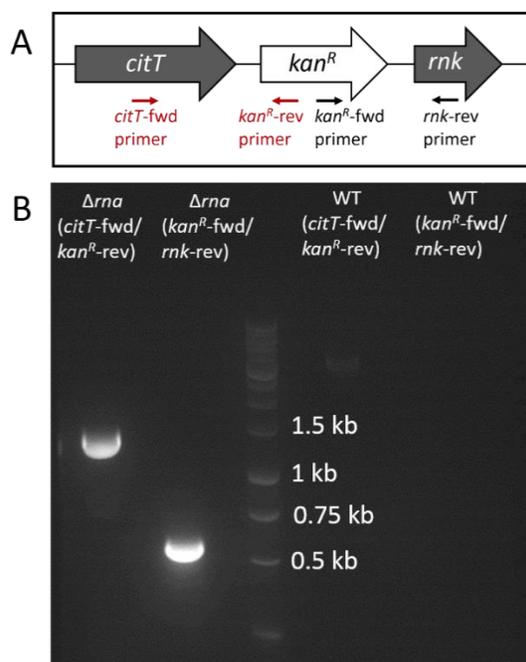
Supplementary Figure S2.13. Validation of *rna*::*kan^R* genotype in the BW25113 Δ*rna* strain. (A) Schematic of PCR-based procedure to validate disruption of the *rna* gene, which lies between *citT* and *rnk* (not to scale). **(B)** Locus-specific primers were utilized to confirm disruption of the *rna* gene with the *kan^R* cassette (encoding kanamycin phosphotransferase[5]), generating products with the expected sizes (first two lanes). Analogous control reactions using WT BW25113 genomic DNA as the template generated no products (last two lanes). For details of the procedure, see **Supplementary Protocol S2.1.**

Supplementary Protocol S2.1: Confirmation of *rna* gene disruption in the BW25113 Δ*rna* strain *via* PCR

 Chromosomal DNA was isolated from Δ*rna* using the GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich) according to the manufacturer's specifications. The genomic DNA was used as the template in two separate PCR reactions to amplify the junctions upstream and downstream of the *rna*::*kan^R* locus, as previously described[5]. PCR was performed using locus-specific primers (for sequences, see **Supplementary Table S2.1**) with Phusion DNA polymerase in Phusion-HF buffer (New England Biolabs), as detailed by the manufacturer. As a negative control, WT BW25113 genomic DNA was subjected to the same PCR amplification conditions. The PCR reactions then were analyzed on a 1.5% agarose gel.

Supplementary Table S2.1. Primer sequences for validation of *rna*::*kan^R* genotype in the BW25113 Δ*rna* strain.

Junction	Primer ID	Sequence (5'-3')	Annealing temperature (°C)
Upstream of *rna*::*kan^R*	*citT*-fwd (upstream gene)	ATT AGC TGG TTG CAG TGG TTC CTC	64
	kan^R-rev	GTC ATA GCC GAA TAG CCT CTC CAC	
Downstream of *rna*::*kan^R*	*kan^R*-fwd	TCG CAG CGC ATC GCC TTC TAT C	64
	rnk-rev (downstream gene)	CAC TTC GCC ATC GCT AAG ATT GC	

Supplementary Table S2.2. Primer sequences for QuikChange site-directed plasmid mutagenesis.

Primer ID	Sequence (5'-3')	Annealing temperature (°C)
CNP-H73L-fwd*	GGGTGTTCTGCTGTGCACCACGAAATTC	65
CNP-H73L-rev*	GAATTCGTGGTGCACAGCAGAACACCC	
CNP-H152L-fwd*	GTGCGCTGGTTACCCTGGGTTGTG	70
CNP-H152L-rev*	CACAACCCAGGGTAACCAGCGCAC	

*CNP numbering is based on the catalytic domain (final 221 amino acid residues; full-length protein accession: UniProtKB-P13233).

Supplementary Protocol S2.2: PCR and thermal cycling conditions for polymerase incomplete primer extension (PIPE) cloning [6]

The pACYC vector and the *noRBS-mRNA* insert were amplified via PCR using the primers in **Supplementary Table S2.3**. Amplification was performed using Phusion DNA polymerase in Phusion-HF buffer (New England Biolabs) according to the manufacturer's protocol, except the final 72°C extension step was omitted to increase the yield of incomplete extension products. The vector PCR reaction mixture subsequently was treated with *Dpn*I (0.8 units per μL of PCR reaction) for 2 h at 37°C to digest the pACYCDuet-1 template DNA. The vector PCR product and the insert PCR product were mixed in a 1:4 ratio and transformed into RbCl-competent DH5α *E. coli*.

***noRBS-mRNA* insert sequence (<mark>T7 promoter</mark>, <mark>T7 terminator</mark>)**

GACTGAGCTC<mark>TAATACGACTCACTATAGG</mark>TATCCGGCGTAGGCGAAGGACG
GGTCCAGTGCGTTCGCGCACTGTTGAGTAGAGTGTGAGCGCCATCTCTACG
GA<mark>CTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTG</mark>A
AGCTTGACT

Supplementary Table S2.3. Primer sequences for polymerase incomplete primer extension (PIPE) cloning.

Primer ID	Sequence (5'-3')	Annealing temperature (°C)
pACYC vector rev	TTCCTAATGCAGGAGTCGCATAAGGGA	61
pACYC vector fwd	CTGAAACCTCAGGCATTTGAGAAGCAC	
noRBS-mRNA insert fwd	TCCCTTATGCGACTCCTGCATTAGGAAGAC TGAGCTCTAATACGACTCAC	65
noRBS-mRNA insert rev	GTGCTTCTCAAATGCCTGAGGTTTCAGAGT CAAGCTTCAAAAAACCCCTC	

Supplementary Table S2.4. Data from microarray analysis of biofilm-related genes.

Symbol	WT 1	WT 2	WT 3	Δ*rna* 1	Δ*rna* 2	Δ*rna* 3	Log$_2$FoldChange	adj.P.Val
csgA	7.92	7.90	7.71	8.17	8.25	8.46	0.449	0.0625
csgB	2.59	2.56	2.76	3.16	3.16	3.26	0.559	0.0197
csgC	6.438	6.44	6.38	6.90	7.39	7.44	0.828	0.0195
csgD	9.50	9.51	9.77	8.08	8.77	8.83	-1.03	0.0255
csgE	10.2	10.2	10.41	8.79	9.25	9.37	-1.11	0.00949
csgF	8.52	8.43	8.63	6.87	7.31	7.31	-1.36	0.00328
csgG	3.44	4.16	4.07	3.69	3.20	3.43	-0.447	0.261
pgaA	7.75	7.94	7.99	6.56	6.26	6.28	-1.53	0.00179
pgaB	5.52	5.66	5.71	4.85	4.31	4.51	-1.07	0.00761
pgaC	6.39	6.12	6.27	5.23	5.53	5.63	-0.792	0.0158
pgaD	10.6	10.47	10.4	8.49	9.22	9.33	-1.48	0.00809

Supplementary Table S2.5. Data from quantitative reverse transcription PCR (RT-qPCR) validation of *pgaA*.

Sample	*pgaA*	Neg. error	Pos. error
WT 1	1.000000	0.104352	0.116510
WT 2	1.068200	0.101733	0.112442
WT 3	1.131980	0.077083	0.082716
Δ*rna* 1	0.600738	0.032671	0.034550
Δ*rna* 2	0.607046	0.058363	0.064572
Δ*rna* 3	0.728109	0.040686	0.043094

Supplementary Table S2.6. Primers for RT-qPCR analysis of *pgaA*.

Primer ID	Sequence (5'-3')
Fwd	CACTGCGCGTAGGCATAAAC
Rev	GCACCGGACAAAGCCAATTT

2.7 Supplementary References

(1) Quan, S.; Hiniker, A.; Collet, J.; Bardwell, J. C. A. In *Bacterial Cell Surfaces: Methods and Protocols*; Delcour, A. H., Ed.; Springer Science+Business Media: New York, 2013; Vol. 966, p 359-366.

(2) O'Toole, G. *J. Vis. Exp.* **2011**, doi: 10.3791/2437, doi: 10.3791/2437.

(3) Stead, M.; Agrawal, A.; Bowden, K.; Nasir, R.; Mohanty, B.; Meagher, R.; Kushner, S. *Nucleic Acids Res.* **2012**, *40*, doi:10.1093/nar/gks1680.

(4) Lopez, P.; Marchand, I.; Yarchuk, O.; Dreyfus, M. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 6067-6072.

(5) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K.; Tomita, M.; Wanner, B.; Mori, H. *Mol. Syst. Biol.* **2006**, *2*, doi: 10.1038/msb4100050.

(6) Klock, H.; Koesema, E.; Knuth, M.; Lesley, S. *Proteins-Structure Function and Bioinformatics* **2008**, *71*, 982-994.

Chapter 3: Chemical tools reveal diverse functions modulated by RNase I and nucleoside 2',3'-cyclic monophosphate pools in *Escherichia coli*

Chapter 3: Chemical tools reveal diverse functions modulated by RNase I and nucleoside 2',3'-cyclic monophosphate pools in *Escherichia coli*

3.1 Introduction

Throughout the kingdoms of life, specialized nucleotides regulate biological functions through second messenger signaling pathways, and primary nucleotide homeostasis influences various processes through direct cellular monitoring of nucleotide concentrations. Among nucleotide second messengers, guanosine 3',5'-cyclic monophosphate (3',5'-cGMP) regulates such processes as vasodilation and visual transduction in eukaryotes,[1] while adenosine 3',5'-cyclic monophosphate (3',5'-cAMP) governs numerous processes in diverse organisms, including steroidogenesis in mammals and carbon catabolism in bacteria.[2] Cyclic dimeric nucleotides also have been discovered, with cyclic dimeric-3':5'-guanosine monophosphate (c-di-GMP) modulating biofilm formation in many bacteria, and other cyclic dimeric nucleotides controlling processes such as sporulation in prokaryotes and innate immunity in eukaryotes.[3] Recent discoveries also have demonstrated that bacterial acyclic nucleotides, including guanosine 3'-diphosphate, 5'-(tri)diphosphate ([p]ppGpp) and P1,P4-di-adenosine 5'-tetraphosphate (Ap4A), induce remodeling of the transcriptome and the proteome[3c] and regulate biofilm production via modulation of c-di-GMP levels,[4] respectively. In addition to the physiological roles of nucleotide second messengers, primary nucleotide metabolism influences a host of physiological processes in bacteria: perturbation of *de novo* nucleotide biosynthesis alters biofilm formation in *Escherichia coli*,[5] exogenous addition of nucleosides and dNTPs promotes a positive chemotactic response in *Vibrio fischerii* [6], and pyrimidine nucleobases function as chemoattractants in *E. coli*.[7] These findings illustrate the multi-faceted nature of nucleotide signal transduction in prokaryotes and

suggest that other aspects of nucleotide synthesis and salvage regulate various cellular processes, warranting further investigation into the biological effects of additional nucleotide pools.

An intriguing class of nucleotides is the nucleoside 2',3'-cyclic monophosphates (2',3'-cNMPs), which were first observed in *E. coli* several decades ago,[8] but limited exploration of their biological relevance has occurred since the initial observation. In eukaryotes, recent studies have demonstrated that organ stress generates extracellular 2',3'-cAMP which is ultimately dephosphorylated to adenosine, with the nucleoside likely activating G-protein coupled receptors to modulate intracellular signaling.[9] In contrast, increased 2',3'-cAMP levels depolarize the mitochondrial membrane through unknown mechanisms, resulting in apoptosis in rat liver cells and oligodendrocytes, further demonstrating the physiological relevance of this 2',3'-cNMP in eukaryotes.[10] In addition, 2',3'-cAMP mediates stress granule assembly in Arabidopsis thaliana,[11] and 2',3'-cAMP, -cGMP, -cIMP, and -cCMP have been quantified in mammalian organs and cells,[12] supporting a physiological role for 2',3'-cGMP, -cIMP and -cCMP in eukaryotes.

While the functions of 2',3'-cNMPs in eukaryotes are only beginning to emerge, even less is known about the biological significance of 2',3'-cNMPs in prokaryotes. Various 2',3'-cNMPs have been identified in different bacterial species (2',3'-cCMP and -cUMP in *Pseudomonas fluorescens*;[13] 2',3'-cAMP in Staphylococcus aureus; 2',3'-cAMP, -cGMP, -cUMP, -cCMP in *E. coli*[14]), with our group reporting the physiological quantification of these cyclic nucleotides in a prokaryote.[15] Within *E. coli*, RNase I (an RNase T2 family member[16]) generates all detectable 2',3'-cNMPs through hydrolysis of mRNA and rRNA, providing the first insight into the biosynthetic origin of these atypical nucleotides in any organism. Additional experiments identified a physiological function for 2',3'-cNMPs and RNase I in biofilm production,[15] with both the RNase I deletion strain (Δ*rna*) and WT cells heterologously expressing a 2',3'-cyclic

nucleotide phosphodiesterase (CNPase) exhibiting a hyper-biofilm phenotype.[15] Taken together, these findings allude to additional 2',3'-cNMP- and RNase I-dependent processes.
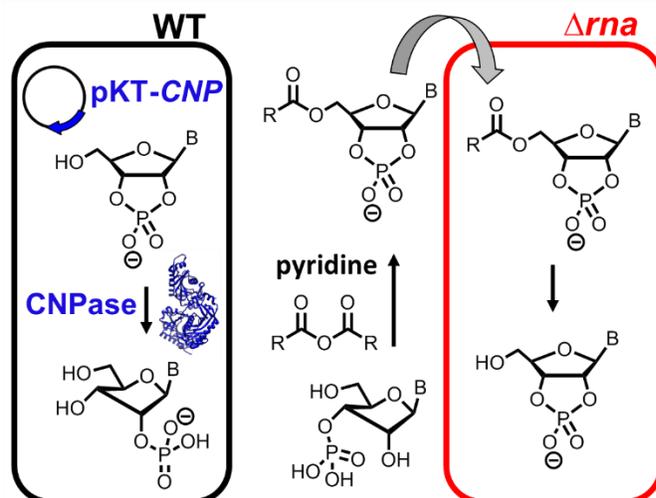

Figure 3.1. (Bio)chemical tools to modulate intracellular 2',3'-cNMP pools. Left: Expression of CNPase in WT *E. coli* enables hydrolysis of 2',3'-cNMPs in the presence of RNase I. Right: Treatment with lipophilic 5'-*O*-ester 2',3'-cNMP derivatives enables up-regulation of 2',3'-cNMP levels in Δ*rna*, which lacks both RNase I and endogenous 2',3'-cNMPs.

To elucidate the range of cellular pathways regulated by RNase I and 2',3'-cNMPs, as well as dissect their differential effects, the present work describes the transcriptome-wide changes in *E. coli* lacking RNase I and *E. coli* expressing CNPase (which have nearly undetectable 2',3'-cNMP levels[15]), enabling identification of cellular processes governed by 2',3'-cNMPs *versus* those regulated by RNase I. In addition, cell-permeable 2',3'-cNMP analogs, inspired by the lipophilic derivatives of 3',5'-cAMP and -cGMP widely employed to study second messenger signaling in eukaryotes,[2a] were synthesized as additional tools to study 2',3'-cNMP-dependent phenotypes in cells lacking both RNase I and endogenous 2',3'-cNMPs. While experiments using lipophilic 3',5'-cNMP analogs in bacteria are limited, cell-permeable analogs of 3',5'-cAMP and -cGMP previously were utilized to characterize 3',5'-cNMP-responsive promoter activation in *E. coli*,[17] demonstrating the feasibility of this strategy to study 2',3'-cNMP-dependent processes in bacteria. Our studies using cell-permeable 2',3'-cNMP derivatives and CNPase have linked 2',3'-cNMPs and RNase I to the regulation of diverse cellular processes, including biofilm formation, motility, acid resistance, *β*-lactam tolerance, and nucleotide metabolism, highlighting key pathways controlled by 2',3'-cNMPs and

suggesting the presence of 2',3'-cNMP-sensing proteins within the cell. These (bio)chemical tools will facilitate further elucidation of the mechanisms governing 2',3'-cNMP- and RNase I-dependent processes in additional bacterial taxa and in distantly related organisms, informing the discovery of novel approaches to control organismal function through modulation of 2',3'-cNMP metabolism and signaling.

3.2 Results

Comparative transcriptomics reveals RNase I- and 2',3'-cNMP-dependent processes

The previous finding that RNase I (encoded by the *rna* gene) produces all measureable 2',3'-cNMPs in *E. coli* necessitated the development of methods to modulate 2',3'-cNMP levels independently of RNase I expression. To this end, we leveraged the catalytic domain of *Rattus norvegicus* cyclic nucleotide phosphodiesterase[18] (CNPase; UniprotKB P13233 for full-length protein) to hydrolyze 2',3'-cNMPs in RNase I[+] cells. Previous work validated that *E. coli* expressing CNPase from plasmid pKT-*CNP* exhibit sub-quantifiable levels of 2',3'-cAMP and -cGMP, along with ~25-fold and ~15-fold lower levels of 2',3'-cCMP and -cUMP, respectively [15]. Therefore, expression of CNPase (from plasmid pKT-*CNP*) enabled comparison of 2',3'-cNMP- and RNase I-linked processes through transcriptomic profiling of RNase I[+] *E. coli* with reduced 2',3'-cNMP levels and *E. coli* deficient for RNase I. These gene expression studies revealed ~800 genes specifically dysregulated in Δ*rna* relative to WT and nearly 700 genes uniquely altered in WT *E. coli* expressing CNPase compared to control cultures expressing an inactive CNPase variant (**Figure 3.2A**). An additional 141 transcripts were mutually altered upon either *rna* deletion or CNPase expression, with 80 of these genes displaying the same regulatory trend in both Δ*rna* and WT pKT-*CNP*, relative to the respective control strains (**Figure 3.2A**). Intriguingly, both RNase I and 2',3'-cNMPs modulate genes encoding diverse protein classes including transcription factors,

transporters, and hydrolases (**Figure 3.2 B,C**), linking 2',3'-cNMPs and RNase I to numerous cellular functions. While mRNA levels of genes involved in many of the same processes were similarly perturbed in both the *rna* mutant and the CNPase expression strain, certain transcripts were differentially regulated by *rna* deletion compared to CNPase expression. For example, biofilm and motility genes were dysregulated both in cells lacking RNase I and in RNase I$^+$ cells with decreased 2',3'-cNMP concentrations, but the absence of both RNase I and 2',3'-cNMPs in Δ*rna* affected a greater number of these transcripts compared to the hydrolysis of 2',3'-cNMPs in cells expressing RNase I (**Figures 3.3F** and **3.4B**).[15] A similar case exists for transcripts encoding components of nucleotide metabolism. Interestingly, *rna* deletion and CNPase expression generally altered the transcription of different genes associated with a particular cellular process, demonstrating that aberrant 2',3'-cNMPs elicit distinct transcriptional effects depending on the presence of RNase I. Due to the expansive role of 3',5'-cAMP in regulating *E. coli* transcription through interaction with the 3',5'-cAMP receptor protein (Crp),[19] the intracellular concentration of 3',5'-cAMP was quantified in WT and Δ*rna* to probe the potential function of 3',5'-cAMP-Crp in regulating the altered transcriptional profile in the absence of RNase I. LC-MS/MS analysis revealed barely detectable levels of 3',5'-cAMP in both strains (limit of detection is ~150 fmol[12]), which is not surprising due to the attenuation of adenylate cyclase activity in the presence of glucose[20] (the carbon source in all experiments) (**Supplementary Figure S3.1**). These data indicate that altered 3',5'-cAMP levels are not modulating the transcriptional changes in *E. coli* lacking RNase I and 2',3'-cNMPs.

2',3'-cNMPs influence flagellar motility

Analysis of the transcriptome in WT/Δ*rna* revealed substantial up-regulation of >30 genes involved in chemotaxis and flagellar motility in *E. coli* lacking RNase I (**Figure 3.3F**). Notably,

these motility-associated genes include transcripts encoding methyl-accepting chemotaxis receptor proteins (MCPs; *tap*, *tar*, *tsr*, *trg*, *aer*),[21] intracellular Che effectors (*cheAW*, *cheRBYZ*),[21] transcriptional activators (*flgM*, *fliAZ*),[22] flagellar biosynthesis/export proteins (*flgN*, *fliST*, *flhBA*, *fliR*),[23] and components of the flagellar motor (*motAB*, *flgEFGHI*, *flgKL*, *fliC*, *fliE*)[23] (**Figure 3.3F**). However, while expression of CNPase increased the mRNA levels of flagellar-associated genes *fliK*, *fliR*, and *fliE*, decreasing the 2',3'-cNMP concentration did not alter abundance of other flagellar- or chemotaxis-related transcripts.
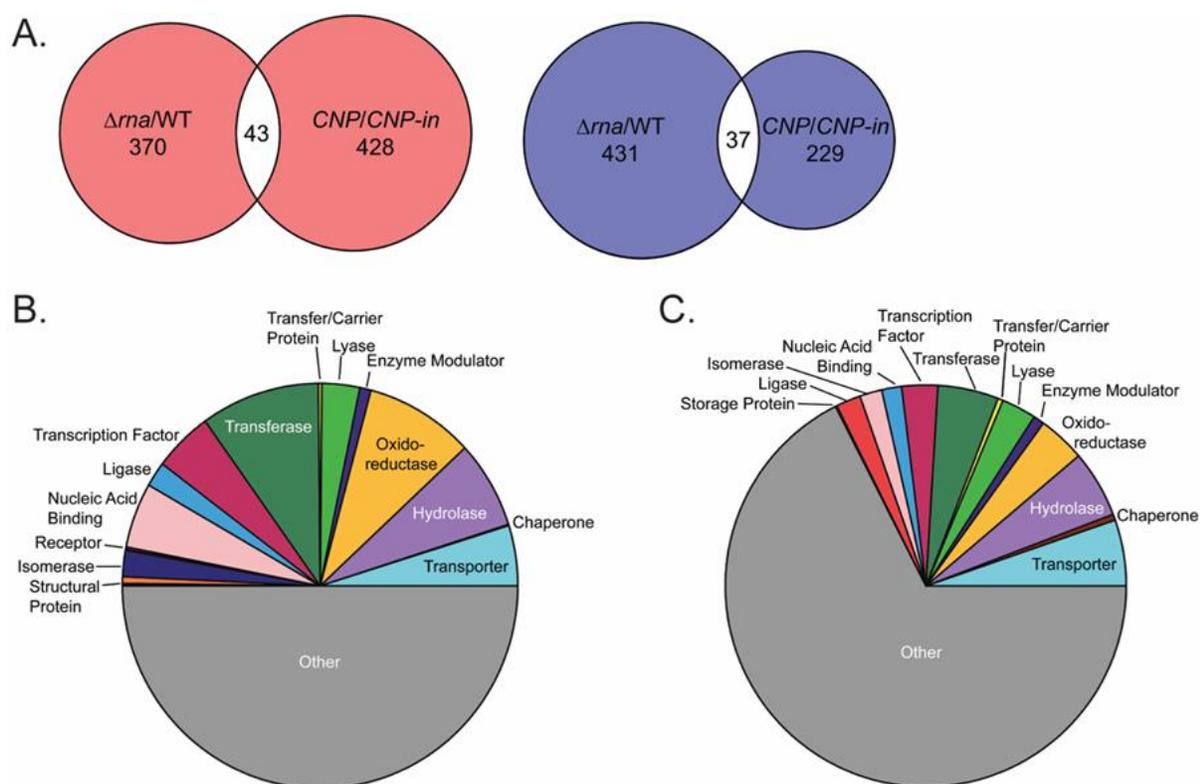

Figure 3.2. Global gene expression analyses reveal RNase I- and 2',3'-dependent cellular processes. (**A**) Venn diagrams depicting quantities of up-regulated (red) and down-regulated genes (blue) in Δ*rna vs.* WT and in *CNP vs. CNP-inact*, as well genes mutually altered under both conditions. Venn diagrams were created using Venn Diagram Plotter software (Pacific Northwest National Laboratory). (**B, C**) Gene ontology pie charts displaying differentially expressed genes according to protein class in Δ*rna vs.* WT (**B**) and in *CNP vs. CNP-inact* (**C**). Pie charts were plotted using the Panther Gene Ontology Database (http://pantherdb.org/).

To investigate the phenotypic consequences of these altered gene expression profiles, the effect of RNase I and 2',3'-cNMPs on flagellar-dependent swimming motility was assayed. In agreement with the increased expression of chemotaxis and motility genes in Δ*rna* relative to WT, the RNase I-deficient mutant was hypermotile (**Figure 3.3**), and complementation of Δ*rna* with plasmid pBAD33-*rna* restored WT swimming behavior (**Supplementary Figure S3.3**), demonstrating that RNase I and 2',3'-cNMPs regulate swimming motility. To further interrogate the role of RNase I and 2',3'-cNMPs in this process, the swimming motility of WT cells expressing CNPase or CNPase-inact was evaluated. Hydrolysis of 2',3'-cNMPs through CNPase expression did not alter motility (**Figure 3.3C**), indicating that the increased expression of *fliK*, *fliR*, and *fliE* in WT cells expressing CNPase is insufficient to confer hypermotility in the presence of RNase I, possibly due to the very low expression levels of motility machinery in even the non-transformed WT strain (**Figure 3.3D, E, F** and **Supplementary Figure S3.7**). In contrast, treatment of Δ*rna* with either Bt-cAMP or Bt-cUMP attenuated motility in a dose-dependent manner without affecting motility of the WT control strain (**Figure 3.3A, B**), demonstrating that 2',3'-cNMPs modulate motility in the absence of RNase I. Importantly, neither Bt-cAMP nor Bt-cUMP inhibited cell growth (**Supplementary Figure S3.2**), and sodium butyrate (NaBt), the product of 5'-*O*-ester hydrolysis, did not alter swimming motility (**Supplementary Figure S3.5**), confirming a function for the 2',3'-cNMPs. As an additional control experiment, motility was assayed in the presence of 5'-*O*-benzoyl 2',3'-cUMP (Bz-cUMP). This compound also impaired the motility of Δ*rna* (**Supplementary Figure S3.6**), further validating the role of 2',3'-cNMPs in this process and ruling out possible confounding effects from the 5'-*O*-ester substituent. In addition, positive control experiments using uracil and ribose, which are known *E. coli* chemoattractants,[7,24]

demonstrated that both the WT and Δrna strains respond normally to these established chemo-modulators (**Supplementary Figure S3.9**).
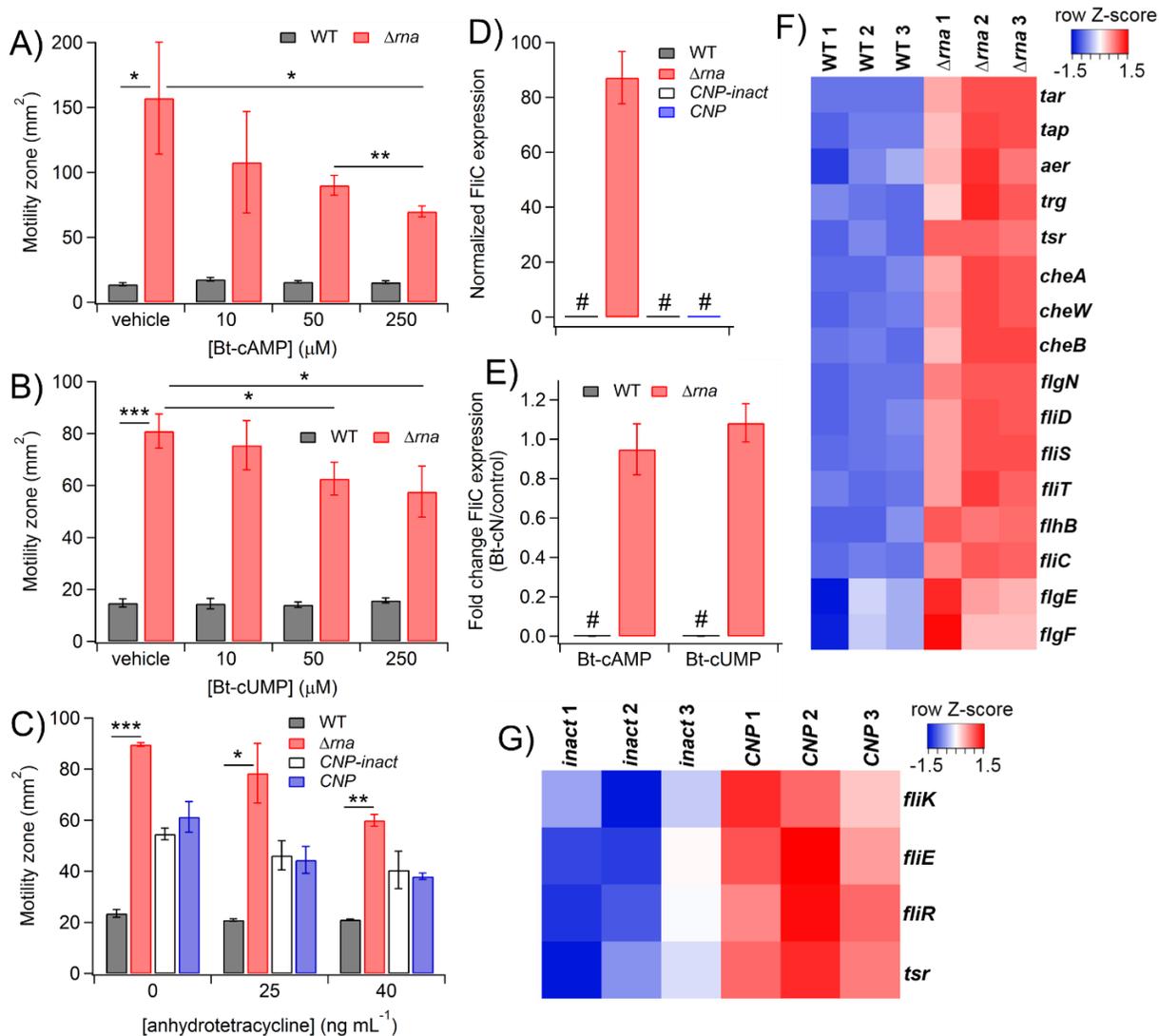

Figure 3.3. RNase I and 2',3'-cNMPs modulate swimming motility. (**A, B, C**) *E. coli* lacking RNase I are hypermotile relative to WT, and treatment with Bt-cAMP (**A**) or Bt-cUMP (**B**) inhibits motility in Δrna. (**C**) Inducing 2',3'-cNMP hydrolysis in WT cells through CNPase expression does not impact motility relative to control cultures expressing CNPase-inact. Motility was assayed in the presence of various anhydrotetracycline concentrations to induce phosphodiesterase expression. (**D**) Expression of flagellar structural protein FliC increases in Δrna relative to WT (quantified by western blot). FliC expression is undetectable in WT, *CNP-inact*, and *CNP* (# denotes undetectable FliC expression). (**E**) Treatment with either Bt-cAMP or Bt-cUMP does not alter FliC expression in WT or Δrna. Heatmap displaying differentially expressed motility-related genes in WT *vs.* Δrna (**D**) and *CNP-inact vs. CNP* (**E**). All data represent the mean ± standard deviation of at least three biological replicates (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

To further probe the molecular effects of increased expression of flagellar export and assembly genes upon deletion of *rna* or expression of CNPase, the abundance of the FliC flagellar filament was quantified by western blot in WT, Δ*rna*, *CNP*, and *CNP-inact* strains. In agreement with the ~8-fold up-regulation of the *fliC* transcript in Δ*rna* compared to WT (**Figure 3.3F**), the FliC protein level was elevated in the RNase I-deficient mutant (**Figure 3.3D** and **Supplementary Figure S3.7**). In contrast, neither CNPase expression nor treatment with Bt-cNMPs altered FliC abundance (**Figure 3.3D, E** and **Supplementary Figure S3.7**). These findings demonstrate that altering 2',3'-cNMP concentrations in RNase I$^+$ *E. coli* does not elicit an observable motility phenotype due to the extremely low abundance of the FliC filament in BW25113, which is known to be hypomotile.[25]

2',3'-cNMPs modulate biofilm formation

Our group previously demonstrated that *E. coli* deficient for RNase I lacks 2',3'-cNMPs and displays a hyper-biofilm phenotype due to increased expression of curli structural genes *csgBAC*.[15] Moreover, biofilm production was demonstrated to increase in WT *E. coli* expressing CNPase,[15] and the present study demonstrates that CNPase expression also up-regulated production of *csgC*, along with several putative fimbrial adhesion transcripts and the *fimA* gene encoding a pilin subunit (**Figure 3.4B**). These data provide insight into the
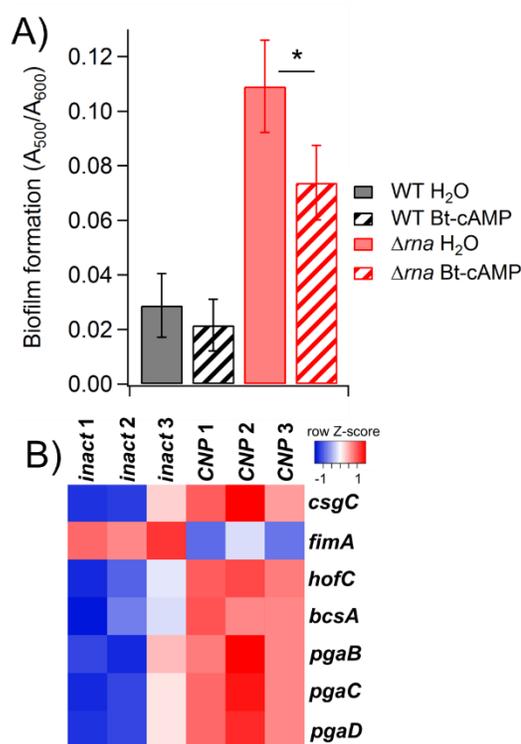
Figure 3.4. RNase I and 2',3'-cAMPs mediate biofilm formation. (**A**) Biofilm production increases in RNase I-deficient *E. coli* (quantified by Congo red staining), and treatment with Bt-cAMP (500 µM) impairs biofilm formation in Δ*rna*. Data represent the mean ± standard deviation of six biological replicates (* $P < 0.05$). (**B**) Heatmap depicting biofilm-related transcript levels in WT cells expressing either CNPase-inact or CNPase.

transcriptional mechanisms underlying the increased biofilm production upon CNPase-mediated 2',3'-cNMP hydrolysis. Intriguingly, CNPase expression induced transcription of the *pgaABCD* operon responsible for poly-*N*-acetyl-*β*-1,6-D-glucosamine (PNAG) biosynthesis in RNase I⁺ *E. coli* (**Figure 3.4B**), whereas *rna* deletion dampened expression of the PNAG biosynthetic cluster.[15] Transcription from the *pga* promoter is repressed by OmpR,[26] and CNPase expression decreased *ompR* transcript abundance, while *rna* deletion did not affect the *ompR* mRNA level. These data suggest that the differential modulation of PNAG biosynthesis in WT cells expressing CNPase *vs.* Δ*rna* cells is regulated, at least in part, by OmpR. To further evaluate the functional role of 2',3'-cNMPs in biofilm formation, cultures of WT and Δ*rna* were treated with Bt-cAMP. Relative to untreated control cultures, the *rna* mutant produced ~30% less biofilm in the presence of Bt-cAMP, while WT biofilm formation was unaffected (**Figure 3.4A**), supporting the conclusion that 2',3'-cNMPs regulate biofilm formation independently of RNase I. As mentioned above, Bt-cAMP did not affect the growth of either strain, even at a concentration of 1 mM (**Supplementary Figure S3.2A**). Moreover, sodium butyrate (NaBt) did not perturb biofilm formation (**Supplementary Figure S3.10**), demonstrating that the observed biofilm inhibition in Δ*rna* results from exogenously increasing the intracellular 2',3'-cAMP concentration.

RNase I regulates 5'-NDP concentrations

The transcriptomic data indicate that expression of *nrdAB*, which encode subunits of a type I aerobic ribonucleotide reductase (RNR),[27] were down-regulated ~1.9- and 2.6-fold, respectively, in Δ*rna* relative to WT (**Figure 3.5D**). Conversely, CNPase expression did not alter *nrdAB* expression. Due to the role of RNR in reducing 5'-NDPs to the corresponding 2'-deoxy 5'-NDPs,[27-28] 5'-NDP pools were quantified to assess the effect of altered RNase I and 2',3'-cNMP levels on 5'-NDP metabolism. In agreement with the gene expression data, 5'-NDPs accumulated

in Δ*rna* relative to WT (**Figure 3.5A**), whereas these pools were not perturbed by CNPase expression relative to inactive phosphodiesterase expression (**Figure 3.5A**). Collectively, these data demonstrate that RNase I indirectly modulates 5'-NDP levels by regulating expression of RNR-encoding genes through as-yet-unknown mechanisms.
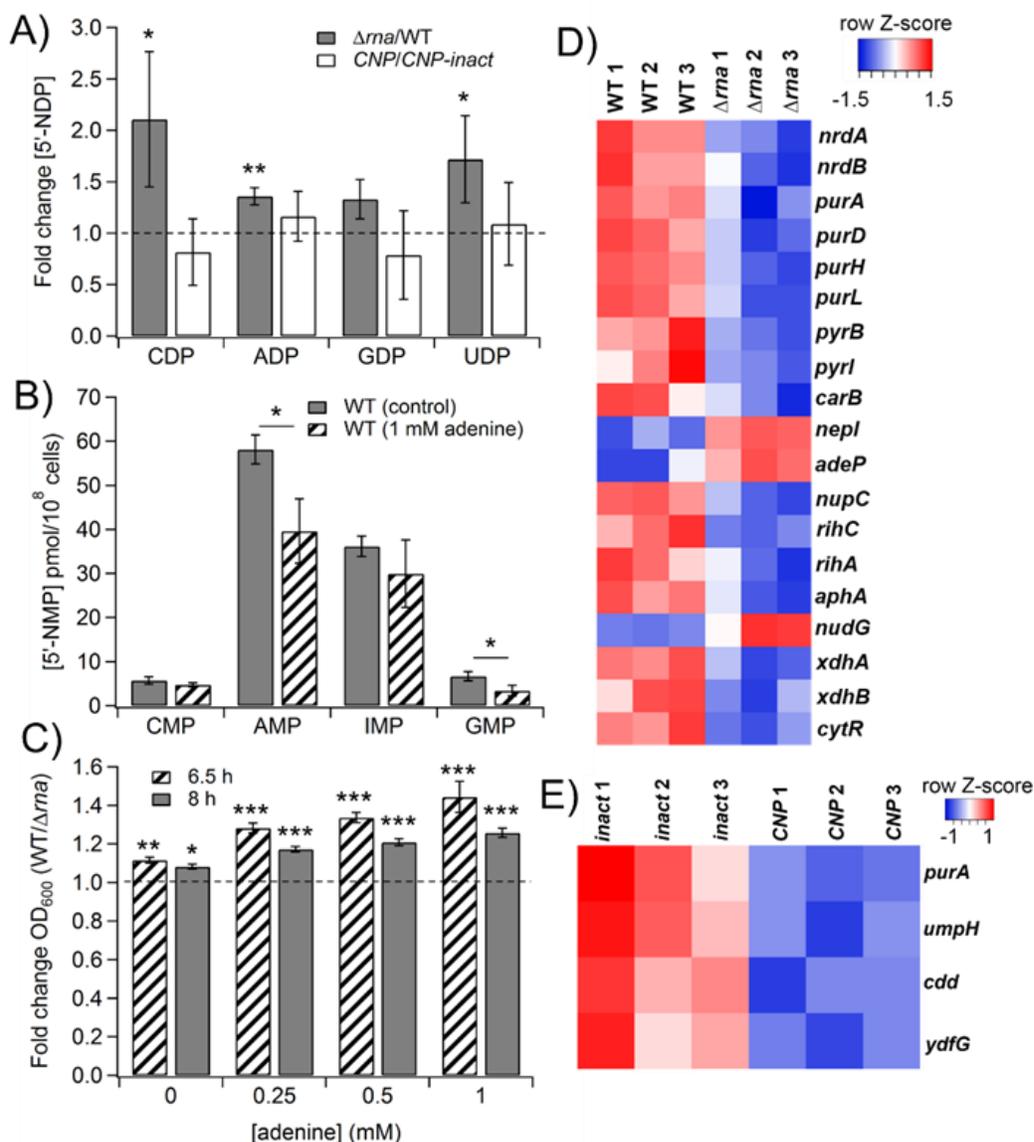

Figure 3.5. RNase I and 2',3'-cNMPs influence nucleotide pool homeostasis. (**A**) Depletion of RNase I increases 5'-NDP levels, while expression of CNPase has no effect. (**B**) *E. coli* lacking RNase I are more sensitive to adenine-induced growth inhibition, as compared to WT. (**C**) Treatment of WT *E. coli* with 1 mM adenine decreases intracellular purine concentrations. Data represent the mean ± standard deviation of at least three biological replicates (* $P < 0.05$; **$P < 0.01$; *** $P < 0.001$) (**D, E**) Heatmap displaying differentially expressed nucleotide metabolic genes in WT *vs.* Δ*rna* (**D**) and *CNP-inact vs. CNP* (**E**).

RNase I and 2',3'-cNMPs broadly impact nucleotide metabolism

Depletion of RNase I down-regulated transcription of numerous genes involved in purine and pyrimidine nucleotide metabolism, including *purDH* and *purL,* as well as *pyrBI* and *carB* (**Figure 3.5D**), all of which function in *de novo* nucleotide synthesis. The Δ*rna* strain also exhibited increased expression of *nepI*, encoding a purine nucleoside efflux pump, and attenuated transcription of nucleoside symporter genes *nupC* and *adeP* (**Figure 3.5D**). Other genes involved in nucleotide salvage also were down-regulated in Δ*rna*, namely those encoding nucleoside hydrolases *rihC*, *rihA*, and *nudG* (**Figure 3.5D**). Expression of CNPase similarly dampened transcription of *carA*, an essential gene in *de novo* pyrimidine synthesis, and attenuated expression of the salvage genes *cdd*, *umpH*, *upp*, *hofP*, and *xdhD*. Furthermore, *purA* transcript levels are reduced both in Δ*rna* and in WT cells expressing CNPase (**Figure 3.5**). Collectively, these data reveal that both RNase I and 2',3'-cNMPs influence nucleotide metabolism, although largely through modulation of distinct genes.

Due to the impact of aberrant RNase I and/or 2',3'-cNMP levels on nucleotide-related transcripts, the effect of *rna* deletion and CNPase expression on adenine sensitivity was interrogated. Adenine-induced growth inhibition often is exacerbated in *E. coli* mutants lacking components of *de novo* purine biosynthesis or catabolism,[29] leading to the hypothesis that RNase I and/or 2',3'-cNMPs may influence adenine toxicity. In support of the hypothesis, adenine more strongly inhibited the growth of RNase I-deficient *E. coli* relative to WT (**Figure 3.5B**). However, CNPase expression did not alter adenine sensitivity compared to control (**Supplementary Figure S3.11C**), indicating that the greater adenine-induced growth defect in Δ*rna* results from the absence of RNase I and/or the complete lack of 2',3'-cNMPs in this strain. Prior studies determined that adenine toxicity in *E. coli* occurs primarily due to guanine nucleotide starvation,[29] and

supplementation with guanosine or other natural purines abrogates the bacteriostatic effect of adenine.[29-30] Therefore, guanosine was added to cultures of WT and Δ*rna* in an attempt to restore normal growth in the presence of adenine; however, guanosine inhibited the growth of both strains in our experiments, both in the presence and absence of adenine (**Supplementary Figure S3.11A, B**). These disparate effects could be due to strain-dependent differences. While guanosine addition did not reverse adenine toxicity, concentrations of 5'-GMP and -AMP were lower in adenine-treated WT cultures relative to untreated control (**Figure 3.5C**). This result further suggests that altered purine metabolism mediates the differential sensitivity of WT and Δ*rna* to adenine toxicity, and corroborates the dysregulated expression of nucleotide metabolic genes in cells lacking RNase I.

RNase I and 2',3'-cNMPs regulate acid resistance

Expression of CNPase induced transcription of several genes involved in acid resistance, including the glutamate decarboxylase system encoded by *gadA* and *gadBC* (**Figure 3.6C**). Therefore, the acid tolerance of WT pKT-*CNP* and WT pKT-*CNP-inact* was evaluated, and the data revealed a 10-fold increase in survival rate at pH 2.5 upon CNPase expression, as compared to the survival rate of cells expressing CNPase-inact (**Figure 3.6A**). Conversely, the absence of RNase I reduced expression of a different set of acid resistance genes (**Figure 3.6B**), with Δ*rna* exhibiting the expected decrease in acid tolerance relative to WT *E. coli* in the phenotypic assay (**Figure 3.6A**).

RNase I influences *β*-lactam sensitivity

Expression of several genes involved in peptidoglycan maturation, such as *ampH*, *mrcB*, and *murEF*, was decreased in Δ*rna vs.* WT cells (**Figure 3.7B**). Notably, the *blr*, *lpoA*, and *yfeW*

transcripts implicated in penicillin binding and β-lactam resistance also were down-regulated in Δrna (**Figure 3.7B**), suggesting that *E. coli* lacking RNase I would exhibit decreased resistance to β-lactam challenge. In agreement with the gene expression data, dose-response assays revealed that Δrna was hyper-sensitive to carbenicillin-induced toxicity relative to WT (**Figure 3.7A**). Intriguingly, expression of CNPase similarly attenuated transcription of genes encoding the penicillin-binding proteins FtsI, LpoB, and DacC (**Supplementary Figure S3.12B**). However, CNPase expression did not alter carbenicillin tolerance relative to expression of CNPase-inact (**Supplementary Figure S3.12A**).
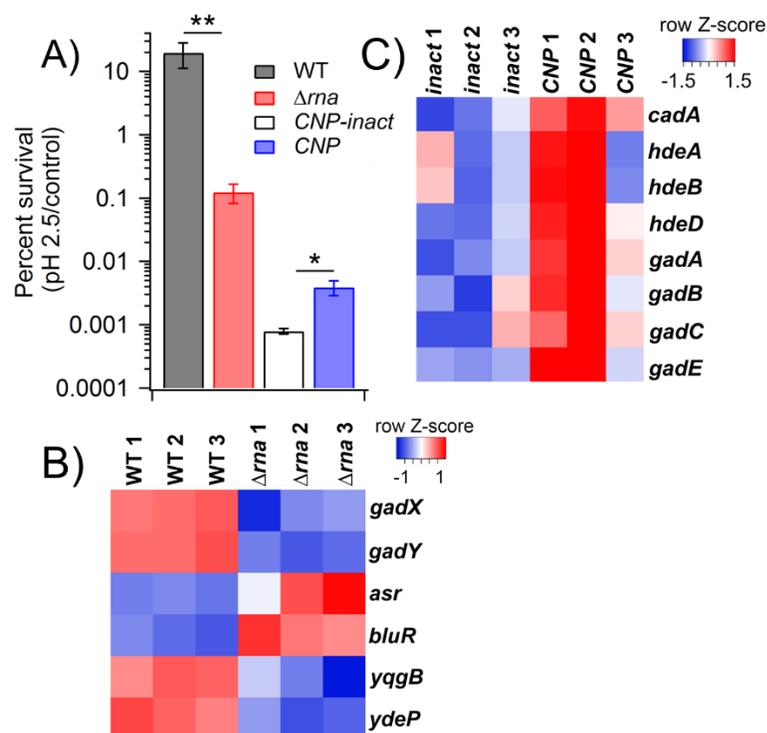

Figure 3.6. RNase I and 2',3'-cNMPs modulate acid tolerance. (**A**) Relative to the WT strain, *E. coli* lacking RNase I are more sensitive to low pH. Conversely, expression of CNPase confers increased acid resistance, as compared to expression of CNPase-inact. Acid sensitivity was assayed by quantifying CFU mL^{-1} for bacteria cultured at pH 2.5 compared to bacteria grown at neutral pH (control). Data represent the mean ± standard deviation of three biological replicates (* $P < 0.05$; ** $P < 0.01$). (**B, C**) Heatmaps depicting transcript abundance of acid resistance genes in WT *vs.* Δrna (**B**) and in CNPase-inact *vs.* CNPase (**C**).

3.3 Discussion

Nucleotide signaling mediates responses to various environmental stimuli in diverse bacterial taxa, orchestrating complex processes such as the motility-sessility transition, sporulation, and virulence factor production.[3c] The present work utilizes (bio)chemical perturbation of *in vivo* 2',3'-cNMP levels along with global transcriptome profiling and phenotypic analyses to characterize the function of 2',3'-cNMPs in bacterial nucleotide signaling and metabolism. Comparative transcriptomics has identified unique roles for RNase I and 2',3'-cNMPs in bacterial physiology, while also elucidating mutual processes altered in the absence of either RNase I or 2',3'-cNMPs. Modulation of nucleotide homeostasis perhaps mediates some of the transcriptional and phenotypic changes observed both in Δ*rna* and in RNase I[+] cells expressing CNPase, as mRNA levels associated with *de novo* nucleotide biosynthesis and salvage genes were dysregulated under both conditions (**Figure 3.5D**). Moreover, primary nucleotide metabolism influences processes such as biofilm formation in *E. coli*[5] (perhaps through modulation of c-di-GMP pools) and cell wall rigidity in *Lactococcus lactis.*[31] A prior study also established that NTP concentrations impact the efficacy of transcription initiation from rRNA promoters,[32] suggesting a potential mechanism through which aberrant nucleotide metabolism could drive transcriptional changes upon depletion of RNase I or expression of CNPase. Additionally (or alternatively), 2',3'-cNMPs potentially govern certain processes through direct interactions with macromolecular effectors; recent work identified the polyadenylate-binding protein Rbp47b in *A. thaliana* as the first 2',3'-cNMP-binding effector in any organism.[11] Rbp47b serves as a scaffold for assembly of the multi-protein stress granule complex, and this assembly is facilitated by 2',3'-cAMP binding to Rbp47b.[11] Intriguingly, Rbp47b also binds mRNA,[33] suggesting that stress granule formation is modulated by the extent of mRNA decay in the cell, as degradation of poly-adenylated mRNA

likely increases the 2',3'-cAMP concentration. These findings allude to the existence of other effectors in diverse organisms that sense RNA-derived 2',3'-cNMPs as reporters of cellular stress.
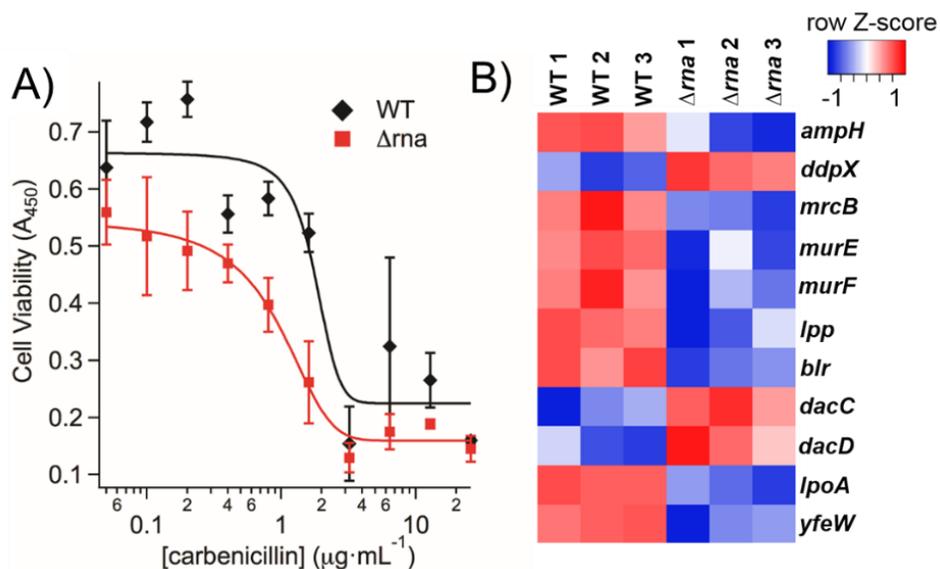

Figure 3.7. RNase I impacts tolerance to *β*-lactam treatment. (**A**) Δ*rna* exhibits increased sensitivity to carbenicillin relative to WT (as determined using a colorimetric tetrazolium-based cell viability assay). (**B**) Heatmap displaying differentially expressed genes involved in peptidoglycan synthesis and β-lactam resistance in WT and Δ*rna*.

The transcriptional profiling data also revealed certain cellular processes for which a greater number of transcripts were altered by RNase I depletion than by CNPase expression, and *vice versa*. For example, while perturbation of either RNase I or 2',3'-cNMP levels dysregulated expression of genes associated with peptidoglycan maturation and lipopolysaccharide biosynthesis, the absence of RNase I perturbed a greater number of these mRNAs (**Figure 3.7B** and **Supplementary Figure S3.12B**). These differences indicate that 2',3'-cNMP production is not the sole mechanism through which RNase I modulates gene expression in *E. coli*, suggesting that other functions of RNase I impact the transcriptome. The *rna* gene produces both a cytoplasmic and periplasmic variant of RNase I; the cytoplasmic version displays a proclivity for the hydrolysis of small (~2-12-residue) oligoribonucleotides (oligoRNAs) *in vitro*, while the

periplasmic form is non-specific.[34] The different variants of RNase I potentially modulate gene expression through distinct mechanisms. For example, the dysregulation of genes involved in peptidoglycan biosynthesis suggests a potential function for periplasmic RNase I in cell wall assembly (**Figure 3.7B**). Importantly, the rate of mRNA decay is not altered in *E. coli* lacking the *rna* gene,[35] demonstrating that RNase I is not directly modulating transcript abundance through mRNA degradation. Alternatively, the *in vitro* preference of cytoplasmic RNase I for the degradation of small RNA substrates suggests that small oligoRNAs could potentially accumulate in the Δ*rna* strain and alter transcription. Interestingly, prior studies in *Pseudomonas aeruginosa* deficient for oligoribonuclease (which degrades short RNA substrates to 5'-NMPs) demonstrated that high levels of oligoRNAs shift transcriptional start sites across the genome.[36] These findings suggest that a similar mechanism may elicit the transcriptional changes in Δ*rna E. coli* relative to WT. In addition, accumulation of oligoRNAs could modulate transcript stability through an anti-sense mechanism. Due to the fact that *E. coli* encodes both RNase I and oligoribonuclease[37] (whereas *P. aeruginosa* lacks a close homolog of RNase I), additional work is necessary to probe the effect of RNase I depletion on oligoRNA levels. It also remains possible that RNase I influences transcription in a catalytically independent fashion, as catalytically inactivated T2 family RNases modulate certain cellular processes in eukaryotes.[16]

Interestingly, CNPase expression dysregulated many transcripts that were not perturbed by *rna* deletion, indicating that low 2',3'-cNMP levels elicit different effects depending on whether RNase I is present in the cell. One striking example is the Trp biosynthetic cluster *trpDCBA*, which was ~5-7-fold down-regulated in RNase I⁺ *E. coli* expressing CNPase but was not altered by *rna* deletion (**Supplementary Figure S3.13**). A possible explanation for these CNPase-specific effects is altered flux through 2',3'-cNMP catabolic pathways, as CNPase expression inevitably perturbs

other nucleotide/side pools due to 2',3'-cNMP hydrolysis. Conversely, catabolism of 2',3'-cNMPs does not occur in Δ*rna*, as this strain lacks detectable 2',3'-cNMP concentrations,[15] which likely alters nucleotide salvage pathways.

In addition to serving as key components of cellular metabolism, certain nucleotides are established modulators of bacterial chemotaxis and motility. For example, (deoxy)ribonucleosides/tides function as chemoattractants for the Gram-negative bacterium *Vibrio fischeri*,[6] and pyrimidines induce a positive chemotactic response in *E. coli*,[7] suggesting a possible role for other nucleotides and related compounds in chemotaxis and motility. The second messenger c-di-GMP also modulates motility through intracellular interactions with various motility-related proteins,[3b,3d] such as YcgR which decreases flagellar velocity in response to c-di-GMP binding.[38] In *E. coli*, chemotaxis signal transduction is initiated upon binding of a chemoattractant or a chemorepellent to a methyl-accepting chemotaxis protein, which ultimately transmits the signal to the multi-protein flagellar apparatus *via* a phosphorelay cascade mediated by a series of kinases and response-regulators known as the Che proteins.[39] Flagellar motility is fundamental to effective host colonization in several Gram-negative taxa, thus illustrating the therapeutic potential of disrupting chemotaxis and motility.[40] This work identifies 2',3'-cNMPs as a novel component of nucleotide signaling in motility regulation. While the mechanistic links between 2',3'-cNMPs and chemotaxis remain elusive, we previously determined that 2',3'-cNMPs cannot enter the *E. coli* cell,[15] and 2',3'-cAMP has no effect on swimming motility (**Supplementary Figure S3.4**), indicating that 2',3'-cNMPs do not function as exogenous chemo-attractants or -repellents, as these processes require transport across the outer membrane and subsequent binding to periplasmic MCPs.[41] Alternatively, the gene expression data described herein link RNase I and 2',3'-cNMPs to transcriptional regulation of intracellular components of

chemotaxis and flagellar assembly (**Figure 3.3F**). For example, CNPase expression up-regulated transcription of *fliK*, which coordinates flagellar export,[42] and perturbation of either RNase I expression or 2',3'-cNMP levels increased production of *fliR* and *fliE*, encoding the transmembrane flagellar export pore and a component of the flagellar base, respectively[43] (**Figure 3.3F**). Immunoblot experiments also revealed greater abundance of the FliC flagellar filament in Δ*rna* relative to WT (**Figure 3.3D, E** and **Supplementary Figure S3.7**), corroborating the transcriptomic and phenotypic studies. However, FliC expression was not affected by CNPase expression or by Bt-cNMP treatment (**Figure 3.3D, E**). These results indicate that Bt-cNMP treatment impairs motility in Δ*rna* without altering FliC abundance, suggesting that perturbation of 2',3'-cNMP levels impacts proper FliC filamentation and/or modulates other facets of flagellar assembly or chemotaxis signaling, such as flagellar motor proteins.

In addition to regulating flagellar motility, nucleotide metabolism and c-di-GMP signaling govern biofilm production. Biofilms are heterogeneous bacterial communities consisting of extracellular nucleic acids, cellulose, amyloid curli fibers, and poly-*N*-acetyl-*β*-1,6-D-glucosamine (PNAG) which are synthesized by the bacteria.[44] In Gram-negative bacteria, the second messenger c-di-GMP interacts with multiple effectors to regulate production of extracellular polymers and modulate the transition from planktonic growth to sessility.[3b,3d] Notably, the biofilm matrix protects the bacteria against antibiotic drugs and other environmental insults.[44] In addition, biofilms are enriched in metabolically dormant cells known as persisters, which are recalcitrant to conventional antimicrobial therapy.[45] Consequently, further investigation into the aspects of biofilm regulation could potentially inform the discovery of novel antibiotic agents. Previous phenotypic assays and global gene expression analyses with WT and Δ*rna E. coli* demonstrated that 2',3'-cNMPs function in biofilm regulation,[15] thus identifying a novel nucleotide pool

involved in biofilm formation. The present work provides further insight into the role of 2',3'-cNMPs in biofilm production through the discovery that decreasing 2',3'-cNMP levels up-regulated expression of curli assembly gene *csgC* (**Figure 3.4B**), and treatment with Bt-cAMP attenuated biofilm production in RNase I-deficient *E. coli* (**Figure 3.4A**). Moreover, perturbation of either 2',3'-cNMP levels or RNase I expression dysregulated transcription of nucleotide metabolic genes (**Figure 3.5F**), establishing a link to prior studies which indicated that perturbation of *de novo* pyrimidine or purine nucleotide biosynthesis alters biofilm morphology in *E. coli*.[5]

Indirect modulation of c-di-GMP biosynthesis functions as one mechanistic link between primary nucleotide metabolism and biofilm formation, as antimetabolite-mediated disruption of *de novo* purine synthesis decreases the concentration of c-di-GMP and impairs biofilm production in *E. coli*.[5c] In addition, exogenous uracil stimulates diguanylate cyclase Q (DgcQ)-dependent cellulose production in *E. coli* pyrimidine auxotrophs, demonstrating that pyrimidine metabolism also influences c-di-GMP signaling through unknown mechanisms.[5b] Therefore, the finding that disrupting RNase I expression or 2',3'-cNMP concentrations dysregulated the expression of purine and pyrimidine nucleotide metabolic genes (**Figure 3.5F**) suggests that 2',3'-cNMPs modulate nucleotide pools and can perturb c-di-GMP signaling, which likely mediates the biofilm phenotype in the Δ*rna* strain and in WT cells expressing CNPase. The expression of several genes encoding DGCs and c-di-GMP phosphodiesterases (PDEs) also was altered by both CNPase expression and *rna* deletion, further suggesting that 2',3'-cNMPs and RNase I influence biofilm production through modulation of distinct c-di-GMP metabolic enzymes. Although previous investigations in our group determined that the total concentration of c-di-GMP is not perturbed in *E. coli* lacking RNase I,[15] modulation of individual DGCs or c-di-GMP PDEs can impact *E. coli* biofilm

production without altering the global abundance of c-di-GMP.[46] The emerging regulatory functions for RNase I and 2',3'-cNMPs in biofilm formation could inform the discovery of novel adjuvants to remedy chronic, biofilm-associated microbial infections.

The role of RNase I in nucleotide homeostasis further is demonstrated by the elevated 5'-NDP concentrations and the increased sensitivity to adenine in Δ*rna* compared to WT (**Figure 3.5A, B**). The increased 5'-NDP levels in Δ*rna* relative to WT likely is caused by down-regulated expression of the aerobic type Ia RNR NrdAB (**Figure 3.5D**), as RNRs reduce 5'-NDPs to 2'-deoxy 5'-NDPs.[27-28] Expression of the *nrdAB* locus is subject to complex regulation by several transcriptional activators and repressors, including Crp,[19] HNS,[47] Fis,[48] NrdR,[49] and DnaA.[48,50] Although DNA microarray analysis indicated that mRNA levels of these transcription factors were not altered in the Δ*rna* strain, transcription of *nrdAB* is governed by ATP binding to the DnaA transcriptional regulator,[50] suggesting that altered nucleotide levels in Δ*rna* could decrease RNR expression by modulating DnaA binding to the *nrdAB* promoter. In contrast to RNase I-deficient *E. coli*, *nrdAB* transcript abundance was not affected in RNase I[+] cells expressing CNPase (**Figure 3.5A**), demonstrating that low 2',3'-cNMP concentrations are insufficient to alter RNR expression in the presence of RNase I, despite the finding that CNPase-mediated hydrolysis of 2',3'-cNMPs perturbed several nucleotide metabolic genes. Interestingly, the lack of RNase I altered a different set of nucleotide synthesis and salvage genes compared to the reduction of 2',3'-cNMP levels in RNase I[+] cells, further indicating that RNase I and 2',3'-cNMPs regulate distinct aspects of nucleotide homeostasis. The altered expression profile of nucleotide metabolic genes upon perturbation of either RNase I or 2',3'-cNMP levels potentially could modulate the enzymatic activity of nucleotide metabolic enzymes, as many proteins in this class are regulated allosterically or orthosterically by nucleotide binding.[51] One such example is RNR, which is allosterically

regulated by (d)NTP concentrations, both in terms of substrate specificity and reductase activity.[28] Proper RNR function is vital to cell survival; *Saccharomyces cerevisiae* expressing a non-natural *RNR1* allele with a mutated allosteric specificity binding site accumulates nearly 20-fold higher levels of dTTP and dCTP relative to wild-type concentrations,[52] resulting in an elevated mutation rate across the genome.[53] Due to its essentiality, RNR has been targeted in anti-cancer and anti-bacterial chemotherapy.[54] Consequently, these emerging regulatory links between RNase I and RNR suggest that T2 family RNases could be modulated to interfere with RNR function in certain disease states.

The therapeutic relevance of RNase I and 2',3'-cNMPs further is demonstrated by the altered sensitivity to acidic conditions and *β*-lactam treatment in cells with aberrant RNase I expression or 2',3'-cNMP concentrations (**Figures 3.6 and 3.7**). Acid tolerance is critical for colonization of the mammalian gastrointestinal tract by both pathogenic and probiotic bacterial species.[55] Additionally, β-lactam antibiotics are among the most widely prescribed drugs to treat bacterial infections,[56] emphasizing the emerging significance of RNase I and 2',3'-cNMPs in microbial pathogenesis. Perturbation of basal RNase I and 2',3'-cNMP levels potentially influences resistance to low pH and β-lactam treatment through modulation of amino acid (AA) homeostasis, as bacterial survival under these stressors is influenced by AA levels due to the role of proton-dependent AA decarboxylases and DD-transpeptidases in acid tolerance and β-lactam resistance, respectively.[55-56] AA homeostasis intersects with *de novo* nucleotide biosynthesis, as certain AAs are substrates of nucleotide anabolic enzymes such as PyrB and CarA.[57] Indeed, *pyrB* expression was attenuated by RNase I depletion, and *carA* transcription was down-regulated upon CNPase expression (**Figure 3.5**), further suggesting that dysregulation of *de novo* nucleotide biosynthesis alters amino acid levels and impacts resistance to acid and β-lactams. In fact, prior

studies demonstrated that down-regulated expression of the pyrimidine biosynthetic gene *pyrB* (encoding aspartate carbamoyl transferase) increases the L-Asp concentration and thus modulates peptidoglycan cross-linking in *Lactococcus lactis*,[31] alluding to similar links between *de novo* nucleotide synthesis, amino acid homeostasis, and cell wall assembly in *E. coli*.

The present work identifies 2',3'-cNMP pools and RNase I as novel components of bacterial signal transduction with implications in pathogenesis, expanding the scope of nucleotide signaling beyond the paradigmatic 3',5'-cNMPs and c-di-NMPs. Transcriptional profiling and phenotypic investigations, in tandem with controlled manipulation of 2',3'-cNMP levels, have identified bacterial processes regulated by 2',3'-cNMPs and RNase I, providing novel links between nucleotide metabolism and virulence-associated phenotypes such as motility, acid tolerance, and β-lactam resistance. Additional gene expression analyses and bioanalytical experiments suggest that these processes are mediated in part by dysregulated nucleotide homeostasis upon perturbation of RNase I or 2',3'-cNMP levels. Future experiments aim to identify potential 2',3'-cNMP-binding effectors and provide additional mechanistic insight the functions of RNase I and 2',3'-cNMP pools in prokaryotic physiology. In addition, the recombinant CNPase and cell-permeable 2',3'-cNMP derivatives developed herein will enable dissection of processes linked to 2',3'-cNMPs and T2 family RNases across the kingdoms of life.

3.4 Materials and Methods

Bacterial strains, plasmids, general culture conditions, commercial chemicals, and statistical analyses

The *E. coli* strain BW25113 (*lacI*^q $rrnB_{T14}$ $\Delta lacZ_{WJ16}$ *hsdR514* $\Delta araBAD_{AH33}$ $\Delta rhaBAD_{LD78}$)[58] and the RNase I-deficient mutant (*rna*::kan^R; Δrna)[59] in the BW25113 strain background were obtained

as stocks from the Keio collection. We previously validated disruption of the *rna* gene in Δ*rna via* locus-specific PCR amplifications.[15] For bacterial culture, discrete colonies of bacteria on Lysogeny Broth (LB)-agar plates were used to inoculate 3-mL cultures in M9 minimal medium supplemented with 0.4% glucose and 0.2% casamino acids in 15-mL plastic culture tubes, and the cultures were incubated at 37°C overnight with orbital shaking at 200-225 rpm, unless otherwise noted. The starter culture then was sub-cultured 1:100 into fresh medium and incubated under the same conditions, unless otherwise specified. Kanamycin, carbenicillin, and chloramphenicol were used at working concentrations of 25, 100, and 30 μg mL^{-1}, respectively. Plasmids pKT-*CNP* and pKT-*CNP-inact* (H73L/H152L) were obtained by restriction enzyme-based subcloning as previously described.[15] Plasmids pBAD33-*rna*, pKT-*CNP* (cam^R), and pKT-*CNP-inact* (cam^R) were constructed using polymerase incomplete primer extension (PIPE) cloning[60] (for detailed procedures and primer sequences, see **Supplementary Protocols S3.1, 3.2** and **Supplementary Tables S3.1, 3.2**). Analytical standards of nucleoside 5'-monophosphates and nucleoside 5'-diphosphates were purchased as sodium salts from Chem-Impex (Wood Dale, IL, USA). Adenosine 3'-monophosphate (free acid) and uridine 3'-monophosphate (disodium salt) were purchased from Sigma-Aldrich and Chem-Impex, respectively. All data depict at least *n*=3 biological replicates, and a two-sample *t*-test was employed to assess statistical significance, where equal or unequal variance was evaluated using an *F*-test. A *P*-value < 0.05 was considered statistically significant.

General conditions for chemical synthesis and characterization

All reactions were conducted in flame-dried glassware under a nitrogen atmosphere using anhydrous solvents (Drisolv®, MilliporeSigma) and magnetic stirring. For thin layer chromatography (TLC), alumina-backed silica gel 60 F$_{254}$ plates were used with a mobile phase of

6/1.5/1.5/1.5/0.25 EtOAc/MeCN/MeOH/H$_2$O/NH$_4$OH; the plates were visualized under UV light. Diethylaminoethyl (DEAE) Sephadex® A-25 resin (GE Healthcare) was utilized for anion-exchange chromatography. ^{1}H, ^{13}C, and ^{31}P nuclear magnetic resonance (NMR) spectra were recorded on Varian INOVA 400 and Mercury 300 spectrometers. Chemical shifts are reported in ppm and are referenced to the residual solvent signal for ^{1}H and ^{13}C NMR (for ^{1}H NMR: DMSO-d6 = 2.50 ppm; for ^{13}C NMR: DMSO-d6 = 39.52 ppm). ^{31}P NMR chemical shifts are referenced externally to the phosphoric acid signal (0 ppm) in DMSO-d6. *J*-coupling values are reported in hertz (Hz). A Thermo LTQ FTMS was used to collect high-resolution mass spectra (HRMS). NMR spectra for all compounds are provided as **Supplementary spectra** within the **Supplementary Material** (section 3.6).

Preparation of DEAE Sephadex® A-25 bicarbonate form and compound purification

DEAE Sephadex® A-25 resin was hydrated in 100 mM NH$_4$HCO$_3$, poured into a flash chromatography column, and washed with several column volumes of 100 mM NH$_4$HCO$_3$, followed by several column volumes of water. The crude material was dissolved in 5-10 mL of water, and the pH was adjusted to ~7-8 (based on pH paper). This solution was loaded onto the column and eluted with water for several fractions. Subsequently, the ionic strength of the eluent was slowly increased by gradual addition of aqueous NH$_4$HCO$_3$ to elute the desired 2',3'-cNMP analogs as the ammonium salts. The product-containing fractions were concentrated to dryness by lyophilization to remove NH$_4$HCO$_3$.

Conversion of 3'-UMP (disodium salt) to 3'-UMP (mixed pyridinium/sodium salt)

To enhance solubility in pyridine, the disodium salt of 3'-UMP was converted to the mixed pyridinium/sodium salt by passing the nucleotide over a column of Amberlite® IR-120 (pyridinium

form). To this end, the cation-exchange resin (free acid form) was hydrated in 10% (v/v) aqueous pyridine, and washed with several column volumes of 10% (v/v) aqueous pyridine, followed by washing with several column volumes of water. 3'-UMP (disodium) was dissolved in water and slowly passed through the column. The UV-active fractions were lyophilized to obtain 3'-UMP (mixed pyridinium/sodium salt). The presence of the pyridinium cation was confirmed by ^{1}H NMR analysis.

Synthesis of 5'-*O*-butyryl adenosine 2',3'-cyclic monophosphate (Bt-cAMP)

To a vigorously stirred solution of 3'-AMP (free acid) (0.144 mmol) and *N,N*'-dicyclohexyl-4-morpholinecarboxamidine (0.136 mmol) in pyridine (6 mL), butyric anhydride (1.15 mmol) was added dropwise. The reaction was stirred at room temperature (RT) and monitored by silica gel TLC. Additional butyric anhydride was added as necessary to drive the reaction forward. Upon completion of the reaction, the solution was cooled in an ice-water bath and quenched by dropwise addition of MeOH (5 mL). After stirring for 3 h, the reaction was diluted with toluene and concentrated *in vacuo*. The residue was subjected to azeotropic distillation with several additional portions of toluene, subsequently partitioned between H_2O and Et_2O, and washed with several portions of Et_2O and CH_2Cl_2. The aqueous phase was concentrated *in vacuo* by azeotropic distillation with EtOH and the crude material was purified by anion-exchange chromatography over DEAE-Sephadex® A-25 resin, as detailed above. The compound was obtained in 59% yield as the ammonium salt (35.6 mg). **^{1}H NMR** (400 MHz, DMSO-d6) δ 8.39 (s, 1H), 8.18 (s, 1H), 7.56 (s, br, 2H), 7.30 (t, br, J = 47.0 Hz, 4H), 6.18 (d, J = 3.7 Hz, 1H), 5.34 (ddd, J = 9.2, 7.1, 3.7 Hz, 1H), 4.87 (ddd, J = 11.5, 7.0, 4.9 Hz, 1H), 4.41 − 4.27 (m, 2H), 4.24 − 4.11 (m, 1H), 2.25 (td, J = 7.3, 2.0 Hz, 2H), 1.48 (q, J = 7.4 Hz, 2H), 0.82 (t, J = 7.4 Hz, 3H). **^{13}C NMR** (100 MHz, DMSO-d6) δ 172.60, 155.68, 152.30, 149.00, 140.10, 119.00, 88.23, 82.06, 78.49, 75.85, 63.20,

35.12, 17.86, 13.42. **³¹P NMR** (162 MHz, DMSO-d6) δ 16.54. **HRMS** (ESI-, *m/z*) [M-H]⁻ calcd for $C_{14}H_{17}N_5O_7P^-$ 398.0871, found 398.0871.

Synthesis of 5'-*O*-butyryl uridine 2',3'-cyclic monophosphate (Bt-cUMP)

The compound was synthesized from 3'-UMP (mixed pyridinium/sodium salt) in analogy to the preparation of Bt-cAMP. The compound was obtained in 72% yield as the ammonium salt (69.5 mg). **¹H NMR** (400 MHz, DMSO-d6) δ 11.44 (s, br, 1H), 7.69 (d, *J* = 8.0 Hz, 1H), 7.25 (s, br, 4H), 5.83 (d, *J* = 3.0 Hz, 1H), 5.65 (d, *J* = 8.0 Hz, 1H), 4.76 (td, *J* = 7.6, 3.1 Hz, 1H), 4.56 (dt, *J* = 12.6, 6.3 Hz, 1H), 4.29 (dd, *J* = 11.3, 3.2 Hz, 1H), 4.26 − 4.10 (m, 2H), 2.30 (t, *J* = 7.3 Hz, 2H), 1.53 (h, *J* = 7.3 Hz, 2H), 0.87 (t, *J* = 7.4 Hz, 3H). **¹³C NMR** (100 MHz, DMSO-d6) δ 172.58, 163.13, 150.23, 142.52, 102.08, 91.41, 81.99, 78.19, 75.32, 63.50, 35.11, 17.85, 13.42. **³¹P NMR** (162 MHz, DMSO-d6) δ 16.24. **HRMS** (ESI-, *m/z*) [M-H]⁻ calcd for $C_{13}H_{16}N_2O_9P^-$ 375.0599, found 375.0597.

Synthesis of 5'-*O*-benzoyl uridine 2',3'-cyclic monophosphate (Bz-cUMP)

The compound was prepared by treating 3'-UMP (mixed pyridinium/sodium salt) with benzoic anhydride, in analogy to the synthesis of Bt-cUMP. The compound was obtained in 15% yield as the ammonium salt (38.4 mg). **¹H NMR** (400 MHz, DMSO-d6) δ 7.99 (dd, *J* = 8.2, 1.0 Hz, 2H), 7.71 (d, *J* = 8.1 Hz, 1H), 7.69 − 7.63 (m, 1H), 7.54 (t, *J* = 7.7 Hz, 2H), 5.87 (d, *J* = 3.0 Hz, 1H), 5.57 (d, *J* = 8.0 Hz, 1H), 4.81 (td, *J* = 7.6, 3.1 Hz, 1H), 4.70 (ddd, *J* = 13.0, 7.2, 5.6 Hz, 1H), 4.54 (dd, *J* = 11.6, 3.8 Hz, 1H), 4.50 − 4.34 (m, 2H). **¹³C NMR** (100 MHz, DMSO-d6) δ 165.50, 163.13, 150.24, 142.63, 133.54, 129.31, 128.83, 102.04, 91.64, 82.02, 78.33, 75.39, 64.33. **³¹P NMR** (162 MHz, DMSO-d6) δ 16.31. **HRMS** (ESI-, *m/z*) [M-H]⁻ calcd for $C_{16}H_{14}N_2O_9P^-$ 409.0442, found 409.0446.

Motility assay procedure

Bacterial swimming motility was assayed according to a published protocol with minor modifications.[61] Motility medium (1% tryptone, 0.5% NaCl, 0.3% agar) was autoclaved and subsequently cooled in a 60°C water bath. The medium then was supplemented with the cell-permeable 2',3'-cNMP derivatives (or other additives), distributed into 60 x 15 mm petri dishes (10 mL of medium per dish), and allowed to solidify at room temperature overnight. Bacteria from an overnight culture in LB were diluted 1:100 into fresh LB and 2.5 µL were plated onto the surface of each plate. The plates were incubated at 30°C for 24 h. The area of the motility zone was quantified by measuring the zone across three diameters using a digital caliper.

Quantitative western blot analysis of flagellar protein expression

Cultures of WT and Δ*rna* (1.5 mL) were grown in 15-mL plastic culture tubes in the presence of Bt-cAMP (1 mM), Bt-cUMP (1 mM), or vehicle (water). Upon reaching OD_{600} ~0.6, 1 mL of culture was harvested by centrifugation at 10,000*g* for 5 min at RT. The pellet was flash frozen in liquid nitrogen and stored at -80°C. For western blot analysis, the cells were lysed into 100 µL of BugBuster® (EMD-Millipore) and the insoluble protein fraction was isolated according to the manufacturer's protocol. The protein was resuspended in 100 µL of sodium phosphate buffer (50 mM, pH 7.4), and the protein concentration of each sample was quantified in triplicate *via* Bradford assay (Bio-Rad Protein Assay Dye) to normalize protein loading for subsequent western blot analysis. The remaining protein sample was diluted 1:1 with 2x Laemmli sample buffer and denatured by heating at 95°C for 10 min. The samples then were diluted ~10-fold (precise dilution factor determined based on the Bradford assay) to normalize for protein concentration and 15 µL were separated by SDS-PAGE on Criterion™ TGX™ precast midi protein gels (4-20% acrylamide, Bio-Rad) at 4°C. Proteins were transferred to nitrocellulose membranes (0.2 µm, Bio-

Rad) using the Mixed MW mode on the Trans-Blot® Turbo™ transfer system according to the manufacturer's instructions. The blots were processed essentially as described in the Opti-4CN™ Substrate Kit (Bio-Rad). Briefly, blots were blocked by incubation in 3% Blocker solution (Bio-Rad) for 2 h at RT prior to 12 h incubation at 4°C with rabbit polyclonal anti-flagellin primary antibody (15000-fold dilution, abcam 93713). The blots then were incubated for 1 h at RT with goat anti-rabbit secondary antibody conjugated to horseradish peroxidase (12000-fold dilution, abcam 205718), and bands were detected by incubation with the Opti-4CN™ substrate for 15 min at RT. Blots were imaged using an Epson Perfection V600 photo scanner operating in the Professional Setting. The band intensity in the resulting 16-bit gray-scale tif image files was quantified using ImageJ (National Institutes of Health). The identity of the FliC band was confirmed by western blot analyses using strain BW25113 *fliC*::kan[R] (Δ*fliC*) (obtained from the Keio collection)[59] as a negative control and strain K12 W3110 overexpessing FliC from plasmid pCA24N-*fliC* (obtained from the ASKA collection)[62] as a positive control. For images of all western blots, see **Supplementary Figure S3.7**. Disruption of the *fliC* gene with the kan[R] resistance cassette in Δ*fliC* was confirmed by PCR (**Supplementary Figure S3.8** and **Supplementary Table S3.3**).

Congo red biofilm assay procedure

The Congo red assay was based on a published protocol.[63] LB (3 mL) was inoculated with a single colony of BW25113 (WT) or Δ*rna* from LB-agar plates and cultured overnight in 15-mL plastic culture tubes. Each overnight culture was used to inoculate (1:50) 10 mL YESCA (1% casamino acids, 0.12% yeast extract) containing 0.0025% Congo red in 50-mL Celltreat® conical tubes (sterile, polypropylene) (lids left loose for gas exchange). After reaching an OD_{600} ~0.3-0.4, 1 mL of each culture was transferred to a 1.6-mL Eppendorf tube and either treated with vehicle (YESCA

medium) or with 500 μM Bt-cAMP. The cultures were incubated for 24 h at room temperature without shaking (lids left open and tubes were covered in plastic wrap and foil). For each culture, 200 μL of supernatant were transferred to a 96-well microplate (Corning® Costar®) following centrifugation at 12000g for 15 min. Biofilm formation was quantified by recording the absorbance at 500 nm. To normalize for cell density, each culture was disturbed by pipetting and 200 μL were transferred to a 96-well microplate prior to recording the OD_{600} using a microplate reader.

Extraction of NDPs

Static cultures of WT and Δrna (100-mL) were incubated at RT in 500-mL Ultra Yield^{TM} flasks (Thomson Instrument Company, Oceanside, CA, USA). Additional cultures of WT pKT-CNP and WT pKT-CNP-$inact$ were incubated under the same conditions, and protein expression was induced with anhydrotetracycline (25 ng mL^{-1}) after 1.5 h of incubation. Upon reaching OD_{600} ~0.7-0.8, cells were harvested from 10-mL of culture by centrifugation at 2800g for 10 min at RT. Cells were flash frozen in liquid nitrogen and kept at -80°C until nucleotide extraction. For NDP extraction, the cell pellet was lysed by sonication on ice into 1 mL of aqueous formic acid (1 M), in analogy to published procedure.[64] The resulting lysate was freeze-dried and resuspended in 300 μL of 50 mM sodium phosphate buffer (pH 7.4) containing 0.5 μM 8-bromo adenosine 3',5'-cyclic monophosphate (8-Br 3',5'-cAMP) as internal standard (IS). The extracts were centrifuged at 12000g for 25 min at 4°C prior to LC-MS/MS analysis.

LC-MS/MS quantification of NDPs

Nucleotides were separated and quantified by LC-MS/MS analogously to published procedure.[65] The mobile phase A consisted of 5 mM triethylammonium acetate (pH 7) in water and mobile phase B consisted of 5 mM triethylammonium acetate (pH 7) in 1:1 water/acetonitrile. The flow

rate and chromatography method have been published previously.[65] Nucleotides were detected in negative-ion mode using the parent mono-anion and quantification was based on fragment ions with m/z 304.1, 305.1, 328.1, and 344.1 for CDP, UDP, ADP, and GDP, respectively. An internal standard (IS) method was employed for nucleotide quantification using 8-Br 3',5'-cAMP (0.5 µM) as the IS. The IS was detected in negative-ion mode using the parent mono-anion and quantified using the fragment ions with m/z 212.0 and 214.0 for 8-^{79}Br 3',5'-cAMP and 8-^{81}Br 3',5'-cAMP, respectively. For each nucleotide analyte a calibration curve was obtained using authentic standards spanning a concentration range from 0.02-20 µM. The calibration data then were fit to a linear regression model to calculate the concentration in the extract. To normalize for differences in bacterial culture density, nucleotide concentrations were adjusted using a value of 11.1 x 10^8 cells mL^{-1} OD$_{600}$$^{-1}$ (where mL and OD$_{600}$ are the volume and 600 nm optical density of the culture sample, respectively); this conversion factor has been determined previously for *E. coli* grown in M9 minimal medium supplemented with glucose [66].

Extraction of NMPs ± adenine treatment

Cultures of WT and Δ*rna* (10-mL) were incubated in 50-mL conical tubes (VWR International) in the presence or absence of 1 mM adenine. Upon reaching OD$_{600}$ ~0.7, cells were harvested by centrifugation at 2800g for 10 min at RT. Cells were flash frozen in liquid nitrogen and kept at -80°C until nucleotide extraction. For NMP extraction, the cell pellet was lysed by sonication on ice into 500 µL of acetonitrile/methanol/H$_2$O (2/2/1, v/v/v) in analogy to published procedure.[15] The resulting lysate was concentrated to dryness using a vacuum centrifuge and resuspended in 250 µL of 50 mM sodium phosphate buffer (pH 7.4) containing 0.5 µM IS. The extracts were centrifuged at 12000g for 25 min at 4°C prior to LC-MS/MS analysis.

LC-MS/MS quantification of NMPs

Quantification of NMPs was performed using published chromatographic conditions.[15] Nucleotides were detected in positive-ion mode using the parent mono-cation and quantification was based on nucleobase fragment ions with m/z 112.1, 113.1, 136.1, 152.1, 212.0 and 214.0 for CMP, UMP, AMP, GMP, 8-^{79}Br 3',5'-cAMP, and 8-^{81}Br 3',5'-cAMP, respectively. Nucleotides were quantified using an IS methods, as detailed above for quantification of NDPs.

Adenine sensitivity assay

Cultures of WT, Δ*rna*, WT pKT-*CNP*, and WT pKT-*CNP-inact* (2-mL) were incubated in 24-well microtiter plates (Corning Costar, sterile, untreated, polystyrene) in the presence or absence of 0.25, 0.50, or 1 mM adenine, and the OD_{600} was recorded hourly to assess the bacteriostatic effect of adenine on each strain. Additional parallel cultures of WT and Δ*rna* containing the same adenine concentrations were prepared in the presence of 0.25 mM guanosine to assess the potential of guanosine to rescue adenine-induced growth inhibition.[29]

Acid sensitivity assay

Acid resistance was assayed essentially as described previously.[67] Cultures of WT, Δ*rna*, WT pKT-*CNP*, and WT pKT-*CNP-inact* (3-mL) were incubated to OD_{600} ~0.2 and treated with anhydrotetracycline (25 ng mL^{-1}) to induce protein expression. Incubation was continued to OD_{600} ~0.6, and the cultures then were inoculated 1:20 into 2 mL of fresh M9 (0.4% glucose, 0.2% casamino acids) at either pH 2.5 (pH adjusted using HCl) or pH 7 (control) in a 24-well microtiter plate (VWR International, sterile, untreated, polystyrene). The cultures were incubated for 2 h and volume-normalized colony-forming units (CFU mL^{-1}) were quantified by 6x6 drop plating[68] to determine the survival rate of each strain at pH 2.5 relative to pH 7.

Carbenicillin sensitivity assay

Carbenicillin sensitivity was determined spectrophotometrically by monitoring reduction of the tetrazolium XTT as a reporter of cell viability, as described previously.[69] Cultures of WT, Δ*rna*, WT pKT-*CNP* (cam^R), and WT pKT-*CNP-inact* (cam^R) (200 μL) were incubated in a 96-well microtiter plate (VWR International, sterile, untreated, polystyrene) in the presence of various carbenicillin concentrations spanning 0.05 to 25.6 μg mL^{-1} in two-fold serial dilutions. Additional cultures were treated with either 0 or 100 μg mL^{-1} carbenicillin as positive and negative growth controls, respectively. The plate was then incubated for 6 hours and cell viability was quantified using the XTT Cell Proliferation Kit II (Roche), essentially as described by the manufacturer. Activated XTT (50 μL) was added to each 200-μL culture and the absorbance at 450 nm was recorded on a microplate reader.

Transcriptome profiling of WT and Δ*rna*

Analysis of mRNA transcript levels were quantified by the Emory Integrated Genomics Core and analyzed by the Emory Integrated Computational Core. Six *E. coli* pellets (three biological replicates of WT and three of Δ*rna*) were submitted for extraction and expression profiling on the Affymetrix *E. coli* Genome 2.0 Array. RNA was extracted using Qiagen miRNEasy kit (with on-column DNase treatment). Cells were lysed using 700 μL Qiazol with 100 mg acid-washed beads (150-600 μm) on the Qiagen TissueLyser at 30 Hz for 5 min. RNA was eluted in 30 μL of nuclease free water and 1 μL was used to determine the concentration on a Nanodrop 1000. An additional 1 μL was used to assess sample profiles on the Agilent 2100 using the RNA 6000 Nano assay.

Whole-Transcript Expression Analysis (Gene ST Arrays) was performed as follows. RNA (10 ng) was processed according to the GeneChip® WT Pico Reagent Kit protocol. Labeled cDNA was hybridized to the *E. coli* Genome 2.0 microarray for 16-18 hours at 45°C. Hybridized microarrays were washed and stained on an Affymetrix GeneChip Fluidics Station 450 using the

appropriate chip dependent fluidics script. Intensity data were extracted using an Affymetrix 7G scanner and the Command Console software suite.

The obtained expression data from the microarray experiment were analyzed using the 'limma' package in R/Bioconductor (http://www.r-project.org). The raw data were $\log_2$ transformed and normalized across the samples by Robust Multi-array Average (RMA) normalization. The differentially expressed genes were identified on the basis of Benjamini-Hochberg (BH) multiple test adjusted *P* values (*i.e.* false-discovery rate; FDR) and fold changes (the increase in number of transcript copies). Genes with an FDR value <0.05 were considered significantly differentially expressed. Heat maps based on the z-score-normalized probe signal were created using Heatmapper (www.heatmapper.ca).[70]

Gene expression data obtained from the microarray experiment have been submitted to ArrayExpress at EMBL-EBI (http://www.ebi.ac.uk/arrayexpress/) under accession number E-MTAB-6095.

Transcriptome profiling of WT pKT-*CNP* and WT pKT-*CNP-inact*

Total RNA (2 µg) was extracted and subjected to ribosomal RNA depletion using the Bacterial Ribo-Zero™ rRNA Removal Kit (EpiCentre, Madison, WI, USA) following the manufacturer's instructions. Complementary DNA (cDNA) libraries were prepared with the ScriptSeq v2 RNA-Seq library preparation kit (Epicentre, Madison, WI, USA). Briefly, the rRNA-depleted total RNA sample was fragmented using an RNA fragmentation solution, and the fragmented RNA was reverse transcribed into cDNA using random hexamer primers containing a tagging sequence at the 5′ end; 3′-tagging was accomplished using a terminal-tagging oligo (TTO) consisting of a random hexamer flanked by a 5'-tag sequence and a blocked 3'-terminus. The di-tagged cDNA

was purified using AMPure™ XP beads (Agencourt, Beckmann-Coulter, USA), and PCR amplified to add index and sequencing adapters. After amplification, the final library was purified using AMPureXP^TM beads, and the final pooled libraries were sequenced on the Illumina HiSeq3000 system in a Single-end (SE) 150 cycle format. Each sample was sequenced to approximate depth of 8-12 million reads. For data processing and statistical analysis, RNA-Seq reads were aligned to the GenBank *E. coli* BW25113 genomic reference sequence (CP009273.1) using the STAR Aligner v2.5.2b [71] and transcript abundance was estimated using htseq-count v0.6.1p1.[72] Differential expression analysis was performed with DESeq2.[73] The NCBI index (ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCA/000/750/555/GCA_000750555.1_ASM75055v1) was utilized for gene annotation.

Gene expression data have been deposited to the NCBI Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/) under accession number GSE114871.

3.5 References

(1) (a) Denninger, J.; Marletta, M. *Biochim. Biophys. Acta* **1999**, *1411*, 334-350; (b) Domek-Lopacinska, K.; Stroksznajder, J. *Journal of Physiology and Pharmacology* **2005**, *56*, 15-34.

(2) (a) Butcher, R. *New Engl. J. Med.* **1968**, *279*, 1378-1384; (b) McDonough, K.; Rodriguez, A. *Nat. Rev. Microbiol.* **2012**, *10*, 27-38.

(3) (a) Danilchanka, O.; Mekalanos, J. *Cell* **2013**, *154*, 962-970; (b) Hengge, R. *Nat. Rev. Microbiol.* **2009**, *7*, 263-273; (c) Kalia, D.; Merey, G.; Nakayama, S.; Zheng, Y.; Zhou, J.; Luo, Y.; Guo, M.; Roembke, B.; Sintim, H. *Chem. Soc. Rev.* **2013**, *42*, 305-341; (d) Romling, U.; Galperin, M.; Gomelsky, M. *Microbiol. Mol. Biol. Rev.* **2013**, *77*, 1-52.

(4) (a) Boehm, A.; Steiner, S.; Zaehringer, F.; Casanova, A.; Hamburger, F.; Ritz, D.; Keck, W.; Ackermann, M.; Schirmer, T.; Jenal, U. *Mol. Microbiol.* **2009**, *72*, 1500-1516; (b) Monds, R.;

Newell, P.; Wagner, J.; Schwartzman, J.; Lu, W.; Rabinowitz, J.; O'Toole, G. *J. Bacteriol.* **2010**, *192*, 3011-3023.

(5) (a) Attila, C.; Ueda, A.; Wood, T. *Appl. Microbiol. Biotechnol.* **2009**, *82*, 525-533; (b) Garavaglia, M.; Rossi, E.; Landini, P. *PLOS ONE* **2012**, *7*, doi:10.1371/journal.pone.0031252; (c) Antoniani, D.; Rossi, E.; Rinaldo, S.; Bocci, P.; Lolicato, M.; Paiardini, A.; Raffaelli, N.; Cutruzzola, F.; Landini, P. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7325-7336.

(6) DeLoney-Marino, C.; Wolfe, A.; Visick, K. *Appl. Environ. Microbiol.* **2003**, *69*, 7527-7530.

(7) Liu, X.; Parales, R. *J. Bacteriol.* **2008**, *190*, 972-979.

(8) Wade, H. *Biochem. J* **1961**, *78*, 457-472.

(9) Jackson, E. *Am. J. Physiol. Renal Physiol.* **2011**, *301*, F1160-F1167.

(10) Azarashvili, T.; Krestinina, O.; Galvita, A.; Grachev, D.; Baburina, Y.; Stricker, R.; Evtodienko, Y.; Reiser, G. *Am. J. Physiol. Cell Physiol.* **2009**, *296*, C1428-C1439.

(11) Kosmacz, M.; Luzarowski, M.; Kerber, O.; Leniak, E.; Gutierrez-Beltran, E.; Beltran, J. C. M.; Gorka, M.; Szlachetko, J.; Veyel, D.; Graf, A.; Skirycz, A. *Plant Physiol.* **2018**, *177*, 411-421.

(12) Jia, X.; Fontaine, B. M.; Strobel, F.; Weinert, E. E. *Biomolecules* **2014**, *4*, 1070-1092.

(13) Bordeleau, E.; Oberc, C.; Ameen, E.; da Silva, A.; Yan, H. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 4520-4522.

(14) Liu, A.; Yu, Y.; Sheng, Q.; Zheng, X.; Yang, J.; Li, P.; Shi, M.; Zhou, B.; Zhang, Y.; Chen, X. *ACS Chem. Biol.* **2016**, *11*, 2414-2419.

(15) Fontaine, B. M.; Martin, K. S.; Garcia-Rodriguez, J. M.; Jung, C.; Southwell, J. E.; Jia, X.; Weinert, E. E. *Biochem. J* **2018**, *478*, 1491-1506.

(16) Luhtala, N.; Parker, R. *Trends Biochem. Sci.* **2010**, *35*, 253-259.

(17) Wheeler, J.; Freihat, L.; Irving, H. *BMC Biotechnol.* **2013**, *13*, doi: 10.1186/1472-6750-1113-1197.

(18) Gravel, M.; Robert, F.; Kottis, V.; Gallouzi, I.; Pelletier, J.; Braun, P. *J. Neurosci. Res.* **2009**, *87*, 1069-1079.

(19) Zheng, D.; Constantinidou, C.; Hobman, J.; Minchin, S. *Nucleic Acids Res.* **2004**, *32*, 5874-5893.

(20) Gorke, B.; Stulke, J. *Nat. Rev. Microbiol.* **2008**, *6*, 613-624.

(21) Sourjik, V. *Trends Microbiol.* **2004**, *12*, 569-576.

(22) Chilcott, G.; Hughes, K. *Microbiol. Mol. Biol. Rev.* **2000**, *64*, 694-+.

(23) Minamino, T.; Imada, K. *Trends Microbiol.* **2015**, *23*, 267-274.

(24) Adler, J.; Hazelbauer, G.; Dahl, M. *J. Bacteriol.* **1973**, *115*, 824-847.

(25) Wood, T.; Barrios, A.; Herzberg, M.; Lee, J. *Appl. Microbiol. Biotechnol.* **2006**, *72*, 361-367.

(26) Oropeza, R.; Salgado-Bravo, R.; Calva, E. *Microbiology* **2015**, *161*, 903-913.

(27) Brignole, E.; Ando, N.; Zimanyi, C.; Drennan, C. *Biochem. Soc. Trans.* **2012**, *40*, 523-530.

(28) Torrents, E. *Front. Cell. Infect. Microbiol.* **2014**, *4*, doi: 10.3389/fcimb.2014.00052.

(29) Levine, R.; Taylor, M. *J. Bacteriol.* **1982**, *149*, 923-930.

(30) Shive, W.; Roberts, E. *J. Biol. Chem.* **1946**, *162*, 463-471.

(31) Solopova, A.; Formosa-Dague, C.; Courtin, P.; Furlan, S.; Veiga, P.; Pechoux, C.; Armalyte, J.; Sadauskas, M.; Kok, J.; Hols, P.; Dufrene, Y.; Kuipers, O.; Chapot-Chartier, M.; Kulakauskas, S. *J. Biol. Chem.* **2016**, *291*, 11323-11336.

(32) Gaal, T.; Bartlett, M.; Ross, W.; Turnbough, C.; Gourse, R. *Science* **1997**, *278*, 2092-2097.

(33) Lorkovic, Z.; Kirk, D.; Klahre, U.; Hemmings-Mieszczak, M.; Filipowicz, W. *RNA* **2000**, *6*, 1610-1624.

(34) Cannistraro, V.; Kennell, D. *J. Bacteriol.* **1991**, *173*, 4653-4659.

(35) Kivity-Vogel, T. E., D. *Biochim. Biophys. Acta* **1967**, *138*, 66-75.

(36) Goldman, S.; Sharp, J.; Vvedenskaya, I.; Livny, J.; Dove, S.; Nickels, B. *Mol. Cell* **2011**, *42*, 817-825.

(37) (a) Ghosh, S.; Deutscher, M. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4372-4377; (b) Niyogi, S.; Datta, A. *J. Biol. Chem.* **1975**, *250*, 7307-7312.

(38) (a) Fang, X.; Gomelsky, M. *Mol. Microbiol.* **2010**, *76*, 1295-1305; (b) Paul, K.; Nieto, V.; Carlquist, W.; Blair, D.; Harshey, R. *Mol. Cell* **2010**, *38*, 128-139.

(39) Terashima, H.; Kojima, S.; Homma, M.; Jeon, K. *Int. Rev. Cell Mol. Biol.* **2008**, *270*, 39-85.

(40) Josenhans, C.; Suerbaum, S. *Int. J. Med. Microbiol.* **2002**, *291*, 605-614.

(41) Alexander, R. P.; Zhulin, I. B. *Proc. Natl. Acad. Sci. USA* **2007**, *104*, 2885-2890.

(42) Kawagishi, I.; Homma, M.; Williams, A.; Macnab, R. *J. Bacteriol.* **1996**, *178*, 2954-2959.

(43) (a) Malakooti, J.; Ely, B.; Matsumura, P. *J. Bacteriol.* **1994**, *176*, 189-197; (b) Muller, V.; Jones, C.; Kawagishi, I.; Aizawa, S.; Macnab, R. *J. Bacteriol.* **1992**, *174*, 2298-2304.

(44) Flemming, H.; Wingender, J. *Nat. Rev. Microbiol.* **2010**, *8*, 623-633.

(45) Wood, T.; Knabel, S.; Kwan, B. *Appl. Environ. Microbiol.* **2013**, *79*, 7116-7121.

(46) Sarenko, O.; Klauck, G.; Wilke, F.; Pfiffer, V.; Richter, A.; Herbst, S.; Kaever, V.; Hengge, R. *MBio* **2017**, *8*.

(47) Cendra, M.; Juarez, A.; Madrid, C.; Torrents, E. *J. Bacteriol.* **2013**, *195*, 4255-4263.

(48) Augustin, L.; Jacobson, B.; Fuchs, J. *J. Bacteriol.* **1994**, *176*, 378-387.

(49) Torrents, E.; Grinberg, I.; Gorovitz-Harris, B.; Lundstrom, H.; Borovok, I.; Aharonowitz, Y.; Sjoberg, B.; Cohen, G. *J. Bacteriol.* **2007**, *189*, 5012-5021.

(50) Olliver, A.; Saggioro, C.; Herrick, J.; Sclavi, B. *Mol. Microbiol.* **2010**, *76*, 1555-1571.

(51) (a) Cho, B.; Federowicz, S.; Embree, M.; Park, Y.; Kim, D.; Palsson, B. *Nucleic Acids Res.* **2011**, *39*, 6456-6464; (b) Long, C.; Pardee, A. *J. Biol. Chem.* **1967**, *242*, 4715-4721; (c) Serina, L.; Blondin, C.; Krin, E.; Sismeiro, O.; Danchin, A.; Sakamoto, H.; Gilles, A.; Barzu, O. *Biochemistry* **1995**, *34*, 5066-5074; (d) Zhou, G.; Smith, J.; Zalkin, H. *J. Biol. Chem.* **1994**, *269*, 6784-6789.

(52) Kumar, D.; Abdulovic, A.; Viberg, J.; Nilsson, A.; Kunkel, T.; Chabes, A. *Nucleic Acids Res.* **2011**, *39*, 1360-1371.

(53) Watt, D. L.; Buckland, R. J.; Lujan, S. A.; Kunkel, T. A.; Chabes, A. *Nucleic Acids Res.* **2016**, *44*, 1669-1680.

(54) (a) Shao, J.; Liu, X.; Zhu, L.; Yen, Y. *Expert Opin. Ther.Targets* **2013**, *17*, 1423-1437; (b) Tholander, F.; Sjoberg, B. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 9798-9803.

(55) Kanjee, U.; Houry, W.; Gottesman, S. *Annu. Rev. Microbiol.* **2013**, *67*, 65-81.

(56) Fisher, J.; Meroueh, S.; Mobashery, S. *Chem. Rev.* **2005**, *105*, 395-424.

(57) (a) Rosenbusch, J.; Weber, K. *J. Biol. Chem.* **1971**, *246*, 1644-1657; (b) Thoden, J.; Raushel, F.; Benning, M.; Rayment, I.; Holden, H. *Acta Cryst. D-Biol. Cryst.* **1999**, *55*, 8-24.

(58) Datsenko, K.; Wanner, B. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 6640-6645.

(59) Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K.; Tomita, M.; Wanner, B.; Mori, H. *Mol. Syst. Biol.* **2006**, *2*, doi: 10.1038/msb4100050.

(60) Klock, H.; Koesema, E.; Knuth, M.; Lesley, S. *Proteins-Structure Function and Bioinformatics* **2008**, *71*, 982-994.

(61) Hussain, M.; Zhang, H.; Xu, J.; Liu, Q.; Jiang, Z.; Zhang, L. *J. Bacteriol.* **2008**, *190*, 1045-1053.

(62) Kitagawa, M.; Ara, T.; Arifuzzaman, M.; Ioka-Nakamichi, T.; Inamoto, E.; Toyonaga, H.; Mori, H. *DNA Res.* **2005**, *12*, 291-299.

(63) Spiers, A.; Bohannon, J.; Gehrig, S.; Rainey, P. *Mol. Microbiol.* **2003**, *50*, 15-27.

(64) Ikuma, H.; Tetley, R. *Plant Physiol.* **1976**, *58*, 320-323.

(65) Chen, P.; Liu, Z.; Liu, S.; Xie, Z.; Aimiuwu, J.; Pang, J.; Klisovic, R.; Blum, W.; Grever, M.; Marcucci, G.; Chan, K. *Pharm. Res.* **2009**, *26*, 1504-1515.

(66) Volkmer, B.; Heinemann, M. *PLOS ONE* **2011**, *6*, doi: 10.1371/journal.pone.0023126.

(67) Lu, P.; Ma, D.; Chen, Y.; Guo, Y.; Chen, G.; Deng, H.; Shi, Y. *Cell Res.* **2013**, *23*, 635-644.

(68) Chen, C.; Nace, G.; Irwin, P. *J. Microbiol. Methods* **2003**, *55*, 475-479.

(69) Roslev, P.; King, G. *Appl. Environ. Microbiol.* **1993**, *59*, 2891-2896.

(70) Babicki, S.; Arndt, D.; Marcu, A.; Liang, Y.; Grant, J.; Maciejewski, A.; Wishart, D. *Nucleic Acids Res.* **2016**, *44*, W147-W153.

(71) Dobin, A.; Davis, C.; Schlesinger, F.; Drenkow, J.; Zaleski, C.; Jha, S.; Batut, P.; Chaisson, M.; Gingeras, T. *Bioinformatics* **2013**, *29*, 15-21.

(72) Anders, S.; Pyl, P.; Huber, W. *Bioinformatics* **2015**, *31*, 166-169.

(73) Love, M.; Huber, W.; Anders, S. *Genome Biology* **2014**, *15*, DOI: 10.1186/s13059-13014-10550-13058.

3.6 Supplementary Material
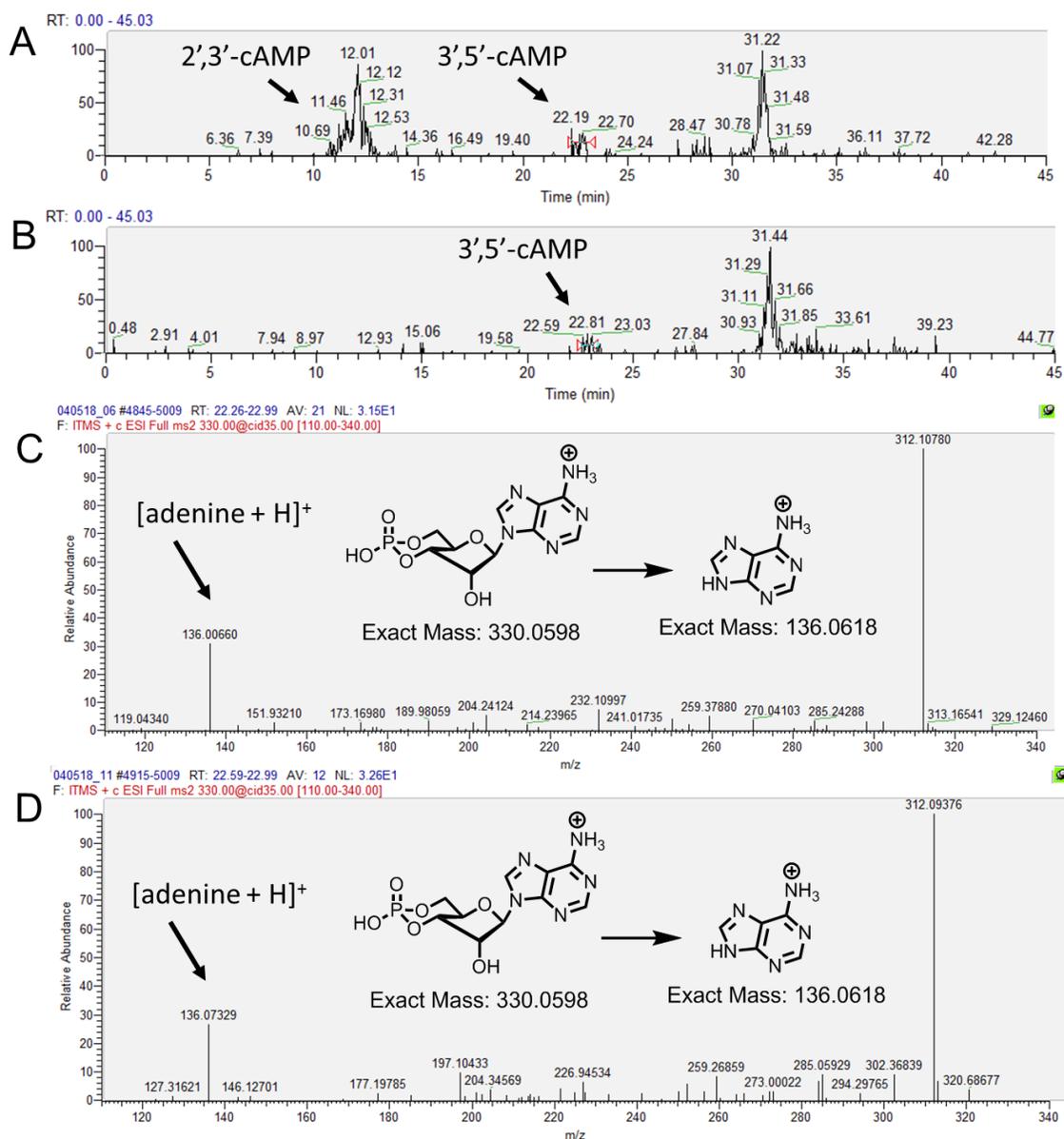
Supplementary Figure S3.1. The intracellular concentration of adenosine 3',5'-cyclic monophosphate (3',5'-cAMP) is barely detectable in both BW25113 (WT) and the RNase I-deficient mutant (Δ*rna*). (**A, B**) LC-MS/MS chromatograms depicting the adenosine 3',5'-cyclic monophosphate (3',5'-cAMP) peak in representative extracts from BW25113 WT (**A**) and Δ*rna* (**B**). 3',5'-cAMP was detected based on the protonated parent cation with *m/z* 330.1 and quantified by the protonated adenine cation with *m/z* 136.1 (for details, see **Materials and Methods; section 3.4**). (**C, D**) MS/MS spectra of the 3',5'-cAMP peak (eluting at ~22 min) which depict the protonated adenine fragment ion with *m/z* 136.1 in WT (**C**) and Δ*rna* (**D**). These

data demonstrate that 3',5'-cAMP is not eliciting the transcriptional changes in Δ*rna* relative to WT.
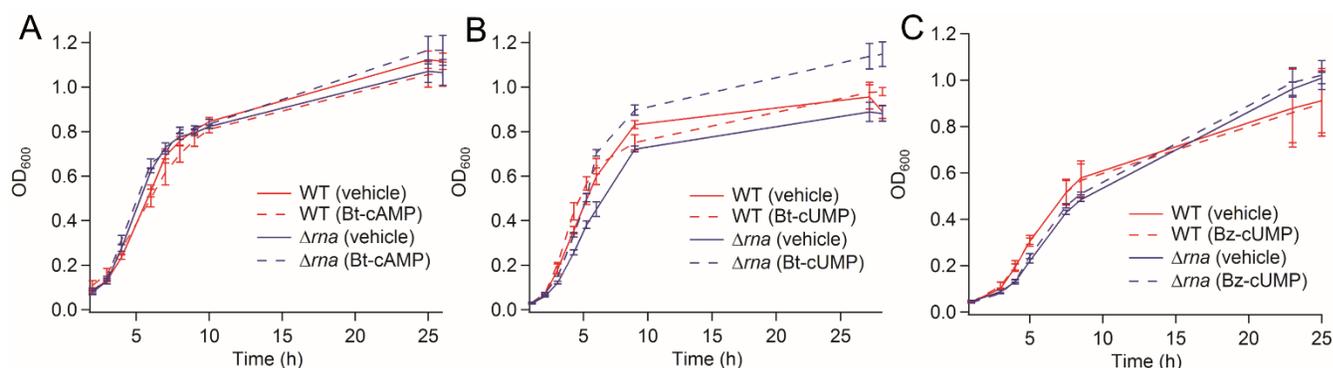
Supplementary Figure S3.2. Treatment with Bt-cNMPs does not affect *E. coli* growth. WT and Δ*rna* were cultured in 200 μL of M9 minimal (0.4% glucose, 0.2% casamino acids) in a 96-well microtiter plate at 37°C with 200 rpm shaking. Cultures were treated with either DMSO (vehicle), 1 mM Bt-cAMP (**A**), 1 mM Bt-cUMP (**B**), or 1 mM Bz-cUMP (**C**) and the 600 nm absorbance (OD$_{600}$) was recorded periodically using a microplate reader. Data represent the mean ± standard deviation of four biological replicates.
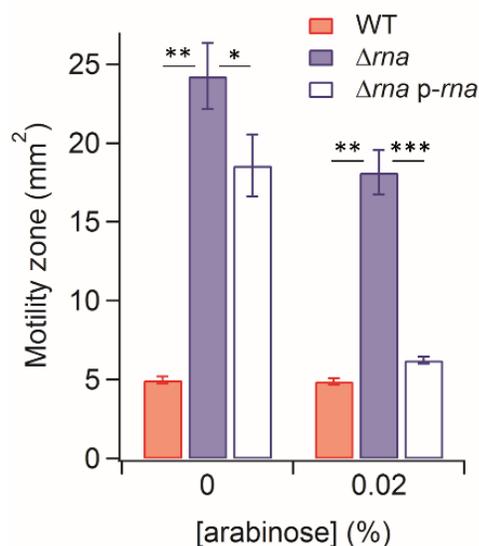
Supplementary Figure S3.3. Complementation of Δ*rna* with plasmid pBAD33-*rna* (p-*rna*) attenuates the hypermotile phenotype in RNase I-deficient *E. coli*. Swimming motility was assayed on M9 motility plates supplemented with 0.4% glycerol and either 0 or 0.02% arabinose to induce RNase I expression from the pBAD33 vector. Data represent the mean ± standard deviation of three biological replicates (* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$). For detailed motility assay procedure, see **Materials and Methods** (**section 3.4**).
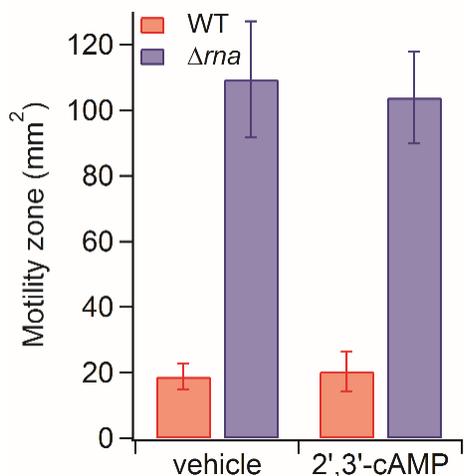
Supplementary Figure S3.4. Treatment with 2',3'-cAMP does not impact swimming motility. Swimming motility was assayed for WT and Δ*rna* in the presence of 2',3'-cAMP (250 µM) or H$_2$O (vehicle). Data represent the mean ± standard deviation of three biological replicates. For detailed motility assay procedure, see **Materials and Methods** (**section 3.4**).
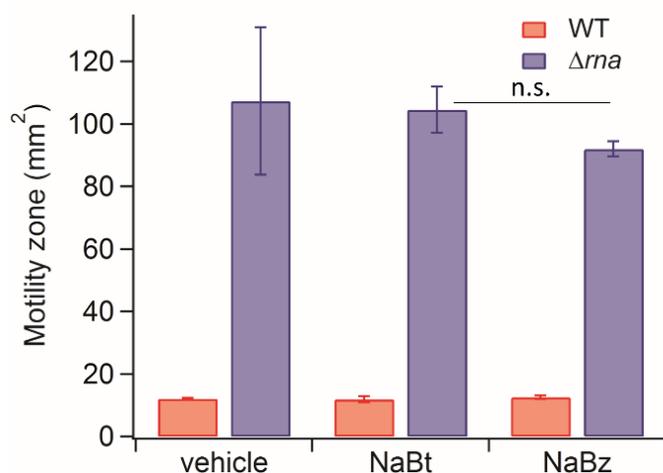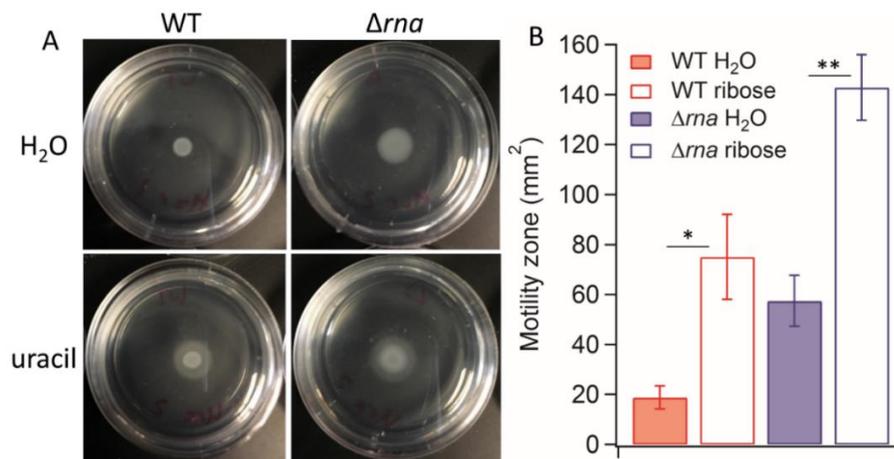

Supplementary Figure S3.5. Treatment with either sodium butyrate or sodium benzoate does not affect *E. coli* swimming motility. Swimming motility was assayed for WT and Δ*rna* in the presence of 500 µM sodium butyrate (NaBt), 500 µM sodium benzoate (NaBz), or H$_2$O (vehicle). Data represent the mean ± standard deviation of four biological replicates (n.s.; not significant). For detailed motility assay procedure, see **Materials and Methods** (**section 3.4**).
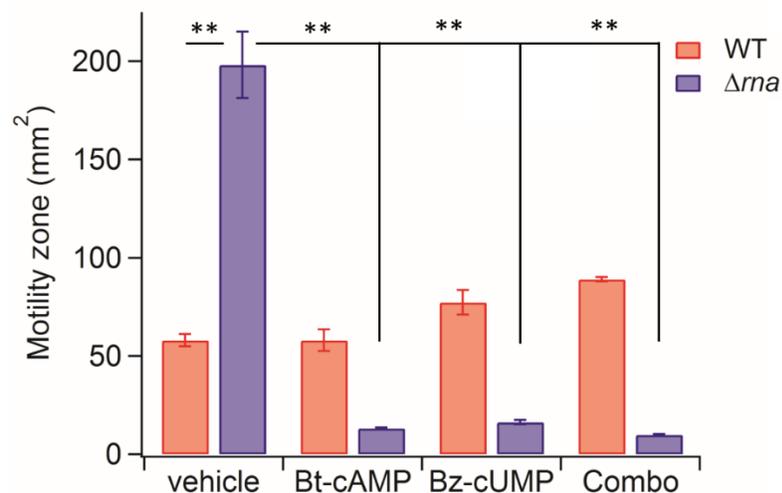
Supplementary Figure S3.6. Treatment with Bt-cAMP (1 mM) and/or Bz-cUMP (0.18 mM) impairs swimming motility in Δ*rna*. Data represent the mean ± standard deviation of three biological replicates (* $P < 0.05$; ** $P < 0.01$). For detailed motility assay procedure, see **Materials and Methods** (**section 3.4**).
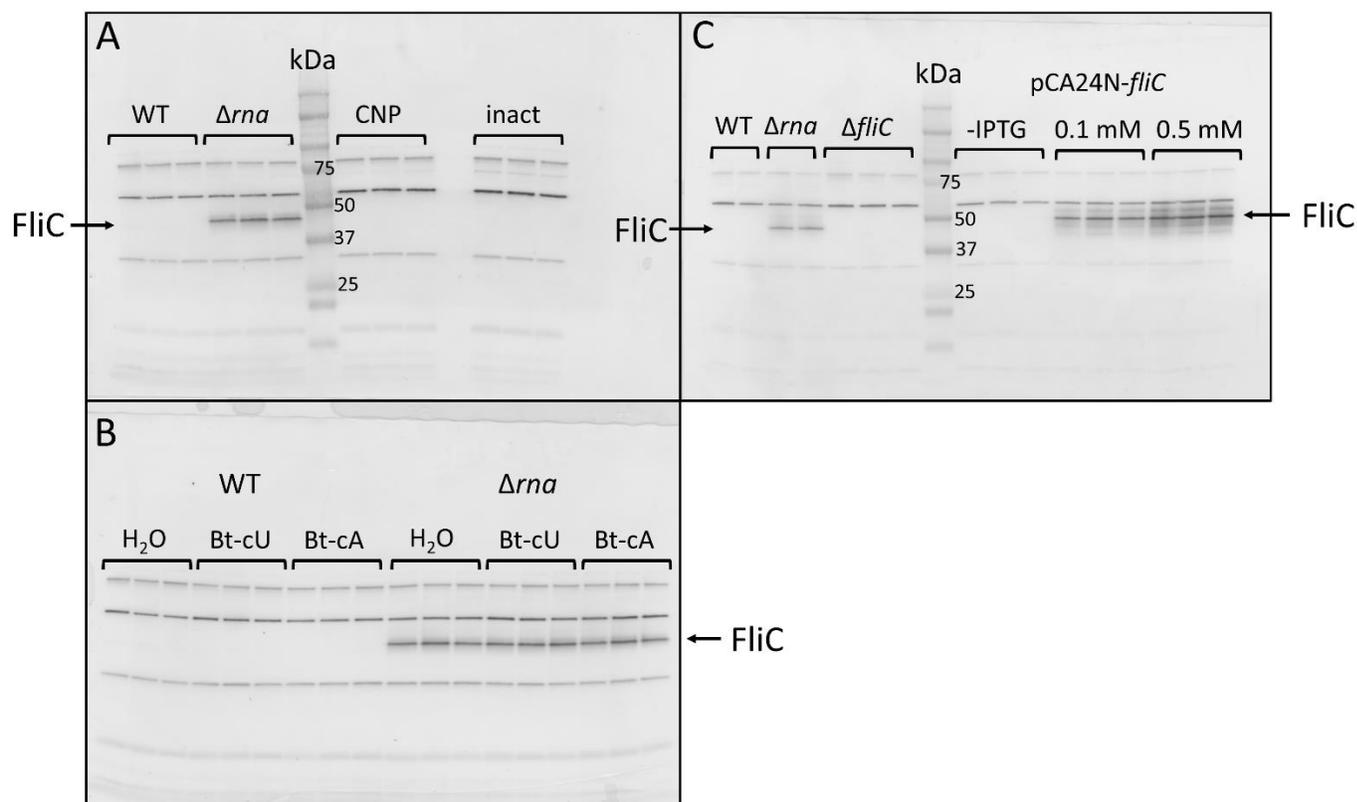
Supplementary Figure S3.7. (**A**) Image of western blot used for quantification of FliC expression in WT, Δ*rna*, WT pKT-*CNP* and WT pKT-*CNP-inact* (see Fig. 2D in main text for quantitative data) (**B**) Image of western blot used to quantify FliC expression in WT and Δ*rna*

upon treatment with either Bt-cAMP (1 mM), Bt-cUMP (1 mM), or H$_2$O (see Fig. 2C in main text for quantitative data). (**C**) Western blot analysis of *E. coli* either lacking *fliC* (Δ*fliC*) or overexpressing FliC from plasmid pCA24N-*fliC* in the presence or absence of isopropyl β-D-thiogalactopyranoside (IPTG; 0.1 or 0.5 mM). These data demonstrate that the band denoted by the arrow corresponds to FliC, indicating that FliC is undetectable in BW25113 (WT) and in WT expressing CNPase or CNPase-inact from plasmid pKT-*CNP* or pKT-*CNP-inact*, respectively. For all conditions assayed, each lane represents an individual biological replicate. For detailed western blot protocol, see **Materials and Methods** (**section 3.4**).
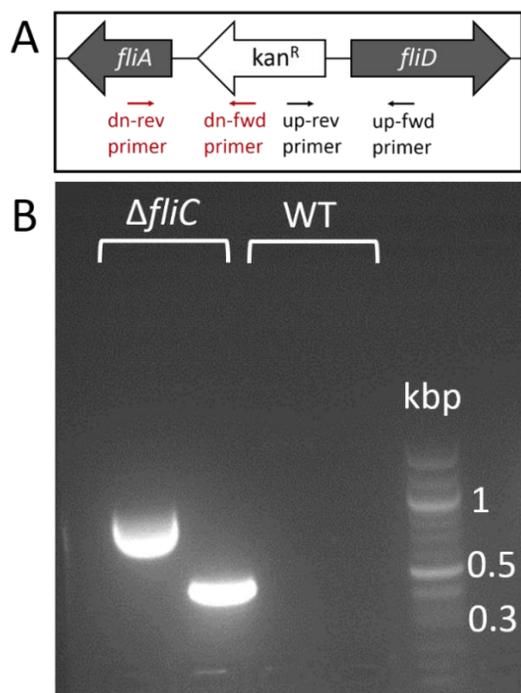
Supplementary Figure S3.8. PCR-based validation of *fliC*::kan^R genotype in BW25113 Δ*fliC*. (**A**) Schematic of the procedure to validate replacement of *fliC* (which lies between *fliD* and *fliA*) with the kan^R resistance cassette (not to scale). Two pairs of primers were designed to amplify across the junctions upstream and downstream of kan^R. (**B**) PCR products were analyzed on a 1.5% agarose gel, revealing products with the expected sizes in reactions containing Δ*fliC* chromosomal DNA as the template (lane 1: upstream junction; lane 2: downstream junction). In contrast, control PCR amplifications using BW25113 genomic DNA as the template produced no detectable products (last two lanes). PCR reactions were performed as described previously to validate the *rna*::kan^R genotype in BW25113 Δ*rna*.[1] For primer sequences, see **Supplementary Table S3.3**.

Supplementary Figure S3.9. WT and Δ*rna* exhibit normal chemotactic behavior in the presence of 1 mM uracil (**A**) and 1.5% w/v ribose (**B**). (**A**) The presence of the faint outer halo in the presence of uracil is indicative of a chemoattractant;[2] plate images are representative of three biological replicates. (**B**) Data represent the mean ± standard deviation of three biological replicates (* $P < 0.05$; ** $P < 0.01$). For detailed motility assay procedure, see **Materials and Methods** (**section 3.4**).
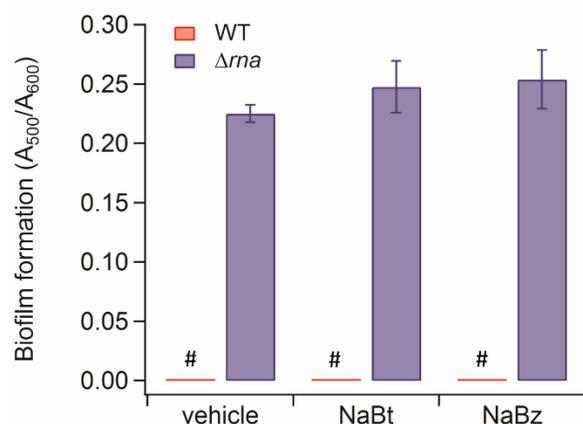
Supplementary Figure S3.10. Treatment with either sodium butyrate or sodium benzoate does not affect *E. coli* biofilm formation. WT and Δ*rna* were assayed for biofilm formation *via* Congo red staining (for details, see **Methods** in the main text) in the presence of 500 μM sodium butyrate (NaBt), 500 μM sodium benzoate (NaBz), or H_2O (vehicle). Data represent the mean ± standard deviation of six biological replicates (# denotes undetectable biofilm formation).
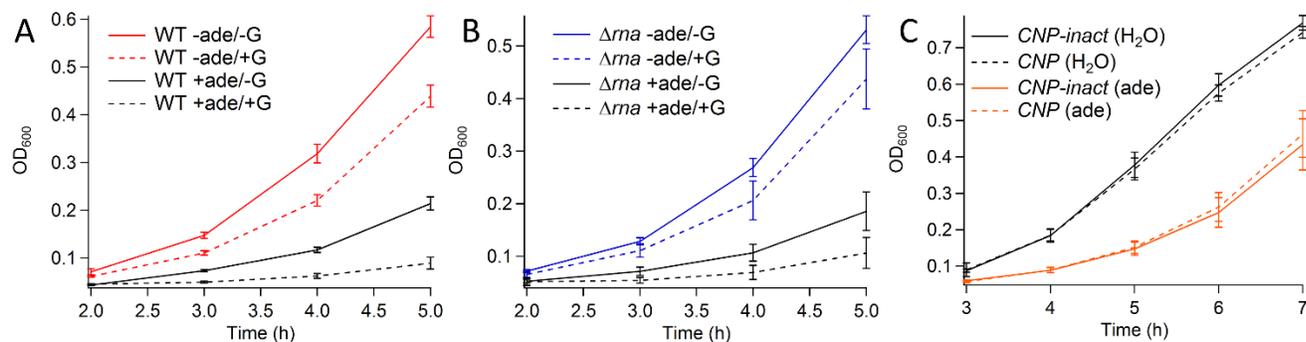
Supplementary Figure S3.11. Exogenous guanosine does not abrogate adenine-induced growth inhibition, in contrast to literature precedent for a different *E. coli* K-strain.[3] Effect of exogenous adenine and/or guanosine on growth of WT (**A**), Δ*rna* (**B**), and *CNP-inact*/*CNP* (**C**). Treatment with either purine elicits a bacteriostatic effect. Data represent the mean ± standard deviation of six biological replicates
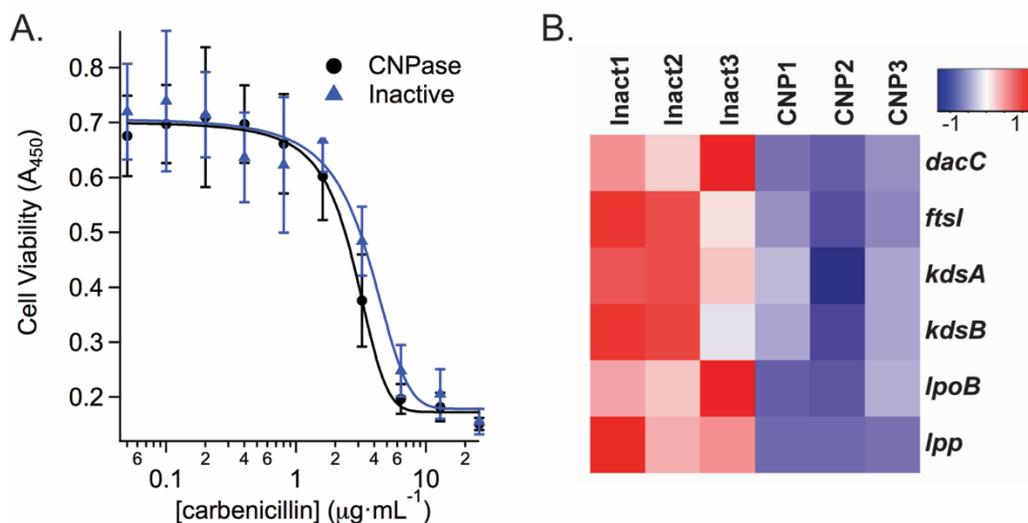
Supplementary Figure S3.12. CNPase expression does not affect *β*-lactam sensitivity. (**A**) Cell viability of cells expressing CNPase and CNPase-inact incubated with varying concentrations of carbenicillin. (**B**) Heatmap of genes potentially involved in *β*-lactam sensitivity.
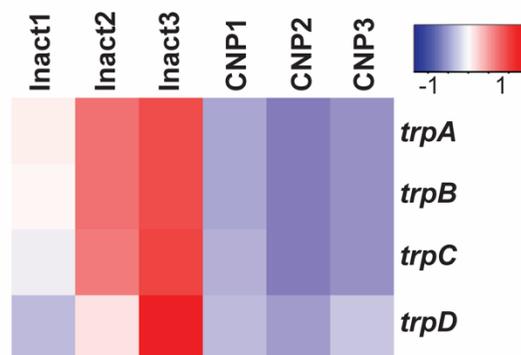
Supplementary Figure S3.13. Genes involved in tryptophan synthesis are downregulated in *E. coli* expressing CNPase.

Supplementary Protocol S3.1: PCR and thermal cycling conditions for construction of plasmid pBAD33-*rna* by polymerase incomplete primer extension (PIPE) cloning.

BW25113 chromosomal DNA was isolated using the GenElute Bacterial Genomic DNA Kit (Sigma-Aldrich). The *rna* gene was amplified from the genomic template using Phusion DNA polymerase in Phusion-HF buffer (New England Biolabs). Thermal cycling was performed essentially as recommended by the manufacturer, but the final extension was omitted to enrich for incomplete extension products. An analogous thermal cycling procedure was used to amplify the pBAD33 vector. Primer sequences are provided in **Supplementary Table S3.1**. Following amplification, the PCR reactions were treated with *Dpn*I (New England BioLabs; 0.8 units per μL of PCR reaction) to digest the template DNA, and the products were gel-purified and extracted using the GenElute Gel Extraction Kit (Sigma-Aldrich). The vector and insert products then were mixed in a 1:3 (v/v) ratio and transformed into RbCl-competent DH5α *E. coli* for plasmid propagation. The insert sequence was confirmed by Sanger sequencing (Eurofins Genomics).

Supplementary Table S3.1: Primer sequences for construction of pBAD33-*rna*.

Primer ID	Sequence (5'-3')	Annealing temperature (°C)
pBAD33 vector rev	ATG GAG CCT CCT CGA ATT CGC TAG	62
pBAD33 vector fwd	CTC GAG AAG CTT GGC TGT TTT GGC	
rna insert fwd	CTA GCG AAT TCG AGG AGG CTC CAT ATG AAA GCA TTC TGG CGT AAC G	64
rna insert rev	GCC AAA ACA GCC AAG CTT CTC GAG TTA ATA ACC CGC TTT ATC AAT CAC AAA GGT	

Supplementary Protocol S3.2: PCR and thermal cycling conditions for replacement of *β*-lactamase gene (amp[R]) with chloramphenicol-acetyl transferase gene (cam[R]) in plasmids pKT-*CNP* and pKT-*CNP-inact* by PIPE cloning.

PIPE cloning was performed essentially as described in **Supplementary Protocol S3.1** using either pKT-*CNP* (amp[R]) or pKT-*CNP-inact* (amp[R]) as the vector template, along with pBAD33 as the template for the cam[R] insert. The agarose gel purification step was omitted. Primer sequences are provided in **Supplementary Table S3.2**.

Supplementary Table S3.2: Primer sequences for construction of pKT-*CNP* (cam[R]) and pKT-*CNP-inact* (cam[R]) from corresponding pKT (amp[R]) plasmids.

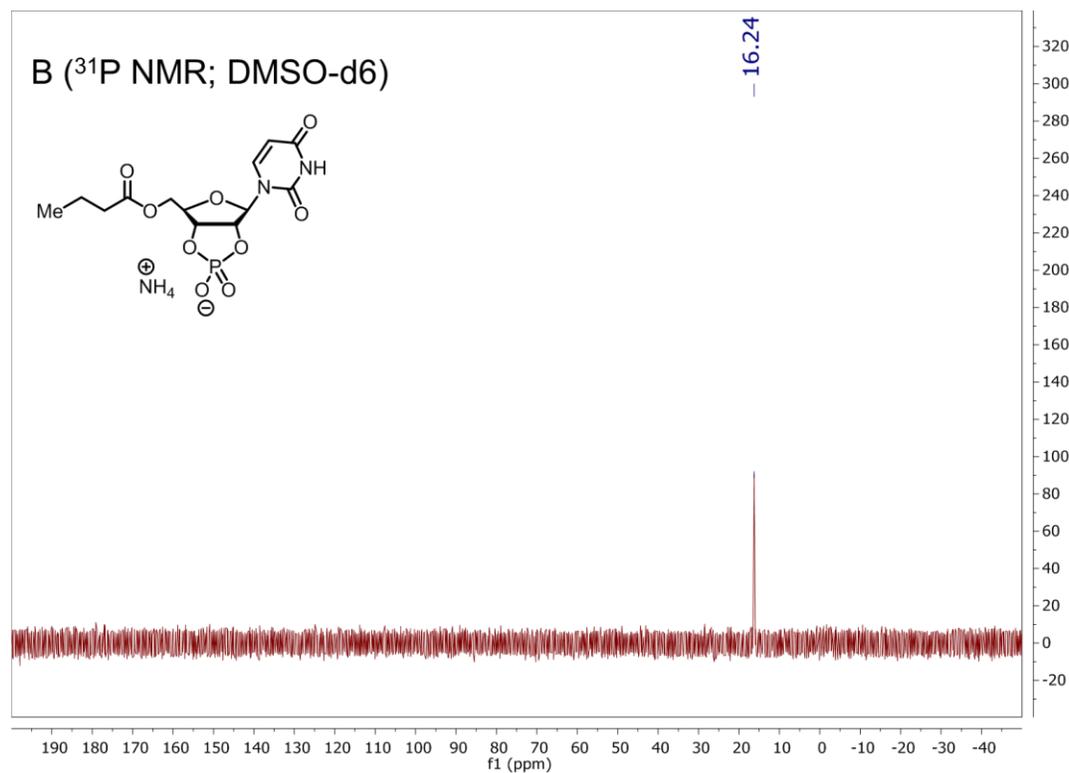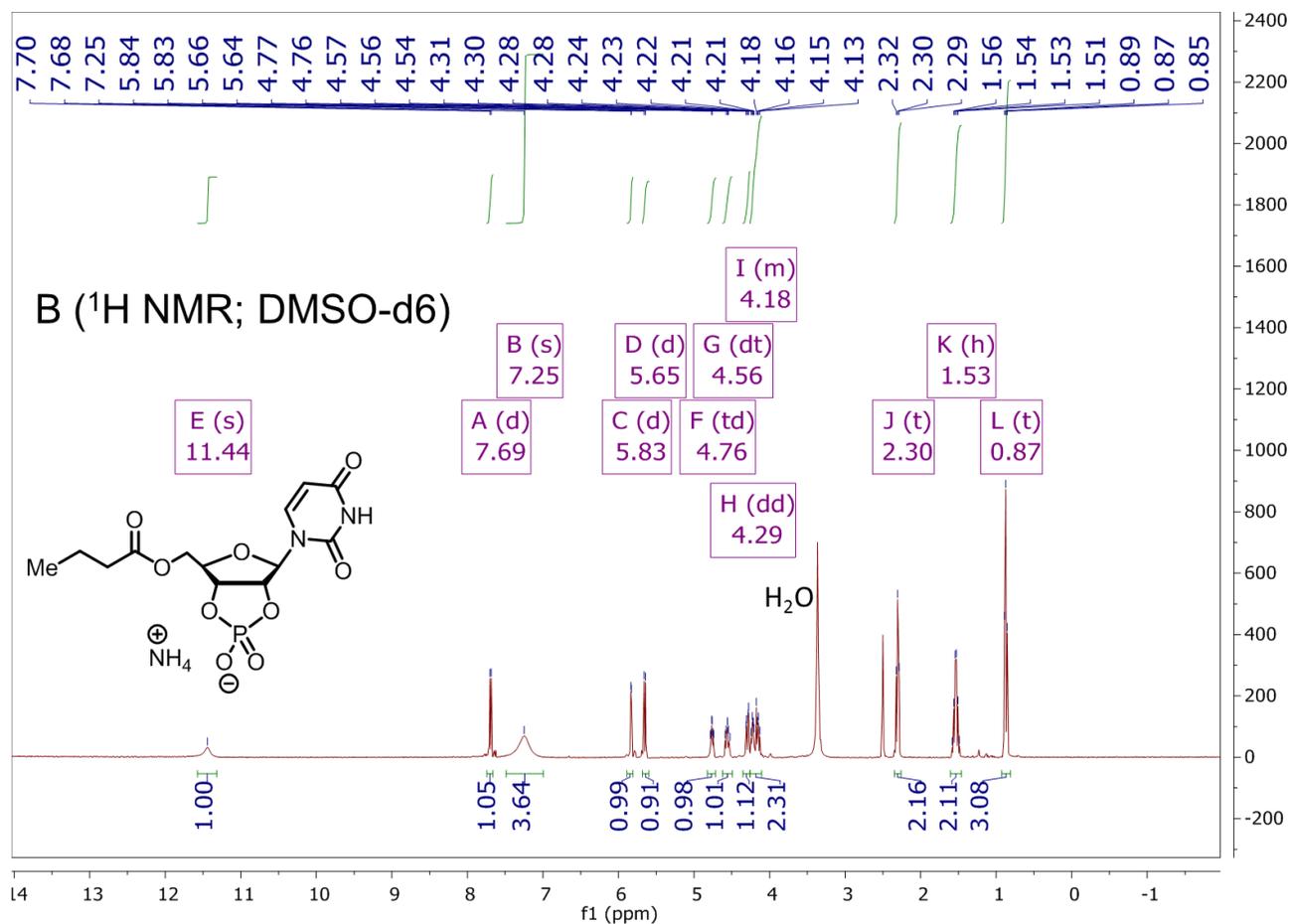| Primer ID	Sequence (5'-3')	Annealing temperature (°C)
pKT vector rev	ACT CTT CCT TTT TCA ATA TTA TTG AAG CAT	57
pKT vector fwd	CTG TCA GAC CAA GTT TAC TCA TAT ATA C	
cam[R] insert fwd	ATG CTT CAA TAA TAT TGA AAA AGG AAG AGT ATG GAG AAA AAA ATC ACT GGA TAT ACC	65
cam[R] insert rev	GTA TAT ATG AGT AAA CTT GGT CTG ACA G TTA CGC CCC GCC CTG	

Supplementary Table S3.3: Primer sequences used for validation of *fliC*::kan[R] genotype in BW25113 Δ*fliC*.
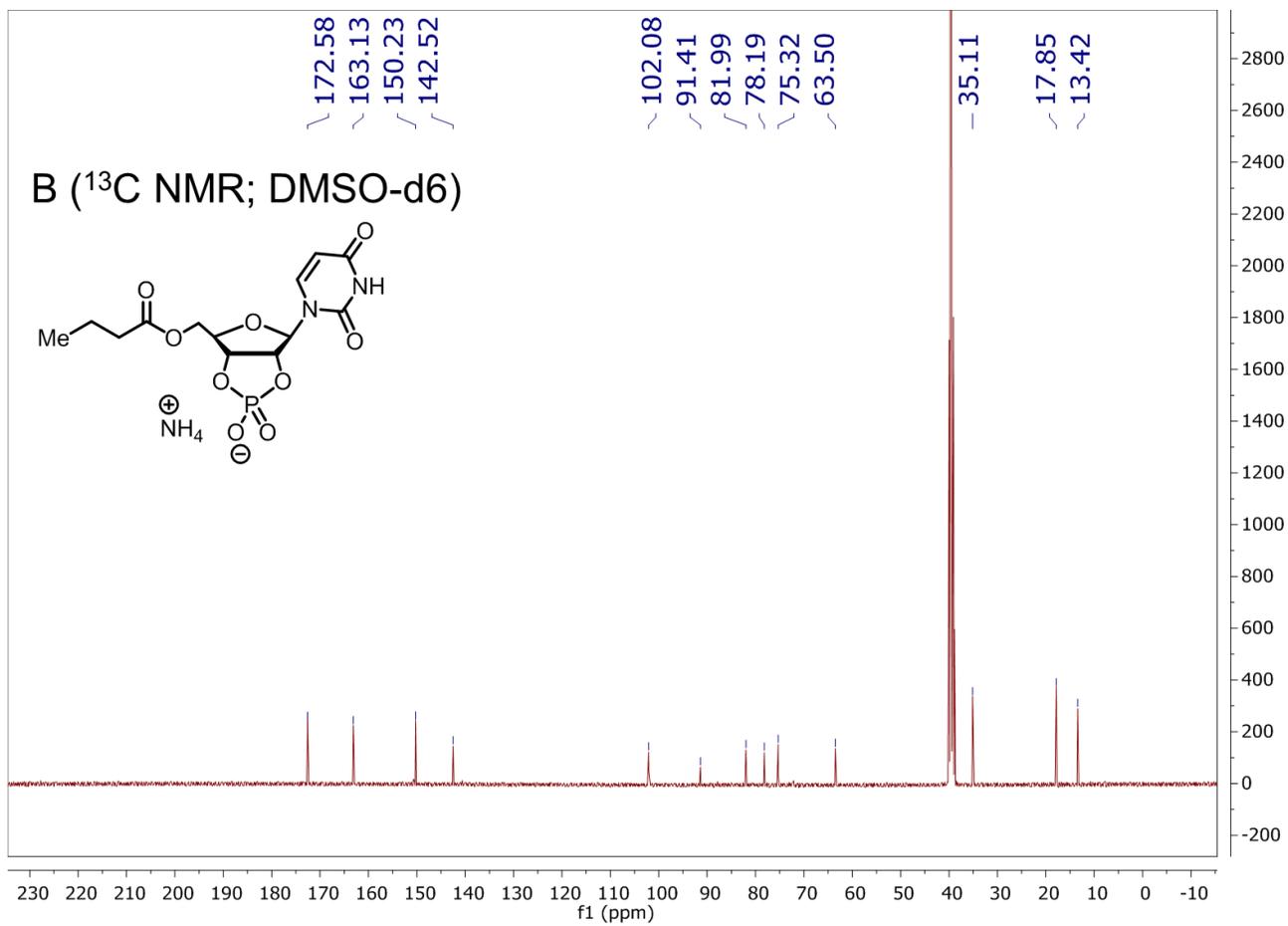
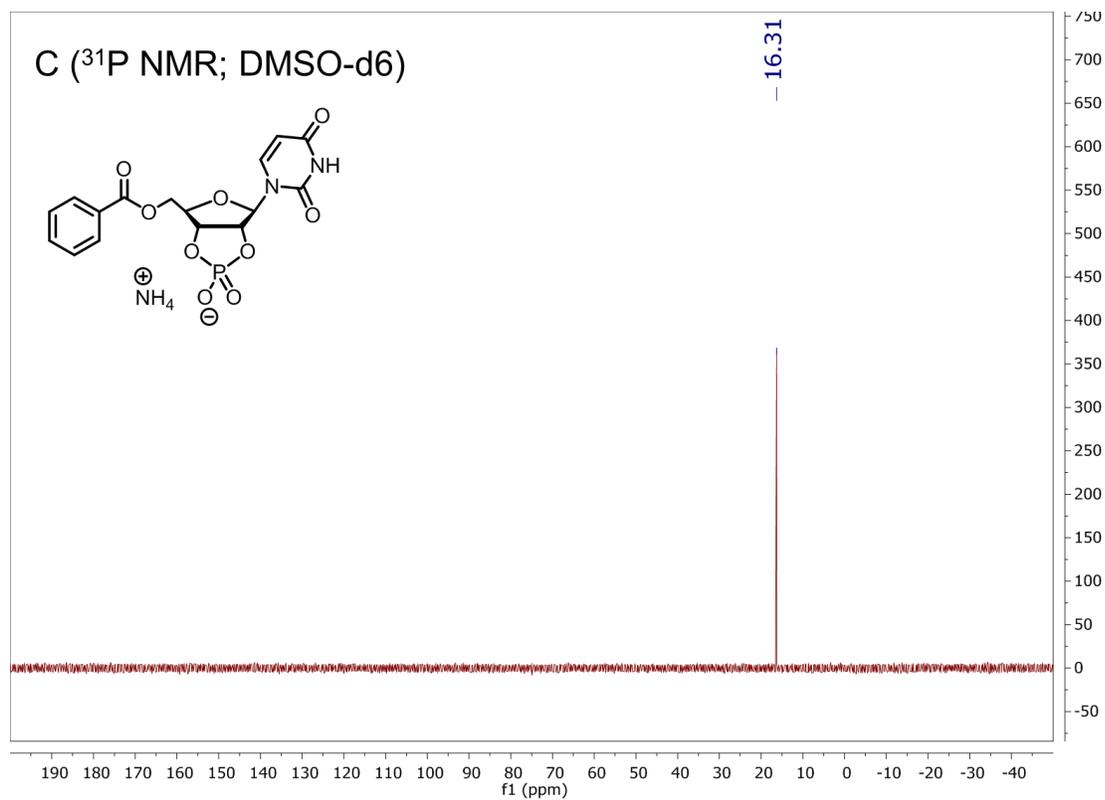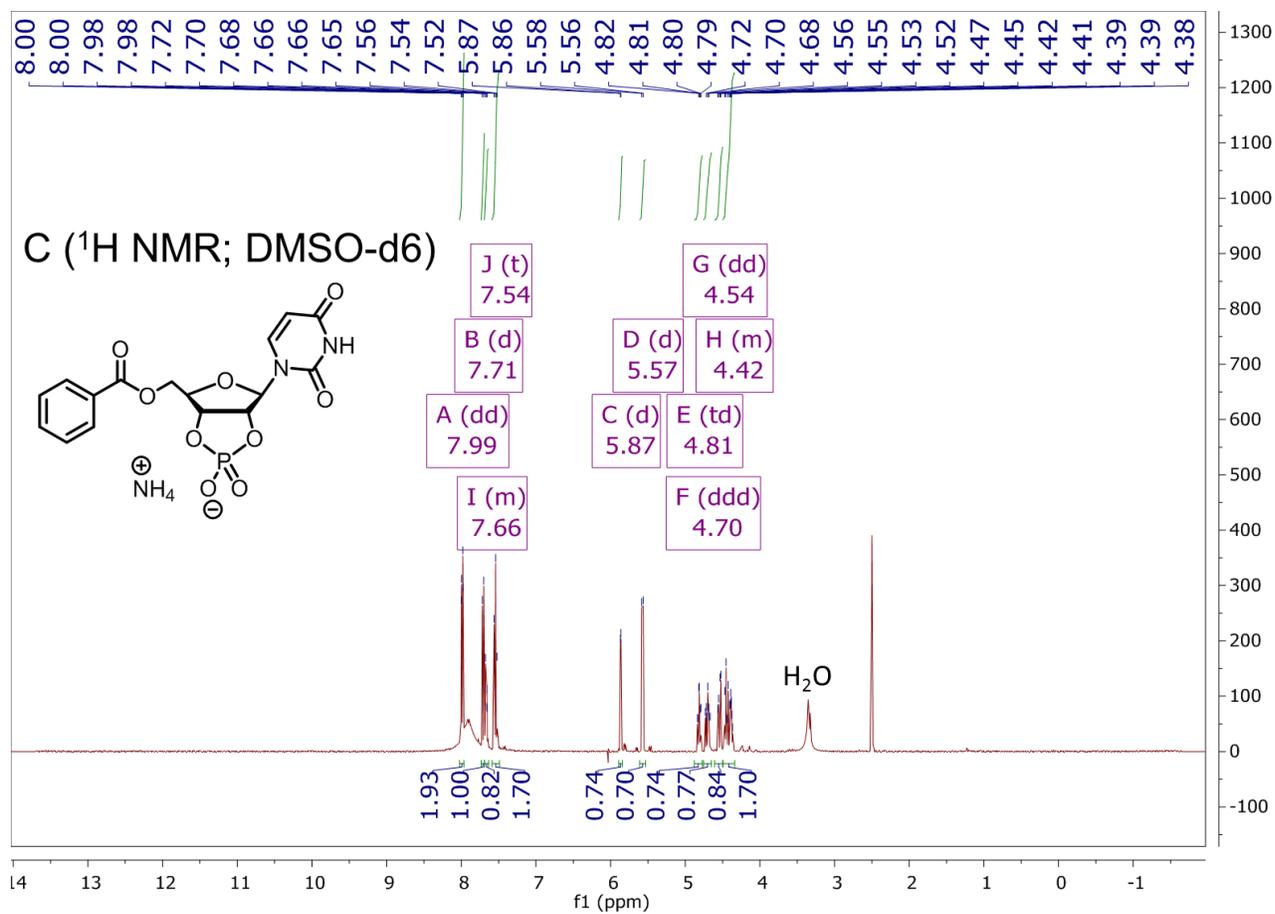| Primer ID	Sequence (5'-3')	Annealing temperature (°C)
Upstream junction (up-fwd)	GTA GGC GCT AAG TTT AGC GGT	64
Upstream junction (up-rev)	GTC ATA GCC GAA TAG CCT CTC CAC	
Downstream junction (dn-fwd)	TCG CAG CGC ATC GCC TTC TAT C	64
Downstream junction (dn-rev)	CAT AAC GCT GCC ACA GCG AGT	

Supplementary Spectra. Nuclear magnetic resonance (NMR) spectra for synthesized compounds. ^{1}H NMR, ^{31}P NMR, and ^{13}C NMR spectra for Bt-cAMP (**A**), Bt-cUMP (**B**), and Bz-cUMP (**C**).
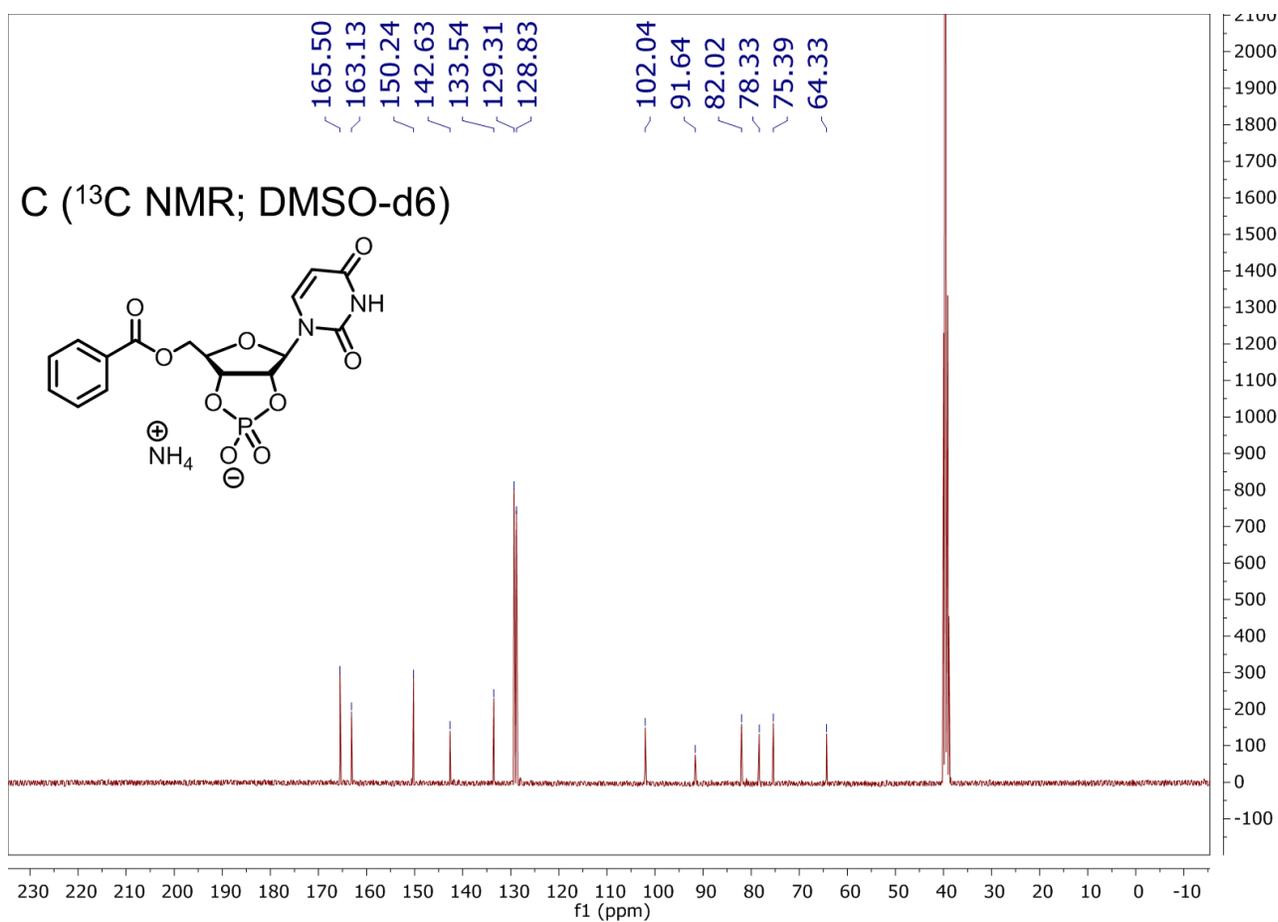
A (^{1}H NMR; DMSO-d6)

8.39, 8.18, 7.56, 7.30, 6.18, 6.18, 5.34, 5.34, 5.33, 5.33, 4.88, 4.87, 4.36, 4.35, 4.35, 4.34, 4.32, 4.32, 4.20, 4.19, 4.18, 4.16, 2.27, 2.26, 2.25, 2.24, 2.23, 2.22, 1.51, 1.49, 1.47, 1.46, 0.84, 0.82, 0.81

B (s) 8.18	F (ddd) 5.34	I (m) 4.17	L (q) 1.48
A (s) 8.39	D (t) 7.30	E (d) 6.18	G (ddd) 4.87	J (td) 2.25	K (t) 0.82
C (s) 7.56	H (m) 4.34

Integrals: 1.00, 1.01, 2.32, 4.20, 1.11, 1.03, 1.16, 2.40, 1.33, 2.22, 2.34, 3.32

A (^{31}P NMR; DMSO-d6)

—16.54

A (^{13}C NMR; DMSO-d6)

172.60
155.68
152.30
149.00
140.10
119.00
88.23
82.06
78.49
75.85
63.20
35.12
17.86
13.42

B (^{1}H NMR; DMSO-d6)

E (s) 11.44

B (s) 7.25

A (d) 7.69

I (m) 4.18

D (d) 5.65

G (dt) 4.56

C (d) 5.83

F (td) 4.76

H (dd) 4.29

K (h) 1.53

J (t) 2.30

L (t) 0.87

H$_2$O

7.70 7.68 7.25 5.84 5.83 5.66 5.64 4.77 4.76 4.57 4.56 4.54 4.31 4.30 4.28 4.28 4.24 4.23 4.22 4.21 4.21 4.18 4.16 4.15 4.13 2.32 2.30 2.29 1.56 1.54 1.53 1.51 0.89 0.87 0.85

1.00 1.05 3.64 0.99 0.91 0.98 1.01 1.12 2.31 2.16 2.11 3.08

B (^{31}P NMR; DMSO-d6)

—16.24

B (^{13}C NMR; DMSO-d6)

172.58
163.13
150.23
142.52
102.08
91.41
81.99
78.19
75.32
63.50
35.11
17.85
13.42

C (^{1}H NMR; DMSO-d6)

J (t) 7.54

B (d) 7.71

A (dd) 7.99

I (m) 7.66

G (dd) 4.54

D (d) 5.57

C (d) 5.87

E (td) 4.81

H (m) 4.42

F (ddd) 4.70

H$_2$O

1.93 1.00 0.82 1.70 0.74 0.70 0.74 0.77 0.84 1.70

C (^{31}P NMR; DMSO-d6)

—16.31

C (^{13}C NMR; DMSO-d6)

165.50
163.13
150.24
142.63
133.54
129.31
128.83
102.04
91.64
82.02
78.33
75.39
64.33

3.7 Supplementary References

(1) Fontaine, B. M.; Martin, K. S.; Garcia-Rodriguez, J. M.; Jung, C.; Southwell, J. E.; Jia, X.; Weinert, E. E. *Biochem. J* **2018**, *478*, 1491-1506.

(2) DeLoney-Marino, C.; Wolfe, A.; Visick, K. *Appl. Environ. Microbiol.* **2003**, *69*, 7527-7530.

(3) Levine, R.; Taylor, M. *J. Bacteriol.* **1982**, *149*, 923-930.

Chapter 4: Work toward elucidating the roles of cytidine 3',5'-cyclic monophosphate in mammals

Author contributions: X.J. and F.S. optimized the methodology for nucleotide extraction and quantitation. X.J. and B.M.F. performed nucleotide extraction and quantitation, conducted protein fractionation experiments, synthesized affinity matrices, and performed pull-down experiments. X.J., B.M.F., and E.E.W. conducted bioinformatic analyses. B.M.F. optimized the assay for cytidylate cyclase activity. E.E.W designed the study.

Chapter 4: Work toward elucidating the roles of cytidine 3',5'-cyclic monophosphate in mammals

4.1 Introduction

Nucleotide second messengers regulate diverse biological processes across the kingdoms of life. In metazoa, the cyclic purines adenosine- and guanosine 3',5'-cyclic monophosphate (3',5'-cAMP; -cGMP) govern a host of endocrine, immunological, cardiological, and neurological functions. Many of these 3',5'-cNMP-dependent pathways are initiated by the interaction of an extracellular hormone or neurotransmitter with a G-protein-coupled receptor (GPCR).[1] The resultant conformational change in the heterotrimeric G-protein complex alters the activity of adenylate or guanylate cyclases (ACs; GCs), which convert ATP or GTP to 3',5'-cAMP or -cGMP, respectively.[2] Modulation of 3',5'-cNMP levels subsequently impacts cell physiology through various mechanisms including phosphorelay signaling and membrane polarization, effectively transducing the initial extracellular signal into the cytosol to influence organismal phenotypes. Intracellular 3',5'-cAMP and -cGMP levels also are governed by 3',5'-cNMP-specific phosphodiesterases (PDEs) which catalyze hydrolysis of the cyclic phosphodiester to afford the acyclic 5'-NMP.[3]

In addition to the paradigmatic nucleotide second messengers discussed above, recent work has identified a role for the dinucleotide cyclic guanosine-(2'-5')-monophosphate-adenosine-(3'-5')-monophosphate (2',3'-cGAMP) in metazoan immunity. The presence of double-stranded viral DNA in the cytosol directly activates the enzyme cyclic GMP-AMP synthase, inducing dimerization of GTP and ATP to produce the atypical dinucleotide 2',3'-cGAMP containing both a 3'-5'- and a 2'-5'-phosphodiester bond, in contrast to the canonical pair of 3',5'-phosphodiester linkages in bacterial c-di-GMP and c-di-AMP.[4] 2',3'-cGAMP subsequently binds the dimeric

STING receptor at the interface of two protomers, inducing phosphorylation of interferon regulatory factor 3 by TANK binding kinase 1, thereby upregulating expression of type I interferon to combat the viral threat.[5] STING also functions in surveillance of bacterial infection by inducing an innate immune response upon c-di-GMP binding,[6] albeit with lower efficacy than the response generated from the interaction with the cognate 2',3'-cGAMP ligand.[7]
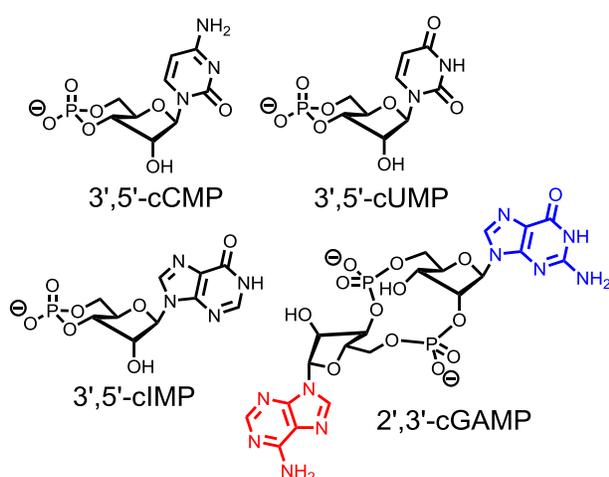

Figure 4.1. Chemical structures of some emerging nucleotide second messengers in eukaryotes.

Bioanalytical investigations of mammalian organs and cells have expanded the collection of known 3',5'-cyclic mononucleotides beyond the established purine second messengers 3',5'-cAMP and -cGMP. Liquid chromatography-tandem mass spectrometry (LC-MS) has identified the cyclic nucleotides 3',5'-cIMP, -cCMP, -cUMP, and -cdTMP in a range of rodent tissues.[8] Recent experiments have begun to elucidate the biological significance of some of these cyclic nucleotides. 3',5'-cIMP is biosynthesized by soluble guanylate cyclase *in vitro*,[9] and further studies demonstrated that this enzyme upregulates 3',5'-cIMP production in hypoxic coronary arteries, suggesting that this purine functions as a *bona fide* second messenger in mammals.[9b] Furthermore, 3',5'-cIMP arises from deamination of 3',5'-cAMP by a selective deaminase in the pathogenic spirochete bacterium *Leptospira interrogans*,[10] constituting a previously unknown facet of 3',5'-cAMP metabolism with potential relevance in other kingdoms of life. Recent studies also suggest that the pyrimidines 3',5'-cCMP and -cUMP govern mammalian physiology. HEK293 and B103 cells exhibit high

concentrations of these cyclic pyrimidines relative to other mammalian cell types;[11] these elevated concentrations are maintained by the bicarbonate-stimulated activity of soluble adenylate cyclase (sAC).[12] Similarly, soluble guanylate cyclase (sGC) synthesizes 3',5'-cCMP and -cUMP *in vitro* and in lung fibroblast cells.[9a,13] Notably, NO promotes the sGC-dependent cytidylate and uridylate cyclase activities,[9a,13] suggesting that production of 3',5'-cCMP and -cUMP is physiologically relevant. In contrast to sAC and sGC, the membrane-associated adenylate cyclase (mAC) and particulate guanylate cyclase (pGC) lack detectable cytidylate and uridylate cyclase activity,[12-13] likely resulting in accumulation of 3',5'-cCMP and -cUMP in the interior of the cytosol, as opposed to the periphery of the cell. This alludes to spatial regulation of these 3',5'-cyclic pyrimidine nucleotide pools.[14] However, an earlier study reported a membrane-associated cytidylate cyclase activity that was inhibited by ATP, ADP, and AMP, suggesting the existence of a previously unknown enzyme involved in 3',5'-cCMP synthesis, in addition to the promiscuous cytidylate cyclase of the characterized cyclases mentioned above. Further experiments demonstrated that the cyclase displayed a unique pH-activity relationship relative to known ACs and GCs, and was unaffected by known activators and inhibitors of AC.[15] Therefore, these findings suggest that eukaryotic enzyme(s) specific for 3',5'-cCMP biosynthesis may exist, but conclusive identification and characterization requires further investigation.

Regarding degradation of the cyclic pyrimidines, several purified metazoan cyclic nucleotide phosphodiesterases (PDEs) representing different families not only accept 3',5'-cAMP and -cGMP as substrates, but 3',5'-cUMP as well.[16] Conversely, 3',5'-cCMP resists hydrolysis by most known cNMP PDEs, with the exception of PDE7A1.[17] Previous reports also identified 3',5'-cCMP-specific PDEs in Leukemia cells,[18] rat liver,[19] and pig liver,[20] but further validation is necessary to place these PDEs within the context of the 11 established 3',5'-cNMP PDE families,

as these experiments were performed 30-40 years ago. Notably, these PDEs displayed a preference for 3',5'-cCMP hydrolysis relative to 3',5'-cAMP and -cGMP hydrolysis. In addition to metabolism by PDEs, intracellular concentrations of these cyclic nucleotides likely are maintained by multi-drug resistance proteins (MRPs), as MRP5 exports 3',5'-cCMP and MRP4 exports both 3',5'-cUMP and -cCMP.[21]

In addition to the identified cyclases and PDEs discussed above, established protein effectors involved in binding 3',5'-cAMP and -cGMP also recognize the cyclic pyrimidine nucleotides. 3',5'-cCMP and -cUMP act as partial activators of regulatory domain RIα of protein kinase G (PKGIα) *in vitro*.[22] Despite the low binding affinity, treatment with cell-permeable N^4,2'-*O*-dibutyryl 3',5'-cCMP (diBt-cCMP) exerts pronounced vascular effects in mice, and a PKGIα knockout abrogates these effects.[23] Furthermore, the cyclic pyrimidines fully activate protein kinase A domains RIα and RIIα *in vitro*, albeit with low binding affinity.[22] Whole cell experiments employing dynamic mass redistribution (DMR), which reports on reorientation of cellular biopolymers in response to chemical perturbation,[24] also implicated membrane-permeable 3',5'-cCMP and -cUMP derivatives containing lipophilic acetoxymethyl phosphate esters in the modulation of biochemical processes.[25] Additional electrophysiology experiments in live cells and nuclear magnetic resonance (NMR) structural studies demonstrated that 3',5'-cCMP and -cUMP partially activate hyperpolarization-activated, 3',5'-cAMP-responsive HCN ion channels 2 and 4 with low potency,[26] further suggesting the potential significance of these cyclic nucleotides in eukaryotic signaling. Notably, the 3',5'-cAMP-responsive exchange factor EPAC1 is virtually unresponsive to 3',5'-cCMP and -cUMP,[25] demonstrating that these pyrimidines differentially modulate known 3',5-cNMP-binding effectors. Previously, a putative cyclic nucleotide-dependent kinase was reported and characterized by fast-atom bombardment mass spectrometry (FAB-MS),

displaying preferential activation by 3',5'-cCMP relative to the established cyclic purine nucleotides.[27] However, this potential 3',5'-cCMP-activated kinase awaits further investigation, similarly to the putative cyclases and PDEs discussed above.

While the purported biological functions of 3',5'-cCMP and -cUMP remain poorly understood, several studies dating back over four decades implicate 3',5'-cCMP in the regulation of cell proliferation. This cyclic nucleotide initially was identified in leukemia cells,[28] and exogenous 3',5'-cCMP eliminated the growth lag in this cell type.[29] Conversely, 3',5'-cAMP and -cGMP displayed anti-proliferative effects.[29] Moreover, the 3',5-cCMP concentration increased markedly following partial hepatectomy, compared to control, suggesting involvement in the response to stress.[30] Further analytical experiments comparing cyclic nucleotide levels in the urine of leukemia patients revealed a marked increase in the 3',5'-cCMP level, compared to relatively modest increases in 3',5'-cAMP and -cGMP concentrations, relative to control samples.[31] Treatment of neutrophilic leukocytes with cell-permeable diBt-cCMP also impacts the production of anionic superoxide in response to chemoattractants, implicating 3',5'-cCMP in regulation of innate immunity. Importantly, diBt-cCMP elicits distinct effects relative to diBt-cAMP and -cGMP.[32] In addition, 3',5'-cCMP appears to regulate cardiovascular function *via* PKG and HCN ion channels,[23,26] as discussed above. Even fewer studies have investigated the role of 3',5'-cUMP, but it agonizes HCN channels 2 and 4, suggesting potential relevance in cardiovascular regulation.[26b] 3',5'-cUMP also appears to function in tandem with 3',5'-cGMP to mediate cell death in neuroblastoma cells, as treatment with a combination of cell-permeable 3',5'-cUMP and -cGMP reduces cell viability.[25] Intriguingly, this exogenous nucleotide addition mimics the toxic effect of the *Pseudomonas aeruginosa* exoenzyme Y (ExoY), a 3',5'-nucleotidyl cyclase that preferentially generates 3',5'-cUMP and -cGMP to damage host lung tissue.[33] The present work

seeks to identify unknown enzymes involved in 3',5'-cCMP-dependent signal transduction and metabolism in eukaryotes.

4.2 Results

Quantification of 3',5'-cCMP and other 3',5'- and 2',3'-cNMPs in tissues and cells

Due to the emerging physiological functions of non-canonical cyclic nucleotides, a LC-MS/MS-based protocol was developed to quantify an array of 3',5'- and 2',3'-cNMPs in various mammalian tissues and cells. Cyclic nucleotide extraction entailed tissue homogenization or cell lysis into a mixture of H_2O/MeOH/MeCN (2:2:1 v/v/v), followed by removal of the solvent and resuspension in aqueous buffer for LC-MS/MS analysis. For analyte quantification, 8-bromo 3',5'-cAMP was utilized as an internal standard. This facile method was employed to quantify physiological concentrations of 3',5'-cAMP, -cGMP, -cIMP, and cCMP, along with the analogous 2',3'-cNMP regioisomers (a reconstructed ion chromatogram obtained using authentic cNMP standards is shown in **Figure 4.2**), in rodent kidney, lung, heart, spleen, liver, and brain (**Figures 4.3** and **4.4**).[8d] Notably, this study marks the first quantification of 3',5'-cIMP in mammalian tissue and constitutes the first discovery of 2',3'-cIMP in any organism, alluding to potential biological roles for these inosine nucleotides. In addition to 3',5'-cIMP, 3',5'-cCMP accumulates in multiple rodent tissues, including kidney, lung, heart, spleen, liver, and brain (**Figure 4.3B**).[8d] Intriguingly, 3',5'-cCMP concentrations in these tissues are similar to basal levels of the paradigmatic second messenger 3',5'-cGMP (**Figure 4.3**); these data suggest functional relevance for 3',5'-cCMP in diverse mammalian organ systems.
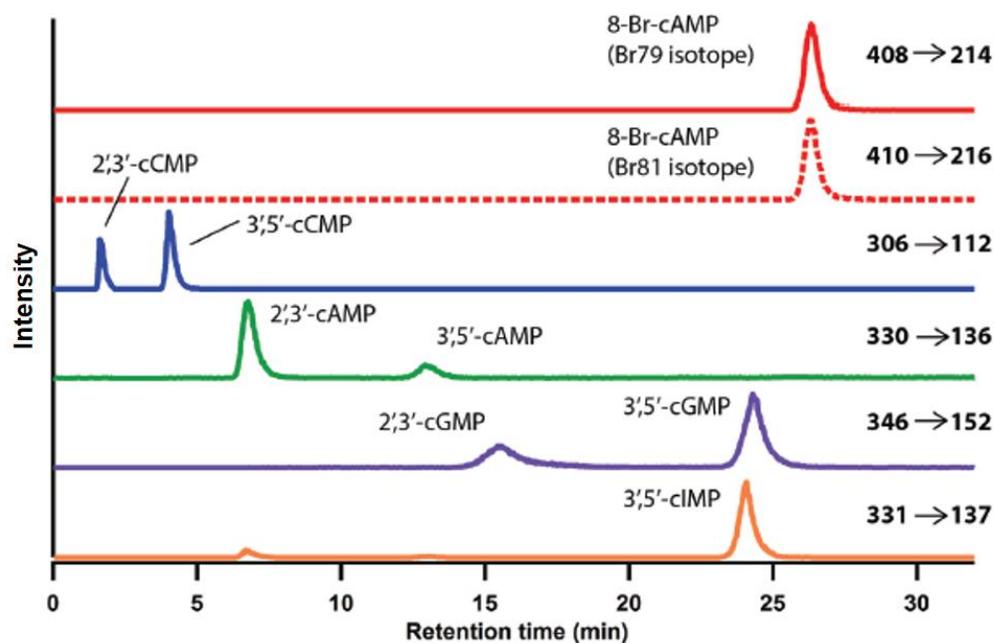
Figure 4.2. Reconstructed ion chromatogram depicting retention times for authentic standards of each cNMP. The *m/z* transitions depict the protonated parent ion [M+H]$^+$ used for nucleotide detection and the protonated nucleobase fragment ion used for quantification.
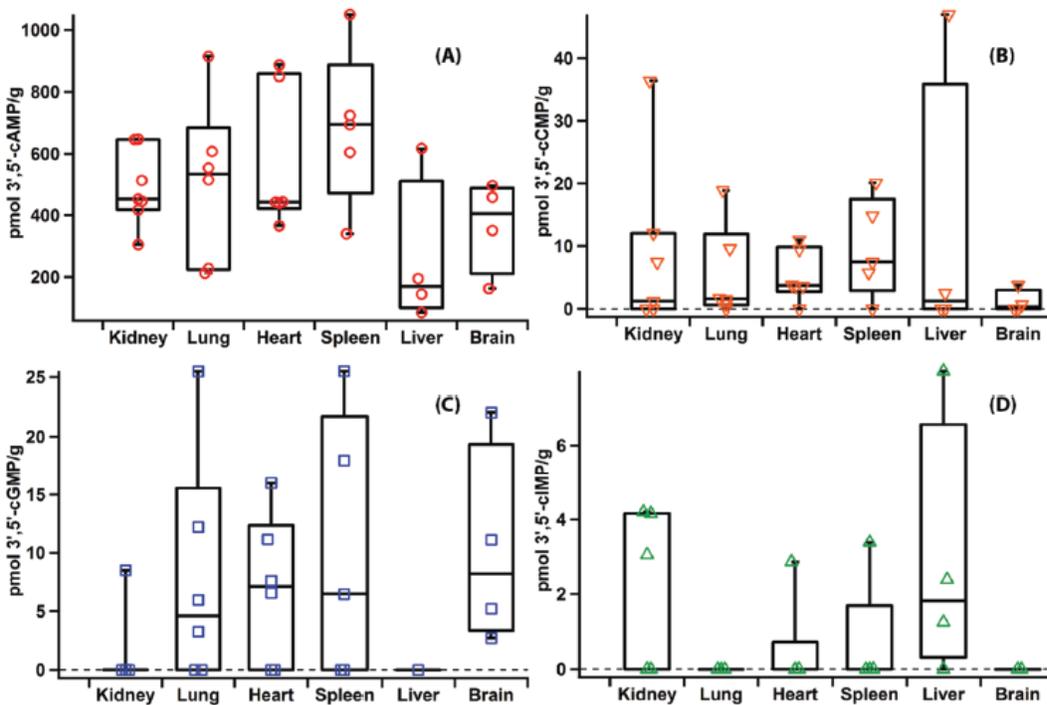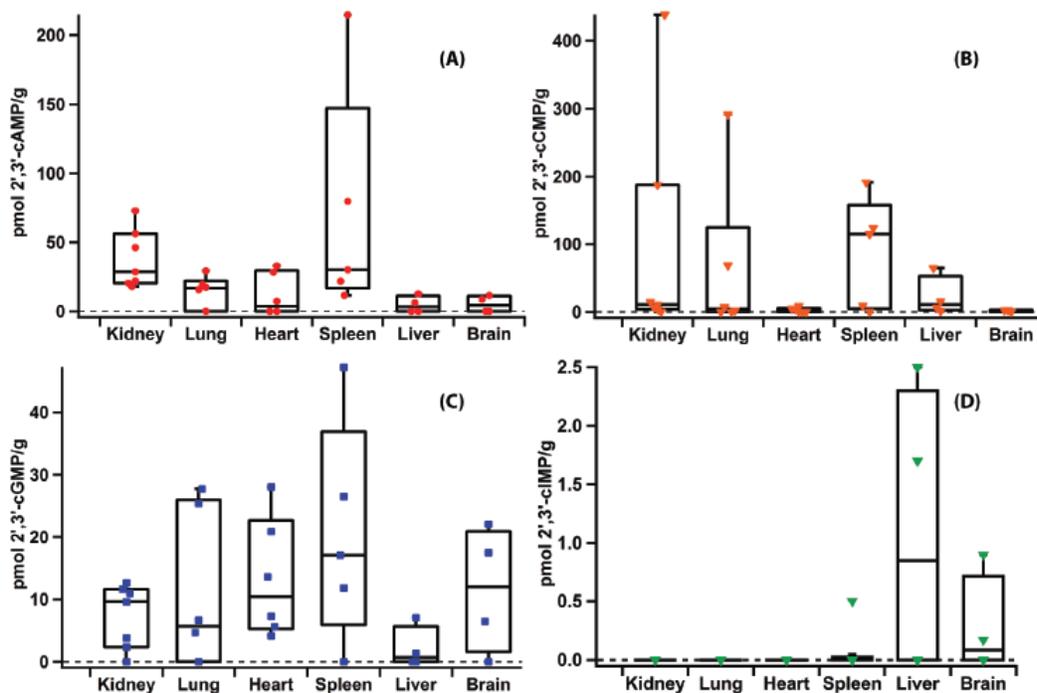

Figure 4.3. Concentrations of 3',5'-cNMPs (pmol/g wet tissue) in various rat organs [**(A)** 3',5'-cAMP **(B)** 3',5'-cCMP **(C)** 3',5'-cGMP **(D)** 3',5'-cIMP]. Each data point depicts the cNMP level quantified in a single biological replicate. The box plot denotes the 75th to 25th percentile concentrations, with the line representing the 50th percentile; the whiskers depict the 90th and 10th percentile.

Figure 4.4. Concentrations of 2',3'-cNMPs (pmol/g wet tissue) in various rat organs [**(A)** 3',5'-cAMP **(B)** 3',5'-cCMP **(C)** 3',5'-cGMP **(D)** 3',5'-cIMP]. Data are depicted as in **Figure 2**.

Attempts to identify putative 3',5'-cCMP-associated metabolic enzymes and effectors

Previous studies have reported the existence of putative mammalian enzymes exhibiting cytidylate cyclase,[15] 3',5'-cCMP-specific phosphodiesterase,[18-19,34] and 3',5'-cCMP-responsive kinase activities.[27] Interestingly, these enzymes exhibit selectivity for the production of 3',5'-cCMP over the canonical purine 3',5'-cNMPs, suggesting that dedicated 3',5'-cCMP-dependent biological functions may exist in mammals. The present work employs activity-guided protein fractionation in tandem with proteomics in an attempt to identify putative 3',5'-cCMP metabolic enzymes and 3',5'-cCMP-binding effector proteins.

Previous reports have described a membrane-associated cytidylate cyclase (mCC) activity in rat brain that exhibits distinct properties compared to known ACs and GCs.[15] Fractionation of this putative mCC was attempted using sucrose density gradient ultracentrifugation followed by

5'-CTP affinity chromatography. The 5'-CTP-agarose matrix **Figure 4.5; compound 4.4)** was prepared by initial activation of the triphosphate moiety with dicyclohexylcarbodiimide,[35] to generate a cyclic trimetaphosphate intermediate (**Figure 4.5; compound 4.2**), which was confirmed by ^{31}P NMR analysis of the reaction mixture. Subsequent amidation with a bis-ethylamine linker afforded the desired 5'-CTP γ-phosphoramidate (**Figure 4.5; compound 4.3**) for coupling to the *N*-hydroxysuccinimide-functionalized agarose resin (**Figure 4.5; compound 4.4**).[35-36]
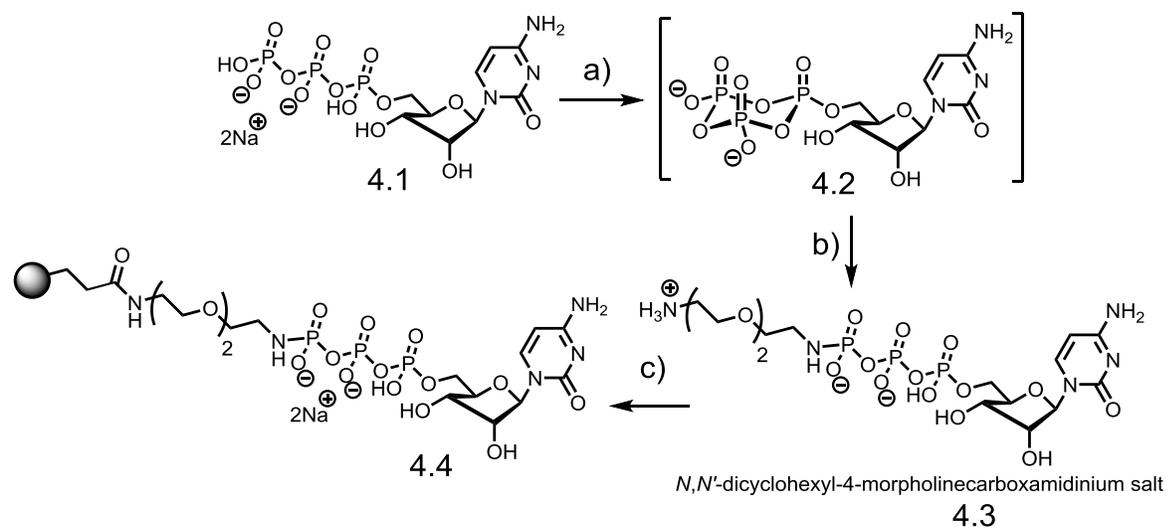
Figure 4.5. Synthesis of 5'-CTP-agarose affinity matrix. **a)** 1) Amberlite IR-120 (pyridinium form); 2) Dicyclohexylcarbodiimide, *N,N'*-dicyclohexyl-4-morpholinecarboxamidine, DMF, 25°C, 24 h. **b)** 2,2'-(Ethylenedioxy)bis(ethylamine), DMF, 25°C, 18 h. **c)** 1) Amberlite IR-120 (sodium form); 30% yield from CTP; 2) *N*-hydroxysuccinimide-agarose, sodium phosphate buffer (0.1 M, pH 7.5), 4°C, 18 h.

Over the course of the fractionation, cyclase activity was monitored using a luminescence-based pyrophosphate detection assay (**Figure 4.6**). Due to the possibility of cyclase-independent pyrophosphate generation, the production of 3',5'-cCMP also was validated chromatographically. The luminescence assay revealed enrichment of CC activity in certain fractions collected from the sucrose density gradient (**Figure 4.7**), but attempted affinity purification over immobilized CTP-agarose was unsuccessful. This failure could be explained by hydrolysis of the CTP ligand from

the agarose matrix (despite the fact that affinity chromatography step was conducted at 4°C). Future efforts will seek to immobilize non-hydrolysable CTP derivatives to circumvent enzymatic cleavage from the solid support. Additionally, fractions containing CC activity will be subjected to LC-MS/MS-based proteomics analysis in an attempt to identify potential uncharacterized nucleotidyl cyclases.
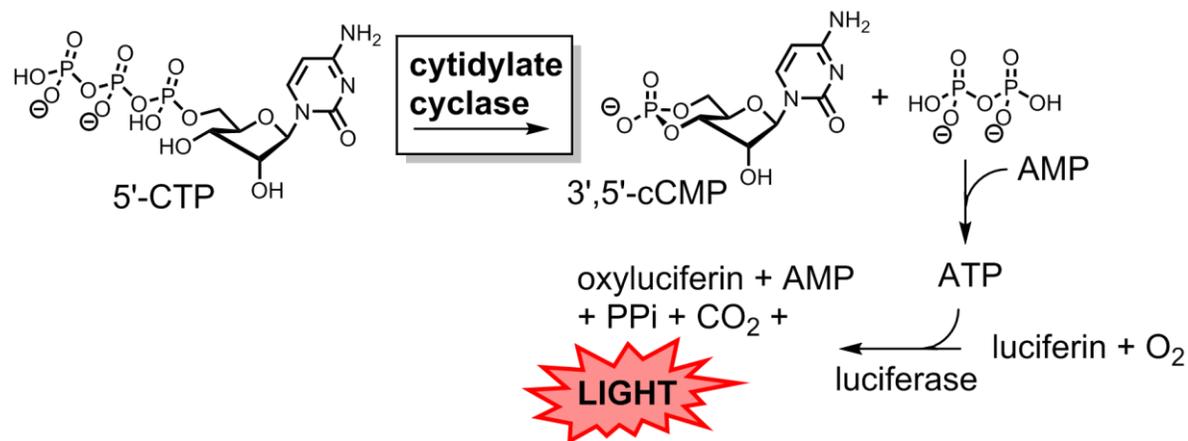

Figure 4.6. Scheme depicting the general mechanism of the luminescence-based nucleotidyl cyclase assay. Light emission is coupled to inorganic pyrophosphate production *via* luciferase-catalyzed oxidation of luciferin to oxyluciferin (Lonza PPiLight Inorganic Pyrophosphate Assay Kit).

In addition to efforts to identify putative CC activities, fractionation of 3',5'-cCMP-specific PDE activity was attempted, as a putative PDE activity previously has been purified from rat liver homogenate.[19b] To this end, a published protocol employing ammonium sulfate precipitation and size exclusion chromatography was utilized to fractionate a reported 3',5'-cCMP PDE.[19b] SDS-PAGE analysis revealed a band with a relative molecular weight (M_r) ~28 kDa (**Figure 4.8A**), which corresponds to the observed M_r of the previously reported rat liver PDE.[19b] The proteins in this band were subjected to trypsin digestion and tandem-MS based proteomic analysis. While several characterized proteins were identified on the basis of multiple peptides, no putative PDEs or uncharacterized proteins were observed. Unfortunately, initial attempts to monitor 3',5'-cCMP PDE activity through the fractionation have been problematic; no PDE

activity has been observed *via* HPLC-based assay in any of the fractions generated during the attempted purification of the putative PDE.
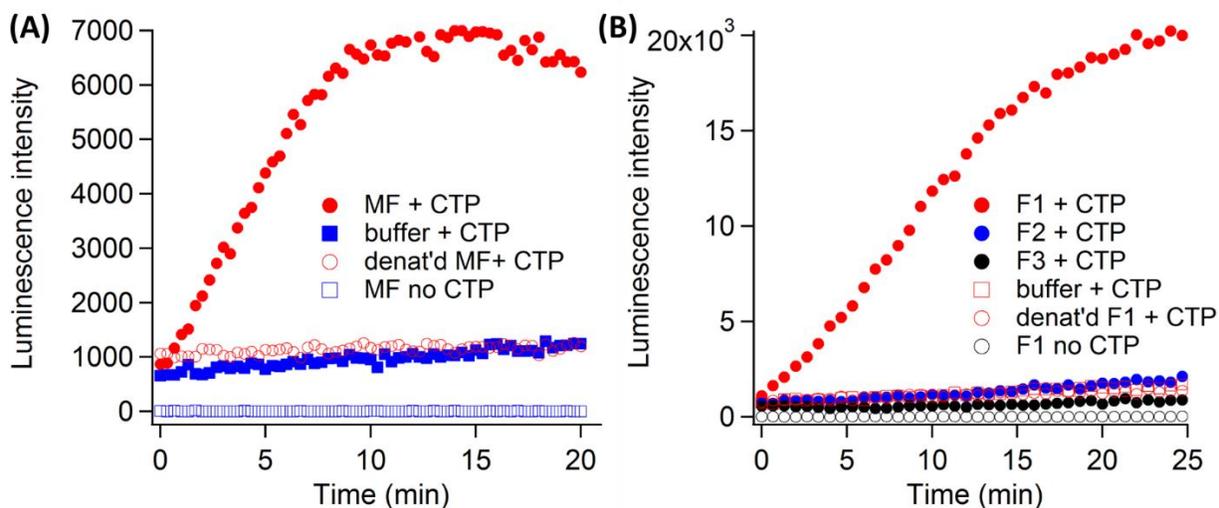

Figure 4.7. Cytidylate cyclase (CC) activity in various fractions obtained during attempted fractionation of the enzyme from rat brain. Activity was monitored using a pyrophosphate detection assay which generates a luminescence signal (Lonza PPiLight Pyrophosphate Assay Kit). **(A)** CC activity in membrane protein fraction (MF), along with control assays (buffer alone, denatured MF, and MF without CTP substrate). **(B)** CC activity in three fractions collected from sucrose density gradient ultracentrifugation (F1-F3) of the total MF, along with control assays (buffer alone, denatured F1, and F1 without CTP substrate).

A published procedure employing ammonium sulfate precipitation also has been adapted in an effort to identify a previously reported 3',5'-cCMP-binding kinase from rat liver.[27] Size-exclusion chromatography afforded a protein with M_r ~50 kDa, corresponding to the reported mass of the putative kinase (**Figure 4.8B**).[27] In addition, a band with M_r ~28 kDa (corresponding to the mass of the putative PDE, as mentioned above[19b]) was observed by SDS-PAGE (**Figure 4.8A**). These protein bands were proteolytically digested and the tryptic peptides were analyzed by LC-MS/MS to identify potential proteins comprising each gel band. Several characterized proteins were identified through the proteomics experiments; however, the data did not include any nucleotide-associated PDEs or kinases.
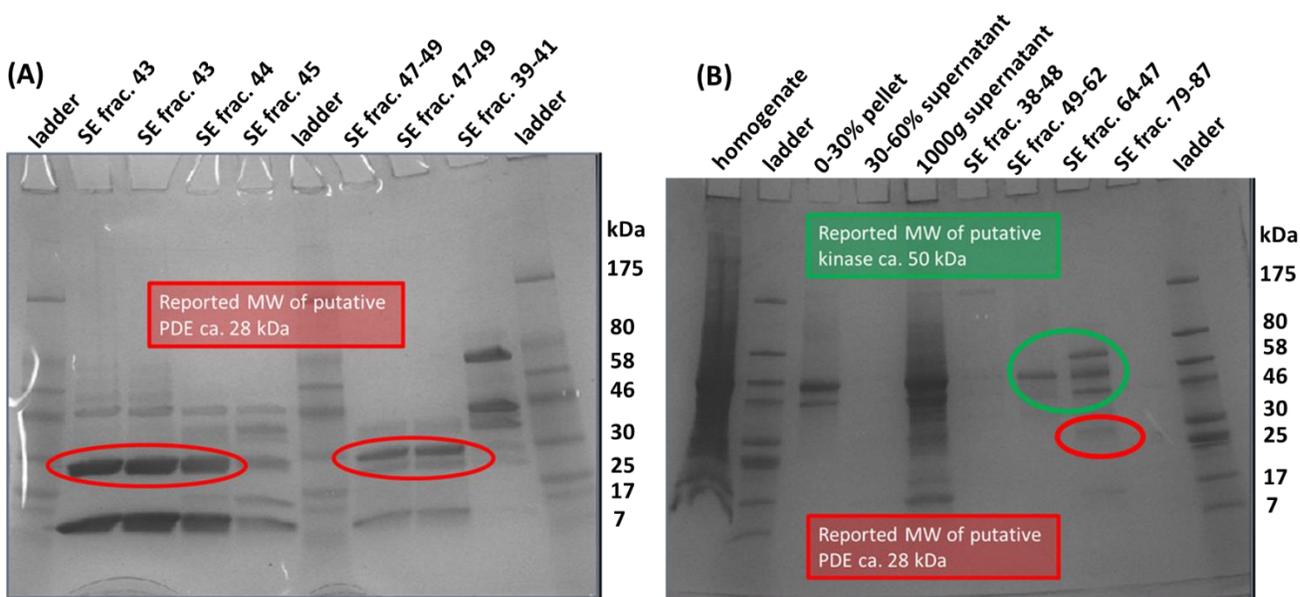
Figure 4.8. (A) SDS-PAGE analysis of size-exclusion (SE) fractions collected during attempted purification of a previously reported 3',5'-cCMP PDE from rat liver.[19b] SE fractionation was performed on the 60-90% ammonium sulfate precipitation fraction. **(B)** SDS-PAGE analysis of various fractions isolated during the attempted purification of previously reported 3',5'-cCMP PDE and kinase activity from rat brain.[19b,27]

The failure to obtain any hits for putative PDEs or cyclic nucleotide-responsive protein kinases through the proteomics approach warranted a bioinformatic search of the *Rattus norvegicus* proteome for predicted and uncharacterized proteins. This was accomplished by performing a BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) of the Gnomon database of rat proteins which contains proteins predicted to exist from *ab initio* models and proteins inferred to exist from homology, in addition to characterized proteins. In an effort to retrieve putative nucleotidyl cyclases, cyclic nucleotide-specific PDEs, and cyclic nucleotide-responsive kinases, BLAST searches were performed using the following *Rattus norvegicus* proteins as query sequences: the catalytic domain of adenylate cyclase type 3 (NCBI: NP_570135.2), the metal-dependent phosphohydrolase motif of 3',5'-cAMP specific PDE 4D isoform 1 (NCBI: NP_001106799.1), and the catalytic α subunit of 3',5'-cAMP-dependent protein kinase A (NCBI: NP_001094392.1). The similarity of the putative protein hits from the BLAST search relative to

the queried protein sequence was then assessed based on two criteria: the Expect (E) value reported by BLAST and the pairwise sequence identity score. The E value describes the likelihood that a protein match with a given similarity to the query sequence is a result of chance, as opposed to a real hit; thus the smaller the E value, the greater the probability that the hit is significant. Since the approximate M_r of the putative 3',5'-cCMP-specific PDE and 3',5'-cCMP-respsonsive kinase have been reported, only putative hits with predicted M_r in the range ± 5 kDa of the reported molecular masses were considered. Due to the difficulty of purifying the putative cytidylate cyclase, which is believed to be membrane-bound, no molecular mass has been approximated.

The BLAST search resulted in several hits when either the catalytic domain of adenylate cyclase type 3 or protein kinase A catalytic subunit α was used as the query sequence (**Table 4.1**). However, a similar BLAST search using the metal-dependent phosphohydrolase motif of 3',5'-cAMP -specific PDE 4D isoform 1 yielded only one hit within the predicted M_r range (**Table 4.1**). These bioinformatic analyses of the putative 3',5'-cCMP -related proteins show that several uncharacterized/predicted proteins potentially exist in the rat proteome with sequence similarity to the active sites of adenylate cyclase type 3 and protein kinase A, though fewer putative proteins containing a metal-dependent phosphohydrolase domain with a molecular mass ~28 kDa were found in the Gnomon database. This again illustrates the necessity of following the target enzymatic activity through the fractionation protocol prior to tandem MS proteomic analysis, as the molecular masses reported in the literature could be inaccurate.

Table 4.1. Putative/predicted proteins identified from a BLAST search of the Gnomon protein database using the specified query sequences (left-hand column). The number of hits is indicated in the right-hand column, along with the accession numbers. *No M_r is reported in the literature for the putative cytidylate cyclase, so the value in column 2 reflects only hits that satisfy the first two criteria.

Query sequence	Number of hits with E value ≤ 1e-24, with pairwise score ≥ 30%, and with Mr ± 5 kDa of query*
Adenylate cyclase type 3 (catalytic domain); NP_570135.2	**8 (gnl|GNOMON|4091412.p, gnl|GNOMON|4817248.p, gnl|GNOMON|4577870.p, gnl|GNOMON|4577877.p, gnl|GNOMON|5157072.p, gnl|GNOMON|5359558.p, gnl|GNOMON|5673064.p, gnl|GNOMON|5152234.p)**
3',5'-cAMP -specific PDE 4D isoform 1 (metal-dep. phosphohydrolase motif); NP_001106799.1	**1 (gnl|GNOMON|7300000.p)**
Protein kinase A, cat. subunit α; NP_001094392.1	**8 (gnl|GNOMON|6518632.p, gnl|GNOMON|4539985.p, gnl|GNOMON|6325693.p, gnl|GNOMON|6241522.p, gnl|GNOMON|5266602.p, gnl|GNOMON|5407924.p, gnl|GNOMON|4735367.p, gnl|GNOMON|7112642.p**

As an additional strategy to identify potential 3',5'-cCMP-binding proteins, 3',5'-cCMP was tethered to an insoluble agarose support to afford matrices for affinity chromatography. Notably, this pull-down approach has been employed to discover effectors of 3',5'-cAMP and -cGMP,[37] and analogous studies using immobilized 3',5'-cCMP have demonstrated that this pyrimidine nucleotide binds the regulatory subunits of protein kinase A.[38] Thus, two 3',5'-cCMP-agarose probes were synthesized in which the agarose resin was coupled to the nucleotide moiety

via the cyclic phosphodiester through either a (poly)propylamine (PPA) linker or a (poly)ethyleneglycol (PEG) linker (**Figure 4.9; compounds 4.8** and **4.9**). The 3',5'-cCMP probes utilized herein were synthesized analogously to published procedures. To this end, 3',5'-cCMP was prepared by refluxing 5'-CMP in the presence of dicyclohexylcarbodiimide, resulting in electrophilic phosphate activation and concomitant intramolecular attack by the 3'-hydroxyl group to afford the cyclic phosphodiester (**Figure 4.9; compound 4.6**).[39] The cyclic phosphodiester then was alkylated with bromobutyric acid in refluxing MeCN to afford the phosphotriester (**Figure 4.9; compound 4.7**).[40] Similar alkylation reactions of 3',5'-cyclic nucleotide phosphodiesters generate predominately the axial stereoisomer (**Figure 4.9; compound 4.7**) due to the Gauche effect,[40] suggesting that formation of the (*R*) phosphorous diastereomer is favored over production of the equatorial (*S*) isomer (**Figure 4.10**). The resulting carboxylate handle was coupled to the terminal amino group of the respective agarose matrices through carbodiimide-mediated amide formation to generate **compounds 4.8 and 4.9** (**Figure 4.9**).
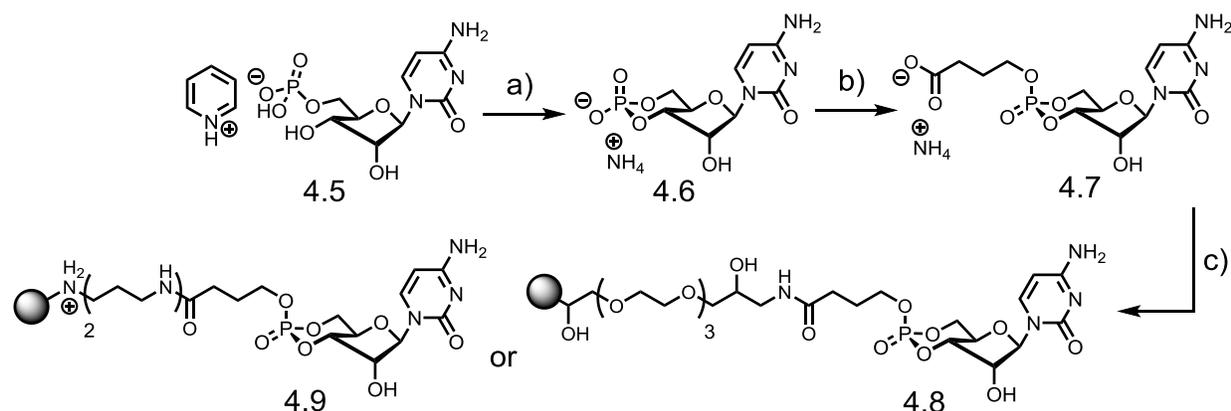
Figure 4.9. Synthesis of 3',5'-cCMP-agarose affinity matrix. **a)** Dicyclohexyl-carbodiimide, *N,N'*-dicyclohexyl-4-morpholinecarboxamidine, pyridine/DMF, reflux, 2.5 h; 25% yield. **b)** 4-Bromobutyric acid, MeCN, reflux, 5 h. **c)** 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride, coupling buffer (0.1 M MES, 0.9% NaCl, pH 7.4), 25°C, 2 h; then Carboxylink PEG agarose resin (to afford **compound 4.8**) or Carboxylink triaminopropyl agarose resin (to afford **compound 4.9**).

Prior to incubation with the immobilized 3',5'-cCMP derivatives, the rat organ homogenates were pre-treated with the unfunctionalized agarose beads in an effort to remove non-specific matrix-binding proteins. Subsequent SDS-PAGE analysis of proteins bound to the 3',5'-cCMP affinity matrices revealed several bands for both the PPA- and PEG-linked derivative. Based on the SDS-PAGE gel, certain proteins appear to display specificity for 3',5'-cCMP-binding (as opposed to non-specific agarose binding), as co-incubation of the immobilized 3',5'-cCMP matrix with unmodified 3',5'-cCMP reduced the intensity of certain bands on the gel. These competition experiments also suggested that the PEG-linked probe promotes more specific interactions with the nucleotide moiety compared to the PPA-linked derivative, as protein bands observed in the 3',5'-cCMP-PEG-agarose samples were more effectively competed away by free 3',5'-cCMP. Unfortunately, proteomic analysis of the tryptic peptides derived from these gel bands was unsuccessful due to poor sequence coverage.

compound 4.7 axial (R_P) diastereomer (favored)

compound 4.7 equatorial (S_P) diastereomer (disfavored)
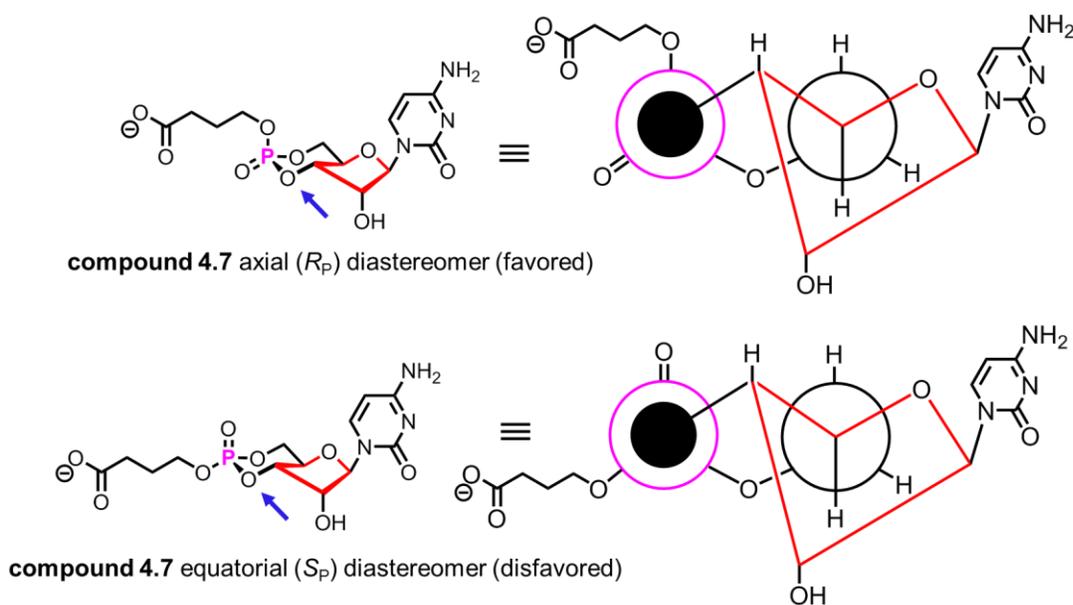
Figure 4.10. Newman projections depicting the axial (*R*) phosphate diastereomer (top) and the equatorial (*S*) phosphate diastereomer (bottom) from the vantage point indicated by the blue arrow, with the phosphorous atom indicated by the pink circle and the oxygen atom indicated by the solid black circle. Literature precedent suggests that the (*R*) isomer is thermodynamically favored due to the Gauche effect, as this orientation places the π* antibonding orbital of the P=O bond (the strongest electron acceptor) *anti* relative to the σ bonding orbital of the C-O bond (the strongest electron donor).[40]

4.3 Discussion

Recent studies have begun to elucidate the physiological functions of 3',5'-cCMP, along with potential metabolic enzymes that regulate intracellular concentrations of this pyrimidine.[14] The finding that intracellular levels of 3',5'-cCMP are similar to concentrations of the canonical second messenger 3',5'-cGMP in several rodent organs suggests the potential of 3',5'-cCMP-dependent processes (**Figure 4.3**). While different rat tissues contain similar 3',5'-cCMP concentrations, previous work has demonstrated that 3',5'-cCMP levels vary in different mammalian cell types,[11] suggesting tissue-specific roles for this cNMP. Moreover, bioanalytic studies suggest that 3',5'-cCMP is perhaps unique to metazoans, as this nucleotide was not detected in prokaryotes, nematodes, plants, fungi, or amoeba.[11] While the factors regulating 3',5'-cCMP concentrations in mammals remain enigmatic, prior studies have discovered that synthesis of 3',5'-cCMP is stimulated by bicarbonate and NO *via* activation of sAC and sGC activity, respectively, in mammalian cells.[12-13] Furthermore, the characterized 3',5'-cAMP phosphodiesterase PDE7A1 hydrolyzes 3',5'-cCMP *in vitro*,[17] suggesting that 3',5'-cCMP levels are governed, at least in part, by characterized enzymes involved in 3',5'-cAMP and -cGMP metabolism. The present work employed enzyme fractionation along with proteomic and bioinformatic analyses in an effort to identify additional, uncharacterized enzymes involved in 3',5'-cCMP biosynthesis and degradation.

A high-throughput luminescence-based assay has been developed and optimized for evaluating cytidylate cyclase activity in crude homogenates (**Figure 4.7**), which will facilitate additional efforts to further fractionate this putative membrane-associated activity for proteomic-based experiments. Future work will employ additional purification steps to supplement sucrose density gradient ultracentrifugation and 5'-CTP affinity chromatography. Additional efforts to

fractionate and characterize previously reported 3',5'-cCMP-specific PDE activity and 3',5'-cCMP-responsive protein kinase activity were largely unsuccessful. However, a BLAST search returned multiple uncharacterized/predicted proteins in the rat proteome with sequence homology to prototypical adenylate cyclases and 3',5'-cAMP-responsive protein kinases (**Table 4.1**), alluding to the possible existence of additional 3',5'-cNMP signaling pathways in mammals. The utilization of a PDE activity assay to guide fractionation would aid future efforts to identify putative PDEs. To this end, a phosphate-detection assay coupling 3',5'-cNMP cleavage with subsequent phosphatase-catalyzed hydrolysis of the acyclic 5'-NMP could be utilized. Identification of putative kinases and other potential 3',5'-cCMP effectors could be achieved through affinity chromatography methods, as immobilized 3',5'-cCMP-agarose has been prepared herein. Future work also will attempt to purify the target enzymes from mammalian cell lines in addition to organs, as this would reduce the complexity of the biological matrix.

4.4 Materials and Methods

Detergent-solubilization of putative cytidylate cyclase

The detergent-solubilization protocol is based on procedures published by Newton *et al*.,[41] Mori *et al*.,[15] and Yamamoto *et al*.[42] A flash-frozen rat liver (Innovative Research) was thawed on ice and homogenized on ice using an OMNI TIP homogenizer in 9 volumes of buffer (50 mM Tris-HCl (pH 7.5), 0.2 M sucrose, 100 mM NaCl, 2 mM DTT, 5 mM theophylline, 10 mM NaF, 1 mM AEBSF). The homogenate was centrifuged at $1000g$ for 5 min at 4°C, and the resulting supernatant further centrifuged at $12000g$ for 25 min. at 4°C. The pellet was resuspended in homogenizing buffer with 0.3% Triton X-100 detergent added

Sucrose density gradient ultracentrifugation

Sucrose density gradient fractionation was based on published procedure.[43] The sucrose density gradient was prepared in a 10-mL ultracentrifge tube by adding 3 mL of a cold solution containing 1.0 M sucrose, 20 mM Tris-HCl (pH 7.4), and 0.06% Triton X-100 to the Ultracentrifuge tube, followed by carefully layering 3 mL of a cold solution containing 0.1 M sucrose, 20 mM Tris-HCl (pH 7.4), and 0.06% Triton X-100 on top of the concentrated sucrose solution. The sucrose-containing buffers were supplemented with 5 mM theophylline, 2 mM DTT, 10 mM NaF, and 1 mM AEBSF. The membrane suspension then was carefully pipetted onto the sucrose gradient, and the tube was centrifuged at 103,000 x g for 3 h at 4°C. Three layers were collected from the sucrose density gradient and assay for cytidylate cyclase activity as described below.

Cytidylate cyclase activity assay

Cyclase activity was assayed in various fractions during the purification procedure (see **Detergent-solubilization of putative cytidylate cyclase** and **Sucrose density gradient ultracentrifugation** above) at 25°C using the PPiLight Pyrophosphate Detection Assay Kit (Lonza) according to the manufacturer's specifications. Reactions contained enzyme fraction (30 µL), 5'-CTP substrate (500 µM), $MnCl_2$ (tetrahydrate) (4 mM), and PPiLight kit mix (20 µL) in a total volume of 60 µL. Luminescence output was monitored in 96-well microtiter plates using a microplate reader (Biotek Synergy). Data were collected using a sensitivity setting of 145 and a reading number of 10. Kinetic readings were recorded every 20-40 s for 20 min using a 350 ms pause for each reading and a 1 ms delay between readings.

Partial purification of putative 3',5'-cCMP-specific phosphodiesterase

The fractionation procedure was based on that of Newton *et al*.[19b] A flash-frozen rat liver (Innovative Research) was homogenized on ice using an OMNI TIP Homogenizer in 9 volumes

of buffer (50 mM Tris-HCl [pH 7.5], 3 mM β-mercaptoethanol). The homogenate was centrifuged at 1000g for 10 min at 4°C. The resulting supernatant was dialyzed for 5 h at 4°C against 12 volumes of the same buffer. Following dialysis, the supernatant was subjected to $(NH_4)_2SO_4$ fractionation by slowly adding the amount of saturated $(NH_4)_2SO_4$ necessary to produce the following fractions: 0-20%, 20-40%, 40-60%, and 60-90%. The 60-90% fraction was further purified by gel filtration (GE Healthcare Life Sciences instrument) using a 26/600 Superdex 200 prep grade column with a 360-mL column volume. The column was eluted with buffer containing 50 mM Tris-HCl [pH 7.5], 50 mM NaCl, and 1 mM DTT at a flow rate of 1 mL/min and a temperature of 4°C.

SDS-PAGE electrophoresis and tryptic digestion in preparation for proteomic analysis

Prior to proteomic analysis, protein fractions were reduced by incubating with 16 mM SDS containing DTT at 90°C for 20 min and alkylated with 100 mM iodoacetamide for 30 min in the dark at 25°C. Denatured proteins then were separated on Criterion mini 4-20% SDS-PAGE gels (Bio-Rad). For digestion and extraction, bands corresponding to the M_r of the previously reported 3',5'-cCMP-specific PDE and kinase were cut into ~1-mm cubes and destained with 25 mM NH_4HCO_3 containing 50% MeCN with shaking at 37°C for 30 min. The destaining solution was removed and this process was repeated three more times. The gel pieces were shrunk by addition of 50 μL MeCN and incubated at 25°C for 15 min. After removal of MeCN, the gel pieces were dried at 25°C for 15 min, digested with trypsin (10 μL) according to the manufacturer's specifications (Thermo Fisher Scientific), and lyophilized to dryness. The lyophilized samples were then reconstituted in loading buffer (1% MeCN, 0.1% formic acid, 0.03% TFA) and loaded onto a C18 reversed-phase analytical column with a 100 μm internal diameter and a 15 cm length. A Water's NanoAcquity UPLC was used to elute the peptides over a linear gradient from 3% to

50% Buffer B (100% MeCN). Buffer A consisted of 1% MeCN and 0.1% formic acid. An LTQ Orbitrap XL mass spectrometer was used to collect 1 MS scan and 10 tandem MS/MS scan per cycle. The tandem MS/MS scans were then searched using a database curated from the NCBI RefSeq project. The database contained both target and decoy sequences that allowed for downstream false discovery rate calculations. The search engine employed was a proprietary version of Sorcerer-Sequest (ver 4.0.03) and the data was filtered using Emory University School of Medicine Proteomics Core software to a 1% false discovery rate.

Additional peptide database searching

The NIST MSPepSearch algorithm was used to search the rat_2011_05_24_it.001 library of tryptic digest tandem MS spectra. The Proteome Discoverer (Thermo Scientific) software suite was used to process the data. The NIST MF score threshold was set to 450 (which approximates a 1% false discovery rate). The constant mods considered were carbamidomethylation of Cys residues (as a result of IAA incubation), oxidation of Met residues, and cyclization of N-terminal Gln to pyro-Gln. In addition, The MASCOT algorithm was used to search the UniprotKB/TrEMBL database of rat proteins. The Proteome Discoverer (Thermo Scientific) software suite was used to process the data. The MOWSE score threshold was set to 20 (which approximates a 1% false discovery rate), and the same set of constant modifications were designated.

Modified procedure for partial purification of putative 3',5'-cCMP-specific PDE and putative 3',5'-cCMP-responsive protein kinase

A flash-frozen rat liver (Innovative Research) was thawed on ice and homogenized on ice using an OMNI TIP homogenizer in 9 volumes of buffer (50 mM TRIS-HCl (pH 7.5), 50 mM NaCl, 3mM β-mercaptoethanol). AEBSF (1 mM) was added immediately prior to homogenization. The

homogenate was centrifuged at 1000 rpm for 1 h min at 4°C. The resulting supernatant was then centrifuged at 35000 rpm (using a Beckman Coulter Ultracentrifuge equipped with a Type 45 Ti rotor) for 1 h at 4°C. The supernatant was subjected to 0-10% $(NH_4)_2SO_4$ precipitation and the soluble fraction subjected to an additional 10-45% $(NH_4)_2SO_4$ precipitation. The resulting soluble fraction was dialyzed against 10 volumes of buffer for 18 h at 4°C and subjected to an additional 45-50% $(NH_4)_2SO_4$ precipitation. The resulting pellet was resuspended in buffer and dialyzed against 10 volumes of buffer for 18 h at 4°C. The 45-90% $(NH_4)_2SO_4$ fraction was further purified by gel filtration (GE Healthcare Life Sciences instrument) using a 26/600 Superdex 200 preparative grade column with a 360 mL column volume. The column was eluted with buffer containing 50 mM Tris-HCl [pH 7.5], 50 mM NaCl, and 1 mM DTT at a flow rate of 1 mL/min and a temperature of 4°C.

Bioinformatic analysis of putative cCMP-related proteins

BLAST searches were performed using adenylate cyclase type 3 (catalytic domain) (NP_570135.2), 3',5'-cAMP specific PDE (NP_001106799.1) (metal-dependent phosphohydrolase motif), and protein kinase A catalytic subunit α (NP_001094392.1) as the query sequences. Sequences were queried against the Gnomon database of rat proteins using the blastp search algorithm. The M_r of the putative proteins was calculated using the Bioinformatics.org Sequence Manipulation Site: Protein Molecular Weight tool (http://www.bioinformatics.org/sms2/protein_mw.html). Pairwise identity scores were determined by aligning the sequence of each hit with the corresponding query sequence using the ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/) Multiple Sequence Alignment tool.[44]

Construction of calibration curves for 3',5'- and 2',3'-cNMP quantification

Calibration curves were generated by plotting the peak area ratio of cNMP:8-Br 3',5'-cAMP (internal standard; IS) against the corresponding concentration ratio using authentic cNMP standards; the concentrations of nucleotide stock solutions was determined by UV-vis spectroscopy. All nucleotide standards were purchased from MP Biomedicals (Solon, OH, USA), TCI America (Portland, OR, USA), Carbosynth (Berkshire, UK), BioLog (Bremen, Germany), or Sigma-Aldrich (St. Louis, MO, USA). Stock solutions of all cNMPs and internal standard 8-Br-cAMP were quantified using a UV-visible spectrophotometer (Cary Series, Agilent Technology, Santa Clara, CA, USA). Calibration curves were constructed by plotting the ratio of peak area for each cNMP/internal standard (IS) against the ratio of concentration of each cNMP/IS in each standard sample. Concentration of IS in each calibration curve is 1 μM, while concentrations of the cNMPs ranged from 0.05–2.5 μM.

Extraction of 3',5'- and 2',3'-cNMPs from rat organ homogenate

A flash-frozen rat liver (Innovative Research) was thawed on ice and homogenized at 4°C into 2:2:1 MeOH/H_2O/MeCN containing 1 mM EDTA and theophylline as PDE inhibitors using an OMNI TIP homogenizer (3 mL extraction solvent per gram of wet tissue). The homogenate was heated at 60°C for 10 min, cooled on ice, and centrifuged at $2427g$ for 90 min. The supernatant was collected and the pellet was washed twice with extraction solvent. The supernatant was collected and concentrated to dryness *in vacuo*. The extracted material was suspended in H_2O and centrifuged at $20000g$ for 30 min. The resulting supernatant was collected, spiked with IS (67 pmol per gram of wet tissue), and lyophilized to dryness. The material was re-dissolved in phosphate buffer (50 mM, pH 7.4) by vigorous vortexing and sonication in a water bath to generate samples for LC-MS/MS.

LC-MS/MS separation and quantification of 3',5'- and 2',3'-cNMPs

Extracts were analyzed by LC-MS/MS using a Thermo Electron LTQ-FTMS equipped with a Dionex Ultimate 3000 dual gradient pump, Shimadzu autosampler, and diode array detector. Instrumentation was controlled using Xcalibur and DCMSlink software (Thermo Fisher Scientific). Chromatographic separation of analytes was achieved using a reversed-phase C18 column (15x2.1 mm, 2.7 μm; Ascentis Express) equipped with a guard column (5x2.1 mm, 2.7 μm; Ascentis Express) using a mobile phase consisting of water (with 0.1% formic acid; A) and methanol (0.1% formic acid; B). The following program was utilized for analyte separation: 0-4 min, 0% B; 4-15 min, 0-1.5% B; 15-20 min, 1.5-8% B; 20-25 min, 8% B; 25-28 min, 8-15% B; 28-35 min, 15% B; 34-45 min, 0% B. The flow rate was 0.3 mL min^{-1}, and extracts were analyzed *via* 20 μL injections. Three independent LC-MS/MS runs were performed for each sample. The column was washed after every four tissue samples using the following protocol: 0-2 min, 0%-100% B, 0% C; 2-10 min, 100% B, 0% C; 10-12 min, 100%-0% B, 0%-100% C; 12-20 min, 0% B, 100% C; 20-25 min, 0% B, 100%-0% C; 25-40 min, 0% B, 0% C. Nucleotides were ionized by electrospray in the positive-ion mode using the following parameters: 5 kV needle voltage, 35 V capillary voltage, 275 °C capillary temperature, and 110 V tube lens voltage. Nucleotide detection was based on the protonated parent ion [M+H]$^+$ and analytes were quantified using the protonated nucleobase fragment ion, which was generated using a normalized collision energy of 35 eV with an activation Q of 0.250 and activation time of 30 ms.

General conditions for chemical synthesis and characterization

All reactions were conducted using anhydrous solvents (Drisolv®, MilliporeSigma) in oven-dried glassware under a nitrogen atmosphere with magnetic stirring. Thin layer chromatography (TLC) was performed on alumina-backed silica gel 60 F$_{254}$ plates; the plates were visualized under UV light. Diethylaminoethyl (DEAE) Sephadex® A-25 resin (GE Healthcare) was utilized for anion-

exchange chromatography. [1]H and ^{31}P nuclear magnetic resonance (NMR) spectra were recorded on Varian INOVA 400 and Mercury 300 spectrometers. Chemical shifts are reported in ppm and are referenced to the residual solvent signal for ^{1}H NMR (D_2O = 4.8 ppm). ^{13}C and ^{31}P NMR chemical shifts are referenced externally to the MeOH signal in D_2O (49.5 ppm) or the trimethylphosphate signal in D_2O (2.5 ppm), respectively. *J*-coupling values are reported in hertz (Hz). A Thermo LTQ FTMS was used to collect high-resolution mass spectra (HRMS).

Preparation of DEAE Sephadex® A-25 bicarbonate form and compound purification

DEAE Sephadex® A-25 resin was hydrated in 100 mM NH_4HCO_3, poured into a flash chromatography column, and washed with several column volumes of 100 mM NH_4HCO_3, followed by several column volumes of water. The crude material was dissolved in 5-25 mL of water, and the pH was adjusted to ~7-8 (based on pH paper). This solution was loaded onto the column and eluted with water for several fractions. Subsequently, the ionic strength of the eluent was slowly increased by gradual addition of aqueous NH_4HCO_3 to elute the desired 2',3'-cNMP analogs as the ammonium salts. The product-containing fractions were concentrated to dryness by lyophilization to remove NH_4HCO_3.

Preparation of nucleotide pyridinium salts from sodium salts

To enhance solubility in pyridine, the sodium cation of the nucleotide salt was exchanged for pyridinium by passing the nucleotide over a column of Amberlite® IR-120 (pyridinium form). To this end, the cation-exchange resin (free acid form) was hydrated in 10% (v/v) aqueous pyridine, and washed with several column volumes of 10% (v/v) aqueous pyridine, followed by washing with several column volumes of water to remove excess pyridine. The nucleotide sodium salt was dissolved in water and slowly passed through the column. The UV-active fractions were

lyophilized to obtain the corresponding pyridinium salt of the nucleotide. The presence of the pyridinium cation was confirmed by ^{1}H NMR spectroscopy.

Synthesis of 5'-CTP γ-phosphoramidate (4.3)

The phosphoramidate was synthesized analogously to published procedure.[35] 5'-CTP (disodium salt) was converted to the pyridinium salt as described above. The resulting pyridinium salt (0.35 mmol) was suspended in DMF (5 mL) and treated with dicyclohexylcarbodiimide (DCC) (2.5 mmol) and *N,N'*-dicyclohexyl-4-morpholinecarboxamidine (0.88 mmol) at 25°C. Formation of the desired cyclic trimetaphosphate intermediate was monitored by ^{31}P NMR (-24 ppm doublet and -23 ppm triplet); additional DCC was added as necessary to drive the reaction toward completion. 2,2'-(Ethylenedioxy)-bis(ethylamine) (360 µL; 2.5 mmol) was added dropwise and the suspension was continued for 1.5 h. ^{31}P NMR analysis revealed new signals at -2 (d), -12 (d), and -23 ppm (t); the -2 ppm resonance is characteristic of the γ-phosphorous of nucleoside phosphoramidates. The reaction was diluted with H_2O and the pH was adjusted to ~7 with aqueous AcOH. The mixture was extracted with three portions of Et_2O and the aqueous phase was concentrated *in vacuo* by azeotropic distillation with EtOH. The residue was purified by DEAE anion-exchange chromatography, as detailed above, to obtain the ammonium salt. The ammonium salt then was converted to the sodium salt using Amberlite IR-120 cation-exchange resin, in analogy to the protocol described above for preparation of the pyridinium salt. The compound was obtained in 30% yield. **^{1}H NMR** (400 MHz, D_2O) δ 7.98 (d, *J* = 7.6 Hz, 1H), 6.15 (d, *J* = 7.6 Hz, 1H), 5.97 (d, *J* = 4.2 Hz, 1H), 4.38 – 4.16 (m, 5H), 3.79 – 3.73 (m, 2H), 3.69 (s, 4H), 3.62 – 3.54 (m, 2H), 3.23 – 3.16 (m, 2H), 3.07 (dt, *J* = 9.7, 5.6 Hz, 2H). **^{13}C NMR** (101 MHz, D_2O) δ 165.91, 141.49, 96.55, 89.28, 82.47, 74.03, 71.07, 69.48, 66.32, 64.75, 40.78, 39.10, 32.19, 24.67, 24.27. **^{31}P NMR** (162 MHz, D_2O) δ -1.32 (d, *J* = 20.6 Hz), -11.51 (d, *J* = 19.1 Hz), -22.81 (t, *J* = 20.2 Hz).

Synthesis of 5'-CTP-agarose (4.4)

The 5'-CTP phosphoramidate **(compound 4.3)** (0.015 mmol) was dissolved in coupling buffer (0.1 M sodium phosphate, pH 7.5 with 0.15 M NaCl) and treated with NHS-agarose resin (330 mg) (Thermo Fisher Scientific). The mixture was rotated end-over-end at 25°C for 1 h; then the reaction was continued at 4°C for 18 h. The reaction was quenched by addition of 3 mL Tris-HCl buffer (1 M, pH 7.4) and the solution was centrifuged at $900g$ for 2 min. The supernatant was collected and the resin was washed three times with coupling buffer.

Synthesis of 3',5'-cCMP (4.6)

The reaction was performed in analogy to published conditions.[39] 5'-CMP (free acid) (0.31 mmol) and *N,N'*-dicyclohexyl-4-morpholinecarboxamidine (0.31 mmol) were suspended in DMF (20 mL) and heated to 80°C to dissolve solids. The mixture was added dropwise over 1.5 h into a refluxing solution of DCC (2.4 mmol) in pyridine (10 mL) and refluxed for 1 h longer. All solvent was removed *in vacuo* and the residue was partitioned between H_2O and Et_2O. The aqueous phase was washed with Et_2O (3x 25 mL) and concentrated to dryness *in vacuo*. The crude product was purified by DEAE anion-exchange chromatography, as described above. The compound was obtained in 25% yield. **^{1}H NMR** (400 MHz, D_2O) δ 8.19 (d, *J* = 8.0 Hz, 1H), 6.24 (d, *J* = 8.0 Hz, 1H), 5.90 (d, *J* = 3.7 Hz, 1H), 4.34 – 4.23 (m, 3H), 4.18 (ddd, *J* = 11.8, 4.4, 1.9 Hz, 1H), 4.05 (ddd, *J* = 11.9, 5.0, 1.9 Hz, 1H). **^{31}P NMR** (162 MHz, D_2O) δ -3.05. **HRMS** (ESI-, *m/z*) [M-H]$^-$ calcd for $C_9H_{11}O_7N_3P^-$ 304.0340, found 304.0339.

Synthesis of 3',5'-cCMP butyryl phosphate ester (4.7)

This procedure is based on literature precedent.[40] 3',5'-cCMP **(compound 4.6)** (0.164 mmol) was dissolved in MeCN (4 mL) at 25°C and treated with 4-bromobutyric acid (0.720 mmol). The

mixture was refluxed in the dark for 5 hours. The solvent was removed by rotary evaporation to obtain an off-white solid, which was washed with H_2O to obtain a light yellow oil. The crude product (as the tetrabutylammonium salt) was partially purified by recrystallization from hot MeOH/CHCl$_3$.

Synthesis of 3',5'-cCMP-PEG-agarose (4.8)

Coupling buffer (0.1 M MES, 0.9% NaCl, pH 4.7) (4 mL) was used to suspend the partially purified 3',5'-cCMP butyryl phosphate ester **(compound 4.7)**. The resulting white suspension was treated with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride (EDC-HCl) (0.268 mmol) and rotated end-over-end for 2 hours at 25°C, resulting in a clear, colorless solution. Carboxylink polyethyleneglycol-agarose resin (Thermo Fisher Scientific) (1.4 mL) was washed 3 times with water and added to the above solution. The white suspension was mixed end-over-end for 3 hours at 25°C. The resin then was washed 3 times with 2 mL of MeOH, followed by 3 times with distilled water. The supernatants from the washes were saved and used to quantify the unbound cNMP *via* UV-visible spectrophotometry using Beer's law (79% yield).

Synthesis of 3',5'-cCMP-PPA-agarose (4.9)

The affinity matrix was synthesized in analogy to **compound 4.8** above by coupling the partially purified 3',5'-cCMP butyryl phosphate ester **(compound 4.7)** with Carboxylink triaminopropyl-agarose resin (Thermo Fisher Scientific) (1.4 mL) (55% yield).

4.5 References

(1) Hardman, J.; Robison, G.; Sutherland, E. *Annu. Rev. Physiol.* **1971**, *33*, 311-336.

(2) Kamenetsky, M.; Middelhaufe, S.; Bank, E.; Levin, L.; Buck, J.; Steegborn, C. *J. Mol. Biol.* **2006**, *362*, 623-639.

(3) Francis, S.; Blount, M.; Corbin, J. *Physiol. Rev.* **2011**, *91*, 651-690.

(4) Kato, K.; Omura, H.; Ishitani, R.; Nureki, O.; Kornberg, R. *Annual Review of Biochemistry, Vol 86* **2017**, *86*, 541-566.

(5) Tanaka, Y.; Chen, Z. *Science Signaling* **2012**, *5*.

(6) McWhirter, S.; Barbalat, R.; Monroe, K.; Fontana, M.; Hyodo, M.; Joncker, N.; Ishii, K.; Akira, S.; Colonna, M.; Chen, Z.; Fitzgerald, K.; Hayakawa, Y.; Vance, R. *J. Exp. Med.* **2009**, *206*, 1899-1911.

(7) Zhang, X.; Shi, H.; Wu, J.; Zhang, X.; Sun, L.; Chen, C.; Chen, Z. *Mol. Cell* **2013**, *51*, 226-235.

(8) (a) Newton, R.; Kingston, E.; Hakeem, N.; Salih, S.; Beynon, J.; Moyse, C. *Biochem. J* **1986**, *236*, 431-439; (b) Newton, R.; Salih, S.; Salvage, B.; Kingston, E. *Biochem. J* **1984**, *221*, 665-673; (c) Bahre, H.; Hartwig, C.; Munder, A.; Wolter, S.; Stelzer, T.; Schirmer, B.; Beckert, U.; Frank, D.; Tummler, B.; Kaever, V.; Seifert, R. *Biochem. Biophys. Res. Commun.* **2015**, *460*, 909-914; (d) Jia, X.; Fontaine, B. M.; Strobel, F.; Weinert, E. E. *Biomolecules* **2014**, *4*, 1070-1092.

(9) (a) Beste, K.; Burhenne, H.; Kaever, V.; Stasch, J.; Seifert, R. *Biochemistry* **2012**, *51*, 2357-2357; (b) Chen, Z.; Zhang, X.; Ying, L.; Dou, D.; Li, Y.; Bai, Y.; Liu, J.; Liu, L.; Feng, H.; Yu, X.; Leung, S.; Vanhoutte, P.; Gao, Y. *American Journal of Physiology-Heart and Circulatory Physiology* **2014**, *307*, H328-H336.

(10) Goble, A.; Feng, Y.; Raushel, F.; Cronan, J. *Acs Chemical Biology* **2013**, *8*, 2622-2629.

(11) Hartwig, C.; Bahre, H.; Wolter, S.; Beckert, U.; Kaever, V.; Seifert, R. *Neurosci. Lett.* **2014**, *579*, 183-187.

(12) Hasan, A.; Danker, K.; Wolter, S.; Bahre, H.; Kaever, V.; Seifert, R. *Biochem. Biophys. Res. Commun.* **2014**, *448*, 236-240.

(13) Bahre, H.; Danker, K.; Stasch, J.; Kaever, V.; Seifert, R. *Biochem. Biophys. Res. Commun.* **2014**, *443*, 1195-1199.

(14) Seifert, R. *Trends Biochem. Sci.* **2015**, *40*, 8-15.

(15) Mori, S.; Yanagida, M.; Kubotsu, K.; Yamamoto, I. *Second Messengers and Phosphoproteins* **1990**, *13*, 1-12.

(16) Reinecke, D.; Burhenne, H.; Sandner, P.; Kaever, V.; Seifert, R. *FEBS Lett.* **2011**, *585*, 3259-3262.

(17) Monzel, M.; Kuhn, M.; Bahre, H.; Seifert, R.; Schneider, E. *FEBS Lett.* **2014**, *588*, 3469-3474.

(18) Cheng, Y.; Bloch, A. *J. Biol. Chem.* **1978**, *253*, 2522-2524.

(19) (a) Kuo, J.; Brackett, N.; Shoji, M.; Tse, J. *J. Biol. Chem.* **1978**, *253*, 2518-2521; (b) Newton, R.; Salih, S. *Int. J. Biochem.* **1986**, *18*, 743-&.

(20) Helfman, D.; Brackett, N.; Kuo, J. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 4422-4425.

(21) Laue, S.; Winterhoff, M.; Kaever, V.; van den Heuvel, J.; Russel, F.; Seifert, R. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2014**, *387*, 893-895.

(22) Wolter, S.; Golombek, M.; Seifert, R. *Biochem. Biophys. Res. Commun.* **2011**, *415*, 563-566.

(23) Wolter, S.; Dove, S.; Golombek, M.; Schwede, F.; Seifert, R. *Naunyn-Schmiedeberg's Arch. Pharmacol.* **2014**, *387*, 1163-1175.

(24) Schroder, R.; Janssen, N.; Schmidt, J.; Kebig, A.; Merten, N.; Hennen, S.; Muller, A.; Blattermann, S.; Mohr-Andra, M.; Zahn, S.; Wenzel, J.; Smith, N.; Gomeza, J.; Drewke, C.; Milligan, G.; Mohr, K.; Kostenis, E. *Nat. Biotechnol.* **2010**, *28*, 943-950.

(25) Beckert, U.; Grundmann, M.; Wolter, S.; Schwede, F.; Rehmann, H.; Kaever, V.; Kostenis, E.; Seifert, R. *Biochem. Biophys. Res. Commun.* **2014**, *451*, 497-502.

(26) (a) Akimoto, M.; Zhang, Z.; Boulton, S.; Selvaratnam, R.; VanSchouwen, B.; Gloyd, M.; Accili, E.; Lange, O.; Melacini, G. *J. Biol. Chem.* **2014**, *289*, 22205-22220; (b) Zong, X.; Krause, S.; Chen, C.; Kruger, J.; Gruner, C.; Cao-Ehlker, X.; Fenske, S.; Wahl-Schott, C.; Biel, M. *J. Biol. Chem.* **2012**, *287*, 26506-26512.

(27) Newton, R.; Khan, J.; Brenton, A.; Langridge, J.; Harris, F.; Walton, T. *Rapid Commun. Mass Spectrom.* **1992**, *6*, 601-607.

(28) Bloch, A. *Biochem. Biophys. Res. Commun.* **1974**, *58*, 652-659.

(29) Bloch, A.; Dutschman, G.; Maue, R. *Biochem. Biophys. Res. Commun.* **1974**, *59*, 955-959.

(30) Murphy, B.; Stone, J. *Proc. Soc. Exp. Biol. Med.* **1980**, *163*, 302-304.

(31) Scavennec, J.; Carcassonne, Y.; Gastaut, J.; Blanc, A.; Cailla, H. *Cancer Res.* **1981**, *41*, 3222-3227.

(32) Ervens, J.; Seifert, R. *Biochem. Biophys. Res. Commun.* **1991**, *174*, 258-267.

(33) (a) Beckert, U.; Wolter, S.; Hartwig, C.; Bahre, H.; Kaever, V.; Ladant, D.; Frank, D.; Seifert, R. *Biochem. Biophys. Res. Commun.* **2014**, *450*, 870-874; (b) Stevens, T.; Ochoa, C.; Morrow, K.; Robson, M.; Prasain, N.; Zhou, C.; Alvarez, D.; Frank, D.; Balczon, R.; Stevens, T. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **2014**, *306*, L915-L924.

(34) Helfman, D.; Kuo, J. *J. Biol. Chem.* **1982**, *257*, 1044-1047.

(35) Suto, R.; Whalen, M.; Bender, B.; Finke, R. *Nucleosides & Nucleotides* **1998**, *17*, 1453-1471.

(36) Haystead, C.; Gregory, P.; Sturgill, T.; Haystead, T. *Eur. J. Biochem.* **1993**, *214*, 459-467.

(37) (a) Corrie, J.; Pizza, C.; Makwana, J.; King, R. *Protein Expression Purif.* **1992**, *3*, 417-420; (b) Robichon, A. *J. Cell. Biochem.* **1991**, *47*.

(38) Hammerschmidt, A.; Chatterji, B.; Zeiser, J.; Schroder, A.; Genieser, H.; Pich, A.; Kaever, V.; Schwede, F.; Wolter, S.; Seifert, R. *Plos One* **2012**, *7*.

(39) (a) Smith, M.; Drummond, G.; Khorana, H. *J. Am. Chem. Soc.* **1960**, *83*, 698-706; (b) Wierenga, W.; Woltersom, J. *Journal of Carbohydrates-Nucleosides-Nucleotides* **1977**, *4*, 189-198.

(40) (a) Eckardt, T.; Hagen, V.; Schade, B.; Schmidt, R.; Schweitzer, C.; Bendig, J. *J. Org. Chem.* **2002**, *67*, 703-710; (b) Kataoka, S.; Isono, J.; Yamaji, N.; Kato, M. *Chem. Lett.* **1986**, 10.1246/cl.1986.1221, 1221-1224.

(41) Newton, R.; Salvage, B.; Hakeem, N. *Biochem. J* **1990**, *265*, 581-586.

(42) Yamamoto, I.; Takai, T.; Mori, S. *Biochim. Biophys. Acta* **1989**, *993*, 191-198.

(43) Swiatek-de Lange, M.; Muller, B.; Ueffing, M. In *Methods in Molecular Biology: Functional Proteomics: Methods and Protocols*; Thompson, J., Ed.; Humana Press: Totowa, NJ, 2008; Vol. 484, p 161-173.

(44) (a) Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. *Nucleic Acids Res.* **2010**, *38*, W695-W699; (b) Larkin, M.; Blackshields, G.; Brown, N.; Chenna, R.; McGettigan, P.; McWilliam, H.; Valentin, F.; Wallace, I.; Wilm, A.; Lopez, R.; Thompson, J.; Gibson, T.; Higgins, D. *Bioinformatics* **2007**, *23*, 2947-2948.

Chapter 5: Development of phyto-inspired phenolic glycosides as inhibitors of

Staphylococcus aureus **biofilm formation**

Adapted from: Fontaine, B. M.*; Nelson, K.*; Lyles, J. T.; Jariwala, P. B.; García-Rodriguez, J. M.; Quave, C. L.; Weinert, E. E. Identification of Ellagic Acid Rhamnoside as a Bioactive Component of a Complex Botanical Extract with Anti-biofilm Activity. *Front. Microbiol.* **2017**, *8*(496), DOI: 10.3389/fmicb.2017.00496 (* contributed equally) (open access).

Chapter 5: Development of phyto-inspired phenolic glycosides as inhibitors of *Staphylococcus aureus* biofilm formation

5.1 Introduction

5.1.1 Antibiotic resistance and biofilm formation

The dawn of the antibiotic era in the early twentieth century revolutionized medicine. However, ease of access to antibiotics and increasing prophylactic use in the clinic and in agriculture threatens to engender an epidemic of antibiotic-resistant bacteria. Reports by the U.S. Centers for Disease Control (CDC) indicate that 23,000 deaths in the US resulted from antibiotic-resistant infections in 2013, placing a $20 billion burden on the healthcare system.[1] Resistant isolates have emerged for every antibiotic scaffold.[2] Among the most serious threats are methicillin-resistant *Staphylococcus aureus* strains (which comprise nearly 50% of U.S. isolates), cephalosporin-resistant *Escherichia coli*, fluoroquinolone-resistant *Pseudomonas aeruginosa*, and vancomycin-resistant enterococci.[1,3] Compounding the threat posed by the increasing prevalence of resistant isolates is a diminished investment by major pharmaceutical companies in antibiotic development.[1] If such trends continue, projections estimate that drug-resistant infections would claim 10 million lives per year by 2050, exceeding the current annual cancer death toll by nearly 2 million.[3c] Another major challenge in eradicating bacterial infections is the formation of surface-attached, heterogeneous microbial communities known as biofilms.[4] These matrices of extracellular polysaccharides, lipids, proteins, and nucleic acids facilitate bacterial evasion of environmental insults such as the host immune system and chemotherapeutics.[5] In addition, biofilm-associated cells exhibit increased tolerance to antibiotics compared to planktonic bacteria due to altered metabolism.[6] Consequently, elucidating additional signaling pathways that govern biofilm production will offer novel strategies to disable essential biofilm-related processes, which

will augment our arsenal against bacterial invaders and facilitate the study of microbial ecology in complex habitats, such as within the microbiome.

5.1.2 Phenolic phytochemicals as modulators of bacterial biofilm production and quorum sensing

In addition to the goal of identifying unknown facets of RNA and nucleotide signaling, this work aims to identify phytochemicals as novel inhibitors of bacterial biofilm formation, with a particular focus on *Staphylococcus aureus*. The presence of methicillin-resistant *S. aureus* biofilm infections on medical devices nearly doubles the treatment cost per infection and the duration of the hospital stay, compared to methicillin-sensitive infections.[7] Moreover, biofilm-associated *S. aureus* cells display increased multi-drug resistance relative to biofilm-deficient strains, illustrating the significance of developing new strategies to perturb biofilm formation.[8] Staphylococci produce a variety of membrane-associated proteins that enable attachment to a myriad of human matrix polymers and abiotic surfaces. Following initial adhesion to the surface, the bacteria replicate, secreting proteins and polysaccharides that facilitate inter-cell attachment,[9] including nucleoid-associated proteins which bind extracellular DNA (eDNA) in the biofilm matrix.[10] Intriguingly, time-resolved confocal microscopy studies recently have identified a novel facet of *S. aureus* biofilm development involving degradation of eDNA via heterogeneous expression of extracellular nuclease Nuc within the biofilm, resulting in the exquisitely controlled release of a cellular subpopulation; the biological implications of this remodeling event remain to be elucidated.[11] Further maturation of the biofilm is orchestrated by heterogeneous metabolic reprogramming which induces the formation of microcolonies and tower-like structures that facilitate cell survival within the matrix.[12] Subsequent dispersal of the established biofilm in response to certain environmental stimuli enables further colonization.[13]
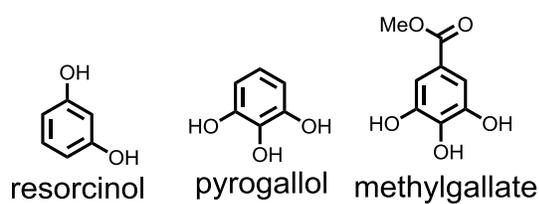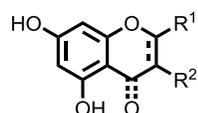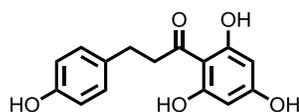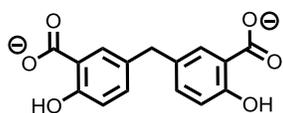
Figure 5.1. Chemical structures of representative phtyo phenols with anti-QS and/or anti-biofilm activity.

The *S. aureus* accessory gene regulator (Agr) quorum sensing system regulates multiple aspects of biofilm development. Staphylococcal Agr QS is mediated by a macrocyclic thiolactone-containing peptide encoded by *agrD*.[14] The ribosomally synthesized AgrD precursor peptide undergoes thiolactonization and excretion by transmembrane peptidase AgrB; this post-translational processing can produce slightly different peptides in individual staphylococci species.[15] Following export from the cell, the AgrD macrocycle binds the sensor histidine kinase AgrC which phosphorylates AgrA, resulting in divergent transcription of the *agrBDCA* operon and the *hld* gene encoding the toxin δ-hemolysin.[16] In addition to encoding this hemolytic protein, the *hld* mRNA (known as RNA-III) post-transcriptionally regulates the expression of certain genes involved in biofilm formation and virulence. More specifically, RNA-III sequesters the Shine-Dalgarno sequence of specific transcripts, including biofilm-related adhesins and the transcription factor Rot (Repressor of toxin), to attenuate biofilm formation and de-repress virulence toxin expression.[15] AgrA also mediates biofilm dispersal *via* upregulation of amphiphilic peptides known as phenol-soluble modulins (PSMs). Consequently, *S. aureus* mutants with a

dysregulated Agr system exhibit increased biofilm production, reduced biofilm dissemination, and diminished virulence.[17]

S. aureus biofilms frequently colonize implanted medical devices and atopic dermatitis lesions. Strikingly, the *S. aureus* inoculum required to initiate infections is reduced by five orders of magnitude in the presence of an implanted device,[18] and over 80% of atopic lesions are colonized by *S. aureus*.[19] Intriguingly, ancient civilizations utilized topical botanical preparations to treat dermatological inflammation, with certain remedies eliciting a statistically significant improvement of the affliction in blinded trials.[20] Modern chromatographic, synthetic, and analytical methods have enabled the isolation and identification of active phytochemicals from complex botanical mixtures. In addition, recent work has identified many secondary metabolites from plants and fungi that impair biofilm formation, virulence, and quorum sensing in Staphylococci and other pathogens[21] (**Figure 5.1**). Importantly, certain phytocompounds inhibit these microbial processes without impacting growth or viability, which presumably hinders the emergence of resistant isolates. For example, flavonoids disrupt bacterial cell-surface adhesion and impair cellular metabolism within the biofilm matrix, displaying anti-biofilm activity against gram-positive *S. aureus*, gram-negative *P. aeruginosa*, and the fungal pathogen *Candida albicans*.[21c] Phenolic flavones also covalently inhibit translocases of the *Salmonella* type III secretion system, protecting mammalian cells against gram-negative bacterial infection.[21h] Phytophenolics modulate virulence-associated processes in gram-positive pathogens as well. The dihydrochalcone phloretin, a component of apples, attenuates production of the α-hemolysin toxin in *S. aureus*,[21g] and the aromatic dicarboxylate 5,5'-methylenedisalicylic acid impairs virulence factor production through inhibition of Stp1, a phosphatase which indirectly upregulates transcription of virulence genes *via* dephosphorylation of kinase Stk1.[21f] Moreover, the clove oil

constituent 4-allyl-2-methoxyphenol impedes initial biofilm colonization and disseminates established *S. aureus* biofilms.[21e]

While the mechanism of action is unclear for most aromatic phytochemicals, polyphenol antioxidants isolated from apples bind several human nucleotide metabolic enzymes *in silico*,[22] and phenolic amides generated from biomass hydrolysis perturb nucleotide metabolism in *E. coli via* competitive inhibition of glutamine amidotransferases essential for *de novo* purine and pyrimidine biosynthesis.[23] These findings suggest that certain phytochemicals perhaps dysregulate biofilm formation, quorum sensing, and virulence through perturbation of cellular nucleotide levels. Additionally, certain flavonoids and other

ellagic acid

potential EA glycosides

Figure 5.2. Chemical structure of ellagic acid and potential structures of ellagic acid (EA) glycosides identified in the 220D-F2 extract prepared from *Rubus ulmifolius.* The regiochemistry and stereochemistry of the EA glycosides are unknown.

phytophenolics function as antioxidants and metal chelators, which could elicit biological effects.[24] Recent work also has elucidated the mechanism of biofilm inhibition by epigallocatechin gallate (EGCG), a major antioxidant in *Camellia sinensis* (green tea). This flavonol ester disrupts *E. coli* biofilm production through two mechanisms: by directly interfering with extracellular assembly of curli subunits CsgB and CsgA and by downregulating expression of *csgBAC* and *dgcC*, which encodes a DGC and promotes cellulose synthesis.[25] The wide-spread occurrence of curli amyloids in biofilms produced by diverse bacterial phyla[26] suggests that disruption of curli amyloidgenesis

could mediate the anti-biofilm activity of EGCG and other phytophenols in other bacterial species. In fact, EGCG also abrogates biofilm production in clinical isolates of *S. aureus* and *S. epidermis*.[27]

Staphylococcal biofilm formation also is disrupted by a phenolic-containing botanical extract prepared from the Elmleaf blackberry plant (*Rubus ulmifolius*).[21d] Rural Italian communities utilize preparations of this Mediterranean shrub to treat dermatological conditions,[28] and activity-based fractionation of *R. ulmifolius* roots identified an extract containing the tetraphenolic dilactone ellagic acid (EA) and various glycosylated derivatives (along with other unidentified molecules) that inhibits *S. aureus* biofilm production[21d] (**Figure 5.2**). However, structural details of the EA glycosides remained elusive, such as the identity of the carbohydrate, the anomeric stereochemistry, and the glycosylation site on the EA component. To elucidate these molecular details, the present work reports the synthesis of EA glycosides and related structural analogs. These studies enabled evaluation of the anti-biofilm activity of EA derivatives present in the *R. ulmifolius* botanical extract, resulting in the identification of an EA rhamnoside derivative as a novel biofilm inhibitor active against *S. aureus*.

5.2 Results

Phenol (**5.1**) and catechol (**5.2**) were glycosylated using per-*O*-acetyl glycosyl donors[29] in the presence of BF$_3$-OEt$_2$.[30] Reactions with acetates of *α*-L-rhamnopyranose and *α*-D-mannopyranose afforded anomerically pure *α*-glycosides (**Figure 5.3; compounds 5.3, 5.4, 5.6, 5.7**), while glycosylation with *β*-D-xylopyranose provided only the *β*- glycoside (**Figure 5.3; compound 5.5**). The resulting protected glycosides were de-*O*-acetylated with triethylamine in methanol,[31] yielding the desired glycosides of phenol (**5.8, 5.10, 5.11**) and catechol (**5.9, 5.12**) (**Figure 5.3**).
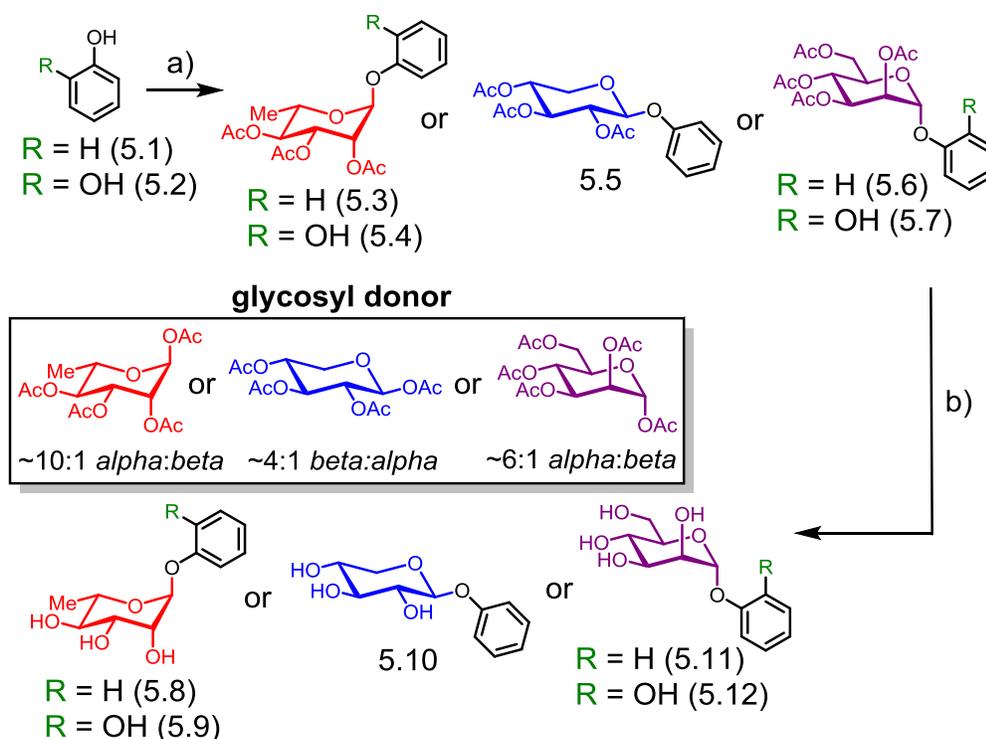
Figure 5.3. Synthesis of phenyl and catechyl glycosides. **a)** Glycosyl donor, BF$_3$-OEt$_2$, CH$_2$Cl$_2$, 25°C; 10-59% yield. **b)** MeOH, Et$_3$N, 25°C; 99% yield.

As EA is extremely insoluble and quite unreactive, synthesis of EA glycosides employed per-O-acetyl α-glycosyl iodide donors,[32] synthesized by treating per-O-acetyl sugars with trimethylsilyl iodide (TMSI) generated *in situ* from iodine and hexamethyldisilane (**Figure 5.4**).[33] To control the regioselectivity of the glycosylation, ellagic acid was first protected as the per-O-TBS silyl ether (**Figure 5.5; compound 5.14**).[34] A previously published X-ray crystal structure indicates that Bu$_4$NF-mediated deprotection of a per-O-silyl EA derivative occurs preferentially at the 3- and 3'-silyl ethers to afford the 3,3'-bisphenolate *in situ*, presumably due to inductive effects of the *meta* oxygen atom of the lactone.[34]

The mild fluoride source tris(dimethylamino)sulfonium difluorotrimethylsilicate (TASF)[35] effected removal of a single silyl ether *in situ* to generate the 3-monophenolate of EA, which

underwent glycosylation in the presence of a per-*O*-acetyl α-glycopyranosyl iodide to furnish the

desired 3-*O*-glycosides as single anomers. Glycosylation with α-L-rhamnopyranosyl iodide or α-

D-xylopyranosyl iodide afforded the α-rhamnopyranoside (**5.15**) or β-xylopyranoside (**5.16**) of

ellagic acid, respectively (**Figure 5.4**). Due to the *cis*-1,2 configuration of the α-xylosyl iodide

donor, addition of tetrabutylammonium iodide (TBAI) was crucial to promote *in situ*

anomerization of the sugar to the more reactive β-iodide.[32d,32e] This facilitated displacement of the

equatorial iodide by the adjacent acetoxy substituent, providing solely the β-xyloside (**5.16**) upon

reaction of the acetoxonium intermediate with the protected ellagic acid acceptor. The protected

EA glycosides then were de-*O*-silylated using K_2CO_3 in wet DMF[36] and subsequently de-*O*-

acetylated using K_2CO_3 in MeOH/H_2O. A final treatment with cation-exchange resin afforded the

desired 3-*O*-glycosides of ellagic acid (**5.17, 5.18**) (**Figure 5.4**).

per-*O*-Ac sugar   a) α–glycosyl iodide
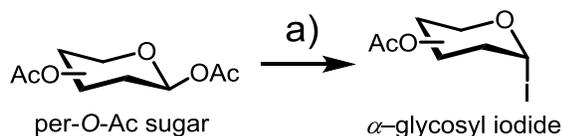
Figure 5.4. Synthesis of α-glycosyl iodides from per-*O*-acetyl sugars. **a)** Iodine, hexamethyldisilane, CH_2Cl_2, 25°C; 81-83% yield.

The anomeric stereochemistry of the β-xylosides (**5.10, 5.18**) was assigned based on the

characteristic vicinal $^3J_{1,2}$ coupling of approximately 10 Hz in the ^{1}H NMR spectrum, indicative of

diaxial coupling. However, the axial 2-hydroxy substituent of rhamnose and mannose precludes

elucidation of anomeric stereochemistry based on $^3J_{1,2}$ coupling, as both the α and β anomers have

$^3J_{1,2}$ coupling constants in the 0-3 Hz range.[37] Alternatively, the α configuration of phenolic

rhamnosides (**5.8, 5.9, 5.17**) and mannosides (**5.11, 5.12**) was deduced from the anomeric $^1J_{CH}$

coupling constant of approximately 170 Hz, which is in accord with the typical magnitude for α-

glycosides, compared to approximately 160 Hz for the corresponding β-glycosides.[37]
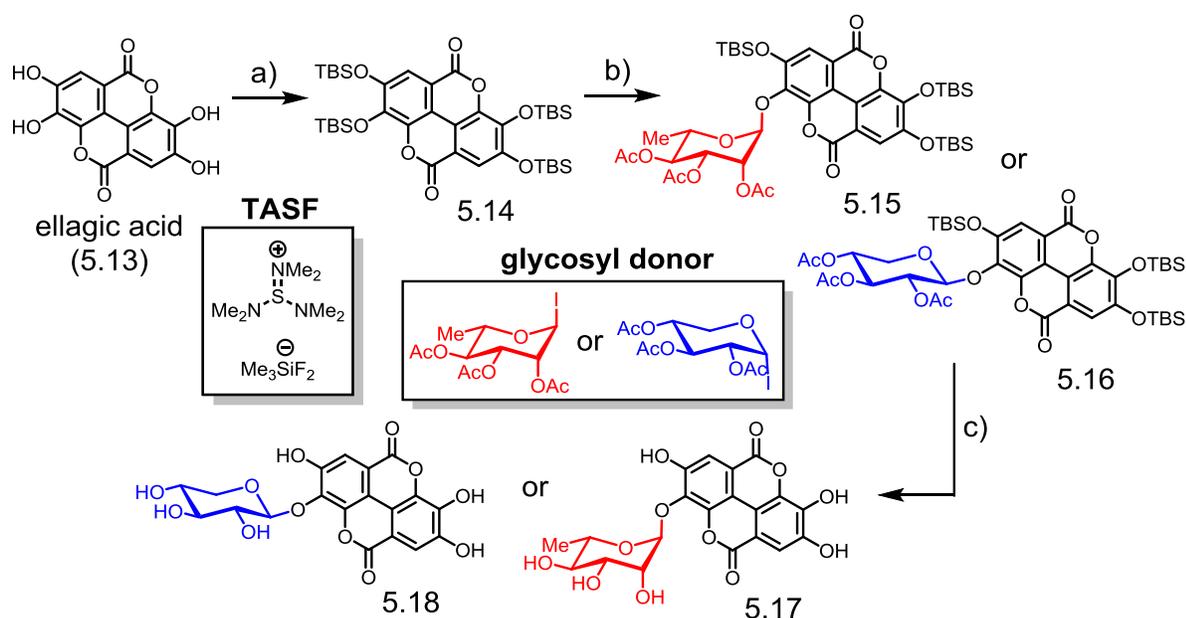
Figure 5.5. Synthesis of ellagic acid glycosides. **a)** TBS-Cl, imidazole, 4-(dimethylamino)-pyridine, CH_2Cl_2/DMF, reflux, 48 h; 71% yield. **b)** Tris(dimethylamino)sulfonium difluorotrimethylsilicate (TASF), CH_2Cl_2, 25°C; *then* glycosyl donor, Bu₄NI (for xylose), reflux; 11% yield (rhamnoside); 15% yield (xyloside). **c)** 1) K_2CO_3, DMF/H_2O (10:1), 25°C; 2) K_2CO_3, MeOH/H_2O (10:1), 25°C; 86% yield (rhamnoside); 92% yield (xyloside).

The synthetic glycosides were utilized as standards in LC-MS experiments to test for the presence of aromatic glycosides in the 220D-F2 extract. The data revealed the presence of both EA xyloside (**5.18**) and EA rhamnoside (**5.17**) in the botanical extract (**Figure 5.6**), based on co-elution with the synthetic standards and the *m/z* of the respective molecular ions. Conversely, the phenyl and catechyl glycosides were not detected.

The EA glycosides and structural analogs subsequently were evaluated as modulators of *S. aureus* growth and biofilm formation (**Table 5.1, Figures S5.1** and **S5.2**). The 220D-F2 and EA also were assayed as controls, as both the extract and EA inhibit *S. aureus* biofilm formation without impacting bacterial growth or survival.[21d] Crystal violet biofilm-staining assays identified phenol (**5.1**), phenyl rhamnoside (**5.8**), catechyl rhamnoside (**5.9**), and ellagic acid rhamnoside (**5.17**) as biofilm inhibitors lacking bacteriostatic or bactericidal properties (**Table 1**, **Figures S5.1**

and **S5.2**). Ellagic acid xyloside (**5.18**) also inhibited biofilm formation, but the $MBIC_{50}$ (64 µg mL^{-1}) was higher than the MIC_{50} (32 µg mL^{-1}), demonstrating that the biofilm inhibition is at least partially due to growth inhibition. Ellagic acid rhamnoside (**5.17**) was the most potent biofilm inhibitor of the panel, and the only compound capable of inhibiting 90% of biofilm production (**Table 5.1**). In addition, confocal laser scanning microscopy analysis further confirmed that EA rhamnoside (**5.17**) decreased *S. aureus* biofilm attachment and tower height (**Figure 5.7**).
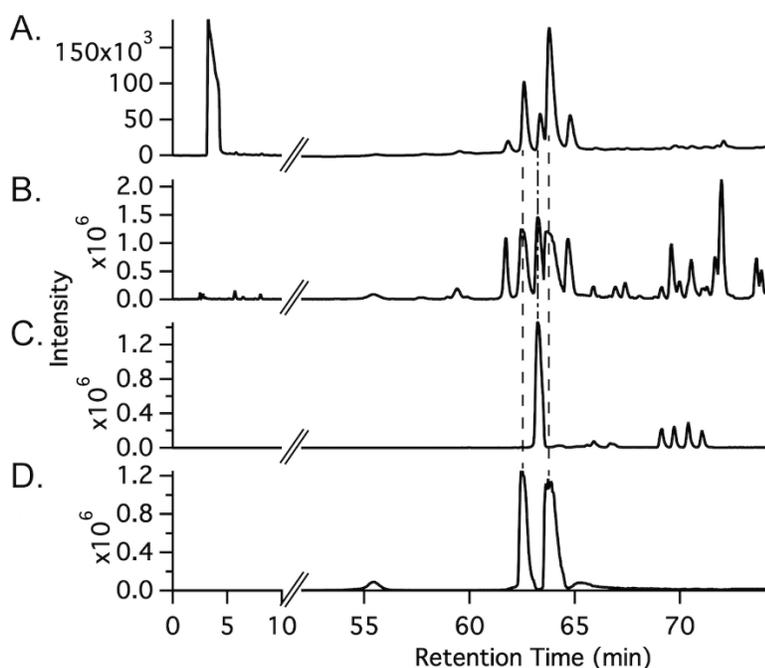
Figure 5.6. Identification of ellagic acid glycosides in the 220D-F2 extract. **A.)** Total UV-visible absorbance chromatogram (210-500 nm) of 220D-F2. **B.)** FT-MS base peak chromatogram of 220D-F2. **C.)** Reconstructed FT-MS chromatogram of ellagic acid rhamnoside (*m/z* window 447-448). **D.)** Reconstructed FT-MS chromatogram of ellagic acid xyloside (*m/z* window 433-434). Comparison of fragmentation patterns from compounds identified in 220D-F2 extract and synthetic standards can be found in Figure S1.

The compounds identified in the extract (EA **[5.13]**, EA rhamnoside **[5.17]**, and EA xyloside **[5.18]**) also were assessed for activity in pair-wise combinations using two-dimensional checkerboard assays, resulting in increased anti-biofilm activity for all three. The increased potency was not due to mere additive effects, as the ΣFIC values for all combinations were <1;

however, no combination was significantly synergistic, as all ΣFIC values were above the synergy cutoff of 0.5 (**Table 5.2**). These results suggest that the 220D-F2 extract elicits increased anti-biofilm activity due to synergy between EA glycosides and other unidentified components of the complex extract.

Table 5.1. Minimum concentrations of extract and synthetic compounds required to inhibit biofilm formation (MBIC) or growth (MIC) of *S. aureus*.

Compound	$MBIC_{50}$ (μg/mL)	$MBIC_{90}$ (μg/mL)	MIC_{50} (μg/mL)	MIC_{90} (μg/mL)
220D-F2	25	100	512	ND
Ellagic Acid (**5.13**)	128	ND	ND	ND
Phenol (**5.1**)	64	ND	ND	ND
Catechol (**5.2**)	ND	ND	ND	ND
Phenyl mannoside (**5.11**)	ND	ND	512	ND
Catechyl mannoside (**5.12**)	ND	ND	ND	ND
Phenyl xyloside (**5.10**)	ND	ND	ND	ND
EA xyloside (**5.18**)	64	ND	32	ND
Phenyl rhamnoside (**5.8**)	64	ND	ND	ND
Catechyl rhamnoside (**5.9**)	128	ND	512	ND
EA rhamnoside (**5.17**)	64	128	ND	ND

ND = not detected; >512 μg/mL

Table 5.2. Fractional inhibitory concentration indices of biofilm inhibition for combinations of compounds identified in extract 220D-F2.

Compound Combination	Lowest concentration with ≤50% biofilm format		
	Ellagic Acid	EA rhamnoside	EA
EA + EA rhamnoside	64	8	
EA + EA xyloside	64	-	
EA rhamnoside + EA xyloside	-	32	

5.3 Discussion

Previous LC-MS/MS analysis of the 220D-F2 extract identified ellagic acid and mono-glycosylated ellagic acid derivatives bearing 5- and 6-deoxypyranose moieties (on the basis of accurate mass and MS/MS spectra).[21d] Therefore, the rhamnoside and xyloside of ellagic acid were synthesized herein, as these deoxypyranoses are found in diverse plant natural products.[38] Analogous glycosides of phenol and catechol were synthesized and included in the screen to assess the structural motifs of the EA core required for biofilm inhibition. Phenol and catechol also were glycosylated with mannose (a non-deoxypyranose) to further interrogate the role of glycosylation on bioactivity, as the sugar component potentially modulates compound solubility and hydrogen bonding. In addition to the glycosides, the corresponding phenolic aglycones were assayed to probe the role of redox chemistry in bioactivity, as the reduction potential of EA lies between the more reducing catechol and the less reducing phenol.[39]
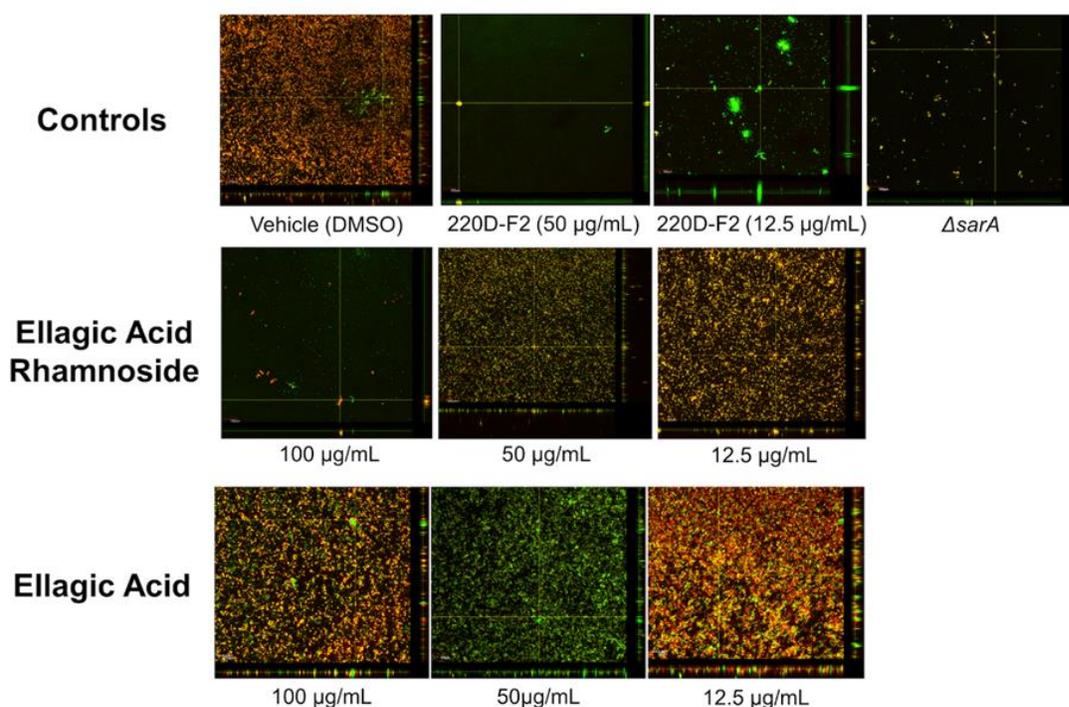
Figure 5.7. Confocal laser scanning microscopy (CLSM) analysis of selected treatments on biofilm production. CLSM experiments were performed with *S. aureus* strain UAMS-1 and its isogenic *sarA* mutant

(UAMS-929) as a control. Bacteria were subjected to various treatments, as indicated in the figure, and biofilm production was analyzed.

A modular synthetic strategy enabled coupling of the appropriate aromatic core to the various carbohydrate components through published glycosylation methodology employing glycosyl per-*O*-acetates[29] or per-*O*-acetyl α-glycosyl iodide donors[32] (for phenol/catechol and ellagic acid, respectively) (**Figures 5.3** and **5.5**). The poor solubility and nucleophilicity of EA necessistated protection as the per-*O*-TBS silyl ether, and subsequent treatment with the mild fluoride source TASF enabled regioselective deprotection of a single silyl group prior to glycosylation (**Figure 5.5**). Notably, these mild de-silylation conditions offer improved regiocontrol compared to prior work which utilized a Bu4NF-mediated deprotection of a per-*O*-silyl EA derivative in the preparation of an EA glucoside.[40] Mild global de-silylation and de-acetylation of the EA glycosides under basic conditions also was optimized to obtain the desired analogs without the need for chromatographic purification (**Figure 5.5**).

The compounds were screened for anti-biofilm and antibiotic activity against a clinical osteomyelitis strain of *S. aureus* (UAMS-1) and an isogenic biofilm-deficient mutant (UAMS-929; Δ*sarA*) as a control. UAMS-1 is commonly utilized in pathogenesis studies due to robust biofilm production and direct clinical relevance. The biofilm inhibition experiments demonstrated that the structure of the carbohydrate component is critical for biological activity, as all rhamnosides (**5.8, 5.9, 5.17**) exhibited anti-biofilm activity. Conversely, the xylosides and mannosides of phenol (**5.10**, **5.11**) and catechol (**5.12**) lacked both anti-biofilm and antibiotic activity, demonstrating that the rhamnosyl moiety is a crucial component of the pharmacophore (Table 1). Furthermore, the rhamnosides (**5.8, 5.9, 5.17**) exhibited increased anti-biofilm potency relative to the corresponding aglycones (**5.1, 5.2, 5.13**), further demonstrating the importance of

the rhamnosyl motif. EA xyloside (**5.18**) also inhibited biofilm formation, but only at concentrations greater than the growth inhibitory concentration (**Table 5.1**). In addition to revealing key features of the sugar component, the screening indicated that biofilm inhibition is not due to redox activity of the phenolic core motifs, as catechol and catechyl glycosides were less potent biofilm inhibitors than the analogous phenol derivatives (**Table 5.1**). Collectively, these data demonstrate that both the phenolic core and the sugar moiety modulate the toxicity and anti-biofilm efficacy against *S. auresus*. Additionally, the sugar component is not simply impacting solubility, as rhamnosyl derivatives elicited increased anti-biofilm potency relative to the xylosides and mannosides (**Table 5.1**).

The rhamnoside and xyloside of EA (**5.17, 5.18**) likely are components of 220D-F2, as the botanical extract contains compounds with retention times, accurate mass spectra, and fragmentation patterns that match those of the synthetic standards (**Figure 5.6**). These data further suggest that EA rhamnoside (**5.17**) is vital to the anti-biofilm activity of 220D-F2 and likely synergizes with other components of the extract.

The effect of EA rhamnoside (**5.17**) on biofilm production further was interrogated by confocal laser scanning microscopy (CLSM) to assess biofilm architecture. Prior studies demonstrated that treatment with 220D-F2 disrupts biofilm morphology and induces the formation of tower-like structures.[21d] Treatment with either EA-rhamnoside (**5.17**) at 100 μg mL^{-1} or 220D-F2 at 50 μg mL^{-1} inhibited biofilm attachment and reduced maximum tower heights to 44 and 48 μm, respectively, in agreement with CLSM analysis of the *sarA* mutant control (**Figure 5.7**).

The present work has identified ellagic acid rhamnoside (**5.17**) as a non-bactericidal inhibitor of *S. aureus* biofilm production which likely contributes to the anti-biofilm activity of

the 220D-F2 botanical extract derived from *Rubus ulmifolius.* Structure-activity studies

elucidated that both the ellagic acid core and the rhamnosyl substituent are critical for biofilm

inhibition. In addition, the development of a synthetic method to prepare EA glycosides will

facilitate future efforts to elucidate the mechanistic details of biofilm inhibition.

5.4 Materials and Methods

220D-F2 Preparation and Analysis

Extract 220D-F2 was prepared from wild harvested samples of *Rubus ulmifolius* and checked for

batch-to-batch reproducibility by HPLC as described previously.[41] LC-FTMS was utilized to

compare the presence of EA glycosides in 220D-F2 using the synthetic EA glycoside standards.

The LC-FTMS analysis was performed using a Thermo Electron LTQ-FTMS equipped with a

Dionex Ultimate 3000 dual gradient pump, Shimadzu autosampler, and diode array detector. The

samples were separated on an Agilent Eclipse XDB-C18 column (4.6 x 250 mm, 5 µm) at

ambient temperature. The 220D-F2 extract was analyzed in 20 µL injections and the EAG

standards were analyzed in 5 µL injection. A linear gradient consisting of 0.1% formic acid in

acetonitrile (A) and 0.1% formic acid in water (B) at a flow rate of 1 mL/min was used for the

chromatographic separations. The initial conditions were 98:2 (A:B) changing to 88:12 (A:B)

over 34 min, this ratio was held until 50 min, and then increased to 75:25 (A:B) at 70 min, then

to 5:95 (A:B) at 82 min and held for 6 min to flush the column before returning to initial

conditions. The HPLC was coupled to a Thermo Scientific LTQ-FT Ultra MS operating in

negative ESI mode. The MS was tuned using ellagic acid and data were acquired in MS^1 mode

scanning from *m/z* of 100–1000 and data-dependent MS^2 mode. Data were processed with

Thermo Scientific Xcalibur 2.2 SP1.48 software (San Jose, CA). The following parameters were

employed: capillary temperature of 275.0°C, N_2 sheath gas of 60, source voltage of 5.0 kV, current of 100.0 μA, and capillary voltage of -41.00 V.

All of the chemically-synthesized phenolic glycosides were subjected to HPLC LC-FTMS analysis using the same conditions described previously (Figures S1-S7). The presence of EA glycosides in 220D-F2 was determined by filtering the LC-FTMS chromatogram of the extract for anions corresponding to the experimentally determined negative ion of the synthetic compounds. The retention time of the resulting peaks was compared to the retention time experimentally determined for the synthesized compounds. Additionally, the MS^1 and MS^2 spectra for the 220D-F2 peaks were compared to that of the synthesized compounds. Peaks from 220D-F2 which had corresponding retention times, parent ions, and MS^2 fragmentation patterns were identified as the indicated compounds.

Growth and Biofilm Inhibition Assays

All test compounds and the 220D-F2 botanical extract control were examined for growth inhibitory and biofilm inhibitory activity following established methods. A well-characterized methicillin sensitive (MSSA) osteomyelitis isolate (UAMS-1) and its isogenic biofilm-deficient *sarA* mutant (UAMS-929) were used in these studies. The *sarA* mutant was selected for use as a control due to its reduced capacity to form a biofilm, which renders it more susceptible to antibiotic treatment.[42] For microbiological tests, strains were grown from freezer stock onto tryptic soy agar (TSA) plates, followed by incubation overnight in tryptic soy broth. All cultures were incubated at 37°C. For growth inhibition studies, MIC_{50} and MIC_{90} values (representing the minimum inhibitory concentration for 50 or 90% of the growth control, respectively) were determined following The Clinical Laboratory Standards Institute (CLSI) M100-S23 guidelines for microtiter broth dilution testing.[43] Briefly, overnight cultures were diluted in cation-adjusted

Mueller Hinton broth (CAMHB) to ~5 x 10^5 CFU mL^{-1} (volume-normalized colony-forming units), which was confirmed by plate counts. Controls included the vehicle (DMSO), along with ampicillin (MP Biomedical) and vancomycin (Sigma-Aldrich). Two-fold serial dilutions were performed on a 96-well microtiter plate to achieve a test range of 4 – 512 µg mL^{-1} for test compounds and 0.5 – 64 µg mL^{-1} for antibiotics. Plates were incubated at 37°C for 18 h. The optical density (OD) at 600 nm (OD$_{600}$) was recorded at 0 and 18 h post-inoculation using a Cytation 3 multimode plate reader (Biotek). The percent inhibition of growth was calculated as described [41].

Growth curve experiments also conducted by recording the OD$_{600}$ at 0, 3, 5, 8, 10, 12 and 18 h post-inoculation. In addition, the volume-normalized CFU was measured by serial dilution and plating onto TSA at 18 hours post-treatment; colonies were counted following 22 h of incubation at 37°C .

Inhibition of biofilm formation was assessed in a static 96-well microtiter plate model with human plasma. Briefly, 20% human plasma diluted in carbonate buffer was added to the biofilm media (tryptic soy broth supplemented with 3.0% NaCl [w/v] and 0.5% dextrose [w/v]) to reach a final concentration of 2% human plasma. Following addition of the test compounds and inoculation with UAMS-1, plates were incubated at 37°C for 22 h. Planktonic cells then were gently aspirated, and the wells were rinsed twice with phosphate-buffered saline to remove non-adherent cells. Adherent biofilms were fixed with 200 µl of ethanol prior to staining for 15 min with 50 µl of 2% (w/v) crystal violet in 20% ethanol (Hardy Diagnostics). The wells were washed with tap water, dried, treated with 100 µL of 2.5% Tween 80 in ethanol, and incubated for 15 min. The solution (20 µL) was transferred to a new plate containing 180 µL PBS/well and the OD$_{595}$ measured by plate reader. The minimum biofilm-inhibiting concentration (MBIC) was

defined as the lowest concentration of extract in which biofilm formation was limited to a level

$\geq$90% (for $MBIC_{90}$) or $\geq$50% (for $MBIC_{50}$) by comparison to the vehicle-treated UAMS-1 strain.

Follow-up assays using the two-dimensional checkerboard method[44] were conducted to

determine any synergistic activities of the compounds found in extract 220D-F2 (ellagic acid,

ellagic acid rhamnoside, and ellagic acid xyloside). The fractional inhibitory concentration (FIC)

index for each compound was calculated as concentration of drug required to inhibit $\geq$50%

biofilm formation in presence of co-drug (measured by OD_{595} of crystal violet eluate), divided by

the $MBIC_{50}$ for that drug. The FIC index is the sum of the two FICs and interactions with a FIC

index ≤ 0.5 were considered synergistic and FIC of 1 additive. The equations are expressed

below.

$FIC_{drug\ 1}$ = [drug 1 in presence of drug 2] $(1/MBIC_{50\ drug\ 1})$

$FIC_{drug\ 2}$ = [drug 2 in presence of drug 1] $(1/MBIC_{50\ drug\ 2})$

FIC index = $\Sigma(FIC_{drug\ 1}, FIC_{drug\ 2})$

Microscopy

Parallel to the above described biofilm assay, biofilm architecture was assessed by confocal laser

scanning microscopy (CLSM). Briefly, biofilms were formed as described above (including

vehicle-treated UAMS-1 and $\Delta sarA$ as controls). After 20 h, the well contents were aspirated and

gently washed three times with 0.85% (w/v) NaCl. The adherent biofilm then was stained with

LIVE/DEAD stain (Invitrogen) at room temperature in the dark for 15 minutes, following the

manufacturer's protocol. Then CLSM images were collected using an Olympus FluoView 1000

confocal scanning system and TIRF inverted microscope. SYTO 9 fluorescence was detected by

excitation/emission at 488/527 nm using a Fluo-3 bandpass filter. Propidium iodide fluorescence

was detected by excitation/emission at 543/612 nm using a Texas Red bandpass filter. All *z*-sections were collected at 4-µm intervals using a 10× objective lens. A 1.27×1.27 mm section of biofilm was selected from the center of the well for each image. Image acquisition and processing was performed using Olympus Fluoview software. Identical acquisition settings were employed for all samples.

Statistical Analysis

All assays performed were analyzed using a two-tailed Student's *t*-test with unequal variance as calculated by Microsoft Excel 2010. DMSO treated (vehicle control) cultures were used as a vehicle control and were compared to those treated with extract for all statistical analyses. *P*-values < 0.05 were considered statistically significant. All assays and other experiments were performed in triplicate or quadruplicate.

Chemical Synthesis

All reactions were performed using dry solvents in flame-dried glassware under a nitrogen atmosphere with magnetic stirring, unless otherwise noted. Reaction solvents were dried over 4 Å molecular sieves, according to a published procedure.[45] Ellagic acid was purchased from Sigma-Aldrich and stored in a vacuum desiccator over Drierite[TM] and phosphorous pentoxide. Phenol and catechol were obtained from Spectrum Chemical (New Brunswick, NJ) and TCI America (Portland, OR). L-rhamnose was purchased from Sigma-Aldrich; D-mannose and D-xylose were obtained from Chem-Impex International Inc. (Wood Dale, IL). All commercially sourced chemicals and reagents were used as received. TLC was performed on aluminum-backed plates coated with silica gel 60 F_{254} and visualized under UV light or by staining with basic $KMnO_4$ solution. Silica gel 60 (40-63 µm) was used for flash column chromatography. [1]H and

^{13}C NMR spectra were recorded on UNITY Plus 600, INOVA 400, and Mercury 300 spectrometers. Chemical shifts are reported in ppm and referenced to the residual solvent signal (for ^{1}H NMR: CDCl$_3$ = 7.26 ppm, CD$_3$OD = 3.31 ppm, DMSO-d6 = 2.50 ppm; for ^{13}C NMR: CDCl$_3$ = 77.16 ppm, CD$_3$OD = 49.00 ppm, DMSO-d6 = 39.52 ppm). High-resolution mass spectra were obtained using a Thermo LTQ FT-MS. NMR spectra are available in the **Supplementary Material** (**Section 5.6**).

1,2,3,4-tetra-*O*-acetyl *α*-L-rhamnopyranoside (10:1 *α:β*)

The compound was synthesized from L-rhamnose according to a published procedure and the spectra match previously published data[29a] (8.60 g in 85% yield).

1,2,3,4,5-penta-*O*-acetyl *α*-D-mannopyranoside (6:1 *α:β*)

The compound was synthesized from D-mannose according to a published procedure and the spectra match previously published data[29c] (1.75 g in 84% yield).

1,2,3,4-tetra-*O*-acetyl *β*-D-xylopyranoside (4:1 *β:α*)

The compound was synthesized from D-xylose according to a published procedure and the spectra match previously published data[29b] (5.53 g in 87% yield).

General glycosylation method:[30]

The appropriate phenol (1 eq) was dissolved in CH$_2$Cl$_2$ containing 4 Å MS. The per-*O*-acetyl glycosyl donor (1 eq) then was added as a solution in CH$_2$Cl$_2$, followed by dropwise addition of BF$_3$-OEt$_2$ (1 eq). The reaction was stirred at 25°C and monitored by silica gel TLC. The reaction was quenched with saturated aqueous NaHCO$_3$ and extracted several times with CHCl$_3$. The combined organic fractions were washed with water and brine, dried over Na$_2$SO$_4$, filtered, and

concentrated *in vacuo*. The residue was purified by silica gel flash column chromatography and all solvent was removed *in vacuo*.

General de-*O*-acetylation method[31]:

The per-*O*-acetyl glycoside (1 eq) was dissolved in MeOH and Et$_3$N (1 eq) was added. The solution was stirred at 25°C and monitored by silica gel TLC. Upon completion of the reaction, the volatiles were removed by azeotropic distillation with toluene *in vacuo*.

1-*O*-Phenyl 2,3,4-tri-*O*-acetyl *α*-L-rhamnopyranoside (5.3)

The compound was synthesized according to the **general glycosylation method.** The residue was purified by silica gel flash column chromatography (325:1 CHCl$_3$/MeOH) to afford 136 mg as white needles in 24% yield. **^{1}H NMR** (400 MHz, CDCl$_3$) δ 7.34 − 7.28 (m, 2H), 7.10 − 7.07 (m, 2H), 7.06 − 7.02 (m, 1H), 5.53 (dd, *J* = 10.1, 3.5 Hz, 1H), 5.47 (d, *J* = 1.8 Hz, 1H), 5.44 (dd, *J* = 3.5, 1.9 Hz, 1H), 5.16 (t, *J* = 10.0 Hz, 1H), 4.01 (dq, *J* = 9.9, 6.2 Hz, 1H), 2.20 (s, 3H), 2.07 (s, 3H), 2.04 (s, 3H), 1.21 (d, *J* = 6.3 Hz, 3H). **^{13}C NMR** (100 MHz, CDCl$_3$) δ 170.17, 170.13, 170.11, 156.04, 129.73, 122.85, 116.53, 95.86, 71.18, 69.92, 69.10, 67.26, 21.02, 20.92, 20.87, 17.58. **HRMS** (ESI$^+$) *m/z*: [M + Na]$^+$ Calcd for C$_{18}$H$_{22}$O$_8$Na$^+$ 389.1207; Found 389.1201.

Phenyl *α*-L-rhamnopyranoside (5.8)

The compound was synthesized according to the **general de-*O*-acetylation method** to afford the desired product as a colorless oil (74 mg, 99% yield). **^{1}H NMR** (400 MHz, CD$_3$OD) δ 7.31 − 7.25 (m, 2H), 7.08 − 7.03 (m, 2H), 7.02 − 6.97 (m, 1H), 5.42 (d, *J* = 1.8 Hz, 1H), 4.00 (dd, *J* = 3.4, 1.8 Hz, 1H), 3.85 (dd, *J* = 9.5, 3.5 Hz, 1H), 3.69 − 3.60 (m, 1H), 3.46 (t, *J* = 9.5 Hz, 1H), 1.22 (d, *J* = 6.2 Hz, 4H). **^{13}C NMR** (100 MHz, CD$_3$OD) δ 157.84, 130.48, 123.17, 117.51, 99.83,

73.85, 72.25, 72.09, 70.57, 18.01. **HRMS** (ESI$^+$) *m/z*: [M + Na]$^+$ Calcd for $C_{12}H_{16}O_5Na^+$

263.0890; Found 263.0885.

1-*O*-(*o*-Hydroxy)phenyl 2,3,4-tri-*O*-acetyl *α*-L-rhamnopyranoside (5.4)

The compound was synthesized according to the **general glycosylation method.** The residue

was purified by silica gel flash column chromatography (60:1 CHCl$_3$/MeOH) to afford 51 mg as

white needles in 10% yield. ^{1}**H NMR** (400 MHz, CDCl$_3$) δ 7.15 (d, *J* = 8.1 Hz, 1H), 6.97 (d, *J* =

4.0 Hz, 2H), 6.82 (dt, *J* = 8.3, 4.5 Hz, 1H), 6.12 (d, *J* = 4.8 Hz, 1H), 5.51 (dd, *J* = 3.5, 1.9 Hz,

1H), 5.47 (dd, *J* = 9.9, 3.5 Hz, 1H), 5.44 (d, *J* = 1.8 Hz, 1H), 5.18 (t, *J* = 9.9 Hz, 1H), 4.06 (dq, *J*

= 9.8, 6.2 Hz, 1H), 2.19 (s, 4H), 2.08 (s, 3H), 2.04 (s, 3H), 1.25 (d, *J* = 6.3 Hz, 3H). ^{13}**C NMR**

(100 MHz, CDCl$_3$) δ 170.43, 170.15, 170.07, 146.10, 143.61, 123.98, 120.54, 116.27, 115.64,

97.14, 70.80, 69.58, 69.16, 67.66, 20.94, 20.88, 20.83, 17.54. **HRMS** (ESI$^+$) *m/z*: [M + Na]$^+$

Calcd for $C_{18}H_{22}O_9Na^+$ 405.1156; Found 405.1164.

o-Hydroxyphenyl *α*-L-rhamnopyranoside (5.9)

The compound was synthesized according to the **general de-*O*-acetylation method** to afford the

desired product as a white solid (34 mg, 99% yield). The spectra match previously published

data.[46]

1-*O*-Phenyl 2,3,4,5-tetra-*O*-acetyl *α*-D-mannopyranoside (5.6)

The compound was synthesized according to the **general glycosylation method.** The residue

was purified by silica gel flash column chromatography (3:2 hexanes/EtOAc) to afford 260 mg

as a white solid in 33% yield. ^{1}**H NMR** (300 MHz, CDCl$_3$) δ 7.36 − 7.28 (m, 2H), 7.13 − 7.03

(m, 3H), 5.58 (dd, *J* = 10.0, 3.5 Hz, 1H), 5.54 (d, *J* = 1.8 Hz, 1H), 5.46 (dd, *J* = 3.5, 1.9 Hz, 1H),

5.38 (t, *J* = 10.1 Hz, 1H), 4.34 − 4.25 (m, 1H), 4.15 − 4.04 (m, 2H), 2.21 (s, 3H), 2.06 (s, 3H),

2.05 (s, 3H), 2.04 (s, 3H). **¹³C NMR** (100 MHz, CDCl₃) δ 170.65, 170.10, 170.05, 169.88, 155.77, 129.77, 123.18, 116.69, 95.99, 69.64, 69.31, 69.09, 66.21, 62.32, 21.01, 20.83, 20.78. **HRMS** (ESI⁺) *m/z*: [M + Na]⁺ Calcd for $C_{20}H_{24}O_{10}Na^+$ 447.1262; Found 447.1265.

Phenyl *α*-D-mannopyranoside (5.11)

The compound was synthesized according to the **general de-*O*-acetylation method** to afford the desired product as a colorless oil (25 mg, 54% yield). The spectra match previously published data.[47]

***o*-Hydroxyphenyl *α*-D-mannopyranoside (5.12)**

The compound was synthesized by glycosylation of catechol with **per-*O*-acetyl α-D-mannopyranoside** according to the **general glycosylation method.** The crude protected glycoside was then deprotected without prior purification according to the **general de-*O*-acetylation method.** The product was isolated *via* silica gel flash column chromatography (10:1 EtOAc/MeOH, 1% AcOH) in 43% yield (17 mg). **¹H NMR** (600 MHz, CD₃OD) δ 7.19 (dd, *J* = 8.1, 1.4 Hz, 1H), 6.90 – 6.85 (m, 1H), 6.83 (dd, *J* = 8.0, 1.7 Hz, 1H), 6.78 – 6.73 (m, 1H), 5.40 (d, *J* = 1.8 Hz, 1H), 4.12 (dd, *J* = 3.4, 1.8 Hz, 1H), 3.99 (dd, *J* = 8.7, 3.3 Hz, 1H), 3.79 – 3.73 (m, 4H). **¹³C NMR** (150 MHz, CD₃OD) δ 148.78, 145.89, 124.43, 120.90, 119.42, 117.29, 101.55, 75.40, 72.33, 71.92, 68.42, 62.67. **HRMS** (ESI⁺) *m/z*: [M + Na]⁺ Calcd for $C_{12}H_{16}O_7Na^+$ 295.0788; Found 295.0788.

Phenyl *β*-D-xylopyranoside (5.10)

The compound was synthesized by glycosylation of phenol with **per-*O*-acetyl β-D-xylopyranoside** according to the **general glycosylation method.** The crude protected glycoside was then deprotected without prior purification according to the **general de-*O*-acetylation**

method. The product was isolated as a white solid *via* silica gel flash column chromatography

(15:1 EtOAc/MeOH, 1% Et$_3$N) in 59% yield (252 mg). The spectra match previously published

data.[48]

2,3,4-tri-*O*-acetyl *α*-L-rhamnopyranosyl iodide

The compound was synthesized by iodination of **1,2,3,4-tetra-*O*-acetyl α-L-rhamnopyranoside**

according to a published procedure and the spectra match previously published data[33] (1.95 g,

81% yield).

2,3,4-tri-*O*-acetyl *α*-D-xylopyranosyl iodide

The compound was synthesized by iodination of **1,2,3,4-tetra-*O*-acetyl *β*-D-xylopyranoside**

according to a published procedure and the spectra match previously published data[33] (4.02 g,

83% yield).

3,3',4,4'-Tetrakis-*O*-*tert*-butyldimethylsilyl ellagic acid (5.14)

The compound was synthesized in analogy to a published procedure with some modifications [34].

Ellagic acid (2 g, 6.62 mmol), 4-(dimethylamino)pyridine (18.5 mg, 0.199 mmol), and imidazole

(2.25 g, 33.1 mmol) were suspended in CH$_2$Cl$_2$/DMF (30 mL/10 mL). A solution of TBDMS-Cl

(5 g, 33.1 mmol) in CH$_2$Cl$_2$ was added and the mixture was stirred at 50°C in the dark for 48 h.

The reaction was cooled to 25°C and quenched with saturated aqueous NH$_4$Cl. The aqueous

phase was extracted with CH$_2$Cl$_2$ and the combined organic phases were dried over Na$_2$SO$_4$,

filtered, and concentrated *in vacuo*. TBDMS-OH was removed by azeotropic distillation with

toluene *in vacuo*. The residue then was adsorbed onto silica gel and purified by silica gel flash

column chromatography (30:1 hexanes/EtOAc) to obtain the desired compound as a pale yellow

solid (2.64 g, 70% yield). **^{1}H NMR** (400 MHz, CDCl$_3$) δ 7.64 (s, 2H), 1.09 (s, 18H), 1.02 (s,

18H), 0.34 (s, 12H), 0.32 (s, 12H). **^{13}C NMR** (100 MHz, CDCl$_3$) δ 159.09, 150.15, 141.21, 140.68, 116.46, 113.69, 110.47, 77.16, 26.15, 26.00, 18.94, 18.86, -3.49, -3.65. **HRMS** (APCI$^+$) *m/z*: [M + H]$^+$ Calcd for C$_{38}$H$_{63}$O$_8$Si$_4^+$ 759.3595; Found 759.3594.

3-*O*-(2",3",4"-Tri-*O*-acetyl *α*-L-rhamnopyranosyl) ellagic acid (5.15)

The synthesis was based on published glycosylation methodology[32b]. **Per-*O*-TBDMS ellagic acid** (930 mg, 1.23 mmol) was dissolved in CH$_2$Cl$_2$ (15 mL) containing 4 Å MS. The solution was stirred at 25°C for 1 h in the dark. A solution of tris(dimethylamino)sulfonium diflurotrimethylsilicate (TASF) (371 mg, 1.35 mmol) in CH$_2$Cl$_2$ was added dropwise and the mixture was stirred at 25°C under a dry nitrogen purge for 10 min (to remove gaseous Me$_3$Si-F). The deprotection was monitored by silica gel TLC. **2,3,4-Tri-*O*-acetyl α-L-rhamnopyranosyl iodide** (1.9 g, 4.9 mmol) was subsequently added as a solution in CH$_2$Cl$_2$ and the temperature was gradually increased to 45°C. After 48 h the reaction was cooled to 25°C and filtered to remove 4 Å and insoluble material. The filtrate was concentrated *in vacuo* and the amber residue was dissolved in CH$_2$Cl$_2$, adsorbed onto Celite (pre-washed with MeOH), and purified by silica gel flash column chromatography (3:1 hexanes/EtOAc). The desired product was obtained as a pale yellow, glassy solid (145 mg, 13% yield). **^{1}H NMR** (400 MHz, CDCl$_3$) δ 7.65 (s, 2H), 5.75 (dd, *J* = 3.3, 1.8 Hz, 1H), 5.67 (d, *J* = 1.5 Hz, 1H), 5.60 (dd, *J* = 10.2, 3.4 Hz, 1H), 5.19 (t, *J* = 10.1 Hz, 1H), 4.78 (dq, *J* = 12.5, 6.3 Hz, 1H), 2.19 (s, 3H), 2.11 (s, 3H), 2.03 (s, 3H), 1.21 (d, *J* = 6.2 Hz, 3H), 1.09 (s, 9H), 1.02 (s, 9H), 1.01 (s, 9H), 0.35 (s, 3H), 0.35 (s, 3H), 0.34 (s, 6H), 0.32 (s, 6H). **^{13}C NMR** (100 MHz, CDCl$_3$) δ 170.24, 170.06, 169.78, 158.67, 158.56, 151.02, 150.50, 143.40, 141.45, 140.60, 139.60, 116.78, 116.54, 114.44, 113.72, 113.33, 110.30, 99.77, 70.56, 69.31, 68.81, 26.12, 25.98, 25.85, 20.98, 20.96, 20.84, 18.94, 18.86, 18.62, 17.33, -3.49, -3.64, -

3.99, -4.05. **HRMS** (ESI$^+$) *m/z*: [M + Na]$^+$ Calcd for $C_{44}H_{64}O_{15}Si_3Na^+$ 939.3445; Found 939.3447.

3-*O*-(2",3",4"-Tri-*O*-acetyl *β*-D-xylopyranosyl) ellagic acid (5.16)

The synthesis was based on published glycosylation methodology[32b]. **Per-*O*-TBDMS ellagic acid** (1.16 g, 1.53 mmol) was dissolved in CH_2Cl_2 (15 mL) containing 4 Å MS. The solution was stirred at 25°C for 1 h in the dark. A solution of TASF (463 mg, 1.68 mmol) in CH_2Cl_2 was added dropwise and the mixture was stirred at 25°C under a dry nitrogen purge for 10 min (to remove gaseous Me_3Si-F). The deprotection was monitored by silica gel TLC. Bu_4NI (565 mg, 1.53 mmol) was subsequently added to the reaction as a solution in CH_2Cl_2, followed by addition of **2,3,4-tri-*O*-acetyl *β*-D-xylopyranosyl iodide** (3.8 g, 6.4 mmol) as a solution in CH_2Cl_2. The temperature was gradually increased to 45°C. After 48 h the reaction was cooled to 25°C and concentrated in vacuo. The residue was then cooled in an ice bath and ice-cold EtOAc was added to precipitate Bu_4NI, which was removed by filtration. The filtrate was concentrated *in vacuo* and the amber residue was dissolved in CH_2Cl_2, adsorbed onto Celite (pre-washed with MeOH), and purified by silica gel flash column chromatography (3:1 hexanes/EtOAc). The desired product was obtained as a pale yellow, glassy solid (209 mg, 15% yield). **^{1}H NMR** (400 MHz, CDCl$_3$) δ 7.62 (s, 1H), 7.60 (s, 1H), 5.74 (d, *J* = 5.0 Hz, 1H), 5.30 (dd, *J* = 7.2, 5.0 Hz, 1H), 5.22 (t, *J* = 6.9 Hz, 1H), 5.04 – 4.98 (m, 1H), 4.41 (dd, *J* = 12.5, 4.2 Hz, 1H), 3.53 (dd, *J* = 12.4, 6.0 Hz, 1H), 2.13 (s, 3H), 2.13 (s, 3H), 2.05 (s, 3H), 1.06 (s, 9H), 1.00 (s, 9H), 0.99 (s, 9H), 0.31 (s, 6H), 0.29 (s, 6H), 0.28 (s, 3H), 0.26 (s, 3H). **^{13}C NMR** (100 MHz, CDCl$_3$) δ 170.03, 169.95, 169.41, 158.65, 151.04, 150.43, 142.63, 141.44, 140.57, 139.39, 116.75, 114.10, 113.69, 113.38, 110.26, 100.03, 69.96, 69.91, 68.57, 62.10, 26.10, 25.95, 25.79, 20.93, 20.91, 20.87,

18.91, 18.84, 18.55, -3.52, -3.66, -4.11, -4.20. **HRMS** (ESI$^+$) *m/z*: [M + Na]$^+$ Calcd for

$C_{43}H_{62}O_{15}Si_3Na^+$ 925.3289; Found 925.3281.

General ellagic acid glycoside deprotection method:

Deprotection of TBDMS silyl ethers was perfomed in analogy to a published procedure.[36] The

appropriate ellagic acid glycoside (1 eq) was dissolved in DMF/H_2O (10:1 v/v) and K_2CO_3 (1.3

eq) was added. The solution was stirred at 25°C in the dark for 5 h. The mixture was diluted with

toluene/MeOH and the pH was adjusted to ~6 with dilute aqueous AcOH. The solvent was

removed *in vacuo* and the residue was subjected to azeotropic distillation with toluene. The

resulting de-silylated glycoside was suspended in MeOH/H_2O (10:1 v/v) and K_2CO_3 (1.5 eq) was

added to effect deprotection of the sugar moiety. The mixture was stirred at 25°C in the dark for

48 h. The reaction then was diluted with MeOH/H_2O and the pH was adjusted to ~4 by gradual

addition of Dowex 50WX8 cation-exchange resin (pre-washed with MeOH/H_2O). The resin was

removed by filtration and washed thoroughly with MeOH/H_2O. The solvent was removed *in*

vacuo and the resulting solid was washed with Et_2O. The resulting material was then dissolved in

10 mM NH_4HCO_3 (pH 7.8) and lyophilized to dryness to afford the desired deprotected ellagic

acid glycosides.

3-*O*-*α*-L-rhamnopyranosyl ellagic acid (5.17)

The compound was prepared according to the **general ellagic acid glycoside deprotection**

method to afford the desired product as a yellow solid (18 mg, 86% yield). **^{1}H NMR** (400 MHz,

DMSO-d6) δ 7.56 (s, 1H), 7.44 (s, 1H), 5.51 (s, 1H), 4.15 (dq, *J* = 12.6, 6.2 Hz, 1H), 4.05 (dd, *J*

= 3.1, 1.7 Hz, 1H), 3.78 (dd, *J* = 9.4, 3.3 Hz, 1H), 3.31 (t, *J* = 9.5 Hz, 1H), 1.07 (d, *J* = 6.2 Hz,

3H). **^{13}C NMR** (100 MHz, DMSO-d6) δ 159.06, 158.71, 152.30, 148.78, 142.40, 136.61, 136.02,

113.26, 112.33, 111.85, 111.10, 109.68, 102.43, 71.40, 70.54, 70.36, 70.13, 17.71. **HRMS** (ESI⁻)

m/z: [M − H+]⁻ Calcd for $C_{20}H_{15}O_{12}^-$ 447.0569; Found 447.0574.

3-*O*-*β*-D-xylopyranosyl ellagic acid (5.18)

The compound was prepared according to the **general ellagic acid glycoside deprotection method** to afford the desired product as a yellow solid (22 mg, 91% yield). **¹H NMR** (400 MHz, DMSO-d6) δ 7.50 (s, 1H), 7.38 (s, 1H), 5.35 (d, *J* = 6.9 Hz, 1H), 3.80 (dd, *J* = 11.4, 5.0 Hz, 1H), 3.47 − 3.39 (m, 2H), 3.27 (t, *J* = 8.3 Hz, 1H), 3.10 (dd, *J* = 11.4, 9.3 Hz, 1H). **¹³C NMR** (100 MHz, DMSO-d6) δ 159.80, 159.20, 151.48, 142.14, 135.78, 135.06, 113.56, 113.06, 112.49, 110.47, 106.64, 102.92, 75.40, 73.16, 69.29, 65.69. **HRMS** (ESI⁻) *m/z*: [M − H+]⁻ Calcd for $C_{19}H_{13}O_{12}^-$ 433.0412; Found 433.0417.

5.5 References

(1) *State of the World's Antibiotics, 2015*. Washington, D.C., 2015.

(2) Serpi, M.; Ferrari, V.; Pertusati, F. *J. Med. Chem.* **2016**, DOI: 10.1021/acs.jmedchem.6b00325.

(3) (a) Infectious Diseases Society of America. *Bad Bugs, No Drugs*. Alexandria, VA, 2004; (b) World Health Organization. *Antimicrobial Resistance: Global Report on Surveillance* WHO: Geneva, Switzerland, 2014; (c) O'Neill, J. *Antimicrobial Resistance: Tackling a Crisis for the Health and Wellness of Nations*. Wellcome Trust and HM Government, London, United Kingdom. **2014**.

(4) Flemming, H.; Wingender, J. *Nat. Rev. Microbiol.* **2010**, *8*, 623-633.

(5) Wu, H.; Moser, C.; Wang, H.; Hoiby, N.; Song, Z. *International Journal of Oral Science* **2015**, *7*, 1-7.

(6) Wood, T.; Knabel, S.; Kwan, B. *Appl. Environ. Microbiol.* **2013**, *79*, 7116-7121.

(7) Parvizi, J.; Pawasarat, I.; Azzam, K.; Joshi, A.; Hansen, E.; Bozic, K. *Journal of Arthroplasty* **2010**, *25*, 103-107.

(8) Neopane, P.; Nepal, H.; Shrestha, R.; Uehara, O.; Abiko, Y. *International Journal of General Medicine* **2017**, *11*, 25-32.

(9) Moormeier, D.; Bayles, K. *Mol. Microbiol.* **2017**, *104*, 365-376.

(10) Goodman, S.; Obergfell, K.; Jurcisek, J.; Novotny, L.; Downey, J.; Ayala, E.; Tjokro, N.; Li, B.; Justice, S.; Bakaletz, L. *Mucosal Immunology* **2011**, *4*, 625-637.

(11) Moormeier, D.; Bose, J.; Horswill, A.; Bayles, K. *Mbio* **2014**, *5*.

(12) Moormeier, D.; Endres, J.; Mann, E.; Sadykov, M.; Horswill, A.; Rice, K.; Fey, P.; Bayles, K. *Appl. Environ. Microbiol.* **2013**, *79*, 3413-3424.

(13) Kavanaugh, J.; Horswill, A. *J. Biol. Chem.* **2016**, *291*, 12556-12564.

(14) Ji, G.; Beavis, R.; Novick, R. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 12055-12059.

(15) Le, K.; Otto, M. *Frontiers in Microbiology* **2015**, *6*.

(16) Koenig, R.; Ray, J.; Maleki, S.; Smeltzer, M.; Hurlburt, B. *J. Bacteriol.* **2004**, *186*, 7549-7555.

(17) Wang, R.; Khan, B.; Cheung, G.; Bach, T.; Jameson-Lee, M.; Kong, K.; Queck, S.; Otto, M. *J. Clin. Invest.* **2011**, *121*, 238-248.

(18) Zimmerli, W.; Waldvogel, F.; Vaudaux, P.; Nydegger, U. *J. Infect. Dis.* **1982**, *146*, 487-497.

(19) Hauser, C.; Wuethrich, B.; Matter, L.; Wilhelm, J.; Sonnabend, W.; Schopper, K. *Dermatologica* **1985**, *170*, 35-39.

(20) Reuter, J.; Merfort, I.; Schempp, C. *American Journal of Clinical Dermatology* **2010**, *11*, 247-267.

(21) (a) Elmasri, W.; Yang, T.; Tran, P.; Hegazy, M.; Hamood, A.; Mechref, Y.; Pare, P. *J. Nat. Prod.* **2015**, *78*, 2-9; (b) Gutierrez-Barranquero, J.; Reen, F.; McCarthy, R.; O'Gara, F. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 3303-3316; (c) Onsare, J.; Arora, D. *J. Appl. Microbiol.* **2015**, *118*, 313-325; (d) Quave, C.; Estevez-Carmona, M.; Compadre, C.; Hobby, G.; Hendrickson, H.; Beenken, K.; Smeltzer, M. *Plos One* **2012**, *7*; (e) Yadav, M.; Chae, S.; Im, G.; Chung, J.; Song, J. *Plos One* **2015**, *10*; (f) Zheng, W.; Liang, Y.; Zhao, H.; Zhang, J.; Li, Z. *ChemBioChem* **2015**, *16*, 1035-1040; (g) Zhou, X.; Liu, S.; Li, W.; Zhang, B.; Liu, B.; Liu, Y.; Deng, X.; Peng, L. *World J. Microbiol. Biotechnol.* **2015**, *31*, 1259-1265; (h) Tsou, L.; Lara-Tejero, M.; RoseFigura, J.; Zhang, Z.; Wang, Y.; Yount, J.; Lefebre, M.; Dossa, P.; Kato, J.; Guan, F.; Lam, W.; Cheng, Y.; Galan, J.; Hang, H. *J. Am. Chem. Soc.* **2016**, *138*, 2209-2218.

(22) Scafuri, B.; Marabotti, A.; Carbone, V.; Minasi, P.; Dotolo, S.; Facchiano, A. *Scientific Reports* **2016**, *6*.

(23) Pisithkul, T.; Jacobson, T.; O'Brien, T.; Stevenson, D.; Amador-Noguez, D. *Appl. Environ. Microbiol.* **2015**, *81*, 5761-5772.

(24) Hassan, S.; Mathesius, U. *J. Exp. Bot.* **2012**, *63*, 3429-3444.

(25) Serra, D.; Mika, F.; Richter, A.; Hengge, R. *Mol. Microbiol.* **2016**, *101*, 136-151.

(26) Larsen, P.; Nielsen, J.; Dueholm, M.; Wetzel, R.; Otzen, D.; Nielsen, P. *Environ. Microbiol.* **2007**, *9*, 3077-3090.

(27) Blanco, A.; Sudano-Roccaro, A.; Spoto, G.; Nostro, A.; Rusciano, D. *Antimicrob. Agents Chemother.* **2005**, *49*, 4339-4343.

(28) Quave, C.; Pieroni, A.; Bennett, B. *Journal of Ethnobiology and Ethnomedicine* **2008**, *4*.

(29) (a) Donahue, M.; Johnston, J. *Bioorg Med Chem Lett* **2006**, *16*, 5602-5604; (b) Camponovo, J.; Hadad, C.; Ruiz, J.; Cloutet, E.; Gatard, S.; Muzart, J.; Bouquillon, S.; Astruc, D. *J Org Chem* **2009**, *74*, 5071-5074; (c) Yu, K.; Kizhakkedathu, J. *Biomacromolecules* **2010**, *11*, 3073-3085.

(30) Jacobsson, M.; Malmberg, J.; Ellervik, U. *Carbohydrate Res* **2006**, *341*, 1266-1281.

(31) Parker, M.; Osidacz, P.; Baldock, G.; Hayasaka, Y.; Black, C.; Pardon, K.; Jeffery, D.; Geue, J.; Herderich, M.; Francis, I. *J Ag Food Chem* **2012**, *60*, 2629-2637.

(32) (a) Bhat, A.; Gervay-Hague, J. *Org Lett* **2001**, *3*, 2081-2084; (b) Du, W.; Gervay-Hague, J. *Org Lett* **2005**, *7*, 2063-2065; (c) Gervay, J.; Hadd, M. J. *J Org Chem* **1997**, *62*, 6961-6967; (d) Gervay, J.; Nguyen, T.; Hadd, M. *Carbohydrate Res* **1997**, *300*, 119-125; (e) Hadd, M. J.; Gervay, J. *Carbohydrate Res* **1999**, *320*, 61-69; (f) Schombs, M.; Park, F.; Du, W.; Kulkarni, S.; Gervay-Hague, J. *J Org Chem* **2010**, *75*, 4891-4898.

(33) Mukhopadhyay, B.; Kartha, K.; Russell, D.; Field, R. *J Org Chem* **2004**, *69*, 7758-7760.

(34) Kobayashi, R.; Hanaya, K.; Shoji, M.; Ohba, S.; Sugai, T. *Biosci Biotechnol Biochem* **2013**, *77*, 810-813.

(35) Scheidt, K.; Chen, H.; Follows, B.; Chemler, S.; Coffey, D.; Roush, W. *J. Org. Chem.* **1998**, *63*, 6436-6437.

(36) Jiang, Z.; Wang, Y. *Tetrahedron Lett* **2003**, *44*, 3859-3861.

(37) Pettit, G.; Mendonca, R.; Knight, J.; Pettit, R. *J Nat Prod* **2011**, *74*, 1922-1930.

(38) Cowan, M. *Clin. Microbiol. Rev.* **1999**, *12*, 564-+.

(39) Steenken, S.; Neta, P. *J. Phys. Chem.* **1982**, *86*, 3661-3667.

(40) Kobayashi, R.; Hanaya, K.; Shoji, M.; Ohba, S.; Sugai, T. *Biosci., Biotechnol., Biochem.* **2013**, *77*, 810-813.

(41) Quave, C. L.; Estevez-Carmona, M.; Compadre, C. M.; Hobby, G.; Hendrickson, H. P.; Beenken, K. E.; Smeltzer, M. S. *PLoS ONE* **2012**, *7*, e28737.

(42) Atwood, D. N.; Beenken, K. E.; Lantz, T. L.; Meeker, D. G.; Lynn, W. B.; Mills, W. B.; Spencer, H. J.; Smeltzer, M. S. *Antimicrob Agents Chemother* **2016**, *60*, 1826-1829.

(43) NCCLS *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria that Grow Aerobically; Approved Standard*; 5th ed.; National Center for Clinical Laboratory Standards: Wayne, PA, 2001; Vol. 20.

(44) Farha, M. A.; Leung, A.; Sewell, E. W.; D'Elia, M. A.; Allison, S. E.; Ejim, L.; Pereira, P. M.; Pinho, M. G.; Wright, G. D.; Brown, E. D. *ACS Chem Biol* **2013**, *8*, 226-233.

(45) Bradley, D.; Williams, G.; Lawton, M. *J. Org. Chem.* **2010**, *75*, 8351-8354.

(46) De Winter, K.; Simcikova, D.; Schalck, B.; Weignerova, L.; Pelantova, H.; Soetaert, W.; Desmet, T.; Kren, V. *Bioresource Technol* **2013**, *147*, 640-644.

(47) Klein, T.; Abgottspon, D.; Wittwer, M.; Rabbani, S.; Herold, J.; Jiang, X.; Kleeb, S.; Luthi, C.; Scharenberg, M.; Bezencon, J.; Gubler, E.; Pang, L.; Smiesko, M.; Cutting, B.; Schwardt, O.; Ernst, B. *J Med Chem* **2010**, *53*, 8627-8641.

(48) Siegbahn, A.; Thorsheim, K.; Stahle, J.; Manner, S.; Hamark, C.; Persson, A.; Tykesson, E.; Mani, K.; Westergren-Thorsson, G.; Widmalm, G.; Ellervik, U. *Org Biomol Chem* **2015**, *13*, 3351-3362.

5.6 Supplementary Material
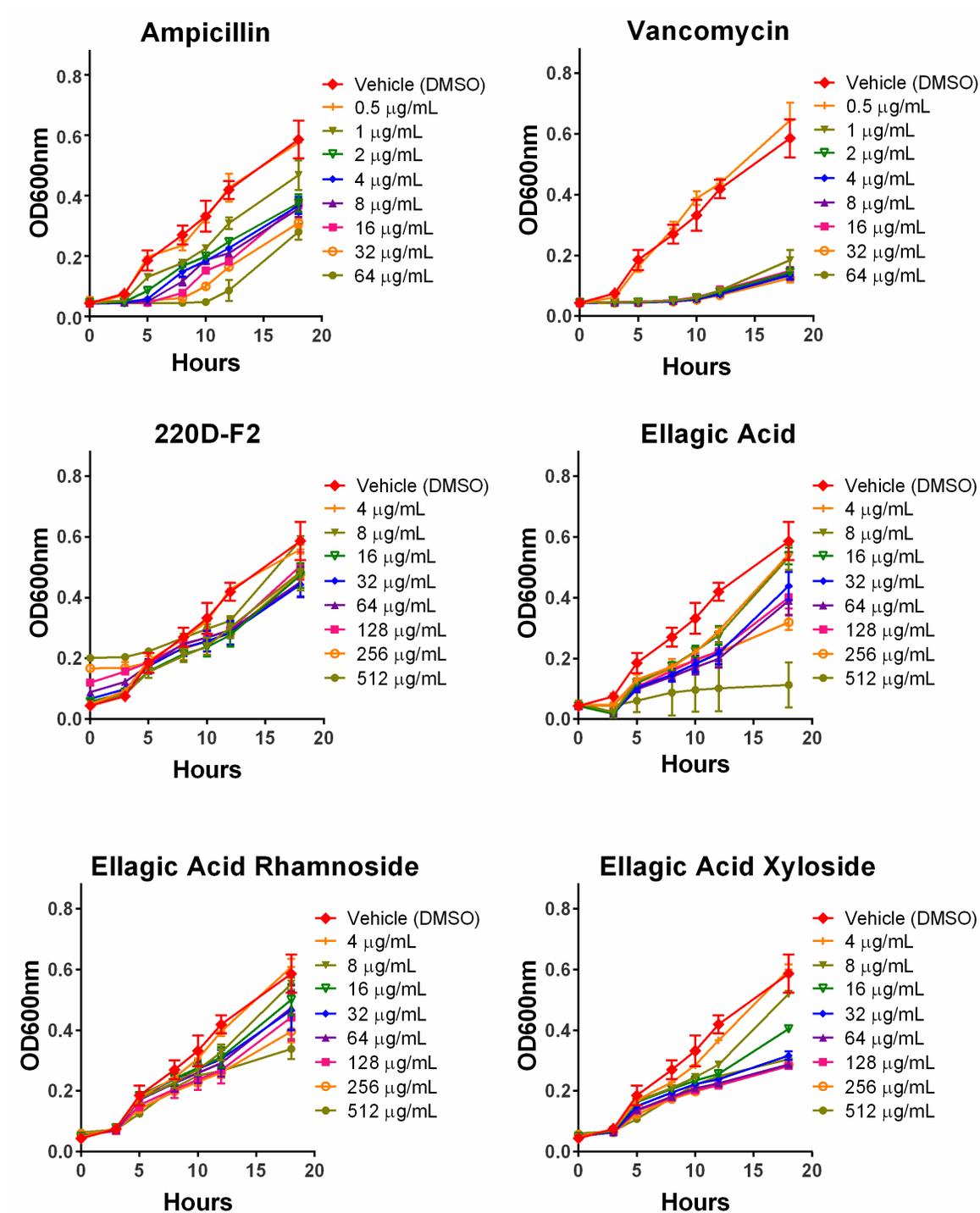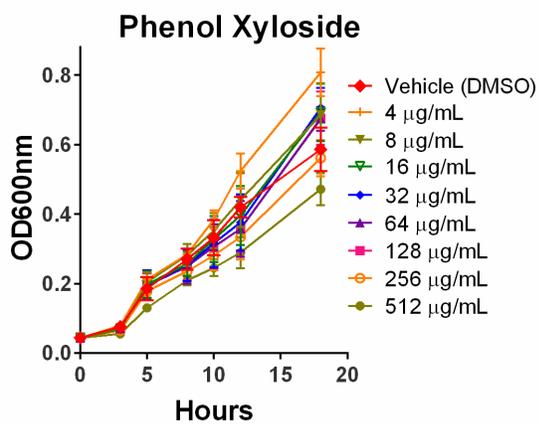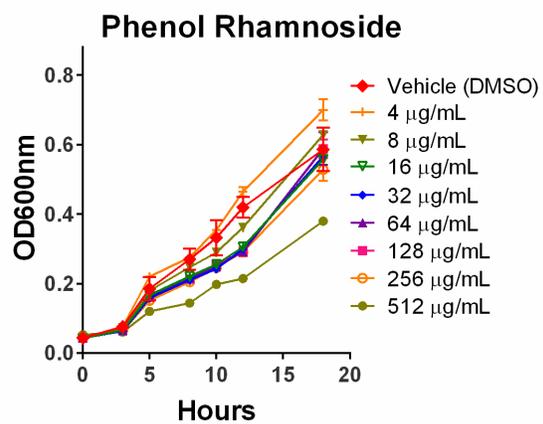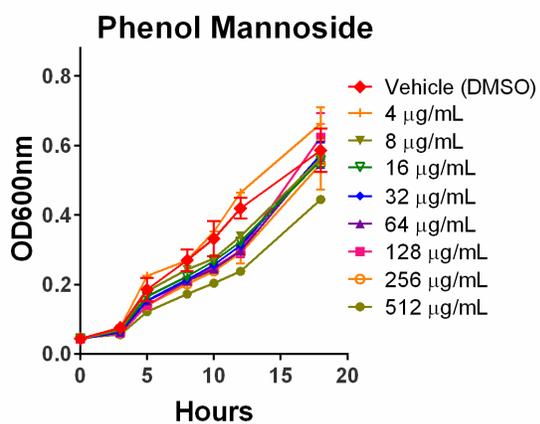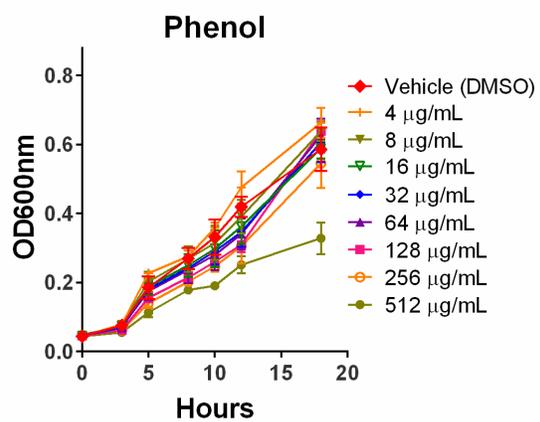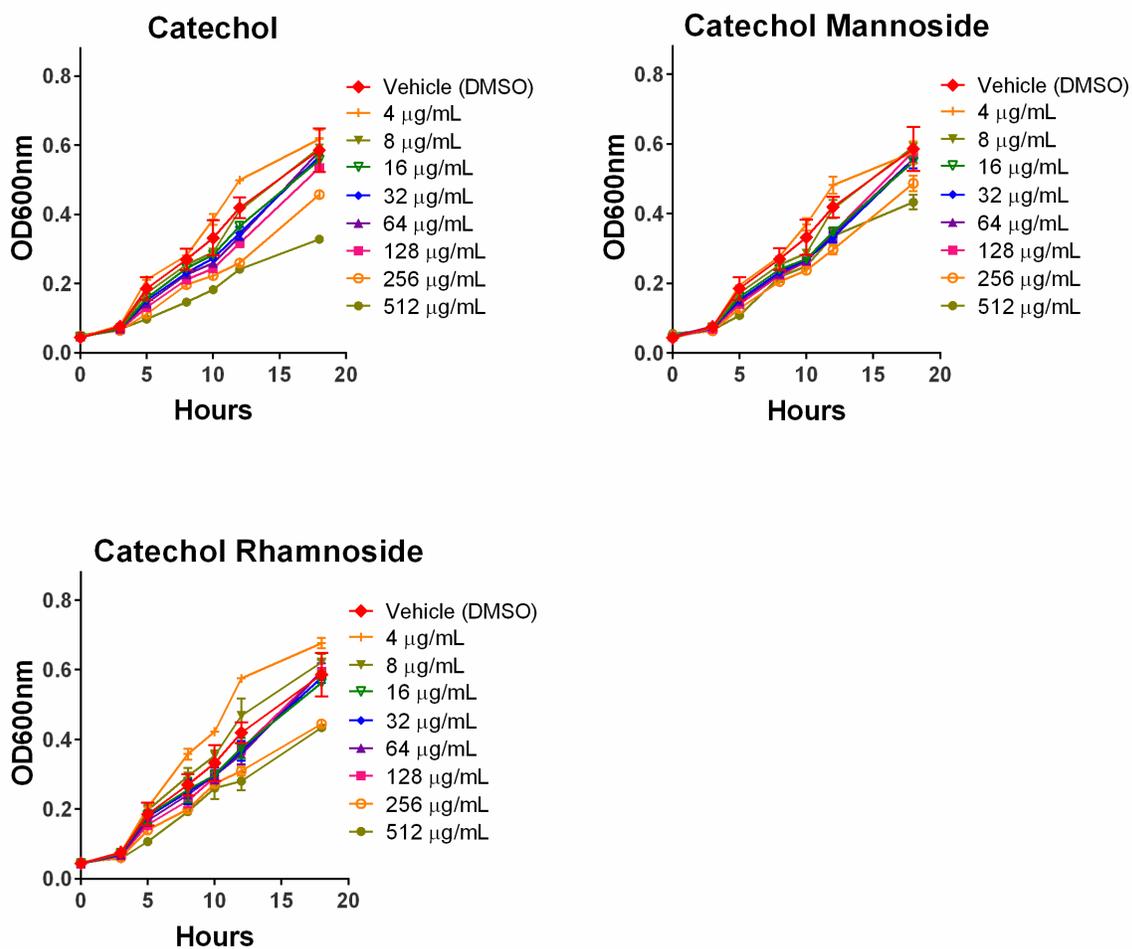
Figure S5.1. Growth curves for antibiotic controls, 220D-F2 extract, and synthetic analogs.
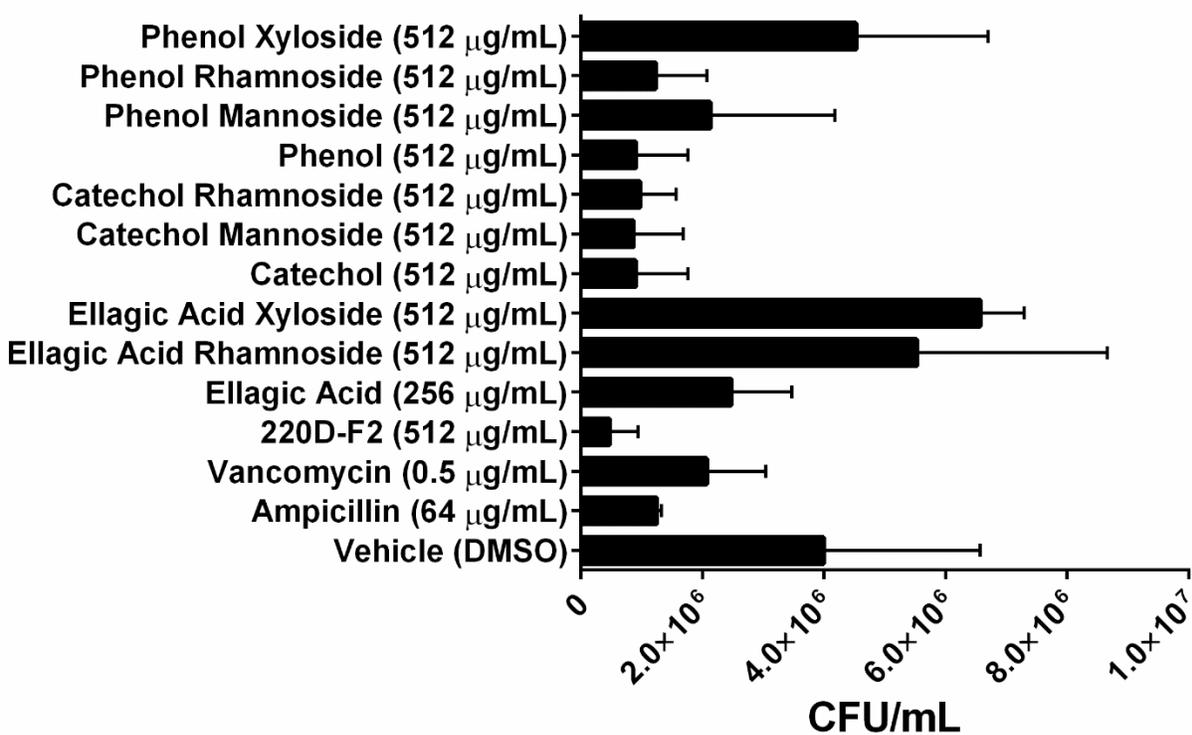
Figure S5.2. Total CFU/mL following 18 h incubation. None of the above represented concentrations exhibited a statistically significant difference from the vehicle (DMSO) control.
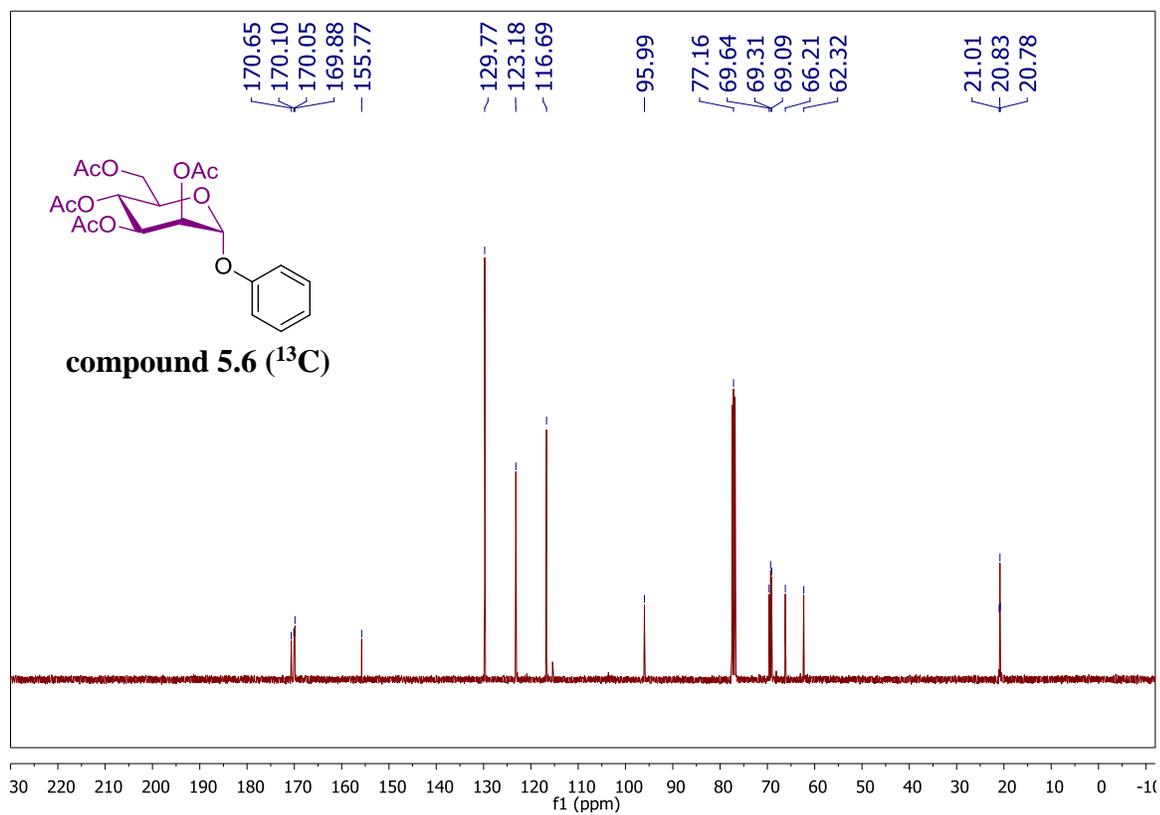
NMR Spectra

compound 5.3 (^{1}H)

compound 5.3 (^{13}C)

compound 5.8 (^{1}H)

compound 5.8 (^{13}C)

compound 5.8 (^{13}C-^{1}H coupled)

anomeric $^1J_{CH}$ = 169.96 Hz

compound 5.4 (^{1}H)

compound 5.4 (^{13}C)

compound 5.9 (^{1}H)

compound 5.9 (^{13}C)

compound 5.6 (^{1}H)

compound 5.6 (^{13}C)

compound 5.11 (^{1}H)

compound 5.11 (^{13}C)

compound 5.12 (^{1}H)

compound 5.12 (^{13}C)

compound 5.12 (^{13}C-^{1}H coupled)

(anomeric $^1J_{CH}$ = 171.46 Hz)

compound 5.10 (^{1}H)

compound 5.10 (^{13}C)

compound 5.14 (^{1}H)

compound 5.14 (^{13}C)

compound 5.15 (^{1}H)

compound 5.15 (^{13}C)

compound 5.15 (^{13}C-^{1}H coupled)

(anomeric $^1J_{CH}$ = 177.75 Hz)

10106.02
9928.27

compound 5.17 (^{1}H)

7.56
7.42
5.51
4.05
3.79
3.79
3.77
3.76
3.34
3.31
2.50
1.07
1.06

1.00
1.02
0.98
1.11
1.08
1.08
1.24
3.13

compound 5.17 (^{13}C)

compound 5.16 (^{1}H)

compound 5.16 (^{13}C)

compound 5.18 (^{1}H)
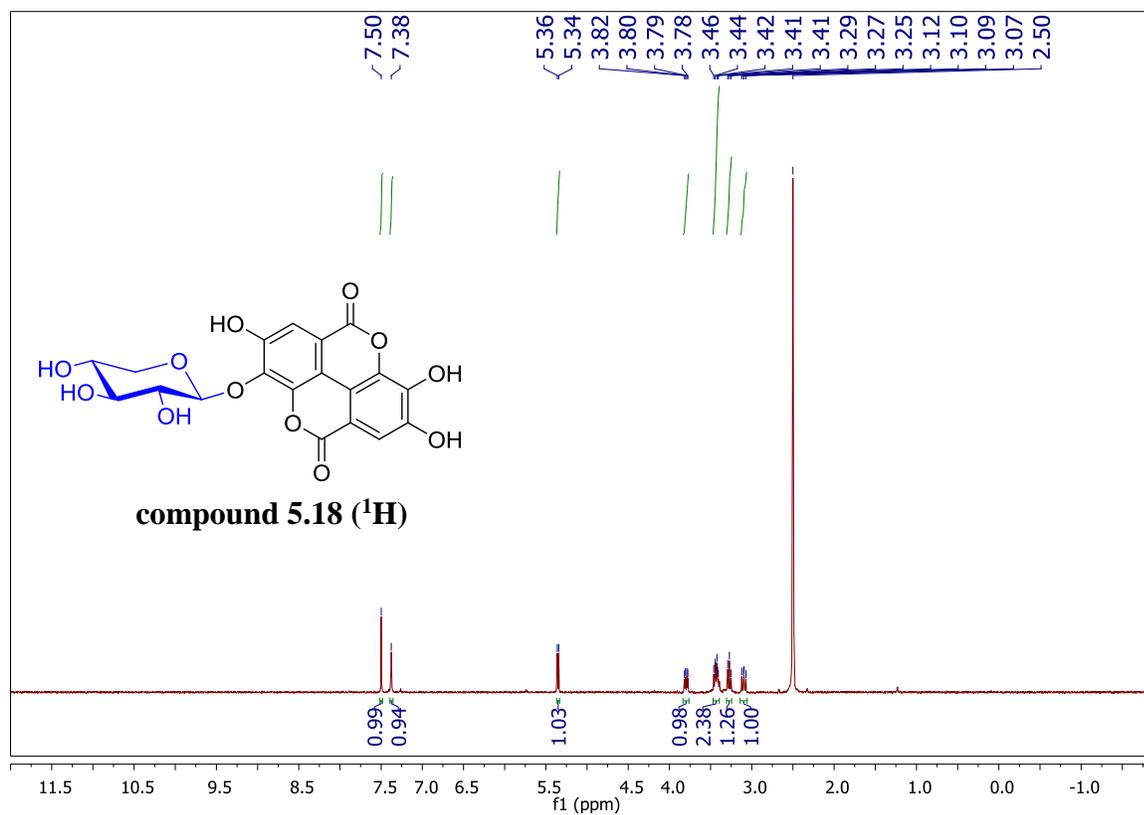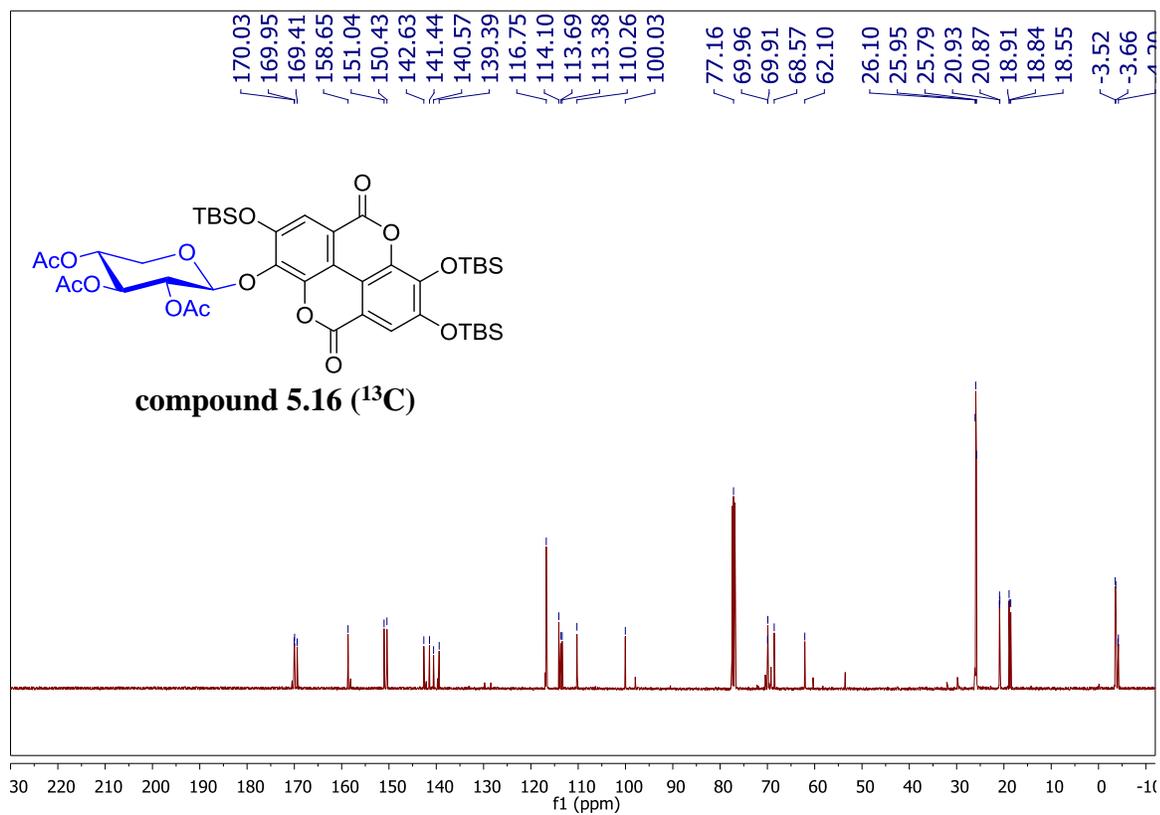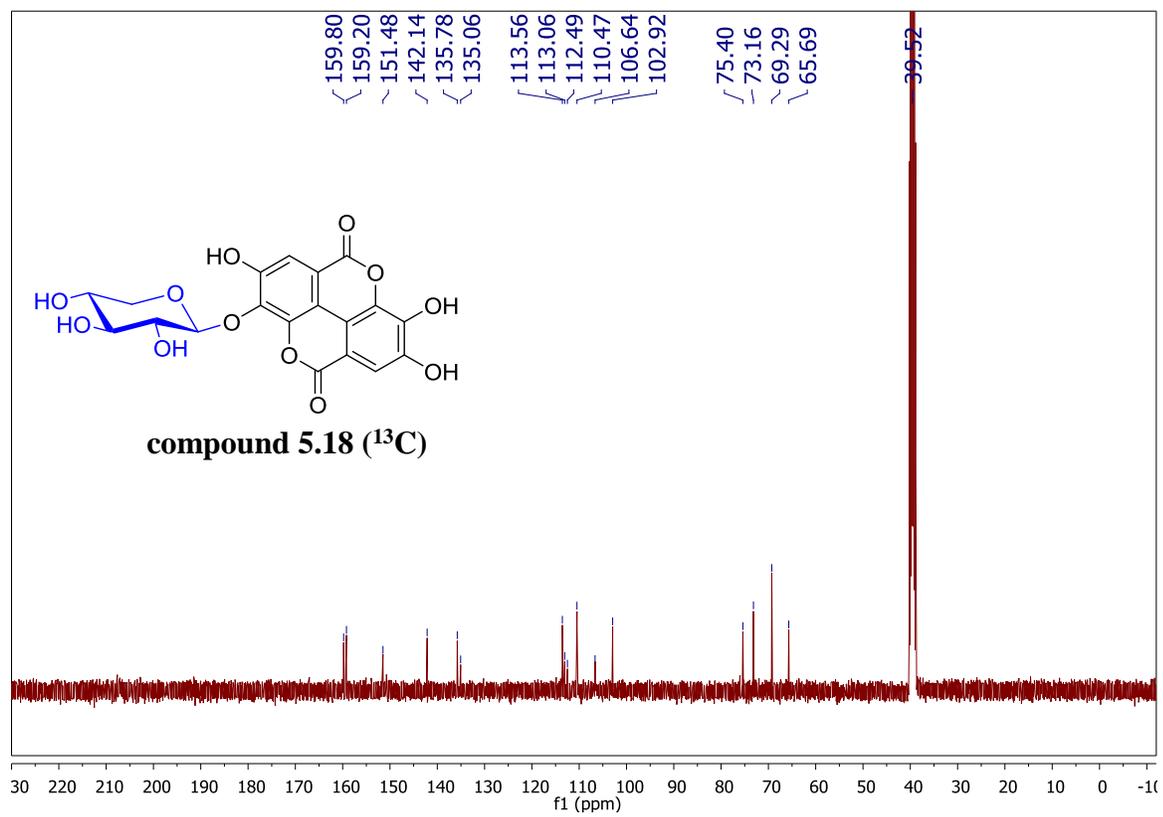
compound 5.18 (^{13}C)

Chapter 6: Conclusions, ongoing work, and future directions

Chapter 6: Conclusions, ongoing work, and future directions

6.1 Conclusions, ongoing work, and future directions

The work reported herein has illuminated overlooked aspects of nucleotide signaling and metabolism through the development of novel (bio)chemical approaches to modulate nucleotide levels *in vivo*, with a focus on unraveling the functions of nucleoside 2',3'-cylcic monophosphates (2',3'-cNMPs) in *Escherichia coli*. Experimental perturbation of RNA decay has demonstrated that RNase I degrades cytoplasmic RNA substrates to produce 2',3'-cNMPs, identifying an unprecedented function for RNase I in mRNA catabolism and ribosome homeostasis (Chapter 2).[1] The importance of *E. coli* RNase I in 2',3'-cNMP production suggests that other members of the widely distributed RNase T2 superfamily generate 2',3'-cNMPs in distantly related organisms.[2] Furthermore, cell-permeable 2',3'-cNMP analogs and a catalytic fragment from a mammalian 2',3'-cNMP phosphodiesterase (CNPase) have been developed to perturb intracellular 2',3'-cNMP concentrations, enabling the discovery of the first RNase I- and 2',3'-cNMP-dependent cellular functions in prokaryotes through a combination of transcriptional profiling and phenotypic analyses (Chapter 3).[3] These studies have revealed roles for 2',3'-cNMPs and RNase I in diverse prokaryotic processes such as motility, biofilm production, and stress resistance. Additionally, targeted nucleotide metabolomic experiments and gene expression analyses have suggested that these processes are mediated by aberrant nucleotide pool regulation upon perturbation of RNase I expression and/or 2',3'-cNMP metabolism. The transcriptional data also suggest that perturbation of either RNase I or 2',3'-cNMP levels impacts c-di-GMP signaling and metabolism, which likely elicits (at least in part) the observed changes in biofilm formation and motility (**Figure 6.1**). Efforts are underway to gain insight into the biochemical effectors linking 2',3'-cNMPs and RNase I to nucleotide metabolism, c-di-GMP transduction, biofilm formation, motility, and other functions.

Furthermore, the collective transcriptional and phenotypic studies have demonstrated that 2',3'-cNMPs elicit different biological effects in *E. coli* depending on whether RNase I is present in the cell, thus identifying unique functions for RNase I *versus* 2',3'-cNMPs. Interestingly, RNase I appears to modulate certain processes independently of 2',3'-cNMP production, as CNPase-catalyzed hydrolysis of 2',3'-cNMP pools does not recapitulate certain phenotypes observed in *E. coli* deficient for both RNase I and 2',3'-cNMPs (Δ*rna*). These findings provide motivation for further exploration of the mechanisms linking RNase I and 2',3'-cNMP pools to physiological regulation in *E. coli*. Furthermore, the development of chemical tools to modulate 2',3'-cNMP levels independently of RNase I expression will promote investigation of these emerging facets of nucleic acid regulation in diverse organisms.
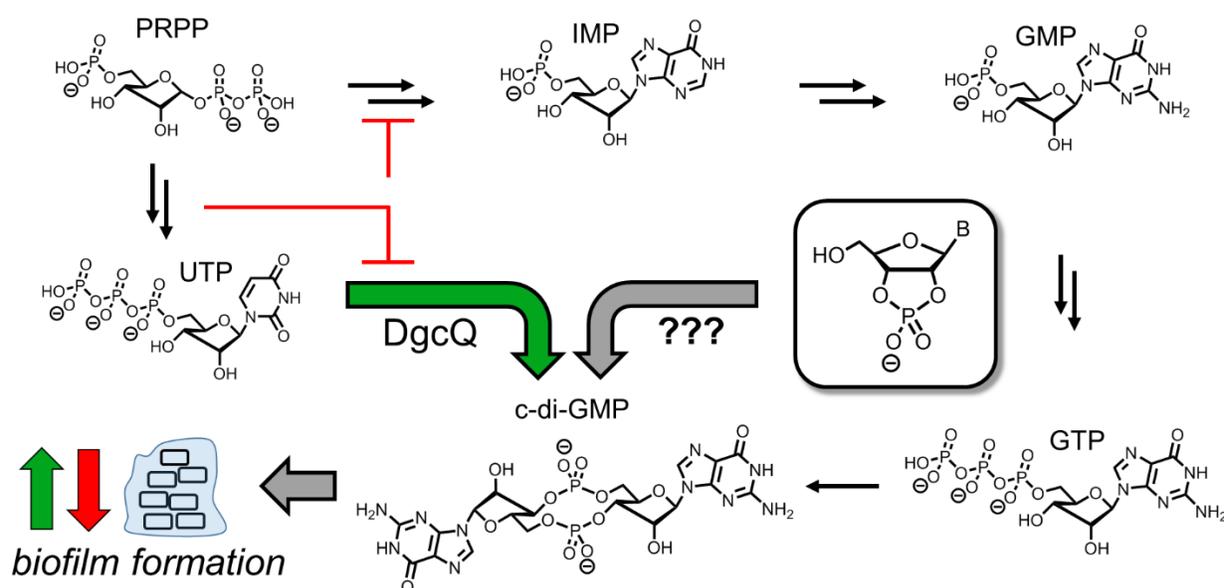

Figure 6.1. Schematic depicting emerging regulatory links between nucleotide metabolism and c-di-GMP-mediated regulation of biofilm formation. As a purine nucleotide, c-di-GMP pools inherently are dependent on *de novo* purine synthesis and salvage. Indeed, recent work revealed that inhibition of *de novo* purine synthesis attenuates biofilm production in *E. coli*, suggesting modulation of downstream c-di-GMP signaling/metabolism.[4] Additional studies discovered that UTP allosterically activates DgcQ, resulting in enhanced production of cellulose. Intriguingly, *N*-carbamoyl L-aspartate, an intermediate in *de novo* pyrimidine synthesis, allosterically inhibits DgcQ activity to suppress cellulose production.[5] These experiments suggest that *E. coli* biofilm formation is tuned by pyrimidine synthesis and degradation. Future work in our group aims to identify the functions of 2',3'-cNMPs and RNase I in the broader context of nucleotide metabolism, which potentially link these atpical cyclic nucleotides to biofilm formation and other functions.

Ongoing efforts seek to identify potential 2',3'-cNMP-binding effectors through affinity-based isolation in combination with proteomics. To this end, the synthesis of immobilized 2',3'-cNMP-agarose matrices is underway in which the agarose support will be linked to either the 5'-hydroxyl group or the nucleobase (**Figure 6.2**). Several methods previously have been developed for 2',3'-cNMP synthesis from nucleoside 2'- or 3'- monophosphates (2'/3'-NMP) or nucleosides. Synthesis from a 2'- or 3'-NMP has been accomplished by dehydrative cyclization of the phosphate in the presence of acylating reagents[6] (**Figure 6.2A**). However, attempts to directly modify the 2',3'-cNMP with the 5'-*O*-alkynyl moiety *via* acylation with propargyl chloroformate were unsuccessful, necessitating preparation of the affinity matrices from the respective nucleosides. To this end, preparation of 2',3'-cNMPs from nucleosides has been accomplished by protecting group manipulation to afford ketal **6.2** as a key intermediate (**Figure 6.2B**).[7] Synthesis of the alkyne-functionalized 2',3'-cUMP was performed from uridine by protecting the 2',3'-diol prior to acylation with propargyl chloroformate to obtain alkynyl 2',3'-cUMP **6.6** (**Figure 6.2C**).[7] Unfortunately, analogous acylation conditions failed to provide the 5'-*O*-propargyloxycarbonyl adenosine derivative due to significant decomposition during aqueous workup. Thus, protected 2',3'-isopropylidene adenosine **6.11** was coupled to *para*-propargyloxyphenol under Mitsunobu conditions to afford the 5'-*O*-phenolic ether **6.12** after ketal cleavage (**Figure 6.2E**),[8] providing a versatile method for modifying nucleosides with phenolic or carboxylic acid fragments.
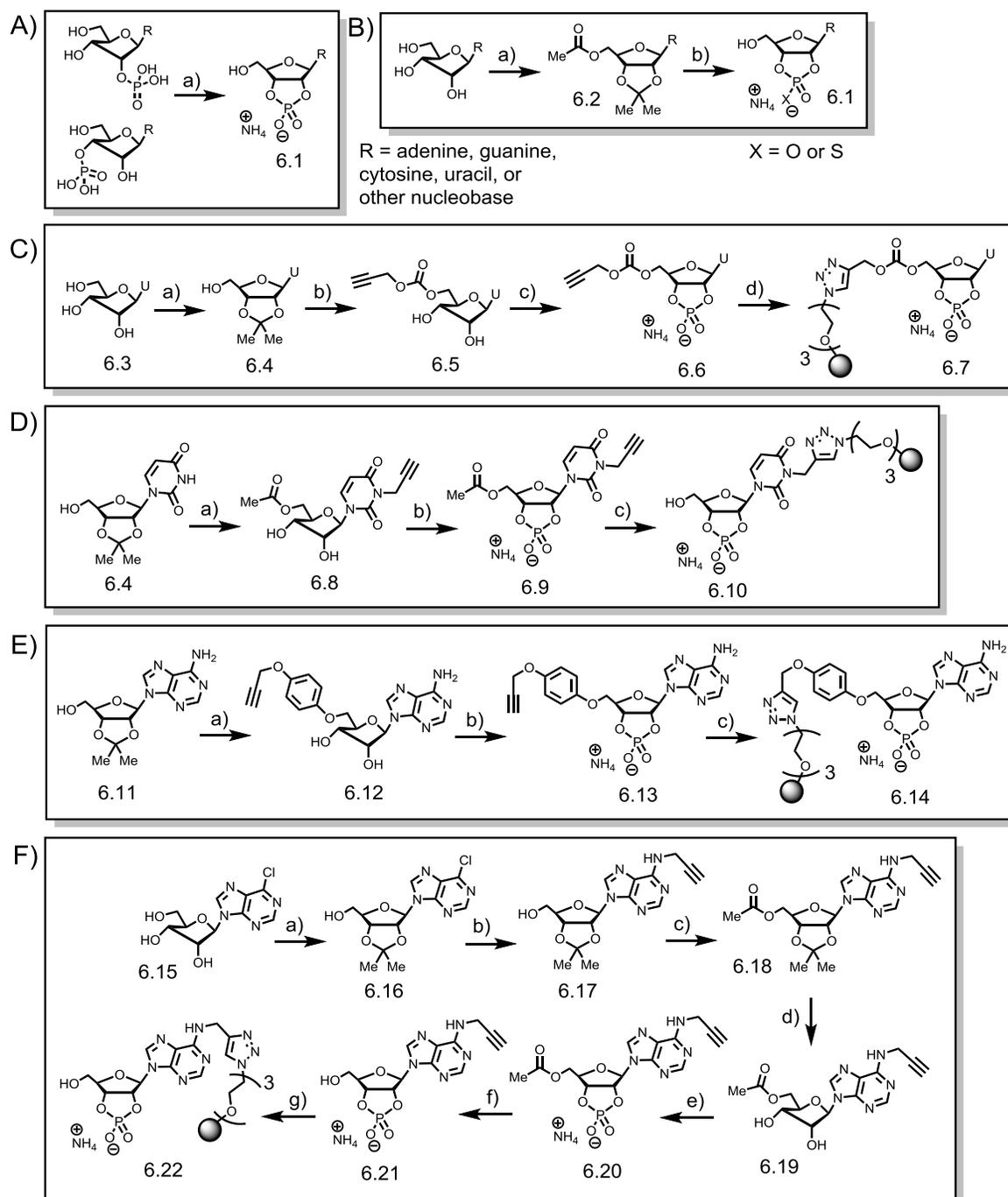
Figure 6.2. Proposed synthetic schemes to prepare 2',3'-cNMPs and analogs. **A) a)** 1) Ac₂O, pyridine; 2) Et₃N, H₂O. **B) a)** 1) Acetone, *para*-toluenesulfonic acid; 2) Ac₂O, pyridine. **b)** 1) 50% aq. HCO₂H; 2) POCl₃ (for X = O) or PSCl₃ (for X = S), pyridine; 3) Et₃N, H₂O. **C) a)** Acetone, *para*-toluenesulfonic acid (*68%*). **b)** 1) Propargyl chloroformate, MeCN, pyridine; 2) 50% aq. HCO₂H (*52% over 2 steps*). **c)** POCl₃, pyridine (*33%*). **d)** Azido-agarose, CuSO₄, sodium ascorbate, ligand, aq. buffer. **D) a)** 1) Propargyl bromide, DBU, THF; 2) Ac₂O, pyridine; 3) 50% aq. HCO₂H. **b)** POCl₃, pyridine. **c)** 1) Et₃N, H₂O. 2) Azido-agarose, CuSO₄, sodium ascorbate, ligand, aq. buffer. **E) a)** 1) THF, PPh₃, diisopropylazodicarboxylate, *para*-propargyloxyphenol (*21%*); 2) 50% aq. HCO₂H (*98%*). **b)** POCl₃, pyridine. **c)** Azido-agarose, CuSO₄, sodium ascorbate, ligand, aq. buffer. **F) a)** 50% aq. HCO₂H. **b)** Propargylamine, K₂CO₃, *n*-PrOH. **c)** Ac₂O,

pyridine. **d)** 50% aq. HCO_2H. **e)** $POCl_3$, pyridine. **f)** Et_3N, H_2O. **g)** Azido-agarose, $CuSO_4$, sodium ascorbate, ligand, aq. buffer.

In addition to preparation of 5'-O-modified derivatives, agarose functionalization at N^3 of 2',3'-cUMP is in progress, and protected N^3-propargyl uridine **6.8** has been prepared to enable phosphorylation (**Figure 6.1D**). Synthesis of N^6-propargyl 2',3'-cAMP will employ amination of 2',3'-O-protected 6'-chloropurine riboside **6.16** with propargylamine,[9] which proceeds in basic DMF. Subsequent 5'-O-protection, 2',3'-O-phosphorylation, and deprotection will afford the alkynyl 2',3'-cyclic phosphodiester **6.21** (**Figure 6.1F**). Analogous amination methodology using 8-bromo purines or 5-bromo pyrimidines also can be employed to further alter the site of agarose attachment.[10] In addition to the preparation of various 2',3'-cNMP affinity matrices, these multiple synthetic approaches enable the preparation of additional (non)-natural analogs (such as 2',3'-cyclic phosphorothioates; **Figure 6.1B**) from commercially available nucleosides.

Current work also aims to identify inhibitors of *E. coli* RNase I due to the emerging role of this nuclease in resistance-related processes, such as copper and iron homeostasis, acid resistance, and *β*-lactam tolerance. A fluorescently labeled RNA oligonucleotide substrate will be utilized to screen for inhibitors *in vitro* using a Förster resonance energy transfer (FRET)-based assay (**Figure 6.2**), in analogy to a published screen which discovered inhibitors of HIV-1 reverse transcriptase-associated RNase H.[11] Prior work also has identified phosphorothioate DNA oligonucleotides as inhibitors of mammalian RNase L;[12] thus the initial screen will focus on disrupting RNase I with non-natural oligonucleotides (**Figure 6.2**), particularly oligo-cytidine and oligo-guanosine analogs due to the preferential affinity of RNase I for these RNA oligomers.[13] Provided that RNase I inhibitors can be identified *in vitro*, subsequent experiments will evaluate their effects on RNase I-dependent phenotypes in *E. coli*.

Along with the initial characterization of RNase I and 2',3'-cNMP pools in *E. coli*, this dissertation has provided a foundation for the future identification of potential enzymes involved in the metabolism and sensing of cytidine 3',5'-cyclic monophosphate (3',5'-cCMP) in eukaryotes (Chapter 4). Towards this end, a nucleotidyl cyclase activity assay has been optimized to facilitate fractionation of putative enzymes involved in 3',5'-cCMP biosynthesis. Moreover, this work has begun to develop fractionation procedures employing synthetic nucleotide affinity matrices to identify 3',5'-cCMP-binding enzymes (such as putative phosphodiesterases and downstream effectors) from mammalian tissue homogenate and cell lysate.
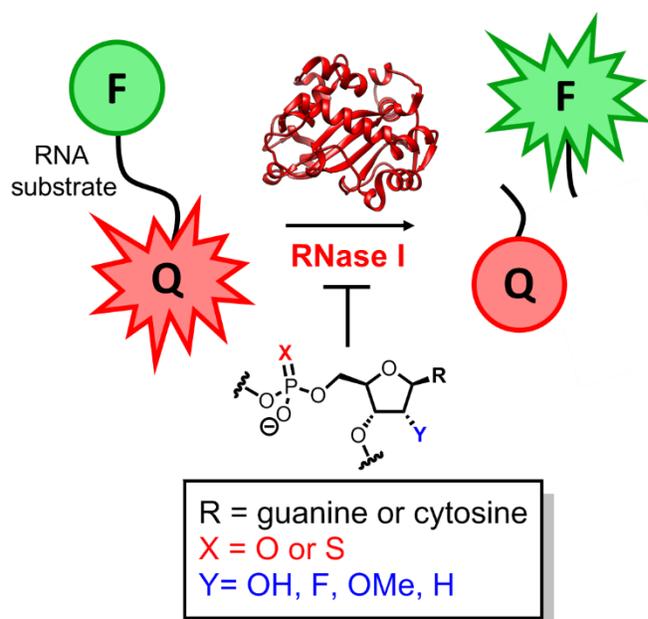

Figure 6.3. High-throughput FRET-based screen for oligonucleotide inhibitors of *E. coli* RNase I. RNase I-mediated cleavage of the substrate RNA will liberate the quencher fluorophore (Q), resulting in increased fluorescence of the donor fluorophore (F). Conversely, inhibition of RNase I activity will attenuate donor fluorescence and increase quencher fluorescence.

In addition to the development of chemical approaches to probe nucleotide signaling in prokaryotic and eukaryotic biology, this work reports the synthesis and biological characterization of small molecule biofilm modulators to further interrogate bacterial physiology (Chapter 5).[14] These compounds were inspired by phenolic glycoside natural products observed in an anti-biofilm extract prepared from *Rubus ulmifolius*, a blackberry shrub native to the Mediterranean. Chemical synthesis enabled identification of ellagic acid glycosides as active constituents of the complex botanical extract, providing novel chemical tools to modulate biofilm formation in *Staphylococcus aureus*. The synthetic strategy leverages a regioselective deprotection

of a per-*O*-silyl ellagic acid aglycone *in situ*, followed by diastereoselective glycosylation with a per-*O*-acetyl glycosyl iodide, ultimately affording the desired ellagic acid glycosides as single anomers after deprotection. Future efforts will interrogate the mechanisms governing biofilm inhibition and evaluate the inhibitory activity against diverse bacterial taxa. Collectively, this dissertation has expanded our arsenal of chemical tools to interrogate biological processes through perturbation and analysis.

6.2 References

(1) Fontaine, B. M.; Martin, K. S.; Garcia-Rodriguez, J. M.; Jung, C.; Southwell, J. E.; Jia, X.; Weinert, E. E. *Biochem. J* **2018**, *478*, 1491-1506.

(2) Luhtala, N.; Parker, R. *Trends Biochem. Sci.* **2010**, *35*, 253-259.

(3) Fontaine, B. M.; Duggal, Y.; Weinert, E. E. *Submitted to Nucl. Acids Res.* **2018**.

(4) Antoniani, D.; Rossi, E.; Rinaldo, S.; Bocci, P.; Lolicato, M.; Paiardini, A.; Raffaelli, N.; Cutruzzola, F.; Landini, P. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 7325-7336.

(5) Rossi, E.; Motta, S.; Aliverti, A.; Cossu, F.; Gourlay, L.; Mauri, P.; Landini, P. *Environ. Microbiol.* **2017**, *19*, 4551-4563.

(6) Wang, Y.; Liu, X.; Yang, Z.; Wang, Q.; Xu, Y.; Wang, Q.; Xu, J. *Nucleic Acids Res.* **1987**, *15*, 4291-4305.

(7) Tatani, K.; Hiratochi, M.; Nonaka, Y.; Isaji, M.; Shuto, S. *Acs Medicinal Chemistry Letters* **2015**, *6*, 244-248.

(8) Lougiakis, N.; Papapetropoulos, A.; Gikas, E.; Toumpas, S.; Efentakis, P.; Wedmann, R.; Zoga, A.; Zhou, Z.; Iliodromitis, E.; Skaltsounis, A.; Filipovic, M.; Pouli, N.; Marakos, P.; Andreadou, I. *J. Med. Chem.* **2016**, *59*, 1776-1790.

(9) Carter-O'Connell, I.; Jin, H.; Morgan, R.; David, L.; Cohen, M. *J. Am. Chem. Soc.* **2014**, *136*, 5201-5204.

(10) (a) Andrei, M.; Bjornstad, V.; Langli, G.; Romming, C.; Klaveness, J.; Tasken, K.; Undheim, K. *Organic & Biomolecular Chemistry* **2007**, *5*, 2070-2080; (b) Maity, J.; Stromberg, R. *Molecules* **2013**, *18*, 12740-12750.

(11) Parniak, M.; Min, K.; Budihas, S.; Le Grice, S.; Beutler, J. *Anal. Biochem.* **2003**, *322*, 33-39.

(12) Player, M.; Torrence, P. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 891-894.

(13) Meador, J.; Cannon, B.; Cannistraro, V.; Kennell, D. *Eur. J. Biochem.* **1990**, *187*, 549-553.

(14) Fontaine, B. M.; Nelson, K.; Lyles, J. T.; Jariwala, P. B.; Garcia-Rodriguez, J. M.; Quave, C. L.; Weinert, E. E. *Front. Microbiol.* **2017**, *8*, DOI: 10.3389/fmicb.2017.00496.