



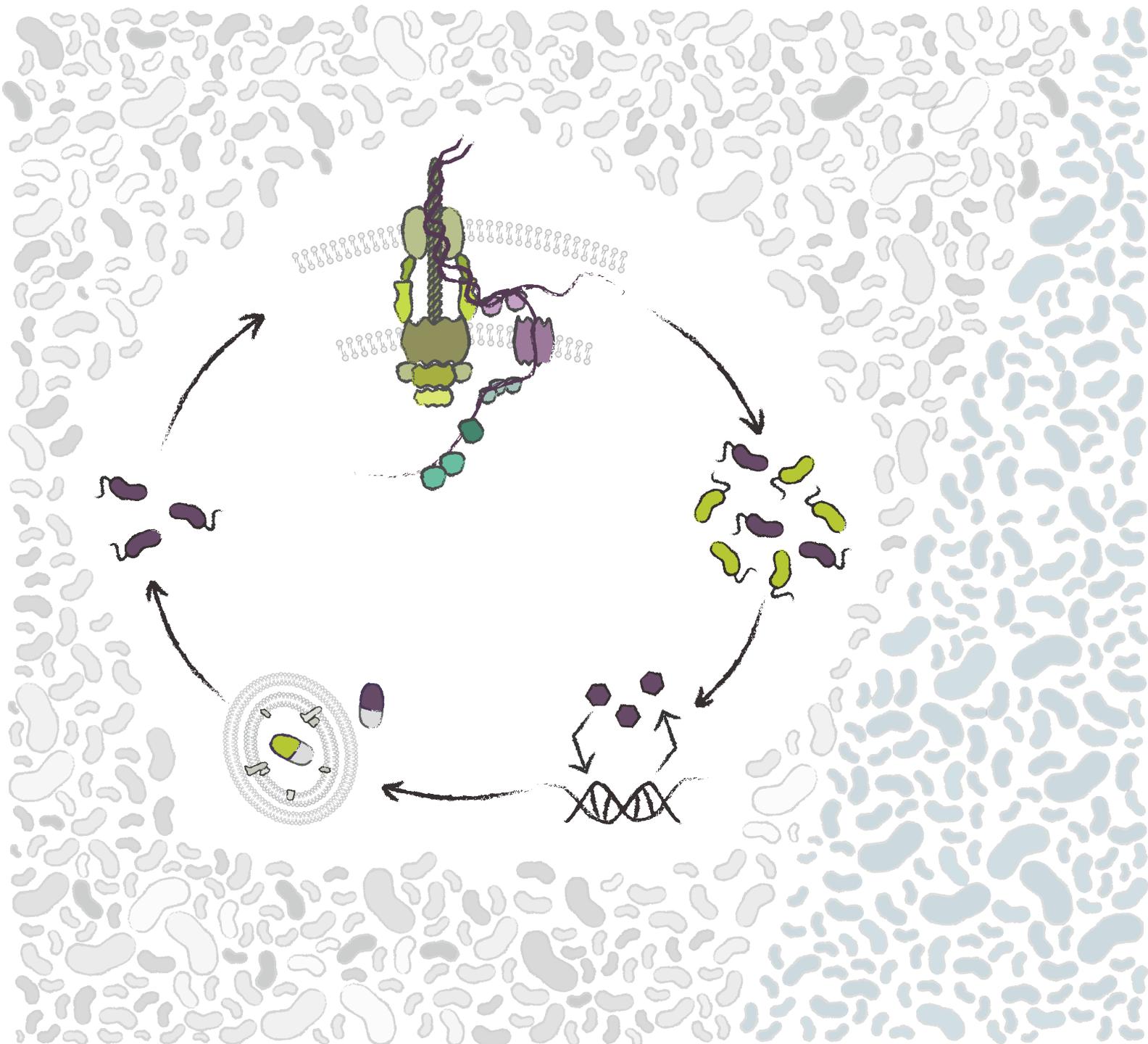
Faculty of Health Sciences

Microbial Adaptation Responses to External Cues

Merete Storflor

A dissertation for the degree of philosophiae doctor

October 2023



A dissertation for the degree of philosophiae doctor

Microbial Adaptation

Responses to External Cues

Merete Storflor



Infection Biology
Department of Pharmacy
Faculty of Health Sciences

UiT- The Arctic University of Norway
Tromsø – October 2023

Acknowledgements

This thesis was carried out at the department of Pharmacy, UiT Arctic University of Tromsø. During my PhD journey I was fortunate enough to attend various insightful conferences, workshops as well as undergo a research stay at Penn State University. This would not have been possible without the financial aid of the PhD schools IBA, and NFIF, as well as UiT travel grants.

I would like to thank my supervisors **Sören** Abel, **Klaus** Harms and **Jacob** Malone. I am grateful for the scientific discussions, and independence. I am also thankful for the enthusiasm, support and lab space provided by **Pål** Johnsen, and **Elizabeth** Fredheim. I would like to thank **Jon Kristian**, **Øystein**, and **Balpreet** for allowing me to work on my PhD in parallel with the exciting work at Chip Nanomaging. I am also grateful to my coworkers, **Ine**, **Anders**, and **Daniel** for the support and laughs shared.

I would like to thank my IFA family, in no specific order. **Fabrizio**, you have greatly impacted my way of thinking not only in terms of scientific approach but also how to evaluate what matters. **Christina**. We bonded in hell-fire and survived. You are truly incredible and your friendship is everything to me. And, lets be honest your daughter, Oliva, is a complete badass. Mr. Batman, **Francios**, you are amazing. I have lost count of the amount of plates and cloning projects you have helped me with. Now, where is that board game? **Chris**, our late-night discussions in the lab, and sippy-sippies about life, love and careers have been invaluable to me. You are truly an amazing friend. **Alex**, your compassion is endless and your smile is never far behind. I am amazed by how you pour your soul into everything you do. **Xana**, you are so beautifully wild, yet constant, and composed. Your sense of humour, and clarity continuously floors me. **Nina**, you always manage to make me laugh, and you constantly brighten the day with your stories. It's hard work, and you do it well. **Jennifer**, your willpower is incredible, waving to you on the other side of that dark corridor was an enormous energy boost. **Joseph** what can I say other than you owe me a beer and I look forward to our next heart-to-heart as well as seeing your amazing family. **João**, you have an incredible compassion and manage to find humour in everything, a true talent. **Beate**, I am so grateful for our hikes, skitrips, and conversations. I have lost count of the amount of times you dragged me out to the wilderness and talked sense in me, and I have appreciated the fun hikes and learning about minecraft from Emil. **Silje**, your compassionate and playful spirit is simply infectious. I have valued our coffee breaks, hikes, adventures and heart-to-hearts, and I look forward to the many more to come! **Miriam** og **Arne** you are both so incredibly thought-full and bursting with creativity. Hanging out with you guys is a pure joy.

I have had the great fortune of knowing incredible individuals whose characteristics and awareness I have been privileged to absorb and process. It is unfortunately too many to thank individually. I thank my entire Tromsø, Oslo, and Bergen family **Julia, Jan Håkon, Øyvind, Jingyi, Vidar, Mikkel, Maria, Rune, Hanne, Anup, Ole Morten, Selenia, Margarita, Pietro, Tejas, Dmytro, Hong, Luis, Deanna, Oddmund, Frode, Linda, Edna, Felix, Ingrid, Duarte, Dominik, Ana, Kathinka,** and **Anine**. As well as the friendships made at Penn State. Specifically **Amrita** and **Jon** - your friendship was invaluable and made the pandemic somewhat manageable.

Lena developing a quicker friendship would be impossible. You are as genuine as you are hilarious. I am also incredibly jealous of your bouldering biceps, and impleccable berry dancing skills. Now, we try our luck with those *uck off mountains!

Anne Caroline, my dear Top-Top. I will never be able to describe how far my love for you extends. You are truly a pure source of love and affection. Truer friendship is hard to find. Thank you for reading through the thesis and correcting my grammar. If that isn't true love I don't know what is. **Leia** is also have the bestest dog in the entire world.

How can I express the love I feel for my parents, **mamma** og **pappa**? Your love is unequivocal, and unconditional. You have supported my every move, and I am beyond lucky to have you as my parents! Jeg er så utrolig glad i dere! **Magnar**, you are the best big brother I could have asked for! Your sense of humour has shaped my personality from birth, making my quirks entirely your fault. Jeg er glad i deg!

I have been fortunate enough to dive into several fields separately. In doing so I have witnessed a small fraction of the awe-inducing spectacle bacteria are capable of, and despite human ingenuity, bacteria have, without fail, managed to adapt. That being said, bacteria can only adapt to what they have already encountered. Whereas, this PhD has prepared me for just about anything.

It's been cholera!

Tromsø, October 2023

A handwritten signature in blue ink, appearing to read "Anders Skjellum". The signature is fluid and cursive, written in a personal style.

*The thesis is dedicated to my parents, Aud and Harry,
two incredible people who offer everything
and show nothing but love and laughter.*

*“Oh, you pretty things
Don't you know [you drive me] insane”
– Modified quote by David Bowie –*

Table of Contents

1	Abstract	1
2	Background	2
3	Aims of the Thesis	4
4	Introduction.....	5
4.1	Population dynamics	5
4.1.1	Selection	5
4.1.2	Heterogeneity.....	7
4.1.2.1	Phenotypic diversity	8
4.1.2.2	Metabolic heterogeneity	9
4.1.2.3	Bet hedging strategy – spread out the risk	11
4.1.3	Horizontal gene transfer.....	12
4.1.3.1	Conjugation	12
4.1.3.2	Transduction.....	12
4.1.3.3	Natural transformation	13
4.1.3.4	Homologous recombination.....	14
4.1.4	<i>V. cholerae</i> transitions between environments.....	15
4.1.5	Infection cycle	16
4.1.6	<i>Vibrio cholerae</i> is adept at natural transformation.....	18
4.1.6.1	Natural Transformation Machinery	19
4.2	Antibiotics	21
4.2.1	The pre-antibiotic era	21
4.2.2	Antibiotic resistance	23
4.2.3	Drug discovery challenges	23
4.2.3.1	Golden age of antibiotics	24
4.2.4	Mechanisms of resistance	25
4.2.4.1	Genetic mutation – resistance	25
4.2.4.2	Horizontal gene transfer – antibiotic resistance.....	26
4.2.4.3	Mobile genetic elements.....	26
4.2.5	Pharmacokinetics and pharmacodynamics PK/PD models	28
4.2.5.1	Drugs in the body – Pharmacokinetics:	29
4.2.5.2	Drugs in bacteria – Pharmacodynamics.....	30

4.2.6	Selection pressures	32
4.2.7	Dosing strategies	33
4.3	The cell envelope	34
4.3.1	Dynamic structure important for protection and adaptation.....	34
4.3.2	The cell envelope.....	34
4.3.3	Membranes.....	36
4.3.3.1	Glycerophospholipids	36
4.3.4	Lipid – protein interaction.....	37
4.3.4.1	Membrane proteins	38
4.3.5	Cardiolipin – MgtA interaction	39
5	Synopses of publications.....	43
	Paper I: Transposon insertion sequencing reveals dependency of natural transformation on amino acid metabolism.....	43
	Paper II: Drug-target binding quantitatively predicts optimal antibiotic dose levels in quinolones.....	44
	Paper III: The bacterial magnesium transporter MgtA reveals highly selective interaction with specific cardiolipin species	45
6	Methodological considerations	47
6.1	Model organisms:	47
6.1.1	<i>Vibrio cholerae</i>	47
6.1.2	<i>Escherichia coli</i>	47
6.2	Transposon insertions sequencing.....	47
6.2.1	Creating the library, conjugation	50
6.2.2	DNA shearing and PCR amplification	50
6.2.3	Size exclusion of DNA fragments.....	51
6.2.4	Sequencing.....	52
6.2.5	Assessing library saturation – proper sampling	53
6.3	Natural transformation homologous recombination	55
6.3.1	Nongenetic inheritance of tolerance.....	56
6.4	MIC determination and time-kill curves.....	57
6.5	Single cell microscopy	58
6.5.1	Fluorescent transcription fusion reporter.....	59
7	Results and Discussion	61

7.1	Membrane Protein localization: Cardiolipin and MgtA	61
7.2	Mechanistic models can limit antibiotic resistance	66
7.3	Amino acid metabolism modulates natural transformation	70
8	Conclusion.....	76
9	Outlook.....	77
10	References	79

Abbreviations

ADME	Absorption, Distribution, Metabolism, and Excretion
AI	Autoinducer
AMR	Antimicrobial Resistance
BFP	Blue Fluorescent Protein
CCR	Carbon Catabolite Repression
CDC	Center of Disease Control
CL	Cardiolipin
DNA	Deoxyribonucleic Acid
DUS	DNA Uptake Sequence
FP	Fluorescent Protein
GFP	Green Fluorescent Protein
HGT	Horizontal Gene Transfer
ICE	Integrative and Conjugative Elements
Mg ²⁺	Magnesium
MGE	Mobile Genetic Element
MIC	Minimal Inhibitory Concentration
MPC	Mutation Prevention Concentration
MSW	Mutant Selection Window
PD	Pharmacokinetics
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
pH	Potential of Hydrogen
PK	Pharmacokinetics
PTS	Phosphoenolpyruvate:Carbohydrate Phosphotransferase System
QS	Quorum Sensing
SOS	Save Our Ship
Ssb	Single Stranded Binding (proteins)
ssDNA	Single Stranded DNA
T6SS	Type VI Secretion System
UV	Ultraviolet
nm	Wavelength
WHO	World Health Organization

List of papers

Paper I: Transposon insertion sequencing reveals dependency of natural transformation on amino acid metabolism

M. Storflor, J. A. Gama, K. Harms, S. Abel
Manuscript

Paper II: Drug-target binding quantitatively predicts optimal antibiotic dose levels in quinolones

F. Clarelli, A. Palmer, B. Singh, **M. Storflor**, S. Lauksund, T. Cohen, S. Abel, P. Abel zur Wiesch

Published - PLOS Computational Biology, 14.08.2020,
<https://doi.org/10.1371/journal.pcbi.1008106>

Paper III: The bacterial magnesium transporter MgtA reveals highly selective interaction with specific cardiolipin species

J. Weikum, J. van Dyck, S. Subramani, D. P. Klebl, **M. Storflor**, S. P. Muench, S. Abel, F. Sobott, J. P. Morth

Manuscript accepted - BBA - Molecular Cell Research

Pre-print: BioRxiv, 27.09.2021. <https://doi.org/10.1101/2021.09.27.462033>

During this thesis I supervised a Master's student, Amie Manzi, publication not included in the thesis.

Author contributions	Experimental design	Experimental execution	Data analysis	Original draft preparation	Visualization	Review and editing
Paper I	MS, KH, SoA	MS	MS, JAG	MS	MS	MS, JAG, KH, SoA
Paper II	F.C, MS, BS, SoA, PAzW	MS, BS, SL	F.C, A.P, MS, BS, AS, PAzW	F.C, SoA, PAzW	F.C, SoA, PAzW	F.C, MS, BS, AP, TC, AoS, PAzW
Paper III	JW, JvD, SaS, DPK, MS, SPM, SoA, FS, JPM	JW, JvD, SS, DPK, MS, SPM, FS, JPM	JW, JvD, SaS, DPK, MS, SPM, SoA, FS, JPM	JW, JPM	JW, JvD, SaS, DPK, MS, SPM, SoA, FS, JPM	JW, JvD, SaS, DPK, MS, SPM, SoA, FS, JPM

Authors: Adam Palmer (AP), Bhupender Singh (BS), David P. Klebl (DPK), Fabrizio Clarelli (FC), Frank Sobott (FS), Jens Preben Morth (JPM), Jeroen van Dyck(JvD), João Alves Gama (JAG), Julia Weikum(JW), Merete Storflor (MS), Pia Abel zur Wiesch (PAzW), Saranya Subramani (SaS), Silje Lauksund (SL), Sören Abel (SoA), Stephen P. Muench (SPM), Ted Cohen(TC).

1 Abstract

Bacteria can adapt to various environments because they efficiently sense and respond to external cues. This is exemplified by antibiotic resistance, a leading cause for death on a global scale. Antibiotics exert a selection pressure on bacteria, where improper use can inadvertently select for antibiotic resistance. Dissemination of antibiotic resistance throughout bacterial populations can occur via horizontal gene transfer. Bacteria can also take up genetic material from the environment and incorporate it into their genome. This is known as natural transformation and is dependent on bacteria entering a state of competence to take up environmental DNA. Induction of competence involves an extensive network of signaling, that involves sensing and responding to external cues such as population density, and nutrient availability. The external cues are then converted into a phenotypic response. The cell surface is critical to regulate this activity and occurs via an intricate tapestry of membrane proteins, lipids and more. The orchestration of these components is vital for appropriate signal processing and proper adaptation.

This thesis explores bacterial adaptation in response to external cues such as antibiotics and environmental DNA. As well as activity and localization of membrane proteins.



Bakterier har en enorm evne til å tilpasse seg ulike miljøer som innebærer å tolke eksterne stressfaktorer. Et eksempel på dette er antibiotikaresistens, som er en ledende dødsårsak på global basis. Antibiotika utsetter bakterier for et seleksjonspress, ved feil bruk vil man kan ubevisst selektere for resistente bakterier. Resistensgener kan så overføres ved horisontal genoverføring. På den måten kan resistens fordeles i populasjonen. Bakterier kan også ta opp fremmed DNA fra miljøet og inkorporere det i sitt genom. Dette kalles naturlig transformering og krever at bakteriene blir kompetente til å kunne ta opp fremmed DNA. Dette innebærer som regel et innviklet nettverk av signalering. Eksterne faktorer som populasjonstetthet og tilgang på næring blir da konvertert til en phenotypisk respons. Celleoverflaten er kritisk for å regulere denne aktiviteten og foregår via ett tettpakket vev av blant annet membranproteiner og lipider. Dette er essensielt for å kunne registrere og videreføre informasjon fra miljøet.

Denne avhandlingen utforsker tilpasningsevnen til bakterier i henhold til stressfaktorer, som antibiotika og DNA fra miljøet, og samspillet mellom aktivering av membranproteiner, som nevnt ovenfor.

2 Background

Bacterial infections were in 2019 ranked as the second leading cause of death⁵. There are several factors, both environmental and behavioral, that contribute to the survival and spread of antibiotic-resistant bacteria. Some key drivers include Improper use of antibiotics, poor water quality, inadequate sanitation, and suboptimal infection prevention. Consequently, these factors limit the effectiveness of antibiotics, reducing the success rate of treatment and limits prevention. Antimicrobial resistance leads to serious illness, prolonged hospital stays and increased healthcare cost. Ironically, the use of antibiotics is a significant contributor to antibiotic resistance. The antibiotic exposes bacteria to a selection pressure that can inadvertently select for resistance, which can then propagates. A bacterial population is also genetically and phenotypically divers^{6,7}, with the added ability to transfer genetic material both vertically and horizontally. These are all factors that influence the ability to adapt to environmental stressors, making it difficult to control infectious disease and imposing additional burden the global public health.

An attractive target for antibiotics is the cell envelope, which functions as the main defense against the immediate environment. It provides a hydrophobic barrier with size exclusion properties and contains components that are indispensable for bacterial survival and growth. This complex architecture comprise of functional layers that differ between Gram-positive and Gram-negative bacteria. Generally, Gram-positive bacteria have two functional layers - a cytoplasmic membrane surrounded by a thick cell wall - while, Gram-negative bacteria have three functional layers - a cytoplasmic or inner membrane, a thin cell wall, and an outer membrane. We will focus on Gram-negatives. There is significant signaling at the bacteria-environment interface, facilitated by transmembrane signaling systems, and membrane proteins⁸. Bacteria have been found to restructure their membrane composition depending on the environment, as observed in *Vibrio cholerae* in the presence of bile⁹. Changes in the bacterial surroundings can be relayed and rapidly converted into cellular signals, allowing the organism to either adapt or relocate. Maintaining the integrity of the envelope is therefore critical for adaptation and survival.

A major route for adaptation is horizontal gene transfer (HGT), and is a recognized evolutionary force that reshapes bacterial genomes. Bacteria can acquire large stretches of genetic material from other organisms that enable them to adapt to their environment. This can result in, for example, increased infectivity as well as improved colonization which can have considerable consequences. An enteric pathogen adept at acquiring new alleles via HGT, is *V. cholerae*. This pathogen holds major socio-economic implications for vulnerable communities, constraining their development¹⁰. It has been seen that a non-toxigenic strain can be converted into a toxigenic strain via natural transformation¹¹, one of the major modes of HGT. This efficient shuffling

or import of genetic material contributes to the success of this pathogen. HGT provides a route for adaptation as well as facilitating the spread resistance.

Antibiotic resistance is ever increasing and remains a global challenge, especially since novel antibiotic discovery has stalled since the 1960s. Bacteria employ several strategies to evade antibiotic effects, such as acquiring genetic mutations that alter the drug target, rendering the drug ineffective. Bacteria can also survive exposure to antibiotics by having acquired genes that confer resistance. Further, they can survive antibiotic exposure by limiting drug uptake, altering the permeability of the outer membrane or generate biofilms that provide an additional “impenetrable” layer of protection¹². Effective drug action requires an interaction between the drug and its target, necessitating an ideal concentration to eradicating the entire bacterial population, limiting resistance development, selection, as well as HGT. However, such a concentration is likely to have toxic effects on the patient. HGT plays an important role in bacterial evolution which includes rapid adaptation to environment as well as reducing the genetic burden of clonal reproduction. HGT can occur between across species allowing for macroevolution^{13,14}. An important mechanism of HGT is known as transformation. This process involves the detection of extracellular DNA, transportation into the cytoplasm and integration into the genome via homologous recombination. However this is only the first step of gene transfer. Given that the variant provides a selective advantage it needs to spread through the population, and there will likely need to be a fine-tuning of the regulator system. While these selective forces take longer¹⁵, the initial gene transfer can aid in local and rapid adaption¹⁶. Which, is why a clear understanding of how bacteria sense and respond to the emmediate environment is crucial. This thesis focuses on understanding microbial responses to external cues such as antibiotic, ions and genetic material. In an effort to gaining an understanding of how bacteria process environmental cues.

3 Aims of the Thesis

This thesis focuses on understanding how bacteria respond to stressors such as extracellular DNA and antibiotics. We were also interested in the role of the cell membrane in modulating the response of protein activity and localization. *Specific parameters we were interested in entail better understanding of how DNA is taken up from the environment, how to avoid resistance to antibiotics, and the interplay of lipids and membrane proteins.*

Specific aims:

Paper I: Gather a better understanding of the mechanisms that facilitate DNA uptake from the environment. To achieve this, we chose *V. cholerae* as a model organism to explore its response to extracellular DNA. Our interest lies in understanding the fitness effects of genes involved in this process, and aimed to find major pathways that contribute to DNA uptake and recombination.

1. Generate a transposon insertion library to assess fitness effects
2. Apply the library to a natural transformation assay
3. Investigate candidate genes/ pathways involved in natural transformation

Paper II: Investigate the bacterial response to antibiotics focusing drug target interactions within the cell. The objective was to understand the impact of antibiotic dosing in connection with target binding, aiming to prevent resistance selection and facilitate the recycling of traditional antibiotics. We opted to use *E. coli* as a model organism, as a great deal of mechanistic information about target binding affinities is available along with intensive studies of specific target binding mutations.

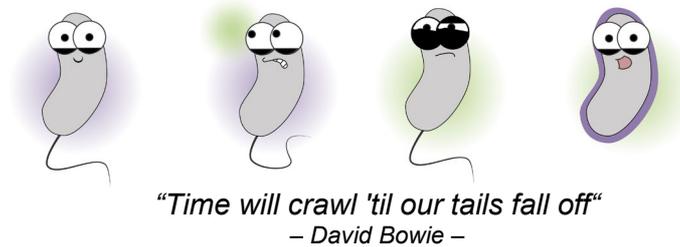
1. Investigate if there is a linear correlation between the minimal inhibitory concentration and target binding affinities
2. Determine MICs
3. Test the effects of altered target sites in terms of drug response

Paper III: Understand how lipid interactions mediate membrane protein activity and function to stabilize complexes and localization. Finding cooperative effects between lipids of the same class and a membrane protein with implications for regulation of infection.

1. Establish an interaction between membrane protein and phospholipid
2. Investigate activity of membrane protein in the context of phospholipid.
3. Localization/anchoring of membrane protein mediated by phospholipid

4 Introduction

4.1 Population dynamics



Bacteria can adapt to a wide range of environments, - *In my opinion, this a direct result of bacteria being 95% free-spirit, and 5% pure spite*¹⁷ - and possess a remarkable capacity to grow. However, they rarely encounter optimal growth conditions outside of the laboratory. Bacteria are often in competition for the limited resources and at the same time must contend with threats from other bacterial species, host defenses, and antibiotics. To navigate these challenges bacteria have developed a wide range of mechanisms that enable them to alter their growth patterns in response to the environment, such as entering a growth arrest or executing population-wide behaviors based on growth density. These complex processes foster bacterial adaptation with mutation rates and inherent heterogeneity contributing to diversity enabling swift responses. Additionally, bacteria utilize horizontal gene transfer (HGT) to acquire genetic material, facilitating rapid adaptation to changing environmental conditions, provided there is a competitive edge. As the fitness-driven effect of natural selection will impact how HGT shapes the microbial community.

4.1.1 Selection

Bacteria can undergo significant change within a relatively short amount of time partly due to the ability to mutate as well as acquire genetic material. This can in turn affect the bacterial generation time e.g. alter the fitness effect. Such fitness effects can be either beneficial, neutral, or detrimental. A beneficial mutation will result in a new allele or gene variant and increases fitness. Eventually the mutation will accumulate within the population and can potentially replace the original allele and reach fixation. Detrimental genes can be outcompeted however there are also interesting effects where a small level of the allele occurrence is retained within the population. Additionally, mutations can accumulate without affecting the overall fitness, which is often referred to as mutational robustness¹⁸.

The environment exerts selection pressures on genetic variants or organisms through natural selection. Mutations drive population variation, forming the basis of adaptation. Bacteria with high fitness will exhibit increased replication rate, influencing allele frequencies based on the principle of fitness-driven selection

(**Figure 1**). For example, in an infinite population without selection pressures the occurrence of all alleles are theorized to be evenly distributed. In such a case any variation will be defined by the rate of mutation, leading to a stochastic distribution. Regardless, an infinite population without selective pressures will not demonstrate large variation¹⁹. Whereas, in the same scenario but with a finite population, allele frequencies will differ more significantly, due to naturally occurring mutations.

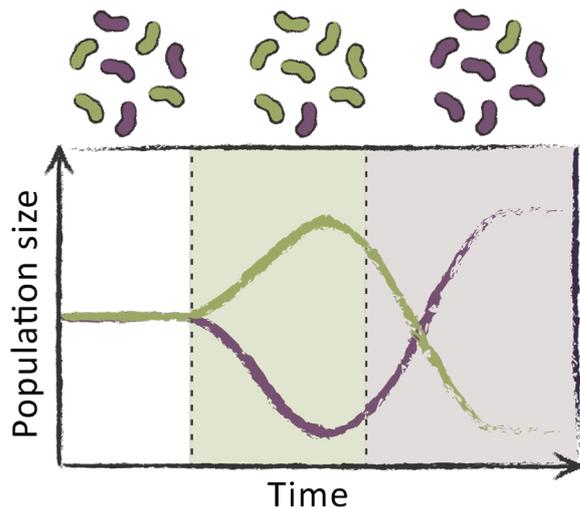


Figure 1 Selection pressures. A bacterial population is depicted to the left consisting of two alleles green and purple. The correspond distribution of alleles or growth curve is directly below. Assume zero selection pressure at the start. After some time the population encounters a selection pressure such as high temperature. This favors the green allele as it has an thicker outer membrane. The purple allele does poorly in this condition. However soon the population encounters a new environmental condition like lower temperatures, promoting the fitness of the purple allele. The oversimplified genetic variability and selection pressures depicted in this figure allows for the population to survival as a whole.

Populations that cycle between environments or have experienced a severe challenge often display an increased mutation rate and maintain mutator alleles. Such organisms are referred to as mutators, often experiencing a higher load of deleterious mutations²⁰. This is because the probability of deleterious mutations occurring is higher than beneficial ones. As a result, the changing environments will apply different selection pressures on the mutator allele. Just as in the presence of a stable environment, mutator alleles will eventually be lost²¹. This demonstrates how genetic variation can aid in transitioning between environment. The allele frequencies in a population will be change over generations simply by chance, this is known as genetic drift. Genetic drift is present in infinite populaitons however will have larger effects in finite populations. The genetic drift can have major effects if the population size experiences a sharp decline, this is referred to as a bottleneck event. It is thus possible for a random event to shape the genetic variability of a population²². If a random event isolates a smaller part of the population, there will be a reduction in genetic variability (**Figure 2**), which can result in “speciation”, given that the population can adapt to the selection pressures, this is known as the founder principle.

In 1968, Motoo Kimura evaluated the rate of mutation for a base change in genetic material. Using a combination of numbers based off comparative studies of hemoglobin, cytochrome C, and human sperm. He found that the mutation rate per base was so frequent that the “*substitutional load [would become] so large that no*

*mammalian species could tolerate it*²³. This implying that most mutations are likely neutral, as the substitutional load is lower and would not impede the rate of evolution. This underlines the importance of genetic drift within a finite population, where evolutionary changes arise from genetic drift acting on neutral substitutions. This principle is now known as the theory of neutrality, which maintains that most mutations have a neutral effect, and any resulting evolutionary benefit is purely due to stochastic processes. Kimura's witty conclusion adds a touch of scientific sass, emphasizing the importance of random genetic drift in finite populations. "To emphasize the founder principle but deny the importance of random genetic drift due to finite population number is, in my opinion, rather similar to assuming a great flood to explain the formation of deep valleys but rejecting a gradual but long lasting process of erosion by water as insufficient to produce such a result." Consequently, novel mutation or pre-existing variation therefore undergoes constant genetic flux, shaping the evolutionary trajectory of the population.

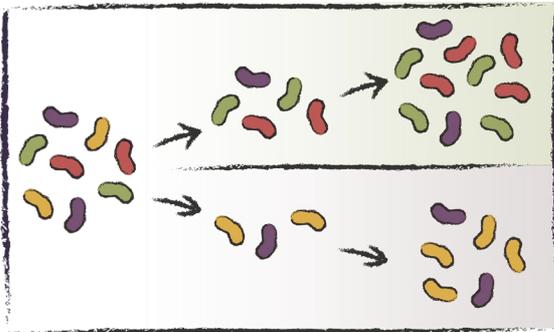


Figure 2 Genetic Drift. Schematic that demonstrates the principle of genetic drift. A bacterial population is depicted on the left. Each color equals an allele. Assume zero selection pressure between allele frequencies. The horizontal black line indicates a drift event that isolates a subpopulation. Subsequent replication and selection pressures will determine if the populations are viable and could result in speciation. A small population with little genetic variability might not be able to adapt and could perish.

4.1.2 Heterogeneity

"Same same, but different"
 – David Skylark, *The Interview* –

Beyond genetic diversity, isogenic populations exhibit notable phenotypic variation despite having an identical genetic make-up. Phenotypic heterogeneity arises from internal factors like individual physiological states and stochastic gene expression, strongly influenced by external factors such as nutrient availability, chemical gradients, cell density, and stress factors²⁴.

Stochastic processes introduce substantial variation, not necessarily stemming from genetic changes but rather from phenotypic effects. The genotype represents the hereditary underlying genetic make-up of the organism, while the phenotype is the observable trait such as morphology and behavior. Consider a marine bacterium attached to chitin, a biopolymer present in crustacean shells. The bacteria can upregulate genes involved in chitin degrading and metabolism. The genotype consists of genes enabling this, and the phenotype is the ability to metabolize chitin. In this aquatic environment, the bacterium faces several selection pressures due to

nutritional scarcity and limited protection. Being able to metabolize an abundant yet insoluble polymer provides nutrition. Further, as chitin is the main component of exoskeletons the bacterium would be attached to a larger organism, with the added benefit of protection from grazing protozoans. In this case, the bacterium has over time adapted to metabolize an insoluble polymer and developed mechanisms to aid in scavenging this nutrient source. While genetic change requires time to develop, phenotypic change occurs by expressing a different preexisting genetic code, and can therefore be more dynamic.

4.1.2.1 Phenotypic diversity

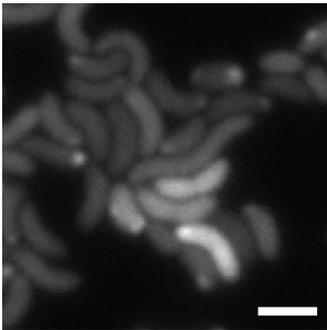


Figure 3 Phenotypic variation. An example of phenotypic variation in gene expression. *V. cholerae* expressing a GFP-fusion protein visualized in grey-scale. Note how some cells are brighter than others. Unpublished. Scalebar 2 μm

Bacteria actively monitor their local environment, switching between phenotypic states in response to environmental cues. The phenotypic state is mainly a representation of expressed genes of translated proteins. **Figure 3** demonstrates phenotypic variation in terms of protein expression. The transition between phenotypic states can be influenced by local environmental shifts, such as concentrations of secreted signaling molecules. This regulatory mechanism operates at both the single-cell and population levels.

Quorum sensing (QS) exemplifies such a phenomenon, where the bacteria regulates expression based on cell density. Autoinducers (AIs) are secreted by most bacteria, including *V. cholerae*, and the accumulating concentration translates into a phenotypic response. Once a threshold is reached, it triggers a signaling cascade that can result in various effects. Generally, in terms of QS there is an over-all repression of replication, which regulates the population density²⁵. However, QS can also affect virulence and, biofilm formation, as well as DNA uptake. In response to an unfavorable environment, bacteria can form biofilms. Biofilms are clusters of bacteria held together by a matrix that can consist of sugars, proteins, and DNA. This state often allows the organism to withstand longer periods in unfavorable environments after for example shifts in temperature, pH, salinity and nutrient availability²⁶. *V. cholerae* is a highly adaptable bacterium and utilizes quorum sensing in multiple ways. At low population densities, the concentration of AI is below the “threshold” promoting biofilm formation, while high concentrations repress biofilm, in favor of dispersal²⁷. Virulence factors and DNA uptake are also cell density dependent^{28,29}. Intriguingly, *V. cholerae* can also rally other bacteria by using autoinducers to promote virulence in other species, such as *E. coli*²⁸.

A major contributor of this is the inherent variability of stochastic gene expression, or noise as it is often referred to³⁰. Transcription of a gene is the process of expressing

a part of the genetic material. Expression is initiated at a promoter sequence, and different promoter sequences will have different expression levels. Translation is the process of “translating” the genetic code into a protein. Even in the same environment individual bacteria of a clonal population will exhibit distinct phenotypes. Both transcriptional and translational noise can result in different concentrations between cells within a population³¹. Random fluctuations of mRNA and protein expression are heavily influenced by this noise³². Variation in intracellular concentrations and binding rates of RNA polymerases, transcription factors, etc., result in fluctuating transcription and translation levels. Additionally, stochastic promoter activity and RNA degradation contribute to variation, even between identical promoters within a cell³³. Additionally, transcription and translation often occurs in bursts which further promote cell-to-cell variability^{34,35}. For example transcriptional bursting, despite its purpose being unknown, contributes to cellular diversity which can aid survival in unpredictable environments³⁶. Population diversity allows for alternative ways to adapt to fluctuating environments, meaning different subsets of the population could be better adapted to different environments³⁷

4.1.2.2 Metabolic heterogeneity

Population diversification can also arise from stochastic metabolic processes³⁸⁻⁴⁰. Inherent molecular variation during gene/protein expression will further promote noise. The stochastic nature of biochemical reactions impacting gene expression will influence the fate of the cell. Heterogeneity has been observed in all forms of bacterial metabolism ranging from central carbon metabolism⁴¹ to amino acid pathways^{42,43}. Biological processes are concentration dependent; and a variety of molecular processes result in heterogeneity (**Figure 4**). Cell division for instance, involves asymmetric partitioning of randomly distributed intracellular components, resulting in daughter cells with different concentrations. Inheriting a complex protein structure provides a fitness advantage, reducing the energy expenditure on protein synthesis⁴⁴. Further there is particularly high noise for transcriptional promoters driving metabolism and stress responses, such as pH, acid, and temperature to name a few. Whereas, promoters that steer gene expression pertaining to sensitive processes e.g. housekeeping genes and remodeling the structure of DNA (methylation and super-coiling), have less noise⁴⁵.

Bacterial populations can readily switch between metabolic networks. Further, bacteria will deal with nutrient limitation and uptake according to the needs of the individual cell²⁴. In the presence of glucose *E. coli* and other bacteria will inhibit the uptake of other carbon sources, this is known as carbon catabolite repression⁴⁶. Once the glucose is used up, the population will switch to another available carbon source, e.g. lactose. This is a classic case of feast-to-famine during growth with two

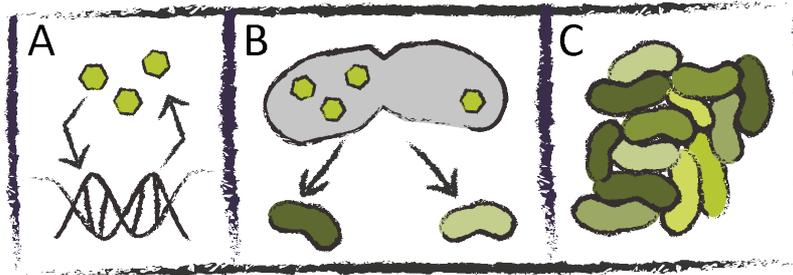


Figure 4 Molecular processes aiding heterogeneity. Examples of molecular processes that generate heterogeneity. (A) Transcriptional activity and feedback loops may result in different expression levels of genes. Promoters of the gene will also play a role for how strongly a gene is expressed. (B) Asymmetric partition can result in an uneven distribution of proteins. (C) Heterogeneity is the resulting sum of molecular processes and environmental influences.

nutrients. During the switch, there is a delay in overall growth for the population. However, a fraction of the bacteria will not experience this growth lag and can start growth immediately. Another fraction experiences a long lag phase, while other bacteria might enter a growth arrest eventually being outcompeted or simply endure at lower levels in the population. This heterogeneous growth response is the result of stochastic biological processes⁴⁷. While bursts of expression can occur, the bacteria need to overcome an expression threshold for the phenotype to remain stable. All genes involved in metabolizing lactose are grouped in an operon called *lac* (An operon is a cluster of genes that are transcribed together). While most of the population represses the *lac* operon, a fraction will have a level of *lac* expression enough for a few molecules of *lac* permease. A permease is a pore that allows transport across the membranes. This allows entry of lactose which in the absence of glucose, generates a positive feedback loop that eventually results in maximal *lac* expression⁴⁷. This low baseline of expression would present itself as a trade-off between fast switching and a corresponding fitness cost while glucose is in abundance. Naturally, stochastic processes are not immune to selection pressures. However, the fitness cost of a biological process can be alleviated by the benefit of the outcome or compensatory mutations. Regardless expressing a system that is not in use is costly and will decrease the overall fitness, especially as fast growth is important when competing for a limited resource. Stochastic processes can therefore have a major impact during this race for resources, as this limits the fitness cost to a fraction of the population. The intricate balance between response and cost within a population then becomes a distribution of risk.

Bacteria are constantly balancing the trade-off between endurance and growth. Planktonic cells exhibit rapid growing but are fragile, while biofilm cells often grow slower yet are robust. As bacteria readily adapt to the environment, environmental factors could impact heterogeneity. Heterogeneity may arise from either responsive or stochastic switching. Responsive switching refers to an environment-dependent response, which would maximize temporal fitness. On the other hand, stochastic

switching refers to a random change in gene expression independent of the environment. The latter would yield multiple maladapted phenotypes of reduced fitness. However this strategy might be better suited in facing unknown environments⁴⁸. This particular form of heterogeneity is known as “bet-hedging”.

4.1.2.3 Bet hedging strategy – spread out the risk

Constant risk factors like scarce nutrition, predation, and hostile conditions make it advantageous to be able to switch between phenotypes. Being too well adapted to one environment may limit survival if conditions change. Bacteria that can modulate their response to environmental change quickly, have been found to have a higher fitness than slow responders³¹. Additionally, heterogeneity enables the lineage of a population to survive in fluctuating environments, given that it is cyclic and not irregular. It can therefore be viewed as an evolutionary “strategy” to avoid a worst-case scenario by distributing the risk throughout the population. The isogenic population introduces stochastic diversification within the pool of phenotypes, leading to varying fitness effects. This phenomenon can offer selective advantages, particularly in dynamic habitats. For instance, when the nutrient source is switched, some cells undergo growth arrest as an adaptive response⁴⁹. Subsequently, these dormant cells face competition from an expanding subpopulation that has the ability to metabolize the new nutrient source⁴¹. This dynamic interplay highlights the evolutionary strategies employed by isogenic populations to adapt and thrive in changing environmental conditions.

Adaptation can either be temporary or permanent. Temporary adaptation to an environmental response can occur by regulating the expression of certain genes to achieve a phenotypic response. An example of temporary adaptation is the ability of orally ingested pathogens, such as *V. cholerae*, to survive gastric acid⁵⁰. This adaptive stress is further suggested to aid initial survival within the intestine⁵¹, promoting positive selection of the adaptive mechanism. Whereas permanent adaptation refers to a lasting change, for example, genetic exchange via viral transduction that increases infectivity and virulence aiding in dissemination of the pathogen. An example of this is the emergence of the toxigenic *V. cholerae* strain after phage (CTX ϕ) interaction, where the event led to the cholera toxin gene being incorporated into the genome⁵². Given the impact horizontal gene transfer (HGT) has on environmental adaptation and spread of antibiotic resistance markers, it is important to understand the machinery that enables the constant flux of genetic material.

4.1.3 Horizontal gene transfer

“The first million is the hardest to make. So, start with the second million.”

- Arnold Schwarzenegger -

HGT can lead to permanent adaptation and there are clear indications that it is a major driver of evolution, exemplified by *E. coli* and its acquisition of indole production and lactose utilization. Most changes in metabolic networks stem from acquiring genes via horizontal gene transfer. Bacterial populations are more often than not experience environments lacking nutrients. Acquiring new metabolic capabilities allows them to utilize a broader range of nutrients, without affecting their core metabolism⁵³, making them stronger contenders. There are three main modes of HGT: conjugation, transduction, and transformation. While the mechanisms differ greatly in transfer strategies, the overall effect is the same i. e. transport the the genetic material across the bacterial envelope and incorporate it into the genome. *While this part of the introduction will focus on transformation, the three modes are discussed further in section 2.2.4.2, within the context of antibiotic resistance.*

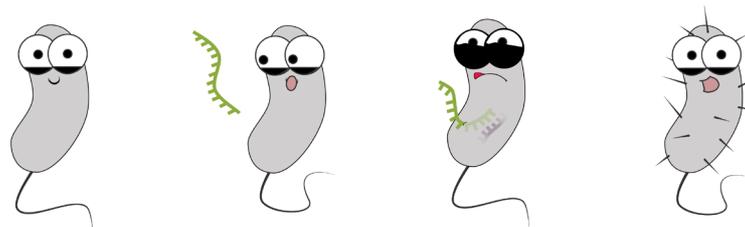
4.1.3.1 Conjugation

In brief, conjugation is a cell-to-cell transfer of a conjugative circular plasmid. The mechanism occurs via a retractable conjugation pilus. The pilus binds to a recipient cell in close proximity. A pore is formed within the membrane of the recipient anchoring the mating bridge. Before transfer can occur the conjugative plasmid DNA is nicked. A single stranded DNA (ssDNA) is then transferred from the donor to the recipient, the complementary sequence is synthesized, and the DNA re-circularizes. The recipient is now a potential donor⁵⁴.

4.1.3.2 Transduction

Transduction occurs via bacteriophages, and occurs by direct injection of the genetic material into the host cell. Upon packaging the viral DNA some of the host DNA can potentially be included. This will then be transferred to the next host, resulting in transduction. Lytic phages for example result in the release of new infectious virions which would then propagate further transfer of the incorporated host DNA.

4.1.3.3 Natural transformation



"Time may change me, but I can't trace time"

– David Bowie –

Transformation differs from conjugation and transduction as it does not require contact with other cells. Competent bacteria can take up exogenous DNA and incorporate it into its genome. While this cuts out the middleman, so to say, there are some potential drawbacks for integrating foreign DNA, such as the uptake of deleterious genes. The likelihood of integrated extracellular DNA having a beneficial effect although present, is low. However, the advantage of recombination in the absence of sexual reproduction could outweigh this drawback, especially if the population is large enough to overcome the fitness interactions of "bad genes"^{55,56}. This could suggest an explanation for why bacteria have evolved complex machinery to take up DNA from the environment. This process is known as natural transformation and often requires coordination of several signaling pathways. Competent bacteria can take up free DNA naturally occurring in the environment, allowing access to a vast genetic pool and enabling transfer across species and strains. The core machinery for natural transformation is fairly conserved across Gram-positive and -negative bacteria. The induction of competence, however, is more specific to species and situations. Some bacteria are constitutively competent and often require specific DNA uptake sequence (DUS), as is the case for *Neisseria gonorrhoeae*⁵⁷. Whereas, other bacteria are transiently competent and requires induction of competence. For instance, competence in *V. cholerae* depends on chitin metabolism, allowing for spatiotemporal isolation of the DNA uptake event⁵⁸.

There is a wide range of DNA uptake machineries and strategies, which implies that the purpose of transformation could very well be different from species to species. While there is evidence of transformation being a driver for evolution the purpose of transformation is debated. Natural transformation is a costly mechanism and requires orchestration of several pathways. It has therefore been suggested that transformation is maintained as it provides a way to utilize DNA as a nutrient or source of nucleotides, providing a slight increase in fitness. However, this could be species-dependent, appearing to be more relevant for *H. influenzae* and *B. subtilis*. In contrast, adding exogenous DNA to *Acinetobacter baylyi* reduces the growth rates of competent bacteria, removing any apparent advantage. However, this does provide the bacteria more time to integrate the DNA into its genome. Furthermore, the most competent bacteria protect DNA from degradation either by downregulating

the secretion of nucleases or by protecting the DNA strand with binding proteins once shuttled into the cell. It is, however, possible that the transforming DNA could be recycled as a nutritional byproduct. An alternative hypothesis suggests that natural transformation aids DNA repair, rendering the integration of exogenous DNA purely coincidental. Meanwhile, another hypothesis suggests that natural transformation is maintained as a route to sample fitness landscapes in neighboring populations⁵⁹. Another possible selection pressure could be the additive effects of competence-induced growth arrest and occasional selection, leading to increased fitness that would eventually outcompete rapidly dividing, non-transforming bacteria. Population heterogeneity could support this theory.

Different bacterial species implement natural transformation with slight variation. For instance, some take DNA indiscriminately, whilst others require a specific sequence. Regardless of flavor and spice, there is some commonality in the primary machinery. In the case of Gram-negatives, a type IV pilus recognizes the DNA, be it double or single stranded, and drags it into the periplasm, where specific proteins bind the DNA, passes the outer membrane via a pore complex. From this point on only one strand is shuttled into the cytoplasm, where single-stranded DNA binding proteins stabilize the strand and protect it from degradation. The DNA can then be integrated into the genome. Gram-positives function very similarly, however they do not have an outer membrane and therefore pull the DNA directly into the cytoplasm via a pore complex in the inner membrane⁶⁰. Once inside the cell, the DNA can be incorporated into the genome to persist in successive generations. While there are several mechanisms for this, homologous recombination is the most efficient and straightforward mode of integration.

4.1.3.4 Homologous recombination

Natural transformation tends to favor linearized ssDNA, and the key to homologous recombination is sequence similarity, as transformation frequencies drop dramatically with increasing sequence divergence. Upon entry into the cell, the ssDNA is bound to single stranded binding proteins, SsB, which keep the DNA from degrading. Depending on your favorite organism, DprA, a DNA processing A protein, either replaces the SsB or binds directly to the ssDNA ultimately recruiting RecA. RecA is a universally conserved DNA repair and maintenance protein that catalyzes strand exchange and is responsible for welding the transforming DNA to the chromosomal DNA resulting in a heteroduplex. The heteroduplex segregates during replication and cell division, yielding a “wild-type” and a “transformed cell”. The heteroduplex is therefore the equivalent of a cellular “save” function.

There are also low frequency recombination events known as illegitimate recombination, where only a short stretch of a few hundred base pairs is enough to serve as an anchor and initiate a RecA-dependent strand exchange. While illegitimate recombination required little sequence similarity it is also possible to

recombine DNA sequences with no sequence similarity as seen in *Acinetobacter baylyi* known as double illegitimate recombination. *This is outside the scope of the thesis.*

4.1.4 *V. cholerae* transitions between environments

Natural transformation promotes the spread of antibiotic resistance⁶¹, but also drives the emergence of new toxigenic strains, as demonstrated in *V. cholerae*¹¹. *V. cholerae* is a human pathogen well-adapted to the human body. Ingestion of the bacteria can lead to the debilitating diarrheal disease known as cholera. While it is difficult to ascertain the true scope of the health risk imposed by the pathogen, the current estimate lies between 1 to 4 million infections per year leading to between 21,000 and 143,000 deaths⁶²⁻⁶⁴. *V. cholerae* is a historic pathogen, without geographical boundaries, and can therefore be found wherever there are people. While there have been outbreaks across the globe, today's outbreaks are concentrated within communities that either lack proper, or experience disrupted sanitation. Outbreaks often follow natural disasters and warfare, further affecting already vulnerable communities.

We are currently experiencing the seventh pandemic, which is the longest lasting so far. There are over 200 serogroups of *V. cholerae* in which there is great genomic diversity within environmental strains however pathogenic strains isolated from patients are relatively clonal. Since the 1800s there have been seven documented pandemics. While the previous pandemics are thought to be caused by O1 strains, the fifth and sixth pandemic were primarily caused by the classical O1 biotype. The current seventh pandemic is caused by El Tor biotype O1 strains. The seventh pandemic began in 1961, making it the longest lasting cholera pandemic to date. The commonality for both toxigenic and pandemic strains are the CTX Φ prophage and the *Vibrio* pathogenicity island. It has been suggested that these toxigenic features were acquired while in the presence of chitin, as this promotes the uptake of DNA.

Standard treatment for *V. cholerae* is oral rehydration therapy, and for severe cases the treatment is supplemented with a course of antibiotics. Antibiotics are recommended in acute cases where massive water loss and severe dehydration is ongoing, for pregnant women, and patients with comorbidities such as severe malnutrition or HIV infection. The antibiotic treatment shortens the duration of the disease with the added benefit of releasing fewer infectious bacteria in the stool. However, there is a steady increase in treatment failure due to multidrug resistant and extensively resistant strains, which is further complicated by HGT and highly mobile genetic elements conferring multi-resistance^{65,66}. The rise in multi-drug resistant bacteria has been documented since the 1970s and there are cases found that are resistant to tetracycline, ampicillin, kanamycin, streptomycin, and trimethoprim-sulfamethoxazole. There are also recent reports of chloramphenicol, furazolidone, and sulphonamides resistance. There is still little information existing in

regards to antimicrobial resistance profiles of clinical samples, mostly due to lengthy culturing for antibiotic susceptibility testing and limited laboratory space. These profiles are simply too time consuming and labor intensive.

4.1.5 Infection cycle

During the infection cycle *V. cholerae* is exposed to very different environmental conditions (**Figure 5**). This is most evident during the transitions between host and aquatic habitat. However, conditions within the host vary considerably and the pathogen has evolved mechanisms to sense and react to the broad spectra of these changes. For example, upon uptake from the aquatic environment the bacteria first encounter the acidic environment of the stomach. *V. cholerae* manages to shield itself from the environment in part due to its ability to form biofilms⁶⁷. Biofilm associated cells have been found to be more resistant to stomach acid and increases survival rates during passage through the gastric environment⁶⁸ (**Figure 5a**). Upon entering the small intestine, host-secreted bicarbonate increases the pH (**Figure 5b**). At the same time bile is secreted, which is a strong chemotactic repellent for *V. cholerae*. Under such conditions biofilms are no longer beneficial for the organism. Concomitantly, flagella-dependent motility is favored and chemotaxis assists in penetrating the mucus layer in order to colonize the epithelial cells^{69 70}. While biofilms are believed to be a hindrance for initial colonization, they appear to be upregulated during the late stages of the disease. In freshly shed diarrhea from human cholera patients, the majority of the pathogen is found in biofilm-associated forms, which – together with the shed mucus – gives the stool its characteristic “rice water like” appearance. It has been speculated that this might prepare the bacteria for its release from the host and exposure to the new, hostile environment, but also for potential uptake into the next host⁷¹. While it is theorized that the pathogen can undergo transformation in the host intestine, it is established that natural transformation occurs in the aquatic environment.

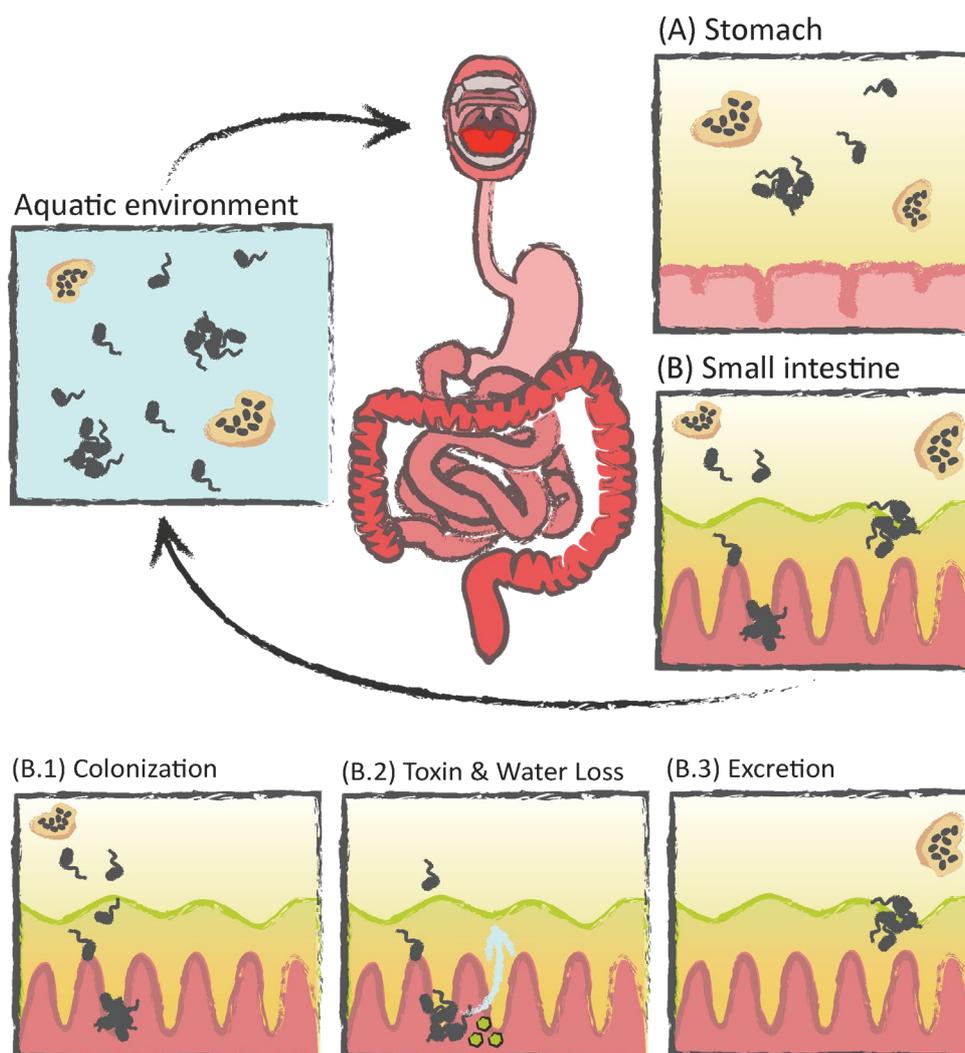
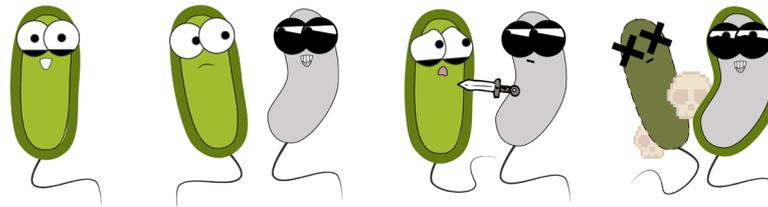


Figure 5 Infection cycle of *V. cholerae*. In its natural aquatic ecosystems *V. cholerae* occurs either as motile planktonic cells or in a sessile form attached to biotic or abiotic surfaces. The facultative pathogen is transmitted to humans via ingestion of contaminated water or food. Once consumed the bacteria transits through the gastrointestinal (GI) tract, past the acidic environment of the stomach (A), to the site of colonization. In humans, the main site of colonization is believed to be the small intestine, where it attaches to the host's epithelial cells and proliferates (B). To reach the epithelium the bacteria has to penetrate the thick mucus layer that lines the GI tract (B.1). It has been proposed that this occurs via a combination of regulated motility, chemotaxis and mucolytic enzyme secretion¹. Eventually ToxT, the master regulator of virulence in *V. cholerae*, is activated². ToxT is important for the expression of the toxin co-regulated pilus (TCP) and cholera toxin (CT). The pilus is important for host colonization, while the secreted cholera toxin binds to the host's cells and increases the concentration of cyclic AMP (cAMP) which causes an efflux of chloride ions and a decreased influx of sodium ions (B.2). The electrolyte imbalance leads to massive efflux of water through the intestinal cells, causing the typical diarrhea that is linked to cholera³. *V. cholerae* exits its host with the diarrhea and reenters the aquatic environment. Especially at the late stage of the infection, this exit is aided by a coordinated "mucosal escape response" in which the bacteria actively detach from the epithelial surface and move into the fluid-filled lumen of the gut⁴ (B.3). Once expelled the pathogen may then restart the infection cycle.

4.1.6 *Vibrio cholerae* is adept at natural transformation



"I don't leave anything up to chance, I make my own luck."

– Harvey Dent, The Dark Knight –

V. cholerae is an aquatic strain often associated with chitin since it provides protection from grazing protozoans and functions as a carbon source. The biopolymer can be found on the shell of crustaceans and zooplankton. Additionally, chitin metabolism induces natural competence in *V. cholerae*⁵⁸. This state is tightly regulated and distinct in both time and space. Association with chitinous material upregulates the expression of the master regulator of competence, TfoX⁷². The chitin-dependent production of TfoX works in unison with carbon catabolite repression (CCR) and is co-regulated by cell density in the form of QS^{73,74}. CCR enables bacteria to regulate their metabolism to better adapt to their environment. In the case of *V. cholerae* this could be transitioning from the nutritious lumen of the intestine to an aquatic chitin containing environment. A major player for CCR is the phosphoenolpyruvate:carbohydrate phosphotransferase system (PTS), which, mediates sugar uptake across the membrane. This eventually leads to an increase in cAMP which is a secondary messenger or signaling molecule. The elevated cAMP levels are eventually recognized by the cAMP-binding protein, CRP. Along with TfoX, the cAMP-CRP complex induces the expression of natural transformation machinery genes. TfoX is the major regulator of natural transformation⁵⁸. However, there is also a need for input from quorum sensing genes via HapR, a regulator that only accumulates at high cell densities. An important function of HapR is the down regulation of a DNA endonuclease, Dns. The bacteria produce two extracellular nucleases, Dns and Xds, Dns being the most efficient and having the most impact on transformability. Exogenous DNA would be consistently degraded if production and excretion of the exonuclease was not shut down⁷⁵. This coordinated response also leads to a dramatic increase in the type VI secretion system (T6SS). T6SS is equivalent to a poison dagger. The bacteria are constantly stabbing their neighbors, injecting toxins. This can result in killing by either attacking the nucleic acid, membrane, or cell wall. If the compound is not neutralized the prey will likely lyse. The main effectors of *V. cholerae* Vrg3 and VasX attach the cell membrane and cell wall. These effects can also kill eukaryotic cells by cross-linking actin structures^{76,77}. *V. cholerae* can implement this killing device to access prey DNA. The mobilized attack lyses the cells, releasing the DNA, which can then be taken up by the natural transformation machinery⁷⁸.

4.1.6.1 Natural Transformation Machinery

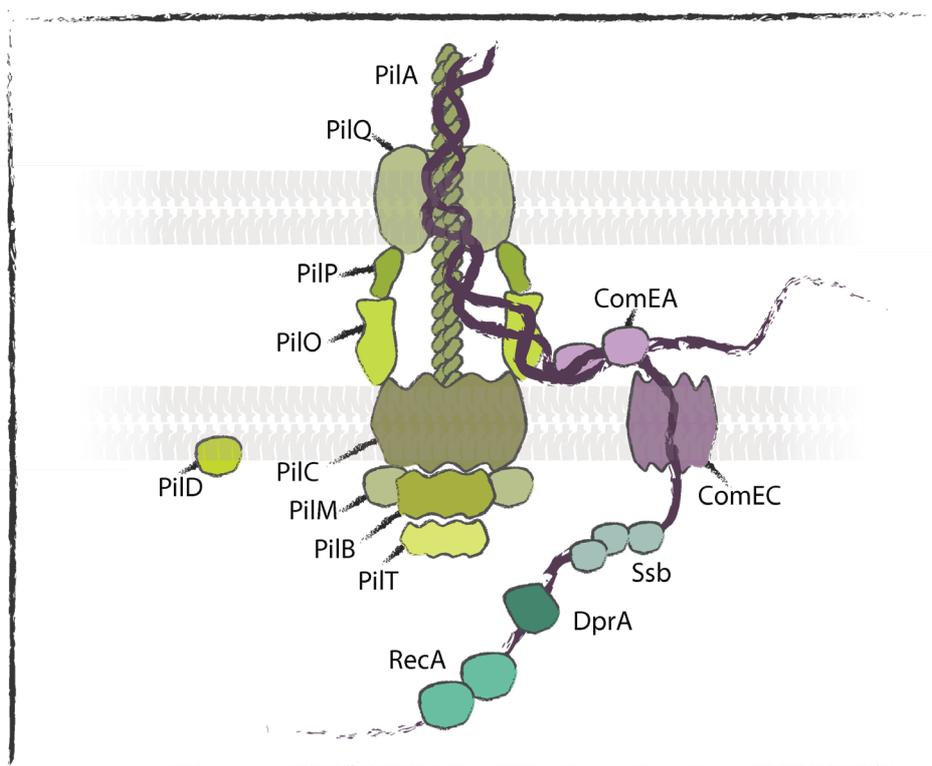


Figure 6 The natural transformation machinery of *V. cholerae*, current model. Prepilins are processed by the peptidase, PiID, and incorporated into the growing transformation appendage. the major pilin subunit is PiIA. the growing pilus is energetically driven by the ATPase PiIB. the pilus extends through the outer membrane pore complex PiIQ, which is possibly stabilized by PiIW. PiIT drives the retraction of the pilus. DNA is pulled into the periplasm and binds ComEA. the DNA is then shuttled into the cytoplasm via an inner membrane channel ComEC. once in the cytoplasm Ssb and DprA binds the ssDNA and facilitates RecA-dependent recombination with the chromosomal DNA of the bacteria.

References for the schematic ^{60,79,80}

Within the aquatic environment, *V. cholerae* can associate with chitin present on various biotic surfaces such as zooplankton and crustaceans. The transmembrane regulator ChiS senses the presences of chitin and induces the expression of competence genes including the chitin utilization genes, the competence pilus, and the type 6 secretion system. This also induces TfoX expression which regulates the expression of pil genes involved in the uptake of DNA. The natural transformation machinery consists of two main parts: a competence pilus, and an inner membrane transporter (**Figure 6**). The competence pilus is a type IV pilus. It is anchored in the inner membrane and extends through the outer membrane. This is the appendage that associates preferentially with dsDNA and pulls it into the periplasm. Once inside the periplasm, the DNA is bound by DNA binding proteins ComEA⁸¹, which shuttle the DNA through the inner membrane transporter ComEC. It is assumed that this occurs as single strand DNA. Once inside the cytoplasm, the DNA is protected from degradation by Ssb, and DprA, which then facilitates RecA loading. RecA is the main recombinase of homologous recombination. Upon finding a homologous region on the chromosome a transformation heteroduplex is initiated. The ssDNA can then be

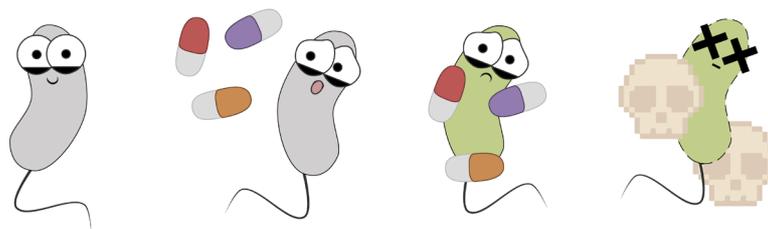
incorporated into one of the strands of the chromosomal DNA and inherited by one of the daughter cells after cell division. Horizontal gene transfer can also result in nongenetic, transient inheritance of resistance to antibiotics⁸². Only one cell will inherit the genetic material. However, expression of the newly integrated DNA can occur immediately. Due to the stochastic nature of cell division (**Figure 4**), the newly expressed protein will be present in both divisional cells. Allowing survival of both. This provides the benefit of transformation without the potential consequence of a genetic fitness cost.

It bares mentioning that isolates from the Haiti outbreak in 2010 contained a conjugative element from the SXT family. This confers resistance to several antibiotics such as: sulfamethoxazole, chloramphenicol, trimethoprim, and streptomycin. Additionally, a gene within this element inhibits natural transformation by expressing an ICE-encoded DNase which localizes in the periplasm⁸³. This DNase would remove the evolvability of the bacteria that would not be beneficial in the long term. However, transmission of pandemic strains is likely to occur between households and not necessarily from ingesting chitin particles carrying the pathogen. That being said, an efficient cholera prevention method is filtering drinking water through a piece of cloth⁸⁴. In fact, this method was found to be so effective that it not only reduces transmission within households that implement the filtration, but also shields neighboring households, purely by limiting the number of cases. This indicates a balance within endemic areas and house-hold transmission. Regardless, a situation where inter-household transmission is the main source of infection could reduce the need for adaptation, as the environments human-to-human are more constant than human-nature-human, reducing the need for adaptation. Of course, it is much more likely that the nuclease is a hitchhiker that is inadvertently selected along with the antibiotic resistance conferred by the SXT element, as presented within a range of isolates acquired in Bangladesh collected in the timespan of a decade, from 2001 to 2011. These strains exhibit an initial prevalence of strains containing IdeA, the first five years. Followed by a transition to strains lacking IdeA, which could then be in favor of transformation. The connectivity of HGT modes and antibiotic selection pressures brings a complexity to pathogenesis that requires more investigation.

“Laughter is the best form of medicine, unless you have diarrhea”

– Source unknown –

4.2 Antibiotics



4.2.1 The pre-antibiotic era

Historically, infections were thought to be a result of miasma. After serious debate and substantial evidence, germ theory was finally accepted and well-established in the early 1900s. While this led to better prevention of disease, there was still a need for effective treatment. The discovery of antibiotics improved quality of life considerably and has saved countless lives from various diseases. Infections we today view as trivial were in the pre-antibiotic era deadly. Strep throat for instance could result in heart failure. Bacterial meningitis was deadly. Ear infections led to hearing impairments, and if spread to the brain caused severe damage. Not to mention the social stigma of disease in a time where infection was unanimous with “punishment from God”. For no other disease, was this truer than for leprosy. The disease caused by *Mycobacterium Leprea*, affecting the skin, and nerves, among other sensitive organs such as the eyes, causing permanent damage if untreated⁸⁵. Survivors of not only leprosy, but countless other diseases, were almost guaranteed a participation trophy presented as a deformity, which guaranteed ostracization. While the development of vaccines play a huge role in safeguarding us from serious disease, as well as provide the possibility for eradication, antibiotics were a genuine game changer. There was finally an efficient weapon against bacterial infections. See **Table 1** for an overview of examples of antibiotics and their main targets.

Where would we be today if a fungal contaminant, *Penicillium notatum*, hadn't landed on an agar plate containing the pathogenic bacterium *Staphylococcus aureus*? By implementing simple and effective antibacterial assays Alexander Fleming noticed, in 1928, the production of a potent substance that inhibited growth of the bacteria. He had discovered the first antibiotic known to modern day medicine. Fleming was of course not the only person involved in the modern day penicillin story. In 1870, Sir John Scott Burdon-Sanderson discovered that mold could inhibit bacterial growth. Joseph Lister detected antibiotic effects on human tissue, in 1871. In 1902, a team of scientists lead by Paul Ehrlich discovered the first known antimicrobial compound of modern day medicine, salvarsan. Which in contrast to naturally occurring substances such as antibiotics, was an arsenic-based chemical effective against syphilis⁸⁶. *There are simple too many contributors to list.* Regardless, a decade after Fleming first discovered the inhibitory properties of *Penicillium notatum*, Chain and Florey manufactured the drug from liquid culture, allowing for large-scale treatment of

bacterial infections^{87,88}. It should be noted that early applications of treating various ailments using moldy bread and soil dates back centuries⁸⁹⁻⁹¹. Additionally traces of antibiotics have been detected in skeletal remains at several archeological sites in Egypt⁹², Sudan^{93,94}, and Jordan⁹⁵. Nevertheless, the discovery of penicillin initiated a hunt for molds and fungi from which one could isolate novel antibiotics. One such story is that of Mattiedna Johnson, who observed just such mold from a jar of soup⁹⁶. She decided to test the spores against a specimen of *Streptococcus haemolyticus*, the causative agent of scarlet fever, a contagious infection primarily affecting children. Upon discovering that the spores were effective at killing the bacteria, she suggested it be mixed with a peppermint syrup to make it more palatable for children infected with the disease. The substance was later patented as terramycin. This is only one of many antibiotics resulting from the hunt for mold, and several more were to come. The serendipitous discovery of antibiotics initiated international collaboration and was the break-through of a century.

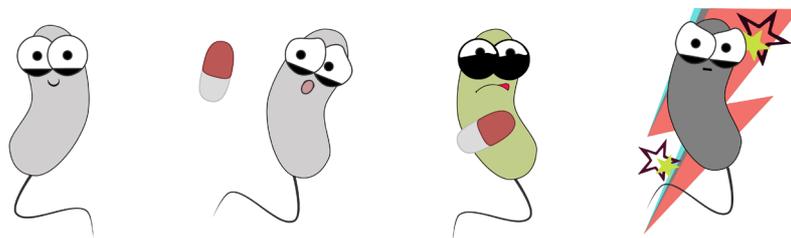
Table 1: Examples of targets of antibiotics⁹⁷⁻⁹⁹:

Mode of action	Antibiotics class	Main targets
DNA replication	Quinolones	Topoisomerases – results in blocking DNA replication
RNA synthesis	Rifamycin	RNA polymerase only when DNA-bound and actively-transcribing
Cell wall synthesis	β -lactams	Blocks the cross-linking of peptidoglycan (PG), via penicillin-binding proteins
	Glycopeptides	Inhibits PG synthesis through binding with PG units
Protein Synthesis	Macrolide, lincosamide, oxazolidinone etc.	Ribosome function, blocking either initiation of complex or translation
	Aminoglycosides	Blocks protein biosynthesis
Metabolic pathways	Trimethoprim, Sulfonamides	Folic acid synthesis
	Isonicotinic acid hydrazide	Mycolic acid synthesis

4.2.2 Antibiotic resistance

“And I’ve been putting out the fire with gasoline”

– David Bowie –



No sooner had antibiotics become common practice than multi-resistant strains became commonplace. In 1940, Ernst Chain demonstrated that some Gram-negative bacteria could destroy penicillin. He named the substance “penicillase”, now known as “ β -lactamase”. The occurrence of treatment failure due to resistant bacteria was increasing. Fleming seeing the fly in the ointment, predicted that these penicillin resistant strains would only become more prevalent in accordance with the theory of natural selection. Stating *“there is probably no chemo-therapeutic drug to which in suitable circumstances the bacteria cannot react by in some way acquiring ‘fastness’ [resistance]”*¹⁰⁰. The prediction came to fruition as soon as penicillin became widely used.

Although antibiotics have considerably improved treatment and health there is a significant increase in antibiotic resistance among common pathogenic bacteria. A recently published systematic analysis evaluated the global burden of antimicrobial resistance (AMR) for 2019⁵. The authors find 1.27 million deaths were a direct result of AMR, additionally estimating that a total of 4.95 million deaths occurred in association with AMR, meaning deaths directly and indirectly associated with AMR. Indirect cases include, but are not limited to, pre-existing morbidities such as AIDS/HIV and malaria. For context, that would remove just about the entire Norwegian population, currently populated at ~5.5 million.

4.2.3 Drug discovery challenges

Aiming to tackle the “silent pandemic”, the World Health Organization (WHO) implemented a global action plan, in 2015. The over-all goal is to mediate and ensure global progress in reducing resistance, with commitments from both G-7 (political) and G-20 (economic) groups. The global action plan has five main objectives¹⁰¹:

- To improve awareness and understanding of antimicrobial resistance.
- To strengthen surveillance and research.
- To reduce the incidence of infection.
- To optimize the use of antimicrobial medicines.
- To ensure sustainable investment in countering antimicrobial resistance.

It should be mentioned that global AMR increased substantially after the COVID-19 pandemic. Rochelle P. Walensky, the director of the center of disease control (CDC) in the U.S., stated: "... as the pandemic pushed healthcare facilities, health departments, and communities near their breaking points in 2020, we saw a significant increase in antimicrobial use, difficulty in following infection prevention and control guidance, and a resulting increase in healthcare-associated, antimicrobial-resistant infections in U.S. hospitals"¹⁰². Poor surveillance, improper use, and outdated dosing regimens all aid in increased AMR.

4.2.3.1 Golden age of antibiotics

The period of 1940 to 1960 is known as the golden age of antibiotic discovery. Initial rigorous screening of soil-derived bacteria (mostly *Actinomycetes* spp.) against susceptible organisms, resulted in the antibiotic streptomycin as well as a framework for antibiotic discovery. This spurred on the discovery of drugs such as vancomycin, clindamycin, rifampin, tetracycline, and daptomycin(ref). All of which were the product of mining soil-derived compounds. About half of the commonly used drugs today were discovered during this era.

Novel drug discovery stagnated after the 1960s. However, while drug discovery faltered, drug design advanced. The pharmaceutical industry began to tailor pre-existing compounds, resulting in greatly improved efficacy. Efficacy refers to the ability of an intervention, in this case a drug, to produce the desired beneficial effect. Unfortunately, the current economic model deems drug discovery and design as simply unsustainable, even with the modern advances in rational drug design and automated high throughput screening. The current estimated cost of developing antibiotics is around US 1.5 billion, with an average revenue of US 46 million per year^{103,104}. Essentially, turning a profit takes decades, meaning the high-risk-to-low-reward ratio simply does not justify the investment for pharmaceutical companies. Especially when disease or chronic conditions that require lengthier treatment periods, have projected sales of US 500 million to over 1 billion¹⁰⁵. Essentially, drug discovery in the looming reality of resistance is costly and inefficient. However, despite drug discovery being bad business, there is substantial evidence of inappropriate use of antibiotics, where about half the prescribed antibiotics are unnecessary, and at times incorrect¹⁰⁶⁻¹⁰⁹. Further, there is accumulating evidence suggesting that there is substantial room for improvement of antibiotic treatment regimes¹¹⁰⁻¹¹². We are to this day counter-selecting susceptible bacteria while promoting survival of resistant strains. Additionally, overuse of antibiotics further propagate resistance.

4.2.4 Mechanisms of resistance

"I brought a lemon to a knife fight"

– The Wombats –

Bacteria have long co-evolved with organisms that produce antibacterial substances. Additionally, bacteria have an impressive genome plasticity along with a wide repertoire of defense mechanisms to overcome threats. The combination of mutational resistance and horizontal gene transfer furthers dissemination of antibiotic resistance genes and has resulted in an astounding diversity of resistance functions¹¹⁶. There are several kinds of resistance mutations. The target can be altered, resulting in reduced binding of the drug to its target, decreased drug uptake, and increased secretion of drug by upregulating efflux pumps, as well as, changes to regulatory networks such as metabolic pathways. See **Table 2** for an overview.

Table 2: Examples of antibiotic resistance mechanisms

Target repair/ target amplification ¹¹³

Target modification/ target bypass

Enzymatic inactivation/ sequestration

Increased/ decreased efflux

Biofilm formation¹¹⁴/ intracellular localization¹¹⁵

4.2.4.1 Genetic mutation – resistance

Bacteria can acquire antibiotic resistance via genetic mutations. Mutations can occur spontaneously or from exposure to mutation-inducing agents. Spontaneous mutations occur as the replication of genetic material is prone to error. Occasionally during DNA replication, a nucleotide can be mismatched, added or omitted, resulting in a mutated allele. In such a case this mutated allele can be propagated through the population via lateral or horizontal gene transfer. Thus, enabling the spread of antibiotic resistance. The frequency of spontaneous mutation depends greatly on the bacteria and the drug in question. Resistance mutations can result in a lower "fitness" compared to the wild-type situation. However, such negative effects, can be alleviated by compensatory mutations, resulting in resistant mutants persisting in the population even in the absence of antibiotics.

The effect of spontaneous mutation is a combination of deleterious mutations and metabolic cost (the fitness effect). Any variant with an increased resistance mutation should then have reduced fitness as the probability of generating a deleterious mutation is higher. However, this is not necessarily the case. The occurrence of strains with higher mutation rates (mutators) are widespread in bacterial populations, suggesting a selective advantage. Mutations that lead to either antibiotic resistance or reduce fitness cost of the resistance mechanism is more likely to appear in cells with higher mutation rates. Weak mutators can more easily cycle between environments as they adapt more readily. Further, when a mutation is linked to resistance, the advantage of being a mutator also increases¹¹⁷, indicating that the mutation rate within a population is constantly being adjusted to levels that favor

adaptation¹¹⁸⁻¹²⁰. Mutational change provides a broad range of resistance mechanisms that is further complicated by transferable genetic material.

4.2.4.2 Horizontal gene transfer – antibiotic resistance

*The mechanisms of HGT are described previous in **section 2.1.3**, therefore the following will discuss HTG of mobile genetic elements within the context of acquired resistance.*

Transferable resistance was first identified by Kunitaro Ochiai¹²¹, and Tomoichiro Akiba¹²², independently, during an outbreak of dysentery. The 1950s outbreak was caused by *shigella*. While working on clinical isolates from dysentery patients, they found *E. coli* isolates demonstrating multi-resistance similar to *shigella* isolates, as well as resistance to drugs that had not yet been used to treat the infection. The researchers found it implausible for so many mutations to have arisen spontaneously, eventually demonstrating the transfer of multi-resistance between *E. coli* and *Shigella* strains and additionally showing that it occurred via a non-reproductive rendezvous, otherwise known as conjugation. The bacteria could swap genes between chromosome(s) and plasmids. It was also hypothesized that the site of transfer occurred within human bowels. Since then, it has been shown that the intestine is a hotspot for horizontal gene transfer^{123,124}. Commensal organisms, *E. coli* in particular, often function as intermediaries promoting the spread of immunity as well as serving as reservoirs for resistance genes either obtained from, or passed on to, bacteria in transit. There is increasing evidence that antibiotics increase all three modes of HGT¹²⁵. To recapitulate HGT, the three major modes that can rearrange the genetic material via homologous recombination are:

- Transformation of competent bacteria by free DNA
- Transduction by bacteriophages
- Conjugation of plasmids and integrative conjugative elements (ICEs)

4.2.4.3 Mobile genetic elements

Briefly, antimicrobials may affect transduction by promoting prophage excision and host cell lysis. While, in terms of conjugation antimicrobials can also enhance the transfer frequency of various mobile genetic elements (MGEs). Transferable resistance is mediated by MGEs such as transposons. Transposons were discovered in search for understanding the promiscuity of certain resistance genes. Transposons are stretches of DNA that have high mobility. Take for instance an *E. coli* strain carrying a transposon within its chromosome. In this case, the transposon confers resistance against streptomycin. The transposon can then "jump" to a conjugative plasmid within the cytoplasm. During a conjugation attempt the transposon can then integrate into the chromosome. Transposons are repetitive DNA sequences that can move (transpose) from one location to another within a genome. They are elements

that carry transposases and dispensable cargo genes, while insertion sequences only carry a transposase. A key aspect of HTG is that it allows for transfer of virulence factors and antibiotic resistance markers, which raises socio-economic concerns, as the spread of antibiotic resistance renders current drugs inefficient. *This thesis will focus on transformation.*

Transformation requires that bacteria become competent, a state which can be naturally or artificially induced. It has been seen that certain drugs can induce as well as increase competence, e.g. transformation frequencies increased in *Streptococcus pneumoniae* under antibiotic stress¹²⁶. The state of competence is transient in *S. pneumoniae* and is controlled by the *com* regulon. Fluoroquinolone antibiotics induced the expression of the *com* regulon. The authors of the paper investigating this phenomenon further suggested competence as a general stress response to changes in the environment¹²⁶, drawing a parallel to the SOS response of *E. coli* as a result of the DNA damage induced by the antibiotic. *Legionella pneumophila* has a similar response to genotoxic stress. DNA-damaging agents, such as fluoroquinolone and UV radiation, increased competence expression, as well as increasing the bacterium's ability to acquire and integrate DNA. Antibiotic stress also induced genetic exchange in *Helicobacter pylori*. The resulting DNA damage increased the DNA uptake machinery in favor of DNA repair systems. It is fascinating, that *H. pylori* was also found to increase the secretion of an enzyme that releases DNA from other cells (lys, HPG27_320). The authors hypothesized that this protein lyses neighboring cells in order to free up DNA for uptake¹²⁷, ultimately expanding its genetic pool, which could increase fitness.

HGT provides a platform for fast-paced adaptation appearing to have an overall beneficial outcome to the organism. However, investigation of the totally genetic diversity within a species, otherwise known as the pan-genome, reveals HGT is not always beneficial within the context of evolutionary. Pan-genome analysis indicates HGT could be unfavorable to large genomes¹²⁸. If transferred genes are on average advantageous, they would be retained. The term pan-genome refers to all genes within a species, and includes both the core and dispensable genomes. For instance, bacteria have genes that are vital for survival and function, which will be classified as the core genome. Other genes can be more flexible. The genome of *V. cholerae* for example makes up ~95% core genomes and ~5% genes belonging to an acquired flexible gene pool¹⁷. The pan-genome would reflect this by being similar in size to the average genome. However, HGT is only favorable if the acquired gene is different from genes already present in the genome. While larger genomes could maintain HGT, this is only up to a certain point, beyond which duplication events could increase and become problematic. It bears mentioning that duplication event can be beneficial and are under constant selection pressure. For instance, the majority of *V. cholerae* strains have roughly 150 duplication events. Most of these involve regulatory functions of chemotaxis and solute transport¹⁷, suggesting a selection

pressure for retaining these abilities. This could be related to the many environments the pathogen encounters during its infection cycle. There is evidence to suggest that HGT has occurred in all eukaryotic lineages. However, it appears to be a disadvantage within organisms that have high fidelity genome replication. This would imply that an organism undergoing genomic flux due to HGT has in a sense predetermined its genome size. That being said, HGT could also be positively selected as a defense mechanism against mobile genetic elements that are selfishly more promiscuous. Regardless a small genome ultimately limits the number of plausible combinations for adaptation, meaning it should be possible to determine an appropriate tactical countermeasure that efficiently combats resistance.

HGT enables rapid adaptation, with dwindling novel drug discoveries and increasing occurrence of resistance. A better understanding of this mechanism is critical as subpar treatment can promote HGT. We need better treatment strategies. One route that can mediate this is pharmacokinetics, and pharmacodynamic modeling.

4.2.5 Pharmacokinetics and pharmacodynamics PK/PD models

The following section will briefly touch upon pharmacokinetics (PK), how the drug functions inside a person. Move on to describe pharmacodynamic (PD) profiles, how the drug interacts within bacteria. Finally, discuss how PK/PD modeling can inform dosing regimens, and overcome some of the obstacles genetic diversity and bacterial adaptation provides. The purpose of this section is to discuss PK and PD in general terms and will therefore not go into any mathematical depth.

While HGT contributes to broad distribution of resistance genes and a rapid evolutionary response to treatment, resistance can also occur as a result of the selection pressure imposed by the antibiotic; this is why finding the optimal dose is vital. The overall goal of the drug discovery is to find the best possible dosing regimen that maximizes treatment. Drug discovery prior to the 1960s was based on blindly screening thousands of compounds. Once a promising drug compound was detected, hundreds of related compounds would be synthesized to find the safest, most effective drug. A central limitation for traditional drug screening was that it struggled to reveal why a drug candidate failed. The 1980s experienced significant advances in molecular biology, crystallography and computational chemistry allowing for structure-based approaches providing both data and direction. The quality of a compound could be determined by molecular properties such as size, shape, lipophilicity, hydrogen bond capability, and polarity. The structure of the target site could be resolved, the active site of the target could be determined, and the drug-target interaction established. Yet, such structural information is still not fully integrated into PK/PD models. Implementing such information could reduce the cost of drug discovery by predicting whether a novel drug candidate would fail out the gate. Additionally this could help inform the dosing strategies used for clinical trials, minimizing error rates.

Furthermore, drug efficacy will depend on how the body reacts to the dosing and drug itself (PK), as well as how the amount of drug, upon reaching its target, exerts its effect (PD). Simply put drugs only have an effect if they bind their intended target. While target binding is a prerequisite it is not a guarantee. Drug efficacy is also dependent on potency and exposure at the active site of the target, exposure meaning the concentration of the drug that reaches the target and that is present for an appropriate amount of time after administration.

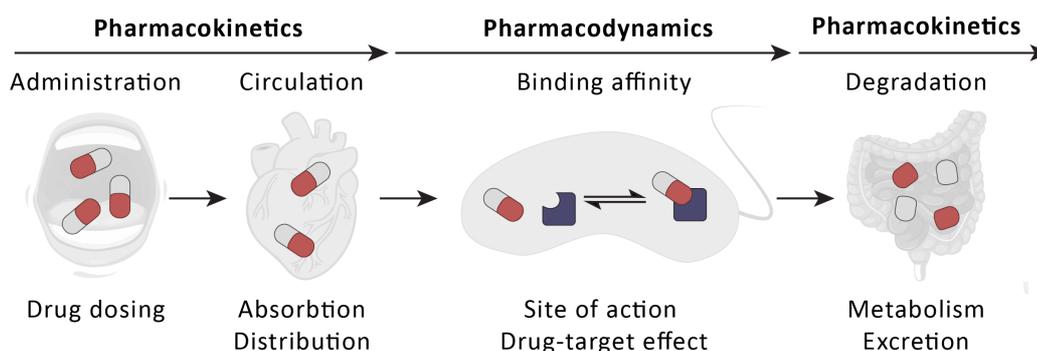


Figure 7 Overview of pharmacokinetics and pharmacodynamics

4.2.5.1 Drugs in the body – Pharmacokinetics:

Table 3: Processes that control the kinetics of the concentration-time profile of a drug in blood, tissue and organs involve:

Absorption: How the drug gets to its target site

- Route of administration, and drug release or liberation
- Additionally, how much of the drug reaches its target (bioavailability).

Distribution: How the drug travels through the bloodstream to various tissues

- Volume of distribution: the amount of drug divided by the amount of drug in the plasma.
- Plasma protein binding: only unbound drug can act on the pharmacological active site. More unbound drugs increases the effect of the drug as well as potential toxicity.

Metabolism: How the drug is degraded

- Drug inactivation, conversion

Excretion: How the drug is removed from the body

- Elimination rate (clearance)
- The time it takes for the serum drug concentration to drop by half (half-life)

Reference material for the summary¹²⁹⁻¹³¹

The core principles of pharmacokinetics comprise of absorption, distribution, metabolism, and excretion (ADME). ADME describes the mass balance, essentially how the drug maneuvers within the body. The human body is often split into compartments, meaning models can range from a single compartment to several hundred. The number of compartments depends on the scientific question. A visual example of a single compartment model could be pouring gin into a glass. The glass will for this purpose function as the human body, while gin is the drug. Filling the glass with gin would demonstrate absorption while emptying the glass would demonstrate clearance in the form of excretion. The instantaneous change in amount

of gin within the glass would be a combination of "the rate of gin filling the glass" and the "the rate of gin being emptied from the glass".

$$\frac{\Delta A}{\Delta t} = [\text{rate of gin entering}] - [\text{rate of gin leaving}] \quad (1)$$

- dA/dt describes the instant rate change for the amount of gin within the glass.

In a clinical sense instantaneous absorption and distribution of a drug is unlikely. Perfusion and permeation will likely need to be considered. These terms deal with drug distribution in a tissue and how the drug moves between compartments. However, intravenous injection (infusion) is one of the fastest drug deliveries, taking a more direct route. In this case, the rate of change between drug and compartment is defined by the elimination rate and the drug concentration. From there the change in amount of drug within the compartment and time can be simplified into a log-linear relationship. This is ideal as the drug amount can then be determined along the slope of the curve¹³². Orally administered drugs are absorbed into the bloodstream via the gastrointestinal tract. The efficacy of the drug can then become encumbered by plasma binding proteins, this is also an issue for intravenous injection. Bacterial infections can occur at any site, however the interstitial fluid of tissue is more common¹³³, meaning the drug needs to leave the bloodstream at one point. This is not possible if bound to plasma proteins. Additionally, the drug loses its effect if bound to plasma proteins, however it is unclear if this is a result of inactivation or is simply preoccupied with plasma proteins¹³⁴.

Transport of the antibiotic to its target site is vital and can be a major obstacle¹³⁵. Fortunately, computational modeling can help overcome some of these obstacles by predicting drug activity, investigating dynamic drug interactions, and evaluate how the drug effects act within a bacterial population.

4.2.5.2 Drugs in bacteria – Pharmacodynamics

Eventually, the administered drug reaches the bacterial infection. At this time the drug needs to overcome the resistance mechanism of bacteria to exert its function. Depending on the mode of action, the drug can either act directly on the membrane or will need to cross the permeability barrier. An additional obstacle is the potential for secretion by efflux pumps. *This will not be discussed further.*

Antibiotics impose a selection pressure by killing susceptible bacteria, while resistant strains remain less affected. A high antibiotic concentration likely leads to effective killing. However, low concentrations give bacteria time to respond. A non-lethal, or sub-inhibitory, concentration of antibiotics promotes several interesting phenomenon, such as high-level target resistance, selection of an intrinsically resistant sub-population and accelerated spread of antibiotic resistance¹³⁶. Evaluation of drug efficacy and dosing is normally described by the minimum inhibitory concentration

(MIC). The minimal inhibitory concentration is defined as the lowest concentration of drug that inhibits growth of bacteria, e.g. the lowest lethal dose. The MIC is a static value and does not consider drug concentrations that fluctuate above or below. This is a typical occurrence for patients prescribed an antibiotic regimen, where one takes an oral dose daily over the course of a defined period. This can be problematic, as clinically relevant high-level resistance emerges from exposure to antibiotic concentrations higher than the minimal inhibitory concentration (MIC). Determining MIC can be as simple as a broth microdilution; however, this does not provide mechanistic information. The outcome of the dose is also dependent on the competition that occurs between the susceptible and resistant populations causing infection.

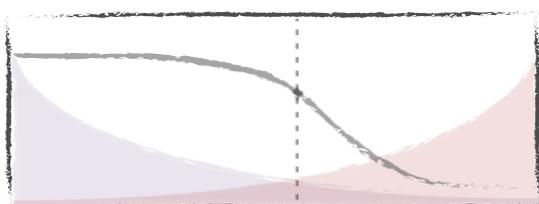


Figure 8a Pharmacodynamic curve. A general pharmacodynamic curve demonstrating net growth (Grey line) of several time kill experiments. The corresponding replication rate and elimination rates are shown in purple, and red respectively.

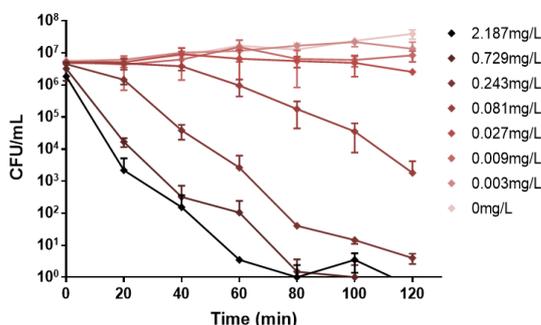


Figure 8b Time-kill experiment. MIC determination for wild-type E.coli and ciprofloxacin. There is partial net growth at 0.081mg/L, while there is slight net death at 0.027mg/L. The true MIC is likely to be somewhere between these values. However, we have decent resolution in this assay and 0.027mg/L is a decent estimate. If you were to plot the net growth of each of these samples, the resulting curve would be similar to the general pharmacodynamic curve.

A more useful MIC determination can be obtained via time-kill experiments. The resulting pharmacodynamic curve will show the net change in the population, including both replication and elimination (**Figure 8a**). A time-kill experiment looks at how the bacterial population responds to a serial dilution of the drug as a function of time. The resulting curve will demonstrate how an isogenic population is affected by the drug concentrations or dose (**Figure 8b**). Once the experiment is performed, and the data is plotted in a graph, the slope of the kill curve demonstrates the effect of the drug. The MIC is defined as zero net growth. This is also where the growth and elimination rates intercept. Time-kill experiments are time-consuming and expensive, since they require multiple drug concentrations and several sampling events. Further, the experiments might need to be repeated to ensure proper sampling of all killing events.

Recently, drug-target binding affinities were determined to be a functional proxy for MIC^{137,138}. As the goal is to optimize dosing regimens of antibiotics that have been in circulation for a long period, there is a substantial amount of information regarding the biochemical properties of the antibiotics. By substituting time-kill curves with readily available biochemical parameters, an entire pharmacodynamic profile can be assessed, and used to guide proper dosing (**paper II**). The administered dose effect

will depend on whether the drug has a bactericidal or bacteriostatic action. Bacteriostatic is defined as inhibiting replication and will not impact the elimination rate, while bactericidal drugs reduce replication rates to zero while increasing elimination rates. Ciprofloxacin used in **Figure 8a** therefore exhibits both a bactericidal and bacteriostatic effect.

4.2.6 Selection pressures

The drug concentration required to ensure that no bacteria recover, is the mutation prevention concentration (MPC). The MIC and MPC values define the boundaries of the mutant selection window (MSW). The main selection for resistance mutants occurs within this range of drug concentrations (**Figure 9**). However selection can also occur at sub-MIC concentrations, which is known as the sub-MIC mutation selection window¹³⁹. Sub-MIC concentrations are not only an issue in active infections, but can also lead to resistance selection within the environment due to various pollutants from the industry and improper disposal of drugs. This is of major concern as even trace amounts of drugs, within the environment, aid resistance¹⁴⁰.

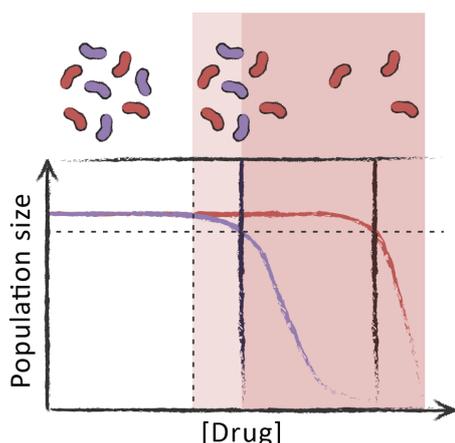


Figure 9: Depicts a kill curve of a bacterial population, where half are susceptible (purple bacteria) and the remaining half is resistant (red bacteria). The bacterial population is depicted above the graph and declines according to the dosing. Where the purple curve intercepts with the horizontal dotted line indicated zero net growth and equals the MIC for the susceptible population. Where the red curve intercepts the horizontal line is the MIC for the resistant bacteria. Notice the susceptible population will start to decline before the dose reaches MIC, indicated by a vertical dashed line. The light shade of red indicates the sub-MIC mutation selection window. The traditional MSW is depicted with dark red shading.

Drug concentrations above the selection window avoid resistance development and selection. Concentrations below the selection window do not kill the susceptible strain however this also avoids favoring the resistant population. Thus, to achieve proper dosing the selection window needs to be properly characterized for the wild-type strain as well as the majority of expected resistance mutants. In other words, this required a complete assessment of how the drug interacts with the target for several drug concentrations within a susceptible strain as well as several resistance mutants, a complete pharmacodynamic profile. It is experimentally not feasible to do this for every drug-target interaction in combination with all potential resistance mutants. Additionally, the drug would need to be administered to the patient, introducing issues such as accumulation which opens a completely different can of worms. Fortunately, such extensive profiles can be achieved by computational modeling.

4.2.7 Dosing strategies

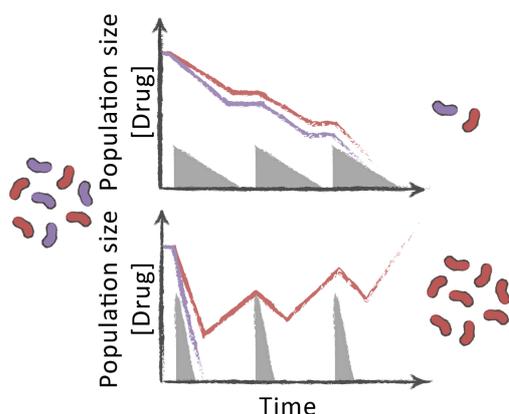
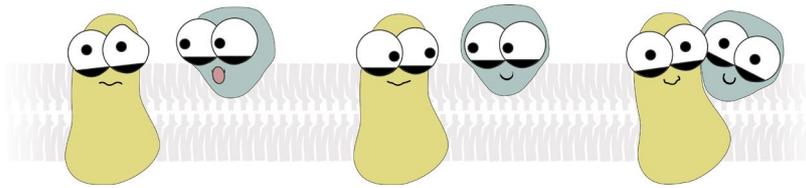


Figure 10: Dosing strategies. The effects of dosing on a bacterial population when frequency remains the same while concentration and half-life differ. Low concentrations with a long half-life result in slower killing but does not preferentially select for resistance. High concentrations with a shorter half-life efficiently kills the susceptible strain but is not enough to remove the resistant population.

To stay above the MIC threshold for a longer period, one can increase the dose and/or shorten the dose interval, as well as increase the infusion time, for instance by implementing intravenous infusion. This has been proven to minimize mortality, though it does require hospitalization¹⁴¹. As an example, high doses over a short period of time could be enough to kill the susceptible population immediately. However, the drug is removed before eliminating the resistant population, whereas lower concentrations over a longer period would eventually clear the infection (**Figure 10**). Altering the chemical properties of a drug can aid absorption and activity. Only the amount of drug that reaches its target and remains

active determines the antibacterial effect, meaning it is equally important to optimize the dosing. Dosing can be optimized by frequency and concentration, which are PK parameters. The route of administration will impact these properties. While using PK/PD to inform dosing is not novel, most current dosing regimens are still based on clinical experience¹⁴². However, computational PK/PD models are gaining traction for dose optimization. PK/PD models can help minimize resistance and hamper HGT. As several drugs need to cross the cell envelope to reach its corresponding target a solid understanding of bacterial membranes could inform such models.

4.3 The cell envelope



4.3.1 Dynamic structure important for protection and adaptation

“That’s the kind of protection, everyone is shouting about”
– David Bowie –

Bacteria have developed a complex permeability barrier that compartmentalizes activity and protects against exposure to various compounds in the environment. The cell envelope is a complex and dynamic structure that protects that bacteria and is important for adaptation. The cell envelope of Gram-positive bacteria consists a cytoplasmic membrane and the cell wall. While, Gram-negative bacteria have an additional outer membrane. This envelope allows for selective uptake of nutrients and other molecules, as well as secretion of signaling compounds, and metabolic waste products. Membranes also play an important role for antibiotic resistance, as they limit molecules from entering the cytoplasm, thus providing intrinsic antibiotic resistance¹⁴³. This is further complicated by the fact that bacteria can also excrete the antibiotics that have bypassed the initial barrier through efflux pumps. Biological membranes consist of a phospholipid bilayer and various membrane proteins. The lipid bilayer consists of two lipid leaflets, both of which often contain an asymmetric distribution of lipid and protein content. Because the cell envelope acts as a barrier between the cell and the environment, bacteria have evolved complex strategies to transduce signals across the membrane to enable cells to sense their environment. Furthermore, the membrane organizes large molecular complexes such as the DNA uptake machinery discussed earlier.

4.3.2 The cell envelope

Gram-positive bacteria, have two functional layers, a cytoplasmic membrane, and a thick cell wall. As, the bacteria lack an outer membrane this makes them more permissible than Gram-negative bacteria. While, Gram-negative bacteria, have an inner membrane and an outer membrane. Between the two membrane layers is the periplasm, which also contains a thin cell wall. The main differences between the architecture of the cell envelopes further allow for characterization of bacteria. Using a simple staining process, known as the Gram stain, one can distinguish between Gram-positive and -negative bacteria. The staining process consists of crystal violet, and a counterstain e.g. safronin. Crystal violet is purple and will bind specific components of the cell wall, while the counterstain is red/pink and stains cell membranes. Gram-positives have a thick cell wall and will retain the purple dye well. During the staining process the outer membrane of Gram-negatives is dissolved.

Additionally, the thin cell of Gram negatives will not retain the purple stain, exposing the pink color of the inner membrane.

The cell wall is composed of peptidoglycans, which are a polymer consisting of polysaccharides and peptide chains. Peptidoglycans are constantly being synthesized, hydrolyzed and modified allowing for growth and cell division. The cell wall has a mesh-like structure, which is robust yet flexible and is to an extent permissible as it contains several large pores, or gaps within the mesh, that allow large molecules entry to the cell. A key structural component of Gram-positive cells are carbohydrate based anionic polymers. They provide stability, permeability, and function as anchor points for extracellular enzymes involved in biosynthesis and degradation of the cell wall, further influencing the integrity of the membrane and interacting with extracellular components. These polymers exist in all Gram-positive bacteria and the most common are teichoic acids¹⁴³. It is also worth mentioning that some bacteria do not have a cell wall, a wide spread example of this is *Mycoplasma*. For most bacteria the cell wall is important for regulating osmosis. However, *Mycoplasma* species survive as intracellular pathogens, where the host maintains a controlled osmotic environment.

The biophysical and biochemical properties of membranes enable compartmentalization. Cell membranes are composed of lipids that contain a hydrophilic headgroup and a hydrophobic tail. As a result, the two lipid leaflets are arranged in a tail-to-tail manner. The lipids are arranged as an inner and outer leaflet, not to be confused with the inner and outer membranes. For instance, the outer membrane of Gram-negatives is comprised of an inner and outer leaflet. The inner leaflet is oriented towards the cytoplasm, while the outer leaflet is either facing the periplasm or the environment. The cell envelope of Gram-negative bacteria can be viewed as four compartments; an inner membrane, a periplasm, with thin cell wall, and an outer membrane, see **Figure 11** for a schematic overview. The outer membrane is highly selective, with size exclusion properties. It also exposes antigenic determinants such as lipopolysaccharides to the environment. It is an asymmetric bilayer of phospholipids and lipopolysaccharides¹⁴⁴. In particular, lipopolysaccharides are only present in the outer leaflet¹⁴⁵. The composition of phospholipids is similar between the inner leaflet of the outer membrane and the cytoplasmic membrane. The outer membrane also contains porins, which are large channels allowing for passive diffusion of large hydrophilic molecules. The main function of porins is to allow entry of ions and nutrients into the periplasm. Lipoproteins are equally important components of the outer membrane as they are involved in transport of molecules, maintaining the envelope integrity, signal transduction and pathogenesis¹⁴³. The periplasm is a viscous aqueous cellular compartment that isolates potentially harmful enzymes, such as alkaline phosphatases. The periplasm also contains proteins important for environmental sensing as well as the sensing of outer membrane damage¹⁴⁶. The inner, or

cytoplasmic, membrane is the primary permeability barrier of the cell, and contains specific transport proteins and permeases. Additionally, the membrane contains enzymes involved in energy and phospholipid production¹⁴⁷.

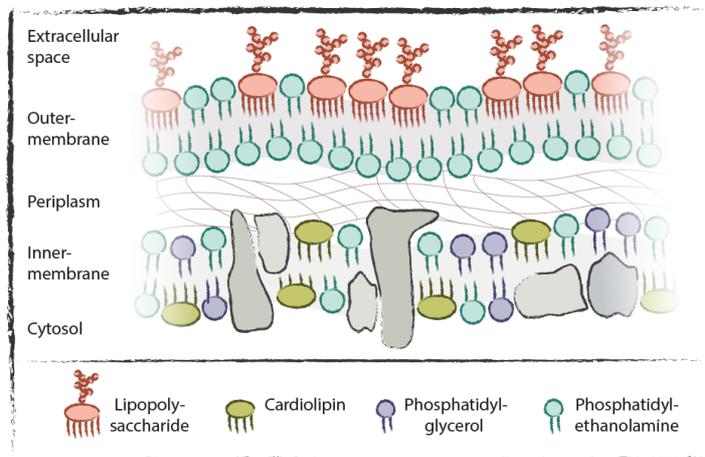


Figure 11 Schematic overview of a Gram-negative cell envelope. On the top is the extracellular space or outside of the bacteria. The first barrier is the outer membrane consisting of an asymmetric lipid bilayer, with lipopolysaccharides on the outer leaflet. Gram-negatives have a thin cell wall in the aqueous periplasmic space, indicated by the mesh of grey lines. The inner membrane is the final layer and surrounds the cytosol. This is a vital compartment, as many of the membrane proteins generate energy, lipid biosynthesis, and mediate protein secretion.

4.3.3 Membranes

Cell function is highly dependent on biological membranes, which enables selective permeation, energy production and environmental sensing. Additionally, membranes form the border between the cytoplasm and the environment. The membrane has previously been thought of as a “fluid mosaic”, where the composition is sparse and randomly distributed. However, membranes are now thought to be non-randomly organized in functional domains crowded by intercalating membrane proteins, several lipid species and other complex structures¹⁴⁸. The bilayers of membranes have an asymmetric distribution of both lipid and protein content, which is most prominent for the outer membrane. The main components of cell membranes are glycerophospholipids and membrane proteins.

4.3.3.1 Glycerophospholipids

As previously stated, glycerophospholipids can structurally be divided into two functional segments, a hydrophilic headgroup and a hydrophobic tail. The hydrophilic part points to the outside of the leaflet, and the hydrophobic part points to the inside of the leaflet. The tail group consists of hydrocarbon chains of various lengths, known as acyl chains. The headgroup can interact with membrane proteins and regulate protein function, while variations in the acyl chain affect membrane fluidity and crowding. Desaturation of acyl will produce a kink in the carbon chain resulting in looser distribution while saturated chains produce more rigid membrane regions. The major lipid components are phosphatidyl-ethanolamine (PE), phosphatidylglycerol (PG), and cardiolipin (CL). PE is the most abundant phospholipid. See **Figure 12** for a structural overview. *V. cholerae* membranes, for example contains ~84% PE, ~10% PG, ~4.0% CL, and 2% of uncharacterized phospholipids⁹. Meanwhile, the phospholipidome of *E. coli* comprises ~70-80% PE, 10-20% PG, and ~5-10% CL¹⁴⁹. The composition depends on the growth phase and environmental conditions.

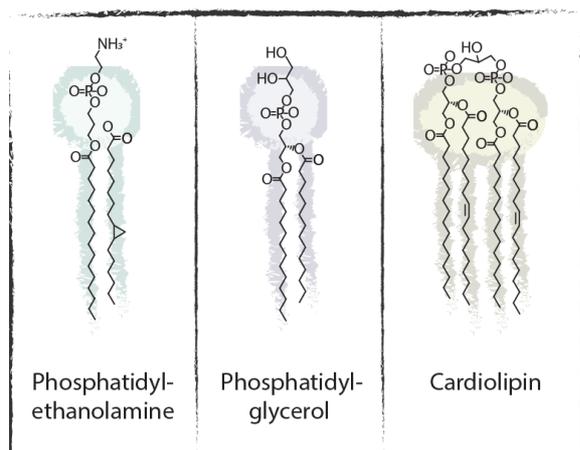


Figure 12 Structure of common phospholipids in bacterial membranes. The polar head group is depicted as a circle with the acyl chains pointing downwards.

The main difference between PE and PG lies within the polar head, either an ammonium group or hydroxyl group, respectively¹⁵⁰, while CL is a unique phospholipid dimer. The polar head group of CL consists of two phosphatide moieties instead of one as is for PE, and PG, with a total of four acyl chains. Due to its structure, the lipid is conical in shape. CL is predominately localized in the inner membrane¹⁵¹ of bacteria and often clusters at the cell poles and cell division septum, where it is thought to relieve curvature stress imposed on the membrane. The synthesis of CL is

regulated by the *clsABC* genes. *clsA* encodes the main synthase during exponential growth, while *clsB*, and *clsC* are more active in the stationary growth phase^{152,153}. The synthases condense two PG producing CL and glycerol as a byproduct.

4.3.4 Lipid – protein interaction

Membranes are sensitive to external stimuli and will remodel in response to stress factors, such as temperature. Cold temperatures will lead to rigidification, while higher temperature result in fluidization. As, changing temperatures impose stress on the membrane, bacteria often respond with various strategies. Bacteria may modify the size of the polar head groups in addition to charge. As, protein-lipid interactions are dependent on the polar head group, and proteins-lipid interactions can act to stabilize the membrane, this will impact fluidity. However, the main impact on membrane fluidity is likely by acyl chain modulation. Desaturation will generate kinks in the chain resulting in fluidization¹⁵⁴, and isomerization of the chain can also affect fluidization. Membrane fluidity is critical to cell function and is constantly being monitored by the cell.

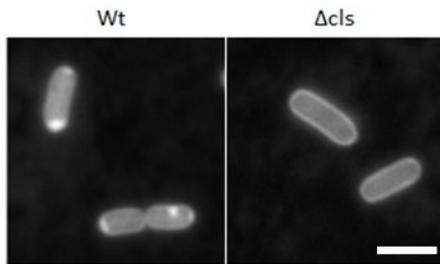


Figure 13 CL stained with NAO. The dye NAO preferentially binds to CL, and accumulates at the cell poles in wild-type *E. coli*. NAO staining of a *clsABC* mutant that cannot produce CL shows diffuse localization throughout the cell perimeter. This is likely due to non-specific binding between NAO and PG. Scalebar is 2 μm .

The concentration of phospholipids can also influence protein localization. ProP is a membrane protein involved in monitoring osmotic stress in *E. coli*. Investigation of the localization of the membrane protein revealed a threshold requirement of CL in order to establish proper polar localization¹⁵⁵. See **Figure 13** for CL membrane distribution. If the cardiolipin content was below the threshold, the protein was spread out through the membrane. The study also correlated the activity of the protein to its CL dependent localization. A recent publication shows that polar localization of cardiolipin is

independent of ProP, and suggests that the localization of ProP diffuses freely through the membrane. However it is retained in the natural CL rich regions of the membrane, which could adhere to a “diffuse - and - capture” model¹⁵⁶. Further CL concentrations fluctuate depending on the growth phase¹⁵⁷. While the synthase is more active during the exponential phase¹⁵⁸, the concentration of CL accumulates throughout the growth cycle and peaks in the stationary phase. It is possible that the accumulation of CL could be a mechanism to regulate ProP activity and restrict it to phases with high CL abundance. While this is conjecture, co-accumulation of ProP and CL has been observed at the cell poles as a result of increasing osmolarity¹⁵⁹.

4.3.4.1 Membrane proteins

Membrane proteins can be classified by how they are positioned within the lipid bilayer as either integral or peripheral. Integral proteins are embedded in the membrane, while peripheral proteins are attached to the membrane surface. When a protein or protein complex traverses the cell envelope, it is referred to as a transmembrane protein. Whereas, peripheral membrane proteins are not embedded, but instead attached to the membrane surface. There are also a subset of membrane proteins known as, lipoproteins. These proteins are anchored to the membrane by a lipid moiety. Most lipoproteins are located in the inner leaflet of the outer membrane¹⁶⁰. An example of a lipoprotein is PilW in *Neisseria meningitidis*. PilW is an outer membrane lipoprotein associated with the type IV pilus of DNA uptake. It was found to be involved in stabilizing the pilus formation, as well as the stability of PilQ, the outer membrane pore complex that allows pilus extension and contraction and DNA uptake into the periplasm. In a *N. meningitidis pilW* mutant piliation could still occur¹⁶¹.

The membrane localization and construction of the natural transformation machinery has yet to be fully characterized. However, Metzinger et al. found that localization seem to involve the membrane secretins, PilQ, and PilB, as well as an ATPase

required for pilus elongation¹⁶². During imaging experiments, these proteins appeared to have an even distribution throughout the outer membrane. The PilB ATPase appeared to be more mobile than PilW. Once in the presence of PilQ, pilus extension occurs. The authors further noticed several co-occurring foci of PilQ and PilB, yet only one pilus formation¹⁶². They hypothesized that this results from a lag in the complex assembly. A few of the complexes could be more complete than others, however PilA would not assemble until after PilB had attached⁷⁹. This suggests there are semi-assembled structures moving around in the membrane that fully assemble once signaled to do so.

Proteins embedded in the membrane can carry out a broad range of cellular function, such as, determining membrane curvature. *V. cholerae* is a highly curved microbe. It was recently discovered that this curvature was a result of a periplasmic filament generated by CrvAB, binds to the peptidoglycan cell wall. CrvAB are homologues. However, CrvB has a domain that is not present in CrvA, which is required to promote curvature, while CrvA is needed from CrvB assembly¹⁶³. This synergy indicates that both proteins are specialized and required for efficient assembly of the periskeletal filament. The curvature is generated by asymmetric insertion of PG. PG was found to integrate quicker into the “outer face”, e.g. the longer side, compared to the “inner face” of *V. cholerae*¹⁶⁴. This curvature was found to aid in host colonization and hydrogel penetration, while straightened cells had reduced motility on low concentration agar plates¹⁶⁴.

4.3.5 Cardiolipin – MgtA interaction

Several membrane proteins localize towards the cell pole¹⁶⁵, where they are thought to interact with CL¹⁶⁶. One such protein is MgtA¹⁶⁷. This is one of two magnesium ion (Mg^{2+}) transporters in *E. coli*. The transporter is required for magnesium uptake from the periplasm into the cytosol, and is also important for growth in *E. coli*¹⁶⁸ (**Figure 14**). The other transporter is CorA. Mg^{2+} is required for bacterial growth and has several roles within bacteria, such as acting as a cofactor for ATPase dependent reactions and acting as a signaling molecule for virulence¹⁶⁹. MgtA is in the inner membrane and is a P-type ATPase. P-type ATPases are ubiquitous primary ion pumps. Subramani et al. found the association with the headgroup of CL was important for MgtA function. Further, finding evidence that suggests the activity is regulated by the local concentration of CL¹⁶⁷. The transporter also demonstrates higher sensitivity to Mg^{2+} concentrations than most other ATPase *in vitro*¹⁶⁷. The authors suggest a model for MgtA activation at low Mg^{2+} concentrations. The expression is upregulated, and the transporter localizes to the cell poles in the inner membrane where it associates with CL and PG. Interaction with CL is critical for MgtA activity, ultimately translocating Mg^{2+} from the periplasm into the cytosol. Once the cytosolic concentration of Mg^{2+} reaches a threshold, the ATPase will inhibit its own expression and its activity within the membrane. MgtA expression and activity is

tightly regulated by the level of available Mg^{2+} . While Mg^{2+} is critical to cell function, MgtA is not the main transporter in *E. coli* and the functional role remains unclear. However, MgtA in *Salmonella enterica* has been linked to host survival during infection¹⁷⁰.

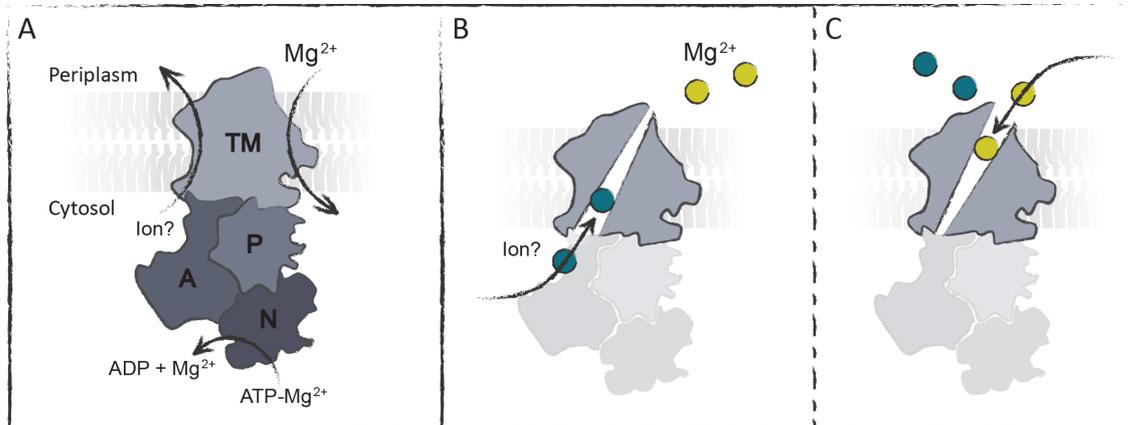


Figure 14 Schematic overview of MgtA structure and function: **A.** P-type ATPase facilitates ATP-dependent ion transport across the inner membrane. The transported consists of four domains. Transmembrane helices (TM), a nucleotide binding domain (N), a phosphorylation domain (P) and an actuator domain (A). The protein complex cycles between two main intermediary conformational states. Each conformational state has different affinities for Mg^{2+} . **B.** The first state has high affinity allowing for binding of ions in the cytosol. **C.** ATP is hydrolyzed and rearranges the domains into an open conformation facing the periplasm. The ion is released in exchange for Mg^{2+} . This occurs as the different states have different binding affinities. Mg^{2+} is also referred to as a counterion in this case. Binding of Mg^{2+} induces ATP hydrolysis and the structural rearrangements open towards the cytosol and releases Mg^{2+} .

Adapted, and printed with permission¹⁶⁷

In summary, the cell envelope is highly variable and adaptive. It is highly regulated, and selectively permeable. The envelope provides protection from the environment while at the same time ensuring that vital components are scavenged. Protein complexes organize within the membrane allowing for uptake of nutrients, such as chitin, in a nutrient poor environment, securing immediate survival. As well as taking up genetic material, such as DNA, which can provide an adaptive advantage. At the same time membranes limit cytoplasmic entrance of substances such as antibiotics. Further, drugs that pass the membrane will also effectively lower the external concentration, which can aid in protecting a fraction of the population. Membranes are more than just lipids, they are the foundation of bacterial adaptation and responsiveness. *I would like to conclude this introduction with an illustration done by Prof. David Goodsell which captures the complexity of the cell compartment and protein interactions with mesmerizing beauty* **Figure 15.**

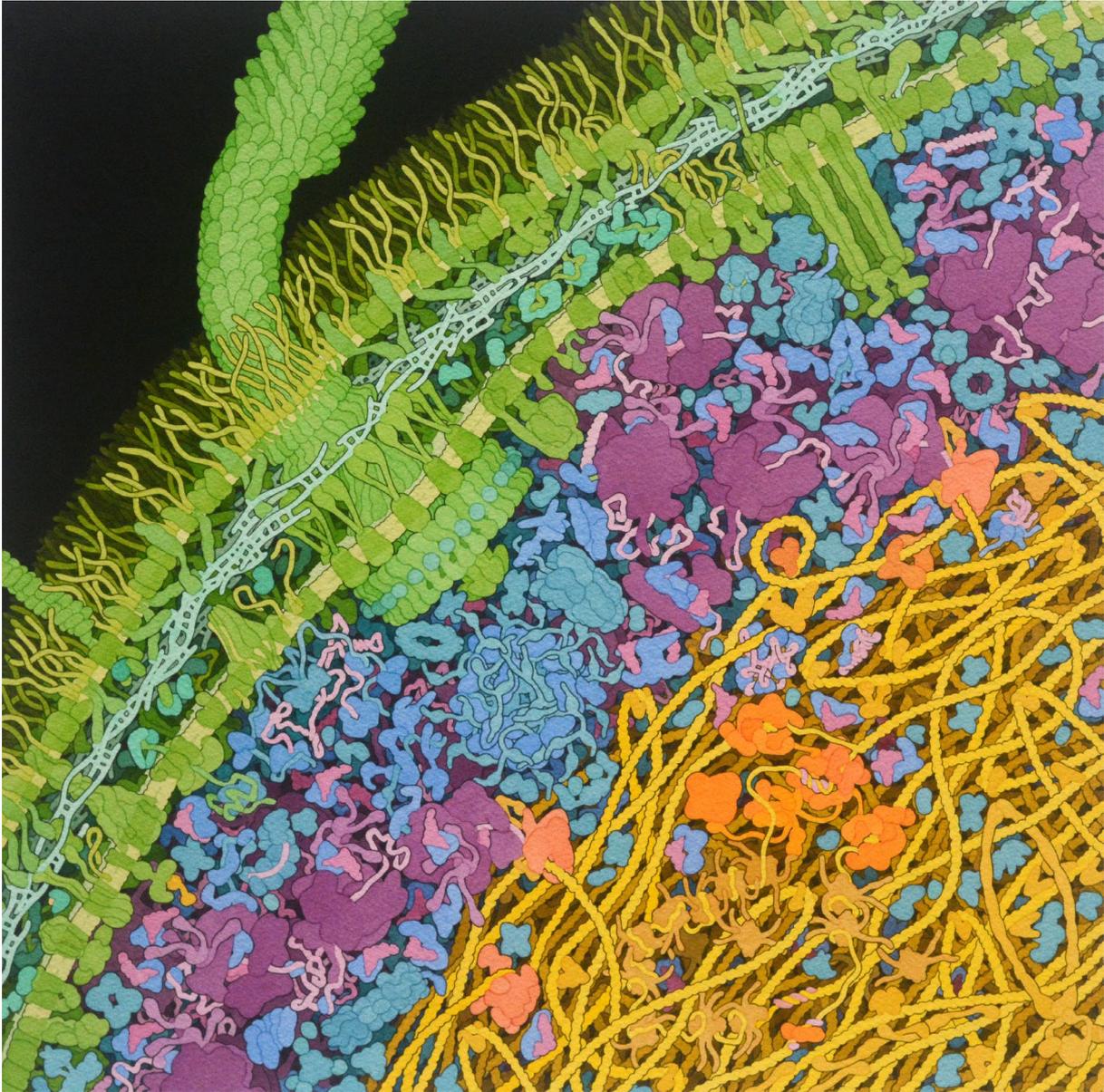


Figure 15 Cross section of an *E. coli*. The artwork depicts the complex organization of a cell. The flagellar complex of *E. coli* crosses the outer membrane, cell wall and anchors in the inner membrane. Notice how cluttered the membranes are. These structures are shown in green. Lipopolysaccharide chains extend out from the outer membrane, followed by a matrix of cross-linked peptidoglycans, before reaching the inner membrane. The cytoplasm is depicted in blue. It is easy to assume this space is a simple void, however seeing the cellular clutter within this context gives a broader appreciation for not only the complex nature of regulating the proteins activity in regard to time and space, but also needing a finely tuned binding affinity to avoid erroneous processes. Protein production is depicted in purple, where ribosomes are interacting with tRNA in maroon to produce white strands of amino acid chains that will eventually be folded into proteins. At the heart of the cell lies the loosely packaged nucleoid depicted in orange and yellow. Illustration printed with permission from Prof. David Goodsell¹⁷¹.

5 Synopses of publications

This thesis aimed to investigate bacterial response to stressors such as extracellular DNA and antibiotics. As well as the complex interplay of lipids and membrane proteins. The following section summarizes the three articles included in the thesis.

Paper I: Transposon insertion sequencing reveals dependency of natural transformation on amino acid metabolism

Natural transformation is a highly orchestrated response to the environment. A multitude of factors influence the diverse process. Extending the known factors of natural transformation generates a more complete picture of the pathogens ability to adapt. To identify additional genes involved in natural transformation we implemented a top-down approach by inducing natural transformation within a high-density transposon insertion population, and find amino acid metabolism genes as important for natural transformation.

In this study we created a high complexity transposon insertion screen within *V. cholerae* O1 El Tor isolate A1552. The transposon insertion sequencing screen was applied to a natural transformation assay. A validation set of mutants confirm the screen as appropriate for investigating genes involved in DNA uptake. Genes involved with amino acid transport and metabolism were most prominent. Addition of external amino acids to wild-type demonstrated a time-dependent effect on transformation frequencies. In contrast, supplementing the specific amino acid to its corresponding amino acid knockout recovered the DNA uptake back to wild-type levels. Fluorescence reporter constructs of competence induction within our mutant panel show a drop in $[P_{comEA}]dsRed$ expression indicating the gene deletion functions upstream of competence induction.

We find that a large portion of the genome is dedicated towards natural transformation and show that the majority involved amino acid transport and metabolism. We argue that amino acid specific metabolism is a critical baseline for natural transformation and is prioritized in a smaller part of the population resulting in the heterogenous distribution of competence.

Paper II: Drug-target binding quantitatively predicts optimal antibiotic dose levels in quinolones

Antibiotic resistance is one of the biggest threats to public health today. At the same time drug discovery has stalled, meaning we need to optimize the treatment strategies currently available. Strategies for optimization focus on optimizing the concentration and frequency of dosing as well as duration of treatment. The initial guide for these values is the minimal inhibitory concentration (MIC). While finding the MIC is straight forward, this method lacks predictability and is further complicated by fluctuating antibiotic concentrations within a patient. It is therefore suggested to use more complex dose-response measurements to obtain a full pharmacodynamic profile, where the effect of the drug is assessed as a kill rate at several concentrations. This process is extensive and time-consuming. There is currently a need for models that can predict the efficacy of modified antibiotics as well as resistance mutations.

We generated a model by finding a linear correlation between MIC and binding affinities (K_D). The model was extended to include binding and unbinding to the target and takes bacterial replication into consideration. The model was then calibrated to bacterial time-kill curves using the quinolone ciprofloxacin. In combination with the known binding parameters the ciprofloxacin, the time-kill experiment allowed for estimating the action and the effects of ciprofloxacin on *E. coli*. Further this data was used to estimate how ciprofloxacin alters the growth, and death rate of bacteria. The model was tested on published time-kill curves involving the effects of ampicillin on *E. coli*. Minor adjustments were made to the model but overall gave a good fit. This is a mechanistic modeling, where all parameters can be changes. Meaning we can simulate, bacterial resistance mutations or modifications of the antibiotics themselves, for single drug administration.

In this study, we accurately predicted antibiotic action based on readily available parameters for drug-target interactions. This model can aid in optimized dosing regimens which is critical to combating the rise in antibiotic resistance.

Paper III: The bacterial magnesium transporter MgtA reveals highly selective interaction with specific cardiolipin species

The bacterial membrane is more than a protective barrier. Proteins and lipids undergo dynamic interactions within the membrane, the complexity of which has yet to be defined, especially in terms of regulatory effects. Using a transmembrane magnesium transporter, MgtA, within the context of *E.coli*, we examined the interplay between transporter and lipid species. We aimed to further the knowledge of protein lipid interactions. And found a preference for two specific CL species. Additionally, demonstrating that not only was this interaction important for transporter activity, but that the two CL species generated a synergistic activation of the transporter. Additionally impacting its localization within the membrane. CL promotes membrane curvature and is present within the bacterial membrane, clustering at the poles of *E. coli*. *V. cholerae* has a similar membrane composition as *E. coli*, however it does not contain an endogenous MgtA transporter. Upon comparing the localization pattern of MgtA between *E. coli* and *V. cholerae*, we found that MgtA localizes to the membrane in both organisms. However, it presents a diffuse distribution within *V. cholerae*, while clustering as polar foci in *E. coli*. This result was then corroborated in an *E. coli* strain lacking the ability to synthesize CL. This indicates that CL is not involved in localization. We speculate that it could still have a potential role for anchoring in the membrane. This study contributes insight into the complex interplay between proteins and lipids. A deeper understanding of the biophysical properties of membranes is important to understand bacterial adaptation.

6 Methodological considerations

6.1 Model organisms:

6.1.1 *Vibrio cholerae*

V. cholerae O1 El Tor (Inaba) strain A1552, originally named 92A1552¹⁷², was isolated in California, USA, and has been linked to the Peruvian pandemic in the 1990s. It is fully sequenced¹⁷³ and can undergo natural transformation¹⁷⁴, and does not contain a MgtA transporter homolog. The pathogen is an excellent adaptor and transitions rapidly between aquatic and host environments. The pathogen shares morphology and physiology with *E. coli*, *Salmonella enterica*, and *Pseudomonas aeruginosa*. The structural mechanisms share similarities allowing for an extended repertoire of methods as well as an extensive experimental foundation that can be built on. While there are similarity, there are small differences that can greatly affect the experimental outcome and tiny tweaks such as codon optimization could be beneficial. Additionally, *V. cholerae* can exert peculiar effects in response to plasmids, sugars, and some antibiotics^{175,176}, which might need to be evaluated. The organism is robust, easy to cultivate, and the reliance on natural transformation simplifies cloning.

6.1.2 *Escherichia coli*

E. coli is a renowned model organism. Our studies implement a K-12 derivative known as MG1655. The genome is fully sequenced and as it has been used within a laboratory setting for over 100 years, there is a huge repertoire of molecular and biotechnical methods that can be directly applied to the organism. While many species of *E. coli* are symbiotic inhabitants of the human small intestine, pathogenic *E. coli* can cause severe disease in form of diarrhea, sepsis, and meningitis^{177,178}.

6.2 Transposon insertions sequencing

Transposon insertion sequencing (TIS) is a powerful tool to investigate the genetic response to a specific condition. It allows for high-throughput screening of a large number of mutants, by using transposons. Transposons are genetic elements that are “mobile by nature”, relocating from one genetic location to another¹⁷⁹. There are several types of transposons, however to keep it simple: autonomous transposable elements encode a transposase that will recognize the inverted repeats in the DNA sequence, cut out the transposon, and insert it into another location. Hence the transposon can keep “jumping”. While TIS methods use nonautonomous transposons where the transposase is on a non-replicative plasmid¹⁸⁰. The transposon insertion occurs semi-automatically, via conjugation, allowing for large-scale mutagenesis. The insertion can then be identified by sequencing. In **paper I**, once we had constructed and evaluated the input library, it was applied to both selective and non-selective conditions run in parallel. Afterwhich, we compared the

insertions of the two conditions with each other using a bioinformatic pipeline. Ultimately searching for what was lost along the way and comparing differences. There are several variations of TIS methods, yet the main differences lies in the transposon, library construction and sample preparation. Please, refer to the in-depth reviews of the methods that highlight similarities and differences^{181,182}.

A major benefit of transposon insertion sequencing is that it allows for a direct link between a genetic input to a phenotypic readout. Additionally, it allowing for large-scale mutagenesis, as well as a rapid and comprehensive assessment of the contribution of genes. The experimental noise will impact the results significantly, mainly referring to genetic drift and sampling error¹⁸³. The genetic drift and sampling will directly impact the complexity and saturation of the screen with further implications for the analysis. All experimental bottlenecks, such as sample preparation and assay conditions, should be evaluated, and sampled thoroughly. In **paper I**, our study encountered several bottlenecks both during the library construction, as well as the experimental setup. The library is constructs via conjugation, this will directly impact the number of insertions you begin with. Whereas, the major bottleneck for the experimental setup was the low transformation frequencies of natural transformation. This was improved by optimizing the transformation assay conditions as well as increasing the sampling.

The TIS pipeline is depicted in **Figure 16**. In brief, the initial construction of the library occurs via conjugation. A non replicative plasmid, carrying a selection maker within the inverted repeats is transferred from donor, *E. coli*, to recipient, *V. cholerae*. Recipient cells where the transposase has integrated the transposon into the genome are selected for. The transposase sequence is not integrated. The transformants are harvested and genomic DNA is extracted (**Section 6.2.1**). The genomic DNA is then fragmented. The fragments are end-repaired to avoid further degradation and adapters are ligated. These become anchor points for PCR primers that bring in sequences important for identification, such as barcodes, and proper sequencer functionality, such as spacer sequences aiding base diversity which is critical for the instrument (**Section 6.2.2**). Before sequencing the fragments need to be uniform in size with a high purity as the sequencer will preferentially bind shorter fragments (**Section 6.2.3**). Once the fragments are sequenced they are referred to as reads. The quality of the reads need to be assess and poor quality reads can be filtered out bioinformatically. The reads are mapped to the genome of interest and the insertion density is evaluated. A high density transposon library consists of multiple insertions in every locus. High saturation refers to a low probability of missed insertions. The term sequencing depth refers to the number of times a “sequence/nucleotide” was detected. While, the coverage is a broader terms referring to the fraction of potential sites detected. These factors will determine the resolution of your screen. After mapping you can evaluate how many times a gene was disrupted by an insertion as well as detected by sequencing. This is an important distinction because

TIS screens fail to evaluate truly essential genes, as these will disappear from the screen if interrupted. *This is why there is some controversy within the community regarding the terms “essential” and non-essential” to denote important genes involved in your test vs neutral genes.* The insertion distribution will allow you to evaluate the quality of your library, which can then be used to test specific conditions. We aimed to test a selective condition and compare it to a non-selective condition in order to find conditionally essential genes. This means we used the non-selective condition as a “control” to remove insertions that are involved in the growth condition and not directly linked to the selective condition we tested.

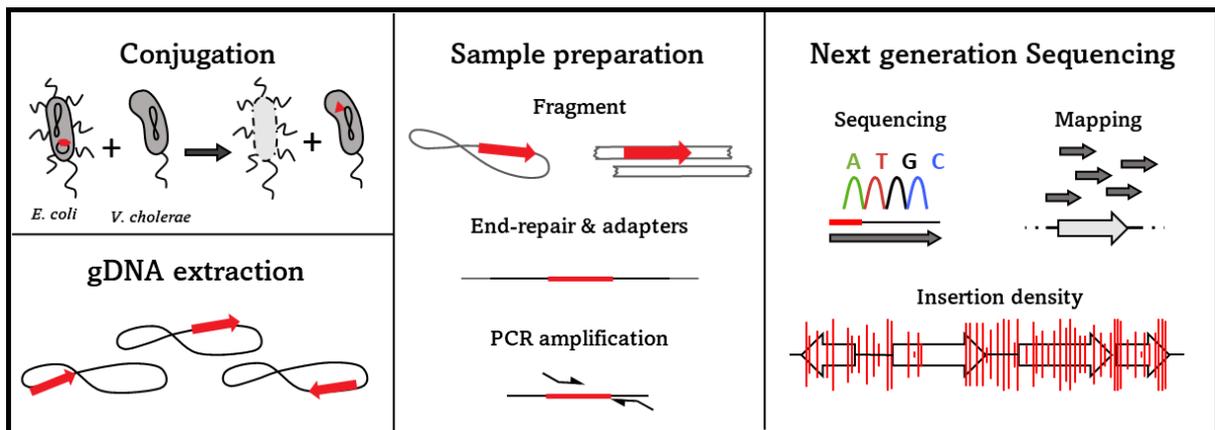


Figure 16 Schematic representation of the transposon insertion sequencing pipeline. The panels are orientated in a top-down, left to right manner. The library is constructed by mating *E. coli* and *V. cholerae*. The *E. coli* is pilated and contains an 8-shaped genome and a plasmid with a red transposon. The *V. cholerae* is depicted with a single pilus and an 8-shaped genome. After mating the *E. coli* perishes through a perforated membrane while the *V. cholerae* now has a red transposon within its genome. The genomic DNA (gDNA) is extracted. Inserted transposons are depicted as red arrows at different areas of the genome, as this is a stochastic process. Sample preparation for the sequencer consists of fragmenting the gDNA, end-repairing the fragments and adding adapters. Adapter sequences are added to the fragments during PCR amplification, and finally applied to next generation sequencing. A small part of the red transposon is sequenced along with flanking gDNA from *V. cholerae*. The reads are depicted as dark gray arrows that overlap with a gene sequence depicted in light gray. Once the reads are mapped to the genome, the insertion density can be assessed. This is shown as red lines of different lengths crossing a segment of genome depicted as four white arrows. The height of the arrow demonstrated the number of times that insert was detected. Notice one of the genes has less red lines or insertions. This could mean the gene was not covered or that it is “essential”. The probability of it being essential will increase with a high complexity, high saturation library. Regions with frequent insertions will indicate non-essential genes, as the interruption did not significantly impact the survival of the bacteria.

The experimental setup and analysis pipeline can be both cumbersome and expensive which is why the following section on TIS is mainly intended for someone wanting to apply the method. While different problems can occur, hopefully, this section results in some issues being avoided. However, in short, it is necessary to evaluate the potential effects of the test condition. When creating a TIS library it is important to ensure maximum efficiency and minimal growth during mating, by considering growth conditions such as temperature, and timing, to limit affecting the disruption frequencies unnecessarily. All impurities need to be removed from the sample by the end of the sample preparation, and the fragment sizes should be uniform. It is also worth integrating specific barcodes to allow pooling libraries on the same MiSeq chip, as this will dramatically cut down on cost. Finally, the sequencing

reads are a list of probabilities and the screen will need to be experimentally validated.

6.2.1 Creating the library, conjugation

The genome of *V. cholerae* consists of two chromosomes with approximately 200,000 regions, where the two nucleosides thymine (T) and adenine (A) are seated together (TA sites). The TA sites are evenly distributed throughout both chromosomes, and the majority of genes have at least seven TA sites. In **paper I**, we used a Himar-based mariner transposon to estimate the fitness effects of genes involved in natural transformation. This transposon will insert itself at TA sites within the genome with low insertional bias¹⁸⁴. The TnSeq library is generated by the mating of a vector containing the transposon. The vector has a low copy number, and the transposon will be inserted permanently in front of a “TA site”. Further, the vector contains a R6K mutant origin of replication requiring the lambda pir gene to maintain the plasmid, making it non-replicative in *V. cholerae*. The density of the library is defined by how many TA sites are interrupted. A high-density library indicates that the majority of dispensable insertion sites are disrupted by a transposon. A high-density library is desirable as transposon insertion data is a list of probabilities where the insert is either lost by chance or as a direct consequence of the test condition.

While constructing the library, careful consideration was given to minimizing competition and random loss of insertions due to the fitness decrease of the transposons and space. We, therefore, optimized the amount of bacteria, duration of conjugation, and incubation conditions to yield high frequency while minimizing growth. During conjugation, the enriched media was also used throughout the experiment to limit the loss of genes important for metabolism and growth. The populations were also plated on large Petri dishes to avoid overcrowding. This was to avoid any limiting the effect of the antibiotic selection for the transposon, but also reduce competition for space and nutrition.

6.2.2 DNA shearing and PCR amplification

The genomic DNA was sheared using the Covaris M220 focused-ultrasonicator. This sonication method is acoustic shearing, while other methods can be chemical or enzymatic. The resulting DNA fragments varied in lengths, however within a range close to 600bp. While the supplier promises strict shearing of the DNA there were smaller fragments not visualized by gel electrophoresis that needed to be removed, because the sequencer preferentially binds shorter fragments. Some of these fragments will be lost during the column purification steps however bead exclusion or gel extractions might be necessary. One way to evaluate the fragments quality is by using a bioanalyzer, while this is not critical it can greatly improve the sequencing quality and yield.

After shearing, the fragmented DNA ends were repaired using the NEBNext Ultra II End Repair/dA-Tailing module, and A-tails (NEB/Bionordika) were added to allow the ligation of T-tailed adapters. Transposon-junctions were amplified in two rounds of PCR. The first round used a primer that recognized the transposon end, and a primer that recognizes the adapter which at the same time introduced an index for sequencing. The second round of PCR introduced the requisite sequence for binding to the Illumina sequencing platform: P5, P7 and six different spacers to ensure variability for the sequencer, each read had different nucleotides at any given cycle.

The PCR amplification is a size exclusion step and the first round of PCR removes the majority of un-useable transposon sequences. Additionally, PCRs can result in over-representation of certain sequences either due to insertion bias, due to a growth advantage before harvesting, or early favoritism of amplicons, otherwise known as a PCR jackpot event. Transposon junctions were amplified in two rounds of PCR. The first round used primers that recognize the transposon end and the ligated flanking adapter. At the same time, this step introduced an index for sequencing. The second round of PCR introduces the requisite sequence for binding to the Illumina sequencing platform, and a variety of spacer sequences ensured variability for the sequencer. This was critical for the autofocus of the MiSeq. Illegitimate sequences were removed by stringent mapping criteria of the sequences that were unique reads. This was done by using “cutadapt”, and “bowtie2”. The reads were also normalized based on the position of the read, due to biased amplification at the origin of replication. Because of the rolling ball replication, reads that flank the *ori* were more represented than at the terminus. Unrealistically represented amplicons and jackpot events were further addressed by normalizing the mapped reads. This is a part of the ARTIST pipeline.

6.2.3 Size exclusion of DNA fragments

To estimate the true density of the constructed library, we needed to optimize the sample preparation for the sequencer. This consisted of ensuring minimal loss during genomic DNA (gDNA) extraction and downstream sample preparation while also maintaining sample purity. DNA purity is a critical factor for the sequencing quality. Extraction of gDNA was therefore performed using a modified version of the Promega gDNA extraction kit. We found that increasing the number of washes and centrifugation steps was a defining factor for purity. Additionally, the yield was dramatically improved by allowing resuspension of the gDNA pellets overnight at 4°C. While only pico grams of DNA is required for next-generation sequencing, our downstream sample preparation involved acoustic shearing of DNA and gel extraction. While acoustic shearing is both stringent and reproducible fragmentation of gDNA, the fragmentation is random and we lose transposon insertions. After shearing the transposon junctions need to be end-repaired and adapters are ligated by homopolymer nucleotide extension DNA. This is required for PCR amplification.

Initially, we used bead exclusion to select the correct fragment size. Even using the most stringent wash, 0.5x left side exclusion performed twice, we were unable to fully remove short sequences. This became an issue due to truncated reads. This was detected after sequencing in the form of a low sequence yield and low-quality scores. As bead exclusion was not enough to remove the smaller fragments from our sample preparation, we decided to physically remove the majority of short fragments by gel electrophoresis and excising the appropriate size. Gel extraction is not as pure as bead size exclusion and requires an additional column purification, meaning additional loss of DNA. For these reasons larger amounts of initial gDNA were used to ensure that enough DNA was retained between the extraction and purification steps. We excised a band of roughly 300-600pb. Note, this is a consistent smear and not too defined fractions of DNA sizes. This method was, after optimization, stringent enough to remove the shorter sequences that were clogging the sequencer, and we could proceed with our sample and evaluate the quality of the library.

6.2.4 Sequencing

The DNA fragments were sequenced using an Illumina MiSeq, which provides some basic statistics that help evaluate the success of the “run”, such as cluster density, clusters passing filter, and estimated yield. These values essential tell you how dense the sequences are, a rough estimate of the overall quality of sequences or reads and an estimate of how many fragments were sequenced. A read is simply the sequence of one DNA fragment. The sequences were then assessed based on quality of reads using the FastQC software. This enables visualization of read lengths and some more indepth statistics regarding the quality of the reads. Low quality sequences can be filtered out using cutadapt as well as non-specific sequences that did not contain the index we used for sample identification. Afterwhich the reads were mapped to the fully annotated *V. cholerae* A1552 genome, using the bowtie2 aligner.

After mapping, the transposon insertion sites were tallied and assigned to either genes or intergenomic regions using a combination of python and MATLAB scripts provided by the Waldor lab. The ARTIST pipeline¹⁸³⁻¹⁸⁵ was used to assess the fitness effects of the high density mutagenesis. A benefit of the pipeline was that it is annotation independent and considers the variability between two samples due to stochastic biological bottlenecks. It consisted of two analysis arms called: EL-ARTIST (Essential loci analysis) and CON-ARTIST (Conditionally essential loci analysis). EL-ARTIST defined the fitness effects of different loci within the TIS library for a single growth condition. While, CON-ARTIST found which loci were required for survival by comparing two transposon libraries that experienced different growth conditions.

The EL-ARTIST arm started by normalizing the data to compensate for incomplete DNA replication, and a sliding window approach was used to find underrepresented regions in the mapped reads. This output was used to train a Hidden Markov model (HMM) and estimated how dispensible each transposon insertion site was for the

growth condition, for a single library. Moving on to CON-ARTIST, an independent control dataset was generated that represented sample errors that could occur due to chance. The resulting control libraries then underwent simulated genetic drift. The reads for the test sample were then compared to the simulated control dataset by using a Mann-Whitney U (MWU) statistical test. This was to avoid making any assumptions about the distribution of reads in each dataset making it less sensitive to experimental bias such as PCR jackpot events. The output was then used to train an annotation-independent HMM. The HMM predicted which genomic regions were required for the growth condition. Consequently, the pipeline classified genes that were both reproducibly and significantly different from the control dataset, and denoted the genes as either conditionally depleted, enriched or unchanged. A more traditional annotation would be conditionally essential, non essential and neutral, respectively.

6.2.5 Assessing library saturation – proper sampling

In this section I refer to the saturation of the output library and not the initial constructed library. The natural transformation assay was performed in three independent experiments. Where the output libraries for the selective and non-selective conditions were sequenced, evaluated and compared before altering the experimental setup.

The first output library of the transformation population resulted in low complexity. The control sample had 1.9M reads and hit 84% TA sites, while the DNA-receiving sample has 3.1M reads hitting only 32% TA sites. We had sequenced the selective condition deeper than the non-selective and ended up with a large amount of insertions disappearing. To help evaluate if this was due to a bottleneck or sampling we evaluated the disruption frequencies (**Figure 17**). As the genomic and intergenomic regions within the genome contain a known amount of potential insertion sites, it is possible to look at the distribution of detected insertions gene. After the reads are mapped this distribution was tallied and the disruption ratio for each region of the genome was evaluated and plotted. The fraction of disrupted insertion sites for all regions can then be distributed into bins that increase with 0.5 increments (**Figure 17a**). Genes sorted into the “0” bin indicate none of the potential insertion sites were detected in the screen, while genes sorted into “1” on the far right side of the graph mean all potential insertion sites were disrupted. This is a rough estimate of essential or under-represented vs neutral insertion effects on genes. In such a scenario one would expect a negative skew of the data, while a positive skew would indicate several insertions are missing from the screen. As the dataset has a positive skew it indicated this was due to an experimental bottleneck, and we increased the sample size.

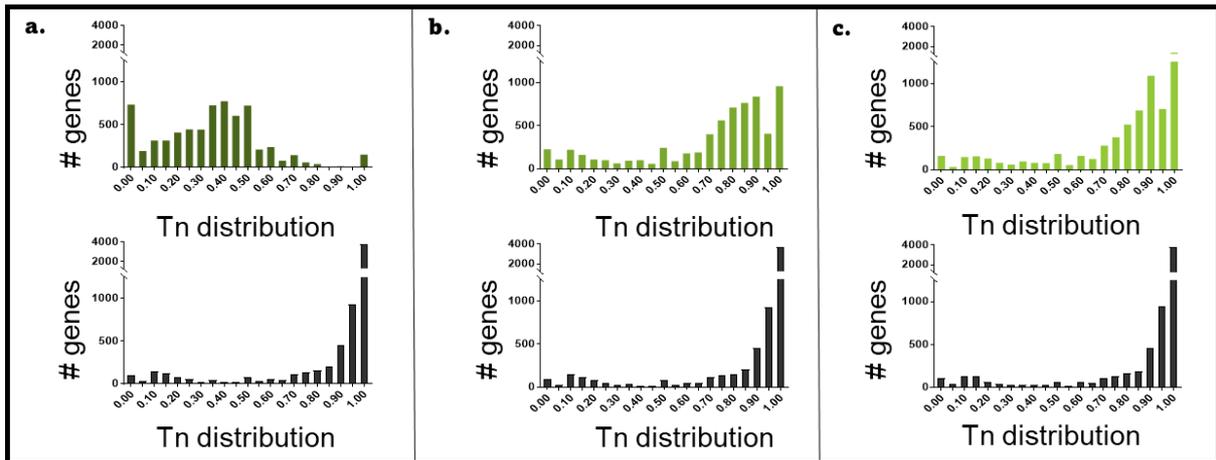


Figure 17 Evaluate library saturation. The graphs demonstrate the distribution of insertions per gene. The green-scale graphs on the top row indicate the test output library. While the corresponding control outputs library can be found below in black. (a.) nr. 1 resulted in a 5-fold library saturation compared to potential sites. (b.) nr. 2 produced a 10- fold library saturation. (c.) nr. 3 resulted in a 50-fold library saturation.

The first output library resulted in a 5-fold saturation of the available TA sites in the genome. The assay was repeated two more times, resulting in 10- and 50- fold saturation of potential TA sites. After sequencing the new output samples, we saw that increasing the amount of transformants resulted in a higher saturation. This can be seen by the peak of the graph shifting towards the right. Sample nr. 2 resulted in 66% of TA sites detected, meaning 66% of potential insertion sites were both disrupted and detected (**Figure 17b**). While, sample nr. 3 detected 72% of the available TA sites (**Figure 17c**). Further sequencing deeper did not result in a major impact. Upon evaluating the distribution of detected insertion sites per genes, we also found a negatively skewed dataset for both output samples nr. 2 and nr. 3. We were confident the output sample was near saturation, and had hit a plateau for insertions detected within our natural transformation assay. Further increasing the sample size would not gain any significant information. We therefore opted to continue the analysis using sample nr. 3, along with the corresponding control sample. Additionally, when evaluating the candidate lists from each sample we could detected important regulatory genes that have been shown to impact natural transformation in other studies. This indicates that our experimental bottleneck was not biased and the loss of insertions was random.

During this evaluation it also became apparent that the transition from enriched media to chitin and back to enriched media did not cause a significant selection pressure. As we saw no dramatic change in insertion frequencies between the input library and the control output sample so we ended up lowering the number of cells used for the control sample. This was minimize the workload significantly while at the same time being confident there was no loss of insertions. This also highlights an important point. When comparing the input library to the control output sample we maintained a high complexity, meaning our experimental setup such as media changes, and other sample preparation did not impact the distribution of insertions.

This is an important point as we are comparing the difference between two ratios meaning any loss will impact the analysis.

6.3 Natural transformation homologous recombination

Natural transformation of *V. cholerae* was tested by incubating bacteria on chitin in minimal media. As this assay is a major bottleneck for our transposon insertion library we needed to maximize the number of transformants. The transformation assay was based on the optimized protocol by the Blokesch laboratory¹⁸⁶, with slight modifications made to the media due to solubility issues (**Figure 18**). The transforming material was a PCR fragment, amplified from the genomic DNA of *V. cholerae* O1 El Tor c6706. The strain contains a point mutation within the *rpsL* gene that confers resistance to streptomycin, this allowed use to select for cells that has taken up and integrated the donor DNA. We avoided using gDNA as donor DNA, since this could lead to the uptake of DNA that would be outside of our detection. We tested different concentrations of our donor DNA, as well as varying stretches of DNA flanking the point mutation (500bp 1000 bp, and 2000 bp). The donor DNA with flanking 2000 bp had the highest recombination frequency. the total size of the fragment was ~ 4000 bp. We also tested several concentrations of DNA; 0.5 ug, 1 ug, 2 ug, 4 ug, and 10 ug. 2 ug gave a sufficiently high transformation frequency, while being experimentally efficient. The transformation assay was performed using chitin flakes from shrimp shells. The frequencies were calculated by dividing the number of streptomycin-resistant transformants by the total number of CFUs.

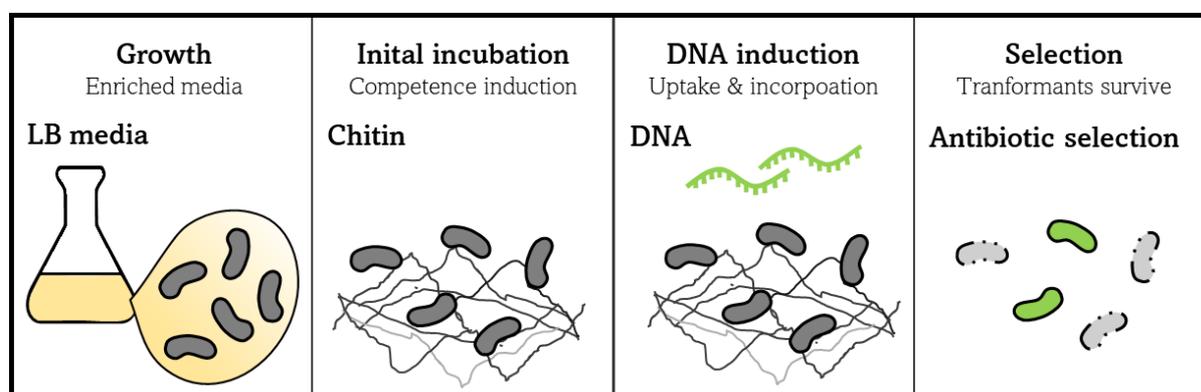


Figure 18 Schematic representation of the natural transformation pipeline. *V. cholerae* is depicted as inflated commas in gray. LB media is depicted in yellow, demonstrating the initial growth phase. Chitin used to induce competence is depicted by jagged black lines. The donor DNA used to confer antibiotic resistance is depicted in green. Bacteria that have taken up and integrated the donor DNA are shown with a green fill. While bacteria that do not survive the antibiotic selection pressure are shown with light gray fill and a perforated outline.

The donor DNA was PCR amplified from genomic DNA of a *V. cholerae* c6706 strain. This is a non-transforming strain due to a frameshift that results in insufficient expression of LuxO and ultimately impairs quorum sensing¹⁷⁴ that carries a single point mutation in the *rpsL* conferring resistance to streptomycin. Point mutations have been seen to be poorly integrated, as mismatch components can block the progression of homologous recombination¹⁸⁶⁻¹⁸⁸. However, our transformation

frequency was optimized close to saturation by increasing the amount of DNA. During homologous recombination, exchange of DNA strands can occur resulting in a heteroduplex. If a mismatched nucleotide occurs within this structure it can be corrected by mismatch repair¹⁸⁹. Since homologous recombination and mismatch repair are entwined, the fact that we used a single point mutation in the donor DNA could result in a lower transformation frequency¹⁹⁰.

While spontaneously arising resistance to streptomycin was never detected, the possibility of nongenetic inheritance cannot be overlooked. Upon replication, after exogenous DNA has been integrated, only one of the cells will contain the newly acquired sequence. However, it has been demonstrated that the newly acquired material is transcribed before cell division, allowing for the random distribution of the gene product¹⁸⁷. It is therefore probable that our screen contains some false positives. The pipeline known as ARTIST was originally optimized to evaluate the fitness landscape of *V. cholerae* genes within a mouse model^{183,184}. The creators of the analysis pipeline found that modeling stochastic loss of disruption in the context of genetic drift minimizes false positives. Additionally, only a small subset of cells will have enough newly synthesized protein to survive antibiotic selection. While we did not encounter any false positives out of the list of candidates we tested, we did not test every gene and it cannot be excluded. However, the effects of which should be minimum as our genetic screen was stringently evaluated and validated.

6.3.1 Nongenetic inheritance of tolerance

Nongenetic inheritance of tolerance within the natural transformation assay lasted long enough to produce long-lasting and potential genetic resistance. Experimentally, this was a pain as it resulted in false positives which in turn increased costs due to additional testing and repetitions. This was also not resolved by simply increasing the antibiotic concentration or swapping out the antibiotic cassette. I tested the MIC of the *V. cholerae* on tetracycline finding 0.75mg/mL to be sufficient for killing. However, after two days on chitin, this was no longer adequate. The same issue appeared when switching to a different antibiotic cassette. This indicated that the chitin flakes were the main problem, either by absorbing the antibiotic or inducing a state of tolerance. I tested this by harvesting bacteria from chitin flakes after 24 h, plating a serial dilution on agar plates, and immediately exposing it to UV radiation for 40s before incubating at 30°C overnight. The control assay was a culture grown to an early exponential phase in LB. The amount of bacteria was matched to that grown on chitin, and underwent the same plating and UV radiation as the chitin assay. The chitin assay survived the UV radiation better than the LB assay, with more than a log increase in survival. While the implications for a chitin-induced state of tolerance are unclear, it merits investigation, and I hypothesize it will also impact acid tolerance.

6.4 MIC determination and time-kill curves

MIC suffers from poor reproducibility. This is especially problematic within the field. Novel nonMIC approaches can potentially alleviate this. Regardless, a diagnostic setting requires a rapid, robust, and accurate measure to inform proper dosing and minimize resistance selection. The MIC is also an endpoint measurement, meaning the dynamics of the drug-target interactions are lost. Implementing a more robust measure such as binding affinity could potentially mediate novel drug development.

A more dynamic approach to evaluate drug-target interactions involved time-kill curves. However, this is extremely time-consuming. A benefit of generating a mechanistic PD model based on extensive time-kill curves removes the need to repeat the time-kill curves and one gains information regarding how the resistance can impact the drug-target interaction, as well as how resistance development can be avoided, within that setting. While some information can be inferred, a major downside to the PD model is that you will need to recalibrate it every time you change the strain or drug, by performing a time-kill curve. You may also need to have an advanced understanding of the model to do so, as it is not simply inputting new values. However, the amount of additional information and interactive data gained by such a model does justify the experimental and computational effort.

Time-kill curve experiments are extensive and impose time pressure on the experimentalist. It is necessary to complete the sample collection, dilution, and plating, as well as prepare for the next sample, as quickly as possible without exceeding the incubation time. For example, 8 drug concentrations, sampled every 20 min for 3 h and a final time point at 18 h, in triplicates, results in 264 samples. For proper evaluation of the colony forming units (CFU), each sample was diluted further and plated on agar plates for counting. The amount of plates needed for such an experiment was 2112 agar plates. This equates ~70 L of liquid agar that needs to be autoclaved and poured. This does not include preliminary tests. It is also likely that the time-kill curve will need to be repeated if peculiar effects appear, such as tolerance or persistence. Depending on the duration and amount of sampling, as well as sleep schedule, such an experiment can take up to a week including preparation time. This will only provide you with the killing effect for the wild-type. You would need to repeat this experiment for every resistant mutant as well. See **Figure 19** for examples of time-kill experiments for a bacteriostatic (erythromycin) drug. The bacteriostatic drug was tested for the wild-type situation as well as within the context of a resistance mutant. However, it was not included in **Paper II** due to the mutant strain having too high resistance compared to the drug solubility. *As a side note, the time-kill curve experiments resulted in my dominant arm becoming exactly 6kg stronger than the other.*

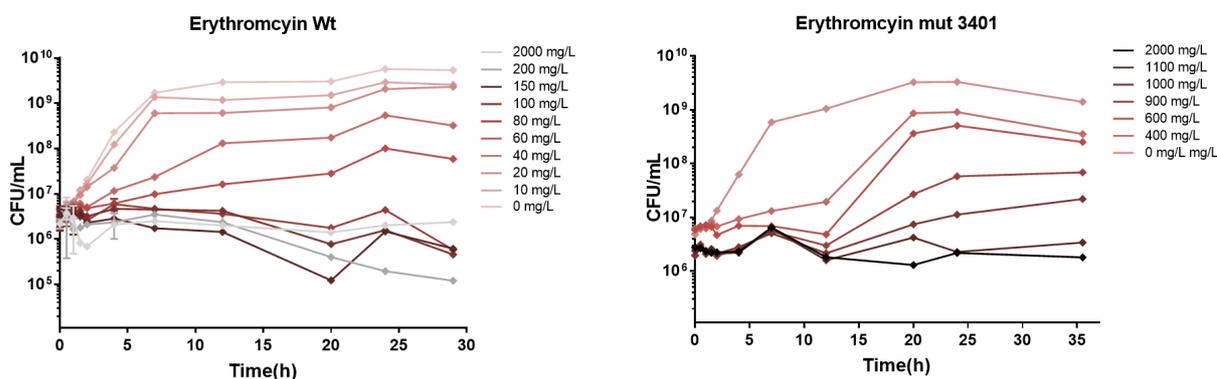


Figure 19 Erythromycin time kill curves, unpublished data. Time kill curves of *E. coli* wild-type and a resistance mutant subjected to erythromycin.

6.5 Single cell microscopy

Microscopy is a valuable tool and provides positional information that can guide interpretations of protein interactions. However, there are some caveats to be mindful of. One of the most critical steps is to consider the imaging modality. This essentially refers to how the light path is manipulated. The different modalities available provide a vast repertoire, however this can be confusing. The light can either excite the entire sample volume, such as with EPI-fluorescence, or the light-path can be manipulated to provide an optical section of the sample such as for confocal imaging or Total Internal Reflective Microscopy (TIRF). Confocal imaging is excellent for imaging complexes within the cell, while TIRF provides a fine optical sectioning, illuminating the membrane surface. I mention this because a basic understanding of the light path aids the biological question. All microscopes are not created equal. However, they work well for the intended applications. There is also a trade-off concerning resolution and the field of view. This should also be considered within the context of the biological question. A localization experiment does not necessarily require super-resolution, as this can limit the imaging area. That been said there are novel imaging modalities that will allow you to obtain a high resolution over an extended field-of-view^{191,192}. However, this might not be necessary, and a simple setup could be sufficient to investigate the cellular process.

In **Papers I and III** we implemented EPI-fluorescence. This is much cheaper than a confocal and the slight resolution gained by using a confocal would be less critical for single cell bacteria. A confocal might be more useful when looking at 3D structures like biofilms or bacterial infections in tissue sample. That been said acquiring a z-stack of images and processing the images with a deconvolution software could give comparable if not more information. Widefield microscopes also have faster image acquisition than confocals, making it more suitable for live cell imaging. Additionally

we needed a high contrast method to visualize the cells. We opted for phase contrast as the imaging technique is well suited for distinguishing between transparent structures with different refractive indexes. as this gave a better contrast of the cells while imaging without fluorescence.

6.5.1 Fluorescent transcription fusion reporter

The main biological considerations concerning fluorescent fusion proteins involved functionality. The activity of the fusion protein should be carefully evaluated to avoid artifacts. When picking the fluorescent protein, it should be taken into consideration that protein folding, brightness, protein stability, and filter settings will also impact the study. If the fusion protein is carried by a plasmid under an inducible promoter, the expression profile should be evaluated to avoid overexpression, but should also reach detectable levels. This should be considered within the context of what is biologically relevant as well. An example of this relevant to **paper III** is the presence of CL peaks at the stationary phase. However, we opted to acquire images in the mid-exponential phase. While we did detect lower levels of CL at this stage, we detected proper localization and avoided peculiar effects associated with the stationary phase.

In **Paper I**, we performed a co-localization or co-occurrence study, where more than one FP is involved. In such a case, one should aim for similar maturation rates. The maturation rates for dsRED.T3 and GFP-mut3* correspond to 78 min and 4 min respectively, which is not ideal and resulted in an uneven expression ratio. This could explain the low expression of competence genes compared to the housekeeping gene *gyrA*. Nevertheless, the expression was examined after 24 h when maximal accumulation is expected to have been achieved. Ideally, we would have recloned this construct. However, due to time constraints and previous validation for localization, we opted to include additional controls where the genes were instead swapped, so the P_{gyrA} would express dsRed, while promoters of competence genes expressed GFPmut3. No difference in overall trend was detected (unpublished data). Red fluorescent proteins often exhibit complex maturation kinetics. GFP maturation is often referred to as a one-step process, while RFP is a two-step process¹⁹³. Therefore, it may be beneficial to implement CFP, GFP, and YFP variants if one needs to track induction curves or is interested in rapid responses. It is important to note that even with a fast maturation rate, it will still take time before a maximum expression is observed.

7 Results and Discussion

This thesis approaches microbial response to external cues within the context of adaptation. The first contact between bacteria and the immediate environment is the cell envelope, which for Gram-negative bacteria includes an outer membrane. The cell membrane provides a hydrophobic barrier with size-exclusion properties and contains components indispensable for bacterial survival and growth. As such, the cell membrane has become increasingly attractive as an antimicrobial target. The composition of the membrane is ever-changing and allows for environmental sensing and initiating an appropriate response. Additionally extracellular DNA is actively taken up from the environment and incorporated into the genome, allowing for rapid adaptation to novel environmental. Moreover, the cell membrane is directly involved in regulating membrane protein activity.

7.1 Membrane Protein localization: Cardiolipin and MgtA

MgtA is an inner membrane P-type ATPase, linked to virulence in *E. coli*. The ATPase is known to interact with CL. In **paper III** this was further characterized in depth. Maximum activity was demonstrated by the ATPase when in the presence of CL 18:1 and 16:0. This nomenclature describes the structure of the acyl chain and refers to 18 carbons and 1 double bond, or 16 carbons without a double bond. CL 18:1 is therefore monounsaturated in cis making the chain kinked while 16:0 is saturated and straight. It should be noted that biological membranes contain more than 50 different types of CL species and my collaborators only tested three different kinds. The CL species tested were CL16:0, CL16:1, and CL18:1. They were selected based on high prevalence in biological membranes.

Monounsaturated lipids increase fluidity, which mediates dynamic transmembrane proteins that undergo a conformational change. The catalytic cycle of MgtA transitions between several states and might require the fluid environment provided by the CL18:1 kink. MgtA activity and stability also required 50% saturated CL 16:0. This could be an artifact for the experimental setup, seeing as these lipid moieties form more ordered domains. The activity was also tested in an artificial membrane, which have different constraints than biological membranes. The enzymatic assays were performed using detergent-lipid micelles instead of liposomes. Micelles consist of lipid monolayers while liposomes are bilayers. MgtA is located in the inner membrane and does not traverse the outer membrane, so a micelle setup is sufficient. However, the assay was performed with only CL species and MgtA. The optimal temperature required to ensure that CL is in a liquid crystalline form is higher than the optimal temperature for MgtA activity and stability. Further, they notice variations in the activity profile of MgtA, likely linked to the detergent used in the

setup, glycol mono-dodecyl ether (C12E8). There is a clear connection between MgtA activity and CL species, which creates a foundation for further studies. However, the activity profile of MgtA should be tested in a lipid bilayer without the presence of detergent, and with various CL species as well as other phospholipids that would allow the membrane to be in a crystalline liquid phase at temperatures that match the transition temperatures for MgtA activity.

Previous experiments showed MgtA localized to the cell poles¹⁶⁷. In that study they evaluated localization patterns of MgtA wild-type and a nonphosphorylatable mutant, MgtA D373N, fused with a C-terminal fluorescent protein (blue fluorescent protein, BFP). They found MgtA-BFP localized to the cell poles, while the MgtA D373N mutant was diffusely spread through the cell. The constructs were expressed in an overexpression model useful for protein isolation, however this can lead to incorrect localization profiles. Overexpressed protein can form aggregates at the cell pole¹⁹⁴, and would result in artifacts. In the case of overexpression, the localization pattern depends on the rate of production. A slow rate of addition favors a single pole, while fast addition results in bipolar localization. Further, my collaborators discovered the C-terminus was involved in Mg²⁺ sensing (unpublished data), indicating that the C-terminal fusion of MgtA would have reduced activity. We, therefore, created an N-terminal fusion of Clover within pBAD33.1, under the control of an inducible promoter, P_{ara}, that allowed for a graded response of protein expression. We opted to avoid BFP as the fluorescent protein (FP) is excited at 381nm and has a peak emission at 445nm. The FP is known to form dimers, which would be disadvantageous to a localization study. Clover, on the other hand, is an optimized version of GFP (green fluorescent protein). It is monomeric providing minimal clustering and undergoes rapid maturation. Additionally, the excitation/emission profiles are tighter than BFP, and matched our filter settings better. The ex/em peaks for Clover at 505/515nm. Further, the FP is brighter and has increased photostability allowing for a lower expression of the construct. Several factors influence the brightness of FPs, such as the quantum yield, extinction coefficient, and the protein folding or maturation rate. The quantum yield shows the ratio of how efficiently incoming light is converted into an output, e.g. the amount of absorbed light vs emitted light. Meanwhile, the extinction coefficient informs the user how strongly the molecule absorbs light at a particular wavelength (nm). These parameters are both theoretical and are void if the protein does not fold properly. Depending on the imaging setup various maturation rates can be used, however for localization studies a faster maturation rate is desirable. A longer maturation rate can be also be used, but should be properly evaluate within the context of timing. The maturation rate for Clover is 22 min at 37°C¹⁹⁵. Once the construct was optimized, we needed to decide which strain backgrounds would provide useful information.

Originally the MgtA-BFP construct was expressed in an *E. coli* C43(DE3) which is an overexpression strain. C43(DE3) is a derivative of the BL21 (DE3) strain that can

better withstand the toxic expression of proteins¹⁹⁶. The overexpression is mediated by the bacteriophage T7 RNA polymerase which has a higher rate of protein expression than the corresponding RNA polymerase in wild-type *E. coli*. We therefore utilized *E. coli* K-12 MG1655 since is derived from a descendent of K-12. K-12 was collected from a fecal isolate from a diphtheria patient in 1922, and has since been maintained in a laboratory setting with minimal alterations to the genome. MG1655 is fully sequenced and widely used as wild-type *E. coli*. This strain was also the parent strain used by the Weibel lab to create a CL synthase triple knockout, Δ clsABC¹⁹⁷. Additionally, *V. cholerae* lacks a native homolog of MgtA, while having a relatively similar membrane composition. These strain backgrounds allowed us to evaluate MgtA localization patterns within a wild-type, a strain with minimal to no CL, and an unfamiliar wild-type background.

The localization patterns for wild-type *E. coli* demonstrated peripheral localization with preferential bipolar accumulation and an occasional minor cluster close to the center of the cell along the membrane. The CL synthase mutant demonstrated a similar localization pattern, although less protein was localized to the membrane and one of the cell poles demonstrated a preference for accumulation. This pattern is not detected in the *V. cholerae* background, although minor detection was observed diffusely spread throughout the membrane and did not exhibit a polar preference. Whereas, MgtA localized to the cell pole in both wild-type and Δ clsABC *E. coli* strains. This indicates that while CL is involved in the activity and possibly stabilizing the MgtA complex, it is not necessarily important for localization.

There could be an unknown membrane protein that mediates polar localization, meaning MgtA could implement a diffusion-capture process. The protein is unlikely to be able to sense membrane curvature as it would result in polar aggregates in *V. cholerae* as well. Further, there are likely traces of CL in the Δ clsABC. Even though the synthases are removed, some redundancies could aid in CL accumulation. Phosphatidylserine synthase, for instance, can result in a minor formation of CL¹⁹⁸. Cumulative redundancies could result in a threshold amount of required CL. A low level of CL could aggregate at one pole, and within the context of asymmetric cell division this could result in some cells having more CL than others. This could be enough CL enabling MgtA clustering to one cell pole, especially as not all cells exhibited polar MgtA localization in the Δ cls mutant. This is, however, purely speculative from my part. Another possibility could be the presence of PG. PG association has been found to promote MgtA activity as well, similar to CL but to a lesser extent¹⁶⁷. The same study that generates the Δ clsABC knockout also found that PG localized to the poles in the absence of CL. It is currently not possible at this time to state definitively whether CL impacts MgtA localization, and further experiments are required.

There are many unanswered questions when it comes to protein localization. While several proteins have been found to localize to the cell pole and some interact with CL, the actual role of CL remains unclear. CL was previously thought to play a role in shaping the bacterial cell due to its conical shape. If this were the case, a triple knockout for the main synthases of CL would result in an observable phenotype, given that multiple redundancies do not mask the effect. The cell only requires 5% CL and the amount of CL accumulates throughout the cell cycle, being more prominent in the stationary phase. Therefore most CL experiments require the bacteria to be above an optical density of 1. In **paper III** we avoided the stationary phase and only investigated early exponential (OD:0.3) as several peculiar effects occur in the stationary phase. Further, the exponential phase is better suited for protein expression studies and reproducibility. As a result, we experienced weak detection of CL (unpublished). Additionally, knocking out the main synthase for PG, *pgsA*, resulted in a slow-growing and fragile cell¹⁹⁷, whereas Δ *clsABC* was unaffected. These results, indicate that CL is not as involved in the cell shape as previously thought. Additionally, CL was recently found to be localized preferentially to membrane curvature after curvature was artificially induced¹⁹⁹. While, there is a lack of consensus within the field regarding CL and curvature, it is possible that the role of CL is to increase fluidity and that it is not specifically involved in shaping the cell.

Further, there is increasing evidence that bacteria utilize a wide variety transcription factors and regulatory RNAs, as well as small proteins to regulate gene expression and environmental response. Small proteins have previously been ignored due to the size cutoffs for genome annotation. However, they are unique in the sense that they consist of 16 – 50 amino acids in the absence of processing, as they have their own open reading frames (ORFs)²⁰⁰. About 30% of the identified small proteins in *E. coli* have been predicted to localize to the membrane, with some even having transmembrane domains. Small proteins can either have subcellular localization or be associated with other proteins. They have been found to integrate into the membrane and influence cellular processes by interacting with or stabilizing other membrane proteins. The small protein, MgtS, is expressed by *E. coli* and contains a hydrophobic α -helix that inserts into the inner membrane^{201,202}. It was previously named YneM before the interaction with MgtA was observed. Interaction with the 33 amino acid protein occurs at the membrane and increases MgtA activity when internal levels of Mg^{2+} are low. Further, it protects MgtA degradation from the membrane-bound protease, FtsH. Interestingly, the small protein was also seen to inhibit PitA when Mg^{2+} is low. PitA is a phosphate/metal transporter, also known as a cation symporter, present in the membrane, and it can dramatically decrease internal levels of Mg^{2+} ^{202,203}. However, MgtS limits Mg^{2+} efflux by PitA²⁰⁴. The increased stability, protection, and activity provided by MgtS could also mediate localization at

the cell pole. The protein – protein interaction would also require more space aided by CL.

While CL might not be responsible for inducing curvature, it can increase membrane fluidity, which mediates interactions between essential membrane-bound protein complexes and could aid in complex clustering. Several proteins are found to be recruited to the cell pole and yet how this occurs is still not fully understood. Two main models for polar localization have been proposed. Firstly, the ‘stochastic nucleation’ model suggests that receptors nucleate at the mid-cell rather than merging in polar clusters^{205,206}, eventually becoming the new cell pole after division. Another model suggests the slightly curved geometry of receptor clustering is attracted to the curves of the cell poles and aids in reducing curvature mismatch²⁰⁶. However, these models are directed at chemoreceptors that traverse the cell envelope and often demonstrate a conical shape where the widest part is in the outer membrane. The receptors are also often found as trimer dimerization. However, these models do not include the considerable amount of membrane protein traffic that occurs within the membrane. The rate of diffusion for eukaryotes is roughly $0.3 \text{ um}^2/\text{sec}$ ²⁰⁷. A study assuming a similar rate for *B. subtilis* estimated that it would take less than 1 min for a septal protein to diffuse to the cell pole^{208,209}. Further, MinD preferentially binds the inner membrane and shuttles between the two cell poles in an ATP-dependent manner. The pole-to-pole oscillation takes roughly 50 s to complete a full cycle¹⁴⁸. There is an enormous amount of traffic occurring within the membrane, and yet protein complexes are organized with preferential localization, where some proteins are septal while others are found at the pole. This could very well be a result of septal insertion and stochastic diffusion. However, signaling events and assembly of protein complexes occur rapidly, indicating an assist. Membrane organization is increasingly understood to be complex. Various lipids accumulate according to the growth phase and environment conditions. Additionally, protein complexes can be sorted based on cell division and protein activity is modulated by lipid interactions and other membrane proteins. The membrane is not simply a scaffold that holds proteins in place, but instead plays an active role in bacterial survival and environmental responses and as such merits further investigation. Further, the cell membrane plays a major role in antibiotic resistance. The cell can adjust its membrane composition to reduce uptake, increase efflux or generate biofilms that provide an additional “impenetrable” layer of protection¹².

7.2 Mechanistic models can limit antibiotic resistance

While the cell envelope is a major obstacle for antibiotic treatment, the bacteria can also inactivate the drug, and alter the target of the drug, as well as acquiring functionalities that mediate drug evasion. Antibiotic resistance is steadily increasing and remains a global challenge. In addition to the absence of novel antibiotic discovery, a proper understanding of drug-target interactions can alleviate some immediate obstacles for antibiotic treatment. Appropriate antibiotic treatment can limit the bacterial response by maximizing bacterial killing and reducing selection for resistance, as well as minimizing the opportunity for resistance to emerge and spread. Proper application of PK/PD knowledge can circumvent such unwanted antibacterial effects.

The success of the antibiotic treatment will depend on the concentration at the target site. Traditionally, the MIC has been used to describe the “potency” of the drug. However, this is a crude measurement with room for error. The MIC informs at what concentrations of drug results in the inhibition of the bacterial growth. The Clinical and Laboratory Standards Institute (CLSI) recommends an inoculum of 5×10^5 CFU/mL, or within the range of $2 \times 10^5 - 8 \times 10^5$ CFU/mL²¹⁰, and the antibiotic is often diluted in 2-fold increments. The standardized measurement is observed after 18 h \pm 2h at 34 – 37°C²¹¹. There is often lab-to-lab variation, and the outcome is dependent on the inoculum. If the inoculum is greater or the duration is longer, the MIC will likely be higher as well. The value is also not directly comparable to *in vivo* concentration during treatment and does not provide any information above or below the MIC.

The MIC is primarily used by diagnostic laboratories to determine resistances as well as *in vitro* activity of new antimicrobials. The MIC value is determined at a relatively low inoculum. Further, as the MIC is a single end-point measurement, the value does not consider important characteristics of bacterial growth and death under antibiotic pressure. This can, for instance, result in treatment failure during clinical decision-making due to improper characterization of the strain and ineffective dosing regimens. Time-kill curves, however, can capture the dynamic relationship between bacteria and drug concentration²¹². The effect of the drug is evaluated at several constant concentrations as a function of time and compared to the growth curve in the absence of the drug. Bactericidal drugs will have a negative growth rate, as there is a killing effect while pure bacteriostatic drugs will reduce the growth rate until it becomes zero. A downside to time-kill curves is that the concentration is constant, while the concentration fluctuates for *in vivo* conditions, limiting the predictive power²¹². Yet, subsequent PK/PD analysis of time-kill data does provide a meaningful understanding of the dynamic drug-target interaction, providing useful information during clinical trials²¹².

PK/PD models can benefit all phases of clinical trials, including preclinical trials. During phase I and II of clinical trials such models reduce the time and cost by informing proper dosing regimens, predicting activity, and mediating with determining the best-suited candidates. Our model is an initial framework intended to mediate phase I and II of clinical trials, as proper evaluation and initial stages can reduce future costly drug failure²¹³.

Determining the MIC is highly dependent on the test conditions, and can therefore differ greatly between labs. It is also a static measurement. However, performing time-kill curves is time-consuming. In **paper II** we proposed an alternative framework for characterizing pharmacodynamic profiles, grounded in drug-target interactions. To this end, we utilized readily available information regarding traditional drug compounds to construct a mechanistic model with predictive capability. During drug development, specific parameters that characterize the drug-target interaction are defined. One such biochemical property is the binding affinity of the drug to its target, which is typically displayed as a disassociation constant, K_D . A small value indicates a strong affinity between the drug and target. By evaluating different binding affinities against the corresponding MICs, my coauthors found a linear correlation between the two values, indicating a strong relationship between the two.

The model, generated by my coauthors in **paper II**, describes the action of the drug on the bacteria as a function of time. The drug's action is based on binding kinetics, i.e. the dynamics of drug-targets complexes using binding and unbinding rates of drug molecules with their targets. Initially, we assume that all of the drug molecules are unbound as well as all the targets inside the bacteria. A growing number of bound targets is related to a decrease of the growth rate (bacteriostatic effect) and/or to an increase of the killing rate (bactericidal effect). The model also considered bound target inheritance upon cell division, as well as a slowing of the growth rate as the population reaches the environmental carrying capacity. The model assumes the concentration outside and inside the bacteria is the same. As, the drug binds its target the concentration of the available drug will decrease, and when the drug unbinds its target the concentration can increase. The drug molecule will then have a probability of interaction with its, e.g. rate of interaction, and is defined by the binding rate. This is essentially a probability that demonstrates that it takes time for the drug to bind to its target. Due to the randomness of drug binding, the bacterial population will demonstrate different target occupancy. The more target that is bound, the stronger the effect the drug will have, until a threshold is reached. This percentage of target occupancy will also change in accordance to time.

The model was calibrated by time-kill curves where wild-type *E. coli* MG1655 was challenged in the presence of ciprofloxacin. This is a commonly used mixed-action drug, meaning it has both bactericidal and bacteriostatic effects. There are two known molecular targets of the drug: DNA gyrase (GyrA₂B₂) and DNA topoisomerase

IV (ParC₂E₂). However, the drug does preferentially bind the DNA gyrase complex in Gram-negative bacteria²¹⁴. The inoculum of the time-kill experiments, as well as each serial diluted time point, was performed in large volumes and triplicates. While this is laborious it does greatly diminish error. To best fit the model, several concentrations ranging from at least 0.5 – 100x MIC, as well as several time points, were required. According to the Eucast table, the MIC breakpoint for susceptible strains is 0.25 mg/L, if ciprofloxacin is used to treat meningitis, and 0.125 mg/L if it is used for a different treatment. In this case, resistance must be excluded. We expected resistance to be excluded for our wild-type strain, and evaluated the MIC by a 3-fold dilution of drug range from 0 – 2 mg/L ciprofloxacin. The MIC was found to be somewhere between 0.0081 mg/L and 0.027 mg/L. (The MIC was later determined to be 0.0139 mg/L, by the model.) This indicated that the solubility limit of the drug could potentially be just above 100 times the MIC required for the model. We also needed to perform a pilot time-kill to determine the appropriate time points to ensure proper sampling of the curve, to evaluate how the drug affected the change in growth rate. Lastly, some final time points were excluded due to the appearance of growth at maximum concentrations. This could be due to either experimental error, tolerance, or resistance however further characterization of this was not performed.

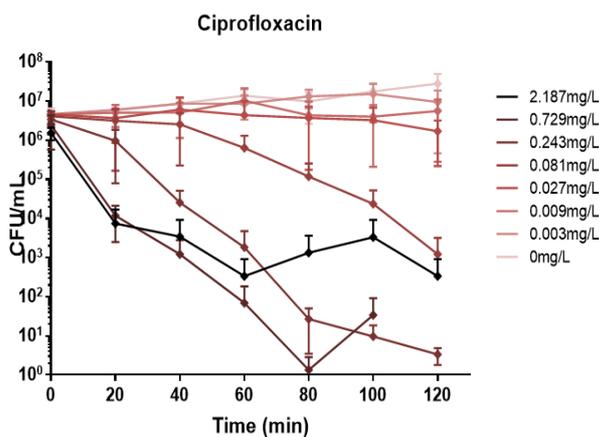


Figure 20 Time-kill curves challenging *E. coli* with ciprofloxacin

The data from the time-kill curves were used to inform how the drug affects the growth and death rates of the bacteria. **Paper II, Figure 3a** shows the model fit to experimental data while **Figure 20**, in this section, shows the raw data obtained from the time-kill curves. As the model is fully mechanistic, i.e. all parameters have a physical meaning, it can predict the time it takes for drug target occupancy to exert an observable effect. As the drug binds to its target this will impact the growth and death rate of the bacteria. The model estimates the

growth and death rate of the bacteria based on the effect of the drug used. The shape of this curve will describe the effect of the drug whether it is bacteriostatic or bactericidal. If the growth is zero the effect of the drug slows the growth rate to zero. If the death rate is increasing, the drug is bactericidal, confirming the drug has mixed action. The parameters for growth and death were kept separate in the model. There is a growth rate ranging between zero and the maximum growth, obtained with bacterial growth with no drug. While the death rate is between zero and a maximum death rate, describing how many bacteria perish per unit of time. While the maximum

death rate was determined by the highest drug concentration used in the kill curves. Therefore, the maximum concentration used should be much higher than the MIC (e.g. 50 or 100 times the MIC) to ensure the maximum death rate was achieved, however, one would still need to be able to collect a few samples. The concentrations in-between should provide a proper sampling of the final growth and death rates as a function of the drug concentration or bound targets (**Paper II, Figure 3b**).

The most commonly used PK/PD indices are; the amount of drug that stays within the therapeutic window (C_{max}/MIC), the unbound fraction of the drug (AUC/MIC), and the dosing interval where the amount of free drug in e.g. the serum, exceeds the MIC ($T > MIC$). The PK/PD indices are still highly reliant on the MIC, meaning the variability and drawbacks of MIC determination will propagate. We showed it is possible to evaluate drug efficacy by using other parameters than the MIC. The model provides information about resistance development given that it arises from an altered K_D . Different drug concentrations will affect the bacteria differently. While a high concentration will likely inhibit growth or ensure killing, such a dose can be toxic to the patient. Further, an appropriate drug concentration for *in vitro* killing may not be high enough to reach the site of infection *in vivo*. A rational dose regimen should therefore effectively treat the infection as well as prevent the emergence of resistance. My coauthors also coupled the model to a PK model and were able to inform dosing regimens (**Paper II, Figure S14**). The model allows for the prediction of the mutant selection window. Further showing that MSW is sensitive to timing (**Paper II, Figure 6**), as there will be a slight replication rate at initial dose administration making the drug effect less prominent. However, this changes as the drug associates with its target.

Mechanistic models have enormous potential for reducing the cost of drug development as well as optimizing the care of individual patients, by for example informed precision dosing²¹⁵. This is a MIC-independent approach that uses readily available drug parameters²¹⁶. The potential of such models can further be seen by a burst of new bacterial susceptibility testing using non-MIC methods²¹⁷. By modeling the dynamic relationship between the drug and bacteria, all the information is retained in a mechanistic PD model. To apply the PK/PD models to clinical practice more research is required. However, antibiotic treatments can benefit from a more complex evaluation and precise dosing approaches²¹⁶. Mechanistic models aimed to streamline new drug development and inform dosing regimens of old antibiotics can address the rise and spread of antibiotic resistance and treatment failure²¹⁸.

7.3 Amino acid metabolism modulates natural transformation

An important route for adaptation is HGT. Bacteria can acquire large stretches of genetic material that enable them to adapt to their environment. This can lead to emergence of new traits, which for pathogenic bacteria can increase infectivity as well as improve colonization, all of which can have devastating consequences for the human host. A mode of HGT is natural transformation, which entails active uptake of free DNA from the environment and subsequent genomic integration. In order for DNA uptake to occur, the bacteria must first enter the physiological state of competence. *V. cholerae* is able at acquiring new genetic information via natural transformation. Further, efficient shuffling or import of genetic material contribute to the success of the pathogen. From a medical perspective this has major socio-economic implications on vulnerable human communities that limit further development⁶²⁻⁶⁴. The current cholera pandemic is caused by the El Tor strains and began in 1961²¹⁹. The El Tor biotype is much more efficient at acquiring exogenous DNA via chitin-induced natural transformation than the classical O1 biotype. The prototypical classical strain lacks this ability due to a frameshift mutation in the regulatory gene *hapR*²²⁰. It has been speculated that the classical strain was outcompeted by El Tor strains, as a result of the reduced capacity to undergo natural transformation²²¹.

Natural transformation is widespread across species but also within the *Vibrio* genus. While it is not the sole driver it does contribute to the evolution of *Vibrio* species, resulting in genomic flexibility²²². While the natural transformation machinery has been partially characterized through genetic studies and orthology analysis, its full complexity remains elusive. To gain deeper insights, we employ transposon insertion sequencing. This high-throughput technique facilitates the evaluation of fitness contributions of genomic regions. The method is well suited to assess fitness landscapes as it directly links phenotype to genotype¹⁸².

In **paper I**, we find that amino acid metabolism genes modulate the progression of natural transformation. In short, we constructed a high-complexity near-saturation transposon insertion library in *V. cholerae*. The library was applied to an optimized natural transformation assay in the presence of chitin (to induce competence), and we compare how the disruption frequencies change in the presence and absence of donor DNA. This procedure allowed us to determine which genes were conditionally essential for DNA uptake and recombination. Analysis of the sequencing data showed that amino acid metabolism was a major pathway involved in DNA uptake and recombination. Deletion mutants of amino acid metabolism genes, *metE*, and *asnB*, resulted in severely reduced transformation frequencies. However, transformation proficiency was recovered by adding the corresponding amino acids, methionine, and asparagine, respectively. By extending our investigation to include fluorescence reporter constructs we could assess competence induction within these

strains. We showed the defective amino acid mutants were impacted in four of the major pathways associated with competence development: chitin metabolism, nucleotide scavenging, carbon catabolite repression (CCR), and quorum sensing, clearly indicating a regulatory role for the amino acid metabolism genes to achieve successful DNA uptake and chromosomal integration. We also find a greater variation of expression levels of genes critical for competence induction within the defective amino acid metabolism strains.

Another transposon insertion study investigating the lifestyle switches of *V. cholerae* between environments found an increased need for amino acid metabolism in pond water⁷¹. Other transposon insertion studies have also found amino acid metabolism to be enriched for *V. cholerae* in animal models^{183,223,224}. These studies further indicate that the host environment is a nutritionally limited environment, and the pathogen would need to switch metabolic focus. Pritchard et al. (2014) compare the input library created in enriched media with the limiting environment of a rabbit gastrointestinal tract, specifically passaging three samples through a rabbit infection model and pooling the sequencing output¹⁸³. In the study, they find several amino acid metabolism genes to be “essential”, which they suggest could be due insufficient supply of nutrients within the host. However, they compare two vastly different conditions: enriched media against a limiting environment, the host. According to their findings, amino acid metabolism is likely linked to a metabolic switch. Our study compares two conditions that experience the same environment, with the main differences being the presence of exogenous DNA and antibiotic selection. If amino acid metabolism was critical for survival on chitin in a limiting environment, the respective genes would be apparent in the control condition and removed from our screen. Instead, the result of our genetic screen suggests there is a connection between amino acid metabolism and natural transformation.

Interestingly, specific amino acids have been found to increase the levels of a major virulence regulator in *V. cholerae* known as ToxR^{225,226}. ToxR is a transmembrane transcription activator that through an intricate signaling cascade is involved in activating ToxT, a central virulence gene regulator^{227,228}. Furthermore the activity of the transcription factor increases in the presence of bile. Mey et al. (2015) noted that supplementing specifically with a mix of asparagine, arginine, glutamate, and serine increased the levels of ToxR expression beyond that of bile induction. This result indicates a potential role for amino acids in terms of environmental sensing and virulence induction, and the authors suggested this induction occurred via a different pathway than that of bile. They further pointed out that the protein composition of the outer membrane changed in response to amino acids and bile, likely as a result of the increased levels of ToxR.

They suggest that the response to amino acids and bile could be linked to a time-dependent mechanism that plays a role in safeguarding against antimicrobial

peptides in the small intestine. This suggests a wider range of functionality for amino acids beyond that of a conventional metabolic role.

Amino acids have also been seen to regulate natural transformation in the Gram-positive organism, *Micrococcus luteus*²²⁹. Lichev et al. (2019), found amino acids to act as nutritional factors or environmental cues, where individual amino acids had both stimulatory and inhibitory effects on competence induction²²⁹. They noticed a threshold-dependent response and observed a slight increase in transformation at low amino acid concentrations. They assume this stimulatory effect was a result of amino acids depletion (e.g. by metabolic processes) before transformation could occur, as the duration of their assay was 20 h. Furthermore, they demonstrated a link between natural transformation and the stringent response. The stringent response is a conserved stress response that enables bacteria to react towards nutrient deprivation²³⁰. Essentially, an alarmone, (p)ppGpp, downregulates factors involved in growth and upregulates stress response genes²³¹. In the *M. luteus* study they find that the alarmone partly directs natural transformation. The addition of external amino acids results in low levels of alarmone and decreased transformation frequency, whereas auxotrophy leads to reduced transformation as well due to a high level of alarmone.

Stimulatory and inhibitory effects of amino acids on competence development have also been described in *B. subtilis*, by Wilson et al. (1968)²³². Glutamic acid and alanine were found to be detrimental to competence development, while the amino acids histidine, tryptophan, arginine, valine, lysine, threonine, glycine, aspartic acid, and methionine had stimulatory effects. The highest level of competence induction when all the stimulatory amino acids were added as a mixture. Competence development in *B. subtilis* is growth dependent and occurs in the late exponential phase/stationary phase²³³. Competence is further a result of bistable expression of *comK*, the master regulator for competence for *B. subtilis*. Fluctuating levels of its expression will determine the fate of the cell. The population has generally low expression levels of *comK*, as it is under the control of a weak promoter. However, due to intrinsic noise, the cell-to-cell variation is high²³⁴. In *B. subtilis*, competence development is likely not linked to catabolite repression to the same extent as competence development and CCR in *V. cholerae*. At the same time, the stimulatory effect of amino acids on competence induction within *B. subtilis* would suggest a role at least in parallel with metabolic activity. Additionally, upon supplementing *B. subtilis* with hydrolyzed casein, or casamino acids, the authors also detected high variation within competence development²³².

Lichev et al. and Wilson et al. are the only studies we could find that had investigated the effects of amino acids on natural transformation, and both studies were performed in Gram-positives. In **Paper I** we found a concentration- and time-dependent reduction of transformation frequencies upon addition of external amino

acids. We argue that optimal amino acid metabolism is required for competence development and find that the competence induce bacteria will rapidly internalize DNA and proceed to recombination before inhibitory substances exert an effect. This could be a strategy to sequestering DNA as a resource. Further, we find a cell-to-cell variation that could indicate there is a specific threshold for gene expression to become competent and that this process is stabilized by amino acid metabolizing genes. This could be a way to subdivide the population where only a fraction will attempt DNA internalization and recombination.

In **Paper I** we found substantial cell-to-cell variation for competence expression for the defective amino acid metabolism mutants tested, $\Delta metE$ and $\Delta asnB$. This could suggest that amino acid metabolism in *V. cholerae* has a role in competence induction, potentially explaining why the transformation frequencies were severely impaired. Interestingly the fluorescent reporter constructs showed a similar trend in promoter activity as the transformation frequencies. The $\Delta tfoX$ mutant demonstrated background signaling close to our detection limit with minimal cell-to-cell variation. Other mutants, $\Delta metE$ and $\Delta asnB$, displayed greater variation within the promoter activity of the competence genes. Intriguingly, the $\Delta asnB$ strain only just managed to achieve transformation within the detection limit of our natural transformation screen, where the promoter activity was just above that of the $\Delta tfoX$ mutant. The $\Delta asnB$ mutant demonstrated only a few cells with strong *comEA* promoter activity. The cell-to-cell variation could be what enables the remaining transformation frequencies. We therefore interpreted these results as a threshold requirement for competence induction, aided by functional amino acid metabolism.

When adding casamino acids to the wild-type we found a slight increase in transformation frequencies between the 24 h and 30 h timepoints. Casamino acids were found to inhibit transformation frequencies at early stages of our assay. We associate this to competence development; it is therefore interesting that there is a minor increase in transformation frequencies at the later stages. We did not see any major growth differences between these timepoints and have previously not detected increased CFUs between initial and final timepoints during the natural transformation assay, so the increase in transformation frequency is likely not due to cell growth. When we added DNase at different time points post DNA induction, we found the majority of DNA was taken up during the first hour however there was a slight increase in transformation frequencies for the later time points as well. We suggest this is a result of recombination. We show that DNA uptake is rapid, however recombination takes longer²³⁵. Dalia et al. demonstrated that *V. cholerae* can undergo several recombination attempts if the first fails, as most attempt fail⁸². To reach peak transformation frequencies several integration attempts may be required. This could be the reason we find a slight increase during the casamino and DNase experiments.

Interestingly, Dalia et al. find that the newly incorporated DNA is synthesized before cell division and test this further by implementing antibiotic resistance markers as transforming DNA. They find the newly synthesized antibiotic resistance gene product is synthesized and stochastically distributed upon cell division. After cell division one cell will have the transformed DNA incorporated into the genome, while the other will have the parental DNA. However, both will have a random distribution of gene product, allowing for nongenetic transfer of gene products. It is thought that competence development and natural transformation is a bet-hedging strategy because the competence state results in a fitness cost^{59,236}, and is a way to safeguard the population from deleterious mutations.

In our reporter assay only half of the wild-type population exhibits detectable *comEA* promoter activity, while the transformation frequencies are less than 5% of the population in optimal conditions. In such a scenario either all cells have the potential to become competent or some cells are predisposed to become competent. The probability would be either based on reaching a threshold for gene expression relevant for competence or an inherent difference. We suggest in **Paper I** that there is a subpopulation of competent cells where a threshold connected to amino acid metabolism, could act as a regulatory checkpoint. I would like to build on this scenario and place it in an evolutionary context. Is it conceivable that natural transformation is a bet-hedging strategy to limit the fitness cost of competence development and potentially detrimental transforming DNA. Additionally, amino acid metabolism has inherent heterogeneity which results in differences within an isogenic population. Within the context of intrinsic noise, it is tempting to envision a more direct role for amino acids in terms of competence induction and natural transformation.

A speculative interpretation of the results in **paper I** could explain why only a small fraction of the *V. cholerae* population become competent, not simply suggesting a bet-hedging strategy but also implying an underlying mechanism. A stochastic availability of amino acids could prime a subset of the population for DNA uptake, before induction of competence by chitin actually leads to expression of the DNA uptake machinery. Ultimately the underlying metabolic foundation determines the cell fate. The regulatory circuitry of natural transformation has in broad strokes been characterized for *V. cholerae*. La Scrudato et al. developed a model for the regulation of competence and natural transformation²³⁷. In that study the authors used plasmids with transcriptional reporter constructs of (red) fluorescent protein genes under control of promoters of competence genes, one of which being the promoter for the *comEA* competence gene, included an internal control construct, a green fluorescent protein under control of a constitutive housekeeping (*gyrA*) promoter. The authors employed both flow cytometry and microscopy for all reporter constructs included in the paper. Based on the representative images there is clear cell-to-cell variation present for the [P_{comEA}]-dsRED expression. This is to be expected as our study uses

the same strain and reporter construct, generously donated by M. Blokesch (Lausanne). Our results in part align with this study. In the paper, they interpreted the heterogeneity as a response to the disorganized topography of chitin, even though, they used chitin-coated beads which would be more uniform than environmentally available chitin. Moreover, the authors indicate a homogenous competence induction upon artificial induction of the master regulator *tfoX*. While the artificial induction of *tfoX* does greatly increase competence and transformation frequencies the representative images still display a slight variation in the [P_{comEA}]-dsRED expression, albeit diminished. It is possible that overexpression of *tfoX* would bypass the required threshold for underlying amino acid metabolism. It is possible that *V. cholerae* could benefit from constant surveillance for environmental DNA, as it experiences a range of environments. However, it would be energetically costly and ultimately inefficient for the entire population to acquire DNA. This consideration implies a need for spatiotemporal regulation provided by asymmetric cell division and stochastic metabolic noise. As we find an involvement of amino acid metabolism, this could add a level of regulation and result in simple fate determination.

Altogether our findings suggest amino acid metabolism could prepare cells for DNA uptake. This assumption is dependent on varied metabolic strategies within a population²³⁸. These considerations would align with a bet-hedging strategy where only a fraction of the population actually samples the available DNA, any adverse effects would be outcompeted, while potentially beneficial mutations provide improved fitness²³⁹. Amino acid metabolism could maintain stochastic signaling pathways allowing for cells to overcome this threshold. Isogenic populations consistently feature phenotypic variations, even in homogenous environments³⁸. This is generally a result of propagated molecular noise (transcriptional variance), transcriptional feedback, and asymmetric partitioning, all of which produce metabolic heterogeneity^{38,39,44}. Furthermore, essential genes are often under strict control whereas metabolic genes exhibit greater heterogeneity in expression⁴⁵, allowing for increased variation within metabolic activity³⁸. We therefore, suspect amino acid metabolism could be the determinant for which bacteria become competent and undergo successful DNA uptake and recombination. Our results hint towards there being a subpopulation that will attempt DNA uptake and integration regardless of nutritional availability in the presence of donor DNA. Our results suggest amino acid metabolism stabilizes the progression of natural transformation. Successful DNA integration is dependent on overcoming a competence threshold. In strains with defective amino acid metabolism, transformation was still partially detected which could be aided by the visible cell-to-cell variation. Intrinsic noise within amino acid metabolism could explain this variation, resulting in a fraction of cells that breach the threshold, as they are predisposed, achieving successful DNA uptake and integration.

Of course, as always: Further experiments are required.

8 Conclusion

Bacteria are complex systems that integrate feedback from various components. Eventually, the sum of the inputs will reach a threshold and a different process will be implemented. Signal feedback and intercepting pathways make up the regulatory network of bacteria, and the adaptive processes of bacteria tend to overlap. Further, regulatory cascades and proteins often have pleiotropic effects. By connecting pathways through positive and negative feedback, shared regulators can coordinate and initiate an appropriate response to the immediate environment. Natural transformation is an excellent example of this as the machinery is dependent on the interplay of several regulatory pathways, such as quorum sensing, chitin-dependent competence induction, and carbon catabolite repression²⁴⁰. While this can be efficient enable fast change, it also maintains a metabolic budget.

Dynamic transporters present in the membrane, such as MgtA, interact with its lipid environment in a flexible and adaptive manner. Understanding how the membrane components interact and modulate each other has implications for drug development, not only by identifying novel drug targets but also inhibiting core complexes. The bacterial membrane is the foundation of survival and is a major barrier to overcome. While there are several ways for bacteria to circumvent drug therapies, this can be bypassed by a proper understanding of drug-to-target dynamics. By modeling how the bacterial population will interact with the drug, it is possible to avoid selection pressure that promote resistance. Due to the intrinsic noise within a population, implementing more robust parameters such as binding kinetics could further improve dosing and minimize the amount of failed clinical trials.

While natural transformation is not as efficient as mobile genetic elements or conjugation in terms of spreading resistance, it does allow for the inheritance of both genetic resistance and nongenetic tolerance. This potential phenomenon makes natural transformation a contender that needs to be dealt with. Due to the reliance on amino acid metabolism within bacteria, it has been suggested that drugs targeting metabolic pathways would promote the killing of antibiotic-resistant bacteria. While it has not been shown that natural transformation can occur within the small intestine, it is hypothesized, as there is overlap in gene expression involved in DNA uptake. Targeting metabolic pathways such as amino acid synthesis could also repress transformation as an added benefit. Proper knowledge of bacterial machinery is the foundation of novel drug discovery and proper treatment strategies, and aids in understanding the adaptive evolutionary mechanisms of bacteria.

9 Outlook

Further investigation is needed to address the impact of amino acid metabolism on natural transformation. Amino acid metabolism is connected to the central metabolism, and has a multitude of redundancies. However, there are some flanking branches that could allow for the production of some additional auxotrophs. While there is little overall growth on chitin, it is difficult to disentangle dormant bacteria as a result of metabolic inactivity and a true effect. It would therefore be interesting to investigate the transcriptome by implementing RNAseq on the wild-type strain, and potentially one of the mutants, either $\Delta metE$ or $\Delta asnB$. It would also be possible to perform a transposon insertion sequencing on one of the mutants to investigate how other pathways compensate. Additionally, it would be interesting to quantify internal levels of amino acids by comparing an enriched media with a minimal media in the presence of chitin, using mass spectroscopy. Further, it could be beneficial to investigate whether transformation within *B. subtilis* is also dependent on amino acid metabolism. As this organism is better characterized than *V. cholerae*, a similar dependence could help guide the direction of the study.

Localization of the natural transformation machinery has gained some attention in recent years, however anchoring, functionality, and the role of the membrane are yet to be defined. The individual components of the natural transformation machinery complex have been suggested to follow a “diffusion and capture” model. It would be interesting to test this hypothesis. In the natural transformation study, we generate a *pilW* knockout mutant, without which transformation drops dramatically. This protein has homology with the *N. gonorrhoeae pilW*, which functions in both pilus formation and DNA uptake. Additionally stabilizing *pilQ* in the membrane. An additional gene *pilF* is suggested to be a piloline that can move between the inner and outer membrane. There is a greater role for membrane dynamics in terms of DNA uptake which deserves a closer look. Additionally, the organism also defines its environment through the combination of the type VI secretion system and the DNA uptake pilus can function in kin recognition²⁴¹. Are the two machineries located in proximity of each other so that once a foreign neighbor is killed the DNA can be taken up more readily? There are also indications of the machinery being primed or semi-assembled in the membrane, however, what is the actual triggering event for assembly? Is DNA sensed, somehow? *V. cholerae* is rather promiscuous when it comes to DNA uptake as it does not require specific sequences, and only minor homology is required for integration. Is there a characteristic of DNA that is recognized?

It would further be interesting to characterize the observation that *V. cholerae* demonstrated tolerance after incubation on chitin, both in regards to UV light and low antibiotic concentrations. While this was a great source of frustration, it was also intriguing that this tolerance to low-level antibiotics in some cases resulted in

permanent resistance to high levels of antibiotics. While the low level of antibiotics was enough to kill the wild-type bacteria when passaged in the absence of chitin, the same concentration was insufficient for killing after bacteria has been incubated in the presence of chitin. Based on my experience, I would expect *V.cholerae* to exhibit an increased tolerance to several stress factors after being incubated on chitin. Especially, as my preliminary experiments investigating “chitin-dependent” tolerance, demonstrated more that a log higher survival in CFUs, compared to the LB cultured control, upon UV-radiation. While the implications for a chitin-induced state of tolerance is unclear it merits investigation, and I hypothesize that it will apply to survival in low pH, which could have implications for acid tolerance as the pathogen can be associated with chitin upon ingestion.

10 References

- 1 Almagro-Moreno, S., Pruss, K. & Taylor, R. K. Intestinal Colonization Dynamics of *Vibrio cholerae*. *PLoS Pathog* **11**, e1004787 (2015). <https://doi.org:10.1371/journal.ppat.1004787>
- 2 Abuaita, B. H. & Withey, J. H. Bicarbonate Induces *Vibrio cholerae* virulence gene expression by enhancing ToxT activity. *Infect Immun* **77**, 4111-4120 (2009). <https://doi.org:10.1128/IAI.00409-09>
- 3 Bharati, K. & Ganguly, N. K. Cholera toxin: a paradigm of a multifunctional protein. *Indian J Med Res* **133**, 179-187 (2011).
- 4 Nielsen, A. T. *et al.* RpoS controls the *Vibrio cholerae* mucosal escape response. *PLoS Pathog* **2**, e109 (2006). <https://doi.org:10.1371/journal.ppat.0020109>
- 5 Antimicrobial Resistance, C. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet* **399**, 629-655 (2022). [https://doi.org:10.1016/S0140-6736\(21\)02724-0](https://doi.org:10.1016/S0140-6736(21)02724-0)
- 6 Lidstrom, M. E. & Konopka, M. C. The role of physiological heterogeneity in microbial population behavior. *Nat Chem Biol* **6**, 705-712 (2010). <https://doi.org:10.1038/nchembio.436>
- 7 Ackermann, M. A functional perspective on phenotypic heterogeneity in microorganisms. *Nat Rev Microbiol* **13**, 497-508 (2015). <https://doi.org:10.1038/nrmicro3491>
- 8 Jung, K., Fabiani, F., Hoyer, E. & Lassak, J. Bacterial transmembrane signalling systems and their engineering for biosensing. *Open Biol* **8** (2018). <https://doi.org:10.1098/rsob.180023>
- 9 Giles, D. K., Hankins, J. V., Guan, Z. & Trent, M. S. Remodelling of the *Vibrio cholerae* membrane by incorporation of exogenous fatty acids from host and aquatic environments. *Mol Microbiol* **79**, 716-728 (2011). <https://doi.org:10.1111/j.1365-2958.2010.07476.x>
- 10 Winter, M., Buckling, A., Harms, K., Johnsen, P. J. & Vos, M. Antimicrobial resistance acquisition via natural transformation: context is everything. *Curr Opin Microbiol* **64**, 133-138 (2021). <https://doi.org:10.1016/j.mib.2021.09.009>
- 11 Sinha-Ray, S., Alam, M. T., Bag, S., Morris, J. G., Jr. & Ali, A. Conversion of a recA-Mediated Non-toxigenic *Vibrio cholerae* O1 Strain to a Toxigenic Strain Using Chitin-Induced Transformation. *Front Microbiol* **10**, 2562 (2019). <https://doi.org:10.3389/fmicb.2019.02562>
- 12 Carey, A. B., Ashenden, A. & Köper, I. Model architectures for bacterial membranes. *Biophysical Reviews* **14**, 111-143 (2022). <https://doi.org:10.1007/s12551-021-00913-7>
- 13 Schönknecht, G. *et al.* Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. *Science* **339**, 1207-1210 (2013). <https://doi.org:10.1126/science.1231707>
- 14 Gogarten, J. P. & Townsend, J. P. Horizontal gene transfer, genome innovation and evolution. *Nat Rev Microbiol* **3**, 679-687 (2005). <https://doi.org:10.1038/nrmicro1204>
- 15 Lercher, M. J. & Pál, C. Integration of Horizontally Transferred Genes into Regulatory Interaction Networks Takes Many Million Years. *Molecular Biology and Evolution* **25**, 559-567 (2007). <https://doi.org:10.1093/molbev/msm283>
- 16 Woods, L. C. *et al.* Horizontal gene transfer potentiates adaptation by reducing selective constraints on the spread of genetic variation. *Proc Natl Acad Sci U S A* **117**, 26868-26875 (2020). <https://doi.org:10.1073/pnas.2005331117>
- 17 Heidelberg, J. F. *et al.* DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**, 477-483 (2000). <https://doi.org:10.1038/35020000>

- 18 Wagner, A. Distributed robustness versus redundancy as causes of mutational robustness. *Bioessays* **27**, 176-188 (2005). <https://doi.org:10.1002/bies.20170>
- 19 Wakeley, J. The limits of theoretical population genetics. *Genetics* **169**, 1-7 (2005). <https://doi.org:10.1093/genetics/169.1.1>
- 20 Drake, J. W., Charlesworth, B., Charlesworth, D. & Crow, J. F. Rates of spontaneous mutation. *Genetics* **148**, 1667-1686 (1998). <https://doi.org:10.1093/genetics/148.4.1667>
- 21 Tröbner, W. & Piechocki, R. Selection against hypermutability in Escherichia coli during long term evolution. *Mol Gen Genet* **198**, 177-178 (1984). <https://doi.org:10.1007/bf00328720>
- 22 Litvak, Y. A.-O. & Bäumlner, A. A.-O. The founder hypothesis: A basis for microbiota resistance, diversity in taxa carriage, and colonization resistance against pathogens.
- 23 Kimura, M. Evolutionary rate at the molecular level. *Nature* **217**, 624-626 (1968). <https://doi.org:10.1038/217624a0>
- 24 Gasperotti, A., Brameyer, S., Fabiani, F. & Jung, K. Phenotypic heterogeneity of microbial populations under nutrient limitation. *Curr Opin Biotechnol* **62**, 160-167 (2020). <https://doi.org:10.1016/j.copbio.2019.09.016>
- 25 Papenfort, K. & Bassler, B. L. Quorum sensing signal-response systems in Gram-negative bacteria. *Nature Reviews Microbiology* **14**, 576-588 (2016). <https://doi.org:10.1038/nrmicro.2016.89>
- 26 Epstein, P. R. Cholera and the environment. *Lancet* **339**, 1167-1168 (1992).
- 27 Bridges, A. A. & Bassler, B. L. The intragenus and interspecies quorum-sensing autoinducers exert distinct control over Vibrio cholerae biofilm formation and dispersal. *PLoS Biol* **17**, e3000429 (2019). <https://doi.org:10.1371/journal.pbio.3000429>
- 28 Gorelik, O. *et al.* Vibrio cholerae autoinducer-1 enhances the virulence of enteropathogenic Escherichia coli. *Scientific Reports* **9**, 4122 (2019). <https://doi.org:10.1038/s41598-019-40859-1>
- 29 Suckow, G., Seitz, P. & Blokesch, M. Quorum sensing contributes to natural transformation of Vibrio cholerae in a species-specific manner. *J Bacteriol* **193**, 4914-4924 (2011). <https://doi.org:10.1128/jb.05396-11>
- 30 Taniguchi, Y. *et al.* Quantifying E. coli Proteome and Transcriptome with Single-Molecule Sensitivity in Single Cells. *Science* **329**, 533-538 (2010). <https://doi.org:10.1126/science.1188308>
- 31 Acar, M., Mettetal, J. T. & van Oudenaarden, A. Stochastic switching as a survival strategy in fluctuating environments. *Nature Genetics* **40**, 471-475 (2008). <https://doi.org:10.1038/ng.110>
- 32 Xie, X. S., Cai, L. & Friedman, N. Stochastic protein expression in individual cells at the single molecule level. *Nature* **440**, 358-362 (2006). <https://doi.org:10.1038/nature04599>
- 33 Baudrimont, A., Jaquet, V., Wallerich, S., Voegeli, S. & Becskei, A. Contribution of RNA Degradation to Intrinsic and Extrinsic Noise in Gene Expression. *Cell Rep* **26**, 3752-3761.e3755 (2019). <https://doi.org:10.1016/j.celrep.2019.03.001>
- 34 Jones, D. & Elf, J. Bursting onto the scene? Exploring stochastic mRNA production in bacteria. *Curr Opin Microbiol* **45**, 124-130 (2018). <https://doi.org:10.1016/j.mib.2018.04.001>
- 35 Chong, S., Chen, C., Ge, H. & Xie, X. S. Mechanism of transcriptional bursting in bacteria. *Cell* **158**, 314-326 (2014). <https://doi.org:10.1016/j.cell.2014.05.038>

- 36 Wolf, D. M., Vazirani, V. V. & Arkin, A. P. Diversity in times of adversity: probabilistic strategies in microbial survival games. *J Theor Biol* **234**, 227-253 (2005). <https://doi.org:10.1016/j.jtbi.2004.11.020>
- 37 Kussell, E. & Leibler, S. Phenotypic Diversity, Population Growth, and Information in Fluctuating Environments. *Science* **309**, 2075-2078 (2005). <https://doi.org:10.1126/science.1114383>
- 38 Evans, T. D. & Zhang, F. Bacterial metabolic heterogeneity: origins and applications in engineering and infectious disease. *Curr Opin Biotechnol* **64**, 183-189 (2020). <https://doi.org:10.1016/j.copbio.2020.04.007>
- 39 Şimşek, E. & Kim, M. The emergence of metabolic heterogeneity and diverse growth responses in isogenic bacterial cells. *The ISME Journal* **12**, 1199-1209 (2018). <https://doi.org:10.1038/s41396-017-0036-2>
- 40 Spratt Madison, R. & Lane, K. Navigating Environmental Transitions: the Role of Phenotypic Variation in Bacterial Responses. *mBio* **13**, e02212-02222 (2022). <https://doi.org:10.1128/mbio.02212-22>
- 41 Solopova, A. *et al.* Bet-hedging during bacterial diauxic shift. *Proc Natl Acad Sci U S A* **111**, 7427-7432 (2014). <https://doi.org:10.1073/pnas.1320063111>
- 42 Mustafi, N., Grünberger, A., Kohlheyer, D., Bott, M. & Frunzke, J. The development and application of a single-cell biosensor for the detection of l-methionine and branched-chain amino acids. *Metab Eng* **14**, 449-457 (2012). <https://doi.org:10.1016/j.ymben.2012.02.002>
- 43 Hernandez-Valdes, J. A., van Gestel, J. & Kuipers, O. P. A riboswitch gives rise to multi-generational phenotypic heterogeneity in an auxotrophic bacterium. *Nature Communications* **11**, 1203 (2020). <https://doi.org:10.1038/s41467-020-15017-1>
- 44 Soltani, M., Vargas-Garcia, C. A., Antunes, D. & Singh, A. Intercellular Variability in Protein Levels from Stochastic Expression and Noisy Cell Cycle Processes. *PLoS Comput Biol* **12**, e1004972 (2016). <https://doi.org:10.1371/journal.pcbi.1004972>
- 45 Silander, O. K. *et al.* A genome-wide analysis of promoter-mediated phenotypic noise in *Escherichia coli*. *PLoS Genet* **8**, e1002443 (2012). <https://doi.org:10.1371/journal.pgen.1002443>
- 46 Magasanik, B. Catabolite repression. *Cold Spring Harb Symp Quant Biol* **26**, 249-256 (1961). <https://doi.org:10.1101/sqb.1961.026.01.031>
- 47 Boulineau, S. *et al.* Single-cell dynamics reveals sustained growth during diauxic shifts. *PLoS One* **8**, e61686 (2013). <https://doi.org:10.1371/journal.pone.0061686>
- 48 Morawska, L. P., Hernandez-Valdes, J. A. & Kuipers, O. P. Diversity of bet-hedging strategies in microbial communities-Recent cases and insights. *WIREs Mech Dis* **14**, e1544 (2022). <https://doi.org:10.1002/wsbm.1544>
- 49 Kotte, O., Volkmer, B., Radzikowski, J. L. & Heinemann, M. Phenotypic bistability in *Escherichia coli*'s central carbon metabolism. *Mol Syst Biol* **10**, 736 (2014). <https://doi.org:10.15252/msb.20135022>
- 50 Merrell, D. S. & Camilli, A. The *cadA* gene of *Vibrio cholerae* is induced during infection and plays a role in acid tolerance. *Mol Microbiol* **34**, 836-849 (1999). <https://doi.org:10.1046/j.1365-2958.1999.01650.x>
- 51 Angelichio, M. J., Merrell, D. S. & Camilli, A. Spatiotemporal analysis of acid adaptation-mediated *Vibrio cholerae* hyperinfectivity. *Infect Immun* **72**, 2405-2407 (2004). <https://doi.org:10.1128/iai.72.4.2405-2407.2004>
- 52 Waldor, M. K. & Mekalanos, J. J. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**, 1910-1914 (1996). <https://doi.org:10.1126/science.272.5270.1910>

- 53 Pál, C., Papp B Fau - Lercher, M. J. & Lercher, M. J. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer.
- 54 Sun, D. Pull in and Push Out: Mechanisms of Horizontal Gene Transfer in Bacteria. *Front Microbiol* **9**, 2154 (2018). <https://doi.org/10.3389/fmicb.2018.02154>
- 55 Moradigaravand, D. & Engelstädter, J. The Evolution of Natural Competence: Disentangling Costs and Benefits of Sex in Bacteria. *Am Nat* **182**, E112-E126 (2013). <https://doi.org/10.1086/671909>
- 56 Redfield, R. J., Schrag, M. R. & Dean, A. M. The evolution of bacterial transformation: sex with poor relations. *Genetics* **146**, 27-38 (1997). <https://doi.org/10.1093/genetics/146.1.27>
- 57 Goodman, S. D. & Scocca, J. J. Identification and arrangement of the DNA sequence recognized in specific transformation of *Neisseria gonorrhoeae*. *Proc Natl Acad Sci U S A* **85**, 6982-6986 (1988). <https://doi.org/10.1073/pnas.85.18.6982>
- 58 Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C. Y. & Schoolnik, G. K. Chitin induces natural competence in *Vibrio cholerae*. *Science* **310**, 1824-1827 (2005). <https://doi.org/10.1126/science.1120096>
- 59 Johnsen, P. J., Dubnau, D. & Levin, B. R. Episodic selection and the maintenance of competence and natural transformation in *Bacillus subtilis*. *Genetics* **181**, 1521-1533 (2009). <https://doi.org/10.1534/genetics.108.099523>
- 60 Johnston, C., Martin, B., Fichant, G., Polard, P. & Claverys, J. P. Bacterial transformation: distribution, shared mechanisms and divergent control. *Nat Rev Microbiol* **12**, 181-196 (2014). <https://doi.org/10.1038/nrmicro3199>
- 61 Blokesch, M. Natural competence for transformation. *Current Biology* **26**, R1126-R1130 (2016). [https://doi.org:https://doi.org/10.1016/j.cub.2016.08.058](https://doi.org/https://doi.org/10.1016/j.cub.2016.08.058)
- 62 Ali, M. *et al.* The global burden of cholera. *Bull World Health Organ* **90**, 209-218a (2012). <https://doi.org/10.2471/blt.11.093427>
- 63 Ali, M., Nelson, A. R., Lopez, A. L. & Sack, D. A. Updated global burden of cholera in endemic countries. *PLoS Negl Trop Dis* **9**, e0003832 (2015). <https://doi.org/10.1371/journal.pntd.0003832>
- 64 Ganesan, D., Gupta, S. S. & Legros, D. Cholera surveillance and estimation of burden of cholera. *Vaccine* **38 Suppl 1**, A13-a17 (2020). <https://doi.org/10.1016/j.vaccine.2019.07.036>
- 65 Fuesslin, V. *et al.* Prediction of Antibiotic Susceptibility Profiles of *Vibrio cholerae* Isolates From Whole Genome Illumina and Nanopore Sequencing Data: CholerAegon. *Front Microbiol* **13**, 909692 (2022). <https://doi.org/10.3389/fmicb.2022.909692>
- 66 De, R. Mobile Genetic Elements of *Vibrio cholerae* and the Evolution of Its Antimicrobial Resistance. *Frontiers in Tropical Diseases* (2021). <https://doi.org/10.3389/fitd.2021.691604>
- 67 Teschler, J. K. *et al.* Living in the matrix: assembly and control of *Vibrio cholerae* biofilms. *Nat Rev Microbiol* **13**, 255-268 (2015). <https://doi.org/10.1038/nrmicro3433>
- 68 Zhu, J. & Mekalanos, J. J. Quorum sensing-dependent biofilms enhance colonization in *Vibrio cholerae*. *Dev Cell* **5**, 647-656 (2003).
- 69 Freter, R., O'Brien, P. C. & Macsai, M. S. Effect of chemotaxis on the interaction of cholera vibrios with intestinal mucosa. *Am J Clin Nutr* **32**, 128-132 (1979).
- 70 Butler, S. M. & Camilli, A. Going against the grain: chemotaxis and infection in *Vibrio cholerae*. *Nat Rev Microbiol* **3**, 611-620 (2005). <https://doi.org/10.1038/nrmicro1207>

- 71 Kamp, H. D., Patimalla-Dipali, B., Lazinski, D. W., Wallace-Gadsden, F. & Camilli, A. Gene fitness landscapes of *Vibrio cholerae* at important stages of its life cycle. *PLoS Pathog* **9**, e1003800 (2013). <https://doi.org:10.1371/journal.ppat.1003800>
- 72 Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C.-Y. & Schoolnik, G. K. Chitin Induces Natural Competence in *Vibrio cholerae*. *Science* **310**, 1824-1827 (2005). <https://doi.org:10.1126/science.1120096>
- 73 Blokesch, M. Chitin colonization, chitin degradation and chitin-induced natural competence of *Vibrio cholerae* are subject to catabolite repression. *Environ Microbiol* **14**, 1898-1912 (2012). <https://doi.org:10.1111/j.1462-2920.2011.02689.x>
- 74 Metzger, L. C. & Blokesch, M. Regulation of competence-mediated horizontal gene transfer in the natural habitat of *Vibrio cholerae*.
- 75 Blokesch, M. & Schoolnik, G. K. The extracellular nuclease Dns and its role in natural transformation of *Vibrio cholerae*. *J Bacteriol* **190**, 7232-7240 (2008). <https://doi.org:10.1128/jb.00959-08>
- 76 Dong, T. G., Ho, B. T., Yoder-Himes, D. R. & Mekalanos, J. J. Identification of T6SS-dependent effector and immunity proteins by Tn-seq in *Vibrio cholerae*. *Proc Natl Acad Sci U S A* **110**, 2623-2628 (2013). <https://doi.org:10.1073/pnas.1222783110>
- 77 Russell, A. B., Peterson, S. B. & Mougous, J. D. Type VI secretion system effectors: poisons with a purpose. *Nat Rev Microbiol* **12**, 137-148 (2014). <https://doi.org:10.1038/nrmicro3185>
- 78 Borgeaud, S., Metzger, L. C., Scignari, T. & Blokesch, M. The type VI secretion system of *Vibrio cholerae* fosters horizontal gene transfer.
- 79 Seitz, P. & Blokesch, M. DNA-uptake machinery of naturally competent *Vibrio cholerae*. *Proc Natl Acad Sci U S A* **110**, 17987-17992 (2013). <https://doi.org:10.1073/pnas.1315647110>
- 80 Matthey, N. & Blokesch, M. The DNA-Uptake Process of Naturally Competent *Vibrio cholerae*. *Trends Microbiol* **24**, 98-110 (2016). <https://doi.org:10.1016/j.tim.2015.10.008>
- 81 Seitz, P. *et al.* COMEA is essential for the transfer of external DNA into the periplasm in naturally transformable *Vibrio cholerae* cells.
- 82 Dalia, A. B. & Dalia, T. N. Spatiotemporal Analysis of DNA Integration during Natural Transformation Reveals a Mode of Nongenetic Inheritance in Bacteria. *Cell* **179**, 1499-1511 e1410 (2019). <https://doi.org:10.1016/j.cell.2019.11.021>
- 83 Dalia, A. B., Seed, K. D., Calderwood, S. B. & Camilli, A. A globally distributed mobile genetic element inhibits natural transformation of *Vibrio cholerae*. *Proc Natl Acad Sci U S A* **112**, 10485-10490 (2015). <https://doi.org:10.1073/pnas.1509097112>
- 84 Huq, A. *et al.* Simple sari cloth filtration of water is sustainable and continues to protect villagers from cholera in Matlab, Bangladesh. *mBio* **1** (2010). <https://doi.org:10.1128/mBio.00034-10>
- 85 Santacroce, L., Del Prete, R., Charitos, I. A. & Bottalico, L. Mycobacterium leprae: A historical study on the origins of leprosy and its social stigma. *Infez Med* **29**, 623-632 (2021). <https://doi.org:10.53854/liim-2904-18>
- 86 Gould, K. Antibiotics: from prehistory to the present day. *Journal of Antimicrobial Chemotherapy* **71**, 572-575 (2016). <https://doi.org:10.1093/jac/dkv484>
- 87 Chain, E. *et al.* Penicillin as a chemotherapeutic agent. 1940. *Clin Orthop Relat Res*, 3-7 (1993).
- 88 Ligon, B. L. Sir Howard Walter Florey--the force behind the development of penicillin. *Semin Pediatr Infect Dis* **15**, 109-114 (2004). <https://doi.org:10.1053/j.spid.2004.04.001>
- 89 Wainwright, M. Moulds in Folk Medicine. *Folklore* **100**, 162-166 (1989).

- 90 Nelson, M. L., Dinardo, A., Hochberg, J. & Armelagos, G. J. Brief communication: Mass spectroscopic characterization of tetracycline in the skeletal remains of an ancient population from Sudanese Nubia 350-550 CE. *Am. J. Phys. Anthropol* **143**, 151-154 (2010). <https://doi.org:10.1002/ajpa.21340>
- 91 Falkinham, J. O., 3rd *et al.* Proliferation of antibiotic-producing bacteria and concomitant antibiotic production as the basis for the antibiotic activity of Jordan's red soils. *Appl Environ Microbiol* **75**, 2735-2741 (2009). <https://doi.org:10.1128/aem.00104-09>
- 92 Cook, M., Molto, E. & Anderson, C. Fluorochrome labelling in Roman period skeletons from Dakhleh Oasis, Egypt. *Am J Phys Anthropol* **80**, 137-143 (1989). <https://doi.org:10.1002/ajpa.1330800202>
- 93 Bassett, E. J., Keith, M. S., Armelagos, G. J., Martin, D. L. & Villanueva, A. R. Tetracycline-labeled human bone from ancient Sudanese Nubia (A.D. 350). *Science* **209**, 1532-1534 (1980). <https://doi.org:10.1126/science.7001623>
- 94 Hummert, J. R. & Van Gerven, D. P. Tetracycline-labeled human bone from a medieval population in Nubia's Batn el Hajar (550-1450 A.D.). *Hum Biol* **54**, 355-363 (1982).
- 95 Grauer, A., Armelagos, GJ. *Skeletal biology of Hesban: a biocultural interpretation.*, (Berrin Springs, 1998).
- 96 Johnson, M. & Kelley, B. L. *Tots Goes to Gbarnga: The Life of Mattiedna Johnson, R.N., M.T.M.A. Div. : a Biography.* (Mattiedna Johnson Family Company, 1988).
- 97 *10.2 Mechanisms of Antibacterial Drugs.*
- 98 *Mode of Actions and Targets for Antibacterial Drugs.* (Creative Biolabs).
- 99 Shelby Watford, S. W. *Bacterial DNA Mutations.* (StatPearls, 2022).
- 100 Fleming, A. *Chemotherapy: Yesterday, To-Day, and To-Morrow.*, 39 (Cambridge University Press, 1946).
- 101 WHO. *Global Action Plan on Antimicrobial Resistance.* (WHO Document Production Services, 2015).
- 102 CDC. *2022 SPECIAL REPORT: COVID-19 U.S. Impact on Antimicrobial Resistance.* (Department of Health and Human Services, CDC, 2022).
- 103 Towse, A. *et al.* Time for a change in how new antibiotics are reimbursed: Development of an insurance framework for funding new antibiotics based on a policy of risk mitigation. *Health Policy* **121**, 1025-1030 (2017). <https://doi.org:10.1016/j.healthpol.2017.07.011>
- 104 Plackett, B. Why big pharma has abandoned antibiotics. *Nature (London)* **586**, S50-S52 (2020). <https://doi.org:10.1038/d41586-020-02884-3>
- 105 Fernandes, P. & Martens, E. Antibiotics in late clinical development. *Biochemical Pharmacology* **133**, 152-163 (2017). <https://doi.org:https://doi.org/10.1016/j.bcp.2016.09.025>
- 106 Aslam, B. *et al.* Antibiotic resistance: a rundown of a global crisis. *Infect Drug Resist* **11**, 1645-1658 (2018). <https://doi.org:10.2147/IDR.S173867>
- 107 Bianco, A., Papadopoli, R., Mascaro, V., Pileggi, C. & Pavia, M. Antibiotic prescriptions to adults with acute respiratory tract infections by Italian general practitioners. *Infect Drug Resist* **11**, 2199-2205 (2018). <https://doi.org:10.2147/IDR.S170349>
- 108 Dekker, A. R., Verheij, T. J. & van der Velden, A. W. Inappropriate antibiotic prescription for respiratory tract indications: most prominent in adult patients. *Fam Pract* **32**, 401-407 (2015). <https://doi.org:10.1093/fampra/cmz019>

- 109 Sulis, G. *et al.* Antibiotic overuse in the primary health care setting: a secondary data analysis of standardised patient studies from India, China and Kenya. *BMJ Glob Health* **5** (2020). <https://doi.org:10.1136/bmjgh-2020-003393>
- 110 Boeree, M. J. *et al.* A dose-ranging trial to optimize the dose of rifampin in the treatment of tuberculosis. *Am J Respir Crit Care Med* **191**, 1058-1065 (2015). <https://doi.org:10.1164/rccm.201407-1264OC>
- 111 Roord, J. J., Wolf, B. H., Gossens, M. M. & Kimpen, J. L. Prospective open randomized study comparing efficacies and safeties of a 3-day course of azithromycin and a 10-day course of erythromycin in children with community-acquired acute lower respiratory tract infections. *Antimicrob Agents Chemother* **40**, 2765-2768 (1996). <https://doi.org:10.1128/AAC.40.12.2765>
- 112 Van Deun, A., Salim, M. A., Das, A. P., Bastian, I. & Portaels, F. Results of a standardised regimen for multidrug-resistant tuberculosis in Bangladesh. *Int J Tuberc Lung Dis* **8**, 560-567 (2004).
- 113 Sandegren, L. & Andersson, D. I. Bacterial gene amplification: implications for the evolution of antibiotic resistance. *Nat Rev Microbiol* **7**, 578-588 (2009). <https://doi.org:10.1038/nrmicro2174>
- 114 Sharma, D., Misba, L. & Khan, A. U. Antibiotics versus biofilm: an emerging battleground in microbial communities. *Antimicrob Resist Infect Control* **8**, 76 (2019). <https://doi.org:10.1186/s13756-019-0533-3>
- 115 Kamaruzzaman, N. F., Kendall, S. & Good, L. Targeting the hard to reach: challenges and novel strategies in the treatment of intracellular bacterial infections. *Br J Pharmacol* **174**, 2225-2236 (2017). <https://doi.org:10.1111/bph.13664>
- 116 Munita, J. M. & Arias, C. A. Mechanisms of Antibiotic Resistance. *Microbiol Spectr* **4** (2016). <https://doi.org:10.1128/microbiolspec.VMBF-0016-2015>
- 117 Tenaille, O., Toupance, B., Le Nagard, H., Taddei, F. & Godelle, B. Mutators, population size, adaptive landscape and the adaptation of asexual populations of bacteria. *Genetics* **152**, 485-493 (1999). <https://doi.org:10.1093/genetics/152.2.485>
- 118 Denamur, E. & Matic, I. Evolution of mutation rates in bacteria. *Mol Microbiol* **60**, 820-827 (2006). <https://doi.org:10.1111/j.1365-2958.2006.05150.x>
- 119 Bridier-Nahmias, A. *et al.* Escherichia coli Genomic Diversity within Extraintestinal Acute Infections Argues for Adaptive Evolution at Play. *mSphere* **6** (2021). <https://doi.org:10.1128/mSphere.01176-20>
- 120 Swings, T. *et al.* Adaptive tuning of mutation rates allows fast response to lethal stress in Escherichia coli. *Elife* **6** (2017). <https://doi.org:10.7554/eLife.22939>
- 121 Ochiai, K., Yamanaka, T., Kimura, K. & Sawada, O. Studies on inheritance of drug resistance between Shigella strains and Escherichia coli strains. *Nippon Iji Shimpo* **1861**, 34-46 (1959).
- 122 Akiba, T., Koyama, K., Ishiki, Y., Kimura, S. & Fukushima, T. On the mechanism of the development of multiple drug-resistant clones of Shigella. *Japanese journal of microbiology* **4**, 219-227 (1960).
- 123 Lambrecht, E. *et al.* Commensal E. coli rapidly transfer antibiotic resistance genes to human intestinal microbiota in the Mucosal Simulator of the Human Intestinal Microbial Ecosystem (M-SHIME). *Int J Food Microbiol* **311**, 108357 (2019). <https://doi.org:10.1016/j.ijfoodmicro.2019.108357>
- 124 Tawfick, M. M., Elshamy, A. A., Mohamed, K. T. & El Menofy, N. G. Gut Commensal Escherichia coli, a High-Risk Reservoir of Transferable Plasmid-Mediated Antimicrobial Resistance Traits. *Infect Drug Resist* **15**, 1077-1091 (2022). <https://doi.org:10.2147/idr.S354884>

- 125 Liu, G., Thomsen, L. E. & Olsen, J. E. Antimicrobial-induced horizontal transfer of antimicrobial resistance genes in bacteria: a mini-review. *J Antimicrob Chemother* **77**, 556-567 (2022). <https://doi.org:10.1093/jac/dkab450>
- 126 Prudhomme, M., Attaiech, L., Sanchez, G., Martin, B. & Claverys, J. P. Antibiotic stress induces genetic transformability in the human pathogen *Streptococcus pneumoniae*. *Science* **313**, 89-92 (2006). <https://doi.org:10.1126/science.1127912>
- 127 Dorer, M. S., Fero, J. & Salama, N. R. DNA damage triggers genetic exchange in *Helicobacter pylori*. *PLoS Pathog* **6**, e1001026 (2010). <https://doi.org:10.1371/journal.ppat.1001026>
- 128 Vogan, A. A. & Higgs, P. G. The advantages and disadvantages of horizontal gene transfer and the emergence of the first species. *Biol Direct* **6**, 1 (2011). <https://doi.org:10.1186/1745-6150-6-1>
- 129 Grogan, S. & Preuss, C. V. in *StatPearls* (2022).
- 130 Udy, A. A., Roberts, J. A. & Lipman, J. *Antibiotic Pharmacokinetic/Pharmacodynamic Considerations in the Critically Ill*. (Singapore: Springer Singapore Pte. Limited, 2017).
- 131 Reichel, A. & Lienau, P. Pharmacokinetics in Drug Discovery: An Exposure-Centred Approach to Optimising and Predicting Drug Efficacy and Safety. *Handb Exp Pharmacol* **232**, 235-260 (2016). https://doi.org:10.1007/164_2015_26
- 132 Talevi, A. & Bellera, C. L. in *The ADME Encyclopedia: A Comprehensive Guide on Biopharmacy and Pharmacokinetics* (ed Alan Talevi) 718-725 (Springer International Publishing, 2022).
- 133 Zeitlinger, M. A. *et al.* Protein binding: do we ever learn? *Antimicrob Agents Chemother* **55**, 3067-3074 (2011). <https://doi.org:10.1128/aac.01433-10>
- 134 Beer, J., Wagner, C. C. & Zeitlinger, M. Protein binding of antimicrobials: methods for quantification and for investigation of its impact on bacterial killing. *Aaps j* **11**, 1-12 (2009). <https://doi.org:10.1208/s12248-008-9072-1>
- 135 Alagga Abdulrahman, G. V. *Drug Absorption*. (StatPearls Publishing, Treasure Island (FL), 2022).
- 136 Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics.
- 137 Abel Zur Wiesch, P. *et al.* Classic reaction kinetics can explain complex patterns of antibiotic action. *Sci Transl Med* **7**, 287ra273 (2015). <https://doi.org:10.1126/scitranslmed.aaa8760>
- 138 Clarelli, F. *et al.* Drug-target binding quantitatively predicts optimal antibiotic dose levels in quinolones. *PLoS Comput Biol* **16**, e1008106 (2020). <https://doi.org:10.1371/journal.pcbi.1008106>
- 139 Gullberg, E. *et al.* Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathog* **7**, e1002158 (2011). <https://doi.org:10.1371/journal.ppat.1002158>
- 140 Andersson, D. I. & Hughes, D. Microbiological effects of sublethal levels of antibiotics. *Nat Rev Microbiol* **12**, 465-478 (2014). <https://doi.org:10.1038/nrmicro3270>
- 141 P Connors, K. Optimizing Antibiotic Pharmacodynamics for Clinical Practice. *Pharmaceutica analytica acta* **4** (2013). <https://doi.org:10.4172/2153-2435.1000214>
- 142 Monogue, M. L., Kuti, J. L. & Nicolau, D. P. Optimizing Antibiotic Dosing Strategies for the Treatment of Gram-negative Infections in the Era of Resistance. *Expert Rev Clin Pharmacol* **9**, 459-476 (2016). <https://doi.org:10.1586/17512433.2016.1133286>
- 143 Benedetto Tiz, D., Kikelj, D. & Zidar, N. Overcoming problems of poor drug penetration into bacteria: challenges and strategies for medicinal chemists. *Expert*

- Opin Drug Discov* **13**, 497-507 (2018).
<https://doi.org:10.1080/17460441.2018.1455660>
- 144 Paulowski, L. *et al.* The Beauty of Asymmetric Membranes: Reconstitution of the Outer Membrane of Gram-Negative Bacteria. *Front Cell Dev Biol* **8**, 586 (2020).
<https://doi.org:10.3389/fcell.2020.00586>
- 145 Raetz, C. R. Enzymology, genetics, and regulation of membrane phospholipid synthesis in *Escherichia coli*. *Microbiol Rev* **42**, 614-659 (1978).
<https://doi.org:10.1128/mr.42.3.614-659.1978>
- 146 Miller, S. I. & Salama, N. R. The gram-negative bacterial periplasm: Size matters. *PLoS Biol* **16**, e2004935 (2018). <https://doi.org:10.1371/journal.pbio.2004935>
- 147 Silhavy, T. J., Kahne, D. & Walker, S. The bacterial cell envelope. *Cold Spring Harb Perspect Biol* **2**, a000414 (2010). <https://doi.org:10.1101/cshperspect.a000414>
- 148 Engelman, D. M. Membranes are more mosaic than fluid. *Nature* **438**, 578-580 (2005). <https://doi.org:10.1038/nature04394>
- 149 Jeucken, A., Helms, J. B. & Brouwers, J. F. Cardiolipin synthases of *Escherichia coli* have phospholipid class specific phospholipase D activity dependent on endogenous and foreign phospholipids. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1863**, 1345-1353 (2018).
<https://doi.org:https://doi.org/10.1016/j.bbali.2018.06.017>
- 150 Murzyn, K., Róg, T. & Pasenkiewicz-Gierula, M. Phosphatidylethanolamine-phosphatidylglycerol bilayer as a model of the inner bacterial membrane. *Biophys J* **88**, 1091-1103 (2005). <https://doi.org:10.1529/biophysj.104.048835>
- 151 van Meer, G., Voelker, D. R. & Feigenson, G. W. Membrane lipids: where they are and how they behave. *Nature Reviews Molecular Cell Biology* **9**, 112-124 (2008).
<https://doi.org:10.1038/nrm2330>
- 152 Guo, D. & Tropp, B. E. A second *Escherichia coli* protein with CL synthase activity. *Biochim Biophys Acta* **1483**, 263-274 (2000). [https://doi.org:10.1016/s1388-1981\(99\)00193-6](https://doi.org:10.1016/s1388-1981(99)00193-6)
- 153 Tan, B. K. *et al.* Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. *Proc Natl Acad Sci U S A* **109**, 16504-16509 (2012). <https://doi.org:10.1073/pnas.1212797109>
- 154 Shivaji, S. & Prakash, J. S. How do bacteria sense and respond to low temperature? *Arch Microbiol* **192**, 85-95 (2010). <https://doi.org:10.1007/s00203-009-0539-y>
- 155 Romantsov, T. *et al.* Cardiolipin promotes polar localization of osmosensory transporter ProP in *Escherichia coli*. *Mol Microbiol* **64**, 1455-1465 (2007).
<https://doi.org:10.1111/j.1365-2958.2007.05727.x>
- 156 Romantsov, T. *et al.* Cardiolipin synthase A colocalizes with cardiolipin and osmosensing transporter ProP at the poles of *Escherichia coli* cells. *Molecular microbiology* **107**, 623-638 (2018). <https://doi.org:10.1111/mmi.13904>
- 157 Hiraoka, S., Matsuzaki, H. & Shibuya, I. Active increase in cardiolipin synthesis in the stationary growth phase and its physiological significance in *Escherichia coli*. *FEBS Lett* **336**, 221-224 (1993). [https://doi.org:10.1016/0014-5793\(93\)80807-7](https://doi.org:10.1016/0014-5793(93)80807-7)
- 158 Tunaitis, E. & Cronan, J. E. Characterization of the cardiolipin synthetase activity of *Escherichia coli* envelopes. *Archives of Biochemistry and Biophysics* **155**, 420-427 (1973). [https://doi.org:https://doi.org/10.1016/0003-9861\(73\)90132-X](https://doi.org:https://doi.org/10.1016/0003-9861(73)90132-X)
- 159 Romantsov, T., Battle Andrew, R., Hendel Jenifer, L., Martinac, B. & Wood Janet, M. Protein Localization in *Escherichia coli* Cells: Comparison of the Cytoplasmic Membrane Proteins ProP, LacY, ProW, AqpZ, MscS, and MscL. *Journal of Bacteriology* **192**, 912-924 (2010). <https://doi.org:10.1128/JB.00967-09>

- 160 Narita, S.-i. & Tokuda, H. Bacterial lipoproteins; biogenesis, sorting and quality control. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **1862**, 1414-1423 (2017). [https://doi.org:https://doi.org/10.1016/j.bbalip.2016.11.009](https://doi.org/10.1016/j.bbalip.2016.11.009)
- 161 Carbonnelle, E., Helaine, S., Prouvensier, L., Nassif, X. & Pelicic, V. Type IV pilus biogenesis in *Neisseria meningitidis*: PilW is involved in a step occurring after pilus assembly, essential for fibre stability and function. *Mol Microbiol* **55**, 54-64 (2005). [https://doi.org:10.1111/j.1365-2958.2004.04364.x](https://doi.org/10.1111/j.1365-2958.2004.04364.x)
- 162 Metzger, L. C. & Blokesch, M. Composition of the DNA-uptake complex of *Vibrio cholerae*. *Mob Genet Elements* **4**, e28142 (2014). [https://doi.org:10.4161/mge.28142](https://doi.org/10.4161/mge.28142)
- 163 Martin, N. R. *et al.* CrvA and CrvB form a curvature-inducing module sufficient to induce cell-shape complexity in Gram-negative bacteria. *Nat Microbiol* **6**, 910-920 (2021). [https://doi.org:10.1038/s41564-021-00924-w](https://doi.org/10.1038/s41564-021-00924-w)
- 164 Bartlett, T. M. *et al.* A Periplasmic Polymer Curves *Vibrio cholerae* and Promotes Pathogenesis. *Cell* **168**, 172-185.e115 (2017). [https://doi.org:10.1016/j.cell.2016.12.019](https://doi.org/10.1016/j.cell.2016.12.019)
- 165 Maddock, J. R. & Shapiro, L. Polar Location of the Chemoreceptor Complex in the *Escherichia coli* Cell. *Science* **259**, 1717-1723 (1993). [https://doi.org:10.1126/science.8456299](https://doi.org/10.1126/science.8456299)
- 166 Laloux, G. & Jacobs-Wagner, C. How do bacteria localize proteins to the cell pole? *J Cell Sci* **127**, 11-19 (2014). [https://doi.org:10.1242/jcs.138628](https://doi.org/10.1242/jcs.138628)
- 167 Subramani, S., Perdreau-Dahl, H. & Morth, J. P. The magnesium transporter A is activated by cardiolipin and is highly sensitive to free magnesium in vitro. *eLife* **5**, e11407 (2016). [https://doi.org:10.7554/eLife.11407](https://doi.org/10.7554/eLife.11407)
- 168 Ishijima, S. *et al.* Magnesium uptake of Arabidopsis transporters, AtMRS2-10 and AtMRS2-11, expressed in *Escherichia coli* mutants: Complementation and growth inhibition by aluminum. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1848**, 1376-1382 (2015). [https://doi.org:https://doi.org/10.1016/j.bbamem.2015.03.005](https://doi.org/10.1016/j.bbamem.2015.03.005)
- 169 Groisman, E. A. *et al.* Bacterial Mg²⁺ homeostasis, transport, and virulence. *Annu Rev Genet* **47**, 625-646 (2013). [https://doi.org:10.1146/annurev-genet-051313-051025](https://doi.org/10.1146/annurev-genet-051313-051025)
- 170 Bourret, T. J., Liu, L., Shaw, J. A., Husain, M. & Vázquez-Torres, A. Magnesium homeostasis protects *Salmonella* against nitrooxidative stress. *Sci Rep* **7**, 15083 (2017). [https://doi.org:10.1038/s41598-017-15445-y](https://doi.org/10.1038/s41598-017-15445-y)
- 171 Zardecki, C. A.-O. *et al.* PDB-101: Educational resources supporting molecular explorations through biology and medicine.
- 172 Yildiz, F. H. & Schoolnik, G. K. Role of rpoS in stress survival and virulence of *Vibrio cholerae*. *J Bacteriol* **180**, 773-784 (1998). [https://doi.org:10.1128/jb.180.4.773-784.1998](https://doi.org/10.1128/jb.180.4.773-784.1998)
- 173 Matthey, N., Drebes Dörr Natália, C. & Blokesch, M. Long-Read-Based Genome Sequences of Pandemic and Environmental *Vibrio cholerae* Strains. *Microbiology Resource Announcements* **7**, e01574-01518 (2018). [https://doi.org:10.1128/MRA.01574-18](https://doi.org/10.1128/MRA.01574-18)
- 174 Stutzmann, S. & Blokesch, M. Circulation of a Quorum-Sensing-Impaired Variant of *Vibrio cholerae* Strain C6706 Masks Important Phenotypes. *mSphere* **1** (2016). [https://doi.org:10.1128/mSphere.00098-16](https://doi.org/10.1128/mSphere.00098-16)
- 175 Jaskólska, M., Adams, D. W. & Blokesch, M. Two defence systems eliminate plasmids from seventh pandemic *Vibrio cholerae*. *Nature* **604**, 323-329 (2022). [https://doi.org:10.1038/s41586-022-04546-y](https://doi.org/10.1038/s41586-022-04546-y)
- 176 Espinosa, E. *et al.* l-Arabinose Induces the Formation of Viable Nonproliferating Spheroplasts in *Vibrio cholerae*. *Applied and Environmental Microbiology* **87**, e02305-02320 (2021). [https://doi.org:10.1128/AEM.02305-20](https://doi.org/10.1128/AEM.02305-20)

- 177 Kaper, J. B., Nataro, J. P. & Mobley, H. L. T. Pathogenic Escherichia coli. *Nature Reviews Microbiology* **2**, 123-140 (2004). <https://doi.org:10.1038/nrmicro818>
- 178 Leimbach, A., Hacker, J. & Dobrindt, U. E. coli as an all-rounder: the thin line between commensalism and pathogenicity. *Curr Top Microbiol Immunol* **358**, 3-32 (2013). https://doi.org:10.1007/82_2012_303
- 179 Hayes, F. TRANSPOSON-BASED STRATEGIES FOR MICROBIAL FUNCTIONAL GENOMICS AND PROTEOMICS. *Annu Rev Genet* **37**, 3-29 (2003). <https://doi.org:10.1146/annurev.genet.37.110801.142807>
- 180 Pray, L. Transposons: The jumping genes. *Nature Education* **1**, 204 (2008).
- 181 van Opijnen, T. & Camilli, A. Transposon insertion sequencing: a new tool for systems-level analysis of microorganisms. *Nat Rev Microbiol* **11**, 435-442 (2013). <https://doi.org:10.1038/nrmicro3033>
- 182 Cain, A. K. *et al.* A decade of advances in transposon-insertion sequencing. *Nat Rev Genet* **21**, 526-540 (2020). <https://doi.org:10.1038/s41576-020-0244-x>
- 183 Pritchard, J. R. *et al.* ARTIST: high-resolution genome-wide assessment of fitness using transposon-insertion sequencing. *PLoS Genet* **10**, e1004782 (2014). <https://doi.org:10.1371/journal.pgen.1004782>
- 184 Chao, M. C. *et al.* High-resolution definition of the Vibrio cholerae essential gene set with hidden Markov model-based analyses of transposon-insertion sequencing data. *Nucleic Acids Res* **41**, 9033-9048 (2013). <https://doi.org:10.1093/nar/gkt654>
- 185 Chao, M. C., Abel, S., Davis, B. M. & Waldor, M. K. The design and analysis of transposon insertion sequencing experiments. *Nat Rev Microbiol* **14**, 119-128 (2016). <https://doi.org:10.1038/nrmicro.2015.7>
- 186 Marvig, R. L. & Blokesch, M. Natural transformation of Vibrio cholerae as a tool--optimizing the procedure. *BMC Microbiol* **10**, 155 (2010). <https://doi.org:10.1186/1471-2180-10-155>
- 187 Dalia, A. B. & Dalia, T. N. Spatiotemporal Analysis of DNA Integration during Natural Transformation Reveals a Mode of Nongenetic Inheritance in Bacteria. *Cell* **179**, 1499-1511.e1410 (2019). <https://doi.org:10.1016/j.cell.2019.11.021>
- 188 Tham, K. C. *et al.* Mismatch repair inhibits homeologous recombination via coordinated directional unwinding of trapped DNA structures. *Mol Cell* **51**, 326-337 (2013). <https://doi.org:10.1016/j.molcel.2013.07.008>
- 189 Spies, M. & Fishel, R. Mismatch repair during homologous and homeologous recombination. *Cold Spring Harb Perspect Biol* **7**, a022657 (2015). <https://doi.org:10.1101/cshperspect.a022657>
- 190 Lorenz, M. G. & Wackernagel, W. Bacterial gene transfer by natural genetic transformation in the environment. *Microbiol Rev* **58**, 563-602 (1994). <https://doi.org:10.1128/mr.58.3.563-602.1994>
- 191 Diekmann, R. *et al.* Chip-based wide field-of-view nanoscopy. *Nature Photonics* **11**, 322-328 (2017). <https://doi.org:10.1038/nphoton.2017.55>
- 192 Tinguely, J.-C., Helle, Ø. I. & Ahluwalia, B. S. Silicon nitride waveguide platform for fluorescence microscopy of living cells. *Opt. Express* **25**, 27678-27690 (2017). <https://doi.org:10.1364/OE.25.027678>
- 193 Hebisch, E., Knebel, J., Landsberg, J., Frey, E. & Leisner, M. High variation of fluorescence protein maturation times in closely related Escherichia coli strains. *PLoS One* **8**, e75991 (2013). <https://doi.org:10.1371/journal.pone.0075991>
- 194 Scheu, K., Gill, R., Saberi, S., Meyer, P. & Emberly, E. Localization of aggregating proteins in bacteria depends on the rate of addition. *Front Microbiol* **5**, 418 (2014). <https://doi.org:10.3389/fmicb.2014.00418>

- 195 Balleza, E., Kim, J. M. & Cluzel, P. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature Methods* **15**, 47-51 (2018). <https://doi.org/10.1038/nmeth.4509>
- 196 Dumon-Seignovert, L., Cariot, G. & Vuillard, L. The toxicity of recombinant proteins in *Escherichia coli*: a comparison of overexpression in BL21(DE3), C41(DE3), and C43(DE3). *Protein Expr Purif* **37**, 203-206 (2004). <https://doi.org/10.1016/j.pep.2004.04.025>
- 197 Oliver Piercen, M. *et al.* Localization of Anionic Phospholipids in *Escherichia coli* Cells. *Journal of Bacteriology* **196**, 3386-3398 (2014). <https://doi.org/10.1128/JB.01877-14>
- 198 Nishijima, S. *et al.* Disruption of the *Escherichia coli* *cls* gene responsible for cardiolipin synthesis. *J Bacteriol* **170**, 775-780 (1988). <https://doi.org/10.1128/jb.170.2.775-780.1988>
- 199 Beltrán-Heredia, E. *et al.* Membrane curvature induces cardiolipin sorting. *Communications Biology* **2**, 225 (2019). <https://doi.org/10.1038/s42003-019-0471-x>
- 200 Hemm, M. R., Paul, B. J., Schneider, T. D., Storz, G. & Rudd, K. E. Small membrane proteins found by comparative genomics and ribosome binding site models. *Mol Microbiol* **70**, 1487-1501 (2008). <https://doi.org/10.1111/j.1365-2958.2008.06495.x>
- 201 Fontaine, F., Fuchs, R. T. & Storz, G. Membrane localization of small proteins in *Escherichia coli*. *J Biol Chem* **286**, 32464-32474 (2011). <https://doi.org/10.1074/jbc.M111.245696>
- 202 Wang, H. *et al.* Increasing intracellular magnesium levels with the 31-amino acid MgtS protein. *Proc Natl Acad Sci U S A* **114**, 5689-5694 (2017). <https://doi.org/10.1073/pnas.1703415114>
- 203 Jackson, R. J. *et al.* Expression of the PitA phosphate/metal transporter of *Escherichia coli* is responsive to zinc and inorganic phosphate levels. *FEMS Microbiology Letters* **289**, 219-224 (2008). <https://doi.org/10.1111/j.1574-6968.2008.01386.x>
- 204 Yin, X. *et al.* The small protein MgtS and small RNA MgrR modulate the PitA phosphate symporter to boost intracellular magnesium levels. *Mol Microbiol* **111**, 131-144 (2019). <https://doi.org/10.1111/mmi.14143>
- 205 Thiem, S. & Sourjik, V. Stochastic assembly of chemoreceptor clusters in *Escherichia coli*.
- 206 Thiem, S., Kentner D Fau - Sourjik, V. & Sourjik, V. Positioning of chemosensory clusters in *E. coli* and its relation to cell division.
- 207 Rudner, D. Z., Pan Q Fau - Losick, R. M. & Losick, R. M. Evidence that subcellular localization of a bacterial membrane protein is achieved by diffusion and capture.
- 208 Meinhardt, H. & de Boer, P. A. Pattern formation in *Escherichia coli*: a model for the pole-to-pole oscillations of Min proteins and the localization of the division site.
- 209 Wettmann, L. & Kruse, K. A.-O. The Min-protein oscillations in *Escherichia coli*: an example of self-organized cellular protein waves. LID - 10.1098/rstb.2017.0111 [doi] LID - 20170111.
- 210 Standard, A. & Edition, N. CLSI document M07-A9. *Wayne, PA: Clinical and Laboratory Standards Institute* (2012).
- 211 Andrews, J. M. Determination of minimum inhibitory concentrations. *J Antimicrob Chemother* **48 Suppl 1**, 5-16 (2001). https://doi.org/10.1093/jac/48.suppl_1.5
- 212 Mueller, M., de la Peña, A. & Derendorf, H. Issues in pharmacokinetics and pharmacodynamics of anti-infective agents: kill curves versus MIC. *Antimicrob Agents Chemother* **48**, 369-377 (2004). <https://doi.org/10.1128/aac.48.2.369-377.2004>

- 213 Sun, D., Gao, W., Hu, H. & Zhou, S. Why 90% of clinical drug development fails and how to improve it? *Acta Pharmaceutica Sinica B* **12**, 3049-3062 (2022). [https://doi.org:https://doi.org/10.1016/j.apsb.2022.02.002](https://doi.org/10.1016/j.apsb.2022.02.002)
- 214 Drlica, K., Malik, M., Kerns, R. J. & Zhao, X. Quinolone-mediated bacterial death. *Antimicrob Agents Chemother* **52**, 385-392 (2008). [https://doi.org:10.1128/aac.01617-06](https://doi.org/10.1128/aac.01617-06)
- 215 Rayner, C. R. *et al.* Model-Informed Drug Development for Anti-Infectives: State of the Art and Future. *Clin Pharmacol Ther* **109**, 867-891 (2021). [https://doi.org:10.1002/cpt.2198](https://doi.org/10.1002/cpt.2198)
- 216 Wicha, S. G. *et al.* From Therapeutic Drug Monitoring to Model-Informed Precision Dosing for Antibiotics. *Clin Pharmacol Ther* **109**, 928-941 (2021). [https://doi.org:10.1002/cpt.2202](https://doi.org/10.1002/cpt.2202)
- 217 van Belkum, A. *et al.* Innovative and rapid antimicrobial susceptibility testing systems. *Nature Reviews Microbiology* **18**, 299-311 (2020). [https://doi.org:10.1038/s41579-020-0327-x](https://doi.org/10.1038/s41579-020-0327-x)
- 218 Landersdorfer, C. B. & Nation, R. L. Limitations of Antibiotic MIC-Based PK-PD Metrics: Looking Back to Move Forward. *Front Pharmacol* **12**, 770518 (2021). [https://doi.org:10.3389/fphar.2021.770518](https://doi.org/10.3389/fphar.2021.770518)
- 219 Hu, D. *et al.* Origins of the current seventh cholera pandemic.
- 220 Joelsson, A., Liu, Z. & Zhu, J. Genetic and phenotypic diversity of quorum-sensing systems in clinical and environmental isolates of *Vibrio cholerae*. *Infect Immun* **74**, 1141-1147 (2006). [https://doi.org:10.1128/iai.74.2.1141-1147.2006](https://doi.org/10.1128/iai.74.2.1141-1147.2006)
- 221 Lloyd, C. J., Mejia-Santana, A., Dalia, T. N., Dalia, A. B. & Klose, K. E. Natural Transformation in a Classical-Biotype *Vibrio cholerae* Strain. *Appl Environ Microbiol* **87** (2021). [https://doi.org:10.1128/aem.00060-21](https://doi.org/10.1128/aem.00060-21)
- 222 Ke, H. M. *et al.* Tracing Genomic Divergence of *Vibrio* Bacteria in the Harvey Clade. *J Bacteriol* **200** (2018). [https://doi.org:10.1128/jb.00001-18](https://doi.org/10.1128/jb.00001-18)
- 223 Fu, Y., Waldor, M. K. & Mekalanos, J. J. Tn-Seq analysis of *Vibrio cholerae* intestinal colonization reveals a role for T6SS-mediated antibacterial activity in the host. *Cell Host Microbe* **14**, 652-663 (2013). [https://doi.org:10.1016/j.chom.2013.11.001](https://doi.org/10.1016/j.chom.2013.11.001)
- 224 Warr, A. R. *et al.* Transposon-insertion sequencing screens unveil requirements for EHEC growth and intestinal colonization. *PLoS Pathog* **15**, e1007652 (2019). [https://doi.org:10.1371/journal.ppat.1007652](https://doi.org/10.1371/journal.ppat.1007652)
- 225 Mey, A. R., Craig, S. A. & Payne, S. M. Effects of amino acid supplementation on porin expression and ToxR levels in *Vibrio cholerae*. *Infect Immun* **80**, 518-528 (2012). [https://doi.org:10.1128/iai.05851-11](https://doi.org/10.1128/iai.05851-11)
- 226 Mey, A. R., Butz, H. A. & Payne, S. M. *Vibrio cholerae* CsrA Regulates ToxR Levels in Response to Amino Acids and Is Essential for Virulence. *mBio* **6**, e01064 (2015). [https://doi.org:10.1128/mBio.01064-15](https://doi.org/10.1128/mBio.01064-15)
- 227 Miller, V. L., Taylor, R. K. & Mekalanos, J. J. Cholera toxin transcriptional activator toxR is a transmembrane DNA binding protein. *Cell* **48**, 271-279 (1987). [https://doi.org:10.1016/0092-8674\(87\)90430-2](https://doi.org/10.1016/0092-8674(87)90430-2)
- 228 Krukonis, E. S., Yu, R. R. & Dirita, V. J. The *Vibrio cholerae* ToxR/TcpP/ToxT virulence cascade: distinct roles for two membrane-localized transcriptional activators on a single promoter. *Mol Microbiol* **38**, 67-84 (2000). [https://doi.org:10.1046/j.1365-2958.2000.02111.x](https://doi.org/10.1046/j.1365-2958.2000.02111.x)
- 229 Lichev, A., Angelov, A., Cucurull, I. A.-O. & Liebl, W. A.-O. Amino acids as nutritional factors and (p)ppGpp as an alarmone of the stringent response regulate natural transformation in *Micrococcus luteus*.
- 230 Boutte, C. C. & Crosson, S. Bacterial lifestyle shapes stringent response activation.

- 231 Magnusson, L. U., Farewell A Fau - Nyström, T. & Nyström, T. ppGpp: a global regulator in *Escherichia coli*.
- 232 Wilson Ga Fau - Bott, K. F. & Bott, K. F. Nutritional factors influencing the development of competence in the *Bacillus subtilis* transformation system.
- 233 Hamoen, L. W., Venema, G. & Kuipers, O. P. Controlling competence in *Bacillus subtilis*: shared use of regulators. *Microbiology (Reading)* **149**, 9-17 (2003).
<https://doi.org/10.1099/mic.0.26003-0>
- 234 Maamar, H., Raj, A. & Dubnau, D. Noise in Gene Expression Determines Cell Fate in *Bacillus subtilis*. *Science* **317**, 526-529 (2007).
<https://doi.org/10.1126/science.1140818>
- 235 Kickstein, E., Harms, K. & Wackernagel, W. Deletions of *recBCD* or *recD* influence genetic transformation differently and are lethal together with a *recJ* deletion in *Acinetobacter baylyi*. *Microbiology (Reading)* **153**, 2259-2270 (2007).
<https://doi.org/10.1099/mic.0.2007/005256-0>
- 236 Wylie, C. S., Trout, A. D., Kessler, D. A. & Levine, H. Optimal strategy for competence differentiation in bacteria. *PLoS Genet* **6**, e1001108 (2010).
<https://doi.org/10.1371/journal.pgen.1001108>
- 237 Lo Scudato, M. & Blokesch, M. The regulatory network of natural competence and transformation of *Vibrio cholerae*. *PLoS Genet* **8**, e1002778 (2012).
<https://doi.org/10.1371/journal.pgen.1002778>
- 238 Passalacqua, K. D., Charbonneau, M. E. & O'Riordan, M. X. D. Bacterial Metabolism Shapes the Host-Pathogen Interface. *Microbiol Spectr* **4** (2016).
<https://doi.org/10.1128/microbiolspec.VMBF-0027-2015>
- 239 Moradigaravand, D. & Engelstädter, J. The impact of natural transformation on adaptation in spatially structured bacterial populations. *BMC Evol Biol* **14**, 141 (2014).
<https://doi.org/10.1186/1471-2148-14-141>
- 240 Sun, Y., Bernardy, E. E., Hammer, B. K. & Miyashiro, T. Competence and natural transformation in vibrios. *Mol Microbiol* **89**, 583-595 (2013).
<https://doi.org/10.1111/mmi.12307>
- 241 Adams, D. W., Stutzmann, S., Stoudmann, C. & Blokesch, M. DNA-uptake pili of *Vibrio cholerae* are required for chitin colonization and capable of kin recognition via sequence-specific self-interaction. *Nat Microbiol* **4**, 1545-1557 (2019).
<https://doi.org/10.1038/s41564-019-0479-5>

Paper I

1 Transposon insertion sequencing reveals dependency 2 of natural transformation on amino acid metabolism

3 Merete Storflor¹, João Alves Gama¹, Klaus Harms¹, Sören Abel¹

4 ¹ Department of Pharmacy, UiT - The Arctic University of Norway, 9037 Tromsø, Norway

5 *Correspondence should be addressed to soren.abel@fhi.no

6 Abstract

7 Natural transformation is a major mode of horizontal gene transfer that shapes the evolution of bacteria.
8 Here we identify genes important for natural transformation with a transposon insertion sequencing
9 approach. By using a highly saturated transposon library in *Vibrio cholerae* we find that an impaired
10 amino acid metabolism inhibits transformation. Transformation of mutants in amino acid metabolizing
11 genes could be recovered back to wild-type level upon supplementation with amino acids. Such
12 supplementation had a concentration- and time- dependent effect on competence induction. Reporter
13 assays for competence regulators demonstrate considerable heterogeneity in amino acid metabolism
14 mutants. This study indicates that amino acid metabolism regulates induction of competence for natural
15 transformation.

16 Background

17 Horizontal gene transfer (HGT) aids in the transfer of virulence factors and antibiotic resistance markers
18 in prokaryotes. This raises global concerns, as the spread of antibiotic resistance renders current drugs
19 ineffective against pathogenic bacteria. A major mechanism of horizontal gene transfer is natural
20 transformation. Which, not only promotes the spread of antibiotic resistance¹, but also drives the
21 emergence of new pathogenic strains, as demonstrated by the conversion of non-toxigenic to toxigenic
22 *Vibrio cholerae*². The current cholera pandemic is caused by *V. cholerae* O1 El Tor. This is an intestinal
23 pathogen that, upon ingestion, can cause debilitating, even life-threatening, diarrhea. Like many other
24 bacteria, *V. cholerae* can take up exogenous DNA from neighboring bacteria using a mechanism known
25 as natural transformation. Previous pandemics have been caused by the classical biotype, a strain with
26 reduced natural transformation. While the machinery is present, a mutation in *hapR* inhibits its
27 expression³. The current cholera pandemic is known as the 7th and began in 1961, whoever the strain
28 persists today making this pandemic the longest of its kind⁴. It has been theorized that this is due to the
29 increased adaptability of the strain promoted by natural transformation as well as mobile genetic
30 elements. Current studies estimate that the pathogen causes one to four million infections per year
31 leading to 21.000 to 143.000 deaths⁵⁻⁷. While, rehydration therapy is an efficient treatment, outbreaks
32 typically occur in regions with limited access to care and clean water. Therefore in 2014, the Global
33 Task Force on Cholera Control (GTFCC) launched a strategy for ending cholera outbreaks by 2030,
34 however there are at present no clear indications of decreasing outbreaks. Proper sanitation and access
35 to clean water and food is critical in controlling the progression of this disease, as well as understanding
36 the regulatory systems that control the pathogens environmental adaptation.

37 Naturally competent bacteria can take up free DNA present in the environment, allowing access to a
38 vast genetic pool and enabling transfer across species and strains. Specifically for *V. cholerae*, the
39 competence state is induced upon association with chitin, and enables natural transformation. Which in

40 combination with competence development is a highly orchestrated response to the environment. The
41 activation of natural transformation requires an intricate network of receptors and response regulators
42 that rapidly sense and react to the immediate environment. The presence of chitin is critical for
43 competence induction and natural transformation, as this increases the expression of the master regulator
44 of natural transformation, *tfoX*^{8,9}. The activity of TfoX is enhanced by TfoS, which is a membrane-bound
45 transcriptional regulator linking chitin to DNA uptake¹⁰. The chitin-dependent production of TfoX
46 works in unison with carbon catabolite repression (CCR) and is co-regulated by cell density in the form
47 of quorum sensing (QS)^{11,12}. CCR enables bacteria to regulate their metabolism to better adapt to their
48 environment, and acts to inhibit natural transformation in the presence of glucose⁸. CyaA will in the
49 absence of a preferred carbon source indirectly enable TfoX to induce transcription of several natural
50 transformation genes¹³. This interplay is under the control of a nucleotide scavenging repressor, CytR,
51 which functions as a global regulator¹⁴. However, the exact interactions are not fully understood.
52 Additionally, natural transformation has been found to be regulated by amino acids in *Micrococcus*
53 *luteus*¹⁵. Lichev et. al, found amino acids to function as environmental cues, where amino acid starvation
54 resulted in competence induction¹⁵. Further, stimulatory, and inhibitory effects of amino acids on
55 competence development have also been detected in *B. subtilis*, by Wilson et. al¹⁶. Glutamic acid and
56 alanine were found to be detrimental to competence development, while other amino acids demonstrated
57 stimulatory effects. Meanwhile, the quorum sensing regulator, HapR, is important to down regulate the
58 extracellular DNA endonuclease, Dns and Xds¹⁷. Exogenous DNA is recognized by a type IV pilus. The
59 DNA is pulled into the periplasm via PilQ, an outer membrane pore complex. ComEA binds the DNA
60 and shuttles one strand into the cytoplasm via an inner membrane channel ComEC¹⁸. Once in the
61 cytoplasm single-stranded DNA binding proteins (Ssb) and DprA binds the ssDNA and facilitates RecA-
62 dependent recombination with the chromosomal DNA. RecA is a universally conserved DNA repair and
63 maintenance protein that catalyzes strand exchange and is responsible for inserting the transforming
64 DNA to the chromosomal DNA resulting in a heteroduplex. Resolving the heteroduplex results in the
65 ssDNA being incorporated into one of the strands of the chromosomal DNA and can then be inherited
66 by one of the daughter cells after cell division.

67 The natural transformation machinery is costly and requires precise orchestration of several pathways.
68 How this machinery originated and is maintained has long been debated. It has been suggested that DNA
69 can be utilized as a nutrient, as well as for DNA repair. Another theory suggests that natural
70 transformation is maintained to sample the genetic diversity of neighboring populations. In this theory,
71 naturally competent populations are always genetically heterogenic, and a small fraction samples the
72 genetic fitness landscape. This could be why only a fraction of a population is transformable and even
73 fewer successfully take up DNA. With this bet-hedging strategy, the potentially detrimental effects of
74 the newly acquired genes would be restricted to a minority of the population, while the pathogen is still
75 sampling potentially beneficial genes. Heterogeneity is often greater for metabolic genes¹⁹. Further
76 several metabolic systems have been acquired via horizontal gene transfer²⁰. It has also been observed
77 that non-pathogenic *V. cholerae* can become pathogenic by taking up a cluster of virulence factors,
78 referred to as the pathogenicity island^{21,22}.

79 To evaluate the fitness contributions of genomic features under specific growth conditions, a high
80 throughput method known as transposon insertion sequencing (TIS) can be implemented. The method
81 is well suited to assess fitness landscapes as it directly links phenotype to genotype²³. This is
82 accomplished by creating a population of bacteria where a single transposon disrupts genomic features,
83 e.g. loci. Simplified, if a disrupted locus is maintained in the population, it is classified as “non-

84 essential”, while an uninterrupted locus may indicate an “essential” function. To determine the position
85 of the disruption, the method relies on next-generation sequencing. In the population a locus may be
86 interrupted several times, and at several positions. The mutants that pass the condition carry either
87 disrupted elements that produce a neutral effect or disrupted elements that would have been
88 disadvantageous, and if uninterrupted would have inhibited the bacteria from passing the condition. The
89 regions of the genome where transposon insertions are missing or reduced can be identified by
90 sequencing a part of the transposon as well as a short stretch of the flanking genomic DNA that bridges
91 the insertion site. In so doing, one can detect loci critical for survival under the condition the library was
92 subjected to. These mutants are denoted as “essential”, because of their high fitness costs. By comparing
93 two different growth conditions with distinct selection pressures, it is possible to identify genetic
94 elements critical for optimal fitness under these specific conditions. The method compares the disruption
95 frequencies of loci of a specific selection pressure to a control condition, as a proxy for fitness. This
96 allows for the identification of “conditionally essential” genetic elements, i.e., disrupted genetic
97 elements with strong fitness defects under the test conditions.

98 In this study, we generated a high-density transposon library in *V. cholerae* and exposed it to a strong
99 selection pressure, selecting for natural uptake and recombination of free DNA into its genome. This
100 setup allowed us to identify a large set of genes involved in natural transformation and recombination.
101 In addition to known regulators like TfoX, we identified MetE and AsnB with strong phenotypes and
102 confirmed their effects by reproducing the results of the transposon screen with independent, genetically
103 engineered disruption mutants. Strikingly, we detected a large set of genes involved in amino acid
104 transport and metabolism that affected natural transformation and further investigated the impact on
105 DNA uptake and recombination.

106 Results

107 Generating a TIS library to investigate factors involved in natural transformation

108 To investigate the genes that are important for natural transformation, a high-density transposon library
109 was created in *V. cholerae*. Random insertion of the transposon into the genome occurred upon delivery
110 of a non-replicative plasmid containing the transposable element through conjugation. The transposable
111 element is a modified Tn *Mariner* containing a kanamycin (Km) resistance gene flanked by inverted
112 repeats²⁴. This transposon can insert at specific DNA sites where the two nucleosides thymine (T) and
113 adenine (A) follow each other (TA sites). By selecting for the transposable element after conjugation, a
114 library of *V. cholerae* was generated where, every bacterium has one transposon randomly inserted
115 within its genetic material. Here it, potentially disrupts the function of a genetic locus, e.g., a gene. After
116 conjugation, the input library yielded ~5.7M CFUs, indicating a number of independently generated
117 insertion mutants in the same order of magnitude. After inserting the transposon, we evaluated the
118 quality of the library. We aimed for high complexity, which means that the majority of the available
119 insertion sites have been disrupted.

120 Genomic DNA was isolated from an aliquot of the library and PCR was performed, so that the resulting
121 PCR product spanned the transposon insertion site. The PCR products were then sequenced with an
122 Illumina MiSeq sequencer using the MiSeq V3, 150-cycle kits. The output sequences were filtered to
123 avoid false positives. The sequencer yielded 21.7 M reads, out of which 19.4 M passed filtering criteria
124 and 17.6M reads mapped once to the genome. Approximately 84% of the potential insertion sites,

125 including intergenic regions, were hit (165'020 TA sites). See **Table 1** for a summary of the input
126 library. We concluded that the input library was of high complexity and near-saturation.

127

Table 1: Sequencing statistics of the input library

128

Reads	17.6 M
-------	--------

129

Sites Hit	165'020
-----------	---------

130

TAs hit	84%
---------	-----

131

Average Read Count	90
--------------------	----

132 TIS analysis reveals a role for amino acid metabolism genes in natural transformation

133 To investigate genes involved in natural transformation, we used our transposon library in a
134 transformation assay that selects against all bacteria that did not take up and chromosomally recombine
135 the donor DNA. If a gene with a function in natural transformation, or recombination, is disrupted by
136 the transposon, it will be less frequently represented, or not at all, within the pool of transformants
137 isolated after the assay and its absence can be detected by sequencing.

138 For the transformation assays, we first induced competence. Our library was grown to early exponential
139 phase in LB media, washed in M9⁺ media, and then transferred to fresh M9⁺ media supplemented with
140 chitin flakes, two sample populations were created (**Figure 1**). After 24 h, one population received donor
141 DNA denoted as the "TnNT" population. The other, "Ctrl" population, did not receive donor DNA. The
142 donor DNA is a PCR product of the *rpsL* gene with a point mutation, A128G. Successful transformation
143 and recombination confers resistance to streptomycin (Sm) allowing for selection of transformants.
144 After another 24 h, the bacteria were detached from chitin by vortexing. A fraction of both populations
145 was used for determining colony-forming units (CFUs), by plating a serial dilution on both selective
146 (Sm) and non-selective LB plates. The ratio of the TnNT population sizes in absence and presence of
147 Sm was used to determine the transformation frequency. The number of colonies from the Ctrl sample
148 on selective plates was used to determine the frequency of spontaneous resistance to Sm. The remaining
149 liquid cultures from both conditions were grown out on 15 mm Petri dishes. The Ctrl sample was plated
150 on non-selective plates, while the TnNT was plated on selective plates containing Sm. All colonies were
151 harvested the next day and prepared for whole genome sequencing. In the Ctrl sample, we expect to find
152 a first set of genes required for survival under the assay conditions, presented as missing insertions. In
153 the TnNT sample, we expect to find in addition a second set of genes required for natural transformation
154 and recombination, and survival on Sm. By comparing the sequencing results from the TnNT sample to
155 the Ctrl sample, we aim to identify the second set of genes. A total of 4.2 M CFUs were harvested for
156 the Ctrl assay, while 1.3 M CFUs were harvested for the TnNT assay. Throughout the assay, spontaneous
157 resistance to Sm was not encountered (**Figure 2a**). The transformation efficiency was 5.5×10^{-2} (**Figure**
158 **2b**). Each sample was sequenced twice, and the results were merged. The sequence statistics are
159 summarized in **Table S. 1**. The Ctrl sample yielded 28.8M reads covering 85% of the genome-wide
160 available TA sites and the TnNT output had 30.5M reads covering 72% of the available TA sites. On
161 average, each site was sequenced about 150 times. As expected with such a high coverage, in the
162 majority of genes, most intragenic TA sites were hit by a transposon at least once (**Figure 2c**). This
163 reflects that most genes are non-essential under the assay conditions. A more detailed analysis of the
164 TnNT TIS screen using the ARTIST pipeline²⁵ identified 263 genes out of a total of 3751 to be

165 conditionally essential and therefore likely to be involved in taking up and incorporating exogenous
 166 DNA. Previously discovered factors involved in competence and transformation were detected as well
 167 and functioned as positive controls to validate the screen. For an overview of all annotated genes found
 168 to be essential for TnNT, the donor DNA receiving sample, refer to **Table S. 2**. The complete list of
 169 annotated genes along with the statistics and the corresponding fitness assigned by the analysis pipeline
 170 can be found in **Table S. 2**.

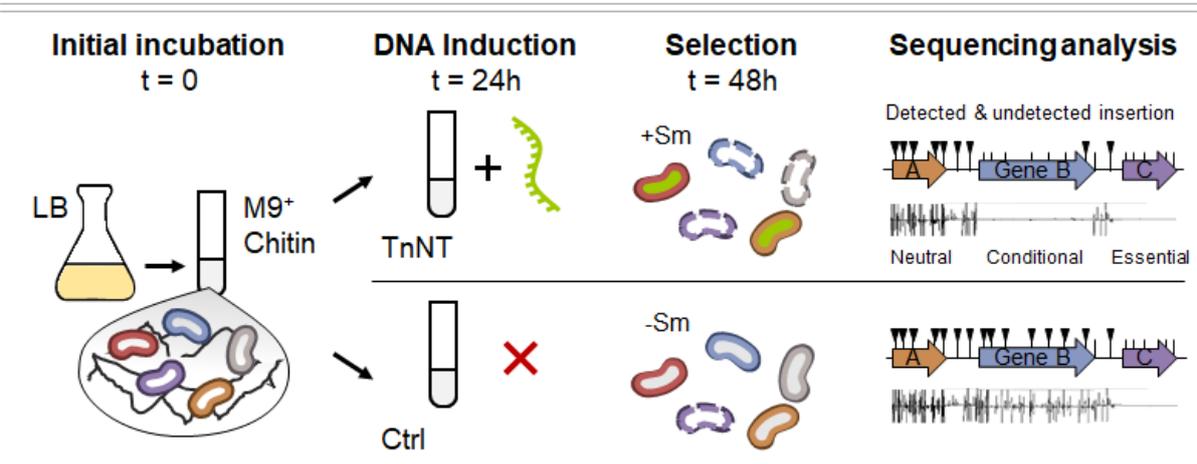


Figure 1 Schematic representation of the transposon insertion experiment to identify genes involved in natural transformation. Bacteria are depicted as curved cylinders with different colors that represent different transposon mutants. The chitin that is used to induce natural competence is depicted with black jagged lines. Donor DNA is depicted as a green line and bacteria that successfully integrated donor DNA have a green fill. The integrated donor DNA provides resistance to streptomycin (Sm). For the control assay, no DNA was added, indicated by a red cross. Transposon mutants that do not survive are indicated by a dotted line. They either die because the transposon mutant is not able to survive the general growth conditions of the assay (violet cell) or because they have not taken up and integrated the donor DNA (blue and grey cells). After sequencing the surviving bacteria, the sequencing results can be visualized graphically. Genes are schematically depicted as arrows with black lines on top representing potential transposon insertion sites (TA-sites). If transposons have inserted at this site, this is indicated by a black triangle. Under each gene, an actual data set depicting transposon sequencing results is shown. Where, each black line depicts a transposon insertion, and the height of the black line represents how often a transposon insertion at this site was sequenced, either in the forward orientation (lines pointing upwards) or the reverse orientation (lines pointing downwards). Depending on the transposon integration pattern, we can classify a genetic locus as “neutral” - when the locus is frequently found disrupted by transposons in the TnNT and the Ctrl assay, “essential” - when it is not disrupted in the TnNT and Ctrl assay, and “conditional” - when it is disrupted in the control condition but not in the TnNT assay.

171

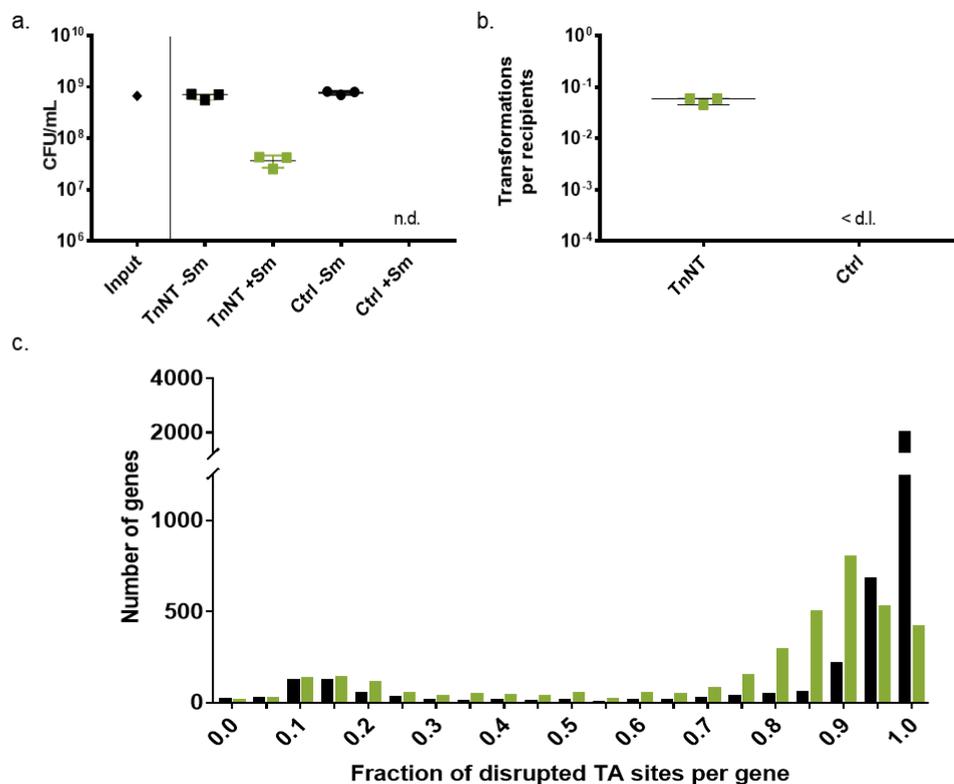


Figure 2 Natural transformation of TIS library and sequence analysis. (a) Bacterial numbers during the transformation assay. CFU/mL were measured for the TnNT sample (squares) that received donor DNA and the Ctrl sample (circles) that did not receive donor DNA with (black) and without (green) selection for streptomycin (Sm). The Ctrl sample remained sensitive to the antibiotics and did not survive, < d.l = below detection limit. (b.) Transformation frequencies of the TnNT sample (green), the transformation frequency was not applicable for the Ctrl sample. (c) Distribution of relative disruptions of TA sites within genes. The graph displays fraction of disrupted insertion sites (TA-sites) per gene plotted by the number of genes for both the Ctrl sample (black bars), and the TnNT sample (green bars).

172 Cluster orthologous groups and enrichment analysis

173 To identify important processes for natural transformation, we categorized the detected genes by
 174 functional groups. The protein sequences of the 263 hits from our screen were mapped and sorted into
 175 cluster orthologous groups (COGs), using the eggNOG sequence mapper²⁶. When cross-referencing a
 176 random subset of sequences with other online databases a small discrepancy was observed for comEA.
 177 KEGG and Uniprot states the molecular function is DNA binding with involvement of the type II
 178 secretion system. This is a periplasmic protein that is critical for shuttling external DNA into the
 179 periplasm¹⁸. However, it was sorted into the “replication and repair” COG. While the protein binds DNA
 180 it is not directly involved in replication or repair. To avoid biasing the screen we did not diverge from
 181 the COG classifications assigned by eggNOG. An enrichment analysis was performed by comparing the
 182 amino acid sequences of all detected genes to the *V. cholerae* A1552 proteome. The relative abundance
 183 of COGs in the conditionally essential group was compared to the relative abundance of the
 184 corresponding COG for the entire proteome (**Figure 3**). Three COGs have an enrichment index above
 185 one, intracellular trafficking, amino acid metabolism and cell motility. Nucleotide metabolism has an
 186 enrichment index close to one. Looking into the genes assigned to this COG we find the majority of the
 187 genes are involved in amino acid synthesis. We therefore focused on genes from the three categories
 188 above 1. The intracellular trafficking COG consisted of 32 enriched sequences, there were 58 enriched
 189 sequences in amino acids metabolism, and 28 enriched sequences in the cell motility COG. Several of
 190 the genes sorted into the intracellular trafficking and the cell motility COGs have a clear link to natural

191 transformation. For example, several pili proteins are involved in the type IV MSHA pilus is required
 192 for colonizing chitin²⁷, and the type IV competence pilus is required for DNA uptake²⁸. As, these groups
 193 overlap we also investigated the nucleotide metabolism COG where we found several amino acids
 194 metabolism genes with dual functions in nucleotide metabolism as well as, *cyoA*, an adenyl cyclase.
 195 A $\Delta cyoA$ mutant is unable to synthesis cAMP and does not transform on chitin^{11,13}. The amino acid
 196 metabolism COG involves biosynthesis and catabolism of amino acids. Amino acid metabolism is the
 197 second greatest enriched COG and is yet to be investigated in the context of natural transformation in
 198 Gram-negatives. To our knowledge only one study in *Micrococcus luteus*, mentions amino acids within
 199 the context of natural transformation¹⁵. We also found that the largest group of our hits from our TIS
 200 screen were involved in amino acid metabolism and transport, indicating a prominent and novel role for
 201 amino acids in natural transformation.

202

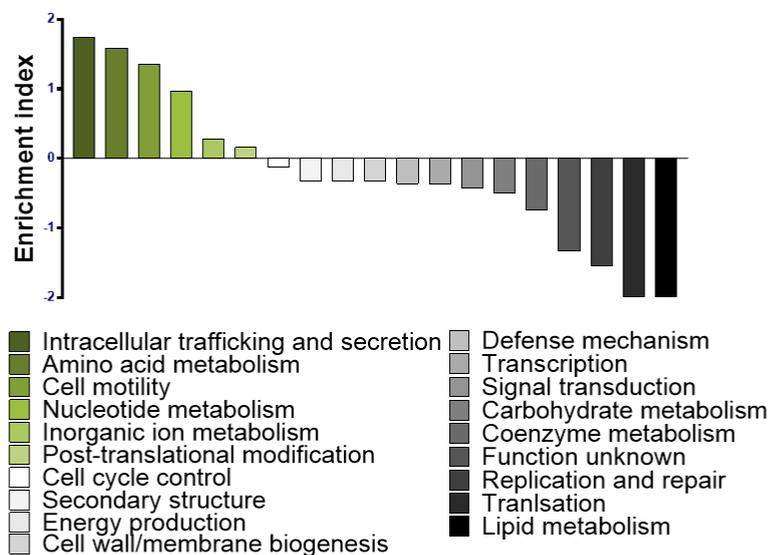


Figure 3 COG analysis of conditionally essential genes important for natural transformation. Y-axis defined as COG enrichment index by $\log_2(\text{COG conditionally essential} / \text{COG whole genome})$, with COG categories on the x-axis. Categories above zero are colored in green-scale, while categories below zero are depicted in gray-scale.

203

204 Verification of TIS screen and transformation frequencies

205 We validated the quality of our TIS screen using a two-pronged approach. Firstly, we compiled a list of
 206 genes from the literature with confirmed roles in natural transformation and recombination. By
 207 comparing this list with our hits, we evaluated how comprehensive our screen was and assessed the
 208 number of false negatives. Secondly, we selected a set of genes from our hit list, disrupted the gene
 209 function, and tested the transformation efficiency of these mutants. This enabled us to evaluate the
 210 selectiveness of our screen and assess the number of false positives.

211 We systematically searched for articles describing genes involved in natural transformation in *V.*
 212 *cholerae*. To be included in our reference list, genes had to fulfill at least one of the following criteria:
 213 i) investigated directly for their role in natural transformation, or ii) available knowledge from
 214 homologues genes (**Table 2**). Based on the literature review we did not find any false negatives and

215 assume a low rate of false positives. Important genes described in the literature to be required for natural
 216 transformation, such as *tfoX*, were classified in our screen as conditionally essential for the donor DNA
 217 receiving sample. In addition, *recA* and *hapR/luxO* were described in the literature as critical for natural
 218 competence these genes were found to be essential in both conditions, and likely required for survival
 219 within our assay conditions. Further, genes known to inhibit natural transformation, such as *dns*, were
 220 confirmed to be non-essential in both conditions as well. Our screen further overlaps with the same set
 221 of genes as studies using different methodologies^{29,30}.

222 To test for false negative, a set of 12 genes from our list of detected genes were selected (**Table 3**).
 223 These genes were chosen to represent the major COGs, including some previously identified genes from
 224 our positive control list such as *tfoX* (competence activator gene). Knockouts were created by deleting
 225 the majority of the open reading frame and replacing it by a tetracycline resistance cassette by allelic
 226 exchange. These deletion strains were used as recipients of donor DNA in our standard natural
 227 transformation assay. The respective transformation frequencies are presented in **Figure 4**. We detect a
 228 range of impaired transformation frequencies. $\Delta mshL$, $\Delta mshQ$, Δpmg , $\Delta carB^{31}$, $\Delta lapD^{32,33}$, and Δppc
 229 had moderately impaired transformation frequencies. $\Delta asnB^{34}$, and $\Delta metE^{35}$ were severely impaired,
 230 close to the limit of detection. While no transformants were observed for $\Delta tfoX^{36}$, $\Delta tsaP^{37,38}$, $\Delta pilW$, and
 231 $\Delta tfoS^{10}$. We could accurately reproduce the published phenotypes from the previously characterized
 232 genes in our validation set (*tfoX*, *mshL*, *mshQ*, *pilW*, and *tfoS*). The only exception is *ttsaP*. In our
 233 analysis, the deletion mutant of this gene had a transformation frequency below the detection limit, while
 234 in Jaskólska, et al. this deletion had a milder phenotype with only a 10-fold reduction³⁷. The reason for
 235 this is unknown. The genes *pmg*, *asnB*, *metE*, *carB*, *lapD*, and *ppc* have to our knowledge not been
 236 characterized in terms of natural transformation. Together, these results indicate that our screen struck
 237 a good balance between strictness (false negatives) and inclusiveness (false positives) and we are
 238 confident that we identified a large number of genes involved in natural transformation, without
 239 including too many false positives. Furthermore, we noticed that the two amino acid metabolism genes
 240 had severely reduced transformation frequencies. This reinforces the results from the COG analysis and
 241 strengthens the hypothesis that amino acid metabolism plays an important role in natural transformation
 242 of *V. cholerae*.

Table 2 TIS screen validation, genes used as positive and negative controls				
Only essential for the donor DNA receiving sample, positive controls				
Gene	Product description	Locus tag c6706	Locus tag A1552	References
<i>tfoX</i>	Master competence regulator	VC1153	A1552VC_00900	36
<i>comEA</i>	Competence protein	VC1917	A1552VC_01694	8,30
<i>comEC</i>	Competence protein	VC1879	A1552VC_01659	8,30
<i>qstR</i>	Quorum sensing and TfoX-dependent regulator	VC0396	A1552VC_00153	39
<i>comM</i>	Competence protein	VC0032	A1552VC_00023	30
<i>comF</i>		VC2718	A1552VC_02473	30,40
<i>pilA</i>	Type IV pilus assembly protein, major subunit	VC2423	A1552VC_02186	8
<i>pilQ</i>	Type IV pilus biogenesis protein	VC2630	A1552VC_02390	8
Essential in both conditions, positive controls				
<i>recA</i>	DNA repair and recombination protein	VC0543	A1552VC_00306	28
<i>hapR/luxO</i>	master global transcription factors, directs downstream gene expression in response to cell density	VC1021	A1552VC_00769	41
Not essential in any condition, negative control				
<i>dns</i>	Extracellular deoxyribonuclease	VC0470	A1552VC_00234	40

243

Locus tag c6706	Locus tag A1552	gene	Product description	COG category
VC0047	A1552VC_00037	<i>tsaP</i>	Uncharacterized protein with LysM domain	Energy production & conversion
VC0402	A1552VC_00159	<i>mshL</i>	MSHA biogenesis protein	Intracellular trafficking & secretion
VC0414	A1552VC_00171	<i>mshQ</i>	MSHA biogenesis protein	Intracellular trafficking & secretion
VC0611	A1552VC_00374	<i>pgm</i>	Phosphoglucomutase	Carbohydrate metabolism & transport
VC0860	A1552VC_00617	<i>pilW</i>	Type IV fimbrial biogenesis protein PilW	Intracellular trafficking & secretion
VC0992	A1552VC_00741	<i>asnB</i>	Asparagine synthetase B [glutamine-hydrolyzing]	Amino acid metabolism & transport
VC1154	A1552VC_00900	<i>tfoX</i>	Master competence regulator	Transcription
VC1704	A1552VC_01494	<i>metE</i>	5-methyltetrahydropteroyltriglutamate--homocysteine methyltransferase	Amino acid metabolism & transport
VC2080	A1552VC_01851	<i>tfoS</i>	HTH-type transcriptional regulator CdhR	Transcription
VC2389	A1552VC_02152	<i>carB</i>	Carbamoyl-phosphate synthase large chain	Nucleotide metabolism and transport
VC2646	A1552VC_02405	<i>ppc</i>	Phosphoenolpyruvate carboxylase	Coenzyme metabolism
VCA1082	A1552VC_A03758	<i>lapD</i>	Membrane bound c-di-GMP receptor, pdeB 2	Signal Transduction

244

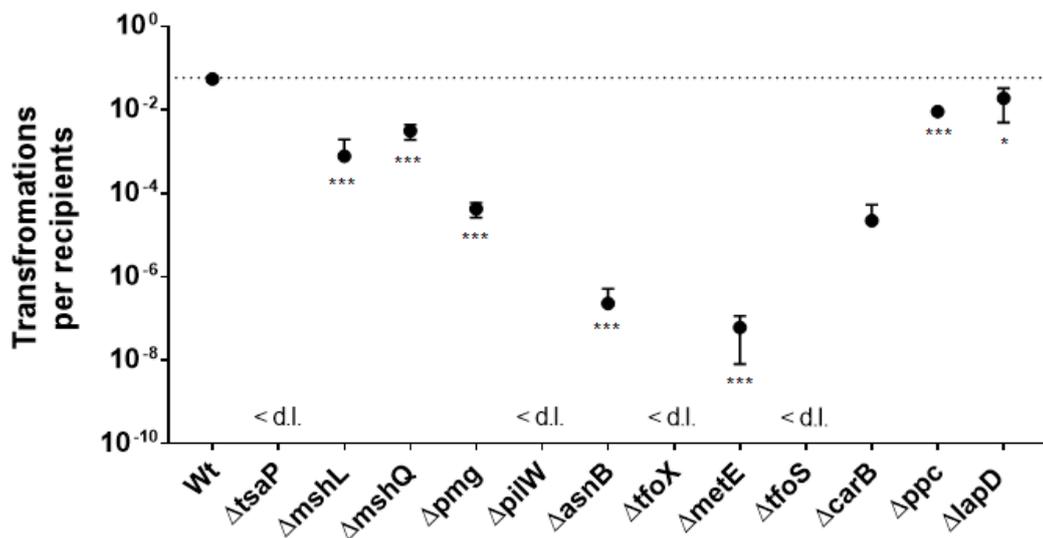


Figure 4 Transformation frequency of the knockout mutants selected from the TnNT screen. Mutants: $\Delta tfoX$, $\Delta tsaP$, $\Delta pilW$, and $\Delta tfoS$ are below detection limit (< d.l.). All samples were tested in triplicates, except from $\Delta pilW$, $\Delta carB$ n=2. The dotted line indicates the median transformation efficiency of the wild-type, the dots denote the median of the experiments, and the error bars denote the standard deviation. * indicates $p < 0.05$, $\Delta carB$ was not included in this analysis due to sample size. (Unpaired two-tailed t-test).

245 External addition of casamino acids inhibit competence initiation

246 We identified a large number of genes involved in amino acid metabolism in our TIS screen.
247 Additionally, the amino acid metabolism COG group was highly enriched. To get an initial impression
248 of how amino acids affect natural transformation, we supplemented *V. cholerae* with casamino acids
249 (CA), a mixture of amino acids obtained through hydrolyzing casein. CA was added in different
250 concentrations at different times during the natural transformation protocol, and the resulting
251 transformation frequencies were recorded. See **Figure 5a** for a schematic overview of the experimental
252 setup. We found a concentration dependent decrease in transformation frequency (**Figure 5b**). CA had
253 the greatest impact on transformation frequency when added during the initial stages of the experiment,
254 when the bacteria were exposed to chitin in the absence of DNA. There was a significant drop in
255 transformation frequencies when CA was added during the initial stages. If CA were added at the start
256 of the experiment all concentrations above 0.2% w/v significantly decreased transformation. After 6 h,
257 all CA concentrations above 0.1% w/v resulted in significantly reduced transformation frequencies. This
258 effect increased with concentration. If CA were added along with or after donor DNA addition, the
259 transformation-inhibiting effect was attenuated. At 24 h, 1 % and 4% CA resulted in a significantly
260 reduced transformation, none of the other concentrations effected the transformation frequency once
261 DNA was added after 24h. By determining the total number of *V. cholerae*, i.e., viable transformed and
262 untransformed cell, at the end of the experiment, we found that CFUs increased slightly with increasing
263 concentrations of CA (**Figure S 1**). Supplementing the natural transformation reaction with casamino
264 acids demonstrated a more prominent inhibitory effect during the initial stages of our natural
265 transformation assay before DNA induction. According to literature the DNA uptake rate for
266 *Acinetobacter calcoaceticus* was calculated, based on experimental evidence, to be a minimum of 60
267 nucleotides per second⁴². Although competence induction for this Gram-negative species differs from
268 *V. cholerae*, we theorized most of the DNA would be taken up efficiently within the first few hours. To
269 test this hypothesis, we introduced DNase into our natural transformation assay at specific timepoints
270 after DNA induction (1 h, 3 h, and 5 h), see **Figure 5c** for the experimental setup. The addition of DNase
271 efficiently degrades exogenous DNA, preventing further DNA uptake⁴². Samples subjected to DNase
272 treatment were harvested 30min after the addition of DNase, while control samples, not subjected to
273 DNase, were harvested the next day. **Figure 5d** shows the transformation frequencies of the
274 corresponding timepoints. While the positive control in this experiment has a wider error bar, there was
275 no statistically significant difference between the DNase treated samples and the positive control. This
276 indicates the majority of DNA is taken up within the first hour and could explain why adding casamino
277 acids post-DNA induction does not impact the transformation frequency. These results indicate that
278 DNA is taken up efficiently within the first hour of induction. Further the growth curve shown in **Figure**
279 **S. 1** show that that CA was used by *V. cholerae* as a source of nutrition. Since the nutritional status has
280 been shown to be important for induction of competence⁹, we speculate one of two possibilities. Either
281 casamino acid availability functions as a nutrient source affecting competence induction, or the
282 combination of amino acids have counteracting inhibitory and stimulatory effects that ultimately limit
283 competence development, as suggested by Lichev et. al¹⁵.

284

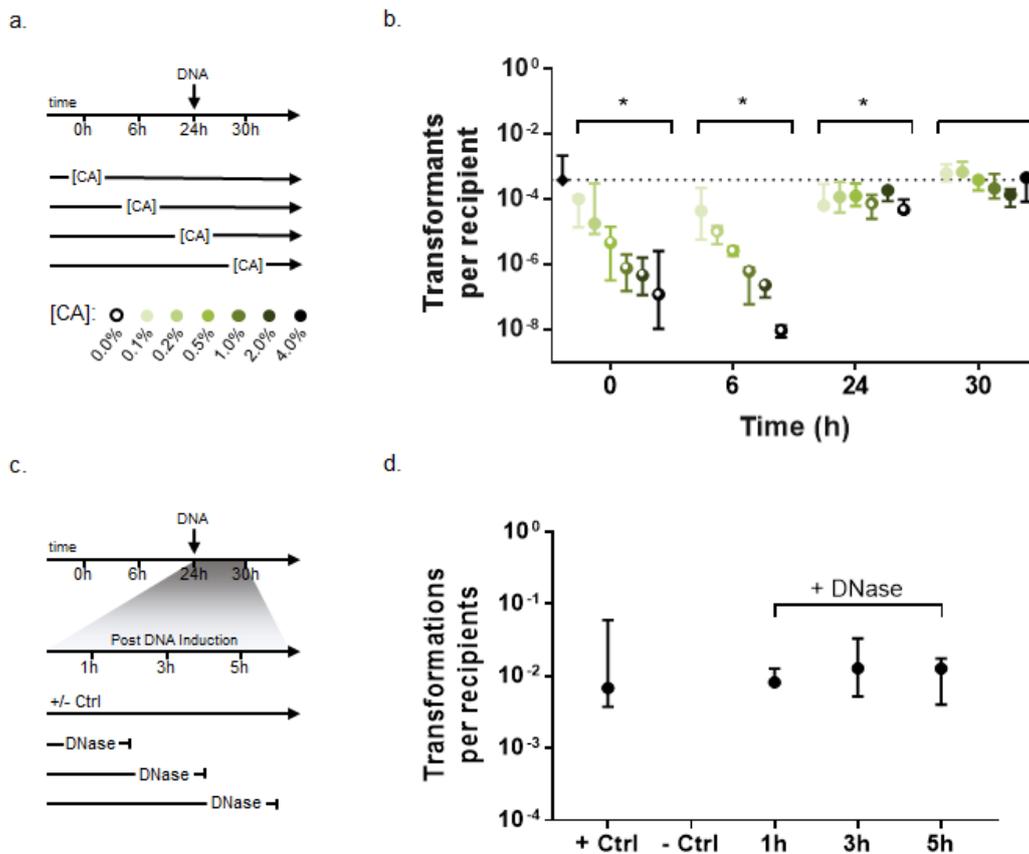


Figure 5 Amino acid supplementation reduces transformation efficiency. (a) Schematic of the experimental setup. At time point 0 h, a culture of wild-type *V. cholerae* in mid-exponential growth phase was exposed to chitin. At 24 h, DNA was added to the culture. Casamino acids (CA) were added to the culture at different concentrations and times during the natural transformation assay. CA concentrations ranged from 0 - 4% w/v and are depicted by increasing saturation of green. Timepoints assessed were 0 h, 6 h, 24 h, and 30 h. (b) Transformation frequencies with different CA concentrations added at the indicated time points. As positive control, no CA was added (black diamond). A dotted line extends from this point to illustrates the difference between the other samples. As negative control, no DNA was added, not depicted in graph as zero colonies grew on the selective media. An asterisk above the timepoint indicates CA addition resulted in a significant decrease in transformation for that timepoint. A white center of the dot indicates which transformation frequencies were significantly impacted. (c) Schematic overview of experimental setup with DNase treatment, no casamino acids were added. Wild-type *V. cholerae* was grown to mid-exponential growth phase and exposed to chitin for 24 h after which DNA was added to all cultures. DNase was added to the cultures at different times during the natural transformation assay. Timepoints assessed were 1 h, 3 h and 5 h post DNA induction. DNase treated cultures were incubated 30 min before cells were harvested. The positive (DNA receiving sample) and negative (non-DNA receiving sample) controls were not subjected to DNase treatment, and where harvest the following day. All conditions were performed in triplicates, the median and 95% CI is displayed. The statistics used was an unpaired two-tailed T-test performed on log transformed data.

287 An optimal concentration of amino acids promotes natural competence

288 There was a substantial transformation defect observed for the knockouts involving amino acid
289 metabolism genes, $\Delta metE$ and $\Delta asnB$ (**Figure 4**). The two genes are both catabolic. *metE* encodes for a
290 homocysteine methyltransferase involved in methionine formation. While, *asnB* encodes an asparagine
291 synthase catalyzing the formation of asparagine⁴³. External addition of casamino acids resulted in a
292 reduction of transformation frequencies for the wild-type. Specific amino acids, however, could rescue
293 the reduced transformation frequency seen for the defective amino acid metabolic genes. To investigate
294 this effect, we supplemented the $\Delta metE$ mutant and the $\Delta asnB$ mutant with methionine and asparagine
295 respectively, using wild-type *V. cholerae* as the control strain. Competence development was evaluated
296 by measuring transformation frequencies, which were assessed by transformants divided by total CFU
297 count. The effect on bacterial growth was measured by total CFU (**Figure 6**). As observed previously,
298 the wild-type had a transformation frequency of approximately 10^{-2} (**Figure 6a&c**). This did not change
299 with the addition of low concentrations of either methionine or asparagine. However, at concentrations
300 in the molar range (0.3 M methionine and 0.19 M asparagine), the transformation efficiency dropped
301 substantially to approximately 10^{-7} and is in a similar range as observed for high CA concentrations
302 added at the same time point (**Figure 5b**). As also expected, the deletion of *metE* and *asnB* substantially
303 decreased the transformation frequency to approximately 10^{-7} . However, in contrast to the wild-type,
304 the addition of small amounts of the respective amino acid increased the transformation efficiency. In
305 the case of $\Delta metE$, 0.3 mM methionine increased the transformation efficiency to wild-type levels at the
306 same amino acid concentration, thereby completely rescuing the phenotype. In the case of $\Delta asnB$, 0.19
307 nM, 0.19 μ M and 0.19 mM asparagine all increased the mutant's transformation frequency substantially
308 to approximately 10^{-5} . For both mutants, the addition of amino acids in the molar range decreased the
309 transformation frequency to a similar extent than the wild-type. The effect of the amino acids on total
310 bacterial growth are minimal, except for the highest tested concentration. Here, methionine reduced the
311 total CFUs of the wild-type and the $\Delta metE$ mutant, while asparagine increased the total CFUs of the
312 wild-type and the $\Delta asnB$ mutant (**Figure 6b&d**). Because both amino acids reduce the transformation
313 frequency at this high concentration, there is no correlation between growth and competence. The
314 mutants also had different growth rates compared to the wild-type, but both effects did not correlate
315 with transformation frequencies (**Figure 6b&d**). These results indicate that it is the lack of available
316 amino acid that reduces competence in contrast to the deletion of the gene itself. They also indicate an
317 optimal range of amino acid availability to promote natural competence, where the lack of an amino
318 acid reduces competence to a similar degree as the overabundance of amino acids.

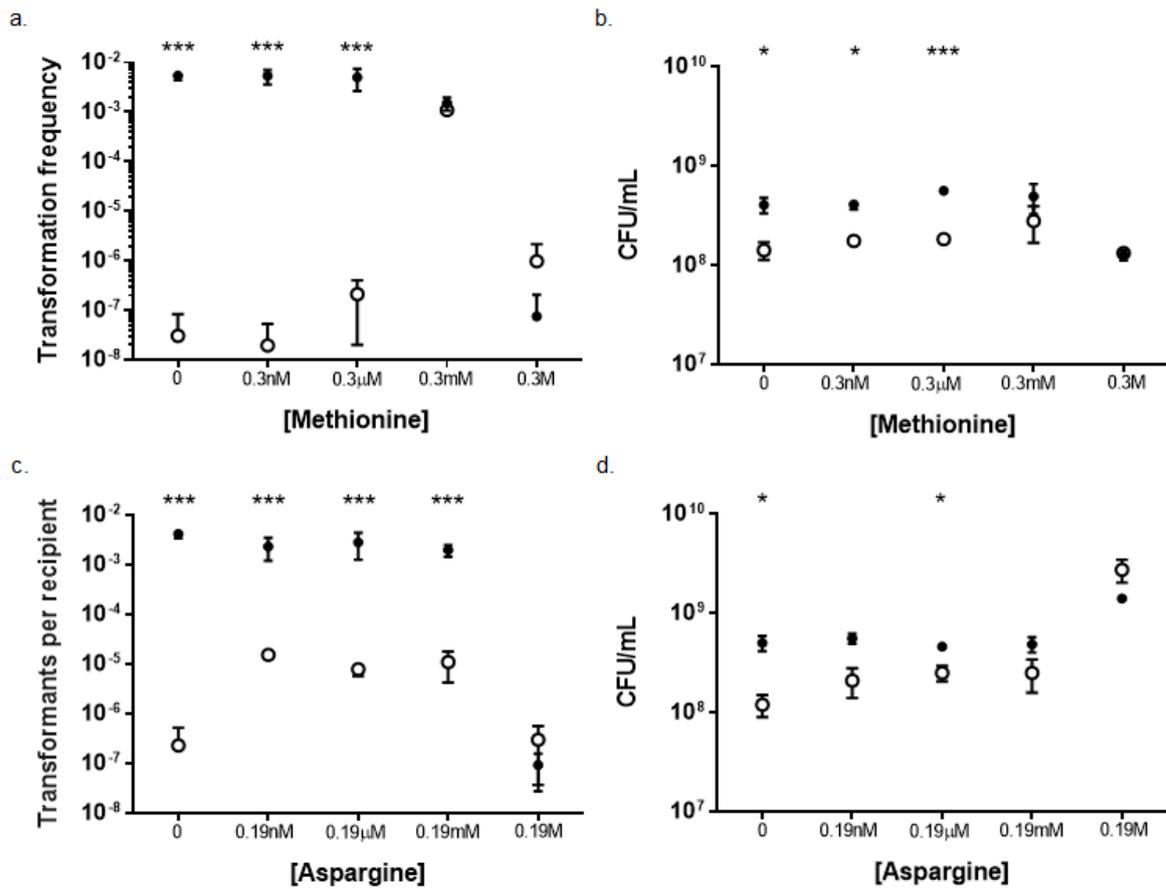


Figure 6 Addition of specific amino acids can rescue the transformation defect of amino-acid metabolism gene deletions. Amino acids were added (0 h) at different concentrations spanning four orders of magnitude to a standard transformation assay of the wild-type (filled black circles) or an amino acid metabolism gene deletion mutant (open black circles). The transformation frequencies and the total CFUs were recorded. (a) Transformation frequencies of wild-type and $\Delta metE$, supplemented with methionine. (b) Total CFUs of the same strains and conditions. (c) Transformation frequencies of wild-type and $\Delta asnB$, supplemented with asparagine. (d) Total CFUs of the same strains and conditions. All tests were performed in triplicates. The circles denote the mean; the error bars indicated the standard deviation. * denotes a p-value <0.05, *** denotes a p-value <0.01 in a Students t-test.

320

321

322

323

324

325

326 Redundancies within the amino acid pathways suggest why defective amino acid
 327 genes still undergo transformation
 328 Natural transformation is a multifaceted phenomenon that involved various pathways. In the case of *V.*
 329 *cholerae*, transformation can only occur when induced by chitin, indicating its niche-specific and
 330 metabolically distinct nature. Our research finding suggest that active amino acids metabolism is
 331 necessary for the integration of DNA. Consequently, an amino acid auxotroph, which relies on external
 332 sources of specific amino acids, would be unable to undergo transformation. To examine the impact of
 333 *metE* and *asnB* deletions in *V. cholerae*, we compared colony formation between mutant and wild-type
 334 strains on plates with and without amino acids. Cells were spread on M9⁺ agar plates, and a 10 µl amino
 335 acid solution was applied at the center of each plate separately. All amino acids, expect for methionine
 336 and asparagine, were dissolved in water at a final concentration of 1M and sterile filtered. Methionine,
 337 and asparagine, on the other hand, have a limited solubility above 0.3M, and 0.19 M, respectively. the
 338 plates were then incubated at 30°C, and growth zones were monitored over two days. The results,
 339 presented in **Table 4**, indicate the no auxotrophic strains grew without the presence of amino acids.
 340 However, the $\Delta metE$ exhibited growth in the presence of all tested amino acids, suggesting redundant
 341 pathways for methionine biosynthesis. It is important to note that methionine metabolism is highly
 342 conserved and interconnected with the tricarboxylic acid (TCA) cycle. Conversely, the $\Delta asnB$ mutant
 343 was unable to grow on minimal media, and other amino acids failed to rescue the phenotype. Neither
 344 the $\Delta metE$ nor the $\Delta asnB$ mutants displayed impaired growth in minimal media. *MetE* serves as a one
 345 of two-methionine synthase, with the other being *metH*. According to the literature, disruption of *metE*
 346 function renders *V. cholerae* incapable of growing in minimal media, while disruption of *metH* does not
 347 exhibit a phenotypic effect³⁵. Additionally, our TIS screen did not identify *metH* as an essential
 348 component. As for asparagine synthesis, *V. cholerae* has two, namely *asnA* and *asnB*. *AsnA* catalyzes
 349 the ammonia-dependent synthesis of asparagine, while *asnB* converts aspartic acid into asparagine,
 350 utilizing either glutamine or ammonia as the nitrogen source in an ATP-dependent process. In our TIS
 351 screen, *asnA* was found to be non-essential and thus was not tested. Although the $\Delta metE$ mutant strains
 352 were rescued by redundant mechanisms, the same was not observed for the $\Delta asnB$ mutant. This suggests
 353 that *asnB* primarily functions as the main synthase involved in asparagine biosynthesis under these
 354 conditions, leading to an auxotrophic phenotype in minimal media. This finding explains why the $\Delta asnB$
 355 mutant exhibited lower transformation frequencies compared to the $\Delta metE$ mutant (**Figure 4**).

356

357
 358
 359
 360

Table 4: Amino acids drop test

	NA	Met	Gly	Ans	Gln	Pro	Ser
Wt	×	↑	↑	↑	↑	↑	↑
$\Delta metE$	×	↑	↑	↑	+	↑	↑
$\Delta asnB$	×	×	×	↑	×	×	×

× no growth, + slight growth, ↑ growth

364

365 Reporter constructs suggest heterogeneity determines competence development

366 Mutants with defective amino acid metabolism exhibited reduced transformation frequencies. The data
367 provided suggest the observed decrease in transformation frequencies could be attributed to impaired
368 competence induction. We therefore opted to investigate the transcription of *comEA*. ComEA is essential
369 to natural transformation and is a DNA binding protein found in the periplasm⁸. As the induction of
370 natural transformation relies on coordination of several pathways, we opted to investigate in parallel
371 carbon catabolite repression (CCR), nucleotide scavenging and quorum sensing (QS)¹³. We investigated
372 the involvement of these pathways among the amino acid metabolism defective mutants, the wild-type
373 and the *tfoX* knockout. TfoX, and CytR are master regulators and co-regulate natural competence
374 induction^{8,14}. TfoX transcription is induced on chitin and facilitates expression of both *comEA* and *cytR*.
375 The induction of *comEA* therefore serves as a biomarker for successful competence induction as well as
376 functional chitin metabolism. The transcriptional repressor, CytR, regulates genes involved in
377 scavenging and metabolizing nucleotides. The expression of TfoX is reduced in a Δ *cyaA* mutant
378 defective in synthesizing the secondary messenger cyclic adenosine monophosphate (cAMP).
379 Supplementation of cAMP restores the expression of TfoX⁴⁴. Altogether this demonstrates a link
380 between CyaA and competence induction. Finally, bacterial populations can collectively control gene
381 expression via QS. In *V. cholerae*, high cell densities relieve repression of the QS master regulator,
382 HapR. HapR then positively regulates the expression of *comEA*, which is required for successful
383 competence induction⁴⁵ via a cofactor QstR³⁹. A direct indicator of HapR activity is *hapA* promoter
384 activity, as HapR directly derepresses the *hapA* promoter⁴⁶. By using the promoter regions of *comEA*,
385 *cyaA*, *cytR* and *hapA* as reporters, we aimed to compare if the reduced transformation frequencies were
386 related to either an impaired competence inducitor, nucleotide scavenging, CCR, or QS.

387 We implemented the previously published transcriptional reporter fusion construct (pBR-[P_{*gyrA*}]GFP-
388 [P_{*comEA*}]DsRed)⁴⁷. This plasmid expresses the fluorescent proteins DsRed.T3[DNT]⁴⁸ under control of the
389 *comEA* promoter, and GFP.mut3⁴⁹ the under control of the promoter of *gyrA* (housekeeping gene),
390 respectively. We created additional constructs by replacing the promoter region of *comEA* with either
391 *cyaA*, *cytR*, or *hapA*. The cells were incubated on chitin for 24 h, harvested, and imaged. The
392 fluorescence of each cell was quantified for both channels. The ratio between the reporter (red
393 fluorescent protein, RFP) and internal control (green fluorescent protein, GFP) is displayed for each
394 reporter in each strain background, see **Figure 7**.

395 For the *comEA* promoter activity the [P_{*comEA*}]dsRed expression relative to [P_{*gyrA*}]GFP expression are
396 depicted in **Figure 7a**; **Figure 7b** for representative images. The wild-type population demonstrated the
397 highest ratio between RFP and GFP signal (wild-type: median = 0.54, Δ *metE*: median = 0.26, Δ *asnB*:
398 median = 0.13, Δ *tfoX*: median = 0.06). In addition to a weaker signal, only a fraction of the population
399 was RFP positive. This is in accordance with previous results⁴⁷. All mutant backgrounds expressed
400 significantly lower intensity ratios relative to the wild-type population (Kruskal-Wallis test followed by
401 Dunn's test set with wild-type as control group, all p-values <0.0001). Both Δ *metE* and Δ *asnB* displayed
402 greater interquartile ranges than the wild-type and Δ *tfoX*. This suggests greater variation in expression
403 levels for Δ *metE* and Δ *asnB*. The Δ *asnB* population showed a more skewed distribution than Δ *metE*,
404 because most cells had very low P_{*comEA*} activity, while a small fraction of cells had high activity (**Figure**
405 **7b**). The wild-type population had lower variation, and the signal variation in the Δ *tfoX* population was
406 even lower. This is in accordance with the literature as TfoX is strictly required for *comEA* expression^{8,17}.

407 For the *cyaA* promoter activity relative to the control is displayed in **Figure 7c**, representative images
408 in **Figure 7d**. The RFP/GFP ratios differed significantly between mutant background and wild-type
409 (Kruskal-Wallis test followed by Dunn's test set with wild-type as control group, all p-values $\Delta metE$:
410 <0.0001 , $\Delta asnB$: 0.0391, $\Delta tfoX$: 0.0002). Interestingly, the $\Delta metE$ population exhibited a lower ratio
411 (median = 1.87) compared to the wild-type (median = 6.20), but with greater variation. In contrast,
412 $\Delta asnB$ (median = 6.59), and $\Delta tfoX$ (median = 6.47) populations had a slightly higher ratio compared to
413 the wild-type.

414 The RFP/GFP ratios for the *cytR* promoter constructs were noticeably higher in the mutant backgrounds
415 compared to the wild-type (Kruskal-Wallis test followed by Dunn's test set with wild-type as control
416 group, all p-values <0.0001). Similar to the $\Delta tfoX$, the amino acid metabolism defective mutants appear
417 to have derepressed *cytR* (**Figure 7e**, representative images in **Figure 7f**).

418 The *hapA* promoter constructs demonstrated significantly lower ratios for the mutants compared to the
419 wild-type (Kruskal-Wallis test followed by Dunn's test set with wild-type as control group, all p-values
420 <0.0001). Additionally, the mutants exhibited lower variation (**Figure 7g**, representative images in
421 **Figure 7h**).

422 These overall results demonstrate dysregulation of several pathways in the mutants, ultimately failing
423 to induce competence at wild-type levels. This could explain the reduced transformation frequencies
424 observed previously for the amino acid metabolism defective mutants. These findings indicate that an
425 intracellular balance of amino acids is important for natural transformation by aiding competence
426 induction in *V. cholerae*.

427

428

429

430

431

432

433

434

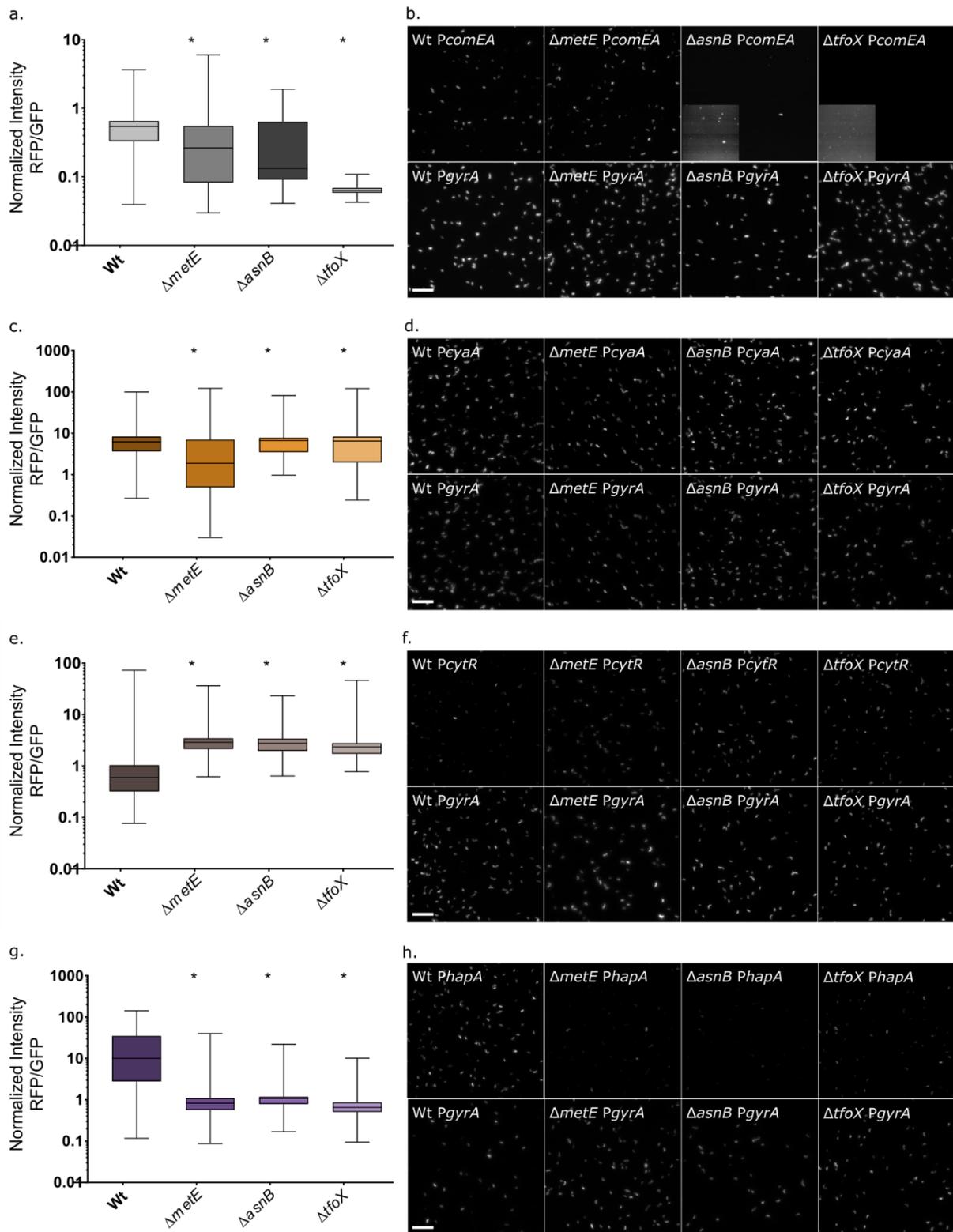


Figure 7 Visualization of the expression of associated pathways required for competence expression within deletion strains. (a., c., e., g.) Shows boxplots of the RFP/GFP ratios for the *comEA*, *cyA*, *cytR*, and *hapA* promoter constructs, respectively. (b., d., f., h.) Depicts representative images of each promoter construct *comEA*, *cyA*, *cytR*, *hapA* respectively. The top row shows the fluorescent channels for RFP signal for wild-type, $\Delta metE$, $\Delta asnB$, and $\Delta tfoX$. The bottom row shows the corresponding GFP images, the [*P_{gyrA}*]GFP internal control. All images were leveled to the wild-type image of each dataset. * indicates a statistical significance, $p < 0.05$. The scalebar is 2 μm .

436 Discussion

437 Natural transformation is an elaborate and energy-consuming process that for *V. cholerae* occurs under
438 specific conditions. Further, the current prevalent El Tor biotype is more efficient in DNA uptake than
439 its predecessor and it is speculated that this could be the reason for why the classical strain was
440 replaced⁵⁰. To investigate the complexity of natural transformation, we generated a robust transposon
441 library within *V. cholerae* A1552, the analysis of which can be a useful resource for future studies
442 involving DNA uptake and recombination (**Table S.3**). Our analysis revealed a novel and prominent
443 role for amino acid metabolism, where the individual genes modulated the natural transformation
444 frequencies. Deletion of amino acid synthases result in reduced transformation frequencies, whereas the
445 addition of specific amino acids either fully or partially recovered the phenotype. Further we found a
446 cell-to-cell variation within the transcriptional activity of the competence-associated gene, *comEA*. Our
447 results show that amino acid metabolism genes modulate the progression of natural transformation.

448 Transposon insertion sequencing is a powerful tool to investigate the genetic response to a specific
449 condition. A drawback when using this method is experimental noise, mainly referring to genetic drift
450 and sampling error²⁵. The genetic drift and sampling will directly impact the complexity and saturation
451 of the screen with further implications for the analysis. Any experimental bottlenecks, such as sample
452 preparation, should therefore be considered and minimized and the biological bottleneck should be
453 sampled thoroughly. Our major bottleneck was low transformation frequencies, as only a fraction of
454 cells takes up DNA and recombine. However, this resulted a random loss on insertion. We therefore
455 optimized the transformation assay to obtain a higher number of transformants, sampled in excess and
456 compared the donor DNA receiving sample to a control population. In addition to identifying a role for
457 amino acid metabolism, our genetic screen confirmed a large number of genes involved in both DNA
458 uptake and recombination. Additionally, the TIS output should be validated by other approaches to
459 assess the expected gene fitness derived from the screen⁵¹. Our TIS screen identified 263 genes to be
460 conditionally essential. Out of which we generated a validation set of knockouts that all resulted in
461 reduced transformation frequencies (**Table S. 2**, and **Figure 4**), and corroborated corresponding
462 published phenotypes. It should be noted that in our screen deletion of *tsaP* abolished transformation
463 frequencies, however in Jaskólska, et al. this deletion demonstrated only a 10-fold reduction compared
464 to wild-type³⁷, likely due to slight variations in experimental conditions. To our knowledge, the genes
465 *pmg*, *asnB*, *metE*, *carB*, *lapD*, and *ppc* have not been characterized in terms of natural transformation.
466 Furthermore, our screen was conservatively assessed, found to be robust and can be a useful resource
467 for future studies involving DNA uptake and recombination.

468 We find that the defective amino acids metabolism dysregulates pathways upstream of competence
469 development (**Figure 7**). The *comEA* promoter activity is the only pathway that displays a graded
470 response depending on the mutant background (**Figure 7a**). Whereas the defective amino acids' impact
471 on competence development also appears to affect the other pathways nucleotide scavenging, CCR, and
472 QS. The overall irregular *cyaA* expression detected in our reporter assay could impact nucleotide
473 scavenging potentially by affecting the levels of cAMP which would have repercussions for *tfoX*
474 expression⁴⁴, further inhibiting competence development. The amino acid metabolism mutants also
475 seem to have derepressed CytR promoter activity, it is thus conceivable that the CCR pathway is not
476 functioning optimally in the mutants. Further, the result for the *hapA* promoter activity indicates that
477 HapA is downregulated suggesting an inhibitory effect on HapR in the mutants, which would have
478 implications for QS. The amino acid metabolism mutants demonstrate throughout statistically different
479 promoter activity from wild-type following instead a similar trend to that of the $\Delta tfoX$. TfoX is a master

480 regulator for natural transformation. It is therefore intriguing that the amino acid mutants behave
481 similarly within all flanking regulatory pathways expect for the *comEA* promoter activity, that instead
482 demonstrates a graded response, more comparable to the transformation frequencies. Further we see
483 more variation in the transformation frequency for $\Delta metE$ as compared to $\Delta asnB$ (**Figure 4**). We also
484 find that the *asnB* mutant is a true auxotroph while the *metE* mutant can be rescued by redundancies
485 within the amino acid pathway (**Table 4**). Likely contributing to the variation we detect indicating more
486 heterogeneity within the defective amino acid metabolism mutants.

487 The defective amino acid metabolism mutants demonstrated lower *comEA* promoter activity than wild-
488 type levels, further demonstrating a graded response within the amino acid metabolism genes. This
489 indicates that the reduced transformation frequencies could be explained by a role of amino acid
490 metabolism in competence induction and or chitin metabolism, as the transformation frequencies follow
491 a similar trend (**Figure 4**). The defective amino acid metabolism genes have varying responses to the
492 DNA uptake and integration while the transformation frequencies for $\Delta tfoX$ and *comEA* reporter activity
493 is inhibited. Further we observe heterogeneity within the defective amino acid metabolism genes, is
494 more prominent within $\Delta asnB$, compared to $\Delta metE$. A fraction of the $\Delta asnB$ population demonstrated
495 comparable *comEA* promoter activity at or above wild-type levels. It is therefore plausible that these
496 cells are competent and capable of taking up and integrating external DNA and could explain the barely
497 detected transformation frequencies we observe for the amino acid metabolism mutants. On a population
498 basis we find that the defective amino acid metabolism genes severely reduce transformation
499 frequencies, while on a single cell level they appear to be a mixture of varying levels of competence.
500 We, therefore, hypothesis that there is an optimal threshold for *comEA* promoter activity to achieve
501 competence induction, this being established by an optimal amino acid metabolism, comparable to
502 competence induction in *Bacillus subtilis*⁵². Where inherent molecular noise results in a bistable gene
503 expression resulting in competence development⁵³.

504 Addition of external amino acids had an inhibiting effect on transformation frequencies (**Figure 5b**). As
505 far as we know the only studies that investigates the effect of amino acids on natural transformation was
506 performed by Lichev et. al¹⁵, and Wilson et. al¹⁶. Both studies were performed in Gram-positive bacteria
507 *Micrococcus luteus*, and *Bacillus subtilis*, respectively. In the Lichev study they propose a regulatory
508 model, where amino acids have both stimulatory and inhibitory roles impacting competence
509 development. Wilson et. al detect stimulatory and inhibitory effects of amino acids on competence
510 development when added during early stationary phase of growth. Stating that amino acids play a
511 defining role in competence development, and they could detect either an stimulatory or inhibitory
512 response within 5 min of addition of amino acids. In our study we show that casamino acids significantly
513 reduce transformation frequencies in both a concentration- and time- dependent manner. It is probable,
514 casamino acids can function as nutrition⁵⁴, inhibiting competence development via the catabolite
515 repression pathway as this is a prerequisite for competence induction¹³. Interestingly we find this to be
516 alleviated in a time-dependent manner. At the 24 h and 30 h time points we find that the transformation
517 frequencies are impaired to a much lesser extent, and only for the higher concentrations. At the 24 h
518 timepoint when DNA and casamino acids are added at the same time, it is feasible the effect of casamino
519 acids takes far longer to exert an effect while DNA internalization is immediate. This is supported by
520 our finding demonstrating that maximum transformation frequencies are achieved within 1 h post DNA
521 induction (**Figure 5d**). It has been hypothesize that there are several pre-assembled or semi-assembled
522 DNA uptake machinery present within the cell⁵⁶, indicating that DNA internalization is rapid as the cell
523 machinery appear to be somewhat primed for assembly. Further, there is evidence of immediate DNA

524 uptake within *V. cholerae* A1552 Δ *dns* when donor DNA is simultaneously inoculated on chitin¹⁷. In the
525 same study Blokesch et al. further show that *V. cholerae* needs to be on chitin for at least 6 h before
526 transformation frequencies can be detected. As we find distinct fractions or subpopulations that exhibit
527 wild-type level *comEA* promoter activity after 24 h on chitin, it is likely that competence is already
528 induced and the cells are ready for DNA uptake minimizing the effect of casamino acids. These cells
529 are likely to be competent and achieve transformation frequency. Recombination, however, takes longer
530 than internalization and has been found to require several attempts⁵⁷.

531 Within the context of our study, additional amino acids could function as nutrition hindering competence
532 induction. However, as the concurrent addition of amino acids and donor DNA did not significantly
533 impact transformation frequencies, likely due to rapid internalization of DNA uptake and possibly a
534 readiness for DNA that trumps amino acids. Further we also found less growth in samples receiving
535 amino acids concurrently or after donor DNA (**Figure S 1**), suggesting that the presence of donor DNA
536 could initiate a growth arrest and promote DNA uptake. This has been seen to occur in both *Bacillus*
537 *subtilis*⁵⁸, and *Acinetobacter baylyi*⁵⁹. Taken together with our results this could suggest competence
538 induction is a binary mechanism aided by optimal amino acid metabolism, and further that once the
539 threshold for competence is breached, the cells will attempt DNA internalization. This could be an
540 intriguing strategy to sequester DNA as a resource, somewhat similarly to the efficient uptake and
541 utilization of chitin that *V. cholerae* exhibits^{60,61}.

542 Our findings indicate that amino acid metabolism plays a stabilizing role in the progression of
543 competence induction and natural transformation. Further, we speculate that functional amino acid
544 metabolism act as a checkpoint. However, the specific function of this checkpoint remains unclear.
545 Further investigation is required to precisely determine the role of amino acid metabolism within the
546 process of natural transformation.

547

548

549 Acknowledgment

550 We would like to thank Melanie Blockesch and the Laboratory of Molecular Microbiology for the
551 donation of the strain *V. cholerae* A1552 Wild type, O1 El Tor Inaba, RifR, and reporter constructs:

- 552 - pBR-GFP_dsRED_Kan
- 553 - pBR-GFP-[PcomEA]dsRED-Kan
- 554 - pBR-GFP PgyrA_PcomEA dsRED_Kan

555

556

557 Declaration of Interests

558 The authors declare no Competing Financial or Non-Financial Interests.

559 Methods

560 Strains, media, and culture conditions

561 Wildtype *Vibrio cholerae* A1552 was grown on Luria Broth agar plates supplemented with 50 ug/mL
562 rifampicin (Rif), while the *V. cholerae* transposon mutant library was selected on LB/Rif/Km,
563 kanamycin (Km) concentration was 50ug/mL. *Escherichia coli* SM10 lambda pir, carrying the
564 conjugative suicide vector pSC186, was grown in LB supplemented with 50ug/mL carbenicillin (Carb).
565 All *V. cholerae* and *E. coli* strains were grown overnight at 37°C unless otherwise indicated. See **Table**
566 **5** for complete strain list.

567 Transposon insertion sequencing

568 Library construction by conjugation

569 Random insertion of the transposon occurred upon delivery of a non-reproducible plasmid containing a
570 transposable element through conjugation. The transposable element is a modified
571 Tn *Mariner* containing a kanamycin (Km) resistance gene flanked by inverted repeats. By selecting for
572 the transposable element after conjugation, a library of *V. cholerae* was generated. Every single
573 bacterium has one transposon randomly inserted within its genetic material, which potentially disrupts
574 a gene function. pSC186 was conjugated into *V. cholerae* A1552 in independent reactions. For each
575 conjugation reaction 250 µl overnight culture *E. coli* SM10 λ pir carrying pSC186 and 250 µl overnight
576 culture of *V. cholerae* A1552 was washed in PBS (1min x 13'000xg), pelleted and finally resuspended
577 in a total of 50 µl PBS. The final resuspension was then spotted onto 0.45mm HA filter (Sigma) placed
578 on LB plates and incubated for 1.5h at 37°C to allow for conjugation, along with single strain controls
579 treated similarly. This yielded 41'117 transposon mutants per conjugation. 139 conjugation reaction
580 were used for the library to ensure high density and maximum coverage of potential insertion sites. After
581 conjugation the filters were transferred to Eppendorf tubes containing 1 mL LB and resuspended by
582 vortex for 20 s and pipetting. Resuspended cells from all filters were pooled pelleted and resuspended
583 in 255 ml LB, the entire volume was plates. 100 µl distributed amongst 251 LB/ Rif/ Kan plate (100
584 mm x 15 mm) and grown at 30°C overnight, ~18 h. The input library yielded ~5.7 M CFUs, where each
585 colony represents one insertion mutant. The remaining culture was plated on LB Sm to check for
586 spontaneous.

587 Input transposon library applied to optimized natural transformation assay

588 The TIS library was grown to early exponential phase, transferred to supplemented M9⁺ media, and
589 incubated on chitin flakes for natural competence induction. The next day, the populations were divided
590 into two conditions, as follows: One population received donor DNA denoted as the "TnNT" population.
591 The second " control " population did not receive donor DNA. Both populations became competent
592 during overnight incubation due to exposure to chitin, however only the population receiving donor
593 DNA could take up the DNA and undergo chromosomal integration. The following day, the TnNT
594 population was removed from chitin by vortexing. A fraction of the sample (450µl) was used for colony
595 determination by serial dilution and plated on both selective (Sm) and non-selective LB plates. The ratio
596 of the population sizes in absence and presence of Sm was used to determine the transformation
597 frequency. The control assay was plated on non-selective LB agar plates. Additionally, the control
598 sample was also tested for spontaneous resistance to Sm by plating selective plates. The remaining liquid
599 culture of both conditions was plated on 15mm Petri dishes for outgrowth and downstream sequencing
600 preparation. The control sample was plated on non-selective plates, while the TnNT was plated on plates

601 containing Sm. All colonies were harvested the next day and prepared for whole genome sequencing. A
602 total of 4.2×CFUs were harvested for the control assay. While the TnNT assay consisted of 1.3×CFUs.

603 Transposon insertions sequencing analysis after natural transformation assay

604 Before the sequences could be statistically analyzed the sequences were trimmed and unwanted
605 sequence were removed, using cutadapt⁶⁸. All sequences lacking the adapter were discarded. The 5' end
606 of all sequences were trimmed, and short sequences were discarded, to avoid low quality reads. The
607 sequences were mapped to the genome using Bowtie2⁶⁹. The Bowtie2 default mapping criteria were
608 used, as this was stringent enough. The sequence statistics are summarized in **Table 2**. The control
609 sample yielded 28.8M reads with 85% TA sites detected. The TnNT output had 30.5M reads and reached
610 a TA insertion of 72%. Both the TnNT and control output were the result of two merged runs. The
611 control sample had almost twice the amount of reads as compared to the initial input library. We
612 sequenced the sample much deeper and therefore assume the input is slightly higher than 84%. However,
613 85% is likely the maximum as the total number of hits does not differ significantly even though the
614 number of reads almost doubled. The number of hit sites between TnNT output sequencing runs # 1,
615 and # 2 are slightly different as the merged dataset has a slight increase in the percentage of TA sites
616 hit. This is minor and does not affect the main results, therefore the analysis was performed on the
617 concatenated sequencing data. The two datasets were compared using the ARTIST pipeline²⁵. In brief,
618 the datasets are normalized to limit replication bias. The analysis then finds the individual likelihood of
619 each TA site being disrupted for each dataset separately. By implementing a sliding window analysis
620 that is used to inform a Hidden Markov Model (HMM). The resulting datasets are then used to simulate
621 several experimental bottleneck events. The disruption ratios are run through the simulated bottlenecks
622 to assess the probabilities of the transposon insertions and evaluate whether the locus is essential. This
623 allows us to see which genes are important for natural transformation. Due to the resolution of our input
624 data and the statistical analysis of the pipeline, we can also assess intergenomic regions as well as detect
625 genetic domains that are important. However, due to the amount of data obtained, this was excluded
626 from our analysis.

627 TIS sample preparation and sequencing

628 The entire culture of TnNT and outgrowth was used for gDNA extraction. The leftover purified gDNA
629 was stored at -20 degrees. Genomic DNA was extracted following the Promega protocol. At least 20ml
630 of each sample underwent lysis, protein removal and DNA precipitation via 100% isopropanol. The
631 gDNA pellet was carefully washed in 75% ethanol and left to dry at RT for 15-30min. Each gDNA
632 pellet was resuspended in 100ul nuclease free water and incubated overnight at 4°C. We avoided flicking
633 and mixing the liquid in the tube as this resulted in dissolving impurities. The next day the purity and
634 concentration of each sample was examined by nanodrop. Only the highest concentrations and purest
635 samples were used for downstream sample preparation. gDNA was randomly fragmented to 600-300 bp
636 pieces by sonication with a Covaris M220 focused-ultrasonicator. 25-30ug gDNA was pooled and
637 aliquoted into covaries tubes (microTUBE-50 AFA Fiber Screw-Cap). Each tube could take 5ug gDNA
638 each, supplemented with 5µl low EDTA buffer (TE Buffer pH 8.0: 10mM Tris-HCl pH8.0, 1mM EDTA
639 pH 8.0.) the ends of the DNA fragments were adapter ligated (using previously published sequences)
640 and subjected to two rounds of PCR amplification with the main purpose being to enrich transposon
641 insertions flanked by genomic DNA and as well as attaching the hybridization sequences P5 and P7,
642 including a barcode and spacers. PCR reactions were carried out with Phusion or Onetaq, according to
643 the manufacturer's instructions, primers are listed in **Table 7**. Before sequencing the fragments were
644 purified by gel extraction 2% agarose in TBE buffer, 300-600bp was excised. Sequencing was

645 performed with an Illumina MiSeq sequencer using the MiSeq V3, 150-cycle kits. The resulting reads
646 were evaluated using FastQC, processed with cutadapt, and mapped to the genome with Bowtie2. A
647 published pipeline known as ARTIST pipeline was followed^{25,70,71}.

648 Natural transformation assay: optimizing the transforming DNA size and 649 concentrations

650 The natural transformability of *V. cholerae* was tested by incubating the bacterial on chitin in minimal
651 media. The transformation assay was based on the optimized protocol from Blokesch⁷², with slight
652 modifications made to the media due to solubility issues. The transforming material was a PCR fragment
653 amplified from genomic DNA of c6706. This is a non-transforming strains due to a frameshift that
654 results in insufficient expression of LuxO and ultimately impairs quorum sensing (Stutzman and
655 Blokesch, 2016). The strain contains a point mutation within the *rpsL* gene that confers resistance to
656 streptomycin. We avoided gDNA as this could lead to uptake of DNA that would be outside of our
657 detection. We also tested different concentrations of our transformation DNA (tDNA), as well as varying
658 stretched of DNA flanking the point mutation (500bp 1000bp, and 2000bp). the tDNA with 2000bp had
659 the highest recombination frequency. We also tested the amount of DNA 0.5, 1µg, 2µg, 4µg, and 10µg.
660 2µg gave a sufficiently high transformation frequency, close to saturation, while being experimentally
661 efficient. The transformation assay was performed using chitin flakes from shrimp shell (Sigma). The
662 frequencies were calculated by dividing the number of streptomycin resistant transformants on the total
663 number of CFUs. the standard deviation was shown as error bars. and are a result of three biological
664 replicates if not otherwise specified.

665 Computation analysis, COG

666 Function annotation was performed using the eggNOG-mapper v5.0²⁶. The sequences were uploaded to
667 the eggNOG database and blasted the *V. cholerae* genome. The COG assignments for then used for an
668 enrichment analysis, which was calculated as the fraction of conditionally essential group found in each
669 COG category divided by the fraction of the whole genome with that COG⁷³. A random subset of 36
670 hits were chosen and cross-referenced to ensure the COG annotations were congruent with other
671 databases available online: KEGG⁷⁴, PATRIC⁷⁵, and UniProt were used. We found no discrepancies
672 between the COG annotations and these databases. Some sequences were associated with several COGs,
673 aiming to minimize bias in the analysis we included all assigned COGs. An example of such a sequence
674 is PilU. It was assigned both cell motility, and intracellular trafficking. The protein is involved in the
675 type IV pilus and functions as a PilT-dependent retraction ATPase⁷⁶. The pilus is important for cell
676 surface anchoring and extends and retracts through a pore in the outer membrane of *V. cholerae*. Several
677 other pilus proteins also had this “double” classification, indicating there is functional overlap between
678 the two COGs. Regardless, the minor differences detected between the COG annotations provided by
679 eggNOG and other databases did not impact the screen as a whole. The COG enrichment analysis
680 revealed that amino acid metabolism was a major functional group.

681 Construction of *V. cholerae* knockouts

682 Knockouts were generated by allelic exchange vector. PCR products obtained from primers listed in
683 **Table 6** were ligated into pCVD442 via isothermal assembly (ITAC) cloning, generating deletion
684 plasmid for allelic exchange. This vector does not replicate in *V. cholerae* and allows for sucrose counter
685 selection, as described previously⁷⁷. The deletion constructs disrupt the gene of interest with a TetAR,
686 conferring resistance to tetracycline, creating in frame deletions leaving 30bp of the start and end of the

687 gene. The tetracycline resistance marker was flanked by 2kb homology regions. The entire 6kb region
688 was PCR amplified and used as donor DNA in our optimized natural transformation assay, transformants
689 were selected on tetracycline 0.75ug/ml. This is above the MIC obtained for our A1552 strain.
690 Traditional allelic exchange did not function as the tetracycline resistance mechanism in concert with
691 levansucrase, encoded by *sacB*, increases the counter-selection by orders of magnitude⁷⁸, and likely had
692 a toxic effect on *V. cholerae*. While the plasmids were already created we attempted homologous
693 recombination using the natural transformation protocol^{72,79}.

694 Addition of casamino acids and addition carbon sources to the natural transformation 695 assay.

696 Three independent ON cultures of wild-type *V. cholerae* were diluted 1:100 in LB and grown with
697 aeration at 37°C until the cultures reached an OD₆₀₀ of 0.3. After which, 800 µl of each culture was
698 divided into Eppendorf tubes, spun down (5600 x g 5 min, RT), washed in M9+ media spun down again,
699 resuspended in M9+ media either containing CA or not. 18 tubes across triplicates were resuspended in
700 M9+ media containing CA, this would be timepoint zero. Each CA concentration was performed for
701 each triplicate and were as follows: 0.1 %, 0.2 %, 0.5 %, 1.0 %, 2.0 %, 4.0 % w/v CA, final volume
702 equaled 1mL for all samples. Each resuspended pellet was then transferred to a culture tube containing
703 chitin flakes per triplicate, and incubated statically at 30°C. After 6 h 18 tubes containing M9+ media
704 were supplemented with concentrated CA, stock was 20%, to reach the final concentration of the
705 previously mentioned concentrations and placed back into 30°C. This was the 6h timepoint. After 24h
706 all tubes were carefully decanted and fresh M9+ media was added. All tubes previously containing CA
707 received the same concentration again, and an additional 18 tubes received the CA concentrations
708 mentioned above, this is the 24 h timepoint. Before placing the tubes back into the incubator all tubes
709 received 2 µg freshly prepared donor DNA. 6 h later another 18 samples received concentrated CA to
710 reach the previously mentioned CA concentrations, the 30h time point 48 h after the experiment initiated
711 the samples were harvested by vortex for 20 s RT and plated for CFU count on LB agar both with and
712 without Sm.

713 DNase treatment

714 Three independent ON cultures of wild-type *V. cholerae* were diluted 1:100 in LB and grown with
715 aeration at 37°C until the cultures reached an OD₆₀₀ of 0.3. After which, 800 µl of each culture was
716 divided into Eppendorf tubes, spun down (5600 x g 5 min, RT), washed in M9+ media, spun down, and
717 resuspended in M9+ media. The suspensions were transferred to chitin flakes and incubated statically at
718 30°C, all samples were prepared in triplicates. After 24 h donor DNA was added to all samples except
719 for the negative controls. 1 µg DNase was added to the sample tubes after 1, 3, and 5 h. The DNase was
720 incubated for 30min in each sample before immediate removal from chitin by vortex for 20 s and plated
721 for CFU count on LB agar both with and without Sm. Both the positive and negative controls were left
722 on chitin until the next day, followed our standard natural transformation assay.

723 Amino acid drop test

724 ON cultures of wild-type, $\Delta metE$, and $\Delta asnB$ strains were diluted and grown to an OD of 0.1. cells were
725 evenly spread on LB agar plates. A line was drawn down the middle of each plate. On one side, the
726 wild-type cells were spread evenly on the other half either $\Delta metE$ or $\Delta asnB$ was spread, special care was
727 taken to avoid the two strains touched. Fresh 1 M stocks of amino acids were prepared of glycine,
728 proline, and serine. Methionine 0.3M, asparagine 0.19 M and glutamine 0.2 M. Amino acids tested

729 were and methionine. 10µl of each amino acid solution was placed at the center of the plate between the
730 two strains, wild-type worked as a growth control. Each plate was prepared in triplicate, using the same
731 amino acid stock. Plates were incubated at 30°C for 18 h and checked for growth. After which the plates
732 stayed at RT and were checked for growth every day for three days total.

733 Promoter reporter constructs, epi fluorescent microscopy settings

734 The reporter construct pBR-[P_{gyrA}]GFP-[P_{comEA}]dsRed, along with an empty plasmid lacking promoter
735 regions was gifted by Blokesch⁴⁷. Using the empty plasmid we created three additional reporter
736 constructs: pBR-[P_{gyrA}]GFP-[P_{cyaA}]dsRed, pBR-[P_{gyrA}]GFP-[P_{cytR}]dsRed, and pBR-[P_{gyrA}]GFP-
737 [P_{hapA}]dsRed. The constructs are transcriptional fusions of promoter regions, P_{cyaA}, P_{cytR}, and P_{hapA}, and
738 genes encoding a fluorescent protein (FPs). The FPs used were GFP.mut3⁴⁹, and DsRed.T3[DNT]⁴⁸.
739 This is a slightly outdated FP pairing as the protein folding for the GFP variant is much faster than
740 DsRED, however, this did not impact our assay as the timeframe for induction was 24h and protein
741 stability was more important. Additionally, we decided to keep this pairing to allow for comparison with
742 the original study⁴⁷. The co-expression of these FPs is also well-suited as they have distinct peaks. The
743 channels were imaged sequentially, and no spectral bleed-through was detected. Each of the reporter
744 constructs were transferred to the wild-type, $\Delta metE$, $\Delta asnB$, and $\Delta tfoX$ strains, respectively. As, *V.*
745 *cholerae* has a plasmid defense system, it is likely to remove plasmids even under selective conditions.
746 To ensure detection of the reporter constructs was correctly evaluated the promoter region of the
747 constitutively expressed housekeeping gene, *gyrA* ([P_{gyrA}]GFP), functioned as an internal control for each
748 plasmid was used. As such only [P_{gyrA}]GFP positive cells were included from the analysis. The strains
749 were incubated on chitin for 24h, harvested, and imaged with a 1% agar patch.

750 ON cultures of wild-type, $\Delta metE$, $\Delta asnB$, $\Delta carB$, and $\Delta tfoX$ were prepared as normal for competence
751 induction on chitin. After 24 h on chitin the samples were harvested by vortex and images one by one.
752 Epifluorescence images were acquired using a DeltaVision Elite Imaging system with Olympus IX-71
753 base (GE healthcare), EDGE / sCMOS camera, solid state illumination source. Filters used were
754 GFP/FITC 475/28, and TRITC 542/27 along with an oil immersion objective Olympus U-Plan Fluorite
755 100X/1.30. Images were acquired using software “Softworx” built-in panel function, as well as the
756 hardware-based, and image-based autofocus unit. Exposure time for both GFP and TRITC channels
757 were set to 600 ms. No foci were detected and the fluorescent signal was evenly spread throughout the
758 cells. The GFP channel was used to generate masks, which were then overlaid on both GFP and RFP
759 channels to extract the corresponding intensity profiles. The ratio between channels was calculated on a
760 single- cell level and plotted as a boxplot, no background subtraction or normalization was performed.
761 The RFP reporter signal was divided by the corresponding GFP internal control signal for each
762 individual cell.

763

764

765

766

767

Table 5 Strain list			
Name used in study	Genotype	Relevant description	Source
A1552	A1552 Rif ^R	<i>Vibrio cholerae</i> O1 El Tor biotype Inaba isolate containing a mutation that makes it resistant to rifampicin (Rif). Used as wild-type in this study.	Melanie Blokesch https://bacdiv.dsmz.de/strain/139869
C6706	C6706 Sm ^R	<i>Vibrio cholerae</i> O1 El Tor biotype Inaba isolate containing a mutation that makes it resistant to streptomycin (Sm). Used as source to transfer the Sm resistance.	Matthew Waldor ⁸⁰
SM10 λ pir	thi thr leu tonA lacY supE recA::RP4-2-Tc::Mu, λ pir phage lysogen	<i>E. coli</i> carrying the transfer genes of the broad host range IncP type plasmid RP4 integrated in its chromosome. Can mobilizing mobRP4 plasmids. The presence of λ pir allows the replication of a defective R6K suicide plasmid and derivatives thereof.	Matthew Waldor
DH5 α λ pir	sup E44, Δ lacU169 (Φ lacZ Δ M15), recA1, endA1, hsdR17, thi-1, gyrA96, relA1, λ pir phage lysogen	<i>E. coli</i> standard strain for cloning and propagation of plasmids. The presence of λ pir allows the replication of a defective R6K suicide plasmid and derivatives thereof.	Matthew Waldor
Δ tfoX	A1552 tfoX::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the tfoX gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ tsaP	A1552 tsaP::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the tsaP gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ mshL	A1552 mshL::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the mshL gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ mshQ	A1552 mshQ::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the mshQ gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ pmg	A1552 pmg::tetR	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the pmg gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ pilW	A1552 pilW::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the pilW gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ asnB	A1552 asnB::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the asnB gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ metE	A1552 metE::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the metE gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ tfoS	A1552 tfoS::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the tfoS gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ carB	A1552 carB::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the carB gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study
Δ lapD	A1552 lapD::tetRA	Strain A1552 Rif ^R harboring a regulator and the tetracycline resistance cassette replacing the lapD gene of <i>V. cholerae</i> O1 El Tor strain A1552.	This study

769 Table 6 Plasmid list

Plasmids	Description	Source
pSC189	R6K suicide plasmid containing the mobRP4 origin of transfer, a ampicillin/carbenicillin resistance gene (<i>bla</i>) and a <i>Tn</i> -mariner transposon. The transposon transfers a kanamycin/neomycin (<i>nptII</i>) resistance gene.	²⁴
pCVD442	R6K suicide plasmid containing the mobRP4 origin of transfer, a ampicillin/carbenicillin resistance gene (<i>bla</i>) and a sucrose counter-selectable marker (<i>sacB</i>).	⁸¹
pCVD442-tfoX	Upstream and downstream fragments flanking <i>tfoX</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-tsaP	Upstream and downstream fragments flanking <i>tsaP</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-mshL	Upstream and downstream fragments flanking <i>mshL</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-mshQ	Upstream and downstream fragments flanking <i>mshQ</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-pmg	Upstream and downstream fragments flanking <i>pmg</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-pilW	Upstream and downstream fragments flanking <i>pilW</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-asnB	Upstream and downstream fragments flanking <i>asnB</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-metE	Upstream and downstream fragments flanking <i>metE</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-tfoS	Upstream and downstream fragments flanking <i>tfoS</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-carB	Upstream and downstream fragments flanking <i>carB</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pCVD442-lapD	Upstream and downstream fragments flanking <i>lapD</i> and cloned to flank a <i>tetRA</i> cassette on pCVD442 cut by XbaI using the ITAC method.	This study
pBR-GFP-dsRed-Kan	pBR322 derivative, <i>aph</i> gene from pBKdsGFP, Ampr ^R , Kan ^R .	Melanie Blokesch ⁴⁷
pBR-[PgyrA]GFP-[PcomEA]dsRed-Kan	Upstream region of <i>comEA</i> cloned into StuI site of pBR-GRP-dsRed-Kan.	Melanie Blokesch ⁴⁷
pBR-[PgyrA]GFP-[PcyaA]dsRed-Kan	Upstream region of <i>PgyrA</i> gene and <i>PcyaA</i> (~200bp each) cloned into NdeI/HindIII site for pBR-GFP-dsRed-Kan.	This study
pBR-[PgyrA]GFP-[PcytR]dsRed-Kan	Upstream region of <i>PgyrA</i> gene and <i>PcytR</i> (~200bp each) cloned into NdeI/HindIII site for pBR-GFP-dsRed-Kan.	This study
pBR-[PgyrA]GFP-[PhapR]dsRed-Kan	Upstream region of <i>PgyrA</i> gene and <i>PhapA</i> (~200bp each) cloned into NdeI/HindIII site for pBR-GFP-dsRed-Kan.	This study

770

771

772

773

Table 7 Primer list		
Primer	Sequence	Description
P0709	TACCACGACCA[3d_A]	Makes the adaptor, Chain terminator
P710	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGTCGTGGTAT	Makes the adaptor, Index Fork
P711	CGCCTTCTTGACGAGTT	Himmer3 out primer. Binds transposon
P712	GTGACTGGAGTTCAGACGTGTG	IndexRPrimer, binds to the adaptor after adaptor ligation, brings in an index sequence
P713	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTGACTTATCAGCCAACCTGT	P5 nr.1, amplifies from the transposon and includes a spacer sequence
P714	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTGACTTATCAGCCAACCTGT	P5 nr.2 amplifies from the transposon and includes a spacer sequence
P715	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTATGACTTATCAGCCAACCTGT	P5 nr.3 amplifies from the transposon and includes a spacer sequence
P716	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTTGTCGACTTATCAGCCAACCTGT	P5 nr.4 amplifies from the transposon and includes a spacer sequence
P717	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTTCGACGACTTATCAGCCAACCTGT	P5 nr.5 amplifies from the transposon and includes a spacer sequence
P718	AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTGCAGCGACGACTTATCAGCCAACCTGT	P5 nr.6 amplifies from the transposon and includes a spacer sequence
P719	CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATC	P7nr1 brings in the eP7 chip attachment site as well as a barcode for identification. Binds the index sequence
P0024	CTCTTATGAACGTGGCGAGC	Amplify donor DNA, RpsI including 2kb flank
P0025	ACCTGACTGACGAACGAAC	Amplify donor DNA, RpsI including 2kb flank
P256	TCAGCGATCGGCTCGTTG	Amplifies tetracycline cassette
P257	TCAATCGTACCCTTTCTCG	Amplifies tetracycline cassette
P0277	caacgagccgatcgtga	Test primer binds inside the Tet cassette, rev
P0278	cgagaagggtgacgattga	Test primer binds inside the Tet cassette, for
P0192	caacgagccgatcgtgattggtaggcaccaacttag	Upstream fragment <i>tfoX</i> flank 2kb UptfoxtetR
P0193	cgagaagggtgacgattgataagttaagcgttagccag	Upstream fragment <i>tfoX</i> flank 2kb tetDowntfoxF
P0194	CGGGAGAGCTCGATATCGCATGCGGTACCTaccggccaatagactcag	Downstream fragment <i>tfoX</i> flank 2kb UpnewtfoxF
P0195	GTGATGGTTAAAAAGGATCGATCCTCTAGatgaggccatgatccaactc	Downstream fragment <i>tfoX</i> flank 2kb DwNewTfoXR
P0279	gtgctcaggctttcataatc	<i>tfoX</i> , gDNA test primers, for
P0280	gcgcggtactgaaaccttat	<i>tfoX</i> , gDNA test primers, rev

P0262	CGGGAGAGCTCGATATCGCATGCGGTACCTcgctgatagttacgcagcac	Upstream fragment <i>tsaP</i> flank 2kb
P0263	CGCAGGGCAACGAGCCGATCGCTGAcacgaccttcttcttgg	Upstream fragment <i>tsaP</i> flank 2kb
P0264	CGAGAAAGGGTGACGATTGAaccgagcagccagatgtagt	Downstream fragment <i>tsaP</i> flank 2kb
P0265	GTGATGGGTTAAAAAGGATCGATCCTgttcgataaggtggcctttg	Downstream fragment <i>tsaP</i> flank 2kb
P0300	TTCGAGCAGGGCTTGATACT	<i>tsaP</i> , gDNA test primers, for
P0301	ATGCCTGCTGCACCTTTCTT	<i>tsaP</i> , gDNA test primers, rev
P0228	CGGGAGAGCTCGATATCGCATGCGGTACCTaaacgtaatcgatggcgaac	Upstream fragment <i>mshL</i> flank 2kb
P0229	CGCAGGGCAACGAGCCGATCGCTGAtttgacgactgttctctgz	Upstream fragment <i>mshL</i> flank 2kb
P0232	CGAGAAAGGGTGACGATTGAacttggcaaaaagagctgga	Downstream fragment <i>mshL</i> flank 2kb
P0233	GTGATGGGTTAAAAAGGATCGATCCTctcgcaaacgagacaatt	Downstream fragment <i>mshL</i> flank 2kb
P0292	TTGTTTGAACAGGTGCTTCG	<i>mshL</i> , gDNA test primers mshLgDNASeqF
P0293	GGGTTCACTTGCACCTGATT	<i>mshL</i> , gDNA test primers mshLgDNASeqR
P0253	CGGGAGAGCTCGATATCGCATGCGGTACCTccttcgagtaatgctcgta	Upstream fragment <i>mshQ</i> flank 2kb
P0254	CGCAGGGCAACGAGCCGATCGCTGAgagagcaatgcaaacatccc	Upstream fragment <i>mshQ</i> flank 2kb
P0407	CGGGAGAGCTCGATATCGCATGCGGTACCTGaatactctgcgctcttgc	Downstream fragment <i>mshQ</i> flank 2kb
P0408	atcaatcacaccaagagccccctggccttgaacaacactg	Downstream fragment <i>mshQ</i> flank 2kb
P0286	CCTTGCAGTAATGCTGCGTA	<i>mshQ</i> , gDNA test primers
P0287	ACATTCGGAACCCAGACTTC	<i>mshQ</i> , gDNA test primers
P0246	CGGGAGAGCTCGATATCGCATGCGGTACCTtaccgcaccttggactaac	Upstream fragment <i>pmg</i> flank 2kb
P0247	CGCAGGGCAACGAGCCGATCGCTGAtgatgttggctaacgcttga	Upstream fragment <i>pmg</i> flank 2kb
P0250	CGAGAAAGGGTGACGATTGAaatactcgggagcgtgttctt	Downstream fragment <i>pmg</i> flank 2kb
P0251	GTGATGGGTTAAAAAGGATCGATCCTgctgcgctgtactttatc	Downstream fragment <i>pmg</i> flank 2kb
P0298	CAGCCGGGACAATTTACCTA	<i>pmg</i> , gDNA test primers
P0299	ACTGTTTGGGTACCGTCAG	<i>pmg</i> , gDNA test primers
P0210	CGGGAGAGCTCGATATCGCATGCGGTACCTaaggcacctcagcacaact	Upstream fragment <i>piIW</i> (VC0860) flank 2kb
P0211	CGCAGGGCAACGAGCCGATCGCTGAcctaattgcaattgaaagcaga	Upstream fragment <i>piIW</i> (VC0860) flank 2kb
P0214	CGAGAAAGGGTGACGATTGAcgttggaaagaaggaagctg	Downstream fragment <i>piIW</i> (VC0860) flank 2kb
P0215	GTGATGGGTTAAAAAGGATCGATCCTccgataccggttgattcac	Downstream fragment <i>piIW</i> (VC0860) flank 2kb
P028	cggtgacgtgtttggtgata	<i>piIW</i> , gDNA test primers
P0282	aacggtgattctagcggaga	<i>piIW</i> , gDNA test primers

P0274	CGGGAGAGCTCGATATCGCATGCGGTACCTctttctgcccatttttagc	Upstream fragment <i>asnB</i> flank 2kb
P0275	CGCAGGGCAACGAGCCGATCGCTGAaggctgtgcgacctattta	Upstream fragment <i>asnB</i> flank 2kb
P0276	CGAGAAAGGGTGACGATTGAtccaaccaccaagaaggtt	Downstream fragment <i>asnB</i> flank 2kb
P0277	GTGATGGGTAAAAAGGATCGATCCTcgtgtgctagggtgaggttt	Downstream fragment <i>asnB</i> flank 2kb
P0311	GCTTTTGCTATGATCAAACGTG	<i>asnB</i> , gDNA test primers
P0312	CAGTGAAGTTGGCATTGGGA	<i>asnB</i> , gDNA test primers
P0266	CGGGAGAGCTCGATATCGCATGCGGTACCTgtatctgccatctcactgc	Upstream fragment <i>metE</i> flank 2kb
P0267	CGCAGGGCAACGAGCCGATCGCTGActtttccaatgccagttt	Upstream fragment <i>metE</i> flank 2kb
P0268	CGAGAAAGGGTGACGATTGAcagcgtttgtgggtcaac	Downstream fragment <i>metE</i> flank 2kb
P0269	GTGATGGGTAAAAAGGATCGATCCTcggtaaagagctcaccttgc	Downstream fragment <i>metE</i> flank 2kb
P0302	GTTTAGCGCTGGCTTACGTC	<i>metE</i> , gDNA test primers
P0303	TGTTTGGTGGCATGACCTTA	<i>metE</i> , gDNA test primers
P0411	CGGGAGAGCTCGATATCGCATGCGGTACCTgccatcaccatcagtttct	Upstream fragment <i>tfoS</i> flank 2kb
P0412	gttgtatggctgcttctcatattgagtgggcgaaaac	Upstream fragment <i>tfoS</i> flank 2kb
P0413	gattttcgccccactcaatatgagcaagcagccatacaac	Downstream fragment <i>tfoS</i> flank 2kb
P0414	ATATGTGATGGGTAAAAAGGATCGATCCTcgaagtctgaccagcatca	Downstream fragment <i>tfoS</i> flank 2kb
P0284	AAAGCCAACGTCACAAAACC	<i>tfoS</i> , gDNA test primers
P0285	GTATTCCGCGCTCTTCTG	<i>tfoS</i> , gDNA test primers
P0270	CGGGAGAGCTCGATATCGCATGCGGTACCTgagcataccgcatgttgc	Upstream fragment <i>carB</i> flank 2kb
P0271	CGCAGGGCAACGAGCCGATCGCTGAgccatttctgggtcagtc	Upstream fragment <i>carB</i> flank 2kb
P0272	CGAGAAAGGGTGACGATTGAaactacaccaccactgaacg	Downstream fragment <i>carB</i> flank 2kb
P0273	GTGATGGGTAAAAAGGATCGATCCTaaacgatagaaagcactgga	Downstream fragment <i>carB</i> flank 2kb
P309	TGAAGCTCATCATGCCATA	<i>carB</i> , gDNA test primers
P310	TCATCCTATCAACGCATAGCC	<i>carB</i> , gDNA test primers
P0234	CGGGAGAGCTCGATATCGCATGCGGTACCTccaaggccgacaaaataaga	Upstream fragment <i>lapD</i> flank 2kb
P0235	CGCAGGGCAACGAGCCGATCGCTGActaaagccaaaccaccgta	Upstream fragment <i>lapD</i> flank 2kb
P0238	CGAGAAAGGGTGACGATTGAcacacgggtgaaacacag	Downstream fragment <i>lapD</i> flank 2kb
P0239	GTGATGGGTAAAAAGGATCGATCCTtgggaagcggaagatattgg	Downstream fragment <i>lapD</i> flank 2kb
P0294	GCTTATGGGCCAGTCTATC	<i>lapD</i> , gDNA test primers
P0295	ACGTCGGTAAAGGCATCATC	<i>lapD</i> , gDNA test primers

P0483	tataggacatggcggaattcAGAGCCATTATCCCTCTATAGTTTG	promoter region of gyrA used to drive GFP expression
P0484	gatatctcccgggagaattcTTTAAGACGCAACCAAGGTCAC	promoter region of gyrA used to drive GFP expression
P0485	tataggacatggcggaattcGATAGACCCTATTTGGTTGTTG	promoter region of comEA used to drive GFP expression
P0486	gatatctcccgggagaattcCTGGCGGGTGTGATCAGTGT	promoter region of comEA used to drive GFP expression
P0487	gaattctcccgggagatattATGATGTTCTATCCATCTGTTT	promoter region of cyaA used to drive GFP expression
P0488	ttcctgttgaaggcctgatCACACAACTAAACACTTATCAAC	promoter region of cyaA used to drive GFP expression
P0489	gaattctcccgggagatattCACATCGCGTTTTATAACGC	promoter region of cytR used to drive RFP expression
P0490	ttcctgttgaaggcctgatGATTTTTACCTCTTTTCTCTATCGAG	promoter region of cytR used to drive RFP expression
P0491	gaattctcccgggagatattGCTTTCTATCGAGTATTCCTGC	promoter region of hapR used to drive RFP expression
P0492	ttcctgttgaaggcctgatTTCTCAATCCTAGAGATGTTGAATG	promoter region of hapR used to drive RFP expression
P0493	gaattctcccgggagatattTTAAGACGCAACCAAGGTCAC	promoter region of gyrA used to drive RFP expression
P0494	ttcctgttgaaggcctgatAGAGCCATTATCCCTCTATAGTTTG	promoter region of gyrA used to drive RFP expression
P0495	GTCGATTATAGGGACATGGC	pBRs1 test and sequencing primers
P0496	TCATAGCTGTTTCTGTGTGA	pBRs2 test and sequencing primers

775 Supplementary figures:

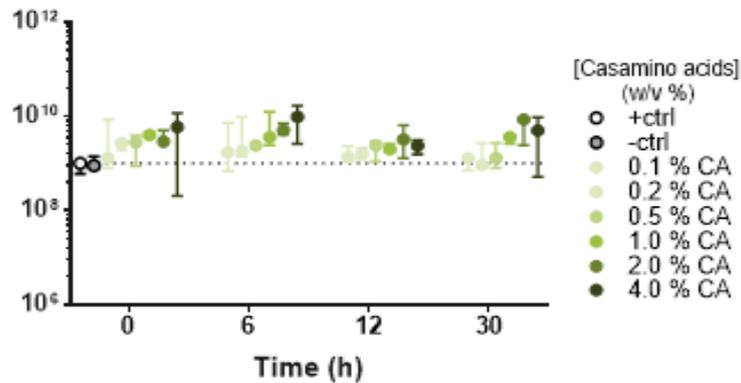


Figure S 2 Total bacterial count at the end of amino acid supplementation. At time point 0 h, a culture of wild-type *V. cholerae* in mid-exponential growth phase was exposed to chitin. At 24 h, DNA was added to the culture. Casamino acids (CA) were added to the culture at different concentrations and times during the natural transformation assay. CA concentrations ranged from 0 - 4% w/v and are depicted by increasing saturation of green. Timepoints assessed were 0 h, 6 h, 24 h, and 30 h. Positive (no CA; filled black circles) and negative (no DNA; open black circles) controls have very similar CFU on the non-selective plates. The dotted line extends from the CFU count of the positive control. All conditions were tested in triplicates.

776

777 Supplementary table:

Table S. 1 Sequencing statistics	TnNT			Ctrl		
	Miseq run #	1	2	Merged	1	2
Reads	14.8M	15.9M	30.7M	13.0M	15.8M	28.8M
Sites Hit	134'195	134'807	141'498	152'218	161'511	165'420
Percent TAs hit	69	69	72	78	83	85
Average Read Count	75	81	157	66	80	147

778

779 Table S. 2 Extended list of candidates

Table S. 2 list of candidates, tested candidates in bold				
Gene (A1552)	Gene (c6706)	Gene name		p-value from MWU
A1552VC_01851	VC2078	cdhR (TfoS)	HTH-type transcriptional regulator CdhR	4.18E-57
A1552VC_00384	VC0622	torS_1	Sensor protein TorS	1.16E-40
A1552VC_01659	VC1878	comEC	ComE operon protein 3	3.99E-34
A1552VC_00383	VC0620	appA_1	Oligopeptide-binding protein AppA	9.99E-30
A1552VC_00171	VC0414	mshQ	hypothetical protein	2.17E-28
A1552VC_00618	VC0861	-	hypothetical protein	1.02E-24
A1552VC_02405	VC2645	ppc	Phosphoenolpyruvate carboxylase	1.67E-24
A1552VC_01420	VC1628	ftsX	hypothetical protein	2.01E-24
A1552VC_02152	VC2388	carB	Carbamoyl-phosphate synthase large chain	2.07E-22

A1552VC_00741	VC0991	asnB	Asparagine synthetase B [glutamine-hydrolyzing]	5.36E-22
A1552VC_01494	VC1703	metE	5-methyltetrahydropteroyltriglutamate-	2.23E-21
A1552VC_00159	VC0402	xcpQ	Type II secretion system protein D	2.70E-21
A1552VC_A03758	VCA1081	lapD	Cyclic di-GMP phosphodiesterase PdeB	3.85E-21
A1552VC_00724	VC0972	chiP	Chitoporin	5.98E-21
A1552VC_02390	VC2629	pilQ	Type IV pilus biogenesis and competence protein	5.81E-19
A1552VC_01835	VC2062	cheA_2	Chemotaxis protein CheA	6.21E-19
A1552VC_00132	VC0384	cysJ_1	Sulfite reductase [NADPH] flavoprotein	6.48E-19
A1552VC_02245	VC2482	ilvI	Acetolactate synthase isozyme 3 large subunit	1.74E-18
A1552VC_00923	VC1174	trpE	Anthranilate synthase component 1	6.03E-18
A1552VC_02187	VC2423	epsE_2	Type II secretion system protein E	8.76E-18
A1552VC_00019	VC0028	ilvD	Dihydroxy-acid dehydratase	1.04E-16
A1552VC_00156	VC0399	mshI	hypothetical protein	3.49E-16
A1552VC_00432	VC0665	zraR_1	Transcriptional regulatory protein ZraR	7.65E-16
A1552VC_02545	VC0126	lysA	Diaminopimelate decarboxylase	8.63E-16
A1552VC_00037	VC0047	tsaP	hypothetical protein	1.32E-15
A1552VC_02759	VC0163	ilvC	Ketol-acid reductoisomerase (NADP(+))	2.53E-15
A1552VC_00133	VC0385	cysI	Sulfite reductase [NADPH] hemoprotein	3.33E-15
A1552VC_00374	VC0611	pgm_1	Phosphoglucomutase	6.18E-15
A1552VC_02188	VC2424	gspF	Putative type II secretion system protein F	6.27E-15
A1552VC_00468	VC0705	pheA	P-protein	6.39E-15
A1552VC_00162	VC0405	epsE_1	Type II secretion system protein E	6.88E-15
A1552VC_02251	VC2489	leuA	2-isopropylmalate synthase	1.04E-14
A1552VC_00163	VC0406	epsF_1	Type II secretion system protein F	1.96E-14
A1552VC_00169	VC0412	mshO	hypothetical protein	2.28E-14
A1552VC_02757	VC0165	bepE_2	Efflux pump membrane transporter BepE	2.32E-14
A1552VC_00038	VC0048	dprA	DNA processing protein DprA	6.24E-14
A1552VC_02243	VC2480	serA	D-3-phosphoglycerate dehydrogenase	1.38E-13
A1552VC_02186	VC2422	pilA	Fimbrial protein	2.21E-13
A1552VC_01790	VC2017	pabC	Aminodeoxychorismate lyase	2.89E-13
A1552VC_00623	VC0869	mltF	Membrane-bound lytic murein transglycosylase F	3.07E-13
A1552VC_00881	VC1134	hisC	Histidinol-phosphate aminotransferase	3.21E-13
A1552VC_00378	VC0615	-	hypothetical protein	3.48E-13
A1552VC_00382	VC0619	dppB_1	Dipeptide transport system permease protein	5.82E-13
A1552VC_02401	VC2641	argG	Argininosuccinate synthase	1.41E-12
A1552VC_02045	VC2272	proA	Gamma-glutamyl phosphate reductase	1.53E-12
A1552VC_02400	VC2640	argH	Argininosuccinate lyase	2.41E-12
A1552VC_02394	VC2633	pilM	hypothetical protein	2.88E-12

A1552VC_02043	VC2270	ribD	Riboflavin biosynthesis protein RibD	6.57E-12
A1552VC_02123	VC2361	thrC	Threonine synthase	1.10E-11
A1552VC_00880	VC1133	hisD	Histidinol dehydrogenase	1.59E-11
A1552VC_00381	VC0618	dppC_1	Dipeptide transport system permease protein	1.59E-11
A1552VC_00376	VC0613	exo I	Beta-hexosaminidase	1.63E-11
A1552VC_00798	VC1050	rssB	Regulator of RpoS	1.73E-11
A1552VC_01159	VC1423	/putrescine-binding periplasmic	Spermidine	2.09E-11
A1552VC_00930	VC1182	trxB	Thioredoxin reductase	2.66E-11
A1552VC_02253	VC2491	leuC	3-isopropylmalate dehydratase large subunit	2.94E-11
A1552VC_02125	VC2363	/homoserine	Bifunctional aspartokinase	4.26E-11
A1552VC_01403	VC1610	metA	Homoserine O-succinyltransferase	5.24E-11
A1552VC_02107	VC2344	serB	Phosphoserine phosphatase	5.26E-11
A1552VC_01390		pleD_3	Response regulator PleD	5.80E-11
A1552VC_00460	VC0696	tyrA	T-protein	6.12E-11
A1552VC_00617	VC0860	-	hypothetical protein	8.60E-11
A1552VC_00379	VC0616	oppF_1	Oligopeptide transport ATP-binding protein OppF	8.93E-11
A1552VC_00161	VC0404	mshN	hypothetical protein	1.35E-10
A1552VC_02319	VC2558	cysN	Sulfate adenylyltransferase subunit 1	1.53E-10
A1552VC_00900	VC1153	tfoX1_1	DNA transformation protein TfoX1	1.64E-10
A1552VC_A02833	VCA0039	-	hypothetical protein	1.75E-10
A1552VC_01806	VC2035	asd1	Aspartate-semialdehyde dehydrogenase 1	1.88E-10
A1552VC_02083	VC2314	argA	Amino-acid acetyltransferase	3.50E-10
A1552VC_A02944	VCA0164	cph2_6	Phytochrome-like protein cph2	3.97E-10
A1552VC_02403	VC2643	argC	N-acetyl-gamma-glutamyl-phosphate reductase	5.40E-10
A1552VC_00018	VC0027	ilvA	L-threonine dehydratase biosynthetic IlvA	7.18E-10
A1552VC_02274	VC2507	argF	Ornithine carbamoyltransferase	9.29E-10
A1552VC_00153	VC0396	csgD_1 qstR	CsgBAC operon transcriptional regulatory	1.15E-09
A1552VC_00380	VC0617	oppD_1	Oligopeptide transport ATP-binding protein OppD	1.18E-09
A1552VC_01676	VC1897	mcpP_4	Methyl-accepting chemotaxis protein McpP	1.42E-09
A1552VC_02392	VC2631	pilO	hypothetical protein	1.56E-09
A1552VC_00023	VC0032	comM	Competence protein ComM	1.59E-09
A1552VC_01035	VC1291	aspC	Aspartate aminotransferase	1.75E-09
A1552VC_00134	VC0386	cysH	Phosphoadenosine phosphosulfate reductase	1.76E-09
A1552VC_00895	VC1148	hexR	HTH-type transcriptional regulator HexR	1.84E-09
A1552VC_02035	VC2261	glnD	Bifunctional	2.19E-09
A1552VC_02437	VC2682	metB	Cystathionine gamma-synthase	2.95E-09
A1552VC_00980	VC1233	sbcB	Exodeoxyribonuclease I	3.08E-09

A1552VC_02252	VC2490	leuB	3-isopropylmalate dehydrogenase	3.27E-09
A1552VC_00615	VC0858	fimA	Fimbrial protein	3.33E-09
A1552VC_02404	VC2644	argE	Acetylornithine deacetylase	3.38E-09
A1552VC_00917	VC1170; overlaps another	trpB_2	Tryptophan synthase beta chain	4.06E-09
A1552VC_00166	VC0409	mshA	hypothetical protein	4.10E-09
A1552VC_02042	VC2269	ribE	Riboflavin synthase	4.65E-09
A1552VC_00168	VC0411	mshD	hypothetical protein	5.92E-09
A1552VC_01613	VC1833	cpoB	Cell division coordinator CpoB	6.05E-09
A1552VC_01654	VC1873; Stress response UPF0229 protein		hypothetical protein	6.47E-09
A1552VC_01404	VC1611	bepA_1	Beta-barrel assembly-enhancing protease	1.05E-08
A1552VC_00616	VC0859	-	hypothetical protein	2.04E-08
A1552VC_01701	VC1922	tig	Trigger factor	2.28E-08
A1552VC_00167	VC0410	mshC	hypothetical protein	3.44E-08
A1552VC_01462	VC1670	metC	Cystathionine beta-lyase MetC	3.78E-08
A1552VC_01750	VC1976	alaA	Glutamate-pyruvate aminotransferase AlaA	4.45E-08
A1552VC_02214	VC2450	relA	GTP pyrophosphokinase	4.90E-08
A1552VC_00933	VC1185	ydaM_1	putative diguanylate cyclase YdaM	6.05E-08
A1552VC_02046	VC2273	proB	Glutamate 5-kinase	6.66E-08
A1552VC_02320	VC2559	cysD	Sulfate adenylyltransferase subunit 2	7.27E-08
A1552VC_00020	VC0029	ilvE	Branched-chain-amino-acid aminotransferase	9.90E-08
A1552VC_02439	VC2684	metF	5,10-methylenetetrahydrofolate reductase	1.45E-07
A1552VC_01855	VC2083	sucD	Succinate--CoA ligase [ADP-forming] subunit	1.46E-07
A1552VC_01856	VC2084	sucC	Succinate--CoA ligase [ADP-forming] subunit	1.95E-07
A1552VC_00879	VC1132	hisG	ATP phosphoribosyltransferase	2.22E-07
A1552VC_02124	VC2362	thrB	Homoserine kinase	3.07E-07
A1552VC_02393	VC2632	pilN	hypothetical protein	3.61E-07
A1552VC_01419	VC1627	-	hypothetical protein	4.40E-07
A1552VC_A03701	VCA1019	gltR_2	HTH-type transcriptional regulator GltR	4.86E-07
A1552VC_00157	VC0400	mshJ	hypothetical protein	5.55E-07
A1552VC_00922	VC1173	trpG_1	Anthranilate synthase component 2	6.63E-07
A1552VC_01006	VC1262	ribA	GTP cyclohydrolase-2	9.10E-07
A1552VC_00302	VC0539	cysT_1	Sulfate transport system permease protein CysT	1.03E-06
A1552VC_00619	VC0863	-	hypothetical protein	1.22E-06
A1552VC_01102	VC1362	cysG_1	Siroheme synthase	1.49E-06
A1552VC_00921	VC1172	trpGD	Bifunctional protein TrpGD	1.73E-06
A1552VC_00170	VC0413	mshP	hypothetical protein	1.82E-06
A1552VC_01161	VC1425	ydcV	Inner membrane ABC transporter permease protein	1.91E-06

A1552VC_02181	VC2417	dsbC	Thiol:disulfide interchange protein DsbC	1.99E-06
A1552VC_00919	NA; overlaps another	trpC_2	Tryptophan biosynthesis protein TrpCF	2.93E-06
A1552VC_00303	VC0540	cysW_1	Sulfate transport system permease protein CysW	2.97E-06
A1552VC_02402	VC2642	argB	Acetylglutamate kinase	3.33E-06
A1552VC_02473	VC2718	comF	hypothetical protein	3.72E-06
A1552VC_01789	VC2016	mltG	Endolytic murein transglycosylase	3.80E-06
A1552VC_02321	VC2560	cysG_2	Siroheme synthase	4.52E-06
A1552VC_00882	VC1135	hisB	Histidine biosynthesis bifunctional protein	5.81E-06
A1552VC_02391	VC2630	pilP	hypothetical protein	8.22E-06
A1552VC_01162	VC1426	/putrescine transport system permease	Spermidine	1.15E-05
A1552VC_A03756	VCA1079	prsE	Type I secretion system membrane fusion protein	1.50E-05
A1552VC_00304	VC0541	/thiosulfate import ATP-binding protein	Sulfate	1.62E-05
A1552VC_01660	VC1879	-	hypothetical protein	1.93E-05
A1552VC_00516	VC0756	ndk	Nucleoside diphosphate kinase	3.35E-05
A1552VC_02547	VC0123	cyaA	Adenylate cyclase	3.52E-05
A1552VC_00916	NA; overlaps another	trpB_1	Tryptophan synthase beta chain	3.66E-05
A1552VC_00097	VC0346	miaA		5.56E-05
A1552VC_01558	VC1774	ybbH	putative HTH-type transcriptional regulator	5.99E-05
A1552VC_01694	VC1916	comEA	ComE operon protein 1	6.14E-05
A1552VC_00915	VC1169	trpA	Tryptophan synthase alpha chain	7.86E-05
A1552VC_00377	VC0614	gspK_1	Glucosamine kinase GspK	8.63E-05
A1552VC_02318	VC2556	cysC	Adenylyl-sulfate kinase	0.000123
A1552VC_00920	NA; overlaps another	trpC_3	Tryptophan biosynthesis protein TrpCF	0.000131
A1552VC_01653	VC1872	-	hypothetical protein	0.00017
A1552VC_02254	VC2492	leuD1	3-isopropylmalate dehydratase small subunit 1	0.000197
A1552VC_02232	VC2469	-	hypothetical protein	0.000203
A1552VC_00158	VC0401	mshK	hypothetical protein	0.000321
A1552VC_02244	VC2481	ilvH	Acetolactate synthase isozyme 3 small subunit	0.000583
A1552VC_00918	VC1171; overlaps another	trpC_1	Tryptophan biosynthesis protein TrpCF	0.000646
A1552VC_02756	VC0166	mdtA_2	Multidrug resistance protein MdtA	0.000989
A1552VC_01421	VC1629	-	putative ABC transporter ATP-binding protein	0.001148
A1552VC_A03702	VCA1020	-	hypothetical protein	0.001171
A1552VC_02233	VC2470	sdhE	FAD assembly factor SdhE	0.00971
A1552VC_00160	VC0403	mshM	hypothetical protein	0.13502
A1552VC_01836	VC2063	-	hypothetical protein	0.970461

781 Table S. 3 list of annotated output from analysis pipeline.

Con-ARTIST numbers defined				
			1 Domain is conditionally essential	
			2 Entirely conditionally essential	
			3 Domain conditionally enriched	
			4 Entirely conditionally enriched	
			5 No difference between conditions	
MWU statistics defined				
MWU p-values		Average p-value for each locus across all MWU simulations		
c6706	gene	A1552	Con-ARTIST	MWU p-values
NC002505_VC0027	ilvA	A1552VC_00018	1	7.18E-10
NC002505_VC0028	ilvD	A1552VC_00019	2	1.04E-16
NC002505_VC0029	ilvE	A1552VC_00020	2	9.90E-08
NC002505_VC0030	ilvM	A1552VC_00021	2	0.587632
NC002505_VC0032	comM	A1552VC_00023	1	1.59E-09
NC002505_VC0034	dsbA	A1552VC_00025	1	4.58E-06
NC002505_VC0047	tsaP	A1552VC_00037	2	1.32E-15
NC002505_VC0048	dprA	A1552VC_00038	2	6.24E-14
NC002505_VC0051	purK	A1552VC_00041	1	0.007791
NC002505_VC0052	purE	A1552VC_00042	2	9.19E-06
NC002505_VC0056	aroE	A1552VC_00045	2	2.47E-10
NC002505_VC0057		A1552VC_00046	1	0.340881
NC002505_VC0346	miaA	A1552VC_00097	1	5.56E-05
NC002505_VC0384	cysJ_1	A1552VC_00132	2	6.48E-19
NC002505_VC0385	cysI	A1552VC_00133	2	3.33E-15
NC002505_VC0386	cysH	A1552VC_00134	2	1.76E-09
NC002505_VC0396	csgD_1	A1552VC_00153	2	1.15E-09
NC002505_VC0399	mshI	A1552VC_00156	2	3.49E-16
NC002505_VC0400	mshJ	A1552VC_00157	2	5.55E-07
NC002505_VC0401	mshK	A1552VC_00158	2	0.000321
NC002505_VC0402	xcpQ	A1552VC_00159	2	2.70E-21
NC002505_VC0403	mshM	A1552VC_00160	2	0.13502
NC002505_VC0404	mshN	A1552VC_00161	2	1.35E-10
NC002505_VC0405	epsE_1	A1552VC_00162	2	6.88E-15
NC002505_VC0406	epsF_1	A1552VC_00163	2	1.96E-14
NC002505_VC0407	mshF	A1552VC_00164	2	0.584881
NC002505_VC0408	pilE	A1552VC_00165	2	0.094115
NC002505_VC0409	mshA	A1552VC_00166	2	4.10E-09
NC002505_VC0410	mshC	A1552VC_00167	2	3.44E-08
NC002505_VC0411	mshD	A1552VC_00168	2	5.92E-09
NC002505_VC0412	mshO	A1552VC_00169	2	2.28E-14
NC002505_VC0413	mshP	A1552VC_00170	2	1.82E-06
NC002505_VC0414	mshQ	A1552VC_00171	2	2.17E-28
NC002505_VC0462	pilT_1	A1552VC_00226	2	4.84E-08

NC002505_VC0538	cysP	A1552VC_00301	2	0.008327
NC002505_VC0539	cysT_1	A1552VC_00302	2	1.03E-06
NC002505_VC0540	cysW_1	A1552VC_00303	2	2.97E-06
NC002505_VC0541	/thiosulfate import ATP-binding protein	A1552VC_00304	2	1.62E-05
NC002505_VC0574	petB	A1552VC_00341	2	0.019233
NC002505_VC0575	/c1	A1552VC_00342	2	0.018749
NC002505_VC0596	dksA	A1552VC_00361	2	0.000368
NC002505_VC0604	acnB	A1552VC_00367	1	2.41E-14
NC002505_VC0611	pgm_1	A1552VC_00374	2	6.18E-15
NC002505_VC0613	exo I	A1552VC_00376	1	1.63E-11
NC002505_VC0614	gspK_1	A1552VC_00377	2	8.63E-05
NC002505_VC0615		A1552VC_00378	2	3.48E-13
NC002505_VC0616	oppF_1	A1552VC_00379	2	8.93E-11
NC002505_VC0617	oppD_1	A1552VC_00380	2	1.18E-09
NC002505_VC0618	dppC_1	A1552VC_00381	2	1.59E-11
NC002505_VC0619	dppB_1	A1552VC_00382	2	5.82E-13
NC002505_VC0620	appA_1	A1552VC_00383	2	9.99E-30
NC002505_VC0622	torS_1	A1552VC_00384	2	1.16E-40
NC002505_VC0623	rsmC	A1552VC_00385	1	0.299418
NC002505_VC0665	zraR_1	A1552VC_00432	1	7.65E-16
NC002505_VC0696	tyrA	A1552VC_00460	1	6.12E-11
NC002505_VC0705	pheA	A1552VC_00468	2	6.39E-15
NC002505_VC0756	ndk	A1552VC_00516	2	3.35E-05
NC002505_VC0829	tcpB	A1552VC_00584	1	0.072654
NC002505_VC0831	tcpC	A1552VC_00586	1	0.02039
NC002505_VC0832	tcpR	A1552VC_00587	1	0.269931
NC002505_VC0835	tcpT	A1552VC_00590	1	0.002058
NC002505_VC0837	tcpF	A1552VC_00592	2	0.005772
NC002505_VC0840	pctA_1	A1552VC_00595	1	0.001339
NC002505_VC0845	acfD	A1552VC_00601	1	0.001243
NC002505_VC0852	bamE	A1552VC_00609	2	0.005835
NC002505_VC0858	fimA	A1552VC_00615	2	3.33E-09
NC002505_VC0859		A1552VC_00616	2	2.04E-08
NC002505_VC0860		A1552VC_00617	2	8.60E-11
NC002505_VC0861		A1552VC_00618	2	1.02E-24
NC002505_VC0863		A1552VC_00619	2	1.22E-06
NC002505_VC0869	mltF	A1552VC_00623	1	3.07E-13
NC002505_VC0870	purL	A1552VC_00624	1	5.35E-08
NC002505_VC0893	pomA	A1552VC_00645	2	2.01E-08
NC002505_VC0894	motB	A1552VC_00646	1	1.81E-05
	treA	A1552VC_00664	1	5.38E-07
NC002505_VC0966	ptsI	A1552VC_00718	2	3.69E-11
NC002505_VC0968	ptsH	A1552VC_00719	2	0.000861
NC002505_VC0973	chiP	A1552VC_00724	2	5.98E-21
NC002505_VC0992	asnB	A1552VC_00741	2	5.36E-22
NC002505_VC1051	rssB	A1552VC_00798	2	1.73E-11

NC002505_VC1099	ackA_1	A1552VC_00844	2	0.003887
NC002505_VC1133	hisG	A1552VC_00879	2	2.22E-07
NC002505_VC1134	hisD	A1552VC_00880	2	1.59E-11
NC002505_VC1135	hisC	A1552VC_00881	2	3.21E-13
NC002505_VC1136	hisB	A1552VC_00882	2	5.81E-06
NC002505_VC1137	hisH	A1552VC_00883	2	1.72E-05
NC002505_VC1138	hisA	A1552VC_00884	2	0.000133
NC002505_VC1139	hisF	A1552VC_00885	2	1.27E-05
NC002505_VC1140	hisI	A1552VC_00886	1	8.19E-05
NC002505_VC1142	icd2	A1552VC_00888	1	5.15E-14
NC002505_VC1149	hexR	A1552VC_00895	2	1.84E-09
NC002505_VC1153		A1552VC_00899	1	3.05E-08
NC002505_VC1154	tfoX1_1	A1552VC_00900	2	1.64E-10
NA; overlaps another	trpA	A1552VC_00915	1	7.86E-05
NC002505_VC1170; overlaps another	trpB_1	A1552VC_00916	2	3.66E-05
NC002505_VC1171; overlaps another	trpB_2	A1552VC_00917	2	4.06E-09
NA; overlaps another	trpC_1	A1552VC_00918	2	0.000646
NA; overlaps another	trpC_2	A1552VC_00919	2	2.93E-06
NC002505_VC1172	trpC_3	A1552VC_00920	2	0.000131
NC002505_VC1173	trpGD	A1552VC_00921	2	1.73E-06
NC002505_VC1174	trpG_1	A1552VC_00922	2	6.63E-07
NC002505_VC1177	trpE	A1552VC_00923	2	6.03E-18
NC002505_VC1183	trxB	A1552VC_00930	1	2.66E-11
NC002505_VC1186	ydaM_1	A1552VC_00933	1	6.05E-08
NC002505_VC1199		A1552VC_00945	2	1.46E-20
NC002505_VC1234	sbcB	A1552VC_00980	1	3.08E-09
NC002505_VC1263	ribA	A1552VC_01006	2	9.10E-07
NC002505_VC1293	aspC	A1552VC_01035	2	1.75E-09
NC002505_VC1363	cysG_1	A1552VC_01102	1	1.49E-06
NC002505_VC1408	yfiR	A1552VC_01142	2	5.43E-05
NC002505_VC1421	/homoserine lactone efflux protein	A1552VC_01154	1	0.001395
NC002505_VC1424	/putrescine-binding periplasmic	A1552VC_01159	2	2.09E-11
NC002505_VC1426	ydcV	A1552VC_01161	1	1.91E-06
NC002505_VC1427	/putrescine transport system permease	A1552VC_01162	2	1.15E-05
NC002505_VC1428	/putrescine import ATP-binding protein	A1552VC_01163	2	0.001459
NC002505_VC1435		A1552VC_01170	2	0.046456
NC002505_VC1436		A1552VC_01171	2	0.429994
NC002505_VC1437	copA_1	A1552VC_01172	2	0.000278
NC002505_VC1438		A1552VC_01173	2	0.20913
NC002505_VC1439	ccoP2	A1552VC_01174	2	0.61108
NC002505_VC1440		A1552VC_01175	2	0.176755
NC002505_VC1441		A1552VC_01176	2	0.071642
NC002505_VC1442	ccoN1	A1552VC_01177	2	0.043858
NC002505_VC1496	prc	A1552VC_01290	2	1.93E-06
NC002505_VC1539		A1552VC_01330	2	0.000348

NC002505_VC1599	pleD_3	A1552VC_01390	2	5.80E-11
NC002505_VC1611	metA	A1552VC_01403	2	5.24E-11
NC002505_VC1612	bepA_1	A1552VC_01404	1	1.05E-08
NC002505_VC1617	cynR_2	A1552VC_01409	1	0.1235
NC002505_VC1618	ynfM	A1552VC_01410	1	0.001684
NC002505_VC1628		A1552VC_01419	1	4.40E-07
NC002505_VC1629	ftsX	A1552VC_01420	2	2.01E-24
NC002505_VC1630		A1552VC_01421	1	0.001148
NC002505_VC1671	metC	A1552VC_01462	2	3.78E-08
NC002505_VC1704	metE	A1552VC_01494	2	2.23E-21
NC002505_VC1706	metR	A1552VC_01495	2	1.36E-06
NC002505_VC1721	purR	A1552VC_01510	2	0.000449
NC002505_VC1775	ybbH	A1552VC_01558	2	5.99E-05
NA		A1552VC_01584	2	0.032148
NC002505_VC1834	cpoB	A1552VC_01613	2	6.05E-09
NC002505_VC1847	ruvC	A1552VC_01626	2	0.111372
NC002505_VC1871		A1552VC_01651	1	0.511792
NC002505_VC1872		A1552VC_01652	2	3.98E-08
NC002505_VC1873		A1552VC_01653	2	0.00017
NC002505_VC1874		A1552VC_01654	1	6.47E-09
NC002505_VC1879	comEC	A1552VC_01659	2	3.99E-34
NC002505_VC1880		A1552VC_01660	2	1.93E-05
NC002505_VC1898	mcpP_4	A1552VC_01676	1	1.42E-09
NC002505_VC1902	dsbB	A1552VC_01680	2	1.01E-07
NC002505_VC1911	pyrF	A1552VC_01688	1	0.000327
NC002505_VC1917	comEA	A1552VC_01694	2	6.14E-05
NC002505_VC1923	tig	A1552VC_01701	2	2.28E-08
NC002505_VC1977	alaA	A1552VC_01750	2	4.45E-08
NC002505_VC2017	mltG	A1552VC_01789	2	3.80E-06
NC002505_VC2018	pabC	A1552VC_01790	2	2.89E-13
NC002505_VC2033	adhE_1	A1552VC_01804	1	1.39E-06
NC002505_VC2036	asd1	A1552VC_01806	2	1.88E-10
NC002505_VC2048	mlaA	A1552VC_01820	2	6.29E-06
NC002505_VC2049	ccmH_1	A1552VC_01821	2	0.061099
NC002505_VC2050	ccmH_2	A1552VC_01822	2	0.009943
NC002505_VC2051	dsbE	A1552VC_01823	2	0.055212
NC002505_VC2052	ccmF	A1552VC_01824	2	0.00177
NC002505_VC2053	ccmE	A1552VC_01825	2	0.114495
NC002505_VC2054		A1552VC_01826	2	0.427757
NC002505_VC2055	ccmC	A1552VC_01827	2	0.089981
NC002505_VC2056	ccmB	A1552VC_01828	1	0.061912
NC002505_VC2063	cheA_2	A1552VC_01835	1	6.21E-19
NC002505_VC2064		A1552VC_01836	2	0.970461
NC002505_VC2065	cheY_4	A1552VC_01837	2	1.29E-05
NC002505_VC2080	cdhR	A1552VC_01851	2	4.18E-57
NC002505_VC2084	sucD	A1552VC_01855	2	1.46E-07

NC002505_VC2085	sucC	A1552VC_01856	2	1.95E-07
NC002505_VC2092	gltA	A1552VC_01863	2	1.46E-14
NC002505_VC2154		A1552VC_01919	2	0.085653
NC002505_VC2156	bamC	A1552VC_01920	2	1.26E-06
NC002505_VC2164	bepA_3	A1552VC_01927	2	1.35E-07
NC002505_VC2198	flgD	A1552VC_01972	2	5.42E-05
NC002505_VC2262	glnD	A1552VC_02035	2	2.19E-09
NC002505_VC2270	ribE	A1552VC_02042	1	4.65E-09
NC002505_VC2271	ribD	A1552VC_02043	2	6.57E-12
NC002505_VC2273	proA	A1552VC_02045	2	1.53E-12
NC002505_VC2274	proB	A1552VC_02046	2	6.66E-08
NC002505_VC2316	argA	A1552VC_02083	1	3.50E-10
NC002505_VC2345	serB	A1552VC_02107	2	5.26E-11
NC002505_VC2362	thrC	A1552VC_02123	2	1.10E-11
NC002505_VC2363	thrB	A1552VC_02124	2	3.07E-07
NC002505_VC2364	/homoserine	A1552VC_02125	2	4.26E-11
NC002505_VC2389	carB	A1552VC_02152	2	2.07E-22
NC002505_VC2390	carA	A1552VC_02153	1	3.72E-07
NC002505_VC2418	dsbC	A1552VC_02181	2	1.99E-06
NC002505_VC2419	xerD_1	A1552VC_02182	1	0.015063
NC002505_VC2422	nadC	A1552VC_02185	1	0.158101
NC002505_VC2423	pilA	A1552VC_02186	2	2.21E-13
NC002505_VC2424	epsE_2	A1552VC_02187	2	8.76E-18
NC002505_VC2425	gspF	A1552VC_02188	2	6.27E-15
NC002505_VC2441		A1552VC_02205	1	0.051528
NC002505_VC2451	relA	A1552VC_02214	2	4.90E-08
NC002505_VC2453	barA_2	A1552VC_02216	1	3.30E-13
NC002505_VC2470		A1552VC_02232	1	0.000203
NC002505_VC2471	sdhE	A1552VC_02233	2	0.00971
NC002505_VC2481	serA	A1552VC_02243	2	1.38E-13
NC002505_VC2482	ilvH	A1552VC_02244	2	0.000583
NC002505_VC2483	ilvI	A1552VC_02245	2	1.74E-18
NC002505_VC2490	leuA	A1552VC_02251	2	1.04E-14
NC002505_VC2491	leuB	A1552VC_02252	2	3.27E-09
NC002505_VC2492	leuC	A1552VC_02253	2	2.94E-11
NC002505_VC2493	leuD1	A1552VC_02254	2	0.000197
NC002505_VC2508	argF	A1552VC_02274	2	9.29E-10
NC002505_VC2510	pyrB	A1552VC_02275	2	3.61E-05
NC002505_VC2515	ibaG	A1552VC_02280	1	0.567681
NC002505_VC2516	yrbB	A1552VC_02281	2	0.436607
NC002505_VC2517	mIaC	A1552VC_02282	2	0.002056
NC002505_VC2518	mIaD	A1552VC_02283	2	0.061351
NC002505_VC2519	mIaE	A1552VC_02284	2	0.003503
NC002505_VC2520	mIaF	A1552VC_02285	1	0.057171
NC002505_VC2529	rpoN	A1552VC_02292	2	2.33E-06
NC002505_VC2558	cysC	A1552VC_02318	1	0.000123

NC002505_VC2559	cysN	A1552VC_02319	2	1.53E-10
NC002505_VC2560	cysD	A1552VC_02320	2	7.27E-08
NC002505_VC2561	cysG_2	A1552VC_02321	2	4.52E-06
NC002505_VC2618	/acetylornithine	A1552VC_02378	2	1.34E-06
NC002505_VC2630	pilQ	A1552VC_02390	2	5.81E-19
NC002505_VC2631	pilP	A1552VC_02391	2	8.22E-06
NC002505_VC2632	pilO	A1552VC_02392	2	1.56E-09
NC002505_VC2633	pilN	A1552VC_02393	2	3.61E-07
NC002505_VC2634	pilM	A1552VC_02394	2	2.88E-12
NC002505_VC2641	argH	A1552VC_02400	2	2.41E-12
NC002505_VC2642	argG	A1552VC_02401	2	1.41E-12
NC002505_VC2643	argB	A1552VC_02402	2	3.33E-06
NC002505_VC2644	argC	A1552VC_02403	2	5.40E-10
NC002505_VC2645	argE	A1552VC_02404	2	3.38E-09
NC002505_VC2646	ppc	A1552VC_02405	2	1.67E-24
NC002505_VC2683	metB	A1552VC_02437	2	2.95E-09
NC002505_VC2685	metF	A1552VC_02439	2	1.45E-07
NC002505_VC2686	zapB	A1552VC_02440	1	0.968132
NC002505_VC2719	comF	A1552VC_02473	2	3.72E-06
NC002505_VC0125	lysA	A1552VC_02545	2	8.63E-16
NC002505_VC0123	cyaY	A1552VC_02546	2	0.829217
NC002505_VC0122	cyaA	A1552VC_02547	1	3.52E-05
NC002505_VC0217	radC	A1552VC_02696	2	2.80E-06
NC002505_VC0211	pyrE	A1552VC_02702	2	2.59E-05
NC002505_VC0186	gor	A1552VC_02725	2	2.24E-05
NC002505_VC0165	mdtA_2	A1552VC_02756	2	0.000989
NC002505_VC0164	bepE_2	A1552VC_02757	2	2.32E-14
NC002505_VC0163		A1552VC_02758	2	0.68648
NC002505_VC0162	ilvC	A1552VC_02759	2	2.53E-15
NC002506_VCA0013	malP	A1552VC_A02807	1	0.000874
NC002506_VCA0015-VCA0016	glgB	A1552VC_A02809	1	3.52E-15
NC002506_VCA0040		A1552VC_A02833	2	1.75E-10
NC002506_VCA0078		A1552VC_A02865	1	0.805847
NC002506_VCA0079	mepM_4	A1552VC_A02866	2	2.20E-07
NC002506_VCA0165	cph2_6	A1552VC_A02944	2	3.97E-10
NC002506_VCA0166	cspV_1	A1552VC_A02946	2	0.013533
NC002506_VCA0454	sbp_1	A1552VC_A03177	1	0.060177
NC002506_VCA0802	ltaS2	A1552VC_A03501	1	0.000768
NC002506_VCA0862		A1552VC_A03554	2	0.001133
NC002506_VCA0863	volA	A1552VC_A03555	1	0.002356
NC002506_VCA0925	pyrC	A1552VC_A03613	1	0.000348
NC002506_VCA1020	gltR_2	A1552VC_A03701	2	4.86E-07
NC002506_VCA1021		A1552VC_A03702	2	0.001171
NC002506_VCA1025	nagB	A1552VC_A03705	2	1.36E-06
NC002506_VCA1080	prsE	A1552VC_A03756	1	1.50E-05
NC002506_VCA1082	pdeB_2	A1552VC_A03758	2	3.85E-21

782 References:

- 783 1 Blokesch, M. Natural competence for transformation. *Current Biology* **26**, R1126-R1130
784 (2016). [https://doi.org:https://doi.org/10.1016/j.cub.2016.08.058](https://doi.org/10.1016/j.cub.2016.08.058)
- 785 2 Sinha-Ray, S., Alam, M. T., Bag, S., Morris, J. G., Jr. & Ali, A. Conversion of a recA-
786 Mediated Non-toxigenic *Vibrio cholerae* O1 Strain to a Toxigenic Strain Using Chitin-
787 Induced Transformation. *Front Microbiol* **10**, 2562 (2019).
788 [https://doi.org:10.3389/fmicb.2019.02562](https://doi.org/10.3389/fmicb.2019.02562)
- 789 3 Joelsson, A., Liu, Z. & Zhu, J. Genetic and phenotypic diversity of quorum-sensing systems in
790 clinical and environmental isolates of *Vibrio cholerae*. *Infect Immun* **74**, 1141-1147 (2006).
791 [https://doi.org:10.1128/iai.74.2.1141-1147.2006](https://doi.org/10.1128/iai.74.2.1141-1147.2006)
- 792 4 Hu, D. *et al.* Origins of the current seventh cholera pandemic.
793 5 Ali, M. *et al.* The global burden of cholera. *Bull World Health Organ* **90**, 209-218a (2012).
794 [https://doi.org:10.2471/blt.11.093427](https://doi.org/10.2471/blt.11.093427)
- 795 6 Ali, M., Nelson, A. R., Lopez, A. L. & Sack, D. A. Updated global burden of cholera in
796 endemic countries. *PLoS Negl Trop Dis* **9**, e0003832 (2015).
797 [https://doi.org:10.1371/journal.pntd.0003832](https://doi.org/10.1371/journal.pntd.0003832)
- 798 7 Ganesan, D., Gupta, S. S. & Legros, D. Cholera surveillance and estimation of burden of
799 cholera. *Vaccine* **38 Suppl 1**, A13-a17 (2020). [https://doi.org:10.1016/j.vaccine.2019.07.036](https://doi.org/10.1016/j.vaccine.2019.07.036)
- 800 8 Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C.-Y. & Schoolnik, G. K. Chitin Induces
801 Natural Competence in *Vibrio cholerae*. *Science* **310**, 1824-1827 (2005).
802 [https://doi.org:10.1126/science.1120096](https://doi.org/10.1126/science.1120096)
- 803 9 Meibom, K. L., Blokesch, M., Dolganov, N. A., Wu, C. Y. & Schoolnik, G. K. Chitin induces
804 natural competence in *Vibrio cholerae*. *Science* **310**, 1824-1827 (2005).
805 [https://doi.org:10.1126/science.1120096](https://doi.org/10.1126/science.1120096)
- 806 10 Dalia, A. B., Lazinski, D. W. & Camilli, A. Identification of a membrane-bound
807 transcriptional regulator that links chitin and natural competence in *Vibrio cholerae*. *mBio* **5**,
808 e01028-01013 (2014). [https://doi.org:10.1128/mBio.01028-13](https://doi.org/10.1128/mBio.01028-13)
- 809 11 Blokesch, M. Chitin colonization, chitin degradation and chitin-induced natural competence of
810 *Vibrio cholerae* are subject to catabolite repression. *Environ Microbiol* **14**, 1898-1912 (2012).
811 [https://doi.org:10.1111/j.1462-2920.2011.02689.x](https://doi.org/10.1111/j.1462-2920.2011.02689.x)
- 812 12 Metzger, L. C. & Blokesch, M. Regulation of competence-mediated horizontal gene transfer
813 in the natural habitat of *Vibrio cholerae*.
- 814 13 Sun, Y., Bernardy, E. E., Hammer, B. K. & Miyashiro, T. Competence and natural
815 transformation in vibrios. *Mol Microbiol* **89**, 583-595 (2013).
816 [https://doi.org:10.1111/mmi.12307](https://doi.org/10.1111/mmi.12307)
- 817 14 Watve, S. S., Thomas, J. & Hammer, B. K. CytR Is a Global Positive Regulator of
818 Competence, Type VI Secretion, and Chitinases in *Vibrio cholerae*. *PLoS One* **10**, e0138834
819 (2015). [https://doi.org:10.1371/journal.pone.0138834](https://doi.org/10.1371/journal.pone.0138834)
- 820 15 Lichev, A., Angelov, A., Cucurull, I. A.-O. & Liebl, W. A.-O. Amino acids as nutritional
821 factors and (p)ppGpp as an alarmone of the stringent response regulate natural transformation
822 in *Micrococcus luteus*.
- 823 16 Wilson Ga Fau - Bott, K. F. & Bott, K. F. Nutritional factors influencing the development of
824 competence in the *Bacillus subtilis* transformation system.
- 825 17 Blokesch, M. & Schoolnik, G. K. The extracellular nuclease Dns and its role in natural
826 transformation of *Vibrio cholerae*. *J Bacteriol* **190**, 7232-7240 (2008).
827 [https://doi.org:10.1128/jb.00959-08](https://doi.org/10.1128/jb.00959-08)
- 828 18 Seitz, P. *et al.* ComEA is essential for the transfer of external DNA into the periplasm in
829 naturally transformable *Vibrio cholerae* cells.
- 830 19 Silander, O. K. *et al.* A genome-wide analysis of promoter-mediated phenotypic noise in
831 *Escherichia coli*. *PLoS Genet* **8**, e1002443 (2012).
832 [https://doi.org:10.1371/journal.pgen.1002443](https://doi.org/10.1371/journal.pgen.1002443)
- 833 20 Pál, C., Papp B Fau - Lercher, M. J. & Lercher, M. J. Adaptive evolution of bacterial
834 metabolic networks by horizontal gene transfer.
- 835 21 Kumar, A., Bag, S. & Das, B. Novel Genetic Tool to Study the Stability of Genomic Islands.

- 836 22 Kumar, A., Das, B. & Kumar, N. Vibrio Pathogenicity Island-1: The Master Determinant of
837 Cholera Pathogenesis.
- 838 23 Cain, A. K. *et al.* A decade of advances in transposon-insertion sequencing. *Nat Rev Genet* **21**,
839 526-540 (2020). <https://doi.org/10.1038/s41576-020-0244-x>
- 840 24 Chiang, S. L. & Rubin, E. J. Construction of a mariner-based transposon for epitope-tagging
841 and genomic targeting. *Gene* **296**, 179-185 (2002). [https://doi.org/10.1016/s0378-1119\(02\)00856-9](https://doi.org/10.1016/s0378-1119(02)00856-9)
- 842
843 25 Pritchard, J. R. *et al.* ARTIST: high-resolution genome-wide assessment of fitness using
844 transposon-insertion sequencing. *PLoS Genet* **10**, e1004782 (2014).
845 <https://doi.org/10.1371/journal.pgen.1004782>
- 846 26 Huerta-Cepas, J. *et al.* eggNOG 5.0: a hierarchical, functionally and phylogenetically
847 annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*
848 **47**, D309-d314 (2019). <https://doi.org/10.1093/nar/gky1085>
- 849 27 Adams, D. W., Stutzmann, S., Stoudmann, C. & Blokesch, M. DNA-uptake pili of *Vibrio*
850 *cholerae* are required for chitin colonization and capable of kin recognition via sequence-
851 specific self-interaction. *Nat Microbiol* **4**, 1545-1557 (2019). <https://doi.org/10.1038/s41564-019-0479-5>
- 852
853 28 Ellison, C. K. *et al.* Retraction of DNA-bound type IV competence pili initiates DNA uptake
854 during natural transformation in *Vibrio cholerae*. *Nat Microbiol* **3**, 773-780 (2018).
855 <https://doi.org/10.1038/s41564-018-0174-y>
- 856 29 Dalia, A. B., Lazinski, D. W. & Camilli, A. Identification of a Membrane-Bound
857 Transcriptional Regulator That Links Chitin and Natural Competence in *Vibrio cholerae*.
858 *mBio* **5**, 10.1128/mbio.01028-01013 (2014). <https://doi.org/doi:10.1128/mbio.01028-13>
- 859 30 Seitz, P. & Blokesch, M. DNA-uptake machinery of naturally competent *Vibrio*
860 *cholerae*. *Proceedings of the National Academy of Sciences* **110**, 17987-17992 (2013).
861 <https://doi.org/doi:10.1073/pnas.1315647110>
- 862 31 Beyhan, S., Bilecen, K., Salama, S. R., Casper-Lindley, C. & Yildiz, F. H. Regulation of
863 Rugosity and Biofilm Formation in *Vibrio cholerae*: Comparison of VpsT and VpsR
864 Regulons and Epistasis Analysis of *vpsT*, *vpsR*, and *hapR*. *Journal of*
865 *Bacteriology* **189**, 388-402 (2007). <https://doi.org/doi:10.1128/JB.00981-06>
- 866 32 Kitts, G. *et al.* A Conserved Regulatory Circuit Controls Large Adhesins in *Vibrio cholerae*.
867 *mBio* **10** (2019). <https://doi.org/10.1128/mBio.02822-19>
- 868 33 Kitts, G. *et al.* A Conserved Regulatory Circuit Controls Large Adhesins in *Vibrio cholerae*.
869 LID - 10.1128/mBio.02822-19 [doi] LID - e02822-19.
- 870 34 Fresquet, V., Thoden Jb Fau - Holden, H. M., Holden Hm Fau - Raushel, F. M. & Raushel, F.
871 M. Kinetic mechanism of asparagine synthetase from *Vibrio cholerae*.
- 872 35 Bogard, R. W., Davies, B. W. & Mekalanos, J. J. MetR-regulated *Vibrio cholerae* metabolism
873 is required for virulence. *mBio* **3** (2012). <https://doi.org/10.1128/mBio.00236-12>
- 874 36 Yamamoto, S. *et al.* Identification of a chitin-induced small RNA that regulates translation of
875 the *tfoX* gene, encoding a positive regulator of natural competence in *Vibrio cholerae*. *J*
876 *Bacteriol* **193**, 1953-1965 (2011). <https://doi.org/10.1128/jb.01340-10>
- 877 37 Jaskólska, M., Stutzmann, S., Stoudmann, C. & Blokesch, M. QstR-dependent regulation of
878 natural competence and type VI secretion in *Vibrio cholerae*.
- 879 38 Siewering, K. *et al.* Peptidoglycan-binding protein Tsap functions in surface assembly of type
880 IV pili. *Proc Natl Acad Sci U S A* **111**, E953-961 (2014).
881 <https://doi.org/10.1073/pnas.1322889111>
- 882 39 Lo Scudato, M. & Blokesch, M. A transcriptional regulator linking quorum sensing and
883 chitin induction to render *Vibrio cholerae* naturally transformable. *Nucleic Acids Research* **41**,
884 3644-3658 (2013). <https://doi.org/10.1093/nar/gkt041>
- 885 40 Jaskólska, M., Stutzmann, S., Stoudmann, C. & Blokesch, M. QstR-dependent regulation of
886 natural competence and type VI secretion in *Vibrio cholerae*. *Nucleic Acids Res* **46**, 10619-
887 10634 (2018). <https://doi.org/10.1093/nar/gky717>
- 888 41 Ball, A. S., Chaparian, R. R. & van Kessel, J. C. Quorum Sensing Gene Regulation by
889 LuxR/HapR Master Regulators in *Vibrios*. LID - 10.1128/JB.00105-17 [doi] LID - e00105-17.

890 42 Palmen, R., Vosman, B., Buijsman, P., Breek, C. K. & Hellingwerf, K. J. Physiological
891 characterization of natural transformation in *Acinetobacter calcoaceticus*. *J Gen Microbiol*
892 **139**, 295-305 (1993). <https://doi.org:10.1099/00221287-139-2-295>

893 43 Fresquet, V., Thoden, J. B., Holden, H. M. & Raushel, F. M. Kinetic mechanism of asparagine
894 synthetase from *Vibrio cholerae*. *Bioorg Chem* **32**, 63-75 (2004).
895 <https://doi.org:10.1016/j.bioorg.2003.10.002>

896 44 Wu, R. *et al.* Direct regulation of the natural competence regulator gene *tfoX* by cyclic AMP
897 (cAMP) and cAMP receptor protein (CRP) in *Vibrios*. *Scientific Reports* **5**, 14921 (2015).
898 <https://doi.org:10.1038/srep14921>

899 45 Suckow, G., Seitz, P. & Blokesch, M. Quorum Sensing Contributes to Natural Transformation
900 of *Vibrio cholerae* in a Species-Specific Manner. *Journal of Bacteriology* **193**, 4914-4924
901 (2011). <https://doi.org:doi:10.1128/jb.05396-11>

902 46 Tsou, A. M., Cai, T., Liu, Z., Zhu, J. & Kulkarni, R. V. Regulatory targets of quorum sensing
903 in *Vibrio cholerae*: evidence for two distinct HapR-binding motifs. *Nucleic Acids Res* **37**,
904 2747-2756 (2009). <https://doi.org:10.1093/nar/gkp121>

905 47 Lo Scudato, M. & Blokesch, M. The regulatory network of natural competence and
906 transformation of *Vibrio cholerae*. *PLoS Genet* **8**, e1002778 (2012).
907 <https://doi.org:10.1371/journal.pgen.1002778>

908 48 Bevis, B. J. & Glick, B. S. Rapidly maturing variants of the *Discosoma* red fluorescent protein
909 (DsRed). *Nat Biotechnol* **20**, 83-87 (2002). <https://doi.org:10.1038/nbt0102-83>

910 49 Cormack, B. P., Valdivia, R. H. & Falkow, S. FACS-optimized mutants of the green
911 fluorescent protein (GFP). *Gene* **173**, 33-38 (1996). [https://doi.org:10.1016/0378-
912 1119\(95\)00685-0](https://doi.org:10.1016/0378-1119(95)00685-0)

913 50 Lloyd, C. J., Mejia-Santana, A., Dalia, T. N., Dalia, A. B. & Klose, K. E. Natural
914 Transformation in a Classical-Biotype *Vibrio cholerae* Strain. *Appl Environ Microbiol* **87**
915 (2021). <https://doi.org:10.1128/aem.00060-21>

916 51 van Opijnen, T. & Camilli, A. Transposon insertion sequencing: a new tool for systems-level
917 analysis of microorganisms. *Nat Rev Microbiol* **11**, 435-442 (2013).
918 <https://doi.org:10.1038/nrmicro3033>

919 52 Dubnau, D. The regulation of genetic competence in *Bacillus subtilis*. *Molecular*
920 *Microbiology* **5**, 11-18 (1991). [https://doi.org:https://doi.org/10.1111/j.1365-
921 2958.1991.tb01820.x](https://doi.org:https://doi.org/10.1111/j.1365-2958.1991.tb01820.x)

922 53 Maamar, H., Raj, A. & Dubnau, D. Noise in Gene Expression Determines Cell Fate in
923 *Bacillus subtilis*. *Science* **317**, 526-529 (2007). <https://doi.org:10.1126/science.1140818>

924 54 Rojo, F. Carbon catabolite repression in *Pseudomonas* : optimizing metabolic versatility and
925 interactions with the environment. *FEMS Microbiol Rev* **34**, 658-684 (2010).
926 <https://doi.org:10.1111/j.1574-6976.2010.00218.x>

927 55 Giuliodori, A. M., Gualerzi, C. O., Soto, S., Vila, J. & Tavio, M. M. Review on Bacterial
928 Stress Topics. *Ann N Y Acad Sci* **1113**, 95-104 (2007).
929 <https://doi.org:10.1196/annals.1391.008>

930 56 Metzger, L. & Blokesch, M. Composition of the DNA-uptake complex of *Vibrio cholerae*.
931 *Mobile genetic elements* **4**, e28142 (2014). <https://doi.org:10.4161/mge.28142>

932 57 Kickstein, E., Harms, K. & Wackernagel, W. Deletions of *recBCD* or *recD* influence genetic
933 transformation differently and are lethal together with a *recJ* deletion in *Acinetobacter baylyi*.
934 *Microbiology (Reading)* **153**, 2259-2270 (2007). <https://doi.org:10.1099/mic.0.2007/005256-0>

935 58 Johnsen, P. J., Dubnau, D. & Levin, B. R. Episodic selection and the maintenance of
936 competence and natural transformation in *Bacillus subtilis*. *Genetics* **181**, 1521-1533 (2009).
937 <https://doi.org:10.1534/genetics.108.099523>

938 59 Utnes, A. L. *et al.* Growth phase-specific evolutionary benefits of natural transformation in
939 *Acinetobacter baylyi*. *Isme j* **9**, 2221-2231 (2015). <https://doi.org:10.1038/ismej.2015.35>

940 60 Li, X. & Roseman, S. The chitinolytic cascade in *Vibrios* is regulated by chitin
941 oligosaccharides and a two-component chitin catabolic sensor/kinase. *Proceedings of the*
942 *National Academy of Sciences* **101**, 627-631 (2004).
943 <https://doi.org:doi:10.1073/pnas.0307645100>

944 61 Meibom, K. L. *et al.* The *Vibrio cholerae* chitin utilization program. *Proceedings of the*
945 *National Academy of Sciences* **101**, 2524-2529 (2004).
946 <https://doi.org/doi:10.1073/pnas.0308707101>

947 62 Passalacqua, K. D., Charbonneau, M. E. & O'Riordan, M. X. D. Bacterial Metabolism Shapes
948 the Host-Pathogen Interface. *Microbiol Spectr* **4** (2016).
949 <https://doi.org:10.1128/microbiolspec.VMBF-0027-2015>

950 63 Moradigaravand, D. & Engelstädter, J. The impact of natural transformation on adaptation in
951 spatially structured bacterial populations. *BMC Evol Biol* **14**, 141 (2014).
952 <https://doi.org:10.1186/1471-2148-14-141>

953 64 Evans, T. D. & Zhang, F. Bacterial metabolic heterogeneity: origins and applications in
954 engineering and infectious disease. *Curr Opin Biotechnol* **64**, 183-189 (2020).
955 <https://doi.org:10.1016/j.copbio.2020.04.007>

956 65 Şimşek, E. & Kim, M. The emergence of metabolic heterogeneity and diverse growth
957 responses in isogenic bacterial cells. *The ISME Journal* **12**, 1199-1209 (2018).
958 <https://doi.org:10.1038/s41396-017-0036-2>

959 66 Soltani, M., Vargas-Garcia, C. A., Antunes, D. & Singh, A. Intercellular Variability in Protein
960 Levels from Stochastic Expression and Noisy Cell Cycle Processes. *PLoS Comput Biol* **12**,
961 e1004972 (2016). <https://doi.org:10.1371/journal.pcbi.1004972>

962 67 Dalia, A. B. & Dalia, T. N. Spatiotemporal Analysis of DNA Integration during Natural
963 Transformation Reveals a Mode of Nongenetic Inheritance in Bacteria. *Cell* **179**, 1499-1511
964 e1410 (2019). <https://doi.org:10.1016/j.cell.2019.11.021>

965 68 Cutadapt v. 2.5 (2019).

966 69 Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**,
967 357-359 (2012). <https://doi.org:10.1038/nmeth.1923>

968 70 Chao, M. C., Abel, S., Davis, B. M. & Waldor, M. K. The design and analysis of transposon
969 insertion sequencing experiments. *Nat Rev Microbiol* **14**, 119-128 (2016).
970 <https://doi.org:10.1038/nrmicro.2015.7>

971 71 Chao, M. C. *et al.* High-resolution definition of the *Vibrio cholerae* essential gene set with
972 hidden Markov model-based analyses of transposon-insertion sequencing data. *Nucleic Acids*
973 *Res* **41**, 9033-9048 (2013). <https://doi.org:10.1093/nar/gkt654>

974 72 Marvig, R. L. & Blokesch, M. Natural transformation of *Vibrio cholerae* as a tool--optimizing
975 the procedure. *BMC Microbiol* **10**, 155 (2010). <https://doi.org:10.1186/1471-2180-10-155>

976 73 Shields, R. C., Zeng, L., Culp, D. J. & Burne, R. A. Genomewide Identification of Essential
977 Genes and Fitness Determinants of *Streptococcus mutans* UA159. *mSphere* **3** (2018).
978 <https://doi.org:10.1128/mSphere.00031-18>

979 74 Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. & Ishiguro-Watanabe, M. KEGG for
980 taxonomy-based analysis of pathways and genomes. *Nucleic Acids Res* **51**, D587-d592 (2023).
981 <https://doi.org:10.1093/nar/gkac963>

982 75 Wattam, A. R. *et al.* Improvements to PATRIC, the all-bacterial Bioinformatics Database and
983 Analysis Resource Center. *Nucleic Acids Res* **45**, D535-d542 (2017).
984 <https://doi.org:10.1093/nar/gkw1017>

985 76 Adams, D. W., Pereira, J. M., Stoudmann, C., Stutzmann, S. & Blokesch, M. The type IV
986 pilus protein PilU functions as a PilT-dependent retraction ATPase. *PLoS Genet* **15**, e1008393
987 (2019). <https://doi.org:10.1371/journal.pgen.1008393>

988 77 Donnenberg, M. S. & Kaper, J. B. Construction of an *eae* deletion mutant of enteropathogenic
989 *Escherichia coli* by using a positive-selection suicide vector. *Infect Immun* **59**, 4310-4317
990 (1991). <https://doi.org:10.1128/iai.59.12.4310-4317.1991>

991 78 Li, X. T., Thomason, L. C., Sawitzke, J. A., Costantino, N. & Court, D. L. Positive and
992 negative selection using the *tetA-sacB* cassette: recombining and P1 transduction in
993 *Escherichia coli*. *Nucleic Acids Res* **41**, e204 (2013). <https://doi.org:10.1093/nar/gkt1075>

994 79 Dalia, A. B. Natural Cotransformation and Multiplex Genome Editing by Natural
995 Transformation (MuGENT) of *Vibrio cholerae*. *Methods Mol Biol* **1839**, 53-64 (2018).
996 https://doi.org:10.1007/978-1-4939-8685-9_6

997 80 Häse, C. C. *et al.* Construction and characterization of recombinant *Vibrio cholerae* strains
998 producing inactive cholera toxin analogs. *Infect Immun* **62**, 3051-3057 (1994).
999 <https://doi.org:10.1128/iai.62.8.3051-3057.1994>
1000 81 Sonnenberg, M. S. & Kaper, J. B. Construction of an *eae* deletion mutant of enteropathogenic
1001 *Escherichia coli* by using a positive-selection suicide vector. *Infection and Immunity* **59**, 4310-
1002 4317 (1991). <https://doi.org:doi:10.1128/iai.59.12.4310-4317.1991>
1003

Paper II

RESEARCH ARTICLE

Drug-target binding quantitatively predicts optimal antibiotic dose levels in quinolones

Fabrizio Clarelli^{1,2,3}, Adam Palmer⁴, Bhupender Singh¹, Merete Storflor^{1,5}, Silje Lauksund¹, Ted Cohen⁶, Sören Abel^{1,3,5,7}, Pia Abel zur Wiesch^{1,2,3,7}*

1 Department of Pharmacy, Faculty of Health Sciences, UiT—The Arctic University of Norway, Tromsø, Norway, **2** Department of Biology, Eberly College of Science, The Pennsylvania State University, University Park, PA, United States of America, **3** Center for Infectious Disease Dynamics, Huck Institutes of the Life Sciences, The Pennsylvania State University, University Park, PA, United States of America, **4** Department of Pharmacology, Computational Medicine Program, University of North Carolina at Chapel Hill, Chapel Hill, NC, United States of America, **5** Department of Veterinary and Biomedical Sciences, College of Agricultural Sciences, The Pennsylvania State University, PA, United States of America, **6** Department of Epidemiology of Microbial Diseases, Yale School of Public Health, New Haven, CT, United States of America, **7** Centre for Molecular Medicine Norway, Nordic EMBL Partnership, Oslo, Norway

☯ These authors contributed equally to this work.

* pkz5060@psu.edu



OPEN ACCESS

Citation: Clarelli F, Palmer A, Singh B, Storflor M, Lauksund S, Cohen T, et al. (2020) Drug-target binding quantitatively predicts optimal antibiotic dose levels in quinolones. *PLoS Comput Biol* 16(8): e1008106. <https://doi.org/10.1371/journal.pcbi.1008106>

Editor: Peter M. Kasson, University of Virginia, UNITED STATES

Received: April 3, 2020

Accepted: June 30, 2020

Published: August 14, 2020

Copyright: © 2020 Clarelli et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Data and computer code will be available at <https://www.abel-zur-wiesch-lab.com/> upon acceptance of the manuscript.

Funding: This work was funded by Bill and Melinda Gates Foundation Grant OPP1111658 (to T.C. & P. AzW.), Research Council of Norway (NFR) Grant 262686 (to P.AzW.) and 249979 (to S.A.), and Helse-Nord Grant 14796 (to S.A.). The funders had no role in study design, data collection and

Abstract

Antibiotic resistance is rising and we urgently need to gain a better quantitative understanding of how antibiotics act, which in turn would also speed up the development of new antibiotics. Here, we describe a computational model (COMBAT-COMputational Model of Bacterial Antibiotic Target-binding) that can quantitatively predict antibiotic dose-response relationships. Our goal is dual: We address a fundamental biological question and investigate how drug-target binding shapes antibiotic action. We also create a tool that can predict antibiotic efficacy a priori. COMBAT requires measurable biochemical parameters of drug-target interaction and can be directly fitted to time-kill curves. As a proof-of-concept, we first investigate the utility of COMBAT with antibiotics belonging to the widely used quinolone class. COMBAT can predict antibiotic efficacy in clinical isolates for quinolones from drug affinity ($R^2 > 0.9$). To further challenge our approach, we also do the reverse: estimate the magnitude of changes in drug-target binding based on antibiotic dose-response curves. We overexpress target molecules to infer changes in antibiotic-target binding from changes in antimicrobial efficacy of ciprofloxacin with 92–94% accuracy. To test the generality of our approach, we use the beta-lactam ampicillin to predict target molecule occupancy at MIC from antimicrobial action with 90% accuracy. Finally, we apply COMBAT to predict antibiotic concentrations that can select for resistance due to novel resistance mutations. Using ciprofloxacin and ampicillin as well defined test cases, our work demonstrates that drug-target binding is a major predictor of bacterial responses to antibiotics. This is surprising because antibiotic action involves many additional effects downstream of drug-target binding. In addition, COMBAT provides a framework to inform optimal antibiotic dose levels that maximize efficacy and minimize the rise of resistant mutants.

analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Author summary

Antibiotic efficacy is traditionally assessed with a single value, the minimal inhibitory concentration, despite its limited predictive power. Dose-response curves that assess efficacy over a range of antibiotic concentrations are promising, but extremely work intensive. Using the well-characterized antibiotics ciprofloxacin and ampicillin as proof-of-concept, we provide a mathematical framework that allows predicting optimal dosing of antibiotics based on drug-target binding parameters. These could be measured early in drug development and can also change when bacteria become resistant. Thus, we can predict the probability that resistant strains are selected at a given dosing regimen without isolating and testing a multitude of resistant strains. Predicting optimal dosing with early accessible experimental data can reduce the prevalent trial-and-error approaches and thereby the number of antibiotic candidates that fail late in drug development.

Introduction

The rise of antibiotic resistance represents an urgent public health threat. In order to effectively combat the spread of antibiotic resistance, we must optimize the use of existing drugs and develop new drugs that are effective against drug-resistant strains. Accordingly, methods to improve antibiotic dose levels to i) maximize efficacy against susceptible strains and ii) minimize resistance evolution play a key role in our defense against antibiotic resistant pathogens.

It is noteworthy that dosing strategies for treatment of susceptible strains (e.g., dosing level [1], dosing frequency [2], and treatment duration [3–5]) have recently been substantially improved, even for antibiotic treatments that have been standard of care for decades. This suggests that there likely remains significant room for optimization in our antibiotic treatment regimens. It also highlights the difficulty in identifying optimal dosing levels for new antibiotics. Indeed, optimizing dosing is one of the biggest challenges in drug development. Traditionally, antibiotic efficacy was mainly described by a single value, the minimal inhibitory concentration (MIC). While correlations between treatment success and MIC have been demonstrated there is limited predictive power [6, 7]. When susceptibility is assessed by MIC, not all patients infected with “susceptible” bacteria are successfully treated with antibiotics. Additionally, a large majority of patients with a “resistant” infection can be successfully treated with antibiotics even when the underlying infection is serious and untreated patients are unlikely to recover [8]. This was e.g. shown for patients with complicated intraabdominal infections [7]. Reasons for this mismatch may include that the MIC only gives the minimal concentration to suppress bacterial growth and contains no information on antibiotic efficacy above or below MIC [9]. This makes the MIC ill-suited to describe efficacy of the strongly fluctuating antibiotic concentrations in patients. This has led to an increase in more sophisticated dose-response measurements where bacteria are exposed to multiple antibiotic concentrations and the kill rate is assessed at each concentration individually (pharmacodynamic profiles). However, these approaches require orders of magnitude more experimental effort than simple MIC measurements because they involve a multitude of antibiotic concentrations and time points. This process is too time-consuming when testing new drug candidates.

It is even more challenging to optimize dose levels to minimize the emergence of antibiotic resistance, both for existing and novel antibiotics. Typically, not only the MIC changes when a strain acquires resistance, also other properties such as the steepness of the dose-response curve and the maximal kill rate at very high concentrations change [10]. Predicting the changes in the dose-response curve is therefore not trivial. Thus, a full pharmacodynamic

profile should be assessed for each potential resistant strain. To this end, resistant strains must be isolated and due to the amount of different resistance mechanisms, a good saturation of the mutational target must be achieved. This requires substantial and lengthy evolutionary experiments. In addition, there remains substantial debate about which dosing strategies best prevent the emergence of resistance during treatment [11–13]. In this context, a useful concept that links antibiotic concentrations with resistance evolution is the resistance selection window (mutant selection window) that ranges from the lowest concentration at which the resistant strain grows faster than the wild-type, usually well below the wild-type MIC, to the MIC of the resistant strain [14–16]. Antibiotic concentrations above the resistance selection window safeguard against *de novo* resistance emergence. Antibiotic concentrations below the resistance selection window do not kill the susceptible strain, but also do not favor the resistant strain and therefore do not promote emergence of resistance. To limit resistance, it is therefore important to identify the resistance selection window and optimize dosing accordingly. However, this again requires obtaining a full pharmacodynamic profile of a majority of the expected resistant mutants and is therefore not feasible as a standard assessment in drug development.

The next challenge to successfully designing antibiotic treatment arises when the experimental information is integrated into mathematical pharmacodynamic models that then predict efficacy under realistic, fluctuating concentrations in patients. Pharmacodynamic models from 1910 (E_{\max} or Hill-models) [17] are still widely used despite assuming instantaneous equilibria of antibiotic-target binding and therefore being often inaccurate when antibiotic concentrations fluctuate. Recently described models that relax these assumptions have been useful in gaining a better qualitative understanding of realistic dosing and complicated drug effects, such as post-antibiotic effects, inoculum effects, and bacterial persistence [18–21]. However, to speed the development of new antibiotics or to inform practices which minimize resistance, we require quantitative predictions for antibiotics or resistant bacterial strains that do not exist yet. Models which permit quantitative predictions of changes in drug efficacy as a function of modification of antibiotic molecules (i.e. new drugs) or novel resistance mutations would be invaluable. Such tools would advance our general mechanistic understanding of antibiotic action, could guide dosing trials of new drugs, and suggest better dosing of existing drugs.

In this report, we describe a mechanistic computational modeling framework (COMBAT-Computational Model of Bacterial Antibiotic Target-binding) that allows us to predict full pharmacodynamic profiles based solely on accessible biochemical parameters describing drug-target interaction. These parameters can be determined early in drug development. We use this framework to investigate how changes in drug target binding, either due to improvements in existing antibiotics or due to resistance mutations in bacteria, affect antibiotic efficacy. We first show that COMBAT accurately predicts bacterial susceptibility as a function of drug-target binding and, conversely, allows inference of these biochemical parameters on the basis of observed patterns of bacterial growth suppression or killing. We then use COMBAT to predict the susceptibility of newly arising resistant variants based on the molecular mechanism of resistance and determine the resistance selection window.

Results

Quinolone target affinities correlate with antibiotic efficacy

To investigate how biochemical changes in antibiotic action modifies bacterial susceptibility, we explored how the affinity of antibiotics to their target affects the MIC. We compared the MICs of quinolones, an antibiotic class in which individual antibiotics have a wide range of affinities to one of their targets, gyrase ($K_D \sim 10^{-4}$ – 10^{-7} M) but are of similar molecular sizes

and have a similar mode of action [22]. This choice allowed us to isolate the effects of differences in drug-target affinity on the MIC.

We obtained binding affinities of quinolones to their gyrase target in *Escherichia coli* from previous studies [23–27]. We then retrieved MIC data for several quinolones from clinical Enterobacteriaceae isolates collected before 1990 [28], i.e., before the widespread emergence of quinolone resistance [22]. We assume that quinolone affinities obtained from clinical Enterobacteriaceae isolates collected before the emergence of resistance correspond to those measured in wild-type *E. coli*.

To make qualitative predictions of MICs, we employed a simplified model based on the assumptions that i) drug-target binding occurs much more quickly than bacterial replication, ii) the antibiotic concentration remains constant and iii) that during the 18 hours of an MIC assay, the concentration gradient of the drug inside and outside the cell has equilibrated. Under these assumptions, the MIC can be expressed as

$$MIC = K_D \frac{f_c}{1 - f_c} \quad (1)$$

where K_D represents the affinity constant and f_c the fraction of the target bound at the MIC [29]. Accordingly, this model predicts that the MIC is linearly correlated with K_D .

Fig 1 shows the correlations between drug-target affinities and MICs for seven quinolones and clinical isolates of 11 different Enterobacteriaceae species. We observed a significant ($p < 0.018$) linear correlation between MIC and K_D in all species, confirming the qualitative model prediction.

A quantitative model to predict antibiotic efficacy

While it was encouraging that our model can qualitatively predict MIC changes, our aim was to quantitatively predict antibiotic treatment performance. The simplified model assumes that the binding kinetics are much faster than bacterial replication, which may not be true in all cases. To expand the generalizability of the model, we extended the modeling framework to allow that bacterial replication may occur in a similar time frame as drug-target binding events.

The full model (COMBAT- Computational Model of Bacterial Antibiotic Target-binding) describes the binding and unbinding of antibiotics to their targets and predicts how such binding dynamics affects bacterial replication and death (Fig 2A). In previous work linking drug-target binding kinetics with bacterial replication [21], we described a population of bacteria with θ target molecules per cell with a system of $\theta + 1$ (bacteria with 0, 1, . . . , θ bound target molecules) ordinary differential equations (ODEs). This system increases in complexity with the number of target molecules and makes fitting the model to data computationally too demanding for most settings. To simplify this prior approach, we developed new mathematical models based on partial differential equations (PDEs), where a single equation describes all bacteria simultaneously. The sum of bacteria within all target occupancy states over time can be described by a time kill curve (Fig 2B), during which the bacterial population is characterized by the distribution of bacterial cells with different levels of target occupancies at each time-step (Fig 2C). This curve can be visualized as a two-dimensional surface in a three-dimensional coordinate system where the number of bacteria is represented on the z-axis, the percent of bacteria with the fraction of bound target molecules on the x-axis, and time on the y-axis (Fig 2D).

Antibiotic action is described by rates of binding (k_f) and unbinding (k_r) to bacterial target molecules (Fig 2A and 2E). The binding of an antibiotic to a target results in the formation of an antibiotic-target molecule complex x , where x ranges between 0 and θ .

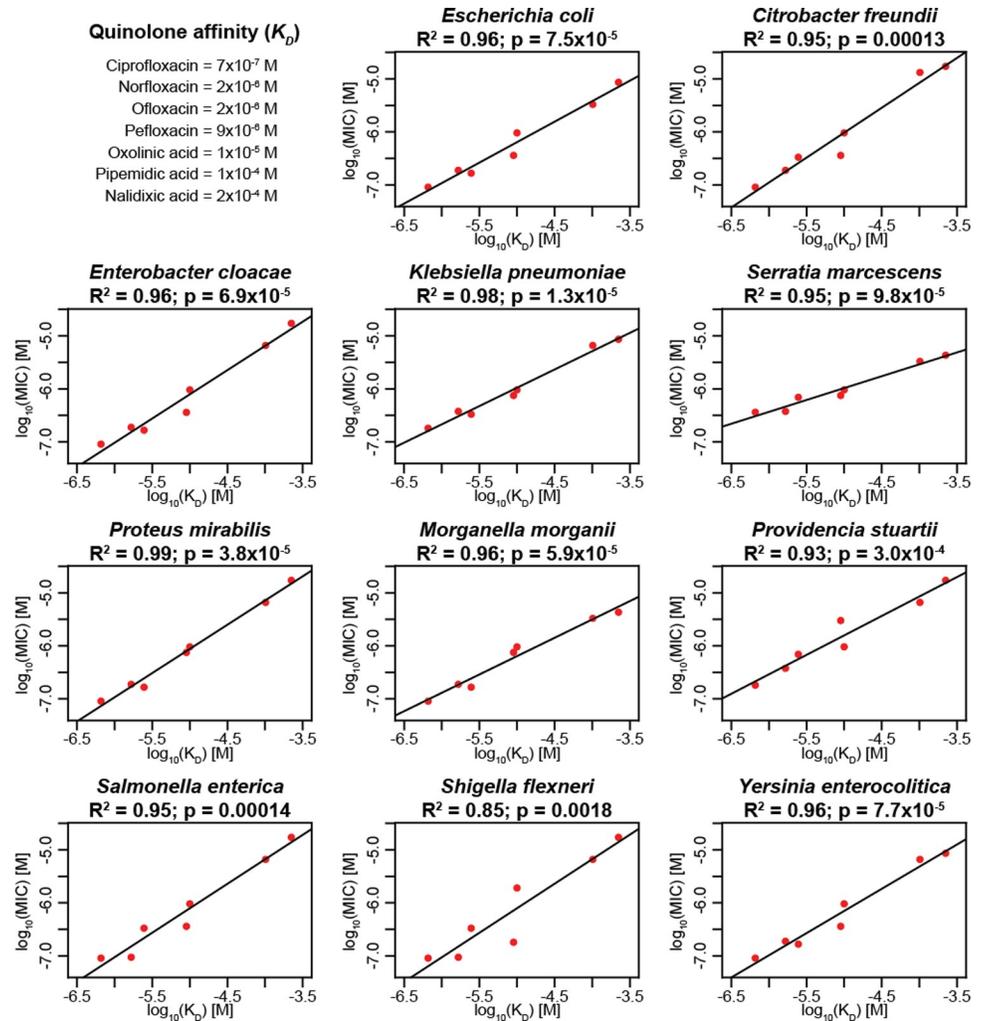


Fig 1. Clinical data confirm linear correlation between MICs and affinities of quinolones to gyrase. We analyzed MIC and drug-target affinity data from 11 Enterobacteriaceae isolates and seven different quinolones. The x-axes show the affinities (K_D) as reported in the literature [24, 26–28], and the y-axes show the MICs, both in mol/L. The adjusted R^2 and p-value of each correlation are given. In cases where there was more than one K_D value reported in the literature, we used the mean for this analysis. The tested MIC values are the median of several clinical isolates described previously [28].

<https://doi.org/10.1371/journal.pcbi.1008106.g001>

COMBAT consists of two mass balance equations: Eq 2 describing bacterial numbers as a function of bound targets and time and Eq 3 describing antibiotic concentration as a function of time (see Methods).

$$\frac{\partial B(x, t)}{\partial t} + \overbrace{\frac{\partial}{\partial x} (v_B(x, t)B(x, t))}^{\text{Binding kinetics}} = \overbrace{-r(x)B(x, t)F_{lim}(t) + S_B(x, t)F_{lim}(t)}^{\text{Replication and its effects on binding}} - \overbrace{\delta(x)B(x, t)}^{\text{Death}} \tag{2}$$

$$\frac{dA(t)}{dt} = -\hat{k}_f A(t) \int_0^\theta (\theta - x)B(x, t)dx + k_r \int_0^\theta xB(x, t)dx \tag{3}$$

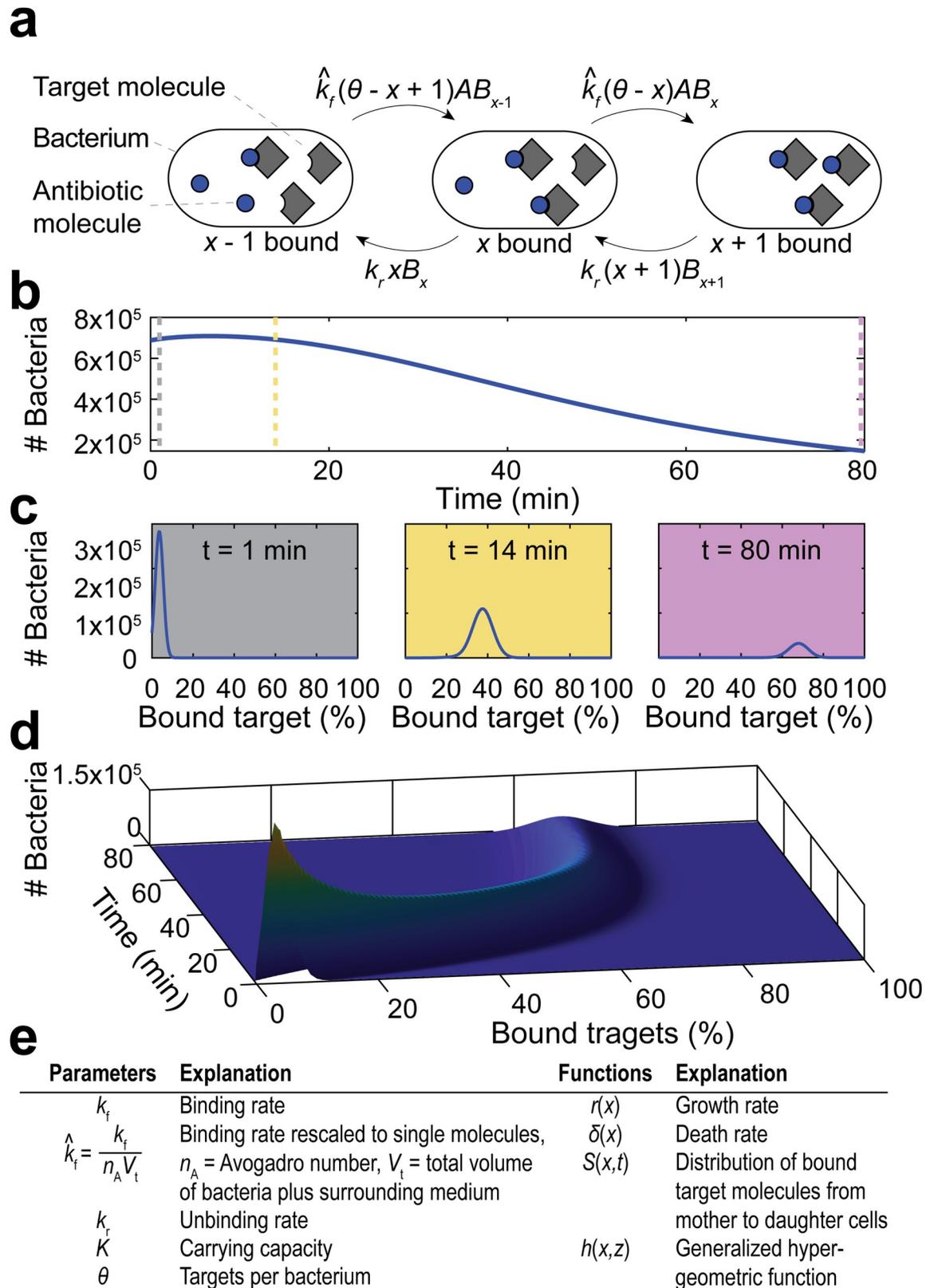


Fig 2. Illustration of modeling approach. a, Schematic illustration of binding kinetics (adapted from [59]). The grey squares depict the drug target molecules, and the blue circles represent antibiotic molecules within bacteria. The arrows indicate individual binding and

unbinding events of the antibiotic to its target molecule in the cell. \hat{k}_f is the adjusted forward reaction rate, k_r is the reverse reaction rate, A is the concentration of antibiotics inside the bacterium, x is the number of bound targets, θ is the number of targets and B_x is the number of bacteria with x bound targets. **b**, Modeled sample time-kill curve, in which the sum of bacteria in all binding states (i.e., the entire population of living bacteria) is followed over time after exposure to antibiotics. The vertical dotted lines indicate the time points depicted in (c); 1 min (grey), 14 min (yellow), and 80 min (purple). **c**, The percentage of bound antibiotic targets in the bacterial population at indicated time points. **d**, Illustration of how the partial differential equation describes the bacterial population as a surface in a three-dimensional coordinate system, the dimensions of which represent percent bound target (x -axis), time (y -axis), and number of bacteria (z -axis). The three time points shown in (c) represent two-dimensional cross-sections at different points of the y -axis. **e**, Overview of used parameters and functions.

<https://doi.org/10.1371/journal.pcbi.1008106.g002>

where $v_B = v_f - v_r$, $v_f = \hat{k}_f A(t)(\theta - x)$ and $v_r = k_r x$. v_B , v_f and v_r can be seen as a generalized velocity $v = \frac{dx}{dt}$.

Eq 4 (part of the replication term in Eq 2) describes how daughter cells inherit bound target molecules from the mother cell during replication:

$$S_B(x, t) = 2 \int_x^\theta h(x, z) r(z) B(z, t) dz; \forall x \in [0, \theta] \quad (4)$$

Eq 5 (part of the replication term in Eq 2) is a logistic growth model describing reduced bacterial replication as the carrying capacity is approached:

$$F_{lim} = \left(1 - \frac{\int_0^\theta B(x, t) dx}{K} \right) \quad (5)$$

Model fit to ciprofloxacin time-kill data

We used the quinolone ciprofloxacin to quantitatively fit bacterial time-kill curves, since this is a commonly used antibiotic for which binding parameters have been directly measured. S1 Table gives an overview of the known parameters used for fitting; S2 Table gives the parameters resulting from our fit.

The functional relationship between the levels of bacterial replication and death on the fraction of bound target molecules is extremely hard to obtain experimentally. We therefore treated the relationships between the fraction of bound target and bacterial replication and death as free parameters in our model fitting. Ciprofloxacin is considered to have both bacteriostatic and bactericidal action (mixed action) [30, 31], and we fitted functions for a monotonically decreasing replication and a monotonically increasing killing with each successively bound target molecule (see Methods & S1 Fig).

Overall, we found that COMBAT could fit the time-kill curves well ($R^2 = 0.93$, Fig 3A). Fig 3B shows the predicted bacterial replication $r(x)$ and death as a function of target occupancy $\delta(x)$ based on the fit obtained in Fig 3A. After model calibration, we simulated bacterial replication during exposure to different antibiotic concentrations for 18 h. For this simulation, positive values indicate an increase in the number of bacteria, and negative values indicate a decrease in the number of bacteria. We estimated a MIC of 0.0139 mg/L (Fig 3C), a value that is within the range of MIC determinations for wt *E. coli* (0.01 mg/L, 0.015 mg/L, 0.017 mg/L and 0.023 mg/L [15, 32–34]).

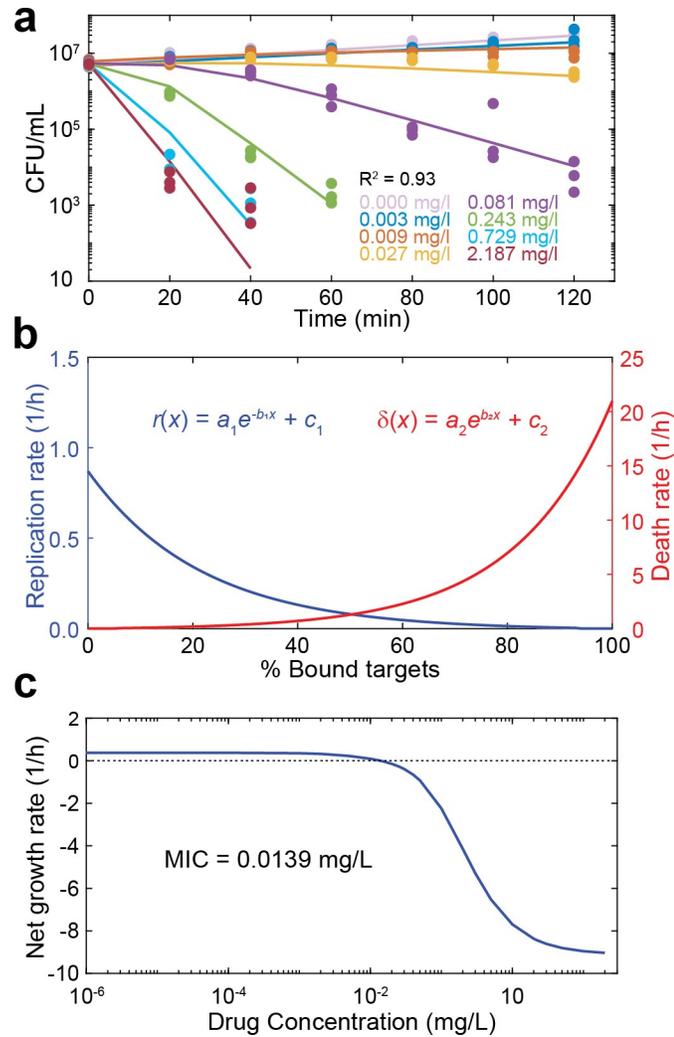


Fig 3. Model predictions for the MIC and the bacteriostatic and bactericidal effects of ciprofloxacin. **a**, Model fit to experimental time-kill curves. The points indicate the experimental data of three independent replicates, and the lines indicate the model fit. Each color indicates a ciprofloxacin concentration as reported in the figure. **b**, The blue line indicates the bacteriostatic effect ($r(x)$, replication rate) of ciprofloxacin and the red line the bactericidal effect ($\delta(x)$, death rate) as a function of the number of bound targets predicted by the model fit in (a). The values of the fitted parameters are listed in S2 Table. **c**, The net growth rate as determined by the slope of a line connecting the initial bacterial density and the final bacterial density of a time-kill curve at 18 h on a logarithmic scale, is given as function of the drug concentration (blue). The dotted horizontal line indicates zero net growth, and the intersection with the blue line predicts the MIC (0.0139 mg/mL).

<https://doi.org/10.1371/journal.pcbi.1008106.g003>

Accurate prediction of target overexpression from time-kill data

Having shown that COMBAT can quantitatively fit experimental data on antibiotic action within biologically plausible parameters, we continued to test the predictive ability of the model. Given our hypothesis that modifications in antibiotic-target interactions lead to predictable changes in bacterial susceptibility, we experimentally induced changes in the antibiotic-target interaction of ciprofloxacin in *E. coli*. We then quantified these biochemical changes by fitting COMBAT to corresponding time-kill curves and compared them to the experimental results. Ciprofloxacin acts on gyrase A₂B₂ tetramers [22]. We used an *E. coli* strain for which both gyrase A and gyrase B are under the control of a single inducible

promoter (P_{lacZ}), such that the amount of gyrase A_2B_2 tetramer can be experimentally manipulated [35]. We measured net growth rates for this strain at different ciprofloxacin concentrations in the presence of 10 μM isopropyl β -D-1-thiogalactopyranoside (IPTG; mild overexpression) and 100 μM IPTG (strong overexpression) and compared it to the wild-type in the absence of the inducer (Fig 4A).

Like previously reported, we find that increasing gyrase content makes *E. coli* more susceptible to ciprofloxacin [35]. We fitted net growth rates allowing the target molecule content, i.e. gyrase A_2B_2 , to vary. We assumed that the only change between the different conditions was the amount of target. We further assumed that the relationship between bound target and bacterial replication or death did not differ between the control strain containing a mock plasmid (no IPTG) and the experiments with overexpression (Fig 4B, between 0% and 100%). Finally, we assumed that the maximal kill rate at very high antibiotic concentrations was accurately measured in our experiments and forced the function describing bacterial death through the measured value when all target molecules are bound. We found the best fit for a 1.31x increase in GyrA_2B_2 target molecule content for bacteria grown in the presence of 10 μM IPTG and a 2.02x increase in GyrA_2B_2 target molecule content for those grown in the presence of 100 μM IPTG.

We subsequently tested these predictions experimentally by analyzing Gyrase A and B content by western blot (Figs 4C and S2). Using realistic association and dissociation rates for biological complexes [36], we predicted a range of functional tetramers based on the relative amount of Gyrase A and B proteins (Fig 4D). S3 Table details the individual measurements, and the procedure to estimate tetramers is provided in the methods section. We found that the observed overexpression was very close to our theoretical prediction, with 1.43x [95% CI 1.19–1.81] overexpression (model prediction = 1.31x overexpression) in the presence of 10 μM IPTG and 2.15x [95% CI 1.73–2.87] overexpression in the presence of 100 μM IPTG (model prediction = 2.02x overexpression).

Accurate prediction of target occupancy at MIC from time-kill data

Next, we tested whether COMBAT can be applied to the action of the beta-lactam ampicillin, a very different antibiotic with a distinct mode of action from quinolones. Using published pharmacodynamic data of *E. coli* exposed to ampicillin [34] also allowed us to compare COMBAT predictions to established pharmacodynamic approaches. Most of the biochemical parameters for ampicillin binding to its target, penicillin-binding proteins (PBPs), have been determined experimentally (S1 Table). Ampicillin is believed to act as a bactericidal drug [37], and this mode of action is supported by findings from single-cell microscopy [29]. We therefore assume that ampicillin binding does not affect bacterial replication. In order to model the consumption of beta-lactams at target inhibition and eventual target recovery, we made small adjustments to Eq 13 (see Methods, description of beta-lactam action).

We fitted COMBAT to published time-kill curves of *E. coli* exposed to ampicillin (Fig 5A). Again, COMBAT provides a good fit to the experimental data between 0 min and 40–60 min. After that time, observed bacterial killing showed a characteristic slowdown at high ampicillin concentrations which is often attributed to persistence [21] (Fig 5A). For the sake of simplicity, we chose to omit bacterial population heterogeneity in this work and therefore cannot describe persistence, even though COMBAT can be adapted to capture this phenomenon [21]. Because ampicillin acts in an entirely bactericidal manner, we assume a constant replication rate (see Methods & S1 Fig) and fitted bacterial death as a function of target binding, $\delta(x)$ (Fig 5B, fitted parameters in S4 Table). Fig 5C shows the predicted net growth rate over a range of drug concentrations. We estimated a MIC of 2.6 mg/L. This MIC is based on the Clinical & Laboratory

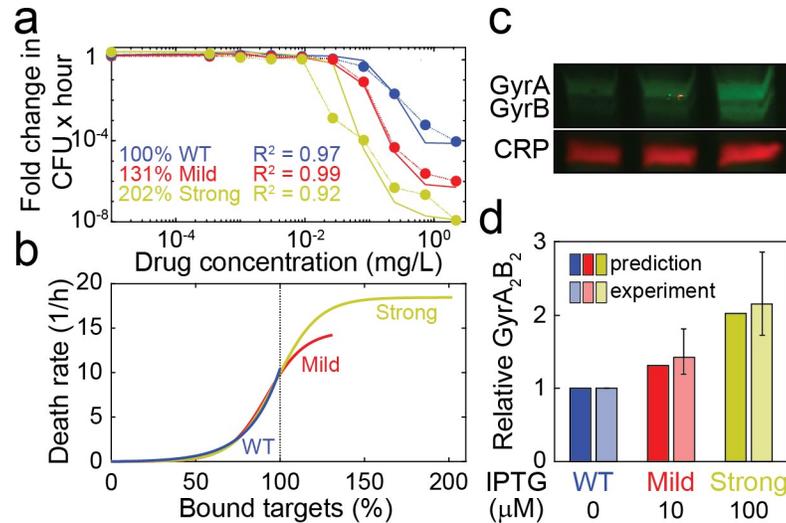


Fig 4. Prediction of relative antibiotic target molecule content from time-kill curves. **a**, Dose-response curves of *E. coli* expressing *gyrA* and *gyrB* under the same IPTG-inducible promoter (SoA3329) grown in the presence of 10 μM IPTG (mild overexpression; red) and 100 μM IPTG (strong overexpression; yellow). A control strain (SoA3330), which expresses wild-type GyrAB levels and contains a mock plasmid, is grown in the absence of inducer (blue). The x-axis indicates the ciprofloxacin concentration, and the y-axis indicates the fold change in colony forming units over time. The dotted lines indicate experimental data, and the solid lines indicate the model fit. The best model fit was obtained for relative target molecule contents of 131% (mild overexpression) and 202% (strong overexpression) relative to the control strain (WT). **b**, Death rates of *E. coli* expressing different levels of GyrAB. The colors represent GyrAB expression conditions as in (a). The x-axis shows the percentage of bound antibiotic target normalized to the control strain; the y-axis shows the death rate $\delta(x)$. Each line represents the best fit for $\delta(x)$. **c**, Western blot analysis of GyrA&B in the strains/conditions shown in (a). CRP (cAMP receptor protein) was used as loading control. A representative example of six replicates is shown; see S2 Fig for full blots. **d**, comparison of theoretical prediction (from (b), solid colors) and GyrA₂B₂ tetramer levels estimated from relative GyrA&B monomer levels (quantified in (c), translucent colors). For the experimental measurements, the bars indicate the mean, and the whiskers represent the 95% confidence interval.

<https://doi.org/10.1371/journal.pcbi.1008106.g004>

Standards Institute definition of the MIC determined at 18 h. The original source of the MIC, which was based on experimental data and a pharmacodynamic model [34] determined an MIC of 3.4 mg/L at 1 h. If we change our prediction to 1 h, our estimated MIC is 3.32 mg/L, which is within 2.5% of the reported value [34].

Having established that COMBAT can also adequately capture the pharmacodynamics of ampicillin, we next tested whether we can estimate experimentally determined target occupancy at the MIC. Our estimated mean occupancy considering both living and dead bacteria is 89% (Fig 5B), a value within previously reported experimental estimates from *Staphylococcus aureus* (84–99%)[38].

Sensitivity of antibiotic efficacy to parameters of drug-target binding

It is possible to vary all parameters in COMBAT and explore their effect. We used this to test how hypothetical chemical changes to ampicillin or ciprofloxacin would affect antibiotic efficacy (S3–S11 Fig). These changes could reflect either bacterial resistance mutations or modifications of the antibiotics themselves. We predict that changes in drug-target affinity, K_D , have more profound effects than changes in target molecule content, bacterial reaction to increasingly bound target (i.e. $\delta(x)$ and $r(x)$), or changes in target molecule content. We also predict that the individual binding rates k_r and k_f , and not just the ratio of these terms, the K_D , are important factors in efficiency. The faster a drug binds, the more efficient we predicted it will be. One intuitive explanation for the observation that k_f drives efficacy is that a slow binding

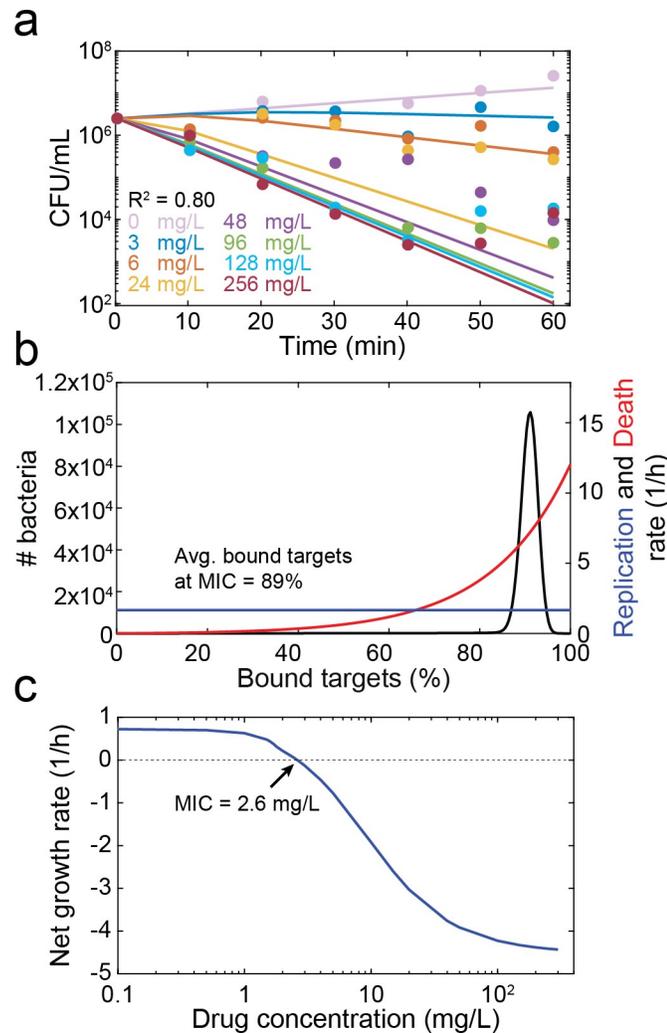


Fig 5. Model prediction of MIC and target occupancy at MIC for ampicillin. **a**, Model fit to previously published time-kill curves [34]. The points represent experimental data, and the lines represent the fit of the model. Each color indicates a single ampicillin concentration, as described in the legend. **b**, Replication (blue) and death (red) rates as a function of the number of bound targets predicted by the model fit in (a). The black line indicates the predicted distribution of target occupancies in a bacterial population (both living and dead cells) exposed to ampicillin at the MIC for 18 h. **c**, The net growth rate, as determined by the slope of a line connecting the initial bacterial density and the bacterial density at 18 h on a logarithmic scale predicted from the model fit in (a), is shown as function of the drug concentration (blue). The dotted horizontal line indicates zero net growth, and the intersection with the blue line predicts the MIC (2.6 mg/mL).

<https://doi.org/10.1371/journal.pcbi.1008106.g005>

fails to rapidly interfere with bacterial replication, which may allow for the production of additional target molecules and thereby reduce the ratio of free antibiotic to target molecules.

Forecasting the resistance selection window

Finally, we illustrate how COMBAT can be used to explore how the molecular mechanisms of resistance mutations affect antibiotic concentrations at which resistance can emerge, i.e., the resistance selection window. We compared predicted net growth rates as a function of ciprofloxacin concentrations for a wild-type strain and an archetypal resistant strain. For this analysis, we assumed that the resistant strain has a 100x slower drug-target binding rate (i.e. ~100x

increased MIC, realistic for novel point mutations [39]) and that the maximum replication rate of the resistant strain is 85% of the wild type strain [40]. We then predicted the antibiotic concentrations at which resistance would be selected. Interestingly, when comparing COMBAT to previous pharmacodynamics models (Fig 5), we observed that estimates of replication rates depend on the selected time frame (Fig 6A). When the timeframe for MIC determination is set to 18 h as defined by CLSI [41], the “competitive resistance selection window”, i.e., the concentration range below the MIC of both strains where the resistant strain is fitter than the wild type, ranges from 0.002 mg/L to 0.014 mg/L for ciprofloxacin (Fig 6A) and 1 mg/L to 2.6 mg/L for ampicillin (S12 Fig), respectively. This corresponds well with previous observations that ciprofloxacin resistance is selected for well below MIC [15]. However, when measuring after 15 min or 45 min, the results are substantially different. The reason for this is illustrated in Fig 6B. COMBAT reproduces non-linear time kill curves where bacterial replication continues until sufficient target is bound to result in a negative net growth rate. This compares well with experimental data around the MIC in Figs 3A and 5A. In Fig 6B, we show model predictions for ciprofloxacin concentrations corresponding to a zero net growth (i.e. same population size) after 15 min, 45 min and 18 h ($MIC_{\text{Resistant}; 15 \text{ min}}$, $MIC_{\text{Resistant}; 45 \text{ min}}$, $MIC_{\text{Resistant}; 18 \text{ h}}$). In all cases, the bacterial population first increases and then decreases slowly. This may have consequences for the selection of resistant strains. Fig 6C illustrates how the resistance selection windows depending on the observed time frame. This suggests that even at concentrations above the 18 h MIC of the resistant strain, there may be initial growth of the resistant strain. In this case, the resistant strain could continue growing at concentration of up to 7 mg/L ciprofloxacin at 15 min, even though the MIC at 18 h is 1.27 mg/L.

Thus, we estimate that sustained levels of 1.27-7mg/L would safeguard against resistance. While ciprofloxacin plasma concentrations typically reach concentrations of 2mg/L after oral uptake and 6mg/L after intravenous administration [42], levels of 2.6 mg/L and above were shown to be chondrotoxic in young animals [43] and concentrations of 40 mg/L are toxic to mitochondria [44]. Clearly, toxicity and risk of resistance must be carefully weighed when deciding on dosing.

Discussion

Optimizing dosing levels of antibiotics is important for maximizing drug efficacy against wild-type strains as well as for minimizing the rise of resistant mutants. Antibiotic efficacy is traditionally described by a single value, the minimal inhibitory concentration (MIC), which has limited predictive power [7, 8]. In more sophisticated dose-response measurements, bacteria are exposed to multiple antibiotic concentrations and the kill rate is assessed at each concentration individually in dose-response curves (pharmacodynamic profiles). However, this approach requires substantial experimental effort and is too time-consuming when testing large libraries of new drug candidates. Limited predictive power of standard measures of pharmacodynamics is not only a problem for antibiotic development, drug attrition in general is mainly due to insufficient predictions of pharmacodynamics rather than pharmacokinetics [45].

Because of the experimental effort, pharmacodynamic profiles for either novel drug candidates or novel resistant strains are often not obtained. Thus, we need a transferrable framework that allows quantitative predictions based on parameters that can be determined *a priori*. Recent studies have reported methods to predict MICs from whole genome sequencing data [46, 47]. However, these methods require transfer of prior knowledge on how the resistance mutations affect MICs in other organisms. There are no methods that could predict *a priori* how chemical changes to an antibiotic structure or novel resistance mutations affect bacterial growth at a given antibiotic concentration.

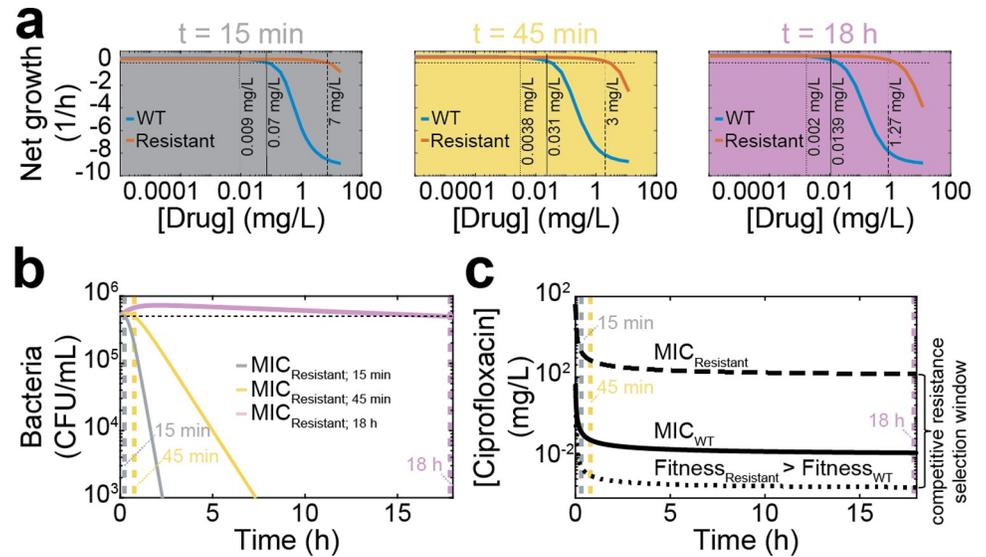


Fig 6. Predicted mutation selection windows for *E. coli* exposed to ciprofloxacin. **a**, The drug concentration of ciprofloxacin is shown on the x-axes, and the average bacterial net growth rate in the first 15 min (grey panel), 45 min (yellow panel), and 18 h (purple panel) of exposure is given on the y-axes. The blue line represents the wild-type strain based on the fits shown in Fig 3, and the red line represents a strain with a hypothetical resistance mutation that decreases the binding rate (k_b) 100-fold and imparts a 15% fitness cost. The horizontal dotted line indicates no net growth. The vertical dotted line indicates where the resistant strain becomes more fit than the wild-type, the solid vertical line indicates the MIC of the wild-type, and the dashed vertical line indicates the MIC of the resistant strain. **b**, Modeled time kill curves of the resistant strain for ciprofloxacin concentrations at which there is no growth at 15 min (grey; MIC_{15 min} = 7 mg/L), 45 min (yellow; MIC_{45 min} = 3 mg/L) and 18 h (purple; MIC_{18 h} = 1.27 mg/L). The horizontal dotted line indicates the initial population size; the vertical dotted lines represent the time points at which the initial and final population size is the same. **c**, The mutation selection window depends on the time at which bacterial growth is observed. The x-axis shows the observed time at which replication rates were determined, the y-axis shows ciprofloxacin concentrations. The dotted curve shows the ciprofloxacin concentration at which the resistant becomes fitter than the WT (Fitness_{Resistant} > Fitness_{WT}), the solid line the MIC of the WT (MIC_{WT}), and the dashed line the MIC of the resistant strain (MIC_{Resistant}). The area between the dotted and dashed line indicates the competitive resistance selection window.

<https://doi.org/10.1371/journal.pcbi.1008106.g006>

Here, we accurately predict antibiotic action on the basis of accessible biochemical parameters of drug-target interaction. Our computational model, COMBAT provides a framework to predict the efficacy of compounds based on drug-target affinity, target number, and target occupancy. These parameters may change both when improving antibiotic lead structures as well as when bacteria evolve resistance. Importantly, they can be measured early in drug development and may even be a by-product of target-based drug discovery [48]. When these data are available, COMBAT makes only one assumption: that the rate of bacterial replication decreases and/or the rate of killing increases with successive target binding. While fitting, we allow this relationship to be gradual or abrupt and select the best fit. This means we do not model specific molecular mechanisms down-stream of drug-target binding, but their effects are subsumed in the functions that connect the kinetics of drug-target binding to bacterial replication and death.

In previous work, for example on antipsychotics [18], antivirals [19] and antibiotics [20, 21], models of drug-target binding kinetics have been used to improve our qualitative understanding of pharmacodynamics. Our study substantially advances this work by making quantitative predictions across antibiotics and bacterial strains when measurable biochemical characteristics change with extremely high accuracy. This is possible because COMBAT employs an efficient and versatile mathematical approach, based on partial differential

equations, that makes it computationally feasible to fit the model to a large range of data. Importantly, we are not only able to predict antibiotic action from biochemical parameters, but can also vice versa use COMBAT to accurately predict biochemical changes from observed patterns of antibiotic action. We have confirmed the excellent predictive power of COMBAT with clinical data as well as experiments with antibiotics with very different mechanisms of action. The high predictive power makes it possible to use modeling to guide dosing. This gives us confidence that biochemical parameters are major determinants of antibiotic action in bacteria and that COMBAT helps to make rational decisions about antibiotic dosing.

In drug development, our mechanistic modeling approach provides insight into which chemical characteristics of drugs may be useful targets for modification. For example, our sensitivity analyses indicate that antibiotics with a similar affinity but faster binding inactivate bacteria more quickly and therefore prevent replication and production of more target molecules, which would change the ratio of antibiotic to target. Furthermore, because e.g. antibiotic binding and unbinding rates can be determined early in the drug development process, such insight can help the transition to preclinical and clinical dosing trials. This may contribute to reducing bottlenecks between these phases of drug development and thereby save money and time.

Avoiding antibiotic concentrations that select for resistance is challenging for two reasons. First, the differences in the pharmacodynamic curves of wild-type and resistant strains are not trivial. Resistance can affect not only the MIC, but also the maximal kill rate at high drug concentrations and the steepness of the dose-response curve [10]. Therefore, one would need to record full pharmacodynamic profiles rather than just MICs to assess the mutation selection window for resistant mutants. Second, this process would have to be repeated for all (or at least a representative set of) potential emerging resistant mutants. This makes it extremely time- and resource-consuming to safeguard against resistance by determining the resistance selection window.

COMBAT offers insight into determinants of the resistance selection window and builds transferrable knowledge that allows estimating useful dose ranges. In concordance with a recent meta-analysis of experimental data [49], our sensitivity analyses predict that changes in drug target binding and unbinding have a greater impact on susceptibility than changes in target molecule content or down-stream processes. Thus, a more comprehensive characterization of the binding parameters of spontaneous resistant mutants would allow an overview of the maximal biologically plausible levels of resistance that can arise with one mutation. Dosing above this level should then safeguard against resistance. This is especially useful for compounds for which it is difficult to saturate the mutational target for resistance, or for safeguarding against resistance to newly introduced antibiotics for which we do not yet have a good overview of resistance conferring mutations.

Good quantitative estimates on the dose-response relationship of new drugs would also help defining the therapeutic window, i.e. the range of drug concentrations at which the drug is effective but not yet toxic. For example in ciprofloxacin, the doses found necessary to prevent resistance after marketing were found to be toxic [50], and an early assessment of doses that might become necessary after resistance is wide-spread might preserve antibiotic utility. If toxicity, solubility or other constraints do not allow dosing above the MIC of expected resistant strains, COMBAT can also predict the concentration range at which resistance is less strongly selected. This could guide decisions on treating with low versus high doses, which is currently controversially debated [11, 12]. COMBAT therefore offers new promise to reduce the failure rates of candidate compounds late in the drug development process when resistance is observed in patients and substantial resources have been invested.

Our quantitative work can help to identify optimal dosing strategies at constant antibiotic concentrations for homogeneous bacterial populations. These measures are commonly used to

assess antibiotic efficacy. In addition, previous work has demonstrated that drug-target binding models outperform traditional pharmacodynamic models for the fluctuating concentrations that actually occur in patients [29, 51]. To illustrate this, we coupled a pharmacokinetic model describing different modes ampicillin administration in patients and predict the pathogen load in infected tissues based on the realistic, fluctuating antibiotic concentrations (S1 Text, S13 and S14 Figs and S5 Table). They can also explain complicated phenomena such as biphasic kill curves, the post-antibiotic effect, or the inoculum effect [20, 21, 52] that often complicate the clinical phase of drug development. COMBAT has similar characteristics that allow capturing these complex phenomena. Therefore, employing COMBAT may be useful for guiding drug development to maximize antibiotic efficacy and minimize *de novo* resistance evolution.

Materials and methods

Mathematical model

COMBAT incorporates the binding and unbinding of antibiotics to their targets and describes how target binding affects bacterial replication and death. This work extends the model developed in [21]. COMBAT consists of a system of two mass balance equations: one PDE for bacteria (describing replication and death as a function of both time and target binding) and one ODE for antibiotic molecules (describing the concentrations as function of time).

In the most basic version of COMBAT, we ignored differences between extracellular and intracellular antibiotic concentrations and only followed the total antibiotic concentration A , assuming that the time needed for drug molecules to enter bacterial cells is negligible. We model ciprofloxacin (to which there is a limited diffusion barrier [53]) and ampicillin (where the target is not in the cytosol, even though the external membrane in gram negatives has to be crossed to reach PBPs). We therefore believe that this assumption is justified in wild-type *E. coli*. This basic version of COMBAT is therefore more accurate for describing antibiotic action where the diffusion barrier to the target is weak.

Binding kinetics

We describe the action of antibiotics as a binding and unbinding process to bacterial target molecules [21]. For simplicity, we assume a constant number of available target molecules θ . The binding process is defined by the formula $A + T \rightleftharpoons x$, where the intracellular antibiotic molecules A react with target molecules T at a rate k_f and form an antibiotic-target molecule complex x , where values for x range between 0 and θ . If the reaction is reversible, the complex dissociates with a rate k_r .

In [21], the association and dissociation terms are described by the following terms

$$\frac{dB_i(t)}{dt} = \overbrace{\hat{k}_f A(t)((\theta - i + 1)B_{i-1}(t) - (\theta - i)B_i(t))}^{\text{Association term}} - \overbrace{k_r(iB_i(t) - (i + 1)B_{i+1}(t))}_{\text{Dissociation term}}; \quad i \in [0, \theta] \quad (6)$$

where $\hat{k}_f = \frac{k_f}{V_{tot}n_A}$, k_f is the association rate, V_{tot} is the volume in which the experiment is performed, n_A is Avogadro's number, k_r is the dissociation rate, B_i is the number of bacteria with i bound targets, and θ is the total number of targets.

This approach requires the use of a large number of ordinary differential equations, $(0 + 1)$ for the bacterial population and one for the antibiotic concentration. To generalize this approach, we assume that the variable of bound targets is a real number $x \in \mathcal{R}$. Under this

continuity assumption, we consider the bacterial cells as a function of x and the time t , thereby reducing the total number of equations to two.

Under the continuity approximation ($x \in \mathcal{R}$), we can rewrite the binding kinetics in the form

$$\frac{\partial B(x, t)}{\partial t} = \overbrace{\frac{\partial}{\partial x} \left(\hat{k}_f A(t)(\theta - x)B(x, t) \right)}^{\text{Association term}} - \overbrace{\frac{\partial}{\partial x} \left(k_r x B(x, t) \right)}^{\text{Dissociation term}} \tag{7}$$

or simply

$$\frac{\partial B(x, t)}{\partial t} = \frac{\partial}{\partial x} \left(v_f(x, t)B(x, t) - v_r(x, t)B(x, t) \right) \tag{8}$$

where $v_f = \hat{k}_f A(t)(\theta - x)$ and $v_r = k_r x$ can be considered as two velocities, i.e., the derivative of the bound targets with respect to the time $\frac{dx}{dt}$.

Replication rate. We assume that the replication rate of bacteria, $r(x)$, is dependent on the number of bound target molecules x . The function $r(x)$ is a monotonically decreasing function of x , such that fewer bacteria replicate as more target is bound. $r(0)$ is the maximum replication rate, corresponding to the replication rate of bacteria in absence of antibiotics. Thus, $r(x)$ describes the bacteriostatic action of the antibiotics, i.e., the effect of the antibiotic on bacterial replication.

Carrying capacity. Replication ceases as the total bacterial population approaches the carrying capacity K . At that point, the replication term of the equation is

$$\frac{\partial B(x, t)}{\partial t} = r(x)B(x, t) \frac{K - \int_0^\theta B(x, t) dx}{K} = r(x)B(x, t)F_{lim} \tag{9}$$

where $F_{lim} = \frac{K - \int_0^\theta B(x, t) dx}{K}$ is the replication-limiting term due to the carrying capacity K , and $0 \leq F_{lim} \leq 1$.

Distribution of target molecules upon division. We assume that the total number of target molecules doubles at replication, such that each daughter cell has the same number as the mother cell. We also assume that the total number of drug-target complexes is preserved in the replication and that the distribution of x bound target molecules of the mother cell to its progeny is described by a hypergeometric sampling of n molecules from x bound and $2\theta - x$ unbound molecules. Under the continuity assumption, we generalize the concept of hypergeometric distribution. Because the hypergeometric distribution is a function of combinations and because a combination is defined as function of factorials, we can use Γ functions in place of factorials and redefine a continuous hypergeometric distribution as a function of Γ functions. A Γ function is

$$\Gamma(\zeta) = \int_0^\infty x^{\zeta-1} e^{-x} dx; \text{Re}(\zeta) > 0 \tag{10}$$

where ζ is a complex number. In this way, the distribution can be expressed as a probability density function of continuous variables. The amount of newborn bacteria is given by the term $r(x)B(x, t)F_{lim}(t)$. We assume that bound target molecules are distributed randomly between mother and daughter cells, with each of them inheriting 50% upon division on average. This means that twice the amount of newborn cells must be redistributed along x to account for the

random distribution process. For example, if a mother cell with 4 bound targets divides, we have two daughter cells, each with a number of bound targets between 0 and 4 (their sum has to be 4), following the generalized hypergeometric distribution. For simplicity, we define $S(x,t)$ to be a function related to the replication rate that depends on the number of bacteria with a number of bound target molecules ranging between x and θ , their specific replication rate $r(x)$, and the fraction of their daughter cells expected to inherit x antibiotic-target complexes $h(x,z)$:

$$S(x, t) = 2 \int_x^\theta h(x, z)r(z)B(z, t)dz \tag{11}$$

Death rate. The death rate function $\delta(x)$ depends on the number of bound target molecules. The function $\delta(x)$ is assumed to be a monotonically increasing function of x , where $\delta(\theta)$ is the maximum death rate, when all targets in the bacteria have been bound by antibiotics. The shape of this function describes the bactericidal action of the antibiotic.

Bacteriostatic and bactericidal effects. We consider several potential functional forms of the relationship between the percentage of bound targets and replication and death rates, because the exact mechanisms how target occupancy affects bacteria is unknown (S1 Fig). We use a sigmoidal function that can cover cases ranging from a linear relationship to a step function. When the inflection point of a sigmoidal function is at 0% or 100% target occupancy, the relationship can also be described by an exponential function. We assume that replication in bactericidal and death in bacteriostatic drugs is independent of the amount of bound target. With sufficient experimental data, the replication rate $r(x)$ and/or the death rate $\delta(x)$ can be obtained by fitting COMBAT to time-kill curves of bacterial populations after antibiotic exposure. The sigmoidal shape of $r(x)$ and $\delta(x)$ can be written as:

$$r(x) = \frac{r_0}{1 + e^{\gamma_r(x-x_{rth})}}; \delta(x) = \frac{d_{max}}{1 + e^{-\gamma_d(x-x_{dth})}} \tag{12}$$

where x_{rth} is the replication rate threshold, x_{dth} is the death rate threshold, and both represent the point where the sigmoidal function reaches 1/2 of its maximum. γ_r and γ_d represent the shape parameters of the replication and death rate functions, respectively. These factors determine the steepness around the inflection point. When they are extreme, the relationship approaches a linear or a step function.

Full equation describing bacterial population. Putting these components together, the full equation describing a bacterial population is:

$$\frac{\partial B(x, t)}{\partial t} + \overbrace{\frac{\partial}{\partial x}(v_f(x, t)B(x, t) - v_r(x, t)B(x, t))}^{\text{Binding kinetics}} = \overbrace{-r(x)B(x, t)F_{lim}(t) + S_B(x, t)F_{lim}(t)}^{\text{Replication and its effects on binding}} - \overbrace{\delta(x)B(x, t)}^{\text{Death}} \tag{13}$$

where $B(x,t)$ is the number of bacteria.

Equation describing antibiotic concentration. The free antibiotic concentration results from mass conservation, i.e., all antibiotic molecules associating with their target are subtracted and all dissociating antibiotic molecules are added. Eq 3 in the results section describes the dynamics of the antibiotic concentration.

Description of beta-lactam action. Beta-lactams acetylate their target molecules (PBPs) and thereby inhibit cell wall synthesis. The acetylation of PBPs consumes beta-lactams. However, PBPs can recover through deacetylation. We modified the term of drug-target dissociation in the equation describing antibiotic concentrations (Eq 3), and set the unbinding rate $k_r = 0$. To reflect the recovery of target molecules, we substituted the dissociation rate k_r in the equation describing the bacterial population with the deacetylation rate k_a , as described in [29].

Initial and boundary conditions. At $t = 0$, we assume that all bacteria have zero bound targets ($x = 0$), and the initial concentration of bacteria is $B(x, 0) = 0$, $x > 0$, and $B(0,0) = B_0$.

At the boundaries of the partial differential equation ($x = 0$, $x = \theta$), we specify that the outgoing velocities are zero. For $x = 0$, i.e. no bound target molecules, the unbinding velocity $v_r(0, t) = 0$, and in $x = \theta$, i.e. all targets are bound, the binding velocity $v_f(\theta, t) = 0$. When the replication term at $x = 0$ and the death term at $x = \theta$ are known, we can solve the partial differential equation with two ordinary differential equations at the boundaries. They are similar to the equations at $x = 0$ and at $x = \theta$ described by Abel zur Wiesch et al. [21], but taking into account that x is a continuous variable instead of a natural number.

Numerical schemes. To solve our system of differential equations, we used a first-order upwind scheme. Specifically, we used the spatial approximation $u'_- = \frac{u(i) - u(i-1)}{\Delta x}$ for the binding term ($v_f > 0$) and the spatial approximation $u'_+ = \frac{u(i+1) - u(i)}{\Delta x}$ for the unbinding term ($v_r < 0$). For the time approximation of both the PDEs and the ODEs, we used the forward approximation $\frac{\Delta B}{\Delta t} = \frac{B^{n+1} - B^n}{\Delta t}$ [54]. We also verified that the Courant-Friedrichs-Lewy condition is satisfied. For fitting the experimental data of bacteria exposed to ciprofloxacin and ampicillin, we used the particle swarm method (“particleswarm” function in Matlab, MathWorks software).

Concentrations of gyrase A₂B₂ tetramers. We assumed that gyrases A and B first homodimerize to A₂ and B₂, respectively, which in turn bind to each other to form the tetramer TR [55]. The following system of equations describes their binding kinetics:

$$\left\{ \begin{array}{l} \frac{dA}{dt} = -2k_1A^2 + 2k_{-1}A_2 \\ \frac{dB}{dt} = -2k_2B^2 + 2k_{-2}B_2 \\ \frac{dA_2}{dx} = k_1A^2 - k_{-1}A_2 - k_3A_2B_2 + k_{-3}TR \\ \frac{dB_2}{dt} = k_2B^2 - k_{-2}B_2 - k_3A_2B_2 + k_{-3}TR \\ \frac{dTR}{dt} = k_3A_2B_2 - k_{-3}TR \end{array} \right. \quad (14)$$



First, we calibrated the model to ensure that we obtain the correct number of gyrase A₂B₂ tetramers (~100) per wild type bacterial cell [56, 57]. This results in an average of each 206

gyrase A and B monomers. Because the association and dissociation rates of the dimers and tetramers are unknown, we sampled 10^4 sets of six parameters in Eq 14 (k_{-3}, \dots, k_3) in a Latin hypercube approach from a biologically plausible range where the association rates are between 10^7 – 10^9 $\text{M}^{-1} \text{s}^{-1}$ and the dissociation rates between 10^{-3} – 10^{-1} s^{-1} [36]. This results in 10^4 estimates for each of the six experimental replicates quantifying gyrase A and B (Figs 4 and S2 and S3 Table).

Experimental methods

Strains, growth conditions and strain construction. *Escherichia coli* strain BW25113 [58] (SoA2740) was transformed with plasmids pCA24N-SC101-gyrAB [35] and pCA24N-SC101- Δ P-YFP [35] using electroporation, resulting in strains BW25113/pCA24N-SC101-gyrAB (SoA3329) and BW25113/pCA24N-SC101- Δ P-YFP (SoA3330), respectively. pCA24N-SC101-gyrAB encodes the *E. coli* gyrAB genes under control of the IPTG inducible LacZ promoter. pCA24N-SC101- Δ P-YFP encodes a promoterless copy of YFP and was used as a control. Bacteria were grown at 30°C on either LB agar or in LB broth, both supplemented with 10 $\mu\text{g}/\text{mL}$ chloramphenicol (Cm) and 10 μM (mild induction) or 100 μM (strong induction) of isopropyl β -D-1-thiogalactopyranoside (IPTG) (43714 5X, VWR Chemicals) when necessary.

Time-kill curves. Overnight cultures of BW25113 or SoA3329 and SoA3330 were diluted 1:1000 in pre-warmed LB or LB-Cm and LB-Cm-IPTG, respectively, and grown with shaking to $\text{OD}_{600} \sim 0.5$. A 1:3 dilution series of ciprofloxacin was made and added to the cultures at indicated concentrations. Additional cultures without antibiotics and with a very high concentration of ciprofloxacin (2.187 mg/L) were used to determine the minimal and maximal kill rate, respectively. Samples were taken immediately prior to addition of the antibiotic and in ~ 20 min intervals or after 45 min, respectively. Samples were washed once in phosphate buffered saline (PBS) before colony forming units (CFUs) were determined for each sample by plating a 1:10 dilution series in PBS on LB agar plates.

GyrAB quantification. To quantify the relative amount of GyrAB, samples of SoA3329 and SoA3330 were collected after 45 min of drug treatment as described above. An equal number of cells corresponding to 1 mL culture at $\text{OD}_{600} = 1$ were harvested by centrifugation. Pelleted cells were lysed at room temperature for 20 min using B-PER bacterial protein extraction reagent (90078, Thermo Scientific) supplemented with 100 $\mu\text{g}/\text{mL}$ lysozyme, 5 units/mL DNaseI (all part of B-PER with Enzymes Bacterial Protein Extraction Kit, 90078, Thermo Scientific) and 100 $\mu\text{M}/\text{mL}$ PMSF (52332, Calbiochem). Samples were stored at -80°C until further use.

Samples were heated to 70°C for 10 min after addition of 1x Bolt sample reducing agent (B0009, Life Technologies) and 1x fluorescent compatible sample buffer (LC2570, Invitrogen). Proteins in whole-cell lysates were separated on 4–15% Mini-Protean TGX Precast gels (456–1085, Bio-Rad) and transferred to 0.2 μm Nitrocellulose membranes (1704158, Bio-Rad).

Membranes were blocked in Odyssey blocking buffer-TBS (927–50000, Li-Cor) for at least one hour at room temperature. Primary antibodies raised against GyrA (Rabbit α -Gyrase A, PA005, Inspiralis), GyrB (Rabbit α -Gyrase B, PB005, Inspiralis), and CRP (Mouse α -*E. coli* CRP, 664304, Nordic Biosite antibodies) were diluted 1:250, 1:250, and 1:2,000 in Odyssey blocking buffer-TBS, respectively. The blocked membranes were incubated with the appropriate primary antibodies overnight at 4°C , washed 4x for 15 min each in TBS-T solution (Tris buffered saline supplemented with Tween20: 0.138 M sodium chloride, 0.0027 M potassium chloride, 0.1% Tween20, pH 8.0 at 25°C), and incubated for 2 h at room temperature with fluorescent labelled secondary antibodies (1:10,000 of IRDye 680RD Goat anti-Mouse IgG, P/N 925–68070, Li-Cor and 1:5000 of IRDye 800CW Goat anti-Rabbit IgG, P/N 925–32211, Li-Cor) in Odyssey blocking

buffer-TBS. Finally, the membranes were washed 4x for 15 min each in TBS-T solution and imaged at 700 nm and 800 nm using a Li-Cor Odyssey Sa scanning system.

Band intensities were quantified from unmodified images using the record measurement tool of Photoshop CS6, normalized to the CRP loading control after background subtraction, and reported relative to SoA3330. For clarity, the “levels” tool of Photoshop CS6 was used to enhance the contrast of shown Western blot images.

Supporting information

S1 Fig. Functions connecting successive antibiotic target binding with bacterial replication and death in bacteriostatic and bactericidal drugs. These graphs show which functions were used to fit the dependence of bacterial replication, $r(x)$, and death, $\delta(x)$, on target occupancy for antibiotics with bacteriostatic, bactericidal or mixed action. Solid lines indicate a sigmoidal relationship, dotted lines indicate a step function, dashed lines indicate a linear relationship, and dash-dotted lines indicate independence, i.e. a constant replication or death rate. The left panels, show the replication rates (blue), the right panels show the death rates (red). The top panels, show rates for a bacteriostatic drug, the middle panels, show rates for a bactericidal drug, and the bottom panels, show rates for a drug with mixed effects. The sum of the replication and death rates at certain target occupancies gives the net growth or decline rate of the bacterial population.

(TIF)

S2 Fig. Western blots of GyrA, GyrB and CRP in *E. coli* whole cell lysates. **a, b,** *E. coli* expressing *gyrA* and *gyrB* under control of the same IPTG-inducible promoter (SoA3329) grown in the presence of 10 μ M IPTG (mild overexpression) and 100 μ M IPTG (strong overexpression). A control strain containing a mock plasmid (SoA3330), representing wild-type GyrAB levels, was grown in the absence of inducer. Whole cell lysates were separated on a SDS-PAGE gel, blotted, and detected with specific fluorescent antibodies against GyrA (green), GyrB (green) and CRP (red). CRP (cAMP receptor protein) was used as loading control.

(TIF)

S3 Fig. Sensitivity analysis of ciprofloxacin fit: Changes in the binding rate k_f . We use the model fitted to experimental data to explore the sensitivity of our results to changes in k_f (0.01x, 0.1x, 1x, 10x, and 100x original value). **a,** Net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h}) / 18 \text{ h}$) as function of drug concentration for different values of the binding rate k_f (see legend). The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the corresponding MICs. **b,** Sensitivity of the MIC to k_f obtained from simulations in (a). The color code indicates the MIC corresponding to the simulation with the same color in (a).

(TIF)

S4 Fig. Sensitivity analysis of ciprofloxacin fit: Changes in the dissociation rate k_r . We use the model fitted to experimental data to explore the sensitivity of our results to changes in k_r (0.01x, 0.1x, 1x, 10x, and 100x original value). **a,** Net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h}) / 18 \text{ h}$) as function of drug concentration for different values of the binding rate k_r (see legend). The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the respective MICs. **b,** Sensitivity of the MIC to k_r obtained from simulations in (a). The color code indicates the MIC corresponding to the simulation with the same color in (a).

(TIF)

S5 Fig. Sensitivity analysis of ciprofloxacin fit: Changes in the drug-target turnover rate.

We use the model fitted to experimental data to explore the sensitivity of our results to changes in the turnover rate of the drug-target complex. We changed k_r and k_f (0.01x, 0.1x, 1x, 10x, 100x, and 1000x original value) while keeping the ratio between k_f and k_r , the affinity K_D , constant. **a**, shows the net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h}) / 18 \text{ h}$) as function of drug concentration for different values of the turnover rate. The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the respective MICs. **b**, Sensitivity of the MIC to turnover rate obtained from simulations in (a). The color code indicates the MIC corresponding to the simulation with the same color in (a).

(TIF)

S6 Fig. Sensitivity analysis of ciprofloxacin fit: Changes in the replication rate with increasingly bound target $r(x)$.

We use the model fitted to experimental data to explore the sensitivity of our results to changes in the replication rate with increasingly bound target $r(x)$. We change the value of bound target at which we obtain a half-maximal replication rate, $x_{1/2}$. **a**, Functions connecting bacterial replication rates $r(x)$ to percentage of bound target molecules with different half-maximal replication rates. **b**, Net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h}) / 18 \text{ h}$) as function of drug concentration for different values of $x_{1/2}$ (see legend). The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the respective MICs. **c**, Sensitivity of the MIC to $r(x)$ obtained from simulations in (b). The color code indicates the MIC corresponding to the simulation with the same color in (a&b).

(TIF)

S7 Fig. Sensitivity analysis of ciprofloxacin fit: Changes in the death rate with increasingly bound target $\delta(x)$.

We use the model fitted to experimental data to explore the sensitivity of our results to changes in the death rate with increasingly bound target $\delta(x)$. We change the value of bound target at which we obtain a half-maximal death rate, $x_{1/2}$. **a**, Functions connecting bacterial death rates $\delta(x)$ to percentage of bound target molecules with different half-maximal death rates. **b**, Net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h}) / 18 \text{ h}$) as function of drug concentration for different values of $x_{1/2}$ (see legend). The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the respective MICs. **c**, Sensitivity of the MIC to $\delta(x)$ obtained from simulations in (b). The color code indicates the MIC corresponding to the simulation with the same color in (a&b).

(TIF)

S8 Fig. Sensitivity analysis of ampicillin fit: Changes in the binding rate k_f .

We use the model fitted to experimental data (Fig 5) to explore the sensitivity of our results to changes in k_f (0.01x, 0.1x, 1x, 10x, and 100x original value). **a**, Net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h}) / 18 \text{ h}$) as function of drug concentration for different values of the binding rate k_f (see legend). The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the respective MICs. **b**, Sensitivity of the MIC to k_f obtained from simulations in (a). The color code indicates the MIC corresponding to the simulation with the same color in (a).

(TIF)

S9 Fig. Sensitivity analysis of ampicillin fit: Changes in the deacetylation rate k_a .

We use the model fitted to experimental data (Fig 5) to explore the sensitivity of our results to changes

in k_a (0.01x, 0.1x, 1x, 10x, and 100x original value). **a**, Net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h})/18 \text{ h}$) as function of drug concentration for different values of the binding rate k_a (see legend). The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the respective MICs. **b**, Sensitivity of the MIC to k_a obtained from simulations in (a). The color code indicates the MIC corresponding to the simulation with the same color in (a). (TIF)

S10 Fig. Sensitivity analysis of ampicillin fit: Changes in the drug-target turnover rate. We use the model fitted to experimental data (Fig 5) to explore the sensitivity of our results to changes in the turnover rate of the drug-target complex. We changed values for k_a and k_f (0.01x, 0.1x, 1x, 10x, and 100x original value) while keeping the ration of k_a/k_f constant. **a**, Net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h})/18 \text{ h}$) as function of drug concentration for different values of the turnover rate (see legend). The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the respective MICs. **b**, Sensitivity of the MIC to turnover rate obtained from simulations in (a). The color code indicates the MIC corresponding to the simulation with the same color in (a). (TIF)

S11 Fig. Sensitivity analysis of ampicillin fit: Changes in the death rate with increasingly bound target $\delta(x)$. We use the model fitted to experimental data (Fig 5) to explore the sensitivity of our results to changes in the death rate with increasingly bound target $\delta(x)$. We change the value of bound target at which we obtain a half-maximal death rate, $x_{1/2}$ (see legend). **a**, Functions connecting bacterial death rates $\delta(x)$ to percentage of bound target molecules with different half-maximal replication rates $x_{1/2}$ (see legend). **b**, Net growth rate ($\log_{10}(\text{bacterial number at 18 h}) - \log_{10}(\text{bacterial number at 0 h})/18 \text{ h}$) as function of drug concentration for different values of $\delta(x)$ (see legend). The dotted horizontal line indicates zero net growth. The intersections of the simulated dose-response curves with this line indicate the respective MICs. **c**, Sensitivity of the MIC to $\delta(x)$ obtained from simulations in (b). The color code indicates the MIC corresponding to the simulation with the same color in (a&b). (TIF)

S12 Fig. Predicted mutation selection windows for *E. coli* exposed to ampicillin. The drug concentration of ampicillin is shown on the x-axes, and the average bacterial net growth rate over 18 h is given on the y-axes. The blue line represents the wild-type strain based on the fits shown in Fig 5, and the red line represents a strain with a theoretical resistance mutation that decreases the binding rate (k_f) 100-fold and imparts a 15% fitness cost. The dotted horizontal line represents no net growth. The first vertical dotted line indicates where the resistant strain becomes fitter than the wild-type (the start of the competitive resistance selection window), the solid vertical line indicates the MIC of the wild-type (the start of the classical resistance selection window), and the dashed vertical line indicates the MIC of the resistant strain, above which selection for resistance should be minimal because both growth of the wild-type and the resistant strain is inhibited. (TIF)

S13 Fig. Schematic of pharmacokinetic model. We simulate plasma and tissue concentrations of ampicillin with a two-compartment pharmacokinetic model. This model described intravenous drug input into the “plasma” compartment, which has an apparent volume of V_1 . From there, it can enter the peripheral “tissue” compartment, characterized by the apparent volume V_2 , with a rate k_{12} . Conversely, the drug can also re-enter the plasma compartment

with a rate k_{21} . From the plasma compartment, the drug is eliminated with a rate k_{10} .
(TIF)

S14 Fig. Coupling a pharmacokinetic model to COMBAT. Two modes of drug administrations, 2 g of drug per day given as single, 5 min i.v. infusions (blue line) and 2 g of drug per day given as continuous i.v. infusion (red line), are simulated. **a**, Simulated drug concentrations in the tissue (i.e., infected) compartment of a two compartment pharmacokinetic model over two days. The dotted black line indicates the MIC of the pathogen (2.6 mg/L). **b**, Pathogen load in the tissue compartment in response to fluctuating drug concentrations predicted by COMBAT over the same timeframe.

(TIF)

S1 Table. Kinetic parameters for used antibiotics. To our knowledge, the association rate for ciprofloxacin has not been determined directly. Because the values of the ratio of dissociation rate k_r and association rate k_f , K_D , diverge by more than an order of magnitude in the literature, we chose to fit the association rate k_f as a free parameter in our model while constraining K_D to remain within the published range (the resulting value of k_f is given together with other fitted parameters in S2 Table).

(DOCX)

S2 Table. Parameters resulting from model fit to experimental time-kill curves of ciprofloxacin in wild-type *E. coli*. Fig 3A shows the model fit and Fig 3B shows the resulting functions for the replication rate $r(x)$ and the death rate $\delta(x)$ as a function of the number of bound targets.

(DOCX)

S3 Table. Results of gyrase level determination and estimated GyrA₂B₂ tetramer levels.

(m) indicates mild overexpression, (s) indicates strong overexpression. Columns headed GyrA and GyrB show the experimentally determined overexpression as fold expression compared to the wild type. The columns headed GyrA₂B₂ show the estimated tetramer levels resulting from each measurement. For the GyrA₂B₂ tetramer estimation, we sampled 10^4 sets association and dissociation rates from a uniform distribution within their reported limits (Latin hypercube approach). We report the standard deviation for each estimate. We give summary estimates in the last row of the table.

(DOCX)

S4 Table. Parameters resulting from model fit to experimental time-kill curves of ampicillin in wild-type *E. coli*. Fig 5B shows the resulting death rate $\delta(x)$ as an exponential function of the number of bound targets $\delta(x) = a_3 e^{b_3 x} + c_3$.

(DOCX)

S5 Table. Parameters used in pharmacokinetic model. The parameters were obtained from [60].

(DOCX)

S1 Text. Predicting pathogen load in patient by coupling COMBAT with a pharmacokinetic model.

(DOCX)

Acknowledgments

We thank Jingyi Liang, Vi Tran, Antal Martinecz, Giovanni Montani, Forrest W. Crawford, Roberto Natalini, Klaus Harms, Christoph Zimmer, Angelo Vannozzi and Rafal Mostowy for

helpful discussions and feedback on the manuscript. We would like to acknowledge R. Regoes for providing previously published raw data.

Author Contributions

Conceptualization: Pia Abel zur Wiesch.

Data curation: Fabrizio Clarelli, Pia Abel zur Wiesch.

Formal analysis: Fabrizio Clarelli, Adam Palmer, Pia Abel zur Wiesch.

Funding acquisition: Ted Cohen, Sören Abel, Pia Abel zur Wiesch.

Investigation: Fabrizio Clarelli, Adam Palmer, Bhupender Singh, Merete Storflor, Silje Lauksund, Pia Abel zur Wiesch.

Methodology: Fabrizio Clarelli, Adam Palmer, Bhupender Singh, Merete Storflor, Silje Lauksund.

Project administration: Ted Cohen, Sören Abel, Pia Abel zur Wiesch.

Resources: Ted Cohen, Sören Abel, Pia Abel zur Wiesch.

Software: Fabrizio Clarelli, Pia Abel zur Wiesch.

Supervision: Ted Cohen, Sören Abel, Pia Abel zur Wiesch.

Validation: Fabrizio Clarelli, Pia Abel zur Wiesch.

Visualization: Fabrizio Clarelli, Sören Abel, Pia Abel zur Wiesch.

Writing – original draft: Fabrizio Clarelli, Sören Abel, Pia Abel zur Wiesch.

Writing – review & editing: Fabrizio Clarelli, Adam Palmer, Bhupender Singh, Merete Storflor, Silje Lauksund, Ted Cohen, Sören Abel, Pia Abel zur Wiesch.

References

1. Boeree MJ, Diacon AH, Dawson R, Narunsky K, du Bois J, Venter A, et al. A dose-ranging trial to optimize the dose of rifampin in the treatment of tuberculosis. *Am J Respir Crit Care Med*. 2015; 191(9):1058–65. <https://doi.org/10.1164/rccm.201407-1264OC> PMID: 25654354.
2. Lan AJ, Colford JM, Colford JM Jr. The impact of dosing frequency on the efficacy of 10-day penicillin or amoxicillin therapy for streptococcal tonsillopharyngitis: A meta-analysis. *Pediatrics*. 2000; 105(2):E19. <https://doi.org/10.1542/peds.105.2.e19> PMID: 10654979.
3. Roord JJ, Wolf BH, Gossens MM, Kimpen JL. Prospective open randomized study comparing efficacies and safeties of a 3-day course of azithromycin and a 10-day course of erythromycin in children with community-acquired acute lower respiratory tract infections. *Antimicrob Agents Chemother*. 1996; 40(12):2765–8. <https://doi.org/10.1128/AAC.40.12.2765> PMID: 9124837; PubMed Central PMCID: PMC163618.
4. Van Deun A, Salim MA, Das AP, Bastian I, Portaels F. Results of a standardised regimen for multidrug-resistant tuberculosis in Bangladesh. *Int J Tuberc Lung Dis*. 2004; 8(5):560–7. PMID: 15137531.
5. WHO. The shorter MDR-TB regimen. http://www.who.int/tb/Short_MDR_regimen_factsheetpdf?ua=1. 2016.
6. Colangeli R, Jedrey H, Kim S, Connell R, Ma S, Chippada Venkata UD, et al. Bacterial Factors That Predict Relapse after Tuberculosis Therapy. *N Engl J Med*. 2018; 379(9):823–33. Epub 2018/08/30. <https://doi.org/10.1056/NEJMoa1715849> PMID: 30157391; PubMed Central PMCID: PMC6317071.
7. Goldstein EJ, Solomkin JS, Citron DM, Alder JD. Clinical efficacy and correlation of clinical outcomes with in vitro susceptibility for anaerobic bacteria in patients with complicated intra-abdominal infections treated with moxifloxacin. *Clin Infect Dis*. 2011; 53(11):1074–80. Epub 2011/10/15. <https://doi.org/10.1093/cid/cir664> PMID: 21998288.

8. Doern GV, Brecher SM. The Clinical Predictive Value (or Lack Thereof) of the Results of In Vitro Antimicrobial Susceptibility Tests. *Journal of Clinical Microbiology*. 2011; 49(9 Supplement):S11–S4. <https://doi.org/10.1128/jcm.00580-11>
9. Onufrak NJ, Forrest A, Gonzalez D. Pharmacokinetic and Pharmacodynamic Principles of Anti-infective Dosing. *Clin Ther*. 2016; 38(9):1930–47. Epub 2016/07/28. <https://doi.org/10.1016/j.clinthera.2016.06.015> PMID: 27449411; PubMed Central PMCID: PMC5039113.
10. Foerster S, Unemo M, Hathaway LJ, Low N, Althaus CL. Time-kill curve analysis and pharmacodynamic modelling for in vitro evaluation of antimicrobials against *Neisseria gonorrhoeae*. *BMC Microbiol*. 2016; 16:216. Epub 2016/09/19. <https://doi.org/10.1186/s12866-016-0838-9> PMID: 27639378; PubMed Central PMCID: PMC5027106.
11. Kouyos RD, Metcalf CJ, Birger R, Klein EY, Abel zur Wiesch P, Ankomah P, et al. The path of least resistance: aggressive or moderate treatment? *Proc Biol Sci*. 2014; 281(1794):20140566. <https://doi.org/10.1098/rspb.2014.0566> PMID: 25253451; PubMed Central PMCID: PMC4211439.
12. Read AF, Day T, Huijben S. The evolution of drug resistance and the curious orthodoxy of aggressive chemotherapy. *Proc Natl Acad Sci U S A*. 2011; 108 Suppl 2:10871–7. <https://doi.org/10.1073/pnas.1100299108> PMID: 21690376; PubMed Central PMCID: PMC3131826.
13. Colijn C, Cohen T. How competition governs whether moderate or aggressive treatment minimizes antibiotic resistance. *Elife*. 2015; 4. <https://doi.org/10.7554/eLife.10559> PMID: 26393685; PubMed Central PMCID: PMC4641510.
14. Drlica K, Zhao X. Mutant selection window hypothesis updated. *Clin Infect Dis*. 2007; 44(5):681–8. <https://doi.org/10.1086/511642> PMID: 17278059.
15. Gullberg E, Cao S, Berg OG, Ilback C, Sandegren L, Hughes D, et al. Selection of resistant bacteria at very low antibiotic concentrations. *PLoS Pathog*. 2011; 7(7):e1002158. <https://doi.org/10.1371/journal.ppat.1002158> PMID: 21811410; PubMed Central PMCID: PMC3141051.
16. Drlica K. The mutant selection window and antimicrobial resistance. *J Antimicrob Chemother*. 2003; 52(1):11–7. Epub 2003/06/14. <https://doi.org/10.1093/jac/dkg269> PMID: 12805267.
17. Gesztelyi R, Zsuga J, Kemeny-Beke A, Varga B, Juhasz B, Tosaki A. The Hill equation and the origin of quantitative pharmacology. *Arch Hist Exact Sci*. 2012; 66(4):427–38. <https://doi.org/10.1007/s00407-012-0098-5> PubMed PMID: WOS:000305840700003.
18. Sykes DA, Moore H, Stott L, Holliday N, Javitch JA, Lane JR, et al. Extrapyramidal side effects of antipsychotics are linked to their association kinetics at dopamine D2 receptors. *Nat Commun*. 2017; 8(1):763. Epub 2017/10/04. <https://doi.org/10.1038/s41467-017-00716-z> PMID: 28970469; PubMed Central PMCID: PMC5624946.
19. Shen L, Rabi SA, Sedaghat AR, Shan L, Lai J, Xing S, et al. A critical subset model provides a conceptual basis for the high antiviral activity of major HIV drugs. *Sci Transl Med*. 2011; 3(91):91ra63. <https://doi.org/10.1126/scitranslmed.3002304> PMID: 21753122; PubMed Central PMCID: PMC3488347.
20. Walkup GK, You Z, Ross PL, Allen EK, Daryaei F, Hale MR, et al. Translating slow-binding inhibition kinetics into cellular and in vivo effects. *Nat Chem Biol*. 2015; 11(6):416–23. Epub 2015/04/22. <https://doi.org/10.1038/nchembio.1796> PMID: 25894085; PubMed Central PMCID: PMC4536915.
21. Abel zur Wiesch P, Abel S, Gkotzis S, Ocampo P, Engelstadter J, Hinkley T, et al. Classic reaction kinetics can explain complex patterns of antibiotic action. *Sci Transl Med*. 2015; 7(287). doi: ARTN 287ra73 <https://doi.org/10.1126/scitranslmed.aaa8760> PMID: 25972005 PubMed PMID: WOS:000354433900001.
22. Aldred KJ, Kerns RJ, Osheroff N. Mechanism of quinolone action and resistance. *Biochemistry*. 2014; 53(10):1565–74. Epub 2014/03/01. <https://doi.org/10.1021/bi5000564> PMID: 24576155; PubMed Central PMCID: PMC3985860.
23. Shen LL, Pernet AG. Mechanism of inhibition of DNA gyrase by analogues of nalidixic acid: the target of the drugs is DNA. *Proc Natl Acad Sci U S A*. 1985; 82(2):307–11. <https://doi.org/10.1073/pnas.82.2.307> PMID: 2982149; PubMed Central PMCID: PMC397026.
24. Shen LL, Mitscher LA, Sharma PN, O'Donnell TJ, Chu DW, Cooper CS, et al. Mechanism of inhibition of DNA gyrase by quinolone antibacterials: a cooperative drug–DNA binding model. *Biochemistry*. 1989; 28(9):3886–94. <https://doi.org/10.1021/bi00435a039> PMID: 2546585.
25. Andriole VT. The quinolones. 3rd ed. San Diego: Academic Press; 2000. XX, 517 S. p.
26. Jungkind DL, American Society for Microbiology Eastern Pennsylvania Branch. Antimicrobial resistance a crisis in health care. New York etc.: Plenum Press; 1995. X, 248 S. p.
27. Kampranis SC, Maxwell A. The DNA gyrase-quinolone complex. ATP hydrolysis and the mechanism of DNA cleavage. *J Biol Chem*. 1998; 273(35):22615–26. Epub 1998/08/26. <https://doi.org/10.1074/jbc.273.35.22615> PMID: 9712890.

28. Siporin C, Heifetz CL, Domagala JM. The New generation of quinolones. New York: M. Dekker; 1990. xii, 347 p., 2 p. of plates p.
29. Abel Zur Wiesch P, Clarelli F, Cohen T. Using Chemical Reaction Kinetics to Predict Optimal Antibiotic Treatment Strategies. *PLoS Comput Biol*. 2017; 13(1):e1005321. <https://doi.org/10.1371/journal.pcbi.1005321> PMID: 28060813; PubMed Central PMCID: PMC5257006.
30. Elliott TS, Shelton A, Greenwood D. The response of *Escherichia coli* to ciprofloxacin and norfloxacin. *J Med Microbiol*. 1987; 23(1):83–8. Epub 1987/02/01. <https://doi.org/10.1099/00222615-23-1-83> PMID: 3546699.
31. Silva F, Lourenco O, Queiroz JA, Domingues FC. Bacteriostatic versus bactericidal activity of ciprofloxacin in *Escherichia coli* assessed by flow cytometry using a novel far-red dye. *J Antibiot (Tokyo)*. 2011; 64(4):321–5. Epub 2011/02/18. <https://doi.org/10.1038/ja.2011.5> PMID: 21326251.
32. Garoff L, Yadav K, Hughes D. Increased expression of Qnr is sufficient to confer clinical resistance to ciprofloxacin in *Escherichia coli*. *J Antimicrob Chemother*. 2018; 73(2):348–52. Epub 2017/11/07. <https://doi.org/10.1093/jac/dkx375> PMID: 29106520; PubMed Central PMCID: PMC5890660.
33. Sulavik MC, Houseweart C, Cramer C, Jiwan N, Murgolo N, Greene J, et al. Antibiotic susceptibility profiles of *Escherichia coli* strains lacking multidrug efflux pump genes. *Antimicrob Agents Chemother*. 2001; 45(4):1126–36. Epub 2001/03/21. <https://doi.org/10.1128/AAC.45.4.1126-1136.2001> PMID: 11257026; PubMed Central PMCID: PMC90435.
34. Regoes RR, Wiuff C, Zappala RM, Garner KN, Baquero F, Levin BR. Pharmacodynamic functions: a multiparameter approach to the design of antibiotic treatment regimens. *Antimicrob Agents Chemother*. 2004; 48(10):3670–6. <https://doi.org/10.1128/AAC.48.10.3670-3676.2004> PMID: 15388418; PubMed Central PMCID: PMC521919.
35. Palmer AC, Kishony R. Opposing effects of target overexpression reveal drug mechanisms. *Nat Commun*. 2014; 5:4296. Epub 2014/07/02. <https://doi.org/10.1038/ncomms5296> PMID: 24980690; PubMed Central PMCID: PMC4408919.
36. Fontana W. Lecture Notes: Continuous-Time Monte-Carlo of Reaction Systems.
37. Ocampo PS, Lazar V, Papp B, Arnoldini M, Abel zur Wiesch P, Busa-Fekete R, et al. Antagonism between bacteriostatic and bactericidal antibiotics is prevalent. *Antimicrob Agents Chemother*. 2014; 58(8):4573–82. Epub 2014/05/29. <https://doi.org/10.1128/AAC.02463-14> PMID: 24867991; PubMed Central PMCID: PMC4135978.
38. Chambers HF, Sachdeva MJ, Hackbarth CJ. Kinetics of penicillin binding to penicillin-binding proteins of *Staphylococcus aureus*. *Biochem J*. 1994; 301 (Pt 1):139–44. <https://doi.org/10.1042/bj3010139> PMID: 8037661; PubMed Central PMCID: PMC1137153.
39. Al-Emran HM, Heisig A, Dekker D, Adu-Sarkodie Y, Cruz Espinoza LM, Panzner U, et al. Detection of a Novel gyrB Mutation Associated With Fluoroquinolone-Nonsusceptible *Salmonella enterica* serovar Typhimurium Isolated From a Bloodstream Infection in Ghana. *Clin Infect Dis*. 2016; 62 Suppl 1:S47–9. Epub 2016/03/05. <https://doi.org/10.1093/cid/civ790> PMID: 26933021.
40. Vogwill T, MacLean RC. The genetic basis of the fitness costs of antimicrobial resistance: a meta-analysis approach. *Evol Appl*. 2015; 8(3):284–95. Epub 2015/04/11. <https://doi.org/10.1111/eva.12202> PMID: 25861386; PubMed Central PMCID: PMC4380922.
41. Jorgensen JH, Ferraro MJ. Antimicrobial susceptibility testing: a review of general principles and contemporary practices. *Clin Infect Dis*. 2009; 49(11):1749–55. Epub 2009/10/28. <https://doi.org/10.1086/647952> PMID: 19857164.
42. Lipman J, Scribante J, Gous AG, Hon H, Tshukutssoane S. Pharmacokinetic profiles of high-dose intravenous ciprofloxacin in severe sepsis. The Baragwanath Ciprofloxacin Study Group. *Antimicrob Agents Chemother*. 1998; 42(9):2235–9. Epub 1998/09/16. <https://doi.org/10.1128/AAC.42.9.2235> PMID: 9736541; PubMed Central PMCID: PMC105792.
43. Stahlmann R, Kuhner S, Shakibaei M, Schwabe R, Flores J, Evander SA, et al. Chondrotoxicity of ciprofloxacin in immature beagle dogs: immunohistochemistry, electron microscopy and drug plasma concentrations. *Arch Toxicol*. 2000; 73(10–11):564–72. Epub 2000/02/09. <https://doi.org/10.1007/s002040050009> PMID: 10663388.
44. Hangas A, Aasumets K, Kekalainen NJ, Paloheina M, Pohjoismaki JL, Gerhold JM, et al. Ciprofloxacin impairs mitochondrial DNA replication initiation through inhibition of Topoisomerase 2. *Nucleic Acids Res*. 2018; 46(18):9625–36. Epub 2018/09/01. <https://doi.org/10.1093/nar/gky793> PMID: 30169847; PubMed Central PMCID: PMC6182158.
45. Waring MJ, Arrowsmith J, Leach AR, Leeson PD, Mandrell S, Owen RM, et al. An analysis of the attrition of drug candidates from four major pharmaceutical companies. *Nat Rev Drug Discov*. 2015; 14(7):475–86. Epub 2015/06/20. <https://doi.org/10.1038/nrd4609> PMID: 26091267.
46. Metcalf BJ, Chochua S, Gertz RE Jr., Li Z, Walker H, Tran T, et al. Using whole genome sequencing to identify resistance determinants and predict antimicrobial resistance phenotypes for year 2015 invasive

- pneumococcal disease isolates recovered in the United States. *Clin Microbiol Infect.* 2016; 22(12):1002 e1- e8. Epub 2016/08/21. <https://doi.org/10.1016/j.cmi.2016.08.001> PMID: 27542334.
47. Eyre DW, De Silva D, Cole K, Peters J, Cole MJ, Grad YH, et al. WGS to predict antibiotic MICs for *Neisseria gonorrhoeae*. *J Antimicrob Chemother.* 2017; 72(7):1937–47. Epub 2017/03/24. <https://doi.org/10.1093/jac/dkx067> PMID: 28333355; PubMed Central PMCID: PMC5890716.
 48. Payne DJ, Gwynn MN, Holmes DJ, Pompliano DL. Drugs for bad bugs: confronting the challenges of antibacterial discovery. *Nat Rev Drug Discov.* 2007; 6(1):29–40. Epub 2006/12/13. <https://doi.org/10.1038/nrd2201> PMID: 17159923.
 49. van der Putten BC, Pasquini G, Remondini D, Janes VA, Matamoros S, Schultz C. Quantifying the contribution of four resistance mechanisms to ciprofloxacin minimum inhibitory concentration in *Escherichia coli*: a systematic review. 2018:372086. <https://doi.org/10.1101/372086> %J bioRxiv
 50. Bennett AC, Bennett CL, Witherspoon BJ, Knopf KB. An evaluation of reports of ciprofloxacin, levofloxacin, and moxifloxacin-association neuropsychiatric toxicities, long-term disability, and aortic aneurysms/dissections disseminated by the Food and Drug Administration and the European Medicines Agency. *Expert Opin Drug Saf.* 2019; 18(11):1055–63. Epub 2019/09/11. <https://doi.org/10.1080/14740338.2019.1665022> PMID: 31500468.
 51. Copeland RA. The drug-target residence time model: a 10-year retrospective. *Nat Rev Drug Discov.* 2016; 15(2):87–95. Epub 2015/12/19. <https://doi.org/10.1038/nrd.2015.18> PMID: 26678621.
 52. Srimani JK, Huang S, Lopatkin AJ, You L. Drug detoxification dynamics explain the postantibiotic effect. *Mol Syst Biol.* 2017; 13(10):948. Epub 2017/10/25. <https://doi.org/10.15252/msb.20177723> PMID: 29061668; PubMed Central PMCID: PMC5658699.
 53. Cramariuc O, Rog T, Javanainen M, Monticelli L, Polishchuk AV, Vattulainen I. Mechanism for translocation of fluoroquinolones across lipid membranes. *Biochim Biophys Acta.* 2012; 1818(11):2563–71. Epub 2012/06/06. <https://doi.org/10.1016/j.bbame.2012.05.027> PMID: 22664062.
 54. Hirsch C. Numerical computation of internal and external flows: introduction to the fundamentals of CFD. Oxford: Butterworth-Heinemann.; 2007. Available from: <http://app.knovel.com/web/toc.v/cid:kpNCIEFFC2>.
 55. Wu J, Zhang Z, Mitchenall LA, Maxwell A, Deng J, Zhang H, et al. The dimer state of GyrB is an active form: implications for the initial complex assembly and processive strand passage. *Nucleic Acids Res.* 2011; 39(19):8488–502. Epub 2011/07/13. <https://doi.org/10.1093/nar/gkr553> PMID: 21745817; PubMed Central PMCID: PMC3201873.
 56. Maier T, Schmidt A, Guell M, Kuhner S, Gavin AC, Aebersold R, et al. Quantification of mRNA and protein and integration with protein turnover in a bacterium. *Mol Syst Biol.* 2011; 7:511. Epub 2011/07/21. <https://doi.org/10.1038/msb.2011.38> PMID: 21772259; PubMed Central PMCID: PMC3159969.
 57. Malmstrom J, Beck M, Schmidt A, Lange V, Deutsch EW, Aebersold R. Proteome-wide cellular protein concentrations of the human pathogen *Leptospira interrogans*. *Nature.* 2009; 460(7256):762–5. Epub 2009/07/17. <https://doi.org/10.1038/nature08184> PMID: 19606093; PubMed Central PMCID: PMC2723184.
 58. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A.* 2000; 97(12):6640–5. Epub 2000/06/01. <https://doi.org/10.1073/pnas.120163297> PMID: 10829079; PubMed Central PMCID: PMC18686.
 59. Martinecz A, Abel Zur Wiesch P. Estimating treatment prolongation for persistent infections. *Pathog Dis.* 2018; 76(6). Epub 2018/08/15. <https://doi.org/10.1093/femspd/fty065> PMID: 30107522; PubMed Central PMCID: PMC6134427.
 60. Ripa S, Ferrante L, & Prenna M (1990) Pharmacokinetics of subactam/ampicillin in humans after intravenous and intramuscular injection. *Chemotherapy* 36(3):185–192. <https://doi.org/10.1159/000238765> PMID: 2338029

Paper III

1 **The bacterial magnesium transporter MgtA reveals highly selective**
2 **interaction with specific cardiolipin species**

3 Julia Weikum ^{1,2}, Jeroen van Dyck ³, Saranya Subramani ¹, David P. Klebl ⁴, Merete Storflor ⁵,
4 Stephen P. Muench ⁴, Sören Abel ⁵, Frank Sobott ^{*6}, J. Preben Morth ^{*1,2,7}

5 ¹ Membrane Transport Group, Centre for Molecular Medicine Norway (NCMM), Nordic EMBL
6 Partnership, University of Oslo, P.O. Box 1137 Blindern, 0318 Oslo, Norway.

7 ² Enzyme and Protein Chemistry, Section for Protein Chemistry and Enzyme Technology,
8 Department of Biotechnology and Biomedicine, Technical University of Denmark, Søtofts Plads,
9 2800, Kgs. Lyngby, Denmark.

10 ³ Department of Chemistry, University of Antwerp, Campus Groenenborger, Groenenborgerlaan
11 171, G.V. 418, 2020 Antwerpen, Belgium.

12 ⁴ School of Biomedical Sciences & The Astbury Centre for Structural Molecular Biology,
13 University of Leeds, Woodhouse Lane, Leeds LS2 9JT, United Kingdom.

14 ⁵ The Abel Lab, Department of Pharmacy, UiT-The Arctic University of Norway, 9037 Tromsø,
15 Norway.

16 ⁶ School of Molecular and Cellular Biology & The Astbury Centre for Structural Molecular
17 Biology, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, United Kingdom.

18 ⁷ Institute for Experimental Medical Research (IEMR), Oslo University Hospital, Ullevål PB 4956
19 Nydalen, NO-0424 Oslo, Norway.

20

21 * correspondence should be addressed to f.sobott@leeds.ac.uk or premo@dtu.dk

22

23 **Abstract**

24 A significant challenge today within protein lipidology is to understand the relationship between
25 cell structure, lipid membrane integrity, ion homeostasis and the embedded membrane proteins.
26 The bacterial magnesium transporter A (MgtA) is a specialized P-type ATPase important for Mg²⁺
27 import into the cytoplasm; disrupted magnesium homeostasis is linked to intrinsic ribosome
28 instability and nitro-oxidative stress in *Salmonella* strains. MgtA's function is highly dependent
29 on anionic lipids, particularly cardiolipin, and further co-localization of cardiolipin with MgtA at

30 the *E. coli* cell poles has been revealed. Here, we show that MgtA has functional specificity for
31 cardiolipin 18:1, but it reaches maximum activity only in combination with cardiolipin 16:0,
32 equivalent to the major components of native cardiolipin found in *E. coli* membranes. This is the
33 first time it has been shown experimentally that two different lipid species from the same class,
34 individually promoting low activity, can enhance activity in combination. Native mass
35 spectrometry verifies the presence of two binding sites for cardiolipin and kinetic studies reveal
36 that a cooperative relationship likely exists between the two different cardiolipin variants
37 (cardiolipin 16:0 and cardiolipin 18:1). This is the first experimental evidence of cooperative
38 effects between lipids of the same class, with only minor variations in their acyl chain composition,
39 acting on a membrane protein. In summary, our results reveal that MgtA exhibits a highly complex
40 interaction with one cardiolipin 18:1 and one cardiolipin 16:0, affecting protein activity and
41 stability, and contributing to our understanding of the particular interactions between lipid
42 environment and membrane proteins. Further, a better understanding of Mg^{2+} homeostasis in
43 bacteria, due to its role as a virulence regulator, will provide further insights into the regulation
44 and mechanism of bacterial infections.

45

46 **Introduction**

47 Biological membranes function as a dynamic interplay between proteins and lipids. The plasma
48 membrane contains a specific lipid composition which can differ locally, and these regions can
49 accommodate functional specialization based on non-random lipid-lipid and lipid-protein
50 interactions (1). The lipid ensemble reveals a high degree of molecular variation in both the head
51 group structure which presents various functional groups at the membrane surface, and in the acyl
52 chains with their variable length and degree of saturation (2), located within the transmembrane
53 segment. This combinatorial space opens the possibility of an additional regulatory function by
54 lipids of the same type which differ only marginally in their acyl chain composition, but have
55 unique structural and functional properties (3). Significant progress has been made in elucidating
56 the intricate and complex interplay between membrane proteins and lipids from experimental
57 studies (4, 5). For several P-type ATPase family members, a role of specific lipid interactions has
58 been elucidated, specifically the plasma membrane Ca^{2+} -ATPase (PMCA) and dimyristoyl
59 phosphatidylcholine-dependent activity variation between PMCA2 and PMCA4 (6), and lipid-

60 protein stoichiometric variations dependent on the functional state of the PMCA (6). The
61 prominent Na⁺/K⁺-ATPase is strongly stimulated and inhibited by lipid-protein interactions (7); of
62 particular importance is the modulatory effect of cholesterol (8) and the stimulatory effect of
63 polyunsaturated lipids such as phosphatidylethanolamine (PE) and phosphatidylserine (PS) (9). The
64 stimulus mainly accelerates the E1P–E2P conformational transition and is mediated by two
65 independent lipid binding sites, with a PS lipid bound between αTM8–10 and the FXVD
66 transmembrane helix, and a PE bound between αM2, 4, 6, and 9 (10).

67 Membrane proteins from bacterial species have also been investigated and while many, however
68 not all, show selective lipid binding (4), their function role often remains poorly understood.
69 Typically, structural studies have revealed the molecular basis of various lipid binding sites in
70 membrane proteins. Specific binding pockets for lipids are referred to as «non-annular lipid sites»,
71 characterized through longer lipid residence times (11). They are stabilized through polar
72 interactions between head groups and charged amino acids at the water-lipid interface or
73 hydrophobic interactions between acyl chains and hydrophobic residues in the transmembrane
74 segments (11). Interestingly, the lipid composition of the membrane, which determines the bulk
75 properties of the lipid bilayer, such as membrane fluidity and bilayer thickness, also influences
76 membrane protein function, as shown for certain P-type ATPases (12). Lipids that associate
77 weakly with transmembrane segments exhibit faster lipid exchange rates. They have been termed
78 «annular lipids/annular lipid shell», as they often form a ring-like structure around the membrane-
79 embedded protein section (11). This study mainly focuses on non-annular, i.e. tightly
80 bound/structural lipid interactions between MgtA and specific cardiolipin species. The magnesium
81 transporter A (MgtA) is a primary active Mg²⁺ transporter, transporting Mg²⁺ from the periplasm
82 to the cytoplasm with orthology widely distributed throughout eubacteria, fungi and plants (13,
83 14). The transporter ensures Mg²⁺ homeostasis which is essential for stabilising membranes and
84 ribosomes, ensuring correct folding of oligonucleotides, and functions as a co-factor for enzymatic
85 reactions. Additionally, Mg²⁺ has been revealed as a bacterial virulence regulator (15) and can also
86 protect against hostile environments, such as withstanding nitro-oxidative stress in professional
87 phagocytes during infection (16). Furthermore, a recently discovered role for MgtA and especially
88 its homologue MgtB suggests that the transporters promote survival in macrophages during
89 infections, by preventing Mg²⁺ deprivation (17).

90 MgtA is a P-type ATPase transporter which utilizes energy from the hydrolysis of adenosine
91 triphosphate (ATP) for Mg^{2+} transport (15, 18) and exhibits a classical P-type ATPase structural
92 composition with three cytoplasmic domains, termed nucleotide-binding (N), phosphorylation (P)
93 and actuator (A) domain, and a transmembrane core with ten transmembrane helices (19, 20). The
94 expression of MgtA is tightly regulated by the two-component system PhoP/PhoQ, which
95 upregulates MgtA levels under low extracellular Mg^{2+} concentrations, low pH and in the presence
96 of antibacterial peptides (15). The PhoP/PhoQ-induced sigma factor RpoS accumulates as a
97 response to low cytoplasmic Mg^{2+} levels in *Salmonella*, tied to the presence of functional Mg^{2+}
98 homeostasis mediated via MgtA and MgtB in *Salmonella* and other pathogens (21-23). Additional
99 regulation of MgtA expression is present through the *mgtA* mRNA leader region, *mgtL*, which
100 senses intracellular Mg^{2+} levels and only promotes MgtA translation under low intracellular Mg^{2+}
101 concentrations (24-26). The enzymatic activity of MgtA is also under tight regulation through
102 substrate inhibition by Mg^{2+}_{free} , which allows MgtA activation only in the concentration range of
103 10 μ M to 1 mM Mg^{2+}_{free} (14). Mg^{2+}_{free} describes Mg^{2+} ions that are not coordinated by ATP but
104 are available for transport by MgtA. However, the mechanism of MgtA inhibition by Mg^{2+}_{free}
105 remains unclear. While most of our understanding of bacterial Mg^{2+} homeostasis derives from
106 *Salmonella*, considerable similarities between *E. coli*, which only has MgtA and not its homologue
107 MgtB, have been described (27).

108 MgtA is located in the inner membrane of *E. coli*, which contains only approximately ~5 %
109 cardiolipin (CL), with the remainder being ~75 % phosphatidylethanolamine (PE) and ~20 %
110 phosphatidylglycerol (PG) (28). Previously, we have shown that CL performs an essential function
111 for MgtA activation, and that it co-localizes at the inner membrane of *E. coli* poles and septal
112 regions (14). Cardiolipin has four acyl chains in comparison to classical glycerophospholipids
113 which only have two. In contrast to eukaryotes, which often have only one or two major
114 cardiolipins (29), *E. coli* contains a large variety of CL species with variable hydrophobic tails
115 with palmitic acid (16:0), palmitoleic acid (16:1) and oleic acid (18:1) as the most abundant acyl
116 chains (30). It has been proposed previously that the anionic head group of cardiolipin is most
117 critical for MgtA ATPase activity, while the acyl chain composition of the lipid only contributes
118 marginally to the enzyme activation (14). Therefore, a function of cardiolipin as a Mg^{2+} or H^+
119 chaperone or reservoir for MgtA-mediated Mg^{2+} -transport was hypothesized (14). The data
120 presented here however indicates a more complex interaction between MgtA and cardiolipin. In

121 this study, we show the specific interaction of MgtA with selected cardiolipin species. Enzyme
122 kinetics studies and native mass spectrometry determined that these lipid-protein interactions are
123 likely mediated through site-specific interactions and biophysical properties of the lipid bilayer
124 environment. We reveal high selectivity of MgtA for CL species 18:1 and CL 16:0, and
125 cooperativity of binding of these two CL molecules to MgtA. Additionally, we show thermal
126 stabilization of MgtA by CL 18:1 through nano differential scanning fluorimetry, while no effect
127 on MgtA stabilization by CL 16:0 was detected. These results highlight the complexity of lipid-
128 membrane protein interactions, with highly selective roles found even for closely related lipid
129 species varying only slightly in acyl chain structure. The implications of this exquisite fine tuning
130 of membrane protein function might indeed be suggestive of a much greater role of lipid
131 composition than previously anticipated.

132

133 **Material and Methods**

134 Lipids were purchased from Avanti Polar Lipids and solubilized in chloroform. *E. coli* cardiolipin
135 (cat no. 841199), cardiolipin 18:1 (cat no. 710335) and cardiolipin 16:0 (cat no. 710333) were
136 purchased as stock solutions of 10 mg/ml. Cardiolipin 14:1 (cat no. 710337) and cardiolipin 16:1
137 (cat no. 710339) were obtained as stock solutions of 5 mg/ml. Total *E. coli* lipid extract (cat no.
138 100500) and 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine (POPE) (cat no. 850757)
139 were purchased as 25 mg/ml stock solution. Detergents were obtained from either Chemical Point
140 UG, Deisenhofen, Germany for n-dodecyl- β -D-maltopyranoside (DDM) or from Nikko Chemicals
141 for octaethylene glycol monododecyl ether (C₁₂E₈). All other chemicals were purchased as grade
142 BioUltra from Sigma-Aldrich.

143

144 **Heterologous expression and purification of MgtA**

145 Cloning of *E. coli* MgtA into pETM11 vector (EMBL), heterologous expression in *E. coli*
146 C43(DE3) and purification were performed as described in Subramani *et al.* (2016) (14, 31). In
147 brief, isolated *E. coli* membranes with overexpressed MgtA were solubilized in 1% β -dodecyl
148 maltoside (β -DDM) prior to Histrap HP (GE Healthcare) purification equilibrated in (25 mM
149 HEPES pH 7.0, 100 mM K₂SO₄, 5% glycerol, 1 mM PMSF, 5 mM β -mercaptoethanol, 3x CMC

150 β -DDM) and eluted in (25 mM HEPES pH 7.0, 100 mM K₂SO₄, 5% glycerol, 1 mM PMSF, 5 mM
151 β -mercaptoethanol, 300 mM imidazole pH 7.6). The elution was followed by size exclusion
152 chromatography, using a HiLoad 16/600 Superdex 200 pg column equilibrated with 1.5 CV of
153 buffer E buffer E (25 mM HEPES, pH 7.0, 100 mM K₂SO₄, 5% glycerol, 1 mM dithiothreitol, 3x
154 CMC β -DDM).

155

156 **ATPase enzymatic assay**

157 Lipid stocks for ATPase enzymatic assays were prepared by tempering indicated lipids, dissolved
158 in chloroform, to room temperature, or heating them to 40 °C until completely dissolved. The
159 lipids were dried under an argon stream with vigorous shaking. The dried lipid film was
160 resuspended in Milli-Q water and 20 mM C₁₂E₈ with vigorous shaking for at least 1 hour at RT
161 and sonication intervals of 5-10 min in between. A final lipid concentration of 6.12 mM or
162 12.24 mM for lipid profiles or a final concentration of 12 mM for Mg²⁺ profiles, respectively, was
163 obtained. Fresh lipid stocks were prepared prior to enzymatic assays. For enzymatic assays in the
164 presence of increasing lipid concentrations (lipid profiles), dilutions to indicated concentrations
165 were prepared in Milli-Q water with 20 mM C₁₂E₈. The same volume (1 μ L) of lipid dilutions was
166 added to the reaction volume to keep detergent levels constant (final concentration: 3x CMC
167 C₁₂E₈), independent of the lipid concentration.

168 ATPase enzymatic assays measuring ATP hydrolysis by detecting released inorganic phosphate
169 were performed according to the protocol described earlier (14). The online web service MAXC
170 (maxchelator.stanford.edu) was used to calculate Mg²⁺_{free} levels in the presence of given
171 concentrations of ATP at pH 7.0 at 37 °C. Apparent V_{max} and K_m were calculated by fitting the
172 curves (specific activity) as a function of Mg²⁺_{free} using a modified uncompetitive substrate
173 inhibition model (equation 1)

$$Y = \frac{V_{max}X}{K_m + X(1 + \frac{X^n}{K_i})} \quad (1)$$

174

175 with X referring to the substrate concentration, Y to the enzyme activity and n to the number of
176 uncompetitive inhibitor sites, provided by GraphPad Prism8 (www.graphpad.com). k_{cat} was
177 estimated using the formula $k_{cat} = V_{max}/[MgtA]$.

178 The Hill coefficient for CL binding was obtained by fitting the curves using the allosteric sigmoidal
179 equation 2

$$Y = \frac{V_{max}X^h}{K_{half}^h + X^h} \quad (2)$$

180
181 provided by GraphPad Prism8 with X referring to the substrate concentration, Y to the enzyme
182 activity, K_{half} to the substrate concentration that produces half-maximal enzyme velocity and h to
183 the Hill slope.

184

185 **Native Mass Spectrometry**

186 Native mass spectrometry (native MS) was performed on purified MgtA protein, subsequently
187 incubated with lipids depending on the experiment. For lipid selectivity of MgtA, total *E. coli* lipid
188 extract (consisting of 67 % PE, 23 % PG and 10 % CL) was used, while selected lipids, POPE and
189 CL, were included for competition assays. Chloroform, in which lipids were dissolved, was
190 removed under an argon stream until the lipid film was completely dried. The lipid film was
191 resuspended in buffer (25 mM HEPES, pH 7.0; 200 mM KCl; 5 % Glycerol) and 6x CMC
192 (1.02 mM) DDM for a final concentration of 1.47 mg/ml for total *E. coli* lipid (for lipid selectivity)
193 or 400 μ M for CL and 2.4 mM for POPE (for competition assays). For the competition assays, CL
194 and POPE were mixed in the indicated ratio. MgtA was diluted in the buffer to a final concentration
195 of 20 μ M. 25 μ l MgtA (20 μ M) were mixed with 25 μ l lipid solution and incubated for 16 h at
196 4 °C in a rotator (5 rpm).

197 MgtA protein samples had to be desalted prior to the native MS measurement. This was
198 accomplished by washing the sample several times in a 100 kDa cut-off concentrator column, with
199 the cut-off small enough to remove most of the unbound lipids and protein-free detergents micelles
200 as well as the non-volatile salts. The protein micelle mass was expected to be around 170 kDa
201 (102.5 kDa for MgtA + ca. 65-70 kDa for the DDM micelle). MgtA was washed with freshly
202 prepared aqueous ammonium acetate (AmAc) at a concentration of 200 mM (adjusted with

203 ammonium hydroxide to pH 7.0), including 3x CMC DDM or C₁₂E₈. 25 µL MgtA was added to
204 475 µL washing solution in the cut-off concentrator column. The sample was concentrated back
205 to 25 µL (ca. 20 µM protein, without Mg²⁺) by spinning it at 10,000 g at 4 °C. When concentrated,
206 the flow-through was discarded and 475 µL fresh wash solution was added. Wash steps were
207 repeated four times, after which the sample was placed in an Eppendorf tube on ice until
208 measurement.

209 Native MS measurements were performed a Synapt G2 HD Q-TOF mass spectrometer (Waters,
210 Wilmslow, UK) with travelling wave (T-wave) ion mobility, using nano-electrospray ionization
211 (ESI) with gold-coated borosilicate capillaries prepared in-house. Instrument parameters in TOF
212 mode were chosen to preserve noncovalent interactions while ensuring optimum spectral
213 resolution: capillary voltage 2.0 kV; source temperature 30 °C; sample cone 200 V; extraction
214 cone 2.0 V; trap collision energy 150 V; transfer collision energy 150 V; trap bias 45 V. Gas
215 pressures in the instrument were: source 6.4 mbar; trap and transfer cell 0.02 mbar; IM cell
216 $3.7 \cdot 10^{-4}$ mbar. Settings were optimized to provide desolvated signals while maintaining important
217 non-covalent interactions.

218

219 **Nano Differential Scanning Fluorometry (nanoDSF)**

220 *E. coli* cardiolipin, specific cardiolipin species or POPE dissolved in chloroform were dried under
221 an argon stream. The lipid film was resuspended in buffer (25 mM HEPES, pH 7.0; 200 mM KCl;
222 5 % glycerol and 21x CMC (1.9 mM) C₁₂E₈) to a final concentration of 1.3 mM. MgtA (final
223 concentration: 14 µM) was mixed with lipid in a 1:100 molar ratio in a final detergent
224 concentration of 3x CMC C₁₂E₈ and filled up with buffer to 40 µL sample volume. Samples were
225 incubated at 4 °C for 16 h, and afterwards excess lipids were removed by centrifugation (20000 g,
226 15 min, 4 °C). nanoDSF was performed using Prometheus NT.48 (Nanotemper) in a temperature
227 range of 15 - 95 °C with a temperature gradient of 1 °C per minute. A minimum of triplicates was
228 used for all experiments.

229 Protein melting temperatures were determined as maxima of the first derivative of the 350 nm/330
230 nm ratio determined by GraphPad Prism8 (www.graphpad.com).

231

232 **In vivo localization**

233 A plasmid carrying *clover-mgtA* was created by gene synthesis (Eurofin Genomics). A start codon
234 free *mgtA* was connected to the 3' end of a stop codon free *clover* by a linker
235 (GGATCCGCTGGCTCCGCTGCTGGTTCTGGCGAATTCggatattc). The sequence is a Gly and
236 Ser rich flexible linker previously designed for GFP fusions (32).

237 This construct was inserted behind the P_{araBAD} promoter of pBAD33.1 (pBAD33.1 was a gift from
238 Christian Raetz; Addgene plasmid #36267; <http://n2t.net/addgene:36267>; RRID:Addgene_36267)
239 (33), so that the ATG of the NdeI restriction site in pBAD33.1 is replaced by the clover start codon
240 to create pSoA280. pSoA280 was transformed in *E. coli* MG1655, MG1655 *clsABC::FRT*
241 (MG1655 Δ *cls*; PO10), and *Vibrio cholerae* c6706. All bacteria were grown at 37 °C in Luria-
242 Bertani (LB) medium with aeration. The media were supplemented with 50 µg/mL or 5 µg/mL
243 chloramphenicol for *E. coli* or *V. cholerae*, respectively, to maintain the plasmid. Expression of
244 Clover-MgtA fusion proteins was induced by addition of 0.2 % L-arabinose (Sigma) in exponential
245 growth phase for 2 h. A Deltavision Elite (GE Healthcare) microscope equipped with Deltavision
246 CMOS camera, a climate chamber set to 37 °C, and an UPlanFLN 100x PH NA1.30 objective
247 (Olympus) was used to image bacteria. 2 µL induced bacterial culture was spotted on a 1 % agar
248 patches immediately prior to microscopy and images in the phase contrast as well as the GFP
249 channel were acquired. The software Oufiti (2015 Oufiti, Jacobs-Wagner Lab) was used to detect
250 bacteria in the phase contrast channel and overlay their outlines with a mesh that subdivides each
251 cell in segments (34). This mesh was used to quantify the fluorescent intensity in each subdivision
252 in the GFP channel. Intensity values and cell lengths were extracted from the Oufiti output files
253 using Matlab (R2019a, Mathworks) and used to plot the intensity profiles of cells. For better
254 visualization, the microscopic image were adjusted using the levels tool in Photoshop CS6 (Adobe)
255 by setting the same black- and white-point values for all fluorescent images.

256

257 **SDS-PAGE and immunoblot**

258 20 ml LB supplemented with appropriate concentrations of arabinose and chloramphenicol were
259 inoculated 1:100 with overnight culture. Cultures were grown for 2 h at 37 °C with aeration. The
260 OD₆₀₀ of the culture was measured and similar cell amounts were harvested by centrifugation at
261 ~20000 x g at 4 °C. Pellets were re-suspended in SDS loading buffer (50 mM Tris-HCl (pH 6.8);
262 2 % sodium dodecyl sulfate (SDS); 100 mM dithiothreitol (DTT); 10 % glycerol, 0.02 % NaN₃)

263 and heated for 10 min at 90 °C. The samples were separated by SDS-polyacrylamide gel
264 electrophoresis (PAGE) on 10 % gels and transferred to nitrocellulose membranes (BioRad).
265 Clover was detected using monoclonal antibodies (anti-GFP 1:5000, Santa Cruz Biotechnologies)
266 and monoclonal anti mouse secondary antibodies conjugated to horseradish peroxidase (1:5000,
267 Bio-Rad). After incubation with ECL substrate (BioRad) chemiluminescence was recorded on a
268 Bio-Rad ChemiDoc MP Imaging System.

269

270 **Flow cytometry**

271 5 ml LB, supplemented with appropriate concentrations of arabinose and chloramphenicol, were
272 inoculated 1:100 with overnight culture. Cultures were grown for 2 h at 37 °C with aeration. The
273 OD₆₀₀ of the culture was measured and the same amount of cells was harvested by centrifugation
274 at 20000 x g at 4 °C. Pellets were resuspended in PBS and FSC, SSC and fluorescence (EX 488
275 nm; EM 533 +/-30 nm) was rescored on a BD Accuri C6 Plus Flow Cytometer (BD Biosciences).
276 PBS was used as control to distinguish bacterial cells from other particles in FSC and SSC.

277

278 **Results**

279 **Native mass spectrometry reveals selective binding of two cardiolipin molecules**

280 Native mass spectrometry (MS) has recently emerged as a powerful tool for the investigation of
281 dynamic and heterogeneous membrane protein structures and lipid interactions (35, 36). This
282 approach was used in this study to characterize the stoichiometry of MgtA-lipid interactions, and to
283 investigate the selectivity of specific cardiolipin species binding preferentially to the protein.

284 Lipid interaction experiments were performed on purified protein, but full delipidation was not
285 always achieved with the protocol (described in Material and Methods) with some samples
286 showing low levels of retained native lipids. Two types of experiments were carried out: (i) a full
287 *E. coli* lipid extract (Avanti Polar Lipids) was added to the purified protein in the presence of 3x
288 CMC DDM in order to examine lipid binding selectivity, and (ii) a titration with cardiolipin (CL)
289 was carried out to investigate binding affinity, using POPE as a nonspecific binding control. The
290 *E. coli* lipid extract, used for MgtA lipidation, contains only 9.8 % CL, while the majority of the
291 lipid content consists of PE (57.5 %), PG (15.1 %) and uncharacterized components (17.6 %)

292 (Avanti Polar Lipids). Cardiolipin used for titration experiments was extracted from *E. coli*
293 membranes (Avanti Polar Lipids), containing a large variety of CL species with different acyl
294 chain composition.

295 Figure 1a shows two charge states (m/z peaks) in the native mass spectrum of delipidated MgtA
296 (as purified) and without addition of lipids; the corresponding full mass spectrum is shown in
297 Supplementary figure 1. From this, the protein mass is determined experimentally as 102,585 +/-
298 8 Da, which matches the calculated sequence mass for MgtA, i.e. 102,665.0 Da with and 102,533.8
299 Da without the N-terminal methionine, reasonably well. Figure 1b shows the same charge states,
300 using the same experimental conditions, after incubation of purified MgtA with *E. coli* lipid extract
301 and subsequent detergent wash. Additional peaks are apparent now which suggests that lipid
302 binding is detected, with the apparent mass shift for the first binding event calculated from these
303 charge states as ca. 1380 +/- 200 Da, and for the second as ca. 1440 +/- 125 Da. These masses
304 broadly correspond to the molecular weight of one and two cardiolipin molecules, respectively. It
305 should be noted that other lipids present in the extract, in particular PE (719.3 Da) and PG (761.07
306 Da), have masses which are roughly half the mass of CL. It is therefore possible, at the limited
307 mass resolution obtained on this instrument, that the observed lipid adduct peaks might be instead
308 due to two or four bound PE or PG molecules, respectively. However, in that case the mass spectra
309 would also be expected to show the 1x and 3x bound states, which is not the case. The identification
310 of the bound lipids as cardiolipin species is also further corroborated by the following data.

311
312 We additionally performed a competition experiment in which binding of cardiolipin was studied
313 in the presence of another lipid, POPE, added in large excess (Figure 2 and Supplementary figure
314 2). POPE is a major lipid component present in the *E. coli* membrane and, therefore, was used as
315 a negative control representing the native *E. coli* lipid environment. Additionally, POPE was
316 included to evaluate if hydrophobic exposed patches on the protein show any nonspecific lipid
317 interactions in our setting. Lipid binding to MgtA was studied in the absence and presence of CL
318 with increasing concentrations up to 30 μM , and correspondingly decreasing concentrations of
319 POPE from 1000 to 970 μM , respectively. This leaves the total lipid concentration constant at
320 1000 μM , which corresponds to a 1:100 molar ratio between MgtA and lipid. In this experiment
321 samples were obtained from a different protein batch, and even the control sample (first two panels
322 in Figure 2a, Purified MgtA after lipid removal, and 1000 μM POPE) revealed already a low

323 amount of bound cardiolipin, which was most likely retained during the purification process in this
324 case. These co-purified cardiolipins were not observed in the delipidated sample of the previous
325 experiment (Figure 1), presumably due to differences in delipidation efficiency. Nevertheless, the
326 addition of increasing amounts of CL showed a corresponding increase in the relative intensity of
327 adduct peaks which correlate to those seen in Figure 1, with the relative intensity of the 1x bound
328 state in particular increasing slightly from ca. 25% to ca. 30%, for both the 25+ and 24+ charge
329 states. Interestingly the 2x lipid-bound peak intensity seems to not change significantly within the
330 limited concentration range studied here (Figure 2b and Supplementary figure 2). While the data
331 demonstrate qualitatively that the binding of the first CL lipid occurs with higher affinity and
332 increases with increasing CL concentration, the second lipid binding event is of lower affinity and
333 can only be seen clearly above 20 μ M CL (i.e. at a 2:1 ratio to the protein).

334 These results support that MgtA selectively binds up to two cardiolipin molecules, and not any
335 other lipids present in the native *E. coli* membrane. They also show that the first binding event
336 happens with higher affinity, and the mass of this cardiolipin species appears to be somewhat lower
337 than that of the second one binding, although the achievable mass resolution and peak shapes do
338 ultimately not allow to make firm conclusions. With a more powerful spectrometer, the masses of
339 the individual lipids bound could be determined more accurately, and the lipid titration series
340 presented here could also be extended to higher CL : POPE ratios in order to quantitatively
341 determine binding affinities and possible cooperative effects.

342

343 **No specificity of cardiolipin binding to selected conformational states of MgtA was detected**

344 As a P-type ATPase, MgtA transports Mg^{2+} according to the catalytic cycle described by the Post-
345 Albers-scheme (35). During the catalytic cycle MgtA likely alternates between different
346 conformational states as described in the introduction for the Na^+/K^+ -ATPase. The two major
347 conformational states are termed E1 and E2, which exhibit large structural differences and
348 different affinities to the transported ion and counter ion (35) (Supplementary figure 3A). To
349 investigate whether MgtA has a higher affinity for cardiolipin in one specific conformational state,
350 which would indicate a role of cardiolipin for a specific transport step, the protein was incubated
351 with inhibitors AlF_4^- and $ADP-AlF_4^-$. These inhibitors essentially mimic phosphate and ATP at the

352 binding sites on the enzyme and lock MgtA in the two different conformational states, E1 and E2,
353 respectively. After samples were incubated with inhibitors, CL extracted from the *E. coli*
354 membrane (*E. coli* CL) was added to MgtA as previously described. Native MS spectra showed
355 binding of cardiolipin in all samples (Supplementary figure 3B). This indicates that there is no, or
356 only a very small, preference of cardiolipin to a selected conformational state of MgtA. In
357 summary, we find that MgtA contains two specific binding sites for lipids, and selectively binds
358 cardiolipin among other major lipids present in *E. coli* membranes.

359

360 **The composition of the hydrophobic cardiolipin tail plays a major role in MgtA activation**

361 Cardiolipins are anionic lipids with a unique dimeric structure, with two phosphate ester groups
362 and four acyl chains (Figure 3A). While the head group defines the lipid class, large molecular
363 diversity is present through variations in the length and saturation degree of the fatty acid chains
364 (Figure 3A). Previously, we have shown that cardiolipin is required for MgtA activation (11).
365 However, all enzymatic assays were performed in the presence of cardiolipin extracted from *E.*
366 *coli* membranes consisting of a mixture of different cardiolipin species (11). To investigate
367 whether specific cardiolipins stimulate MgtA activity, cardiolipin species with varying acyl chain
368 length and saturation were used in the ATPase assay. In *E. coli*, a large variety of cardiolipin
369 species have been detected with over 50 different types present ranging in acyl carbon chain length
370 from 12 to 19 containing different acyl chain combinations (30). Commercially only cardiolipins
371 with four identical acyl chains are available at Avanti Polar Lipids, but these do not mimic the
372 most common natural cardiolipin species in the *E. coli* membrane which contain mixed acyl
373 chains. However, we used the CL species consisting of palmitic acid (16:0), palmitoleic acid (16:1)
374 and oleic acid (18:1) as the most common acyl chains present in *E. coli* cardiolipin. Large
375 differences in MgtA ATPase activity levels were detected in the presence of these different types
376 of cardiolipin molecules (Figure 3B). For instance, activity was reduced to approximately 50 % in
377 the presence of CL 18:1 and to barely 10 % in the presence of CL 14:1 in comparison to *E. coli*
378 CL. In the presence of any single cardiolipin species, MgtA activity never reached the same
379 activation level as in the presence of *E. coli* extracted cardiolipin.

380 Interestingly, an equimolar combination of two species, CL 18:1 and CL 16:0, exhibited higher
381 MgtA ATPase activity levels relative to *E. coli* CL at the same protein:lipid ratio. It is noteworthy

382 that although MgtA exhibits lower activity in the presence of CL 18:1 alone, CL 16:0 by itself
383 does not seem to activate MgtA at all. This corresponds to previous results comparing MgtA
384 activation in the presence of cardiolipin extracted from either *E. coli* or bovine heart membranes,
385 which showed a decreased activity in the latter (14). While cardiolipins share the same head group,
386 *E. coli* cardiolipin has a diverse fatty acid composition with 16:0-18:1 fatty acids at positions 1
387 and 2 of sn-glycerol 3-phosphate as the most dominant species (35 %), whereas bovine heart CLs
388 contain almost entirely linoleate chains (18:1, 18:2, 18:3) and no stearic chains (16:0) (14).

389 As activity studies shown in figure 3B revealed the importance of two different cardiolipin species,
390 CL 18:1 and CL 16:0 for MgtA activity, we hypothesized the specific interactions of these two
391 lipid species with individual binding sites in MgtA. ATPase activity in the presence of selected
392 cardiolipin species followed a sigmoidal shaped curve, which indicates cooperative binding at
393 more than one binding site (Figure 3B). Cooperativity in protein-ligand interaction refers to
394 increased affinity for additional ligands upon binding of the first ligand in the case of positive
395 cooperativity, or decreased ligand affinity of the protein for additional ligands for negative
396 cooperativity (37). Cooperativity of ligand binding can be quantified by the Hill coefficient (n_H),
397 which measures the sigmoidal character of the activation curve (38). The reference state is a
398 hyperbolic curve with a single monomeric binding site, characterized by $n_H=1$. Therefore, binding
399 curves with $n_H>1$ are a direct measurement for cooperativity. Although, n_H is not a direct
400 measurement of ligand binding sites, it is a useful index for the theoretical upper limit of the
401 number of binding sites (38). Only recently, the first case of allosteric modulation and positive
402 cooperative lipid binding has been detected for lipid-lipid interaction on *E. coli* ammonium
403 channel AmtB (39). Fitting of the cardiolipin-dependent activation curve of MgtA to a sigmoidal
404 allosteric model revealed positive cooperative binding of cardiolipin (Figure 3B; Table 1). In our
405 analysis only least squares fits (R-square) of 0.95 or above were considered; i.e. only activation
406 curves in the presence of *E. coli* CL and the CL 18:1 / CL 16:0 mixture in equimolar ratio. MgtA
407 activation curves in the presence of CL 14:1, CL 18:1, CL 16:1 and CL 16:0 were excluded based
408 on the poorer fit. Plotting of MgtA activation curves in the presence of *E. coli* CL, or equimolar
409 amounts of CL 18:1 and CL 16:0 against an allosteric sigmoidal fit, showed a Hill coefficient of
410 2.4 - 3. This supports the hypothesis of more than one cardiolipin binding site with cooperative
411 binding of multiple CL species.

412

413 **Cardiolipin 16:0 is required for optimal MgtA activation**

414 To better understand the effect of CL 18:1 and 16:0 and their combination on MgtA activation,
415 ATPase activity profiles in the presence of these CL species with increasing Mg^{2+} concentrations
416 were performed. The total lipid concentration was kept constant at 12 mM in all experiments.
417 Results confirmed previous observations that equimolar amounts of both lipid species lead to
418 maximum activation (Figure 4A), mimicking the activity levels achieved previously in the
419 presence of native *E. coli* CL(14). In the presence of CL 18:1 alone, MgtA exhibited lower levels
420 of activity, while no measurable ATPase activity was observed in the presence of CL 16:0 alone
421 (Figure 4A). To assess whether the molar ratio between CL 18:1 and CL 16:0 affects activation
422 levels, ATPase assays were performed with molar ratios of 3:1 and 1:3, with either CL 18:1 or CL
423 16:0 as the major lipid. Additionally a molar ratio of 1:1 between CL 18:1 and CL 16:0 (as above)
424 was included again. Maximum enzyme activity was obtained only when 50 % or more of the lipid
425 content consisted of CL 16:0. Enzymatic assays with higher percentages of CL 18:1, e.g. in the
426 3:1 ratio, resembled the activity levels obtained in the presence of CL 18:1 alone (Figure 4B).

427 To verify whether only either CL 18:1 or CL 16:0, mixed with another cardiolipin species, were
428 needed for optimal activation, ATPase assays in the presence of CL 18:1 or CL 16:0 mixed in a
429 1:1 molar ratio with CL 16:1 were performed (Figure 4C). Activation levels did not reach their
430 maximum for samples containing CL 16:1 alone or in either combination with CL 16:0 or 18:1.
431 This highlights the importance of both species, CL 18:1 and CL 16:0, for MgtA activation. It
432 should be noted though that only activating effects according to cardiolipin species were
433 investigated, whereas possible variations in binding affinities between cardiolipin species were not
434 assessed. As all experiments were however performed in large excess of CL at the same total
435 concentration of 12 mM, we assume that if CL 16:0 would exhibit a much higher binding affinity
436 than CL 18:1, it would occupy both binding sites, even in a sample with the majority of lipid
437 represented by CL 18:1, and therefore MgtA would exhibit similar activity levels as in samples
438 containing CL 18:1 and CL 16:0 in a 3:1 and 1:1 molar ratio.

439 Calculated V_{max} , K_m and k_{cat} in the presence of different cardiolipin species are given in Table 2.
440 Interestingly, higher V_{max} and k_{cat} values were observed in the presence of CL 18:1 and CL 16:0
441 in equimolar amounts (V_{max} : $19.3 \mu\text{mol min}^{-1} \text{mg}^{-1}$, k_{cat} : 32.8 s^{-1}) or in excess of CL 16:0 (V_{max} :

442 24.6 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, k_{cat} : 41.8 s^{-1}) in comparison to previously obtained kinetic parameters in the
443 presence of *E. coli* CL (V_{max} : 13.7 $\mu\text{mol min}^{-1} \text{mg}^{-1}$, k_{cat} : 23 s^{-1})(14). However, K_m values also
444 increased for both samples (CL 18:1/16:0 1:1, 67 μM ; CL 18:1/16:0 3:1, 120.9 μM) in comparison
445 to *E. coli* CL (10 μM), indicating decreased affinity for $\text{Mg}^{2+}_{\text{free}}$. It should be noted that K_m
446 increased for all samples containing CL 16:0 or CL 16:1. Overall, results indicate an intricate
447 interplay of lipid species CL 18:1 and CL 16:0 for maximum activity of MgtA.

448

449 **MgtA reveals increased thermal stability in the presence of specific cardiolipin species**

450 To determine the effect of cardiolipin on MgtA thermal stability, nano differential scanning
451 fluorimetry (nanoDSF) was performed in the presence of *E. coli* CL or POPE in a 1:100 protein-
452 lipid molar ratio, matching native MS experiments (Figure 5A). Melting temperatures revealed an
453 increase of ca. 7 °C in the presence of *E. coli* CL (T_m : 49.7 °C) relative to purified MgtA without
454 lipids. The presence of POPE (T_m : 42.2 °C) did not show any upwards temperature shift and
455 resembled the control of MgtA without lipids (T_m : 41.8 °C) (Table 3). This indicates that
456 cardiolipin specifically stabilizes MgtA, while the presence of any other lipid environment,
457 represented by POPE as a major *E. coli* lipid component, does not lead to significant protein
458 stabilization. To analyze whether specific cardiolipin species affect MgtA stability differently,
459 nanoDSF was performed with CL 18:1, CL 16:0 and mixtures of both in different ratios (Figure
460 5B and C). In the presence of CL 18:1, the melting temperature of MgtA was significantly
461 increased by ca. 10 °C, while no thermal stabilization was detectable in the presence of CL 16:0
462 alone (Figure 5B). Melting curves in the presence of both, CL 18:1 and CL 16:0, revealed increased
463 stabilization when the higher proportion of CL 18:1 was present in the mixture (Figure 5C). While
464 a 1:3 mixture, with CL 16:0 as the major component, or 1:1 CL 18:1/16:0 showed melting
465 temperatures of 42.5 and 43.2 °C, respectively, in a 3:1 mixture with CL 18:1 as the major
466 component, a melting temperature of 47 °C was determined. This reflects previous results from
467 the enzymatic studies, which revealed activation in the presence of CL 18:1 and CL 16:0, while in
468 the presence of CL 16:0 alone no activity was detectable. Interestingly, thermal stabilization of
469 MgtA is detectable in the presence of CL 18:1 alone, while this single cardiolipin species was not
470 enough for maximal enzymatic activation.

471 As a control, thermal stabilization of MgtA in the presence of only the inhibitors, AlF_4^- and ADP-
472 AlF_4^- , was investigated, which revealed no effect (Figure 5D). It should be mentioned that melting
473 curves in the presence of *E. coli* CL, CL 18:1 and CL 18:1/16:0 mix showed a flattened curve in
474 comparison to the control without lipids. MgtA was incubated with the indicated lipids in excess
475 overnight (1:100 molar ratio between MgtA and lipid) and excess lipid was removed prior to the
476 experiment by centrifugation. In the case of *E. coli* CL, CL 18:1 and CL 18:1/16:0, some of the
477 protein was removed during the centrifugation step, likely pulled down together with the excess
478 lipid. This affected the final protein concentration in the experiment. The different curve shapes
479 might therefore be due to lower protein concentrations or due to the presence of lipid vesicles
480 forming around the protein, which could interfere with the absorbance measured at 350 and 330
481 nm upon unfolding. Interestingly, the removal of protein by centrifugation of excess lipids already
482 indicates binding of MgtA to the *E. coli* CL and CL 18:1. In summary, MgtA is selectively
483 stabilized by a specific cardiolipin species, CL 18:1.

484

485 **Cardiolipin is not sole signal for localization of MgtA to cell poles**

486 Previously, we showed co-localization of a fluorescently-tagged MgtA to *E. coli* C43(DE3) cell
487 poles (14). As C43(DE3) is an expression strain with genetic modifications in comparison to wild
488 type *E. coli* (40), we verified here that fluorescently-tagged MgtA under the arabinose promoter
489 (Figure 6A) also localizes to the cell poles in wild-type strain *E. coli* MG1655 (Figure 6B).
490 Additionally, we could detect localization of MgtA to cell poles in cardiolipin-deficient strain
491 MG1655 Δ *cls* that contains a deletion of all three cardiolipin synthases (Figure 6B). To test whether
492 the cardiolipin functions as a localization signal and the presence of cardiolipin is sufficient for
493 MgtA localization to the cell poles, its localization was analysed in *Vibrio cholerae*. *V. cholerae*
494 is a Gram-negative bacterium with a similar membrane composition as *E. coli* (41). However, it
495 does not encode an endogenous MgtA homologue. Fluorescently-tagged MgtA did not localize to
496 *V. cholerae* cell poles but distributed equally along the entire cell membrane (Figure 6C,D). To
497 verify that the fluorescent signal is constituted by Clover-MgtA, a western blot directed against
498 GFP investigating the expression of Clover-MgtA in *E. coli* and *V. cholerae* was performed. This
499 control revealed that the fluorescence mainly derives from full-length Clover-MgtA in both
500 bacterial strains, and degradation products or free GFP do not contribute to the fluorescent signal

501 (Supplementary figure 4). In summary, these results indicate that MgtA is localized to bacterial
502 cell poles through an unknown mechanism, which is, at least not solely, dependent on CL.

503

504 **Discussion**

505 Recent research has shown the importance of lipids for membrane protein function, affecting
506 protein activity, stability, localization and oligomerization (42). However, one must distinguish
507 effects through site-specific lipid-protein interactions in comparison to effects mediated through
508 chemical and physical properties of the bilayer, e.g. membrane fluidity, tension and thickness or
509 interfacial curvature (42). Lipid-protein interactions have long focused on different lipid classes,
510 varying in their hydrophobic head group. Only recently, the high selectivity of proteins binding
511 only to specific lipids with defined hydrophobic fatty acid tails has been detected.

512 In our work, we investigated lipid-membrane protein interactions of bacterial magnesium
513 transporter MgtA and its highly specific binding of cardiolipin. We identified high specificity of
514 MgtA for two cardiolipin species, CL 18:1 and CL 16:0. They are essential for enzymatic activity
515 and thermal stabilization of MgtA. Further, native MS revealed two specific cardiolipin binding
516 sites. These results highlight for the first time positive cooperative binding of two different species
517 of the same lipid type leading to an activating effect on protein activity. However, as solely CL
518 18:1 mediated MgtA thermal stabilization, differences in the interaction between MgtA and both
519 CL species and, therefore, potentially different roles of both lipid species for MgtA function can
520 be hypothesized. As native MS revealed two cardiolipin binding sites and CL 18:1 has been shown
521 to affect thermal stability, one can hypothesize that both binding sites are preferentially occupied
522 by CL 18:1 leading to MgtA stabilization (Figure 7). CL 16:0 plays an essential role during MgtA
523 ATPase activity in combination with CL 18:1. As activity assays are performed in large excess of
524 cardiolipin, CL 16:0 might not interact directly with MgtA, but play an unknown role in the
525 protein-lipid bilayer environment. However, this is purely speculative and the interaction between
526 MgtA and both CL species requires further investigation.

527 Interestingly, *E. coli* membranes contain a large variety of cardiolipins with the most abundant
528 species containing acyl chains 18:1, 16:1 and 16:0 (30). We highlight the importance of the
529 combination of specific species CL 18:1 and CL 16:0 for MgtA activity, as their combination with

530 another cardiolipin type, CL 16:1, exhibited severely decreased MgtA activity levels. Although
531 higher V_{\max} and k_{cat} was determined in the presence of cardiolipin species CL 18:1 and CL 16:0 in
532 equimolar amounts in comparison to *E. coli* CL, K_m values also increased. This indicates lower
533 affinity of MgtA to $\text{Mg}^{2+}_{\text{free}}$ in the presence of the specific cardiolipin species. However, higher
534 K_m values were always detected in the presence of CL 16:0 or CL 16:1. The more CL species with
535 acyl chain length of 16 were present in the lipid mixture, the higher K_m values were determined.
536 Meanwhile, activity studies performed in the presence of CL 18:1 alone showed a K_m of 12 μM ,
537 similar to the K_m determined in the presence of *E. coli* CL (10 μM). As an anionic lipid cardiolipin
538 binds Mg^{2+} and one might speculate that different cardiolipin species vary in their binding
539 properties to Mg^{2+} . CL has a propensity to form non-bilayer structures, proposed to form dynamic
540 protein-lipid membrane domains (43). The phase behavior of CL has been shown to be dependent
541 on divalent cations, including Mg^{2+} (44), and further acyl chain length and composition can affect
542 lipid phase behavior (45). Therefore, it can be speculated that tested CL species exhibit small
543 differences in their phase behavior, leading to variations in their interaction with Mg^{2+} .

544 Although we have shown the importance of CL 18:1 and CL 16:0 for MgtA activity, the interaction
545 between MgtA and these lipids is not completely understood yet. The molecular shape of acyl
546 chains affect lipid packing and, therefore, membrane bilayer properties (46). Monounsaturated
547 lipids such as CL 18:1 exhibit a kinked shape and tend to form fluid bilayers at physiological
548 temperatures. Therefore, they exhibit higher flexibility and fluidity in comparison to lipids with
549 saturated acyl chains (46). P-type ATPases undergo large conformational changes during their
550 catalytic cycle, requiring flexibility regarding their lipid environment (47). Other members of the
551 P-type ATPase family, including bacterial Cu^{2+} transporter CopA or heavy metal transporter ZntA,
552 have shown higher activity levels upon increased lipid disorder and fluidity in the presence of
553 unsaturated lipids (48-50). Additionally, MgtA exhibited activation in the presence of CL extracted
554 from bovine heart membranes which contain mainly CL 18:2 (29). In summary, a preference of
555 MgtA for CL 18:1 as an abundant monounsaturated lipid of the *E. coli* membrane corresponds to
556 lipid interactions detected for other P-type ATPases.

557 In comparison, the interplay of CL 16:0 and MgtA is more elusive and needs further investigation.
558 As a lipid with saturated acyl chains, CL 16:0 exhibits different properties in comparison to
559 monounsaturated lipids such as CL 18:1. Lipids with saturated acyl chains pack with higher

560 densities and tend to form non-fluid gel phases (46), so one might speculate that CL 16:0 allows
561 close packing of MgtA in lipid-detergent vesicles. *In vivo* the lipid bilayer exerts lateral pressure
562 that supports structural integrity. However, during the purification process, the detergent
563 solubilization decreases lateral pressure and likely increases conformational freedom of the protein
564 (51). Enzymatic assays were not performed in liposomes, but instead in detergent-lipid micelles.
565 Therefore, tighter packing of CL 16:0 might be beneficial for *in vitro* MgtA activation, but its
566 function *in vivo* needs further investigation. Unfortunately, no additional saturated cardiolipin
567 species could be included to investigate the effect of saturated lipids in MgtA-mediated ion
568 transport, as only a limited selection of cardiolipins were commercially available at the time.

569 Further, cooperative binding of cardiolipin to MgtA was revealed, and the determined Hill
570 coefficient of 2.4-3 in the presence of equimolar concentrations of CL 18:1 and CL 16:0 (Table 1)
571 corresponds to the presence of two cardiolipin binding sites revealed by native MS. Allosteric
572 modulation of protein-protein interactions by lipids has been previously described (39), with the
573 *E. coli* ammonium channel AmtB binding two different lipid types, PE and CL (52). Additionally,
574 it was shown that individual lipid binding affected allosteric interaction between AmtB and a
575 soluble regulatory protein, GlnK. The allosteric modulation was highly selective regarding lipid
576 head group and acyl chain composition. In both cases, it was proposed that binding of certain lipids
577 stabilizes a specific conformation of AmtB which exhibits a higher binding affinity to PE or GlnK,
578 respectively. One can speculate a similar mechanism for positive allosteric modulation of MgtA-
579 CL binding as described for AmtB, as CL has been shown to stabilize MgtA. CL interaction at a
580 specific binding site potentially stabilizes MgtA in a conformational state which exhibits increased
581 affinity for the second CL molecule.

582 Interestingly, other members of the P-type ATPase family have shown high specificity towards
583 lipids with a defined acyl chain composition. Their high lipid specificity has been linked to their
584 requirement of high flexibility in the lipid bilayer due to the large conformational changes during
585 the catalytic cycle (47). SERCA adapts to membranes of different hydrophobic thickness by
586 inducing local deformations and a mismatch between hydrophobic thickness of the bilayer and its
587 membrane embedded part for optimal flexibility (53) is required, with 1-Palmitoyl-2-oleoyl-sn-
588 glycerol-3-phosphocholine (16:0-18:1-PC) as the optimal lipid environment for SERCA. Shorter
589 PC species (di-14:1-PC) were shown to surround the protein like a straitjacket, which prevented

590 rotational and translational movement of the transmembrane helices (53). MgtA also exhibited
591 decreased activation in the presence of the shorter cardiolipin species CL 14:1, likely induced
592 through decreased flexibility required for conformational changes. It should be noted that,
593 although we propose CL 16:0 plays a more indirect role for MgtA function in the protein-lipid
594 bilayer environment, it cannot be excluded that the two identified cardiolipin binding sites are
595 occupied with one from each cardiolipin species, one CL 18:1 and one CL 16:0. Small mass
596 differences were identified between the first and second bound cardiolipin in the native MS
597 spectrum (Figure 1), with the second bound CL revealing a slightly higher mass (ca. 1440 +/- 125
598 Da) in comparison to the first one (1380 +/- 200 Da). The calculated masses for CL 18:1 which
599 has four identical oleic acid chains, and CL 16:0 with four palmitic acid tails, are 1501.0 Da and
600 1396.9 Da, respectively, making it tempting to suggest that the first binding site is occupied by CL
601 16:0, while CL 18:1 binds preferentially to the second site. One could speculate that each
602 cardiolipin species could match the hydrophobic thickness of a specific conformational state of
603 MgtA during the catalytic cycle, therefore exhibiting different binding affinities in different states.
604 However native MS did not reveal changes in cardiolipin binding of MgtA locked in different
605 conformations in the presence of inhibitors (Supplementary Figure 3). The resolution of the native
606 MS experiments was too low to distinctively distinguish different cardiolipin species. Therefore,
607 additional native MS studies with higher resolution could allow identification of specific
608 cardiolipin species bound to the two sites on MgtA. Further, native MS of lipid titration of CL 18:1
609 in the presence CL of 16:0, and vice versa, would allow determination of binding affinities of each
610 cardiolipin species and could shed further light on the cooperative interaction between the lipids
611 and the enzyme.

612 Other P-type ATPases also revealed site-specific lipid binding sites. For Na⁺/K⁺-ATPase, two
613 distinct lipid binding sites have been detected (10). Interestingly, distinct properties were assigned
614 for both sites. Lipid binding site A was connected to stabilization of the pump while lipid binding
615 site B did not affect stabilization, but stimulated activity. Both effects are independent. We have
616 no indication of distinct functions associated with the identified two cardiolipin binding sites on
617 MgtA. As removal of cardiolipin strongly affected MgtA stability making native MS spectrum
618 acquisition difficult, a role of both lipid sites for MgtA stability is likely. It should be noted that
619 native MS was performed in DDM, whereas other studies replaced DDM with UDM which has a
620 higher critical micelle concentration, and therefore allows acquisition of the spectra at lower

621 activation energy (10). This might also allow obtaining spectra of less stable MgtA mutants.
622 Although native MS measurements revealed binding of two cardiolipin molecules, no indication
623 of the localization of the lipid binding site on MgtA could be obtained. In future work, structural
624 characterization of lipid-protein interaction by *in silico* or experimental methods, such as X-ray
625 crystallography or cryo-EM, should be conducted to provide answers to the interaction sites and
626 role of the lipids for MgtA function.

627 We also investigated co-localization of MgtA with cardiolipin at bacterial cell poles. Lipids have
628 been linked to membrane protein localization, e.g. polar localization of osmosensory transporter
629 ProP correlates with proportion and localization of cardiolipin and is less pronounced in *cls*⁻ cells
630 (28). Therefore, cardiolipin has been proposed to promote ProP localization (28). Previous studies
631 also showed co-localization of MgtA to cell poles with cardiolipin, proposing a promoting role of
632 cardiolipin for MgtA localization (14). Our work confirmed polar localization of MgtA in *E. coli*
633 wild-type strain MG1655. MgtA also localized at the cells poles in a cardiolipin synthase knockout
634 strain, *E. coli* MG1655 Δ *clsABC*. In this mutant strain no CL was detectable independent of the
635 bacterial growth phase (55). However, full depletion of CL by genetic approaches induced
636 increased PG concentrations at the cell poles. As PG has previously revealed activation of MgtA,
637 although to a lesser extent than CL, MgtA could exhibit promiscuously a relatively high affinity
638 for PG and display a pattern of cellular localization that is not influenced by depletion of CL (56).
639 Therefore, no conclusion on the importance of cardiolipin for MgtA localization in *E. coli* can be
640 assumed. Interestingly, in *V. cholerae*, which contains a similar membrane composition as *E. coli*,
641 but no MgtA homologue, MgtA did not exhibit polar localization, but localized over the entire cell
642 membrane. This highlights that cardiolipin does not function as the sole signal and promoter of
643 polar MgtA localization. Further research must be performed to identify the mechanism for MgtA
644 localization. A previously predicted signal peptide at the N-terminus, enriched in positively
645 charged residues and hypothesized to interact with cardiolipin, was shown not to be involved in
646 MgtA localization at the cell poles (14), but recent results imply a role of the N-terminus as a lipid
647 anchor for MgtA (57).

648 In summary, our results reveal a complex interplay between MgtA and selective cardiolipin
649 species, affecting MgtA activity, stability and localization. Insights obtained here contribute to our
650 understanding of lipid specificity and lipid interaction of membrane proteins in general.

651

652 **Acknowledgment**

653 We would like to thank John Crook and the Weibel lab for the donation of the *E. coli* strains
654 MG1655 and MG1655 Δ *cls*.

655

656 **Declaration of Interests**

657 The authors declare no Competing Financial or Non-Financial Interests.

658

659 **References**

- 660 1. Engelman DM (2005) Membranes are more mosaic than fluid. *Nature* 438(7068):578-
661 580.
- 662 2. van Meer G, Voelker DR, & Feigenson GW (2008) Membrane lipids: where they are and
663 how they behave. *Nat Rev Mol Cell Biol* 9(2):112-124.
- 664 3. Corradi V, *et al.* (2019) Emerging Diversity in Lipid-Protein Interactions. *Chem Rev*
665 119(9):5775-5848.
- 666 4. Reading E, Laganowsky A, Allison TM, & Robinson CV (2014) Membrane Proteins
667 Bind Lipids Selectively To Modulate Their Structure And Function. *Protein Science*
668 23:231-231.
- 669 5. Laganowsky A, *et al.* (2014) Membrane proteins bind lipids selectively to modulate their
670 structure and function. *Nature* 510(7503):172-175.
- 671 6. Mangialavori IC, *et al.* (2012) Autoinhibition mechanism of the plasma membrane
672 calcium pump isoforms 2 and 4 studied through lipid-protein interaction. *Biochem J*
673 443(1):125-131.
- 674 7. Cornelius F, Habeck M, Kanai R, Toyoshima C, & Karlisch SJ (2015) General and
675 specific lipid-protein interactions in Na,K-ATPase. *Biochim Biophys Acta* 1848(9):1729-
676 1743.
- 677 8. Cornelius F (2001) Modulation of Na,K-ATPase and Na-ATPase activity by
678 phospholipids and cholesterol. I. Steady-state kinetics. *Biochemistry-Us* 40(30):8842-
679 8851.
- 680 9. Habeck M, *et al.* (2015) Stimulation, inhibition, or stabilization of Na,K-ATPase caused
681 by specific lipid interactions at distinct sites. *J Biol Chem* 290(8):4829-4842.
- 682 10. Habeck M, Kapri-Pardes E, Sharon M, & Karlisch SJ (2017) Specific phospholipid
683 binding to Na,K-ATPase at two distinct sites. *Proc Natl Acad Sci U S A* 114(11):2904-
684 2909.
- 685 11. Contreras FX, Ernst AM, Wieland F, & Brugger B (2011) Specificity of intramembrane
686 protein-lipid interactions. *Cold Spring Harb Perspect Biol* 3(6).

- 687 12. Hossain KR & Clarke RJ (2019) General and specific interactions of the phospholipid
688 bilayer with P-type ATPases. *Biophys Rev* 11(3):353-364.
- 689 13. Pohland AC & Schneider D (2019) Mg²⁺ homeostasis and transport in cyanobacteria - at
690 the crossroads of bacterial and chloroplast Mg²⁺ import. *Biol Chem* 400(10):1289-1301.
- 691 14. Subramani S, Perdreau-Dahl H, & Morth JP (2016) The magnesium transporter A is
692 activated by cardiolipin and is highly sensitive to free magnesium in vitro. *Elife* 5.
- 693 15. Groisman EA, *et al.* (2013) Bacterial Mg²⁺ homeostasis, transport, and virulence. *Annu*
694 *Rev Genet* 47:625-646.
- 695 16. Bourret TJ, Liu L, Shaw JA, Husain M, & Vazquez-Torres A (2017) Magnesium
696 homeostasis protects Salmonella against nitrooxidative stress. *Sci Rep* 7(1):15083.
- 697 17. Cunrath O & Bumann D (2019) Host resistance factor SLC11A1 restricts Salmonella
698 growth through magnesium deprivation. *Science* 366(6468):995-+.
- 699 18. Apell HJ (2004) How do P-type ATPases transport ions? *Bioelectrochemistry* 63(1-
700 2):149-156.
- 701 19. Bublitz M, Poulsen H, Morth JP, & Nissen P (2010) In and out of the cation pumps: P-
702 Type ATPase structure revisited. *Current Opinion in Structural Biology* 20(4):431-439.
- 703 20. Morth JP, *et al.* (2011) A structural overview of the plasma membrane Na⁺,K⁺-ATPase
704 and H⁺-ATPase ion pumps. *Nat Rev Mol Cell Biol* 12(1):60-70.
- 705 21. Ford DC, Joshua GWP, Wren BW, & Oyston PCF (2014) The importance of the
706 magnesium transporter MgtB for virulence of Yersinia pseudotuberculosis and Yersinia
707 pestis. *Microbiology (Reading)* 160(Pt 12):2710-2717.
- 708 22. Park M, Nam D, Kweon DH, & Shin D (2018) ATP reduction by MgtC and Mg(2+)
709 homeostasis by MgtA and MgtB enables Salmonella to accumulate RpoS upon low
710 cytoplasmic Mg(2+) stress. *Mol Microbiol* 110(2):283-295.
- 711 23. Lin Z, *et al.* (2017) Virulence and Stress Responses of Shigella flexneri Regulated by
712 PhoP/PhoQ. *Front Microbiol* 8:2689.
- 713 24. Chadani Y, *et al.* (2017) Intrinsic Ribosome Destabilization Underlies Translation and
714 Provides an Organism with a Strategy of Environmental Sensing. *Mol Cell* 68(3):528-+.
- 715 25. Park SY, Cromie MJ, Lee EJ, & Groisman EA (2010) A bacterial mRNA leader that
716 employs different mechanisms to sense disparate intracellular signals. *Cell* 142(5):737-
717 748.
- 718 26. Gall AR, *et al.* (2016) Mg²⁺ regulates transcription of mgtA in Salmonella Typhimurium
719 via translation of proline codons during synthesis of the MgtL peptide. *Proc Natl Acad*
720 *Sci U S A* 113(52):15096-15101.
- 721 27. Moncrief MB & Maguire ME (1999) Magnesium transport in prokaryotes. *J Biol Inorg*
722 *Chem* 4(5):523-527.
- 723 28. Romantsov T, *et al.* (2007) Cardiolipin promotes polar localization of osmosensory
724 transporter ProP in Escherichia coli. *Mol Microbiol* 64(6):1455-1465.
- 725 29. Schlame M (2008) Cardiolipin synthesis for the assembly of bacterial and mitochondrial
726 membranes. *J Lipid Res* 49(8):1607-1620.
- 727 30. Garrett TA, O'Neill AC, & Hopson ML (2012) Quantification of cardiolipin molecular
728 species in Escherichia coli lipid extracts using liquid chromatography/electrospray
729 ionization mass spectrometry. *Rapid Commun Mass Spectrom* 26(19):2267-2274.
- 730 31. Subramani S & Morth JP (2016) Heterologous Expression and Purification of the
731 Magnesium Transporter A (MgtA) in Escherichia coli. *Bio-protocol* 6(22):e2001.

- 732 32. Waldo GS, Standish BM, Berendzen J, & Terwilliger TC (1999) Rapid protein-folding
733 assay using green fluorescent protein. *Nat Biotechnol* 17(7):691-695.
- 734 33. Chung HS & Raetz CR (2010) Interchangeable domains in the Kdo transferases of
735 *Escherichia coli* and *Haemophilus influenzae*. *Biochemistry-Us* 49(19):4126-4137.
- 736 34. Paintdakhi A, *et al.* (2016) Oufiti: an integrated software package for high-accuracy, high-
737 throughput quantitative microscopy analysis. *Mol Microbiol* 99(4):767-777.
- 738 35. Konijnenberg A, van Dyck JF, Kailing LL, & Sobott F (2015) Extending native mass
739 spectrometry approaches to integral membrane proteins. *Biol Chem* 396(9-10):991-1002.
- 740 36. van Dyck JF, Konijnenberg A, & Sobott F (2017) Native Mass Spectrometry for the
741 Characterization of Structure and Interactions of Membrane Proteins. *Methods Mol Biol*
742 1635:205-232.
- 743 37. Neet KE (1995) Cooperativity in enzyme function: equilibrium and kinetic aspects.
744 *Methods Enzymol* 249:519-567.
- 745 38. Edelstein SJ & Le Novere N (2013) Cooperativity of allosteric receptors. *J Mol Biol*
746 425(9):1424-1432.
- 747 39. Patrick JW, *et al.* (2018) Allostery revealed within lipid binding events to membrane
748 proteins. *Proc Natl Acad Sci U S A* 115(12):2976-2981.
- 749 40. Miroux B & Walker JE (1996) Over-production of proteins in *Escherichia coli*: mutant
750 hosts that allow synthesis of some membrane proteins and globular proteins at high
751 levels. *J Mol Biol* 260(3):289-298.
- 752 41. Giles DK, Hankins JV, Guan Z, & Trent MS (2011) Remodelling of the *Vibrio cholerae*
753 membrane by incorporation of exogenous fatty acids from host and aquatic environments.
754 *Mol Microbiol* 79(3):716-728.
- 755 42. Corradi V, *et al.* (2018) Lipid-Protein Interactions Are Unique Fingerprints for
756 Membrane Proteins. *ACS Cent Sci* 4(6):709-717.
- 757 43. Mileykovskaya E & Dowhan W (2009) Cardiolipin membrane domains in prokaryotes
758 and eukaryotes. *Biochim Biophys Acta* 1788(10):2084-2091.
- 759 44. Vail WJ & Stollery JG (1979) Phase changes of cardiolipin vesicles mediated by divalent
760 cations. *Biochim Biophys Acta* 551(1):74-84.
- 761 45. Sankaram MB, Powell GL, & Marsh D (1989) Effect of acyl chain composition on salt-
762 induced lamellar to inverted hexagonal phase transitions in cardiolipin. *Biochim Biophys*
763 *Acta* 980(3):389-392.
- 764 46. Ernst R, Ejsing CS, & Antonny B (2016) Homeoviscous Adaptation and the Regulation
765 of Membrane Lipids. *J Mol Biol* 428(24 Pt A):4776-4791.
- 766 47. Thogersen L & Nissen P (2012) Flexible P-type ATPases interacting with the membrane.
767 *Curr Opin Struct Biol* 22(4):491-499.
- 768 48. Autzen HE, *et al.* (2018) Interactions of a Bacterial Cu(I)-ATPase with a Complex Lipid
769 Environment. *Biochemistry-Us* 57(28):4063-4073.
- 770 49. Zimmer J & Doyle DA (2006) Phospholipid requirement and pH optimum for the in vitro
771 enzymatic activity of the *E. coli* P-type ATPase ZntA. *Biochim Biophys Acta*
772 1758(5):645-652.
- 773 50. Moore BM, Lentz BR, & Meissner G (1978) Effects of sarcoplasmic reticulum Ca²⁺-
774 ATPase on phospholipid bilayer fluidity: boundary lipid. *Biochemistry-Us* 17(24):5248-
775 5255.
- 776 51. Palsdottir H & Hunte C (2004) Lipids in membrane protein structures. *Biochim Biophys*
777 *Acta* 1666(1-2):2-18.

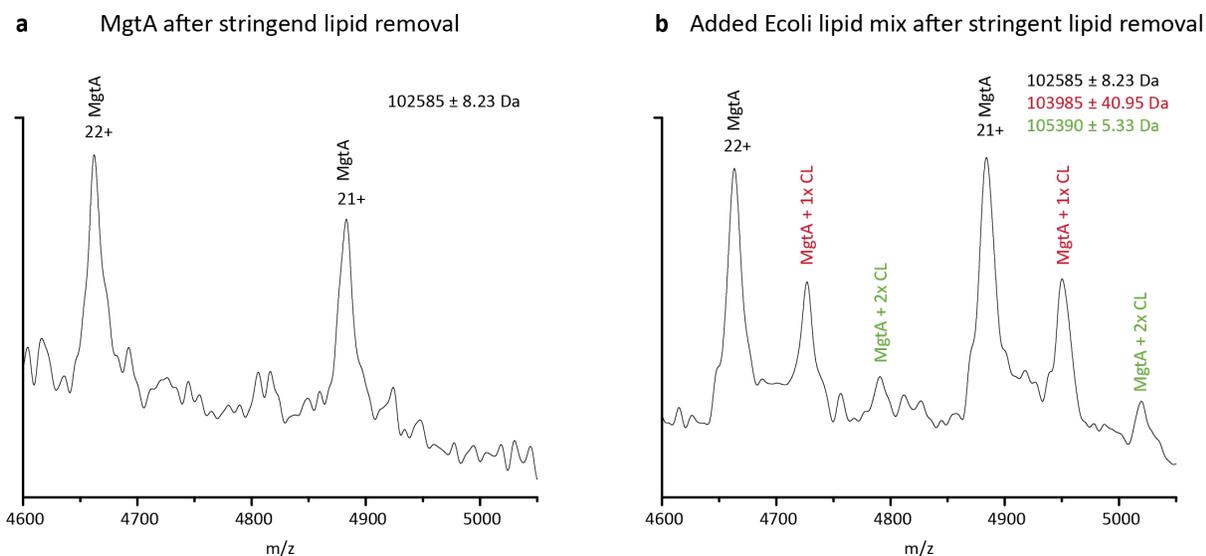
- 778 52. Cong X, Liu Y, Liu W, Liang X, & Laganowsky A (2017) Allosteric modulation of
 779 protein-protein interactions by individual lipid binding events. *Nat Commun* 8(1):2203.
 780 53. Sonntag Y, *et al.* (2011) Mutual adaptation of a membrane protein and its lipid bilayer
 781 during conformational changes. *Nat Commun* 2:304.
 782 54. Cohen E, *et al.* (2005) Purification of Na⁺,K⁺-ATPase expressed in *Pichia pastoris*
 783 reveals an essential role of phospholipid-protein interactions. *J Biol Chem*
 784 280(17):16610-16618.
 785 55. Ryabichko S, *et al.* (2020) Cardiolipin is required in vivo for the stability of bacterial
 786 translocon and optimal membrane protein translocation and insertion. *Sci Rep*
 787 10(1):6296.
 788 56. Oliver PM, *et al.* (2014) Localization of anionic phospholipids in *Escherichia coli* cells. *J*
 789 *Bacteriol* 196(19):3386-3398.
 790 57. Jephthah S, Mansson LK, Belic D, Morth JP, & Skepo M (2020) Physicochemical
 791 Characterisation of KEIF-The Intrinsically Disordered N-Terminal Region of Magnesium
 792 Transporter A. *Biomolecules* 10(4).
 793

794

795 Figures

796

797 Figure 1



	m/z difference	
	22+	21+
MgtA > MgtA + 1x CL	62.54 ± 7.35	66.17 ± 10.83
MgtA + 1x CL > MgtA + 2x CL	65.32 ± 5.03	68.88 ± 6.97

	mass difference (Da)	
	22+	21+
MgtA > MgtA + 1x CL	1375 ± 161	1389 ± 238
MgtA + 1x CL > MgtA + 2x CL	1437 ± 105	1446 ± 146

798

799 **Figure 1: MgtA shows selective binding of cardiolipin in native mass spectrometry**

800 Purified MgtA solubilized in 3x CMC DDM was analysed by native MS (a) in the absence or (b) presence
801 of total *E. coli* lipid extract. Total *E. coli* lipid extract contains only 9.8 % CL, while the remaining lipids
802 are 57.5 % PE, 15.1 % PG and 17.6 % unknown (Avanti Polar Lipids). The left spectrum shows the most
803 intense charge states 22+ and 21+ of purified MgtA after lipid removal by high CMC DDM. The right
804 spectrum exhibits MgtA as seen after incubation with *E. coli* lipid mixture. This sample was washed with
805 3x CMC DDM to remove weakly bound lipids leaving only the strongest interacting lipids. The spectrum
806 on the right reveals additions of approximately 1400 and 2800 Da (red and green) which correspond to the
807 molecular weight of one and two cardiolipin molecules, respectively. The spectra were recorded in TOF
808 mode and the key MS settings are the voltages (sampling cone, trap collision, transfer collision and trap
809 bias) and pressures (backing pressure, trap cell pressure). These values were 200 V, 150 V, 150 V, 45 V,
810 6.41 mbar and 2.00×10^{-2} mbar respectively.

811

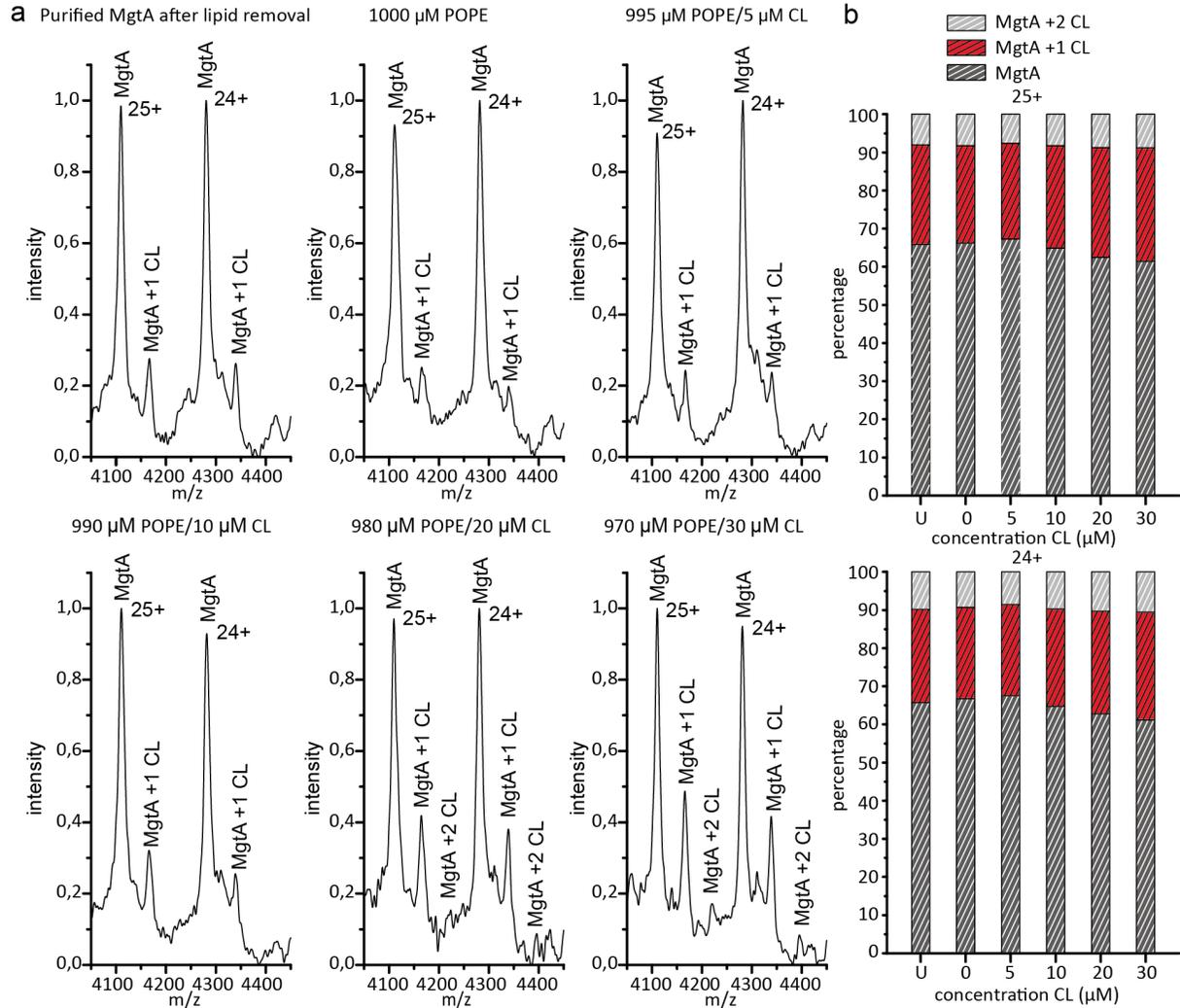
812

813

814

815

816 **Figure 2**



817

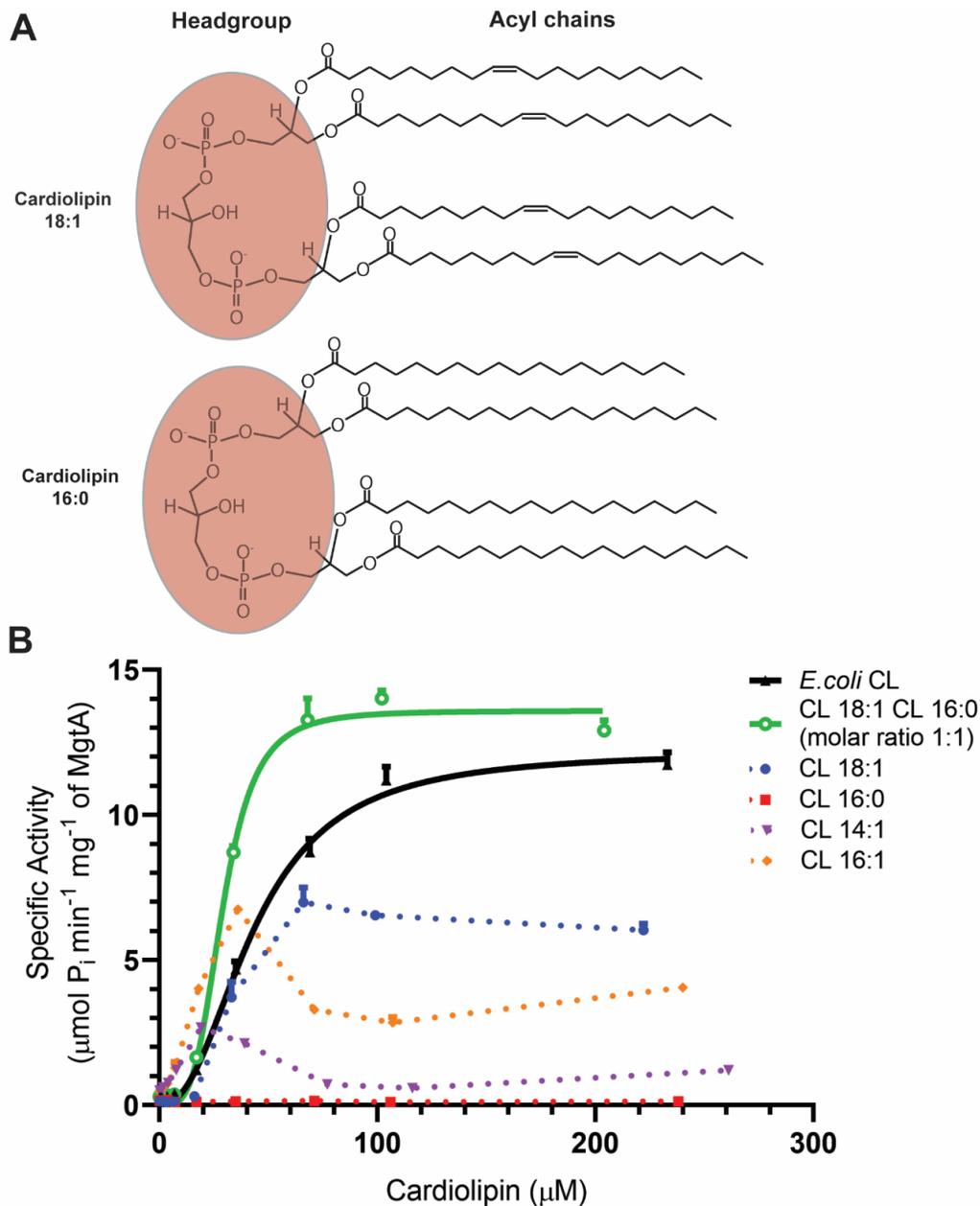
818 **Figure 2: Cardiolipin competition assay of MgtA in the presence of POPE reveals increase in CL**
 819 **binding**

820 **a)** MS spectra showing a lipid competition assay of MgtA with increasing CL concentrations (0 to 30 μM)
 821 and 1000 to 970 μM POPE. In the spectra 1x and 2x CL adducts are assigned. In **b)** the increase in 1x or
 822 2x CL molecules bound to MgtA is shown as percentages according to the increase in CL added to the
 823 sample. The spectra were recorded in TOF mode and the key MS settings are the voltages: sampling cone,
 824 trap collision, transfer collision, trap bias and pressures: backing pressure, trap cell pressure. These values
 825 where 150 V, 200 V, 5 V, 45 V, 8.41 mbar and 2.66×10^{-2} mbar.

826

827

828 **Figure 3**



829

830 **Figure 3: *E. coli* MgtA shows cooperative binding of specific cardiolipin species**

831 (A) Cardiolipins 18:1 and 16:0 contain the same polar hydrophilic headgroup (red), but exhibit differences
832 in the length and saturation degree of their acyl chains.

833 (B) Concentration-dependent activation of MgtA ATPase activity by *E. coli* extracted cardiolipin (black),
834 a 1:1 mixture of cardiolipin 18:1 and 16:0 (green) or individual cardiolipin species. Lipids are prepared as
835 described in the Materials and Methods section; specific activity is determined by measuring phosphate
836 release. Curves are representatives from three independent experiments, showing mean \pm SD. Enzymatic
837 activation curves were fitted to an allosteric sigmoidal model (GraphPad, Prism8). Fitting curves are shown

838 (lines) for data with least squares fit of 0.95 or above with parameters described in table 1. Dotted lines
839 represent activation curves that could not be fitted with the allosteric sigmoidal model.

840

841 **Table 1:** Allosteric fit of purified MgtA in the presence of different cardiolipin species. The V_{\max} and Hill
842 coefficient were determined by least squares fit of the data from Figure 1B, as described in materials and
843 methods. Only data from least squares fit of 0.95 or above was included in the analysis.

	V_{\max} ($\mu\text{mol min}^{-1} \text{mg}^{-1}$)	Hill coefficient
<i>E. coli</i> CL	11.0 +/- 0.8	2.4 +/- 0.6
CL 18:1 CL 16:0 (molar ratio 1:1)	14.1 +/- 1.9	3.0 +/- 0.8

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

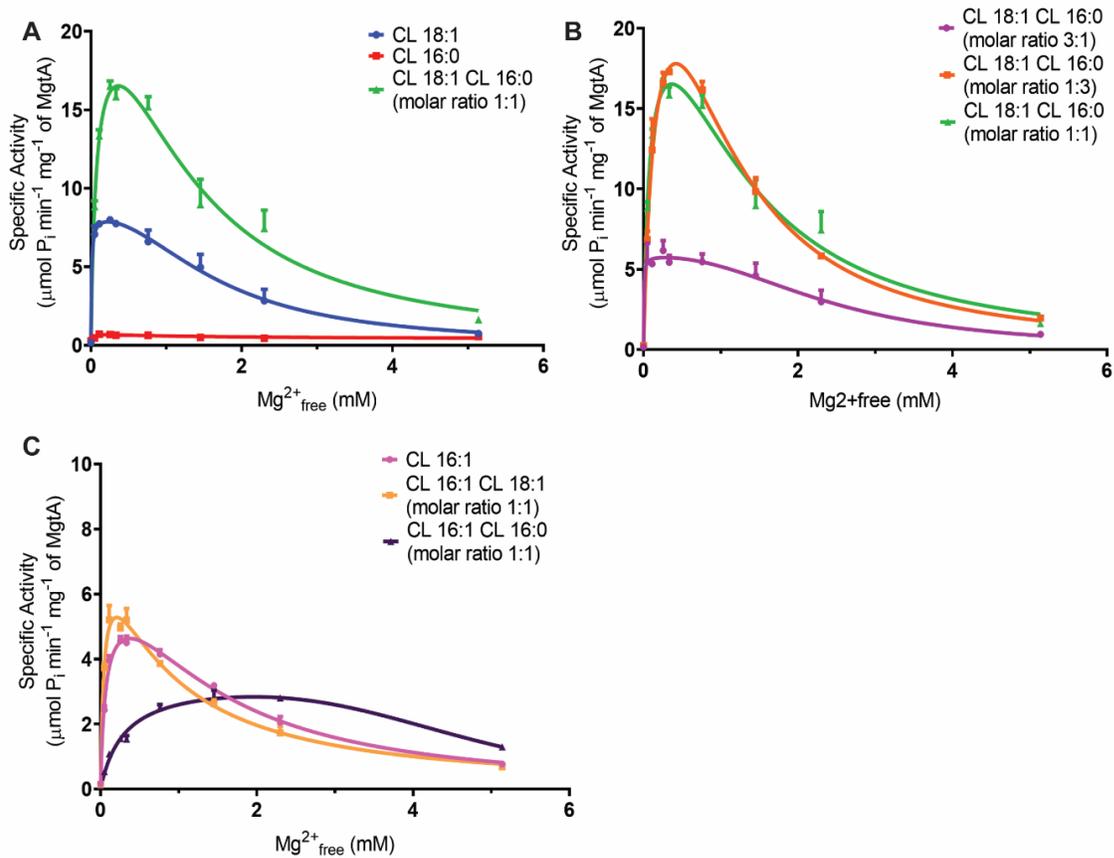
859

860

861

862

863 **Figure 4**



864

865 **Figure 4: *E. coli* MgtA shows selective activation by cardiolipin species 18:1 and 16:0 in specific molar**
866 **ratios**

867 ATP hydrolysis was measured in the presence of specific cardiolipin species with increasing concentrations
868 of Mg^{2+} at 3 mM ATP. Lipid concentration was kept constant at 12 mM (molar ratio of MgtA to lipid
869 1:5000). Curves are representatives from three independent experiments, showing mean \pm SD. Enzymatic
870 data was fitted as a function of Mg^{2+}_{free} using Michaelis-Menten equation (Graphpad Prism 8) as described
871 in the method section. Kinetic parameters are described in table 2. Mg^{2+}_{free} profiles of (A) CL 18:1, CL 16:0
872 and an equimolar mixture of both or (B) CL 18:1 and CL 16:0 in different molar ratios was assessed. (C)
873 As a control Mg^{2+}_{free} profiles of CL 18:1 or CL 16:0 in combination with CL 16:1 were performed.

874

875

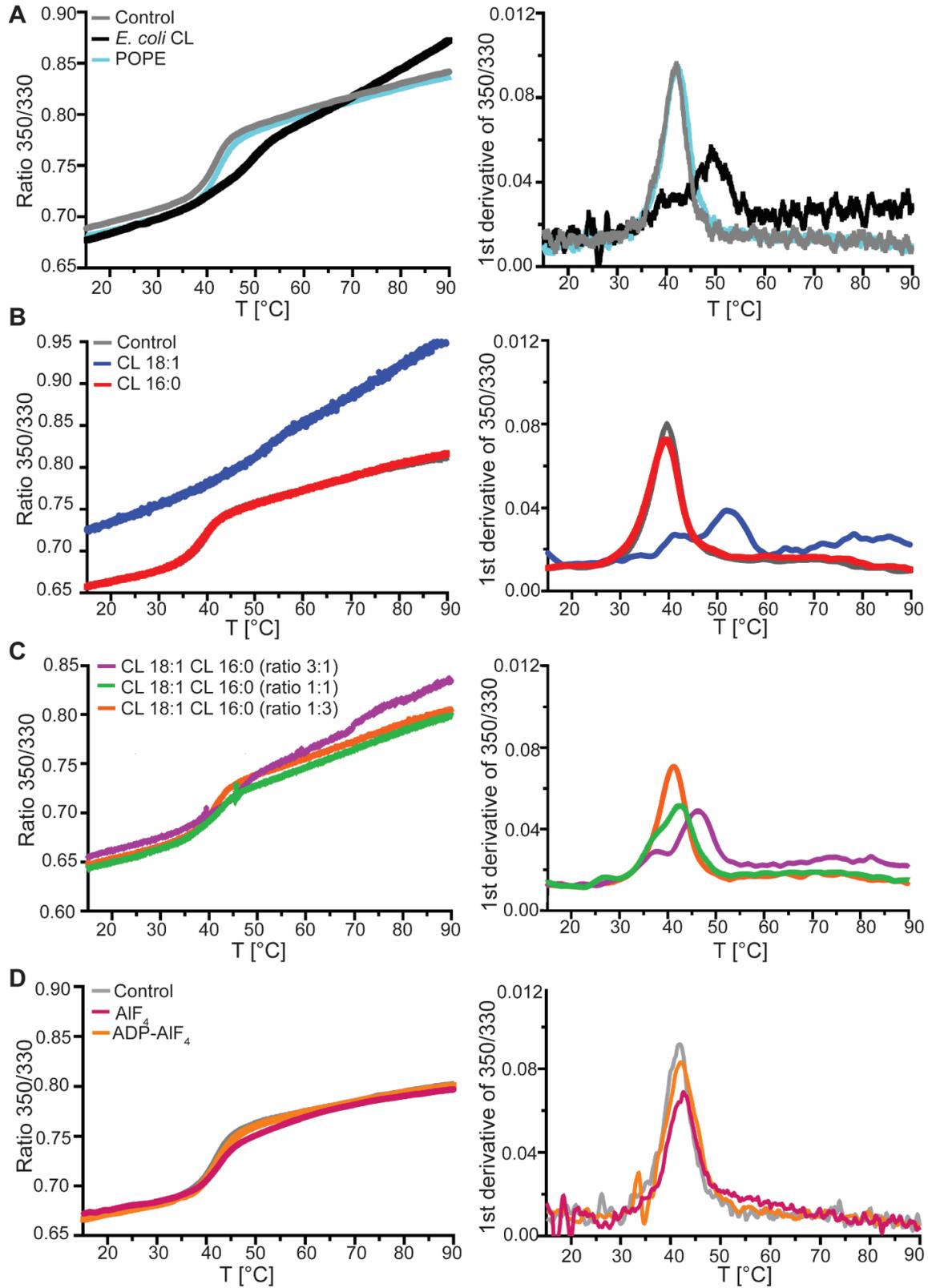
876 **Table 2: Kinetic property of purified MgtA in the presence of different cardiolipin species.** The
877 apparent V_{\max} , K_m and the turnover number (K_{cat}) values were determined by least squares fit of the data
878 from figure 2A, 2B and 2C and put in references to previously obtained kinetics (Subramani *et al.* (2016)).

	V_{\max} ($\mu\text{mol min}^{-1} \text{mg}^{-1}$)	K_m (μM)	K_{cat} (s^{-1})
<i>E. coli</i> CL (Subramani <i>et al.</i> (2016))	13.7 +/- 0.2	10 +/- 0.6	23 +/- 0.3
CL 18:1	7.6 +/- 0.7	12 +/- 7.0	12.8 +/- 1.2
CL 16:0	N/A	N/A	N/A
CL 18:1 CL 16:0 (ratio 1:1)	19.3 +/- 2.4	67 +/- 3.0	32.8 +/- 4.1
CL 18:1 CL 16:0 (ratio 3:1)	7.1 +/- 0.9	4.4 +/- 2.6	12.06 +/- 1.5
CL 18:1 CL 16:0 (ratio 1:3)	24.6 +/- 6.7	120.9 +/- 48.9	41.8 +/- 11.4
CL 16:1	6.0 +/- 0.5	103 +/- 5	10.2 +/- 0.8
CL 18:1 CL 16:1 (ratio 1:1)	6.6 +/- 0.8	40 +/- 12	11.3 +/- 1.4
CL 16:0 CL 16:1 (ratio 1:1)	3.9 +/- 0.5	285 +/- 44	6.7 +/- 0.8

879

880

881



882

883 **Figure 5: MgtA shows increased temperature stabilization in the presence of the specific cardiolipin**
 884 **species, cardiolipin 18:1**

885 Nano differential scanning fluorimetry (nanoDSF) curves (left) and their first derivative (right) of purified
 886 MgtA in the presence of (A) cardiolipin and 1-Palmitoyl-2-oleoyl-sn-glycero-3-phosphoethanolamine
 887 (POPE) or specific cardiolipin species CL 18:1 and CL 16:0 (B) alone and (C) in different molar ratios
 888 were obtained. MgtA was incubated overnight with indicated lipids in a 1:100 molar ratio and samples were
 889 centrifuged to remove excess of lipids. Melting curves were performed on Prometheus NT.48 (Nanotemper)
 890 in a temperature range of 15 °C – 95 °C and a temperature gradient of 1 °C per minute. A paired, one-sided
 891 t-test was applied. ***p<0.001. (D) Additionally, melting curves were performed on MgtA in the presence
 892 of inhibitors, ADP-AlF₄⁻ and AlF₄⁻. MgtA was incubated with inhibitors in 1:500 molar ratio (2.95 mM) for
 893 two hours at 4 °C before melting curves were performed. Curves are representatives from three independent
 894 experiments, showing mean
 895 +/- SD. Obtained melting temperatures are described in table 3.

896

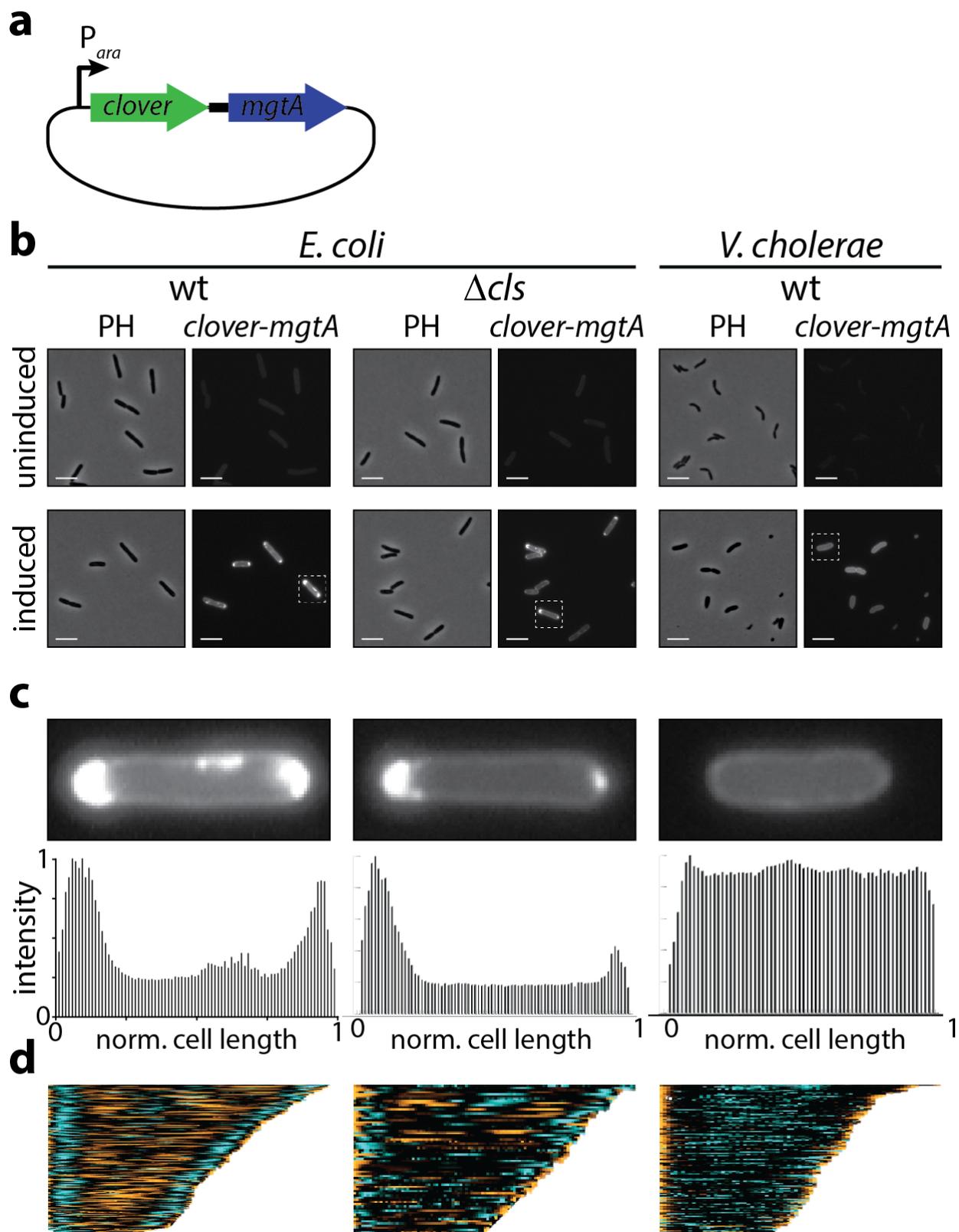
897

898 **Table 3: Melting temperatures of MgtA in the presence of lipids and inhibitors determined by**
 899 **nanoDSF**

	Melting temperature [°C]
MgtA (control)	41.8 +/- 0.4
MgtA in presence of E. coli cardiolipin	49.7 +/- 1.4 ***
MgtA in presence of POPE	42.2 +/- 0.4
MgtA in presence of cardiolipin 18:1	52.5 +/- 1.2 ***
MgtA in presence of cardiolipin 16:0	40.3 +/- 1.0
MgtA in presence of cardiolipin 18:1 16:0 (molar ratio 3:1)	47.0 +/- 0.8
MgtA in presence of cardiolipin 18:1 16:0 (molar ratio 1:1)	43.2 +/- 1.2
MgtA in presence of cardiolipin 18:1 16:0 (molar ratio 1:3)	42.5 +/- 1.2
MgtA in presence of AlF₄⁻	41.8 +/- 0.6
MgtA in presence of ADP-AlF₄⁻	41.6 +/- 0.6

900

901



902

903

904 **Figure 6: MgtA localization to the pole is not solely cardiolipin-dependent**

905 To assess whether cardiolipin functions as the main signal for polar MgtA localization in *E. coli*, in vivo
906 imaging studies were performed using a GFP-tagged MgtA construct.

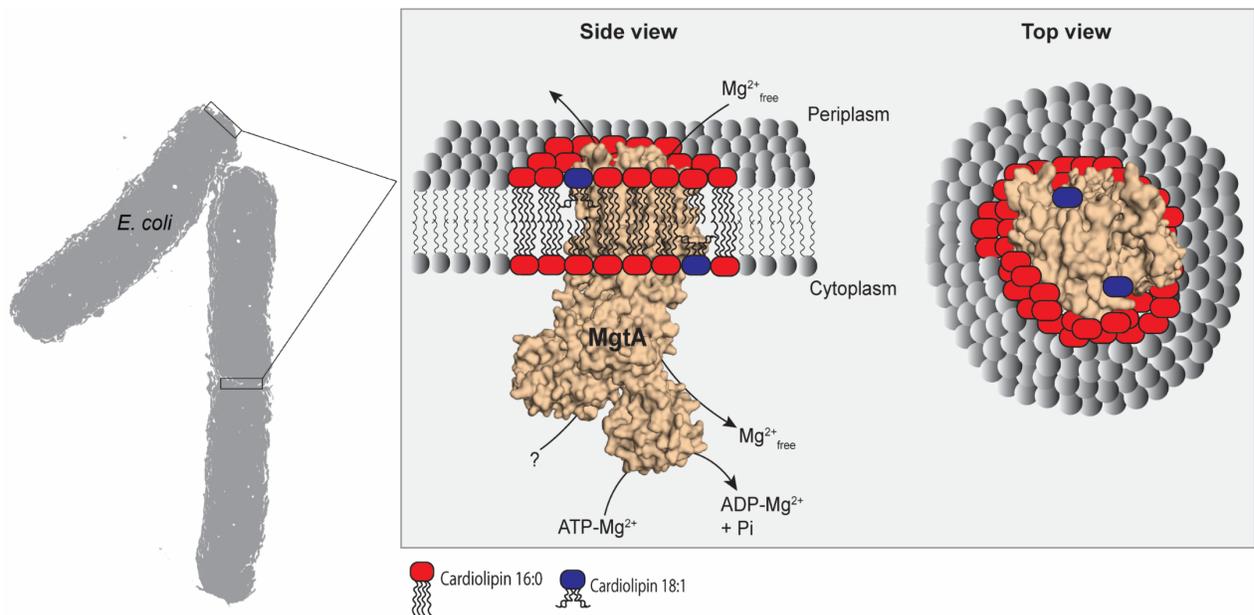
907 (A) Schematic representation of the Clover-MgtA expression plasmid. The expression is controlled by the
908 arabinose-inducible P_{ara} promoter.

909 (B) The localization of MgtA-GFP was assessed in wild-type *E. coli* strain MG1655, cardiolipin-deficient
910 *E. coli* MG1655 Δcls , a knockout strain of cardiolipin synthase (Δcls), and wild type *Vibrio cholerae* strain
911 C6706. The phase contrast channel (PH) and GFP channel (Clover-MgtA) of bacteria grown in the presence
912 of 0.2 % arabinose (induced) or in its absence (uninduced) are shown. White scale bars are 5 μm .

913 (C) Intensity profiles of single, representative cells (highlighted by a dotted white box in (B)). The
914 intensities are normalized to the brightest area in each cell, respectively.

915 (D) Clover-MgtA localization profiles of cells sorted by their cell length. This representation approximates
916 the localization patterns through the cell cycle with small cells approximating newborn cells and long cells
917 approximating cells immediately prior to division. The cell lengths are normalized to the longest cell in
918 each analysis and the fluorescent intensities are normalized to the brightest area in all cells from each strain.
919 *E. coli* MG1655 (n=171 cells), *E. coli* MG1655 Δcls (n=75 cells), and *V. cholerae* C6706 (n=125 cells).

920
921
922



923
924
925

926 **Figure 7: Model of cardiolipin interaction with MgtA**

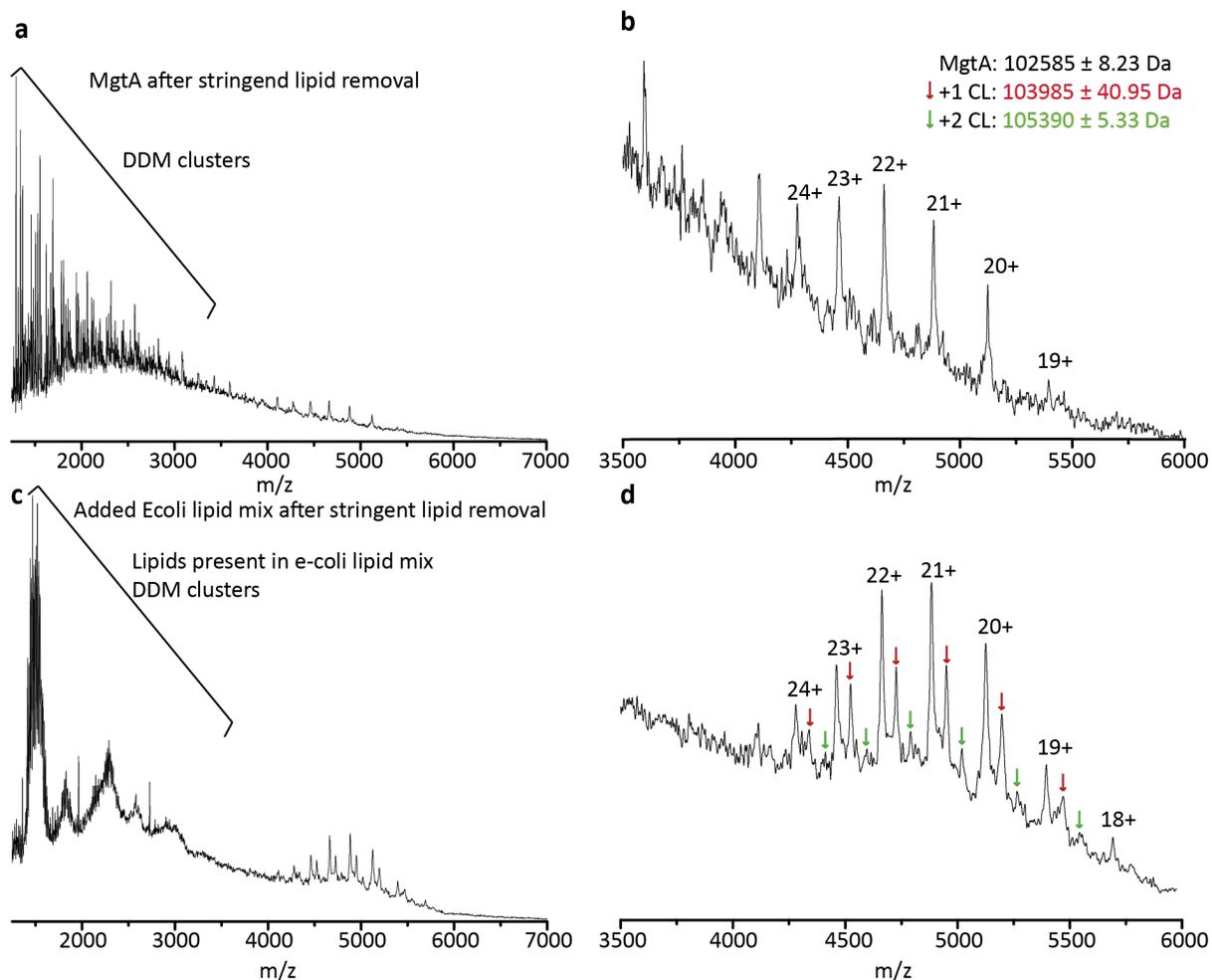
927 MgtA protein localizes in the inner membrane at the *E. coli* cell poles, which represent cardiolipin-rich
928 regions. Association of MgtA with cardiolipin is essential for MgtA activity and stability, but MgtA shows
929 high selectivity for specific cardiolipin species. We propose specific interaction between MgtA and two CL
930 18:1 molecules, affecting MgtA thermal stabilization. Further, CL18:1 is required for MgtA ATPase

931 activity. CL 16:0 is needed for maximal MgtA activation but does not affect protein stabilization
932 significantly.

933

934

935 Supplementary figures

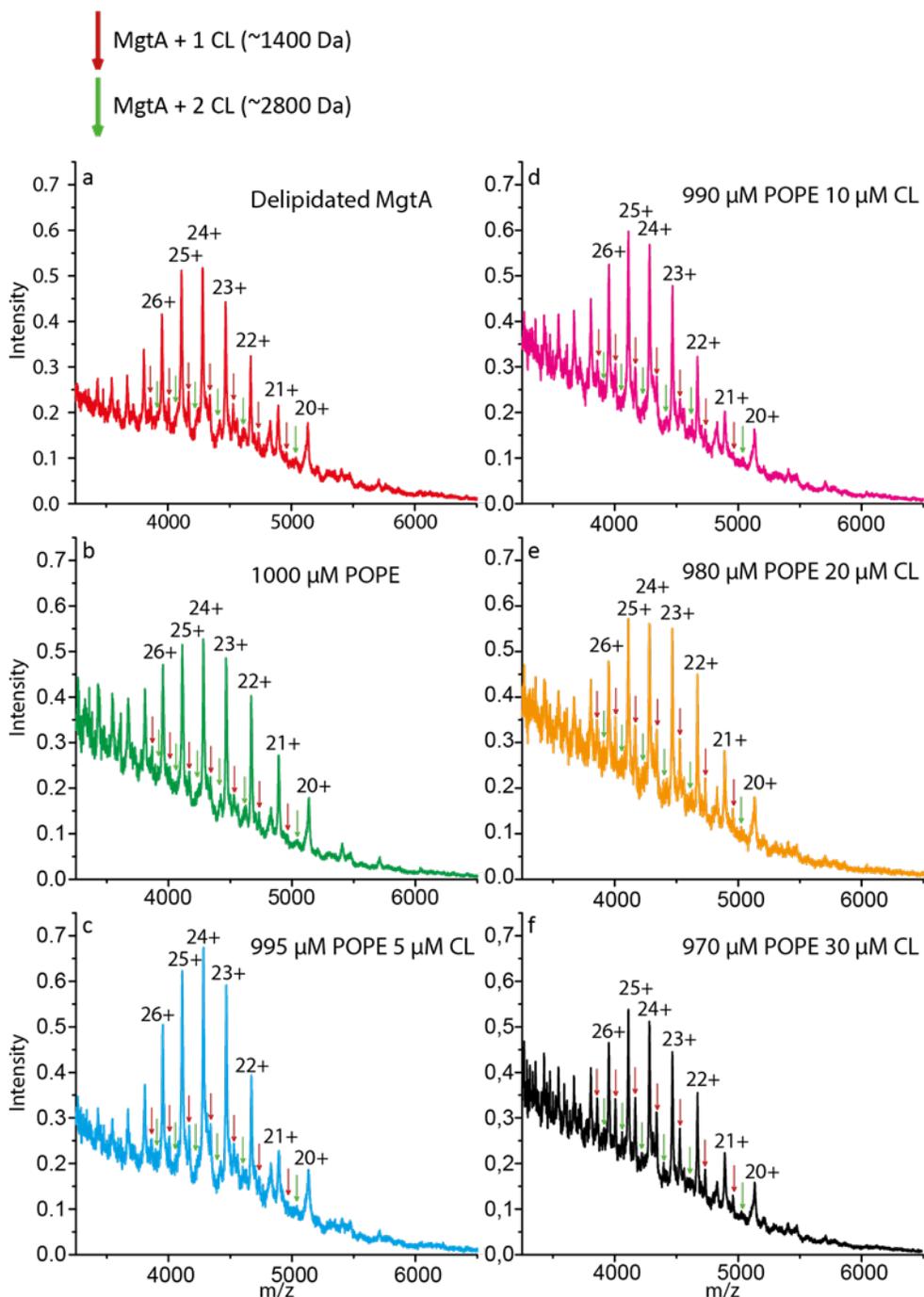


936

937 Supplemental figure 1: MgtA shows selective binding of cardiolipin in native mass spectrometry

938 Purified MgtA, solubilized in 3x CMC DDM, was analysed by native MS (**a**, **b**) in the absence or (**c**, **d**)
939 presence of total *E. coli* lipid extract. Total *E. coli* lipid extract contains only 9.8 % CL, while the remaining
940 lipids are 57.5 % PE, 15.1 % PG and 17.6 % unknown (Avanti Polar Lipids). The upper spectrum shows
941 delipidated MgtA in DDM (19+ to 24+) with DDM clusters in the lower m/z region, while the lower spectra
942 exhibits MgtA with *E. coli* lipid mixture (18+ to 24+) with a mixture of DDM clusters and lipids in the
943 lower m/z region. Close-up on MgtA spectra (right) reveals additions of approximately 1400 and 2800 Da
944 (red and green arrows) which correspond to the molecular weight of one and two cardiolipin molecules,
945 respectively. The spectra were recorded in TOF mode and the key MS settings used to record these spectra
946 are the voltages (sampling cone, trap collision, transfer collision and trap bias) and pressures (backing

947 pressure, trap cell pressure). These values were 200 V, 150 V, 150 V, 45 V, 6.41 mbar and 2.00×10^{-2} mbar
948 respectively.



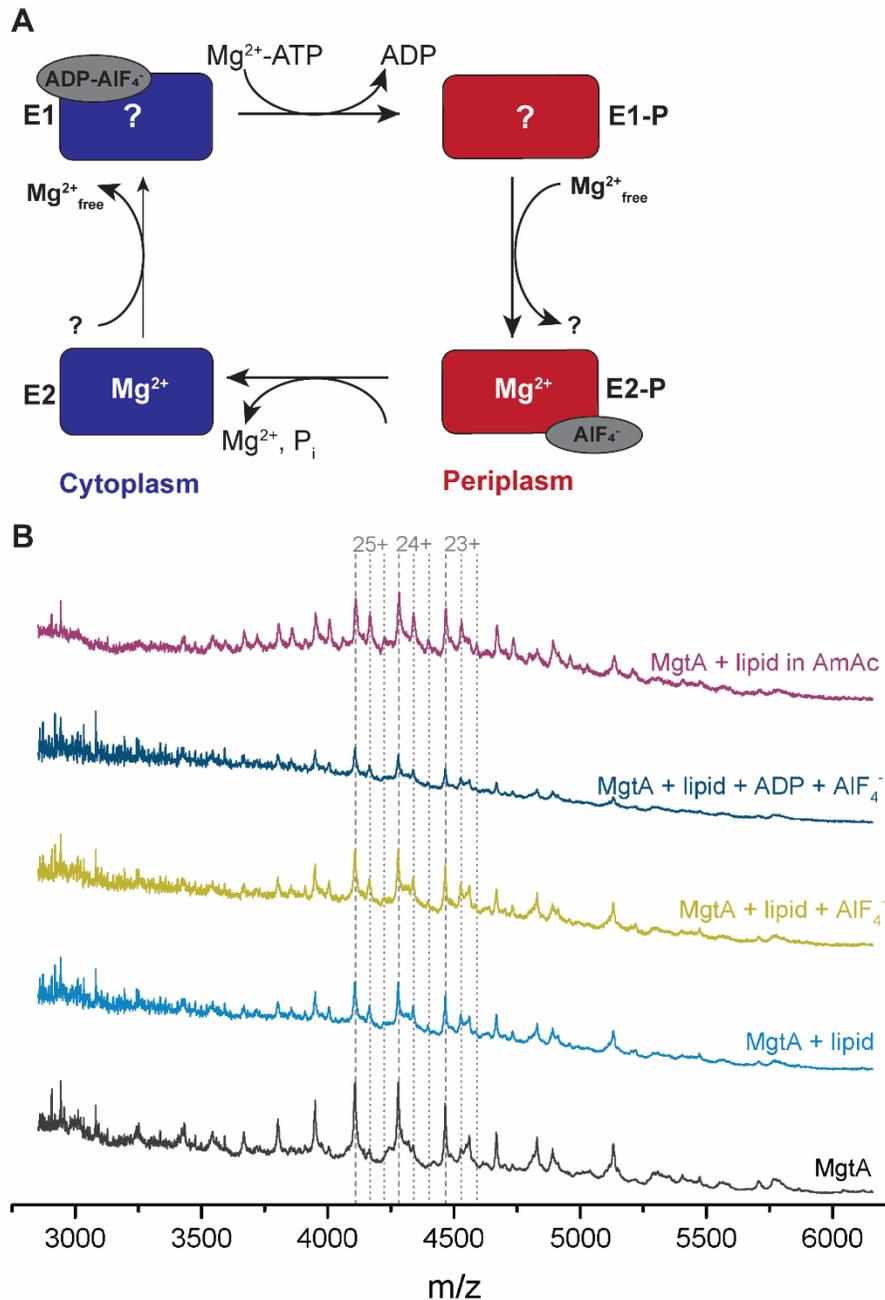
949

950 **Supplemental figure 2:** MS spectra showing a competition assay of MgtA with POPE and increasing CL
951 concentrations measured under activating conditions.

952 The concentrations in the spectra a to f of CL used is 0 to 30 μ M CL and 1000 to 970 μ M POPE lipid. The
953 red arrow is the +1 CL (~1400 Da) and the green arrow +2 CL (~2800 Da). The spectra were recorded in
954 TOF mode and the key MS settings are the voltages: sampling cone, trap collision, transfer collision, trap

955 bias and pressures: backing pressure, trap cell pressure. These values were 150 V,
956 200 V, 5 V, 45 V, 8.41 mbar and 2.66×10^{-2} mbar.

957



958

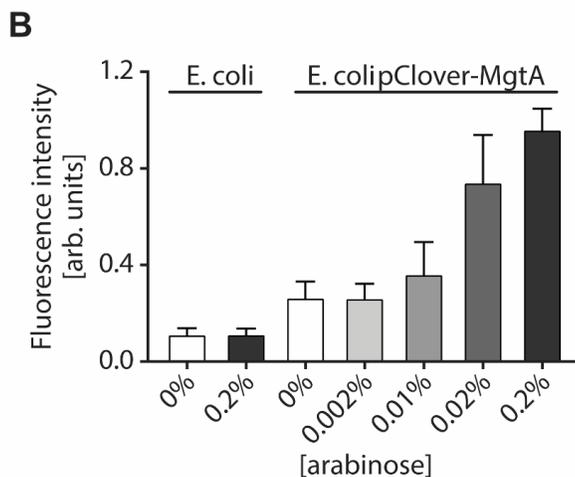
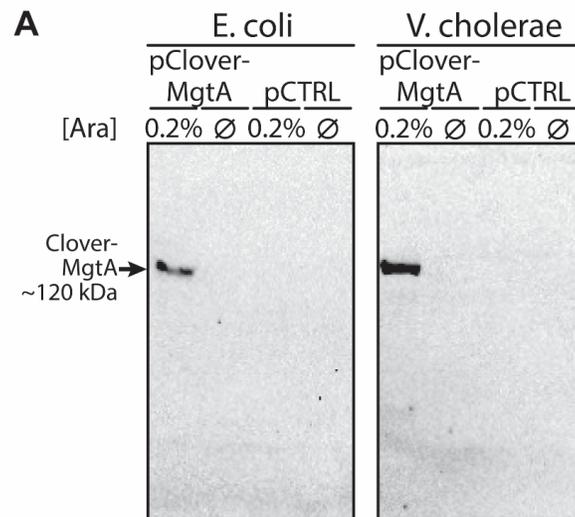
959 **Supplementary figure 3: Cardiolipin binding in the presence of MgtA inhibitors**

960 To determine binding of cardiolipin to MgtA dependent on the stages of transport cycle, lipid binding of
961 MgtA was determined in the presence of AIF_4^- and ADP-AIF_4^- , which lock MgtA in the E2 and E1 state,
962 respectively.

963 (A) Mg^{2+} -transport cycle of by MgtA is described by the Post-Albers cycle, in which MgtA alternates
 964 between different conformational states, termed E1 and E2 state. The E1 state has high affinity and is open
 965 for ions binding from the cytoplasm. The ion transported from the cytoplasm to the periplasm in the case
 966 of MgtA remains unknown. Upon ion binding, autophosphorylation of MgtA by ATP hydrolysis is induced,
 967 forming the E1P state. The phosphorylation induces domain rearrangements leading to the E2P state, which
 968 is now open to the periplasmic side. E2P has low affinity for the bound ions and high affinity for the counter
 969 ion, Mg^{2+} . The exchange of counter ions dephosphorylates the enzyme and forms the E2 state. Upon further
 970 conformational changes the enzyme returns to the E1 state and the counterions are released into the
 971 cytoplasm. MgtA inhibitors, AlF_4^- and $ADP-AlF_4^-$, lock the enzyme in the indicated conformational steps.
 972 (B) MgtA was incubated with inhibitors in 1:100 molar ratio for two hours, following incubation with total
 973 *E. coli* lipid extract in 1:100 molar ratio. Spectra of MgtA lipidated with cardiolipin in the absence (violet,
 974 blue) and in the presence of inhibitors (dark blue, yellow) are shown. A control spectrum of unlipidated
 975 MgtA (black) is included.

976

977



978

979

980 **Supplementary figure 4: Expression of Clover-MgtA**

981

982 (A) Western blot analysis of *E. coli* MG1655 and *V. cholerae* C6706 expressing Clover-fusion to MgtA.
983 *E. coli* and *V. cholerae* containing either a plasmid encoding a translational fusion of Clover-MgtA under
984 control of the arabinose inducible BAD promoter (pClover-MgtA) or the empty pBAD33.1 plasmid
985 (pCTRL) were grown in the presence (0.2%) or absence (Ø) of arabinose. Samples of these cultures were
986 analyzed by Western blot with a monoclonal antibody directed against GFP. (B) Quantification of Clover-
987 MgtA fluorescence by flow cytometry. *E. coli* containing the MgtA-Clover fusion plasmid (pClover-MgtA)
988 or no plasmid were induced with different arabinose concentrations. The fluorescence intensity was
989 quantified after excitation at 488 nm. The bars depict the mean fluorescence, the error bars the standard
990 deviation of at least 450,000 cells.

991

